Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 5567–5571, July 1989
Immunology

Complete physical map of the human immunoglobulin heavy chain constant region gene complex

(antibody genes/long-range restriction mapping)

MARTEN H. HOFKER*, MICHAEL A. WALTER*†, AND DIANE W. COX*†‡

*The Research Institute, The Hospital For Sick Children, Toronto, ON, Canada M5G 1X8; and †Departments of Paediatrics and Medical Genetics, University of Toronto, Toronto, ON, Canada M5S 1A8

Communicated by L. Luca Cavalli-Sforza, April 20, 1989 (received for review December 27, 1988)

ABSTRACT We have found by pulsed-field gel electrophoresis that the human immunoglobulin heavy chain constant region gene complex maps entirely to a 350-kilobase (kb) *Mlu* I fragment. The enzyme *Eag* I was used with pulsed-field gel electrophoresis alone and in double digests with *Spe* I to map the region. $C_\gamma 3$, of the $C_\gamma 3$–$C_\gamma 1$–$C_{\psi\varepsilon} 1$–$C_\alpha 1$ cluster, maps 60 kb to the 3′ side of C_δ; $C_\gamma 2$ of the $C_\gamma 2$–$C_\gamma 4$–C_ε–$C_\alpha 2$ cluster maps 80 kb to the 3′ side of $C_\alpha 1$, where $C_\gamma 3$ encodes the constant region of the immunoglobulin $\gamma 3$ chain, $C_\gamma 1$ encodes the constant region of the immunoglobulin $\gamma 1$ chain, etc. $C_{\psi\gamma}$ maps 35 kb to the 3′ side of $C_\alpha 1$ and is in the same transcriptional orientation as the other genes. Although in the cloned DNA many CpG-containing restriction sites were identified, most of these were methylated in peripheral blood leukocytes. The sites that were not methylated were predominantly found in three clusters, or *Hpa* I tiny fragment islands. One was found on the 5′ side of C_μ; the other two lie 30 kb to the 3′ side of each of the C_α genes and could indicate the presence of regulatory sequences or genes. A region showing strong linkage disequilibrium between all C_γ genes spans at least 160 kb. The 70-kb C_μ–$C_\gamma 3$ region, however, shows no linkage disequilibrium, possibly indicating a recombination hot spot. The immunoglobulin heavy chain constant region has been almost entirely cloned and mapped, and thus most rearrangements occurring in this region should be detectable.

The immunoglobulin heavy chain gene complex maps to chromosome 14q32.3 (1) and is comprised of some 200 variable (V) gene segments, at least 20 diversity (D) elements, and six joining (J) segments, a constant (C) region of nine genes, and two pseudogenes. During B-cell development, gene segments of the two regions recombine together, mediated by the J and D segments that lie between the V and C regions. The variable repertoire determines the binding specificity of the antibody. The C region genes mediate effector functions, such as membrane binding or secretion, or specialized class-dependent functions, such as the ability to activate complement or to cross the placenta (for review, see refs. 2 and 3). The total size of the germ-line immunoglobulin gene complex has been estimated to be about 2500 kilobases (kb), based on the summation of *Not* I fragments as revealed with a set of heavy chain V and C region probes (4). However, an accurate map of the gene segments is still lacking for both the V region and the C region genes.

The organization of portions of the V region has been determined by mapping of cloned regions and has revealed interspersed organization of V_H families (5). A V_H6 fragment maps 90 kb on the 5′ side of the C region of the μ chain (C_μ) (4, 6). There is some evidence that V_H6 is used preferentially in early development (7) and was, therefore, expected to be one of the most 3′ V_H genes, according to the mouse model (8). Identification of 3′ V_H segments may be important for our understanding of the generation and regulation of both the normal and the disease antibody repertoire.

The immunoglobulin heavy chain constant region gene complex (IGHC) is more completely characterized. Apparently a large ancestral duplication has yielded two similar clusters. The $C_\gamma 3$–$C_\gamma 1$–$C_{\psi\varepsilon} 1$–$C_\alpha 1$ cluster maps on the 5′ side of the $C_\gamma 2$–$C_\gamma 4$–C_ε–$C_\alpha 2$ cluster (where $C_\gamma 3$ is the constant region of the immunoglobulin $\gamma 3$ chain, $C_\gamma 1$ is the constant region of the immunoglobulin $\gamma 1$ chain, etc). Gene organization within each cluster has been determined by cosmid cloning (9). These clusters are located on the 3′ side of the C_μ and C_δ genes (10). However, to the best of our knowledge prior to this report, distances between C_δ and $C_\gamma 3$, and between $C_\alpha 1$ and $C_\gamma 2$ were unknown. The $C_{\psi\gamma}$ gene maps between $C_\alpha 1$ and $C_\gamma 2$ as established by linkage analysis (11) and deletion analysis (12, 13).

