Mouse Vk gene classification by nucleic acid sequence similarity

Robert Strohal, Arno Helmberg, Guido Kroemer, and Reinhard Kofler

Institute for General and Experimental Pathology, University of Innsbruck Medical School, Innsbruck, Austria

Abstract. Analyses of immunoglobulin (Ig) variable (V) region gene usage in the immune response, estimates of V gene germline complexity, and other nucleic acid hybridization-based studies depend on the extent to which such genes are related (i.e., sequence similarity) and their organization in gene families. While mouse Igh heavy chain V region (V_H) gene families are relatively well-established, a corresponding systematic classification of Igk light chain V region (V_k) genes has not been reported. The present analysis, in the course of which we reviewed the known extent of the V_k germline gene repertoire and V_k gene usage in a variety of responses to foreign and self antigens, provides a classification of mouse V_k genes in gene families composed of members with >80% overall nucleic acid sequence similarity. This classification differed in several aspects from that of V_H genes: only some V_k gene families were as clearly separated (by >25% sequence dissimilarity) as typical V_H gene families; most V_k gene families were closely related and, in several instances, members from different families were very similar (>80%) over large sequence portions; frequently, classification by nucleic acid sequence similarity diverged from existing classifications based on amino-terminal protein sequence similarity. Our data have implications for V_k gene analyses by nucleic acid hybridization and describe potentially important differences in sequence organization between V_H and V_k genes.

Introduction

The ability of the immune system to recognize virtually any antigen is mediated by the enormous sequence variability in the amino-terminal region of immunoglobulin (Ig) heavy and light chains. Among other mechanisms, this diversity is generated by somatic juxtaposition of gene segments that are separated in the germline, termed variable (V), diversity (heavy chain only), and joining (J) gene segments (reviewed by Tonegawa 1983, Alt et al. 1986). V genes contribute all residues of the first and second complementarity determining region (CDR) of both heavy and light chains, as well as part of the light chain CDR-3, and hence contribute the majority of antigen contact residues (Kabat et al. 1987). In mice, several hundred V_H and V_k (over 90% of all serum Ig is of the Igk isotype) gene segments exist in the germ line (Brodeur and Riblet 1984, Livant et al. 1986, Cory et al. 1981, Kofler et al. 1989). These genes can be very similar or may differ by over 40% nucleotides, and V region classifications based on nucleic and/or amino acid sequence similarity have been proposed (Brodeur and Riblet 1984, Dildrop 1984, Potter et al. 1982). Thus, mouse V_H genes have been grouped in 11 V_H gene families in which members generally share >80% of their nucleic acid sequence within, and <70–75% between, families (Brodeur and Riblet 1984, Winter et al. 1985, Kofler 1988, Reininger et al. 1988). Individual members of a given family cross-hybridize in nucleic acid hybridization assays only with members of their own family. These V_H gene families correspond well with a V_H region classification based on similarities at the protein level (Dildrop 1984). Understanding V_H gene relatedness on the nucleic acid sequence level has greatly facilitated studies regarding the expression of different V_H gene families during ontogeny (Yancopoulos et al. 1984, Perlmutter et al. 1985) and in response to foreign and self antigens (Manser et al. 1987b, Kofler et al. 1987a). These studies have thus provided an important insight into B-cell repertoire generation.

V_k classifications reported to date are confined to the protein level. One attempt to systematically classify Vk proteins was based on the partial amino acid sequence up to the invariant cysteine in position 23 (Cys23), leading to 26 Vk subgroups, designated Vk_Cys (Potter 1977). A
modified classification, based on the length and similarity of the amino termini up to the invariant tryptophan 35 (Trp35) of 79 Vk proteins, was introduced in 1982 (VkTrp subgroups; Potter et al. 1982). Four of the VkCys subgroups were condensed and two new groups were added, resulting in a total of 24 Vk subgroups, six of which are still defined only by sequences up to Cys23. This classification has now been generally accepted and, although an extended comparison at the nucleic acid level has never been reported, the corresponding Vk protein subgroups have been widely used synonymously with Vk gene families. More recently, we have performed a detailed restriction fragment length polymorphism (RFLP) analysis with DNA probes corresponding to 16 Vk protein subgroups, and obtained evidence that such protein groups may not necessarily correspond to gene families analogous to those described for Vk genes (Kofler et al. 1989). Since a large number of full-length Vk nucleic acid sequences has been reported, it is now possible to address, by direct sequence comparison, the matter of whether Vk genes can be organized into gene families, as has been accomplished with Vh genes, and how such Vk gene families relate to the existing Vk protein groups. This issue is of considerable interest for Vk gene usage determinations, repertoire estimates, genomic mapping, and similar studies using nucleic acid hybridization, since such procedures depend on relatedness between Vk groups, gene families, and corresponding DNA probes.

We compiled 248 full-length Vk nucleic acid sequences from the literature and several databases, and assigned them to existing Vk protein classifications with subsequent grouping into gene families comprised of members with >80% overall nucleic acid sequence similarity. Our analysis revealed that the current classification in Vk protein groups or subgroups frequently did not reflect relatedness on the nucleic acid sequence level. Furthermore, Vk gene family organization differed in important aspects from that of Vh gene families; only some of the Vk gene families were clearly separated by sequence dissimilarity of >25%, as is usually observed in Vh gene families. The remaining families were more similar to each other and, in several instances, large portions of genes from different families shared >80% of their sequences, leading to cross-hybridization between those families in hybridization analyses. In addition, although ancillary to the primary aim of this study, we reviewed the specificities encoded by the various Vk gene families and estimated their germline gene complexity.

Methods and nomenclature

Vk nucleic acid sequence bank. A database was constructed consisting of Vk nucleic acid sequences from the Genetic Sequence Data Bank (GenBank, Los Alamos, New Mexico), E. A. Kabat's collection (Kabat et al. 1987), and other publications. Only sequences encoding the entire mature Vk protein were included in the database. If applicable, sequence portions encoding untranslated region, leader sequence, introns, or J segments were removed prior to comparisons. This primary database of 248 full-length Vk sequences was then condensed to a final database of 109 (Fig. 1) by deleting duplicate sequences and those differing by only 1 to 4 base pairs (bp).

Vk protein groups and subgroups. All nucleic acid sequences were translated into amino acids and organized into Vk protein groups and subgroups. Assignment to Vk protein groups (labeled I to VII) was based on the length of the amino-terminal sequence up to the invariant Trp35 (41, 40, 39, 36, 35, 34, and 33 residues, respectively; Kabat et al. 1987). Organization into Vk protein subgroups was based on <13 substitutions up to Trp35 (VkTrp subgroups; Potter et al. 1982). Sequences meeting assignment criteria for more than one subgroup were assigned to the subgroup with the best match.

Vk gene families. Analogous to Vk gene families, we defined a "Vk gene family" as a group of nucleic acid sequences that exhibit >80% overall sequence similarity with every member of this family, and <80% with Vk genes from other families. In nucleic acid hybridization analyses under defined stringency conditions (Brodeur and Riblet 1984), all members of a gene family can be expected to cross-hybridize with each other. The Vk gene family nomenclature proposed in this study was adjusted as far as possible to that used for Vk protein subgroups, in order to minimize confusion in the literature; when Vk protein subgroups and Vk gene families corresponded to each other (e.g., Vk21), the Vk subgroup designation was used for the Vk gene family as well. Vk gene families comprising two or more Vk protein subgroups were given the designation of the respective subgroups (e.g., the V4/5 family comprised V4 and V5 protein subgroups). Addition of capital letters to the designation indicates that a Vk protein subgroup included members from two distinct Vk gene families (e.g., the V9 protein subgroup comprised members from two distinct Vk gene families, termed V9A and V9B, respectively). V3RF and (tentatively) V38C were two new gene families that could not be related unambiguously to any Vk protein subgroup and, hence, were named after a prototypic sequence.

