*Am. J. Hum. Genet. 46:902–905, 1990*

# Linkage of the *Pvu*II Polymorphism with the Common Jewish Mutation for Gaucher Disease

Ari Zimran, Terri Gelbart, and Ernest Beutler

Scripps Clinic and Research Foundation, Research Institute of Scripps Clinic, La Jolla, CA

## Summary

**We have localized the *Pvu*II polymorphism of the glucocerebrosidase gene complex to intron 6 of the active gene. Using the polymerase chain reaction (PCR) to amplify intron 6 of DNA samples from Pv1.1$^-$/Pv1.1$^+$ individuals, we defined the mutation causing this polymorphism as a G→A single-base substitution at position 3931 of the active gene. By analyzing 54 unrelated Gaucher patients we show strong linkage disequilibrium between the Pv1.1$^-$ genotype and the common Jewish mutation 1226 causing the adult type of this disease. Gaucher disease patients heterozygous for the 1226 allele and one unidentified allele (1226/?), particularly those of Jewish ancestry, were predominantly of the Pv1.1$^-$/PV1.1$^+$ genotype. This suggests that one of the unknown alleles may be relatively common and linked to the Pv1.1$^+$ genotype.**

## Introduction

Gaucher disease is a glycolipid storage disorder caused by a deficiency of the enzyme glucocerebrosidase. After cloning the cDNA for this enzyme, we were able to discover an RFLP with the enzyme *Pvu*II by digesting genomic DNA with 20 different restriction endonucleases (Sorge et al. 1985). This polymorphism was characterized by the presence or absence of a 1.1-kb fragment. Genes that generated the 1.1 fragment after digestions were designated Pv1.1$^+$; those from which it was absent were designated Pv1.1$^-$.

Determining the complete sequence of the glucocerebrosidase gene and pseudogene (Horowitz et al. 1989) has now allowed us to precisely localize the base change that causes the *Pvu*II RFLP. We also show that the most common Gaucher disease allele, the substitution at nucleotide 1226, is invariably associated with the Pv1.1$^-$ genotype.

Received August 17, 1989; revision received November 7, 1989.

Address for correspondence and reprints: Ernest Beutler, M.D., Research Institute, Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037.


0002-9297/90/4605-0006$02.00

## Material and Methods

### *Patients Studied*

Blood samples were obtained from patients with Gaucher disease that had been evaluated in the General Clinical Research Center (37 samples) or were provided by their physicians (15 samples). Two samples were of cell lines GM0877 and GM1260 from the NIGMS (Human Genetic Mutant Cell Repository, Camden, NJ).

### *DNA Preparation*

High-molecular-weight DNA was extracted from white blood cells or from cultured cell lines (skin fibroblasts or lymphoblasts) according to standard methods (Maniatis et al. 1982).

### *Localization of the PvuII Polymorphism Site*

The prediction of the location of the *Pvu*II polymorphism site was made by examining the recently published nucleotide sequence of the glucocerebrosidase gene and pseudogene (Horowitz et al. 1989) using the University of Wisconsin Genetics Computer Group's MAPSORT program (Devereux et al. 1984) and comparing the predicted lengths of the gene fragments between the *Pvu*II sites to the ones from those observed in our original investigation, which was carried out with a 1,039-bp cDNA probe (Sorge et al. 1985).

### *DNA Analysis*

The polymerase chain reaction (PCR; Saiki et al. 1988) was utilized to amplify DNA from the glucocerebrosidase gene and pseudogene from each patient. The genotype of each patient was determined by previously described techniques (Zimran et al. 1989). The oligonucleotide primers used for the amplifications and sequencing were based on the published gene and pseudogene sequence (Horowitz et al. 1989) and were synthesized on an automatic DNA synthesizer (ABI, Foster City, CA).

The *Pvu*II polymorphism was studied by amplifying the region between genomic glucocerebrosidase nucleotides 3466 and 4259 with these primers and digesting the samples with *Pvu*II restriction endonuclease (New England Biolabs) according to the directions of the manufacturers. The specific change causing the polymorphism was identified by direct sequencing of the amplified samples. This was accomplished by cutting the amplified band from the gel, eluting the DNA in 250 μl 5mM Tris, 0.1 mM EDTA buffer pH 7.4 overnight at room temperature, reamplifying it with the 5′ primers and the nested $^{32}$P-ATP end-labeled oligonucleotide (Maxam and Gilbert, 1977).

