*Proc. Natl. Acad. Sci. USA*
Vol. 79, pp. 1984–1988, March 1982
Immunology

# Linkage and sequence homology of two human immunoglobulin γ heavy chain constant region genes

(evolution/gene duplication/unequal crossover/gene conversion)

JAY ELLISON AND LEROY HOOD

Division of Biology, California Institute of Technology, Pasadena, California 91125

*Communicated by Norman Horowitz, November 23, 1981*

**ABSTRACT** **We report the nucleotide sequence of a gene encoding a human immunoglobulin $C_{\gamma 2}$ region. Comparison with the previously determined $C_{\gamma 4}$ sequence reveals that these two genes share extensive (≈95%) homology in the three $C_H$ domain exons and adjacent noncoding regions. In contrast, hinge exons have diverged to a much greater degree, implying that natural selection has favored the generation of diversity in these coding regions. We have used the noncoding nucleotide differences to estimate that approximately 6–7 million years have elapsed since the occurrence of the gene duplication or correction event which generated the two identical ancestral genes. In addition we show that the two $C_\gamma$ genes are arranged in human chromosomal DNA in the configuration 5′-$C_{\gamma 2}$-17 kilobase pairs -$C_{\gamma 4}$-3′.**

IgG is the major class of antibody molecule in the serum of mammals, representing 70–80% of the total serum immunoglobulin in humans. It is distinguished from other classes of antibodies by a γ heavy chain constant region ($C_\gamma$). The $C_\gamma$ regions of human IgG molecules are divided into four subclasses ($C_{\gamma 1}$, $C_{\gamma 2}$, $C_{\gamma 3}$, and $C_{\gamma 4}$) encoded by distinct germ-line genes (1). Protein sequence studies (2–5) have shown that the subclasses are highly homologous, indicating that the corresponding genes derive from a common ancestral $C_\gamma$ gene.

Subclasses of IgG have been observed in several other mammals, although the number varies for different species (6). This observation suggested that $C_\gamma$ gene duplications occurred independently in various mammalian evolutionary lines after their divergence from a common ancestor. The model assumes that members of a $C_\gamma$ gene family evolve independently and that the accumulated differences in the $C_\gamma$ genes reflect the time elapsed since the duplication event. Molecular analyses of mouse $C_\gamma$ genes suggest that this simple model may not be correct (7, 8). These studies indicate that, during evolution, genetic information has been exchanged between nonallelic mouse $C_\gamma$ genes. This implies the existence of mechanisms that prevent the $C_\gamma$ genes from freely diverging from one another, so that sequence homology is continually renewed within the gene family. This postulated type of mechanism has been termed "gene correction" (9).

We are interested in determining the structural characteristics of human $C_\gamma$ genes to provide insights into the evolution of the $C_\gamma$ gene family. We previously determined the complete nucleotide sequence of one human $C_\gamma$ gene, that encoding the $C_{\gamma 4}$ region (10). In this paper we report the sequence of a human $C_{\gamma 2}$ gene and compare it to the $C_{\gamma 4}$ sequence. In addition we provide molecular evidence that these genes lie adjacent to one another in human chromosomal DNA.

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.

## MATERIALS AND METHODS

**Materials.** The human genomic DNA library was obtained from T. Maniatis. Sources of nucleic acid enzymes, reagents for DNA sequence analysis, bacteria, and the phage M13mp2 were those described by Steinmetz *et al.* (11).

**Isolation and Subcloning of Cloned Human Chromosomal Fragments.** All cloning experiments were carried out in accordance with the recommended National Institutes of Health guidelines for recombinant DNA research. Isolation and restriction mapping of DNA fragments from a human genomic DNA library cloned in λ Charon 4A bacteriophage were done as described (10). Either the entire Charon 4A recombinant or the $C_\gamma$-containing 6.4-kilobase-pair (kb) *Hin*dIII fragment of clone 5A was subcloned into the phage M13mp2. The DNA was first digested with *Hin*dIII plus either *Ava* II or *Alu* I. *Eco*RI-cleaved M13mp2 DNA and the fragments to be cloned were made flush-ended by treatment with T4 DNA polymerase and then were blunt-end ligated (12). The ligation mixture was used to transform *Escherichia coli* strain JM101, and $C_\gamma$-containing clones were isolated after screening of plaques (13) with the subcloned $C_{\gamma 4}$ gene from clone 24B (see Fig. 1 and ref. 10).

