Comparison of mouse immunoglobulin $\gamma 2a$ and $\gamma 2b$ chain genes suggests that exons can be exchanged between genes in a multigenic family

(nucleotide sequences/domains/intervening and untranslated sequences/unequal crossing-over/rate of divergence)

ROGER OLLO, CHARLES AUFRAY, CORINNE MORCHAMPS*, AND FRANÇOIS ROUGEON

Unité de Génétique et Biochimie du Développement, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

Communicated by François Jacob, December 19, 1980

ABSTRACT A 23-kilobase *EcoRI* DNA fragment coding for the BALB/c immunoglobulin $\gamma 2a$ chain was cloned from mouse embryo DNA in the cosmid pJC74, and a nucleotide sequence of 1904 bases was determined for the entire constant region (C_H1 , C_H2 , and C_H3), the three intervening sequences (IVS 1, IVS 2, and IVS 3) and the 5' and 3' flanking sequences. When the $\gamma 2a$ chain nucleotide sequence was compared with the $\gamma 2b$ chain nucleotide sequence, the percent homology of corresponding segments (excluding deletion and insertion) was 82% for the 5' flanking sequence, 87% for C_H1 , 84% for IVS 1, 96% for the hinge, 95% for IVS 2, 94.6% for C_H2 , 86% for IVS 3, 74% for C_H3 , 89% for the 3' untranslated region, and 92% for the 3' flanking region. These findings show that different domains of $\gamma 2a$ and $\gamma 2b$ have independent rates of evolution and that some of the noncoding segments of the gene are more conserved than are adjacent coding segments. Hypotheses on the possible role of IVS in gene evolution and expression are discussed.

A multigene gene family can be defined as a group of genes exhibiting four properties: multiplicity, close linkage, sequence homology, and related phenotypic functions (1). Genes in higher cells are characterized by the presence of intervening sequences (IVS) or introns located between the protein-coding DNA segments called the exons. Structural analyses of the genes from multigenic families, such as the globin family and the ovalbumin family, led to the hypothesis that split genes evolve in two ways: exons are under a selective pressure at the protein level and evolve slowly, and introns evolve more rapidly, both in length and in nucleotide sequence by deletion, insertion, and point mutations (2, 3). The genes coding for the constant regions of heavy chains (C_H) of immunoglobulins exhibit the four properties of a multigenic family. The C_H gene loci are linked on the chromosome in an order that has been determined by analyzing the pattern of C_H deletions in different plasmocytoma cells and by cloning experiments (4-6), which have established that the gene order for the γ subclass is $\gamma 3$, $\gamma 1$, $\gamma 2b$, $\gamma 2a$.

Analysis of the structure of human and mouse γ subclass chains has shown that these proteins are more closely related to one another than to other heavy chains. Genes encoding mouse $\gamma 2a$ and $\gamma 2b$ chains share high homology, probably resulting from a recent duplication of the three domain structures (7-10).

Immunoglobulin heavy chain genes are interrupted by introns located at homologous positions (11-14). It is likely that all of the C_H genes are derived from a common ancestor having a similar structural organization consisting of exonic segments alternating with introns with the last exonic segment linked to the 3' untranslated region.

We have reported recently the complete sequence of the structural gene for the constant region of a $\gamma 2a$ BALB/c myeloma protein (15). Comparison with the corresponding BALB/c $\gamma 2b$ sequence shows that the different domains have different rates of evolution. Furthermore, the 3' untranslated regions have the same length and exhibit a degree of homology higher than the homology in the adjacent C_H3 domain.

In this paper, we report the complete sequence of the $\gamma 2a$ gene. This sequence has been compared with the sequence of the adjacent $\gamma 2b$ gene (13), and the complex pattern of homology of mouse γ chain genes is discussed.

MATERIALS AND METHODS

Enzymes and Chemicals. T4 polynucleotide kinase; terminal deoxynucleotidyl transferase; and *Pst* I, *EcoRI*, *Bam*HI, and *Alu* I restriction endonucleases were purified according to published procedures (16-18).

T4 DNA ligase and all other restriction endonucleases used were from New England BioLabs. [γ - 32 P]ATP and 3'-[α - 32 P]dATP (cordycepin triphosphate) were obtained from the Radiochemical Centre (Amersham, England) and from New England Nuclear, respectively.

