

RAPID COMMUNICATION

Mutations in Jewish Patients With Gaucher Disease

By Ernest Beutler, Terri Gelbart, Wanda Kuhl, Ari Zimran, and Carol West

DNA from 100 unrelated patients, 97 of whom were Jewish and three half-Jewish, was analyzed for 22 mutations known to cause Gaucher disease. All but seven of the alleles were identified as having previously described mutations. Five of the unidentified mutations proved to be a previously undescribed nucleotide substitution in a splice junction (IVS2+1) that causes skipping of exon 2. Thus, only 2 of 197 alleles remained unidentified. Homozygotes for the most common mutation, that at nucleotide (nt) 1226, manifested, on average, the mildest disease and the latest age of onset. The mutation at nt 84 and the newly described IVS2+1 mutation,

which do not produce any enzyme, were associated with earlier onset and more severe disease. Five of the mutations were considered to be "public," in the sense that they were found in more than one unrelated individual. Screening for these five mutations permitted detection of 97.5% of all Gaucher alleles in this patient population. Because the mutation at nt 1226 is underrepresented in the patient population and because not all homozygotes come to medical attention, screening the Ashkenazi population using DNA analysis should detect approximately 99% of all heterozygotes.

© 1992 by The American Society of Hematology.

GAUCHER DISEASE is the most common glycolipid storage disorder. Although it is panethnic in distribution it is most common by far in the Jewish population, where the gene frequency has been reported to be on the order of .035 to .040.² Although DNA-based analysis for preselected mutations can never detect all of the mutations causing a disease, an understanding of the frequency of the various disease-producing alleles makes it possible to establish the most rational possible approach to screening. Therefore, we have now examined 197 Gaucher disease alleles from unrelated Jewish patients or the Jewish parent of half-Jewish patients. These studies indicate that mutations at nucleotide (nt) 1226, 84, 1448, 1297, and in intron 2 are found repeatedly in unrelated Jewish subjects and account for 97.5% of all the mutations that cause Gaucher disease in this population.

MATERIALS AND METHODS

Patients. One hundred unrelated Gaucher disease patients, 97 Ashkenazi Jewish and three half-Jewish, were investigated. In the case of the half-Jewish patients, DNA from the Jewish parent was available for study. About two thirds of the patients were seen personally by the investigators. Material from the others was submitted by their physicians. In most cases the diagnosis had been established on the basis of histologic examination of marrow, liver, or spleen, and in all patients seen by the investigators the diagnosis had been confirmed by the measurement of peripheral blood

lymphocyte β -glucosidase activity.³ Clinical severity was assessed using a scoring system that we have described previously,⁴ modified so that age at diagnosis is not a component of the severity but rather is treated separately.^{5,6}

Detection of previously described mutations. High-molecular-weight DNA was isolated from peripheral blood leukocytes using an Applied Biosystems (Foster City, CA) DNA extractor. All samples were haplotyped for the polymorphic *Pvu* II site⁷ using the polymerase chain reaction (PCR) as previously described.⁸ Each sample was examined first for the mutation at nt 1226, and except as noted, DNA from individuals who were classified as 1226G/1226G homozygotes was not examined further. All other samples were examined for the 84GG mutation using either mismatched PCR⁹ or allele-specific oligonucleotide hybridization, and for the 1448C mutation using allele-specific oligonucleotide hybridization or previously described methods.⁴ Samples that still had unidentified alleles were examined for all other published mutations using the techniques summarized in Table 1.

