CONCISE REPORT

Concordance of a Point Mutation 5′ to the $^{G}\gamma$ Globin Gene With $^{G}\gamma\beta^{+}$ Hereditary Persistence of Fetal Hemoglobin in the Black Population

By Francis S. Collins, Corinne D. Boehm, Pamela G. Waber, Christian J. Stoeckert, Jr, Sherman M. Weissman, Bernard G. Forget, and Haig H. Kazazian, Jr

Hereditary persistence of fetal hemoglobin (HPFH) is a genetically heterogeneous and clinically benign condition characterized by persistent expression of fetal hemoglobin (Hb F) into adulthood. In the $^{G}\gamma\beta^{+}$ type, no major deletions in the globin gene cluster occur; adult heterozygotes produce approximately 20% Hb F, which results from overproduction of $^{G}\gamma$ chains, with no apparent increase in production from the adjacent $^{A}\gamma$ gene. We have recently described a point mutation 202 base pairs 5′ to the cap site of the $^{G}\gamma$ gene in an individual with $^{G}\gamma\beta^{+}$ HPFH. This mutation abolishes a normal *Apa*I restriction endonuclease site, and thus can be detected by blotting of genomic DNA. We present here further data on the *Apa*I mutation: (1) It occurs in six of seven families with $^{G}\gamma\beta^{+}$ HPFH. (2) In three families, detailed haplotype analysis using 11 polymorphic restriction sites in the β globin cluster has been done. The two that carry the missing *Apa*I site are identical but the third, which has a normal *Apa*I pattern, differs from the other two in at least two sites, one of which is a new polymorphic *Nco* I site between the δ and β globin genes. This suggests the possibility of a different HPFH mutation in the third family. (3) The haplotype of the $^{G}\gamma\beta^{+}$ HPFH chromosome carrying the *Apa*I mutation is different from that of 108 β^{A} chromosomes of black individuals that have been tested. (4) The $^{G}\gamma$ *Apa*I site is normal in 61 β^{A} and 109 β^{S} alleles from non-HPFH black individuals, including 22 who share the same haplotype for the intragenic $^{G}\gamma$, $^{A}\gamma$ *Hind*III polymorphisms. These data add support to the possibility that the −202 mutation is actually causative of the $^{G}\gamma\beta^{+}$ HPFH phenotype.

HEREDITARY PERSISTENCE of fetal hemoglobin (HPFH) is a clinically benign and genetically heterogeneous condition, characterized by production of increased amounts of fetal hemoglobin (Hb F) in adult life in the absence of erythropoietic stress.[1,2] Such mutations affecting the normal fetal to adult globin-switching program are of considerable interest, since they provide an accessible tool for studying the differential control of gene expression, and particularly because the reversal of this switch is likely to be of clinical benefit to individuals with hemoglobinopathies affecting the β globin chain.[3,4]

Many individuals with HPFH are found by genomic blotting techniques to carry large deletions of the β globin cluster.[5-9] The location of these deletions has led to various theories about the location of control elements responsible for the normal switching process (reviewed[10]). The size of these deletions (greater than 70 kilobases [kb]) makes it difficult, however, to discern subtle features of the mechanism. In this regard, individuals with HPFH who do not carry such deletions may provide important clues. We have recently reported the cloning of the β globin cluster in the condition known as $^{G}\gamma\beta^{+}$ HPFH found in blacks.[11] Heterozygotes for this condition produce 15% to 25% Hb F, which is almost exclusively of the $^{G}\gamma$ type. Segregation of the $^{G}\gamma\beta^{+}$ HPFH trait in families suggests that the HPFH determinant lies within the β gene cluster.[12-16] We confirmed that there are no detectable rearrangements in a 40-kb fragment containing the fetal ($^{G}\gamma$ and $^{A}\gamma$) and adult (δ and β) genes from the $^{G}\gamma\beta^{+}$ HPFH chromosome, but detected a point mutation 202 base pairs (bp) 5′ to the $^{G}\gamma$ gene. This mutation (Fig 1) destroys an *Apa*I restriction site, and so it can be detected by genomic blotting techniques.

