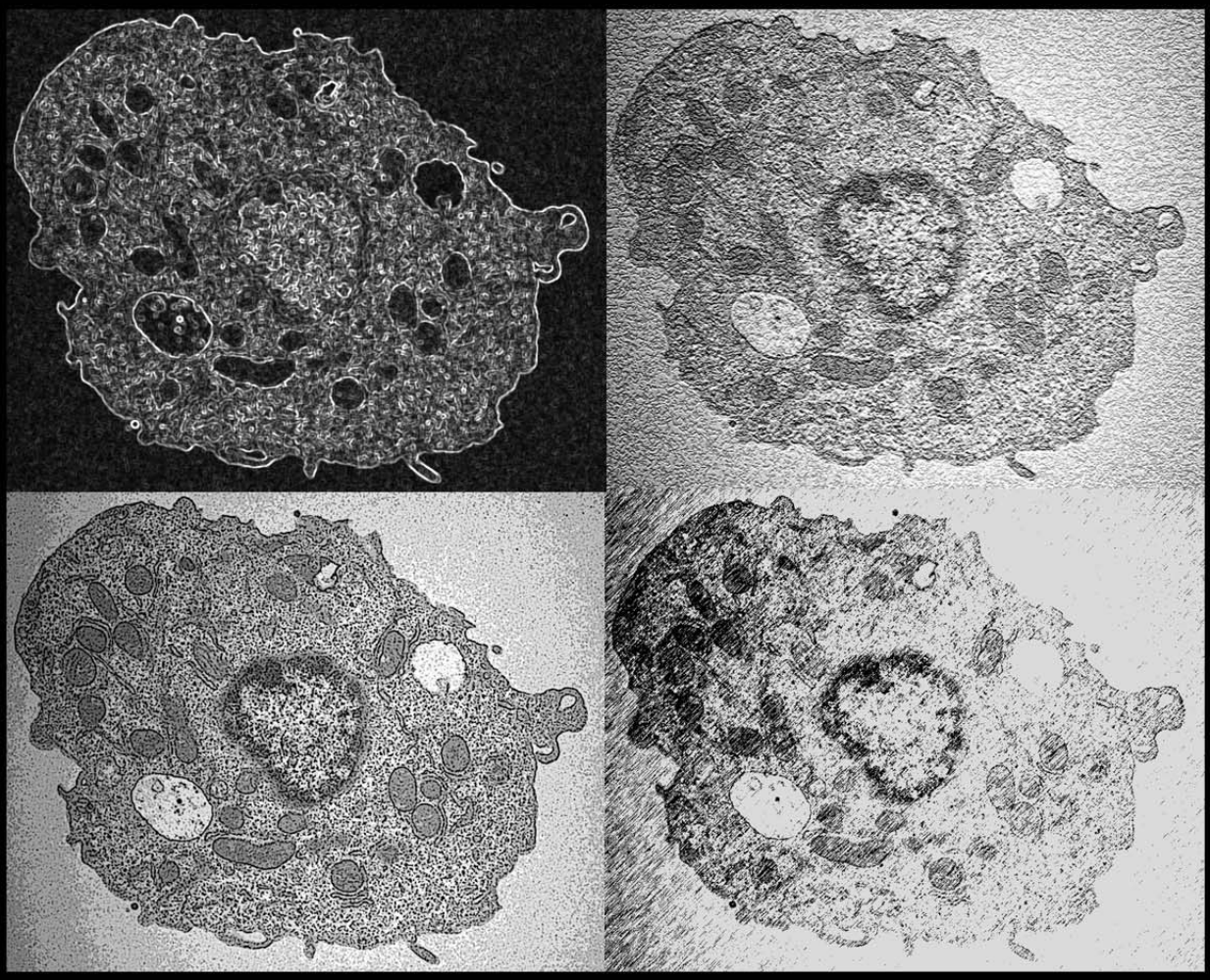
Caracterització de mutacions causants de la malaltia de Gaucher. Aproximació a una teràpia gènica.



Anna Diaz Font

2006

Addenda

Particiació en altres treballs d'investigació durant la realització d'aquesta tesi

Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

New Insights into the Origin of the Gaucher Disease-Causing Mutation N370S: Extended Haplotype Analysis Using the 5GC3.2, 5470 G/A, and ITG6.2 Polymorphisms

Submitted 09/17/01
(Communicated by E. Beutler, M.D., 09/25/01)

Adriana Rodríguez-Marí,[1,2] Anna Díaz-Font,[1,2] Amparo Chabás,[3] Gregory M. Pastores,[4] Daniel Grinberg,[1] and Lluïsa Vilageliu[1]

ABSTRACT: Gaucher disease is a lysosomal storage disorder inherited as an autosomal recessive trait. It is highly prevalent among Ashkenazi Jews but also present in other populations. Mutations in the glucocerebrosidase gene are the main cause of the disorder. One of these gene defects, N370S, is the most prevalent disease allele in the Ashkenazi Jewish patient population and also frequent in others, such as the Spanish and Portuguese Gaucher disease populations. Previous results based on haplotype analysis support the hypothesis of a single origin for this mutation. We have extended the haplotype analysis to include three newly described polymorphisms, 5GC3.2, ITG6.2 (very close to the gene), and 5470 G/A (in intron 7 of the *GBA* gene) in a sample of Spanish and Ashkenazi Jewish patients. The results confirm the single origin of the mutation in these two populations. The 5470A allele is only found in N370S chromosomes and was believed to be limited to the Portuguese population. Here we describe that it is also present with a similar frequency in Spain. Moreover, most of the 5470A alleles are found within particular haplotypes, which have some differences from the common N370S haplotype.

Key Words: Gaucher disease; haplotype analysis; origin of mutation N370S; GBA gene.

INTRODUCTION

Gaucher disease is a recessively inherited lysosomal storage disorder characterized by decreased hydrolysis of glucocerebroside due to deficient activity of glucocerebrosidase (D-glucosyl acylsphingosine glucohydrolase, EC 3.2.1.45). It is mainly due to mutations in the structural gene encoding glucocerebrosidase (*GBA*), mapped to chromosome 1q21. Gaucher disease is particularly frequent in the Ashkenazi Jewish population, with a disease incidence of around 1 in 850 (1). It is also found in other populations, albeit with lower frequency, ranging from 1 in 40,000 to 1 in 60,000 (2). Among Ashkenazi Jewish Gaucher patients, approximately 70% of the disease alleles carry the N370S (1226A→G) mutation (3–5). Approximately two-thirds of the individuals homozygous for this mutation escape detection because of the very mild clinical manifestation. Thus, it is estimated that the frequency of the N370S mutation in the Ashkenazi Jewish population may, in fact, be higher (90% of all Gaucher disease mutations) (6, 7). This mutation is also frequent in other Gaucher disease patient populations, particularly in Spain and Portugal. In Spain it accounts for more than 40% of the mutant alleles (or for as much as 57% if only type I Gaucher disease patients are considered) (8, 9) and among Portuguese patients, it accounts for 53.7% (10).

Twelve polymorphic sites have been detected within the *GBA* gene (11). These markers are all

Correspondence and reprint requests to: Lluïsa Vilageliu. Fax: 34-93-411-0969. E-mail: lluisa@porthos.bio.ub.es.
[1] Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, E-08028 Barcelona, Spain.
[2] These authors contributed equally to this work.
[3] Institut de Bioquímica Clínica, Corporació Sanitària, Mejia Lequerica s/n, Edificio Helios III, 08028 Barcelona, Spain.
[4] Neurogenetics Program, Department of Neurology and Pediatrics, NYU School of Medicine, New York, New York 10016.

1079-9796/01 $35.00

Rodríguez-Marí *et al.* Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

in linkage disequilibrium and constitute only two major haplotypes. The N370S mutation has always been found associated with the "−" haplotype (12), suggesting a single origin for this gene defect.

We mapped the *GBA* gene to a 3.2-cM interval between markers D1S305 and D1S2624 (13), with no recombination detected between the gene and markers D1S2777, D1S2721, and D1S2140. The results for the Ashkenazi Jewish and Spanish Gaucher disease patients in our previous haplotype analysis study support the hypothesis of a single origin for all N370S chromosomes (14, 15). Moreover, similar results have been obtained for other European populations (16).

Winfield *et al.* (17) and Lau *et al.* (18) recently described two new polymorphisms (ITG6.2 and 5GC3.2) in the GBA gene region that could be very useful for a more detailed analysis to establish the origin of the N370S chromosomes. Besides, Amaral *et al.* (19) described an intronic polymorphism (5470G→A) that has only been detected in Portuguese Gaucher disease patients, with the exception of one patient described by Lau *et al.* (18). Here we include these three polymorphisms to extend our earlier haplotype analysis. Moreover, we show that the 5470G→A polymorphism has a similar frequency in Spain than in Portugal.

SUBJECTS AND METHODS

Subjects

Samples from 61 unrelated Ashkenazi Jewish Gaucher disease patients, 43 relatives and 40 Ashkenazi Jewish controls were collected with informed consent by Dr. G. M. Pastores (New York University Medical Center). Similarly, samples from 50 unrelated Spanish Gaucher disease patients, 38 relatives and 48 controls, were collected with informed consent by Dr. A. Chabás (Institut de Bioquímica Clínica, Barcelona, Spain).

DNA Preparation

Genomic DNA was prepared from harvested skin fibroblasts or peripheral blood leukocytes, using a standard method (20).

Analysis of Microsatellites 5GC3.2 and ITG6.2

FAM-fluorescently labeled primers (18) were used to amplify the microsatellites 5GC3.2 and ITG6.2 (17). PCRs were performed in a 25-μl volume, containing 100 ng of DNA, 0.4 μM of each primer, 0.2 mM dNTPs, 1 U *Taq* polymerase (Promega) in 1× PCR buffer (Promega), on a GeneAmp PCR System 9700 thermal cycler (Perkin–Elmer). Conditions were as follows: 5GC3.2: 1 min denaturation at 94°C, and 35 cycles of 94°C for 40 s, 52°C for 30 s, followed by 5 min extension at 72°C; ITG6.2: 1 min denaturation at 94°C, and 35 cycles of 94°C for 40 s, 56°C for 30 s, followed by 5 minute extension at 72°C. Genotyping was performed in an ABI PRISM 310 genetic analyzer from PE Applied Biosystems. A volume of 0.5 μl from the PCR product was mixed with 11.5 μl formamide and 0.5 μl of the ROX-labeled Genescan 400 HD size standard (PE Applied Biosystems), and denatured for 3 min at 94°C. Data were analyzed with a Genescan 3.1 analyzer and Genotyper 2.5 software. Sizes of alleles were adjusted to fit data from that of Lau *et al.* (18).

Analysis of the Polymorphism 5470G→A

The region containing the 5470G>A polymorphism (19) was PCR amplified using the same conditions as described for the ITG6.2 microsatellite, with the following primers: forward, 5′-TGGGCTTCACCCCTGAACAT-3′; and reverse, 5′-TCTCCCGGGTTCAACTGATT-3′. The polymorphism was detected by *Bsu*36I (Promega) digestion (5 U of restriction enzyme, 8 μl of the PCR products in a final volume of 20 μl, at 37°C for 3 h). Digestion products were resolved in 2% agarose gels, and visualized with UV light after ethidium bromide staining.

RESULTS

Analysis of the 5GC3.2 Polymorphism in Ashkenazi Jewish and Spanish Patients

The analysis of the 5GC3.2 polymorphism is shown in Fig. 1. Five different alleles have been

Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

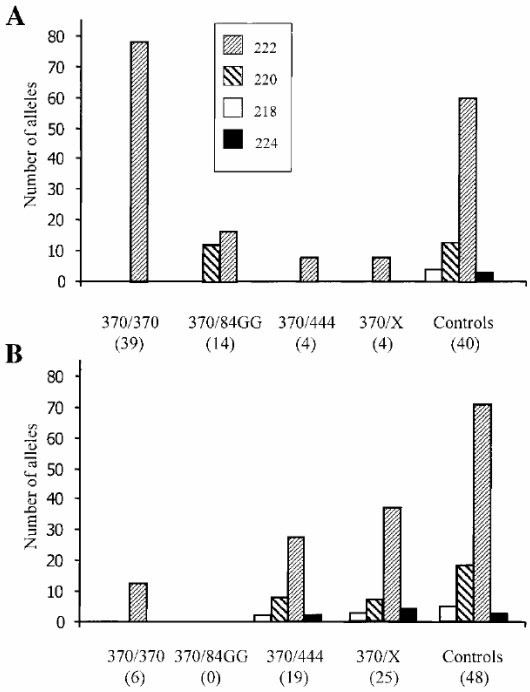


FIG. 1. Frequencies for the different alleles for the 5GC3.2 polymorphism in groups of Gaucher disease patients with the indicated genotypes and controls (number of patients in parentheses). Allele codes are shown within the box. (A) Ashkenazi Jewish individuals. (B) Spanish individuals.

previously described: 218, 220, 222, 224, and 226 (18). We detected four of them. Allele 226 (only described in one individual so far) was not found. In the Ashkenazi Jewish patients, all N370S chromosomes seem to bear the 222 allele for this polymorphism (Fig. 1A). This is evident for the 47 individuals with genotypes N370S/N370S (39), N370S/L444P (4) or N370S/X (4) as they are all homozygous for the 222 allele (X means a mutation other than N370S, L444P, and 84GG). Of the fourteen N370S/84GG individuals, one is homozygous for the 222 allele and the rest have the genotype 220/222, which is consistent with the N370S–222 association. Accordingly, the 84GG mutation seems to be associated with 220 allele at 5GC3.2 in all but one case. In total, the N370S–222 haplotype was confirmed in 87 out of 100 chromosomes. In the rest, the results are consistent with an N370S–222 association, although phase could not be established. Allele frequencies for the different alleles in 40 Ashkenazi Jewish control individuals were 218, 0.05; 220, 0.16; 222, 0.75; and 224, 0.04.

In Spanish Gaucher disease patients (Fig. 1B), all 12 chromosomes of individuals homozygous for the N370S mutation bear the N370S–222 haplotype. In total, 56 N370S chromosomes were studied and phase was established for 43 of them. Forty present the N370S–222 haplotype whereas three bear the 224 allele. These chromosomes also bear the 5470A variant (see below). Allele frequencies in 48 Spanish control individuals were 218, 0.05; 220, 0.20; 222, 0.74; and 224, 0.01.