The IGHC region is highly polymorphic. Allotypes of IgG, the Gm variants (for review, see ref. 14), as well as restriction fragment length polymorphisms (RFLPs) (11, 15–17) have been demonstrated. Most disease-association studies are based on the Gm system and have yielded low correlations between haplotypes and disease. Since this may be due to a limited number of alleles, RFLPs may prove to be extremely useful in extending the polymorphisms of the Gm system to a much larger haplotype set. As many as 82 haplotypes have been described (16). The two C_γ clusters show a tight linkage disequilibrium extending at least from $C_\gamma 3$ to $C_\gamma 4$ (16). However no such linkage disequilibrium exists between the μ-switch RFLPs and $C_\gamma 3$ (15). Optimal use of the immunoglobulin genes in disease-association studies can only be made when the entire region has been mapped by physical and genetic means. In particular the region between C_μ and $C_\gamma 3$ requires more study, since the lack of association between these loci must be considered in seeking disease associations. The presence of regulatory elements or even potential genes in this region cannot be excluded.

The physical map of the IGHC, which we present here, will allow a better understanding of recombination frequencies in this region of the genome. In addition, the identification of rare-cutting sites has led to the detection of three *Hpa* I tiny fragment (HTF) islands, an indication of nonmethylated regions that are CpG-rich and most likely evolutionarily conserved (18, 19). Understanding the nature of HTF islands is important in view of their usefulness in identifying genes.

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

Abbreviations: IGHC complex, immunoglobulin heavy chain constant region gene complex; V, D, J, C, variable, diversity, joining, and constant region genes, respectively; PFGE, pulsed-field gel electrophoresis; HTF, *Hpa* I tiny fragment; $C_\gamma 1$, $C_\gamma 3$, etc., constant region of the immunoglobulin $\gamma 1$ chain, constant region of the immunoglobulin $\gamma 3$ chain, etc.; RFLP, restriction fragment length polymorphism; V_H, heavy chain V gene segment.

‡To whom reprint requests should be addressed at the Department of Genetics, The Hospital For Sick Children, Toronto, ON, Canada M5G 1X8.

MATERIALS AND METHODS

Cosmid Clones and Probes. Cosmids cosIg1, cosIg8, cosIg10, and cosIg13, provided by T. H. Rabbitts (Medical Research Council, Cambridge), were described (9). Cosmids c3p1 and c17p3 (provided by H. W. Schroeder, Jr., University of Washington, Seattle) and cosμ6 have been described (6). Cosmid CH630A was provided by J. Ellison (University of California at Los Angeles). Probe 24BRH is a 2.0-kb *Hin*dIII–*Eco*RI fragment of clone 24B (20), containing the $C_\gamma 4$ gene. Probe α5.6 is a 5.6-kb *Sst* I–*Bam*HI fragment containing the $C_\alpha 1$ gene plus a 3′ portion and was derived from the phage Huα1 (21). Probe CE1.6 (M.A.W. and N. Norman, unpublished data) is a 1.6-kb *Sst* I fragment that maps on the 5′ side of the C_ε gene but is deleted in the $C_{\psi\varepsilon}1$ gene and was derived from the phage Huε (22). Probe sμ is a 2.2-kb *Sst* I fragment derived from the switch region of C_μ, isolated from phage h18, and provided by R. Wall (National Institutes of Health, Bethesda). Probe PCW101, provided by M. Belle White (University of Wisconsin, Madison), maps just on the 5′ side of C_δ (23). Probe Cla17 is a 17-kb *Cla* I fragment of cosmid c17p3 (6) containing the 9-kb repeat of the D region.

Restriction Mapping. For conventional electrophoresis, genomic DNA was isolated as described (24). DNA was digested with restriction enzymes obtained from Pharmacia, New England Biolabs, and Boehringer Mannheim. Incubations were carried out in the five-buffer system of Boehringer, except for rare-cutting enzymes that were used in the recommended buffers as specified by New England Biolabs. Separation of relatively large DNA fragments (15 kb–100 kb) was obtained by electrophoresis of 3 μg of the digested samples through a 0.35% agarose gel for 88 hr at 0.5 V/cm in TEA buffer (25). Gels were denatured in 0.4 M NaOH/1.5 M NaCl and transferred in the same solution to GeneScreen*Plus* (DuPont). Probes were labeled with [α-^{32}P]dCTP (Amersham, 3000 Ci/mmol; 1 Ci = 37 GBq) to a specific activity of 2×10^9 dpm/μg by using a random-priming kit (Boehringer Mannheim). Hybridizations were carried out at 65°C in 7% (wt/vol) SDS/0.5 M sodium phosphate, pH 7.2/10 mM EDTA/1% bovine serum albumin (26).

Pulsed-field gel electrophoresis (PFGE) was performed as described (6). DNA used for this analysis was derived from the peripheral blood cells of a normal Caucasian and was the same as that used for conventional electrophoresis.

Spe I sites were mapped to cloned regions of the IGHC by linearization of the cosmids with *Sal* I and subsequent partial digestion with *Spe* I and hybridization with pBR322 fragments (27).