In this paper, we report the further investigation of non-HPFH blacks by genomic blotting techniques, and demonstrate that the specific *Apa*I site is present 5′ to the $^{G}\gamma$ gene in all of their β globin clusters. This increases the likelihood that the −202 $^{G}\gamma$ mutation is actually responsible for the HPFH phenotype, rather than representing a simple polymorphism.

MATERIALS AND METHODS

Subjects

DNA was derived from (1) American black couples, referred for prenatal diagnosis of sickle cell anemia, and their offspring; (2) black individuals homozygous for sickle disease; and (3) three individuals with $^{G}\gamma\beta^{+}$ HPFH, two of which (A and C) are heterozygous for both $^{G}\gamma\beta^{+}$ HPFH and Hb S. Patient A is individual II-2 described in our previous report.[11] Patient B is individual II-1 in the $^{G}\gamma\beta^{+}$ HPFH family described by Farquhar et al,[14] whose leukocyte DNA was kindly provided by Dr George Stamatoyannopoulos

From the Departments of Human Genetics and Medicine, Yale University School of Medicine, New Haven, Conn, and the Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore.

Supported in part by National Institutes of Health grants No. GM32156, AM28376, and AM13983, by the Cooley's Anemia Foundation, and by the Charles E. Culpeper Foundation.

Submitted Sept 4, 1984; accepted Sept 19, 1984.

Address reprint requests to Dr Francis S. Collins, Division of Medical Genetics, Department of Internal Medicine, 4708 Medical Sciences II, Box 015, University of Michigan Medical School, Ann Arbor, MI 48109.

0006-4971/84/6406-0024$03.00/0

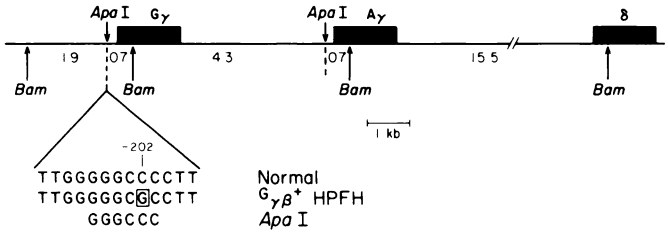

Fig 1. Point mutation at position −202 of the $^{G}\gamma$ gene in $^{G}\gamma\beta^{+}$ HPFH. The mutation abolishes a normal *Apa*I restriction site at this location. The location of other *Apa*I and *Bam*HI sites and the sizes of fragments (in kilobases) expected from a double digest with both of these enzymes are shown.

(University of Washington, Seattle). Patient C is the proposita in the family described by Schwartz[13,16]; DNA was kindly provided by Drs E. Schwartz and S. Surrey, Children's Hospital, Philadelphia.

Definition of Haplotypes Based on Restriction Site Polymorphisms

The polymorphic restriction endonuclease sites used to define haplotypes were studied as previously described for this cohort of patients.[17] Sites examined were: *Hinc*II 5′ to the ϵ globin gene, *Hind*III in IVS-2 of the $^{G}\gamma$ and $^{A}\gamma$ genes, *Hinc*II in the $\psi\beta$ gene and 3′ to it, *Taq*I 5′ to the δ globin gene, *Hinf*I 5′ to the β globin gene, *HgiA*I in the first exon of the β globin gene, *Ava*II in IVS-2 of the β globin gene, and *Hpa*I, *Hind*III, and *Bam*HI 3′ to the β globin gene. In addition, a newly discovered *Nco*I polymorphic site between the δ and β genes was ascertained in some of the alleles. This *Nco*I site is 3.4 kb 5′ to the β gene and its presence results in a 3.4-kb fragment, while its absence yields a 7.4-kb fragment. The polymorphism is detected by use of the 1.9-kb *Bam*HI fragment, which includes the 5′ end of the β gene.

