

## Myeloma proteins from NZB and BALB/c mice: Structural and functional differences

(protein sequence/genealogical analysis/NZB myeloma immunoglobulins)

E. LOH\*†, B. BLACK\*, R. RIBLET‡, M. WEIGERT‡, J. M. HOOD\*§, AND L. HOOD\*

\*Division of Biology, California Institute of Technology, Pasadena, California 91125; and †Institute of Cancer Research, Philadelphia, Pennsylvania 19111

Communicated by Ray D. Owen, September 25, 1978

**ABSTRACT** Structural and functional analyses of myeloma immunoglobulins from inbred BALB/c mice and humans have provided important insights into the structure of the antibody molecule and the expression and evolution of antibody genes. One important question concerning these analyses is whether the myeloma process selects, in a nonrandom manner, the lymphocytes to be transformed. The availability of myeloma tumors in a second inbred strain of mouse, NZB, permits us to approach this question. In this respect, the NH<sub>2</sub>-terminal amino acid sequences of 27  $\kappa$  light chains as well as data relating to the antigen-binding properties and immunoglobulin class distribution of NZB myeloma proteins are presented and compared with similar data from the BALB/c mouse. These studies suggest that the myeloma proteins from the BALB/c and NZB mice constitute two populations of immunoglobulins with distinct functional and structural properties. The implications of this observation are discussed.

Myeloma tumors may be artificially induced by injecting mineral oil into the peritoneal cavity of inbred BALB/c mice (1). Chemical analysis of the homogeneous immunoglobulins derived from these myeloma tumors has provided important insights into the structure, genetics, and evolution of antibody molecules and genes (2, 3). Moreover, the observations that immunoglobulin variable regions can be divided into related sets or *subgroups* of sequences and that there are regions of extreme variability, termed "hypervariable regions," have placed important constraints on the modern theories of antibody diversity (4). Indeed, our estimates as to the complexities of the various immunoglobulin gene families have been derived from the nature and extent of the heterogeneity seen in the corresponding myeloma proteins. Thus, it is important to determine whether the highly artificial process of myeloma induction provides a random sampling of lymphocyte diversity in the mouse. For example, if only 10% of the lymphocyte population were capable of being transformed by the myeloma process, then diversity estimates would be too low by a factor of 10. The availability of a second inbred mouse strain, NZB, in which myeloma tumors can be induced gave us an opportunity to examine this question.

It should be noted that three observations suggest that BALB/c myeloma proteins may represent only a selected subset of the antibody variable (V) regions. First, approximately 5-10% of the BALB/c myeloma proteins bind a restricted spectrum of simple haptens, including dinitrophenol, phosphorylcholine, and various simple carbohydrate moieties (5). This frequency is much greater than would be expected from the frequency of lymphocytes binding the same haptens in normal individuals (6). Accordingly, the BALB/c myeloma proteins appear to represent a restricted sample of the potential functional repertoire of the BALB/c mouse. Second, nearly all

of the BALB/c myeloma heavy chains screened thus far have unblocked NH<sub>2</sub>-terminal groups, whereas only about 20% of the serum heavy chains are unblocked (7). The presence of blocked and unblocked NH<sub>2</sub>-terminal groups is indicative of different V subgroups (8). Thus, the BALB/c myeloma proteins and serum immunoglobulins reflect different distributions of V subgroups and, presumably, different antibody specificities. Finally, when the residue alternatives at certain positions of myeloma sequences are compared to their counterparts from the normal serum pool, it is clear that some residues in the myeloma pool are not found in the normal pool (9). Conversely, residue alternatives found in the normal pool are missing from the myeloma pool. These differences again suggest that the normal and myeloma pools of sequences express somewhat different subsets of the antibody repertoire.

Selection also appears to operate at the level of myeloma constant regions. Indeed, two-thirds of BALB/c myeloma immunoglobulins are of the IgA class whereas less than 10% of the circulating immunoglobulins are of the IgA class (10). Thus, there is evidence from both V and constant regions that the BALB/c myeloma proteins do not represent a random selection of normal serum antibody molecules.

