Am. J. Hum. Genet. 53:1186–1197, 1993

Nature and Frequency of Mutations in the α-Galactosidase A Gene That Cause Fabry Disease

Christine M. Eng, Lois A. Resnick-Silverman, Dana J. Niehaus, Kenneth H. Astrin, and Robert J. Desnick

Department of Human Genetics, Mount Sinai School of Medicine, New York

Summary

Fabry disease, an X-linked inborn error of glycosphingolipid catabolism, results from mutations in the α-galactosidase A (α-Gal A) gene at Xq22.1. To determine the nature and frequency of the molecular lesions causing the classical and milder-variant Fabry phenotypes, and for precise carrier detection in Fabry families, the α-Gal A transcripts or genomic sequences from unrelated Fabry hemizygotes were analyzed. In patients with the classical phenotype, 18 new mutations were identified: N34S, C56G, W162R, R227Q, R227X, D264V, D266V, S297F, D313Y, G328A, W340X, E398X, IVS2+2, IVS5Δ-2,3, 773Δ2, 954Δ5, 1016Δ11, and 1123Δ53. Unrelated asymptomatic or mildly affected patients with symptoms confined to the heart had a missense mutation, N215S, that expressed residual enzymatic activity. Related, moderately affected patients with late-onset cardiac and pulmonary manifestations had a small deletion, 1208Δ3, that predicted the in-frame deletion of arginine 404 near the terminus of the 429 residue enzyme polypeptide. In addition, five small gene rearrangements involving exonic sequences were identified in unrelated classically affected patients. Two small deletions and one small duplication had short direct repeats at their respective breakpoint junctions and presumably resulted from slipped mispairing. A deletion occurred at a potential polymerase α arrest site, while the breakpoints of another deletion occurred at an inverted tetranucleotide repeat. Screening of unrelated Fabry patients with allele-specific oligonucleotides for seven mutations revealed that these were private, with the notable exception of N215S, R227Q, and R227X, which were each found in several unrelated families from different ethnic backgrounds. The CpG dinucleotide at codon 227 was the most common site of mutation, having been altered in 5% of the 148 unrelated Fabry alleles. These studies revealed that most α-Gal A lesions were private, that codon 227 was a mutational hot spot, and that certain mutations predicted a milder disease phenotype.

Introduction

Fabry disease is an X-linked inborn error of glycosphingolipid catabolism resulting from the deficient activity of the lysosomal exogalactohydrolase, α-galactosidase A (α-Gal A; E.C.3.2.1.22; Desnick and Bishop 1989). Affected hemizygous males accumulate neutral glycosphingolipids with terminal α-linked galactosyl moieties primarily in the plasma and in vascular endothelial lysosomes. In classically affected hemizygotes, the major disease manifestations include angiokeratoma, acroparesthesias, and vascular disease of the heart, kidneys, and brain, leading to early demise in adulthood. Milder or asymptomatic variants with residual α-Gal A activity have been described (for review, see Desnick and Bishop 1989). Recently, the full-length cDNA and entire genomic sequence encoding human α-Gal A were isolated and characterized (Bishop et al. 1986, 1988*a*, 1988*b*; Kornreich et al. 1989*a*, 1989*b*). The 14-kb genomic sequence contains seven coding exons and 12 intronic *Alu* repetitive elements. The processed α-Gal A transcript of ~1.45 kb encodes a 31-residue leader sequence and a 398-amino-acid mature lysosomal subunit that is glycosylated and dimerized to form the ac-

Received July 23, 1993; revision received September 1, 1993.

Address for correspondence and reprints: R. J. Desnick, Ph.D., M.D., Mount Sinai School of Medicine, Fifth Avenue at 100th Street, New York, NY 10029-6574.

0002-9297/93/5306-0005$02.00

tive enzyme glycoprotein. Stable overexpression of the α-Gal A cDNA in Chinese hamster ovary cells resulted in the high-level, selective secretion of active glycosylated and phosphorylated enzyme (Ioannou et al. 1992), permitting future trials of enzyme replacement in this disease.

The availability of the α-Gal A cDNA and genomic sequences has facilitated the identification of the first molecular lesions causing the classical and milder variant disease phenotypes. Southern analysis of 165 unrelated Fabry hemizygotes detected 6 classically affected hemizygotes with different gene rearrangements, including five partial gene deletions and a partial duplication (Bernstein et al. 1989). Subsequent determination of the rearrangement breakpoints demonstrated that one involved an *Alu-Alu* recombination, while five had short direct repeats at their respective breakpoints (Kornreich et al. 1990). Sequence analysis of the α-Gal A gene from other classically affected hemizygotes identified a 13-bp deletion in exon 1 (Ishii et al. 1991), a 5′ donor splice-site mutation in intron 6 (Sakuraba et al. 1992), a nonsense mutation in exon 1 (Sakuraba et al. 1990), and missense mutations in exons 1, 2, and 7 (Koide et al. 1990; Ishii et al. 1992). In contrast, characterization of the gene from mildly affected or asymptomatic variants with residual enzymatic activity revealed several missense mutations in exons 6 and 7 (Q279E, M296V, R301Q, and R356W [Bernstein et al. 1989; Sakuraba et al. 1990; von Scheidt et al. 1991; Ishii et al. 1992]). Thus, 17 specific lesions causing Fabry disease have been reported in the unrelated families studied to date, with one family having an unusual allele with two missense lesions (Ishii et al. 1992).

To further evaluate possible genotype/phenotype correlations, to permit precise heterozygote detection, and to facilitate selection of potential candidates for recombinant enzyme–replacement trials, an investigation was undertaken to determine the nature and frequency of the α-Gal A lesions causing Fabry disease. In this communication, we report the identification of 20 new α-Gal A mutations that cause Fabry disease. These mutations were detected either by reverse transcription of total cellular RNA, followed by amplification and sequencing of the α-Gal A cDNA, or by amplification and sequencing of α-Gal A regions from genomic DNA. These lesions include 10 missense, 3 nonsense, and 2 splice-junction mutations, as well as four small deletions, and one 5-bp insertion. Of these, N215S was identified in unrelated asymptomatic or mildly affected variants with symptoms primarily confined to the heart, and 1208Δ3 was identified in related, moderately affected hemizygotes who had late-onset cardiac and pulmonary manifestations. Although most mutations were private, N215S, R227Q, and R227X each occurred in several unrelated Fabry families from different ethnic backgrounds.