Bacteria and Plasmids. The bacterial strains used for the cloning experiments were *Escherichia coli* 1400 or *E. coli* ED 8767 (*Sup* E, *Sup* F, *met*⁻, *rec* A56) lysogenized by phage λ L512 (C_{I857}) and *E. coli* C600 (r_k^- , m_k^- , *rec* BC).

The lysogens used for preparation of packaging mixtures BHB 2688 N205 *rec* A⁻ (λ imm. 434 b2 red3 Eam 4 Sam 7)/ λ and BHB 2690 N205 *rec* A⁻ (λ imm. 434 cI ts red3 Dam 15 Sam 7)/ λ were obtained from B. Hohn.

Plasmids pBR322 and pJC74 (obtained from J. Collins) (19) were prepared according to Katz *et al.* (20). Plasmid pG2a-10-21, which contains a full-length transcript of a $\gamma 2a$ chain mRNA, has been described (21).

Partial Purification of the $\gamma 2a$ Chain Gene and Construction of Recombinant Cosmids. High molecular weight DNA from 15-day-old BALB/c mouse embryos was prepared according to Maniatis *et al.* (22), completely digested by restriction endonuclease *EcoRI*, and fractionated on a 10-40% sucrose gradient (22). DNA fragments in the 19- to 25-kilobase (kb) range were recombined with *EcoRI*-linearized cosmid pJC74 DNA. After *in vitro* packaging, infection of *E. coli* 1400 was carried out (22). The efficiency of packaging was 10^5 recombinant bacteria per μ g of mouse DNA. The bacteria were screened by *in situ* hybridization (23) with a 32 P-labeled nick-translated $\gamma 2a$ probe prepared according to Rigby and coworkers (24). The $\gamma 2a$ probe

Abbreviations: C_H , constant region of the heavy chain; IVS, intervening sequence; kb, kilobase(s); bp, base pair(s).

* Present address: Institut National de la Santé et de la Recherche Médicale U78, 543 Chemin de la Breteque, 76230 Bois-Guillaume, France.

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.

was a *Pst* I fragment (1000 nucleotides long) prepared from plasmid pG2a-10-21 (21).

Subcloning of *Bam*HI and *Sac* I Restriction Endonuclease Fragments. The cosmid DNA containing the γ 2a chain gene was digested with *Bam*HI, and the fragments were subcloned in the *Bam*HI site of pBR322 (25) and propagated in *E. coli* C600. The cosmid γ 2a was digested by *Sac* I, and the fragments were cloned in the *Pst* I site of pBR322 by dC-dG tailing (21).

Screening of recombinant clones was carried out according to Grunstein and Hogness (26); identification was by the rapid alkaline lysis method (27) and by the Southern gel-blotting procedure according to Breathnach *et al.* (28).

DNA Sequence Determination. The nucleotide sequence of the γ 2a chain gene was determined by using the partial chemical degradation method of Maxam and Gilbert (29) in conditions described previously (15).

RESULTS

Isolation of a Recombinant Cosmid Containing the γ 2a Chain Gene. When a nick-translated cDNA γ 2a clone is used as a hybridization probe (21), three *Eco*RI fragments are detected by the Southern gel-blotting procedure: 6.6 kb, 19 kb, and 23 kb. The 6.6-kb *Eco*RI fragment and the 19-kb *Eco*RI fragment have been isolated in a shotgun experiment using the Charon 4A vector phage, and identified as the γ 2b and γ 3 genes already described (30, 31). The sequence of the γ 3 gene will be reported elsewhere. Because of its size, the 23-kb fragment cannot be packaged in the Charon 4A. To concentrate the γ 2a gene sequence, *Eco*RI fragments were fractionated by sucrose-gradient centrifugation. Fractions ranging in size from 20 to 25 kb were pooled and packaged *in vitro* in the cosmid pJC74. By *in situ* hybridization, 2×10^4 recombinant cosmids were screened (23). One recombinant clone showing a strong hybridization with the nick-translated γ 2a insert was selected for further studies. Blotting experiments showed that the γ 2a gene was present in only two overlapping *Sac* I and *Bam*HI fragments. To obtain large amounts of DNA for nucleotide sequence determination, the *Bam*HI fragments were inserted into the *Bam*HI site of pBR322 and the *Sac* I fragments were inserted in the *Pst* I site of pBR322 by dC-dG tailing. This method allows the reconstruction of the *Sac* I and the *Pst* I sites (unpublished results). The strategy for DNA sequence determination is shown in Fig. 1.

