

Identification of the second common Jewish Gaucher disease mutation makes possible population-based screening for the heterozygous state

E. BEUTLER*, T. GELBART, W. KUHL, J. SORGE, AND C. WEST

The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037

Contributed by E. Beutler, August 23, 1991

ABSTRACT Gaucher disease is an autosomal recessive glycolipid storage disease characterized by a deficiency of glucocerebrosidase. The disease is most common in persons of Ashkenazi Jewish ancestry and the most common mutation, accounting for about 75% of the mutant alleles in this population, is known to be an A → G substitution at cDNA nucleotide (nt) 1226. Screening for this disease has not been possible because nearly 25% of the mutant alleles had not been identified, but linkage analysis led to the suggestion that most of these could be accounted for by a single mutation. We now report the discovery of this mutation. The insertion of a single nucleotide, a second guanine at cDNA nt 84 (the 84GG mutation), has been detected in the 5' coding region of the glucocerebrosidase gene. The amount of mRNA produced is shown to be normal but since the frameshift produces early termination, no translation product is seen. This finding is consistent with the virtual absence of antigen found in patients carrying this mutation. The 84GG mutation accounts for most of the previously unidentified Gaucher disease mutations in Jewish patients. The common Jewish mutation at nt 1226, the 84GG mutation, and the less-common mutation at nt 1448 accounted for 95% of all of the Gaucher disease-producing alleles in 71 Jewish patients. This now makes it possible to screen for heterozygotes on a DNA level with a relatively low risk of missing couples at risk for producing infants with Gaucher disease.

Gaucher disease is the most common glycolipid storage disease. It is due to a deficiency of the enzyme glucocerebrosidase. The disease is most prevalent in the Jewish population with a heterozygote frequency that we have estimated to approach 9% (1). In Jewish patients with clinically significant Gaucher disease, about 75% of the disease-causing alleles contain a characteristic A → G mutation at cDNA nucleotide (nt) 1226 (designated the 1226G mutation) (2, 3). The same mutation is also common in the non-Jewish population, where it is found to account for approximately 25% of the disease-producing alleles. A second, much less common mutation is at cDNA nt 1448, a T → C mutation. It can occur either as a point mutation (4, 5) or as the result of a cross-over between the glucocerebrosidase gene and pseudogene, producing a nonfunctioning fusion gene (6). The 1448C mutation accounts for only about 2% of Jewish Gaucher disease-producing alleles and for about 40% of the alleles in non-Jewish patients. Thus, in both Jewish and non-Jewish patients many of the Gaucher disease alleles have remained unidentified and have been designated “?”.

Although part of this residue of unidentified mutations undoubtedly represents multiple sporadic mutations and, indeed, some such mutations unique to one or two families have been identified (7–9), linkage data suggested that there was an additional common mutation that had eluded detection. A restriction fragment length polymorphism with the enzyme *Pvu* II exists at

genomic nt 3931 in intron 6 of the glucocerebrosidase gene (10, 11). This restriction polymorphism can be used to classify glucocerebrosidase genes as $Pv1.1^+$ and $Pv1.1^-$. The latter genotype is the most common, accounting for some 65% of glucocerebrosidase genes in Western populations (11). The 1226G mutation is invariably found in the context of the $Pv1.1^-$ genotype (10). Many Jewish and non-Jewish Gaucher disease patients are heterozygous for the 1226G mutation and a “?” mutation. Since the 1226G mutation is always linked to $Pv1.1^-$, one would predict that sporadic mutations in the other alleles would have the same distribution as the $Pv1.1^+$ and $Pv1.1^-$ polymorphism does in the population as a whole. Thus, 65% of patients with the 1226G/? genotype would be $Pv1.1^-/Pv1.1^-$ and 35% would be $Pv1.1^-/Pv1.1^+$. In point of fact, this ratio is reversed (10) suggesting the existence of a common Gaucher disease mutation in linkage disequilibrium with the *Pvu* II restriction polymorphism. Accordingly, we have made a special effort to find the putative second common Gaucher disease mutation in patients with the 1226G/? $Pv1.1^-/Pv1.1^+$ genotype. We have now identified this mutation as an insertion of an extra guanine at cDNA nt 84, producing a frame shift in the leader sequence and consequent absence of the enzyme protein. The results of the analysis of the DNA of 71 Jewish and 38 non-Jewish patients for this and 10 other mutations are presented.