Detection of a new mutation. When the mutation in a patient had not been identified as one of those previously described and cultured lymphoblasts were available, total RNA was isolated from the cultured cells.²¹ cDNA was made from 2 to 4 μ g of total RNA using reverse transcriptase for first-strand synthesis and the PCR with primers designed to amplify the cDNA in two overlapping fragments. To create the 5' fragment, first-strand synthesis was performed using an antisense primer made to cDNA nt 1298* through 1317 as previously described.²³ This was followed by 35 cycles of the PCR using sense primer nt -46 through -27 and antisense primer 1271 through 1290. Single-stranded DNA was made using the PCR with sense primer nt -25 through -6 for the sense strand and antisense primer nt 891 through 910 for the antisense strand. First-strand synthesis for the 3' fragment was primed with antisense to cDNA nt 2011 through 2030 and PCR was performed with sense primer nt 468 through 487 and antisense primers nt 1821 through 1840. The sense strand for sequencing was primed with oligonucleotide nt 871 through 890 and antisense was primed at nt 1621 through 1640. Sequencing was performed using T7 DNA polymerase with primers spaced about 200 nucleotides apart along the cDNA.

Genomic sequencing. Genomic DNA was amplified using 50 ng each of each of sense primer genomic 917† through 936 and antisense primer genomic 1558 through 1576 for 15 cycles followed by 35 cycles of amplification with antisense primer genomic 1317

From the Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA; and Shaare Zedek Medical Center, Jerusalem, Israel.

Submitted December 20, 1991; accepted January 17, 1992.

Supported by National Institutes of Health Grants DK36639 and RR00833 and the Sam Stein and Rose Stein Charitable Trust Fund.

This is manuscript 7154-MEM from The Scripps Research Institute.

Presented, in part, at the 32nd Annual Meeting of the American Society of Hematology, Dec 7-10, 1991, Denver, CO.¹

Address reprint requests to Ernest Beutler, MD, Department of Molecular and Experimental Medicine, The Scripps Research Institute, 10666 N Torrey Pines Rd, La Jolla, CA 92037.

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. section 1734 solely to indicate this fact.

© 1992 by The American Society of Hematology.

0006-4971/92/7907-0036\$3.00/0

*cDNA nucleotides are numbered starting with the upstream ATG as no. 1.²²

†Genomic nucleotides are numbered according to Horowitz et al.²⁴

Table 1. Mutations Detected in the Present Study and Method of Detection

cDNA No.	Genomic No.	nt SUB	Sense Primer Genomic No.	Antisense Primer Genomic No.	Restriction Enzyme	Normal Pattern bp	Mutant Pattern bp	Reference
84	1035	G → GG	979-997	1035-1053	(<i>BsaBI</i>)	74	56, 19	9
Intron 2	1067	g → a*	979-997	1316-1335	<i>HphI</i>	141, 117, 99	240, 117	This report
476	3060	G → A	2978-2997	3121-3140	<i>BstNI</i>	163	82, 81	10
481	3065	C → T	2978-2997	3474-3493	<i>KpnI</i>	428, 88	516	Unpublished
535	3119	G → C	3052-3071	3474-3493	<i>BspHI</i>	442	375, 67	11
586	3170	A → C	3052-3071	3474-3493	<i>ScrFI</i>	220, 166, 48, 6	166, 111, 109, 48, 8	12
751	3545	T → C	3465-3484	3791-3810	<i>DraIII</i>	346	280, 86	Unpublished
754	3548	T → A*	3528-3547	3791-3810	(<i>NsiI</i>)	283	262, 21	13
764	4113	T → A	3465-3484	4240-4259	<i>KpnI</i>	795	648, 147	14
1043	5259	C → T	5183-5202	5307-5326	<i>BanI</i>	86, 58	144	12
1053	5269	G → T	5183-5202	5307-5326	<i>KpnI</i>	90, 54	144	12
1090	5306	G → A*	5156-5175	5490-5510	<i>Bsu36I</i>	355	205, 150	15
1093	5309	G → A	5183-5202	5307-5326	<i>BsmI</i>	121, 23	144	11
1141	5357	T → G	5156-5175	5490-5510	<i>StuI</i>	199, 156	355	15
1208	5424	G → C	5405-5423	5490-5510	(<i>AccI</i>)	107	90, 17	12
1226	5841	A → G	5821-5840	5906-5925	(<i>XhoI</i>)	105	89, 16	16
1297	5912	G → T	5865-5885	5914-5933	(<i>BanI</i>)	47, 22	69	17
1342	5957	G → C*	5878-5897	6468-6487	<i>StyI</i>	342, 156, 75, 37	342, 156, 112	15
1343	5958	A → T	5878-5897	6031-6050	<i>AflIII</i>	92, 81	173	17
1361	5976	C → G	5878-5897	6031-6050	<i>HhaI</i>	173	98, 75	18
1448	6433	T → C*	5183-5202	6516-6535	<i>NciI</i>	1353	1250, 103	19
1504	6489	C → T	6463-6482	6516-6535	<i>MspI</i>	47, 26	73	20
1604	6683	G → A	6632-6651	6782-6801	<i>HphI</i>	170	128, 42	Unpublished