Nucleotide Sequence of the γ 2a Chain Gene and Comparison with the γ 2b Chain Gene. The complete nucleotide sequence of the γ 2a chain gene is shown in Fig. 2. The sequence starts 152 nucleotides before the first nucleotide of the C_H1 and

ends 128 nucleotides after the polyadenylation site. When the sequences of the DNA segments encoding the γ 2a chain were compared with the sequence of the γ 2a heavy chain mRNA (15), no differences were found in the three domains, the hinge, and the 3' untranslated region.

The nucleotide sequence shows that, as in the case of all heavy chain genes so far studied (11–14), gene segments coding for the domains and the hinge are separated by IVS and there are no introns at the junction of the C_H3 domain with the 3' untranslated region.

The splicing points are in accord with the GT-AG rule proposed by Breathnach *et al.* (33). The lengths of the introns are IVS 1 = 312 base pairs (bp), IVS 2 = 109 bp, and IVS 3 = 112 bp. These values are similar to those found for the corresponding IVS segments of γ 1 and γ 2b—i.e., 356 and 316, 98 and 107, and 120 and 112, respectively.

The sequences of the γ 2a and γ 2b chain genes have been aligned to maximize the homology. As previously shown, the DNA sequences corresponding to the C_H1 , C_H2 , and C_H3 domains and the 3' untranslated region can be aligned without introducing deletions or insertions, although two deletions must be introduced in the hinge region. The percent homology that has been calculated is the ratio of the number of unchanged nucleotides to the sum of positions compared. When deletions or insertions exist in the segment, the nucleotides are excluded from the comparison. The homology map of the γ 2a and γ 2b chain genes is given in Table 1.

Homologies in the Intervening Sequences. The C_H1 domains and the hinges of γ 2a and γ 2b are, respectively, 87% and 96% homologous at the nucleotide level. These two segments (C_H1 and hinge) are interrupted by IVS 1, which has 84% homology. An analysis of the distribution of the differences along the sequence shows that, unlike other pairs of duplicated genes, homology is not found in the regions bordering the coding sequences. For instance, IVS 1 differed by 1 deletion, 1 insertion, and 3 nucleotide substitutions in the first 15 positions. In contrast, a sequence of 40 bp having 92% homology is found in the second moiety of IVS 1. The nucleotide sequences in IVS 2 are highly preserved. They differ in length by 2 bp, due to one insertion and two deletions. Both IVS 2 sequences show small differences, appearing as homologous as the adjacent exons in γ 2a and γ 2b (Table 1). IVS 3 has 86% homology and exactly the same length (112 nucleotides). It is noteworthy that 6 of the 15 base substitutions are found in the first 17 nucleotides. Although the C_H3 domains have only 74% homology, the adjacent 3' untranslated regions are 89% homologous. Thus, the exonic segment C_H3 has a higher degree of divergence than the adjacent IVS 3 and the 3' untranslated sequence.