Analysis of the ITG6.2 Polymorphism in Ashkenazi Jewish and Spanish Patients

We detected the five alleles previously described (314, 318, 322, 326, and 330). In the Ashkenazi Jewish patients, nearly all with the N370S mutation are homozygous for the 318 allele (Fig. 2A). The only exception is one N370S/L444P patient with the 318/322 genotype for the ITG6.2 polymorphism. Thus, the phase N370S-318 was confirmed for all N370S chromosomes but one, which is also consistent with that association. The frequencies in 40 Ashkenazi Jewish control individuals were 0.01, 0.52, 0.36, 0.08, and 0.03 for alleles 314, 318, 322, 326, and 330, respectively.

Among Spanish Gaucher disease patients, all six N370S homozygotes are also homozygous for the 318 allele of the polymorphism. All compound heterozygotes carrying the N370S mutation also carry one 318 allele except for one individual with genotype N370S/L444P who is homozygous for the 322 allele. In 30 of 31 N370S chromosomes for which phase was established, the haplotype N370S-318 was confirmed. The remaining chromosome, as described above, bears the N370S-322 combination. In 48 Spanish controls, the allele frequencies were 0.01, 0.36, 0.55, and 0.08 for alleles 314, 318, 322, and 326, respectively. No allele 330 was detected in Spanish individuals.

Rodríguez-Marí *et al.*

Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

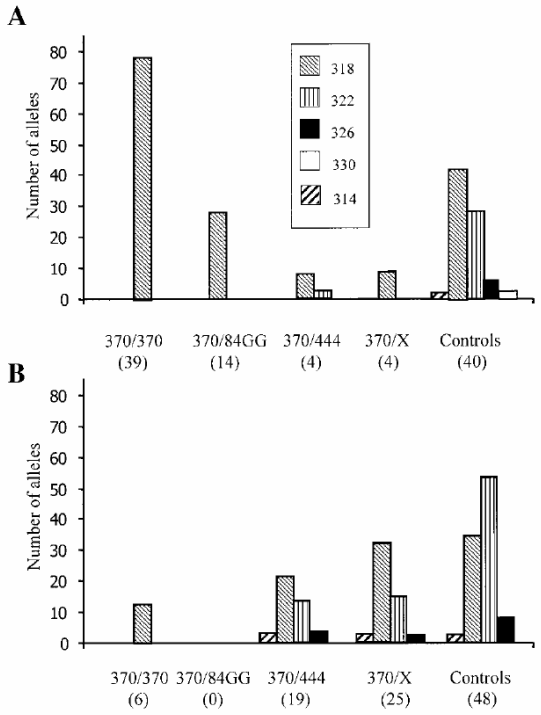


FIG. 2. Frequencies for the different alleles for the ITG6.2 polymorphism in groups of Gaucher disease patients with the indicated genotypes and controls (number of patients in parentheses). Allele codes are shown within the box. (A) Ashkenazi Jewish individuals. (B) Spanish individuals.

Analysis of the 5470 G→A Polymorphism in Ashkenazi Jewish and Spanish Patients

None of the 61 unrelated Ashkenazi Jewish patients analyzed carries the 5470A allele. Twenty-two of them were heterozygous and 39 homozygous for the N370S mutation, which makes a total of 100 N370S chromosomes. The polymorphism was also analyzed in 40 Ashkenazi Jewish control individuals and the 5470A variant was not found either.

In contrast, the 5470A allele was found in six out of 56 N370S chromosomes (10.7%) which correspond to 50 Spanish Gaucher disease patients, six homozygous and 44 heterozygous for the N370S mutation. The 5470A variant was also found once in the 48 Spanish control individuals analyzed. Interestingly, further analysis showed that this individual was a carrier of the N370S mutation. In all cases in which phase was established (4 of 7), the A variant and the N370S mutation are in the same chromosome. This association may also be present in the remaining three cases but it could not be shown.

Refinement of the Genetic Map around the GBA Gene

To perform haplotype analysis markers were ordered according to the last online available version—July 2001—of the human genome sequence draft (21, 22). The detailed map of this region was recently corrected (http://www.ncbi.nih.gov/). The updated map of the GBA gene and close markers is shown on Fig. 3. Note that marker D1S1595, included in our previous work (14), was shown to be the same as marker D1S2140 and so only the latter was kept in the Fig. 3 and used throughout this work. Marker D1S2721 is the only one whose position has not been determined. Transcription orientation (indicated in Fig. 3 by an arrow) was deduced as follows: the GBA sequence is within the working draft sequence segment NT 021889. In that segment, which is 224,394 bp long, several markers are also included. For example, markers stSG15801 and stSG27737 (which are shown in Fig. 3) are located downstream of the GBA gene. The map positions of these markers (distances from pter) are 150,600 and 150,665 kb, respectively. If these data are correct, they are centromeric to GBA, and the transcription orientation is from telomere to centromere.

Haplotype and Linkage Disequilibrium Analyses

We have previously performed haplotype analysis using several polymorphic markers close to the GBA gene (14). We now extend this analysis to include three additional polymorphisms (5GC3.2, 5470G→A, and ITG6.2) as shown in Fig. 4. Markers are ordered according to their map position described above. The map position of marker D1S2721 is still unknown.

The haplotype analysis in Ashkenazi Jewish Gaucher disease patients (Fig. 4A) shows that 61.6% (45/73) of all the N370S chromosomes for

Blood Cells, Molecules, and Diseases (2001) 27(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

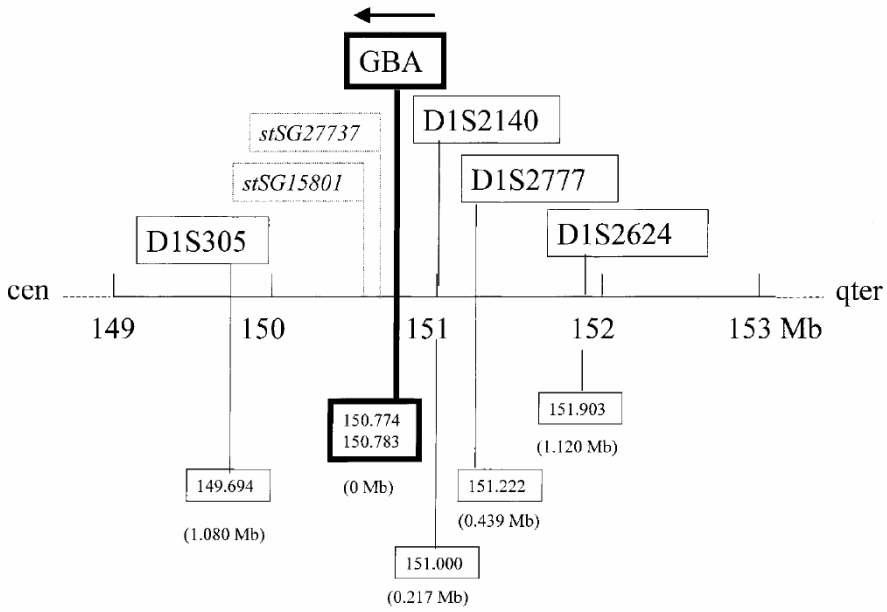


FIG. 3. Map of the 1q21 region containing the GBA gene (wide lines). Microsatellite markers used for haplotype analysis are indicated (narrow lines). STS markers used to determine transcription orientation are shown (dotted lines; see text for details). Map positions are indicated in the lower part of the figure. Map positions are indicated as distances to 1 pter in megabases within the boxes. Distances to the GBA gene in megabases are shown in brackets. The arrow above the GBA gene indicates transcription direction.

which phase could be established for all seven markers, bear the same haplotype. If the distal marker D1S2624 (1.1 Mb apart) is not considered, the consensus haplotype is borne by 97.3% (71/73) of the chromosomes.

Among Spanish Gaucher disease patients (Fig. 4B), the whole haplotype is not conserved; only 2 of 16 N370S chromosomes for which phase could be established for all seven markers bear the haplotype found in Ashkenazi Jewish Gaucher disease patients. However, if the distal marker D1S2624 is not taken into account, 11 of 16 chromosomes (68.7%) bear the same "inner" haplotype, also found in more than 97% of the Ashkenazi Jewish N370S chromosomes. The marker D1S2721 present a different allele in four out of the 5 chromosomes with variants in the inner haplotype, while D1S2140 present a non-consensus allele in the remaining chromosome and D1S2777 present the same allele in all the chromosomes.

Regarding the two chromosomes bearing the 5470A allele, both bear allele 7 for D1S2721 (instead of allele 3 present in the consensus haplotype). Moreover one of the two bears the allele 224 for the 5GC3.2 polymorphism, instead of allele 222 present in all other chromosomes. There are other chromosomes that carry the 5470A allele in phase with the allele 224 for the 5GC3.2 polymorphism (see below) but were not included in Fig. 4 because phase was not established for the other markers.

Alleles 222 of marker 5GC3.2 and 318 of marker ITG6.2 are associated with the N370S mutation in the Ashkenazi Jewish Gaucher disease patient population. Linkage disequilibrium values are highly significant ($P < 0.001$ for all the markers). Among Spanish Gaucher disease patients, linkage disequilibrium values are also significant but the P value is lower for the 222 allele of marker 5GC3.2 ($P = 0.018$). Allele 5470A, only present in the Spanish Gaucher disease population, is also associated with the N370S mutation ($P < 0.001$).

Rodríguez-Marí *et al.* Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

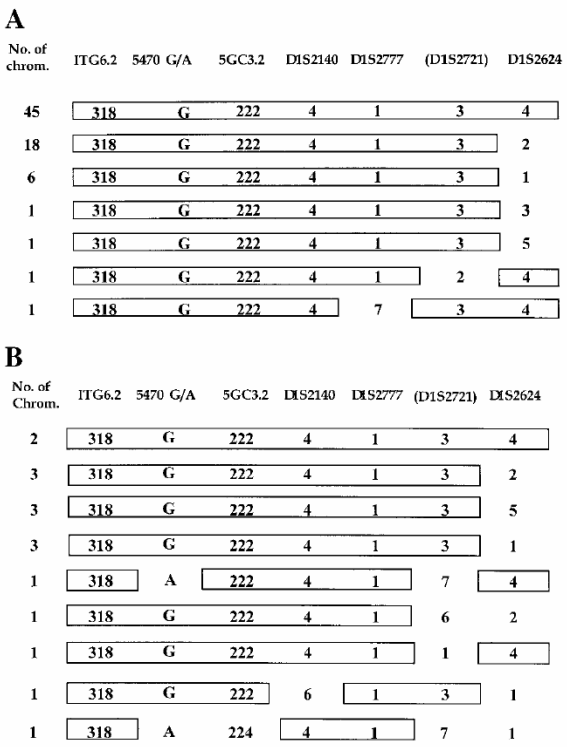


FIG. 4. Haplotype analysis for 1q21 markers. Chromosomes from Ashkenazi Jewish (A) and Spanish (B) Gaucher disease patients bearing the N370S mutation were analyzed for markers ITG6.2, 5470G/A, 5GC3.2, D1S2140, D1S2777, D1S2721, and D1S2624. The conserved parts of the haplotypes are boxed.

Genotypes for Different Polymorphisms in Spanish Individuals Carrying the 5470A Allele

Table 1 shows the genotypes for the GBA mutations and for markers 5GC3.2, ITG6.2, and D1S2140 of the six Spanish Gaucher disease patients and a control individual (see above) carrying the 5470A allele and the N370S mutation. Patients I.6, I.13, I.19 and I.29 were described elsewhere (9). For patients I.6 and I.19, phases have been established and their haplotypes have been included in Fig. 4.

As previously stated, most of the N370S chromosomes from Spanish patients bear the 222 allele for the 5GC3 polymorphism, whereas the 224 allele is only present in three out of the 43 chromosomes for which phase was established (7%). However, when considering the seven individuals bearing the 5470A allele, a very different result is obtained. Three of them—I.19, I.49, and I.53—have a confirmed N370S–224 haplotype (these are the 3 of 43 mentioned above). Three others—I.13, I.29, and the control individual—for whom phase was not established, also bear the 224 allele. The remaining case has the most prevalent N370S–222 phase (I.6). The allele frequency of the 224 allele in Spanish Gaucher disease patients bearing the N370S mutation is 5/100 (5.0%). However, for those carrying the 5470A variant, the 224 allele frequency is 5/12 (41.7%). Regarding N370S chromosomes for which phase is known, the frequency of the N370S–224 haplotype rises from 7% (3/43) to 75% (3/4) if all N370S chromosomes or only those bearing the 5470A allele are considered.

For the ITG6.2 and D1S2140 markers, the 5470A allele seems to present the same strong association with the 318 and 4 alleles, respectively, as the rest of the N370S chromosomes.

TABLE 1

Genotypes for Different Polymorphisms in Spanish Individuals Carrying the 5470A Allele

Individuals[a]	Genotypes	5470G → A[b]	5GC3.2[b]	ITG6.2[b]	D1S2140[b]
I.6	**N370S**/Y313H	**A**/G	**222**/220	**318**/322	**4**/7
I.13	**N370S**/RecNciI	*A/G*	*224/222*	*318/326*	**4**/4
I.19	**N370S**/L444P	**A**/G	**224**/220	**318**/322	**4**/3
I.29	**N370S**/G113E	*A/G*	*224/222*	*318/322*	*4/7*
I.49	**N370S**/?	**A**/G	**224**/220	*318/322*	**4**/4
I.53	**N370S**/?	**A**/G	**224**/222	**318**/322	**4**/4
C	**N370S**/wt	*A/G*	*224/218*	*318/322*	**4**/4

[a] I, type I Gaucher disease patients; C, control individual.
[b] Genotypes in italics mean that phase could not be established. Alleles in bold are in phase with the N370S mutation.

Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

DISCUSSION

We previously showed a common origin for N370S chromosomes in Ashkenazi Jewish and Spanish Gaucher disease patients by haplotype analyses using several microsatellite markers (14). In the present study we have extended this analysis to include three newly described polymorphisms, two closely flanking the gene and the third within intron 7 of the GBA gene.

Marker 5GC3.2 was described by Winfield *et al.* (17) and by Lau *et al.* (18). In our sample of Ashkenazi Jewish Gaucher disease patients, we found that the N370S mutation was associated with allele 222 of marker 5GC3.2 in all of the cases in which phase were established (87 of the 100 chromosomes). For the 13 Gaucher disease patients with unknown phase, all of whom had the N370S/84GG genotype (the remaining individual is homozygous for the 222 allele), the results are consistent with an N370S–222 haplotype. For mutation 84GG, all but one of our Ashkenazi Jewish patients bear the N370S/84GG genotype for the GBA gene and 220/222 for the marker. So the most probable phases would be N370S–222 and 84GG–220. These results are similar to those of Lau *et al.* (18) who also found the N370S–222 and 84GG–220 associations. However, we found one patient who is N370S/84GG but homozygous for the 222 allele. This is the first case of an 84GG–222 haplotype. It shows that the 84GG mutation is not always found in the same genomic context. In the Spanish patients we also found the N370S–222 haplotype as the main one, but we found at least three chromosomes that carry the N370S mutation together with the 224 allele of the 5GC3.2 marker. This allele has a population frequency as low as 0.01. This is the first time in which mutation N370S was not found in association with the 222 allele. Interestingly, all the 224 chromosomes also carry the 5470A variant (see below).

