

# The Complete Nucleotide Sequences of the Heavy Chain Variable Regions of Six Monospecific Rheumatoid Factors Derived from Epstein-Barr Virus-transformed B Cells Isolated from the Synovial Tissue of Patients with Rheumatoid Arthritis

## Further Evidence That Some Autoantibodies Are Unmutated Copies of Germ Line Genes

Virginia Pascual,\* Ingrid Randen,<sup>‡</sup> Keith Thompson,<sup>‡</sup> Mouldy Sioud,<sup>‡</sup> Oystein Forre,<sup>‡</sup> Jacob Natvig,<sup>‡</sup> and J. Donald Capra\*

\*Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9048; <sup>‡</sup>University of Oslo, Institute of Immunology and Rheumatology and the Rheumatism Hospital, Oslo 1, Norway; and <sup>§</sup>Medical Research Council, MIP Unit and Department of Immunology; Institute of Animal Physiology and Genetics; Cambridge CB2 4AT, United Kingdom

### Abstract

Structural studies of human monoclonal rheumatoid factors (RF) derived from patients with monoclonal gammopathies have revealed a restriction in the usage of heavy and light chain variable regions. These studies have suggested that germline genes with little if any somatic mutation can generate RF specificity. However, there is no information presently available regarding the primary structure and genetic origin of RF in rheumatoid arthritis. In this study, we have isolated and sequenced the VH regions of six monoclonal RF derived from the synovial membranes of two patients with rheumatoid arthritis and one with the juvenile polyarticular form of the disease. We found the same VH families as previously reported among monoclonal paraproteins with RF activity. However, our sample was diverse regarding the VH, DH, and JH gene segments used. Among VHI RF there was conservation in the length of CDRIII as well as restriction in the amino acid generated at the V-D junction, as opposed to VHIII RF and non-RF VHI molecules that are highly heterogeneous in these two aspects. We also found that different JH gene segments may contribute to RF specificity. The VH, DH, and JH elements of one RF in our study all had clearly identifiable germline counterparts. This RF displays a nearly germline configuration throughout its entire heavy chain and represents another example of an autoantibody encoded by one of the VH gene segments from the preimmune fetal repertoire. (*J. Clin. Invest.* 1990. 86:1320-1328.) Key words: human autoantibody • rheumatoid factors

### Introduction

Rheumatoid factors (RFs)<sup>1</sup> are autoantibodies that react with epitopes in the Fc region of IgG molecules (1). Even though RF activity has been found distributed among all classes of secreted immunoglobulins, IgM RF are the most frequently detected in serum by standard techniques. RF are found in normals, in acute or chronic infections, and after certain immunizations (1-3). Furthermore, the B cell precursor fre-

quency for these autoantibodies is quite significant especially when specific subsets like CD5<sup>+</sup> B cells are considered (4, 5). This has led several investigators to propose a physiological role for RF in the clearance of circulating immune complexes. RF are also present in autoimmune and/or hyperglobulinemic states, although in most situations it is not clear that they play a pathogenic role. In rheumatoid arthritis (RA), these autoantibodies are produced by peripheral and synovial B cells (6-8), and a patient's RF titer, especially when the RF is an IgG, generally correlates with systemic manifestations like rheumatoid nodules and vasculitis and with a more aggressive, erosive articular disease (6, 9). These findings support the view that RF in RA may be structurally unique or under different regulatory control.

Most of our current knowledge concerning the structure and genetic origin of human RF derives from studies on monoclonal IgM paraproteins isolated from patients with B cell dyscrasias, particularly mixed cryoglobulinemia (10). Based on cross-reactive idiotypes (CRI), these RF were initially classified into three different groups: Wa, Po, and Bla, the last being specific for antibodies with RF and anti-DNA histone activity (11). Early amino acid sequence studies revealed a marked restriction among the light chains used by these molecules (12). For example, the light chains of the Wa group have been shown by several groups to belong to the VKIIIb subfamily and are the conserved product of a germline VK gene, HumKv 325 (13, 14). Another group of monoclonal RF uses VKIIIa light chains and bear a CRI identified by the monoclonal antibody 6B6.6 (15, 16) and are thought to derive from the germline HumKv 328 gene (17). The heavy chains of these monoclonal paraproteins have been extensively studied either by amino acid sequence or serologically by using antibodies directed against structural determinants specific for the human VH families (18-21). Some of these studies have revealed a striking pattern of combinatorial association between RF light and heavy chains. While most of the VKIIIb light chains combine with members of the VHI family that express the CRI identified by the monoclonal antibody G6 (22), the VKIIIa subgroup is preferentially associated with members of the VHIV family (23).

The relationship between the RF in patients with mixed cryoglobulinemia and the RF present in patients with RA was first addressed by Forre et al. (24). They reported that although both Wa- and Po-like CRI patterns were present among the polyclonal RF of RA patients, additional CRI were also found. Recently, Mageed et al. (25) demonstrated that only a small fraction of the polyclonal population of RF present in the sera of RA patients reacts with mAb specific for the two previously

Received for publication 27 December 1990 and in revised form 2 April 1990.

1. Abbreviations used in this paper: CRI, cross-reactive idiotypes; RA, rheumatoid arthritis; RF, rheumatoid factor.

*J. Clin. Invest.*

© The American Society for Clinical Investigation, Inc.

0021-9738/90/10/1320/09 \$2.00

Volume 86, October 1990, 1320-1328

described light chains present in monoclonal paraproteins. These findings imply either that these two groups of RF have different genetic origins, or, as has been shown in murine models of autoimmunity (26), RF in autoimmune disorders undergo extensive somatic diversification.

Structural analyses of RF in RA have been limited by the difficulty in isolating monoclonal RF from these patients. Recently, as part of an ongoing project attempting to understand the genetic basis of RF specificity and its implications in the pathogenesis of autoimmunity, we have obtained monoclonal RF from the synovial tissue of patients with RA (27). Since this is the central site of pathological activity in this disease and is known to be a rich source of RF-producing B cells (7, 8, 28, 29), we reasoned that monoclonal RF derived from synovial B lymphocytes would be particularly relevant to understanding the genetic origin of the autoantibody response in this disorder.

To this end, we have studied the molecular structures of the heavy chains of six monoclonal IgM RF derived from the synovial tissue of two patients with RA and one with the polyarticular form of juvenile rheumatoid arthritis (JRA). We found considerable heterogeneity among the VH, DH, and JH gene segments expressed in these molecules compared to the RF-M components. We confirm the previously reported generation of a glutamic acid residue at the V-D junction in RF belonging to the VHI family. Finally, while some of the molecules we studied could have arisen by somatic mutation, remarkably one RF is a nearly direct copy of germ line VH-DH-DH-JH rearrangement. Additionally, this RF uses a VH gene segment found expressed early in ontogeny (30). These data argue that at least some human antibodies do not require somatic mutation in their heavy chains to become RF.