FIG. 1. Strategy for sequence determination of γ 2a heavy chain gene. pSG2a-150 (1-kb) and pBG2a-4 (6.5-kb) overlapping fragments containing the γ 2a gene subcloned from cosmid G2a-2 are shown by solid lines. Cutting sites of restriction endonucleases used are indicated at bottom. Arrows show direction and extent of nucleotide reading. *Sac* I and *Pst* I fragments were labeled at their 3' end; all other fragments were labeled at their 5' end. The segments of the gene are delineated by vertical bars. H, hinge; 3'UT, 3' untranslated region.

```

1      GGATACACCATCAAGAGGAGGAAGGTGCTT( ACCCTGCCTGTCTCAAAGCAATTGAAGAAGACACAGTCTTTTCGAGCAAAGCTAAGACC
      ----A( T--CTG--A----T--C----G-T-----T-----G--TC-AA-----A-

      A K T T A P S V Y P L A
89     AGAGCCTCTCCTAAATATCTGAGGCCACAGATAACAGAAAAGCTCACACATCCTCCTCTCTTGCAGGCCAAAACAAGCCCCATCGGTCTATCCACTGGCC
      ---AT-AT---T-----C-----T-----T--(GA-----T-----C-----A-----
      P V C G D T T G S S V T L G C L V K G Y F P E P V T L T W N S G S
189    CCTGTGTGTGGAGATACAACCTGGCTCCTCGGTGACTCTAGGATGCTGGTCAAGGGTTATTTCCCTGAGCCAGTGACCTTGACCTGGAAGTCTGGATCCC
      ---G-----T-----C-----G-----C-----C-----T-----TG-----T-----
      L S S G V H T F P A V L Q S D L V T L S S S V T V T S S T W P S Q S
289    TGTCAGTGGTGTGCACACCTTCCCAGCTGTCTGAGTCTGACCTCTACACCTCAGCAGCTCAGTGAACCTCGAGCACCTGGCCAGCCAGTCTC
      ---CA-----C-----GA-----TA-G-----CC---C---T---A---T---A-
      I T C N V A H P A S S T K V D K K I
389    CATCACCTGCAATGTGGCCACCCGGCAAGCAGCACCAAGGTGGACAAGAAAATTGGTGAGGAAAACM )AGGGGAGTAGAGGTTACAAAGTGATTAGT
      -G-----GC--T--T---A--C-----C-----A---C-----AGG( )--TTC-----G--G-A---C--A-T-G--G
486    CTAAGGCCTTAGCCTAGCTAGACCAGCCAGGATCAG( CAGCCATCACCAAAAATGGGAAACCTTGGCCAGAGAGAAGGAGATACTGACTGTGACTCCC
      -A---TA-----T---A-----C)T-GG-T-----C)GGA-TG---A-----GGA---CA-----T--C--C-----
585    TCTTGAAACTTCTAACTATGACCACCTACCTTCAAGGTGCATGATCCTCTAGGATAGATGTCCTTGTCTATTCCAGGATCATCTCG(ACCTAAGCCCAT
      --C--G--A--G-----AC-----A--T-----G-----G--G-----G-A-----
      E P R G P
684    ACCAGGGACAAAG( TCCCTGGTTTGGTGCCTTTTCTCCTTCAAACCTTGAGTAACCCCCAGCCTTCTCTCTGCAGAGCCAGAGGGGCC( )
      ( )--( )-----CTTCTCTC-T-----C-----AACC-----AT-----C-----ATTCA
      T I K P C P P C K C P
777    ACAATCAAGCCCTGTCTCCATGCG( AAGTCCCCAGGTAAGTCACTAGACCAGAGCTCCACTCC( CGGGAGAATGGTAAGTGTGTAAAC
      ---C-----C-----AAGGAGTGTAC-----C)-----A--( )-----A-----
      A P N L L G G P S V F I F P P
858    ATCCCTGCACTAGAGGATAAGCCATGTACAGATCCATTTCCATCTCTCCTCATCAGCACCTAACCTCTTGGGTGGACCATCCGCTTCTCATCTTCCCTCCA
      -----T-A-G-----A-----T-----GA-----
      K I K D V L M I S L S P I V T C V V V D V S E D D P D V Q I S W F
958    AAGATCAAGGATGTACTCATGATCTCCCTGAGCCCCATAGTCACATGTGTGGTGGTGGATGTGAGCGAGGATGACCCAGATGTCCAGATCAGCTGGTTTG
      --T-----CA---AG---G-----C-----
      V N N V E V H T A Q T Q T H R E D V N S T L R V V S A L P I Q H Q D
1058   TGAACAACGTGGAAGTACACACAGCTCAGACACAAACCCATAGAGAGGATTACAACAGTACTCTCCGGGTGGTCACTGCCCTCCCATCCAGCACCAGGA
      -----A-----CA-----
      W M S G K E F K C K V N N K D L P A P I E R T I S K P K
1158   CTGGATGAGTGGCAAGGAGTTCAAATGCAAGGTCAACAACAAAGACCTCCAGCGCCCATCGAGAGAACCATCTCAAACCCAAAGGTGAGAGCTGCAGC
      -----T-A-----ATT-----G--C-----G
      G
1258   CTGACTGCATGGGGCTGGGATGGGCATAAGGATAAAGGTCTGTGTGGACAGCCTTCTGCTTACGCCATGACCTTTGTGTATGTTTCTACCTCACAGGG
      ACA-----A-----T---A-----CA-----C-A-----A--C-----
      S V R A P Q V V V L P P P E E E M T K K Q V T L T C M V T D F M P
1358   TCAGTAAGAGCTCCACAGGTATATGCTTGCCTCCACAGAGAAGAGATGACTAAGAAACAGGTCACTCTGACCTGCATGGTCCACAGCTTCATGCCTG
      CT---C-----A---CA-----G-----C---GC--T--T-C-G---G-T---G---C---T---C---GTG-G---AC---
      E D I V V E W T N N T K T E L N Y K N T E P V L D S D G S Y F M V S
1458   AAGACATTTACGTGGAGTGGACCAACAACGGGAAAACAGAGCTAAACTACAAGAACACTGAACCACTGCTGGACTCTGATGGTTCTTACTTCATGTACAG
      G-----CAGT-----G--T---C-T-----GAG-----G---C-C---T---A-----C-----A--T---
      K L R V E K K N W V E R N S Y S C S V V H E G L H N H H T T K S F
1558   CAAGCTGAGAGTGGAAAAGAAGAACTGGGTGGAAAGAAATAGCTACTCTGTTCACTGGTCCACGAGGGTCTGCACAATCACCACACGACTAAGAGCTTC
      ---C-ATA--A---CA-GC--G---A-A---C-G--TC--T---A---CAAC---AGA-----A-A---T---T---CT--AG---C-A---
      S R T P G K
1658   TCCCGGACTCCGGGTAATGAGCTCAGCACCACAAAACCTCTCAGGTCCAAAGAGACACCCACACTCATCTCCATGCTTCCCTTGTATAATAAAGCACC
      -----T-----G-----T-----TGG--C---A-----A-----T---
      * poly(A)
1758   CAGCAATGCCTGGGACCATGTAAACTGTCTGTTCTTTCCAAGTATAGAGCATAGCCACAGGCTGATATTTCTGGCCAGGGTTGGAGGACAGCCTT
      -----A-----T-----T-----G-----GGG-----
1858   GTCTATAGGAAGAGAATGAGGTTTTTGCACCTGCAGGACTCAGAGCTC