The second polymorphism analyzed was ITG6.2 (18). Mutation N370S was associated with the 318 allele of this polymorphism in all the known-phase Ashkenazi Jewish Gaucher disease patients and in all but one known-phase Spanish Gaucher disease patients. The only exception is a Spanish patient who bears the N370S mutation together with allele 322. Lau *et al.* (18) also found the N370S-318 association except for one Ashkenazi Jewish patient, homozygous for N370S who bears the 318/322 genotype for the marker.

Regarding population frequencies, the 5GC3.2 alleles present no differences between Ashkenazi Jewish and Spanish control populations. Our data are similar to those of Lau *et al.* (18). However, for marker ITG6.2 we found a marked difference between allele frequencies in Ashkenazi Jewish and Spanish control individuals. The 318 allele is the most prevalent among Ashkenazi Jewish individuals (52.5%) while allele 322 accounts for 36%, whereas the most frequent allele in the Spanish population is 322 (55%) and the 318 allele is the second one (35%).

The 5470 G→A polymorphism was identified by Amaral *et al.* (19) who detected it only in chromosomes carrying the N370S mutation (frequency of 0.11). Because of the negative results obtained in a limited search in other populations, they suggested that the polymorphism could be confined to Portuguese Gaucher disease patients. In their study, eight Spanish N370S chromosomes were included. Lau *et al.* (18) reported the first finding of a 5470A allele outside Portugal, in an Ashkenazi Jewish Gaucher disease patient homozygous for N370S. In our study we found this variant in 6 of 56 N370S chromosomes of Spanish Gaucher disease patients. Interestingly we also found the 5470A allele in a control individual who, after further analysis, was shown to be an N370S carrier. This makes a total of seven 5470A alleles out of 57 N370S chromosomes, representing a frequency of 0.12, similar to that found in the Portuguese population. So, unlike the previous hypothesis (19), the 5470 G→A polymorphism does not appear to be confined to Portugal, but, at least, extends to the Iberian Peninsula. In relation to the Ashkenazi Jewish patients, our results agree with previous data (18, 19), in the sense that the variant seems to be absent or very rare in this group of patients, as it was not detected in a sample of 100 N370S Ashkenazi Jewish chromosomes.

Figure 5 shows a possible pathway for the generation of the different haplotypes found in the

Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

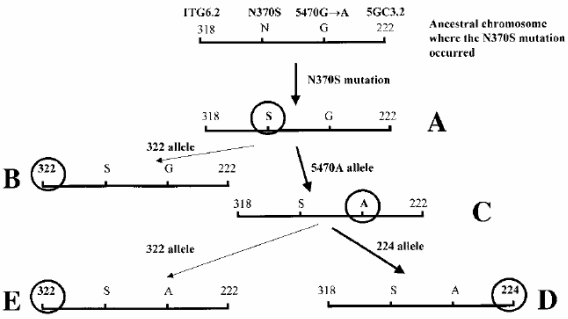


FIG. 5. Putative pathway for the generation of the different haplotypes found in N370S chromosomes. Haplotype (A) was found in most N370S chromosomes. Haplotypes (B) and (C) were found once each in our series of Spanish Gaucher disease patients. Haplotype (D) was confirmed in three chromosomes belonging to Spanish patients. Haplotype (E) was described by Lau *et al.* (18).

chromosomes that carry the 5470A allele together with the N370S mutation. It seems quite clear that the ancestral chromosome in which mutation N370S occurred bore the 318 and 222 alleles for polymorphisms ITG6.2 and 5GC3.2, respectively, since nearly all N370S chromosomes have these two alleles. The 5470A allele should have arisen on a chromosome 318/370S/222. Two independent events could have generated two haplotypes derived from that chromosome: the change of allele 318 to 322 at ITG6.2 and the change of allele 222 to 224 at 5GC3.2. Haplotype (A) corresponds to most of the N370S chromosomes, which do not bear the 5470A allele. Haplotype (B) and haplotype (C) were only found in one Spanish Gaucher disease patient each. Haplotype (D) is the one that was detected in three of the Spanish Gaucher disease patients. The compound genotypes for the remaining three individuals bearing the 5470A allele (two Gaucher disease patients and a control individual) are consistent with the presence of haplotype (D) as well. An alternative path could have taken place if the change 222→224 at 5GC3.2 occurred before the G→A transition at 5470 position. If this is so, haplotype (C) should have derived from haplotype (D) as a revertant at the 5GC3.2 site. Finally, haplotype (E) was only found once by Lau *et al.* (18).

Previous data on the N370S origin (14, 15) from our group and others (16), suggest that the mutation occurred first in non-Jewish and then passed into the Jewish population. Linkage disequilibrium data presented here are in agreement with this hypothesis: linkage disequilibrium values are higher in Ashkenazi Jewish than in Spanish Gaucher disease chromosomes, indicating a more recent introduction of the mutation in the Ashkenazi Jewish population. Haplotype (A) appears to have been passed on from non-Jewish to Jewish individuals, and is now present in the great majority of the Ashkenazi Jewish Gaucher disease chromosomes studied so far. Haplotype (E) is quite unusual because it was described in an Ashkenazi Jewish patient, being the only Ashkenazi Jewish chromosome bearing the 5470A out of nearly 300 N370S chromosomes (18, 19, this study). Besides, it is the only N370S chromosome bearing the ITG6.2-322 allele. Lau *et al.* suggested that this chromosome could have a non-Jewish origin. This would fit better in the pathway shown in Fig. 5: haplotypes (B) and (E) could have arisen by two independent mutations at ITG6.2.

It should be very interesting to analyze the haplotype of the Portuguese patients who bear the 5470A, particularly to see whether they also bear the 224 allele at 5GC3.2. This would validate the pathway presented here. Moreover, we have shown that the 5470G/A polymorphism is not confined to Portugal. Genotyping this variant in other population would determine whether its distribution is limited to the Iberian Peninsula.

ACKNOWLEDGMENTS

We thank Dr. Bru Cormand, Dr. Laura Gort, and Dr. Marta Pascual for their help and suggestions and Robin Rycroft for revising the English. We also thank Serveis Científico-Tècnics, Universitat de Barcelona, for automated sequencing resources. This study was supported by CICYT (SAF 97-0074 and SAF 2000-0200), by Fundació Marató de TV3 (Project Nos. 98-1220 and 1221), and by Genzyme.

REFERENCES

1. Beutler, E., and Grabowski, G. A. (2001) Gaucher disease. *In* The Metabolic and Molecular Bases of

Blood Cells, Molecules, and Diseases (2001) 27(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.), pp. 3635–3668. McGraw-Hill, New York.

2. Grabowski, G. A. (1993) Gaucher Disease—Enzymology, Genetics, and Treatment. *In* Advances in Human Genetics (Harris, H., and Hirschhorn, K., Eds.), pp. 377–441. Plenum Press, New York.
3. Beutler, E., Gelbart, T., Kuhl, W., Zimran, A., and West, C. (1992) Mutations in Jewish patients with Gaucher disease. *Blood* **79,** 1662–1666.
4. Horowitz, M., Tzuri, G., Eyal, N., Berebi, A., Kolodny, E. H., Brady, R. O., Barton, N. W., Abrahamov, A., and Zimran, A. (1993) Prevalence of 9 mutations among Jewish and non-Jewish Gaucher disease patients. *Am. J. Hum. Genet.* **53,** 921–930.
5. Sibille, A., Eng, C. M., Kim, S. J., Pastores, G., and Grabowski, G. A. (1993) Phenotype/genotype correlations in Gaucher disease type-I—Clinical and therapeutic implications. *Am. J. Hum. Genet.* **52,** 1094–1101.
6. Beutler, E., Nguyen, N. J., Henneberger, M. W., Smolec, J. M., Mcpherson, R. A., West, C., and Gelbart, T. (1993) Gaucher Disease—Gene frequencies in the Ashkenazi Jewish population. *Am. J. Hum. Genet.* **52,** 85–88.
7. Grabowski, G. A. (1997) Gaucher's disease: Gene frequencies and genotype/phenotype correlations. *Genet. Test.* **1,** 5–12.
8. Cormand, B., Vilageliu, L., Burguera, J. M., Balcells, S., González-Duarte, R., Grinberg, D., and Chabás, A. (1995) Gaucher disease in Spanish patients: Analysis of eight mutations. *Hum. Mutat.* **5,** 303–309.
9. Cormand, B., Grinberg, D., Gort, L., Chabás, A., and Vilageliu, L. (1998) Molecular analysis and clinical findings in the Spanish Gaucher disease population: Putative haplotype of the N370S ancestral chromosome. *Hum. Mutat.* **11,** 295–305.
10. Amaral, O., Pinto, E., Fortuna, M., Lacerda, L., and Miranda, M. (1996) Type 1 Gaucher disease: Identification of N396T and prevalence of glucocerebrosidase mutations in the Portuguese. *Hum. Mutat.* **8,** 280–281.
11. Beutler, E., West, C., and Gelbart, T. (1992) Polymorphisms in the human glucocerebrosidase gene. *Genomics* **12,** 795–800.
12. Zimran, A., Gelbart, T., and Beutler, E. (1990) Linkage of the *Pvu*II polymorphism with the common Jewish mutation for Gaucher disease. *Am. J. Hum. Genet.* **46,** 902–905.
13. Cormand, B., Montfort, M., Chabás, A., Vilageliu, L., and Grinberg, D. (1997) Genetic fine localization of the beta glucocerebrosidase (GBA) and prosaposin (PSAP) genes: Implications for Gaucher disease. *Hum. Genet.* **100,** 75–79.
14. Díaz, A., Montfort, M., Cormand, B., Zeng, B. J., Pastores, G. M., Chabás, A., Vilageliu, L., and Grinberg, D. (1999) Gaucher disease: The N370S mutation in Ashkenazi Jewish and Spanish patients has a common origin and arose several thousand years ago. *Am. J. Hum. Genet.* **64,** 1233–1238.
15. Díaz, A., Montfort, M., Cormand, B., Zeng, B. J., Pastores, G. M., Chabás, A., Vilageliu, L., and Grinberg, D. (2000) On the age of the most prevalent Gaucher disease-causing mutation, N370S. *Am. J. Hum. Genet.* **66,** 2014–2015.
16. Díaz, G. A., Gelb, B. D., Risch, N., Nygaard, T. G., Frisch, A., Cohen, I. J., Miranda, C. S., Amaral, O., Maire, I., Poenaru, L., Caillaud, C., Weizberg, M., Mistry, P., and Desnick, R. J. (2000) Gaucher disease: The origins of the Ashkenazi Jewish N370S and 84GG acid beta-glucosidase mutations *Am. J. Hum. Genet.* **66,** 1821–1832.
17. Winfield, S. L., Tayebi, N., Martin, B. M., Ginns, E. I., and Sidransky, E. (1997) Identification of three additional genes contiguous to the glucocerebrosidase locus on chromosome 1q21: Implications for Gaucher disease. *Genome Res.* **7,** 1020–1026.
18. Lau, E. K., Tayebi, N., Ingraham, L. J., Winfield, S. L., Koprivica, V., Stone, D. L., Zimran, A., Ginns, E. I., and Sidransky, E. (1999) Two novel polymorphic sequences in the glucocerebrosidase gene region enhance mutational screening and founder effect studies of patients with Gaucher disease. *Hum. Genet.* **104,** 293–300.
19. Amaral, O., Marcao, A., Pinto, E., Zimran, A., and Sa Miranda, M. C. (1997) Distinct haplotype in non-Ashkenazi Gaucher patients with N370S mutation. *Blood Cells Mol. Dis.* **23,** 415–416.
20. Miller, S. A., Dyke, D. D., and Polesky, H. F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **16,** 1215.
21. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins,

Blood Cells, Molecules, and Diseases (2001) **27**(5) Sept/Oct: 950–959
doi:10.1006/bcmd.2001.0468, available online at http://www.idealibrary.com on IDEAL®

M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., and Yan, C. (2001) The sequence of the human genome. *Science* **291,** 1304–1351.

22. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., and Frazier, M. (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921.

Research Signpost
37/661 (2), Fort P.O., Trivandrum-695 023, Kerala, India

Recent Res. Devel. Human Genet., 1 (2002): 223-235 ISBN: 81-7736-107-4

Molecular basis of Gaucher disease

Daniel Grinberg[1], Amparo Chabás[2], Bru Cormand[1], Laura Gort[2]
Magda Montfort[1], Anna Díaz-Font[1] and Lluïsa Vilageliu[1]
[1]Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645
08028 Barcelona, Spain; [2]Institut de Bioquímica Clínica, Corporació Sanitària, Mejia
Lequerica s/n, Edificio Helios III, 08028 Barcelona, Spain

Abstract

Gaucher disease is the most prevalent sphingolipidosis. It is due mainly to mutations in the GBA gene. Here, we describe the work carried out in our laboratory and discuss general features on the molecular basis of the disease. Mutations and genotype-phenotype correlation were analysed in several populations. The fine mapping of the GBA and PSAP genes and studies on the origin of the most frequent mutation, N370S, are also reported.

1. Introduction

Gaucher disease is a storage disorder due to lysosomal dysfunction. Lysosomal diseases are classified according to the metabolic pathway affected and to the nature of

Correspondence/Reprint request: Dr. Daniel Grinberg, Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, E-08028 Barcelona, Spain. E-mail: danielr@porthos.bio.ub.es

the substrate that accumulates. In Gaucher disease, the deficiency affects sphingolipid catabolism. Thus, it is classed as a sphingolipidosis. It is, in fact, the most frequent sphingolipidosis.

The main Gaucher disease symptoms are haematologic abnormalities, enlargement of liver and spleen (hepatosplenomegaly), progressive substitution of bone marrow by macrophages loaded with lipids (Gaucher cells), osteolytic degeneration of the skeleton that can cause bone crises and, in a few patients, gradual degeneration of the central nervous system (for a review see Ref. 1).

The clinic manifestations of Gaucher disease are very heterogeneous. Three clinical subtypes have been defined according to the absence (type I) or presence and severity (types II and III) of neurological symptoms [1].

