*Am. J. Hum. Genet. 65:77–87, 1999*

# Structure of the *GM2A* Gene: Identification of an Exon 2 Nonsense Mutation and a Naturally Occurring Transcript with an In-Frame Deletion of Exon 2

Biao Chen,[1,2] Brigitte Rigat,[1] Cynthia Curry,[3] and Don J. Mahuran[1,2]

[1]Research Institute, Hospital for Sick Children, and [2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto; and [3]Genetic Medicine Valley Children's Hospital, University of California San Francisco, Madera, CA

## Summary

**Deficiency of the GM2 activator protein, encoded by *GM2A*, results in the rare AB-variant form of GM2 gangliosidosis. Four mutations have been identified, but the human gene structure has been only partially characterized. We report a new patient from a Laotian deme whose cells are deficient in both GM2-activator mRNA and protein. However, reverse transcription (RT)-PCR detected some normal-sized cDNA and a smaller cDNA species, which was not seen in the RT-PCR products from normal controls. Sequencing revealed that, although the patient's normal-sized cDNA contained a single nonsense mutation in exon 2, his smaller cDNA was the result of an in-frame deletion of exon 2. Long PCR was used to amplify introns 1 and 2 from patient and normal genomic DNA, and no differences in size, in 5′ and 3′ end sequences, or in restriction-mapping patterns were observed. From these data we developed a set of four PCR primers that can be used to identify *GM2A* mutations. We use this procedure to demonstrate that the patient is likely homozygous for the nonsense mutation. Other reports have associated nonsense mutations with dramatically reduced levels of mRNA and with an increased level of skipping of the exon containing the mutation, thus reestablishing an open reading frame. However, a recent article has concluded that, in some cases, the latter observation is caused by an artifact of RT-PCR. In support of this conclusion, we demonstrate that, if the competing, normal-sized cDNA is removed from the initial RT-PCR products, from both patient and normal cells, by an exon 2–specific restriction digest; a second round of PCR produces similar amounts of exon 2–deleted cDNA.**

Received February 12, 1999; accepted for publication May 6, 1999; electronically published June 9, 1999.

Address for correspondence and reprints: Dr. Don J. Mahuran, Research Institute, The Hospital For Sick Children, 555 University Avenue, Toronto, Ontario, Canada, M5G 1X8. E-mail: hex@sickkids.on.ca


0002-9297/99/6501-0012$02.00

## Introduction

The correct synthesis and interaction of three gene products are needed to hydrolyze GM2 ganglioside (GM2). Two of these are the $\alpha$ and $\beta$ subunits of the lysosomal glycosidase $\beta$-hexosaminidase A (Hex A; E.C.3.2.1.52), which irreversibly dimerize in the endoplasmic reticulum. The third is the Hex A substrate (GM2)–specific cofactor, the GM2 activator protein (Activator), which forms a complex with GM2 and then reversibly interacts with Hex A in the lysosome (Gravel et al. 1995; Mahuran 1998; Mahuran, in press). These three proteins are encoded by the *HEXA, HEXB,* and *GM2A* genes, respectively. Although $\geqslant$87 and $\geqslant$23 mutations in *HEXA* and *HEXB* have been associated with various forms of Tay-Sachs disease and Sandhoff disease, respectively, only four mutations in *GM2A* have been described, and all are likely to occur in the homozygous form in four patients with the AB-variant form of acute GM2 gangliosidosis (MIM 272750; reviewed by Gravel et al. 1995; Mahuran, in press). Although the size, placement, and partial nucleotide sequence at each end of the introns in human *HEXA* and *HEXB* have been determined (Proia and Soravia 1987; Neote et al. 1988; Proia 1988), those comprising *GM2A* have been incompletely characterized (Klima et al. 1991).

It is known that, in vivo, the *GM2A* mRNA is rare and requires the use of poly $A^+$ RNA to be detected by northern blotting techniques, which have estimated its length to be ~2.5 kb. The Activator's coding region comprises only 582 nucleotides, with most of the remainder forming a long 3′ untranslated end sequence (Klima et al. 1991; Xie et al. 1991; Nagarajan et al. 1992). *GM2A* has been mapped to chromosome 5q 31.3-33.1 (Heng et al. 1993; Swallow et al. 1993), and there is also a nonfunctional processed *GM2A* pseudogene on chromosome 3 (Xie et al. 1992*a*; Swallow et al. 1993). *GM2A* is a small gene, $\leqslant$16 kb in length. Exons 2–4 have been identified in genomic clones, with the existence of exon 1 being extrapolated to account for the remaining 81 bp of 5′ coding sequence found in cloned cDNA. Among the three introns, only the 0.38-kb intron

3 has been fully sequenced. The 5′ and 3′ end sequences of intron 1 and 2 remain incomplete (Klima et al. 1991).

In this article, we identify a fifth mutation in *GM2A,* a nonsense mutation in exon 2 ($Glu^{54}STOP$). As in the other four documented AB-variant cases, the mutation is likely present in the homozygous form (the possible presence of a second large *GM2A* deletion allele could not be completely ruled out, in any of these cases), and the patient exhibits the acute GM2 gangliosidosis phenotype, similar to acute Tay-Sachs and Sandhoff disease (reviewed by Gravel et al. 1995; Mahuran, in press). In the course of characterizing the biochemical effects of this nonsense mutation on mRNA splicing and stability, we determined the size, partial-restriction map, and 5′ and 3′ end sequences of *GM2A*'s three introns. We also designed four pairs of primers that can be used with common PCR conditions for the amplification of the coding sequence plus each of the intron/exon junctions of *GM2A*. These four PCR reactions, done simultaneously, make it possible for other investigators to efficiently screen *GM2A* for new mutations.

Finally, in the context of our novel *GM2A* nonsense mutation, we examine the controversy (reviewed by Maquat 1996) surrounding the hypothesis that nonsense mutations promote the skipping of the exon in which they are contained, to reestablish the reading frame (Dietz et al. 1993; Mazoyer et al. 1998). This hypothesis would require the presence in the nucleus of a mechanism for reading the frame prior to the splicing out of introns (reviewed by Dietz and Kendzior 1994).

## Patient and Methods

### *Case Report*

The patient, a child of Laotian Hmong ancestry (i.e., from a geographically isolated, small Laotian hill tribe), was first examined at age 2.5 years for evaluation of his neurodegenerative course. He was the product of a term pregnancy complicated only by decreased fetal activity late in gestation. Two older siblings and one younger sibling were clinically normal. Careful evaluation of the family pedigree revealed no consanguinity. The child was thought to be normal until age ~5 mo, when he was noted to have delayed motor milestones and increasing weakness. At age 9 mo, magnetic resonance imaging showed increased signal density in the periventricular white matter and altered signal density in the basal ganglia. Ophthalmologic evaluation showed bilateral macular cherry-red spots. Lysosomal enzyme studies showed normal enzyme activity in lymphocytes for Hex A, arylsulfatase A, α-L iduronidase, sphingomyelinase, galactocerebrosidase, and mannosidase. Additional laboratory evaluation included normal very-long–chain fatty acid and phytanic acid levels, normal quantitative test results on amino and organic acids, normal lactate and pyruvate levels, and normal findings of mitochondrial DNA studies.