(), Mismatched to create restriction site.

*Sequence matches glucocerebrosidase pseudogene at this position.

through 1336. The product was sequenced with T7 DNA polymerase using as an oligonucleotide primer genomic 989 through 1008.

RESULTS

Discovery of a new splicing mutation. Sequencing of cDNA from a patient heterozygous for the 1226G mutation, whose other allele did not contain any of the known mutations, showed that exon 2 was missing. Accordingly, the genomic sequence of the area including the splice donor of intron 2 was determined. A point mutation was found at genomic nucleotide 1067, the first nucleotide in intron 2, which was changed from g to a. This mutation is designated IVS2+1, to designate that the first nucleotide of the 5' end of the intervening sequence is mutated. Samples from those patients in whom both mutations had not been found were screened for this mutation as indicated in Table 1.

Incidence of various mutations. The results of genotyping are summarized in Tables 2 and 3. Both of the disease-producing alleles were identified in 95 of the 97 Jewish patients. In the three half-Jewish patients, the Jewish allele was identified in all cases by analysis of DNA from the parents. Only 2 of 197 Jewish alleles remained unidentified, and 90.4% of the Gaucher disease alleles were either the 1226G or 84GG mutations. The most common genotype was, as expected, the 1226G/1226G homozygous state, and the second was the 1226G/84GG heterozygous state.

Genotype/phenotype correlation. The relationship between clinical severity, age at evaluation, and the genotype is summarized in Fig 1. The average severity of clinical manifestations in the patients homozygous for the 1226G

mutation was the least of any genotype. The average severity was considerably greater for the patients who were heterozygous for the 1226G, and either the 84GG, IVS2, or 1448C mutations. The median, mean, and range of age of onset for patients with different genotypes is summarized in Table 4.

DISCUSSION

When a disease-producing gene exists at polymorphic levels in a population, it may be presumed to confer some advantage, usually among heterozygotes. When such a selective advantage does exist, more than one mutation often achieves a high population frequency. This seems to be the case with Gaucher disease where it has previously been found that two mutations, 1226G and 84GG, both achieve relatively high gene frequencies in the Jewish population.⁹ Although the vast majority of mutations may

Table 2. Frequency of Jewish Gaucher Disease Mutations in 97 Jewish and 3 Half-Jewish Patients

Mutation	No. of Alleles	%
1226G	152	77.16
84GG	26	13.20
1448C and XOVR*	6	3.05
Intron 2	5	2.54
1297T	3	1.52
1604A	1	0.51
1504T	1	0.51
751C	1	0.51
?	2	1.02

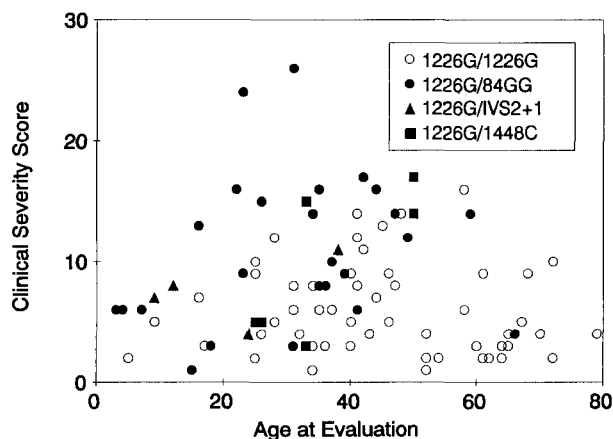
*XOVR is a fusion gene.²⁵

Table 3. Haplotypes of 97 Jewish Patients With Gaucher Disease (half-Jewish patients excluded)