The observation that myeloma tumors can be induced in a second inbred strain of mice, NZB, permits us to compare two populations of myeloma immunoglobulins induced within a single species, the mouse (11). In this paper we report on the antigen-binding properties and immunoglobulin class distribution of NZB myeloma immunoglobulins as well as the NH<sub>2</sub>-terminal sequences from 27  $\kappa$  chains. These studies suggest that the myeloma proteins from the BALB/c and NZB mice constitute two populations of immunoglobulins with distinct functional and structural properties.

### MATERIALS AND METHODS

**Myeloma Tumors.** The plasmacytomas used for these studies were induced in NZB/NIH mice.

**Immunoglobulin Class Identification and Antigen-Binding Assays.** The serological techniques for immunoglobulin class identification and the antigen-binding assays will be described elsewhere.

**Immunoglobulin Production.** Plasmacytomas were passaged subcutaneously in (NZB  $\times$  BALB/c)F<sub>1</sub> hybrid mice. These F<sub>1</sub> mice were produced at the Institute for Cancer Research (ICR) from matings of NZB/NIH and BALB/c ICR mice. The tumors secreting myeloma proteins were passaged intraperitoneally to 30-50 (NZB  $\times$  BALB/c)F<sub>1</sub> mice primed

Abbreviations: P<sub>i</sub>/NaCl, phosphate-buffered saline; V region, variable region.

† Present address: Department of Biochemistry, Stanford Medical School, Stanford, CA 94305.

§ Present address: Mathematics Department, California Polytechnic State University, San Luis Obispo, CA 93407.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

1 week prior to passage with 0.5 ml of pristane (2,6,10,14-tetramethylpentadecane; Aldrich). The ascites fluid from these mice was collected and pooled.

**Immunoglobulin Purification.** The ascites fluids obtained from tumor-bearing mice were clarified by centrifugation at  $15,000 \times g$  for 10–18 min. An equal volume of phosphate-buffered saline ( $P_i/NaCl$ ) with azide and EDTA (0.15 M  $NaCl/0.01$  M  $PO_4/1$  mM Na azide/1 mM EDTA, pH 7.4) was added and diluted 1:1 with a neutral solution of saturated ammonium sulfate. Volumes containing about 200 mg of protein were applied to Sephadex G-200 columns equilibrated  $P_i/NaCl/azide/EDTA$ .

**Reduction/Alkylation.** The peak corresponding to the myeloma protein was concentrated to 20 mg/ml and dialyzed against 0.15 M Tris-HCl, pH 8/0.15 M  $NaCl/2$  mM EDTA, pH 7.0. Then, 0.01 vol of 1 M dithiothreitol was added and the solution was stirred at 37°C for 1.5 hr. The sample was then placed in an ice bath and 0.022 vol of 1 M iodoacetamide was added. The alkylation reaction was terminated by the addition of dithiothreitol to a molarity equal to that of the iodoacetamide added. The solution was dialyzed against 8 M urea in propionic acid for 2 hr.

**Separation of Heavy and Light Chains.** Reduced and alkylated proteins were fractionated on Sephadex G-150 columns equilibrated in urea/propionic acid. Pooled light chains were dialyzed against water and lyophilized.

**Automated Sequence Analysis.** The amino acid sequence analysis of the light chains was carried out on a Beckman 890A Sequencer using a dimethylbenzylamine program as described (12–14). Briefly, aliquots of the phenylthiohydantoin derivatives of the amino acids were identified by gas chromatography and thin-layer chromatography. The amino acid derivatives were also hydrolyzed to free amino acids and analyzed on a Durrum D-500 amino acid analyzer as described (13). A high-pressure liquid chromatography system (Waters Associates, Inc., Milford, MA) using a pH 4.27 buffer system (sodium acetate/water/methanol) was used to identify the amino acid derivatives from some of the proteins. Generally, each protein was analyzed a single time. Repetitive yields ranged from 85–92%.

**Mathematical Analyses.** Fifty published  $NH_2$ -terminal sequences of  $\kappa$  chains from BALB/c mice were compared with the 27 NZB  $\kappa$  chains presented in this paper. Mathematical techniques have been developed which allow us to ask several questions of these two populations of  $\kappa$  sequences (15). First, given one set of proteins (e.g., the BALB/c  $\kappa$  chains), what is the likelihood that a second set of proteins (e.g., the NZB  $\kappa$  chains) were drawn from the same population of sequences? Second, in quantitative terms, how much diversity exists within each population of sequences? These mathematical techniques can be summarized briefly as follows.