Material and Methods

Establishment of Lymphoblasts and Isolation of Genomic DNA and Total RNA

Whole-blood samples were obtained from affected hemizygotes and heterozygous females, with informed consent. The enzymatic diagnosis of affected hemizygotes and heterozygotes was established by determination of the 4-methylumbelliferyl-α-Gal A activity, as described elsewhere (Desnick et al. 1973). The phenotypic manifestations of each patient were evaluated by two of us (R.J.D. and/or C.M.E.). Lymphoblast cell lines were established and maintained as described elsewhere (Bernstein et al. 1989). Genomic DNA was isolated from whole blood or lymphoblasts by standard techniques (Sambrook et al. 1989). Total RNA was prepared from lymphoblasts by the acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi 1987).

Identification of Gene Rearrangements

Initial screening for major gene rearrangements (≥50 bp) was accomplished by multiplex amplification of α-Gal A sequences by using 500 μg of genomic DNA, according to the method of Kornreich and Desnick (1993).

Reverse Transcription and Amplification (RT-PCR) of the α-Gal A Transcript and Sequencing

For RT-PCR, first-strand complementary cDNA was reverse-transcribed from ~10 μg of total RNA isolated from lymphoblasts of unrelated Fabry hemizygotes. The entire α-Gal A cDNA was amplified in two overlapping fragments by using two sets of primers, and then each amplification product was subcloned into pGem 4Z, and double-stranded dideoxy sequencing was performed using α-Gal A–specific primers, as described elsewhere (von Schiedt et al. 1991).

Genomic Amplification of α-Gal A Exons, Subcloning, and Double-Stranded Sequencing

For genomic amplification, the complete coding region including intron/exon boundaries was amplified from genomic DNA (500 ng) in four fragments by PCR (Saiki et al. 1985). The respective sense (P1–P4) and

Table 1

Oligonucleotide Primers for Amplification of the α-Gal A Gene

Exon Amplified	Primer No.	Primer Sequence	Orientation	α-Gal A Gene Coordinates[a]	Fragment Size (bp)
Amplification and subcloning:					
1	P1	5′ ACTACTAAGCTTGGATCACTAAGGTGCCGC 3′	+	926–9446	558
	P5	5′ GCTGCTGAATTCAACTGTTCCCGTTGAGACTC 3′	−	1441–1460	
2	P2	5′ GCTGCTGAATTCTTGTGATTACTACCACACT 3′	+	4968–4987	390
	P6	5′ GCTAACAAGCTTCTGTACAGAAGTGC 3′	−	5312–5334	
3 and 4	P3	5′ TCTTGCTAAAGCTTCATGTACTTAA 3′	+	6750–6774	1,784
	P7	5′ GCTGCTGAATTCGTAAGTAACGTTGGACTTTG 3′	−	8491–8510	
5–7	P4	5′ ATCATCGAGCTCACAAGGATGTTAGT 3′	+	10078–10094	1,247
	P8	5′ ACTACTGAATTCAGGAAGTAGTAGTTGG 3′	−	11285–11301	
Amplification for direct solid-phase sequencing:					
1	B1	GGATCACTAAGGTGCCGC[b]	+	926–943	534
	P9	AACTGTTCCCGTTGAG	−	1441–1460	
2	B2	CTTGTGATTACTACCACACT[b]	+	4968–4987	366
	P10	AACAAGCTTCTGTACAGAAGTGC	−	5312–5334	
3 and 4	B3	TCTTGCTAAAGCTTCATGTACTTAA	+	6750–6774	1,760
	P11	b-GTAAGTAACGTTGGACTTTG	−	8491–8510	
5–7	B4	CTCACAAGGATGTTAGT[b]	+	10078–10094	1,223
	P12	CAGGAAGTAGTAGTTGG	−	11285–11301	

[a] Kornreich et al. (1989*b*).
[b] Biotinylated primer.

antisense (P5–P8) primers, which included *Eco*RI and *Hin*dIII (exon 1, exon 2, and exons 3 plus 4) or *Eco*RI and *Sac*I (exons 5–7) restriction sites for directional subcloning, were synthesized on an Applied Biosystems 380B DNA synthesizer (Applied Biosystems, Foster City, CA) (table 1). The amplification reaction (100 μl) contained a final concentration of ~1 μM of each primer, 2 mM of each dNTP, 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% (w/v) gelatin, and 2 units of *Taq* DNA polymerase (Perkin Elmer Cetus, Norwalk, CT). The reaction mixture was incubated at 94°C for 7 min to denature the DNA and inactivate proteases. Amplifications were performed for 30 cycles, each cycle consisting of denaturation at 92°C for 1 min, annealing at 61°C for 2 min, and extension at 72°C for 2 min. The purified fragments were double-digested with the appropriate restriction enzymes and were subcloned into pGEM-4Z (Promega, Madison, WI). Minipreps (Sambrook et al. 1989) of multiple ampicillin-resistant colonies were sequenced by the dideoxy chain termination method (Sanger et al. 1977) using SP6 or T7 sequencing primers. In order to rule out mutations resulting from amplification errors, several independent products were sequenced to confirm the putative mutation.

Genomic Amplification and Single-Stranded Solid-Phase Direct Sequencing

Four sense biotinylated oligonucleotide primers (designated "B1"–"B4") and four antisense oligonucleotide primers (denoted "P9"–"P12") (table 1) were synthesized to amplify exons 1, 2, 3 plus 4, and 5–7 (each product included ~50 bp of flanking intronic sequences). Genomic DNA (500 ng) from each of the probands was PCR amplified (Saiki et al. 1985) with the respective primer sets (final concentration of each primer was ~1 μM), as described above. Amplifications were performed for 30 cycles, with denaturation at 94°C for 30 s, annealing at 56°C for 45 s, and extension at 72°C for 60 s, allowing for a final extension step

of 7 min at 72°C. The α-Gal A amplification products from unrelated Fabry patients were compared with those of the normal gene by agarose gel electrophoresis. Double-stranded PCR products (40 μl) were denatured by standard techniques, and the biotinylated single strands were isolated by affinity capture using streptavidin-coated magnetic beads (Dynabeads M-280; DYNAL, New York) (Hultman et al. 1989). The biotinylated single-stranded PCR products were subjected to dideoxy chain termination sequencing using α-Gal A–specific sequencing primers at a final concentration of 0.5 pmol.