**DNA Sequence Analysis.** Individual M13 subclones were analyzed by the dideoxy technique (14, 15) essentially as described by Steinmetz *et al.* (11), except that [α-$^{32}$P]dATP was used as the labeled precursor. Alignment of the analyzed fragments yielding the composite $C_\gamma$ DNA sequence was determined by either overlaps of *Ava* II and *Alu* I fragments or by homology of the translated DNA sequence to existing sequence data for a human γ2 protein (3).

## RESULTS

**Human $C_{\gamma 2}$ and $C_{\gamma 4}$ Genes Are Linked.** Human genomic DNA clones hybridizing to a human $C_{\gamma 3}$ cDNA probe were isolated as described (10). Restriction maps for the inserts of five of these clones indicate that the corresponding chromosomal fragments overlap (Fig. 1). From clone blot hybridization experiments with the $C_{\gamma 3}$ probe, we deduce that two separate regions on the composite stretch of human DNA contain $C_\gamma$ sequences. We have previously determined the nucleotide sequence of the gene on the right of Fig. 1 and found that it encodes a $C_{\gamma 4}$ region (10). The sequence analysis of the $C_{\gamma 4}$ gene indicated that it is transcribed from left to right in Fig. 1, allowing us to orient the mRNA synonymous strand as shown. Below we show that the gene lying 17 kb 5′ to $C_{\gamma 4}$ is a $C_{\gamma 2}$ gene and that it also is transcribed from left to right in Fig. 1. This intergenic distance is comparable to the distances found be-

Abbreviations: $C_H$, constant region of heavy chain; kb, kilobase pair(s).

tween mouse $C_\gamma$ genes, which have been shown to range from 17 to 34 kb (16).

**Primary Structure of a Human $C_{\gamma 2}$ Gene.** The nucleotide sequence of the $C_\gamma$ gene contained in clone 5A is presented in Fig. 2. Translation of the coding regions indicates that this gene is transcribed in the same direction as the $C_{\gamma 4}$ gene and that it encodes the constant region of a γ2 protein. This conclusion is based on comparison with the complete sequence of the myeloma γ2 protein Til (3). The predicted protein sequence agrees with the Til sequence at all but three positions. Two of these are differences in amide assignment: we found a glutamic acid codon for position 20 of the $C_H1$ coding region and an asparagine codon for position 84 of the $C_H2$ exon, whereas Wang *et al.* (3) found glutamine and aspartic acid, respectively, for these residues. The other difference is a serine ↔ alanine interchange at position 60 of the $C_H1$ domain. The hinge segment, which most clearly distinguishes the human $C_\gamma$ subclasses, is identical for the putative $C_{\gamma 2}$ gene and protein sequences. Thus we feel confident in classifying this gene as a $C_{\gamma 2}$ gene.

Our sequence of the $C_{\gamma 2}$ gene, which begins 214 nucleotides 5′ to the $C_H1$ coding region and continues 207 residues past the termination codon, contains the same general structural features that we previously observed for the human $C_{\gamma 4}$ gene (10). The constant region and hinge exons are separated from one another by intervening DNA sequences (whose lengths are virtually identical in the two genes), and characteristic residues are present at the intron–exon junctions which presumably play a role in determining the proper splicing of the coding segments in the nuclear RNA precursor (17). The hexanucleotide A-A-T-A-A-A, which has been implicated as a signal sequence for the polyadenylylation of eukaryotic structural gene transcripts (18), is centered 103 nucleotides 3′ to the translation stop codon. As observed for the human $C_{\gamma 4}$ gene (10) and mouse $C_\gamma$ genes (8, 19, 20), a lysine residue not present in the mature protein is encoded at the COOH terminus of the $C_H3$ exon.