Genotype	No. of Patients			
	Total	Pv1.1 ⁺ /Pv1.1 ⁺	Pv1.1 ⁺ /Pv1.1 ⁻	Pv1.1 ⁻ /Pv1.1 ⁻
1226G/1226G	54	0	0	54
1226G/84GG	25	0	25	0
1226G/IVS2+1	5	0	0	5
1226G/1448C	4	0	3	1
1226G/1297T	3	0	0	3
1226G/XOVR	1	0	0	1
1226G/751C	1	0	1	0
1448C/1504T	1	0	1	0
1604A/84GG	1	1	0	0
1226G/?	2	0	0	2

be regarded as being "public" in that they are found repeatedly in unrelated individuals, there is a baseline frequency of new ("private") mutations, such as will occur with any gene in any population. Thus, disease diagnosis and/or screening the heterozygotes can never be 100% accurate if based on the detection of specific, prevalent mutations. Therefore, it is essential to know which disease alleles exist in a population and to assess their frequency, an undertaking that requires the examination of large numbers of alleles.

In the present study, we have been able to obtain DNA samples from 100 patients with Gaucher disease, representing 197 alleles of Jewish origin. The mutations in all but two of these alleles have been identified at the DNA level. Most of these mutations have been described previously, but the mutation in intron 2 was detected for the first time in this study, although the same mutation has apparently been detected independently by G. He and G.A. Grawbowski (personal communication, December 1991). Like several other mutations that cause Gaucher disease, this mutation represents the normal pseudogene sequence. Like the pseudogene-derived transcript, the transcript from the mutant gene lacks exon 2, the usual consequence of a mutation in the 5' splice consensus.²³

**Fig 1. The relationship between clinical severity, age of evaluation, and the Gaucher disease genotype of Jewish patients.****Table 4. Age of Onset of Symptoms in Patients With the Four Most Common Genotypes**

Genotype	Age of Onset (yr)		
	Median	Mean	Range
1226G/1226G	26	28.9	2-70
1226G/84GG	5	12.8	1-66
1226G/1448C	15	16.5	3-33
1226G/IVS2+1	4	9.4	1.5-2.8

Four of the mutations may be regarded as being "public" in that they occurred in more than one family in this population, always in a single haplotype. A fifth mutation, that at nt 1448, is public in the sense that it occurs repeatedly but is actually probably sporadic in nature because it occurs both in the Pv1.1⁺ and Pv1.1⁻ haplotype, as we have previously shown.⁸ This substitution seems to be in a mutational hot spot, and may be formed repeatedly either by point mutations or by gene conversion, since the pseudogene contains this sequence.

Together, these five mutations accounted for 192 (97.5%) of the disease-producing alleles. The mutations exist in two clusters. At the 5' end of the gene are the 84GG and IVS2+1 mutations; near the 3' end of the gene are the 1226G, 1297T, and 1448C mutations. The clustering of the mutations is fortunate from a practical point of view, because it facilitates screening for all five based on PCR amplification of merely two segments of the gene. This, in point of fact, is readily accomplished simultaneously in one step using 4 amplimers (unpublished data).

The present study confirms, in a larger population of Ashkenazi Jewish patients, that the most common mutations are the ones at nt 1226 and 84.^{2,4,6} It is also apparent that the great majority of alleles that cause Gaucher disease in the Jewish population have now been identified. The 1226G and 84GG mutations accounted for 90.4% of the mutations in the present series. Including the other three public mutations, the percentage of mutations identified in Jewish Gaucher disease patients increases to 97.5%.

It is important to recognize that the proportion of the different mutations to each other in the population of Gaucher disease patients, viz, 152:26:6:5:3 (1226G:84GG:1448C:IVS2+1:1297T), does not constitute an unbiased estimate of the relative frequency of these mutations in the population as a whole. This is the case because the clinical manifestations of the common 1226G/1226G genotype may be minimal, or not exist at all,² so that the frequency of the 1226G mutation is bound to be underestimated with respect to the other, more severe mutations.