Organization of mouse Vk sequences on the protein and nucleic acid level

The major goal of this study was to investigate the organization of mouse Vk genes in terms of nucleic acid sequence similarity, and to determine the relationship of such organization to existing Vk protein classifications. To this end, we first compiled 109 distinct (i.e., >4 bp different), full-length Vk nucleic acid sequences that were used as a database for subsequent analyses (Fig. 1). The sequences were translated into amino acids (Fig. 2) and assigned to protein groups and subgroups (Table 1).

Classification into protein groups was based on the number of residues up to the invariant Trp35 and, hence, was unambiguous in all instances. However, this classification was of limited practical value, since it frequently did not reflect structural relatedness (i.e., sequence similarity) between Vk sequences. For example, group V included members of several, sometimes quite dissimilar, Vk gene families (V23, V12/13, V3RF, V11, V9A, V9B, V10, V38C, V19/28). On the other hand,
similar members from a single V \(_k\) gene family (V\(_k\)4/5) were present in different groups (IV and VI).

Organization of \(V_k\) proteins into subgroups using <13 mismatches up to Trp25 as a criterium (Potter et al. 1982) better reflected primary structure similarities, although such organization frequently led to multiple assignments, in which cases only a single assignment for the sequence representing the best match was included (Table 1). Moreover, as will be shown below, this classification repeatedly failed to adequately reflect overall similarity at the nucleic acid sequence level. Finally, some sequences (discussed below) could not be assigned unambiguously to any existing \(V_k\)Trp subgroup.

We then determined whether \(V_k\) nucleic acid sequences could be organized into gene families (analogous to \(V_H\) genes), and how such families related to \(V_k\) protein groups and subgroups. For this purpose, all \(V_k\) genes in the data bank were arranged in groups of >80% sequence similarity, which were termed \(V_k\) gene families. The characteristics of these families and their relationship to \(V_k\) protein groups and subgroups are detailed below. A quick summary outlining how the different classifications correspond to each other is presented in Table 2.

\(V_{21}\) gene family. All \(V_{21}\) genes fulfilled the criteria for a typical \(V\) gene family, i.e., all members were >80% similar (mostly >90%) and differed from all other \(V_k\) sequences by at least 25%. This gene family corresponded completely to protein subgroup \(V_{21}\) which, in turn, coincided with \(V_k\) protein group III. Five germline genes have been cloned (Heinrich et al. 1984) and approximately ten expressed sequences have been published. \(V_{21}\) genes were used in response to influenza hemagglutinin (Clarke et al. 1985, Meek et al. 1989) and major histocompatibility complex class II antigens (Devaux et al. 1985), and encoded some lupus-associated autoantibodies (Shlomchik et al. 1987c), which differed from all known \(V_{21}\) germline genes by >30 bp and may have derived from an unknown germline gene, all other expressed sequences were very similar to, and hence probably derived from, known \(V_{21}\) germline genes. RFLP (Kofler et al. 1989) and gene cloning analyses (Heinrich et al. 1984) suggested an estimated 6 to 13 \(V_{21}\) germline genes in the genome of most inbred strains of mice.

Finally, an incomplete \(V_{21}\) sequence (VM201, Meek et al. 1989), which was therefore not included in our data bank, should be mentioned as it lacked two codons in CDR-1 in comparison to other \(V_{21}\) sequences. Unless caused by somatic events, this would make the corresponding germline gene the only \(V_k\) gene with 37 codons up to Trp35.

\(V_{23}\) gene family. Similar to \(V_{21}\), \(V_{23}\) sequences were well separated from all other \(V_k\) sequences, and formed a gene family that corresponded entirely to its protein counterpart, the \(V_{23}\) subgroup (protein group V). One germline gene has been reported (Pech et al. 1981) that was subsequently observed in RFs from BALB/c mice (Shlomchik et al. 1987a), and that probably encoded an (NZB \(\times\) NZW)F\(_1\) RNA-specific autoantibody (Eilat et al. 1988).

Additional \(V_{23}\) genes, more distant from the above germline gene but closely related to each other, possibly derived from a second \(V_{23}\) germline gene and encoded nitrophenyl-specific anti-idiotypes (Sablitzky and Rajewsky 1984) and a creatine-kinase-specific antibody (Buckel et al. 1987). A nonfunctional \(V_{23}\) member was cloned from an MRL/n RF-producing hybridoma and might correspond to another \(V_k\) (pseudo) gene (Kofler et al. 1989). Our previous RFLP analyses suggested the presence of four to eight \(V_{23}\) germline genes in the genome of most inbred strains of mice. However, this may represent an over-estimate due to cross-hybridization of the more conserved 3' portion of the \(V_{23}\) probe with \(V_k\) sequences (Kofler et al. 1989, and below).
Fig. 2. Amino acid sequences deduced from 109 $V_k$ nucleic acid sequences contained in the $V_k$ database. Dots have been introduced to maximize homology; $X$, undetermined amino acids. Remainder of legend as for Figure 1.
**Table 1. V_{k}** nucleic acid sequence database*

| Code | V_{k} | Group | Subgroup | Spec | Strain | Class | Ref |
|------|-------|-------|----------|------|--------|-------|-----|
| 001-005 21 | III | 21 | G | BALB/c | N/A | (1) |
| 006 | | | HA | BALB/c | IgG | (2) |
| 007,008 | RF | MRL/lpr | IgG | (3) |
| 009 | nf | BALB/c | N/A | (4) |
| 010 | 23 | V | G | BALB/c | N/A | (5) |
| 011 | nf | N/R | N/A | (6) |
| 012 | RNA | (NZB × W)F_{1} | F_{1} | IgG | (7) |
| 013 | nf | MRL/hi | N/A | (8) |
| 014,015 | Anti-ID | C57BL/6 | IgG | (9) |
| 016 | CK | BALB/c | IgG | (10) |
| 017-019 4/5 IV | 4, 5 | G | BALB/c | N/A | (11) |
| 020 | G | BALB/c | N/A | (12) |
| 021 | histone | MRL/lpr | IgG | (13) |
| 022 | DNA | MRL/lpr | IgG | (14) |
| 023,024 | OX | BALB/c | IgG | (15) |
| 025 | OX | BALB/c | IgG | (16) |
| 026 | RF | BALB/c | IgG | (17) |
| 027 | ALP | BALB/c | IgG | (18) |
| 028-032 VI | 4 | G | BALB/c | N/A | (11) |
| 033-038 | CaAg | BALB/c | IgG | (19) |
| 039 | unknown | BALB/c | IgG | (20) |
| 040 | RF | MRL/lpr | IgG | (3) |
| 041 | CD20 | BALB/c | IgG | (21) |
| 042,043 4/5 VI | 4, OX | BALB/c | IgG | (22) |
| 044-048 | OX | BALB/c | IgG | (23) |
| 049-054 | OX | BALB/c | IgG | (16) |
| 055 | OX | BALB/c | IgG | (15) |
| 056 | OX | BALB/c | IgG | (15) |
| 057 | OX | BALB/c | IgG | (15) |
| 058,059 12/13 V | 12-13 | G | BALB/c | N/A | (24) |
| 060,061 | Anti-C57BL/6 | IgG | (9) |
| 062 | RF V | ambiguous | RF | MRL/lpr | IgM | (13) |
| 063 | 11 | V | 11 | nf | NZB | N/A | (25) |
| 064 | 9A | V | 9 | G | BALB/c | N/A | (26) |
| 065 | G | BALB/c | N/A | (27) |
| 066 | lysosome | BALB/c | IgG | (28) |
| 067 | DNA | (NZB × W)F_{1} | IgM | (14) |
| 068 | 9B | V | 9 | G | BALB/c | N/A | (5) |
| 069 | degoxin | A/J | IgG | (29) |
| 070 | BrBRC | CBA/J | IgM | (30) |
| 071 | 10 | V | 10 | G | A/J | N/A | (31) |
| 072 | nf | NZB | N/A | (25) |
| 073 | 38C V | ambiguous | unknown | C3H/HeN | IgM | (32) |
| 074 | HA | BALB/c | IgG | (33) |
| 075 | 24/25 II | 24 | G | BALB/c | N/A | (34) |
| 076,077 | G | BALB/c | N/A | (35) |
| 078,079 | RF | C3H/He | IgM | (3) |
| 080 | RF | BALB/c | IgM | (17) |
| 081 | GAC | A/J | IgM | (36) |
| 082-084 1 II | 1 | G | BALB/c | N/A | (37) |
| 085 | dextran | BALB/c | IgA | (38) |
| 086 | GAT | BALB/c | IgG | (39) |
| 087 | RF | BALB/c | IgM | (17) |
| 088 | RF | Anti-ID | BALB/c | IgG | (40) |
| 089 | 2 II | 2 | DNA | (NZB × W)F_{1} | IgM | (14) |
| 090 | 8 I | 8 | DNP | BALB/c | IgA | (41) |
| 091 | HEL | BALB/c | IgG | (28) |
| 092 | RF | MRL/lpr | IgA | (3) |
| 093,094 | RF | BALB/c | IgM | (17) |
| 095 | RF | 129/1v | IgM | (17) |
| 096 | HA | BALB/c | IgM | (33) |
| 097 | 22 | I | 22 | PC | BALB/c | IgG | (42) |
| 098 | 19/28 V | 28 | G | BALB/c | N/A | (43) |
| 099 | RF | MRL/lpr | IgM | (8) |
| 100 | 14-15-19 | 15-19 | CEA | BALB/c | IgG | (44) |
| 101 | TNP | BALB/c | IgM | (45) |
| 102 | CASA | N/R | IgG | (46) |
| 103 | CEA | N/R | IgG | (47) |
| 104-108 | RF | BALB/c | IgM | (17) |
| 109 | RF | MRL/lpr | IgG | (3) |