Allele-specific Oligonucleotide (ASO) Hybridization

ASO hybridization was performed using standard techniques (Wang et al. 1990) to screen unrelated Fabry patients for the presence of the common and other mutations, as well as to screen normal individuals to determine whether these lesions were common polymorphisms. Appropriate genomic regions were amplified using the respective oligonucleotide primers (table 1), as described above; the amplification products were denatured and applied to a nitrocellulose membrane (Zetabind; AMF-CUNO, Meriden, CT); and the normal or mutant sequence was detected by using end-labeled 17-bp oligonucleotide primers specific for the normal or mutant sequence. The ASOs used included N215S—5′-AAAGCCCAATTATACAG-3′ (normal) and 5′-AAAGCCCAGTTATACAG-3′ (mutant); R227Q—5′AATCACTGGCGAAATTTTGCT-3′ (normal) and 5′AATCACTGGCAAAATTTTGCT-3′ (mutant); R227X—5′-ATCACTGGCGAAATTTT-3′ (normal) and 5′-ATCACTGGTGAAATTTT-3′ (mutant); W162R—5′-TTTGCTGACTGGGGAGT-3′ (normal) and 5′-TTTGCTGACCGGGGAGT-3′ (mutant); D264V—5′-TGGAATGACCCAGATAT-3′ (normal) and 5′-TGGAATGTCCCAGATAT-3′ (mutant); S297F—5′-TATTCATGTCTTAATGAC-3′ (normal) and 5′-TATTCATGTTTTAATGAC-3′ (mutant); and G328A—5′-AAGCAAGGGTACCAGC-3′ (normal) and 5′-AAGCAAGCGTACCAGC-3′ (mutant).

Results

Three methods for α-Gal A mutation detection were compared: (1) total lymphoblast RNA was isolated and reverse-transcribed to cDNA, and the α-Gal A transcript was amplified in two fragments (designated "RT-PCR"), and then each amplification product was subcloned for double-stranded sequencing; (2) individual α-Gal A exons and flanking intronic regions were amplified from genomic DNA by using oligonucleotide primers with different restriction sites at their ends, and then the amplified products were subcloned for single-stranded or double-stranded sequencing (designated "genomic amplification and sequencing"); and (3) exons including flanking intronic regions were amplified from genomic DNA by using primer sets containing one biotinylated primer, which permitted Streptavidin-mediated capture of single-stranded PCR products for direct solid-phase sequencing (designated "genomic amplification and direct sequencing"). Each identified mutation was the only alteration in the α-Gal A coding and/or intron/exon sequence found in the respective proband. The mutations were confirmed by sequence analysis of independent PCR products and, when available, by pedigree analysis of affected individuals and obligate heterozygotes in each family. These analyses identified 20 new mutations (fig. 1 and table 2), consisting of 4 small (2–53-bp) exonic deletions, 1 exonic 5-bp insertion, and 15 single-base substitutions, including 10 missense, 3 nonsense, and 2 splice-junction mutations. In the three methods used, 3 lesions were detected by RT-PCR, 7 by genomic amplification, subcloning, and sequencing, and 10 by genomic amplification and direct sequencing.

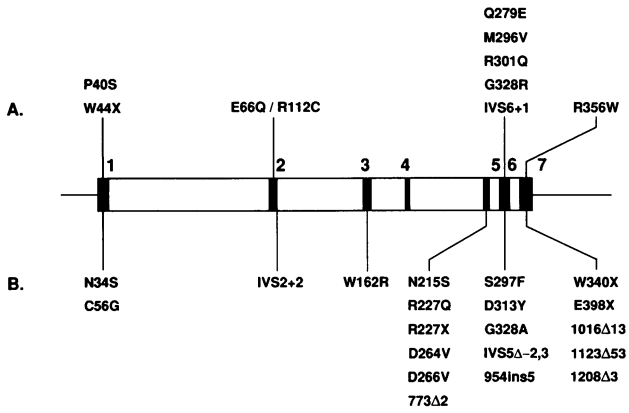

Figure 1 α-Gal A gene mutations causing Fabry disease. Numbered exons are indicated by solid rectangles. *A*, Previously published point mutations. Partial deletions and a duplication (Kornreich et al. 1990; Ishii et al. 1991) are not shown. *B*, 20 Mutations reported in the present communication.

α-Gal A Missense and Nonsense Mutations

Of the 13 coding region mutations found in the α-Gal A gene, only 3 (~25%) were identified in exons 1–4 (encoding residues 1–213), while 10 were found in exons 5–7 (encoding residues 214–429). Of these sin-

Table 2

Mutations in the Human α-Gal A Gene That Cause Fabry Disease

Location	Allele Designation	Nucleotide Change[a]	Alteration	Restriction Site[b]	Ethnic Origin	Phenotype
Missense/nonsense mutations:						
Exon 1	N34S	AAT→AGT	Asn→Ser		Polish	Classic
	C56G	TGC→GGC	Cys→Gly	+ (*Hae*III)	English	Classic
Exon 3	W162R	TGG→CGG	Trp→Arg	+ (*Msp*I and *Scr*FI)	Italian	Classic
Exon 5	N215S[c]	AAT→AGT	Asn→Ser	+ (*Bsr*I)	Italian	Variant
	R227Q	CGA→CAA	Arg→Gln		Common; see text	Classic
	R227X	CGA→TGA	Arg→Stop	+ (*Hph*I)	Common; see text	Classic
	D264V	GAC→GTC	Asp→Val		Scottish/English	Classic
	D266V	GAT→GTT	Asp→Val	+ (*Bsr*I)	African-American	Classic
Exon 6	S297F	TCT→TTT	Ser→Phe	+ (*Mse*I)	Italian	Classic
	D313Y	GAT→TAT	Asp→Tyr		German	Classic
	G328A	GGG→GCG	Gly→Ala	− (*Kpn*I)	Scottish/Irish	Classic
Exon 7	W340X	TGG→TGA	Trp→Stop		African-American	Classic
	E398X	GAA→TAA	Glu→Stop	+ (*Ase*I and *Mse*I)	Hispanic	Classic
Splicing mutations:						
Intron 2	IVS2^{+2}	gt→gg	5′ Donor site		Sephardic Jewish	Classic
Intron 5	IVS5-Δ2,3	tcag→t--g	3′ Acceptor site		Irish	Classic
Coding-region insertion/ deletions:						
Exon 5	773Δ2	2 bp			Portuguese	Severe
Exon 6	954ins5	dup 5 bp			German	Severe
Exon 7	1016Δ11	11 bp			German	Severe
	1123Δ53	53 bp			Dutch	Severe
	1208Δ3[c]	3 bp			English	Variant[c]

[a] An underline denotes CpG.

[b] + = Created by mutation; and − = obliterated by mutation.