**Comparison of $C_{\gamma 2}$ and $C_{\gamma 4}$ Gene Sequences.** In Fig. 2 the nucleotide sequences of the human $C_{\gamma 2}$ and $C_{\gamma 4}$ genes are aligned for direct comparison. The substitutions leading to coding differences are indicated by listing the distinguishing amino acids of the $C_{\gamma 4}$ gene. The $C_{\gamma 2}$ region contains one less amino acid than its $C_{\gamma 4}$ counterpart; the apparent deletion is located near the $NH_2$ terminus of the $C_H2$ domain [we refer to the difference as a deletion because the $C_H2$ domains of the human γ1 and γ3 heavy chains contain the same number of amino acids as does the corresponding domain of the $C_{\gamma 4}$ region (2, 4)]. The nucleotide sequence alignment giving maximal homology in this region suggests that the coding difference arose from deletion or insertion events at two sites.

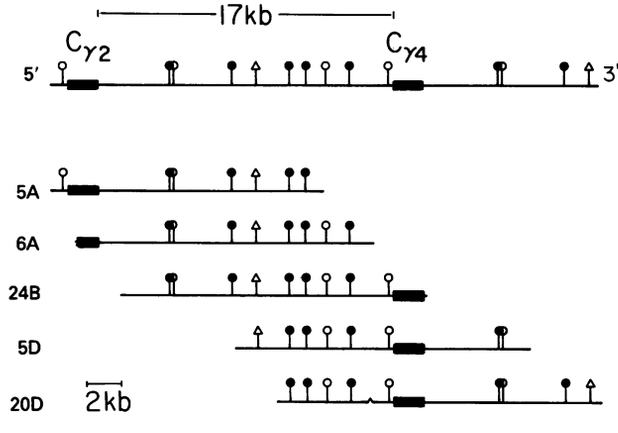


FIG. 1. Restriction map of a chromosomal region containing two human $C_\gamma$ genes. The five lowest lines represent the human DNA inserts of recombinants cloned in Charon 4A. Positions of cleavage of the restriction enzymes *Hin*dIII (♀), *Bam*HI (♦), and *Eco*RI (♙) are indicated. The restriction maps indicate that the cloned fragments overlap as shown. Solid bars represent the regions encompassed by $C_\gamma$ coding sequences; introns are not shown. A composite map of the chromosomal region spanned by the cloned fragments appears at the top. Arrows indicate the 5′-to-3′ orientation of the mRNA synonymous strands of the two $C_\gamma$ genes (see text). The *Bam*HI–*Hin*dIII fragment immediately 5′ to the $C_{\gamma 4}$ gene in clone 20D is approximately 100 nucleotides longer than the corresponding fragment in clones 24B and 5D. This conceivably could represent allelic variation.

Table 1 lists the nucleotide differences in the various coding and noncoding segments of the two human genes. Two notable features of the homology relationship are evident from the data. First, the noncoding regions show nearly as much homology as do the $C_H$ domain exons (≈95%). Second, the hinge exons are only about 70% homologous and thus are far more divergent than any of the other coding or noncoding regions.

In studies of recently diverged genes, Perler *et al.* (21) determined that noncoding nucleotide substitutions appear at a rate of approximately $7 \times 10^{-9}$ nucleotide substitutions per site per year. Assuming that these substitutions are phenotypically silent and thus not subjected to natural selection, this rate approximates the actual mutation rate and is presumed to be linear over a relatively short evolutionary period [about 100 million years according to these authors (21)]. Thus we should be able to use the data of Table 1 to estimate the time of divergence of the human $C_{\gamma 2}$ and $C_{\gamma 4}$ genes from a common ancestral sequence. Using the total percentage divergence in noncoding regions (4.6%) and the above substitution rate, we estimate that approximately 6.6 million years have elapsed since divergence of the human $C_{\gamma 2}$ and $C_{\gamma 4}$ genes.

## DISCUSSION

**Human $C_\gamma$ Coding Sequences.** We previously observed that the predicted protein sequence encoded in our $C_{\gamma 4}$ gene differs by a single residue from the partially determined sequence of a human γ4 protein (10). The $C_{\gamma 2}$ region encoded by the gene reported here is seen to differ by three amino acid residues from the complete sequence determined for the human γ2 heavy chain Til (3). Most of the sequence of the constant region of another human γ2 chain has been determined (22), and comparison of the three $C_{\gamma 2}$ sequences (two protein and one DNA) reveals three to four interchanges between all pairs of compared sequences, most of which are due to differences in amide assignment. We cannot be certain that any of these differences reflect genetic polymorphisms rather than technical artifacts. Thus, the protein polymorphisms seen in this small sample of $C_{\gamma 2}$ and $C_{\gamma 4}$ sequences are quite limited and possibly nonexistent. This observation is not surprising, given the paucity of different allotypes observed for human γ2 and γ4 chains (23). In contrast, the human γ1 and γ3 chains exhibit a large number of distinct genetic variants (24).

