Diversification, Not Use, of the Immunoglobulin V_H Gene Repertoire Is Restricted in DiGeorge Syndrome

By R. N. Haire, R. D. Buell, R. T. Litman, Y. Ohta, S. M. Fu,* T. Honjo,‡ F. Matsuda,§ M. de la Morena, J. Carro, R. A. Good, and G. W. Litman

*From the Department of Pediatrics, University of South Florida, All Children's Hospital, St. Petersburg, Florida 33701; the *University of Virginia, School of Medicine, Charlottesville, Virginia 22908; the ‡Department of Medical Chemistry, and the §Center for Molecular Biology and Genetics, Kyoto University, Kyoto 606, Japan*

Summary

Immunoglobulin (Ig) genes were isolated from unamplified conventional as well as polymerase chain reaction–generated cDNA libraries constructed from the peripheral blood cells of a patient with complete DiGeorge syndrome. Comparison of the sequences of 36 heavy chain clones to the recently expanded database of human V_H genes permitted identification of the germline V_H genes that are expressed in this patient as well as placement of 19 of these genes in a partially resolved 0.8-mb region of the human V_H locus. The pattern of V_H gene use does not resemble the fetal (early) repertoire. However, as in the fetal repertoire, there are a number of cDNAs derived from germline genes that previously have been identified as autoantibodies. Two $D\mu$ sequences also were identified, as was another sequence resulting from a unique recombination event linking J_H to an unidentified sequence containing a recombination signal sequence–like heptamer. All of the DiGeorge cDNAs are closely related to germline V_H genes, showing little or no evidence of somatic mutation. In contrast, comparably selected IgM V_H sequences derived from normal adult and age-matched human libraries, and from a second DiGeorge syndrome patient in whom the degree of thymic dysfunction is much less severe, exhibit considerable evidence of somatic mutation. The absence of somatic mutation is consistent with the atypical development of functional antibody responses associated with complete DiGeorge syndrome and implicates a role for T cells in the generation of diversity within the B cell repertoire.

Analyses of the Ig repertoire, as represented in peripheral blood, have been the focus of recent investigations. Owing to the relative ease of PCR amplification, most of the efforts have been directed at the CDR3 junction and the generation of diversity through D_H element (as well as J_H) usage and modification (1–3). Studies of the human fetal and cord blood repertoires have been carried out using conventional cDNA libraries (4–7), and while the normal human V_H repertoire has not been studied extensively in terms of expression of specific V_H genes, recent advances in the enumeration and mapping of germline V_H genes (8–10) now allow such investigations. These findings have had considerable impact in terms of our understanding of the developmental expression of Ig genes as well as the relationship of autoimmunity to specific genes. The nature of Ig gene usage in human primary immunodeficiency diseases, including X-linked agammaglobulinemia (11–13) common variable immunodeficiency (14) and severe combined immunodeficiency (15), has been examined. The present report addresses the Ig gene repertoire of patients with DiGeorge syndrome, a developmental field defect that involves dysmorphogenesis of the third and fourth pharyngeal pouches and is associated with varying degrees of fascial dysmorphia, hypoparathyroidism, anomalies of the great cardiac vessels, and partial development of the thymus. Abnormal B cell function in these patients is associated with the thymic abnormality and results in susceptibility to microbial pathogens, despite the maintenance of low to near normal levels of circulating Ig. The nature of the Ig genes that are transcribed in these patients is significant in terms of their impaired antibody function. In this regard, we have produced cDNA libraries from the PBMC of a DiGeorge patient with near complete failure of development of the thymus as well as the T cell compartment of immunity. Gene transcripts in this patient are compared with those expressed by normal neonates, a normal adult, and a second patient with a much less complete form of the disease, so-called partial DiGeorge syndrome.

Materials and Methods

Patients. At the time of study, the patient was an 8-mo-old female born full term after an uneventful pregnancy to unrelated

J. Exp. Med. © The Rockefeller University Press • 0022-1007/93/09/0825/10 $2.00
Volume 178 September 1993 825–834

parents. The family history was unremarkable. The patient presented with hypocalcemia and hypophosphatemia in the neonatal period and had undetectable levels of parathyroid hormone (PTH = <1 pmol/liter). Since the age of 3 mo, the patient suffered from recurrent respiratory tract infections, pneumonia, persistent oral candidiasis, and failure to thrive. A vascular ring (right aortic arch and aberrant left subclavian artery) was evident on cardiac workup. At 2 mo, the patient received oral polio vaccine and diptheria, pertussis, tetanus vaccine and at 6 mo, the child developed diarrhea, and polio virus was isolated from the stool. At 8.5 mo, immunological parameters were: T3, <1% (normal range [N] = 63–85%); T4, <1% (N = 37–57%); T8, 2% (N = 18–36%); T11, 10% (N = 71–89%); NK, 8% (N = 2–14%); B1, 86% (N = 7–19%); B4, 86% (N = 8%); B6, 10% (N = 2%); TCR γ/δ, <1% (N = 4–10%); TCR α/β, <1% (N = 64–68%); and CD5, 54% (N = 80%). The T cell numbers and proportions were exceedingly low and the B cell numbers and proportions were markedly elevated. Since the patient had very few $CD3^+$ T cells, the $CD5^+$ cells were likely to be B cells. Thus, ~70% of the patient's B cells were $CD5^+$ and this percentage is similar to that of normal infants of a comparable age (16). The total lymphocyte count was 2,128/mm^3 (10 wk earlier, 4,088/ mm^3). Thymulin level was not detectable, at <1:4 dilution (N = 1:16–128), mitogen-induced lymphocyte transformations and mitogen-induced production of Ig-secreting cells were decreased. At 3 mo, IgG = 187 mg/dl, IgA = <6 mg/dl, IgM = 30 mg/dl, and anti-B isohemagglutinin titer = 1:2.

The Ig V_H gene repertoire was also studied in a 31-mo-old male child with partial DiGeorge syndrome who had normal facial features, history of hypoparathyroidism, hypocalcemia in the neonatal period, cardiac abnormalities, and reduced T cell numbers, to approximately one-half of the normal values for age-matched controls. The normals used for comparison were a healthy adult volunteer and a 9-mo-old child.

PBMC were prepared from 20 ml of the patient's blood by the standard Ficoll-Hypaque method. RNA was prepared by the RNAzol B method (Cinna/Biotecx, Houston, TX). For conventional cDNA library synthesis, poly(A)$^+$ RNA was selected by oligo(dT)-conjugated magnetic beads (Dynal, Oslo, Norway). A cDNA kit (Pharmacia Fine Chemicals, Piscataway, NJ) was used with only minor modification to synthesize cDNA that was ligated into λ gt11.