*DNA Analysis at the Apa*I *Site*

It was necessary to perform double digestions with *Apa*I and another enzyme, *Bam*HI, to determine the presence or absence of the *Apa*I site 5′ to the $^{G}\gamma$ gene, since the next upstream *Apa*I site is at least 30 kb away.[10] Digestion was done according to the manufacturer's recommendations (New England Biolabs, Beverly, Mass). DNA isolation, electrophoresis of DNA fragments in agarose gels, transfer to nitrocellulose filters, hybridization with radioactive probes, washing of filters, and autoradiography were carried out as previously described.[17] The probe was JW151, a γ cDNA–containing plasmid.[18]

RESULTS

The normal *Apa*I and *Bam*HI sites surrounding the fetal globin genes are shown in Fig 1. The *Bam*HI sites have not been found to be polymorphic. Thus, if both *Apa*I sites are present, double digestion with *Apa*I and *Bam*HI and Southern blotting with a γ cDNA probe should result in bands of 15.5, 4.3, and 0.7 kb. If the $^{G}\gamma$ *Apa*I site is absent, a 2.6-kb band will appear, and if the $^{A}\gamma$ *Apa*I site is absent, the 4.3-kb band will be replaced by one of 5.0 kb.

Figure 2 shows a representative genomic blot including several non-HPFH individuals and patient A, who shows the presence of the 2.6-kb band corresponding to loss of the $^{G}\gamma$ *Apa*I site. Patient B also showed this 2.6-kb band, but patient C did not (data not shown). The haplotype of the $^{G}\gamma\beta^{+}$ HPFH allele from these patients at 11 polymorphic sites in the β gene cluster is shown in Fig 3; A and B are identical, but C differs at at least two sites. It is noteworthy that the haplotype of patients A and B appears to be unique and has not been detected in 108 β^{A} chromosomes from other black individuals. The association of unique haplotypes with mutant chromosomes has been previously described in the case of the sickle mutation.[17]

A total of 61 β^{A} and 109 β^{S} alleles from black individuals were analyzed by this approach, and all showed the normal *Apa*I and *Bam*HI digestion pattern. Four β^{A} chromosomes examined had a haplotype identical to that of patients A and B at ten of the 11 polymorphic sites, differing only at the 5′-ϵ *Hinc*II site. A fifth β^{A} chromosome was identical in haplotype to the HPFH chromosome except that it lacked the *Hinf*I polymorphic site 5′ to the β gene. This *Hinf*I site is within a region of sequence randomization,[19] and we would consider this chromosome to be closely related to the haplotype of the HPFH chromosomes. It is also noteworthy that 22 of the 61 β^{A} alleles studied had the same pattern at the $^{G}\gamma$ and $^{A}\gamma$ *Hind*III sites (− −) as the $^{G}\gamma\beta^{+}$ HPFH chromosome.

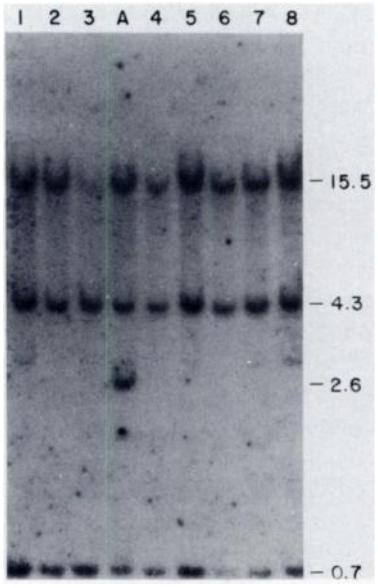

Fig 2. Southern blot, using a γ cDNA probe and digesting genomic DNA with *Apa*I and *Bam*HI. The expected 15.5-, 4.3-, and 0.7-kb bands are seen for eight normal patients (lanes 1–3, 4–8). Patient A is heterozygous for $^{G}\gamma\beta^{+}$ HPFH and yields an additional 2.6-kb band attributable to the loss of the *Apa*I site 5′ to $^{G}\gamma$ (Fig 1). A faint band of unknown origin at about 3.0 kb is seen in all lanes.