One striking feature of the coding sequences compared here is the extensive divergence in the hinge exons. Fig. 2 and Table 1 reveal that the hinge exons are situated between two introns showing the same high degree of homology as the $C_\gamma$ exons. Assuming that changes in intron sequences are not acted on by natural selection, we expect the rate of appearance of these changes to reflect the true mutation rate. The hinge sequences, however, show a much greater rate of genetically fixed change than the observed presumed basal level of the introns. We believe that the high rate of base substitution in hinge exons is due to rapid fixation of the substituted nucleotides by natural selection. There apparently is a selective advantage to generating diversity in the hinge coding region. The nature of this advantage is not obvious, although others have speculated about dif-

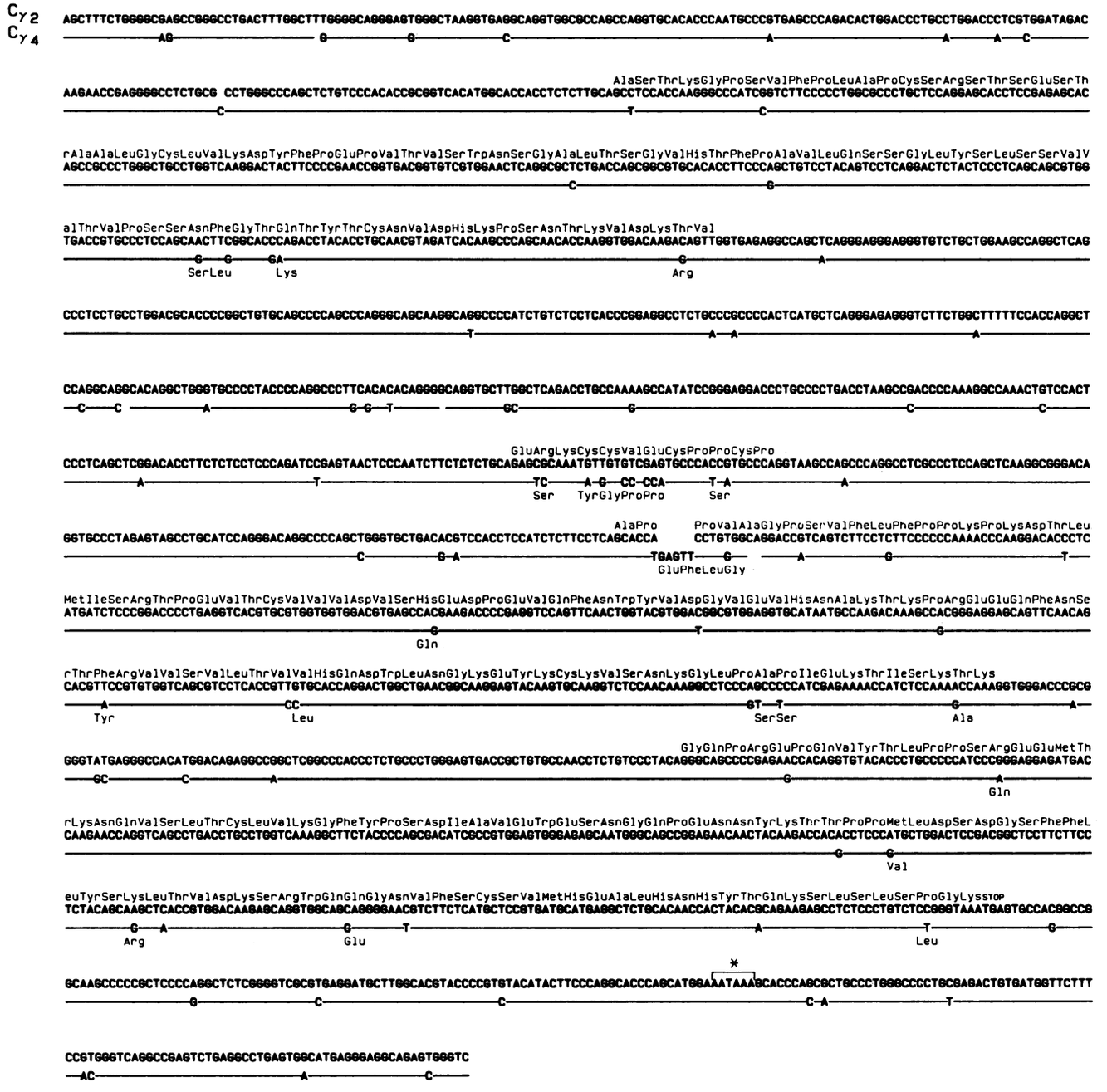

FIG. 2. Nucleotide sequences of the human $C_{\gamma 2}$ and $C_{\gamma 4}$ genes. The complete $C_{\gamma 2}$ protein sequence is listed above the corresponding DNA sequence. Solid lines represent identity of the $C_{\gamma 4}$ gene to the $C_{\gamma 2}$ sequence; where differences occur, the $C_{\gamma 4}$ sequence is indicated. A total of 2015 nucleotide positions are compared, including sequence gaps introduced into both genes to maintain sequence homology. Amino acid differences encoded in the $C_{\gamma 4}$ sequence are indicated by listing the $C_{\gamma 4}$-specific amino acids below the corresponding codons. Note that seven nucleotides near the 5′ end of the $C_H2$ exon are read in different translational reading frames in the two genes. "Stop" indicates the termination codon UGA. The presumptive poly(A) addition signal sequence is marked by an asterisk.

ferences in hinge sequences leading to different effector functions for the IgG subclass molecules (3, 25–28).