[c] N215S results in an extremely mild phenotype confined to the heart. Asymptomatic at age 43 years when diagnosed in the course of renal evaluation for mild proteinuria, the phenotype of 1208Δ3 was moderate, with cardiac involvement in two cousins.

gle-base substitutions, 10 were missense mutations, and 3 were nonsense mutations (table 2). The missense mutations identified in classically affected hemizygotes included (1) an A-to-G transition in codon 34 of exon 1, predicting an asparagine-to-serine substitution (N34S) at the third residue of the mature enzyme subunit; (2) a T-to-G transversion of codon 56 in exon 1, resulting in a cystine-to-glycine substitution (C56G); (3) a T-to-C transition in codon 162 of exon 3, predicting a tryptophan-to-arginine replacement (W162R); (4) a G-to-A transition of codon 227 (fig. 2*A*), resulting in an arginine-to-glutamine substitution (R227Q); (5) an A-to-T transversion in codon 264 of exon 5, predicting an aspartic acid-to-valine substitution (D264V); (6) an A-to-T transversion in codon 266, also predicting an aspartic acid-to-valine substitution (D266V); (7) a C-to-T transition in codon 297, resulting in a serine-to-phenylalanine substitution (S297F); (8) a G-to-T transversion in codon 313, predicting an aspartic acid-to-tyrosine substitution (D313Y); and (9) a G-to-C transversion in codon 328, resulting in a glycine-to-alanine substitution (G328A). In addition, an A-to-G transition in codon 215 of exon 5 (fig. 2*C*), resulting in the substitution of an asparagine by a serine (N215S), was identified in an atypical variant with manifestations confined to the heart (Bishop et al. 1982). The N215S mutation occurred at a highly conserved N-glycosylation consensus site (A. M. Wang, personal communication) and expressed ~5% and 25% of normal α-Gal A enzymatic activity in peripheral leukocytes and lymphoblasts, respectively (C. M. Eng, unpublished results). The three nonsense mutations, all identified in classically affected hemizygotes, included (1) a C-to-T transition in codon 227 of exon 5 (R227X) (fig. 2*B*), predicting the deletion of the terminal 202 residues of the enzyme polypeptide; (2) a G-to-A transition in codon 340 of exon 7

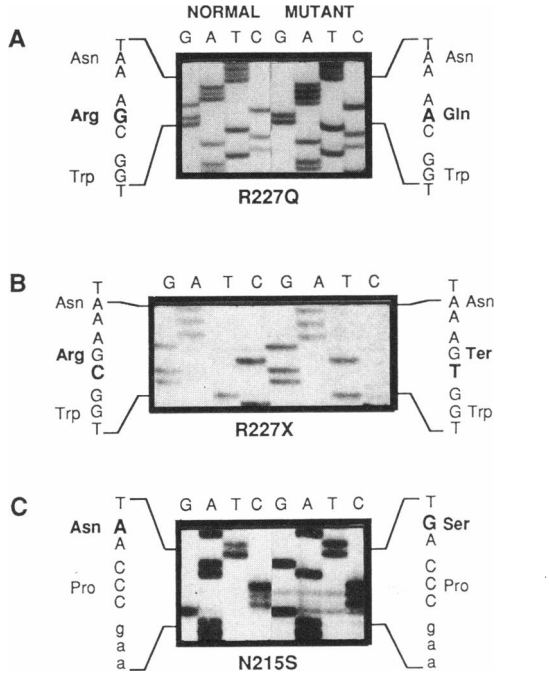

Figure 2 Nucleotide sequence showing three common Fabry mutations in the α-Gal A gene. Nucleotide substitution is in boldface. The normal sequence appears on the left, and the mutant sequence is on the right. *A*, G-to-A transition in codon 227 (R227Q), detected by genomic amplification and sequencing. *B*, C-to-T transition in codon 227 (R227X), detected by genomic amplification and sequencing. *C*, A-to-G transition in codon 215 (N215S), detected by RT-PCR.

(W340X), predicting the deletion of the terminal 89 amino acids; and (3) a G-to-T transversion in codon 398 of exon 7 (E398X), predicting an enzyme polypeptide missing the 32 carboxy-terminal residues.

α-Gal A mRNA Processing Defects

By genomic amplification and sequencing and by solid-phase direct sequencing of amplified α-Gal A genomic products, two mutations that altered processing of the α-Gal A transcript were identified. A gt-to-gg transversion at the 5′ donor splice site of intron 2 (designated "IVS2+2") was found in a classically affected hemizygote (fig. 3*A*). A 2-bp deletion, tcag-to-t--g (fig. 3*B*), that disrupted the 3′ acceptor splice site of intron 5 also was identified in a classically affected hemizygote.

α-Gal A Coding Sequence Insertion and Deletions

A five-base duplication and four small deletions that were not detected by the α-Gal A multiplex PCR technique (Kornreich and Desnick 1993) were identified either by RT-PCR and sequencing or by genomic amplification and solid-phase direct sequencing (table 2). All of these gene rearrangements resulted in frameshift mutations that led to premature chain termination, and these gene rearrangements were detected in classically affected hemizygotes, with the exception of a three-base deletion (1208Δ3) in exon 7 in a patient with a later-onset moderate disease phenotype (see below). The five-base insertion, 954ins5, in exon 6 occurred at codon 318 and involved the duplication of five bases (TTGCC) that were flanked by the short direct repeat AT (i.e., AT TGCC TTGCC AT) (figs. 3*C* and 4*A*). This insertion caused a frameshift from codon 319 to a premature termination codon at position 349, with amino acid residues 319–348 all different from the corresponding normal sequence. The 2-bp deletion, 773Δ2,

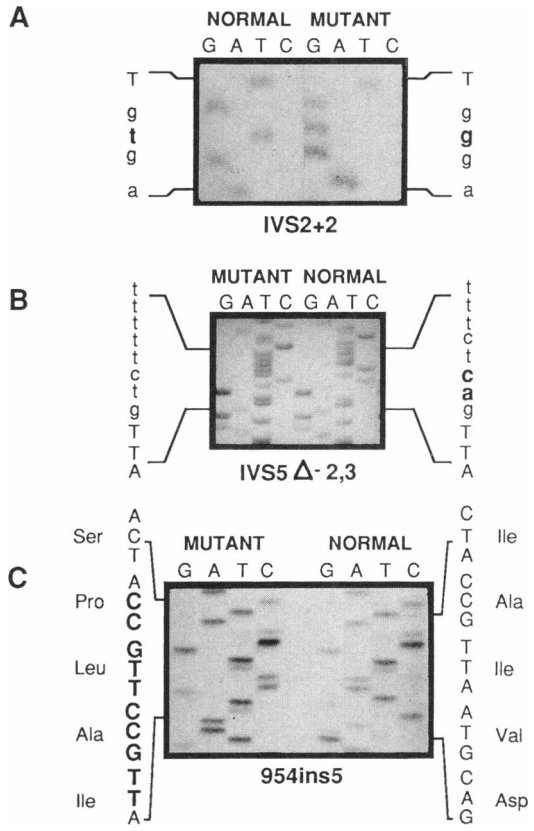

Figure 3 Nucleotide sequence showing three mutations in the α-Gal A gene. *A*, gt-to-gg transversion at the 5′ splice donor site of intron 2 (IVS2+2), detected by genomic amplification and direct sequencing. *B*, tcag-to-t--g deletion in the 3′ acceptor site of intron 5 (IVS5-Δ2,3), detected by genomic amplification and sequencing. *C*, 5-bp duplication at codon 318 (954ins5), detected by RT-PCR.

