Volume 12 Number 18 1984 Nucleic Acids Research

Two novel arrangements of the human fetal globin genes: $G\gamma$-$G\gamma$ and $A\gamma$-$A\gamma$

Patricia A.Powers, Cigdem Altay[1], Titus H.J.Huisman[2] and Oliver Smithies

Laboratory of Genetics, University of Wisconsin, Madison, WI 53706, USA, [1]Children's Medical Center, Hacettepe University, Ankara, Turkey, and [2]Medical College of Georgia, Department of Cell and Molecular Biology, Augusta, GA 30912, USA

Received 2 July 1984; Revised and Accepted 24 August 1984

ABSTRACT

We describe two novel arrangements of the human fetal globin gene region: one chromosome with two linked $A\gamma$ genes ($A\gamma$-$A\gamma$) and two chromosomes with two linked $G\gamma$ genes ($G\gamma$-$G\gamma$). The γ genes of these three chromosomes were cloned and the unusual 5' $A\gamma$ gene and one of the unusual 3' $G\gamma$ genes were partially sequenced. Both of these unusual genes differ from the genes normally found at their respective locations by a nucleotide substitution at the site of the single coding region difference between normal $G\gamma$ and $A\gamma$ genes. In both cases, the substitution is identical to the nucleotide found at that position in the normal neighboring gene. The unusual 3' $G\gamma$ gene also differs from normal $A\gamma$ genes at two other nucleotide positions, but both differences appear to be "private" or exclusive to this particular gene. These unusual fetal globin gene arrangements could have arisen from point mutations or from gene conversions of limited extent, the boundaries of which have been determined for all three chromosomes.

INTRODUCTION

Intergenic gene conversion, the localized exchange of information between related or allelic genes, appears to be an important mechanism in the evolution of multigene families. It has been implicated in both the maintenance of identity between related genes (1, 2, 3), and the generation of diversity between allelic genes (4, 5).

The first evidence for gene conversion between human genes was described by Slightom et al. (1). DNA sequence analysis of three human fetal globin genes from one individual suggested that, in one of the genes, part of the DNA sequence had been replaced by the corresponding region of the linked related gene. The resulting chromosome retains the normal arrangement of genes.

The striking homology between the human fetal globin genes ($G\gamma$ and $A\gamma$) makes them an excellent system for the study of gene conversion. Although the $G\gamma$ and $A\gamma$ genes probably arose from a common ancestor about 34 million years ago (2), the resulting γ chains differ at only a single amino acid residue having a glycine ($G\gamma$) or an alanine ($A\gamma$) at position 136 (6). Furthermore,

© IRL Press Limited, Oxford, England.

the coding regions of the two genes have only the single nucleotide difference which causes the amino acid difference.

In an attempt to further investigate the occurrence of gene conversion between the fetal globin genes, we have looked for fetal globin gene regions having either two linked $^{G}\gamma$ or two linked $^{A}\gamma$ fetal globin genes since they should be excellent candidates for having been formed via gene conversion between the human fetal globin genes. We describe here the result of a successful search for individuals having either two linked $^{G}\gamma$ or two linked $^{A}\gamma$ fetal globin genes and we characterize their unusual γ globin genes.

MATERIALS AND METHODS

DNA Preparation

Large molecular weight DNA was prepared from fibroblasts (563, A. R., C. G.) by the method of Blattner et al. (10) and from peripheral blood leukocytes from M. P. using minor modifications of the procedure described by Poncz et al. (11).

Restriction Enzyme Mapping of Genomic DNA

Restriction enzyme mapping of genomic DNAs was performed using minor modifications of the transfer techniques of Southern (12) and the hybridization protocol of Jeffreys and Flavell (13). Probes (illustrated in Figure 1C) used to map the genomic DNAs were γIVS2, a 457 bp Bam HI-Pvu II fragment which contains the 5' portion of the large intervening sequence of the $^{A}\gamma$ globin gene and inter-γ, a 700 bp Eco RI fragment derived from the intergenic region between the γ globin genes. After hybridization, filters hybridized with γIVS2 were washed at 68°C for 1.5 hours in 3xSSC, 0.5%SDS, and those hybridized with the inter-γ probe were washed in 0.1xSSC, 0.5% SDS at 68° for 1.5 hours.

Cloning the Fetal Globin Genes of the $^{G}\gamma$-$^{G}\gamma$ and $^{A}\gamma$-$^{A}\gamma$ Chromosomes

The two human fetal globin genes are normally contained on a 13 kbp Bgl II fragment. Libraries were constructed by digesting DNA from A. R., C. G. and M. P. to completion with Bgl II, fractionating the digested DNA on 5-20% sodium chloride gradients, ligating the 12-16 kbp DNA fragments to arms of the bacteriophage vector Charon 30 (14), assembling the ligated recombinant molecules into phages by in vitro packaging (15) and screening the resulting unamplified library with the γIVS2 probe. Each of the fetal globin genes was subcloned into the plasmid vector pAT153 (16).