We have previously pointed out that patients with 1226G/1226G homozygotes have a milder disease course than do patients with the other mutations,^{4,5,26} and this observation has been confirmed by others.^{27,28} The milder nature of disease produced by the 1226G allele was also evident in the present study, although as before, there was considerable variability in the severity of the disease within each genotype. Not only do patients homozygous for the 1226G genotype tend to have milder disease, but it is not an

infrequent experience to find siblings of patients with the 1226G homozygous genotype who are totally unaware of the fact that they have Gaucher disease. Moreover, 1226G homozygotes have been encountered in a survey of the normal Ashkenazi Jewish population.² Therefore, it follows that the population of patients with Gaucher disease seen by physicians and referred to centers is biased toward the mutations that produce the most severe disease genotype, namely the 84GG mutation, which makes no protein product, and the IVS2+1 mutation which, presumably, is also incapable of directing the production of any enzyme. Therefore, the percentage of 1226G alleles in any Ashkenazi Jewish patient population will always represent an underestimation of the relative population frequency of this mutation. We have recently examined over 1,000 blood samples from healthy Ashkenazi Jewish subjects for the 1226G and 84GG alleles. Fifty-nine 1226G alleles were found among 1,145 samples and only 6 84GG alleles in 1,053 samples examined, a ratio of 9:1, which is as predicted, higher than the 5.8:1 ratio found in the Gaucher

disease patient population. From the point of view of population screening, 97.4% of alleles detected among patients represents an underestimate and the percentage of Gaucher disease alleles represented by one of the public alleles described here would probably detect about 99% of all Gaucher disease producing alleles in the Ashkenazi Jewish population.

The present investigation has been limited to the population in which Gaucher disease is most common, Ashkenazi Jews. In other populations, where Gaucher disease is not polymorphic, the incidence of the "public" alleles is much lower. It appears that less than 80% of the Gaucher disease producing alleles in the non-Jewish population fall into this category, and most of these are the 1448C mutation. The latter mutation is not really public in the sense of being the result of a founder effect, but represents the result of a mutational hot spot. The representation of more "private" alleles in such populations is to be expected because proportionately a much larger number of mutations represent rare sporadic events.