MATERIALS AND METHODS

cDNA Sequencing After the PCR. Poly(A)⁺ mRNA was isolated from cultured lymphoblasts or fibroblasts obtained from two unrelated patients with the 1226G/? $Pv1.1^-/Pv1.1^+$ genotype. First-strand cDNA was made using oligo(dT) and reverse transcriptase as described (12). DNA extending from cDNA nt -46 to cDNA nt 1722 was amplified as a single segment by 35 PCR cycles using 250 ng of a sense primer (nt -46 to nt -27) and antisense primer (nt 1703–1722). The PCR product was purified with phenol/chloroform, ethanol-precipitated, and redissolved in water. Single-stranded DNA was then produced by 35 cycles of an unbalanced PCR with 5–10% of the first PCR product serving as template using 300 ng of a sense primer (nt -25 to nt -6) or antisense primer (nt 1621–1640). Sequencing of the single-stranded PCR-generated DNA was accomplished with cDNA primers spaced about 200 nt apart along the cDNA.

In addition, amplified cDNA was cloned into Bluescript SK⁺ (Stratagene) after the PCR using a sense primer with an *Eco*RI site and an antisense primer with a *Xba* I site. This construct was used to prepare RNA for translational assays and was also sequenced.

Genomic Sequences. High molecular weight DNA was isolated from a patient with the 1226G/? $Pv1.1^-/Pv1.1^+$ genotype, digested to completion with *Eco*RI, and cloned into λ Dash. Glucocerebrosidase clones were identified using a 2.23-kilobase cDNA probe (13) and were distinguished from

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.

*To whom reprint requests should be addressed.

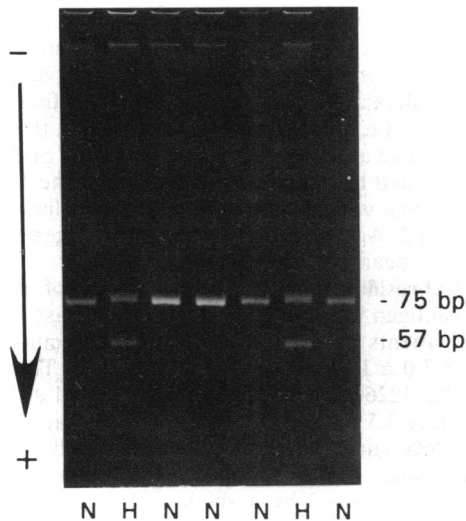


FIG. 1. Detection of the 84GG mutation by a mismatched PCR (see text). Amplified DNA with the 84GG mutation was cleaved by *Bsa*BI, whereas DNA with the normal sequence is not cleaved. The results of gel electrophoresis of amplified and digested DNA from two heterozygotes (lanes H) and five normals (lanes N) are shown.

the pseudogene clones by the fact that they did not react to an oligonucleotide probe in the 55-base-pair (bp) coding region deletion in exon 9 of the pseudogene. We used oligonucleotide-specific probes to identify clones with and without the 1226G mutation. The clones were digested with *Xba* I and subcloned into Bluescript. Double-stranded sequencing was carried out by the chain-termination technique using appropriate primers based on the known sequence (14). Ambiguities were resolved by sequencing the opposite strand.

Quantitation of mRNA. Since one of the two alleles of the patient contained the 1226G mutation and the other did not, it was possible to determine the relative amounts of mRNA produced from each of the alleles of patients heterozygous for the 1226G mutation (15).

In Vitro Translation. The cDNA clones that had been prepared in Bluescript SK⁺ (see above) were identified with respect to whether they contained the 1226G or the 84GG mutations. They were then linearized with *Xba* I. One microgram of each linearized clone was added to the transcription system we have described (16), except that the concentration of nucleotides was increased to 0.8 mM. One-fifth of each RNA transcript preparation was translated in a rabbit reticulocyte lysate system (16) and subjected to Laemmli SDS/PAGE on a 10% polyacrylamide gel.

