*Proc. Natl. Acad. Sci. USA*
Vol. 87, pp. 2092–2096, March 1990
Genetics

# Mutations and a polymorphism in the factor VIII gene discovered by denaturing gradient gel electrophoresis

(hemophilia A)

SCOTT KOGAN AND JANE GITSCHIER*

Howard Hughes Medical Institute and Department of Medicine, University of California, U-426, San Francisco, CA 94143-0724

*Communicated by Leonard S. Lerman, December 1, 1989 (received for review June 18, 1989)*

ABSTRACT **Hemophilia A results from mutations in the gene coding for coagulation factor VIII. We used denaturing gradient gel electrophoresis to screen for mutations in the region of the factor VIII gene coding for the first acidic domain. Amplification primers were designed employing the MELTMAP computer program to optimize the ability to detect mutations. Screening of amplified DNA from 228 unselected hemophilia A patients revealed two mutations and one polymorphism. Rescreening the same population by making heteroduplexes between amplified patient and control samples prior to electrophoresis revealed one additional mutation. The mutations include two missense and one 4-base-pair deletion, and each mutation was found in patients with severe hemophilia. The polymorphism, located adjacent to the adenine branch site in intron 7, is useful for genetic prediction in some cases where the *Bcl* I and *Xba* I polymorphisms are uninformative. These results suggest that DNA amplification and denaturing gradient gel electrophoresis should be an excellent strategy for identifying mutations and polymorphisms in defined regions of the factor VIII gene and other large genes.**

Hemophilia A, an X chromosome-linked bleeding disorder affecting 1 in 5000 males worldwide, results from defects in clotting factor VIII. The wide range of clinical severity exhibited by hemophiliacs plus the high incidence of sporadic cases suggests that hemophilia A results from a heterogeneous collection of mutations. Identification of these mutations has led to accurate diagnosis in some families and has provided insights into function of the factor VIII protein as well as into mutagenesis in humans (1–4). However, mutations have been identified in only about 10% of patients thus far. Progress has been slow due to the large size of the gene (186 kilobases) and to the tissue specificity, low abundance, and large size of the mRNA (5–7). In this report we apply two techniques that help to overcome these problems: DNA amplification (8) and denaturing gradient gel electrophoresis (DGGE) (9). Although DNA amplification is now a routine first step for analyzing DNA, DGGE has been exploited only recently for analyzing globin variants (10, 11) and for detecting hemophilia B mutations in a selected population (12).

DGGE relies on the property of DNA to denature in domains rather than all at once, resulting in a branched structure. The melting temperature ($t_m$) of a domain is dependent on both its base composition and its nucleotide sequence. In practice, melting is achieved by electrophoresis into a polyacrylamide gel containing a concentration gradient of two denaturants, formamide and urea. A DNA fragment advances rapidly until it reaches a denaturant concentration equivalent to the $t_m$ of the first domain, at which point it branches, greatly slowing its progress. The great utility of DGGE is that DNA fragments differing by only one nucleotide will usually exhibit differences in their melting behavior (9, 13). The change in sequence causes the fragments to melt at different denaturant concentrations and thus to be retarded at different positions in the gel.

This study differs from previously published work involving DGGE in that a large unselected patient population was screened for mutations. We searched for mutations in the sequence coding for one of the two acidic domains of factor VIII. These small domains have been implicated as important for factor VIII function because they are the epitopes for neutralizing anti-factor VIII antibodies (14–16), are flanked by cleavage sites for factor VIII activation and inactivation (17, 18), and are not present in the sister protein of factor VIII, factor V (19). An important aspect of this work is the use of the MELTMAP computer program to rationally design amplification primers.

## MATERIALS AND METHODS

**DNA Samples.** DNA was prepared from peripheral blood leukocytes of hemophilia A patients and carriers with unknown factor VIII gene mutations. Samples from a total of 228 individuals were examined, including 155 from the United States, 69 from England, 3 from the People's Republic of China, and 1 from Finland. Of these, 158 suffered from severe hemophilia A, 10 had moderately severe disease, and 25 had mild hemophilia. The severity of the disease was not known in 35, and 26 of the patients were known to produce inhibitor. Five of the 228 samples were female hemophiliacs and two were known carriers of hemophilia. DNAs were also prepared from two control cell lines (5), 4X and AL7, derived from individuals unaffected by hemophilia A.

