Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 5442–5445, August 1985
Genetics

Heterogeneity in type I Gaucher disease demonstrated by restriction mapping of the gene

(cloning/polymorphism/selection/population genetics)

J. SORGE, T. GELBART, C. WEST, B. WESTWOOD, AND E. BEUTLER

Department of Basic and Clinical Research, Scripps Clinic and Research Foundation, 10666 North Torrey Pines Road, La Jolla, CA 92037

Contributed by E. Beutler, April 10, 1985

ABSTRACT **A cloned fragment of human glucocerebrosidase cDNA has been used as a probe to study restriction polymorphisms in the region of the gene for Gaucher disease. Variability in the size of fragments produced by digestion with the restriction endonucleases *Pvu* II and *Kpn* I was discovered. The *Pvu* II polymorphism was found to be a very prevalent one with a gene frequency of 0.65 for the Pv1.1$^-$ allele and 0.35 for the Pv1.1$^+$ allele. Similar frequencies were encountered among diverse ethnic groups. Five of eight Jewish patients with Gaucher disease were found to be heterozygous for the *Pvu* II restriction polymorphism. One non-Jewish patient with type I Gaucher disease was heterozygous for the *Kpn* I variant. The existence of Gaucher disease genes in association with either allele of the ancient *Pvu* II polymorphism clearly indicates that, even within the Jewish population, the Gaucher disease mutation has occurred independently more than once. Presumably, different mutations have also occurred in the non-Jewish population.**

Gaucher disease is an autosomal recessive disorder characterized by an accumulation of the sphingoglycolipid glucocerebroside. It is due to a deficiency of the enzyme glucocerebrosidase (1, 2). The most common form of this disorder, designated type I, is prevalent in those of Eastern European Jewish ethnic origin. The gene frequency in this group may be as high as 0.02–0.04 (3, 4). Type I disease also occurs sporadically in non-Jewish populations. More severe forms of the disease, designated types II and III, are associated with central nervous system involvement and do not seem to have any predilection for those of Jewish ancestry. It has been proposed (3) that the existence of Gaucher disease at polymorphic frequencies in the Eastern European Jewish population might be due to a single gene mutation in "an extended kindred" (3), as it were. However, we have previously suggested, on the basis of the severity of the defect as found in different family groupings (5, 6) and of immunologic examination of the residual enzyme (7), that even within this population Gaucher disease was genetically heterogeneous.

We have now succeeded in cloning human glucocerebrosidase cDNA. Using a large cloned fragment of this gene as a probe, we have discovered a restriction polymorphism that clearly establishes that even within the Eastern European Jewish population, gene defects for Gaucher disease have arisen independently more than once.

MATERIALS AND METHODS

The cloning of glucocerebrosidase cDNA in λ phage gt11 and its complete sequence will be described in detail in another paper (8). Sequencing of the DNA probe was accomplished by the method of Maxam and Gilbert (9). The probe used in the present studies was a 1039-base-pair (bp) cDNA fragment (our clone G5A-1Y). It was cut from a pBR322 plasmid and purified by preparative agarose gel electrophoresis. A radioactive probe was prepared by reverse transcribing the denatured DNA in the presence of the following: random calf thymus primers (1 mg/ml), 50 mM Tris·HCl (pH 8.3), 20 mM KCl, 10 mM $MgCl_2$, 5 mM dithiothreitol, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, and 10 μCi of [α-^{32}P]dTTP (600 Ci/mM; 1 Ci = 37 GBq). The probe was isolated by Sephadex G-50 filtration.

Restriction endonucleases were obtained from the commercial sources listed in Table 1. DNA was isolated from peripheral blood leukocytes or from cultured skin fibroblasts by standard techniques (10). Ten micrograms of the purified DNA was incubated with 150–200 units of restriction endonuclease for 2.5 hr at 37°C. After electrophoresis in 0.8% agarose for 16 hr at 1.7 V/cm, the DNA was transferred to nitrocellulose filters, developed by addition of ≈1 μCi of probe per 100 cm^2 of filter, and visualized using XAR x-ray film (11). A *Hin*dIII digest of λ DNA provided molecular weight markers.

Whenever variation in the restriction pattern was seen, digestion of an additional sample of DNA was carried out with twice the amount of enzyme, and samples were taken at time intervals to make certain that DNA digestion by the restriction endonuclease had been complete.

RESULTS

The DNA sequence of the probe used in this study is depicted in Fig. 1. It represents the 5′ end of the glucocerebrosidase cDNA.

