Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 5194–5198, July 1986
Genetics

A gene deletion ending within a complex array of repeated sequences 3′ to the human β-globin gene cluster

(hereditary persistence of fetal hemoglobin/inverted and direct repeats/recombination/restriction fragment length variation)

PAULA S. HENTHORN*†, DIXIE L. MAGER*‡, T. H. J. HUISMAN§, AND O. SMITHIES*¶

*Laboratory of Genetics, University of Wisconsin, Madison, WI 53706; and §Department of Cell and Molecular Biology, Medical College of Georgia, Augusta, GA 30912

Contributed by O. Smithies, March 24, 1986

ABSTRACT **A DNA fragment containing the deletion junction region from an Indian individual with a type of hereditary persistence of fetal hemoglobin has been cloned. Using a probe isolated from this deletion-spanning clone, we located the 3′ breakpoint of the deletion in normal DNA to a region 30 kilobase pairs (kb) downstream of the β-globin gene. The deletion removes 48.5 kb of DNA. Sequences of the deletion junction and of the normal DNA surrounding the 3′ breakpoint were determined and compared to the previously determined sequence of the normal DNA surrounding the 5′ breakpoint. This comparison shows that the deletion was the result of a nonhomologous recombinational event, although there is a 5-base-pair (bp) region of local homology between the normal DNAs at their breakpoints. The 5′ deletion breakpoint occurs in the *Alu* family repeat 3′ to the $^{A}\gamma$-globin gene. The 3′ breakpoint is located within a region that contains the following: a portion of an L1 (*Kpn* I) repeat, a perfect 160-bp palindrome, and a set of 41-bp direct repeats that are found elsewhere in the human genome. A variation in restriction fragment lengths was observed in this region in one family.**

More than 20 different deletions of over 100 base pairs (bp) of DNA have been identified in the human β-globin gene cluster. Many of these deletions remove part or all of the β-globin gene, giving rise to a thalassemia phenotype. However, in some instances of β-globin gene deletion, synthesis of an unusually high percentage of fetal hemoglobin (Hb F) persists throughout adult life and ameliorates the thalassemia phenotype. This condition is known as hereditary persistence of fetal hemoglobin (HPFH). We have undertaken analyses of the HPFH deletions to learn about the mechanisms involved in chromosomal rearrangements in mammalian cells and to identify regions important in the control of β-globin chain expression.

In the present study, we have characterized a deletion in the β-globin gene cluster that is associated with increased Hb F production in an Indian patient variously referred to as $^{G}\gamma^{A}\gamma$-HPFH type III (HPFH-3) or Indian HPFH (1). Patients heterozygous for this deletion have 20–30% Hb F and other hematologic parameters consistent with their classification as HPFH (1, 2). This condition is distinguishable phenotypically from two other forms of HPFH also associated with deletions, HPFH-1 and HPFH-2 (3, 4), primarily on the basis of the ratios of $^{G}\gamma$- and $^{A}\gamma$-globin chains in the Hb F of heterozygotes (1, 2). HPFH-1 and HPFH-2 heterozygotes have $^{G}\gamma$ to $^{A}\gamma$ ratios of about 50:50 and 30:70, respectively, whereas HPFH-3 heterozygotes have a $^{G}\gamma$ to $^{A}\gamma$ ratio of about 70:30 (1, 2, 5). In normal adults, in whom the amount of Hb F averages less than 1%, this ratio is 40:60. Each of these three deletions has been found as a compound heterozygote with β^{+} thalassemia, in which case the HPFH-3 deletion gives rise to a thalassemia phenotype more severe than the phenotypes associated with either the HPFH-1 or HPFH-2 deletions (5). This difference led Wainscoat and coworkers to classify the phenotype associated with the Indian deletion as a form of $(\delta\beta)^{0}$ thalassemia (5).

In previous studies the 5′ breakpoint of the HPFH-3 deletion was mapped to a region 3′ to the $^{A}\gamma$-globin gene but the location of the 3′ breakpoint was not determined (1, 5). In this study we used DNA cloning and sequencing methods to locate the 5′ and 3′ breakpoints precisely and to characterize the deletion event. Localization of both breakpoints of this Indian HPFH deletion also enabled us to determine the length of the deletion and to compare it to other deletions associated with related phenotypes.

MATERIALS AND METHODS

DNA. High molecular weight DNA was prepared from cultured human fibroblasts of a karyotypically normal embryo (563, from Robert DeMars). DNA from the individual B.P.C. (2), who is heterozygous for the HPFH-3 deletion, and from the other individuals studied was isolated from leukocytes by using the procedure of Poncz *et al.* (6) with minor modifications.