Type I, the non-neuronopathic form (MIM 230800), is the most frequent: between 1/40,000 and 1/60,000, in the general population, and between 1/400 and 1/600 in Ashkenazi Jews. There is no central nervous system affection and the age of onset and the severity of the disease are very variable.

Type II, the acute neuronopathic form (MIM 230900), presents an incidence of 1/100,000, without ethnic differences. It is the most severe form because of the rapid neurovisceral accumulation of lipids which causes death within the first two years of life. Oculomotor abnormalities are often the first neurological symptoms.

Type III, the subacute neuronopathic form (MIM 231000), has intermediate severity compared with types I and II. Its progression is slower and the neurological symptoms appear later, which allows the patient to reach adulthood. Its incidence lies between 1/50,000 and 1/100,000 and, although there are no ethnic differences, it is particularly frequent in Norbottnia, in the North of Sweden.

Gaucher disease is inherited as an autosomal recessive trait and it is mostly due to a deficiency in glucocerebrosidase (EC 3.2.1.45) [2]. This enzyme is a lysosomal acid hydrolase that catalyses the hydrolysis of glucosylceramide to ceramide and glucose, in the presence of an activator protein named saposin C or SAP-C. The impaired activity of glucocerebrosidase leads to the accumulation of glucosylceramide (or glucocerebroside) within the lysosomes of the macrophages in the reticuloendothelial system, mainly in the spleen, the liver and the bone marrow.

Glucocerebrosidase is a 497-amino acid long membrane glycoprotein of 65 kDa. The gene that codes for this enzyme (*GBA*) has been cloned and localized to the long arm of chromosome 1, at 1q21 [3]. It is split in eleven exons and ten introns and it spans approximately 7 kb. Introns 2, 4, 6 and 7 contain *Alu* sequences. A highly homologous pseudogene has been found 16 kb downstream from the functional gene. It is transcribed but it is not translated to protein. At the DNA level, it shows a 96% identity with the gene [4]. The pseudogene has a structure similar to that of the gene but it lacks the *Alu* sequences and has several point mutations and deletions that preclude the synthesis of a functional protein.

Thirteen polymorphisms have been described in the *GBA* gene. Twelve of them are in non-coding regions. They are in strong linkage disequilibrium among them and give rise to two major haplotypes [5].

In most cases, Gaucher disease is due to mutations in the *GBA* gene. Only in two cases the defect was found in the gene coding for the activator protein SAP-C [6,7], located in the chromosome region 10q21-22.

More than two hundred mutations causing Gaucher disease have been described so far. They include missense and nonsense point mutations, insertions, deletions, splicing mutations and gene-pseudogene rearrangements [8]. These rearrangements generate complex alleles that bear several mutations as a result of unequal crossover or gene conversion events.

Few *GBA* mutations are frequent. The relative frequency varies from one population to another. Among Ashkenazi Jews, the N370S mutation, a substitution of a serine for asparagine at amino acid 370 of the protein, is the most prevalent and accounts for 68-77% of all mutant alleles. The second most frequent mutation is the insertion of a guanine at position 84 of the cDNA (84GG), accounting for 10-13%. Only four mutations (N370S, 84GG, L444P and IVS2+1) represent 89-96% of the molecular defects in this population [9].

The most prevalent mutations in non-Jewish population of European origin are the substitutions N370S (23-54%) and L444P (13-57%) [10-22].

Some of the mutations and the phenotypic expression of the disease are significantly correlated, which proves very useful for the genetic counselling and the therapy. The presence of mutation N370S, either in homozygosis or heterozygosis, precludes neuropathologic symptoms and, -thus-, it is always associated with diagnosis of type I, non-neuronopathic, Gaucher disease. In contrast, the allele L444P (in the absence of N370S) correlates with the neuronopathic forms of the disease [23,24]. The D409H/D409H genotype is associated with a phenotype characterized by cardiovascular calcifications and oculomotor apraxia [25,26].

Several therapeutic strategies have been developed for Gaucher disease, owing to the high prevalence of a non-neuronopathic form. Bone marrow transplantation was replaced by enzyme replacement therapy, which consists of infusion of a modified glucocerebrosidase (ceredase) specifically targeted to macrophages [27]. This therapeutic approach began in 1991 and, nowadays, it is applied to more than two thousands patients worldwide. Although many of the symptoms improve in most of these patients, the treatment is very expensive and makes patients infusion-dependent for life. Gene therapy is a new strategy for the treatment of Gaucher disease, since the cells mainly involved are macrophages that derive from haematopoietic stem cells. Clinical trials are being performed following approved protocols [28]. However, no clear result has been obtained so far.

In this article, we describe some of the work that we have carried out on Gaucher disease. We performed mutational analyses and studies to establish genotype-phenotype correlations, in Spanish patients. Moreover, these analyses were extended to some Argentinean and Italian patients. We also mapped the *GBA* and *PSAP* genes, which allowed us to design a molecular diagnosis strategy and to perform studies on the origin of the most frequent mutation, N370S. The structure and origin of some complex alleles was also analysed. Some of the mutant alleles were expressed in a baculovirus system to characterise the mutant proteins.

2. Mutation detection

When our study started, no data on the mutations responsible for Gaucher disease in the Spanish population were available and information about other non-Jewish

populations was very scarce. Most reports described a few individuals of non-specific origin or were limited to the analysis of a small group of mutations. The assessment of mutation prevalences in a particular population is a useful tool to set up a molecular diagnosis protocol for the disease.

We initially analysed Spanish Gaucher disease patients for prevalent mutations described for other populations [15], by enzyme digestion or oligonucleotide specific allele hybridization (ASOH) after PCR amplification. Thereafter, an exhaustive mutational analysis of the *GBA* gene was conducted on DNA from 53 Spanish patients. We performed PCR amplification followed by single strand conformation polymorphism (SSCP) of 14 DNA fragments covering the coding region, the flanking intronic junctions and part of the promoter region of the *GBA* gene. We thus identified 101 of the 106 mutated alleles, which account for 95% of the studied chromosomes (Fig. 1).

In total, 24 different mutations were found (16 missense, 3 non-sense, 1 insertion, 2 deletions, 1 complex allele and 1 splicing mutation). Sixteen of the 24 mutations had not been previously described [29-31]. Most of the missense mutations are grouped on exons 5-6 and 8-10 (Fig. 2) in agreement with previous data showing that these exons code for functionally relevant regions of the protein [32].

Many other populations, such as the English [14], the Portuguese [17], the Australasian [33-34], the Greek [16], the Polish [18], the German [20] and the French [19] have been studied. We compared the mutation frequencies in the Spanish population with those in other populations, which revealed two facts: First, the most prevalent mutation in the Spanish population is the N370S (43.4%), as in Ashkenazi Jews (and also in Portugal), but not in other populations. For instance, it accounts for only 27% of the muations in England and 25% in Australasia. Thus, we studied the putative common origin of this mutation in the Spanish and Jewish populations (see below). Second, the third most prevalent mutation in Spain is the D409H, which is much less frequent in other populations.

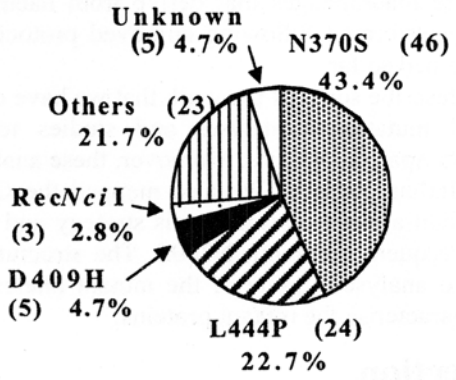


Figure 1. Relative frequencies of the *GBA* mutations in Spanish Gaucher disease patients. Number of alleles are shown in parentheses.

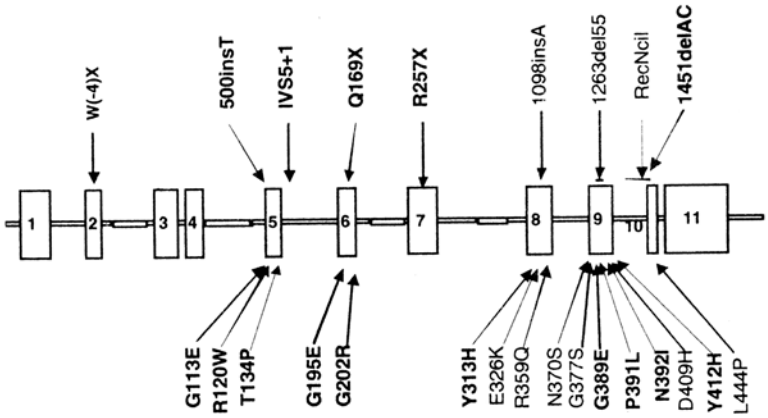


Figure 2. Scheme of the genomic organization of the *GBA* gene showing the position of the different mutations found in Spanish patients.

Missense mutations are depicted below the gene and the rest of them above. New mutations are indicated in bold.

We also performed the mutational analysis of a Gaucher disease population from Argentina [35]. Using the above-described techniques, 31 unrelated patients were analysed and the mutation responsible for the disease was identified in 58 out of the 62 mutated alleles (94%).

Mutation N370S is also the most frequent in this population (46.4%) but, it is followed by the recombinant allele Rec*Nci*I (21%) (Fig. 3). This recombinant allele, which is generated by gene-pseudogene interchanges, was studied in detail as described below. In total, 13 different mutations, three of which are novel, were identified in the Argentinean patients.

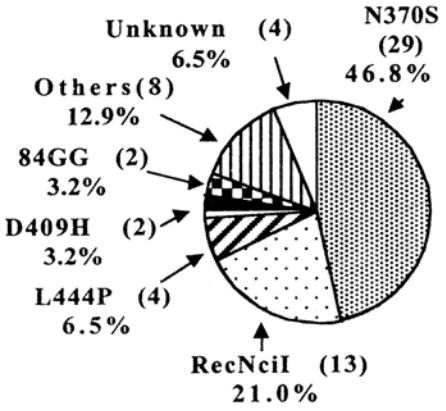


Figure 3. Relative frequencies of the *GBA* mutations in Argentinian Gaucher disease patients. Number of alleles are shown in parentheses.

3. Genotype-phenotype correlations

The assessment of the mutations responsible for the disease in each patient is required for prenatal diagnosis, genetic counselling and for treatment. However, in many cases there are no clear genotype-phenotype correlations. In Gaucher disease, there is only one well established correlation: mutation N370S, in one or both alleles, is always associated with type I, non-neuronopathic, disease. Beutler et al. [36] classified the mutations in the *GBA* gene as mild, severe or lethal alterations. A mild mutation involves no neurological damage, (the main example is the N370S change). A mutation is considered to be severe when it results in neuronopathy, and lethal when it has never been found in homozygosity and no functional protein is produced. Nearly all type I patients bear the N370S mutation. Thus, it is difficult to identify new mild mutations.

We identified two new mild mutations in two Italian patients [37]. Both individuals were type I patients and homozygous for mutations not previously described, I402T and V375L, respectively.

Regarding genotype-phenotype correlations, we would like to highlight the case of three sisters with genotype D409H/D409H, who presented a particular phenotype with cardiovascular calcifications and moderate visceromegaly [25]. Similar symptoms in patients bearing the D409H mutation in homozygosis have been described elsewhere, supporting this newly described genotype-phenotype correlation [26,38,39]. In 1998, we described a new patient homozygous for the D409H mutation, who started enzyme replacement therapy at two months of age [40] and has neither neurological nor cardiovascular symptoms so far. Recently, new cases with this association have been reported in several populations.

4. Mapping of the genes *GBA* and *PSAP*

The *GBA* gene was localised in chromosome 1q21 by in *situ hybridisation* in 1985 [3]. In 1997, we undertook the genetic fine localisation of this gene in relation to highly polymorphic microsatellite markers. Two point and multipoint linkage analyses between an intragenic polymorphism and several microsatellite markers on the short arm of chromosome 1 were performed [41]. Three of these markers (D1S2140, D1S2777 and D1S2721) were found to be at 0 cM from the *GBA* gene, while D1S305 and D1S2624 were at 1.6 cM each, at both sides of the gene. Fig. 4 shows the genetic (cM) and physical (Mb) distances deduced from the Human Genome Map [42,43]. Note that marker D1S2721 is not on Fig. 4, since its physical position is unknown.

We also mapped the *PSAP* gene, which encodes prosaposin, the precursor of Saposin C (and three other saposins), which is responsible for Gaucher disease in a few cases.

The fine genetic mapping of the *GBA* gene allowed us to set up a molecular indirect diagnosis. Moreover, we identified the putative ancestral haplotype in which mutation N370S occurred and investigated the possible common origin of this mutation in the Spanish and Jewish populations. These two issues are discussed below.

5. Indirect molecular diagnosis

The main drawback of mutation analysis for prenatal diagnosis and carrier detection in Gaucher disease is that rare mutations account for more than 25 % of the mutant alleles

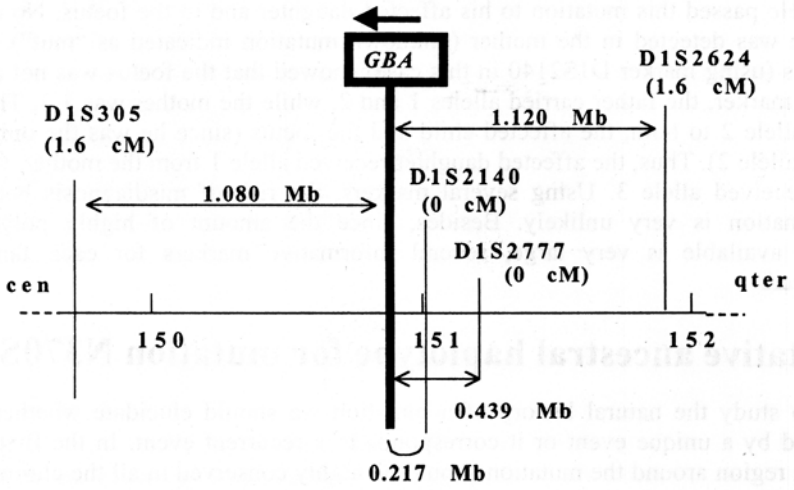


Figure 4. Map of the 1q21 region containing the GBA gene (wide lines). Microsatellite markers close to the gene are indicated (narrow lines). Distances to *GBA* are shown in Mb and cM.

Framework map positions are indicated as Mb from 1pter. The arrow above the *GBA* gene indicates transcription direction.

in most populations. While most laboratories perform the analysis of two or three frequent in most populations. While most laboratories perform the analysis of two or three frequent mutations, the individual detection of rare mutations is too expensive and time consuming for routine use. The fine genetic mapping of the *GBA* gene provided a diagnostic protocol based on cosegregation analysis, using highly polymorphic markers [44]. This protocol is useful when at least one mutated allele does not correspond to the most common mutations. An example is shown in Fig 5.