DNA Sequencing

Sequence analysis was carried out by the method of Maxam and Gilbert (17)

with modifications as described (1). All DNA sequences were determined on both strands.

RESULTS

The Subjects

A. R., C. G. and M. P. were three Black individuals selected by inspecting the records of several hundred newborns routinely tested over the past several years for their neonatal Hb types and $G\gamma/(G\gamma + A\gamma)$ chain ratios. They were suspected of having either two $G\gamma$ or two $A\gamma$ fetal globin genes on the same chromosome because their $G\gamma/(G\gamma + A\gamma)$ chain ratios deviated from normal values (0.6-0.75) in the appropriate directions.

The first individual, A. R., had an unusually low $G\gamma/(G\gamma + A\gamma)$ chain ratio of 0.35 at birth. Over the course of his first year, his $G\gamma/(G\gamma + A\gamma)$ chain ratio declined to 0.05-0.1 while his HbF levels decreased normally. The same situation was present in his first cousin, who at birth had a $G\gamma/(G\gamma + A\gamma)$ chain ratio of 0.45 (7).

The second individual, C. G., had a $G\gamma/(G\gamma + A\gamma)$ chain ratio at birth of almost 1.0. At one year of age his $G\gamma/(G\gamma + A\gamma)$ chain ratio remained essentially 1.0, although his HbF level appeared to have declined normally (7).

The third individual, M. P., described at age 8 by Huisman et al. (8), was first studied as a newborn in 1971. At birth, his blood samples showed, in addition to HbA, the presence of two types of fetal hemoglobin. 13% of his fetal hemoglobin was HbF Port Royal, a γ chain variant with a Glu → Ala substitution at position 125 (9), and 87% was normal HbF with a $G\gamma/(G\gamma + A\gamma)$ chain ratio of 0.83.

DNA prepared from a fibroblast culture from an embryo (563) with the usual arrangement of $G\gamma$ and $A\gamma$ genes on both chromosomes was used as the normal control in our comparisons. The entire fetal globin gene region of both of the chromosomes of this individual has been sequenced (1, 2, and Slightom, Shen and Smithies, unpublished data).

Mapping of Genomic DNA

The overall strategy used for studying both the $G\gamma$-$G\gamma$ and $A\gamma$-$A\gamma$ chromosomes is based on the presence of the single nucleotide difference in the coding region of the two γ genes at codon 136, a G in the $G\gamma$ gene versus a C in the $A\gamma$ gene. This difference generates a *Pst* I restriction site in the $A\gamma$ globin gene that is absent in the $G\gamma$ globin gene. This site, together with a *Pst* I restriction site at the junction of the large intervening sequence and

the third exon of both genes, causes the appearance of three bands when DNA from normal $^{G}\gamma$-$^{A}\gamma$ homozygotes is digested with Pst I and hybridized to the γIVS2 probe (Figure 1A, lane 1). The three bands are: a 4.1 kilobase pair (kbp) fragment containing the 5' flanking sequence and the 5' portion of the $^{G}\gamma$ gene; a 4.9 kbp fragment containing the 3' portion of the $^{G}\gamma$ gene, the intergenic region and the 5' portion of the $^{A}\gamma$ gene; and a 0.8 kbp fragment containing the 3' portion of the $^{A}\gamma$ gene (Figure 1C). Change of a $^{G}\gamma$ gene to an $^{A}\gamma$ gene either by mutation or by gene conversion is expected to introduce a new Pst I site into the third exon of the 5' γ globin gene. Conversely, change of an $^{A}\gamma$ gene to a $^{G}\gamma$ gene is expected to result in the loss of the Pst I site in the third exon of the 3' γ globin gene.

The $^{A}\gamma$-$^{A}\gamma$ Gene Arrangement

Figure 1A illustrates the results obtained when DNA from A. R. (the individual with a low $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratio) is digested with Pst I and hybridized to the γIVS2 probe. His DNA (lane 2) gives the same bands of 4.9, 4.1 and 0.8 kbp as the normal control (lane 1), but the 4.9 kbp band appears to be half as intense and the 0.8 kbp band twice as intense as the bands of the normal control DNA. This suggests that a new Pst I site is present in the third exon of only one of A. R.'s 5' γ globin genes.