**Other Genetic and Evolutionary Considerations.** Yamawaki-Kataoka *et al.* (8) performed a detailed sequence comparison of three of the four mouse $C_\gamma$ genes, those encoding the $C_H$ regions of γ1, γ2a, and γ2b proteins. The homologies among these genes are much less than the homology seen for the human $C_\gamma$ genes reported here [on the order of 50–80%, depending on the genes compared (8)]. The mouse $C_\gamma$ genes, however, are more similar to one another than they are to the human genes (refs. 8 and 10, and this work). One way to explain the greater intraspecies homologies is to postulate that the common ancestor of humans and mice had a single $C_\gamma$ gene and that multiple $C_\gamma$ genes in the two species arose from independent gene duplication events occurring after speciation. Another explanation is that the common ancestor possessed multiple $C_\gamma$ genes and that gene correction mechanisms have operated to maintain sequence homology within the respective $C_\gamma$ gene families. According to these views, because the human $C_\gamma$ genes are more similar to one another than are the mouse genes, duplication or correction events have occurred more recently in the $C_\gamma$ genes of humans than of mice.

Gene correction can be defined, in the context of a multigene family, as a genetic process whereby the sequence of all or part

Table 1. Sequence differences between $C_{\gamma 2}$ and $C_{\gamma 4}$ genes

| Gene segment | Residues compared, no. | Gaps* | Substitutions, no. | | Difference, % | | |
|---|---|---|---|---|---|---|---|
| | | | Noncoding or synonymous coding | Coding replacement | Silent or noncoding | Replacement | Total |
| Coding: | | | | | | | |
| $C_H1$ | 294 | | 5 | 4 | 1.7 | 1.4 | 3.1 |
| Hinge | 36 | | 2 | 9 | 5.6 | 25.0 | 30.6 |
| $C_H2$ | 325† | 7‡ | 9 | 6 | 2.8 | 1.8 | 4.6 |
| $C_H3$ | 321 | | 5 | 5 | 1.6 | 1.6 | 3.2 |
| Noncoding: | | | | | | | |
| 5′ flanking | 213 | 2 | 9 | | 4.2 | | 4.2 |
| $C_H1$–hinge intron | 390 | 2 | 18 | | 4.6 | | 4.6 |
| Hinge–$C_H2$ intron | 118 | | 4 | | 3.4 | | 3.4 |
| $C_H2$–$C_H3$ intron | 97 | | 5 | | 5.1 | | 5.1 |
| 3′ untranslated§ | 125 | | 6 | | 4.8 | | 4.8 |
| 3′ flanking | 82 | | 5 | | 6.1 | | 6.1 |

* These were introduced as noted in Fig. 2; the relevant residues were not compared and do not contribute to the calculation of % difference.