Detection of Mutations in DNA from Gaucher Disease Patients. The 84GG mutation does not create or destroy a restriction endonuclease site. Therefore, we used a mismatched PCR (17), wherein a primer is designed to contain a mismatch creating a restriction site in the mutant but not in

the normal allele. We used, as a 5' primer, 5'-GAATGTC-CCAAGCCTTGA [nt 979-997 of the published genomic sequence (14)]. This primer mismatches the glucocerebrosidase pseudogene at its 17th and 19th position. The 3' primer was 3'-CGTAGATGAAGTCCGTCAC [nt 1035-1053 of the genomic sequence (14)]. The 3rd nucleotide has been changed from an adenine to a thymine to create a *Bsa*BI site if two guanines are present at genomic nt 1035 (cDNA nt 84). A PCR was performed in 0.5× PCR buffer (18) containing all four deoxynucleotides (each at 0.5 mM), 5% (vol/vol) dimethyl sulfoxide, 125 ng of each primer, 0.75 unit of *Taq* polymerase (Boehringer Mannheim), and 0.5 μg of genomic DNA in a 50-μl system. The PCR was performed for 29 cycles consisting of 30 s at 92°C, 30 s at 59°C, and 30 s at 72°C.

The PCR products (15 μl) were incubated in a 50-μl system with 1× New England Biolabs buffer 2 and 20 units of *Bsa*BI at 60°C for 1.5 h. After addition of 2.5 vol of ethanol and chilling, the precipitate was dried and redissolved in 15 μl of loading dye buffer (19). Electrophoresis was carried out on a 12% polyacrylamide gel in 0.5× TBE (1× TBE = 90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3) and the gel was stained with ethidium bromide. The results obtained using this technique are illustrated in Fig. 1.

Examination for the other known mutations was carried out as summarized in Table 1. The method used for the detection of the 1226G mutation (15) and the 1448C mutation (3) has been described. Samples that had the 1448C mutation were further examined for a cross-over (XOVR) gene by Southern blot analysis after digestion with *Ssp* I (6).

Quantitation of Glucocerebrosidase Antigen. Glucocerebrosidase antigen was quantitated with monoclonal antibodies as described (24).

Gaucher Disease Patients. The diagnosis of Gaucher disease was well-established in all patients, either by histopathologic study of the marrow or demonstration of diagnostically lowered levels of acid β-glucosidase in peripheral blood cells (25). Classification by ethnic origin was according to the family history provided by each patient. Four half-Jewish patients were excluded from the analysis. Otherwise all Gaucher disease patients referred to us from which DNA was available for study were included. The clinical severity score was calculated as described (3) except that the age of diagnosis on first symptoms was not incorporated into the calculation.

RESULTS

cDNA Sequences. The sequences determined directly on PCR-amplified DNA from two patients with Gaucher disease with the 1226G/? Pv1.1⁻/Pv1.1⁺ genotype showed two abnormalities. (i) Two bands were found in the sequencing gels at nt 1226 where both the normal adenine and abnormal guanine were found representing the heterozygous state for the 1226G mutation. (ii) An insertion of a guanine was present after nt 84. Two bands were apparent at each position in the sequence gel after the inserted guanine representing the

Table 1. Methods used for detection of known Gaucher disease mutations

cDNA nt	Ref.	Genomic nt	Substitution	Genomic sense primer, nt	Genomic antisense primer, nt	Restriction enzyme	Normal pattern, bp	Mutant pattern, bp
476	20	3060	G → A	2978-2997	3121-3140	<i>Bst</i> NI	163	82, 81
764	9	4113	T → A	3465-3484	4240-4259	<i>Kpn</i> I	795	648, 147
1090	8	5306	G → A	5156-5175	5490-5510	<i>Bsu</i> 36I	355	205, 150
1141	8	5357	T → G	5156-5175	5490-5510	<i>Stu</i> I	199, 156	355
1342	21	5957	G → C	5878-5897	6468-6487	<i>Sty</i> I	342, 156, 75, 37	342, 156, 112
1361	22	5976	C → G	5878-5897	6031-6050	<i>Hha</i> I	173	98, 75
1504	23	6489	C → T	6463-6482	6516-6535	<i>Msp</i> I	47, 26	73
1604	*	6683	G → A	6632-6651	6782-6801	<i>Hph</i> I	170	128, 42

*E.B. and W.K., unpublished data.

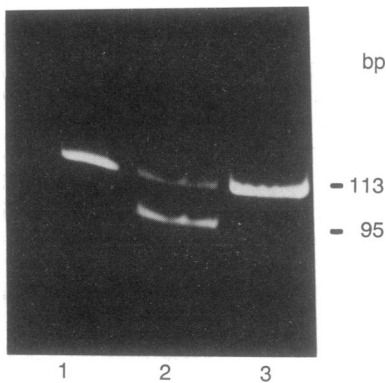


Fig. 2. Quantitation of the amount of glucocerebrosidase mRNA from a normal individual and an individual with the 1226G/84GG genotype of Gaucher disease. cDNA was PCR-amplified with mismatched primers designed to allow cleavage of the 1226G DNA by *Xho* I restriction endonuclease. Lanes: 1, 1226G/84GG heterozygote, undigested; 2, same as lane 1 but digested with *Xho* I; 3, normal control digested with *Xho* I. The amount of mRNA produced by the 1226G allele and the 84GG allele was approximately equal.

sequences of the displaced mutant and the 1226G mutant, which did not have the 84GG insertion.