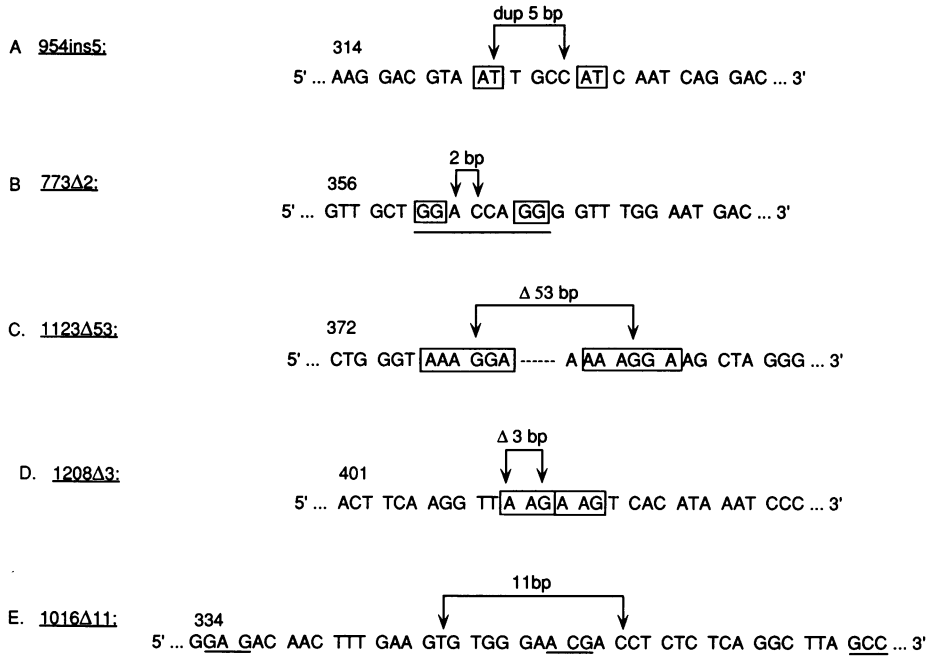

Figure 4 Nucleotide sequence flanking four partial gene deletions and one 5-bp insertion. Arrows indicate the nucleotides that were duplicated (*A*) or deleted (*B–E*). Short direct repeat sequences are boxed. *A*, Sequence flanking the 5-bp insertion (954ins5). The nucleotides TTGCC (indicated by arrows) were duplicated and inserted in tandem between short direct repeats of AT. *B*, Sequence flanking the 2-bp deletion in exon 5 (773Δ2). The inverted tetranucleotide GGAC CAGG is underlined. The 2-bp deletion (AC) is indicated by the arrows. *C*, Sequence flanking the 1123Δ53 deletion. The breakpoints of the deletion are indicated by the arrows. The hexanucleotide repeat sequence that flanked the deletion breakpoints (AAAGGA) is boxed. *D*, Sequence flanking the 1208Δ3 deletion. The tandemly repeated AAG AAG is indicated, and the deletion of one repeat is shown by the arrows. As a result of the deletion, the TTA at codon 403 was preserved, and the AGA at codon 404 was deleted and replaced by the next in-frame AGT. *E*, Sequence flanking the 11-bp deletion in exon 7 (1016Δ11). The 11 bp that were eliminated are between the arrows. The three consensus motifs for polymerase α arrest that occur in proximity to the deletion breakpoints are underlined.

in exon 5 occurred in a sequence containing the inverted tetranucleotides, GGAC CAGG, each of which contained the short direct repeat, GG (fig. 4*B*). This deletion occurred in codons 258 and 259 and predicted a truncated polypeptide of 262 residues. The 53-base deletion, 1123Δ53, in exon 7 had breakpoints that occurred at the hexanucleotide direct repeat AAAGGA (fig 4*C*). This deletion eliminated codons 375–392 and predicted a truncated polypeptide of 379 residues, of which the last five amino acids differed from the normal sequence. A three-base deletion, 1208Δ3, was detected in the terminal portion of exon 7 from a moderately affected patient and resulted from a deletion of one of two contiguous AAG triplet repeats (fig. 4*D*). This in-frame deletion predicted the elimination of arginine 404, while the remainder of the 428 enzyme polypeptide sequence was unchanged. An 11-bp deletion, 1016Δ11, in exon 7 had breakpoints that were near several consensus triplets (e.g., GAG, GCC, GCG, and ACG) for polymerase α arrest (Weaver and DePamphilis 1982). This deletion involved codons 339–343 and predicted a truncated polypeptide of 369 residues, with residues 339–369 different from those in the normal sequence.

Common α-Gal A Mutations

In the course of determining the mutations in 22 unrelated Fabry families, the R227X allele was detected in 2 families, and the R227Q mutation was identified in 2 others, suggesting that the CpG dinucleotide in this codon was a mutational hot spot. To determine whether the R227Q or R227X allele occurred in other Fabry families or whether either base substitution was a polymorphism, ASO hybridizations were performed with amplified genomic DNA from unrelated Fabry patients and normal individuals. Neither mutation was detected in the 100 normal alleles studied. Of the 126 unrelated Fabry patients analyzed, 3 additional hemizy-

Table 3

Frequency of Mutations in Unrelated Fabry Families

Mutation	Unrelated Fabry Alleles Sequenced	ASO Analyses (No. of Alleles)	Total
R227Q	2/22	1/126	3/148
R227X	2/22	3/126	5/148
W162R	1/22	0/37	1/59
D264V	1/22	0/112	1/134
S297F	1/22	0/109	1/131
G328A	1/22	0/100	1/122
N215S	1[a]	2[b]	3

[a] Atypical patient; N215S was detected by sequencing.
[b] Atypical patients; N215S detected by ASO analysis.

gotes had the R227X allele, and 1 had the R227Q allele. When these are added to the two R227X families and the two R227Q families identified independently by sequencing, the R227X and R227Q alleles occurred in 3.4% and 2% of the 148 unrelated Fabry families studied, respectively (table 3).

Because the N215S mutation was identified in an atypical variant with a mild disease phenotype (Bishop et al. 1981), DNAs from other atypical variants were screened by ASO hybridization analysis. Two unrelated atypical variants screened had the N215S allele. Subsequent sequence analyses did not reveal any other alteration of the α-Gal A coding sequences in these variants, nor was the N215S substitution identified in 100 normal alleles examined by ASO analyses.