Restriction Enzyme Mapping of Genomic DNA. Mapping of genomic DNA and hybridizations were as described by Vanin *et al.* (4). The γIVS probe is a 457-bp *Bam*HI–*Pvu* II fragment consisting principally of the 5′ half of the second intervening sequence of the $^{G}\gamma$-globin gene. The 3′ IH probe shown in Fig. 1*B* is a 680-bp *Fok* I–*Bgl* II fragment isolated from a plasmid subclone of phage IH21. Probe pRK28, provided by R. Kaufman, is a 0.7-kilobase-pair (kb) *Bgl* II–*Eco*RI fragment located 17 kb 3′ to the β-globin gene.

Phage Cloning. Preparative agarose gel electrophoresis was used to isolate 9- to 18-kb fragments from a complete *Xba* I digest of DNA from individual B.P.C. These fragments were ligated to *Xba* I arms of DNA from λ phage Charon 35 (7). Assembly of ligated DNA molecules into phage particles *in vitro* was as described by Maniatis *et al.* (8). Resulting phages were screened without amplification.

DNA Sequencing and Analysis. DNA sequencing was performed as described by Maxam and Gilbert (9). Sequences were determined either by sequencing both DNA strands or by sequencing the same DNA strand at least twice. DNA sequences were analyzed by using the software provided by the University of Wisconsin Genetics Computer Group (10).

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: Hb F, fetal hemoglobin; HPFH, hereditary persistence of Hb F; kb, kilobase pair(s); bp, base pair(s).

†Present address: Department of Human Genetics, University of Pennsylvania, Philadelphia, PA 19104.
‡Present address: Terry Fox Laboratory, British Columbia Cancer Research Centre, Vancouver, BC, V5Z 1L3, Canada.
¶To whom reprint requests should be addressed.

RESULTS AND DISCUSSION

Isolation of a Clone Spanning the Indian HPFH Deletion Junction. The genomic mapping data of Kutlar *et al.* (1) identified a novel 11.5-kb *Xba* I fragment that is detected in DNA from individual B.P.C. using a probe (γIVS) that hybridizes to the second intervening sequence of the γ-globin genes. Their data showed that this *Xba* I fragment spans the HPFH deletion. To isolate a clone containing this fragment, recombinant phages containing size-fractionated *Xba* I fragments from the heterozygous Indian HPFH individual were screened with the γIVS probe. The size fractionation step effectively excluded all other fragments that hybridize to this probe. Six hybridizing phages were isolated, all containing the 11.5-kb *Xba* I fragment identified by genomic mapping (1). A restriction map of one of these phages, IH21, is presented in Fig. 1*B* and shows that this phage contains the fragments expected for a clone spanning the deletion.

Fig. 1 compares the published map of a portion of the β-globin gene cluster with the map of the phage IH21. The *Bgl* II site normally found 5 kb downstream of the ^{A}γ-globin gene is absent in phage IH21, indicating that the 5′ deletion breakpoint occurs on the 5′ side of that *Bgl* II site. The furthest 3′ site common to normal and Indian HPFH DNA is an *Eco*RI site approximately 3 kb 3′ to the ^{A}γ-globin gene. The deletion junction is therefore located within the 1.75-kb *Eco*RI–*Bam*HI fragment of phage IH21. This 1.75-kb fragment was subcloned and its map was compared to that of normal DNA from this region to localize the deletion junction to within 600 bp of the 3′ end of the 1.75 *Eco*RI–*Bam*HI fragment (data not shown). The nucleotide sequence of this region was determined and is presented below.

Localization of the Normal DNA Surrounding the 3′ Breakpoint. To locate the 3′ breakpoint, normal genomic DNA surrounding this breakpoint (which we call the 3′ normal DNA) was mapped by using a unique probe from phage IH21 isolated from the 3′ side of the deletion junction. The location of this probe (3′ IH) is shown in Fig. 1*B*. The normal genomic map derived from the use of this probe is shown in Fig. 1*C*. When the map derived from genomic DNA is compared to the map of phage IH21 there is complete agreement with all restriction enzyme sites located on the 3′ side of the deletion junction. This comparison also agrees with the location of the 3′ deletion breakpoint between the *Nco* I and *Bam*HI sites in Fig. 1*C*.