The Genomic Sequence. The exon sequence of the genomic DNA of the allele containing the 1226G mutation and the one that did not contain the mutation were identical except for the insertion of an extra guanine at genomic nt 1035 (14) (cDNA nt 84) in exon 2 of the allele without the 1226G mutation.

Quantitation of mRNA. The relative amounts of mRNA produced by the 1226G and 84GG alleles were quantitated using three different approaches. (i) DNA amplified from a full-length cDNA using mismatched oligonucleotides to create an *Xho* I site in the presence of the 1226G was examined. The amount of amplified material with and without the 1226G mutation was approximately the same (Fig. 2). (ii) These

Table 2. Genotypes of 71 Jewish and 38 non-Jewish Gaucher disease patients

Genotype (mutant cDNA nt)	Patients, no.			
	Total	Pv1.1 +/+	Pv1.1 +/-	Pv1.1 -/-
Jewish patients				
1226G/1226G	37	0	0	37
1226G/84GG	22		22	
1226G/?	6		1	5
1226G/1448C	3		2	1
1226G/XOVR*	1			1
1448C/1504T	1		1	
1604A/84GG	1	1		
Non-Jewish patients				
764A/?	1	1		
764A/Trun†	1		1	
1226G/1226G	2			2
1226G/1448C	5		3	2
1226G/476A	1			1
1226G/?	5		3	2
1226G/XOVR*	1			1
1226G/Trun†	1		1	
1448C/1448C	6	5	1	
1448C/1504T	2	1	1	
1448C/?	3		1	2
1448C/XOVR*	3	1		2
1504T/?	1		1	
1604A/84GG	1	1		
?/?	5	2	2	1

*Crossover between gene and pseudogene (6).

†Gene truncated at the 3' end.

results were confirmed by incorporating radioactive dATP in the PCR mixture, removing samples at 15, 21, and 27 cycles, and showing that in the linear portion of the curve the rate of increase in radioactivity of the *Xho* I-cleavable fragment was approximately the same as that of the fragment that was not cleaved. (iii) The amount of the two species of cDNA could also be estimated by the relative strengths of the 1226A and the 1226G signal when the PCR-amplified full-length cDNA was sequenced. Again, the amount of the two gene products seemed to be nearly the same.

Glucocerebrosidase Antigen. The amount of glucocerebrosidase antigen in five cultured skin fibroblast lines from unrelated patients homozygous for the 1226G mutation was found to be 7.0 ± 1.0 milliunits (mean \pm SEM). Thus a single copy of the 1226G gene would be expected to produce approximately 3.5 milliunits of antigen. Four heterozygotes for the 1226G and 84GG mutations averaged 4.3 ± 0.42 milliunits, a value that does not differ significantly from the 3.5 milliunits produced by a single 1226G gene. One may therefore conclude that little or no antigen was produced by the 84GG mutant allele.

Frequency of the 84GG and 1226G Mutations in Gaucher Disease Patients. Table 2 presents the genotypes of 71 Jewish and 38 non-Jewish Gaucher disease patients that we have studied. It is apparent that the 84GG mutation is a very common one in the Jewish population, its frequency being second only to that of the 1226G mutation. It was encountered only once in the non-Jewish patients. As reported (10), the 1226G mutation was invariably linked to the Pv1.1⁻ genotype. The 84GG mutation, on the other hand, was always found in the context of the Pv1.1⁺ genotype.

In contrast to patients homozygous for the 1226G mutation who characteristically have mild late-onset Gaucher disease (3), patients heterozygous for the 1226G and 84GG mutations had relatively severe disease with early onset. The median age of first symptoms or diagnosis of Gaucher disease was 30.5 years in patients homozygous for the 1226G mutation and only 6 years in the group heterozygous for the 1226G and 84GG mutations. The relationship between the clinical severity score and the age of evaluation is presented in Fig. 3.