At the time of evaluation, the patient was experiencing approximately three major motor seizures and hundreds of myoclonic jerks per day. Hyperacusis was extreme, with an exaggerated startle response. Frequent suctioning was required for pooling of secretions in the oropharynx, and sucking and swallowing were impaired. He was able to indicate hunger by crying and responded minimally to the family.

Physical examination revealed a nondysmorphic, profoundly hypotonic child, who was unresponsive to his environment. He had numerous myoclonic seizures and distinct hyperacusis. Growth parameters included a head circumference of 50 cm (50th percentile), height of 35.5 cm (25th–50th percentile), and weight of 10.2 kg (3d percentile). There was a bell-shaped thorax and pectus excavatum. Abdominal muscles were weak, with diastasis recti. Tone was diffusely and severely decreased. Deep-tendon reflexes were brisk, with bilateral ankle clonus. There was profound head lag and severe generalized hypotonia. Mild contractures were seen at the elbows, knees, and ankles. Roving eye movements with no-gaze palsy were noted. There was minimal response to painful stimuli.

Despite the normal Hex A levels in lymphocytes, the clinical evaluation strongly suggested Tay-Sachs disease. A lumbar puncture was done for ganglioside profile, which revealed a highly elevated total ganglioside content (performed by Dr. David Wenger at Jefferson Medical College, Philadelphia). Skin biopsy was done for fibroblast DNA analysis.

The patient was last examined at age 2 years 8 mo, at which time he had increasing respiratory distress and severely decreased muscle mass. He was subsequently admitted for pneumonia. The family moved out of state after the child was discharged from the hospital. We assume that the patient expired but are unable to confirm his age at death.

### *Cell Lines and Leukocyte Sample*

Two normal fibroblast cell lines, referred to as N1 and N2, were obtained from the tissue-culture facility at the Hospital for Sick Children. The patient's cultured fibroblast cells were obtained from Valley Children's Hospital. These fibroblast cell lines were grown in α-minimal essential media (αMEM; Flow Laboratories) supplemented with 10% (v/v) fetal-calf serum, 100 μg streptomycin/ml, and 100 μg penicillin/ml.

### *Western Blot Analysis*

The confluent fibroblast cells from N1 and N2 and the cells from the Laotian patient were harvested with $NaH_2PO_4$ lysis buffer (10 mM $NaH_2PO_4$ with 5% glycerol) and subjected to five rounds of freeze thawing. The

lysate was extracted with an equal volume of $CCl_4$, and the total protein was determined (Lowry et al. 1951). Fifty micrograms and 100 μg of total lysate protein, from normal or from patient cells, were mixed separately with equal volumes of loading buffer (containing 20 mM DTT and 4% SDS), and were then boiled for 5 min. Western blotting was done as described elsewhere (Xie et al. 1992*b*).

### *Poly A RNA Isolation and Northern Blot Analysis*

Since the activator mRNA is rare (Xie et al. 1991), poly $A^+$ RNA must be used for northern blot analysis. Northern blot analysis of 4 μg poly $A^+$ RNA from normal and from AB-variant cells was done as described elsewhere (Xie et al. 1992*b*), with use of two radiolabeled probes consisting of a full-length GM2 activator cDNA (~600 bp) and, as a control, a ~350 bp cDNA fragment of cytochrome C oxidase complex VI (obtained from Dr. B. H. Robinson, Hospital for Sick Children).

### *Total RNA Isolation and RT-PCR*

Total RNA was isolated from 150 mm–diameter plates of confluent cultured fibroblast cells from N1 and N2 and from the patient, by means of the guanidinium-thiocyanate method (Chomczynski and Sacchi 1987). Two primers, 5′-TTGGATCCCACCCTTCCCGATGCAG (primer 1680, upstream, −15→16 [counting from the first "A" of the Activator initiating ATG]) and 5′-GGATCCGTGGGAGTTTGGCCTTGGC (primer 705, downstream, 666→648) were designed for reverse transcription and PCR (*Bam*HI sites are underlined). Reverse transcription was done in a total volume of 100 μl, containing 2 μg total RNA, 0.6 μg primer 705, 0.5 mM each dNTP, and 200 U MMLV transcriptase (Gibco BRL). The mixture was incubated at 37°C for 1.5 h. Twenty microliters of these transcription solutions were directly used for PCR, done in a Robocycler 40 (Stratagene). After a "hot start," 2.5 U *Taq* DNA polymerase (Promega) were added in a 50-μl reaction volume, with 1.25 μM each of primers 705 and 1680, 0.2 mM each dNTP, and 1.5 mM $MgCl_2$. PCR was done for 35 cycles each of 2 min at 94°C, 30 s at 61°C, and 1 min at 72°C, with a final-addition cycle of 2 min at 94°C, 30 s at 61°C, and 5 min at 72°C.

### *Cloning and Sequencing of Activator cDNA Samples*

Two different sizes of cDNA from the RT-PCR product, obtained from patient cells after separation, were observed by means of agarose-gel electrophoresis. These were purified with the Gene Clean kit (Bio/Can Scientific). One microliter each of the cDNA product from the patient and 1 μl of fresh RT-PCR product from N1 and N2 were mixed separately with 2 μl PCR 2.1 vector and 1 μl T4 ligase from the TA cloning kit (Invitrogen) and then ligated at 14°C overnight. One microliter of the ligation reaction was transformed in INVαF One Shot Competent Cell (Invitrogen). White clones were selected and grown in LB media containing 50 μg ampicillin/ml. Isolation of the plasmid DNA was done with a Qiagen Mini-Preparation kit.

The DNA sequences of each cDNA were determined by the dideoxy chain–termination method (see below). To ensure accuracy, we determined the full-length sequences of three clones from each sample, using sense and antisense strands.

### *Isolation of Genomic DNA*

Confluent fibroblast cells were harvested by scraping with a rubber policeman, resuspended in 2 ml PBS, and centrifuged. The cell pellets (from fibroblasts or leukocytes) were resuspended in lysis buffer (100 mM Tris-Cl [pH 8], 40 mM EDTA [pH 8], 0.2% SDS, and 0.6 mg proteinase K/ml), ground, and incubated at 55°C for 10 h. The mixture was extracted with 1 vol phenol/chloroform, and the DNA was precipitated by addition of 2 vol 100% ethanol. The DNA pellet was rinsed with 70% ethanol and resuspended in 10 mM Tris (pH 8) and 1 mM EDTA (pH 8).