Our hybridization data and the map of the 3′ normal DNA enabled us to locate the 3′ breakpoint to a region 3′ to the β-globin gene already cloned by Grosveld and coworkers. They have isolated overlapping cosmid clones that extend at least 40 kb 3′ to the β-globin gene. Within these cosmids are two adjacent *Bgl* II fragments of 6.1 and 1.72 kb (F. G. Grosveld, personal communication) that correspond in size to the *Bgl* II fragments seen in the normal DNA surrounding the 3′ end of the Indian HPFH deletion (see Fig. 1 *B* and *C*). In addition, probe 3′ IH hybridizes to a cosmid subclone, GSE340 (kindly provided by F. G. Grosveld), that contains the 6.1- and 1.72-kb *Bgl* II fragments as well as the DNA extending to the *Kpn* I site identified by an asterisk in Fig. 1*C*. The restriction enzyme map of the plasmid GSE340 agrees completely with the map of the region generated by mapping genomic DNA with the 3′ IH probe (not shown). We therefore determined the sequence of the region between *Nco* I and *Bam*HI sites in plasmid GSE340, and we present the sequence below.

These hybridization and mapping data indicate that the 3′ HPFH-3 deletion breakpoint is located 30 kb 3′ to the β-globin gene, in a previously cloned region. Consequently the length of the HPFH-3 deletion can be calculated; it is 48.5 kb.

DNA Sequence Analysis of the Deletion Junction and 5′ and 3′ Breakpoints. The nucleotide sequences of the regions encompassing the 5′ and 3′ breakpoints in normal DNA, and of the deletion junction, were determined and used to characterize the nature of the deletion event. Fig. 2*A* presents an alignment of these sequences. The sequence of the normal DNA in the region of the 5′ breakpoint (which we call 5′ normal DNA, or 5′N) is part of the sequence of the *Alu* family repeat that occurs 3′ to the ^{A}γ-globin gene. This sequence had already been determined in our laboratory (John Devereux, personal communication). The sequence of the deletion junction region (DEL), indicated in Fig. 1*B*, is from a subclone of phage IH21. The 3′ normal DNA sequence (3′N), indicated in Fig. 1*C*, was determined from the plasmid GSE340. Homology between the 5′ normal and deletion junction sequences ends between positions 189 and 190. This position is within the *Alu* family repeat approximately 40 bp from the poly(A) sequence found at the 3′ end of *Alu* sequences (11). The deletion junction sequence and 3′ normal DNA sequences are identical from position 185 to the ends of the sequenced regions, indicating that the 3′ breakpoint is

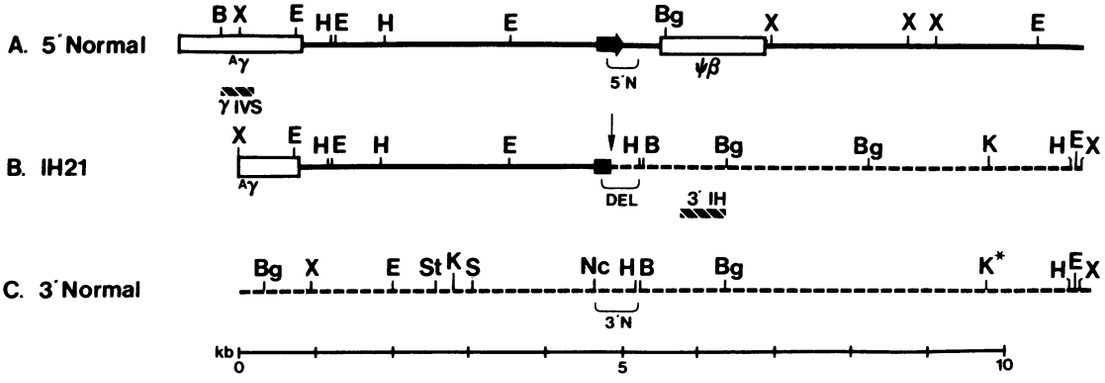

FIG. 1. (*A*) Map of the normal β-globin gene cluster between the ^{A}γ- and δ-globin genes. A solid horizontal arrow shows the location of an *Alu* family repeat. The hatched box shows the region hybridizing to the γIVS probe. The location of the 5′ normal DNA sequence (5′N) presented in Fig. 2*A* is indicated. ψβ indicates a β-globin pseudogene. (*B*) Map of the 11.5-kb *Xba* I fragment insert of phage IH21 that spans the deletion junction (indicated by a vertical arrow). The solid box near the vertical arrow indicates the *Alu* family repeat, part of which has been deleted. The location of the deletion-spanning DNA sequence (DEL) is shown. The location of the 3′ IH probe (see text) is indicated by a hatched box. The portion of the clone 3′ to the deletion junction is represented by a broken line. (*C*) Map of the normal 3′ DNA of the Indian deletion. All restriction sites shown were derived from mapping normal genomic DNA by using the 3′ IH probe. The asterisk indicates a *Kpn* I site discussed in the text. The bracket indicates the 3′ normal sequence (3′N) presented in Fig. 2*A*. Restriction enzyme abbreviations: B, *Bam*HI; Bg, *Bgl* II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn* I; Nc, *Nco* I; S, *Sst* I; St, *Stu* I; X, *Xba* I.