**Computer Analysis.** Computer analyses were performed on an IBM AT computer with a math coprocessor using Melt87 MELTMAP programs, as described (20). These programs were generously provided by L. Lerman (MIT).

**DNA Amplification.** Samples for DGGE were amplified in 35-μl reaction mixtures containing 10–250 ng of genomic DNA, 17 mM $(NH_4)_2SO_4$, 67 mM Tris·HCl (pH 8.8), 6.7 mM $MgCl_2$, bovine serum albumin (170 μg/ml), 1 mM dGTP, 1 mM dATP, 1 mM dTTP, 1 mM dCTP, 10% (vol/vol) dimethyl sulfoxide, 0.2 μM primer 11.2, and 0.2 μM primer 11.6. Samples were prepared, 1 unit of *Thermus aquaticus* DNA polymerase (Cetus) was added and samples were covered with 20 μl of mineral oil (Vi-Jon, San Leandro, CA) before being denatured for 2 min at 94°C and amplified for 30 cycles on a DNA Thermal Cycler (Perkin–Elmer/Cetus) (21). Each cycle consisted of 1 min at 92°C, 1 min at 50°C, and 1 min at 70°C. Amplifications for the heteroduplex analysis differed slightly, as described in Fig. 4.

**DGGE.** Gel apparatus and conditions were as described (22) except that gel spacers were 1 mm. Each amplified

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: DGGE, denaturing gradient gel electrophoresis; $t_m$, melting temperature; I7P, intron 7 polymorphism.

*To whom reprint requests should be addressed.

Table 1. Primers

| Primer | Sequence |
|---|---|
| 11.2 | 5'-TAATGTACCCAAGTTTTAGG |
| 11.6 | 5'-CCGCGGGCGGGCCTCGCGCCGCGGGCCGGGACCCGCGGCCTGCAGAACATGAGCCAATTC |
| 11.13 | 5'-TATAGAACAGCCTAATATAGCAACAGACTC |

In primer 11.13, the underlined bases are different from those in intron 7. The intron 7 sequence in this region is GACACTCT(G/A)ACATTGT, with the bases shown in bold for alignment. The recognition sequence for *Alw*NI, CAGNNNCTG, will be created only when the guanine allele is present.

sample (5 μl) was mixed with 5 μl of nondenaturing loading buffer (22) and electrophoresed in a 6.5% polyacrylamide gel with a linearly increasing gradient from 37.5% (vol/vol) denaturant to 52.5% denaturant [100% denaturant = 7 M urea/40% (vol/vol) formamide] at 150 V for 5 hr. Gels were stained for 5 min in ethidium bromide (1 μg/ml) and photographed by UV transillumination with Polaroid type 084 black and white film.

**Formation of Heteroduplexes.** Heteroduplexes of normal and patient DNAs were formed and analyzed using several different protocols, as conditions were continually improved. For the first 107 samples, a 7.5-μl aliquot of amplified normal DNA and a 7.5-μl aliquot of amplified patient DNA were combined and 1.6 μl of 500 mM EDTA was added. The mixture was denatured for 10 min at 98°C and allowed to reanneal for 10 min at 55°C, thus forming heteroduplexes. Loading buffer (4 μl) was added and the entire sample was loaded onto the gel. For the remaining samples, 1 unit of *Thermus aquaticus* polymerase was added after the amplified normal and patient DNAs were combined, and samples were amplified for an additional five cycles before loading on the gel, allowing heteroduplexes to form between control and patient samples. Finally, for the gel in Fig. 4, 24 μl of amplified normal DNA and 24 μl of patient DNA were mixed and coamplified an additional 5 cycles. The final product (10 μl) was loaded on the gel after 5 μl of loading buffer had been added.

**DNA Sequencing.** Amplified DNA samples were sequenced directly as described (23). Samples were sequenced in both directions, except for H23 for which the 4-base-pair (bp) deletion was confirmed by hybridization of the amplified sample to an oligonucleotide matching the sequence obtained in one direction.