The results of restriction endonuclease mapping of genomic DNA from unrelated individuals with 20 different restriction enzymes is shown in Table 1. Variability in the pattern was observed only with two restriction endonucleases—namely, *Pvu* II and *Kpn* I. Further studies using these two restriction endonucleases were therefore performed. The restriction patterns produced with these enzymes in Gaucher disease patients and in normal subjects are presented in Figs. 2 and 3.

In the case of *Pvu* II, the examined subjects could be fairly evenly divided into those whose digested DNA contained a 1.1-kilobase (kb) fragment and those whose DNA did not. The gene that produced the 1.1-kb fragment was designated Pv1.1$^+$ and the gene that lacked the fragment was designated Pv1.1$^-$. Since even one Pv1.1$^+$ gene would produce a 1.1-kb fragment, a genotype of Pv1.1$^-$/Pv1.1$^-$ was assigned to those who lacked the 1.1-kb fragment and a genotype of Pv1.1$^+$/Pv1.1$^-$ was assigned to those in whom the fragment was present. The validity of this interpretation was confirmed by the discovery of a family in which both parents had the

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. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); kb, kilobase(s).

Table 1. Restriction endonuclease mapping of the glucocerebrosidase locus by using 20 different enzymes

Endonuclease	Source	Number of unrelated subjects examined				Fragment size, kb
		JG	NJG	JN	NJN	
Apa I	1,2	6	4	3	1	10.1 5.0 4.0 3.3 1.6
*Bam*HI	1	5	3	4	4	4.2 2.8
Bgl I	2	1		2		1.5
Bgl II	1	7	2	3	4	12.5 9.5 8.1
*Bst*EII	1,2	5	2	3	1	12.0 9.3 7.8
Cfo I*	2	5	1	2		21.0 18.5 16.0 13.5 9.2 8.6 7.9 5.8 5.3
Cla I	1	3		2	1	22.0 19.0
Dra I	3	3			2	10.5 2.0
*Eco*RI	2,3	7	4	13	8	13.5 12.5
*Eco*RV	1,2	6	3	10	2	17.0
*Hind*III	1,2	5	2	3	4	8.9 7.5 6.5
Hpa I*	1				4	22.0 20.0 13.5 10.0
Kpn I	1,2,3	7	2	13	18	2.3 1.8 1.5 1.4 (1.3) (1.2)
Mbo II	1	5	1	2		1.5 1.4
Msp I*	1				4	1.2 1.0 0.8
Pst I	1,2	7	3	3	1	2.3 2.0 1.8 1.5 1.2 1.0
Pvu II	1,2,3	7	3	13	45	5.7 1.6 1.3 (1.1) 1.0 0.9
Sac I	1,2	6	2	2		6.4 4.7
Xba I	1	3		2	1	4.9 4.6 4.3
Xho I	1,2	3		3	1	24.0 19.5

JG, Jewish Gaucher; NJG, non-Jewish Gaucher; JN, Jewish normal; NJN, non-Jewish normal. Fragments present in only some subjects are shown in parentheses. Sources: 1, New England Biolabs; 2, Boehringer Mannheim; 3, Pharmacia.
*Methylation sensitive.

1.1-kb fragment but the children did not. Eight unrelated individuals with another *Pvu* II restriction pattern were encountered. A very dark 1.1-kb band was present on the autoradiograph and instead of the normally present 1.0-kb band a faint slightly larger band of perhaps 1.02 kb was visualized. These subjects presumably were of the $Pv1.1^+/Pv1.1^+$ genotype. Table 2 shows the population distribution of the *Pvu* II restriction pattern.

DNA from eight Jewish type I Gaucher disease patients was examined. Five were found to be of the $Pv1.1^+/Pv1.1^-$ genotype and three were of the $Pv1.1^-/Pv1.1^-$ type. Two non-Jewish Gaucher patients were both $Pv1.1^+/Pv1.1^-$. One of these was heterozygous for the *Kpn* I polymorphism. Two type II Gaucher disease fibroblast DNA samples were of the $Pv1.1^-/Pv1.1^-$ genotype; both had the normal *Kpn* I restriction pattern.