## Methods

**Generation of monoclonal RF.** The six RF were isolated from EBV-transformed B lymphocytes obtained from the inflammatory synovial tissue of two patients with RA (S.J. and T.S.) and one patient (K.L.) with the juvenile polyarticular form of the disease (JRA). RF-SJ1 and RF-SJ2 are from patient S.J. RF-TS1, RF-TS2, and RF-TS3 are from patient T.S. RF-KL1 was obtained from patient K.L. The EBV-transformed cell lines were fused with a murine myeloma (X63-Ag8.653), and the resulting hybridomas were cloned by limiting dilution. The characterization of the cell lines has been described elsewhere (27).

**Slot blot RNA hybridization.** Total RNA was extracted from  $10^6$  cells according to Chirgwin et al. (31). After denaturation with 12.5% formaldehyde (Aldrich Chemical Co., Milwaukee, WI), 10  $\mu$ g was transferred to nitrocellulose membranes in  $10\times$  SSC, dried under vacuum and hybridized to hexamer-extended DNA probes corresponding to the six known human VH families. Filters were subsequently washed at 65°C in  $0.1\times$  SSC, 0.1% SDS, and exposed overnight on Kodak XAR film with amplifying screens.

**Cloning the productively rearranged VH genes.** 20  $\mu$ g of total RNA was reverse transcribed into first strand cDNA using oligo d(T) priming in a 50- $\mu$ l reaction (32). Second strand synthesis and amplification were carried out via the polymerase chain reaction method on 1  $\mu$ l of the ss-cDNA reaction in a final volume of 100  $\mu$ l containing 200  $\mu$ M of each dNTP, 50 mM KCl, 10 mM Tris-HCl, pH 8.3 at 37°C, 1.5 mM  $MgCl_2$ , 2.5 U of Taq polymerase, and 50 pmols of primers (33). The primers consisted of an 18 mer from the IgM constant region priming the 3' end (5'TGGGGCGGATGCACTCCC 3') and oligonucleotides specific for the 5' end sequences of the VHI and VHIII families (VHI-leader 5'ATGGACTGGACCTGGAGGATC3', VHI-5'AGGTGCCCACTCCCAGGTGCAGCT3', VHIII-leader 5'CTCACCATG-

GAGTTTGGGCTG3'). Amplification of the RF-SJ2 germline gene was done using primers derived from the leader and the 23 nucleotide spacer sequences of the previously published 3005 and hv1.9III genes (3005-leader 5'GGGTTTTCTCGTTGCTCTT3', 3005-sp 5'CTGGGCGCACAATGACTTC3') (34, 35). 40 cycles of amplification were then performed using the following conditions: denaturation step (94°C for 1 min), annealing (55°C for 3 min), and extension (72°C for 3 min). The PCR products were purified with low melt agarose and blunt end ligated into a dephosphorylated, Eco RV-digested plasmid (pBluescript KS<sup>+</sup>). The ligation mixture was used to transform BSJ-72 competent cells. The colonies were screened with internal oligomers whose sequences were derived from the specific framework or hypervariable regions of the previously determined VH families as well as the JH consensus sequence oligomer end-labeled with 32P-gamma-ATP and polynucleotide kinase. Single-stranded DNA was prepared from double positive colonies and sequenced by the dideoxy chain termination procedure (36) using 35-S-alpha-thio ATP and Sequenase (37).

**Direct sequencing of rearranged VH genes.** Two of the clones, RF-TS2 and RF-KL1, were amplified after the synthesis of first strand cDNA following the method described by Gyllenstein and Erlich (38), except for a primer ratio of 100/1 pmols. PCR samples were diluted to 2 ml with 50 mM Tris (pH 8) and purified by spinning at 4,000 rpm for 30 min in Centricon-30 (Amicon Corp., Danvers, MA) microconcentrators. 40–60  $\mu$ l of retentate were dried down, resuspended in 7  $\mu$ l of water, and sequenced with either internal oligonucleotides or the primers used for amplification by the dideoxy chain termination procedure. Alternatively, 1  $\mu$ l of the PCR mixture was reamplified without further purification using 50 pmols of one of the primers and directly sequenced.

**Southern filter hybridization.** Genomic DNA was prepared from the hybridomas, peripheral blood cells as well as from normal liver. DNA was digested with restriction endonucleases, separated on 0.6% agarose gels, and transferred by the method of Southern (39). Hybridization with hexamer-extended probes was accomplished, and the filters were washed in  $0.1\times$  SSC, 0.1% SDS at 65°C with overnight exposure on Kodak XAR films using amplifying screens.

## Results

**Characterization of the human synovial rheumatoid factors and determination of the VH gene family.** The six antibodies studied are monoclonal IgM with RF activity. While RF-TS1, RF-TS3, RF-SJ2, and RF-KL1 react only with human IgG, RF-SJ1 and RF-TS2 also recognize rabbit IgG. The reactivities of these antibodies with different epitopes on the Fc region of IgG were determined by human gamma globulin ELISA using IgG myeloma proteins of known subclass and Gm allotype. As it is shown in Table I and has been published elsewhere (27), the six antibodies can be assigned to three different groups: (a) pan specific (RF-SJ1 and RF-TS1); (b) classical Ga (RF-TS2, RF-TS3 and RF-SJ2); (c) and new Ga-related (RF-KL1). The incidence of RF related cross-reactive idiotypes among these molecules (Table I) has also been published elsewhere (40).

The distribution of VH families and JH segments among the monoclonal RF is shown in Table I. When cytoplasmic RNA from the RF-producing hybridomas was hybridized with VH probes specific for the six human VH families, only VHI and VHIII gave positive results. These results were confirmed by amplifying a fragment of the expected size with leader sequence specific primers and by sequencing the VH regions (see below).