### *Long PCR to Amplify Intron 1 and Intron 2 of* GM2A

Two pairs of primers were designed with RIGHTPRIMER (BioDisk), for the amplification of intron 1 and intron 2 of *GM2A*. The primers for intron 1 were each composed of 18 nt exonic sequence plus 10 nt, to produce an *Xho*I site; that is, 5′-GCGCCTCGAGGACCCACCCTTCCCGATG (primer 2833, upstream, −15→3) and 5′-GCGCCTCGAGCAGGGGGACACTGGTGCT (primer 2834, downstream, 228→211) (the *Xho*I sites are underlined). The primers for the amplification of intron 2 were each composed of 32 nt exonic sequence plus 10 nt, to generate the *Xho*I site; that is, 5′-GCGCCTCGAGTTCCTGGGATAACTGTGATGAAGGGAAGGACC (primer 805, upstream, 102→133) and 5′-GCGCCTCGAGGGCAGGGCTCCCCAGTAGGAATTAACATGTCA (primer 806, downstream, 376→345).

The *rTth* extra-long DNA polymerase (*rTth* XL) was purchased from Perkin-Elmer. Intron 1 was amplified with primers 2833 and 2834, by use of the GeneAmp PCR system 2400 (Perkin-Elmer). An AmpliWax PCR Gem 100 bead (Perkin-Elmer) was added. PCR was done in a total volume of 100 μl containing 0.4 μg template, 0.5 μM primers, 4 U *rTth* XL, 200 μM each dNTP, and 1.1 mM $Mg(C_2H_3O_2)_2$. Denaturation was at 94°C for 1 min, followed by 35 cycles each of 15 s at 94°C, 30 s at 60°C, and 5 min at 72°C; and an extra extension period of 10 min at 72°C.

Intron 2 was amplified with the Advantage KlenTaq polymerase mix (Clontech), which contained KlenTaq-

1 DNA polymerase, TaqStart anti-*Taq* antibody, and minor Deep Vent DNA polymerase. Intron 2 was amplified with primers 805 and 806 in a Robocycler 40 (Stratagene). PCR was done with a total reaction volume of 50 $\mu$l containing 0.4–0.6 $\mu$g templates, 0.5 $\mu$M primers, 1 $\mu$l Advantage KlenTaq polymerase mix, and 200 $\mu$M each dNTP, in a buffer of 40 mM Tricine-KOH (pH 9.2) containing 15 mM $KC_2H_3O_2$, 3.5 mM $Mg(C_2H_3O_2)_2$, and 75 $\mu$g BSA/ml. The cycling parameters were denaturation at 99°C for 35 s, annealing at 67°C for 30 s, and elongation at 68°C—first for 9 min during cycles 1–15, then for 11 min during cycles 16–25, and finally for 13 min during cycles 26–35. Ten microliters each of the PCR products were mixed with DNA-loading dye and were analyzed in a 1% agarose gel.

### *Restriction Analysis of Intron 1 and Intron 2*

The PCR products of intron 1 and intron 2 were digested with the following endonucleases: *Bam*HI, *Sst*I, *Eco*RI, *Xba*I, *Kpn*I, *Hin*dIII, and *Xho*I. The digestion results were analyzed in a 1% agarose gel. Four *Bam*HI-digestion fragments comprising intron 1 were separated with QIAEX II gel extraction kit (Qiagen), and each fragment was digested with the other six endonucleases. Two *Sst*I-digestion fragments comprising intron 2 were also separated, and each fragment was digested with the other six endonucleases.

### *Determination of the Length of Intron 1 of the Patient's* GM2A *by Long PCR*

A fragment containing intron 1 of *GM2A* was amplified from patient fibroblast DNA, by use of primers 2833 and 2834, with *rTth* XL DNA polymerase (see the Long PCR to Amplify Intron 1 and Intron 2 of GM2A subsection, above). Another fragment containing intron 2 of *GM2A* was also amplified from the patient's DNA, by use of primers 805 and 806, with Advantage KlenTaq polymerase (see the Long PCR to Amplify Intron 1 and Intron 2 of GM2A subsection, above). Each PCR product was digested with *Bam*HI, *Eco*RI, and *Bam*HI plus *Eco*RI and then analyzed in a 1% agarose gel.

### *Identification and Subcloning of Both Ends of Introns 1 and 2 from Long PCR of Normal Genomic DNA*

Restriction analysis showed that two fragments of intron 1 digested by *Bam*HI (0.6 kb and 1.2 kb) were the 5′ and 3′ ends of intron 1 (which contains terminal *Xho*I sites). These two fragments were digested independently with *Xho*I and were purified with use of the QIAEX II gel extraction kit. One hundred nanograms of each fragment was then ligated into the *Bam*HI/ *Xho*I sites of 30 ng pBluescript vector (Stratagene), with T4 ligase (Boerhinger). The ligation products were transformed with XL-Blue MRF competent cell (Stratagene). White clones were selected and grown in LB media containing 50$\mu$g ampicillin/ml. Isolation of the plasmid DNA was done with the Qiagen Mini-Preparation kit. Four-kilobase and 2.9-kb fragments of intron 2 digested with *Sst*I were also subcloned into the *Sst*I/*Xho*I site of the same vector and were analyzed by similar methods.

### *Nucleotide Sequencing on the Normal 5′ and 3′ Ends of Introns 1 and 2*

The nucleotide sequences of the preceding inserts were determined by the dideoxy-chain–termination method (Sanger et al. 1977). Plasmid DNA products (1.5 $\mu$g) containing each end of the introns were mixed with 30 ng each of T7 or T3 primers (Stratagene), $\alpha[^{35}S]$-dATP (Amersham), and other components of the Sequenase kit as recommended by the manufacturer (Pharmacia). To confirm the sequences of the introns, we sequenced both antisense and sense strands from PCR products obtained from five different normal genomic DNA samples.

### *Designing an Efficient PCR Method to Amplify the Exons and Exon/Intron Junctions of GM2A*

On the basis of the nucleotide sequences obtained from the 5′ and 3′ ends of the introns (GenBank), four pairs of primers were designed (table 1) with RIGHTPRIMER (BioDisk). Fragments made up of exons and their intronic flanking regions were amplified with each pair of primers (table 1) by use of a GeneAmp PCR System 2400 (Perkin-Elmer). PCR was done in a total volume of 100 $\mu$l containing 0.6 $\mu$g genomic DNA, 0.5 $\mu$M primers, 0.2 mM each dNTP, 2.5 mM $MgCl_2$, and 2.5 U *Taq*-Gold DNA polymerase (Perkin-Elmer). The cycling parameters used were as follows: 10 min at 94°C for denaturation; 43 cycles each of 30 s at 94°C, 30 s at 54°C, and 30 s at 72°C; and then 10 min at 72°C for elongation. With these conditions, PCR amplification of exon 3 and its flanking region produced some nonspecific DNA. Although these nonspecific bands did not interfere with the isolation of the target fragment, it was determined that they could be removed by lowering the concentration of $MgCl_2$ to 2.0 mM. On the basis of exonic and intronic sequences obtained from this and a previous study (Xie et al. 1991), each of the PCR fragments—E1, E2, E3, and E4—was designed to contain a unique restriction site, *Xba*I, *Sst*I, *Pst*I, and *Eco*RI, respectively (table 1). These restriction digestion sites can be used as controls to confirm the identity of each amplified fragment.