Figure 1B illustrates the results obtained when Pst I digests of DNA from A. R. and the normal control are hybridized to the inter-γ probe. DNA from the normal control yields a single band at 4.9 kbp (lane 5). DNA from A. R. yields two bands, the normal 4.9 kbp band and a 4.1 kbp band corresponding to a truncated intergenic Pst I fragment (lane 6). DNA from A. R. digested with the restriction endonucleases, Bam Hl, Eco RI, Bgl I, Acc I, Ava I, Hind III and Pvu II gave normal patterns when hybridized to both the γIVS2 and the inter-γ probes (data not shown). These results, as well as the hematological data presented above, support our conclusion that the 5' γ gene in one of A. R.'s chromosome has been changed from a $^{G}\gamma$ gene into an $^{A}\gamma$ gene. The restriction map corresponding to this unusual fetal globin gene arrangement is summarized in Figure 1C.

The $^{G}\gamma$-$^{G}\gamma$ Gene Arrangements

DNA samples from C. G. and M. P., the individuals with high $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratios suspected of having a $^{G}\gamma$-$^{G}\gamma$ fetal globin gene arrangement, were also compared to the normal DNA sample. Figure 1A shows the results obtained when DNA from the normal control (lane 1), M. P. (lane 3) and C. G. (lane 4) digested with Pst I was hybridized to the γIVS2 probe. A new band of 3.5 kbp is found in the lanes of DNA from M. P. and C. G. (lanes 3 and 4,

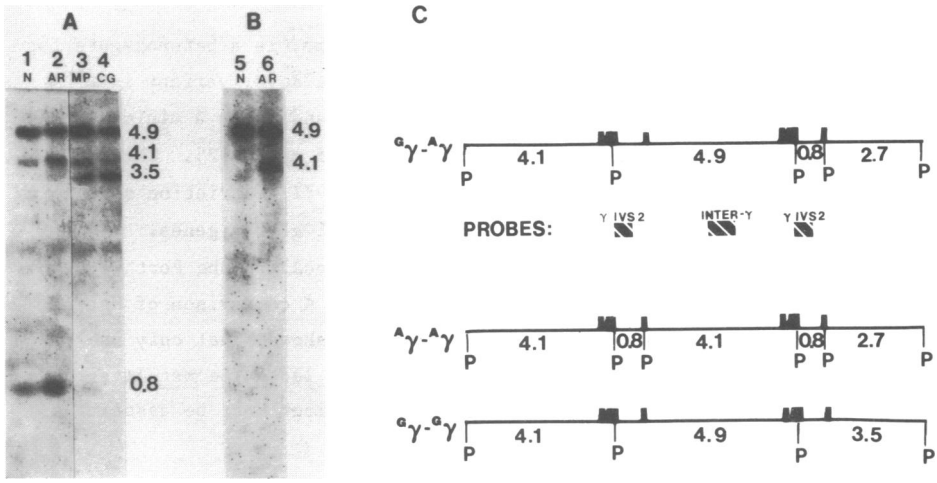

Figure 1. Genomic Mapping of the $^{A}\gamma$-$^{A}\gamma$ and $^{G}\gamma$-$^{G}\gamma$ Gene Arrangement with Pst I.

A. Autoradiograph of DNA samples from the normal control, 563 (lane 1), A. R. (lane 2), M. P. (lane 3) and C. G. (lane 4) after digestion with Pst I, electrophoresis in 0.8% agarose and hybridization to the γIVS2 probe.

B. Autoradiograph of DNA samples from the normal control, 563 (lane 5) and A. R. (lane 6) after digestion with Pst I, electrophoresis in 0.8% agarose and hybridization to the inter-γ probe.

C. Restriction map and fragment sizes in kilobase pairs of the normal ($^{G}\gamma$-$^{A}\gamma$) gene arrangement (upper line), the $^{A}\gamma$-$^{A}\gamma$ gene arrangement (center line) and the $^{G}\gamma$-$^{G}\gamma$ gene arrangement (lower line). Coding regions are represented by heavy bars. The sites of hybridization of the probes are shown by striped bars.

respectively). The lane of DNA from M. P. (lane 3) also has the normal (4.9, 4.1 and 0.8 kbp) bands, although both the unusual 3.5 kbp band and the normal 0.8 kbp band corresponding to the 3' region of the $^{A}\gamma$ gene are of reduced intensity when compared to the other bands. DNA from C. G. (lane 4) contains the normal 4.9, and 4.1 kbp bands while the normal 0.8 kbp band is completely absent and the unusual 3.5 kbp band is of full intensity. These findings suggest that the Pst I restriction site in the third exon of the 3' $^{A}\gamma$ gene is absent in one of the 3' γ globin genes of M. P., and in both 3' γ globin genes of C. G. DNA from M. P. and C. G. digested with the restriction enzymes Eco RI, Pvu II, Sph I, Acc I, and Rsa I gave normal patterns when hybridized to the γIVS2 probe (data not shown). From these observations, and the hematological data presented above, we conclude that a $^{G}\gamma$-$^{G}\gamma$ fetal globin gene arrangement is present in M. P. and that C. G. has this arrangement on both chromosomes. The restriction map corresponding to the $^{G}\gamma$-$^{G}\gamma$ gene arrangement is represented in Figure 1C.

