Am. J. Hum. Genet. 62:53–63, 1998

Genotype-Phenotype Correspondence in Sanfilippo Syndrome Type B

Hong G. Zhao, Elena L. Aronovich and Chester B. Whitley

Gene Therapy Program, Institute of Human Genetics, Department of Pediatrics, University of Minnesota, Minneapolis

Summary

Sanfilippo syndrome type B, or mucopolysaccharidosis type IIIB, results from defects in the gene for α-N-acetylglucosaminidase (NAGLU); only a few mutations have been described. To rapidly identify most NAGLU mutations, an automated sequencing procedure was developed for analysis of the entire coding region, including exon-intron borders. By this method, eight affected families were studied, and the mutations in all 16 alleles were identified, more than doubling the number of published mutations for this gene. Eight mutations were described for the first time: five missense mutations (Y140C, Y455C, P521L, S612G, and R674C), two nonsense mutations (W675X and Q706X), and one 24-nucleotide insertion. Currently, 36% of all point mutations (8 of 22 alleles) involve R674, a codon having a CpG dinucleotide in the critical initial position. Other mutations were found in more than one family, raising the possibility that some may be relatively common and, possibly, ancient mutations. Six new nonpathological mutations were also identified and likely represent polymorphic variants of the NAGLU gene, two of which might alter enzyme level. Establishing genotype-phenotype relationships will be vital in the evaluation of experimental treatments such as gene therapy.

Introduction

The Sanfilippo syndromes are the inborn errors of glycosaminoglycan metabolism that are distinguished by the systemic accumulation of heparan sulfate and its abnormal appearance in urine (Sanfilippo and Good 1964; Neufeld and Muenzer 1995). The lysosomal hydrolase α-N-acetylglucosaminidase (NAGLU; E.C.3.2.1.50) is required for the removal of terminal α-N-acetylglucosamine residues from heparan sulfate (von Figura and Kresse 1972), and its deficiency is one of four enzymatic defects associated with the clinical phenotype of Sanfilippo syndrome. In patients lacking NAGLU activity, partially degraded heparan sulfate fragments accumulate in tissues. Presumably a secondary consequence, there are increased levels of gangliosides in brain (Hara et al. 1984; Haust and Gordon 1986), a finding that parallels the progressive cerebral atrophy of this disorder (Murata et al. 1989).

Affected children are usually diagnosed at 2–6 years of age, presenting with hyperactivity, attention deficit disorder, and developmental delay. However, physical stigmata are typically limited to hirsutism, with coarsening of the facial features in some cases. Neurological deterioration progresses slowly throughout childhood, to include the characteristic features of aphasia, vasomotor instability, seizures, dementia, and, eventually, a near-vegetative state. Death ensues, typically during the 2d or 3d decade of life. Patients with a more attenuated phenotype have been reported (Andria et al. 1979; Ballabio et al. 1984).

Recent isolation and characterization of the NAGLU coding sequence and genomic DNA (Zhao et al. 1994, 1996) have been confirmed by independent partial protein sequencing and subsequent expression of a cDNA (Weber et al. 1996). However, only a handful of mutations have been identified (Zhao et al. 1996), in part because of the difficulty encountered in sequencing the highly GC-rich 5′ end of the gene. Despite the difficulties in characterizing NAGLU mutations, delineating such mutations has tremendous value, such as elucidation of the molecular basis of variable phenotypes and facilitation of carrier identification, as well as selection and evaluation of patients undergoing experimental treatments such as bone-marrow transplantation and gene therapy.

This report describes both a method for rapid sequencing of the entire NAGLU coding region and its application in mutation analysis. With this approach, the specific genetic defects in eight families have been characterized, including one patient with the attenuated form of Sanfilippo syndrome type B.

Received August 4, 1997; accepted for publication November 5, 1997; electronically published January 28, 1998.

Address for correspondence and reprints: Dr. Chester B. Whitley, Box 446, UMHC, University of Minnesota, 420 Delaware Street SE, Minneapolis, MN 55455. E-mail: whitley@maroon.tc.umn.edu

© 1998 by The American Society of Human Genetics. All rights reserved.
0002-9297/98/6201-0011$02.00

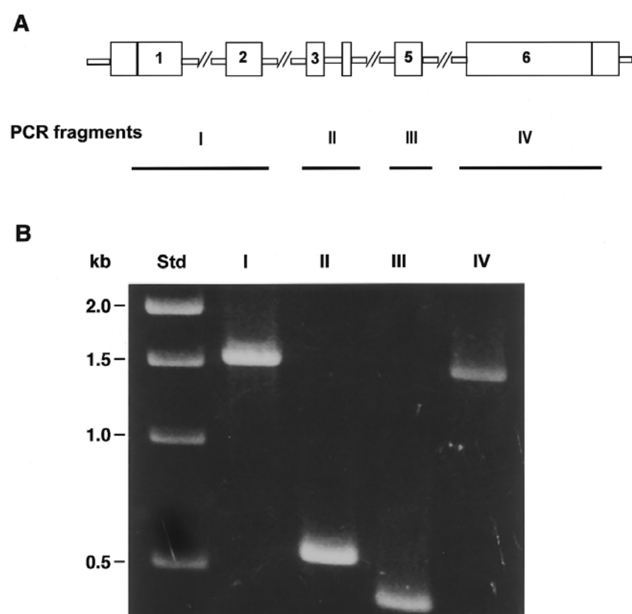

Figure 1 Strategy for PCR amplification of the NAGLU gene coding region, including the exon-intron borders. *A,* Sequence of NAGLU gene, composed of six exons, which were amplified in four PCR fragments. *B,* Products of PCR. Each amplicon was generated as a unique product ready for sequencing.