† The homologous residues read in different reading frames are not compared here.

‡ See Fig. 2.

§ This region extends from the residue immediately 3′ to the termination codon to the site of poly(A) addition. We have tentatively placed this latter site 125 nucleotides 3′ to the stop codon (see figure 5 of ref. 10).

of the corrected gene is replaced by the sequence of a homologous nonallelic gene (9). Two mechanisms for gene correction among tandemly linked genes are homologous unequal crossover and gene conversion. Both models assume that the genetic recombination takes place between tandem gene arrays that are in phase but out of register, so that apposing DNA sequences are homologous but not identical. When unequal crossover events occur in intergenic regions, the result is expansion and contraction of the size of the gene family. Repeated events of this type during evolution can result in fixation of the sequence of a single family member at the expense of the other members (29). The result of an unequal crossover event within nonallelic structural genes is the production of a hybrid gene, the classic example being that of hemoglobin Lepore (30). Unlike unequal crossovers, gene conversion events do not change the size of the gene family. Rather, a given stretch of DNA sequence of one gene is replaced by the sequence of another nonallelic gene through a recombination event.

In light of these considerations and our determination of the $C_{\gamma 2}$ and $C_{\gamma 4}$ gene order, it is interesting to examine the case of a human IgG molecule that apparently resulted from a recombination between the $C_{\gamma 2}$ and $C_{\gamma 4}$ genes. Natvig and Kunkel (31) described a myeloma protein in which the $C_H1$ and $C_H2$ domains were characteristic of a γ4 chain and the $C_H3$ domain resembled that of a γ2 polypeptide. Fig. 3A schematically diagrams an unequal crossover involving $C_{\gamma 2}$ and $C_{\gamma 4}$ genes linked in the order we have determined. Both products of the recombination event contain hybrid genes; chromosome I contains only a Lepore-like gene, whereas chromosome II contains normal $C_{\gamma 2}$ and $C_{\gamma 4}$ genes in addition to a Lepore-like gene. In studies of sera from normal volunteers which apparently contained the normal counterpart of the hybrid myeloma protein, Natvig and Kunkel found that these individuals expressed normal IgG2 molecules, although they lacked IgG4 proteins. Chromosome II of Fig. 3A conceivably could represent the configuration in the DNA of these individuals, if one assumes that the normal $C_{\gamma 4}$ gene present is not expressed. Upstream (5′) to both the normal $C_{\gamma 4}$ gene and the hybrid gene are DNA sequences that mediate class switching. If $C_H$ switching is subclass-specific, the relevant sequences 5′ to the normal $C_{\gamma 4}$ gene may be excluded from the switching event by virtue of their being downstream from the same sequences present near the hybrid gene. If this explanation for the failure to express the normal $C_{\gamma 4}$ gene is correct, it may shed some light on the mechanisms whereby $C_H$ genes are selected for class-switch recombination.