A rapid amplification of cDNA ends [RACE] library was generated from 5 μg of whole RNA by reverse transcription with 200 ng of oligo(dT) primer and 30 U AMV reverse transcriptase (Seikagaku, Rockville, MD). The first-strand product was filtered in an Ultrafree MC 10,000–molecular weight cutoff (Millipore Corp., Bedford, MA) to remove excess primer. The cDNA was tailed with oligo(dG) for 30 min using 30 U of TdT (GIBCO BRL, Gaithersburg, MD), filtered, and PCR amplification (17) was performed using an EcoRI-linked poly(dC) (n = 14) and an internal IgM constant region primer: TTGCACACCACGTGTCG (codon 203–209). The PCR product was digested with EcoRI (using the natural site in the first exon of the IgM heavy chain C region [C_H1) and applied to a 1% NuSieve GTG agarose gel (FMC, Rockland, ME). The gel section containing material >300 bp was heated to liquify the agarose, extracted with phenol, ethanol precipitated, and ligated into λ gt11.

Library screening was done with an oligonucleotide probe complementary to J_H 1-6: 5′-TGAGGAGACGGTGACCAGGGTTCC-TTGGCCCCAG-3′. Positive clones were amplified with an EcoRI-linked J_H primer together with either a right or left λ arm primer and the product was subcloned into M13 for sequencing. Several cDNAs were selected from the RACE library by screening with DNA probes produced by PCR using family-specific framework region 1 (FR1)[1] and FR3 primers (18). To determine the C_H iso-

[1] *Abbreviation used in this paper:* FR, framework region.

Table 1. *DiGeorge cDNAs Are Grouped by Their Library Source, Conventional Library vs. RACE Library*

Clone	V_H family	Length (codon)	Germline parent	Substitutions
		Conventional library		
22	1	CDR1 (34)	71-5 (1-58P)	
27	1	5′ leader	1-18	
37	1	FR3 (75)	DP10* or HV1263(?)	
3	3	5′ leader	DP77 (3-21)	
4	3	CDR2 (60)	DP42 (3-53)*	
13	3	5′ leader	3-7	
15	3	5′ leader	GLSJ2*	
17	3	5′ leader	3-9	
21	3	5′ leader	9-111 (3-30)*	
31	3	5′ leader	DP51 (3-48)	
34	3	5′ leader	DP58	
36	3	5′ FR2 (36)	9-1 (3-15)*	
38	3	5′ leader	DP50	
10	4	FR2 (41)	4-4	
20	4	3′ leader	DP65 (4-31)	
28	4	3′ leader	4.11 (4-59)*	

continued

type of the sequences from the conventional library, another PCR reaction was carried out using C_H-specific primers.

Homology searches to V_H sequences containing open reading frames, extracted from GenBank (8–10), used I/FIND (Intelligenetics, Mountain View, CA).

Results

V_H Family Representation in the Patient with Complete DiGeorge Syndrome. Owing to the potential of PCR-RACE methods (19) for biasing the representative frequency of different V_H genes and gene subsets (see below), a conventional cDNA library was prepared from 70 μg of patient whole RNA in parallel with the equivalent amount of RNA from a normal adult control. 2×10^5 recombinants were generated in the DiGeorge library and 16 Ig J_H^+ clones were recovered. Parallel screenings of the (normal) control library yielded ~1,000 J_H^+ cDNAs and ~10^6 recombinant plaques. All of the cDNAs recovered from the DiGeorge syndrome patient conventional library were sequenced. 11 sequences extend from the leader to J_H and 5 are truncated within V_H (Table 1); all represent a correct open reading frame. Each of the 16 cDNAs recovered from the conventional library is derived from a different germline gene (see below).

The J_H-selected clones (nos. 203 and 205–211) from the PCR-RACE library are all V_HI, suggesting a possible V_HI amplification bias (see below). Therefore, an additional nine clones were selected from this library using V_H family–speci-

Table 1. *(continued)*

Clone	V_H family	Length (codon)	Germline parent	Substitutions
		RACE library		
203	1	5′ leader	1-18	
205	1	5′ leader	1-18	FR3(R) 89GTG→GGG
206	1	5′ leader	1-18	FR3(2R) 75ACG→GCG 84 TCT→CCT
207	1	5′ leader	1-18	
208	1	5′ leader	DP10*	
209	1	5′ leader	1-18	FR1 (S), CDR2 (R) 18GTG→GTA 51ATC→ACC
210	1	5′ leader	DP10*	
211	1	5′ leader	DP10*	
		V_H-selected race		
1.1	1	Mid leader	1-3B	
1.2	1	5′ leader	1-18	
1.4	1	5′ leader	1-18	
1.5	1	5′ leader	1-18	
3.1	3	5′ leader	GLSJ2*	
4.1	4	5′ leader	DP65 (4-31)	
4.3	4	5′ leader	4.21 (4-34)	
4.5	4	5′ leader	4.21 (4-34)	
4.2	4	5′ leader	V2-1 (4-39)	

The clones from the RACE library are either isolated by a V_H-independent screening method (J_H) or by V_H selection. The germline parent gene is determined by homology to the database and, in most cases, is indicated by the original name. Germline DP50 was first reported as 3019b9 (35) and 1-18 was reported as DP14; the alternative designations are used for simplicity. All cDNAs are 100% identical to the first germline gene indicated except where there is a substitution noted (last column). Substitutions refer to the functional region. Silent (*S*), replacement (*R*), and the number of mutations, if more than one, are indicated, followed by codon number (36) and the specific nucleotide substitution relative to gene 1-18(DP14). The mapped gene equivalent (9) is underlined and, in some cases, placed in parenthesis when the original name is different or when the germline gene from the map differs slightly from the first germline gene listed. For example, DP51 (cDNA31) is identified as mapped gene 3-48, DP77 (cDNA 3) as 3-21, DP42 (probable cDNA 4) as 3-53, DP65 (cDNAs 20 and 4.1) as 4-31, and 4.11 (cDNA 28) as 4-59, even though the germline sequences homologous to the cDNAs differ by a single nucleotide from the indicated germline sequence derived from the map. Clone length indicates where, in a functional region of V_H, the sequence begins with the codon number in parentheses (36). 5′ leader means the sequence obtained begins upstream of or near the 5′ end of the V_H leader sequence. All germline genes denoted DP are taken from (10). Gene 71-5 (23) was originally considered a pseudogene and later determined to be a functional gene (10). Other germline genes are HV1263, GLSJ2, 9-III and 9-1,4.11,4.21, and V2-1 (8).
* Genes that are associated with the fetal repertoire.