Patients and Methods

Patients

Subjects were patients of the University of Minnesota whose diagnosis was established on the basis of deficiency of plasma NAGLU enzyme, pathological elevation of heparan sulfate in urine, and the presence of characteristic clinical features.

Molecular-Genetic Methods for Automated Analysis

Extensive work was done to develop a series of oligonucleotide primers and reaction conditions that would yield amplicons and would allow for their sequencing under a uniform set of conditions that have been described, in part, elsewhere (Aronovich et al. 1996) and that are detailed below. Oligonucleotide primers used for PCR amplification and sequencing in this study were synthesized on an Applied Biosystems 391 DNA Synthesizer.

Amplification of NAGLU Genomic DNA

Genomic DNA was isolated, by use of a Puregene DNA isolation kit (Gentra Systems), either from 3 ml whole blood, from 0.5–2 × 10^6 lymphoblastoid cell line (LCL) cells, or from a flask (T-125; Falcon) of cultured fibroblasts. Four PCR reactions were developed to amplify the six exons of the NAGLU gene, including exon-intron junctions (fig. 1). Each PCR was performed on a thermocycler (GeneAmp PCR System 9600; Perkin-Elmer), in a total volume of 50 μl containing 0.6–1 μg genomic DNA, 25 pmol each of a pair of primers (table 1), 800 μM each dNTP, 60 mM Tris-HCl (pH 9.0 at 22°C), 15 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, and 2.5 units thermostable DNA polymerase (*Taq* polymerase; Boehringer Mannheim). For the amplicon extending over exons 1 and 2, the amplification reaction was supplemented to include 10% dimethyl sulfoxide. After denaturation (95°C for 1 min), each of the four separate PCR fragments was generated by 40 cycles of a program comprising denaturation (95°C for 30 s), annealing (55°C for 30 s), and extension (72°C for 30 s). Each of the four PCR products was isolated by use of a QIAquick PCR purification kit (QIAGEN), and an aliquot (3 μl) was evaluated, for concentration and size, by electrophoresis (0.8% agarose gel stained with ethidium bromide), by comparison with a molecular marker (500-bp ladder; Invitrogen).

Sequencing Technique

Automated sequencing was done by cycle sequencing in reactions containing 100 ng DNA template, 3.2 pmol appropriate primer (table 2), and fluorescent dideoxynucleotides (PRISM Ready Reaction Dye Deoxy Terminator cycle sequencing kit; Applied Biosystems). After initial denaturation (96°C for 1 min), cycle sequencing reactions were performed with 30 cycles of a program comprising denaturation (96°C for 10 s), annealing (50°C for 5 s), and elongation (60°C for 4 min). Products of sequencing reactions were isolated by gel-filtration

Table 1

Oligonucleotide Primers for Amplification of the NAGLU Coding Region in Four Segments

			Primer	
Amplicon	Size (bp)	Location (Exon[s])	Upstream (5′ End) (5′→3′)	Downstream (3′ End) (5′→3′)
I	1,540	1 and 2	ACGCCCCCAAGGGAGTATC	GCACGTTGAAAGCACTTCTA
II	560	3 and 4	AGCGCCCAGCACAAAGAA	AAATTCCCTCTCTGAGCTAA
III	433	5	AAACCAGGAGCTGTAGAGAAGT	CTGCCTACCCCTACTGACATCT
IV	1,527	6	GGCCCTCTGTTTCATCACTC	CAAGCGTGGCAGCAGTGACC

Table 2

Oligonucleotide Primers for Automated Sequencing of the NAGLU Coding Region

Primer	Oligonucleotide Sequence (5′→3′)	DNA Strand	Location
1	ACGCCCCCAAGGGAGTAT	Sense	1
2	GTACCCGGAAGCCGCGACTC	Sense	1
3	GCGGCGGTGGGGGTCCTTCT	Sense	1
4	AAGGCAGAGTGGCTGGCTCAA	Antisense	1
5	GCCCAGGCTGTAGGTGTC	Antisense	1
6	CAGAAGGGCCGAGTTTGGAG	Sense	2
7	GCACGTTGAAAGCACTTCTA	Antisense	2
8	AGCACAAAGAAGCAATGAGT	Sense	3
9	GTGTACCTGGCCTTGGGCCT	Sense	3
10	CCTCATCTCCCAGGATACACG	Antisense	3
11	AAATTCCCTCTCTGAGCTAA	Antisense	4
12	CATCTGCTACTCCCTTCTTT	Antisense	5
13	GCTGTTGAACACTATGGTGA	Sense	5
14	CCAAGGCTGGCTCTTCCAGCA	Sense	6
15	TCACAGCCTCTAGGGCTCCAA	Antisense	6
16	CACCATGGTAGGCACGGGCAT	Sense	6
17	TCCTCCGTTCTTCGCGGATGG	Antisense	6
18	TAGCCTCCACGCTGCCCCT	Antisense	6
19	CAACCGATCTGATGTGTTTG	Sense	6
20	TTGGCATAGTCCAGGATGTTG	Antisense	6
21	GGTTGGTGGCCAACTACTACA	Sense	6
22	TGTGATGTCTGTCCAGCTCT	Antisense	6

chromatography (CENTRI-SEP column; Princeton Separations) and were loaded on a 6% denaturing polyacrylamide gel for electrophoresis on an automated sequenator instrument (model 373A; Applied Biosystems). Electronic sequence data were transferred to alignment software (Sequencher 3.0™; Gene Codes Corporation) running on a desktop computer (Power Macintosh model 7600; Apple).