**Analysis of Polymorphism Data.** By combining our data with that of Wion *et al.* (24) and Janco *et al.* (25) on 453 chromosomes, we found that 326 (72%) were positive (+) at the *Bcl* I site and 127 (28%) were negative (−). Thus, the *Bcl* I polymorphism is predicted to be +/− in 2(0.72)(0.28) or 40% of women seeking diagnostic counseling. When *Bcl* I is +, *Xba* I is + on 120 of 169 chromosomes (71%) and − on 49 (29%) (see Table 2, study 4). Therefore, 2(0.71)(0.29) or 41% of women who are +/+ at *Bcl* I are predicted to be +/− at Xba I. These 41% are $(0.72)^2(0.41)$ or 21% of all cases. When *Bcl* I is −, intron 7 polymorphism (I7P) is guanine on 21 of 44 chromosomes tested in this study (48%) and is adenine on 23 (52%) (see Table 2, study 1). Therefore, 2(0.48)(0.52) or 50% of women who are −/− at *Bcl* I are predicted to be +/− at I7P. These 50% represent $(0.28)^2(0.50)$ or 4% of total cases.

## RESULTS

**Preliminary Experiments.** To produce a suitable fragment for analysis by DGGE, a variety of primer choices were examined with the MELTMAP program. Amplification of genomic DNA with the selected primers (11.2 and 11.6, Table 1) produces a 354-bp DNA fragment that includes a high-melting 40-bp GC-clamp at the 5' end [not part of the genomic DNA sequence, as described in Sheffield *et al.* (11)], the intron 7 region necessary for proper splicing, and the portion of exon 8 that codes for amino acids 318–385 (Fig. 1). Genomic sequences amplified with these primers are readily assayed for single-base mutations because the entire factor VIII region melts as a single melting domain, as shown by the melting map in Fig. 1.

Before screening the patient samples, perpendicular and travel-schedule gels were electrophoresed as described in Myers *et al.* (22) to determine empirically the appropriate concentration gradient and running time for detecting mutations in the 354-bp fragment. These conditions were used for all DGGE analyses. In addition, a variety of test fragments (four other amplified fragments of known DNA sequence) were also examined with the MELTMAP program and on perpendicular gels to determine the relationship between percent denaturant and the predicted $t_m$. The relationship, found to be percent denaturant = $(t_m - 50) \times 2.3$, was used to analyze the present data (see *Discussion*).

**Discovery of a DNA Polymorphism in Intron 7.** An initial experiment was performed to test the appropriateness of gel-running conditions for the intron 7/exon 8 region and to check for common polymorphisms. DNA samples from two control human cell lines, known to differ at the *Bcl* I and *Xba* I intragenic polymorphic sites (21), as well as samples from eight hemophilia A patients were amplified and analyzed. As shown in Fig. 2, the two normal samples migrated to different

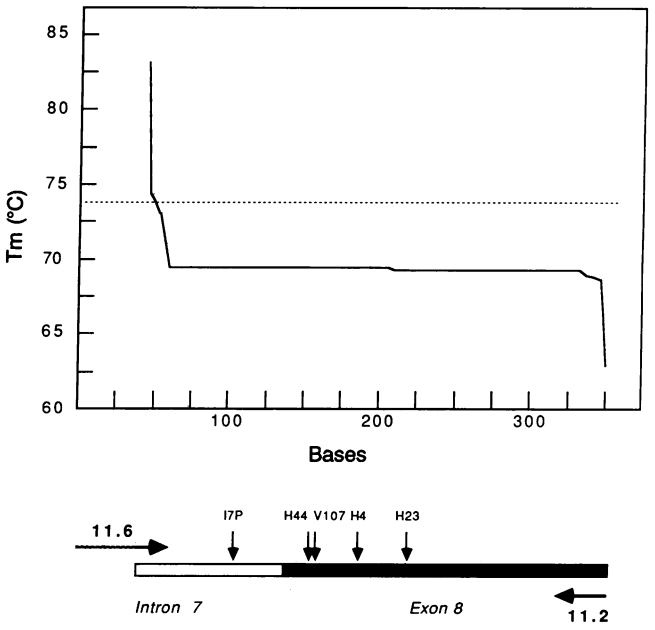


FIG. 1. Thermal stability of amplified sequence coding for the first acidic domain. The $t_m$ is plotted as a function of base position in amplified DNA. Primer 11.6 includes the 40-bp G+C-rich region, indicated by the hatched line, which provides a thermal-stable domain in the resulting DNA fragment. The $t_m$ in the GC-clamp exceeds 85°C, and the $t_m$ for the remainder of the molecule is roughly constant. Values for $t_m$ were determined using the Melt87 program developed by Leonard Lerman (MIT). The broken line indicates the theoretical dissociation temperature for the whole molecule, the temperature at which the dissociation constant for complete strand separation has the value $1 \times 10^{-6}$ M. Also shown are the locations of the polymorphism and four mutations described in this study.