---

* Only sequences encoding the entire mature V_{k} region and differing by >4 bp are contained in this database (see Methods).

**V_{k}4/5** gene family. V_{k}Trp subgroups V_{k}4 (groups IV and VI) and V_{k}5 (group IV) were encoded by highly similar (around 90%) genes forming a gene family, termed V_{k}4/5, that was separated from all other V_{k} sequences by >25% of their nucleotides. This was the largest V_{k} gene family, composed of approximately 25-50 members, as deduced from RFLP (Kofler et al. 1989) and gene cloning (Even et al. 1985) studies. Fourteen germline genes (ten V_{k}4 and four V_{k}5 genes) have been isolated thus far (Even et al. 1985, Höchtl et al. 1982). V_{k}4/5 genes were found in antibodies specific for galactan (Heller et al. 1987), oxazolone (Koartenin and Maekelae 1987, Berek and Milstein 1987), dextran (Skidler et al. 1985, Aolakkar et al. 1987), the lymphocyte surface marker CD20 (Liu 1987b), alpenrolanol (Nahmias et al. 1988), red blood cells (Pennell et al. 1988), and DNA, histone, and Ig self antigens (Shlomchik et al. 1987c, Kofler et al. 1987b, Kofler et al. 1988a, Shlomchik et al. 1987b).

---

**V_{k}12/13 gene family.** The sequences encoding V_{k}12–13 proteins (group V) formed another well-defined family that corresponded to all V_{k}12–13 subgroup proteins (Kabat et al. 1987, Potter et al. 1982). Two germline genes have been published (Nishioka and Leder 1980, Seidman et al. 1978), one of which (K2) may be involved in the nitrophenyl-specific anti-idiotypic response (Sablitsky and Rajewsky 1984). A more distant V_{k}12/13 gene encoded anti-idiotypic light chains in the GAT.
Table 2. Correlation between \( V_k \) gene families and \( V_k \) protein groups
and subgroups.

| \( V_k \) gene family | \( V_{kCys} \) subgroup | \( V_{kTrp} \) subgroup | \( V_k \) protein group |
|----------------------|---------------------|---------------------|----------------------|
| 21                   | 21                  | 21                  | III                  |
| 23                   | 23                  | 23                  | V                    |
| 4/5                  | 4                   | 4                   | IV, VI               |
| 12/13                | 12, 13              | 12-13               | V                    |
| RF ambiguous assignment | V                  | V                   |                      |
| 11                   | 11                  | 11                  | V                    |
| 9A                   | 9                   | 9                   | V                    |
| 9B                   | 9                   | 9                   | V                    |
| 10                   | 10                  | 10                  | V                    |
| 38C ambiguous assignment | V                  | V                   |                      |
| 24/25                | 24                  | 25                  | II, I                |
| 1                    | 1, 3, 26            | 1                   | II                   |
| 2                    | 2                   | 2                   | II                   |
| 8                    | 8                   | 8                   | I                    |
| 22                   | 22                  | 22                  | I                    |
| 19/28                | 14, 15, 19          | 19                  | V                    |

* Relatedness between \( V_k \) gene families and \( V_k \) protein subgroups 20 and 27, and \( V_{kCys} \) subgroups 6, 7, 16, 17, and 18 (for which only partial protein sequences are known), could not be determined.

(\text{Glu}^{60} \text{Ala}^{30} \text{Tyr}^{10}) \text{ system} (Ollier et al. 1985). In RFLP analyses, two strongly and several weakly hybridizing restriction fragments were observed (Kofler et al. 1989). Whether the latter corresponded to additional, more distant, \( V_{k12/13} \) germline genes or are due to high similarity (>80%) in portions of the probe with other \( V_k \) genes (particularly those of \( V_k \) gene families 9A, 9B, 10, and III) remains to be determined.

\( V_{kRF} \) gene family. The MRL-RF24 \( V_k \) protein (Kofler et al. 1987b), a member of the large protein group V, had 12 mismatches up to Trp35 from two \( V_{k12-13} \) proteins (K2 and MOPC129), but differed from the remaining \( V_{k12-13} \) proteins (and all other \( V_k \) genes) by >12 residues. Thus, this protein could not be unambiguously assigned to known \( V_k \) subgroups. Its nucleic acid sequence differed from all \( V_k \) sequences by >25%, thus forming a distinct \( V_k \) gene family, termed \( V_{kRF} \). Used as a probe, this gene identified a single restriction fragment that was absent in haplotype \( Igk^l \) (Kofler et al. 1989). The corresponding (as yet uncloned) germline gene probably also encoded a \( \text{Igk}^l \) (Kofler et al. 1989).

\( V_{k9A} \) gene family. The \( V_{k9A} \) protein subgroup, another member of the large protein group V (Potter et al. 1982), comprised sequences that, at the nucleic acid level, fell into two distinct gene families, termed \( V_{k9A} \) and \( V_{k9B} \). The \( V_{k9A} \) gene family included two germline genes (Seidman et al. 1979, Max et al. 1980), one of which may be expressed in hen egg lysozyme antibodies (Darsley and Rees 1985). Another expressed \( V_{k9A} \) gene from an NZB x NZW F1 anti-DNA IgM (Kofler et al. 1988) was only 88% similar to the other germline gene and probably derived from an unknown \( V_{k9A} \) germline gene. In addition, \( V_{k9A} \) genes have been observed in GAT- idioype-specific antibodies (Ollier et al. 1985).