```

1904

FIG. 2. Nucleotide sequences of $\gamma 2a$ and $\gamma 2b$ heavy chain genes. Only those nucleotides that differ from the $\gamma 2a$ chain are indicated for the $\gamma 2b$ chain; homologous positions are indicated by dashes. Italic letters, amino acids [one-letter code (32)]. Parentheses, deletions required to maximize the homology.

Homology in the 5' and 3' Flanking Sequences. Comparison of the $\gamma 2a$ and $\gamma 2b$ 5' flanking sequences shows that the homology extends to the region preceding the C_H1 domain. Thus,

it appears that the 50 nucleotides preceding the C_H1 have about the same homology as the C_H1 (90%), and these nucleotides can be aligned without introducing deletions or insertions.

Table 1. Comparison of sequences of different segments of the $\gamma 2a$ and $\gamma 2b$ heavy chain genes

Nucleotide position*	$\gamma 2a$ gene segment	Nucleotides, no.	Gaps,† no.	Matching nucleotides, no.	Adjusted total length‡	Homology,§ %
0	5' flanking					
19						
152		133	3	129	136	82 (77.2)
443	C _H 1	291	0	291	291	87 (87)
755	IVS 1	312	11	306	324	84 (78)
803	Hinge	48	2	48	66	96 (70)
912	IVS 2	109	3	106	110	95 (92)
1252	C _H 2	330	0	330	330	94 (94)
1354	IVS 3	119	0	119	119	86 (86)
1675	C _H 3	321	0	321	321	74 (74)
1778	3'UT	103	0	103	103	89 (89)
1835		57	0	57	57	92 (92)
1904	3' flanking					

* In sequence of the $\gamma 2a$ gene (see Fig. 2).

† To maximize the homology (see Fig. 2).

‡ Aligned sequences.

§ Numbers in parentheses refer to homology expressed as the number of matching nucleotides per number of positions compared, including gaps, according to Van Ooyen *et al.* (34).

Comparison of the 3' untranslated regions of the $\gamma 2a$ and $\gamma 2b$ mRNAs shows an unexpectedly high degree of homology, which can be related to a structural requirement at the mRNA level. This homology extends over the poly(A) addition site; the next 57 nucleotides can be aligned and have 92% homology.