**Two RF from the same patient use different VHI gene segments.** RF-TS1 is a classical Wa RF. It bears the antigenic determinants defined by the VKIIIb CRI-specific mAb 17.109

Table I. Source and Characterization of Six Monoreactive Rheumatoid Factors

	Isotype	Specificity	Reactivity with monoclonal antibodies								JH
			C7	C6	17.109	6B6.6	G6	D12	B6	VH family	
RF-TS1	IgM, k	PAN	+	+	+	—	+	—	—	VHI	JH3
RF-TS3	IgM, k	Ga	—	—	—	—	—	—	—	VHI	JH4
RF-KL1	IgM, k	New Ga	—	—	—	—	—	—	—	VHIII	JH4
RF-TS2*	IgM, k	Ga	—	—	—	—	—	—	—	VHIII	JH3
RF-SJ2	IgM, λ	Ga	—	—	—	—	—	+	+	VHIII	JH6
RF-SJ1*	IgM, λ	PAN	—	—	—	—	—	+	+	VHIII	JH6

Specificities: PAN, all IgG subclasses and allotypes. Ga (classical), IgG1, IgG2, IgG4. New Ga-related, IgG1, IgG2, IgG4, and G3m (st). Light chain associated CRI, C7 (VKIII), C6 (VKIIIb), 17.109 (VKIIIb), and 6B6.6 (VKIIIa). Heavy chain associated CRI, G6 (VHI), D12, and B6 (VHIII). \* RF-TS2 and RF-SJ1 also react with rabbit IgG.

as well as the VHI-related CRI identified by the mAb G6 (Table I). Several G6(+) heavy chains have been sequenced at the amino acid and/or nucleotide level, and the striking homology among them suggests a common genetic origin (41, 42). We found that RF-TS1 shares 94.5% identity with the germline gene hv1263, and 95.8% with 51P1, a VH gene isolated from a fetal liver cDNA library (Fig. 1) (30). Since we cannot be certain of the germline counterpart of RF-TS1, we cannot determine whether somatic mutation plays a role in the generation of its RF specificity.

Unlike RF-TS1, RF-TS3 (the other VHI-expressing RF) did not react with any of the idiotype-specific mab tested (Table I). Consistent with this lack of known serologic markers, the nucleotide and deduced amino acid sequence of

RF-TS3 is not particularly similar to previously sequenced G6(+) VHI RF nor to any of the known germline VHI genes (data not shown). However, this gene segment is 95% homologous at the nucleotide level to the VHI gene segment expressed in Ab47, a natural bireactive antibody with RF and anti-ssDNA specificities (43). At the amino acid level, they differ by only five amino acid residues, three of which are concentrated in CDR1, one in CDR2 and another one in the third framework (Figs. 1 and 2). Interestingly, the three substitutions in CDR1 have been found in other germline VH genes, expressed VH genes, and even in pseudogenes. Even more impressive is

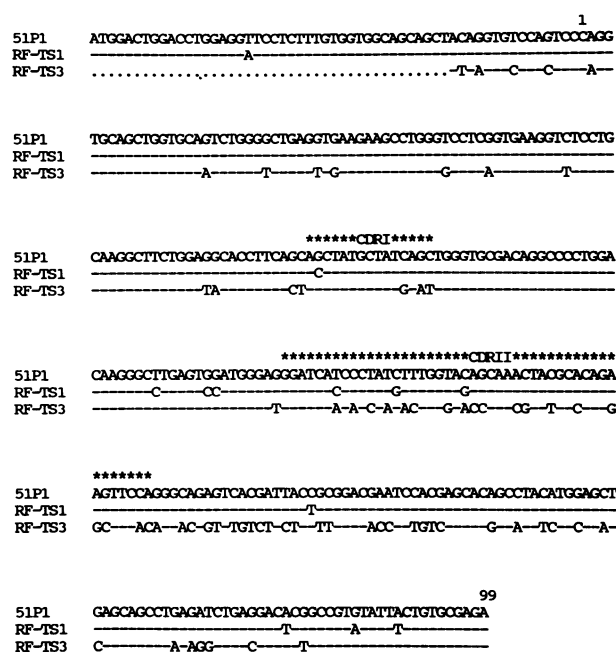


Figure 1. Nucleotide sequences of the VH segments of VHI RF. The top sequence, 51P1, is a cDNA from a fetal liver (30) and is used for comparison as it represents an example of G6(+) VHI gene segment (42). Homologies are indicated with dashes. Deletions or absence of sequence are depicted by dots.

## A

	1	CDR1	50
BOR	EVQLVQSGAE	VKKPGSSVKV	TCRASGDTFS
KAS	—H—	—S—C—	—Y—
SIE	—	—R—T—G—	—GYT—
WOL	—M—	—R—S—T—G—V	—DYKGL—
RF-TS1	L—	—S—	—T—
RF-TS3	Q—S—L—A—	—S—Y—T—	—Y—MN—
Ab47	Q—S—L—A—	—S—Y—	—H—N—

## B

	CDR2	98
BOR	GLIPIFGITPNY	AQKFGGRVTI
KAS	—QA—	—A—N—
SIE	SPAKWIDPFGQ	VYIKWE—V
WOL	Q—PLR—NGEVK	NRGSVV—SV
RF-TS1	—T—M—RA—	—S—V—F—
RF-TS3	W—NINT—N—T—	—G—T—FVF—
Ab47	WININT—N—T—	—G—T—FVF—

## C

	CDR3	100
BOR	EGRRMAINP	FDYWGQGITLVSS
KAS	—YGDYGR—	—F—
SIE	—WKGQVN—	—V—
WOL	—YGFDTSDY	—YY—
RF-TS1	—DPYGDYVNP	—I—M—
RF-TS3	—DSNGYKI	—D—F—

Figure 2. (a) Comparison of the amino acid sequences of the VH regions of four previously reported VHI RF derived from patients with monoclonal gammopathies (BOR, KAS, SIE, and WOL) (19, 20) and the translated amino acid sequences of the two VHI RF derived from RA patients. Ab47 is a natural autoantibody with RF and anti-ssDNA activity (43). (b) Comparison of the amino acid sequences of the CDR3 regions.

the fact that the RF-TS3 CDRI is found as a block in other VHI and VHIII genes (44). This raises the possibility that gene conversion events might contribute to this structure at either the germline or the somatic level.

*Four RF derived from three different patients use different VHIII gene segments.* The VH gene segments of the other four monoclonal RF belong to the VHIII family. RF-SJ2 shows 98.5% identity to a germline VHIII gene, humvh 3005 (34), and 99.5% identity to the VH gene segment of 56P1, a VH region found expressed in the same fetal cDNA library as 51P1 (Fig. 3 a) (30). At the amino acid level, RF-SJ2 and 56P1 are 99% homologous with only one difference in FW3 (Fig. 3 b). The near identity of 56P1 and RF-SJ2 is striking as they were isolated from unrelated humans. The VHIII family has been shown to contain several highly polymorphic members, the VH locus defined by humvh 3005 and/or 56P1 being one of them (45). Using primers specific for the leader and 23 nucleotide spacer region of the 3005 gene, the allelic counterpart of this gene was amplified and sequenced in patient S.J. The sequence, shown in Fig. 4, was 100% identical to the 56P1 VH gene segment (30), implying that either RF-SJ2 has accumulated two somatic mutations in its FW3 or that we have sequenced the allelic variant of the actual donor gene for RF-SJ2. Studies are presently underway to fully characterize these genes. RF-SJ1 derives from the same patient as RF-SJ2. They share 94.6% homology in VH and use very similar DH segments rearranged to JH6.