RESULTS

Long-Range Restriction Map Around the IGHC. An estimate of the total size of the IGHC was obtained by hybridizing probes for C_γ (24BRH), C_α (α5.6), and the switch region of C_μ (sμ) to DNA digested with *Mlu* I, *Eag* I, and *Bss*HII and separated using PFGE. Probe 24BRH yielded a 350-kb *Mlu* I fragment, a 520-kb *Bss*HII fragment, and two smaller *Eag* I fragments (Fig. 1*A*). Probe α5.6 revealed the same pattern (data not shown). The sμ probe also revealed the 350-kb *Mlu* I fragment and the 520-kb *Bss*HII fragment (Fig. 1*B*). Since the most 5′ C gene (C_μ) and the most 3′ gene ($C_\alpha 2$) were on the same 350-kb *Mlu* I fragment, the entire IGHC was contained within this fragment.

The *Eag* I fragments revealed by 24BRH were 180 kb and 130 kb, as observed when shorter switch times were used, and were not cleaved by *Mlu* I (Fig. 1*C*). Hence the *Eag* I fragments were entirely contained within the *Mlu* I fragment. Since no fragment was observed in the size range of 40 kb, the two *Eag* I fragments are most likely contiguous as discussed below.

The 180-kb band contains sμ (Fig. 1*B*). Also a 300-kb band is visible that results from partial digestion. The 300-kb band is much stronger for probe sμ than for probe 24BRH; therefore, the *Eag* I site just on the 5′ side of sμ (see below) is partially digested.

Probe CE1.6, specific for the C_ε gene, maps to the 130-kb *Eag* I fragment (Fig. 1*D*). Also a smaller, 100-kb fragment is visible that is occasionally produced when more enzyme is used. This *Eag* I site maps between C_ε and $C_\alpha 2$, as confirmed by double digests with *Spe* I and *Eco*RI. A summary of the PFGE results is shown in Fig. 2.

Mapping the Rare-Cutting Sites Relative to the IGHC. The mapping of the rare-cutting sites, specifically *Eag* I, relative to the IGHC gene segments was necessary to resolve the distances between IGHC genes and the rare-cutting sites and to obtain the distance between C_μ and $C_\gamma 3$ and between $C_\gamma 2$ and $C_\alpha 2$.

Double digests were used to screen large regions of the IGHC for rare-cutting sites, by using restriction enzymes that generate fragments >20 kb in combination with rare-cutting enzymes. One enzyme, *Spe* I, proved to be particularly useful, since it generates a 110-kb fragment for probes sμ and PCW101 (4) and allowed scanning for sites of a comparable

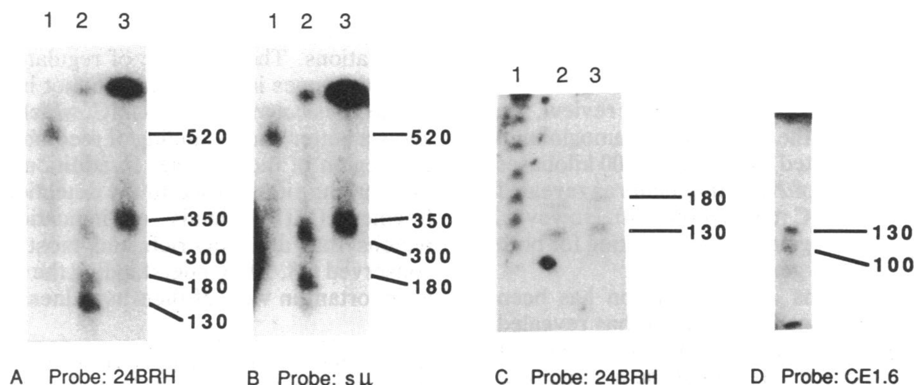

FIG. 1. (*A* and *B*) Autoradiograms of a pulsed-field gel hybridized with probe 24BRH (*A*) and subsequently with probe sμ (*B*). DNA was cleaved with *Bss*HII (lanes 1), *Mlu* I (lanes 2), and *Eag* I (lanes 3). A switch time of 45 sec was used, giving separation in the range of 50–700 kb. Both probes detect *Mlu* I and *Bss*HII fragments of identical size. Probe sμ detects *Eag* I fragments whose sizes coincide with the two largest 24BRH fragments. (*C*) Autoradiogram of a pulsed-field gel hybridized with 24BRH. DNA was cleaved with *Eag* I (lane 2) or *Eag* I and *Mlu* I (lane 3). The 25-sec switch time resolved 50- to 300-kb fragments, allowing sizing of the *Eag* I fragments of 180 kb and 130 kb. The λ ladder (lane 1) has increments of 50 kb; the smallest λ fragment is not visible. (*D*) Autoradiogram of *Eag* I-digested DNA, hybridized with probe CE1.6. In addition to the 130-kb fragment, a smaller fragment is visible, due to different digestion conditions.