Discussion

The recent isolation and characterization of the cDNA and genomic sequences encoding human α-Gal A permitted the identification of a variety of molecular lesions causing Fabry disease. To date, a total of 17 lesions have been identified, including 7 partial gene rearrangements, 1 RNA processing defect, 1 nonsense mutation, and 8 missense mutations (fig. 1*A*) (Bernstein et al. 1989; Kornreich et al. 1990; Sakuraba et al. 1990, 1992; Ishii et al. 1991, 1992; von Scheidt et al. 1991). No common lesions have been found, and each of the identified mutations occurred in unrelated patients, indicating the genetic heterogeneity of the molecular lesions causing this disease. It is notable that all of these mutations were detected in patients with the classical disease phenotype, except for three missense mutations (M296V, Q279E, and R301Q) that were found in atypical variants, whose manifestations were confined to the heart and who did not have the classical symptoms including angiokeratoma, acroparathesias, corneal opacities, hypohidrosis, or renal disease (Sakuraba et al. 1990; von Scheidt et al. 1991; Ishii et al. 1992). In the present communication, 20 new mutations have been described (fig. 1*B*), which further define the heterogeneity of the disease-causing mutations, identify a mutational hot spot in the α-Gal A gene, and delineate a common mutation that results in an atypical mild phenotype.

In the course of these studies, three different strategies were employed to sequence the α-Gal A coding region and flanking intronic sequences. All three techniques were successful in detecting mutations; of the 20 new mutations, 3 were identified by RT-PCR, 7 by genomic amplification and sequencing, and 10 by solid-phase direct sequencing. The latter technique was preferred because it both used genomic DNA instead of RNA and captured single-stranded PCR products for sequencing, thereby eliminating the need to subclone the amplification product. Moreover, RT-PCR may not detect some nonsense mutations, because of either the instability of the mutant transcripts or the alteration of splice-site selection (Dietz et al. 1993).

The majority of mutations detected in Fabry disease are point mutations. Previously, one splicing, one nonsense, and eight missense mutations had been reported (Koide et al. 1990; Sakuraba et al. 1990, 1992; Ishii et al. 1991, 1992; von Scheidt et al. 1991). When the 15 identified base substitutions described here are included, a total of 25 have been detected, including 3 splicing, 4 nonsense, and 18 missense mutations. These base substitutions occurred in six of the seven exons, with exons 5 and 6 having five and seven lesions, respectively, and with exon 4 having none. Several of these mutations were notable. R227Q and R227X, which occurred at a CpG nucleotide (Coulondre et al. 1978; Cooper and Youssoufian 1988), were the most common mutations causing the classical phenotype. Taken together, they were found in 8 (5%) of 148 of unrelated Fabry families studied, including families whose mutant alleles could be traced to Danish, English, German, Indian, Irish, Italian, and Polish ancestries. Two families with the R227Q mutation were of German descent, but a common ancestor or demographic region could not be identified. Further support for their nonrelatedness could be established by the analysis of polymorphic markers flanking the α-Gal A gene (Vetrie et al. 1993). That at least six unrelated families had either mutation

indicates that the CpG in codon 227 is an α-Gal A mutational hot spot causing Fabry disease. In contrast, of the 14 other CpG dinucleotides in the α-Gal A coding sequence, only 3, at codons 112, 301, and 356 (Bernstein et al. 1989; Sakuraba et al. 1990; Ishii et al. 1992), have been involved in a mutation reported to date.

Also notable was the finding that N215S was a common mutation among the rare atypical variants who either are asymptomatic or have mild disease manifestations (Eng et al. 1993; Desnick et al., in press). This lesion was first identified in a patient of Italian origin who at age 42 years had none of the classical Fabry manifestations (angiokeratoma, acroparathesias, corneal or lenticular opacities, hypohidrosis, or renal insufficiency). N215S obliterates a functional N-glycosylation consensus site, but the expressed enzyme retains ~5%–25% of normal activity in various tissues and cultured cells from affected patients (Bishop et al. 1981; L. A. Resnick-Silverman, unpublished results). It is interesting that the N215S mutation was identified by ASO hybridization in two other unrelated atypical patients with similar mild phenotypic manifestations confined to the heart. A second variant, of Czechoslovakian descent (Elleder et al. 1990), died at 63 years of age, of a pulmonary embolism and bilateral cardiac failure, but had no history of renal disease. Autopsy findings revealed marked cardiomegaly and concentric left-ventricular hypertrophy, but there was no histochemical and/or ultrastructural evidence of glycolipid deposition in the coronary or pulmonary arteries, aorta, brain, kidney, liver, or pancreas. The third N215S variant was a 63-year-old male of German descent who was asymptomatic except for hypertrophic cardiomyopathy.

Previously, point mutations had been described in three atypical variants with disease manifestations confined to the heart (M296V [von Scheidt et al. 1991], Q279E [Ishii et al. 1992], and R301Q [Sakuraba et al. 1990]). Most patients with the classical phenotype have no detectable α-Gal A activity and expire in their 30s or 40s, from complications of renal and vascular involvement, before myocardial manifestations become debilitating (Desnick and Bishop 1989). In contrast, patients with the atypical phenotype apparently have sufficient residual α-Gal A activity to protect their kidneys and vascular endothelium. However, the level of residual activity may be inadequate to prevent the progressive glycosphingolipid deposition in myocardial cells. In late adulthood, these atypical patients develop myocardial disease and present as the cardiac variants of Fabry disease. Thus, these missense mutations, which express residual activity, predict a mild phenotype and provide the basis for genotype/phenotype correlations in this disease (also see 1208Δ3 below).

In contrast to the R227Q, R227X, and N215S mutations, which occurred in several unrelated families, ASO analyses indicated that selected other mutations (i.e., W162R, D264V, S297F, and G328A) were private. In fact, the absence of a common mutation(s) suggested that the vast majority of mutations causing Fabry disease are private, being confined to a single pedigree. In the 22 unrelated Fabry families studied here, the two mRNA processing defects, IVS2+2 and IVS5Δ-2,3, were detected in classically affected hemizygotes. The five small exonic gene rearrangements were detected in classically affected patients, caused frameshift mutations, and predicted truncated enzyme polypeptides, with the notable exception of 1208Δ3, which deleted arginine 404 in several moderately affected, related patients. Patients with the 1208Δ3 allele had late (5th-decade) onset of chronic obstructive pulmonary disease, as well as mild cardiac involvement.