fic probes and these represent V_HI, V_HIII, and V_HIV families. Overall the RACE-generated clones are longer than those produced by conventional cDNA synthesis, with all but one having a complete leader sequence. While there are differences between the two groups of sequences, there are some germline elements common to both libraries, specifically V_HI 1–18, V_HIII GLSJ2, and V_HIV DP10. Detailed inspection of the noncoding regions and the 5′ terminus of the most commonly amplified RACE clones in this study indicates that the V_HI genes DP10 and 1–18 are not homologous (data not shown). Thus, the basis for a TdT preference is not evident from comparisons of these sequences. These observations emphasize the difficulty in using PCR libraries to quantitate the relative frequency of different transcript classes (i.e., V_H families or subfamilies) or even to determine their presence or absence (R. N. Haire and G. W. Litman, unpublished observations). However, in some cases, for example the partial DiGeorge syndrome patient described in Table 2, V_H family representation in the PCR-RACE cDNA library appears to be proportional to germline representation and is similar to the results from the conventional library of the DiGeorge syndrome patient and from the normal controls (data not shown). Notwithstanding, these complications, PCR-amplified clones can yield very useful information regarding junctional diversity and somatic mutation.

Specific V_H Gene Use. 30 cDNAs, including no. 28, the single IgA clone (all others are IgM), exactly match germline database sequences (Table 1). Two clones with incomplete CDR2 sequence can be identified from the available sequence, and the unique matching germline parent is indicated. Clone 37, truncated in FR3, is homologous to two known germline genes. Only clone 209, with a replacement substitution in CDR2 and a silent mutation in FR1, is consistent with a somatically modified sequence but could also reflect a polymorphism and/or represent the product of closely related genes, PCR-introduced cloning errors, or sequence determination errors. The substituted nucleotides in clones 205 and 209 were confirmed by recloning and resequencing. The predominance of mismatches in FRs in not usually considered characteristic of somatic mutation, and the mutations are at different positions, suggesting that the FR3 substitutions in clones 205 and 206 are not due to a polymorphic variant of 1-18. Finally, it is noteworthy that the incidence of mismatched nucleotides, while low, is higher in the RACE library than in the conventional library. Taq polymerase error arising during the preliminary RACE amplification cycles may be an explanation for these substitutions.

Eight V_HIII cDNAs from the conventional library extend from the 5′ leader through J_H and are matches with germline sequences: DP77, DP50, DP51, DP58 (10), 3-7, 3-9 (9), GLSJ2 (8), and 9-III (20). Clone 13 has a 9-bp V_H truncation relative to germline 3-7 at the V_H-D_H junction. Clone 36 extends from FR2 and corresponds to germline V_HIII 9-1 (20). Truncated clone 4 exactly matches and probably is derived from germline gene DP42 (10). Transcripts using V_HIV germline genes 4-4 (9), DP65 (10), and 4.11 (21) also were recovered in screening the conventional library. Clone 10 represents the first cDNA described that derives from the newly defined germline gene 4-4. Four additional V_HIV cDNAs derived from the RACE library correspond to germline genes DP65, 4.21 (21), and V2-1 (22).

V_HI cDNAs identified in the conventional library are derived from at least three germline genes. 71–5 originally was considered a pseudogene (23) on the basis of an atypical intron splice sequence as well as an atypical RSS nonamer, but clone 22, while truncated, is homologous with 71-5 throughout its observed length and appears to be a functional transcript. On the basis of FR3 identity, cDNA clone 37 is assigned tentatively to germline gene DP10 or HV1263 (10). In the RACE library, V_HI cDNAs are derived from germline genes 1-3b (24) and DP10. The remaining V_HI genes correspond as perfect matches or with limited FR3 substitu-

Table 2. *A Comparison of Gene Usage in Partial and Full DiGeorge Syndrome*

	DiGeorge (n = 33)	Partial DiGeorge (n = 25)
Germline parents identified	19	11(−14)
Nonmatching sequences	3	7(−11)
Substitutions/V_H	0-2	0-31
Sequences greater than two substitutions	0	7
Expressed germline genes shared with DiGeorge patient	Not applicable	7
Expressed germline genes shared with fetal repertoire	6	2
CDR3 length (*bp*)	23(±10)	25(±8)

Germline parents identified indicates the minimum and probable maximum number of different genes expressed in the cDNA repertoires. The smaller number reflects only the positively identified germline genes while the larger number represents additional substituted sequences that are similar to and probably derived from other known germline V_Hs. The overall range of substitutions per V_H gene is given in the third line of the table. n, number of clones studied. CDR3 length is defined as the total number of nucleotides found between identifiable V_H- and J_H-encoded sequences.

tions relative to germline gene 1-18 (9), except for no. 209, which has single changes in FR1 and CDR2, and therefore may represent a somatically mutated gene (1-18), as discussed above.

With the recently expanded data bases of human V_H sequences (9, 10), the incidence of somatic mutation in the IgM cDNAs of normal individuals can be estimated. Our analysis of the cDNAs from the human cord blood library (6) reveals that 7 of 11 sequences do not match database sequences and only 4 are exact matches. The nonmatching sequences differ by an average of 10 bp each with 50% of those in CDR2, indicating either somatic mutation or expression of other unknown germline genes. In the normal adult library, which we constructed in the course of this study, ~45% of the 23 fully analyzed sequences correspond to germline V_H gene transcripts, similar to the ratio observed for cord blood. In a second control library, derived from the peripheral blood of a 9-mo-old nonimmunodeficient infant, 7 of 11 sequences do not match database germline genes with at least five substitutions per cDNA.

Matching cDNAs to germline sequences is more difficult in the RACE cDNA library prepared from the partial DiGeorge syndrome patient, summarized in Table 2, due to the presence of some sequences that are either highly mutated or derived from unidentified germline genes. The unidentifiable sequences comprise a minimum of seven cDNAs: two V_HI and five V_HIII, with an average of 16 (±10) "randomly" distributed substitutions compared with the nearest germline V_H; i.e., similar cDNAs are substituted at different sites. These appear to be highly somatically mutated transcripts. Unlike the situation in the DiGeorge syndrome patient, the partial DiGeorge syndrome patient cDNAs are IgM, as well as IgA and IgG, but there are both unmutated and highly mutated cDNAs of each isotype.