Furthermore, the V-D and D-J joints are identical in both clones (Fig. 5). The similarity of RF-SJ1 and RF-SJ2 suggests that if they are clonally related, RF-SJ1 represents an example of somatic mutation plus antigenic selection. However, the fact that the closest germ line gene to RF-SJ1 is not humvh 3005 but humvh 1.9III (95.7% identity) (35) argues that RF-SJ1 has a different genetic origin than RF-SJ2 (Fig. 3 a). Alternatively, VH gene replacement with a more distantly related VHIII gene could explain this structure. Indeed, humvh 3005 is an example of a VH gene segment that contains a recombination sequence (heptamer-like) at its 3' end. To address this question, we performed Southern analysis on the cell line DNA using J<sub>H</sub> and IgM switching region probes that confirmed the relatedness of these two clones (data not shown). The fine specificity of RF-SJ1 is different from RF-SJ2 in that RF-SJ1 also recognizes rabbit IgG while RF-SJ2 does not (Table I). The only other RF that reacts with rabbit IgG is RF-TS2 and both RF-TS2 and RF-SJ1 share a glycine substitution in CDRI (Fig. 5).

The closest germ line VH gene segment to RF-TS2 is also humvh 1.9III (97.5% identity at the nucleotide level) (Fig. 3). This germ line gene has been found expressed early in ontogeny (FL2.2) (46) and in the repertoire of a normal adult as an IgM with anti-dsDNA specificity (Kim 4.6) (47). Out of seven nucleotide differences between RF-TS2 and humvh 1.9III, three are concentrated in the CDRII, one in the FW1 and three in the FW3. Two nucleotide substitutions in both the CDRII and FW3 encode amino acid replacements (Fig. 3 c), accounting for a replacement/silent substitution rate of 1.3 in these two regions. This finding suggests that RF-TS2 does not arise from humvh 1.9III, especially if we consider that on a random basis 75% of all substitutions should be replacements. On the other hand, we cannot rule out that RF-SJ1 and RF-TS2 represent allelic variants of the humvh 1.9III germ line gene.

The VH gene segment expressed in RF-KL1 shows 95.1%

identity to VH26 (48) (Fig. 3 d), another example of a VHIII gene expressed in the fetal repertoire (30P1) (30) as well as in an autoantibody with anti-dsDNA activity (18-2), derived in this case from a patient with SLE (49). At the nucleotide level they differ by 17 nucleotides, 14 of which are clustered in CDRI and CDRII. None of the five differences found in CDRI encodes amino acid substitutions and only five out of nine are responsible for amino acid changes in CDRII (R/S = 1.25) (Fig. 3 d). All of these data support the idea that RF-KL1 derives from a germline VH gene segment different from VH26 and preclude our determining the germline versus somatic contribution to this molecule. On the other hand, RF-KL1 is the only VHIII RF in this study that does not express the VHIII-related idiotypes identified by the mab B6 and D12 (40). Substantial differences in FW1, CDRI, and CDRIII compared to the other three molecules could account for this finding (Fig. 5).

*CDRIII regions are heterogeneous among immunoglobulins with RF activity.* As Fig. 5 shows, with the exception of the VHIII RF-SJ2 and RF-SJ1, there is no obvious sequence homology among the D segments expressed in these molecules. RF-SJ2 and RF-SJ1 seem to derive from the germ line DRL-2 gene (50) with either N segment addition at the 5' end or D-D fusion to the germline D gene DHFL.16 (51, 52) (Fig. 4). The D segment of RF-SJ1 differs by nine nucleotides from RF-SJ2, four of which are responsible for amino acid substitutions (Fig. 5). Whether this represents mutation or derivation from different germline genes cannot be answered at this time. Among the known human D segments there are no genomic counterparts for the D segments found in RF-TS3, RF-TS1, RF-TS2 and RF-KL1 (Fig. 6) (50, 51, 53).

Two findings within the CDRIII region of the VHI RF sequenced in this study confirm a previous report from our laboratory (20), one being the conservation in length of this particular region, ranging from 11 to 14 amino acid residues, and the other the V-D junctional glutamic acid (Fig. 2). Considering the randomness of V-D rearrangement, this amino acid is probably important in conferring RF specificity.

*JH usage among RF.* JH usage does not appear to be restricted among VHIII RF. RF-SJ2 and RF-SJ1 use JH6 gene segments without any 5' deletion compared to its reported germ line counterpart (54). RF-SJ1 has a histidine instead of the fifth tyrosine residue. Considering that these two molecules derive from the same patient, this is probably due to somatic mutation although expression of the other allele has not been ruled out. On the other hand, both molecules have at position 6 a tyrosine replacing the glycine present in the reported germ line JH6 gene segment and a lysine instead of the glutamine at position 12 as well as a silent nucleotide substitution (54) (Fig. 7). Using oligonucleotides specific for these differences, we have screened genomic JH6 segments from patient S.J. as well as from normals. We found that the JH6 gene segment expressed in RF-SJ2 is identical to one of the JH6 germline alleles in patient S.J. Indeed, the variation seen in JH6 represents a polymorphism in the general population (data not shown).

RF-TS2 and RF-TS1 use JH3 gene segments and RF-TS3 and RF-KL1 JH4 genes with one and two amino acid substitutions respectively. RF-TS1 is the first example of a Wa (+) RF that does not use a JH4 gene segment. Since RF-TS1 is reactive with the mab G6, these data argue that G6 reactivity must be located within the VH segment.



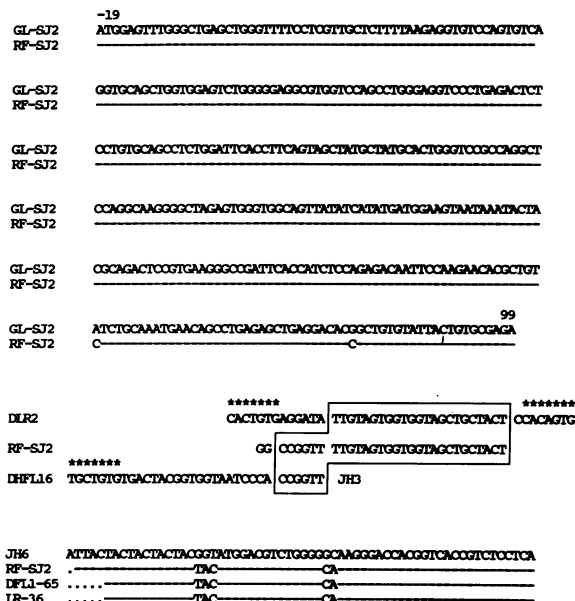


Figure 4. Comparison of the nucleotide sequences of the VH, DH, and JH gene segments of RF-SJ2 and its germ line counterparts. GL-SJ2 was isolated from patient S.J. DLR2 is a germ line D segment (54). DHFL16 is a rearranged D segment that maintains its 5' germline configuration (51). DFL1-65 is a fetal cDNA sequence (I. Sanz, personal communication). LR-36 is an unproductive rearrangement isolated from a leukemia cell line (50).

