Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 8307–8311, October 1991
Genetics

Molecular characterization of mild-to-moderate hemophilia A: Detection of the mutation in 25 of 29 patients by denaturing gradient gel electrophoresis

MIYOKO HIGUCHI*, STYLIANOS E. ANTONARAKIS*, LAURA KASCH*, JOHANNES OLDENBURG†, EFFROSINI ECONOMOU-PETERSEN*, KLAUS OLEK†, MORIO ARAI‡, HIROSHI INABA‡, AND HAIG H. KAZAZIAN, JR.*§

*Center for Medical Genetics, Departments of Pediatrics and Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD 21205; †Institute for Experimental Hematology, University of Bonn, Bonn, Federal Republic of Germany; and ‡Department of Clinical Pathology, Tokyo Medical College, Tokyo, Japan

Communicated by Leonard S. Lerman, May 6, 1991

ABSTRACT To date it has been difficult to characterize completely a genetic disorder, such as hemophilia A, in which the involved gene is large and unrelated affected individuals have different mutations, most of which are point mutations. Toward this end, we analyzed the DNA of 29 patients with mild-to-moderate hemophilia A in which the causative mutation is likely to be a missense mutation. Using computer analysis, we determined the melting properties of factor VIII gene sequences to design primer sets for PCR amplification and subsequent denaturing gradient gel electrophoresis (DGGE). A total of 45 primer sets was chosen to amplify 99% of the coding region of the gene and 41 of 50 splice junctions. To facilitate detection of point mutations, we mixed DNA from two male patients, and both homoduplexes and heteroduplexes were analyzed. With these 45 primer sets, 26 DNAs containing previously identified point mutations in the factor VIII gene were studied, and all 26 mutations were easily distinguishable from normal. After analyzing the 29 patients with unknown mutations, we identified the disease-producing mutation in 25 (86%). Two polymorphisms and two rare normal variants were also found. Therefore, DGGE after computer analysis is a powerful method for nearly complete characterization of disease-producing mutations and polymorphisms in large genes such as that for factor VIII.

Hemophilia A, an X chromosome-linked recessive bleeding disorder that affects ≈1 in 10,000 males in all population groups, is due to a deficiency of factor VIII procoagulant activity (1, 2). A large number of mutations in the factor VIII gene causing hemophilia A (designated *F8C*) have been reported (see ref. 3 for a review). Gross gene rearrangements, such as deletions, insertions, and duplications, account for only 5% of the molecular defects in hemophilia A patients, while point mutations are thought to account for the remaining 95% of defects.

Detection of all possible point mutations in *F8C* is hampered by its large size, its many exons, and the high frequency of *de novo* mutations that result in different mutations in unrelated affected individuals. These factors make hemophilia A a model for the development of methods for the complete characterization of disease-producing mutations in a single gene disorder caused by an almost unlimited variety of point mutations. Recently, a number of screening methods (4–7) have been applied in a limited fashion to identify point mutations in *F8C*. To date, comprehensive mutation analysis has been accomplished only in the factor IX gene, *F9*, defects of which cause hemophilia B (8). However, *F9* contains only eight exons and a coding region of 1.4 kilobases (kb) compared with 26 exons and a 7-kb coding region in *F8C* (2, 9).

Denaturing gradient gel electrophoresis (DGGE) separates DNA fragments according to their melting properties in a gel system that contains a linear gradient of DNA denaturants. Partially melted DNA fragments are required for separation of normal and mutant DNA, where mutations fall in the melted region of the fragments. For example, base changes in all but the highest temperature melting domain are usually detected by DGGE (10, 11). Attachment of a highest melting domain, called a "GC-clamp," to fragments of interest improves the gel system so that base changes in the highest temperature melting domain become accessible to DGGE analysis (12, 13). In addition, the use of heteroduplexes between normal and mutant DNA that contain a mismatch increases the resolution by DGGE so that virtually all possible base changes can be detected, even when a base change does not produce a shift in the mobility of the mutant homoduplex (14, 15).

In this report we demonstrate the usefulness of DGGE analysis to detect all possible mutations in the coding region of *F8C*. We analyzed 29 patients with mild-to-moderate hemophilia A, which accounts for 30–40% of the disease. We describe 20 disease-producing mutations in 25 patients.

MATERIALS AND METHODS

Subjects. The initial group of subjects consisted of 31 patients with a clinical diagnosis of mild-to-moderate hemophilia A. However, we found frameshift mutations in two patients—deletion of two adenosine a residues at codons 2154/2155 in JH-69 and deletion of a single adenosine a at codons 1439–1441 in JH-142. When clotting assays were repeated in these patients, both had no factor VIII activity and were reclassified as severe hemophiliacs. Thus, a total of 29 patients with mild-to-moderate hemophilia A were analyzed, including 15 German patients (16) and 14 Japanese patients (17). When genomic DNA was analyzed by Southern blot analysis with several restriction endonucleases including *Taq* I and three cDNA fragments as probes (18), no molecular defects were found (16, 17). Factor VIII activity in these 29 patients ranged from 2.4% to 38% of normal (Table 1).

Computer Analysis. Computer analysis was carried out with the programs MELT 87 and SQHTX, provided by L. Lerman (Massachusetts Institute of Technology), to aid in the placement of PCR primers (20).