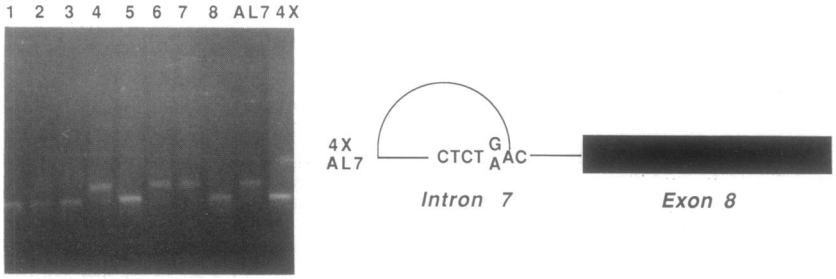


FIG. 2. Polymorphism in intron 7 detected by DGGE. Amplified DNA samples from unselected patients (lanes 1–8) and the AL7 and 4X control DNAs (lanes 9 and 10, respectively) were electrophoresed. The upper band is found in samples containing an adenine at the polymorphic site, and the lower band is found in samples with a guanine. Also depicted is the intron 7 lariat at the adenine branch site, indicating the position of the polymorphism.

positions in the gel. Similarly, the patient samples all migrated to one or the other of these positions, suggesting a DNA polymorphism. Amplified fragments from the two normal samples were sequenced, and a guanine versus adenine polymorphism was discovered 27 nucleotides on the 5′ side of exon 8, within the consensus branch site needed for splicing exons 7 and 8 (Fig. 2). No alteration in gene function is expected as a result of the polymorphism as it conserves the purine normally present at this location (26). Of the 237 independent chromosomes analyzed in this study, 188 (79%) contained the guanine and 49 (21%) contained the adenine. Therefore, if used alone for prenatal diagnosis, the I7P should be informative in 33% (2 × 0.79 × 0.21) of all cases.

We determined that the I7P can give a modest addition to the diagnostic information provided by the most widely used intragenic polymorphisms, *Bcl* I (27) and *Xba* I (24). Some of the 237 chromosomes were analyzed at the *Bcl* I and *Xba* I loci. The results of these pair-wise comparisons are shown in Table 2, studies 1–3. By combining published data with our data on chromosomes not analyzed at the I7P locus (Table 2, study 4), we estimate that the I7P is useful in 50% of cases in which *Bcl* I is uninformative and −/−, or 4% of all cases.

Table 2. Summary of polymorphism data

| Study | *Bcl* I | *Xba* I | *I7P* | *n* |
|---|---|---|---|---|
| 1 | + | | G | 80 |
| | + | | A | 0 |
| | − | | G | 21 |
| | − | | A | 23 |
| | | | Total | 124 |
| 2 | | + | G | 22 |
| | | + | A | 0 |
| | | − | G | 19 |
| | | − | A | 5 |
| | | | Total | 46 |
| 3 | + | + | G | 21 |
| | + | − | G | 12 |
| | − | − | G | 5 |
| | − | − | A | 5 |
| | + | + | A | 0 |
| | + | − | A | 0 |
| | − | + | A | 0 |
| | − | + | G | 0 |
| | | | Total | 43 |
| 4 | + | + | | 120 |
| | + | − | | 49 |
| | − | + | | 1 |
| | − | − | | 48 |
| | | | Total | 218 |

Data in studies 1–3 are from our work. Data in study 4 are from our work combined with that of published reports by Wion *et al.* (24) and Janco *et al.* (25). +, Presence of restriction site; −, absence of restriction site.

Analysis with all three sites should make intragenic diagnosis possible in approximately 65% of cases.

Although I7P does not alter a restriction site, it can be analyzed by restriction digestion using the method of Haliassos *et al.* (28). An amplification primer was designed such that when the guanine allele is present, an *Alw*NI restriction site is created in the amplified product (primer 11.13, Table 1). Thus amplification with primers 11.13 and 11.2, followed by digestion with *Alw*NI, will give fragments of 232 bp for the guanine allele (cleaved) and 260 bp for the adenine allele (uncleaved).

**Screen for Mutations.** The intron 7/exon 8 regions of 226 unselected hemophilia A patients and 2 carriers of hemophilia A were amplified and electrophoresed on denaturing gradient gels. The gels were stained with ethidium bromide and photographed. Also included in this analysis were DNAs from the two cell lines, 4X and AL7, varying in their intron 7 alleles, and a positive control sample, H4, shown to have a nonsense mutation in codon 336 (3) and now found to have a guanine at the I7P. The allelic controls were included in all experiments and samples were judged as having mutations if they varied in mobility from them.