To determine whether the BALB/c and NZB sequences come from the same pool, one calculates a diversity distance index  $[D(B,N)]$  in which  $D$  is the diversity distance,  $B$  is the BALB/c sequences, and  $N$  is the NZB sequences. The diversity distance index is a measure of the similarity of the amino acid distributions in the two sets of sequences. The significance of the distance index is estimated by creating a statistical model that uses a randomly chosen second set of 27 BALB/c sequences drawn from the complete set of BALB/c sequences. Then one calculates the most likely distance  $[\mu(B,27)]$  this second set will have from the original BALB/c set and the standard deviation  $[\delta(3,27)]$  of the distance index from  $\mu(B,27)$  if one repeats the random selection many times. Thus, to ask if BALB/c and NZB  $\kappa$  chains are drawn from the same pool, one calculates the distance index  $D(B,N)$  and compares this to  $\mu(B,27)$  and  $\delta(B,27)$ . If the difference between  $D(B,N)$  and  $\mu(B,27)$  is greater than

$2 \times \delta(B,27)$ , one can conclude that the BALB/c and NZB are drawn from different pools.

The diversity within each set of proteins can be determined by calculating the distance index for each sequence of a population to the population as a whole and averaging all of the individual distances. This gives a quantitative measure of the diversity within a population, designated the variation index of  $W(B)$  for the case of the BALB/c sequences.

**Relatedness Trees.** Amino acid sequence data can be visually compared by drawing relatedness or genealogical trees that depict a theoretical ancestry of the sequences based on the assumption of minimizing mutational events. Distances on the tree are proportional to the number of genetic events (substitutions, deletions, insertions) that are needed to account for sequence differences. Where the tree branches (nodal points), a gene duplication is represented with the sequences below the branch point sharing a common ancestor.

Relatedness trees were prepared by a variation of the method of Fitch and Margoliash (16). Data from the BALB/c  $\kappa$  chain sequences are taken from Hood *et al.* (9). Because of the large number of sequences, the proteins were assigned to one of four subpopulations according to their relatedness, and a detailed tree was constructed for each subdivision. The four trees were joined by making a master tree of selected members of each separate tree.

## RESULTS AND DISCUSSION

**Nature and Extent of V Region Diversity Are Similar in NZB and BALB/c Myeloma  $\kappa$  Light Chains.** Twenty-seven V regions from light chains ( $V_L$ ) of NZB immunoglobulins were analyzed by automatic sequence analysis (Fig. 1). Nineteen of the  $V_\kappa$  sequences are distinct from one another, differing by as many as 13 residues.

By several criteria the diversity of NZB and BALB/c myeloma  $\kappa$  chains is similar even though they appear to be drawn from distinct sets of proteins. (i) The distribution of variation in the first 23 residues is similar (Fig. 1). For example, the pools of myeloma  $\kappa$  chains for both strains show little variation at positions 5, 6, and 16, whereas positions 9–12, 15, 17, and 18 vary extensively. (ii) The degree of variation as measured by the variation indices  $W(B)$  and  $W(N)$  for 23 residues is similar for the BALB/c and NZB pools of  $\kappa$  chains (Table 1). (iii) The BALB/c and NZB  $\kappa$  chains both exhibit size variation in the first hypervariable region. For example, the NZB light chains fall into a minimum of four different size classes with 33, 34, 35, or 38 residues before the invariant tryptophan that terminates the first hypervariable region. Thus, the great diversity of the NZB and BALB/c  $\kappa$  chains in the first 23 residues extends through the first hypervariable region and, accordingly, will be reflected in the diversity of the corresponding antigen-binding sites. (iv) As has been found with BALB/c V regions, there is a correlation between the size of the first  $\kappa$  chain hypervariable region and the  $NH_2$ -terminal sequence (17, 18). Indeed, light chains that are identical in their first 23 amino acids invariably have first hypervariable regions of identical size. For example, one pair of NZB  $\kappa$  chains with 38 residues before the invariant tryptophan have the same  $NH_2$ -terminal sequence (i.e., PC3741 and PC2880). The same is true of a second pair with 35 residues (PC144 and PC2419) and a third pair with 34 residues (PC3609 and PC3612). This correlation is true of other sequences in the literature including several sets of  $\kappa$  chains from antigen-binding proteins and the set of closely related BALB/c  $\kappa$  chains including M321, M70, M653, and T124.