Abbreviations: DGGE, denaturing gradient gel electrophoresis; IVS, intervening sequence(s); CRM, cross-reacting material.

§To whom reprint requests should be addressed at: The Johns Hopkins Hospital, Center for Medical Genetics, CMSC 10-110, 600 North Wolfe Street, Baltimore, MD 21205.

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.

Table 1. Summary of mutations identified

Patient	Factor VIII activity, %	Exon	Codon	Nucleotide change	Amino acid change	Domain	Transition at CpG
JH-139	19	4	185	AAA → ACA	Lys-166 → Thr	A_1	No
JH-143	3.2	5	224	GGG → TGG	Gly-205 → Trp	A_1	No
		IVS-5 donor site −1 AAG/GT → AAT/GT					
JH-87	14–16	7	314	ACT → GCT	Thr-295 → Ala	A_1	No
JH-131	10.5	9	431	TTG → TTT	Leu-412 → Phe	A_2	No
JH-125/-126	2.7/3.5	10	492	TAT → TGT	Tyr-473 → Cys	A_2	No
JH-94/-95/-96	14.5/18/17–38	11	546	CGG → TGG*	Arg-527 → Trp	A_2	Yes
JH-97/-98	6.7/4.2	11	550	CGC → TGC*	Arg-531 → Cys	A_2	Yes
JH-144†	9.2	11	550	CGC → GGC	Arg-531 → Gly	A_2	No
JH-123	6.8	11	584	CAG → AAG	Gln-565 → Lys	A_2	No
		IVS-11 donor site −3 CAG/GT → AAG/GT					
JH-136	14[3]	13	663	GCA → GTA	Ala-644 → Val	A_2	No
JH-132	2.4	14	1057	GAG → AAG	Glu-1038 → Lys	B	No
JH-103	15	16	1844	CCC → TCC[3]	Pro-1825 → Ser	A_3	No
JH-104	9–18	16	1862	CTG → CTA	Leu-1843 → Leu	A_3	No
		IVS-16 donor site −1 CTG/GT → CTA/GT*					
JH-130	1–5	17	1867	CAC → CGC	His-1848 → Arg	A_3	No
JH-137/-138‡	3.4/2.6	19	2016	CGG → TGG	Arg-1997 → Trp	A_3	Yes
JH-133	11	22	2120	TTT → TTG	Phe-2101 → Leu	C_1	No
JH-134	5–8	22	2138	TCC → TAC	Ser-2119 → Tyr	C_1	No
JH-65	5–7	23	2169	CGT → CAT	Arg-2150 → His	C_1	Yes
JH-68	7.4	23	2178	CGC → TGC	Arg-2159 → Cys	C_1	Yes
JH-75	7.5	26	2319	CCG → CTG	Pro-2300 → Leu	C_2	Yes

*These nucleotide substitutions were also independently detected in our lab by single-strand DNA gel electrophoresis (19).
†A second nucleotide substitution G → A was identified at position −19 in intervening sequence (IVS)-2.
‡In JH-138, a Val-2223 → Met change was also identified (see *Discussion*).

Amplification of Genomic DNA. High molecular weight leukocyte DNA was amplified by using *Taq* DNA polymerase (Cetus) (21). PCR was performed in volumes of 40 μl for DGGE analysis and 100 μl for sequence analysis [100 μl of the reaction mixture contained 200–400 ng of genomic DNA, 400 nM each PCR primer, 200 μM each dNTP, 10 mM Tris·HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.02% gelatin, and 2 units of *Taq* DNA polymerase] for 35 cycles, each cycle consisting of 94°C for 20 sec, a thermal transition from 94°C to the annealing temperature over 2 min, 52–60°C for 45 sec, and 72°C for 30–60 sec, followed by extension at 72°C for 10 min. For amplification of exon 14 regions, a first amplification of 30 cycles comprised a step-cycle program (94°C for 20 sec, 55°C for 30 sec, and 72°C for 3 min) in a total volume of 100 μl. The 3.2-kb PCR product containing the entire exon 14 was used as template DNA for amplification of appropriate exon 14 subfragments. See the legend to Fig. 1 for primers.

DGGE Analysis. PCR products (≈80 ng of each) from two patients were combined to form heteroduplexes. After heat denaturation at 95°C for 5 min, the DNA solution was slowly cooled to room temperature (>30 min) and subjected to DGGE under conditions determined empirically for each PCR product (11, 15). DNA was loaded onto a 6.5% polyacrylamide gel (14 cm × 19 cm, 0.75 mm thick) containing a linear gradient of denaturants and electrophoresed at 2–4 V/cm for 16–23 hr. The gradient difference in denaturants used was 20% [100% denaturants = 7 M urea/40% (vol/vol) formamide]. Gels were then stained in ethidium bromide and photographed with a UV transilluminator.

Sequence Analysis. Purified GC-clamped PCR products that showed abnormal migrating patterns on DGGE were directly sequenced as described (4, 21, 23) with non-GC-clamped primers as sequencing primers.