\( V_{k9B} \) gene family. The T1 sequence and its germline counterpart, V-L6 (Pech et al. 1981), both assigned to the \( V_{k9} \) protein subgroup (Potter et al. 1982), differed from \( V_{k9A} \) (and all other \( V_k \) nucleic acid sequences by >20% and, hence, formed a separate family, termed \( V_{k9B} \).\n
\( V_{k10} \) gene family. This family corresponded to the \( V_{k10} \) subgroup (protein group V). RFLP data suggested two to three \( V_{k10} \) germline genes (Kofler et al. 1989), one of which has been cloned (Sanz and Capra 1987, Wysocki et al. 1987) and probably encoded arsonate (Manser et al. 1987a, Meek et al. 1987), oxazolone (Berek et al. 1985), oligosaccharide (Matsuda and Kabat 1989), bromelain-treated red blood cell autoantibodies from lupus and normal mice (Reininger et al. 1987).

\( V_{k11}, 9A, 9B, 10, \) and 38C gene families. The \( V_k \) gene families discussed thus far were clearly separated from all other \( V_k \) genes by >25% overall sequence dissimilarity and in this respect resembled \( V_{kII} \) gene families. The following five gene families, distantly related to \( V_{k12/13} \) and \( V_{kRF} \), were less well separated from one another.

\( V_{k11} \) gene family. For this gene family with four to six germline genes by RFLP analysis (Kofler et al. 1989), a single nucleic acid sequence corresponding to a nonfunctional rearrangement from an NZB myeloma (Kelley et al. 1985) was present in the data bank. This sequence fulfilled protein assignment criteria for \( V_k \) protein subgroups 9, 10, and 11; however, it best matched \( V_{k11} \) proteins. Comparisons with the entire data bank (including some \( V_{k9} \) and \( V_{k10} \) sequences) revealed matches of only 76% or less at the nucleic acid level, making this sequence the prototype for the \( V_{k11} \) gene family. \( V_{k11} \) proteins were observed in the beta 2,1 fructosan response (Kabat et al. 1987).

\( V_{k9A} \) gene family. The \( V_{k9A} \) protein subgroup, another member of the large protein group V (Potter et al. 1982), comprised sequences that, at the nucleic acid level, fell into two distinct gene families, termed \( V_{k9A} \) and \( V_{k9B} \). The \( V_{k9A} \) gene family included two germline genes (Seidman et al. 1979, Max et al. 1980), one of which may be expressed in hen egg lysozyme antibodies (Darsley and Rees 1985). Another expressed \( V_{k9A} \) gene from an NZB x NZW F1 anti-DNA IgM (Kofler et al. 1988) was only 88% similar to the other germline gene and probably derived from an unknown \( V_{k9A} \) germline gene. In addition, \( V_{k9A} \) genes have been observed in GAT- idioype-specific antibodies (Ollier et al. 1985).

\( V_{k9B} \) gene family. The T1 sequence and its germline counterpart, V-L6 (Pech et al. 1981), both assigned to the \( V_{k9} \) protein subgroup (Potter et al. 1982), differed from \( V_{k9A} \) (and all other \( V_k \) nucleic acid sequences by >20% and, hence, formed a separate family, termed \( V_{k9B} \).\n
\( V_{k10} \) gene family. This family corresponded to the \( V_{k10} \) subgroup (protein group V). RFLP data suggested two to three \( V_{k10} \) germline genes (Kofler et al. 1989), one of which has been cloned (Sanz and Capra 1987, Wysocki et al. 1987) and probably encoded arsonate (Manser et al. 1987a, Meek et al. 1987), oxazolone (Berek et al. 1985), oligosaccharide (Matsuda and Kabat 1989), bromelain-treated red blood cell autoantibodies from lupus and normal mice (Reininger et al. 1987).
NZB myeloma (Kelley et al. 1985), and might correspond to one of the uncloned V\textsubscript{10} germine genes.

\textit{V\textsubscript{38C} gene family (tentative).} The very similar (97\%) sequences encoding the 38C13 lymphoma (Campbell 1987) and the VM113 anti-hemagglutinin hybridoma (Meek et al. 1989) light chains, respectively, were >20\% different from all other V\textsubscript{k} nucleic acid sequences in the database and, hence, could not be assigned to any V\textsubscript{k} gene family; the closest matches (77–78\%) were observed with a V\textsubscript{10} germine gene (Sanz and Capra 1987, Wysocki et al. 1987). At the amino acid level, members of four V\textsubscript{k}Trp subgroups (V\textsubscript{k}9, V\textsubscript{k}10, V\textsubscript{k}11, and V\textsubscript{k}12/13) exhibited equally distant relatedness (nine and more residues difference in the N-terminal 35 amino acids), making unambiguous assignment at the protein level impossible. Whether these sequences were the representatives of a new V\textsubscript{k} gene family or corresponded to highly mutated (V\textsubscript{10}) genes remains to be determined.

V\textsubscript{24/25}, V\textsubscript{I}, and V\textsubscript{2} gene families. The next three families were grouped together based on sequence similarity of up to 78\% between V\textsubscript{24/25} members and V\textsubscript{I} and V\textsubscript{2} genes, respectively, and because the overall similarity between V\textsubscript{2} and some V\textsubscript{I} genes exceeded 80\%. The latter observation, i.e., similarity of >80\% between some, but not all, members of two gene families, obviously constitutes a problem in this type of V\textsubscript{k} gene classification (see below).

V\textsubscript{24/25} gene family. Originally, only a single V\textsubscript{24} germine gene (involved in the phosphocholine response; Malipiero et al. 1987, Gearhart and Bogenhagen 1983) had been reported (Selsing and Storb 1981). Other investigators have cloned this, a related pseudogene, and two additional V\textsubscript{24} germine genes (John et al. 1984). The latter were only about 82–83\% similar to the V\textsubscript{24} prototype and may have encoded 	extit{Streptococcus} group A carbohydrate antibody light chains previously assigned to the V\textsubscript{25} subgroup (Lutz and Davie 1988). Hence, these two V\textsubscript{k}Trp subgroups (protein group II) were probably encoded by distant members of a single V\textsubscript{k} gene family. In addition to the four cloned V\textsubscript{24/25} germine genes, evidence was obtained for the presence of at least two more germine genes in this family: firstly, RFs from autoimmune and normal mice (Shlomchik et al. 1987a, 1987c) expressed V\textsubscript{24} genes very similar to each other, but >30 bp different from the closest V\textsubscript{24} germine gene, suggesting an additional germine gene; secondly, since all cloned V\textsubscript{24/25} genes had 40 codons up to Trp35, the germine gene encoding Hy2.5.13 with 41 N-terminal amino acids (Kabat et al. 1987) has yet to be isolated.

V\textsubscript{I} and V\textsubscript{2} gene families. Protein subgroups V\textsubscript{I} (already previously condensed with Cys23 subgroups V\textsubscript{3} and V\textsubscript{26}; Potter et al. 1982) and V\textsubscript{2} were encoded by sequences that, using a stringent family definition, precluded classification into either a single, or two distinct, gene families; all V\textsubscript{I} nucleic acid sequences were >80\% similar, yet the three almost identical V\textsubscript{2} nucleic acid sequences reported (Akolkar et al. 1987, Kofler et al. 1988, Panka et al. 1988) shared up to 81.7\% similarity with some, but only about 75\% with other, V\textsubscript{I} members. Moreover, sequence similarity in the 3’ portion of several V\textsubscript{I} and V\textsubscript{2} genes was around 90\%. These two “gene families” were, therefore, partially overlapping. However, for reasons of clarity, we have retained them as separate V\textsubscript{k} gene families.