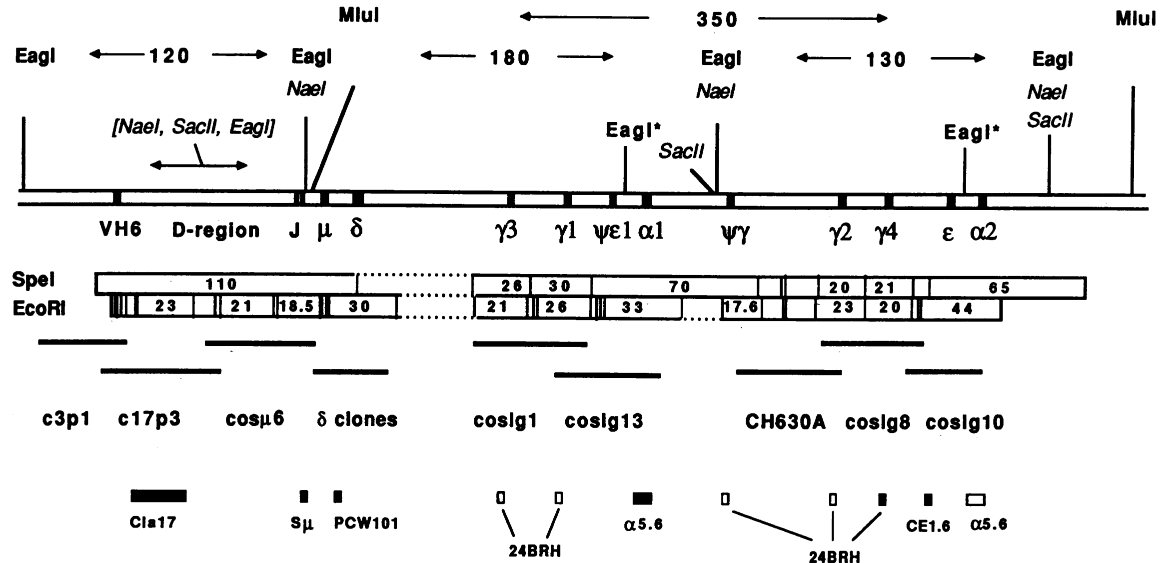

FIG. 2. Restriction map, with fragment sizes in kb, of the IGHC showing the map positions of the V, D, J, and C gene segments. The various gene segments are indicated by solid boxes. Cosmid clones are indicated by solid bars, and their clone numbers are given below. δ clones refer to a series of plasmid clones (23). Position of the cloned probes is indicated by solid boxes at the bottom of the figure. Open boxes show additional regions to which probes 24BRH and α5.6 hybridize. Dotted lines between the *Spe* I and *Eco*RI fragments refer to regions that are not mapped for these enzymes. *Eag* I and *Mlu* I sites (boldface) were obtained from PFGE analysis and by conventional electrophoresis using double digests; *Eag* I* sites were occasionally observed, when an excess amount of enzyme was used; *Eag* I, *Nae* I, and *Sac* II sites (italics) were obtained from conventional electrophoresis using double digests only.

region of the 5′ side of C_δ. $C_\alpha1$ and $C_\alpha2$ were on 70-kb and 65-kb fragments, respectively (Fig. 3*C*). The assignment of $C_\alpha2$ to the 65-kb fragment was obtained by using probe CE1.6, which hybridized only to the smaller band (data not shown). Restriction maps of the cosmids, obtained from published data for *Eco*RI (9) and generated in this study for *Spe* I, showed that the 5′ *Spe* I sites were 5 kb to the 5′ side of C_ε and 4 kb to the 5′ side of $C_{\psi\varepsilon}1$. Hence *Spe* I allowed us to map sites 50 kb to the 3′ side of both C_α genes.

The 70-kb band revealed by probe 24BRH (Fig. 3*A*) is identical to the *Spe* I fragment found for $C_\alpha1$. A probe for the switch region of $C_\gamma1$ (γ1-switch), which does not hybridize to $C_{\psi\gamma}$ (28), does not show this band (Fig. 3*B*); therefore $C_{\psi\gamma}$ was assigned to the 70-kb *Spe* I fragment.

Double digests using *Spe* I and *Eag* I yielded a 50-kb band with probe α5.6 (Fig. 3*C*), showing that *Eag* I sites mapped 30 kb to the 3′ side of each C_α gene. Another *Eag* I site was found 24 kb to the 5′ side of C_δ, as revealed by probe PCW101 (Fig. 3*D*). Double digests with *Eag* I and *Eco*RI (data not shown) mapped the latter *Eag* I site more accurately 2 kb to the 5′ side of the μ switch. The $C_{\psi\gamma}$ gene mapped to a *Spe* I–*Eag* I fragment of 18 kb, which agrees with our finding that the $C_{\psi\gamma}$ and $C_\alpha1$ genes are on identical *Spe* I fragments.

Double digests with *Spe* I plus *Mlu* I or *Bss*HI showed that these sites map on the 3′ side of the *Spe* I fragment of $C_\alpha2$. The 5′ *Mlu* I site was revealed by probe PCW101 (Fig. 3*D*) and maps 1 kb to the 5′ side of the μ switch. Two enzymes, *Nae* I and *Sac* II, had sites very close to the *Eag* I site revealed by $C_\alpha1$ and $C_\alpha2$. These sites were also demonstrated with probe 24BRH and provided additional evidence for linking $C_{\psi\gamma}$ to $C_\alpha1$.