Analysis of Polymorphisms. Amplification of genomic DNA and restriction analysis of PCR products were performed in a reaction mixture volume of 20 μl as described (24). After the initial denaturation at 94°C for 6 min, 40 cycles of amplification consisting of a three-step-cycle (20 sec at 94°C, 30 sec at 55°C, and 30 sec at 72°C) for the *Bcl* I polymorphism (25) and a two-step-cycle (20 sec at 94°C and 45 sec at 65°C) for the *Xba* I polymorphism (26) were followed by a 10-min extension at 72°C.

RESULTS

Preliminary Experiments. To synthesize primers suitable for analysis of the entire coding region and splice junctions of *F8C* (22), we examined various primer choices using computer programs MELT 87 and SQHTX (20). A 40-nucleotide (nt) G+C-rich sequence (GC clamp) was attached to the 3′ or 5′ primer (13). A total of 45 sets of primers was chosen. Computer analysis predicted that mutations should be detected with these 45 sets of primers in 8.0 kb of *F8C*, including ≈6.9 kb (99%) of the coding region, 1.1 kb of (IVS), and 150 base pairs (bp) of 5′ and 3′ untranslated sequences. Sizes of amplified products ranged from 178 bp to 788 bp. We determined empirically the optimal gradient conditions and electrophoresis time for a given PCR product (15). To test the feasibility of the method, we first analyzed 26 known mutations in 11 exons of *F8C* previously detected by other methods (refs. 3 and 4; unpublished data). When heteroduplexes of normal and mutant PCR products were analyzed, all of these known mutations were distinguishable from normal (Fig. 1). Of the 26 mutations, 23 were 1-nt substitutions, 2 were deletion mutations of 2 and 3 nt, and 1 was a 1-nt insertion.

Screening for Mutations. The majority of mutations causing mild-to-moderate hemophilia A should reside within the coding region or splice consensus sequences of *F8C*. After analyzing all 45 regions, we identified disease-producing mutations in 25 of 29 patients studied (Table 1). In addition, 2 rare variants and 2 polymorphisms were also detected. All of these nucleotide alterations were detected by a typical "four-band-pattern" on denaturing gradient gels, which con-