The availability of a partial physical map of the human V_H locus (9) permits mapping of the corresponding germline genes expressed in the complete DiGeorge syndrome patient (shown in Fig. 1) along with the relative map positions of genes identified in fetal libraries (4, 5, 7). Both groups of genes are widely dispersed over the locus and are largely nonoverlapping, with the exception of mapped genes 4-59/58P2, 3-53/63P1, 3-30/1-9 III, and 3-15/20P1. However, two unmapped genes also are shared between the fetal libraries and the DiGeorge syndrome patient, GLSJ2(56P1) and DP10/51P1. Therefore, only 32% of the genes from the DiGeorge syndrome patient are also associated with a "fetal" repertoire, i.e., approximately two-thirds of the expressed genes are not fetal. The partial DiGeorge syndrome patient shares seven germline genes (65%) with the repertoire from the patient with complete DiGeorge syndrome, but only shares two (18%) with the fetal repertoire (data not shown). The sequences obtained from the human cord blood cells are not homologous to any of those from the DiGeorge syndrome patient; however, this comparison is complicated by a high frequency of apparent somatic mutation present in the cord blood sequences.

Germline genes that are associated with autoantibodies also are shown (Fig. 1). Clone 1.1 is derived from germline 1-3b. This gene differs by only a single nucleotide from an anti-DNA autoantibody (24). Note that the germline gene 1-3b was identified as haplotype B in that investigation, while clone 10 from our patient with partial DiGeorge syndrome is derived from gene 4-4, haplotype A. Other specific genes expressed in this patient with the partial DiGeorge syndrome that may be associated with autoantibody include: germline 4.21 (mapped as 4-34), a polyreactive rheumatoid factor, and germline 9-1 (3-15), an anti-DNA antibody (8). Germline 1-18 differs from a reported autoantibody V_H1GRR by a single nucleotide and germline 1.9III (3-30), which also differs by a single nucleotide from the autoantibody Kim 4.6. Finally, germline gene 4.11 is mapped as gene 4-59 and is related to autoantibody Pag-1 (9), although four nucleotide substitutions have occurred. Of the seven germline genes expressed in the complete DiGeorge syndrome patient that are homologous to reported autoantibody sequences, all but one, GLSJ2, are located within the mapped region. Detection of specific

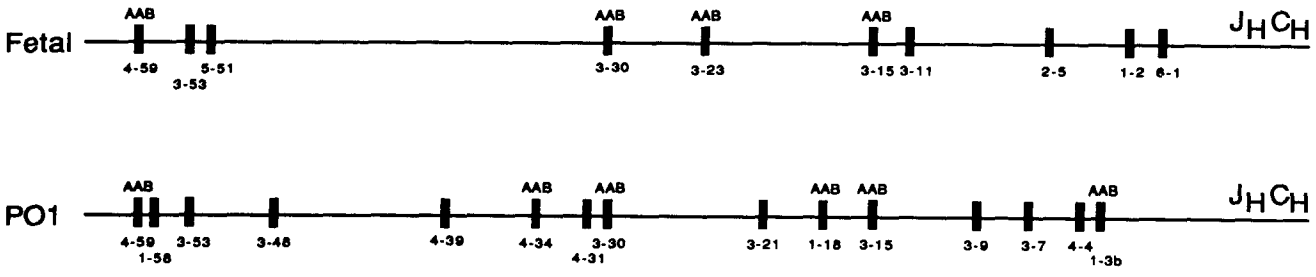

Figure 1. The relative physical position of the mapped germline parents for cDNAs derived from the complete DiGeorge syndrome patient compared with those assigned to the "fetal repertoire" (4, 5, 7). Fetal genes are mapped if they occur at least twice in the references cited, with the exception of 2-5, which was encountered only once. Note that not all germline parents identified in the DiGeorge and fetal repertoires could be assigned to the physical map and that the sequences shown are limited to those identified and mapped in a 0.8-mb J_H-C_H proximal region (9). Numerous other V_H genes have not been mapped as of yet (10) and must lie further upstream from J_H, lie outside of the contiguous portion of the V_H locus, or represent polymorphisms. The relationships between alternative names for the germline gene and designation of the mapped gene are evident in Table 1. The notation AAB indicates a gene coding for a cDNA identified as an autoantibody (9). In most cases, the sequences of the autoantibodies are identical to or differ by a single nucleotide from the mapped germline genes, except in the case of 4-59, which differs by 4 bp. The relative position shown is approximately proportional to the distance between genes on the 0.8-mb map.

autoantibody genes in the normal adult and in age-matched libraries is more difficult because of the large number of somatically modified sequences; however, to date only autoantibody-related genes 4.21, 1.9III, hv3005, and 1-18 have been observed.

CDR3 Region. Sequences of all cDNAs derived from the complete DiGeorge syndrome patients, from the conserved Cys (codon 92 FR3) at the 3′ terminus of V_H through the conserved Trp (codon 103 FR4) of J, are shown in Fig. 2. A high proportion of the cDNAs have a recognizable (germline) D_H element that has been joined in the conventional V-D-J manner; seven different D_H families are represented (2). The most common D_H family is DLR (33% of the sequences), followed by DXP (23%), DN (20%), $D_H$fl (10%), DK (7%), and DM (3%). We did not identify any DA or D_HQ52 members, and only one probable DIR sequence (clone no. 206) is noted. While the DLR family may be proportionally increased, the relative occurrence of most other D families is in good agreement with those reported by others investigating Ig expression in the peripheral blood of normal individuals (1, 2). The most commonly used specific D_H genes are D4 (n = 5), DN1-TG (n = 3; see below), DN1 (or DN4 in some cDNAs; n = 3), and DXP4-D23/7 (n = 4). We did not observe inverted D_H sequences of significant length that were not accountable for by other, noninverted D's. One case of joining of two D_H elements was observed in clone 3.1, in which D2 and D21/9 are joined in a particularly long CDR3.