Nucleotide Sequences at the Intron-Exon Junctions. Examination of the sequences at the intron-exon junctions of eukaryotic organism genes has shown that the G-T and A-G dinucleotides at the 5' and 3' ends of the intron are characteristic of the splice signal required for the expression of the gene (33). Comparison of homologous sequences at the intron-exon junction in the γ subclass shows that the sequence at the junction is preserved for a short distance, similar to other multigenic families. Furthermore, it is possible to characterize a consensus sequence (Fig. 3).

At the 5' end of the introns, the consensus sequence is G-G-T-PuPuG and, like other genes, the intronic sequences contain a pyrimidine-rich region at the 3' end (35). With the exception of the $\gamma 1$ IVS 2, the characteristic 3' end of the intronic sequence is C-A-G. The 5' intronic sequence, G-G-T-G-A-G, found frequently in the γ families, has also been found in the mouse globin families (2). The heptanucleotide C-A-G-C-T-C-C at the 3' end of IVS 2 in $\gamma 2a$ has frequently been found at the 3' end of IVS 2 of the mouse globin genes (2).

DISCUSSION

Comparison of the primary structure of the nonallelic $\gamma 2a$ and $\gamma 2b$ chain genes suggests the following: (i) independent domains have different rates of evolution; (ii) the average homology of nonprotein coding elements is identical to that of coding elements (86%); and (iii) surprisingly, some nonprotein coding elements are more conserved than the adjacent coding elements. In the absence of an understanding of the role of introns in the expression and the evolution of genes in higher cells, only hypotheses can be proposed to explain the pattern of homology.

When the biological properties of $\gamma 2a$ and $\gamma 2b$ proteins are compared, it appears that both have the ability to fix complement but only $\gamma 2a$ has the ability to mediate local anaphylaxis (36). The first property is determined by the C_H2 domain, which is the most homologous, and the property to bind to mastocytes is determined by the C_H3 domain, which is the most divergent (37). If we assume that the pattern of homology between do-

main results only from a functional selection at the protein level, conservation of IVS, especially IVS 2 and IVS 3, could be explained by a selection pressure at the level of the primary transcript of the gene. Nevertheless, when the $\gamma 1$ and $\gamma 2a$ chain genes are compared, it appears that sequence divergence is higher in intronic than in exonic segments (data not shown).

Another hypothesis is to assume that the $\gamma 2a$ and $\gamma 2b$ chain genes result from a recent tandem duplication of a primitive $\gamma 2$ gene, as suggested by the high degree of conservation of non-protein coding elements. As a result of a double unequal crossing-over at the IVS 3-C_H3 domain and the C_H3-3' untranslated segment junctions, a C_H3 domain would be transferred into one of the duplicated forms of the gene. Examination of the amino acid substitutions along the C_H3 domain shows that the changes are not equally distributed along the sequence. This is the case for the segment encoding the amino acids from position 410 to 428. The exact boundaries of the transfer, if any, cannot be precisely assigned if intergenic conversion occurs during the crossing-over. A domain-transfer event has been proposed recently by Miyata *et al.* (38) to explain the fact that the C_H1 domain shows limited divergence among $\gamma 1$ and $\gamma 2b$.

The results presented here and those reported by Miyata *et al.* support Gilbert's hypothesis (35) that the rate of evolution of a gene in a multigenic family can be increased by coding element exchange. Nevertheless, independent of the case of the C_H3 domain, the divergence in the nonprotein coding segment appears limited. It is particularly interesting that the homology extends over the polyadenylation site. This could be due to a structural requirement related to the mechanism by which the alternative modes of processing produce membrane-bound or secreted immunoglobulin (39, 40).

We thank R. Nageotte and B. Chambrud for excellent technical assistance. We thank particularly W. Roskam and also J. L. Sikorav, A. Royal, and B. Cami for helpful discussions. We are grateful to Profs. G. Vyas and P. Baldacci for critical reading of this manuscript. This work was supported by Grants A.T.P. 3663 and 4247 from the Centre National de la Recherche Scientifique and the Fondation pour la Recherche Médicale Française.

1. Hood, L. (1976) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **35**, 2158-2167.
2. Leder, P., Hansen, J. N., Konkel, D., Leder, A., Nishioka, Y., & Talkington, C. (1980) *Science* **209**, 1336-1342.
3. Heilig, R., Perrin, F., Gannon, F., Mandel, J. L., & Chambon, P. (1980) *Cell* **20**, 625-637.