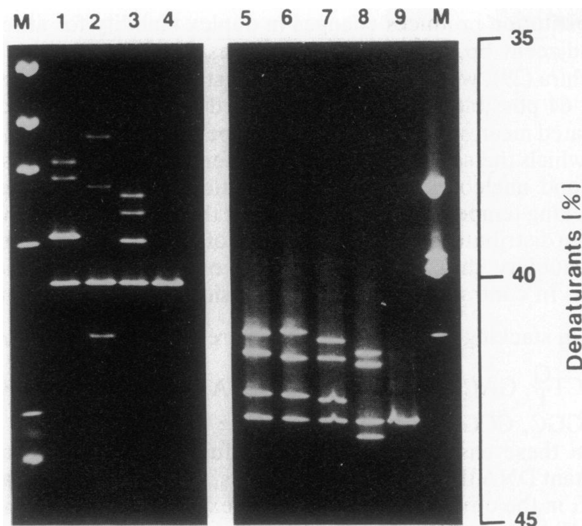

FIG. 1. Characterization of mutations in exons 22 and 23. PCR products containing known mutations or newly discovered mutations were combined with those from a normal male and subjected to DGGE. Lanes M contain ϕX174 DNA digested with *Hae* III. (*Left*) Analysis of exon 22. The 248-bp PCR product was electrophoresed in a 30–50% denaturing gradient gel for 18 hr at 40 V. Seen are PCR products containing a known mutation, C → T at codon 2135 (27) (lane 1); two mutations found in this study, T → G at codon 2120 (lane 2) and C → A at codon 2138 (lane 3); and the PCR product from a normal male (lane 4). The two upper bands are heteroduplexes and the two lower bands are homoduplexes of normal and mutant DNA. (*Right*) Analysis of exon 23. The 257-bp PCR product was electrophoresed in a 30–50% denaturing gradient gel for 20 hr at 60 V. Seen are PCR products containing a known mutation, C → T at codon 2166 (28) (lane 5); three mutations found in this study, G → A at codon 2169 (lane 6), C → T at codon 2178 (lane 7), and a 2-nt deletion at codons 2154 and 2155 (lane 8); and the PCR product from a normal male (lane 9). The following oligonucleotides were used as primers in this study: exon 1, 5′-(GC)$_x$TTTGCTTCTCCAGTTGAACAT-3′ and 5′-CGATCAGACCCTACAGGACA-3′; exon 2, 5′-(GC)$_x$T-TGAAGTGTCCACCAAAATGAACGACT-3′ and 5′-GATAC-CCAATTTCATAAATAGCATTCA-3′; exon 3, 5′-(GC)$_x$GTAC-TATCCCCAAGTAACCTTT-3′ and 5′-TATTCATAGAATGACAG-GACAATAGGA-3′; exon 4, 5′-(GC)$_x$TACAG-TGGATA-TAGAAAGGACAATTTT-3′ and 5′-TGCTTATTTCATCT-CAATCCTACGCTT-3′; exon 5, 5′-(GC)$_x$CCTCCTAGTGACAATT-TCCTACAATGA-3′ and 5′-AGCAGAGGATTTCTTTCAGGAATC-CAA-3′; exon 6, 5′-(GC)$_x$CATGAGACACCATGCTTAGCTGAC-TCTA-3′ and 5′-AACTCTGGTGCTGAATTTGGAAGACCCT-3′; exon 7, 5′-(GC)$_x$TCAGATTCTCTACTTCATAGCCATAGGT-3′ ("F8-69") and 5′-GAAACTGTCCAAGGTCCATCA-3′, also F8-69 and 5′-GTAGGACTGGATATTTATAATATTCATT-3′; exon 8, 5′-(GC)$_x$ATATAGCAAGACACTCTGACATTG-3′ and 5′-AGAGAG-TACCAATAGTCAAA-3′; exon 9, 5′-(GC)$_x$AGAGTTGGATT-TGAGCCTACCTAG-3′ and 5′-CAGACTTTTTCTTCTTACCTGA-CCTT-3′; exon 10, 5′-(GC)$_x$GGATTTGATCTTAGATCTCGCT-TAT-3′ and 5′-GAGCTATAAACGAGGGAATATTTAC-3′; exon 11, 5′-(GC)$_x$ATGGTTTTGCTTGTGGGTAG-3′ and 5′-GGATCCGA-CATACACTGAGAATGAA-3′; exon 12, 5′-(GC)$_x$CTAG-CTCCTACCTGACAACATCAGTAGC-3′ and 5′-CTCAAGCTG-CACTCCAGCTG-3′, also 5′-(GC)$_x$TATTTGATGAGAAC-CGAAGC-3′ and 5′-ATTCACCACCCACTGGACTTAAGTGCTG-3′; exon 13, 5′-(GC)$_x$GATGTGTCTAAATCTCTTTTTCC-CCATTG-3′ and 5′-ATATAATAACTAACCTGGGTTTTCCATC-3′; for exon 14 oligonucleotide primers, see below; exon 15, 5′-(GC)$_x$CACCTAGGAAAATGAGGATGTGAGGC-3′ and 5′-TTCT-TGTAATTCCACTGTCCTTAACT-3′; exon 16, 5′-(GC)$_x$GT-CGTTATTGTTCTACAGGTA-3′ and 5′-GTGGTCAAGCACAATA-GACACCTGC-3′; exon 17, 5′-TGATGAGAAATCCACTC-TGG-3′ and 5′-(GC)$_x$GTGCAATCTGCATTTCACAG-3′; exon 18, 5′-(GC)$_x$GTGGGAGTGGAATCCTCATAGATGTCAG-3′ ("F8-5") and 5′-ACTGATTGTGTTCCCAGTGC-3′; and F8-5 and 5′-AGTGCCATTTTATACTCCTC-3′; exon 19, 5′-(GC)$_x$-TTAGTGAAAAGAAATAATTTCTGTT-3′ and 5′-GTAGGCT-GAGTAGGTAGGGAACCTCTG-3′; exon 20, 5′-(GC)$_x$-ATTTTGTGCACTCTAGTTACTGTGT-3′ and 5′-TATAATCAGC-CCAGGTTCTTGGAGC-3′; exon 21, 5′-ATATAG-CAAGACACTCTGACATTG-3′ and 5′-(GC)$_x$GAGT-GAATGTGATACATTTCCCATCA-3′; exon 22, 5′-(GC)$_x$AAATAG-GTTAAAATAAAGTGTTAT-3′ and 5′-TGACTAATTACATAC-CATTAAGGTT-3′; exon 23, 5′-(GC)$_x$CTCTGTATTCACTTTC-CATG-3′ and 5′-AAGGATATGGGATGACTTGGCACT-3′; exon 24, 5′-(GC)$_x$GCTCAGTATAACTGAGGCTG-3′ and 5′-CTCTGAGT-CAGTTAAACAGT-3′; exon 25, 5′-(GC)$_x$TGACCAAG-AGGCTACTAGTCCA-3′ and 5′-CAGCTTACCTTTACTTTGC-CAT-3′; exon 26, 5′-TTAGCACAAAGGTAGAAGGC-3′ and 5′-AGCGTCTGTGCTTTGCAGTG-3′ ("F8-16"); and 5′-(GC)$_x$CCT-TCACACCTGTGGTGAAC-3′ and F8-16. (GC)$_x$ is 5′-CGCCCGC-CCCGCCCGCCGCCCGCCCCGCCCGCCGCCCGCCCCCGCCCGC-3′ and nucleotide numbers in exon 14 are from ref. 22. Exon 14 oligonucleotide primers were: the 3.2-kb fragment containing the entire exon 14, 5′-GAGAACCTCTAACAGAACGT-3′, and 5′-AGCAGAGCAAAG-GAATAACC-3′ ("F8-30"); 5′ portion of exon 14, 5′-(GC)$_x$-GAGAACCTCTAACAGAACGT-3′ ("F8-111"), and 5′-GTCA-CAACTAGAAACCTTCAG-3′; and F8-111 and 5′-ACTAGAG-GAGACATTTTGTAT-3′; nucleotides 2377–2694, 5′-(GC)$_x$GCACACAGAACACCTATGCCT-3′, and 5′-AGAAACTTTGAAATCAAGTTT-3′; nucleotides 2641–3146, 5′-(GC)$_x$AAACTGGGGACAACTGCAGCA-3′ ("F8-115"), and 5′-ATGATTTAGCCTCAAAGCTGT-3′; nucleotides 2641–2871, F8-115, and 5′-CAAGCTCAGAGGTCCACCAG-3′; nucleotides 3181–3391, 5′-(GC)$_x$GTGACACCTTTGATTCATGAC-3′, and 5′-CAGAGAGTTCTTTCCATGAG-3′; nucleotides 3352–3583, 5′-ATGCTATTCTTGCCAGAATCAGCA-3′, and 5′-(GC)$_x$GAAATAGGTTTCTGCTGCTTGG-3′; nucleotides 3523–3710, 5′-GAATTTACAAAGGACGTAGG-3′, and 5′-(GC)$_x$GT-CACTGTATGTATCTGAGG-3′; nucleotides 3601–3968, 5′-(GC)$_x$TTACATGAAAATAATACACACAATC-3′, and 5′-TGCTGGCTTGTATTAGGAGATATC-3′; nucleotides 3787–4381, 5′-(GC)$_x$TATGCTCCAGTACTTCAAGAT-3′ ("F8-119"), and 5′-TAGAATGGCTAAAGAAAGGTT-3′; nucleotides 3787–4142, F8-119, and 5′-TTCTCCTTCTCATTGTAGTC-3′; nucleotides 4093–4435, 5′-CATTTGACCCCGAGCACCCTC-3′, and 5′-(GC)$_x$TTGAT-CACCAGTCATCTCCAAGGT-3′; nucleotides 4339–4983, 5′-(GC)$_x$GTCCAAGAAAGCAGTCATTTC-3′ ("F8-121"), and 5′-GACTGGTGGGTTTTGAGAGCA-3′; nucleotides 4339–4778, F8-121, and 5′-TGAGTACCATAGTGGTTATCC-3′; 3′ portion of exon 14, 5′-(GC)$_x$CCTGGGCAAAGCAAGGTAGG-3′ ("F8-51"), and F8-30; and F8-51 and 5′-CAACTGATATGGTATCATCA-3′.