A second explanation for the above clinical observations is

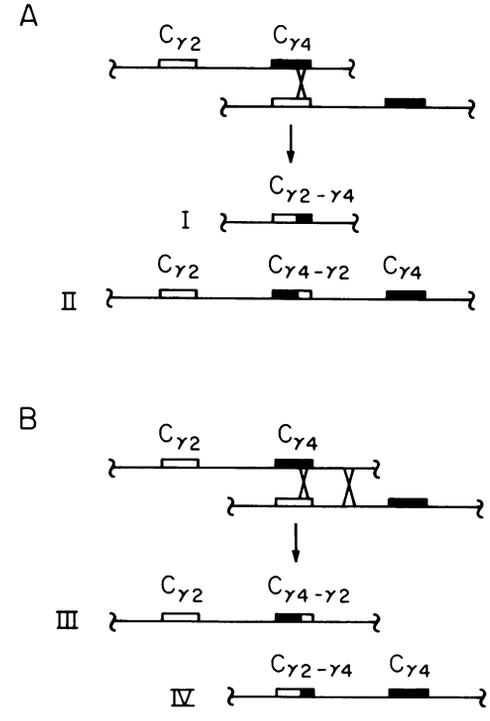


FIG. 3. Two models for the generation of a hybrid $C_{\gamma 4}$–$C_{\gamma 2}$ protein. (A). Unequal crossing over. A crossover occurs between misaligned chromosomes. The site of reciprocal exchange is indicated by the crossed lines. The two resulting chromosomes are labeled I and II. For simplicity we refer to the species participating in the genetic exchange as homologous chromosomes, but sister chromatids could equally well be involved. (B). Gene conversion. Gene conversion results in genetic information from one gene being copied onto a second gene, thus leading to gene correction. The chromosomes resulting from the gene conversion are labeled III and IV.

presented schematically in Fig. 3*B*. Here, a gene conversion process is proposed which involves a genetic exchange between aligned $C_{\gamma 2}$ and $C_{\gamma 4}$ genes. Chromosome III of Fig. 3*B* is consistent with the phenotype expressed by the volunteer individuals mentioned above. The gene conversion model is the simpler of the two proposals in that it does not require an *ad hoc* explanation for the failure to express a normal $C_{\gamma 4}$ gene.

Several examples of apparent gene correction have been reported for mammalian genes (8, 32, 33). In all of these cases, recombination points are proposed to lie within the structural genes rather than in intergenic regions. This conclusion follows from the observation that the levels of homology are different in different parts of the genes involved. Relatively sharp boundaries (representing presumed recombination break points) are observed, on either side of which are regions of greater and lesser homology. This pattern is evident when one examines rodent $C_\gamma$ sequences. For the case of the two most similar mouse genes ($C_{\gamma 2a}$ and $C_{\gamma 2b}$), the percentage of silent site substitutions in the $C_H1$ and $C_H3$ exons is 2–3 times greater than the corresponding neutral changes seen in the region between these exons (8). Comparison of guinea pig IgG subclass sequences also reveals disparate levels of homology in different $C_H$ domains. The $C_{\gamma 1}$ and $C_{\gamma 2}$ protein sequences are >90% identical in $C_H1$ domains, whereas the homology is 64–70% in $C_H2$ and $C_H3$ domains (28). In contrast, the two human $C_\gamma$ genes compared here show a nearly constant level of homology among the different exons and introns (except for the hinge region, as noted above). Thus, if a gene correction mechanism is responsible for the extensive homology between these genes, we would propose that the recombination break points of the most recent correction event lie outside the regions we have analyzed.

Clinical studies have uncovered a number of examples of apparent duplications and deletions of human $C_\gamma$ genes (34–37), as well as several cases of hybrid human $\gamma$ heavy chains (31, 38, 39). Thus, it appears that unequal crossover or gene conversion events occur at a significant frequency among human $C_\gamma$ genes. These types of events can be envisioned to play a part in both gene duplication and gene correction. It seems reasonable to view duplication and correction as two different manifestations of the same fundamental genetic process, one which has played an important role in the evolution of the human $C_\gamma$ gene family.

We thank Karyl Minard for excellent technical assistance and Bernita Larsh for the preparation of the manuscript. This work was supported by Grant GM 20927 from the National Institutes of Health and National Research Service Award 1T32 GM07616 from the National Institute of General Medical Sciences to J.E.