Direct diagnosis of the N370S mutation revealed that the father was a N370S

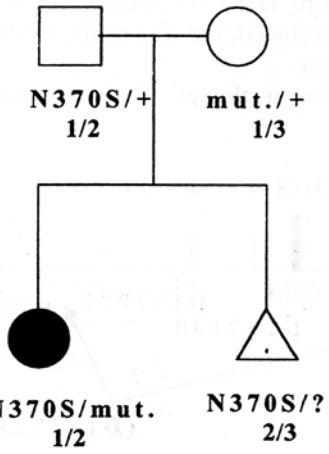


Figure 5. Example of cosegregation anlaysis for indirect molecular diagnosis. See text for details.

carrier. He passed this mutation to his affected daughter and to the foetus. No common mutation was detected in the mother (unknown mutation indicated as "mut"). Indirect diagnosis (using marker D1S2140 in this case) showed that the foetus was not affected. For this marker, the father carried alleles 1 and 2, while the mother was 1/3. The father passed allele 2 to both, the affected child and the foetus (since he was the only parent bearing allele 2). Thus, the affected daughter received allele 1 from the mother, while the foetus received allele 3. Using several markers, the risk of misdiagnosis because of recombination is very unlikely. Besides, since the amount of highly polymorphic markers available is very large, several informative markers for each family are guaranteed.

6. Putative ancestral haplotype for mutation N370S

To study the natural history of a mutation we should elucidate whether it was generated by a unique event or it corresponds to a recurrent event. In the first case, a genomic region around the mutation should be highly conserved in all the chromosomes that carry the mutation because it is inherited from a common ancestor. This conservation is detected by linkage disequilibrium between the mutation and polymorphic markers close to the gene.

To examine identity by descent by haplotype analysis, we first analysed five microsatellite markers closely linked to the *GBA* gene (D1S305, D1S2140, D1S2777, D1S2721 and D1S2624) on chromosomes bearing the common N370S mutation [31]. Initially, 18 N370S chromosomes from Spanish individuals, for which the phase was known (using samples from parents and/or siblings), were analysed. The haplotype 4-1-3, for the three closest markers (Fig. 6), was conserved in 16/18 (88.9%) of the chromosomes.

Further analyses performed on 66 Ashkenazi Jewish patients revealed that most of the Jewish and Spanish N370S chromosomes share alleles for most of these markers [45]. For 73 of the 104 N370S chromosomes from Ashkenazi Jews analysed, the phase was established for all the markers. Forty-three of these 73 (59 %) showed a conserved haplotype for markers D1S2140, D1S2777, D1S2721 and D1S2624 (Fig. 6). If the more distant marker D1S2624 is excluded, the proportion of chromosomes with a conserved haplotype rises to 68/73 (93 %).

In conclusion, a similar haplotype was associated with the common N370S

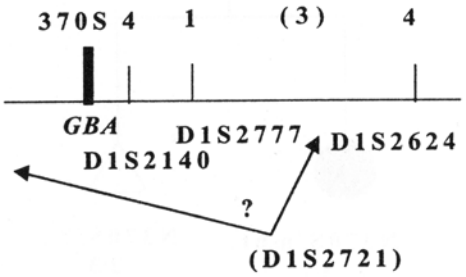


Figure 6. Conserved haplotype around the *GBA* N370S mutation.

mutation. Our data strongly support a founder effect as opposed to the recurrence of the N370S mutation. Although a new occurrence of N370S cannot be ruled out, it seems very unlikely. Assuming a founder effect, both genetic drift and selective advantage can explain the expansion of the mutated allele in the population.

Following Risch et al. [46] we estimated the age of the N370S mutation. Its origin was set between centuries 10^{th} and 13^{th} AD [47]. Since linkage disequilibrium values in Jews are higher than in Spanish patients, the mutation may have originated in a non-Jewish chromosome and entered the Jewish population more recently. These results have been confirmed using new polymorphisms very close to the *GBA* gene [48].

7. Complex alleles

The genomic region around the *GBA* gene in 1q21 includes several genes and pseudogenes. The *GBA* has a highly homologous pseudogene, *GBAP*, located 16 kb downstream. Other genes are metaxin, which also has a pseudogene, and thrombospondin. *GBA* and metaxin genes and pseudogenes are located in a very short genomic region and ordered as follows: *GBA*, pseudo-metaxin, *GBAP*, metaxin. This structure seems to be the result of an ancestral duplication. The presence of homologous regions in a small genomic interval enhances rearrangements, which result in Gaucher disease causing alleles named as "complex alleles" or "Rec" mutations. These alleles are generated by either crossover or gene conversion events between the *GBA* gene and its pseudogene. All were found in heterozygosis, suggesting lethality in homozygosis. The few homozygous cases described to date correspond to abortions or neonatal deaths [49-51].

The first gene-pseudogene fusion allele was described in 1990 [52]. It bears several sequence changes, normally present in the pseudogene: two missense mutations, L444P and A456P, and the silent polymorphism V460V. Since then, alleles with these changes, usually known as Rec*Nci*I, have been reported in various populations, being the most frequent complex allele in GD. Another complex allele, Rec*TL* [53], which bears the same changes as Rec*Nci*I plus D409H, has also been described. Each of these Rec alleles may have arisen from diverse mutational mechanisms like gene conversion or unequal crossing- over. In most cases this issue was not addressed and the name Rec*Nci*I or Rec*TL* is used regardless of the mutational mechanism.

The mutant allele Rec*Nci*I is very prevalent (21%) among Argentinean GD patients [35] (Fig. 3). The Rec*Nci*I allele results either from a gene/pseudogene crossing-over leading to gene fusion or from a gene conversion event [52]. ASO hybridisation or digestion analysis cannot distinguish these two mechanisms, but Southern blot analysis can. When digested with SspI enzyme, normal chromosomes produce a pattern of two bands of 17 and 12 kb, corresponding to the *GBA* gene and pseudogene respectively. Complex alleles caused by gene conversion are assumed to present the same band pattern as controls.

Alternatively, complex alleles produced by an unequal crossover between gene and pseudogene present a band of 14 kb, corresponding to the fusion gene (Fig. 7). Five of the 12 Argentinean patients bearing the Rec*Nci*I allele that were analysed by Southern blot presented a pattern that can be explained by a gene conversion event, while unequal crossover, which produces a fusion gene, was responsible for the complex allele

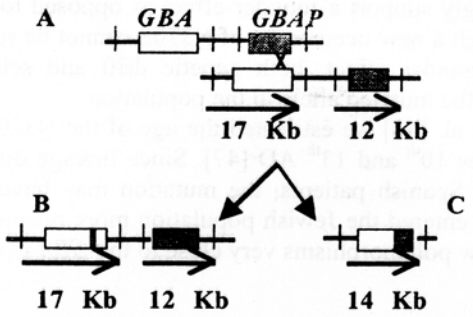


Figure 7. Putative model for the origin of the different Rec*Nci*I alleles. (A) Regions encompassing the *GBA* gene in different chromosomes are mispaired due to the high homology between gene and pseudogene. Arrows correspond to *Ssp*I fragmens: 17 kb (gene) and 12 kb (pseudogene). Rec*Nci*I alleles could be generated either by gene conversion (B), Southern *Ssp*I pattern as in wild type, or by unequal crossover (C), which produces a new *Ssp*I band of 14 kb (fusion gene).

identified in one patient.

Unexpectedly, a third pattern of 17-16-12 kb was found in six unrelated patients. This may correspond to a duplication of the pseudogene due to a crossover downstream of the *GBA* gene, probably between metaxin and its pseudogene, as suggested elsewhere [54]. This rearrangement does not appear to be pathogenic since a copy of the *GBA* gene remains intact. Moreover, it has also been found in control individuals.

In addition to these two relatively frequent complex alleles, some other recombinant alleles have been reported in individual cases. The recombination site for most of them lies in the 3' end of the gene, beyond exon 8 [8,55,56]. Only a few crossovers have been described in the 5' end of the *GBA* gene like those reported by Reissner et al. [57] in intron 3 or Filocamo et al. [58] in intron 6. We have recently described a new recombinant allele, present in heterozygosis in a type I GD patient [59]. An exhaustive molecular characterisation of this allele allowed the refinement of the crossover site to an 18-bp interval in the boundary between intron 2 and exon 3 of the *GBA* gene. This is the most upstream recombination event described so far in a GD mutant allele.

8. Expression of mutant alleles

Mutant alleles are expressed to characterise the mutant protein and compare its biochemical features with those of the normal enzyme. This study shows that a new change is a disease-causing mutation. Moreover, it may be useful for prognosis and therapy. Expression studies using a new baculovirus expression system are being performed. The new method generates recombinant baculovirus relying on site specific transposition in modified *E.coli* cells and allows selection of recombinant baculovirus. The normal and mutated proteins are evaluated by enzyme activity and Western blot. Several different *GBA* mutations have been expressed using a baculovirus system [60-63].

In our study, expression of wild type cDNA increases acid β-glucosidase activity (measured with 4MU- β-glucoside) by 6 to 10-fold the value of control fibroblasts. We are currently expressing cDNAs corresponding to the mutated alleles Y313H, P391L,

N392I, I402T, D409H and E326K+N188S. For some of the mutated alleles, like P391L and N392I mutations, expression studies revealed very low specific activities although the protein was clearly detected by immunoblot. In contrast, alleles I402T, D409H and E326K+N188S had significant glucocerebrosidase activities, ranging from about 5 to 30% of the wild type values.

9. Perspectives

Most of the mutations leading to Gaucher disease have been analysed in many populations, which has allowed molecular diagnosis and the establishment of some genotype-phenotype correlations, very useful for prognosis, treatment and counselling. Moreover, expression studies as mentioned above, help to characterise mutant enzymes and provide data on the structure-function correlation. However, it is now clear that other factors are involved in the clinical manifestation of the disease, since patients sharing the same genotype present very distinct clinical traits.

It has been shown that polymorphisms in several genes may contribute to the susceptibility to common diseases. This is also the case of the so-called monogenic diseases, in which several genes other than that having a main effect, contribute to phenotypic variability. Further research on Gaucher disease should focus on the analysis of these possible modulator genes.

Finally, the main topic to be addressed is the development of a definitive therapy for Gaucher disease patients. Enzyme replacement is valid but implies a life-long treatment and it is very expensive. A new strategy, currently in trial, is based on substrate reduction and may be a reliable alternative or complement. However, neither of them is a definitive solution. Gene therapy may become the final solution if the technical limitations are overcome. New methods, vectors and approaches should be developed to obtain a fully effective therapy for Gaucher disease.

Acknowledgements

We wish to thank Dr. Gregory Pastores, Dr. Néstor Chamoles, Teresa L. Harboe, Mariana Blanco and Adriana Rodriguez-Marí for their valuable contribution. We also thank the physicians who referred the patients and Robin Rycroft for revising the English. This study was supported by CICYT (SAF 97-0074 and SAF 2000-0200), by Fundació Marató de TV3 (Proj. Nos. 98-1220 and 1221), and by Genzyme.

References

1. Beutler, E., Grabowski, G.A. In: Scriver CR, Beaudet AL, Sly WS, Valle D. (Editors) The Metabolic and Molecular Bases of Inherited Disease. Page 3635, McGraw-Hill Inc. New York.
2. Brady, R.O., Kanfer, J.N., Shapiro, D. 1965, Biochem. Biophys. Res. Commun. 18, 221.
3. Ginns, E.I., Choudary, P.V., Tsuji, S., Martin, B., Stubblefield, B., Sawyer, J., Hozier, J.,Barranger, J.A. 1985, Proc. Natl. Acad. Sci. USA 82, 7101.
4. Horowitz, M., Wilder, S., Horowitz, Z., Reiner, O., Gelbart, T., Beutler, E. 1989, Genomics 4, 87.
5. Beutler, E., West, C., Gelbart, T. 1992, Genomics 12, 795.
6. Christomanou, H., Aignesberger, A., Linke, R.P. 1986, Biol. Chem. Hoppe-Seyler 367, 879.
7. Christomanou, H., Chabás, A., Pámpols, T., Guardiola, A. 1989, Klin. Wochenschr. 67, 999.
8. Beutler, E., Gelbart, T. 1998, Blood Cells Mol. Dis. 24, 2.