sists of two upper bands representing the heteroduplexes and two lower bands representing the homoduplexes of mutant and normal DNA (Fig. 1).

Although functional studies are needed to determine the consequences of the mutations listed in Table 1, we concluded that these mutations were likely to be disease-producing for two reasons. First, 24 of 25 patients had only one sequence change within the coding region of *F8C*, while one patient (JH-138) had two changes (see *Discussion*). Second, the frequency of neutral substitutions within *F8C* is extremely low (ref. 4 and the present study).

Mutations Identified. The 25 defined mutations occurred at 20 different sites. These included 1 putative splicing defect and 19 missense mutations of which 4 were found in more than 1 individual (Table 1). All missense mutations led to nonconservative amino acid substitutions.

Three of four recurrent missense mutations were CpG hotspot mutations at codons 546, 550, and 2016. A C → T transition at codon 546 (Arg-527 → Trp) was found in three German patients, JH-94, JH-95, and JH-96. (Note that the difference in codon number and amino acid residue number reflects the fact that codon 1 encodes the first methionine of the translation product and amino acid 1 is the N-terminal residue of the secreted protein.) The polymorphism haplotype in *F8C* of JH-96 differed from that of JH-94 and JH-95, indicating at least two independent origins of this mutation. A C → T transition at codon 550 (Arg-531 → Cys) was also detected in two Japanese patients, JH-97 and JH-98, but extended pedigree analysis revealed that they are distantly

related. The Arg-1997 → Trp mutation was found in JH-137 and JH-138. These substitutions probably occurred independently, since an additional nucleotide substitution in exon 25 was found in JH-138 (see *Discussion*) but not in JH-137, and the factor VIII haplotype of JH-138 differed from that of JH-137. The A → G mutation at codon 492 was found in two patients, JH-125 and JH-126, but they have identical factor VIII haplotypes, and they belong to the same ethnic group. Thus, it is likely that this mutation has a single origin.

The remaining 15 missense mutations were observed once and occurred randomly throughout the coding region of the gene with the exception of exon 14. Although exon 14 contains nearly 40% of the coding region, we identified only one putative disease-producing mutation (JH-132), a G → A change at codon 1057 (Glu-1038 → Lys).

DNA Polymorphisms Identified. The only common polymorphism found within the coding region was a C → G transversion in codon 1241 in exon 14 (Asp-1222 → Glu). This change has been reported as a difference between a cloned cDNA and genomic DNA (22). Among 80 X chromosomes analyzed, 58 (73%) had a C at codon 1241 and 22 (28%) had a G. This polymorphism is in complete linkage disequilibrium with the polymorphism at position −27 in IVS-7 (6, 7), indicating that its use would not add to the diagnostic information obtained by analysis of the *Bcl* I and *Xba* I polymorphisms.

Another polymorphism found was a silent A → C change in codon 1288. The rarer allele (C) was found in 5 of 80 patients, always on a chromosome carrying a G at codon 1241. Because of the low frequency (6.3%) of the rare allele, this polymorphism is not of diagnostic value.