DISCUSSION

We have cloned human glucocerebrosidase cDNA from a λgt11 expression library. Its authenticity has been established by demonstrating (*i*) that it has homology with the amino acid sequence of portions of glucocerebrosidase; (*ii*) that it hybridizes with chromosome 1 DNA (known to contain

```
   1   GG TTC CTG CAT CCT TGT TTT TGT TTA GTG GAT CCT CTA TCC TTC    44
  45  AGA GAC TCT GGA ACC CCT GTG GTC TTC TCT TCA TCT AAT GAC CCT    89
  90  GAG GGG ATG GAG TTT TCA AGT CCT TCC AGA GAG GAA TGT CCC AAG   134
 135  CCT TTG AGT AGG GTA AGC ATC ATG GCT GGC AGC CTC ACA GGT TTG   179
 180  CTT CTA CTT CAG GCA GTG TCG TGG GCA TCA GGT GCC CGC CCC TGC   224
 225  ATC CCT AAA AGC TTC GGC TAC AGC TCG GTG GTG TGT GTC TGC AAT   269
 270  GCC ACA TAC TGT GAC TCC TTT GAC CCC CCG ACC TTT CCT GCC CTT   314
 315  GGT ACC TTC AGC CGC TAT GAG AGT ACA CGC AGT GGG CGA CGG ATG   359
 360  GAG CTG AGT ATG GGG CCC ATC CAG GCT AAT CAC ACG GGC ACA GGC   404
 405  CTG CTA CTG ACC CTG CAG CCA GAA CAG AAG TTC CAG AAA GTG AAG   449
 450  GGA TTT GGA GGG GCC ATG ACA GAT GCT GCT GCT CTC AAC ATC CTT   494
 495  GCC CTG TCA CCC CCT GCC CAA AAT TTG CTA CTT AAA TCG TAC TTC   539
 540  TCT GAA GAA GGA ATC GGA TAT AAC ATC ATC CGG GTA CCC ATG GCC   584
 585  AGC TGT GAC TTC TCC ATC CGC ACC TAC ACC TAT GCA GAC ACC CCT   629
 630  GAT TTC CAG TTG CAC AAC TTC AGC CTC CCA GAG GAA GAT ACC AAG   674
 675  CTC AAG ATA CCC CTG ATT CAC CGA GCC CTG CAG TTG GCC CAG CGT   719
 720  CCC GTT TCA CTC CTT GCC AGC CCC TGG ACA TCA CCC ACT TGG CTC   764
 765  AAG ACC AAT GGA GCG GTG AAT GGG AAG GGG TCA CTC AAG GGA CAG   809
 810  CCC GGA GAC ATC TAC CAC CAG ACC TGG GCC AGA TAC TTT GTG AAG   854
 855  TTC CTG GAT GCC TAT GCT GAG CAC AAG TTA CAG TTC TGG GCA GTG   899
 900  ACA GCT GAA AAT GAG CCT TCT GCT GGG CTG TTG AGT GGA TAC CCC   944
 945  TTC CAG TGC CTG GGC TTC ACC CCT GAA CAT CAG CGA GAC TTC ATT   989
 990  GCC CGT GAC CTA GGT CCT ACC CTC GCC AAC AGT ACT CAC CAC AAT  1034
1035  GTC CG
```

FIG. 1. DNA sequence of probe used in these studies. The 5′ end is at the top left. The initiator ATG is underlined.

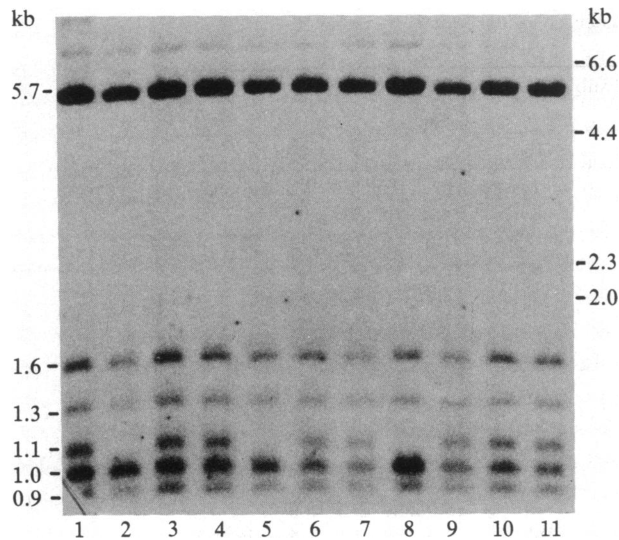

FIG. 2. *Pvu* II restriction endonuclease digests of DNA from 11 unrelated persons, developed with the probe shown in Fig. 1. Size markers are on the right and sizes of the bands are on the left. Lanes: 1 and 2, Jewish controls; 3 and 4, parents of patient with Gaucher disease; 5–11, patients with type I Gaucher disease; 7 and 9, non-Jewish subjects; 5, 6, 8, 10, and 11, Jewish subjects. The *Pvu* II polymorphism is characterized by the 1.1-kb band, absent from lanes 2, 5, and 8 (genotype, $Pv1.1^-/Pv1.1^-$). Samples in lanes 1, 3, 4, 6, 7, 9, 10, and 11 are $Pv1.1^+/Pv1.1^-$.