9. Beutler, E., Gelbart, T., Kuhl, W., Zimran, A., West, C. 1992, Blood 79, 1662.
10. Beutler, E., Gelbart, T. 1992, Am. J. Med. Genet. 44, 389.
11. Horowitz, M., Tzuri, G., Eyal, N., Berebi, A., Kolodny, E.H., Brady, R.O., Barton, N.W., Abrahamov, A., Zimran, A. 1993, Am. J. Hum. Genet. 53, 921.
12. Dahl, N., Lagerstrom, M., Erikson, A., Pettersson, U. 1990, Am. J. Hum. Genet. 47, 275.
13. Tuteja, R., Bembi, B., Agosti, E., Baralle, F.E. 1993, Hum. Mol. Genet. 2, 781.
14. Walley, A.J., Barth, M.L., Ellis, I., Fensom, A.H., Harris, A. 1993, J. Med. Genet. 30, 280.
15. Cormand, B., Vilageliu, L., Burguera, J.M., Balcells, S., Gonzalez-Duarte, R., Grinberg, D., Chabás, A. 1995, Hum. Mutat. 5, 303.
16. Michelakakis, H., Dimitriou, E., Vanweely, S., Boot, R.G., Mavridou, I., Verhoek, M., Aerts, J. 1995, J. Inher. Metab. Dis. 18, 609.
17. Amaral, O., Pinto, E., Fortuna, M., Lacerda, L., Miranda, M. 1996, Hum. Mutat. 8, 280.
18. Tylki-Szymanska, A., Millat, G., Maire, I., Czartoryska, B. 1996, Eur. J. Hum. Genet. 4, 334.
19. Germain, D.P., Puech, J.P., Caillaud, C., Kahn, A., Poenaru, L. 1998, Am. J. Hum. Genet. 63, 415.
20. Lecoutre, P., Demina, A., Beutler, E., Beck, M., Petrides, P. 1997, Hum. Genet. 99, 816.
21. Boot, R.G., Hollak, C., Verhoek, M. et al., 1997, Hum. Mutat. 10, 348.
22. Hodanova, K., Hrebicek, M., Cervenkova, M., Mrazova, L., Veprekova, L.,Zeman, J. 1999, Blood Cells Mol. Dis. 25, 287.
23. 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.
24. Theophilus, B., Latham, T., Grabowski, G.A., Smith, F.I. 1989, Am. J. Hum. Genet. 45, 212.
25. Chabás, A., Cormand, B., Grinberg, D., Burguera, J.M., Balcells, S., Merino, J.L., Mate, I., Sobrino, J.A., Gonzalez-Duarte, R., Vilageliu, L. 1995, J. Med. Genet. 32, 740.
26. Abrahamov, A., Elstein, D., Grosstsur, V., Farber, B., Glaser, Y., Hadashalpern, I., Ronen, S., Tafakjdi, M., Horowitz, M., Zimran, A. 1995, Lancet 346, 1000.
27. Barton, N.W., Brady, R.O., Dambrosia, J.M., et al. 1991, N. Engl. J. Med. 324, 1464.
28. Dunbar, C., Kohn, D., Karlsson, S., Barton, N., Brady, R. et al. 1996, Hum. Gene Ther. 7, 231.
29. Cormand, B., Vilageliu, L., Balcells, S., Gonzalez-Duarte, R., Chabás, A., Grinberg, D. 1996, Hum. Mutat. 7, 272.
30. Chabás, A., Cormand, B., Balcells, S., Gonzalez-Duarte, R., Casanova, C., Colomer, J., Vilageliu, L., Grinberg, D. 1996, J. Inher. Metab. Dis. 19, 798.
31. Cormand, B., Grinberg, D., Gort, L., Chabás, A., Vilageliu, L. 1998, Hum. Mutat. 11, 295.
32. Dinur, T., Osiecki, K.M., Legler, G., Gatt, S., Desnick, R.J., Grabowski, G.A. 1986, Proc. Natl. Acad. Sci. USA 83, 1660.
33. Lewis, B.D., Nelson, P.V., Robertson, E.F., Morris, C.P. 1994, Am. J. Med. Genet. 49, 218.
34. Nelson, P.V., Carey, W.F., Morris, C.P., Lewis, B.D. 1995, Am. J. Med. Genet. 58, 382.
35. Cormand, B., Harboe, T.L., Gort, L., Campoy, C., Blanco, M., Chamoles, N., Chabás, A., Vilageliu, L., Grinberg, D. 1998, Am. J. Med. Genet. 80, 343.
36. Beutler, E., Demina, A., Gelbart, T. 1994, Mol. Med. 1, 82.
37. Cormand, B., Grinberg, D., Gort, L., Fiumara, A., Barone, R., Vilageliu, L., Chabás, A. 1997, Am. J. Med. Genet. 70, 437.
38. Beutler, E., Kattamis, C., Sipe, J., Lipson, M. 1995, Lancet 46, 1637.
39. Uyama, E., Uchino, M., Ida, H., Eto, Y., Owada, M. 1997, J. Med. Genet. 34, 175.
40. Chabás, A., Gort, L., Montfort, M., Castelló, F., Dominguez, M.C., Grinberg, D., Vilageliu, L. 1998, J. Med. Genet. 35, 775.
41. Cormand, B., Montfort, M., Chabás, A., Vilageliu, L., Grinberg, D. 1997, Hum. Genet. 100, 75.
42. Lander, E.S., Linton, L.M., Birren, B. et al., 2001, Nature 409, 860.
43. Venter, J.C., Adams, M.D., Myers, E.W. et al., 2001, Science 291, 1304.
44. Cormand, B., Montfort, M., Chabás, A., Grinberg, D., Vilageliu, L. 1998, Prenat. Diagn. 18, 207.
45. Díaz, A., Montfort, M., Cormand, B., Zeng, B.J., Pastores, G.M., Chabás, A., Vilageliu, L., Grinberg, D. 1999, Am. J. Hum. Genet. 64, 1233.
46. Risch, N., Deleon, D., Ozelius, L., Kramer, P., Almasy, L., Singer, B., Fahn, S., Breakefield, X., Bressman, S. 1995, Nat. Genet. 9, 152.
47. Díaz, A., Montfort, M., Cormand, B., Zeng, B.J., Pastores, G.M., Chabás, A., Vilageliu, L., Grinberg, D. 2000, Am. J. Hum. Genet. 66, 2014.

48. Rodríguez-Marí, A., Díaz-Font, A., Chabás, A., Pastores, G.M., Grinberg, D., Vilageliu, L. 2001, Blood Cells Mol. Dis. 27, 950.
49. Strasberg, P.M., Skomorowski, M.A., Warren, I.B., Hilson, W.L., Callahan, J.W., Clarke, J.T.R. 1994, Biochem. Med. Metab. Biol. 53, 16.
50. Sidransky, E., Tayebi, N., Stubblefield, B.K., Eliason, W., Klineburgess, A., Pizzolato, G.P., Cox, J.N., Porta, J., Bottani, A., Delozierblanchet, C.D. 1996, J. Med. Genet. 33, 132.
51. Finn, L.S., Zhang, M., Chen, S.H., Scott, C.R. 2000, Am. J. Med. Genet. 91, 222.
52. Zimran, A., Sorge, J., Gross, E., Kubitz, M., West, C., Beutler, E. 1990, J. Clin. Invest. 85, 219.
53. Latham, T., Grabowski, A., Theophilus, B.D.M., Smith, F.I. 1990, Am. J. Hum. Genet. 47, 79.
54. Tayebi, N., Park, J., Madike, V., Sidransky, E. 2000, Hum. Genet. 107, 400.
55. Tayebi, N., Stone, D.L., Sidransky, E. 1999, Mol. Genet. Metab. 68, 209.
56. Stone, D.L., Tayebi, N., Orvisky, E., Stubblefield, B., Madike, V., Sidransky, E. 2000, Hum. Mutat. 15, 181.
57. Reissner, K., Tayebi, N., Stubblefield, B.K. et al., 1998, Mol. Genet. Metab. 63, 281.
58. Filocamo, M., Bonuccelli, G., Mazzotti, R., Giona, F., Gatti, R. 2000, Blood Cells Mol. Dis. 26, 307.
59. Cormand, B., Díaz, A., Grinberg, D., Chabás, A., Vilageliu, L. 2000, Blood Cells Mol. Dis. 26, 409.
60. Grace, M.E., Graves, P.N., Smith, F.I., Grabowski, G.A. 1990, J. Biol. Chem. 265, 6827.
61. Grace, M.E., Newman, K.M., Scheinker, V., Berg-Fussman, A., Grabowski, G.A. 1994, J. Biol. Chem. 269, 2283.
62. Choy, F.Y.M., Wei, C., Levin, D. 1996, Am. J. Med. Genet. 65, 184.
63. Amaral, O., Marcao, A., Miranda, M., Desnick, R.J., Grace, M.E. 2000, Eur. J. Hum. Genet. 8, 95.

Available online at www.sciencedirect.com
SCIENCE DIRECT
Blood Cells, Molecules, and Diseases 35 (2005) 253 – 258
BLOOD CELLS, MOLECULES, & DISEASES
www.elsevier.com/locate/ybcmd

Perinatal lethal phenotype with generalized ichthyosis in a type 2 Gaucher disease patient with the [L444P;E326K]/P182L genotype: Effect of the E326K change in neonatal and classic forms of the disease

Amparo Chabás[a,*], Laura Gort[a], Anna Díaz-Font[b], Magdalena Montfort[b], Raül Santamaría[b], Manuel Cidrás[c], Daniel Grinberg[b], Lluïsa Vilageliu[b]

[a]*Institut de Bioquímica Clínica, Corporació Sanitària Clínic, Mejia Lequerica, s/n. Edifici Helios III, planta baixa. 08028 Barcelona, Spain*
[b]*Dept. de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain*
[c]*Hospital Materno-Infantil, Serv. Neonatología, Las Palmas de Gran Canaria, Spain*

Submitted 25 April 2005
Available online 20 June 2005
(Communicated by E. Beutler, M.D., 27 April 2005)

Abstract

Gaucher disease, the most common lysosomal storage disorder, encompasses a wide spectrum of clinical symptoms. The perinatal lethal form is very rare and is considered a distinct form of classic type 2 Gaucher disease. Prominent features of the severe perinatal form are hepatosplenomegaly variable, associated with hydrops fetalis and ichthyosis. Here, we describe a child who presented generalized ichthyosis and died at 25 days of age. Genotype analysis revealed compound heterozygosity for the complex allele [L444P;E326K] and mutation P182L, described for the first time in this patient. Mutations E326K and L444P were on the same chromosome. Expression studies of mutant glucocerebrosidases showed that the double mutant allele had lower activity, 8.5% of wild type, in contrast to the activity of individual E326K and L444P mutant enzymes, 42.7% and 14.1%, respectively. The P182L mutant enzyme showed no glucocerebrosidase activity.

A revision of the genotypes identified in a series of Spanish patients with type 2 Gaucher disease showed that the complex allele [L444P;E326K] accounted for 19.2% of patient alleles and that homozygosity for this allele or its heterozygosity with mutation L444P, or another severe mutation such as P182L, was associated with the perinatal lethal presentation of the disease. In contrast, the [L444P;E326K] allele was not detected in patients with classic type 2 diagnosed when several months old. The high frequency of the E326K substitution observed in patients with type 2 as compared to the general population (0.5%) suggests that this change may have a modulating negative effect on the clinical condition of these Gaucher disease patients when present in combination with mutation L444P.

The relatively high prevalence of the double mutant allele in Spanish patients prompted us to perform a haplotype analysis, using four polymorphic markers, which suggest a common origin for this allele. During the mutational analysis of the series of type 2 patients, a novel mutation, I260T (c.896T > C), was identified.

Keywords: Gaucher disease; GBA; Ichthyosis; Perinatal form; [L444P;E326K] allele; I260T mutation

Introduction

Gaucher disease (GD, OMIM 230800, 230900, 231000), the inherited deficiency of the lysosomal enzyme glucocerebrosidase (GBA, EC 3.2.1.45), presents broad clinical and genetic heterogeneity. Clinically, three phenotypes have been delineated on the basis of the presence and severity of central nervous system involvement: type 1 (non-neuronopathic adult form), type 2 (acute neuronopathic form diagnosed classically in several months old infants), and type 3 (subacute neuronopathic juvenile form) [1].

Severe perinatal presentation of GD is very rare, however, the identification of new neonatal cases with a distinct severe phenotype of GD has continued to increase since the generation of a mouse model with a null allele in the *GBA*

* Corresponding author. Fax: +34 93 2275668.
E-mail address: achabas@clinic.ub.es (A. Chabás).

1079-9796/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bcmd.2005.04.007

gene [2] prompted the search for an analogous presentation associated with glucocerebrosidase deficiency in humans [3].

The description of new cases has led some authors to consider the perinatal lethal phenotype as a distinct form of classic type 2 GD with hepatosplenomegaly as the major sign associated with non-immune hydrops fetalis, ichthyosis, arthrogryposis, and facial dysmorphy [4]. A clinical feature frequently reported in fetal and neonatal cases is gross skin abnormalities, mainly generalized ichthyosis. This dermatological manifestation described as collodion skin [5] was referred to as "eczematous skin" and "seborrheic dermatitis" in early reports such as that of Hernández and Bueno [6] in two Spanish siblings with GD, and according to Mignot et al. [4] may be present in 41% of infants with the perinatal lethal form of GD.

Here, we present the clinical and molecular characterization of a newborn with generalized ichthyosis. The patient was heterozygous for the double mutant allele [L444P;E326K] and mutation P182L. These alleles show very low and no activity, respectively, when expressed in an in vitro system. We also review the genotypes in a series of patients with neonatal or early presentation of the disease and underline the contribution of the E326K change on the phenotypic expression of type 2 GD. Haplotype analysis for the [L444P;E326K] allele, present in several Spanish patients, suggests a single origin. One novel mutation, I260T (c.896T > C), was identified when genotyping Spanish type 2 GD patients.

Case report

The patient, a boy weighing 1690 g, was born at 35 weeks of gestation to a Spanish couple after ovocyte donation. He was admitted to the hospital at 10 days of age because of low weight and generalized ichthyosiform skin. The patient exhibited spontaneous movements, normal cardiorespiratory activity, hepatosplenomegaly and generalized hypotonia. He presented with severe thrombocytopenia that required several transfusions without recovering to normal values. The patient died at 25 days of age.

Materials and methods

DNA isolation

Genomic DNA was prepared from cultured skin fibroblasts and peripheral blood using a standard method [7].

Mutation analysis

The DNA from the proband was initially screened for the most frequent mutations in the Spanish GD population [8]. We then performed PCR and single-strand conformation polymorphism (SSCP) analysis of 14 DNA fragments covering the 11 exons, the flanking intronic junctions, part of the promoter region, and the first polyadenylation signal. Samples showing abnormal SSCP patterns were sequenced by fluorescent dideoxy cycle sequencing (ABI 373A Fluorescent DNA sequencer, Perkin-Elmer, Norwalk, CT).

Screening for the E326K change was performed in genomic DNA from 100 control individuals and 56 GD patients. Samples were PCR-amplified [8] and tested for the presence of the E326K allele by restriction digestion with *Mbo*II. Screening for mutation P182L was performed in genomic DNA from 54 control individuals by SSCP, and for the novel mutation, I260T, was performed in genomic DNA from 50 control individuals by restriction digestion with *Bse*MI.

Expression and characterization of mutant proteins

GBA cDNAs carrying the mutant alleles were expressed using a baculovirus expression system, and the mutant proteins were analyzed by Western blot as previously described [9]. Glucocerebrosidase (acid β-glucosidase) activity was measured using the fluorogenic substrate 4-methyl-umbelliferyl-β-D-glucopyranoside as reported [10].

Polymorphism analyses

Microsatellites 5GC3.2 and ITG6.2 were analyzed by PCR amplification using FAM-fluorescently labeled primers and genotyping performed in an ABI PRISM 310 Genetic analyzer PE Applied Biosystems as previously described [11]. The data were analyzed with the Genescan 3.1 Analyzer and the Genotyper 2.5 software. The 6144A > G polymorphism was analyzed by *Hha*I digestion as described [8]. The D1S1595 (also named D1S2140) belongs to the MapPairs set (Research Genetics, Huntsville, AL) and was analyzed according to the manufacturer's recommendations.

Results

The diagnosis of GD in the proband was established by the severe deficiency of glucocerebrosidase in cultured fibroblasts (0.7% of control cells).