Three V\textsubscript{I} germine genes (Corbet et al. 1987) and approximately 40 expressed V\textsubscript{I} sequences have been reported. With the exception of an anti-dextran V\textsubscript{k} gene (W3129; Borden and Kabat 1987) with >15\% differences from any known V\textsubscript{I} gene, all expressed sequences were highly homologous to one of the above germine genes, suggesting that the total V\textsubscript{I} germine gene number may not exceed four. A more direct complexity estimate in our previous RFLP analysis was hampered by cross-hybridization of the V\textsubscript{I} probe to non-V\textsubscript{I} genes due to >80\% sequence similarity in the 3’ region of V\textsubscript{I} and other V\textsubscript{k} genes (see below and Kofler et al. 1989). V\textsubscript{I} genes were used in a variety of responses to foreign and self antigens (reviewed by Schiff et al. 1988, Kofler et al. 1987a). V\textsubscript{2} germine genes have not yet been reported; the three expressed sequences encoded antibodies to dextran (Akolkar et al. 1987), digoxin (Panka et al. 1988), and DNA (Kofler et al. 1988).

V\textsubscript{8}, V\textsubscript{22}, and V\textsubscript{19/28} gene families. The following three gene families were separated from each other by >20\%, and from all other V\textsubscript{k} genes by >25\%, overall sequence similarity; however, large portions (codons 35 to 94) of their genes had between 80\% and 89\% common nucleotides, leading to extensive cross-hybridizations (Kofler et al. 1989).

V\textsubscript{8} gene family. All sequences encoding V\textsubscript{k}Trp subgroup V\textsubscript{8} (protein group I) were around 90\% similar and shared up to 78\% of their nucleotides with V\textsubscript{19/28} and V\textsubscript{22} genes. Similarity in codons 35–94 was even higher, reaching 87\% with V\textsubscript{28} genes. The complexity of this gene family was difficult to assess by RFLP analyses due to possible cross-hybridization, however, at least half of the 13–20 fragments hybridizing to a V\textsubscript{8} probe probably belonged to this large family (Kofler et al. 1989). V\textsubscript{8} genes encoded antibodies to phosphocholine (Malipiero et al. 1987), dinitrophenyl (Riley et al. 1986), and hen egg lysozyme (Darsley and Rees 1985), as well as RF-like (Shlomchik et al. 1987a, 1987c) and DNA-specific (Eilat et al. 1988) autoantibodies.
$V_{22}$ gene family. The only two, almost identical, $V_{22}$ (protein group I) sequences available for comparison, S107A (Kwan et al. 1981) and HPCA97 (Berek 1984), revealed between 80% and 89% similarity with a large portion (codons 35 to 94) of all $V_{19/28}$ genes. The remaining nucleotides were, however, only <70% similar, resulting in an overall similarity of 72%-75%, thus refuting assignment of $V_{22}$ and $V_{19/28}$ genes to a common gene family. Similarity with ~8 genes was in the range of 75%-77% and mismatches were distributed evenly over the entire gene. RFLP analyses suggested one to two $V_{22}$ germline genes; additional weak restriction fragments hybridizing to a $V_{22}$ probe on Southern blots probably corresponded to genes from the $V_{19/28}$ and $V_{8}$ families (Kofler et al. 1989). $V_{22}$ genes encoded phosphocholine antibodies (Malipiero et al. 1987).

$V_{19/28}$ gene family. Sequences encoding $V_{1}$Trp subgroups 19 (comprising $V_{Cys14}$ and 15 sequences) and 28 were >80% similar among each other and differed from all other $V_{1}$ genes (except $V_{8}$ and $V_{22}$, see above) by >25%. Thus, they were combined to a single $V_{1}$ gene family, which was termed $V_{19/28}$. However, this $V_{1}$ gene family (like some other $V_{1}$ gene families, see below) behaved atypically in nucleic acid hybridization studies as compared to $V_{H}$ gene families: different DNA probes from this family, i.e., a $V_{19}$ and a $V_{28}$ probe, did not hybridize to an identical, but to an overlapping, set of restriction fragments (Kofler et al. 1989). This could be explained by cross-hybridization of the $V_{28}$, but not the $V_{19}$, probe with $V_{8}$ genes.

RFLP data suggested four to six $V_{19/28}$ germline genes (Kofler et al. 1989), one of which, a $V_{28}$ germline gene, also known as $V_{Ser}$, from haplotypes Igk-$V_{Ser}^a$, Igk-$V_{Ser}^b$, Igk-$V_{Ser}^c$, and Igk-$V_{Ser}^d$, has been cloned (Boyd et al. 1986, Ponath et al. 1989). $V_{19/28}$ genes encoded antibodies to trinitrophenyl (Hawley et al. 1982), carcinoembryonic antigen (Cabilly et al. 1984, Beidler et al. 1988), human breast/lung/colon cancer cells (Sahagan 1986), influenza hemagglutinin (Meek et al. 1989), and an RNA-specific (Eilat et al. 1988) and some RF-like autoantibodies (Kofler et al. 1989, Shlomchik et al. 1987a, 1987c).

Relatedness between $V_{1}$ gene families and implications for nucleic acid hybridization assays with $V_{1}$ probes

Figure 3 shows the relatedness between different $V_{1}$ gene families as reflected by overall nucleic acid sequence similarity. A significant difference from $V_{H}$ gene families was apparent, since the latter are generally more distantly...
related by sequence similarity. Obviously, if members from different families are only a few percent less similar than those from within a family, cross-hybridizations might occur, particularly if these differences are not evenly distributed over the entire sequence. As described above, large sequence portions with high degrees of similarity were indeed observed in genes from families 8, 22, and 19/28, and thus explain the previously observed cross-hybridizations between those families. Closer scrutiny of the similarities between portions of Vk sequences from different families revealed that the 3' region (particularly codons 57–88, corresponding to frame work region 3) were generally more closely related than the remaining sequence, and this portion might precipitate unexpected cross-hybridizations, even between otherwise distant Vk gene families. For example, Vk10 and Vk9A genes had a 135 bp 3' sequence with 83% similarity, and VkRF and Vk9B genes shared 84% of 103 nucleotides at the 3' end. As a further complication, different genes from a given family may exhibit more or less cross-hybridizations with genes from other families.

Because of the differences in the organization of Vh and Vk genes, nucleic acid sequence hybridization assays with Vk DNA probes require particular care in the selection of probes and in data interpretation. While in general any member of a Vh gene family used as a probe will recognize its entire family, but will not cross-hybridize with other families, our previous RFLP analyses and the current study strongly suggest that Vh probes may often behave differently. As a rule, probes devoid of the more 'promiscuous' 3' sequences will be more specific; however, such probes may not always hybridize to all members of their gene families, and therefore require the use of two or more genes to probe the entire family.

**Vk germline gene complexity**

Another question addressed in this study regards the total number of Vk genes in the genome of inbred mice. We estimated the complexity of known Vk gene families by using RFLP criteria (Kofler et al. 1989) and by taking into account expressed and germline genes identified for each family. Regarding expressed sequences, we assumed that IgM sequences with > 6, and IgG sequences with > 30 mismatches from known germline genes may have derived from as yet unknown germline genes. Allelic differences were also considered, however this was a minor concern as the majority of sequences in the database (91/109) derived from the same haplotype (Igk1).