nature and pathogenic potential of the immune complexes formed is not clear. (b) Idiotypic analyses of sera and synovial fluid from patients with RA have shown that their RF carry a

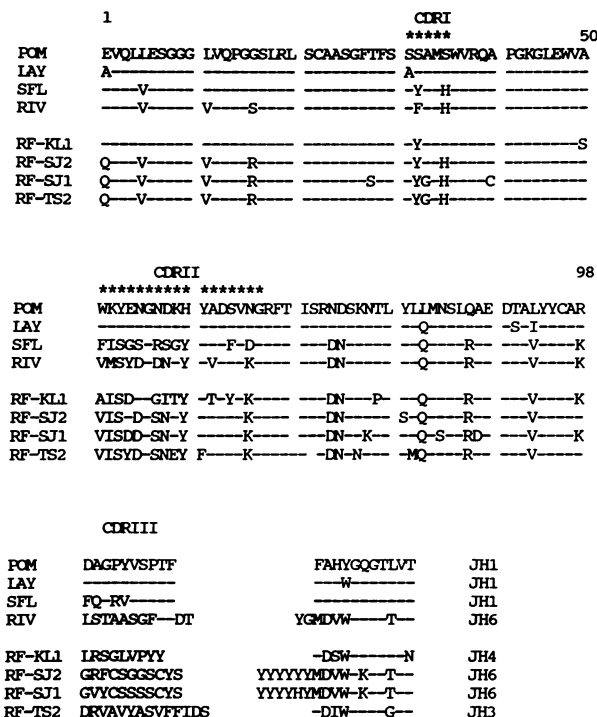


Figure 5. Comparison of the amino acid sequences of the heavy chain of four previously reported VHIII RF (18) and the deduced amino acid sequences of four RF from patients with RA.

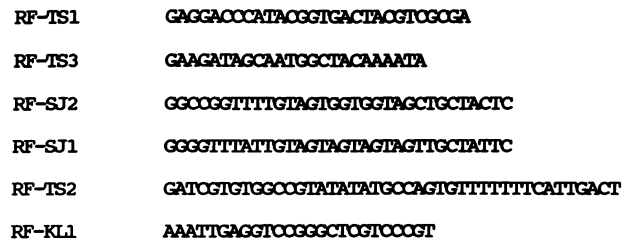


Figure 6. Nucleotide sequence of the D segments expressed in the six monoclonal RF from patients with RA.

private repertoire of idiotypes (59, 60) and that only a small percentage bear the antigenic determinants present in other autoimmune disorders like Sjogren's syndrome or in the monoclonal paraproteins from patients with mixed cryoglobulinemia or Waldstrom's macroglobulinemia (24, 25). Considering the proven germline origin of this latter group of idiotypic markers (13, 17), RF in RA might be the product of a somatically diversified, antigenically selected response.

With regard to the structural characterization of these RA-specific RF (as opposed to lymphoid malignancies-derived RF), our first question addressed their VH family use. We found that the six VH gene segments studied belong to the VHI and VHIII families. This distribution of families agrees with the distribution previously reported among monoclonal paraproteins with RF activity (18-22). Natvig et al. (61) studied, by partial amino acid sequence, a monoclonal RF from a patient with RA that also used a VH segment belonging to the VHIII family. The VHI and VHIII families include from 50 to more than 100 genes (35). Since they are the largest and most heterogeneous families within the human VH locus, the possibility of VH family restriction among this particular set of RF is difficult to evaluate. If a VH family restriction does exist, it could be explained in several ways.

The VH region itself could generate antigenic specificity. Alternatively, as has been proposed for CD5 (+) acute and chronic lymphocytic leukemias, different cellular subsets might preferentially express different VH families (62). A third possibility is that we have introduced an experimental bias by EBV transformation. We consider this possibility less likely because a preferential usage of particular families has been

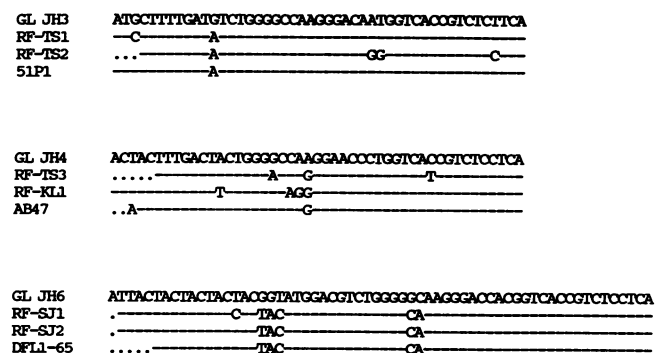


Figure 7. JH segments expressed in the six monoreactive RF. The sequences are compared to the germ line genes (top line) and to expressed JH segments (bottom line) to identify possible polymorphisms. 51P1 is a cDNA from a fetal liver (30). AB47 is a natural bireactive autoantibody (43). DFL1-65 is a fetal liver cDNA (I. Sanz, personal communication).

found in other autoantibody responses like the cold agglutinins and natural polyreactive antibodies, which are also derived from EBV transformed lines, and preferentially rearrange members of the small VHIV family (23, 43).

Monoclonal paraproteins with RF activity often use members of the VHI family recognized by the mAb G6 (22). This phenotypic marker also reacts with the immunoglobulins secreted by the malignant CD5<sup>+</sup> cells of patients with CLL (42). While some of the reported G6<sup>+</sup> molecules are 99% identical in their VH regions to 51P1, a VH gene found expressed in the fetal repertoire (30), others display only 92% similarity, suggesting that mAb G6 identifies the product of a subset of VHI gene products with a certain degree of heterogeneity, either at the germline level or as the result of somatic diversification. In our sample, two out of six RF (RF-TS1 and RF-TS3) use VHI gene segments. We were unable to ascertain their genetic origin since the closest homology was 95% with two expressed VH gene segments (51P1 and Ab47 respectively). On the other hand RF-TS1, a G6<sup>+</sup> molecule, and RF-TS3 are < 90% homologous at the nucleotide level. These findings suggest that at least two different subsets of VHI genes are involved in the generation of RF activity.