the glucocerebrosidase gene); and (*iii*) that protein expressed from the cDNA clone in bacteria reacts with an affinity-purified highly specific anti-glucocerebrosidase antiserum (8). Family studies in patients with Gaucher disease also support the validity of the clone. Cloning of glucocerebrosidase has been reported previously (12), but no DNA sequence data have been disclosed.

DNA from normal controls and from patients with Gaucher disease and their family members was screened for restriction polymorphisms with 20 different restriction endonucleases. Two genetic variants were found, a common variant with *Pvu* II and a less common one with *Kpn* I.

The appearance of a new 1.1-kb band in the *Pvu* II digest

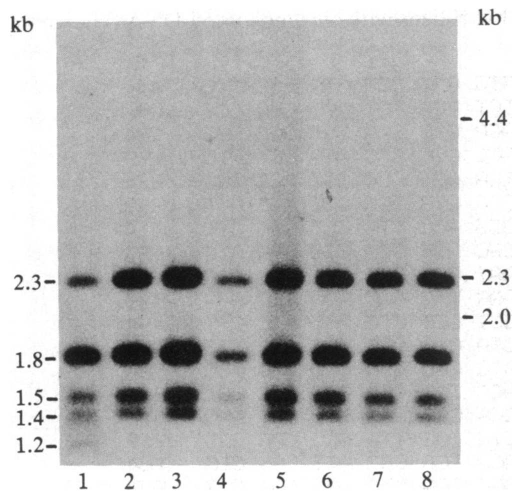

FIG. 3. *Kpn* I restriction endonuclease digests of DNA from eight unrelated persons, developed with the probe shown in Fig. 1. Size markers are on the right and sizes of the bands are on the left. Lanes: 1–3, type I Gaucher disease patients; 4–8, controls; 2–8, Jewish subject; 1, non-Jewish subjects. The extra band at 1.2 kb distinguishes the non-Jewish Gaucher disease patient from all other samples that have been digested with *Kpn* I.

Table 2. *Pvu* II polymorphism in different populations

	$Pv1.1^-/Pv1.1^+$	$Pv1.1^-/Pv1.1^-$	$Pv1.1^+/Pv1.1^+$	Gene frequency $Pv1.1^-$	Gene frequency $Pv1.1^+$
White, non-Jewish	4	3	1	0.625	0.375
White, Jewish	9	9	2	0.675	0.325
Black	16	13	3	0.656	0.343
Oriental	2	2	2	0.500	0.500
Spanish	1	1	0	0.750	0.250
Total	32	28	8	0.647	0.353
Expected*	31.06	28.47	8.47		

*Based on Hardy–Weinberg equilibrium.

prepared from $Pv1.1^+/Pv1.1^+$ homozygotes was accompanied by loss of the normally present 1.0-kb band. It is reasonable to assume that the $Pv1.1^+$ mutation is due to the loss of a *Pvu* II restriction site ≈100 bp from the end of a 1.1-kb segment of DNA between two other *Pvu* II sites. The 100-bp fragment that is formed is not bound to the nitrocellulose filter. The incidence of the *Pvu* II polymorphism appears to be essentially identical in all racial groups investigated. Collectively, the frequency of the $Pv1.1^-$ gene was 0.65 and that of the $Pv1.1^+$ gene was 0.35. The three genotypes fit closely the values predicted from the Hardy–Weinberg equation. It is apparent, therefore, that the *Pvu* II polymorphism is an ancient one. The Gaucher disease mutations, in contrast, have probably arisen relatively recently. Since the homozygous state for this gene produces a markedly adverse effect upon fitness, the high incidence in the Eastern European Jewish population might imply that it confers an advantage to heterozygotes. However, no advantage is known and a "founder effect" has been considered another possible cause for the high gene frequency.