HbF Port Royal

M. P., in adddition to having a $^{G}\gamma$-$^{G}\gamma$ chromosome, is a heterozygote for the structural γ chain variant Port Royal. The Port Royal variant is a Glu → Ala substitution at amino acid position 125, and so as a minimum must include an A to C transversion in the second base of codon 125. This transversion must also result in the loss of an Mst II restriction site which normally occurs in the third exon of the human fetal globin genes. Restriction site mapping can therefore be used to localize the Port Royal variant among the four fetal globin genes of M. P. A comparison of Mst II digests of the cloned γ genes from C. G. and M. P. showed that only the unusual 3' $^{G}\gamma$ gene of the $^{G}\gamma$-$^{G}\gamma$ chromosome of M. P. lacks the Mst II site (data not shown). Consequently, the Port Royal variant must be associated with this unusual 3' $^{G}\gamma$ gene.

DNA Sequence Analysis

As discussed above, the unusual $^{A}\gamma$-$^{A}\gamma$ and $^{G}\gamma$-$^{G}\gamma$ fetal globin gene arrangements must each differ from the normal $^{G}\gamma$-$^{A}\gamma$ gene arrangement by a minimum of a single nucleotide. This single nucleotide difference located in the third exon corresponds to the single coding region difference between the normal $^{G}\gamma$ and $^{A}\gamma$ genes which results in the presence of a glycine ($^{G}\gamma$) or an alanine ($^{A}\gamma$) at amino acid position 136. In order to determine the extent to which the unusual 5' $^{A}\gamma$ gene and 3' $^{G}\gamma$ gene differ from their normal counterparts, we determined the DNA sequence of the region which surrounds the single coding region difference between normal $^{G}\gamma$ and $^{A}\gamma$ genes for both the unusual 5' $^{A}\gamma$ gene of A. R. and the unusual 3' $^{G}\gamma$ gene of M. P. This region includes the 3' portion of the large intervening sequence, the third exon and the 3' untranslated region. Figure 2 presents the sequences obtained for the unusual genes and compares them to the corresponding sequences of two linked $^{G}\gamma$ and $^{A}\gamma$ genes from our normal control DNA (563 chromosome A), and Figure 3 summarizes the nucleotide differences between these four sequences.

As illustrated in Figure 3, the unusual 5' $^{A}\gamma$ gene of A. R. differs from the normal $^{G}\gamma$ gene by only a single nucleotide, yet it differs from the normal $^{A}\gamma$ gene at 16 positions. In this region, the normal $^{G}\gamma$ and $^{A}\gamma$ genes differ at 17 positions, with position 370 being the site of the single coding region difference between the two genes. As shown in Figure 3, the unusual 5' $^{A}\gamma$ gene contains a C at position 370, like the normal $^{A}\gamma$ gene, instead of the G present at this position in the normal $^{G}\gamma$ gene. Thus, the sequence of the unusual 5' $^{A}\gamma$ gene appears to be much more like that of a normal $^{G}\gamma$ gene than of a normal $^{A}\gamma$ gene.

```
                                                                      80
563 Gγ   TGTGTGAGAGCGTGTGTTTCTTTTAACGTTTTCAGCCTACAGCATACAGGGTTCATGGTGGCAAGAAGATAACAAGATTT
AR 5' Aγ       AG                  T           G                   C       A
563 Aγ         TC                  C           A                   G       G
MP 3' Gγ       TC                  C           A                   C       G

                                                                     160
563 Gγ   AAATTATGGCCAGTGACTAGTGCTGCAAGAAGAACAACTACCTGCATTTAATGGGAAAGCAAAATCTCAGGCTTTGAGGG
AR 5' Aγ                       GC  AA                          A
563 Aγ                         TG  GG                          G
MP 3' Gγ                       TG  GG                          G

                                                                     240
563 Gγ   AAGTTAACATAGGCTTGATTCTGGGTGGAAGCTTGGTGTGTAGTTATCTGGAGGCCAGGCTGGAGCTCTCAGCTCACTAT
AR 5' Aγ                           G       T
563 Aγ                             G       G
MP 3' Gγ                           A       G

                                                                     320
                                                  LeuLeuGlyAsnValLeuValThrValLeuAlaIleHisPheGly
563 Gγ   GGGTTCATCTTTATTGTCTCCTTTCATCTCAACAGCTCCTGGGAAATGTGCTGGTGACCGTTTTGGCAATCCATTTCGGC
AR 5' Aγ
563 Aγ
MP 3' Gγ

                                                                     400
                 *Ala                          *Ala
         LysGluPheThrProGluValGlnAlaSerTrpGlnLysMetValThrGlyValAlaSerAlaLeuSerSerArgTyrHi
563 Gγ   AAAGAATTCACCCCTGAGGTGCAGGCTTCCTGGCAGAAGATGGTGACTGGAGTGGCCAGTGCCCTGTCCTCCAGATACCA
AR 5' Aγ                A                             *C
563 Aγ                  A                             *C
MP 3' Gγ                *C                             G

         *Ter                                                        480
563 Gγ   CTGAGCTCACTGCCCATGATGCAGAGCTTTCAAGGATAGGCTTTATTCTGCAAGCAATACAAATAATAAATCTATTCTGC
AR 5' Aγ       TCAC          G
563 Aγ         CTCT          T
MP 3' Gγ       CTCT          T

                                                  526
563 Gγ   TAAGAGATCACACATGGTTGTCTTCAGTTCTTTTTTTTATGTCTTT
AR 5' Aγ  A             G  G       T          TG
563 Aγ    G             A  T       C          CA
MP 3' Gγ  G             A  T       C          CA
```