The other four RF that we studied use gene segments from the VHIII family. Three of them (RF-SJ1, RF-TS2, and RF-KL1) are only 95–97% homologous to two previously reported germ line VHIII genes, precluding us from determining whether they are germ line encoded. However, the structure of RF-SJ2 is remarkable in that this molecule displays 99.5% identity in its VH segment with another of the VH genes reported by Shroeder et al. (30) (56P1) from the preimmune repertoire of a 130-d-old fetus (30) as well as to its putative germline counterpart (Fig. 4). Interestingly, the VH gene segment of RF-SJ2 is rearranged to a D segment that, except for eight nucleotides at the V-D joint, is an exact copy of the germ line DRL2 gene (Fig. 4) (50). These 8 nucleotides could represent N segment addition or they could derive from a direct D-D fusion with another D segment, DHFL16 (Fig. 4) as has recently been described by Meek et al. (52). The JH6 segment expressed in this antibody shows five nucleotide differences compared to the original published germ line counterpart (54). Only one of the substitutions is silent, the other four accounting for two amino acid replacements (tyrosine for glycine and lysine for glutamine at positions 6 and 12, respectively). We have amplified the germ line JH6 segments of two of the patients included in this study as well as several normals and probed them with oligonucleotides that recognize these differences. Our results clearly indicate that the JH6 gene segment expressed in RF-SJ2 is present in the germ line of different individuals including patient S.J. and thus represents a polymorphism rather than the result of somatic mutation. JH6 gene segments sharing the same nucleotide sequence have been found in an aberrant D-J rearrangement (54) and in a fetal cDNA (I. Sanz, personal communication). The conservation of the DLR-2 germ line gene configuration at the 3' end of the D segment expressed in RF-SJ2 allows us to conclude that rather than having deleted a glycine residue, this polymorphic gene has a tyrosine instead. RF-SJ2 is in summary an example of an antibody that uses a VH gene segment from the fetal repertoire and recognizes self while retaining almost completely its germ line structure. Several of the VH gene segments expressed early in ontogeny have already been found in either healthy adults or adults with autoimmune disease as unmutated copies of germ line genes encoding autoantibodies with

different specificities (47, 49, 63). These findings could be explained if, as some authors have proposed, the VH gene segments could by themselves generate binding specificity (64). In the sterile fetal environment, those clones that interact with self antigens in a polyreactive, low affinity manner would get expanded. The same "self-reactive" VH segments recombined with different DH, JH or light chain elements or somatically mutated could either diversify, changing the specificity (mono versus polyspecific; foreign versus self), and/or increase the affinity for self antigens. Thus, autoimmunity could be the result of a dysregulation in the normal mechanisms that control the expression of the self-reacting, high affinity antibodies.

Regarding the genetic origin of the D segments expressed in this group of RA-derived RF, only two of them (RF-SJ2 and RF-SJ1) can be easily explained on the basis of the reported DH germ line genes. Whether the rest are the product of somatic mechanisms of diversification or derive from different, not yet described, germ line D gene segments cannot be answered at this time. Considering the CDRIII regions from all the reported VHI RF, there are two remarkable findings. (a) The relative conservation of length (11–14 residues), which is neither observed in VHIII RF nor in other VHI molecules without RF activity. This length conservation probably has implications in terms of interaction with antigen; and (b) the conservation of the junctional amino acid residues. The CDRIII regions of the four previously reported VHI G6<sup>+</sup> rheumatoid factors BOR, KAS, SIE, and WOL as well as RF-TS1 contain a glutamic acid residue at the V-D and, except for WOL, a proline at the D-J joint. Even RF-TS3, which is encoded by a distantly related member of the VHI family unlike the G6<sup>+</sup> subset, contains a glutamic acid residue at the V-D joint. These findings confirm previous results from our laboratory (20) and suggest that the amino acid residue at the V-D joint is important for RF specificity. Junctional diversity has been shown to be crucial to antibody binding in several systems including the A/J Ars-A immune response through the generation, by intracodonic recombination, of an arginine residue at position 96 in the light chain (65).

We found, on the other hand, a tremendous degree of heterogeneity among the D segments expressed in our sample of RF. Only RF-TS1 and the reported paraprotein KAS share a stretch of four amino acids in the D segment. This high level of diversity could be a reflection of the variety of minor antigenic specificities that can be recognized within the Fc region of the IgG molecule (66).

These studies present the first molecular structures of RF derived from the synovial tissue B lymphocytes of patients with RA. This group of RFs use a diverse array of VH, DH and JH gene segments, and at least some of them are nearly germ line-encoded, implying that these potentially pathogenic autoantibodies might have the same genetic origin as the ones expressed in nonautoimmune situations. Additionally, some VH genes expressed in the fetal preimmune repertoire can encode RF specificity. These results further support the concept that normal and diverse genetic elements are responsible for the generation of RF activity in patients with RA.

## Acknowledgments

The authors are grateful to Kimberly Victor for helpful discussions and Steven Scholl for superb technical assistance.

J. D. Capra is supported in part by a grant from the National Institutes of Health (CA 44016). I. Randen, O. Forre, and J. B. Natvig



are supported by the Norwegian Women's Health Organization, the Norma and Leon Hess and the Grethe Harbitz Legacies.