### *PCR of Genomic DNA Fragments from Patient and Controls*

Exon 1, exon 2, and their flanking regions in *GM2A* were amplified from the genomic DNA of the patient

**Table 1**

**Sequence of the Primers Used to Amplify Exons and Exon/Intron Junctions of *GM2A***

| Exon and Primers (22 nt) | 5′ Location[a] | Melting Temperature (°C) | Size (bp) | Verification Digest | |
|---|---|---|---|---|---|
| | | | | Enzyme | Fragment (bp) |
| E1: | | | | | |
| GGAAGGCATTTAAAGGACCTCT | E1 (−69) | 58.1 | 426 | *Xba*I | 246+180 |
| AAGGCTGTCTGCATTTTCACTC | IVS 1 (256) | 60.0 | | | |
| E2: | | | | | |
| CCCAGGTTCATAGGTATGGAGT | IVS1 (−149) | 56.2 | 528 | *Sst*I | 76+452 |
| GCTGGCGCTGACATCAGAGATT | IVS2 (218) | 56.2 | | | |
| E3: | | | | | |
| GATTATTAATGTCTGCCATGGG | IVS2 (−220) | 54.4 | 716 | *Pst*I | 484+232 |
| GCTAGGATTGCAGCCATGAACT | IVS3 (313) | 58.1 | | | |
| E4: | | | | | |
| CATGTCTCTGGATTTGTAAGCC | IVS3 (−290) | 56.2 | 665 | *Eco*RI | 312+352 |
| GGCTATCAAGAACTGTCCAACT | E4 (802) | 56.2 | | | |

Note.—Part of the 3′ UTR is excluded from the data.

[a] Nucleotide numbering in exonic sequences is done from the "A" of the first ATG as nt 1; nucleotides in intronic sequences are numbered from the first nucleotide of the 5′ end of the intron (plus number) or the last nucleotide of 3′ end of the intron (minus number).

and N1 and N2 with use of the first and second pairs of primers in table 1, as described in the previous sections. Each PCR product was purified with the QIAGEN PCR purification kit and directly sequenced.

### *Direct Sequence of PCR Fragments*

To examine the nucleotide sequences at both ends of introns 1 and 2, we performed direct sequencing of the above PCR fragments, using the Thermosequenase radiolabeled terminator cycle sequencing kit (Amersham). The reactions were done in a Robocycler 40 (Stragagene), with 2 min at 94°C for denaturation and 25 cycles each of 30 s at 94°C, 30 s at 55°C, and 1 min at 72°C.

### *Detection of the Exon 2 Deleted Activator mRNA in Normal Cells*

DNA sequences from *GM2A* showed that only a single *Hin*fI site exists in the cDNA encoding the Activator, and it resides in exon 2. Therefore, the full-length cDNA can be cleaved by *Hin*fI, whereas any cDNA species lacking exon 2 will remain intact after digestion. Five microliters each of RT-PCR products from N1 and from the patient sample were digested with *Hin*fI. Two "nested" primers, ATCGCCCTGGGCTTGCTT (primer 1438, upstream, 31→48) and ACAAAACAGAGGAAAAGG (primer 1968, downstream 642→625), were used for a second round of PCR. Nested PCR amplification was done with 1 $\mu$l/50 $\mu$l digested RT-PCR products or with an equal amount of undigested RT-PCR products, with 0.2 $\mu$M each of primers, 1.0 mM $MgCl_2$, and 0.2 mM each dNTP. The cycling parameters were 10 min at 94°C for denaturation; 35 cycles each of 30 s at 94°C, 30 s at 61°C, and 30 s at 72°C; and then a final 10 min at 72°C for elongation. Ten nanograms of either the wild-type or of the exon 2–deleted cloned cDNA, as well as an $H_2O$ sample, were included as controls for the nested PCR amplification.

## Results

The initial diagnosis of the acute AB-variant form of GM2 gangliosidosis given the patient was done on the basis of his clinical presentation coupled with the presence of normal levels of both Hex A and B in his leukocytes and plasma (see the Case Report subsection, above). Because the four previously described *GM2A* mutations causing disease result in undetectable levels of Activator cross-reacting material (CRM), we first analyzed the lysate from the patient's fibroblasts by using

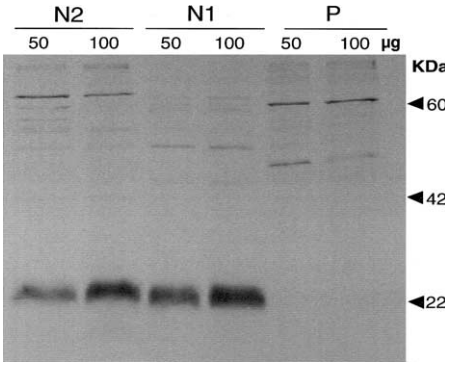



**Figure 1** Western blot analysis, done with a primary rabbit anti-Activator IgG, of fibroblast lysate (either 50 or 100 $\mu$g total protein) from N1, N2, and the patient.

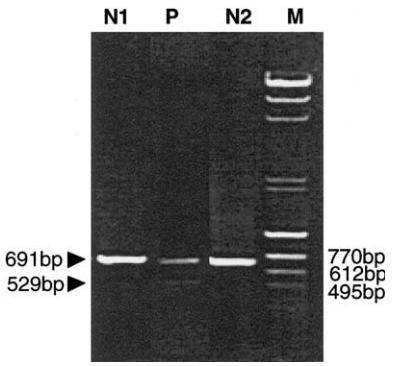



**Figure 2** RT-PCR of total RNA from N1, N2, and the patient. Arrows indicate the normal-sized product (691 bp), the unexpected smaller product (529 bp) found only in the patient's cDNA, and, in lane M containing λDNA/ *Hin*dIII markers, the sizes of the surrounding markers.

western blotting. Whereas lysate from two normal control cell lines produced easily detectable immunoreactive bands corresponding to the expected molecular weight of 22,000, a similar band was not detectable in lysate samples from the patient's cells containing similar levels of total protein (fig. 1). This observation confirmed the patient's diagnosis.

We next examined the level of Activator mRNA from poly $A^+$ RNA isolated from patient and normal cells. No Activator mRNA could be detected in the patient's sample (data not shown). To determine whether any RNA was transcribed from either of the patient's *GM2A* alleles, we performed RT-PCR. Although this procedure resulted in the expected single 691-bp band in samples from two normal cell lines (fig. 2, lanes N1 and N2), two bands were amplified from the patient's sample, corresponding to 691 and 529 bp (fig. 2, lane P). However, the amount of the smaller cDNA was much less than the larger, normal-sized cDNA.