This approach led to a total of about 70–140 genes (Table 3). Obviously, such estimates need to be taken with caution due to the peculiarities of Vk gene probes discussed above, and to inherent limitations of the RFLP technique (discussed by Kofler et al. 1989). Furthermore, possible additional, as yet uncloned, Vk genes and gene families in the mouse genome have not been included. However, although evidence for some additional Vk genes exists, their number might be limited. For two VkTrp subgroups, Vk27 (group I) and Vk20 (group VII), nucleic acid sequences have not been identified, but the corresponding Vk gene families may be small since only a single sequence for each subgroup has been reported to date. D’Hoostelaere published another novel Vk gene family (pc9-26) with approximately six members as suggested by RFLP analyses (D’Hoostelaere et al. 1988), but whether or not this family related to either of the two subgroups above, or to Vk38C, is unknown. Nevertheless, the large number of responses to foreign and self antigens investigated at the nucleic acid sequence level, and repeated isolation of identical sequences, suggest that the majority of the mouse Vk germline repertoire might now be known.

**Table 3. Vk germline gene complexity**

| Vk gene family | Germline genes |
|----------------|----------------|
| Cloned         | Estimated      |
| Vk21           | 5              | 6–13          |
| Vk23           | 1              | 2–4           |
| Vk4/5          | 14             | 25–50         |
| Vk12/13        | 2              | 2–8           |
| VkRF           | –              | 0–1           |
| Vk11           | –              | 4–6           |
| Vk9A           | 2              | 4–9           |
| Vk9B           | 1              | 2             |
| Vk10           | 1              | 2–3           |
| Vk38C          | –              | ?             |
| Vk24/25        | 4              | 6             |
| VkI            | 3              | 4–6           |
| Vk2            | –              | 1–6           |
| Vk8            | –              | 5–16          |
| Vk22           | –              | 1–2           |
| Vk19/28        | –              | 4–6           |

* References to cloned Vk germline genes are given throughout the text. The one-member VkRF family is deleted in haplotype Igk1 mice (Kofler et al. 1989).

**Acknowledgments.** We thank Dr. G. Wick for critically reviewing, and Ms. M. Kat Occhipinti for editing, this manuscript. The work reported herein was supported by a grant from the Austrian Research Council (Project S-41/06). R. Strohal is supported by a fellowship from the Emil Boral Stiftung and G. Kroemer receives an Erwin Schrödinger Stipend (J0307M) from the Austrian Research Council.

**References**

Akolkar, P. N., Sikder, S. K., Bhattacharya, S. B., Liao, J., Gruedo, F., Morrison, S. L., and Kabat, E. A.: Different Vk and Vh germline genes are used to produce similar combining sites with specificity for alpha (1–6) dectrans. J Immunol 138: 4472–4479, 1987
Altenburger, W., Steinmetz, M., and Zachau, H.G.: Functional and
non-functional joining in immunoglobulin light chain genes of a
mouse myeloma. Nature 287: 603-607, 1980

Beidler, C. B., Ludwig, J. R., Cardenas, J., Phelps, J., Papworth,
C. G., Melcher, E., Sierzecki, M., Myers, L. J., Unger, B. W.,
Fisher, M., David, G. S., and Johnson, M. J.: Cloning and high
level expression of a chimeric antibody with specificity for human
carcinomembranetic antigen. J Immunol 141: 4053-4060, 1988

Berek, C.: The D segment defines the T15 idiotope: the im-
munoresponse of A/J mice to Pneumococcus pneumoniae. Eur J
Immunol 14: 1043-1048, 1984

Berek, C. and Milstein, C.: Mutation drift and repertoire shift in
the maturation of the immune response. Immunol Rev 96: 23-41, 1987

Berek, C., Griffiths, G. M., and Milstein, C.: Molecular events during
maturation of the immune response to oxazolone. Nature 316: 412-418, 1985

Berek, C., Jarvis, J. M., and Milstein, C.: Activation of memory and
virgin B cell clones in hyperimmune animals. Eur J Immunol 17:
1121, 1987

Borden, P. and Kabat, E. A.: Nucleotide sequence of the cDNAs en-
coding the variable region heavy and light chains of a myeloma pro-
tein specific for the terminal nonreducing end of alpha (1-6) dextr.
Proc Natl Acad Sci USA 84: 2440-2443, 1987

Boyd, T. M., Goldrick, M. M., and Gottlieb, P. D.: Structural dif-
fferences in a single gene encoding the V\textsubscript{\textgamma}C\textsigma\textvarepsilon group of light chains explain the existence of two mouse light-chain genetic markers.
Proc Natl Acad Sci USA 83: 9134-9138, 1986

Brodeur, P. H. and Riblet, R.: The immunoglobulin heavy chain
variable region (Igh-V) locus in the mouse. I. One hundred Igh-V
genes comprise seven families of homologous genes. Eur J Immunol
14: 922-930, 1984

Bruck, C., Co, M. S., Slauj, M., Gaulton, G. N., Smith, T., Fields,
B. N., Mullins, J. I., and Greene, M. I.: Nucleic acid sequence of
an internal image-bearing monoclonal anti-idiotyp and its com-
parison to the sequence of the external antigen. Proc Natl Acad Sci USA 83:
6578-6582, 1986

Buckel, P., Huebner-Parajsz, C., Mattes, R., Lenz, H., Haug, H., and
Beaucamp, K.: Cloning and nucleotide sequence of heavy- and
light-chain cDNAs from creatine-kinase-specific monoclonal antibody.
Gene 51: 13-19, 1987

Cabilly, S., Rigs, S. A. D., Pande, H., Shively, J. E., Holmes, W. E.,
Rov, M., Perry, L. J., Wetzell, R., and Heyneker, H. L.: Generation
of antibody activity from immunoglobulin polypeptide chains
produced in Escherichia coli. Proc Natl Acad Sci USA 81: 3273-3277, 1984

Campbell, M. J.: Idiotype vaccination against murine B cell lymphoma.
Humoral and cellular responses elicited by tumor-derived immu-
noglobulin M and its molecular subunits. J Immunol 139:
2825-2833, 1987

Clarke, S. H., Huppi, K., Ruezinsky, D., Staudt, L., Gerhard, W., and
Weigert, M.: Inter- and intraclonal diversity in the antibody
response to influenza hemagglutinin. J Exp Med 161: 687-704,
1985

Corbet, S., Milili, M., Fougereau, M., and Schiff, C.: Two V\textsubscript{\textgamma} germ-
line genes related to the GAT idiotype network (AbI and Ab3/AbI') account for the major subfamilies of the mouse V\textsubscript{\textgamma} variability subgroup. J Immunol 138: 932-939, 1987

Cory, S., Tyler, B. M., and Adams, J. M.: Sets of immunoglobulin V\textsubscript{\textgamma} genes homologous to ten cloned V\textsubscript{\textgamma} sequences: implications for the number of germline V\textsubscript{\textgamma} genes. J Mol Appl Genet 1: 103-116, 1981

Darsley, M. J. and Rees, A. R.: Nucleotide sequence of five anti-
lysozyme monoclonal antibodies. EMBO J 4: 393-398, 1985

Devaux, C., Moinier, D., Mazza, G., Guo, X., Marchetto, S.,
Fougereau, M., and Rievers, M.: Preferential expression of V\textsubscript{\textgamma}21E
on Id\textalpha 7 positive monoclonal anti-I-E antibodies. J Immunol
34: 4024-4030, 1985

D’Hoostelaere, L. A., Huppi, K., Mock, B., Mallett, C., and Potter,
M.: The Igk L chain allelic groups among the Igk haplotypes and
Igk crossover populations suggest a gene order. J Immunol 141:
652-661, 1988

Dildrop, R.: A new classification of mouse V\textsubscript{\textgamma} sequences. Immunol To-
day 5: 85-86, 1984