Previous studies of α-Gal A gene rearrangements revealed five partial gene deletions of ~0.4–4.6 kb and one large partial duplication of 8.1 kb, in unrelated classically affected patients (Bernstein et al. 1989). Since the α-Gal A gene contains 12 *Alu* repetitive elements, representing ~30% of the gene, or ~1 *Alu*/kb, it was suggested that the α-Gal A gene might be predisposed to gene rearrangements due to *Alu-Alu* recombination (Kornreich et al. 1989*b*). However, of the previous six partial gene rearrangements, only one was due to an *Alu-Alu* recombinational event, while the others presumably resulted from illegitimate recombination involving the short direct repeats at their breakpoint junctions (Kornreich et al. 1990). Of the five small exonic gene rearrangements described here, the 954ins5 duplication and three deletions—773Δ2, 1123Δ53 and 1208Δ3—had short direct repeats at their breakpoint junctions (figs. 4*B–D*), suggesting that they resulted from slipped mispairing. This mechanism of illegitimate recombination involves the misalignment during DNA replication of short direct repeats of 2–8 bp that flank the deletion breakpoints, which generate a single-stranded loop structure containing one of the repeats and the intervening sequence between the two repeats. The looped sequence is excised by DNA repair enzymes, and the strands are rejoined, resulting in a deletion (Efstratiadis et al. 1980; Kornreich et al. 1990). For example, slipped mispairing of the hexanucleotide direct repeat may be responsible for the generation of the 53-bp deletion in exon 7 (1123Δ53) (fig. 4*C*). Similarly,

the 1208Δ3 deletion may have resulted from misalignment of two tandemly repeated AAG triplets (fig. 4*D*). The 773Δ2 mutation may have resulted from slipped mispairing of the GG dinucleotide direct repeats that flank the deletion site (fig. 4*B*). In addition, the 8 bp in the region of the deletion form a symmetric element with an axis of internal symmetry (GGAC CAGG), which also may play a functional role in the generation of the 2-bp deletion of AC (Krawczak and Cooper 1991).

Slipped mispairing also may explain the generation of the five-base duplication in exon 5 (954ins5) (fig. 4*A*), which was flanked by the dinucleotide AT direct repeat. The generation of an insertion differs from that of a deletion, in that the mispairing occurs when the newly synthesized strand dissociates from the primer strand and base pairs illegitimately with the primer strand. In this case, the slippage resulted in the duplication of five bases, which were inserted prior to the second AT repeat. Several other examples of insertions due to short direct repeats have been reported (e.g., see Bottema et al. 1989; Howells et al. 1990). However, the 954ins5 mutation appears to be one of the most compatible duplications to be formed by the slipped-mispairing mechanism. It is interesting that the relatively infrequent generation of insertions, as opposed to deletions, at sites of direct repeats suggests that the proposed mechanism of insertional mutagenesis between direct repeats may be less favored in terms of the chemical energy required for their formation (Cooper and Krawczak 1991).

It is interesting that the 1016Δ11 deletion may be due to the presence of polymerase α arrest sites proximal to the deletion and at the 3′ breakpoint of the deletion. Polymerase α has been shown to be the only polymerase that catalyzes DNA synthesis at replication forks (DePamphilis and Wassarman 1980). Weaver and DePamphilis (1982) found that specific sequences GAG, GCC, GCG, and ACG (in decreasing frequency) were found at sites of polymerase α arrest. The presence of these sequences may slow the progress of the polymerase, increasing the likelihood that, in replication, an error may occur that could lead to a deletion or duplication event. It is interesting that the most commonly involved motif, GAG, was found 14 bases upstream of the 5′ deletion breakpoint—and that the less frequent motifs, ACG and GCG, were found 2 bases upstream (within the deletion) and 14 bases downstream of the 3′ deletion breakpoint, respectively (fig. 4*E*). The occurrence of these arrest sites, in conjunction with a potential stem-loop formation occurring 3′ to the breakpoint, may have resulted in the formation of this deletion.

In summary, the identification of the precise molecular lesions in families with Fabry disease permits precise carrier detection, delineation of possible genotype/phenotype correlations, and selection of candidates for enzyme-replacement therapy. For this X-linked disease, the ability to identify carrier females by molecular analysis abrogates the need to perform multiple-enzyme assays on individual hair roots (e.g., see Beaudet and Caskey 1978) or to laboriously clone single cells from fibroblast cultures (e.g., see Romeo and Migeon 1970). In fact, recent experience has demonstrated that carriers have been detected by molecular studies of at-risk females who had no clinical findings and normal or near-normal levels of α-Gal A activity in various sources. Clearly, molecular carrier detection should be offered to all at-risk females whenever the mutation in a family is known. In addition, the identification of these mutations should eventually establish the frequency of each Fabry allele and facilitate the recognition of genotype/phenotype relationships. Toward that goal, a Fabry Disease Consortium, analagous to the successful Cystic Fibrosis Consortium, has been established by our laboratory so that investigators can share and communicate identified mutations and phenotypes (contact C.M.E. or R.J.D.). Finally, mutation identification will be important for the selection of patients for clinical evaluation of enzyme replacement in this disease.

Acknowledgments

The authors acknowledge Dr. Yaoming Wang, Dr. Lynn Ierardi, and Annette Enriquez for assistance in DNA sequencing and acknowledge Dr. Thomas Fitzmaurice for ASO screening. We thank Lisa Bland for preparation of the manuscript, and we thank the following physicians for referral of patients and/or diagnostic specimens: Drs. K. David (Brooklyn), A. Hayes (Raleigh), L. J. Elsas II (Atlanta), M. Elleder (Prague), R. R. Frants (Leiden), M. M. Popovtzer (Jerusalem), G. Krueger (Salt Lake City), V. Michaels (Rochester, MN), J. W. Hanson (Iowa City), W. Kruse (Iowa City), and R. E. Pyeritz (Baltimore). This research was supported in part by NIH Merit Award 5 RO1 DK34045, March of Dimes Birth Defects Foundation grant 1-0853, National Center for Research Resources, National Institutes of Health grant 5 MO1 RR00071 to the Mount Sinai General Clinical Research Center, and National Institutes of Health grant 5 P30 HD28822 to the Mount Sinai Child Health Research Center.