As shown in Fig. 3, two mutations were detected in samples from severe hemophilia A patients, V107 and H23. By sequencing these amplified DNAs directly, V107 was found to have a thymine to cytosine mutation that changes the cysteine at codon 329 to an arginine, and H23 to have a 4-nucleotide deletion within the region coding for the first acidic domain. This latter mutation causes a frameshift and thereby a truncated protein product. The deletion occurred in a repetitive AAT and AAG motif. Small deletions within repeat sequences have been observed in bacterial genomes (29) as well as in human globin genes (30) and are thought to occur by a "slipped mispairing" mechanism during DNA replication. Indeed this particular sequence of the factor VIII gene may be prone to deletion, as another patient had a 2-bp deletion in this same area (31).

**Rescreening by Heteroduplex Analysis.** Although it has been estimated that 80–95% (ref. 13; L. Lerman, personal communication) of mutations are detectable by mobility shifts of GC-clamped fragments with denaturing gradient gels, virtu-

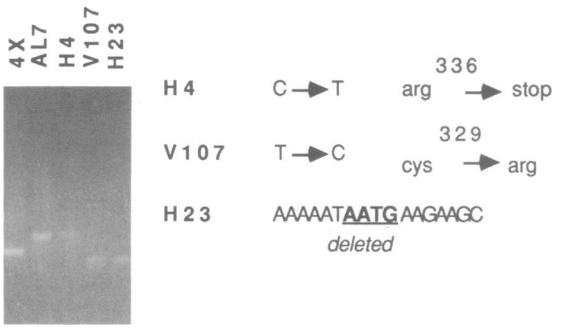


FIG. 3. Mutations identified by DGGE. Samples (as indicated) were amplified and loaded directly on a denaturing gradient gel. The exact bases deleted in H23 cannot be determined because of the repetitive motif.

ally every single-base change should be recognizable when heteroduplexes are formed between patient and control DNAs (32, 33). In heteroduplexed molecules—i.e., DNA molecules in which one strand is derived from a normal control and one strand is derived from a hemophiliac sample—single-base differences destabilize the DNA, causing a significant upward shift on a denaturing gradient gel. The original "homoduplexed" bands are also seen.

Each of the 226 patient samples for which the original DDGE screen did not reveal a mutation was retested by heteroduplexing with the appropriate normal control, and one more mutation was discovered. Fig. 4 shows that the amplified DNA from patient H44 comigrates with the amplified DNA from the 4X control on a denaturing gradient gel. However, when this sample is heteroduplexed with the 4X DNA, new bands appear, indicating a mutation. These bands represent the two combinations of normal and patient strands in the heteroduplexes. Direct sequencing of amplified H44 detected a guanine to cytosine change within codon 326 resulting in a valine to leucine change. Since this is a conservative amino acid change, it is not obvious how this mutation would lead to the patient's severe hemophilia. Interestingly, an unrelated patient of a different ethnic background was observed to have the same change (31).

## DISCUSSION

By screening 228 hemophilia DNA samples for mutations with DGGE, we detected a useful polymorphism and discovered three mutations. Although 3/228 (1.5%) is a small fraction, it should be kept in mind that we screened for mutations in only 2% of the factor VIII mRNA and only 0.2% of the gene. Furthermore, in all previous studies from this laboratory involving most of this patient population, only 15 mutations have been detected (refs. 1, 3, and 34; B. Levinson and J.G., unpublished data).

For each fragment with a sequence variation, we have compared the observed mobility shift (relative to the 4X DNA sample) with that predicted by the MELTMAP computer program. As shown in Table 3, the theoretic $t_m$ is first determined by the MELTMAP program and then compared to that of the 4X sample. The predicted mobility shift can then be determined by the equation in the table legend. Comparison of the predicted and observed mobility shifts indicates that the direction of the predicted shift is correct, although the magnitude of the shift is not consistently comparable. Similar variation was observed in a previous study on globin gene mutations (13). V107 and H23 are predicted to migrate to different positions but were observed to have coincident mobility. Although the predicted and observed mobility shifts for sample H23 are very close, they are not truly comparable since the 4X and H23 samples have a 4-bp size difference, which may also affect the mobilities. In contrast, the predicted $t_m$ for H44 is identical to that of the 4X DNA, consistent with their comigration by homoduplex analysis. Another interesting point is that the H4 sample is observed to denature at a lower concentration of denaturant than AL7, despite having a predicted higher $t_m$. These two samples actually differ in two bases, the H4 mutation as well as the I7P. Since the mobility of fragments with a single-base change differs from that predicted, it is possible that the observed variation is augmented when two base changes are involved.