In JH-144, we identified a G → A change at position −19 in IVS-2. This creates a potential weak splice acceptor site, gtaacctttggcAG ↓ A. The nucleotides underlined (5 of 15) differ from the consensus acceptor sequence, whereas at the normal acceptor site of IVS-2, 2 of 15 nucleotides differ from the consensus sequence. However, a C → G transversion at codon 550 (Arg-531 → Gly) was also found in this patient. This change is more likely to produce disease because another nucleotide substitution at this codon, a C → T transition (Arg-531 → Cys), was the only mutation found in two distantly related patients (JH-97 and JH-98). The nucleotide change in IVS-2 was not seen in 96 normal X chromosomes or in 153 additional hemophilia A patients (unpublished data), indicating that it is a rare normal variant.

The second rare variant was a G → A transition at codon 2242 (Val-2223 → Met) in patient JH-138 (see *Discussion*).

DISCUSSION

DGGE of amplified PCR products with a 40-nt GC-clamp is a powerful tool for mutation-screening of large genes (6, 7). In this study, to increase sensitivity and facilitate screening of a large number of patients, we modified our previous procedures (7). First, genomic PCR primers were chosen after computer analysis to assure that the regions of interest reside within the lowest melting domains (11, 20). Second, we determined the optimal conditions for each test fragment by varying denaturant concentrations and time of electrophoresis to maximize the difference in mobility between homoduplexes of normal and mutant DNA (15). However, certain nucleotide substitutions (see below) do not change the melting properties of the test fragment. Third, to overcome this problem we combined PCR products from two randomly selected patients to observe heteroduplexes. Mutations can be detected by heteroduplexes even when mutant DNA is not distinguishable from normal DNA after DGGE.

To understand the theoretical background of the DGGE technique, we examined changes in local duplex stability due to 1-nt substitutions. Duplex stability is attributable to electronic interactions of the stacked base pairs. Thus, a 1-nt substitution produces changes in duplex stability for at least 2 adjacent bp. Using stacking values from Gotoh and Tagashira (29), we calculated the mean stacking temperature of the 64 possible triplets and examined the difference in calculated mean stacking temperatures between any two triplets in which the second nucleotides differ. All 64 transitions at second nucleotides result in dramatic differences in mean stacking temperature, indicating that these differences, even when distributed in a melting domain of 400 bp, would be seen as mobility shifts in homoduplexes on denaturing gradient gels. In contrast, 22 of 128 transversions do not change the mean stacking temperature. These are: $AA^{G}_{T} \rightleftarrows AT^{G}_{T}$, $CA^{G}_{T} \rightleftarrows CT^{G}_{T}$, GAC ⇄ CTC, TAA ⇄ TTA, $AC^{A}_{T} \rightleftarrows AG^{A}_{T}$, GCC ⇄ GGC, CCG ⇄ CGG, and TCA ⇄ TGA.

In these cases, however, heteroduplexes of normal and mutant DNA that contain mismatches, C-C in one duplex and G-G in the other or A-A in one duplex and T-T in the other, are distinguishable from the homoduplexes. The previously described G → C mutation at codon 345 is a good example (4). This mutation, AGT → ACT, cannot be detected by our analysis of homoduplexes (7), but is identified by heteroduplex analysis (ref. 6 and the present study).

To corroborate further that few, if any, mutations escape detection by our DGGE method, we sequenced three regions, exons 6, 15, and 19, in 47 other patients with hemophilia A because (*i*) no abnormality was found by DGGE in the 47 patients analyzed, and (*ii*) each of these exons contains at least 1 CpG dinucleotide. Sequence analysis uncovered no mutations and confirmed the DGGE result.

After analyzing nearly the entire coding region and 41 of 50 splice sites, disease-producing mutations were identified in 25 (86%) of 29 patients selected for mild-to-moderate hemophilia A. In JH-138, we identified 2-nucleotide substitutions, a C → T transition at codon 2016 (Arg-1997 → Trp) and a G → A transition at codon 2242 (Val-2223 → Met). The latter substitution occurred at nucleotide 1 of exon 25, which is included in the consensus sequence of acceptor splice sites. At this position, G is present in 50% of a large number of acceptor sites analyzed, and A is found in 26% (30), suggesting that this substitution does not affect RNA splicing. It is possible that Val-2223 → Met is deleterious. However, it is more likely that hemophilia A in JH-138 is due to the Arg-1997 → Trp change, since this transition was also identified as the only mutation in an unrelated patient, JH-137, with another factor VIII haplotype.

Among 20 distinct disease-producing mutations characterized in this study, 5 are of particular interest. A G → T transition at codon 224 (Gly-205 → Trp) was identified in JH-143. This substitution also occurred at position −1 of the donor splice site of IVS-5. At this position, T is present in 8% of the donor splice sites analyzed, while G is present in 78% (30). A number of natural mutations have been observed at the −1 position of donor splice sites (31). The amino acid substitution, the splicing defect, or both could produce the hemophilia phenotype.

Two different substitutions were found in a CpG dinucleotide at codon 550 in the A_2 domain, C → T and C → G. The C → T transition occurred in two related patients JH-97 and JH-98, producing Arg-531 → Cys. Biochemical assays in JH-97 show that he has cross-reacting material (CRM)-positive hemophilia (factor VIII activity, 8.2%; factor VIII antigen, 36%). The C → G transversion was found in JH-144, resulting in Arg-531 → Gly. Thus, Arg-531 is of functional importance, because 2 different mutations alter this amino acid among 29 patients studied.