References

Beaudet AL, Caskey CT (1978) Detection of Fabry's disease heterozygotes by hair root analysis. Clin Genet 13:251–257

Bernstein HS, Bishop DF, Astrin KH, Kornreich R, Eng CM, Sakuraba H, Desnick RJ (1989) Fabry disease: six gene rearrangements and an exonic point mutation in the α-galactosidase gene. J Clin Invest 83:1390–1399

Bishop DF, Calhoun DH, Bernstein HS, Hantzopoulos P, Quinn M, Desnick RJ (1986) Human α-galactosidase A: nucleotide sequence of a cDNA clone encoding the mature enzyme. Proc Natl Acad Sci USA 83:4859–4863

Bishop DF, Grabowski GA, Desnick RJ (1981) Fabry disease: an asymptomatic hemizygote with significant residual α-galactosidase activity. Am J Hum Genet 33:71A

Bishop DF, Kornreich R, Desnick RJ (1988*a*) Structural organization of the α-galactosidase A gene: further evidence for the absence of a 3′ untranslated region. Proc Natl Acad Sci USA 85:3903–3907

Bishop DF, Kornreich R, Eng CM, Ioannou YA, Fitzmaurice TF, Desnick RJ (1988b) Human α-galactosidase: characterization and eukaryotic expression of the full-length cDNA and structural organization of the gene. In: Salvayre R, Douste-Blazy L, Gatt S (eds) Lipid storage disorders. Plenum, New York, pp 809–822

Bottema CDK, Ketterling RP, Cho HI, Sommer SS (1989) Hemophilia B in a male with a four-base insertion that arose in the germline of his mother. Nucleic Acids Res 17:10139

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Cooper DN, Krawczak M (1991) Mechanisms of insertional mutagenesis in human genes causing genetic disease. Hum Genet 87:409–415

Cooper DN, Youssoufian H (1988) The CpG dinucleotide and human genetic disease. Hum Genet 78:151–155

Coulondre C, Miller JH, Farabangh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in *Escherichia coli*. Nature 274:775–780

DePamphilis ML, Wassarman PM (1980) Replication of eukaryotic chromosomes: a close-up of the replication fork. Annu Rev Biochem 49:627–666

Desnick, RJ, Allen KY, Desnick SJ, Raman MK, Bernlohr RW, Krivit W (1973) Fabry's disease: enzymatic diagnosis of hemizygotes and heterozygotes: α-galactosidase activities in plasma, serum, urine and leukocytes. J Lab Clin Med 81:157–171

Desnick RJ, Bishop DF (1989) Fabry disease: α-galactosidase A deficiency; Schindler disease: α-*N*-acetylgalactosaminidase deficiency. In: Beaudet A, Sly W, Valle D (eds) The metabolic basis of inherited disease, 6th ed. McGraw-Hill, New York, pp 1751–1796

Desnick RJ, Ioannou YA, Eng CM. Fabry disease: α-galactosidase A deficiency. In: Beaudet A, Sly W, Valle D (eds) The metabolic basis of inherited disease, 7th ed. McGraw-Hill, New York (in press)

Dietz HC, Valle D, Francomano CA, Kendzior RJ Jr, Pyeritz R, Cutting GR (1993) The skipping of constitutive exons in vivo induced by nonsense mutations. Science 259:680–683

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, et al (1980) The structure and evolution of the human β-globin gene family. Cell 21:653–668

Elleder M, Bradová V, Šmid F, Buděšínsky M, Harzer K, Kustermann-Kuhn B, Ledvinová, et al (1990) Cardiocyte storage and hypertrophy as a sole manifestation of Fabry's disease. Virchows Arch A Pathol Anat Histopathol 417: 449–455

Eng CM, Niehaus DJ, Desnick RJ (1993) Molecular analysis of classical and variant phenotypes. Pediatr Res 33:128A

Howells DW, Forrest SM, Dahl H-HM, Cotton RGH (1990) Insertion of an extra codon for threonine is a cause of dihydropteridine reductase deficiency. Am J Hum Genet 47:279–285

Hultman T, Stahl S, Hornes E, Uhlen M (1989) Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. Nucleic Acids Res 17:4937–4946

Ioannou YA, Bishop DF, Desnick RJ (1992) Overexpression of human α-galactosidase A results in its crystallization in lysosomes and selective secretion. J Cell Biol 119:1137–1150

Ishii S, Sakuraba H, Shimmoto M, Minamikawa-Tachino R, Suzuki T, Suzuki Y (1991) Fabry disease: detection of a 13-bp deletion in α-galactosidase A gene and its application to gene diagnosis of heterozygotes. Ann Neurol 29:560–564

Ishii S, Sakuraba H, Suzuki Y (1992) Point mutations in the upstream region of the α-galactosidase A gene exon 6 in an atypical variant of Fabry disease. Hum Genet 89:29–32

Koide T, Ishiura M, Iwai K, Inoue M, Kaneda Y, Okada Y, Uchida T (1990) A case of Fabry's disease in a patient with no α-galactosidase A activity caused by a single amino acid substitution of Pro-40 by Ser. FEBS Lett 259:353–356

Kornreich R, Bishop DF, Desnick RJ (1989*a*) The gene encoding α-galactosidase A and gene rearrangements causing Fabry disease. Trans Assoc Am Phys 102:30–43

——— (1990) α-Galactosidase A gene rearrangements causing Fabry disease: identification of short direct repeats at breakpoints in an *Alu*-rich gene. J Biol Chem 265:9319–9326

Kornreich R, Desnick RJ (1993) Fabry disease: detection of gene rearrangements in the human α-galactosidase A gene by multiplex PCR amplification. Hum Mutat 2:108–111

Kornreich R, Desnick RJ, Bishop DF (1989*b*) Nucleotide sequence of the human α-galactosidase A gene. Nucleic Acids Res 17:3301–3302

Krawczak M, Cooper DN (1991) Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. Hum Genet 86:425–441

Romeo G, Migeon BR (1970) Genetic inactivation of the α-galactosidase locus in carriers of Fabry's disease. Science 170:180–181

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350–1354

Sakuraba H, Eng CM, Desnick RJ, Bishop DF (1992) Invariant exon skipping in the human α-galactosidase A pre-mRNA: a g^{+1} to t substitution in a 5′-splice site causing Fabry disease. Genomics 12:643–650

Sakuraba H, Oshima A, Fukuhara Y, Shimmoto M, Nagao Y, Bishop DF, Desnick RJ, et al (1990) Identification of point mutations in the α-galactosidase A gene in classical and atypical hemizygotes with Fabry disease. Am J Hum Genet 47:784–789

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2d ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Vetrie D, Bobrow M, Harris A (1993) Construction of a 5.2 megabase physical map of the human X chromosome at Xq22 using pulsed-field gel electrophoresis and yeast artificial chromosomes. Genomics 15:631–642

von Scheidt W, Eng CM, Fitzmaurice TF, Erdmann E, Hübner G, Olsen EGJ, Christomanou H, et al (1991) An atypical variant of Fabry's disease confined with manifestations to the heart. N Engl J Med 324:395–399

Wang AM, Schindler D, Desnick RJ (1990) Schindler disease: the molecular lesion in the α-*N*-acetylgalactosaminidase gene that causes an infantile neuroaxonal dystrophy. J Clin Invest 86:1752–1756

Weaver DT, DePamphilis ML (1982) Specific sequences in native DNA that arrest synthesis by DNA Polymerase α. J Biol Chem 257:2075–2086