Elait, D., Webster, D. M., and Rees, A. R.: V region sequences of anti-
DNA and anti-RNA autoantibodies from NZB/NZW F1 mice. J
Immunol 141: 1745-1753, 1988

Even, J., Griffiths, G. M., Berek, C., and Milstein, C.: Light chain
germ-line genes and the immune response to 2-phenyloxazolone.
EMBO J 4: 3439-3445, 1985

Gearhart, P. J. and Bogenhagen, D. F.: Clusters of point mutations are
found exclusively around rearranged antibody variable genes. Proc
Natl Acad Sci USA 80: 3439-3443, 1987

Griffiths, G. M., Berek, C., Kaartinen, M., and Milstein, C.: Somatic
mutation and the maturation of immune response to
2-phenyloxazolone. Nature 312: 271-275, 1984

Hawley, R. G., Shulman, M. J., Murialdo, H., Gibson, D. M., and
Hozumi, N.: Mutant immunoglobulin genes have repetitive DNA
elements inserted into their intervening sequences. Proc Natl Acad
Sci USA 79: 7425-7429, 1982

Heinrich, G., Traunecker, A., and Teneqwa, S.: Somatic mutation
creates diversity in the major group of mouse immunoglobulin
kappa light chains. J Exp Med 159: 417-435, 1984

Heller, M., Owens, J. D., Mushinski, J. F., and Rudikoff, S.: Amino
acids at the site of V\textsubscript{\textgamma}-J\textgamma recombination not encoded by germline
sequences. J Exp Med 166: 637-646, 1987

Höcht, J., Müller, C. R., and Zaczau, H. G.: Recombined flanks of
the variable and joining segments of immunoglobulin genes.
Proc Natl Acad Sci USA 79: 1383-1387, 1982

John, R., Gershenfeld, H., and Weissman, I. L.: Evolution of a
multigene family of V\textsubscript{\textgamma} germ line genes. EMBO J 3: 185-191,
1984

Kaartinen, M. and Macekáelæ, O.: Functional analogues of the V\textsubscript{\textgamma}21 gene in different strains of mice: evolutionary conservation but
diversity based on V-J joining. J Immunol 138: 1613-1617, 1987

Kaartinen, M., Griffiths, G. M., Markham, A., and Milstein, C.: mRNA sequences define an unusually restricted IgG response to
2-phenyloxazolone and its early diversification. Nature 304:
320-324, 1983

Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M., and Got-
tesman, K. S.: Sequences of Proteins of Immunological Interest, 4th
dition, US Department of Health and Human Services, Public Health
Service, National Institutes of Health, Bethesda, 1987

Kearney, J. F., Radbruch, A., Liesegang, B., and Rajewsky, K.: A
mouse myeloma cell line that has lost immunoglobulin expression
but permits the construction of antibody-searching hybrid cell lines.
J Immunol 123: 1548-1550, 1979

Kelley, D. E., Wiedermann, L. M., Pittet, A. C., Strauss, S., Nelson,
K. J., Davis, J., Van Ness, B., and Perry, R. P.: Nonproductive
kappa immunoglobulin genes: recombinational abnormalities and
other lesions affecting transcription, RNA processing, turnover,
and translation. Mol Cell Biol 5: 1660-1675, 1985

Koller, R.: A new murine Ig V\textsubscript{\textgamma} gene family. J Immunol 140:
4031-4034, 1988

Koller, R., Dixon, F. J., and Theofilopoulos, A. N.: The genetic origin of
autoantibodies. Immunol Today 8: 374-380, 1987a

Koller, R., Noonan, D. J., Strohal, R., Balderas, R. S., Moller,
N. P. H., Dixon, F. J., and Theofilopoulos, A. N.: Molecular
analysis of the murine lupus-associated anti-self response: involve-
ment of a large number of heavy and light chain variable region
genes. Eur J Immunol 17: 91-95, 1987
Koffler, R., Strohal, R., Balderas, R. S., Johnson, M. E., Noonan, D. J., Lutz, C. T., and Davie, J. M.: Genetics and primary structure of V_k gene families. *Immunogenetics* 29: 65–74, 1989

Kwan, S.-P., Rudikoff, S., Seidman, J. G., Leder, P., and Scharff, Kofler, R., Duchosal, M. A., and Dixon, F. J.: Complexity, polymorphism and connectivity of mouse V_k gene families. *Immunogenetics* 29: 65–74, 1989

Kwan, S.-P., Rudikoff, S., Seidman, J. G., Leder, P., and Scharff, M. D.: Nucleic acid and protein sequences of phosphocholine-binding light chains. *J Exp Med* 153: 1366–1370, 1981

Liu, A. Y.: Production of a mouse-human chimeric monoclonal antibody to CD20 with potent Fc-dependent biologic activity. *J Immunol* 139: 3521, 1987

Liu, A. Y., Robinson, R. R., Hellström, K. E., Murray, E. D., Jr., Chang, C. P., and Hellström, I.: Chimeric mouse-human IgG1 antibody that can mediate lysis of cancer cells. *Proc Natl Acad Sci USA* 84: 3439–3443, 1987

Livant, D., Blatt, C., and Hood, L.: One heavy chain variable region gene segment encoding antibody to CD20 with potent Fc-dependent biologic activity. *J Immunol* 139: 3521, 1987

Livant, D., Blatt, C., and Hood, L.: One heavy chain variable region gene segment subfamily in the BALB/c mouse contains 500–1000 or more members. *Cell* 47: 461–470, 1986

Lutz, C. T. and Davie, J. M.: Genes and primary structure of V_k gene segments encoding antibody to Streptococcal group A carbohydrate. Comparison of V_k gene structure with idiotope expression. *J Immunol* 140: 641–645, 1988

Malpiiero, U. V., Levy, N. S., and Gearhart, P. J.: Somatic mutation in anti-phosphorylcholine antibodies. *Immunol Rev* 96: 59–74, 1987

Manser, T., Parhami-Seren, B., Margolies, M. N., and Gefter, M. L.: Somatically mutated forms of a major anti-p-azophenylarsonate antibody variable region with drastically reduced affinity for p-azophenylarsonate. By-products of an antigen-drive immune response? *J Exp Med* 166: 1456–1463, 1987a

Manser, T., Wysocki, L. J., Margolies, M. N., and Gefter, M. L.: Evolution of antibody variable region structure during the immune response. *Immunol Rev* 96: 141–162, 1987b

Matsuda, T. and Kabat, E. A.: Variable region cDNA sequences and antigen binding specificity of mouse monoclonal antibodies to isomaltosyl oligosaccharides coupled to proteins. T-dependent analogous of α(1 → 6)dexturan. *J Immunol* 142: 863–870, 1989

Max, E. E., Seidman, J. G., Miller, H., and Leder, P.: Variation in the crossover point of kappa immunoglobulin gene V-J recombination: evidence from a cryptic gene. *Cell* 21: 793–799, 1980

Meek, K., Sanz, I., Rathsam, G., Nisonoff, A., and Capra, J. D.: Identities of the V_{10-Ars-A} gene segments of the A/J and BALB/c strains. *Proc Natl Acad Sci USA* 84: 6244–6248, 1987

Meek, K., Johansson, B., Schulman, J., Bona, C., and Capra, J. D.: Nucleotide changes in sequential variants of influenza virus hemagglutinin genes and molecular structures of corresponding monoclonal antibodies specific for each variant. *Proc Natl Acad Sci USA* 86: 4664–4668, 1989

Nahmias, C., Strosberg, A. D., and Emorine, L. I.: The immune response toward beta-adrenergic ligands and their receptors. VIII. Extensive diversity of V_{32} and V_{21} gene encoding anti-alprenolol antibodies. *J Immunol* 140: 1304–1311, 1988