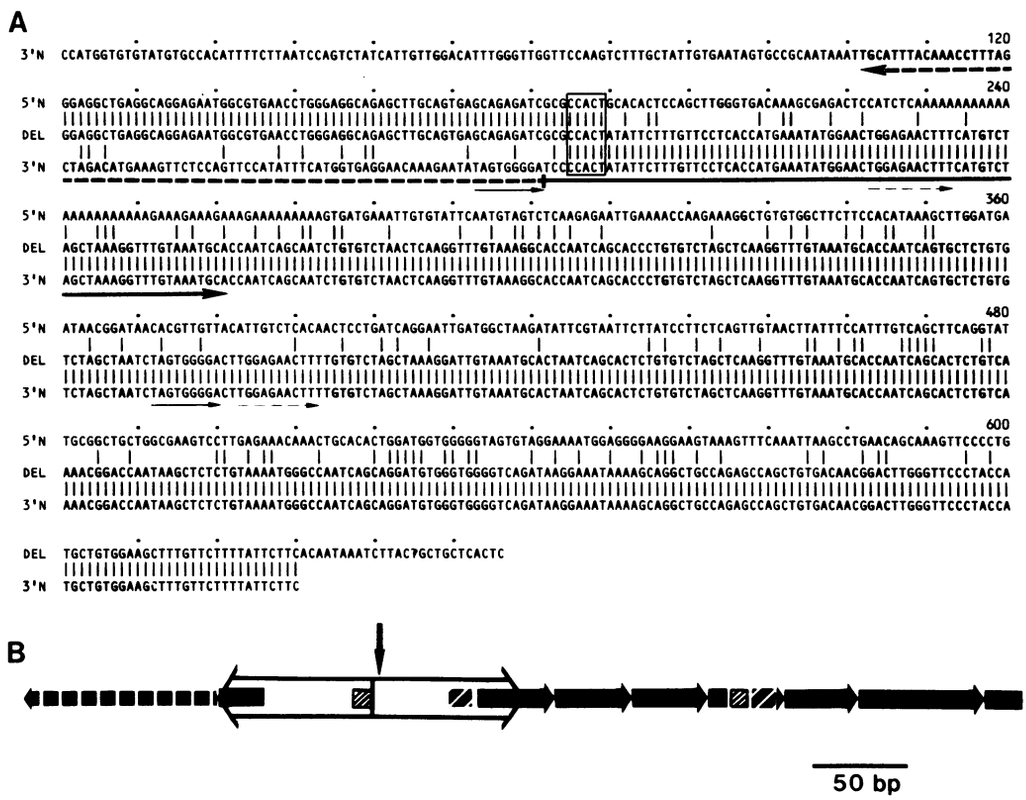

FIG. 2. (*A*) Sequences from the regions of the Indian HPFH deletion junction and of normal DNA surrounding the 5′ and 3′ deletion breakpoints. The 5′ normal sequence (5′N) is from the *Alu* family repeat found 3′ to the ^Aγ-globin gene (see Fig. 1*A*) and was kindly provided by John Devereux. The deletion-spanning sequence (DEL) is from the region of clone IH21 indicated in Fig. 1*B*. The 3′ normal sequence (3′N) is derived from clone GSE340 in the region indicated in Fig. 1*C*. The regions of the deletion junction and 5′ and 3′ breakpoints are enclosed in a box. Vertical lines connect identical bases. The 160-bp palindrome in the 3′N sequence is underlined with heavy solid and broken arrows. Two pairs of short direct repeats discussed in the text are indicated by a pair of light solid and a pair of light broken arrows. (*B*) Representation of the repeated sequences surrounding the HPFH-3 3′ breakpoint (marked with a vertical arrow). The broken bar with an arrow indicates the L1 repeat and its orientation. The large open arrows show the 160-bp palindrome. The solid bars and arrows show sequences belonging to a family of tandem repeats discussed in the text. The hatched boxes indicate sequences that are unrelated to the family of tandem repeats but occur within one copy of these repeats (see text and Fig. 3*A*). These sequences also occur within the 160-bp palindrome.

between positions 184 and 185. Since the five residues, CCACT, between positions 184 and 190 are identical in all three sequences, we assign the deletion junction point and deletion breakpoints to between positions 184 and 190 of the relevant sequences.