From these mapping data, it can be concluded that $C_{\psi\gamma}$ maps to the 3′ side of the 180-kb *Eag* I fragment. Since no additional *Eag* I fragment has been found with 24BRH, $C_{\psi\gamma}$ must map to the 130-kb *Eag* I fragment and hence the two *Eag* I fragments are contiguous.

Orientation of $C_{\psi\gamma}$. $C_{\psi\gamma}$ maps to a 17.6-kb *Eco*RI fragment. *Spe* I recognized a site 2.6 kb from one of the ends of the *Eco*RI fragment, which places $C_{\psi\gamma}$ at the 3′ end of the 70-kb *Spe* I fragment (Fig. 4). In agreement with our map, the 17.6-kb *Eco*RI fragment is not cleaved by *Eag* I. In addition, $C_{\psi\gamma}$ maps to a 10-kb or 8.8-kb *Bam*HI fragment (dependent

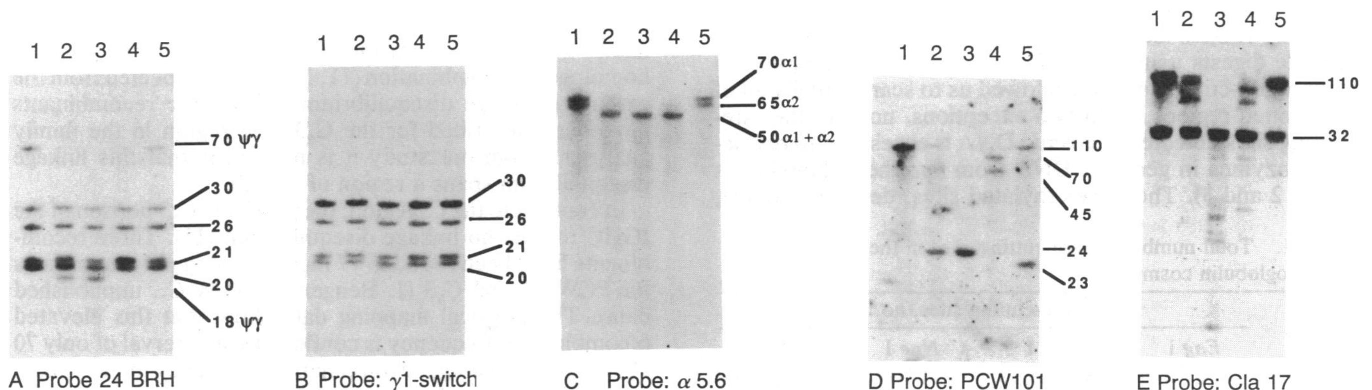

FIG. 3. Autoradiograms of Southern blots of genomic DNA digested with the following enzymes. Lanes: 1, *Spe* I; 2, *Spe* I and *Eag* I; 3, *Spe* I and *Nae* I; 4, *Spe* I and *Sac* II; 5, *Spe* I and *Mlu* I. Identical blots were hybridized with probes 24BRH (*A*), γ1-switch (*B*), α5.6 (*C*), PCW101 (*D*), and Cla17, a 17-kb *Cla* I fragment of c17p3, representing the D region (*E*). The sizes of the α1 and α2 fragments were estimated from their position between the 110-kb *Spe* I fragment for probe PCW101 (*D*) and a 50-kb λ fragment.

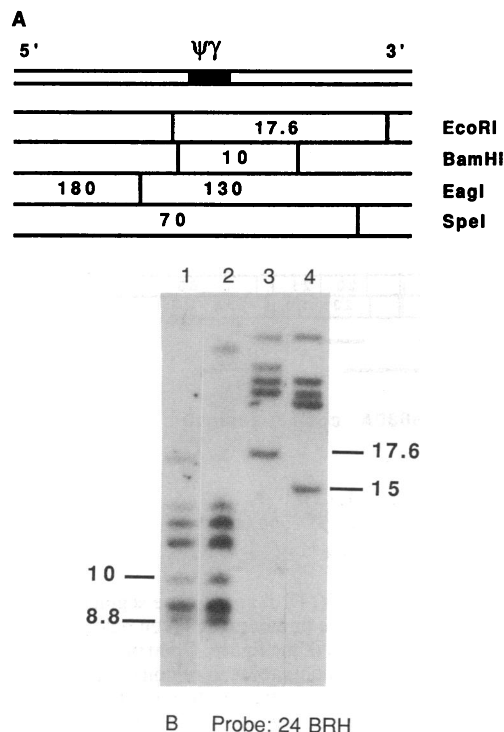

FIG. 4. Orientation of $C_{\psi\gamma}$ in the IGHC. (*A*) Restriction map around $C_{\psi\gamma}$. The 5' *Bam*HI and *Eco*RI sites are from ref. 28; for the 10-kb *Bam*HI allele, the *Bam*HI site is found 500 base pairs to the 3' side of the *Eco*RI site. Other sites are derived from the data of Figs. 3 and 4*B*. (*B*) Autoradiogram of a Southern blot of genomic DNA digested and enzymes as follows and hybridized with probe 24BRH. The 10-kb *Bam*HI fragment for $C_{\psi\gamma}$ does not alter upon *Spe* I digestion, whereas the $C_{\psi\gamma}$ *Eco*RI fragment is cleaved by *Spe* I. Lanes: 1, *Bam*HI and *Spe* I; 2, *Bam*HI; 3, *Eco*RI; 4, *Eco*RI and *Spe* I.