**Pool of NZB Myeloma Proteins Appears To Be Distinct from the Pool of BALB/c Myeloma Proteins by Several**

**Criteria.** The pools of NZB and BALB/c myeloma proteins differ from one another by three criteria: immunoglobulin class distribution, antigen-binding properties, and amino acid sequence. (i) The expression of myeloma immunoglobulin classes in these two strains is quite distinct in that the distribution of NZB myeloma population assessed to date ( $\approx 600$  proteins) is 52% in the IgG class and 21% in the IgA class, whereas the distribution of the BALB/c myeloma population is 22% in the IgG class and 46% in the IgA class (unpublished data; see ref. 4). Thus, the myeloma process amplifies different ratios of immunoglobulin classes in these two inbred strains even though the same method of induction is used. (ii) The NZB myeloma proteins tend to bind different haptens than do the BALB/c myeloma proteins (Table 2). For example, it is striking that none of the NZB myeloma proteins, of >200 screened, bind dinitrophenol or phosphorylcholine, two of the most common haptens bound by BALB/c myeloma proteins. Indeed, 12 NZB myeloma proteins from those screened bind DNA (unpublished data), whereas few BALB/c proteins bind this antigen (19). (iii) The  $V_\kappa$  regions from the NZB and BALB/c myeloma proteins appear to form distinct populations of amino acid sequences.

The third supposition is supported by three observations.

(i) The  $NH_2$ -terminal 23 residues of 17 V sequences of NZB myeloma  $\kappa$  chains have been examined (Fig. 1). Nineteen different  $V_\kappa$  sequences are found among them and only 1 of these sequences is identical to any of the 43 different BALB/c  $V_\kappa$  sequences available for comparison over this region. Identity of but a single  $V_\kappa$  sequence between NZB and BALB/c populations would be expected if the pool of possible V sequences in mice is so large that few repetitions would be seen. However, both the BALB/c and NZB pools of sequences show multiple repeats of V regions examined over their  $NH_2$ -terminal 23 residues. For example, within the NZB pool of  $V_\kappa$  regions, two identical pairs of sequences appear (e.g., PC2367 and PC2316)

Table 1. Statistical comparison of NZB and BALB/c light chains

|                                  |                       |
|----------------------------------|-----------------------|
| Diversity distance               | $D(B,N) = 4.99$       |
| Expected distance                | $D(B,27) = 3.69$      |
| Standard deviation               | $\mu(B,27) = 0.60$    |
| Probability (assuming normality) | <0.015                |
| Internal diversity, BALB/c       | $W(B) = 18.6 \pm 2.8$ |
| Internal diversity, NZB          | $W(N) = 21.5 \pm 1.2$ |

$B = 50$  BALB/c sequences;  $N = 27$  NZB sequences.

and three probably identical triplets appear (e.g., PC144, PC2419, and PC2454). In addition, the BALB/c sequences demonstrate at least 12 pairs of identical  $V_\kappa$  regions (9). Thus, if the NZB and BALB/c myelomas were drawn from the same pool, one would expect more V sequences to be shared by the NZB and BALB/c myeloma  $\kappa$  chains.

(ii) A relatedness analysis of the NZB and BALB/c  $\kappa$  sequences also suggests that they constitute, at least in part, distinct populations. Fig. 2 is a relatedness tree which graphically displays the sequence relationships of the NZB and BALB/c myeloma  $\kappa$  chains in their first 23 residues. Individual immunoglobulin chains are represented by number designations at the terminal twigs of the tree. The NZB V regions are boxed to distinguish them from the BALB/c  $V_\kappa$  counterparts. Many NZB sequences are three or more base changes from their nearest BALB/c counterpart. For example, the PC2316, PC2367, and PC2200 cluster in the upper right and the PC144, PC2419, and PC2454 set in the lower left constitute groups of related NZB sequences that have no BALB/c counterparts. Conversely, sets of related  $\kappa$  chains of the BALB/c set such as the TEPC15 and HOPC8 set have no NZB counterpart.