DISCUSSION

We have previously postulated on the basis of linkage data that one additional common Gaucher disease mutation existed in the Jewish population linked to the Pv1.1⁺ haplotype (10). We have now identified this mutation as an insertion of a guanine after the normal guanine at cDNA nt 84. This mutation is in the portion of the DNA that is translated to the

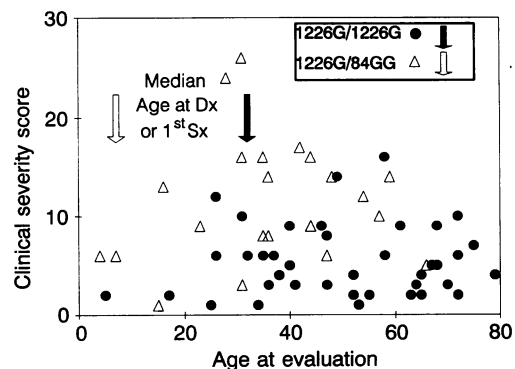


Fig. 3. Relationship between disease severity and age of Gaucher disease patients homozygous for the 1226G mutation and heterozygous for the 1226G and 84GG mutations. The solid arrow indicates the median age at diagnosis (Dx) or first symptoms (Sx) for the 1226G homozygotes and the open arrow indicates the same for the 1226G/84GG heterozygotes.

leader sequence. The frame shift that results from the nucleotide insertion causes a stop codon to appear 18 amino acids downstream from the insertion. The amount of glucocerebrosidase mRNA in such individuals is normal but no translation product could be detected in an *in vitro* translation system and little or no enzyme antigen attributed to this allele could be found in cell lines from patients. Accordingly, the clinical disorder produced in patients heterozygous for the 1226G and the 84GG mutations is more severe than that observed in most homozygotes for the 1226G mutation. Moreover, no homozygotes for the 84GG allele have been encountered, even in the severe infantile form of Gaucher disease. This form of disease is very rare among Jews and it is possible the homozygous state is lethal before birth.

We could identify more than 95% of the disease-producing alleles at the DNA level in 71 Jewish subjects with Gaucher disease. Most of these are mutations at nt 1226, 84, and 1448, and screening for these three mutations accounts for 94.4% of the Gaucher-producing alleles in this group of patients. Among the non-Jewish patients the 1226G and 84GG mutations were less common. Whereas these genes apparently achieve polymorphic frequencies in the Jewish population, they are much less common in non-Jews. As a consequence more sporadic mutations are present in the non-Jewish population and only 75% of the disease-producing alleles were identified.

It would be possible to estimate the various gene frequencies in the population as a whole from the relative gene frequencies in the Gaucher population and from the frequency of the most common gene if all individuals who were either homozygous or compound heterozygotes for Gaucher disease-producing alleles were represented in the patient population. Unfortunately, the gene frequencies observed in the patient population are quite biased. The homozygotes for the 1226G mutation are frequently diagnosed only late in life and may have no significant clinical manifestations of Gaucher disease (1). This results in underestimation of the proportion of 1226G alleles in the general population. Moreover, the 84GG may be lethal in the homozygous state; we have encountered no 84GG/84GG individuals. This, however, would only be a minor perturbing factor because the gene frequency of the 84GG mutation will be relatively low.

Our current best estimate of the frequency of the 1226G mutation in the Ashkenazi Jewish population is about 0.026 [ref. 1 and A. Zimran, personal communication]. If we assume that, in the Ashkenazi Jewish population as a whole, 85% instead of 75% of Gaucher alleles are the 1226G mutation, the estimated population frequency of all Gaucher disease alleles is 0.031. The frequency of alleles other than 1226G, 84GG, and 1448C would be 3.3% of this or 1×10^{-3} . Based on these estimates there is only 1 chance in about 1,000,000 that a Jewish couple in which the 1226G, 84GG, or 1448C alleles have not been found would be at risk for having a child with Gaucher disease. In a couple in which one had been found to have one of these three mutations and the other not the risk would be about 1:1000 if the other partner had been screened only for the three most common mutations and 1:1300 if tests were carried out for the other currently known mutations.

The discovery of the 84GG mutation now makes it possible to screen effectively for Gaucher disease, particularly in Ashkenazi Jewish populations. The implementation of such a screening program would not be inexpensive because of the technology required but, with the cost of treating a patient with Gaucher disease being in the range of \$50,000–\$800,000

per year for enzyme replacement (26), screening should nonetheless prove to be highly cost effective.