## References

- Williams, R. C., and H. G. Kunkel. 1962. Rheumatoid factor, complement and conglutinin aberrations in patients with subacute bacterial endocarditis. *J. Clin. Invest.* 41:666-675.
- Koopman, J., and R. A. Schrohenloher. 1980. In vitro synthesis of IgM rheumatoid factor by lymphocytes from healthy donors. *J. Immunol.* 2:934-939.
- Welch, M. J., S. Fong, J. H. Vaughan, and D. A. Carson. 1983. Increased frequency of rheumatoid factor-precursor B lymphocytes after immunization of normal adults with tetanus toxoid. *Clin. Exp. Immunol.* 51:299-304.
- Hardy, R. R., K. Hayakawa, M. Shimizu, K. Yamasaki, and T. Kishimoto. 1987. Rheumatoid factor secretion from human Leu-1+ B cells. *Science (Wash. DC)*. 23:81-83.
- Casali, P., S. E. Burastero, M. Nakamura, G. Inghirami, and A. L. Notkins. 1987. Human lymphocytes making rheumatoid factor and antibody to ssDNA belong to the Leu-1+ B-cell subset. *Science (Wash. DC)*. 23:77-81.
- Panush, R. S., N. E. Bianco, and P. H. Schur. 1971. Serum and synovial fluid IgG, IgA and IgM antigammaglobulins in rheumatoid arthritis. *Arthritis Rheum.* 14:737-747.
- Natvig, J. B., and E. Munthe. 1975. Self-associating IgG rheumatoid factors represents a major response of plasma cells in rheumatoid inflammatory tissue. *Ann. NY. Acad. Sci.* 256:88-95.
- Panush, R. S., A. K. Bittner, M. Sullivan, P. Katz, and S. Longley. 1985. IgM rheumatoid factor elaboration by blood, bone marrow and synovial mononuclear cells in patients with rheumatoid arthritis. *Clin. Immunol. Immunopathol.* 34:387-391.
- Withrington, R. H., I. Teitsson, H. Valdimarsson, and M. H. Seifert. 1984. Prospective study of early rheumatoid arthritis. II. Association of rheumatoid factor isotypes with fluctuations in disease activity. *Ann. Rheum. Dis.* 43:679-685.
- Kunkel, H. G., V. Agnello, F. G. Joslin, R. J. Winchester, and J. D. Capra. 1973. Cross-idiotypic specificity among monoclonal IgM proteins with anti-gamma globulin activity. *J. Exp. Med.* 137:331-342.
- Agnello, V., A. Arbetter, G. I. de Kasep, R. Powell, E. M. Tan, and F. Joslin. 1980. Evidence for a subset of rheumatoid factors that cross-react with DNA-histone and have a distinct cross-reactive idiotype. *J. Exp. Med.* 151:1514.
- Kunkel, H. G., R. J. Winchester, F. G. Joslin, and J. D. Capra. 1974. Similarities in the light chains of anti-gamma globulins showing cross-idiotypic specificities. *J. Exp. Med.* 139:128-136.
- Radoux, V., P. P. Chen, J. A. Sorge, and D. A. Carson. 1986. A conserved human Vk gene directly encodes rheumatoid factor light chains. *J. Exp. Med.* 164:2119-2124.
- Chen, P. P., K. Albrandt, T. J. Kipps, V. Radoux, F. T. Liu, and D. A. Carson. 1987. Isolation and characterization of human kIII germ line genes: implications for the molecular basis of human kIII light chain diversity. *J. Immunol.* 139:1727-1733.
- Schrohenloher, R. E., and W. J. Koopman. 1986. An idiotype common to rheumatoid factors from patients with rheumatoid arthritis identified by a monoclonal antibody. *Arthritis Rheum.* 29:S28.
- Crowley, J. J., R. D. Goldfien, R. E. Schrohenloher, H. L. Spiegelberg, G. J. Silverman, R. A. Mageed, R. Jefferis, W. J. Koopman, D. A. Carson, and S. Fong. 1988. Incidence of three cross-reactive idiotypes on human rheumatoid factors paraproteins. *J. Immunol.* 140:3411-3418.
- Chen, P. P., D. L. Robbins, F. R. Jirik, T. J. Kipps, and D. A. Carson. 1987. Isolation and characterization of a light chain variable region gene for human rheumatoid factors. *J. Exp. Med.* 166:1900-1905.
- Capra, J. D., and J. M. Kehoe. 1974. Structure of antibodies with shared idiotype: the complete sequence of the heavy chain variable regions of two immunoglobulin M anti-gamma globulins. *Proc. Natl. Acad. Sci. USA.* 71:4032-4036.
- Andrews, D. W., and J. D. Capra. 1981. Complete amino acid sequence of variable domains from two monoclonal human anti-gamma globulins of the Wa cross-idiotypic group: suggestion that the J segments are involved in the structural correlate of the idiotype. *Proc. Natl. Acad. Sci. USA.* 78:3799-3803.
- Newkirk, M. M., R. A. Mageed, R. Jefferis, P. P. Chen, and J. D. Capra. 1987. Complete amino acid sequences of variable regions of two human IgM rheumatoid factors, BOR and KAS of the Wa idiotype family, reveal restricted use of heavy and light chain variable and joining region segments. *J. Exp. Med.* 166:550-564.
- Silverman, G. J., F. Goni, J. Fernandez, P. P. Chen, B. Frangione, and D. A. Carson. 1988. Distinct patterns of heavy chain variable region subgroup use by human monoclonal autoantibodies of different specificity. *J. Exp. Med.* 168:2361-2366.
- Mageed, R. A., M. Dearlove, D. M. Goodall, and R. Jefferis. 1986. Immunogenic and antigenic epitopes of immunoglobulins. XVII. Monoclonal anti-idiotypes to the heavy chain of human rheumatoid factors. *Rheumatol. Int.* 6:179-183.
- Silverman, G., and D. A. Carson. 1990. Structural characterization of human monoclonal cold agglutinins: evidence for a distinct primary sequence-defined VH4 idiotype. *Eur. J. Immunol.* 20:351-356.
- Forre, O., J. H. Dobloug, T. E. Michaelsen, and J. B. Natvig. 1979. Evidence of similar idiotypic determinants on different rheumatoid factor populations. *Scand. J. Immunol.* 9:281-287.
- Mageed, R. A., and R. Jefferis. 1988. Analysis of rheumatoid factor autoantibodies in patients with essential mixed cryoglobulinemia and rheumatoid arthritis. *Scand. J. Rheumatol. Suppl.* 75:172-178.
- Shlomchik, M. J., A. Marshak-Rothstein, B. B. Wolfowicz, T. L. Rothstein, and M. G. Weigert. 1987. The role of clonal selection and somatic mutation in autoimmunity. *Nature (Lond.)*. 328:805-811.
- Randen, I., K. M. Thompson, J. B. Natvig, O. Forre, and K. Waalen. 1989. Human monoclonal rheumatoid factors from the polyclonal repertoire of rheumatoid synovial tissue: production and characterization. *Clin. Exp. Immunol.* 78:13-18.
- Munthe, E., and J. B. Natvig. 1972. Complement-fixing intracellular complexes of IgG rheumatoid factor in rheumatoid plasma cells. *Scand. J. Immunol.* 1:217-229.
- Natvig, J. B., I. Randen, K. Thompson, O. Forre, and E. Munthe. 1989. The B cell system in rheumatoid inflammation. New insights into the pathogenesis of rheumatoid arthritis using synovial B cells hybridoma clones. *Springer Semin. Immunopathol.* 11:301-313.
- Schroeder, H. W., J. L. Hillson, and R. M. Perlmutter. 1987. Early restriction in the human antibody repertoire. *Science (Wash. DC)*. 238:791-793.
- Chirgwin, J. M., A. E. Przbyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294-5299.
- Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene (Amst.)*. 25:263-269.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higushi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC)*. 239:487-491.
- Chen, P. P. 1990. Hypothesis: the early expression of some autoantibody-associated heavy chain variable region genes is controlled by specific regulatory elements. *Scand. J. Immunol.* 31:673-678.
- Berman, J. E., S. J. Mellis, R. Pollock, C. L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, R. C. Cantor, and F. W. Alt. 1988. Content and organization of the human VH locus: definition of three new families and linkage to the Ig CH locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:727-738.
- Sanger, F., A. R. Carlson, B. G. Barrel, A. J. H. Smith, and B.