With the exception of the 5 bp in common at the breakpoints, the sequences surrounding the 5′ and 3′ breakpoints in normal DNAs show no significant homology to each other. Computer searches for homology included comparing the 5′ and 3′ normal sequences by dot matrix analysis and by alignment of both sequences with and without gaps and comparing the number of matches to the number of matches found when one of the sequences was scrambled. Junctions of this type, in which there are a few nucleotides (1 to 5) of identity at the junction but in which there is no extensive homology between the recombining DNAs, are frequently observed in the products of nonhomologous recombinational events in mammalian cells. For example, Roth *et al.* (12) observed that 5% of junctions involving viral DNA analyzed at the nucleotide sequence level contain as many as 5 bp of homologous sequence at the junction.

Repeated Sequences Near the 3′ Breakpoint. The DNA surrounding the Indian HPFH 3′ breakpoint was found by computer analysis to contain a complex array of repeated sequences, including a part of a *Kpn* or L1 repeat (13), a long palindrome, and a set of short tandem repeats. This array is shown schematically in Fig. 2*B*. Bases 1 through 100 of the 3′ normal sequence are 98% identical to bases 5654 through 5555 of the 6000-bp consensus L1 family sequence (M. Singer, personal communication). The portion of the L1 repeat seen here is in the 3′-to-5′ orientation with respect to the previously defined 5′-to-3′ orientation of L1 repeats. Since the sequence upstream of the 3′ normal sequence shown in Fig. 2*A* was not determined, the total length of the L1 repeat is not known.

A perfect palindrome occurs in the 3′ normal DNA spanning a region of 160 bp beginning at position 102 and ending at position 261. The two 80-bp halves of the palindrome, indicated in Fig. 2*B* by large open arrows and in Fig. 2*A* by a heavy solid arrow and a broken arrow under the normal 3′ sequence, could form an 80-bp stem and loop structure. The midpoint of the palindrome repeat is between positions 181 and 182, only 3 bp from the 3′ breakpoint of the IH deletion. Because at least three bases appear to be required in a DNA loop because of conformational constraints (14), the 3′ breakpoint may be within a single-stranded loop at the end of a stem-loop structure in normal DNA.

A set of short tandem repeats also occurs in the 3′ normal DNA, as shown in Fig. 2*B* from position 237 through position 522. These repeats consist of a 41-bp sequence repeated wholly or partially at least seven times, with one interruption of 24 bases of unrelated sequence at positions 370 through 393. The sequences TAGTGGGGA and TGGAGAACTTT from the 24 bases of unrelated sequence are also found elsewhere in the 3′ normal sequence, at positions 173–181 and 223–233, as indicated by light arrows in Fig. 2*A* and by heavy and light hatched boxes in Fig. 2*B*.

Fig. 3*A* shows an alignment of the short repeats illustrated in Fig. 2*B*. Pairwise comparisons show that the sequences of different copies of the direct repeat are 85–100% identical to