upon the allele involved). For the 10-kb allele, the *Eco*RI and *Bam*HI fragments mapped within a 500-base-pair region (28). These two sites are 5' with respect to the putative transcriptional orientation of $C_{\psi\gamma}$. Since the 10-kb *Bam*HI allele was unaltered in *Spe* I/*Bam*HI-digested DNA (Fig. 4), the *Spe* I site maps at the 3' end, outside the *Bam*HI fragment but inside the *Eco*RI fragment. From the map (Fig. 2) and these double-digest results, it can be concluded that $C_{\psi\gamma}$ is in the normal transcriptional orientation with the rest of the genes of the cluster.

Methylation of the IGHC. Analysis of each of the cosmids with *Eag* I, *Bss*HII, and *Sac* II yielded an abundance of restriction sites (Table 1). Several sites were in the exons of the C_γ genes; every C_γ gene contains at least one *Nae* I and two *Sac* II sites (sequences were obtained from GenBank). Double digests using either *Eco*RI or *Spe* I in combination with a rare-cutting enzyme allowed us to scan virtually all of the cloned regions. With two exceptions, none of the rare-cutting sites detected in cosmid DNA is accessible to restriction enzymes in genomic DNA from peripheral blood cells (Figs. 2 and 3). The nonmethylated sites, described above, mapped to three small distinct regions outside the cosmid clones and have characteristics of HTF islands (see *Discussion*).

Interestingly, many rare-cutting enzymes, especially *Nae* I, have accessible sites in the D region. Fig. 3*E* shows the results for a probe containing the 9-kb D region repeat (29). In addition to the 110-kb *Spe* I fragment, a 32-kb fragment can be found containing a D region that is located >1000 kb to the 5' side of the D region (30). The 110-kb fragment contains many accessible *Nae* I sites as shown by the presence of a large number of bands, while the 32-kb fragment appeared to be unaltered upon digestion by *Nae* I or any of the other enzymes used in the double digests.

Table 1. Total number of rare-cutting sites in the immunoglobulin cosmids

	Rare-cutting sites, no.			
Cosmid	*Eag* I (CGGCCG)	*Mlu* I (ACGCGT)	*Nae* I (GCCGGC)	*Sac* II (CCGCGG)
cosIg1	6	1	7	4
cosIg8	9	0	9	6
cosIg10	1	0	3	2
cosIg13	4	1	6	3

DISCUSSION

IGHC Structure. Genes of the human IGHC are contained within 300 kb. The distance between C_δ and $C_\gamma3$ is 60 kb, very similar to the distance in the mouse. In both species, the genes are in the same order 5'-C_μ, C_δ, C_γ, C_ε, C_α-3'. The C_γ gene in particular has been duplicated, giving rise to four copies in the mouse (31). In man, most likely, after an initial C_γ duplication, an entire C_γ–C_γ–C_ε–C_α segment has been duplicated (9). The HTF islands 30 kb to the 3' side of C_α, characterized by the presence of sites for *Eag* I, *Nae* I, and *Sac* II, also appear to be duplicated. Since the regions between C_α and the HTF islands lack *Spe* I sites in both cases (*Spe* I cleaves on average as frequently as *Bam*HI), it is very likely that this duplication has involved at least 90 kb.

It has been suggested from sequence analysis of the C_γ hinge regions that an ancestral $C_{\psi\gamma}$ and $C_\gamma1$ yielded $C_\gamma3$ by unequal crossing-over (28). This cannot be readily explained by the current physical map. Our lack of understanding of the evolutionary relationships may be due to the disappearance of some crucial intermediate alleles from the population. In addition, gene conversion could have contributed to these sequence homologies, as has been described for other C_γ genes (32). Interestingly, since $C_{\psi\gamma}$ is in the normal transcriptional orientation and its sequence has no deleterious mutations (33), except for the lack of a switch region (28), expression may be possible.

The ongoing evolution of the IGHC is evidenced by the detection of many deletion events, especially in populations of Southern Italy (34) and Tunisia (12, 13, 35), and of duplication events (11, 36). Our long-range map, based on peripheral blood cell DNA will allow more rapid characterization of these deletions and duplications.

Unequal Rates of Homologous Recombination in the IGHC. Several studies, using Gm markers as well as RFLPs, demonstrate linkage disequilibrium throughout the C_γ region. Two ancient haplotypes have been observed, and many alleles are derived from those by either point mutation or homologous recombination (11, 15, 16). As expected from the extensive linkage disequilibrium, no meiotic recombinants have been described for the $C_\gamma3$–$C_\gamma4$ region in the family analyses. From our study it is now clear that this linkage disequilibrium spans a region of 160 kb.