Figure 2. Nucleotide Sequences of the Normal $^{G}\gamma$ Gene, the Unusual 5' $^{A}\gamma$ Gene, the Normal $^{A}\gamma$ Gene and the Unusual 3' $^{G}\gamma$ Gene.

The 526 bp of DNA sequence representing the 3' portion of IVS2, the third exon and the 3' untranslated region with some flanking DNA is given in full for the normal $^{G}\gamma$ gene (563 $^{G}\gamma$). Nucleotides which differ in the unusual 5' $^{A}\gamma$ gene of A. R. (AR 5' $^{A}\gamma$), the normal $^{A}\gamma$ gene (563 $^{A}\gamma$) and the unusual 3' $^{G}\gamma$ gene of M. P. (MP 3' $^{G}\gamma$) are listed below the sequence of the normal $^{G}\gamma$ gene. The sequences of both the unusual 5' $^{A}\gamma$ gene of A. R. and the unusual 3' $^{G}\gamma$ gene of M. P. were determined in full in this study. The sequences of the normal $^{G}\gamma$ gene and normal $^{A}\gamma$ gene from chromosome A of 563 were previously determined (1). The amino acids encoded by the third exon are printed above the sequence. The starred nucleotides code for the starred amino acids at position 125 and 136 listed above the amino acid sequence. "Ter" is the terminator codon. The polyA addition signal is underlined.

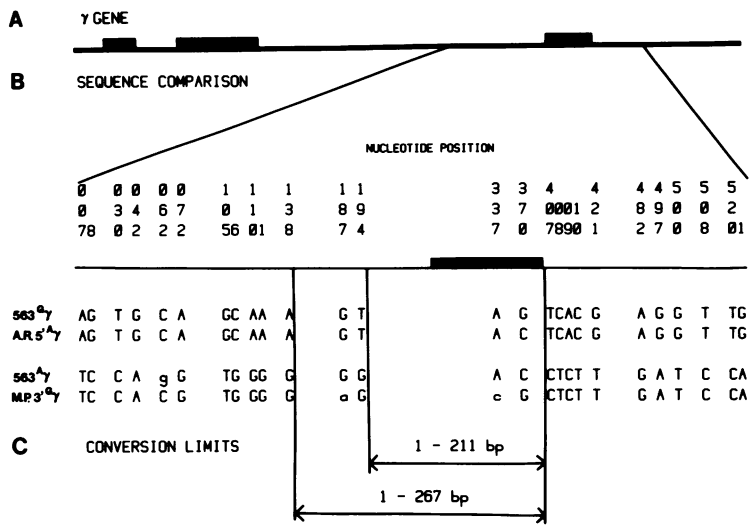

Figure 3. A Comparison of the Normal $^{G}\gamma$ Gene, the Unusual 5' $^{A}\gamma$ Gene, the Normal $^{A}\gamma$ Gene, and the Unusual 3' $^{G}\gamma$ Gene.