Nomenclature

Mutations are reported in terms of standard nomenclature as recently recommended by consensus groups (Beaudet and Tsui 1993; Beutler et al. 1996). Positions in the DNA and amino acid sequences are enumerated from the first nucleotide of the initial ATG codon (i.e., cDNA designation). However, intronic nucleotides are designated with respect to the genomic sequence (GenBank accession number U43572) and are prefixed with a "g." Mutations have been submitted to the MPS Mutation Database (http://www.peds.umn.edu/Centers/gene/).

Results

Identification of NAGLU Mutations

Prior experience had demonstrated that the GC-rich 5′ end of the gene was extremely difficult to amplify and sequence. Extensive work was done to identify primers and reaction conditions for a reliable and reproducible method, part of which has been outlined in a preliminary report (Aronovich et al. 1996) and is detailed here. Four DNA segments were amplified from genomic template in a manner that included exon-intron borders (fig. 1). These amplicons were then analyzed in a series of 22 overlapping dye-terminator cycle sequencing reactions that evaluated each of 4,060 nucleotide positions, each read in both the 5′→3′ direction and the 3′→5′ direction (fig. 2). Sequencing calls and electrophoretofluorogram tracings from an AB373 Sequenator were transferred to a Macintosh computer and were compared, with the assistance of Sequencher 3.0 software, with the patterns of normal DNA. Eight unrelated families were studied, among which eight novel mutations and six other, unreported DNA variations were identified (table 3).

Family A: Y140C (419A→G) and R674C (2020C→T)

Two mutations were identified in a 24-year-old female with end-stage Sanfilippo syndrome (fig. 3). This neurologically devastated woman used no expressive lan-

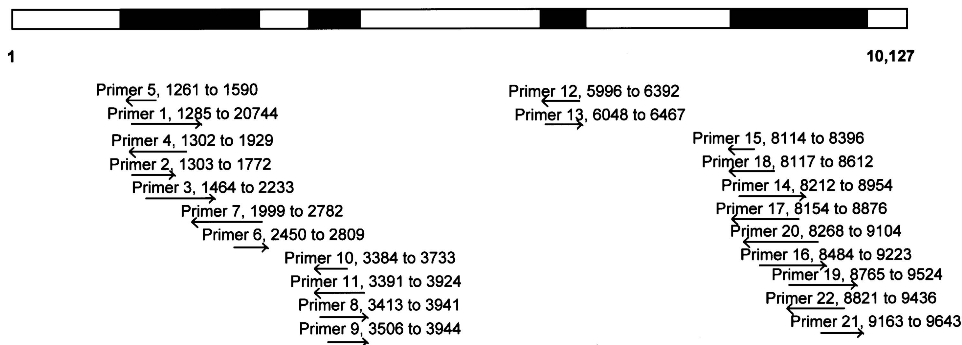

Figure 2 Distribution of 22 sequencing reactions extending over the entire NAGLU coding region. AB373 sequence "base calls" and electrophoretofluorograms are transferred to a free-standing program (Sequencher 3.0), for alignment and analysis. With this strategy, the entire coding region, including exon-intron borders, is sequenced, in both the 5′→3′ direction and the 3′→5′ direction.

Table 3

Summary of Pathological NAGLU Mutations and Clinical Phenotype in Patients with Sanfilippo Syndrome Type B

Family	Allele 1 Mutation (Enzyme Alteration)	Allele 2 Mutation (Enzyme Alteration)	Phenotype	Current Clinical Status
A	419A→G (Y140C)	2020C→T (R674C)	Severe	Living 24-year-old female with end-stage neurological debilitation
B	503delGGAGCGGCCA (truncated protein)	503delGGAGCGGCCA (truncated protein)	Severe	Proband deceased at age 16¼ years
			Severe	Second affected sibling deceased at age 16⅔ years
C	889C→T (R297X)	889C→T (R297X)	Severe	Living, severely demented at age 13 years, despite marrow transplantation
D	1364A→G (Y455C)	2021A→G (R674H)		Living 6-year-old with near-normal development
E	1364A→G (Y455C)	1562C→T (P521L)	Severe	Proband a 28-year-old neurologically devastated female
			Severe	Second affected sibling a 25-year-old neurologically devastated male
			Severe	Third affected sibling a 23-year-old neurologically devastated male
F	2020C→T (R674C)	2021G→A (R674H)	Severe	Living 6-year-old female with moderate mental retardation
G	2024G→A (W675X)	2116C→T (Q706X)	Severe	Living 6-year-old female with moderate mental retardation
H	1834A→G (S612G)	233insGCGGCGCGGCGCGCGTGCGGGTGC	Attenuated	Physically active 32-year-old female with moderate mental retardation
Normal variant	423C→T (S141S)	933C→G (A311A)	...	Child exhibiting reduced NAGLU catalytic activity with normal glycosaminoglycan excretion