The Arg-527 → Trp change was identified in three patients (JH-94, JH-95, and JH-96) of whom at least two were unrelated. Patients JH-94 and JH-96 have CRM^{+} hemophilia

(factor VIII activities, 15–38%; factor VIII antigen levels, 75–100%). Factor VIII antigen was not measured in JH-95. It is interesting to note that two mutations producing CRM^+ hemophilia are located in close proximity in exon 11.

In contrast to activated factor V, activated factor VIII does not contain the A_2 domain (32); however, a number of factor VIII inhibitors bind to the 44-kDa subunit of factor VIII containing the A_2 domain (33, 34). We have found eight independent mutations in the A_2 domain (exons 9–13), suggesting that the 44-kDa subunit has functional importance.

The only putative disease-producing mutation found in the B domain results in a Gln-1038 → Lys change. Functional studies with recombinant factor VIII cDNA in heterologous cells have shown that deletion of all or a large portion of the B domain has no effect on procoagulant activity or binding to von Willebrand factor (35, 36). However, it is possible that the region around amino acid 1038 in the B domain is important for intracellular processing.

No nucleotide alteration was found in four patients. Factor VIII activity of these patients was 38%, 20%, 3.6%, and 3.5% of normal. The combination of factor VIII haplotypes, ethnic backgrounds, and differences in factor VIII activity suggests that each of these patients has a different mutation. Mutations in these patients may be located at a number of unanalyzed sites, including nine splice junctions, possible cryptic splice sites in introns, promoter sequences 5′ to the gene, and the RNA cleavage and polyadenylylation site (31, 37).

Studies of recombinant factor VIII protein (32, 38) and the discovery of natural mutations affecting both activation of factor VIII by thrombin (5, 39–42) and binding of factor VIII to von Willebrand factor (4) have provided significant information regarding processing and some functional elements of factor VIII. However, functional domains essential for factor VIII procoagulant activity such as binding to (*i*) factor X, (*ii*) activated factor IX, and (*iii*) metal ion remain to be established. A large collection of disease-producing mutations will help to delineate these domains. For example, mutations in the C_2 domain are candidates for a study of factor VIII binding to phospholipid (43, 44).

We have demonstrated that the combination of PCR and DGGE after computer analysis is a powerful and sensitive method for detection of nucleotide substitutions in coding regions of large genes. Improvements in the method, including automation, should allow its future use in analysis of hemophilia A carrier status in females at risk.

We are grateful to Dr. L. S. Lerman for providing the computer program and advice, Dr. R. M. Lawn for making available unpublished sequence information from introns, Dr. A. F. Scott for discussions, Drs. E. G. D. Tuddenham and R. Schwaab for factor VIII antigen assay, Dr. M. Fukimaki for patient information, L. Taylor for expert secretarial assistance, and J. Amberger for artwork. This study was supported by Grant HL38165 from the National Institutes of Health to H.H.K. and S.E.A.