A. A representation of a γ globin gene with coding regions represented by heavy bars.

B. The sequence differences between the normal $^{G}\gamma$ gene (563 $^{G}\gamma$), the unusual 5' $^{A}\gamma$ gene of A. R. (AR 5' $^{A}\gamma$), the normal $^{A}\gamma$ gene (563 $^{A}\gamma$) and the unusual 3' $^{G}\gamma$ gene of M. P. (MP 3' $^{G}\gamma$) shown in Figure 2 are tabulated. Nucleotides at positions 62, 187 and 337 represented as lower case letters are "private" mutations (see text). The nucleotide present at position 194 determines the presence (T) or absence (G) of the polymorphic Hind III restriction site in IVS2. Position 337 is the site of the HbF Port Royal variant in the 3' $^{G}\gamma$ gene of M. P. Position 370 is the location of the single coding region difference between the normal $^{G}\gamma$ and $^{A}\gamma$ genes.

C. The two lines represent the limit of the two gene conversions described in the text. For each, the minimal length is a single nucleotide at position 370. The upper line represents the gene conversion of maximal length of 211 bp with maximal borders between positions 195 and 406. The lower line represents the gene conversion of maximal length of 267 bp with maximal borders at positions 139 and 406.

As shown in Figure 3, the unusual 3' $^{G}\gamma$ gene differs from the normal $^{A}\gamma$ gene at 4 positions and from the normal $^{G}\gamma$ gene at 18 positions. The 3' $^{G}\gamma$ gene differs from the normal $^{A}\gamma$ gene at position 370, containing a G like that in normal $^{G}\gamma$ genes instead of the C found at that postion in normal $^{A}\gamma$ genes. The HbF Port Royal variant 3' $^{G}\gamma$ gene contains, as expected, a C at position 337 instead of the A found in normal $^{A}\gamma$ and $^{G}\gamma$ genes, which results in the observed Glu → Ala substitution at amino acid residue 125. The other two differences between the normal $^{A}\gamma$ gene and the unusual 3' $^{G}\gamma$ gene, one at position 62 in the normal $^{A}\gamma$ gene and the other at position 187 in the unusual

3' $^{G}\gamma$ gene, each appear to be due to "private" mutations. At position 62, the normal $^{A}\gamma$ gene used in our comparisons contains a G. All other γ genes sequenced, including the unusual 3' $^{G}\gamma$ gene of M. P. (Figure 3) contain a C at this position (1 and Powers, unpublished data). The simplest explanation is that the normal $^{A}\gamma$ gene used for our comparison has a "private" mutation at position 62. By the same type of reasoning, we conclude that the A at position 187 of the unusual 3' $^{G}\gamma$ gene of M. P. appears to be a "private" mutation since an A has not been seen at this position in eight other genes of both $^{A}\gamma$ and $^{G}\gamma$ types (1 and Powers, unpublished results). In this context, the Port Royal base substitution at position 337 should be regarded as a private mutation, even though its presence in a coding region had led to its previous detection. These comparisons show that the unusual 3' $^{G}\gamma$ gene of M. P. is much more like a normal $^{A}\gamma$ gene than it is like a normal $^{G}\gamma$ gene.

DISCUSSION

We have searched for and found two new arrangements of the human fetal globin genes: the $^{G}\gamma$-$^{G}\gamma$ gene arrangement with two linked $^{G}\gamma$ genes and the $^{A}\gamma$-$^{A}\gamma$ gene arrangement with two linked $^{A}\gamma$ genes. We demonstrate by restriction enzyme mapping and DNA sequence analysis that A. R., the individual with the unusually low $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ globin chain ratio, has two linked $^{A}\gamma$ genes on one chromosome and the normal $^{G}\gamma$-$^{A}\gamma$ gene arrangement on the other chromosome. By similar analysis, we show that C. G. and M. P., both with unusually high $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratios, are homozygous and heterozygous, respectively, for fetal globin regions with two linked $^{G}\gamma$ genes.

We originally described the presence of 4% $^{A}\gamma$ chains in addition to 96% $^{G}\gamma$ chains in the fetal hemoglobin of C. G. (7). However, since we find no evidence for an $^{A}\gamma$ gene in DNA from C. G. by genomic mapping, we presume that the small amount of material originally identified as $^{A}\gamma^{I}$ chain on the HPLC chromatographs was an artifact.