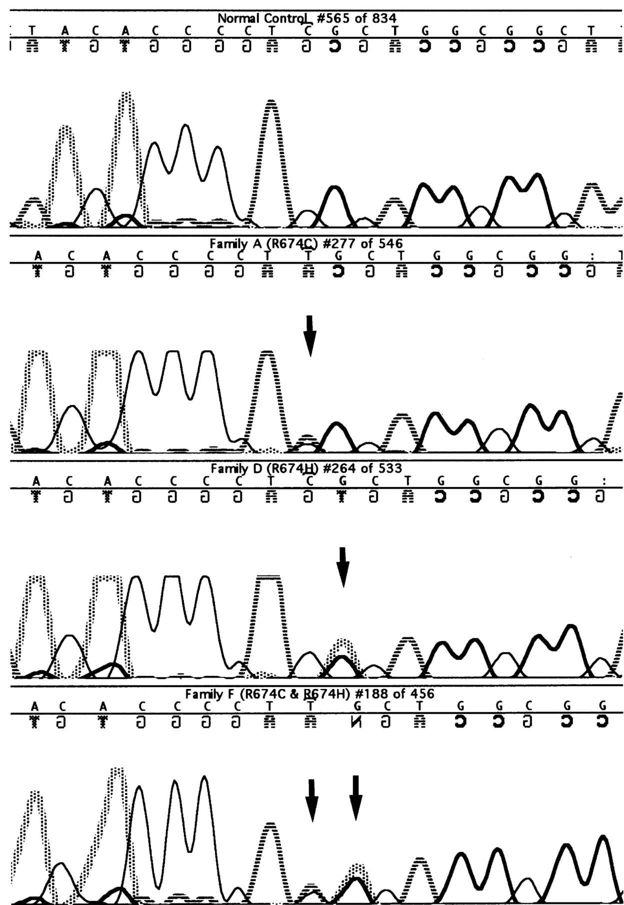

Figure 3 Representative electrophoretofluorographs of an affected individual and probands from families A, D, and F, illustrating the normal and mutant patterns at the R674 hot spot. Arrows indicate mutant nucleotides.

guage and had a spastic neuromuscular examination. During a prolonged period of illness, she had required a gastrostomy tube, which was removed after several months; however, she eats a bulbar diet, which is spoon fed. The proband was found to be heterozygous for two novel mutations, Y140C (419A→G) and R674C (2020C→T). In addition, the proband was found to be heterozygous for a previously unreported silent mutation, 1828C→T (L610L).

Family B: 10-nt Deletion

In an Italian-American family with two deceased, affected children, mutations were sought by analysis of both parents and an unaffected sibling. All three individuals were found to be heterozygous for a 10-nt deletion, 503delGGAGCGGCCA. Homozygosity for this same mutation was previously reported in an independently ascertained proband from Italy (Zhao et al. 1996), thus suggesting the possibility of a common, Italian source.

Family C: R297X (889C→T)

The proband in this family was the first patient in North America to undergo allogeneic bone marrow transplantation for Sanfilippo syndrome type B. As a newborn, the infant had an extremely complicated clinical course. An atonic stomach and related feeding problems required 11 surgical procedures during infancy and led to early recognition of mucopolysaccharidosis (MPS) IIIB, the diagnosis based on unexpected lysosomal inclusions identified in a stomach biopsy. At 20 mo of age, the patient underwent allogeneic bone-marrow transplantation (at the University of Minnesota on November 14, 1986), receiving marrow from his HLA-identical brother. The proband achieved full donor engraftment (>99% donor), as determined by DNA markers. After the transplantation, normal homozygote donor levels of leukocyte NAGLU enzyme activity were achieved (fig. 4*A*). There was a corresponding decrease in urinary glycosaminoglycan excretion (fig. 4*B* and *C*), although levels never reached the normal range. Now, at 13 years of age, the patient is ambulatory, with minimal physical disabilities, but has persistent mixed conductive and sensorineural hearing loss. He remains severely retarded (fig. 4*D*) and has never acquired meaningful expressive language.

From frozen LCL cells stored prior to transplantation, the mutation R297X was observed as the only deviation from the consensus sequence. Although the parents were unaware of any consanguinity, the patient was found to be homozygous throughout the NAGLU coding region.

Family D: Y455C (1364A→G) and R674H (2021G→A)

In a 6-year-old female with essentially normal development, two mutations were identified. The proband was heterozygous for mutation 1364A→G in exon 6, which causes the normal tyrosine to be replaced by cysteine at amino acid 455. This change occurs in a region that is highly conserved in both the human and bovine NAGLU proteins, there being 87.9% identity in a 33-peptide segment centered relative to this position (H. G. Zhao and E. F. Neufeld, unpublished observation). The patient (fig. 3) was also found to be heterozygous for 2021G→A (R674H), a previously reported mutation (Zhao et al. 1996).

Family E: P521L (1562C→T) and Y455C (1364A→G)

In a sibship of two affected males and one affected female with severe disease, the proband was found to be heterozygous for two missense mutations. One mutation, 1364A→G (Y455C), was observed in an unre-