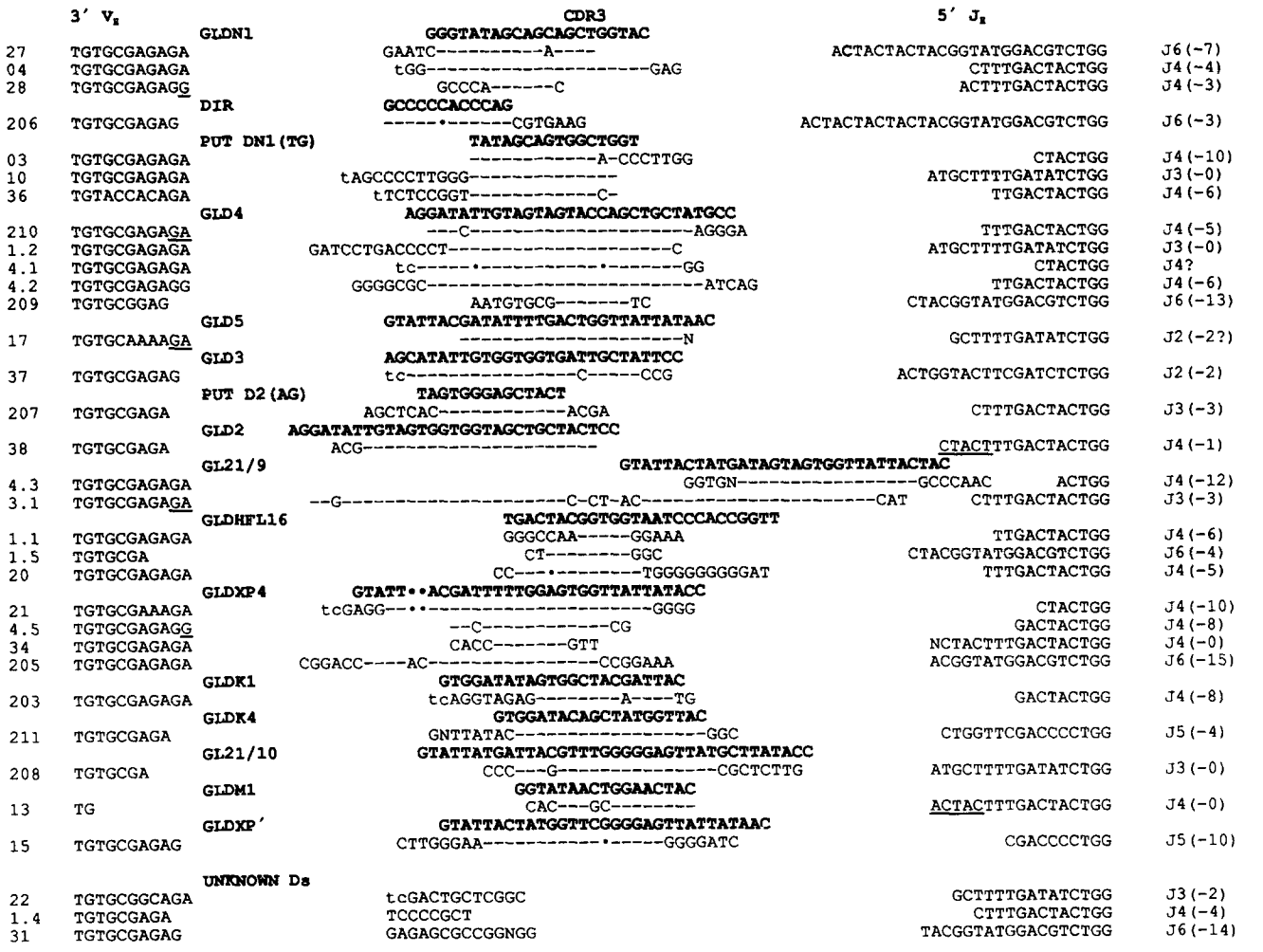

Figure 2. CDR3 sequences are shown under the corresponding complete germline D core sequences (***bold***). At the left are nucleotides attributable to the 3′ end of V_H, determined from homology with germline genes indicated in Table 1, to codon 95. At the right are nucleotides attributable to 5′ J_H indicated to the conserved TGG (codon 102). Nucleotides underlined at the end of V_H and the beginning of J_H are common to both the germline V or J and the germline D_H segment involved. Dashes indicate nucleotide identity with the germline D_H and substitutions are indicated within the dashed regions (except for the V-D1-D2-J clone 3.1, which has probable N region bases between the D's). The nucleotides to the right and left of dashed regions are assumed to be N region additions unless they can be identified as P nucleotides (*lower case*). Note that P nucleotide identity can be extended to the first TC pair within the CDR3 of clone 210 and similarly to the first T in clone 17 CDR3. Dots indicate a space introduced to maximize homology. Putative germline genes DN1-TG (1) and D2-AG are shown as consensus sequences derived from eight and seven cDNAs, respectively, obtained from other patients and normal individuals (data not shown).

The putative D_H element DN1-TG identified recently by Yamada et al. (1) and encountered in three of the cDNAs from this patient also is found in the control and probably represents an undescribed germline element. Similarly a D2 variant, D2-AG, was encountered in the DiGeorge syndrome patient and controls (R. N. Haire and Y. Ohta, unpublished observations). The core sequences for the putative D_H elements are shown in Fig. 2.

In general the CDR3 sequences in the patient with complete DiGeorge syndrome have rather high homology vs. normals to the germline core sequences. Somatic mutation rates within D_H elements are difficult to estimate due to N region additions, but are present at ~3% in the sequences from the DiGeorge syndrome patient vs. ~6–7% in controls (Fig. 2; 1, 3). D_H reading frame use in this study appears to be unrestricted in the normal sequences, in contrast to murine Ig genes. For example, DN1 is used differently in all three reading frames in clones 27, 04, and 28. Similarly, clones 210, 1.2, and 4.2 use D4 in different reading frames.

Overall CDR3 length, defined here as the sum of D, P, and N region (7) contributions, appears to be reduced somewhat in the patients with partial or complete DiGeorge syndrome compared with that of normal individual adults, with a mean of 23 (±10) bp in the DiGeorge syndrome patient, and 25 (±8) in the partial DiGeorge syndrome patient vs. ~33 in the normal adults (2). This latter finding was also confirmed by observations with normal sequences examined in the course of the present study (data not shown). However, mean CDR3 lengths of 24 and 14 have been observed in neonates (2) and in B cells from fetuses (7), respectively, and the CDR3 length seen in the patients' sequences may be normal for individuals at that age. The patient with DiGeorge syndrome exhibits N region nucleotide additions of 8.1 (±4) similar to the size distribution observed by Sanz (2); approximately one-third of the sequences in this study have an addition of >10 bp. With the possible exception of clone 17, none of the CDR3 regions that have identifiable D_H segments appear to lack N region additions. About one-third of the sequences from the DiGeorge syndrome patient have 5′ P nucleotides at the V-D junction, compared with ~50% in cDNAs derived from normal adults (1) and ~78% in those derived from fetal cDNAs (7); 5′ exonuclease activity is evident, especially at the D-J junction, in many of the sequences. J element usage in this complete DiGeorge syndrome patient is: J4 (~50%), J3 (18%), J6 (18%), J2 (6%), J5 (6%), and J1 (not seen). These frequencies correspond well with those from our normal databases and from available sequence data (1, 2).