Our mapping data suggest that the event which led to the formation of these unusual fetal globin gene arrangements is of limited extent. Restriction enzyme mapping of the $^{A}\gamma$-$^{A}\gamma$ gene arrangement shows that the unusual 5' $^{A}\gamma$ gene retains all of the tested restriction sites that occur in the neighborhood of normal 5' γ genes. However, there is an extra *Pst* I restriction site in the unusual 5' $^{A}\gamma$ gene at the same place in the third exon as the *Pst* I restriction site characteristic of normal $^{A}\gamma$ genes. Similarly, the unusual 3' $^{G}\gamma$ genes of the $^{G}\gamma$-$^{G}\gamma$ gene arrangement retain all of the tested restriction sites of the normal $^{A}\gamma$ gene except that they lack the *Pst* I

restriction site associated with the third exon of normal $^{A}\gamma$ genes.

Analysis of the nucleotide sequences of the 3' portion of the large intervening sequence, the third exon and the 3' untranslated region of two of these genes demonstrates that both the unusual 5' $^{A}\gamma$ gene of A. R. and the unusual 3' $^{G}\gamma$ gene of M. P. differ from the gene normally found at these positions (a $^{G}\gamma$ and an $^{A}\gamma$ gene, respectively) by a nucleotide substitution. In both genes, the unusual nucleotide is identical to the nucleotide present at the same position in the normal neighboring γ globin gene which raises the possibility that these unusual genes are the result of transfer of DNA sequences from the linked γ gene. Alternatively, the unusual nucleotide in each of these genes may be the product of a simple point mutation.

There appears to be no possible experimental means of determining whether the formation of these unusual chromosomes occurred as the result of point mutations or by the conversion of each of the unusual genes by DNA derived from the normal neighboring gene. However, although we cannot distinguish between these two mechanisms, we can set limits for any simple conversion event which involves the transfer of a stretch of contiguous nucleotides and "non-patchy" repair. In the region analysed here, the normal $^{G}\gamma$ and $^{A}\gamma$ genes differ at 17 positions. The unusual 5' $^{A}\gamma$ gene contains nucleotides characteristic of the normal $^{G}\gamma$ gene at all of these positions except position 370, where it contains the nucleotide characteristic of a normal $^{A}\gamma$ gene. If the unusual 5' $^{A}\gamma$ gene was formed by the transfer of contiguous nucleotides from a normal $^{A}\gamma$ gene, this gene might lie upon the same chromosome (an intrachromosomal gene conversion) or upon a homologous chromosome (an interchromosomal gene conversion). Comparing the nucleotide differences between the unusual 5' $^{A}\gamma$ gene and the normal $^{A}\gamma$ gene presented in Figure 3, we note that the DNA sequence transferred must lie between positions 139 and 406, in a region where, with the exception of the difference at position 370, the normal $^{G}\gamma$ and $^{A}\gamma$ genes are identical. Thus, the maximal length of a gene conversion would be 267 bp with its borders at positions 139 and 406 (see Figure 3C). Formally, the minimal length of gene conversion is one bp, the single nucleotide at position 370.

The occurrence of a polymorphic Hind III restriction site in the large intervening sequence of both γ globin genes (18) generated by the polymorphic nucleotide at position 194 can be used to further define the limits of any possible gene conversion. Either of two nucleotides, a T or a G, are found at position 194. The presence of a T at this position generates the Hind III restriction site, while a G at this position destroys the Hind III restriction

site. The unusual 5' $^{A}\gamma$ gene contains a T at position 194. The linked normal 3' $^{A}\gamma$ gene lacks the Hind III restriction site (data not shown). If the nucleotide sequence transferred to the unusual 5' $^{A}\gamma$ gene was derived from the linked 3' $^{A}\gamma$ gene (by an intrachromosomal gene conversion), the region transferred can not include this Hind III restriction site sequence. Therefore, the maximal intrachromosomal transfer is reduced to 211 bp with its borders between position 195 and 406 (see Figure 3C).

Other individuals with unusually low $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratios have been described by us (19, 20, 21), although to date, we have found only a single example of the $^{A}\gamma$-$^{A}\gamma$ gene arrangement. Two of the other individuals with low $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratios have been shown to have single γ genes of the $^{A}\gamma$ type (19, 20 and A. Metzenberg, personal communication).

We can also calculate the maximal extent of transfer in the formation of the $^{G}\gamma$-$^{G}\gamma$ gene arrangement in M. P. All γ genes of M. P. lack the polymorphic Hind III restriction site in the large intervening sequence (data not shown). Therefore, the boundaries of maximal transfer of DNA sequence from the normal linked $^{G}\gamma$ gene to the unusual 3' $^{G}\gamma$ gene must be at positions 139 and 406, with a maximal length of transfer of 267 bp and minimal length of transfer of 1 bp. The maximal transfer of DNA by interchromosomal gene conversion would be 267 bp if a G was present at position 194 (absence of the Hind III restriction site) in the donor $^{G}\gamma$ gene, and 211 bp if a T was present at position 194 (presence of the Hind III restriction site) in the donor $^{G}\gamma$ gene.