(iii) We have found a diversity distance index  $D(B,N)$  of 4.99 as a measure of the difference between the distributions of amino acids in the first 23 residues of the BALB/c and NZB

| Light Chain | Position           |   |    |    |    |   |   |     |   |   |   |     |   | Heavy Chain Class |       |       |   |     |   |   |       |   |   |       |       |       |
|-------------|--------------------|---|----|----|----|---|---|-----|---|---|---|-----|---|-------------------|-------|-------|---|-----|---|---|-------|---|---|-------|-------|-------|
|             | 1                  | 5 | 10 | 15 | 20 |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   |       |       |       |
| PC2408      | D                  | I | V  | M  | T  | Q | F | P   | S | S | L | N   | V | S                 | A     | G     | E | S   | V | T | M     | S | C | IgG2b |       |       |
| PC2367      | D                  | I | V  | M  | S  | Q | S | P   | S | S | L | A   | V | S                 | V     | G     | Z | K   | V | T | M     | S | C | IgG2a |       |       |
| PC2316      | -----              |   |    |    |    |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   | IgA   |       |       |
| PC3612      | D                  | I | V  | M  | T  | Q | S | H   | K | F | M | S   | T | S                 | V     | G     | B | R   | V | S | I     | T | C | IgG2b |       |       |
| PC3609      | -----              |   |    |    |    |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   | IgG2b |       |       |
| PC3249      | ----- (H)(K) ----- |   |    |    |    |   |   |     |   |   |   |     |   | D                 | X     | ----- |   |     |   |   | X     | X | X | IgG1  |       |       |
| PC2200      | D                  | I | V  | M  | S  | Q | S | P   | S | S | L | A   | V | S                 | A     | G     | Z | K   | V | T | M     | B | C | IgG2a |       |       |
| PC118       | D                  | V | V  | V  | T  | Q | T | P   | L | S | L | P   | V | S                 | F     | G     | D | Q   | V | S | I     | S | C | IgA   |       |       |
| PC39        | D                  | I | V  | M  | T  | Q | S | Q   | K | F | M | S   | T | S                 | V     | G     | D | R   | V | S | I     | T | C | IgG2a |       |       |
| PC920       | D                  | I | V  | L  | T  | Q | D | E   | I | S | N | P   | V | T                 | S     | G     | E | R   | V | X | I     | S | C |       |       |       |
| PC938       | D                  | I | Q  | M  | T  | Q | S | (S) | S | S | F | (S) | V | S                 | L     | G     | B | R   | V | T | I     | T | C |       |       |       |
| PC144       | E                  | I | V  | L  | T  | Q | S | P   | A | L | M | A   | A | X                 | P     | G     | Z | K   | V | T | I     | T | C | IgA   |       |       |
| PC2419      | -----              |   |    |    |    |   |   |     |   |   |   |     |   | S                 | ----- |       |   |     |   | E | ----- |   |   |       |       | IgG2a |
| PC2454      | -----              |   |    |    |    |   |   |     |   |   |   |     |   | S                 | ----- |       |   |     |   | E | ----- |   |   |       |       | IgG2a |
| PC613       | D                  | I | Q  | M  | T  | Q | S | P   | S | X | L | S   | A | S                 | L     | G     | K | G   | V | T | I     | T | C |       |       |       |
| PC674       | D                  | I | Q  | M  | T  | Q | S | P   | A | S | L | S   | V | S                 | V     | G     | E | T   | V | T | I     | T | C | IgA   |       |       |
| PC2880      | D                  | I | V  | L  | T  | Q | S | P   | A | S | L | A   | V | S                 | L     | G     | Q | R   | A | T | I     | T | C | IgG2b |       |       |
| PC3741      | -----              |   |    |    |    |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   | IgM   |       |       |
| PC1229      | -----              |   |    |    |    |   |   |     |   |   |   |     |   | Z                 | ----- |       |   |     |   |   |       |   |   |       | IgG2a |       |
| PC2155      | E                  | I | V  | L  | T  | Q | S | P   | A | I | T | A   | A | S                 | L     | G     | Z | K   | V | T | I     | T | C | IgG2b |       |       |
| PC2787      | D                  | I | Q  | M  | T  | Q | S | P   | S | S | L | A   | A | S                 | L     | G     | Z | R   | I | S | L     | T | C | IgM   |       |       |
| PC2205      | D                  | V | V  | M  | T  | Q | T | P   | L | S | L | P   | V | S                 | L     | G     | D | Q   | A | S | I     | S | C | IgG2b |       |       |
| PC2567      | -----              |   |    |    |    |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   | Ig3   |       |       |
| PC2954      | D                  | I | Q  | M  | T  | Q | S | M   | A | S | L | S   | A | S                 | V     | G     | Z | T   | V | T | I     | T | C | IgA   |       |       |
| PC3858      | D                  | I | L  | L  | T  | Q | S | P   | A | T | L | S   | V | S                 | P     | G     | E | R   | V | S | F     | S | C | IgA   |       |       |
| PC3936      | D                  | I | Q  | M  | T  | Q | S | P   | S | N | L | S   | A | S                 | L     | G     | E | R   | V | S | L     | T | C | IgA   |       |       |
| PC2279      | D                  | V | Q  | I  | T  | Q | S | P   | S | Y | L | A   | A | S                 | P     | G     | E | (T) | I | T | I     | N | C | IgG2b |       |       |
| PC373       | Blocked            |   |    |    |    |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   |       |       |       |
| PC2426      | Blocked            |   |    |    |    |   |   |     |   |   |   |     |   |                   |       |       |   |     |   |   |       |   |   | IgG2b |       |       |