## Results

### *Localization of the Mutation in the Pvl.l$^+$/Pvl.l$^-$ Polymorphism*

The deduced locations of *Pvu*II sites of the glucocerebrosidase gene and pseudogene are shown in figure 1. With the 1,039-bp probe used in our original investigations, fragments >3.075, 1.983, 1.489, >1.271, 0.983, 0.874, and 0.102 kb long would be predicted, and these correspond reasonably well to the reported fragments (Sorge et al. 1985): *5.7* (>3.075), *1.6* (1.489), *1.3* (>1.271), *1.0* (0.983), and *0.9* (.874) kb.

It seemed likely from the data shown in figure 1 that the site of the mutation was such as to create a new *Pvu*II site at nucleotide 3931, where a potential *Pvu*II site with only one mismatch was identified. Cleavage by *Pvu*II at this site converts a .983-kb fragment to .856-kb and .127-kb fragments.

We therefore amplified this region from Pv1.1$^+$/Pv1.1$^+$, Pv1.1$^+$/Pv1.1$^-$, and Pv1.1$^-$/Pv1.1$^-$ DNA with oligonucleotide primers complementary to the antisense strand at nucleotides 3465 through 3484 and to the sense strand at nucleotides 4240 through 4259. These oligonucleotides are complementary to nucleotides

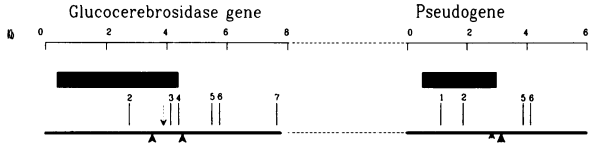


**Figure 1** The *Pvu*II restriction sites deduced from the published sequence (Horowitz et al. 1989), shown as vertical lines. Homologous sites have been assigned the same number. The bars represent the position of the 1,039-bp probe used to detect the polymorphism (Sorge et al. 1985). The dotted arrow between sites 2 and 3 represents the position of the polymorphic *Pvu*II site. The positions of the primers used for PCR are denoted by the arrowheads. A small arrowhead is used to denote the fact that the 5′ primer does not match the pseudogene.

2536 through 2554 of the antisense strand and nucleotides 2990 through 3009 of the sense strand of the pseudogene. Since the region examined was in one of the Alu sequences inserted into the active gene and missing from the pseudogene, the 794-bp fragment amplified represented the active gene. The pseudogene is not amplified because the 5′ primer has three mismatches compared with the pseudogene sequence; furthermore, one of the three base changes is at the 3′ end of the primer. If it had been amplified its product would have been a fragment of only 474 bp, and no such fragment was formed.

Sequencing of the 794-bp amplified fragment showed that the *Pvu*II polymorphic site is a G→A single-base substitution at position 3931. We confirmed this finding by cleavage of the amplified fragment with *Pvu*II, a method also used for rapid determination of the Pv1.1 genotype of the patients (fig. 2).

### *Relationship between Gaucher Disease Mutations and Pvu*II *Genotype*

Table 1 presents the *Pvu*II and the Gaucher disease genotypes of 54 unrelated patients with the disease. All homozygotes for the common Jewish Gaucher mutation at nucleotide 1226 were Pv1.1$^-$/Pv1.1$^-$ homozygotes. In contrast, one homozygote for the mutation at nucleotide 1448 was a Pv1.1$^-$/Pv1.1$^-$ homozygote, and the two others were Pv1.1$^+$/Pv1.1$^+$.

## Discussion

The cloning and sequencing of glucocerebrosidase genes has confirmed the prediction that type I Gaucher disease is heterogeneous on a molecular level (Beutler 1979). Two mutations appear to predominate. The