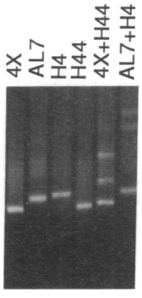


FIG. 4. Mutations identified by heteroduplex analysis on denaturing gradient gels. Differences of note in these amplifications were that they were performed in 50 μl containing 4.5 mM $MgCl_2$, all four dNTPs (each at 0.2 mM), and only 0.15 μM primer 11.2 and 0.15 μM primer 11.6, and did not contain dimethyl sulfoxide. Samples were denatured for 7 min at 95°C prior to the addition of *Thermus aquaticus* DNA polymerase and amplifications consisted of 35 cycles of 15 sec at 94°C, 15 sec at 50°C, and 30 sec at 70°C, using the suggestion of Shiping Cai to diminish background.

Table 3. Comparison of predicted and observed melting differences

| DNA sample | Predicted $t_m$, °C | Predicted $\Delta t_m$, °C | Predicted shift, mm | Observed shift, mm |
|---|---|---|---|---|
| 4X | 69.000 | | | |
| AL7 | 68.773 | −0.227 | −6.8 | −4.5 |
| H4 | 68.824 | −0.176 | −5.3 | −4.7 |
| V107 | 69.288 | +0.288 | +8.6 | +3.0 |
| H23 | 69.066 | +0.066 | +2.0 | +3.0 |
| H44 | 69.000 | 0.0 | 0.0 | 0.0 |

$t_m$ of the domain is represented by the $t_m$ at base 150, and this value is predicted by the Melt87 program. The change in $t_m$ ($\Delta t_m$) is measured relative to the 4X sample. The predicted shift equals the ratio of the change in percent denaturant to the total change in percent denaturant of the gel (15%) times the total gel length (195 mm), where percent denaturant = $(t_m - 50) \times 2.3$.

Our laboratory has experimented with several methods of screening for mutations in genomic DNA of hemophilia A patients, including RNase cleavage, discriminant oligonucleotide hybridization, and DGGE of both total genomic DNA and amplified DNA. Of these, we are particularly encouraged by our progress with denaturing gradient gels using amplified DNA. The procedure is exquisitely sensitive, it does not involve the use of radioactivity, and, in contrast to using DGGE on total human DNA, it allows rapid DNA sequencing once mutations are detected. In addition, we have never observed mistakes introduced by the amplification itself. We have not experimented with the chemical cleavage method of Cotton *et al.* (35), which, although equally sensitive, requires more manipulations and radioactivity.

However, even with the scope of DGGE on amplified DNA, there are choices to be made regarding strategy. First is the issue of screening with homoduplexes versus heteroduplexes. Use of homoduplexes is advantageous in that there are fewer manipulations, the quality of the amplification of each sample need not be assessed prior to DGGE analysis, and multiple regions can be analyzed without confusion on a single gel. However, 5–20% of mutations will be missed. Heteroduplexing has the disadvantages of multiple manipulations and of the need to check the extent of amplification of both control and patient samples prior to forming heteroduplexes (to ensure approximately equal amounts of each). Another possible complication is the presence of polymorphisms, requiring that samples be heteroduplexed with the appropriate normal control (as in this study). Nevertheless, this approach makes possible the detection of virtually all mutations, feasibly in regions as large as 1 kilobase. In addition, the technical obstacles with heteroduplexing might be overcome by analyzing DNA of female carriers of hemophilia, since heteroduplexes form in the amplification reaction. Although two heterozygous carriers were screened in this study because the hemophiliac DNA was not available, heterozygote screening may not be feasible as a general approach. At least 30% of hemophilia patients are sporadic cases, and thus it cannot be assumed that the mother carries the mutation. In addition, the mother's DNA sample is often not available.

A second choice must be made as to whether to amplify and analyze regions of the factor VIII gene or regions of the factor VIII cDNA. Recent work by Sarkar and Sommer (36) has shown it is possible to amplify cDNA from tissue-specific RNAs in peripheral blood leukocytes. This has the advantages of enabling multiple small exons to be analyzed as a

single fragment and of excluding any intronic polymorphisms that could complicate heteroduplex analyses but has the disadvantages of needing to obtain fresh blood samples from hemophiliacs and the inability to discover intronic polymorphisms useful for genetic diagnosis.