FIG. 1.  $NH_2$ -terminal amino acid sequences of NZB  $\kappa$  chains (one-letter amino acid code). X denotes an unknown residue. Brackets indicate an uncertainty of sequence assignment. "Blocked" denotes that the  $\alpha$  amino group is covalently blocked and not free for the Edman reaction. Blanks indicate unassigned immunoglobulin class.

Table 2. Antigen binding spectrum of NZB and BALB/c myelomas

|                        | NZB | BALB/c |
|------------------------|-----|--------|
| DNA                    | ++  | -      |
| Dinitrophenol          | -   | ++     |
| Phosphorylcholine      | -   | ++     |
| $\alpha$ -1,3-Dextran  | -   | +      |
| $\alpha$ -1,6-Galactan | -   | +      |
| $\alpha$ -1,6-Dextran  | +   | +      |

+, 1-10 tumors of that specificity have been found; ++, >10 tumors found.

light chains (Table 1). Randomly chosen sets of sequences taken from the BALB/c population would give a diversity distance index [ $D(B,27)$ ] of only 3.69 with a SD of 0.60. Thus, the NZB sequences deviate from their BALB/c counterparts by more than two SD and, accordingly, the probability these two populations of proteins were chosen from the same pool is small (<0.015).

**Why Does the Myeloma Process in BALB/c and NZB Mice Appear to Transform Distinct Populations of Lymphocytes?** The complete or nearly complete sequences of 19 NZB  $V_k$  regions of one frequently expressed subgroup ( $V_{k21}$ ) have been determined (unpublished data). Several of these  $V_k$  regions are identical to BALB/c  $V_k$  regions. Accordingly, the NZB and BALB/c mice share some identical  $V_k$  region. Thus, two models may account for the apparent transformation of distinct populations of lymphocytes during the induction of myeloma tumors in NZB and BALB/c mice. The BALB/c and NZB mice may have certain distinct V genes or they may have genetic differences, outside the V structural genes, that either produce different antigenic exposures (e.g., susceptibility to viral infections) or operate as distinct control elements to

modulate V gene expression. Obviously, both of these possibilities could be true.

Certain distinct V genes could have been fixed in these strains from polymorphisms that preexisted in their ancestors. If so, different V regions could be expressed in these strains. Thus, some of the sequences found in one strain but not in the other (Fig. 2) may constitute structural gene polymorphisms. This hypothesis does not explain the distinct immunoglobulin class distribution in the myeloma proteins of these two strains, nor does it account for the observation that the myeloma proteins in the two strains bind distinct sets of antigens (Table 2) because a response to any of these haptens can be obtained in immunized animals from either strain.