The present studies unequivocally indicate that selection is the cause of the high frequency of the type I Gaucher disease gene in the Jewish population. In patients with this recessively inherited disorder, both alleles at the glucocerebrosidase locus have a Gaucher disease mutation. Since five Jewish patients with Gaucher disease were heterozygous for the Pv1.1 polymorphism, Gaucher disease genes are commonly linked to both $Pv1.1^+$ and $Pv1.1^-$ genes; the fact that the Gaucher mutation is found in association with both alleles of the ancient *Pvu* II polymorphism indicates that it has arisen more than once in the Jewish population. To be repeatedly amplified in one population, mutations producing glucocerebrosidase deficiency must have provided a selective advantage.

Not surprisingly, further heterogeneity was encountered in non-Jewish Gaucher disease patients. Here, variability at a *Kpn* I site was discovered in a non-Jewish child with Gaucher disease. Thus, the Gaucher disease gene that arose in conjunction with the *Kpn* I site presumably was another independent genetic event. In this case, however, the relative rarity of the *Kpn* I mutation makes it remotely possible that the *Kpn* I mutation was a recent genetic event engrafted upon a preexisting Gaucher disease mutation. It is even possible that one of the Gaucher disease mutations in this patient creates or abolishes a *Kpn* I site.

Since we do not yet have knowledge of the precise base-pair substitution(s) of the $Pv1.1^+$ and the $Pv1.1^-$ linked Gaucher disease mutations, we are not certain whether these independently arising mutational events were identical. In the case of the sickle cell mutation, an identical change in the base-pair sequence appears to have arisen independently on several occasions in different populations (13). However, the sickling phenomenon is subject to a great many molecular constraints, while enzyme deficiencies may be produced by

a myriad of base-pair changes. We consider it unlikely that the mutation found in the various Gaucher disease mutations will be the same. Moreover, studies of cross-reacting immunologic materials in Gaucher disease (7) and family studies (6) have already suggested that Gaucher disease, even within the Jewish population, is heterogeneous.

The cloned glucocerebrosidase gene may, in the future, prove to be useful in developing gene replacement therapy for this serious genetic disease (14). At present, it has enabled us to gain insight into the population genetics of this disorder. In addition, the existence of restriction polymorphisms will make the probe a useful tool, in selected cases, for heterozygote detection and prenatal diagnosis.

This work was supported in part by Grant CA 36448 from the National Institutes of Health and Grant RR 00833 from the National Institutes of Health, Division of Research Resources. This manuscript is publication 3903 BCR from the Research Institute of Scripps Clinic.

1. Desnick, R. J., Gatt, S. & Grabowski, G. A., eds. (1982) *Gaucher Disease: A Century of Delineation and Research* (Liss, New York).
2. Brady, R. O. & Barranger, J. A. (1983) in *The Metabolic Basis of Inherited Disease*, eds. Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. L. & Brown, M. S. (McGraw-Hill, New York), pp. 842–856.
3. Motulsky, A. G. (1979) in *Genetic Diseases Among Ashkenazi Jews*, eds. Goodman, R. M. & Motulsky, A. G. (Raven, New York), pp. 167–168.
4. Kolodny, E. H., Ullman, M. D., Mankin, H. J., Raghavan, S. S., Topol, J. & Sullivan, J. L. (1982) in *Gaucher Disease: A Century of Delineation and Research*, eds. Desnick, R. J., Gatt, S. & Grabowski, G. A. (Liss, New York), pp. 33–65.
5. Beutler, E. (1979) in *Genetic Diseases Among Ashkenazi Jews*, eds. Goodman, R. M. & Motulsky, A. G. (Raven, New York), pp. 157–167.
6. Beutler, E. (1977) *J. Am. Med. Assoc.* **237**, 2529.
7. Beutler, E., Kuhl, W. & Sorge, J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6506–6510.
8. Sorge, J., West, C., Westwood, B. & Beutler, E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, in press.
9. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
10. Beutler, E. & Kuhl, W. (1983) in *Clinical Laboratory Assays: New Technology and Future Directions*, eds. Nakamura, R. K., Dito, W. R. & Tucker, E. S. (Masson, New York), pp. 231–247.
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
12. Ginns, E. I., Choudary, P. V., Martin, B. M., Winfield, S., Stubblefield, B., Mayor, J., Merkle-Lehman, D., Murray, G. J., Bowers, L. A. & Barranger, J. A. (1984) *Biochem. Biophys. Res. Commun.* **123**, 574–580.
13. Kan, Y. W. & Dozy, A. M. (1980) *Science* **209**, 388–391.
14. Beutler, E. (1981) *Trends Biochem. Sci.* **6**, 95–97.