Near, R. I. and Haber, E.: Characterization of the heavy and light chain immunoglobulin variable region genes used in a set of anti-digoxin antibodies. *Mol Immunol* 26: 371–382, 1989

Nishihara, Y. and Leder, P.: Organization and complete sequence of identical embryonic and plasmacytoma kappa V-region genes. *J Biol Chem* 255: 3691–3694, 1980

Ollier, P., Rocca-Serra, J., Somme, G., Theze, J., and Fougeroue, M.: The idiotypic network and the internal image: possible regulation of a germ-line network by paucigene encoded Ab2 (anti-idiotypic) antibodies in the GST system. *EMBO J* 4: 3681–3688, 1985

Panka, D. J. and Margolies, M. N.: Complete variable region sequences of five homologous high affinity anti-digoxin antibodies. *J Immunol* 139: 2385–2391, 1987

Panka, D. J., Mudgett-Hunter, M., Parks, D. R., Peterson, L. L., Herzenberg, L. A., Haber, E., and Margolies, M. N.: Variable region framework differences result in decreased or increased affinity of variant anti-digoxin antibodies. *Proc Natl Acad Sci USA* 85: 3080–3084, 1988

Parslow, T. G., Blair, D. L., Murphy, W. J., and Granner, D. K.: Structure of the 5’ ends of immunoglobulin genes: a novel conserved sequence. *Proc Natl Acad Sci USA* 81: 2650–2654, 1984

Pech, M., Höchtl, J., Schnell, H., and Zachau, H. G.: Differences between germ-line and rearranged immunoglobulin V_k coding sequences suggest a localized mutation mechanism. *Nature* 291: 668–670, 1981

Pennelli, C. A., Arnold, L. W., Haughton, G., and Clarke, S. H.: Restricted Ig variable region gene expression among Ly-1+ B cell lymphomas. *J Immunol* 141: 2788–2796, 1988

Perlmutter, R. M., Kearney, J. F., Chang, S. P., and Hood, L. E.: Developmentally controlled expression of immunoglobulin V_k genes. *Science* 227: 1597–1601, 1985

Porath, P. D., Hillis, D. M., and Gottlieb, P. D.: Structural and evolutionary comparisons of four alleles of the mouse immunoglobulin kappa chain gene, Igk-Vser. *Immunogenetics* 29: 249–257, 1989

Potter, M.: Antigen-binding myeloma proteins of mice. *Adv Immunol* 25: 141–211, 1977

Potter, M., Newell, J. B., Rudikoff, S., and Haber, E.: Classification of mouse V_k groups based on the partial amino acid sequence to the first invariant tryptophan: impact of 14 new sequences from IgG myeloma proteins. *Mol Immunol* 19: 1619–1630, 1982

Reininger, L., Offiler, P., Poncet, P., Kaushik, A., and Jaton, J. C.: Novel V_k genes encode virtually identical variable regions of six murine monoclonal anti-bromelin-treated red blood cell autoantibodies. *J Immunol* 138: 316–323, 1987

Reininger, L., Kaushik, A., Izui, S., and Jaton, J.-C.: A member of a new V_k gene family encodes anti-bromelinized mouse red blood cell autoantibodies. *Eur J Immunol* 18: 1521–1526, 1988

Riley, S. C., Connors, S. J., Klinman, N. R., and Ogata, R. T.: Preferential expression of variable region heavy chain gene segments by predominant 2,4-dinitrophenyl-specific BALB/c neonatal antibody clonotypes. *Proc Natl Acad Sci USA* 83: 2589–2593, 1986

Salitirzky, F. and Rajewsky, K.: Molecular basis of an isogenic anti-idiotypic response. *EMBO J* 3: 3005–3012, 1984

Sahagan, B. G.: A genetically engineered murine/human chimeric antibody retains specificity for human tumor-associated antigen. *J Immunol* 137: 1066, 1986

Sanz, I. and Capra, J. D.: V_k and J_k gene segments of A/J Ars-A antibodies: somatic recombination generates the essential arginine at the junction of the variable and joining regions. *Proc Natl Acad Sci USA* 84: 1085–1089, 1987

Schiff, C., Corbet, S., Millili, M., and Fougeroue, M.: Interstrain conservation of the murine GAT-specific antibody V_{repertoire} as analyzed at the germline gene level. *EMBO J* 2: 1771–1776, 1983

Schiff, C., Corbet, S., and Fougeroue, M.: The Ig germline gene repertoire: economy or wastage? *Immunol Today* 9: 10–14, 1988

Seidman, J. G., Leder, A., Edgell, M. H., Polsky, F., Tilghman, S. M., Tienmeier, D. C., and Leder, P.: Multiple related immunoglobulin variable-region genes identified by cloning and sequence analysis. *Proc Natl Acad Sci USA* 75: 3881–3885, 1978

Seidman, J. G., Max, E. E., and Leder, P.: A kappa-immunoglobulin gene is formed by site-specific recombination without further somatic mutation. *Nature* 280: 370–375, 1979

Selinger, E. and Storb, U.: Somatic mutation of immunoglobulin light-chain variable-region genes. *Cell* 25: 47–58, 1981

Shlomchik, M., Nemazee, D., Van Snick, J., and Weigert, M.: Variable region sequences of murine IgM anti-IgG monoclonal autoan-
tibodies (rheumatoid factors). II. Comparison of hybridomas der-
ived by lipopolysaccharide stimulation and secondary protein
immunization. J Exp Med 165: 970–987, 1987a
Shlomchik, M. J., Aucoin, A. H., Pisetsky, D. S., and Weigert, M. G.: Structure and function of anti-DNA autoantibodies derived from a single autoimmune mouse. Proc Natl Acad Sci USA 84: 9150–9154, 1987b
Shlomchik, M. J., Marshak-Rothstein, A., Wolfowicz, C. B., Roth-
stein, T. L., and Weigert, M. G.: The role of clonal selection and
somatic mutation in autoimmunity. Nature 328: 805–811, 1987c
Sikder, S. K., Akolkar, P. N., Kaladas, P. M., Morrison, S. L., and
Kabat, E. A.: Sequences of variable regions of hybridoma an-
tibodies to alpha (1−6) dextran in BALB/c and C57BL/6 mice.
J Immunol 135: 4215–4221, 1985
Strohal, R., Kroemer, G., Wick, G., and Kohler, R.: Complete variable
region sequence of a non-functionally rearranged kappa light chain
transcribed in the non-secretor P3-X63-Ag8.653 myeloma cell line.
Nucleic Acids Res 15: 2771, 1987

Tonegawa, S.: Somatic generation of antibody diversity. Nature 302:
575–581, 1983
Winter, E., Radbruch, A., and Krawinkel, U.: Members of novel V\textsubscript{H}
gene families are found in VDJ regions of polyclonally activated
B-lymphocytes. EMBO J 4: 2861–2867, 1985
Wysocki, L. J., Gridley, T., Huang, S., Grandeu, A. G., and Gefter,
M. L.: Single germline V\textsubscript{H} and V\textsubscript{L} genes encode predominating an-
tibody variable regions elicited in strain A mice by immunization
with p-azophenylarsonate. J Exp Med 166: 1–11, 1987
Yancopoulos, G. D., Desiderio, S. V., Paskind, M., Kearney, J. P.,
Baltimore, D., and Alt, F. W.: Preferential utilization of the most
J\textsubscript{H}-proximal V\textsubscript{H} gene segments in pre-B-cell lines. Nature 311:
727–733, 1984

Received July 31, 1989; revised version received September 26, 1989.