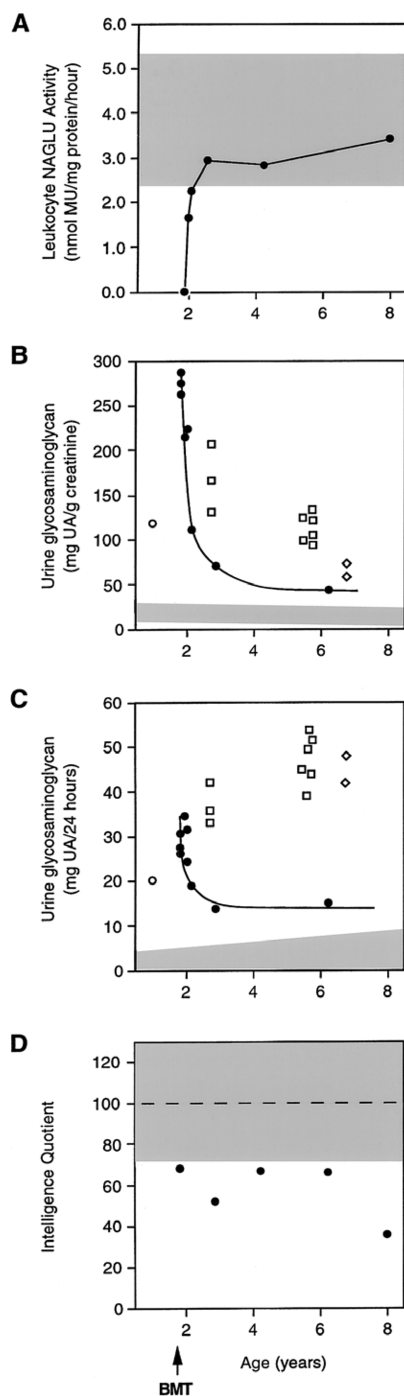

Figure 4 Clinical features of the proband with Sanfilippo syndrome type B, in family C, who underwent bone-marrow transplantation (●). The features compared are leukocyte NAGLU enzyme activity (*A*), urine glycosaminoglycan levels (mg uronic acid/g creatinine [*B*] or mg uronic acid/g creatinine/24 h [*C*]), and intelligence quotient from representative tests or subtests—i.e., either MDI of Bayley Scales or Stanford-Binet Intelligence Scale, 4th ed., composite score (*D*) . Results are plotted against the normal range (*gray shading*) and are compared with values for untreated patients with Sanfilippo syndrome types A (□), B (○), and C (◇).

lated family (family D, discussed above), and the second was a unique mutation, 1562C→T (P521L). In addition, this patient was found to be homozygous for two intronic 1-nt changes, one being a deletion (g2093delA) and the other being an insertion (g2308insA).

Complete analysis of the proband' s unaffected sister demonstrated that she had neither of her sibling's pathological mutations but was homozygous for g2308insA and was also homozygous for another intron mutation, g2744G→C.

Family F: R674C (2020C→T) and R674H (2021G→A)

A 6-year-old girl with moderate mental retardation and typical physical features was found to be heterozygous for two different mutations affecting codon 674 (fig. 3). The patient had the mutation 2020C→T (R674C) seen in the proband of family A (discussed above) and also had 2021G→A (R674H), which previously had been observed in two Arabian patients (Zhao et al. 1996). These mutations were also found in the patient's parents, the former being found in the mother and the latter being present in the father.

Family G: W675X (2024G→A) and Q706X (2116C→T)

In a 6-½-year-old female with moderate mental retardation and typical physical features, two nonsense mutations were identified. The proband was heterozygous for 2024G→A (W675X) and also was heterozygous for 2116C→T (Q706X).

Family H: Attenuated Sanfilippo Syndrome Type B, S612G (1834A→G), and a 24-nt Insertion

Mutations were studied in a unique patient with the attenuated form of Sanfilippo syndrome type B. At age 32 years, the proband was somewhat retarded, as illustrated by her ability to be employed in a sheltered workshop where, on a typical day, she assembled 1,000 cardboard boxes during an 8-h work shift. The proband and her father both were found to have overlapping peaks of guanine and adenine at nucleotide 1834—that is, 1834A→G (S612G). The mother showed the normal sequence, except for a 24-nt in-frame insertion—233insGCGGCGCGGCGCGCGTGCGGGTGC—also found in the proband.

Nonpathological, Reduced NAGLU Activity

A patient of Israeli ethnogeographic background previously had been found to have low NAGLU enzyme activity in the heterozygote range, as measured in plasma and cultured fibroblasts. Mutation analysis demonstrated an entirely normal sequence, except that the patient was heterozygous for two, previously unreported silent mutations—423C→T (S141S) and 933C→G

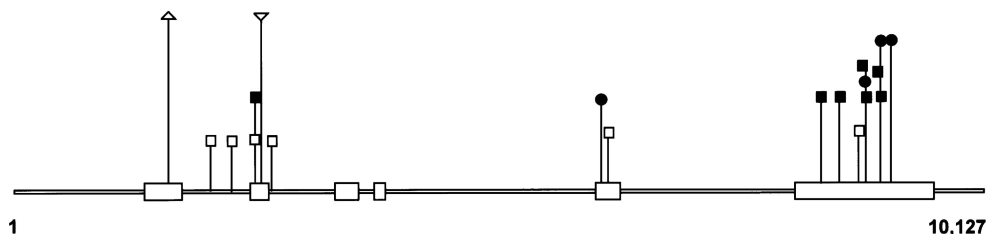

Figure 5 Mutation map of the NAGLU gene. Exons 1–6 are represented by unblackened boxes. The position and clinical effect of the mutations are shown above the gene: ■ = missense mutation; ● = nonsense mutation; □ = silent mutation (presumed polymorphism); Δ = insertion; and ▽ = deletion.

(A311A)—neither of which has been observed in any previous studies. Although these changes would not be expected to have any effect on protein structure, it is possible that one or both could result in reduced translational efficiency because of the use of low-frequency codons. Such changes are known to have significant effects on the biology of other organisms and recently have been demonstrated to be the molecular basis of the Los Angeles biochemical phenotype of galactosemia (Langley et al. 1997).

Discussion

The primary clinical impetus to mutation analysis for Sanfilippo syndrome type B has been to provide more-accurate and specific carrier testing for the relatives of an affected individual. In the past, identification of carriers among relatives has been attempted by quantitative enzyme assays (von Figura et al. 1973, 1975). However, this method is subject to error, because of the variability intrinsic to enzyme-assay methods, differences in polymorphic variants with altered catalytic activity, as well as other potential biological variations (Vance et al. 1980, 1981; Pande et al. 1992).

The automated sequencing method developed here provides a satisfactory method to identify virtually all mutations of the NAGLU gene. Oligonucleotide primers and reaction conditions developed for this study have resolved the difficulties in analysis of the GC-rich 5′ end of the NAGLU gene. To achieve this level of confidence and accuracy, the analytic method was extended beyond the simple acceptance of "base calls" assigned by the sequenator computer; better evaluations were achieved by down-loading electrophoretofluorogram tracings into supplementary alignment and analysis software (i.e., Sequencher), for visual comparison of "test" and "normal" profiles. Analysis is further aided by comparison of multiple sequencing reactions that cover cycle-sequencing reactions reading both the 5′→3′ and 3′→5′ directions of the gene. This bidirectional analysis essentially doubles the number of sequencing reactions required but greatly aides in resolving ambiguities. This redundancy also serves to provide an internal validation of the method.