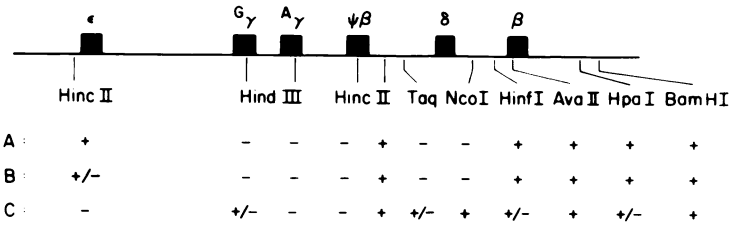

Fig 3. Haplotypes of the $^{G}\gamma\beta^{+}$ HPFH patients. For patients A and B, it was possible to determine the status of most of the various polymorphic sites (−, absent; +, present) specifically for the HPFH allele by family studies and/or analysis of cosmid clones. Patient A was heterozygous at the 5′-ε *Hinc* II and *Hpa*I sites, and the assignments shown reflect the assumption that his Hb S allele is the common sickle haplotype,[17] which is strongly suggested by the status of the other polymorphic sites. Patient B was heterozygous at the 5′-ε *Hinc*II site, as was patient C at several sites; it was not possible to separate alleles at these sites with family studies. Data on the *Hind*III and *Bam*HI sites of patient C have been previously reported[13]; the status of the *Hinc*II, *Ava*II, and *Hpa*I sites was kindly provided by Dr Saul Surrey (personal communication). Patient C must differ from A and B at least at the *Nco* I site and ε *Hinc* II site, and probably also at the $^{G}\gamma$ *Hind*III site, since the only common Hb S haplotype compatible with the ψβ *Hinc*II findings is (− −) at the γ gene *Hind*III polymorphic sites.[17]

Another patient heterozygous for $^{G}\gamma\beta^{+}$ HPFH[12] has been found to carry the same missing *Apa*I site (R.W. Jones and D.J. Weatherall, University of Oxford, England, personal communication), as have 11 more heterozygotes from three families studied by Gilman et al.[20] Haplotype information on these families has not been reported.

*Nco*I Polymorphism

In the course of analysis of clones from the original patients with $^{G}\gamma\beta^{+}$ HPFH, a nucleotide difference which obliterates an *Nco*I site was detected 3.4 kb 5′ to the β globin gene. This polymorphism was important in the study of β^{A}-bearing chromosomes. Upon analysis of genomic DNA of a large number of blacks, Mediterraneans, Chinese, and Asian Indians, it was clear that this DNA polymorphism is limited to blacks. The absence of the *Nco*I restriction site has been found in 13 (24%) of 55 β^{A}-bearing chromosomes and none of 42 β^{S}-bearing chromosomes.

There is no evidence of significant linkage disequilibrium of the *Nco*I site with the polymorphic sites 5′ or 3′ to it. When the *Nco*I site was present, 18 of 35 β^{A} chromosomes examined were positive at the polymorphic *Taq*I site 5′ to the δ gene. When the *Nco*I site was absent, four of 12 β^{A} chromosomes examined were *Taq*I positive. On the 3′ side, when the *Nco*I site was present, 28 of 38 β^{A} chromosomes examined were positive at the polymorphic *Hinf*I site 5′ to the β gene. When the *Nco*I site was absent, ten of 11 β^{A} chromosomes were positive for *Hinf*I. Similarly, no association was found between the presence of the *Nco*I site and the presence of the three different β gene frameworks. While the linkage disequilibrium of the alleles at polymorphic restriction sites in the β globin complex is not as strong in blacks as in Caucasians,[21] it is still statistically significant. The data presented here suggest that the region of recombinational randomization previously described includes not only the *Hinf*I[17,19,21] and *Rsa*I[22] sites 5′ to the β gene, but extends 5′ to at least this *Nco*I site.

DISCUSSION

The HPFH syndromes are particularly important to characterize because of the opportunity they provide to identify *cis*-acting DNA sequences, which are responsible for the normal orderly switch from fetal to adult hemoglobin synthesis that occurs at the time of birth. The cloning and DNA sequence analysis of the $^{G}\gamma\beta^{+}$ HPFH allele has recently revealed a point mutation 202 bp 5′ to the $^{G}\gamma$ gene.[11] The dilemma posed by this discovery will become a common one in the future as more and more sequence data are accumulated on mutant alleles with complex phenotypes. Is the point mutation actually responsible for the phenotype, or is it a polymorphic variant with no functional consequences, which just happens to be linked to the mutant gene? In this particular situation, we are fortunate that the mutation alters a restriction site so that its presence can be sought in genomic DNA without the need for cloning and sequencing of this DNA fragment from multiple individuals.