REFERENCES

1. Beutler E, Sorge JA, Kuhl W, West C, Gelbart T: The modern diagnosis of Gaucher disease: Discovery of a second common Jewish mutation makes DNA-based diagnosis and screening feasible. *Blood* 78:165a, 1991 (abstr, suppl)
2. Zimran A, Gelbart T, Westwood B, Grabowski GA, Beutler E: High frequency of the 1226 mutation for type I Gaucher disease among the Ashkenazi Jewish population. *Am J Hum Genet* 49:855, 1991
3. Beutler E, Kuhl W: Detection of the defect of Gaucher's disease and its carrier state in peripheral-blood leukocytes. *Lancet* 1:612, 1970
4. Zimran A, Sorge J, Gross E, Kubitz M, West C, Beutler E: Prediction of severity of Gaucher's disease by identification of mutations at DNA level. *Lancet* 2:349, 1989
5. Beutler E: Gaucher's disease. *N Engl J Med* 325:1354, 1991
6. Zimran A, Kay AC, Gelbart T, Garver P, Saven A, Beutler E: The natural history of adult type Gaucher disease. Clinical, laboratory, radiologic and genetic features of 53 patients. In preparation, 1992
7. Sorge J, Gelbart T, West C, Westwood B, Beutler E: Heterogeneity in type I Gaucher disease demonstrated by restriction mapping of the gene. *Proc Natl Acad Sci USA* 82:5442, 1985
8. Zimran A, Gelbart T, Beutler E: Linkage of the Pvu II polymorphism with the common Jewish mutation for Gaucher disease. *Am J Hum Genet* 46:902, 1990
9. Beutler E, Gelbart T, Kuhl W, Sorge J, West C: Identification of the second common Jewish Gaucher disease mutation makes possible population based screening for the heterozygote state. *Proc Natl Acad Sci USA* 88:10544, 1991
10. Graves PN, Grabowski GA, Eisner R, Palese P, Smith FI: Gaucher disease type I: Cloning and characterization of a cDNA encoding acid β -glucosidase from an Ashkenazi Jewish patient. *DNA* 7:521, 1988
11. Eyal N, Firon N, Wilder S, Kolodny EH, Horowitz M: Three unique base pair changes in a family with Gaucher disease. *Hum Genet* 87:328, 1991
12. Latham TE, Theophilus BDM, Grabowski GA, Smith FI: Heterogeneity of mutations in the acid β -glucosidase gene of Gaucher disease patients. *DNA Cell Biol* 10:15, 1991
13. Kawame H, Eto Y: A new glucocerebrosidase-gene missense mutation responsible for neuronopathic Gaucher disease in Japanese patients. *Am J Hum Genet* 49:1378, 1991
14. Beutler E, Gelbart T: Gaucher disease associated with a unique *KpnI* restriction site: Identification of the amino acid substitution. *Ann Hum Genet* 54:149, 1990
15. Eyal N, Wilder S, Horowitz M: Prevalent and rare mutations among Gaucher patients. *Gene* 96:277, 1990
16. Tsuji S, Martin BM, Barranger JA, Stubblefield BK, La-Marca ME, Ginns EI: Genetic heterogeneity in type 1 Gaucher disease: Multiple genotypes in Ashkenazic and non-Ashkenazic individuals. *Proc Natl Acad Sci USA* 85:2349, 1988
17. Theophilus BDM, Latham T, Grabowski GA, Smith FI: Comparison of RNase A, chemical cleavage, and GC-clamped denaturing gradient gel electrophoresis for the detection of mutations in exon 9 of the human acid β -glucosidase gene. *Nucleic Acids Res* 17:7707, 1989
18. Wigderson M, Firon N, Horowitz Z, Wilder S, Frishberg Y, Reiner O, Horowitz M: Characterization of mutations in Gaucher patients by cDNA cloning. *Am J Hum Genet* 44:365, 1989
19. Tsuji S, Choudary PV, Martin BM, Stubblefield BK, Mayor JA, Barranger JA, Ginns EI: A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease. *N Engl J Med* 316:570, 1987
20. Hong CM, Ohashi T, Yu XJ, Weiler S, Barranger JA: Sequence of two alleles responsible for Gaucher disease. *DNA Cell Biol* 9:233, 1990
21. Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156, 1987
22. Sorge J, West C, Westwood B, Beutler E: Molecular cloning and nucleotide sequence of the human glucocerebrosidase gene. *Proc Natl Acad Sci USA* 82:7289, 1985
23. Sorge J, Gross E, West C, Beutler E: High level transcription of the glucocerebrosidase pseudogene in normal subjects and patients with Gaucher disease. *J Clin Invest* 86:1137, 1990
24. Horowitz M, Wilder S, Horowitz Z, Reiner O, Gelbart T, Beutler E: The human glucocerebrosidase gene and pseudogene: Structure and evolution. *Genomics* 4:87, 1989

25. Zimran A, Sorge J, Gross E, Kubitz M, West C, Beutler E: A glucocerebrosidase fusion gene in Gaucher disease. Implications for the molecular anatomy, pathogenesis and diagnosis of this disorder. *J Clin Invest* 85:219, 1990
26. Zimran A, Sorge J, Gross E, Beutler E: Prediction of the clinical course of Gaucher disease by identification of mutation at DNA level. *Blood* 72:155a, 1988 (abstr, suppl)
27. Theophilus B, Latham T, Grabowski GA, Smith FI: Gaucher disease: Molecular heterogeneity and phenotype-genotype correlations. *Am J Hum Genet* 45:212, 1989
28. Kolodny EH, Firon N, Eyal N, Horowitz M: Mutation analysis of an Ashkenazi Jewish family with Gaucher disease in three successive generations. *Am J Med Genet* 36:467, 1990