In this series of patients the method of sequence analysis appears to readily identify concurrent heterozygous mutations. This was demonstrated in cases of known point mutations and known deletions and also was the means by which several previously unknown mutations and polymorphisms were identified. Ongoing experience should further assess the sensitivity of this seemingly reliable system.

Mutation analysis is important for other reasons, not the least of which are to gain insights into the molecular pathology of this disease and to establish genotype-phenotype relationships. (Note that the phrase "genotype-phenotype correlation" has been used to describe the association between particular genotypes and specific clinical phenotypes. Despite the fact that statistical analysis would be particularly valuable, there is rarely any attempt to establish the statistical basis for such associations—e.g., with a "correlation coefficient." To encourage the statistical approach, we reserve the terminology "genotype-phenotype correlation" for those associations in which statistical analysis has been accomplished. When an important genotype-phenotype association is suggested but is not statistically evaluated, we describe this as a "relationship" or "correspondence.")

Rare patients with a more attenuated phenotype of Sanfilippo syndrome have been reported (Andria et al. 1979; van de Kamp et al. 1981; Di Natale et al. 1982; von Figura et al. 1984; van Schrojenstein-de Valk and van de Kamp 1987; Di Natale 1991). Historically, such patients, with a more slowly progressive form of the disease, have been described as representing the "mild form" of Sanfilippo syndrome (or some other MPS disorder). However, the neurologically devastating clinical course is not "mild" in any sense of the word, and this term only serves to mislead clinicians and relatives about the horrendous neurological deterioration of such individuals. We have therefore abandoned this misleading

Table 4

Mutations of the NAGLU Gene

Number	Exon	Genotype	CpG[a]	No. of Independent Observations of Alleles[b]	Biochemical Phenotype (NAGLU Protein)	Clinical Phenotype	Reference
1	1	233insGCGGCGCGGCGCGCGTGCGGGTGC		2	8-Amino-acid insertion	Sanfilippo syndrome[c]	Present report
2	Intron 1	g2089delA		2	Normal	Polymorphism[d]	Present report
3	Intron 1	g2304insA		2	Normal	Polymorphism[d]	Present report
4	2	419A→G		1	Y140C	Sanfilippo syndrome	Present report
5	2	423C→T		1	S141S[e]	Polymorphism[d]	Present report
6	2	503delGGAGCGGCCA		5	Truncated at 182	Sanfilippo syndrome	Zhao et al. (1996)
7	Intron 2	g2739G→C		20	Normal	Polymorphism[d]	Present report
8	5	889C→T	+	3	R297X	Sanfilippo syndrome	Zhao et al. (1996)
9	5	933C→G		1	A311A[e]	Polymorphism[d]	Present report
10	6	1364A→G		2	Y455C	Sanfilippo syndrome	Present report
11	6	1562C→T	+	1	P521L	Sanfilippo syndrome	Present report
12	6	1828C→T		2	L610L	Polymorphism[d]	Present report
13	6	1834A→G		2	S612G	Sanfilippo syndrome[c]	Present report
14	6	1876C→T	+	2	R626X	Sanfilippo syndrome	Zhao et al. (1996)
15	6	1928G→A	+	1	R643H	Sanfilippo syndrome	Zhao et al. (1996)
16	6	2020C→T	+	3	R674C	Sanfilippo syndrome	Present report
17	6	2021G→A	+	7	R674H	Sanfilippo syndrome	Zhao et al. (1996)
18	6	2024G→A		1	W675X	Sanfilippo syndrome	Present report
19	6	2116C→T		1	Q706X	Sanfilippo syndrome	Present report
		Other, normal alleles		9			

[a] A plus sign (+) denotes presence of CpG.

[b] A total of 36 alleles were studied.

[c] Associated with the attenuated phenotype of Sanfilippo syndrome in a single compound-heterozygote individual. Additional cases are required in order to establish the clinical phenotype associated with the allele.

[d] Normal variant allele, which is presumed to be a polymorphism, although the numbers of observations are currently insufficient to meet the formal definition.

[e] Possibly associated with reduced NAGLU catalytic activity; however, additional cases are required in order to confirm this association.

terminology, in favor of a more accurate designation—referring to forms with more slowly evolving disease as the "attenuated form."

In cell-fusion studies using fibroblasts from a patient with the attenuated form (Andria et al. 1979) and cells from more typically severe cases, no complementation was observed in heterokaryons, thus indicating that these forms might be ascribed to allelic mutations (Ballabio et al. 1984). In fact, the molecular biology of several lysosomal storage disorders suggests that some mutations result in "residual enzyme activity" responsible for the attenuated phenotype. The present report includes another case (family H) and confirms that such a patient with attenuated disease does have two allelic mutations. Although there are no biological markers or assays for this phenotype (e.g., residual enzyme activity or glycosaminoglycan levels), a catalog of specific mutations may provide a means for early identification of such individuals.

In the course of the present study of several affected individuals, a series of eight patients revealed eight previously unreported mutations. Five missense mutations (Y140C, Y455C, P521L, S612G, and R674C), two nonsense mutations (W675X and Q706X), and one 24-nt insertion (233insGCGGCGCGGCGCGCGTGCGGTGC) have been characterized. In addition, this study has identified six new polymorphisms: g2089delA, g2304insA, 423C→T (S141S), g2739G→C, 933C→G (A311A), and 1828C→T (L610L). Point mutations, deletions, insertions, and polymorphisms have been reported in the NAGLU gene (fig. 5) and are summarized in table 4.