In contrast to the C_γ region, the μ switch–$C_\gamma3$ portion of the IGHC reveals no linkage disequilibrium (15). Three recombinants have been found in 77 meioses using polymorphisms for PCW101 and $C_\gamma3$ (J. Benger and D.W.C., unpublished data). The physical mapping data show that this elevated recombination frequency is confined to an interval of only 70 kb. From the average recombination frequency in the human genome, one would expect only one recombinant in 1400 meioses; in this 70-kb region from μ switch to $C_\gamma3$, the recombination frequency is ≈50 times higher. Probes flanking this region and in linkage disequilibrium with either 3' or 5' RFLPs may improve disease association studies and may

also lead to the characterization of this apparent hot spot for recombination.

HTF Islands. HTF islands, which have many *Hpa* II sites, are defined as nonmethylated, CpG-rich sequences. Since most CpG dimers in mammalian genomes are methylated and have a strong tendency to undergo transition to TpG, nonmethylated CpG dimers are scarce. However, they are commonly found in the 5′ region of housekeeping genes (18, 19). Rare-cutting enzymes, such as those used in this study, cleave predominantly in HTF islands. For the IGHC, the examination of cloned DNA has revealed many rare-cutting sites common to HTF islands, notably for *Sac* II (37). However, in genomic DNA, the only sites detected were confined to a 1-kb region on the 5′ side of the C_{μ} switch, containing sites for *Mlu* I, *Nae* I, and *Eag* I, and to two regions 30 kb to the 3′ side of both $C_{\alpha}1$ and $C_{\alpha}2$, containing sites for *Sac* II, *Nae* I, and *Eag* I. The occurrence of multiple recognition sites for the same enzyme in one of the HTF islands will remain undetected.

The accessibility of rare-cutting sites just on the 5′ side of the μ switch is compatible with the observation that the C_{μ}–J intron and the 3′ portion of the J region contains nonmethylated CpG dinucleotides in granulocyte DNA (38, 39), which is the major fraction of peripheral blood DNA. Hypomethylation is one of the requirements for the production of immunoglobulin transcripts (39). Restriction analysis of cosmid DNA and the sequence data for the J region (10) show that the *Eag* I and *Nae* I sites map in the 3′ portion of the J region on the 5′ side of the μ enhancer (40). Moreover, the sequence for J region reveals 23 *Hpa* II sites in 3.1 kb, characteristic of an HTF island. Our study shows that hypomethylation extends much further 5′, as judged from the presence of many accessible *Nae* I sites and other rare-cutting sites in the D region. In granulocytes, this hypomethylation is probably a remnant of potential activation of the immunoglobulin gene, which is shut off due to conformational changes in the chromosome (39).

In further studies, it will be of interest to determine if HTF islands on the 3′ side of the C_{α} genes are of functional importance. They could be associated with regulatory elements for IGHC or mark the presence of additional genes, such as those found in the mouse major histocompatibility complex (41). The HTF island on the 5′ side of $C_{\psi\gamma}$ could be the 5′ end or the promoter region of a gene (partly) encoded by $C_{\psi\gamma}$.

We are grateful to Drs. T. H. Rabbitts (Medical Research Council, Cambridge), H. W. Schroeder, Jr. (Seattle), and J. Ellison (University of California, Los Angeles) for providing cosmid clones. Probes for this study were generously made available by Drs. L. Hood, P. Leder, R. Wall, and M. Belle White. We thank Sarah Smith for her excellent technical assistance, and Judith Benger and Dr. Gail Fraizer for their helpful discussion. This work was supported by the National Science and Engineering Research Council of Canada (A0212). M.A.W. was supported by a studentship and M.H.H. was supported by a fellowship from the Medical Research Council of Canada.