FIG. 3. Comparison of nucleotide sequences at the exon-intron junctions in $\gamma 2a$, $\gamma 2b$, and $\gamma 1$ heavy chain genes. Nucleotide sequences of $\gamma 2b$ and $\gamma 1$ chains at the splicing sites established by Yamawaki-Kataoka *et al.* (13) were compared with homologous position of the $\gamma 2a$ chain gene. Identical nucleotides are indicated by dashes.

4. Honjo, T. & Kataoka, T. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2140-2144.
5. Cory, S. & Adams, J. M. (1980) *Cell* **19**, 37-51.
6. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2138-2142.
7. Fougereau, M., Bourgois, A., de Preval, C., Rocca-Serra, J. & Schiff, C. (1976) *Ann. Immunol. (Paris)* **127C**, 607-631.
8. Yamawaki-Kataoka, Y., Sato, K., Shimizu, A., Kataoka, T., Mano, Y., Ono, M., Kawakami, M. & Honjo, T. (1979) *Biochemistry* **18**, 490-494.
9. Adetugbo, K. (1978) *J. Biol. Chem.* **253**, 6068-6075.
10. Milstein, C. & Pinck, J. R. L. (1970) *Prog. Biophys. Mol. Biol.* **21**, 209-263.
11. Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 857-861.
12. Sakano, H., Rogers, J. H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. & Tonegawa, S. (1979) *Nature (London)* **277**, 627-632.
13. Yamawaki-Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M. & Honjo, T. (1980) *Nature (London)* **283**, 786-789.
14. Gough, N. M., Kemp, D. J., Tyler, B. M., Adams, J. M. & Cory, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 554-558.
15. Sikorav, J. L., Auffray, C. & Rougeon, F. (1980) *Nucleic Acids Res.* **8**, 3143-3155.
16. Chang, L. M. S. & Bolland, F. J. (1971) *J. Biol. Chem.* **246**, 909-916.
17. Panet, A., VandeSande, J. H., Loewen, P. C., Khorana, H. G., Raaij, A. J., Lillehaug, J. R. & Kleppe, K. (1973) *Biochemistry* **12**, 5045-5049.
18. Greene, P. J., Heyneker, H. L., Bolivar, F., Rodriguez, R. L., Betlach, M. C., Covarrubias, A. A., Backman, K., Russel, D. J., Tait, R. & Boyer, H. W. (1978) *Nucleic Acids Res.* **5**, 2373-2380.
19. Collins, J. & Brünig, H. J. (1978) *Gene* **4**, 85-107.
20. Katz, L., Kingsbury, D. T. & Helinsky, D. R. (1973) *J. Bacteriol.* **114**, 557-591.
21. Auffray, C., Nageotte, R., Chambraud, B. & Rougeon, F. (1980) *Nucleic Acids Res.* **8**, 1231-1241.
22. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C. & Quon, D. (1978) *Cell* **15**, 687-701.
23. Cami, B. & Kourilsky, P. (1978) *Nucleic Acids Res.* **5**, 2381-2390.
24. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237-251.
25. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95-113.
26. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961-3965.
27. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513-1523.
28. Breathnach, R., Mandel, J. L. & Chambon, P. (1977) *Nature (London)* **270**, 314-319.
29. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
30. Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. & Honjo, T. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4240-4244.
31. Cory, S., Jackson, J. & Adams, J. M. (1980) *Nature (London)* **285**, 450-456.
32. Dayhoff, M. D. (1976) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5, pp. 189-190.
33. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853-4857.
34. Van Ooyen, A., Van den Berg, J., Mantei, N. & Weissmann, C. (1979) *Science* **206**, 337-344.
35. Gilbert, W. (1978) *Nature (London)* **271**, 501.
36. Kehoe, J. M. & Fougereau, M. (1969) *Nature (London)* **224**, 1212-1213.
37. Turner, M. W. (1977) *Immunochemistry* **1**, 36-57.
38. Miyata, T., Yasunaga, T., Yamawaki-Kataoka, Y., Obata, M. & Honjo, T. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2143-2147.
39. Early, P., Rogers, J., Davis, M., Kalame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20**, 313-319.
40. Oi, V. T., Bryan, V. M., Herzenberg, L. A. & Herzenberg, L. A. (1980) *J. Exp. Med.* **151**, 1260-1274.