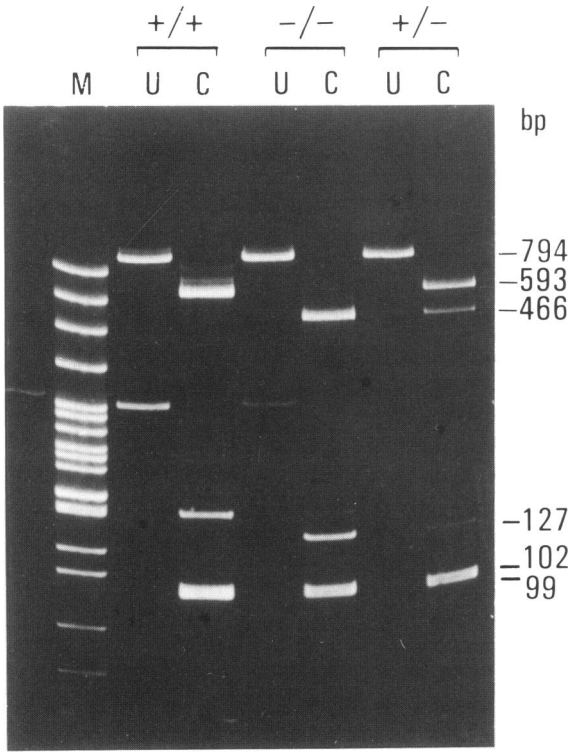


**Figure 2** Electrophoresis of genomic DNA amplified, using the PCR as indicated in the text. Lane M, marker lane containing pBR322 digested with *Msp*I. The three Pv1.1 genotypes are shown at the top. Lanes U, undigested, amplified DNA. Lanes C, DNA digested with the restriction endonuclease *Pvu*II. Electrophoresis was carried out on a 10% polyacrylamide gel at 39 V/cm for approximately 1 h. The bands at approximately 242 and 150 bp, seen clearly in the +/+ lanes and barely in the others, represent contaminating bands presumably unrelated to the glucocerebrosidase locus. *Pvu*II sites 3 and 4 (fig. 1) produce the 99- and 102-bp fragments seen in all the C lanes. These sites are unrelated to the *Pvu*II polymorphism. When the polymorphic restriction site (dotted arrow in fig. 1) is absent the fragment from the 5′ end of the PCR amplification to site 3 (figure 1) produces a 593-bp fragment. When the polymorphic restriction site is present this fragment is cut into fragments of 466 and 127 bp. Note that the "+" genotype, originally so designated because a 1.1-Kb fragment was present, lacks the restriction site which is evident in the "–" allele. Thus, the 794-bp fragment amplified from the "+" gene is cleaved into fragments of 593, 102, and 99 bp. The additional restriction site in the Pv1.1$^-$ allele results in the cleavage of the 593-bp fragment into fragments of 466 and 127 bp. The 127-bp fragment in the right lane is barely visible on the gel, probably because the amount of DNA in the lane is smaller than that in the other lanes, and only one of the two alleles has the polymorphic site.

A→G change at cDNA nucleotide 1226 (Tsuji et al. 1988) results in relatively mild disease and accounts for approximately 78% of the mutant genes in Jewish patients with Gaucher disease (Zimran et al. 1989). The second most common mutation is the T→C change at

**Table 1**

***Pvu*II Polymorphism and Genotype of 54 Gaucher Disease Patients**

| Gaucher Genotype | No. of Patients | *Pvu*II Genotype | | |
|---|---|---|---|---|
| | | + / + | + / − | − / − |
| 1226/1226 . . . . . | 23 | 0 | 0 | 23 |
| 1226/1448 . . . . . | 6 | 0 | 3 | 3 |
| 1226/? . . . . . | 16 | 0 | 13 | 3 |
| 1226/Xovr . . . . . | 1 | 0 | 0 | 1 |
| 1448/1448 . . . . . | 3 | 2 | 0 | 1[a] |
| 1448/1361 . . . . . | 1 | 0 | 0 | 1[b] |
| 1448/? . . . . . | 1 | 0 | 0 | 1 |
| ?/? . . . . . | 3 | 3 | 0 | 0 |

[a] Cell line GM0877.
[b] Cell line GM1260.

nucleotide 1448 (Tsuji et al. 1987), an abnormality that is associated with more severe disease. The *Pvu*II polymorphism of the glucocerebrosidase gene complex is apparently an ancient one: it is represented in all races at similar (Sorge et al. 1985) although not identical (Masuno et al. 1989) levels. We have now localized the site of this RFLP to nucleotide 3931 in the glucocerebrosidase gene. A G→A single-base substitution creates a *Pvu*II site at this location, giving rise to the Pv1.1$^-$ genotype (since the presence of the site cleaves the 1.1 fragment originally detected on Southern blots). Precise analysis of the mutation shows that the conversion is not actually one from a 1.1 to a 1.0 fragment but rther one from a 983-bp fragment to a 856-bp fragment. The 127-kb fragment that is formed is not detected on Southern blots. Although the genotype might more accurately be described as Pv1$^-$ and Pv1$^+$, we will retain the originally proposed nomenclature to prevent confusion.