Note Added in Proof. Since transmitting this paper, we have found that two of the six Jewish 1226G/? patients have a previously described (21) G → T mutation at cDNA nucleotide 1297 and can, therefore, be designated 1226G/1297T. More than 97% of the mutations among the Ashkenazi Jewish patients are now identified, increasing further the power of DNA analysis in heterozygote detection. We have now examined the DNA of 664 normal Ashkenazi Jewish subjects for the 84GG mutation. It was present in 4 or at 0.6%. Although the frequency of the 1226G mutation was 4.6 times as great as the 84GG in the patient population, it is nearly 9 times as frequent in the healthy population, supporting our assumption that it is underrepresented in the patient population.

This work was supported by National Institutes of Health Grants DK36639 and RR00833 and by the Sam Stein and Rose Stein Charitable Trust Fund. This is manuscript 6971-MEM from the Research Institute of Scripps Clinic.

- Zimran, A., Gelbart, T., Westwood, B., Grabowski, G. A. & Beutler, E. (1991) *Am. J. Hum. Genet.* **49**, 855–859.
- Tsuji, S., Martin, B. M., Barranger, J. A., Stubblefield, B. K., LaMarca, M. E. & Ginns, E. I. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2349–2352, and correction (1988) **85**, 5708.
- Zimran, A., Sorge, J., Gross, E., Kubitz, M., West, C. & Beutler, E. (1989) *Lancet* **ii**, 349–352.
- Tsuji, S., Choudary, P. V., Martin, B. M., Stubblefield, B. K., Mayor, J. A., Barranger, J. A. & Ginns, E. I. (1987) *N. Engl. J. Med.* **316**, 570–575.
- Dahl, N., Lagerström, M., Erikson, A. & Pettersson, U. (1990) *Am. J. Hum. Genet.* **47**, 275–278.
- Zimran, A., Sorge, J., Gross, E., Kubitz, M., West, C. & Beutler, E. (1990) *J. Clin. Invest.* **85**, 219–222.
- Latham, T. E., Theophilus, B. D. M., Grabowski, G. A. & Smith, F. I. (1991) *DNA Cell Biol.* **10**, 15–21.
- Eyal, N., Wilder, S. & Horowitz, M. (1990) *Gene* **96**, 277–283.
- Beutler, E. & Gelbart, T. (1990) *Ann. Hum. Genet.* **54**, 149–153.
- Zimran, A., Gelbart, T. & Beutler, E. (1990) *Am. J. Hum. Genet.* **46**, 902–905.
- Sorge, J., Gelbart, T., West, C., Westwood, B. & Beutler, E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5442–5445.
- Sorge, J., Gross, E., West, C. & Beutler, E. (1990) *J. Clin. Invest.* **86**, 1137–1141.
- Sorge, J., West, C., Westwood, B. & Beutler, E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7289–7293.
- Horowitz, M., Wilder, S., Horowitz, Z., Reiner, O., Gelbart, T. & Beutler, E. (1989) *Genomics* **4**, 87–96.
- Beutler, E., Gelbart, T. & West, C. (1990) *Clin. Chim. Acta* **194**, 161–166.
- Sorge, J. A., West, C., Kuhl, W., Treger, L. & Beutler, E. (1987) *Am. J. Hum. Genet.* **41**, 1016–1024.
- Kumar, R. & Dunn, L. L. (1989) *Oncogene Res.* **4**, 235–241.
- Kogan, S. C., Doherty, M. & Gitschier, J. (1987) *N. Engl. J. Med.* **317**, 985–990.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
- Graves, P. N., Grabowski, G. A., Eisner, R., Palese, P. & Smith, F. I. (1988) *DNA* **7**, 521–528.
- Theophilus, B. D. M., Latham, T., Grabowski, G. A. & Smith, F. I. (1989) *Nucleic Acids Res.* **17**, 7707–7722.
- Wigderson, M., Firon, N., Horowitz, Z., Wilder, S., Frishberg, Y., Reiner, O. & Horowitz, M. (1989) *Am. J. Hum. Genet.* **44**, 365–377.
- Hong, C. M., Ohashi, T., Yu, X. J., Weiler, S. & Barranger, J. A. (1990) *DNA Cell Biol.* **9**, 233–241.
- Beutler, E., Kuhl, W. & Sorge, J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6506–6510.
- Beutler, E. & Kuhl, W. (1970) *Lancet* **i**, 612–613.
- Beutler, E., Kay, A. C., Saven, A., Garver, P., Thurston, D., Dawson, A. & Rosenbloom, B. (1991) *Blood* **78**, 1183–1189.