The results of a population survey of 170 non-HPFH alleles in blacks (61 β^{A} and 109 β^{S}) show that none of these harbor the mutation 202 bp 5′ to the $^{G}\gamma$ gene. Had even one such β gene cluster been found, it would have cast serious doubt on the functional significance of the mutation. It is important to note the status of the surrounding polymorphic markers in the non-HPFH alleles, however, as there is strong linkage disequilibrium present in the cluster of polymorphic markers 5′ to the *Nco*I site.[17] Four β^{A} alleles were found that are identical to the two $^{G}\gamma\beta^{+}$ HPFH alleles at ten of the 11 polymorphic sites in the β globin cluster, the one exception being the 5′-ε *Hinc*II site. Only one β^{A} allele was found that was identical to that of the $^{G}\gamma\beta^{+}$ HPFH

chromosome at all sites 5′ to the δ gene. This chromosome had a haplotype identical to that of the HPFH chromosome except at the *Hinf*I site 5′ to the β gene. One would expect that this β^A chromosome would be the most closely related to the $^G\gamma\beta^+$ HPFH allele, and if the *Apa*I mutation were a relatively recent occurrence, with insufficient time for crossing over within the complex to have dispersed it to different haplotypes, it might still be found in this chromosome. Nevertheless, the $-202\ ^G\gamma$ mutation was not present in this chromosome.

The observation that the two $^G\gamma\beta^+$ HPFH alleles from patients A and B are identical at all polymorphic sites suggests that they derive from a common ancestor, although these individuals are not known to be related, and are apparently separated by at least four meiotic events. This linkage of the HPFH phenotype and the β globin cluster markers is further supportive evidence that the $^G\gamma\beta^+$ HPFH locus is located close to or within the β globin cluster. The $^G\gamma\beta^+$ HPFH allele in patient C occurs on a different haplotype background. One explanation would be that the actual $^G\gamma\beta^+$ HPFH mutation lies outside the β globin cluster, with this patient representing a cross-over. The specific elevation of $^G\gamma$ would be hard to explain by a distant mutation, however, and it seems more likely that patient C is an example of genetic heterogeneity in the $^G\gamma\beta^+$ HPFH phenotype, especially since the family analysis suggests that she may be a new mutation.[13] We would predict that she harbors a different mutation in the 5′ flanking region of $^G\gamma$, which does not abolish the *Apa*I site. Obviously, it will be important to clone and sequence this particular allele.

At present, we favor the hypothesis that the −202 mutation is actually causative, although we cannot completely exclude the possibility that it is a simple polymorphism. An example of this sort of situation has been found in β-thalassemia, in which a silent mutation in the β coding region creating a *Pst*I site is associated with another point mutation that causes β-thalassemia.[23] The *Pst*I abnormality, though itself functionally silent, has so far not been found on a nonthalassemic allele. In the $^G\gamma\beta^+$ HPFH situation, however, the specific elevation of $^G\gamma$ synthesis and the absence of other mutations in the $^G\gamma$ 5′ flanking region by sequence analysis argues in favor of a cause and effect relationship. Definitive proof that the −202 mutation produces HPFH will require demonstration of abnormal behavior of the cloned gene in gene transfer systems.

ACKNOWLEDGMENT

We thank Drs G. Stamatoyannopoulos, E. Schwartz, and S. Surrey for DNA samples; Drs R. Jones, D. Weatherall, J. Gilman, and T.H.J. Huisman for permitting us to cite data on their subject families; and Dr S. Antonarakis for helpful suggestions.

REFERENCES

1. Weatherall DJ, Clegg JB: The Thalassemia Syndromes (ed 3). Oxford, England, Blackwell, 1981

2. Wood WG, Clegg JB, Weatherall DJ: Hereditary persistence of fetal hemoglobin (HPFH) and δβ thalassemia. Br J Haematol 43:509, 1979

3. Wood WG, Weatherall DJ, Clegg JB, Hamblin TJ, Edwards JH, Barlow AM: Heterocellular hereditary persistence of fetal hemoglobin (heterocellular HPFH) and its interaction with β thalassemia. Br J Haematol 36:461, 1977

4. Pembrey ME, Wood WG, Weatherall DJ, Perrine RP: Fetal hemoglobin production and the sickle gene in the oases of Eastern Saudi Arabia. Br J Haematol 40:415, 1978

5. Fritsch EF, Lawn RM, Maniatis T: Characterization of deletions which affect the expression of fetal globin genes in man. Nature 279:598, 1979