An alternative and perhaps more likely explanation is that genetic differences outside the V genes lead to the expression of distinct populations of V genes in the lymphocytes of these strains. For example, the physiological differences in these two strains may lead to distinct internal antigenic environments, or distinct regulatory elements in the two strains may lead to the clonal expansion of (or induction of tolerance in) different sets of lymphocytes in the two strains. Hence, the myeloma process would transform distinct populations of lymphocytes in these two strains. This supposition is consistent with the observation that NZB but not BALB/c mice develop autoimmune disease with the concomitant expansion of clones of lymphocytes to self antigens (20) and that these antiself immunoglobulins are seen in the NZB but not the myeloma proteins. For example, 12 different NZB myeloma proteins bind DNA whereas few, if any, of their BALB/c counterparts do (Table 2). This model proposes that the same structural genes for both sets of V sequences exist in both strains but that other genetic differences create distinct internal conditions that lead to the expression of different subsets of V genes. A possible way to differentiate between the models would be through nucleic acid

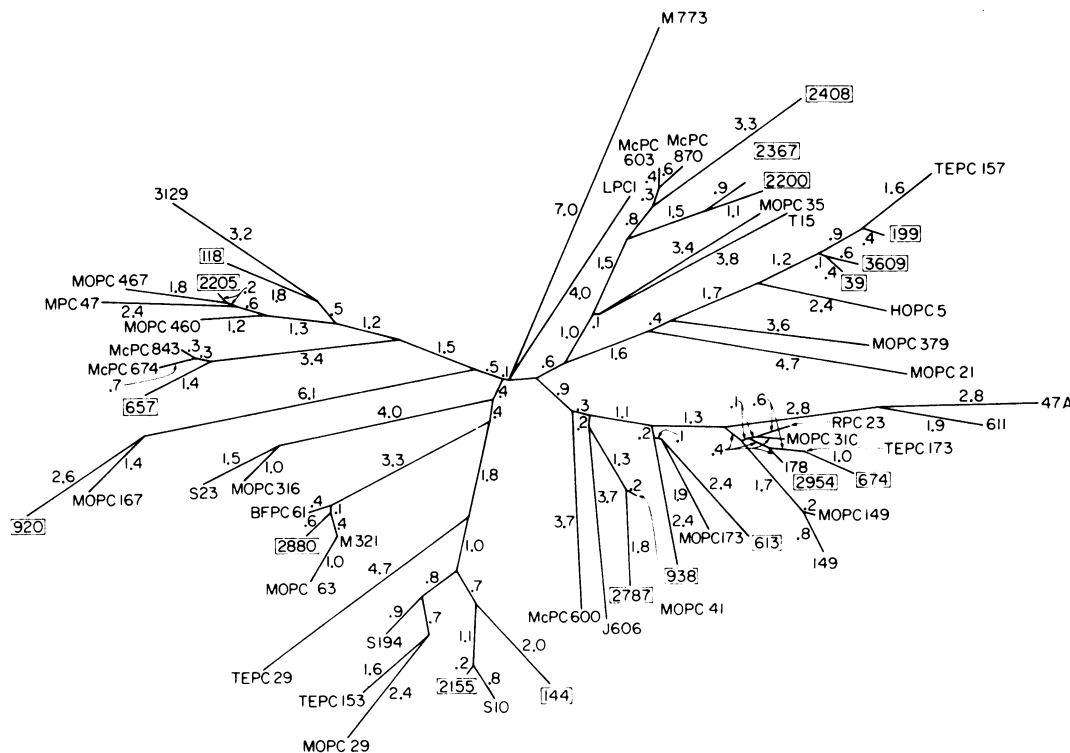


FIG. 2. Relatedness tree of mouse light chains. Boxes enclose NZB light chains. Examples of BALB/c light chain sequences with multiple members are not indicated. Lengths of lines connecting proteins and nodes indicate the minimal number of base substitutions. [A preliminary form of this relatedness tree was published in ref. 4 and has been refined with additional data and computer searches for a tree with minimal branch lengths (see ref. 16).]

chemistry, by determining whether identical structural genes for individual  $\kappa$  chains exist in both strains. The final answer may be that both models are partially correct.