1. Kunkel, H. G., Allen, J. C., Grey, H. M., Martensson, L. & Grubb, R. (1964) *Nature (London)* **203**, 413–414.
2. Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, U. & Waxdal, M. J. (1969) *Proc. Natl. Acad. Sci. USA* **63**, 78–85.
3. Wang, A.-C., Tung, E. & Fudenberg, H. H. (1980) *J. Immunol.* **125**, 1048–1054.
4. Frangione, B., Rosenwasser, E., Prelli, F. & Franklin, E. C. (1980) *Biochemistry* **19**, 4304–4308.
5. Pink, J. R. L., Buttery, S. H., DeVries, G. M. & Milstein, C. (1970) *Biochem. J.* **117**, 33–47.
6. Nisonoff, A., Hopper, J. E. & Spring, S. B. (1975) in *The Antibody Molecule*, eds. Dixon, F. J., Jr., & Kunkel, H. G. (Academic, New York), pp. 314–318.
7. Miyata, T., Yasunaga, T., Yamawaki-Kataoka, Y., Obata, M. & Honjo, T. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2143–2147.
8. Yamawaki-Kataoka, Y., Miyata, T. & Honjo, T. (1981) *Nucleic Acids Res.* **9**, 1365–1381.
9. Hood, L., Campbell, J. H. & Elgin, S. C. R. (1975) *Annu. Rev. Genet.* **9**, 305–353.
10. Ellison, J., Buxbaum, J. & Hood, L. (1981) *DNA* **1**, 11–18.
11. Steinmetz, M., Moore, K. W., Frelinger, J. G., Sher, B. T., Shen, F.-W., Boyse, E. A. & Hood, L. (1981) *Cell* **25**, 683–692.
12. Wartell, R. M. & Reznikoff, W. S. (1980) *Gene* **9**, 307–319.
13. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
14. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161–178.
15. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
16. Roeder, W., Maki, R., Traunecker, A. & Tonegawa, S. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 474–478.
17. Lewin, B. (1980) *Cell* **22**, 324–326.
18. Proudfoot, N. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
19. Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N. & Mano, Y. (1979) *Cell* **18**, 559–568.
20. Tucker, P. W., Marcu, K. B., Newell, N., Richards, J. & Blattner, F. R. (1979) *Science* **206**, 1303–1306.
21. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. & Dodgson, J. (1980) *Cell* **20**, 555–566.
22. Connell, G. E., Parr, D. M. & Hofmann, T. (1979) *Can. J. Biochem.* **57**, 758–767.
23. Schanfield, M. S. (1978) in *Basic and Clinical Immunology*, eds. Fudenberg, H. H., Stites, D. P., Caldwell, J. L. & Wells, J. V. (Lange, Los Altos, CA), p. 59.
24. W.H.O. meeting announcement (1976) *J. Immunol.* **117**, 1056–1058.
25. Brunhouse, R. & Cebra, J. J. (1979) *Mol. Immunol.* **16**, 907–917.
26. Isenman, D. E., Dorrington, K. J. & Painter, R. H. (1975) *J. Immunol.* **114**, 1726–1729.
27. Boackle, R. J., Johnson, B. J. & Caughman, G. B. (1979) *Nature (London)* **282**, 742–743.
28. Cebra, J. J., Brunhouse, R. F., Cordle, C. T., Daiss, J., Fecheimer, M., Ricardo, M., Thunberg, A. & Wolfe, P. B. (1977) *Prog. Immunol.* **3**, 264–277.
29. Smith, G. P. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 507–514.
30. Flavell, R. A., Kooter, J. M., De Boer, E., Little, P. F. R. & Williamson, R. (1978) *Cell* **15**, 25–41.
31. Natvig, J. B. & Kunkel, H. G. (1974) *J. Immunol.* **112**, 1277–1284.
32. Slightom, J. L., Blechl, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
33. Liebhaber, S. A., Goossens, M. & Kan, Y. W. (1981) *Nature (London)* **290**, 26–29.
34. van Loghem, E., Sukernik, R. I., Osipova, L. P., Zegers, B. J. M., Matsumoto, H., de Lange, G. & Lefranc, G. (1980) *J. Immunogenet.* **7**, 285–299.
35. Lefranc, G., Dumitresco, S.-M., Salier, J.-P., Rivat, L., de Lange, G., van Loghem, E. & Loiselet, J. (1979) *J. Immunogenet.* **6**, 215–221.
36. Lefranc, G., Rivat, L., Rivat, C., Loiselet, J. & Ropartz, C. (1976) *Am. J. Hum. Genet.* **28**, 51–61.
37. van Loghem, E. & Natvig, J. B. (1970) *Vox Sang.* **18**, 421–434.
38. Kunkel, H. G., Natvig, J. B. & Joslin, F. G. (1969) *Proc. Natl. Acad. Sci. USA* **62**, 144–149.
39. Arnand, P., Wang, A.-C., Gianazza, E., Wang, I. Y., Lasue, Y., Creyssel, R. & Fudenberg, H. H. (1981) *Mol. Immunol.* **18**, 379–384.