Atypical Transcripts. In the course of screening the RACE cDNA library from the patient with complete DiGeorge syndrome using the J_H probe, we isolated and sequenced three short J_H^+ transcripts that lack V_H regions. Two transcripts are typical of the Dμ type (25): DiG3 consists of D4 joined to J_H6, with evidence of exonuclease activity at the D/J junction, and DiG4 consists of D4 joined to J_H4. Both transcripts begin at the 5′ noncoding (D intron), ~135 bp from D4, but differ in the degree of exonuclease catalyzed elimination of D4 sequence. In both cases, Cμ (IgM) sequences are contiguous with J_H.

A third transcript obtained from the patient with complete DiGeorge syndrome (DiG 5) is unique. J_H6, with 19 bp eliminated 5′, is joined to 209 bp of a sequence that is undetectable in searches of the mammalian database. However, 37 bp upstream from the J_H6 sequence is a perfect RSS heptamer, CACTGTG. This sequence also contains IgM joined to J_H and, as with the Dμ-like sequences, does not represent a spurious PCR priming event. The sequences of these three atypical transcripts from the DiGeorge syndrome patient are shown in Fig. 3. Finally, we have identified a single Dμ se-

DIG3 D4/J6 D_μ

aaaagccccctggaaatcatagtatcagcaggagaactagccagagacagcaagaggggactcagtg

RSS RSS D

actcccgcggggacaggaGGATTTTGTggggctcgtgtCACTGTGAGGATATNGTAGTAGTACCAG

J

CTGCTATcttgCTACTACTACTACGGTATGGACGTJ6 – Cμ

DIG4 D4/J4(?) D_μ

aaaaaccccctggaaatmataggtataagcagggagaactagcaagagacagcaagaggggactcag

RSS RSS D

taacccccgtggggacaggaGGATTTTGGggggctcgtgtCACTGTGAGGATATTGTAGTAGTACC

AGCTGCTATGgggtaJ4 – Cμ

DIG5

gactacaatctaagtctttctgatttcaaagctcctgcttattaccacttcacagcacacctagaaa

gcataaggtgggagcaggttaaaatgcatacttgtctttggttatcccctctactttcctgtccttc

RSS

tctgcctacctgtatctgcatccttggtCACTGTGgaaaattatcttctagggtcctgtggwgacag

J

agggcTATGGACGTJ6 – Cμ

Figure 3. Dμ transcripts (DIG3 and DIG4) and one aberrant transcript (DIG5) are shown from the beginning of recognizable germline D_H sequence up to the J_H oligonucleotide sequence used for subcloning. *m* = c or a, and *w* = t or a. The presence of Cμ was confirmed by PCR with isotype-specific oligonucleotide primers. Recombination signal sequence heptamer and nonamer (*RSS*) are shown in bold, J_H sequence (*J*) is in italics, followed by the J family designates, with the remaining J sequence, that of the PCR primer, omitted. D_H (*D*) coding sequence is underlined.

quence each from the peripheral blood of the partial DiGeorge syndrome patient and from a normal adult control (also from a RACE library) but have detected no counterpart to the aberrant transcript, DiG 5, from the patient with the complete DiGeorge syndrome.

Discussion

Considerable variation in the degree of immunodeficiency is noted in patients with DiGeorge syndrome. The T cell immunodeficiency ranges from mild deficiency in T cell number and function to virtually complete absence of T lymphocytes, absence of thymulin, and very high numbers and proportions of B lymphocytes in the blood. Patients with DiGeorge syndrome have a propensity to develop a high frequency of autoimmune manifestations and autoimmune disease. The profound T cell deficit present in the patient with severe or complete DiGeorge syndrome studied here occurs in <10% of patients afflicted with this primary immunodeficiency disease. A near complete lack of T cells provoked speculation that the B cell population might be comprised of a limited number of clonal lines with attendant restriction in the Ig V_H and possibly D_H as well as J_H repertoires due to presumed abnormalities of T cell–dependent B cell development. However, the repertoire of Ig that has developed in this example of near complete DiGeorge syndrome is quite distinct from the fetal repertoire and does not appear to be clonally restricted in terms of use of germline genes. At most, only six genes isolated from the patient with DiGeorge syndrome were derived from the fetal subset (Table 1).

Overall, a minimum of 19 different germline genes were found to be expressed in these libraries, inconsistent with a limited repertoire (14). Furthermore the RACE clones derived from the same germline parent use multiple, different D_H and J_H segments, and the different junctional N nucleotides confirm their uniqueness, arguing strongly against any underlying defect in B cell clonal expansion as the basis for antibody deficiency. The overlap between germline gene usage in the human cord blood library and that observed in this patient is probably minimal.

$D\mu$ sequences have previously been described in patients with primary immunodeficiency as well as in normal individuals, and their functional significance in the murine system has been defined (26). The apparent increased frequency of these transcripts in some primary immunodeficiencies may relate to reduced numbers of typical B cell transcripts. A unique transcript found in the complete DiGeorge syndrome patient appears to represent a recombination event between J_H and another region possessing a heptamer element. It will be of interest to determine whether a similar transcription product can be found in other patients with equivalent T cell deficiencies.

The availability of the recently expanded human V_H database and the physical map of the V_H locus has greatly facilitated the precise identification of germline genes for our analysis. Previously, separation of cDNAs derived from the various genes of the GLSJ2/3005/9-III complex was not possible and all of these highly related germline genes were termed "fetal." However, it recently has been determined that in the 56P1(−) haplotype, the predominant fetal gene is V26 (7), not the more closely related (to 56P1) genes 9-III or 3005. Furthermore, it has been suggested that the recurrence of certain sequences in different studies is consistent with the normal human V_H repertoire being limited to the expression of ~25 germline genes (14). However, that estimate may be somewhat low due to failure to resolve recently defined, unique germline genes.

It appears that in the DiGeorge syndrome patient a large number of genes associated with autoimmune reactivity may be expressed; such a circumstance may be avoided (or minimized) by normal T cell regulation. The significant expression of such genes (associated with autoimmune reactivity) may be a factor in the secondary pathogenesis of this disorder in which autoreactivity is a frequent, significant clinical feature.