The cDNAs from N1 and N2 and the two different-sized cDNAs from the patient's cells (fig. 2) were cloned and their sequences were determined. The sequence of the cDNA from N1 was consistent with the previously published wild-type sequence. However, the cDNA from N2 contained a single G175→A transition, which would encode a $Val^{59}$Ile substitution in the Activator protein. This substitution was also found in DNA samples from several other normal individuals (data not shown); thus, it must represent a common neutral polymorphism. More significantly, in the larger cDNA from the patient, a G160→T transversion in exon 2 was found (fig. 3), which converts the codon for $Glu^{54}$ to a STOP codon. The smaller cDNA was also sequenced and determined to have an in-frame deletion of the same exon 2 (fig. 4). The complete sequences of both cDNAs were ascertained and determined to have no other changes from the wild type. Identical data were obtained from at least one other clone of each cDNA, from different RT-PCR products (data not shown).

To determine the mechanism responsible for the smaller cDNA found in the patient's RT-PCR product, we first completed the characterization of the normal *GM2A* structure. The size of intron 1 and 2 was determined by long PCR of genomic DNA from a normal individual. The initial primers used to amplify both introns were located on exons, and these exonic sequences allowed us to orient the cDNA fragments in GM2A. The PCR product containing the 3′ end of exon 1, all of intron 1, and the 5′ end of exon 2 was ~6.7 kb in length (fig. 5, lane I1), indicating that intron 1 is ~6.45 kb long. The PCR product containing the 3′ end of exon 2, all of intron 2, and the 5′ end of exon 3 was ~6.9 kb (fig. 5, lane I2), indicating that the length of intron 2 is ~6.60 kb. These long-PCR analyses were also done on DNA samples from the patient's cells and produced identical results (data not shown). Restriction-digestion patterns of the two long-PCR products from the patient's DNA were also identical to those from normal samples (data not shown).

By combining our new data with previously determined sequence data, we were able to generate a restriction map of the entire gene (fig. 6*A*). Also, by cloning and sequencing each intron-exon–junction fragment, we were able to design four pairs of PCR primers (fig. 6*B*; table 1) that can amplify the gene with common PCR conditions (see the Patient and Methods section, above). In addition, each PCR fragment was designed to contain a unique restriction site that allows for an easy confirmation of its identity (fig. 7, table 1).

Using our newly designed primer sets "E1" and "E2" (table 1) to amplify the patient's DNA, we obtained single bands of the expected sizes, 426 and 528 bp, respectively (data not shown). Direct sequencing of these fragments showed that the patient was either homozygous for the nonsense mutation G160→T or heterozygous for that mutation and for a large *GM2A*-deletion

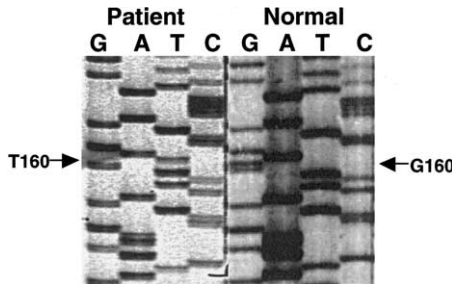



**Figure 3** Partial nucleotide sequence of a cDNA fragment from a normal control and the patient. The arrows indicate the single G160→T transversion found.

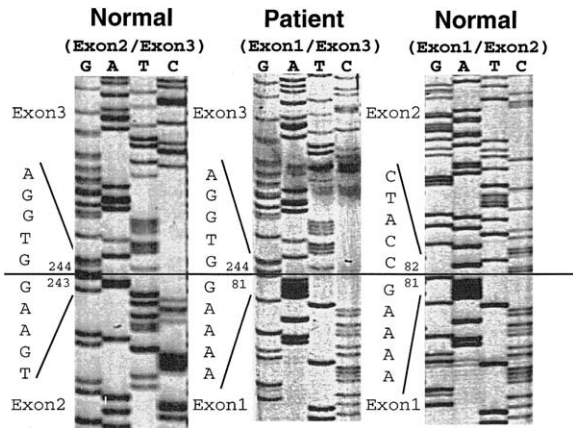



**Figure 4** Partial-nucleotide sequence of the smaller cDNA RT-PCR product from patient samples, and cDNA fragments from normal controls.

allele (fig. 8). All other sequences, including those at the exon 1/ intron 1, intron 1/exon 2, and exon 2/intron 2 junctions, were normal (data not shown). Thus, neither these nor any of the above data explain the presence of the exon 2–deleted cDNA species (ΔE2) seen in the patient's—but not in the normal controls'—RT-PCR products (fig. 2).

Because northern blotting was unable to detect any Activator mRNA (either normal-sized or ΔE2; data not shown) in the patient's cells, we believed it was possible that a similarly low level of the smaller ΔE2 might also be present in normal RT-PCR samples (i.e., a naturally occurring, alternatively spliced form of mRNA), but, because of the relatively high level of normal Activator mRNA in the controls, ΔE2 was not detectable by RT-PCR. To test this theory, we developed a nested PCR procedure specifically for use on the ΔE2 product. Because only one *Hin*fI site exists in normal Activator cDNA and since it is located in exon 2, digestion of the normal RT-PCR fragment with this enzyme should preclude its amplification, while at the same time allowing amplification of ΔE2, with use of nested primers 1438 and 1968 (see the Detection of the Exon 2 Deleted Activator mRNA in Normal Cells subsection, above) for a second round of PCR. Without digestion by *Hin*fI, the second-round PCR product of the normal sample predominantly contained the full-length product, but some of the small ΔE2 became visible (fig. 9, lane 3), although still apparently less than can be seen in the patient's sample (fig. 9, lane 4). However, after the digestion of both the patient's and the normal control's RT-PCR products with *Hin*fI, the second-round nested PCRs produced predominantly the ΔE2 cDNA fragment in about equal amounts (fig. 9, lanes 5 and 6). A small amount of the full-length fragment is also detectable, to reflect the fact that restriction digestions are not 100% efficient (Valentine and Heflich 1997). Significantly, none of the ΔE2 product was detected in the PCR products from cloned normal cDNA (fig. 9, lane 1), indicating that the ΔE2 cDNA was a small, undetectable component in the first-round normal RT-PCR sample (fig. 2, lanes N1 and N2), and its presence in the second round of PCR (fig. 9, lanes 3 and 5) was not caused by contamination. The lack of contamination is also supported by the negative $H_2O$ control (fig. 9, lane 7).

