RAPID COMMUNICATION

Mutations in Jewish Patients With Gaucher Disease

By Ernest Beutler, Terri Gelbart, Wanda Kuhl, Ari Zimmerman, 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 0.035 to 0.040.2 Although DNA-based analysis for preslected 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 confirmed by the measurement of peripheral blood lymphocyte β-glucosidase activity. 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

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 site6 using the polymerase chain reaction (PCR) as previously described. 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 PCR7 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.8 cDNA was made from 2 to 4 μg 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 previously described.9 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 1211 through 1309 and PCR was performed with sense primer nt 468 through 487 and antisense primer nt 1821 through 1840. The sense strand for sequencing was primed with oligonucleotide nt 871 through 890 and antisense primer 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

© 1992 by The American Society of Hematology.

which was changed from g to a. This mutation is designated
merase using as an oligonucleotide primer genomic 989 through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
through
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. Grarowski (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

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 Pvl.1+ and Pvl.1- haplotype, as we have previously shown.23 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 84.23,24 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,3 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,5,23 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
infrquent 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. 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 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


Mutations in Jewish patients with Gaucher disease

E Beutler, T Gelbart, W Kuhl, A Zimran and C West