Genotype analysis

Genomic DNA amplification and digestion with *Nci*I showed that the patient was heterozygous for mutation L444P (c.1448T > C). SSCP analysis of the fragments of the *GBA* gene showed altered patterns in exons 6 and 8 which were due to mutations P182L (c.662C > T; exon 6) and E326K (c.1093G > A; exon 8) as shown by sequencing of the PCR fragments. Analysis of the DNA from the patient's father disclosed that mutations E326K and L444P were on the same chromosome and that P182L mutation was absent. P182L is a mutation that was described for the first time in this patient [9]. One hundred and eight random normal

chromosomes were tested for P182L mutation, and negative results were obtained.

Expression studies of mutant glucocerebrosidases

To determine the effect of these mutations on enzymatic activity, P182L, the double mutant allele [L444P;E326K] and their individual mutations were expressed in Sf9 cells using the baculovirus expression system. The P182L mutant enzyme showed no glucocerebrosidase activity. Activity of the mutant enzyme E326K was 42.7% of wild type while that of L444P mutant was 14.1%. However, activity decreased to 8.5% in the [L444P;E326K] double mutant enzyme. Thus, the residual glucocerebrosidase activity observed in the patient was due to the double mutant allele.

Clinical findings and genotypes in a series of Spanish type 2 GD patients. Frequency of the [L444P;E326K] allele

We retrospectively reviewed 12 patients with perinatal or classic type 2 presentation (Table 1). Cases 1 and 2 showed early presentation without neurological symptoms and died before 4 months of age. Both patients carried genotypes bearing the complex allele [L444P;E326K] in homozygosis and heterozygosis, respectively. This double mutant allele was also present in case 4 who, like cases 5 to 13, had developed the neurological symptoms reported in classic type 2 GD. L444P was the most frequent mutation and was detected in 50% of patient alleles (13/26), either as a single mutation in 30.8% (8/26) or as a complex allele with E326K in 19.2% (5/26) of alleles. Among the point mutations detected in this group of patients, we identified mutation I260T (c.896T > C) in exon 7, which had not been reported previously and was not detected in control individuals.

The E326K substitution among GD patients and healthy controls

The allelic frequency of E326K had previously been reported to be low and similar in the control population (0.9%) and in GD patients (1.3%) [12]. The high frequency of this allele in the type 2 patient genotypes of our series prompted us to screen for this allele in the Spanish general population and in patients with different types of GD. We analyzed the presence of the E326K change in 100 healthy individuals and in 56 GD patients. In controls, only 1 of the 200 alleles screened carried this substitution (frequency 0.5%). No other change was detected in this sample after sequencing most of the *GBA* gene. Analysis of GD patients with the distinct clinical types showed that 13 out of the 112 GD alleles carried the E326K substitution: 6 alleles detected in 37 type 1 patients, 5 in 13 type 2 patients (shown in Table 1), and 2 in 6 type 3 patients (frequencies 8.1%, 19.2%, and 16.6%, respectively).

Among the patients, E326K was found always in a double mutant allele. In 11 out of 13 alleles, L444P was found in the same chromosome. In the other two alleles, the mutation in *cis* with E326K was either N188S or F213I. The genotypes of the patients bearing the [L444P;E326K] allele are shown in Table 2.

Haplotype analysis of [L444P;E326K] alleles

To find out whether the relatively high number of [L444P;E326K] alleles could share a common origin, a haplotype analysis using four polymorphisms was performed. Phases were established, when possible, by family analysis. For type 1 patients for whom family samples were not available, the reported association between particular

Table 1
Perinatal lethal form and classic type 2 Gaucher disease in Spanish patients

Case number[a]	Genotype	Age at onset	Age at death	Clinical data
1 (II.8)[b]	[L444P;E326K]/[L444P;E326K]	Birth	82 d	Hepatosplenomegaly (HS), ascites, cholestatic jaundice, thrombocytopenia. Sibling similarly affected without ascites
2 (II.7) [b]	[L444P;E326K]/L444P	2 mo	3.5 mo	HS, bleeding tendency, growth retardation
3 (II.12)	**[L444P;E326K]/P182L**	**Birth**	**25 d**	**Present report**
4 (II.13)	[L444P;E326K]/G202R	4 mo	12 mo	HS, apneas, hypotonia, leukopenia, brain atrophy
5 (II.1) [b]	L444P/c.1263del55	<5 mo	8.5 mo	HS, pneumonia, seizures, head retroflexion, hyperreflexia, thrombocytopenia. Affected sibling dead at 9.5 mo
6 (II.3) [b]	L444P/R120W	5 mo	9 mo	HS, generalized hypertonia, neck hyperextension, opisthotonos, strabismus, pulmonary fibrosis. Affected sibling dead at 10 mo
7 (II.5) [b]	L444P/N392I	6 mo	10 mo	HS, myoclonic seizures, hypertonia, psychomotor retardation
8 (II.2) [b]	L444P/G195E	8 mo	NA	HS, moderate psychomotor retardation, vomiting, failure to thrive
9 (II.14)	L444P/G202R	4 mo	6 mo	HS, myoclonia, hypertonia, ichthyotic dermatitis
10 (II.15)	L444P/I260T[c]	6 mo	21 mo	HS, peculiar facial appearance, strabismus, trismus, respiratory difficulties, laryngeal spasm
11 (II.11)	L444P/c.203_204insC	5 mo	6 mo	HS, growth retardation, opisthotonos, hypotonia
12 (II.4) [b]	G389E/?	6 mo	10 mo	HS, myoclonic seizures, hypertonia, psychomotor retardation
13 (II.6) [b]	D409H/R120W	11 mo	20 mo	HS, spasticity, psychomotor retardation

NA.—data not available
[a] In brackets, patient number used in other publications or databases.
[b] Patient reported in [8].
[c] Novel mutation.

Table 2
Haplotype analysis of [L444P;E326K] chromosomes in Spanish GD patients

	Genotype	6144A→G	ITG6.2	5GC3.2	D1S1595
Type 1 patients					
I.21[a]	**[L444P;E326K]**/N370S	**G** A	**322** 318	**220** 222	**3** 5
I.39[a]	**[L444P;E326K]**/N370S	**G** A	**326** 318	**220** 222	**5** 5
I.44	**[L444P;E326K]**/N370S	**G** A	**322** 318	**220** 222	**3** 5
I.61	**[L444P;E326K]**/N370S	**G** A	**322** 322	**220** 222	**3** 3
Type 2 patients					
II.7[a] (case 2)	**[L444P;E326K]**/L444P	**G** A	**318/322** 322	**220/222** 220/222	**3/7** 5/7
II.8[a] (case 1)	**[L444P;E326K]/[L444P;E326K]**	**G G**	**322 322**	**220 220**	**3 3**
II.12 (case 3, this study)	**[L444P;E326K]**/P182L	**G** A	**322** 322	**220** 220	**5** 7
II.13 (case 4)[b]	**[L444P;E326K]**/G202R	–	–	–	–
Type 3 patients					
III.3[a]	**[L444P;E326K]**/D409H	**G** A	**318/322** –	**220/222** 222	**3/7** 7
III.4	**[L444P;E326K]**/D409H	**G** G	–	**220** 220	**3/7** 7/3

Alleles in phase with the [L444P;E326K] are indicated in bold.
[a] Patient reported in [8].
[b] No sample available.

marker alleles and the N370S mutation was assumed [11]. As shown in Table 2, a common haplotype G(6144A > G)-322(ITG6.2)-220(5GC3.2)-3(D1S1595) was found for most of the 10 chromosomes bearing the [L444P;E326K] allele analyzed. In particular, the haplotype was confirmed for 5 alleles, while 3 more cases have consistent genotypes (phases could not be established). Moreover, if the D1S1595, the marker that is far away from the gene, is not considered, 90% of the chromosomes bear the same haplotype. The only exception is case I.39 where alleles for ITG6.2 and D1S1595 are different from those in the common haplotype.

Structure protein analysis

The complex allele [L444P;E326K] expressed almost no detectable catalytic activity, while the individual E326K and L444P alleles expressed 42.7% and 14.1% of wild type, respectively. According to the 3D X-ray structure of β-glucosidase reported by Dvir et al. [13], L444P located on domain II may lead to disruption of the protein hydrophobic core and alter folding. The E326 position is located on domain III, in the sixth α-helix of the Tim barrel, at the same face than the arginine at position 329. Thus, the substitution of the basic lysine for glutamic acid (E326K) may destroy the salt bridge between both side chains (E326 and R329) and also may introduce a repulsion state involving a close lysine (K321) side chain. The reduced activity of the double mutant allele suggests that the correct interaction between these two domains may have been disturbed, thereby affecting the stability of the catalytic site or the binding of substrates or other macromolecules necessary for catalysis.

Discussion

Association between ichthyosis and glucocerebrosidase deficiency was first described in two siblings with neonatal GD by Lui et al. [5]. Since then, about 14 cases of GD with generalized ichthyosis have been reported. These cases were diagnosed at birth; most were without neurologic involvement and died within 3 months [4,14–23]. A few cases had neonatal presentation, evolution with neurologic symptoms, and survival rates similar to classic type 2 GD [24]. The case described in the present study should be included in the group of patients with the perinatal lethal form of GD. This patient was heterozygous for the complex allele [L444P;E326K] and for mutation P182L. This mutation was identified in this patient for the first time, and it was reported when several mutant alleles were expressed [9] without description of the clinical findings of the patient. Later, it was also reported in an Italian patient [25]. The expression studies of the P182L mutation showed that it had no detectable glucocerebrosidase activity [9]. A similar result was found by Miocic et al. using a different expression system [25]. Regarding E326K, its presence in the double mutant allele [L444P;E326K] significantly lowered the activity of that of the L444P mutant enzyme, behaving as a modifier variant, as reported by Montfort et al. [9], rather than as a neutral polymorphism, consequently increasing the severity of the disease.

Significant genotypic variability is observed when comparing the genotypes of 9 published cases with generalized ichthyosis. Mutation L444P, the most frequent mutation encountered in these 9 patients, occurred as part of the recombinant allele *RecNciI* [17,18,21] or in combination with E326K [20] but never alone. The double mutant allele [L444P;E326K] identified in the Spanish case presented here had previously been described in another patient with ichthyosis but with the neurologic symptoms of type 2 GD who had inherited this allele from her Mexican mother. The second allele carried mutation E233X [20]. In addition, another patient of Mexican origin bore this complex allele (genotype [L444P;E326K]/[recG;L444P]) [26], but no clinical description is available except for the mention of absence of ichthyosis [27]. This genotypic variability is extensive to classic type 2 GD. Thus, it is difficult to draw

predictions about the phenotype and the lethality of the disease based on the genotype, although generalizations such as that L444P homozygotes do not develop type 2 GD can be inferred [26]. In fact, in the series of Spanish type 2 patients, none carried genotype L444P/L444P, while this genotype has been identified in 2 children with the type 3 form of the disease.

On the basis of the results from our series of type 2 GD patients, we might conclude that homozygosity for the complex allele [L444P;E326K] or heterozygosity for this allele and mutation L444P or another severe mutation, such as P182L, could be associated with the perinatal lethal variant of GD, with early presentation, no neurologic involvement and death within the first 3 months (cases 1–3).

Regarding the remaining patients, cases 5–11 showed classic type 2 disease. Most of these were compound heterozygotes for mutation L444P alone and a missense mutation, a deletion, an insertion or a still unknown allele, with onset around 6 months of age and development of the neurologic symptoms as reported in the acute neuronopathic form of the disease. Case 4 could be considered as part of a continuous spectrum of neurological involvement, despite the scarce clinical data available. Only two patients in this series (cases 12 and 13) did not carry mutation L444P in their genotypes.

Moreover, the double mutant allele [L444P;E326K] has also been identified in two unrelated patients with the subacute type 3 disease, both carrying the genotype D409H/[L444P;E326K].

The relatively high prevalence of the double mutant allele in Spanish patients prompted us to perform a haplotype analysis, using four polymorphic markers. A common haplotype was identified in most (9/10) of the chromosomes bearing the [L444P; E326K]. If only the three closest markers are taken into account, the only exception could be explained by a mutation occurred in the polymorphic marker ITG6.2, changing allele 322 to 326. This would suggest a common origin for all the [L444P;E326K] chromosomes. Interestingly, the two only other reported cases with this double mutant allele, two Mexican patients [20,26], could have a Spanish origin.

The presence of the complex allele [L444P; E326K] found in a high proportion of patients with type 2 GD (19.2%) suggests a modulating action of the E326K substitution when in combination with mutation L444P on the clinical presentation of these patients. This stresses the need for complete genotype analysis when attempting to correlate genotype with the phenotypic expression and to establish appropriate treatment.

Acknowledgments

The authors thank Dr. Sílvia Atrian (Departament de Genètica) for helpful collaboration in the structural protein analysis. The technical assistance of J. Jarque, H. Sellés, and T. Soler is gratefully acknowledged. A.D.F. is a recipient of a fellowship from the Spanish Ministerio de Educación y Cultura, and R.S. is a recipient of a fellowship from the Spanish Ministerio de Ciencia y Tecnologia. This work was supported in part by FIS (Redes Temáticas REDEMETH G03/054) and Plan Nacional I+D (SAF 2003-00386).

References

[1] E. Beutler, G.A. Grabowski, Gaucher disease, in: C.R. Scriver, W.S. Sly, B. Childs, A.L. Beaudet, D. Valle, K.W. Kinzler, B. Volgstein, (Eds.),.The metabolic and molecular bases of inherited disease, vol. III, McGraw-Hill, New York, 2001, pp. 3635–3668.

[2] V.L.J. Tybulewicz, M.L. Tremblay, M.E. LaMarca, R. Willemsen, B.K. Stubblefield, S. Winfield, B. Zablocka, E. Sidransky, B.M. Martin, S.P. Huang, K.A. Mintzer, H. Westphal, R.C. Mulligan, E.I. Ginns, Animal model of Gaucher's disease from targeted disruption of the mouse glucocerebrosidase gene, Nature 357 (1992) 407–410.

[3] E. Sidransky, D.M. Sherer, E.I. Ginns, Gaucher disease in the neonate: a distinct Gaucher phenotype is analogous to a mouse model created by targeted disruption of the glucocerebrosidase gene, Pediatr. Res. 32 (1992) 494–498.

[4] C. Mignot, A. Gelot, B. Bessieres, F. Daffos, M. Voyer, F. Menez, C.F. Bianco, S. Odent, D. Le Duff, P. Loget, P. Fargier, J. Costil, P. Josset, J. Roume, M.T. Vanier, I. Maire, T. Billette de Villemeur, Perinatal–lethal Gaucher disease, Am. J. Med. Genet. 120A (2003) 338–344.

[5] K. Lui, C. Commens, R. Choong, R. Jaworski, Collodion babies with Gaucher's disease, Arch. Dis. Child. 63 (1988) 854–856.