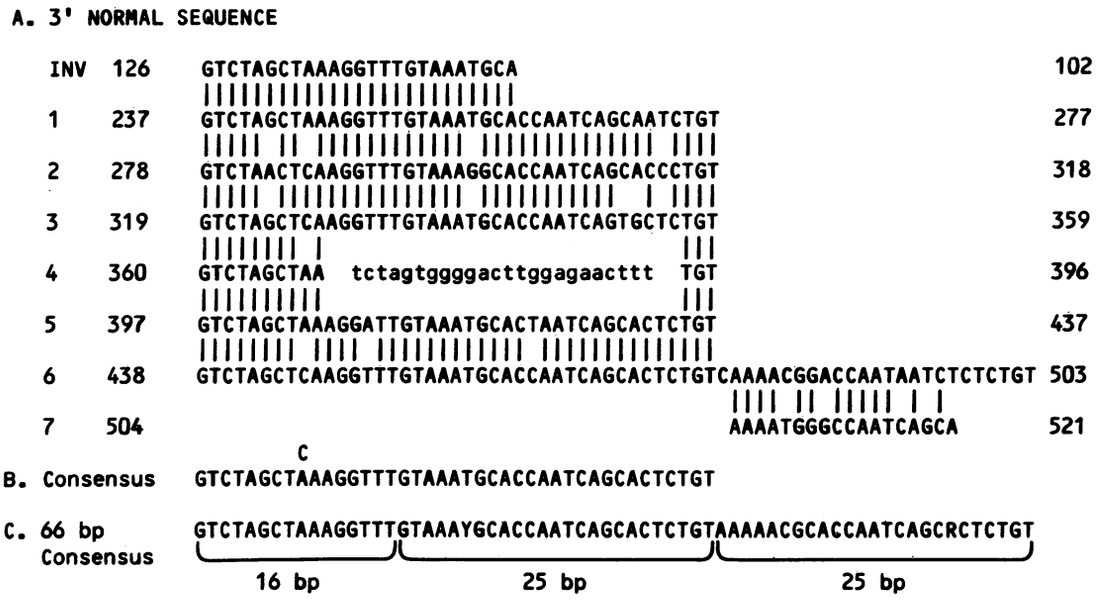

FIG. 3. (*A*) Alignment of the direct repeats in the 3′ normal DNA. The copies of the repeated sequence are aligned to show their homologies to each other and to the previously identified repeat shown below in *C*. All bases from positions 237 through 521 are shown. Sequence segments written in lowercase show no similarity to the repeats. A portion of the repeat found in inverted orientation, positions 126 through 102, is labeled INV. (*B*) A 41-bp consensus sequence of the repeating unit. (*C*) A consensus sequence derived from the 66-bp repeating unit reported by Yang and coworkers (15). The sequence shown is the reverse complement of the sequence they report. The limits of the duplicated 25-bp unit and the 16-bp unit are shown.

each other. The consensus of the 41-bp repeat unit is shown in Fig. 3*B*. Note that the ends of the 160-bp-long palindrome contain 25 bp of this 41-bp repeat. In total there are eight complete or partial copies of this 41-bp repeated sequence; the seven copies between positions 237 and 521 are in the same orientation and the copy between positions 102 and 126 is in the opposite orientation. This complex array suggests that the region has undergone multiple rearrangements.

The short repeats found near the Indian HPFH deletion 3′ breakpoint are closely related to a set of repeats described previously by Yang and coworkers (15). These workers identified an array of eight 66-bp tandem repeats in a clone isolated from human DNA. Their 66-bp element consists of two almost identical 25-bp sequences and an unrelated 16-bp region. A comparison of the 41-bp consensus sequence of the repeats described here (Fig. 3*B*) to the consensus sequence of their 66-bp repeat unit shown in Fig. 3*C* shows 100% homology between them. Although the basic repeat unit characterized here consists of alternating 25-bp and 16-bp segments, rather than the two 25-bp units alternating with one 16-bp segment seen by Yang *et al.* (15), the two repeats are closely related.

The short repeats 3′ to the Indian HPFH deletion are repeated elsewhere in the human genome. A DNA fragment containing these repeated sequences was labeled and hybridized to a Southern transfer of restriction enzyme digests of human genomic DNA. Hybridization of the repeated sequence to genomic DNA gave a smear throughout the lanes (data not shown). No specific hybridizing bands [such as are seen in DNA hybridized to *Kpn* or L1 family repeats (16)] were apparent. Restriction enzymes used included *Hin*dIII, *Bgl* II, *Eco*RI, *Xba* I, *Kpn* I, and *Bam*HI.

Variations in the Lengths of Genomic and Cloned 3′ Normal DNA. The region surrounding the Indian HPFH 3′ breakpoint appears to be particularly susceptible to recombinational events both *in vivo* and during cloning. To analyze genomic DNA arrangements in this region in the general population, DNA from 16 individuals of Black, Caucasian, or Asian origin was digested with *Nco* I and *Bgl* II and hybridized to the 3′ IH probe. The map of the 3′ normal DNA (Fig. 1*C*) predicts that this probe should detect a 1.7-kb *Nco* I–*Bgl* II fragment containing the region of inverted and direct repeats. In 14 individuals only the expected fragment was detected. However, DNA from two related Black individuals (A.R. and T.R.) was found to contain both 1.7- and 1.6-kb *Nco* I–*Bgl* II fragments (Fig. 4). A.R. and T.R. thus appear to be heterozygous for different length *Nco* I–*Bgl* II fragments in this region. Additional mapping of DNA from A.R. and T.R. confirms that the length difference results from a deletion of DNA between the *Nco* I and *Bam*HI sites in this region. In *Sst* I–*Bgl* II digests two bands of 3.3 and 3.2 kb are seen in A.R. and T.R. DNA, but only a single 1.1-kb band is seen in