Indeed one of the unexpected and beneficial outcomes of this study was the discovery of a two-allele polymorphism in intron 7. Most intragenic polymorphisms have been found to be in linkage disequilibrium with the *Bcl* I polymorphism, including the *Msp* I and *Hin*dIII polymorphisms, as well as two sequence polymorphisms (21, 37, 38). The *Bgl* I polymorphism also exhibits substantial linkage disequilibrium with the *Bcl* I site (39). One exception, the *Xba* I polymorphism, allows analysis in approximately half of those cases for which the proband is uninformative by virtue of being homozygous for the presence of the *Bcl* I site. The I7P is complementary in that it is predicted to be informative in half of those cases for which the proband lacks the *Bcl* I site on both X chromosomes. Although this adds a modest 4% to the overall level of intragenic diagnosis, it can be selectively analyzed in appropriate cases. Such analysis can be performed using DGGE or as also described herein with the use of amplification and digestion.

**Note Added in Proof.** (*i*) A pedigree analysis indicates that the missense mutation in patient H44 is not present in his mother or sisters. As his hemophilia is a sporadic occurrence, the coincident appearance of hemophilia and the mutation suggests that the valine to leucine substitution is responsible for severe hemophilia. (*ii*) A similar study of hemophilia A mutations has been performed (40).

We are extremely grateful to Richard Myers for his continued encouragement and interest in this project, for suggesting the use of GC-clamp oligonucleotides, and for expert commentary on this manuscript. We greatly appreciate Leonard Lerman's advice and provision of MELTMAP programs. We also thank Marie Doherty for oligonucleotide synthesis, Mark Gray for detailed protocols, Shiping Cai for suggesting the heteroduplex conditions shown in Fig. 4, William Wood for communicating unpublished factor VIII sequence, and Susan Kenwrick, Barbara Levinson, Aida Metzenberg, and Kevin Shannon for careful reading of the manuscript. This project was made possible by Edward Tuddenham, Eleanor Goldman, Robert Janco, John Phillips III, Pamela Orlando, Marion Koerper, Marc Shuman, Susan Karp, and Shiping Cai who have collected patient blood samples. J.G. is an assistant investigator with the Howard Hughes Medical Institute.