## Discussion

The Activator is a small monomeric protein of 193 amino acids, and the first 23 residues are predicted to serve as a signal peptide, which is removed cotranslationally. In the lysosome, further processing of the N terminus results in a mature chain of 162 residues (reviewed by Mahuran 1998). The deficiency of the Activator results in the AB-variant form of GM2 gangliosidosis, which is a very rare disease (Gravel et al. 1995; Sandhoff et al. 1995). Only four disease-associated mutations have been described: ΔAAG264 (ΔLys[88]), ΔA410 (33 new amino acids, with a loss of 24; Schepers et al. 1996), and T412→C (Cys[138]Arg) (Schröder et al. 1991; Xie et al. 1992*b*) in exon 3; and G506→C (Arg[169]Pro) (Schröder et al. 1993) in exon 4. These were all presumed to be present in the homozygous form in the four respective patients. Whereas none of these mutations result in detectable Activator CRM in patient cells, all of them have been reported to result in normal steady-state levels of Activator mRNA. Cells from the new AB-variant pa-

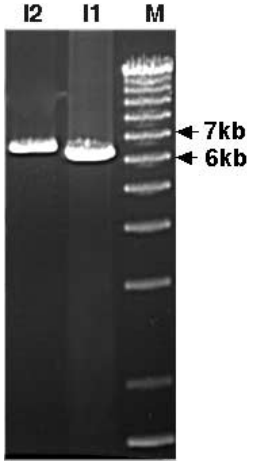
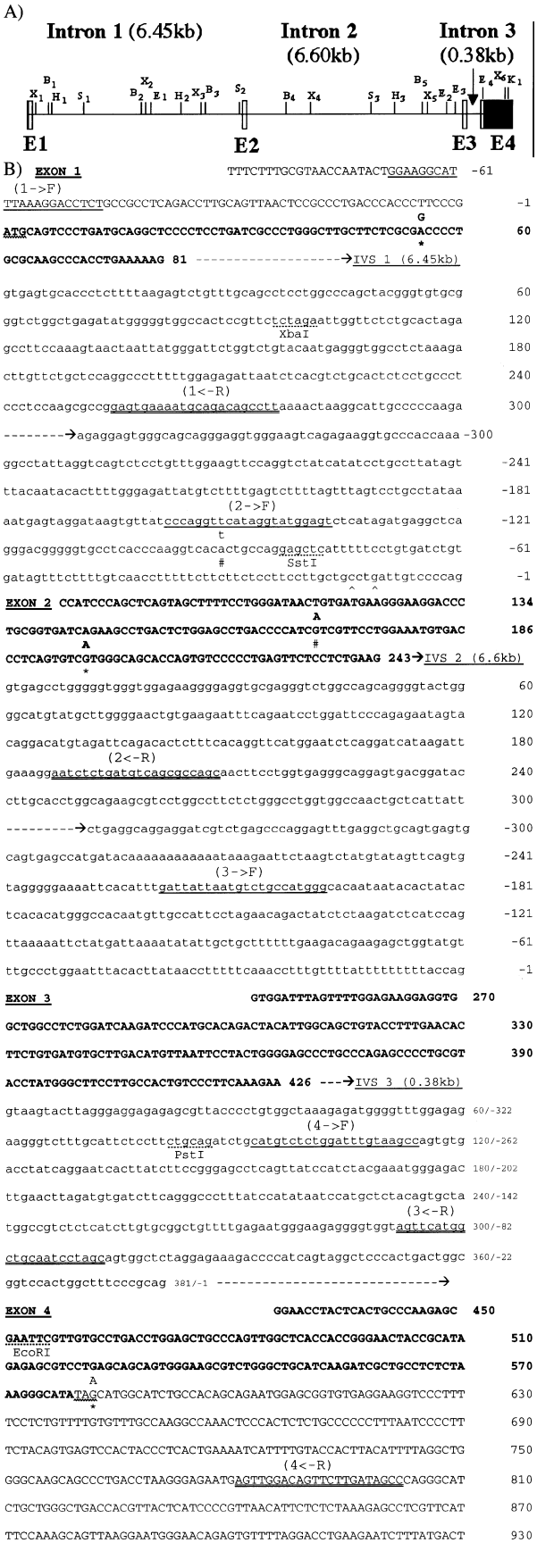



**Figure 5** Determination of the length of intron 1 (~6.45 kb) and intron 2 (~6.60 kb) of *GM2A*. Intron-containing fragments were amplified by long PCR with use of exonic primers. Each fragment contained ~150 bp exonic sequence at each end. Size markers ("M") are also shown.




tient we report here show no detectable steady-state mRNA or any CRM, by northern (data not shown) or western blotting (fig. 1), respectively. Only when we performed RT-PCR could Activator mRNA of normal size be detected in the patient's cells, but we also observed a smaller cDNA species present at an even lower level (fig. 2). Nucleotide sequencing of the cDNAs uncovered a nonsense mutation in exon 2 at the codon for $Glu^{54}$ (encoding the loss of 140 residues) in the normal size (fig. 3), and an in-frame deletion of exon 2 in the smaller cDNA, ΔE2 (fig. 4), which retains the original open reading frame. Direct sequencing of genomic DNA indicated that the patient's intron/exon junctions and flanking sequences of his *GM2A* gene were normal (data not shown) and that he appeared to be homozygous for the nonsense mutation (fig. 8). These data eliminated the possibility that the low level of the smaller ΔE2, detected by RT-PCR in the patient's cells, was a product of abnormal splicing from a second mutant allele but cannot totally exclude the possible presence of a second very large *GM2A* deletion allele (i.e., the patient is hemizygous).

Premature STOP codons can be generated directly by a nonsense point mutation or indirectly by a frame-shift mutation. The latter can be generated by abnormal mRNA splicing, deletions, or insertions. There are numerous cases of these types of mutations causing GM2 gangliosidosis (Gravel et al. 1995; Mahuran, in press). Of those in *HEXA* or *HEXB,* where steady-state mRNA levels have been reported, only 1 of 13 have not been associated with a dramatically reduced amount of transcript (Gravel et al. 1995; Mahuran, in press). This lone exception also produces the most–C-terminal STOP codon, ΔC1510, in exon 13 of 14 in *HEXA* and results

**Figure 6** *A,* Restriction map of *GM2A;* B = *Bam*HI, E = *Eco*RI, S = *Sst*I, X = *Xba*I, K = *Kpn*I, and H = *Hin*dIII. No *Xho*I sites were present in intron 1 or 2. The numbers with each restriction site indicate the number of times it has appeared, reading from 5′ to 3′ in the gene structure. Intron lengths are given in brackets. The unblackened boxes refer to exons, and the blackened boxes refer to 5′ and 3′ UTRs of the cDNA. *B,* Sequence of the exons and exon/intron junctions of *GM2A;* intronic sequences are given in lowercase, the encoding exonic sequences in bold uppercase, and the untranslated exonic sequences in regular uppercase. The translation initiation (exon 1) and termination (exon 4) sites are wavy-underlined. The forward and reverse primers for amplification of the 4 exons and their flanking regions are single underlined and double underlined, respectively. The number followed by an arrow and either an "F" (forward) or "R" (reverse) in brackets above each primer indicates the exon being amplified and also the primer's direction (see table 1). The unique restriction site in each fragment, used to confirm its identity, is dotted-underlined. Three reported exonic polymorphisms are indicated with an asterisk (*). The two new polymorphisms identified during this study, G175A and IVS1 a-92t, and are indicated with a pound sign (#). We also report the correction of previously published sequences at IVS1–12 and IVS1–15 (Klima et al. 1991), which are indicated by a caret (^).