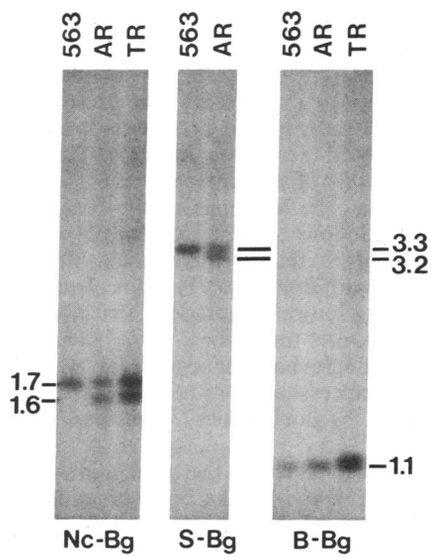

FIG. 4. Autoradiograms demonstrating the restriction fragment length variation detectable with the 3′ IH probe. Lanes labeled 563 contain DNA that does not show variation. Lanes labeled AR and TR contain DNA from two related individuals who are heterozygous for the length variation. Restriction enzymes: Nc, *Nco* I; Bg, *Bgl* II; S, *Sst* I; B, *Bam*HI. Fragment sizes are given in kb.

*Bam*HI–*Bgl* II digests from these individuals (Fig. 4). The simplest explanation of these data is that a small deletion has occurred between the *Nco* I and *Bam*HI sites in one chromosome of these two individuals. Such a deletion would alter the sizes of both the *Nco* I–*Bgl* II and *Sst* I–*Bgl* II fragments but would not affect the size of the *Bam*HI–*Bgl* II fragments (see Fig. 1*C*).

The molecular event that generated this length variation may have involved the 160-bp palindrome, the shorter direct repeats, or both. When we attempted to isolate the normal DNA surrounding the 3′ breakpoint we found that several clones from this region had rearranged during cloning (data not shown). For example, one of these clones had suffered a deletion of 82 bp within the palindrome from position 149 to 230 but was otherwise identical to the normal 3′ sequence presented in Fig. 2*A*. A similar deletion in the inverted repeat in genomic DNA could account for the length difference in individuals A.R. and T.R. Alternatively, rearrangements of the tandem repeats in this region may be responsible for the small deletion. Recombination between directly repeated sequences has been shown in several instances to be a source of length polymorphisms in genomic DNA (17–19).

Effect of the HPFH-3 Deletion on Hemoglobin Expression. Three hypotheses have been proposed to account for the effects of various deletions on the expression of fetal hemoglobin in affected individuals. (For reviews see refs. 20 and 21.) One hypothesis is that a regulatory region exists between the $^{A}\gamma$- and δ-globin genes, the deletion of which results in increased Hb F synthesis (22). This hypothesis is compatible with the phenotype associated with the HPFH-3 deletion, since the deletion removes the region of DNA previously implicated as the control region. A second hypothesis proposes that fetal and adult globin gene chromatin domains have distinct 5′ and 3′ boundaries and that persistent expression of Hb F occurs when the adult domain is partly or totally deleted, leaving the fetal domain functional (23). The HPFH-3 deletion is also compatible with this hypothesis because it could be assumed to remove 48.5 kb of DNA from the domain containing the adult δ- and β-globin genes, leaving the fetal globin domain intact. The third hypothesis suggests that the expression of the fetal globin genes is at least in part controlled by the "new" DNA that has been brought adjacent to them by the deletion event (3). Because the DNA sequences brought near to the fetal globin genes in the HPFH-3 case are different from the "new" sequences in any other deletion described to date, the validity of the third hypothesis cannot be evaluated by the present case.

The three deletions, HPFH-1, -2, and -3, are each associated with different $^{G}\gamma$ to $^{A}\gamma$ chain ratios. These differences could be the result of the different extents of these three deletions; that is, the locations of the breakpoints or the lengths of the deletions could in some way lead to differential expression of the two fetal γ-globin genes. Alternatively, it is possible that the different $^{G}\gamma$ to $^{A}\gamma$ ratios observed are due at least in part to additional mutations that are linked to the deletions, which alter the ratio of $^{G}\gamma$ to $^{A}\gamma$ chain expression. For instance, single base changes that apparently affect γ-globin gene expression have recently been described for both the $^{G}\gamma$- and $^{A}\gamma$-globin genes (for reviews see refs. 21 and 24).