1. Cox, D. W., Markovic, V. D. & Teshima, I. E. (1982) *Nature (London)* **297**, 428–430.
2. Tonegawa, S. (1983) *Nature (London)* **302**, 575–581.
3. Honjo, T. (1983) *Annu. Rev. Immunol.* **1**, 499–528.
4. Berman, J. E., Mellis, S. J., Pollock, R., Smith, C. L., Suh, H., Heinke, B., Kowal, C., Surti, U., Chess, L., Cantor, R. & Alt, F. W. (1988) *EMBO J.* **7**, 727–738.
5. Kodaira, M., Kinashi, T., Umemura, J., Matsuda, F., Noma, T., Ono, Y. & Honjo, T. (1986) *J. Mol. Biol.* **195**, 761–768.
6. Schroeder, H. W., Jr., Walter, M. A., Hofker, M. H., Ebens, A., Van Dijk, K. W., Liao, L. C., Cox, D. W., Milner, E. C. B. & Perlmuter, R. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8196–8200.
7. Schroeder, H. W., Jr., Hillson, J. L. & Perlmuter, R. M. (1987) *Science* **238**, 791–793.
8. Yancopoulos, G. D., Desiderio, S. V., Paskin, M., Kearny, J. F., Baltimore, D. & Alt, F. W. (1984) *Nature (London)* **311**, 727–733.
9. Flanagan, J. G. & Rabbitts, T. H. (1982) *Nature (London)* **300**, 709–713.
10. Ravetch, J. V., Siebenlist, U., Korsmeyer, S., Waldman, T. & Leder, P. (1981) *Cell* **27**, 583–591.
11. Bech-Hansen, N. T., Linsley, P. & Cox, D. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6952–6956.
12. Lefranc, M. P., Lefranc, G., de Lange, G., Out, T. A., van den Broek, P. J., van Nieuwkoop, J., Radl, J., Helal, A. N., Chaabani, H., van Loghem, E. & Rabbitts, T. H. (1983) *Mol. Biol. Med.* **1**, 207–217.
13. Chaabani, H., Bech-Hansen, N. T. & Cox, D. W. (1985) *Am. J. Hum. Genet.* **37**, 1164–1171.
14. Schanfield, M. & van Loghem, E. (1986) in *Handbook of Experimental Immunology*, ed. Weir, D. M. (Blackwell, Oxford), Vol. 3, pp. 94.1–94.18.
15. Migone, N., de Lange, G., Piazza, A. & Cavalli-Sforza, L. L. (1985) *Am. J. Hum. Genet.* **37**, 1146–1163.
16. Johnson, M. J., de Lange, G. & Cavalli-Sforza, L. L. (1986) *Am. J. Hum. Genet.* **38**, 617–640.
17. Jazwinska, E. C., Duncley, H., Propert, D. N., Gatenby, P. A. & Serjeantson, S. W. (1988) *Am. J. Hum. Genet.* **43**, 175–181.
18. Bird, A. P. (1986) *Nature (London)* **321**, 209–213.
19. Bird, A. P. (1987) *Trends Genet.* **3**, 342–347.
20. Ellison, J., Buxbaum, J. & Hood, L. (1981) *DNA* **1**, 11–18.
21. Ravetch, J. V., Kirsch, I. R. & Leder, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6734–6738.
22. Max, E. E., Battey, J., Ney, R., Kirsch, I. R. & Leder, P. (1982) *Cell* **29**, 691–699.
23. Belle White, M., Shen, A., Word, C. J., Tucker, P. W. & Blattner, F. R. (1985) *Science* **228**, 733–737.
24. Hofker, M. H., Wapenaar, M. C., Goor, N., Bakker, E., van Ommen, G. J. B. & Pearson, P. L. (1985) *Hum. Genet.* **70**, 148–156.
25. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
26. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
27. Hofker, M. H., van Ommen, G. J. B., Bakker, E., Burmeister, M. & Pearson, P. L. (1986) *Hum. Genet.* **74**, 270–274.
28. Takahashi, N., Ueda, S., Obata, M., Nikaido, T., Nakai, S. & Honjo, T. (1982) *Cell* **29**, 671–679.
29. Siebenlist, U., Ravetch, J. V., Korsmeyer, S., Waldman, T. & Leder, P. (1981) *Nature (London)* **294**, 631–635.
30. Matsuda, F., Lee, K. H., Nakai, S., Sato, T., Kodaira, M., Zong, S. Q., Ohno, H., Fukuharta, S. & Honjo, T. (1988) *EMBO J.* **7**, 1047–1051.
31. Shimizu, A., Takahashi, N., Yaoita, I. & Honjo, T. (1982) *Cell* **28**, 499–506.
32. Lefranc, M. P., Helal, A. N., de Lange, G., Chaabani, H., van Loghem, E. & Lefranc, G. (1986) *FEBS Lett.* **196**, 96–102.
33. Bensmana, M., Huck, S., Lefranc, G. & Lefranc, M. P. (1988) *Nucleic Acids Res.* **16**, 3108.
34. Migone, N., Oliviero, S., de Lange, G., Delacroix, D. L., Boschis, D., Altruda, F., Silengo, L., DeMarchi, M. & Carbonara, A. O. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5811–5815.
35. Lefranc, G., Lefranc, M. P. & Rabbitts, T. H. (1982) *Nature (London)* **300**, 760–762.
36. Bech-Hansen, N. T. & Cox, D. W. (1986) *Am. J. Hum. Genet.* **38**, 67–74.
37. Lindsay, S. & Bird, A. P. (1987) *Nature (London)* **326**, 336–338.
38. Bianchi, N. O., Peltomaki, P., Bianchi, M. S., Knuutila, S. & de la Chapelle, A. (1987) *Biochim. Biophys. Acta* **909**, 245–250.
39. Ford, A. M., Watt, S. M., Furley, A. J. W., Molgaard, H. V. & Greaves, M. F. (1988) *EMBO J.* **7**, 2393–2399.
40. Rabbitts, T. H., Forster, A., Baer, R. & Hamlyn, P. H. (1983) *Nature (London)* **306**, 806–809.
41. Abe, K., Wei, J.-F., Wei, F.-S., Hsu, J.-C., Uehara, E., Artzt, K. & Bennett, D. (1988) *EMBO J.* **7**, 3441–3449.