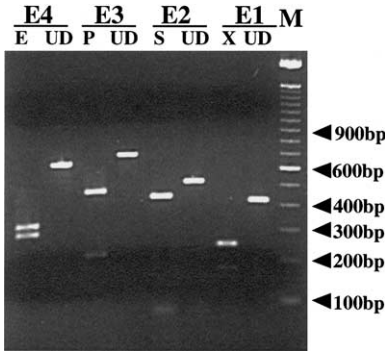



**Figure 7** PCR fragments containing exons E1–E4 and their flanking intronic sequences, amplified with the primers in table 1. Each fragment is shown intact undigested (UD) or digested with its unique restriction site. X = *Xba*I, S = *Sst*I, P = *Pst*I, and E = *Eco*RI (table 1). Size markers (M; 100 bp) are also shown.

in the loss of only 22 residues (Zokaeem et al. 1987). In contrast to the effect of early STOP codons on steady-state levels of mutant mRNA, the four in-frame deletions or insertion mutations that have been characterized in *HEXA* and *HEXB* all produce normal levels of mutant mRNA (Mahuran, in press). An example of this type of mutation is the major *HEXA* mutation among Japanese Tay-Sachs patients, in which a "$g^{-1}t$" substitution in intron 5 results in the in-frame deletion of exon 6 (Tanaka et al. 1993). Identical observations have been made for similar mutations in other genes (Maquat 1996).

In light of these examples, we believe that it is not surprising that the premature STOP we report in exon 2 at codon 54 results in no detectable Activator mRNA, by northern blotting, in the patient's cells, whereas a previously reported STOP in exon 3 at codon 170 does (Schepers et al. 1996). However, the presence of an even lower steady-state level of the smaller ΔE2 product detected by RT-PCR in the patient's cells, but not in normal cells, remained to be explained. If this was caused by a second unidentified deletion allele, the deletion would have to encompass all of exon 2 and either >149 bp of the 3′ end of IVS1 or >218 bp of the 5′ end of IVS2; otherwise, the deletion would have been detected by long PCR done with our exonic primers (see the Long PCR to Amplify Intron 1 and Intron 2 of GM2A subsection, above) or by conventional PCR done with the intronic primers we subsequently designed to amplify exon 2 and its flanking sequences (see table 1, exon E2). It is unlikely that such a deletion would be able to transcribe an RNA product that was missing only exon 2. Furthermore, because no early STOP codon is present in the ΔE2 RNA, its stability should be similar to that of the wild-type transcription product (Maquat 1996). Thus, the very low level of ΔE2 RNA strongly indicates that it is not transcribed by a second unidentified deletion allele but is an alternatively spliced form of RNA transcribed from the homozygous or hemizygous Glu[54]STOP *GM2A* allele(s) of the patient (our data do not eliminate the possible presence of a second, very large deletion allele that does not transcribe Activator RNA detectable by RT-PCR).

Our present data do not prove conclusively that the patient is homozygous for the nonsense mutation, and his pedigree revealed no consanguinity. However, two observations indicate that it is still likely that both of his *GM2A* alleles contain the Glu[54]STOP mutation. First, the AB-variant form of GM2 gangliosidosis is a very rare disease, and all four of the previously characterized patients were also likely homozygous for their distinct mutations (hemizygosity also could not be ruled out in these cases). Second, the patient's family belongs to a deme, the Hmong, a small Laotian hill tribe, and it is unlikely that two *GM2A* mutant alleles would exist in such a small population. For example, the *HEXB* mutations responsible for Sandhoff disease (also a rare disease, but still much more common than the AB-variant) in an Argentinean deme have been characterized. It was concluded that there was only a single novel high-frequency splice-site mutation present in this population (Kleiman et al. 1994).

There has been a great deal of previously reported data concluding that shortened reading frames (i.e., early STOP codons), can lead not only to mRNA instability, but also to the in-frame skipping of the constitutive exon in which the mutation is found (Dietz et al. 1993; Dietz and Kendzior 1994; Ronce et al. 1997; Mazoyer et al. 1998). There remains a controversy as to whether this is caused by a mechanism in the nucleus that can sense the lack of an open reading frame and effect normal splicing (reviewed by Maquat 1996). All but the data we present in figure 9 would appear to support the presence of such a mechanism.

The procedure we used to demonstrate the presence of similar amounts of ΔE2 in both normal and mutant cells (see fig. 9) was based on an experiment reported

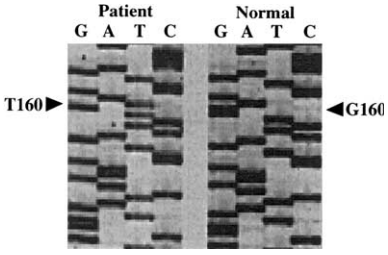



**Figure 8** Direct nucleotide sequence of a PCR fragment (fig. 6*B* and fig. 7 E2) of genomic DNA from the patient and a normal control. Arrows indicate the homozygous G160T mutation in the patient's DNA.

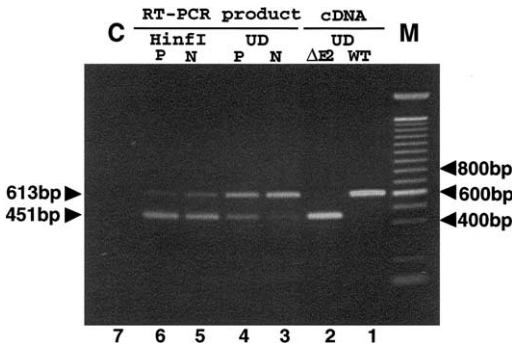



**Figure 9** Nested PCR–amplified RT-PCR product from the patient and a normal control. "ΔE2" and "WT" refer to the cDNAs missing or containing exon 2, respectively. Undigested RT-PCR product (UD) or the RT-PCR products digested with *Hin*fI prior to the second nested PCR procedure, *Hin*fI, are indicated. Cloned WT and ΔE2 cDNA, as well as a $H_2O$ control sample (C) were also subjected to the nested PCR procedure. Size markers (M) are identified with arrows.

by Valentine and Heflich (1997). They examined a homozygous nonsense mutation associated with exon skipping in *hprt* mRNA of Chinese hamster ovary cells and concluded that the apparent increase in exon skipping was an RT-PCR artifact. This artifact was caused by the instability, and thus very low steady-state levels, of the normal-sized RNA containing the nonsense mutation, coupled with the normal stability of the smaller ΔE-RNA, in which the nonsense mutation had been deleted and the reading frame had been restored. This small ΔE-RNA was found to be produced constitutively at very low levels in normal as well as in mutant cells. However, it could be amplified to a level of detectability in normal cells only after the removal of the full-length wild-type RT-PCR–generated cDNA by digestion with a specific restriction enzyme, followed by a second nested PCR. Thus, our data presented in figures 2 and 9 are in full agreement with the conclusion made by Valentine and Heflich and do not support the presence of a nuclear mechanism that can sense the frame of precursor mRNA and affect final splicing events.