In terms of V_H family representation, the composition of the conventional library is similar to that accepted for normal peripheral blood, i.e., V_HIII >V_HI, IV, and in this sense too, a restriction is not apparent in this patient. V_HV and V_HVI genes, which are prominent in the fetal repertoire, were not observed in this patient. Although V_HII, V_HV, and V_HVI genes were not detected in the conventional or RACE libraries, these families typically would be expected to account for only 10–15% of total Ig clones in a normal individual and seem to be reduced to undetectable levels in this patient. Screening of a large number of recombinants in the RACE library failed to reveal members of these families, although they have been detected both in RACE libraries of the age-matched control and in those of other immunodeficiency patients (our unpublished results) as well as in the conventional library from the normal adult. D_H segment usage in this DiGeorge syndrome patient primarily involves DXP and DLR segments and normal J_H element usage, whereas the fetal repertoire frequently uses DQ52 and J_H 1, 2, and 3 (5).

The CDR3 region length in this DiGeorge syndrome patient is typical for age-matched normal controls. The relatively limited occurrence of P nucleotides also argues against the use of a fetal repertoire in the complete DiGeorge patient. However, the presence of a high proportion of sequences identified with autoantibody activity, the $D\mu$ transcripts, and the absence of somatic mutation in the cDNAs are consistent with a fetal or antigenetically primitive repertoire.

The repertoire observed in the patient with partial DiGeorge syndrome, a condition intermediate between normal and complete DiGeorge syndrome, shares a significant fraction (50–64%) of expressed germline genes with that seen in the blood of the patient with the complete DiGeorge syndrome. There appear to be two classes of sequences in the patient with partial DiGeorge syndrome, those that are unmutated and a few highly mutated transcripts. Perhaps this phenomenon reflects a T cell defect that results in both a reduced incidence of somatic mutation and subsequent deficient regulation in selection of B cell clones.

Although the majority of the patient's B cells were $CD5^+$, we do not consider it to be atypical, as similar percentages of $CD5^+$ B cells are found in normal infants. The

question as to whether B cells are derived from distinct lineages remains controversial (27, 28). It is apparent that $CD5^-$ B cells can be induced in both human and mouse to express CD5. In the mouse, there is a preferential usage of proximal V_H segments in $CD5^+$ B cells (29, 30). Our data show no preferential usage of V_H segments and thus are inconsistent with preferential usage in human $CD5^+$ cells.

The unique repertoire observed in the patient with complete DiGeorge syndrome reflects the influence of abnormal thymic differentiation and development of the T cell system on B cell ontogenetic development (31). However, the T cell effect does not result in a fetal repertoire, but rather is reflected in a general failure of B cell diversification as reflected in the level of somatic mutation. In the 9,000 informative nucleotides identified in the present study, only five differ from data on existing germline sequences derived from unrelated individuals, suggesting not only a marked reduction in somatic mutation but a very low level of V_H polymorphism in humans. The number of differences encountered from germline sequences is equivalent to the rate of error in PCR amplification, which we estimate as ~1:1,700 nucleotides. Evidence for limited occurrence of somatic mutation in this patient is based on comparison both to adult and age-matched controls, as well as to analysis of cord blood data (6).

Somatic mutation is already well established by the time of birth in humans and continues to represent a source of Ig diversification within the peripheral B cells throughout adult life (3). Absence of somatic mutation in the complete DiGeorge syndrome patient accounts for the inability of such patients to mount an effective immune response to microbial pathogens. Thus, despite low to near normal levels of Ig, the quality of antibody produced as judged by both specificity and affinity must be markedly reduced. Infusion of normal Ig (intravenous gamma globulin) is a standard course of therapy in patients with DiGeorge syndrome.

The absence of somatic mutation may relate to T cell dependence of the ontogeny of normal B cell functional maturation. Mouse B lymphocytes acquire the capacity to give a normal adult-like heterologous response between 7 and 10 d of age (32, 33), and this functional maturation is independent of antigenic drive since it is observed in germ-free mice (34). In a model using irradiated mice, reconstituted with adult thymocytes and B cells from fetal liver and adult spleen, the maturation of B cells to produce antibodies with heterogenous affinities is shown to depend on cells derived from thymus. Based on the findings reported here, diverse usage of V_H genes in B cells is independent of T-B interaction(s). However, somatic maturation is dependent on T cells that are absent (or greatly reduced) in complete DiGeorge syndrome patients and in neonatal mice, but present to some degree in patients with partial DiGeorge syndrome, developed in human neonates, and developed to some degree in human fetuses, which have normal thymus functions. Thus, the generation of a B cell repertoire can be considered a two-step process, i.e., diversification of a V repertoire and somatic mutation. In patients with complete DiGeorge syndrome who have profound failure of thymus development, the V repertoire diversified but somatic mutation is compromised. This two-stage hypothesis may explain the functional antibody deficiency noted in patients with other immunodeficiency syndromes and significant levels of circulating Ig.

We thank Barbara Pryor for editorial assistance and J. D. Capra for providing human V_H sequence data.

G. W. Litman is the Hines Professor at the University of South Florida College of Medicine/All Children's Hospital and is supported by a grant from the National Institutes of Health (RO1 GM-38656). S. M. Fu is supported by a grant from the National Institutes of Health (RO1 CA-34546).

Address correspondence to G. W. Litman, Department of Pediatrics, University of South Florida, All Children's Hospital, 801 Sixth Street South, St. Petersburg, FL 33701.

Received for publication 16 March 1993 and in revised form 19 May 1993.

References

1. Yamada, M., R. Wasserman, B.A. Reichard, S. Shane, A.J. Caton, and G. Rovera. 1991. Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes. *J. Exp. Med.* 173:395.
2. Sanz, I. 1991. Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions. *J. Immunol.* 147:1720.
3. Huang, C., A.K. Stewart, R.S. Schwartz, and B.D. Stollar. 1992. Immunoglobulin heavy chain gene expression in peripheral blood B lymphocytes. *J. Clin. Invest.* 89:1331.
4. Schroeder, H.W., J.L. Hillson, and R.M. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science (Wash. DC).* 238:791.
5. Schroeder, H.W., and J.Y. Wang. 1990. Preferential utilization of conserved immunoglobulin heavy chain variable gene segments during human fetal life. *Proc. Natl. Acad. Sci. USA.* 87:6146.
6. Mortari, F., J.A. Newton, J.Y. Wang, and H.W. Schroeder. 1992. The human cord blood antibody repertoire. Frequent usage of the VH7 gene family. *Eur. J. Immunol.* 22:241.
7. Hillson, J.L., I.R. Oppliger, E.H. Sasso, E.C.B. Milner, and