We have precisely located the breakpoints of the deletion associated with HPFH-3. The 5′ breakpoint is in the *Alu* family repeat found between the $^{A}\gamma$- and $\psi\beta$-globin genes. The 3′ breakpoint is found in a region containing several different types of repeated DNA. This breakpoint is located very close to a possible single-stranded loop of a stem-loop structure formed by a palindrome in the area. The possible secondary structure in this region may have facilitated the breakage event. The palindrome lies between an L1 (or *Kpn* I) repeat and a set of tandemly repeated sequences found elsewhere in the genome. We have detected a length difference in the DNA containing these repeats in one family. Clones encompassing this region are difficult to isolate and they frequently rearrange during growth. Taken together, these results suggest that this region is a highly recombinogenic part of the human genome.

We thank Paul Breyer and Natalie Borenstein for invaluable technical assistance and John Devereux for permission to quote some of his unpublished DNA sequences. We also thank Frances Mann, Carol Jokinen, and Sheryl Logan for typing, and members of our laboratory for their comments on this manuscript. We are grateful to Dr. Frank Grosveld for communicating his unpublished work. This work was supported by National Institutes of Health Grants GM20069 and AM20120 to O.S. and by National Heart, Lung, and Blood Institute, U.S. Public Health Service Research Grant HLB-05168 to T.H.J.H. This is paper no. 2878 from the Laboratory of Genetics, University of Wisconsin, Madison, and contribution no. 0950 from the Department of Cell and Molecular Biology, Medical College of Georgia, Augusta, GA.

1. Kutlar, A., Gardiner, M. B., Headlee, M. G., Reese, A. L., Cleek, M. P., Nagle, S., Sukumaran, P. K. & Huisman, T. H. J. (1984) *Biochem. Genet.* **22,** 21–35.
2. Schroeder, W. A., Huisman, T. H. J. & Sukumaran, P. K. (1973) *Br. J. Haematol.* **25,** 131–135.
3. Tuan, D., Feingold, E., Newman, M., Weissman, S. M. & Forget, B. G. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6937–6941.
4. Vanin, E. F., Henthorn, P. S., Kioussis, D., Grosveld, F. & Smithies, O. (1983) *Cell* **35,** 701–709.
5. Wainscoat, J. S., Old, J. M., Wood, W. G., Trent, R. J. & Weatherall, D. J. (1984) *Br. J. Haematol.* **58,** 353–360.
6. Poncz, M., Solowiejczyk, D., Harpel, B., Mory, Y., Schwartz, E. & Surrey, S. (1982) *Hemoglobin* **6,** 27–36.
7. Loenen, W. A. M. & Blattner, F. R. (1983) *Gene* **26,** 171–179.
8. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 264–268.
9. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
10. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
11. Schmid, C. W. & Jelinek, W. R. (1982) *Science* **216,** 1065–1070.
12. Roth, D. B., Porter, T. N. & Wilson, J. H. (1985) *Mol. Cell. Biol.* **5,** 2599–2607.
13. Singer, M. F. & Skowronski, J. (1985) *Trends Biochem. Sci.* **10,** 119–122.
14. Tinoco, I., Uhlenbeck, O. C. & Levine, M. D. (1971) *Nature (London)* **230,** 362–367.
15. Yang, R., Fristensky, B., Deutch, A. H., Huang, R. C., Tan, Y. H., Narang, S. A. & Wu, R. (1983) *Gene* **25,** 59–66.
16. Manuelidis, L. & Biro, P. A. (1982) *Nucleic Acids Res.* **10,** 3221–3239.
17. Stoker, N. G., Cheah, K. S. E., Griffin, J. R., Pope, F. M. & Solomon, E. (1985) *Nucleic Acids Res.* **13,** 4613–4622.
18. Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985) *Nature (London)* **314,** 67–73.
19. Goodbourn, S. E. Y., Higgs, D. R., Clegg, J. B. & Weatherall, D. J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5022–5026.
20. Collins, F. S. & Weissman, S. M. (1984) *Prog. Nucleic Acid Res. Mol. Biol.* **31,** 315–458.
21. Karlsson, S. & Nienhuis, A. W. (1985) *Annu. Rev. Biochem.* **54,** 1071–1108.
22. Huisman, T. H. J., Schroeder, W. A., Efremov, G. D., Duma, H., Mladenovski, B., Hyman, C. B., Rachmilewitz, E. A., Bouver, N., Miller, A., Brodie, A., Shelton, J. R., Shelton, J. B. & Apell, G. (1974) *Ann. N.Y. Acad. Sci.* **232,** 107–124.
23. Bernards, R. & Flavell, R. A. (1980) *Nucleic Acids Res.* **8,** 1521–1534.
24. Gilman, J. G. & Huisman, T. H. J. (1985) *Blood* **66,** 783–787.