To date, ~36% of all point mutations (8 of 22 alleles) involve a substitution for arginine at the CGC codon 674. Although the number of subjects is limited to 13 published individuals, this locus appears to be a "hot spot" for 1-bp substitution. Two of the nine most likely 1-nt substitutions at this codon have been observed, both occurring at the CpG dinucleotide: the C→T transition at nucleotide 2020 (CGC→TGC), causing R674C, and the G→A transition at nucleotide 2021 (CGC→CAC), resulting in R674H. Except for this codon, the majority of the currently known pathological mutations in the NAGLU gene are heterogeneous, typically unique to the affected family and usually not found in unrelated patients.

Concomitantly with this study of pathological mutations, six apparently silent mutations were identified (table 4). Of the six variants, each occurred in a single family, except for g2739G→C, which previously had been reported in an unrelated individual (Zhao et al. 1996). Although these mutations are not likely to have any effect on molecular or clinical phenotype, the lack of population studies constrains formal recognition of these changes as "polymorphisms." Two of these mutations, 423C→T (S141S) and 933C→G (A311A), occurred in an individual who had half-normal levels of NAGLU enzyme in plasma and cultured fibroblasts. Such mutations might subject NAGLU mRNA to unfavorable codon utilization and thus reduce the level of enzyme that is synthesized, resulting in a reduced-activity variant. Such biased codon utilization may have broad significance in evolution (Powell and Moriyama 1997) and recently has been proposed as the molecular basis of the Los Angeles biochemical phenotype of galactosemia (Langley et al. 1997). These two NAGLU mutations are probably not pathological, insofar as this individual was found to have normal glycosaminoglycan excretion.

For some patients with MPS diseases, such as Hurler syndrome (MPS type I), very early bone marrow transplantation can result in a "normal" long-term intellectual outcome even in children with "severe" mutations (Whitley et al. 1993*a,* 1993*b*). However, the outcome for children with Sanfilippo syndrome has been less well defined. An earlier report of transplantation in two children with Sanfilippo syndrome type B was equivocal, indicating that neither was as handicapped as an untreated sibling of the same age, although hyperactivity and behavioral problems were present in one of them (Vellodi et al. 1992). In contrast, the outcome for a patient (in family B) in the present study was clearly poor; although the patient achieved donor levels of leukocyte enzyme and reduced glycosaminoglycan excretion, urine heparan sulfate did not reach the normal range, nor was the intellectual outcome good. The establishment of genotype-phenotype relationships would markedly improve the interpretation of such clinical trials, perhaps offering insight into the variations. For example, there is growing concern that the immunogenicity of foreign, albeit therapeutic, proteins may be a limiting factor in cellular therapies. Recent gene-marking studies have observed adverse infusion reactions (Selvaggi et al. 1997), which may be due to introduction of a bovine antigen during cell culture (Muul et al. 1997). In a clinical trial of gene therapy, inclusion of the bacterial neomycin-resistance gene in retroviral vectors has sensitized patients, resulting in T–lymphocyte-mediated elimination of cells (Muul et al. 1997), thereby reducing therapeutic efficacy. Introduction of the normal protein into a patient with a mutant enzyme has the potential to result in similar problems. For example, 13% of patients with Gaucher disease who receive β-glucocerebrosidase develop antibodies against the normal enzyme (Richards et al. 1993), rarely causing anaphylaxis. For this reason, it may be important to know the precise mutations of each patient receiving such therapies, to identify which mutations are likely to be associated with immune elimination/inactivation and other adverse effects.

On the basis of the present and previous studies, there

is an emerging view of the normal NAGLU gene and its variants. Several mutations are recurrent at the "hot spot" R674. However, most others have been observed in at least two unrelated individuals and likely reflect ancient mutational events. Further work is needed in diverse ethnogeographic populations; however, it is becoming clear that the vast majority of mutations result in a relatively severe phenotype, with early and progressive mental retardation. Only one individual with a relatively attenuated phenotype has been studied. However, even this patient is reported to have had progression of intellectual impairment at age 32 years. In this case, two mutations merit further study, and we speculate that at least one of them results in some small level of residual enzyme activity, thus accounting for the attenuated phenotype. Such information may be helpful, both in the provision of prognostic information for newly diagnosed patients and in the evaluation of therapies.

Acknowledgments

We thank Dr. Eric Eccleston (Institute of Human Genetics, University of Minnesota) for operation of the Sequenator. This work was supported by the Children's Medical Research Foundation (Susan and Brad Wilson, cofounders) (Chicago).

References

Andria G, Di Natale P, Del Giudice E, Strisciuglio P, Murino P (1979) Sanfilippo B syndrome (MPS III B): mild and severe forms within the same sibship. Clin Genet 15: 500–504

Aronovich EL, Zhao HG, Neufeld EF, Whitley CB (1996) Mutation analysis in Sanfilippo syndrome type B by automated sequencing of the NAGLU coding region. Am J Hum Genet Suppl 59:A246

Ballabio A, Pallini R, Di NP (1984) Mucopolysaccharidosis III B: hybridization studies on fibroblasts from a mild case and fibroblasts from severe patients. Clin Genet 25: 191–195

Beaudet AL, Tsui LC (1993) A suggested nomenclature for designating mutations. Hum Mutat 2:245–248

Beutler E, McKusick VA, Motulsky AG, Scriver CR, Hutchinson F (1996) Mutation nomenclature: nicknames, systematic names, and unique identifiers. Hum Mutat 8: 203–206

Di Natale P (1991) Sanfilippo B disease: a re-examination of a particular sibship after 12 years. J Inherit Metab Dis 14:23–28