6. Tuan D, Murnane MJ, deRiel JK, Forget BG: Heterogeneity in the molecular basis of hereditary persistence of fetal hemoglobin. Nature 258:335, 1980

7. Bernards R, Flavell RA: Physical mapping of the globin gene deletion in hereditary persistence of foetal hemoglobin (HPFH). Nucleic Acids Res 8:1521, 1980

8. Tuan D, Feingold E, Newman M, Weissman SM, Forget BG: Different 3′ end points of deletions causing δβ-thalassemia and hereditary persistence of fetal hemoglobin: Implications for the control of γ-globin gene expression in man. Proc Natl Acad Sci USA 80:6937, 1983

9. Vanin EF, Henthorn PS, Kioussis D, Grosveld F, Smithies O: Unexpected relationships between four large deletions in the human β-globin gene cluster. Cell 35:701, 1983

10. Collins FS, Weissman SM: The molecular genetics of human hemoglobin. Prog Nucleic Acid Res Mol Biol 31:315, 1984

11. Collins FS, Stoeckert CJ, Serjeant GR, Forget BG, Weissman SM: $^G\gamma\beta^+$ Hereditary persistence of fetal hemoglobin: Cosmid cloning and identification of a specific mutation 5′ to the $^G\gamma$ gene. Proc Natl Acad Sci USA 81:4894, 1984

12. Jones RW, Old JM, Wood WG, Clegg JB, Weatherall DJ: Restriction endonuclease maps of the β-like globin gene cluster in the British and Greek forms of HPFH, and for one example of $^G\gamma\beta^+$ HPFH. Br J Haematol 50:415, 1982

13. Balsley JF, Rappaport E, Schwartz E, Surrey S: The γ-δ-β globin gene region in $^G\gamma$-β^+-hereditary persistence of fetal hemoglobin. Blood 59:828, 1982

14. Farquhar M, Gelinas R, Tatsis B, Murray J, Yago M, Mueller R, Stamatoyannopoulos G: Restriction endonuclease mapping of γ-δ-β-globin region in $^G\gamma(\beta)^+$ HPFH and a Chinese $^A\gamma$ HPFH variant. Am J Hum Genet 35:611, 1983

15. Huisman THJ, Miller A, Schroeder WA: A $^G\gamma$ type of hereditary persistence of fetal hemoglobin with β chain production in cis. Am J Hum Genet 27:765, 1975

16. Friedman S, Schwartz E: Hereditary persistence of foetal hemoglobin with β-chain synthesis in *cis* position ($^G\gamma$-β^+-HPFH) in a Negro family. Nature 259:138, 1976

17. Antonarakis SE, Boehm CD, Serjeant GR, Theisen CE, Dover GJ, Kazazian HH: Origin of the β^S globin gene in blacks: The

contribution of recurrent mutation or gene conversion or both. Proc Natl Acad Sci USA 81:853, 1984

18. Wilson JT, Wilson LB, deRiel JK, Villa-Komaroff L, Efstratiadis A, Forget BG, Weissman SM: Insertion of synthetic copies of human globin genes into bacterial plasmids. Nucleic Acids Res 5:563, 1978

19. Orkin SM, Antonarakis SE, Kazazian HH: Polymorphism and molecular pathology of the human beta-globin gene. Prog Hematol 13:49, 1983

20. Gilman JG, Harano T, Nakatsuji T, Bakioglu I, Reese AL, Gardiner MB, Huisman THJ: The ratio of the $^{G}\gamma$ and $^{A}\gamma$ chains: Variations due to anomalies at the molecular level. Ann NY Acad Sci (in press)

21. Chakravarti A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH: Non-uniform recombination within the β-globin gene cluster. Am J Hum Genet (in press)

22. Semenza GL, Malladi P, Surrey S, Delgrosso K, Poncz M, Schwartz E: Detection of a novel DNA polymorphism in the β-globin cluster. J Biol Chem 259:6045, 1984

23. Kazazian HH, Orkin SH, Antonarakis SE, Sexton JP, Boehm CD, Goff SC, Waber PG: Molecular characterization of seven β-thalassemia mutations in Asian Indians. EMBO J 3:593, 1984