- Roe. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143:161-178.
37. Tabor, S., and C. C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* 84:4767-4771.
38. Gyllenstein, U. B., and H. A. Erlich. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA.* 85:7652-7656.
39. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 545 pp.
40. Thompson, K., I. Randen, J. B. Natvig, R. A. Mageed, R. Jefferis, D. A. Carson, H. Tighe, and O. Forre. 1990. Human monoclonal rheumatoid factors derived from the polyclonal repertoire of rheumatic synovial tissue: incidence of cross-reactive idiotypes. *Eur. J. Immunol.* 20:863-868.
41. Chen, P. P., M. F. Liu, C. A. Glass, S. Sinha, T. J. Kipps, and D. A. Carson. 1989. Characterization of two immunoglobulin VH genes that are homologous to human rheumatoid factors. *Arthritis Rheum.* 32:72-76.
42. Kipps, T. J., E. Tomhave, L. F. Pratt, S. Duffy, P. P. Chen, and D. A. Carson. 1989. Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 86:5913-5917.
43. Sanz, I., P. Casali, J. W. Thomas, A. L. Notkins, and J. D. Capra. 1989. Nucleotide sequences of eight human natural autoantibody VH regions reveals apparent restricted use of VH families. *J. Immunol.* 142:4054-4061.
44. Dersimonian, H., K. P. W. J. McAdam, C. Mackworth-Young, and B. D. Stollar. 1989. The recurrent expression of variable region segments in human IgM anti-DNA autoantibodies. *J. Immunol.* 142:4027-4033.
45. Chen, P. P., K. A. Siminovich, N. J. Olsen, R. A. Erger, and D. A. Carson. 1989. A highly informative probe for two polymorphic VH gene regions that contain one or more autoantibody-associated VH genes. *J. Clin. Invest.* 84:706-710.
46. Nickerson, K. G., J. Berman, E. Glickman, L. Chess, and F. W. Alt. 1989. Early human IgH gene assembly in Epstein-Barr virus-transformed fetal B cell studies. Preferential utilization of the most JH-proximal D segment (DQ52) and two unusual VH-related rearrangements. *J. Exp. Med.* 169:1391-1403.
47. Cairns, E., P. C. Kwong, V. Misener, P. Ip, D. A. Bell, and K. A. Siminovich. 1989. Analysis of variable region genes encoding a human anti-DNA antibody of normal origin. *J. Immunol.* 143:685-691.
48. Matthyssens, G., and T. H. Rabbitts. 1980. Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region. *Proc. Natl. Acad. Sci. USA.* 77:6561-6565.
49. Dersimonian, H., R. S. Schwartz, K. J. Barrett, and B. D. Stollar. 1987. Relationship of human variable region heavy chain germ line gene to genes encoding anti-DNA autoantibodies. *J. Immunol.* 139:2496-2501.
50. Siebenlist, U., J. V. Ravetch, S. Korsmeyer, T. Waldman, and P. Leder. 1981. Human immunoglobulin D segments encoded in tandem multigenic families. *Nature (Lond.)* 294:631.
51. Sonntag, D., B. Weingartner, and R. Grutzmann. 1989. A member of a novel human DH gene family: DHFL16. *Nucleic Acids Res.* 17:1267-1268.
52. Meek, K. D., C. A. Hasemann, and J. D. Capra. 1989. Novel rearrangements at the immunoglobulin D locus. Inversions and fusions add to IgH somatic diversity. *J. Exp. Med.* 170:39.
53. Ichihara, Y., H. Matsuoka, and Y. Kurosawa. 1988. Organization of human immunoglobulin heavy chain diversity loci. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:4141-4150.
54. Ravetch, J. V., U. Siebenlist, S. Korsmeyer, T. Waldman, and P. Leder. 1981. Structure of the human immunoglobulin u locus: characterization of embryonic and rearranged J and D genes. *Cell.* 27:583-591.
55. Slaughter, L., D. A. Carson, F. C. Jensen, T. L. Holbrook, and J. H. Vaughan. 1978. In vitro effects of Epstein-Barr virus on peripheral blood mononuclear cells from patients with rheumatoid arthritis and normal subjects. *J. Exp. Med.* 148:1428-1434.
56. Steinitz, M., G. Izak, S. Cohen, M. Ehrenfeld, and I. Flechner. 1980. Continuous production of monoclonal rheumatoid factor by EBV-transformed lymphocytes. *Nature (Lond.)* 287:443-445.
57. Nakamura, M., S. E. Burastero, A. L. Notkins, and P. Casali. 1988. Human monoclonal rheumatoid factor-like antibodies from CD5 (Leu-1)+ B cells are polyreactive. *J. Immunol.* 140:4180-4186.
58. Burastero, S., P. Casali, R. L. Wilder, and A. L. Notkins. 1988. Monoreactive, high affinity rheumatoid factors are produced by CD5+ B cells from patients with rheumatoid arthritis. *J. Exp. Med.* 168:1979-1992.
59. Nelson, J. L., F. A. Nardella, I. R. Oppliger, and M. Mannik. 1987. Rheumatoid factors from patients with rheumatoid arthritis possess private repertoires of idiotypes. *J. Immunol.* 138:1391-1396.
60. Ruiz-Arguelles, A., and M. Presno-Bernal. 1989. Demonstration of a cross-reactive idotype (IdRQ) in rheumatoid factors from patients with rheumatoid arthritis but not in rheumatoid factors from healthy, aged subjects. *Arthritis Rheum.* 32:134-138.
61. Natvig, J. B., K. Thompson, I. Randen, M. Steinitz, M. Tausig, D. Beale, P. Barker, K. Sletten, K. Waalen, and O. Forre. 1988. Partial amino acid sequence and variable subgroup determination (VH and VL) of a monoclonal rheumatoid factor derived from a rheumatoid arthritis patient. *Scand. J. Immunol.* 75:127-132.
62. Shen, A., C. Humphries, P. Tucker, and F. Blattner. 1987. Human heavy chain variable region family nonrandomly rearranged in familial chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 84:8563-8567.
63. Sanz, I., H. Dang, M. Takei, N. Talal, and J. D. Capra. 1989. VH sequence of a human anti-Sm autoantibody: evidence that autoantibodies can be unmutated copies of germ line genes. *J. Immunol.* 142:883-887.
64. Schroeder, H. W., J. L. Hillson, and R. M. Perlmutter. 1989. Structure and evolution of mammalian Vh families. *Int. Immunol.* 2:41-50.
65. Sanz, I., and J. D. Capra. 1987. V $\kappa$  and J $\kappa$  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.
66. Natvig, J. B., P. I. Gaarder, and M. W. Turner. 1972. IgG antigens of the C gamma-2 and C gamma-3 homology regions interacting with rheumatoid factors. *Clin. Exp. Immunol.* 12:177-184.