1. Gitschier, J., Wood, W. I., Tuddenham, E. G. D., Shuman, M. A., Goralka, T. M., Chen, E. Y. & Lawn, R. M. (1985) *Nature (London)* **315,** 427–430.
2. Antonarakis, S. E., Waber, P. G., Kittur, S. D., Patel, A. S., Kazazian, H. H., Mellis, M. A., Counts, R. B., Stamatoyannopoulos, G., Bowie, E. J. W., Fass, D. N., Pittman, D. D., Wozney, J. M. & Toole, J. J. (1985) *N. Engl. J. Med.* **313,** 842–848.
3. Gitschier, J., Kogan, S., Levinson, B. & Tuddenham, E. G. D. (1988) *Blood* **72,** 1022–1028.
4. Gitschier, J. (1988) *Am. J. Hum. Genet.* **43,** 274–279.
5. Wood, W. I., Capon, D. J., Simonsen, C. C., Eaton, D. L., Gitschier, J., Keyt, B., Seeburg, P. H., Smith, D. L., Hollingshead, P., Wion, K. L., Delwart, E., Tuddenham, E. G. D., Vehar, G. A. & Lawn, R. M. (1984) *Nature (London)* **312,** 330–337.
6. Toole, J. J., Knopf, J. L., Wozney, J. M., Sultzman, L. A., Buecker, J. L., Pittman, D. D., Kaufman, R. J., Brown, E., Shoemaker, C., Orr, E. C., Amphlett, G. W., Foster, W. B., Coe, M. L., Knutson, G. J., Fass, D. N. & Hewick, R. M. (1984) *Nature (London)* **312,** 342–347.
7. Wion, K. L., Kelly, D., Summerfield, J. A., Tuddenham, E. G. D. & Lawn, R. M. (1985) *Nature (London)* **317,** 726–729.
8. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230,** 1350–1354.
9. Fischer, S. G. & Lerman, L. S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1579–1583.
10. Myers, R. M., Lumelsky, N., Lerman, L. S. & Maniatis, T. (1985) *Nature (London)* **313,** 495–498.
11. Sheffield, V. C., Cox, D. R., Lerman, L. S. & Myers, R. M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 232–236.
12. Attree, O., Vidaud, D., Vidaud, M., Anselem, S., Lavergne, J. M. & Goosens, M. (1989) *Genomics* **4,** 266–272.
13. Myers, R. M., Fischer, S. G., Lerman, L. S. & Maniatis, T. (1985) *Nucleic Acids Res.* **13,** 3131–3145.
14. Ware, J., Toomey, J. R. & Stafford, D. W. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3165–3169.
15. Foster, P. A., Fulcher, C. A., Houghten, R. A., Mahoney, S. G. & Zimmerman, T. S. (1988) *J. Clin. Invest.* **82,** 123–128.
16. Shima, M., Fulcher, C. A., Houghten, R. A., de Graaf Mahoney, S. & Zimmerman, T. S. (1988) *J. Biol. Chem.* **263,** 10198–10203.
17. Eaton, D., Rodriquez, H. & Vehar, G. A. (1986) *Biochemistry* **25,** 505–512.
18. Pittman, D. D. & Kaufman, R. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2429–2433.
19. Kane, W. H. & Davie, E. W (1988) *Blood* **71,** 539–555.
20. Lerman, L. S. & Silverstein, K. (1987) *Methods Enzymol.* **155,** 482–501.
21. Kogan, S. C., Doherty, M. & Gitschier, J. (1987) *N. Engl. J. Med.* **317,** 985–990.
22. Myers, R. M., Maniatis, T. & Lerman, L. S. (1987) *Methods Enzymol.* **155,** 501–527.
23. Wong, C., Dowling, C. E., Saiki, R. K., Higuchi, R. G., Erlich, H. A. & Kazazian, H. H., Jr. (1987) *Nature (London)* **330,** 384–386.
24. Wion, K. L., Tuddenham, E. G. D. & Lawn, R. M. (1986) *Nucleic Acids Res.* **14,** 4535–4542.
25. Janco, R. L., Phillips, J. A., III, Orlando, P. J., Woodard, M. J., Wion, K. L. & Lawn, R. M. (1987) *Blood* **69,** 1539–1541.
26. Ruskin, B., Krainer, A. R., Maniatis, T. & Green, M. R. (1984) *Cell* **38,** 317–331.
27. Gitschier, J., Drayna, D., Tuddenham, E. G. D., White, R. L. & Lawn, R. M. (1985) *Nature (London)* **314,** 738–740.
28. Haliassos, A., Chomel, J. C., Tesson, L., Baudis, M., Kruh, J., Kaplan, J. C. & Kitzis, A. (1989) *Nucleic Acids Res.* **17,** 3606.
29. Farabaugh, P. J., Schmeissner, U., Hofer, M. & Miller, J. H. (1978) *J. Mol. Biol.* **126,** 847–863.
30. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21,** 653–668.
31. Wong, C., Inaba, H., Antonarakis, S. E., Fujimaki, M. & Kazazian, H. H. (1988) *Am. J. Hum. Genet.* **43,** 100 (abstr.).
32. Myers, R. M. & Maniatis, T. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51,** 275–284.
33. Myers, R. M., Sheffield, V. C. & Cox, D. R. (1988) in *Genomic Analysis: A Practical Approach*, ed. Davies, K. E. (IRL, Oxford), pp. 95–138.
34. Levinson, B., Janco, R., Phillips, J., III, & Gitschier, J. (1987) *Nucleic Acids Res.* **15,** 9797–9805.
35. Cotton, R. G. H., Rodrigues, N. R. & Campbell, R. D. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4397–4401.
36. Sarkar, G. & Sommer, S. S. (1989) *Science* **244,** 331–334.
37. Ahrens, P., Kruse, T. A., Schwartz, M., Rasmussen, P. B. & Din, N. (1987) *Hum. Genet.* **76,** 127–128.
38. Youssoufian, H., Phillips, D. G., Kazazian, H. H., Jr., & Antonarakis, S. E. (1987) *Nucleic Acids Res.* **15,** 6312.
39. Janco, R. L., Phillips, J. A., III, Orlando, P., Davies, K. E., Old, J. & Antonarakis, S. E. (1986) *Lancet* **i,** 148–149.
40. Traystman, M. D., Higuchi, M., Kasper, C. K., Antonarakis, S. E. & Kazazian, H. H. (1990) *Genomics* **6,** 293–301.