## Acknowledgments

We thank David Wenger for facilitating the collaboration between C.C. and D.M., C.-C. Hui for help with RNA-DNA hybridization, Raymond Tellier for his help with long PCR, and Amy Leung for her technical assistance in western blot analysis. This work was supported by a Medical Research Council of Canada grant (to D.M.)

## Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Enzyme Commission, http://expasy.proteome.org.au/enzyme (for Hex A [E.C.3.2.1.52])

GenBank, http://www.ncbi.nlm.nih.gov/Web/Genbank/index.html (for GM2A exons and flanking intronic sequences [AF124717, AF124718, and AF124719])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for AB-variant form of GM2 gangliosidosis [MIM 272750])

## References

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Dietz HC, Kendzior RJ Jr (1994) Maintenance of an open reading frame as an additional level of scrutiny during splice site selection. Nat Genet 8:183–188

Dietz HC, Valle D, Francomano CA, Kendzior RJ Jr, Pyeritz RE, Cutting GR (1993) The skipping of constitutive exons in vivo induced by nonsense mutations. Science 259: 680–683

Gravel RA, Clarke JTR, Kaback MM, Mahuran D, Sandhoff K, Suzuki K (1995) The $G_{M2}$ gangliosidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease. Vol 2. McGraw-Hill, New York, pp 2839–2879

Heng HHQ, Xie B, Shi X-M, Tsui L-C, Mahuran DJ (1993) Refined mapping of the $G_{M2}$ activator protein (GM2A) locus to 5q 31.3-33.1, distal to the spinal muscular atrophy locus. Genomics 18:429–431

Kleiman FE, De Kremer RD, De Ramirez AO, Gravel RA, Argaraña CE (1994) Sandhoff disease in Argentina: high frequency of a splice site mutation in the HEXB gene and correlation between enzyme and DNA-based tests for heterozygote detection. Hum Genet 94:279–282

Klima H, Tanaka A, Schnabel D, Nakano T, Schröder M, Suzuki K, Sandhoff K (1991) Characterization of full-length cDNAs and the gene coding for the human GM2 activator protein. FEBS Lett 289:260–264

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem 193:265–275

Mahuran DJ (1998) The GM2 activator protein: its roles as a co-factor in GM2 hydrolysis and as a general glycolipid transport protein. Biochem Biophys Acta 1393:1–18

Mahuran DJ. Biochemical consequences of mutations causing the GM2 Gangliosidoses. Biochem Biophys Acta (in press)

Maquat LE (1996) Defects in RNA splicing and the consequence of shortened translational reading frames. Am J Hum Genet 59:279–286

Mazoyer S, Puget N, Perrin-Vidoz L, Lynch HT, Serova-Sinilnikova OM, Lenoir GM (1998) A BRCA1 nonsense mutation causes exon skipping. Am J Hum Genet 62:713–715

Nagarajan S, Chen HC, Li SC, Li YT, Lockyer JM (1992) Evidence for two cDNA clones encoding human GM2-activator protein. Biochem J 282:807–813

Neote K, Bapat B, Dumbrille-Ross A, Troxel C, Schuster SM, Mahuran DJ, Gravel RA (1988) Characterization of the hu-

man *HEXB* gene encoding lysosomal β-hexosaminidase. Genomics 3:279–286

Proia RL (1988) Gene encoding the human β-hexosaminidase β-chain: Extensive homology of intron placement in the α- and β-genes. Proc Natl Acad Sci USA 85:1883–1887

Proia RL, Soravia E (1987) Organization of the gene encoding the human β-hexosaminidase a chain. J Biol Chem 262: 5677–5681

Ronce N, Moizard M-P, Robb L, Toutain A, Villard L, Moraine C (1997) A C2055T transition in exon 8 of the ATP7A gene is associated with exon skipping in an occipital horn syndrome family. Am J Hum Genet 61:233–238

Sandhoff K, Harzer K, Fürst W (1995) Sphingolipid activator proteins. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic basis of inherited disease. Vol 2. McGraw-Hill, New York, pp 2427–2441

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Schepers U, Glombitza G, Hoffmann A, Chabas A, Ozand P, Sandhoff K (1996) Molecular analysis of a GM2-activator deficiency in two patients with GM2-gangliosidosis AB variant. Am J Hum Genet 59:1048–1056

Schröder M, Schnabel D, Hurwitz R, Young E, Suzuki K, Sandhoff K (1993) Molecular genetics of GM2-gangliosidosis AB variant: a novel mutation and expression in BHK cells. Hum Genet 92:437–440

Schröder M, Schnabel D, Suzuki K, Sandhoff K (1991) A mutation in the gene of a glycolipid-binding protein (GM2 activator) that causes GM2-gangliosidosis variant AB. FEBS Lett 290:1–3

Swallow DM, Islam I, Fox MF, Povey S, Klima H, Schepers U, Sandhoff K (1993) Regional localization of the gene coding for the GM2 activator protein (GM2A) to chromosome 5q32-33 and confirmation of the assignment of GM2AP to chromosome 3. Ann Hum Genet 57:187–193

Tanaka A, Sakuraba H, Isshiki G, Suzuki K (1993) The major mutation among Japanese patients with infantile Tay-Sachs disease: a G-to-T transversion at the acceptor site of intron 5 of the beta-hexosaminidase alpha gene. Biochem Biophys Res Commun 192:539–546

Valentine CR, Heflich RH (1997) The association of nonsense mutations with exon-skipping in hprt mRNA of Chinese hamster ovary cells results from an artifact of RT-PCR. RNA 3:660–676

Xie B, Kennedy JL, McInnes B, Auger D, Mahuran D (1992*a*) Identification of a processed pseudogene related to the functional gene encoding the GM2 activator protein: localization of the pseudogene to human chromosome 3 and the functional gene to human chromosome 5. Genomics 14:796–798

Xie B, McInnes B, Neote K, Lamhonwah A-M, Mahuran D (1991) Isolation and expression of a full-length cDNA encoding the human GM2 activator protein. Biochem Biophys Res Comm 177:1217–1223

Xie B, Wang W, Mahuran DJ (1992*b*) A $Cys_{138}$-to-Arg substitution in the $G_{M2}$ activator protein is associated with the AB variant form of $G_{M2}$ gangliosidosis. Am J Hum Genet 50:1046–1052

Zokaeem G, Bayleran J, Kaplan P, Hechtman P, Neufeld EF (1987) A shortened β-hexosaminidase α-chain in an Italian patient with infantile Tay-Sachs disease. Am J Hum Genet 40:537–547