M.H. Wener. 1992. Emerging human B cell repertoire: influence of developmental stage and interindividual variation. *J. Immunol.* 149:3741.
8. Pascual, V., and J.D. Capra. 1991. Human immunoglobulin heavy-chain variable region genes: organization, polymorphism, and expression. *Adv. Immunol.* 49:1.
9. Matsuda, F., E.K. Shin, H. Nagoaka, R. Matsumura, M. Haino, Y. Fukita, S. Taka-ishi, T. Imai, J.H. Riley, R. Anand, E. Soeda, and T. Honjo. 1993. Structure and physical mapping of 64 variable segments in the 3′ 0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nature Genetics.* 3:88.
10. Tomlinson, I.M., G. Walter, J.D. Marks, M.B. Llewelyn, and G. Winter. 1992. The repertoire of human germline V_H sequences reveals about fifty groups of V_H segments with different hypervariable loops. *J. Mol. Biol.* 227:776.
11. Anker, R., M.E. Conley, and B.A. Pollok. 1989. Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia. *J. Exp. Med.* 169:2109.
12. Timmers, E., M. Kenter, A. Thompson, M.E.M. Kraakman, J.E. Berman, F.W. Alt, and R.K.B. Schuurman. 1991. Diversity of immunoglobulin heavy chain gene segment rearrangement in B lymphoblastoid cell lines from X-linked agammaglobulinemia patients. *Eur. J. Immunol.* 21:2355.
13. Mortari, F., H.D. Ochs, R.J.P. Wedgwood, and H.W. Schroeder. 1991. Immunoglobulin variable heavy chain cDNA sequence from a patient with X-linked agammaglobulinemia. *Nucleic Acids Res.* 19:673.
14. Braun, J., L. Berberian, L. King, I. Sanz, and H.L. Govan III. 1992. Restricted use of fetal VH3 immunoglobulin genes by unselected B cells in the adult. *J. Clin. Invest.* 89:1395.
15. Schwarz, K., T.E. Hansen-Hagge, C. Knobloch, W. Friedrich, E. Kleihauer, and C.R. Bartram. 1991. Severe combined immunodeficiency (SCID) in man: B cell-negative (B^-) SCID patients exhibit an irregular recombination pattern at the J_H locus. *J. Exp. Med.* 174:1039.
16. Ibegbu, C.C., A.J. Nahmias, T.J. Spira, D.J. Stoll, B. Jones, N. Symbas, S. Nesheim, H. Mendez, H. Keyserling, and F.K. Lee. 1992. $CD5^+$ B cells in normal newborns and infants, and in those with HIV and intrauterine infections. *Ann. NY Acad. Sci.* 651:572.
17. Innis, M.A., D.H. Gelfand, J.J. Snisky, and T.J. White. 1990. PCR Protocols. Academic Press, Inc., San Diego. 482 pp.
18. Schroeder, H.W., J.L. Hillson, and R.M. Perlmutter. 1990. Structure and evolution of mammalian V_H families. *Int. Immunol.* 20:41.
19. Loh, E.Y., J.F. Elliott, S. Cwirla, L.L. Lanier, and M.M. Davis. 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor δ chain. *Science (Wash. DC).* 243:217.
20. Berman, J.E., S.J. Mellis, R. Pollock, C.L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C.R. Cantor, and F.W. Alt. 1988. Content and organization of the human Ig V_H locus: definition of three new V_H families and linkage to the Ig C_H locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:727.
21. Sanz, I., P. Kelly, C. Williams, S. Scholl, P. Tucker, and J.D. Capra. 1989. The smaller human V_H gene families display remarkably little polymorphism. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3741.
22. van Es, J.H., M. Heutink, H. Aanstoot, and T. Logtenberg. 1992. Sequence analysis of members of the human Ig VH4 gene family derived from a single V_H locus. *J. Immunol.* 149: 492.
23. Kodaira, M., T. Kinashi, I. Umemura, F. Matsuda, T. Noma, Y. Ono, and T. Honjo. 1986. Organization and evolution of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* 190:529.
24. Shin, E.K., F. Matsuda, H. Nagaoka, Y. Fukita, T. Imai, K. Yokoyama, E. Soeda, and T. Honjo. 1991. Physical map of the 3′ region of the human immunoglobulin heavy chain locus: clustering of autoantibody-related variable segments in one haplotype. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3641.
25. Reth, M.G., and F.W. Alt. 1984. Novel immunoglobulin heavy chains are produced from DJ_H gene segment rearrangements in lymphoid cells. *Nature (Lond.).* 312:418.
26. Gu, H., D. Kitamura, and K. Rajewsky. 1991. B cell development regulated by gene rearrangement: arrest of maturation by membrane-bound Dμ protein and selection of D_H element reading frames. *Cell.* 65:47.
27. Herzenberg, L.A., and A.B. Kantor. 1993. B-cell lineages exist in the mouse. *Immunol. Today.* 79:79.
28. Haughton, G., L.W. Arnold, A.C. Whitmore, and S.H. Clarke. 1993. B-1 cells are made, not born. *Immunol. Today.* 79:84.
29. H.H. Wortis. 1992. Surface markers, heavy chain sequences and B cell lineages. *Int. Rev. Immunol.* 8:235.
30. T.J. Kipps. 1989. The CD5 B cell. *Adv. Immunol.* 47:117.
31. Blackwell, T.K., and F.W. Alt. 1989. Mechanism and developmental program of immunoglobulin gene rearrangement in mammals. *Ann. Rev. Genet.* 23:605.
32. Sherr, D.H., M.R. Szewczuk, and G.W. Siskind. 1978. Ontogeny of B-lymphocyte function. V. Thymus cell involvement in the functional maturation of B-lymphocytes from fetal mice transferred into adult irradiated hosts. *J. Exp. Med.* 147:196.
33. Szewczuk, M.R., D.H. Sherr, and G.W. Siskind. 1978. Ontogeny of B-lymphocyte function. VI. Ontogeny of thymus cell capacity to facilitate the functional maturation of B lymphocytes. *Eur. J. Immunol.* 8:370.
34. Goidl, E.A., and G.W. Siskind. 1974. Ontogeny of B-lymphocyte function. I. Restricted heterogeneity of the antibody response of B lymphocytes from neonatal and fetal mice. *J. Exp. Med.* 140:1285.
35. Olee, T., P.-M. Yang, K.A. Siminovitch, N.J. Olsen, J. Hillson, J. Wu, F. Kozin, D.A. Carson, and P.P. Chen. 1991. Molecular basis of an autoantibody-associated restriction fragment length polymorphism that confers susceptibility to autoimmune diseases. *J. Clin. Invest.* 88:193.
36. Kabat, E.A., T.T. Wu, C. Foeller, H.M. Perry, and K. Gottesman. 1991. Sequences of Proteins of Immunological Interest. U.S. Dept. Health and Human Services, Washington, D.C.