The question arises whether the HbF Port Royal variant (A → C at position 337) arose before or after the $^{G}\gamma$-$^{G}\gamma$ gene arrangement. Since this variant has only been described in association with an elevated $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratio (8, 9, 21, 22), we think it likely that the A → C substitution occurred after the formation of the $^{G}\gamma$-$^{G}\gamma$ gene arrangement.

Individuals with increased $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ globin chain ratios are relatively common. We have observed increased $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratios (> 80%) in 4-8% of Black newborns (21, 22, 23). In general, our data suggest that the high $^{G}\gamma/(^{G}\gamma + ^{A}\gamma)$ chain ratios in Black newborns is the consequence of the presence of at least the two types of $^{G}\gamma$-$^{G}\gamma$ gene arrangements seen in M. P. and C. G.

ACKNOWLEDGEMENTS

We thank Herman F. Harris, Marsha G. Headlee and Alice L. Reese for assistance in obtaining blood and fibroblast samples from the individual described here, Natalie Borenstein for her technical assistance, and Dr. Paula

Henthorn and Debra Endean for helpful discussions. This work was supported by NIH grants GM20069 and AM20120 to O. S and NIH grant HLB-05168 to T. H. J. H. This is paper No. 2703 from the Laboratory of Genetics, University of Wisconsin-Madison and No. 0855 from the Department of Cell and Molecular Biology, Medical College of Georgia.

REFERENCES

1. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Cell 21, 627-638.
2. Shen, S.H., Slightom, J.L. and Smithies, O. (1981) Cell 26, 191-203.
3. Leibhaber, S.A., Gossens, M., and Kan, Y.W. (1981) Nature 290, 26-29.
4. Engel, R. (1981) Nature 290, 191-192.
5. Baltimore, D. (1981) Cell 24, 592-594.
6. Schroeder, W.A., Huisman, T.H.J., Shelton, J.R., Shelton, J.B., Kleihauer, E.F., Dozy, A.M. and Robberson, B. (1968) Proc. Natl. Acad. Sci. USA 60, 537-544.
7. Huisman, T.H.J. and Altay, C. (1981) Blood 58, 491-500.
8. Huisman, T.H.J., Reese, A.L., Gravely, M.E., Harris, H. and Wilson, J.B. (1980) Am. J. Hematology 9, 137-150.
9. Brimhall, B., Vedvick, T.S., Jones, R.T., Ahern, E., Palomino, E. and Ahern, V. (1974) Br. J. Haematology 27, 313-318.
10. Blattner, F.R., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Richards, J.E., Slightom, J.L., Tucker, P.W. and Smithies, O. (1978) Science 202, 1279-1284.
11. Poncz, M., Solowiejczyk, D., Harpel, B., Mory, Y., Schwartz, E., and Surrey, S. (1982) Hemoglobin, 6, 27-36.
12. Southern, E. (1975) J. Mol. Biol. 98, 503-507.
13. Jeffreys, A.J., and Flavell, R.A. (1977) Cell 12, 429-439.
14. Rimm, D.L., Horness, D., Kucera, J., and Blattner, F.R. (1980) Gene 12, 301-309.
15. Hohn, B. (1979) Methods in Enzymology 68, 299-309
16. Twigg, A.J. and Sheratt, D. (1980) Nature 283, 216-218.
17. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
18. Jeffreys, A.J. (1979) Cell 18, 1-10.
19. Sukuman, P.K., Nakatsuji, T., Gardiner, M.B., Reese, A.L., Gilman, J.G. and Huisman, T.H.J. (1983) Nucl. Acids Res. 11, 4635-4643.
20. Nakatsuji, T., Ohba, Y. and Huisman, T.H.J. (1984) Am. J. Hematol. 16, 189-192.
21. Huisman, T.H.J., Schroeder, W.A., Felice, A., Powars, D., and Ringelhann, B. (1977) Nature, 265, 63-65.
22. Huisman, T.H.J., Reese, A.L., Gardiner, M.B., Wilson, J.B., Lam, H., Reynolds, A., Nagle, S., Trowell, P., Zeng, Y-T., Huang, S-Z., Sukumaran, P.K., Miwa, S., Efremov, G.D., Petkov, G., Sciarratta, G.V., and Sansone, G. (1983) Am. J. Hematol., 14, 133-148.
23. Nakatsuji, T., Carver, J., Wilson, J. B., Lam, H., Reese, A.L., Nagle, S., Miwa, S., and Huisman, T.H.J. (1983) Am. J. Hematol., 14, 121-132.