It seems likely from our data that the 1448 mutation arose more than once, since two of the homozygotes for this mutation were homozygous for Pv1.1$^+$ and the other was homozygous for Pv1.1$^-$. Unless this is a very ancient mutation that has persisted in humans, allowing linkage equilibrium to occur, there must have been at least two separate mutational events that gave rise to this base-pair change. It may be significant from this point of view that the 1448 mutation produces a sequence in the active gene that is identical to the sequence normally present in the pseudogene. Thus, this sequence can be produced in the glucocerebrosidase gene not only by a random point mutational event, but also by "gene conversion" (Baltimore 1981) or by an

unequal crossing-over with the pseudogene, occurring at a point upstream from cDNA nucleotide 1448. Such crossing-over has, indeed, been documented by sequencing cDNA and analyzing the genomic DNA from a patient with Gaucher disease (Zimran et al. 1990).

In contrast to the situation with the 1448 mutation, 46 genes at risk in the 23 homozygotes for the 1226 mutations studied all were Pv1.1$^-$. The frequency of the Pv1.1$^-$ allele in the general population, including the Jewish population, has been found to be 0.65. Thus, at equilibrium the probability that 23 of 23 individuals would be homozygous for this allele is only $2.48 \times 10^{-9}$. Moreover, at least one glucocerebrosidase gene from all of the 23 patients who were heterozygous for this mutation had the Pv1.1$^-$ genotype. Thus, our findings are most consistent with the hypothesis that the 1226 mutation arose only once, and then in a Pv1.1$^-$ allele.

It is especially notable that 13 of 16 patients who were of the 1226/? genotype were Pv1.1$^+$/Pv1.1$^-$. Since the Pv1.1$^-$ allele can reliably be assigned to the 1226 mutation it follows that the unknown mutation, designated "?" occurred in a Pv1.1$^+$ allele. Since the Pv1.1$^+$ allele has a frequency of only 0.35 this is a most unusual finding, expected by binomial probability with a $P$ of only $2 \times 10^{-4}$. Even more impressive is the fact that of the 13 Pv1.1$^+$ ? alleles, 12 were derived from a Jewish parent; of the 3 putative Pv1.1$^-$ ? alleles, only one was derived from a Jewish parent. Such a finding would be expected in the event that there is yet a third relatively common Jewish Gaucher disease allele which happens to be linked to the Pv1.1$^+$ genotype.

## Acknowledgments

This work was supported by NIH grants DK36639 and RR00833, March of Dimes grant 1-955, and the Sam Stein and Rose Stein Charitable Trust. This is publication 5984 MEM from the Research Institute of Scripps Clinic.

## References

Baltimore D (1981) Gene conversion: some implications for immunoglobulin genes. Cell 24:592–594

Beutler E (1979) Gaucher disease. In: Goodman RM, Motulsky AG (eds) Genetic diseases among Ashkenazi Jews. Raven, New York, pp 157–169

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Horowitz M, Wilder S, Horowitz Z, Reiner O, Gelbart T, Beutler E (1989) The human glucocerebrosidase gene and pseudogene: structure and evolution. Genomics 4:87–96

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Masuno M, Orii T, Sukegawa K, Taga T (1989) Restriction fragment length polymorphism analysis in healthy Japanese individuals and Japanese families with Gaucher disease. Acta Paediatr Jpn 31:158–162

Maxam AM, Gilbert W (1977) A new method for sequencing DNA. Proc Natl Acad Sci USA 74:560–564

Saiki RK, Gelfand, DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–494

Sorge J, Gelbart T, West C, Westwood B, Beutler E (1985) Heterogeneity in type I Gaucher disease demonstrated by restriction mapping of the gene. Proc Natl Acad Sci USA 82:5442–5445

Tsuji S, Choudary PV, Martin BM, Stubblefield BK, Mayor JA, Barranger JA, Ginns EI (1987) A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease. N Engl J Med 316:570–621

Tsuji S, Martin BM, Barranger JA, Stubblefield BK, LaMarca ME, Ginns EI (1988) Genetic heterogeneity in type 1 Gaucher disease: multiple genotypes in Ashkenazic and non-Ashkenazic individuals. Proc Natl Acad Sci USA 85: 2349–2352, 5708

Zimran A, Sorge J, Gross E, Kubitz M, West C, Beutler E (1989) Prediction of the clinical course of Gaucher disease by identification of mutations at the DNA level. Lancet 2:349–352

——— (1990) A glucocerebrosidase fusion gene in Gaucher disease: implications for the molecular anatomy, pathogenesis and diagnosis of this disorder. J Clin Invest 85: 219–222