1. Mann, K. G. (1984) *Prog. Hemostasis Thromb.* **7**, 1–23.
2. Sadler, J. E. & Davie, E. W. (1987) in *The Molecular Basis of Blood Disease*, eds. Stamatoyannopoulos, G., Nienhuis, A. J., Leder, P. & Majerus, P. W. (Saunders, Philadelphia), pp. 575–630.
3. Antonarakis, S. E. & Kazazian, H. H., Jr. (1988) *Trends Genet.* **4**, 233–237.
4. Higuchi, M., Wong, C., Kochhan, L., Olek, K., Aronis, S., Kasper, C. K., Kazazian, H. H., Jr., & Antonarakis, S. E. (1990) *Genomics* **6**, 65–71.
5. Gitschier, J., Kogan, S., Levinson, B. & Tuddenham, E. G. D. (1988) *Blood* **72**, 1022–1028.
6. Kogan, S. & Gitschier, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2092–2096.
7. Traystman, M. D., Higuchi, M., Kasper, C. K., Antonarakis, S. E. & Kazazian, H. H., Jr. (1990) *Genomics* **6**, 293–301.
8. Gianelli, F., Green, P. M., High, K. A., Lozier, J. N., Lillicrap, D. P., Ludwig, M., Olek, K., Reitsma, P. H., Goossens, M., Yoshioka, A., Sommer, S. & Brownlee, G. G. (1990) *Nucleic Acids Res.* **18**, 4053–4059.
9. Yoshitake, S., Schach, B. G., Foster, D. C., Davie, E. W. & Kurachi, K. (1985) *Biochemistry* **24**, 3736–3750.
10. Fisher, S. G. & Lerman, L. S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1579–1583.
11. Myers, R. M., Maniatis, T. & Lerman, L. S. (1987) *Methods Enzymol.* **155**, 501–527.
12. Myers, R. M., Fischer, S. G., Lerman, L. S. & Maniatis, T. (1985) *Nucleic Acids Res.* **13**, 3131–3145.
13. Sheffield, V. C., Cox, D. R., Lerman, L. S. & Myers, R. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 232–236.
14. Myers, R. M., Lumelsky, N., Lerman, L. S. & Maniatis, T. (1985) *Nature (London)* **313**, 495–498.
15. Myers, R. M., Sheffield, V. & Cox, D. R. (1988) in *Genomic Analysis: A Practical Approach*, ed. Davies, K. (IRL, Oxford), pp. 95–139.
16. Higuchi, M., Kochhan, L., Schwaab, R., Egli, H., Brackmann, H. H., Horst, J. & Olek, K. (1989) *Blood* **74**, 1045–1051.
17. Inaba, H., Fujimaki, M., Kazazian, H. H., Jr., & Antonarakis, S. E. (1989) *Hum. Genet.* **81**, 335–338.
18. Toole, J. J., Knopf, J. L., Wozney, J. M., Sultzman, L. A., Beucker, J. L., Pittman, D. D., Kaufman, R. J., Brown, E., Shoemaker, C., Orr, E. C., Amphlett, G. W., Foster, B., Coe, M. L., Knutson, G. J., Fass, D. N. & Hewick, R. M. (1984) *Nature (London)* **312**, 342–347.
19. Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2766–2770.
20. Lerman, L. S. & Silverstein, K. (1987) *Methods Enzymol.* **155**, 482–501.
21. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
22. Gitschier, J., Wood, W. I., Goralka, T. M., Wion, K. L., Chen, E. Y., Eaton, D. H., Vehar, G. A., Capon, D. J. & Lawn, R. M. (1984) *Nature (London)* **312**, 326–330.
23. Wong, C., Dowling, C. D., Saiki, R. K., Higuchi, R. G., Erlich, H. A. & Kazazian, H. H., Jr. (1987) *Nature (London)* **330**, 384–386.
24. Kogan, S. C., Doherty, M. & Gitschier, J. (1987) *N. Engl. J. Med.* **317**, 985–990.
25. Gitschier, J., Drayna, D., Tuddenham, E. G. D., White, R. L. & Lawn, R. M. (1985) *Nature (London)* **314**, 738–740.
26. Wion, K. L., Tuddenham, E. G. D. & Lawn, R. M. (1986) *Nucleic Acids Res.* **14**, 4535–4542.
27. Youssoufian, H., Kazazian, H. H., Jr., Phillips, D. G., Aronis, S., Tsiftis, G., Brown, V. A. & Antonarakis, S. E. (1986) *Nature (London)* **324**, 380–382.
28. Youssoufian, H., Antonarakis, S. E., Bell, W., Griffin, A. M. & Kazazian, H. H., Jr. (1988) *Am. J. Hum. Genet.* **42**, 718–725.
29. Gotoh, O. & Tagashira, Y. (1981) *Biopolymers* **20**, 1033–1042.
30. Shapiro, M. B. & Senapathy, P. (1987) *Nucleic Acids Res.* **15**, 7155–7174.
31. Kazazian, H. H., Jr. (1990) *Semin. Hematol.* **27**, 209–229.
32. Pittman, D. D. & Kaufman, R. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2429–2433.
33. Fay, P. J. (1988) *Arch. Biochem. Biophys.* **262**, 525–533.
34. Fulcher, C. A., de Graaf Mahoney, S., Roberts, J. R., Kasper, C. K. & Zimmerman, T. S. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7728–7732.
35. Toole, J. J., Pittman, D. D., Orr, E. C., Murtha, P., Wasley, L. C. & Kaufman, R. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5939–5942.
36. Eaton, D. L., Wood, W. I., Eaton, D., Hass, P. E., Hollingshead, R., Wion, K., Mather, J., Lawn, R. M., Vehar, G. A. & Gorman, C. (1986) *Biochemistry* **25**, 8342–8347.
37. Hobbs, H. H., Russell, D. W., Brown, M. S. & Goldstein, J. L. (1990) *Annu. Rev. Genet.* **24**, 133–170.
38. Kaufman, R. J., Pittman, D. D., Wasley, L. C., Foster, B. W., Amphlett, G. W. & Giles, A. R. (1987) *Thromb. Haemostasis* **58**, 537 (abstr.).
39. Arai, M., Inaba, H., Higuchi, M., Antonarakis, S. E., Kazazian, H. H., Jr., Fujimaki, M. & Hoyer, L. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4277–4281.
40. Shima, M., Ware, J., Yoshioka, A., Fukui, H. & Fulcher, C. A. (1989) *Blood* **74**, 1612–1617.
41. Arai, M., Higuchi, M., Antonarakis, S. E., Kazazian, H. H., Jr., Phillips, J. A., III, Janco, R. L. & Hoyer, L. W. (1990) *Blood* **75**, 384–389.
42. Pattinson, J. K., McVey, J. H., Boon, M., Ajani, A. & Tuddenham, E. G. D. (1990) *Br. J. Haematol.* **75**, 73–77.
43. Arai, M., Scandella, D. & Hoyer, L. W. (1989) *J. Clin. Invest.* **83**, 1978–1983.
44. Foster, P. A., Fulcher, C. A., Houghton, R. A. & Zimmerman, T. S. (1990) *Blood* **75**, 1999–2004.