**Because Myeloma Proteins Appear to Represent Only a Subset of the Total Lymphocyte Repertoire That the Mouse Can Express, V Gene Counting from Myeloma Data Provide Only Minimal Estimates of Antibody Gene Diversity.** There are several reasons for believing that the myeloma populations in BALB/c and NZB mice represent only a subset of the total immunoglobulin diversity these mice are capable of producing. (i) The antigen-binding spectra are quite distinct (Table 2) even though both strains can, by immunization, produce specificities not seen in the myeloma population. (ii) The normal pool of  $\kappa$  chains has residue alternatives not seen in the myeloma pool and vice versa (9). (iii) The  $V_H$  subgroup distributions in the myeloma population and in normal immunoglobulins are quite distinct. Accordingly, the myeloma process provides a window through which we can glimpse the antibody repertoire, but the window probably views just a fraction of the total diversity. Because it is impossible to say how small this fraction is, it is difficult to estimate what fraction of the antibody repertoire is expressed by the myeloma process in each of the inbred strains. Estimates of the number of V subgroups in the antibody families, based on the sequence analysis of myeloma proteins, are therefore minimal estimates and could be far too low. Thus, the number of V genes in many antibody families may be substantially higher than previously estimated.

This work was supported by grants from the National Institutes of Health. E.L. was supported by a National Institutes of Health Graduate Research Training Grant.

1. Potter, M. (1972) *Physiol. Rev.* **52**, 631-719.
2. Gally, J. A. & Edelman, G. M. (1970) *Nature (London)* **227**, 341-348.
3. Hood, L., Campbell, J. H. & Elgin, S. C. R. (1975) *Annu. Rev. Genet.* **9**, 305-353.
4. Hood, L., Loh, E., Hubert, J., Barstad, P., Eaton, B., Early, P., Fuhrman, J., Johnson, N., Kronenberg, M. & Schilling, J. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 817-835.
5. Potter, M. (1970) *Ann. N. Y. Acad. Sci.* **190**, 306-321.
6. Press, J. L. & Klinman, N. (1974) *Eur. J. Immunol.* **4**, 155-159.
7. Capra, J. D., Wasserman, R. L. & Kehoe, J. M. (1973) *J. Exp. Med.* **138**, 410-427.
8. Capra, J. D. & Kehoe, J. M. (1975) *J. Immunol.* **114**, 678-681.
9. Hood, L., Barstad, P., Loh, E. & Nottenburg, C. (1974) in *The Immune System: Genes, Receptors, Signals*, eds Sercarz, E. E., Williamson, A. R. & Fox, C. F. (Academic, New York), pp. 119-139.
10. Barth, W. F., McLaughlin, C. L. & Fahey, J. (1965) *J. Immunol.* **95**, 781-790.
11. Warner, N. L. & Varry, Z. (1970) *J. Immunol.* **105**, 812-817.
12. Hood, L., McKean, D., Farnsworth, V. & Potter, M. (1973) *Biochemistry* **12**, 741-749.
13. Smithies, O., Gibson, D., Fanning, E. M., Goodfliesch, R. M., Gilman, J. G. & Ballantyne, D. M. (1971) *Biochemistry* **10**, 4912-4921.
14. Pisano, J. & Bronzert, T. (1972) *Anal. Biochem.* **45**, 43-59.
15. Hood, J. M., Loh, E. & Hood, L. (1976) *Biochem. Genet.* **14**, 467-479.
16. Fitch, W. & Margoliash, E. (1967) *Science* **155**, 279.
17. Potter, M., Padlan, E. & Rudikoff, S. (1976) *J. Immunol.* **117**, 626-629.
18. Barstad, P., Hubert, J., Hunkapiller, M., Goetze, A., Schilling, J., Black, B., Eaton, B., Richards, J., Weigert, M. & Hood, L. (1978) *Eur. J. Immunol.* **8**, 497-503.
19. Schubert, D., Jobe, A. & Cohn, M. (1968) *Nature (London)* **220**, 882-885.
20. Howie, J. B. & Helyer, B. J. (1968) *Adv. Immunol.* **9**, 215-266.