Di Natale P, Murino P, Pontarelli G, Salvatore D, Andria G (1982) Sanfilippo B syndrome (MPS III B): altered residual alpha-N-acetylglucosaminidase activity in an unusual sibship. Clin Chim Acta 122:135–143

Hara A, Kitazawa N, Taketomi T (1984) Abnormalities of glycosphingolipids in mucopolysaccharidosis type III B. J Lipid Res 25:175–184

Haust MD, Gordon BA (1986) Ultrastructural and biochemical aspects of the Sanfilippo syndrome—type III genetic mucopolysaccharidosis. Connect Tissue Res 15: 57–64

Langley SD, Lai K, Dembure PP, Hjelm LN, Elsas LJ (1997) Molecular basis for Duarte and Los Angeles variant galactosemia. Am J Hum Genet 60:366–372

Murata R, Nakajima S, Tanaka A, Miyagi N, Matsuoka O, Kogame S, Inoue Y (1989) MR imaging of the brain in patients with mucopolysaccharidosis. Am J Neuroradiol 10:1165–1170

Muul LM, Soenen SL, Walker R, Ram Z, Oldfield EH, Blaese RM (1997) Studies of the immune response in patients receiving gene therapy. In: Blau HM, Wilson JM (eds) Molecular and cellular biology of gene therapy. Keystone Symposia, Snowbird, UT, pp 52

Neufeld EF, Muenzer J (1995) The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular basis of inherited disease, 7th ed. McGraw-Hill, New York, pp 2465–2494

Pande H, Chester A, Lie H, Thorsby E, Stormorken H (1992) Concomitant occurrence of mucopolysaccharidosis IIIB and Glanzmann's thrombasthenia: further evidence of a hyperactive alpha-N-acetylglucosaminidase-producing allele. Clin Genet 41:243–247

Powell JR, Moriyama EN (1997) Evolution of codon usage bias in *Drosophila*. Proc Natl Acad Sci USA 94: 7784–7790

Richards SM, Olson TA, McPherson JM (1993) Antibody response in patients with Gaucher disease after repeated infusion with macrophage-targeted glucocerebrosidase. Blood 82:1402–1409

Sanfilippo SJ, Good RA (1964) A laboratory study of the Hurler syndrome. Am J Dis Child 102:140

Selvaggi TA, Walker RE, Fleisher TA (1997) Development of antibodies to fetal calf serum with arthus-like reactions in human immunodeficiency virus-infected patients given syngeneic lymphocyte infusions. Blood 89:776–779

Vance JM, Conneally PM, Wappner RS, Yu PL, Brandt IK, Pericak-Vance MA (1981) Carrier detection in Sanfilippo syndrome type B: report of six families. Clin Genet 20: 135–140

Vance JM, Pericak-Vance MA, Elston RC, Conneally PM, Namboodiri KK, Wappner RS, Yu PL (1980) Evidence of genetic variation for alpha-N-acetyl-D-glucosaminidase in black and white populations: a new polymorphism. Am J Med Genet 7:131–140

van de Kamp J, Niermeijer MF, von Figura K, Giesberts MA (1981) Genetic heterogeneity and clinical variability in the Sanfilippo syndrome (types A, B, and C). Clin Genet 20:152–160

van Schrojenstein-de Valk H, van de Kamp J (1987) Follow-up on seven adult patients with mild Sanfilippo B-disease. Am J Med Genet 28:125–129

Vellodi A, Young E, New M, Pot-Mees C, Hugh-Jones K (1992) Bone marrow transplantation for Sanfilippo disease type B. J Inherit Metab Dis 15:911–918

von Figura K, Hasilik A, Steckel F, van de Kamp J (1984) Biosynthesis and maturation of α-*N*-acetylglucosaminidase in normal and Sanfilippo B-fibroblasts. Am J Hum Genet 36:93–100

von Figura K, Kresse H (1972) The Sanfilippo B corrective

factor: a N-acetyl-alpha-D-glucosamindiase. Biochem Biophys Res Commun 48:262–269

von Figura K, Logering M, Kresse H (1975) Serum alpha-N-acetylglucosaminidase: determination, characterization, and corrective activity in Sanifilippo B fibroblasts. Z Klin Chem Klin Biochem 13:285–289

von Figura K, Logering M, Mersmann G, Kresse H (1973) Sanfilippo B disease: serum assays for detection of homozygous and heterozygous individuals in three families. J Pediatr 83:607–611

Weber B, Blanch L, Clements PR, Scott HS, Hopwood JJ (1996) Cloning and expression of the gene involved in Sanfilippo B syndrome (mucopolysaccharidosis III B). Hum Mol Genet 5:771–777

Whitley CB, Belani KG, Chang PN, Summers CG, Blazar BR, Tsai MY, Latchaw RE, et al (1993) Long-term outcome of Hurler syndrome following bone marrow transplantation. Am J Med Genet 46:209–218

Whitley CB, Krivit W, Ramsay NKC, Kersey JH, Chang PN, Latchaw RE, Blazar BR, et al (1993) Mutation analysis and clinical outcome of patients with Hurler syndrome (mucopolysaccharidosis type I-H) undergoing bone marrow transplantation. Am J Hum Genet Suppl 53:101

Zhao HG, Li HH, Bach G, Schmidtchen A, Neufeld EF (1996) The molecular basis of Sanfilippo syndrome type B. Proc Natl Acad Sci USA 93:6101–6105

Zhao HG, Lopez R, Rennecker J, Neufeld EF (1994) Sanfilippo syndrome type B: cDNA and gene encoding human α-N-acetyl glycosaminidase. Am J Hum Genet Suppl 55: A252