[6] F. Hernández, M. Bueno, Infantile neurological Gaucher's disease in three siblings, Virchows Arch. (Pathol. Anat.) 360 (1973) 27–32.

[7] S.A. Miller, D.D. Dyke, H.F. Polesky, A simple salting out method for extracting DNA from nucleated cells, Nucleic Acids Res. 16 (1988) 1215.

[8] B. Cormand, D. Grinberg, L. Gort, A. Chabás, L. Vilageliu, Molecular analysis and clinical findings in the Spanish Gaucher disease population: putative haplotype of the N370S ancestral chromosome, Hum. Mutat. 11 (1998) 295–305.

[9] M. Montfort, A. Chabás, L. Vilageliu, D. Grinberg, Functional analysis of 13 GBA mutant alleles identified in Gaucher disease patients: pathogenic changes and "modifier" polymorphisms, Hum. Mutat. 23 (2004) 567–575.

[10] A. Chabás, B. Cormand, D. Grinberg, J.M. Burguera, S. Balcells, J.L. Merino, I. Mate, J.A. Sobrino, R. Gonzalez-Duarte, L. Vilageliu, Unusual expression of Gaucher's disease: cardiovascular calcifications in three sibs homozygous for the D409H mutation, J. Med. Genet. 32 (1995) 740–742.

[11] A. Rodríguez-Marí, A. Díaz-Font, A. Chabás, G.M. Pastores, D. Grinberg, L. Vilageliu, New insights into the origin of the Gaucher disease-causing mutation N370S: extended haplotype analysis using the 5GC3.2, 5470G/A, and ITG6.2 polymorphisms, Blood Cells Mol. Diseases 27 (2001) 950–959.

[12] J.K. Park, N. Tayebi, B.K. Stubblefield, M.E. LaMarca, J.J. MacKenzie, D.L. Stone, E. Sidransky, The E326K mutation and Gaucher disease: mutation or polymorphism?, Clin. Genet. 61 (2002) 32–34.

[13] H. Dvir, M. Harel, A.A. McCarthy, L. Toker, I. Silman, A.H. Futerman, J.L. Sussman, X-ray structure of human acid-beta-glucosidase, the defective enzyme in Gaucher disease, EMBO Rep. 4 (2003) 704–709.

[14] A.H. Lipson, M. Rogers, A. Berry, Collodion babies with Gaucher's disease: a further case, Arch. Dis. Child. 66 (1991) 667.

[15] D.M. Sherer, L.A. Metlay, R.A. Sinkin, C. Mongeon, R.E. Lee, J.R. Woods, Congenital ichthyosis with restrictive dermopathy and

Gaucher disease: a new syndrome with associated prenatal diagnostic and pathology findings, Obstet. Gynecol. 81 (1993) 842–844.
[16] Z. Ince, A. Coban, O. Peker, U. Ince, G. Can, Gaucher disease associated with congenital ichthyosis in the neonate, Eur. J. Pediatr. 154 (1995) 418–422.
[17] E. Sidransky, N. Tayebi, B.K. Stubblefield, W. Eliason, A. Klineburgess, GP. Pizzolato, J.N. Cox, J. Porta, A. Bottani, C.D. DeLozier-Blanchet, The clinical, molecular, and pathological characterisation of a family with two cases of lethal perinatal type 2 Gaucher disease, J. Med. Genet. 33 (1996) 132–136.
[18] S. Rowlands, H. Murray, Prenatal ultrasound findings in a fetus diagnosed with Gaucher's disease (type 2) at birth, Prenat. Diagn. 17 (1997) 765–769.
[19] D.L. Stone, O.P. van Diggelen, J.B.C. de Klerk, J.L.J. Gaillard, M.F. Niermeijer, R. Willemsen, N. Tayebi, E. Sidransky, Is the perinatal lethal form of Gaucher disease more common than classic type 2 Gaucher disease? Eur. J. Hum. Genet. 7 (1999) 505–509.
[20] M.E. Grace, P. Ashton-Prolla, G.M. Pastores, A. Soni, R.J. Desnick, Non-pseudogene-derived complex acid beta-glucosidase mutations causing mild type 1 and severe type 2 Gaucher disease, J. Clin. Invest. 103 (1999) 817–823.
[21] L.S. Finn, M. Zhang, S.H. Chen, C.R. Scott, Severe Type II Gaucher disease with ichthyosis, arthrogryposis and neuronal apoptosis: Molecular and pathological analyses, Am. J. Med. Genet. 91 (2000) 222–226.
[22] D. Van Gysel, R.L.P. Lijnen, S.S. Moekti, P.C.J. de Laat, A.P. Oranje, Collodion baby: a follow-up study of 17 cases, J. Eur. Acad. Dermatol. Venereol. 16 (2002) 472–475.
[23] H. Zhao, M. Keddache, L. Bailey, G.L. Arnold, G. Grabowski, Gaucher's disease: identification of novel mutant alleles and genotype–phenotype relationships, Clin. Genet. 64 (2003) 57–64.
[24] A. Fujimoto, N. Tayebi, E. Sidransky, Congenital ichthyosis preceding neurologic symptoms in two sibs with type 2 Gaucher disease, Am. J. Med. Genet. 59 (1995) 356–358.
[25] S. Miocic, M. Filocamo, S. Dominissini, A.L.E. Montalvo, K. Vlahovicek, M. Deganuto, R. Mazzotti, R. Cariati, B. Bembi, M.G. Pittis, Identification and functional characterization of five novel mutant alleles in 58 Italian patients with Gaucher disease type 1, Hum. Mutat. 25 (2005) 100.
[26] D.L. Stone, N. Tayebi, E. Orvisky, B. Stubblefield, V. Madike, E. Sidransky, Glucocerebrosidase gene mutations in patients with type 2 Gaucher disease, Hum. Mutat. 15 (2000) 181–188.
[27] N. Tayebi, K.J. Reissner, E.K. Lau, B.K. Stubblefield, A.C. Klineburgess, B.M. Martin, E. Sidransky, Genotypic heterogeneity and phenotypic variation among patients with type 2 Gaucher's disease, Pediatr. Res. 43 (1998) 571–578.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Genetics and Metabolism 86 (2005) 206–211
Molecular Genetics and Metabolism
www.elsevier.com/locate/ymgme

Clinical and mutational characterization of three patients with multiple sulfatase deficiency: Report of a new splicing mutation

Anna Díaz-Font [a], Raül Santamaría [a], Mònica Cozar [a], Mariana Blanco [b], Néstor Chamoles [b,†], Maria Josep Coll [c], Amparo Chabás [c], Lluïsa Vilageliu [a], Daniel Grinberg [a,*]

[a] *Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain*
[b] *Fundación para el Estudio de las Enfermedades Neurometabólicas, Buenos Aires, Argentina*
[c] *Institut de Bioquímica Clínica, Corporació Sanitària Clínic, Barcelona, Spain*

Received 12 July 2005; accepted 12 July 2005
Available online 26 August 2005

Abstract

Multiple sulfatase deficiency (MSD) is a rare autosomal recessive lysosomal storage disease characterized by impaired activity of all known sulfatases. The gene *SUMF1*, recently identified, encodes the enzyme responsible for post-translational modification of a cysteine residue, which is essential for the activity of sulfatases. Fewer than 30 MSD patients have been reported to date and 23 different mutations in the *SUMF1* gene have been identified. Here, we present the characterization of the mutant alleles of two Spanish and one Argentinean MSD patients. While the two Spanish patients were homozygous for the previously described mutations, c.463T>C (p.S155P) and c.1033C>T (p.R345C), the Argentinean patient was homozygous for the new mutation IVS7 + 5 G>T. A minigene approach was used to analyze the effect of the splice site mutation identified, due to the lack of sample from the patient. This experiment showed that this change altered the normal splicing of the RNA, which strongly suggests that this is the molecular cause of the disease in this patient.

Keywords: Multiple sulfatase deficiency; *SUMF1* gene; Minigene analysis; Splicing mutation

Introduction

Multiple sulfatase deficiency (MSD, OMIM #272200) is a rare autosomal recessively inherited lysosomal storage disorder characterized by the accumulation of sulfated lipids and acid mucopolysaccharides [1]. MSD combines the enzyme deficiency and phenotypic features of at least six entities [2]: metachromatic leukodystrophy, Maroteaux-Lamy syndrome, X-linked ichthyosis, Hunter syndrome, Sanfilippo A syndrome, and Morquio syndrome. The defect involves a modification process common to seven or more sulfatases [3]: the post-translational conversion of a cysteine to Cα-formylglycine (FGly), which is required for generating catalytically active sulfatases [4].

Dierks et al. [5] and Cosma et al. [6] identified the gene responsible for the disease, *SUMF1*, which encodes the FGly-generating enzyme (FGE) that catalyses the Cys to FGly post-translational modification of different sulfatases in the endoplasmic reticulum (ER). The human *SUMF1* gene has nine exons, spans 105 kb, and maps to chromosome 3p26 [5].

Only 27 MSD patients have been published to date. Dierks et al. [5] analyzed seven of them and identified

* Corresponding author. Fax: +34 93 403 4420.
E-mail address: dgrinberg@ub.edu (D. Grinberg).
† The authors are sad to inform that Dr. Chamoles passed away during the revision process of this manuscript.

1096-7192/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2005.07.004

nine different mutant alleles in the *SUMF1* gene, while Cosma et al. [6,7] analyzed 20 patients and characterized 38 mutant alleles, including 22 different mutations. In total, 23 different mutations and 49 mutant alleles were reported. Recently, the crystal structure of FGE has been determined [8]. In this work, the active site has been defined, a structure-based mechanism of the enzyme has been proposed and the effect of all missense mutations found in MSD patients has been explained by the FGE structure.

Here, we present clinical data and the mutation analysis of three MSD patients, two from Spain and one from Argentina, and report a new mutation, the IVS7 + 5G>T, which was shown by minigene analysis to affect the splicing process.

Subjects and methods

Patients

The *SUMF1* gene was analyzed in two Spanish patients (MSD1 and MSD2), diagnosed at the Institut de Bioquímica Clínica (Barcelona), and one Argentinean patient (MSD3), diagnosed at the Fundación para el estudio de las Enfermedades Neurometabólicas (Buenos Aires, Argentina).

Control DNA samples were obtained, anonymously, from the blood bank at the Hospital de la Vall d'Hebron, Barcelona.

Patient MSD1 was a 27-month-old boy who was admitted to hospital for evaluation of developmental delay. His parents were consanguineous. On medical examination he presented with coarse facial features (macrocephaly, straight forehead, and low-set ears). He had slight abnormalities of lower anterior ribs and feet, but no hepatosplenomegaly. Ichthyosis affected areas of scalp, neck, trunk, and legs. He was mentally retarded and was unable to walk or to stand without support. At the age of 7 years, the child was bedridden and quadriplegic. Seizures were frequent.

Patient MSD2, a boy, was referred at the age of 11 months because of delay in psychomotor development and myoclonia. His parents were consanguineous and had two other sons who had died when 3.5 years and 2.5 months old. The patient had coarse facies, moderate deafness, hepatosplenomegaly, and ichthyosis. He was hypotonic and unable to stand.

The Argentinean patient, MSD3 (Fig. 1), was a girl who was seen at the hospital because of failure to thrive and respiratory problems. She was born to non-consanguineous parents. Examination at 4 years of age revealed hypertrichosis, gingival hypertrophy, hepatomegaly, skeletal abnormalities, and erythematous lesions in scalp and arms. Muscle tone was loose. Psychomotor retardation was severe and her speech and manual skills were those of a 12–15-month-old child.

DNA preparation

Genomic DNA was prepared from harvested skin fibroblasts or peripheral blood leukocytes using the *Wizard Genomic DNA Purification Kit* (Promega, Madison, WI, USA).

PCR amplification and sequencing

The nine exons of the *SUMF1* gene were amplified by PCR using the primers described in Table 1. The PCR was performed using 1 U of *Taq DNA Polymerase* (Promega, Madison, WI, USA), 10 pmols of each primer, 1.5 mM $MgCl_2$, 200 μM dNTPs, and 100 ng of genomic DNA in the recommended buffer, as follows: initial

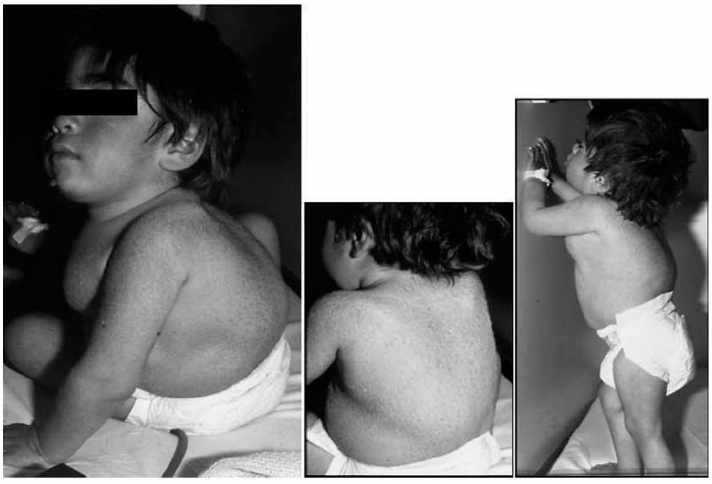


Fig. 1. Patient MSD3 presenting skin lesions particularly visible on the back and arms.

Table 1
Primers used for the amplification of *SUMF1* exons and flanking regions

Exon	Forward 5′ → 3′	Reverse 5′ → 3′
1	SUMF1-1F: GCTCAAAATTCTGCCCCTTT	SUMF1-1R: CTACTCCAACCCCGTCCAG
2	SUMF1-2F: CAGGTACACCTTGGCTGTGA	SUMF1-2R: TCTGCCACAGAATGCAGTAAA
3	SUMF1-3F: AGAGAAAAGGGCACACATGC	SUMF1-3R: CACCCAAACCCTTTTCAATG
4	SUMF1-4F: CATTGAGATTTTTGCATGT	SUMF1-4R:TTGTCAACTGGGGAAAGAGG
5	SUMF1-5F: CCAGCTCCTGCTTCTTGTTC	SUMF1-5R: GACCAACATATTGTCAGGCAAA
6	SUMF1-6F: CAAAACCTGTGCCTGTTTGA	SUMF1-6R: GAAGGGAACACCGTGTGAGT
7	SUMF1-7F: TCTTGCTGCATTTCTGCACT	SUMF1-7R: AGAGGACAGATGCCACCATT
8	SUMF1-8F: AGCTGGTTGGATCAAAGTCG	SUMF1-8R: TGGGGTTAGGTAGGCATTTG
9	SUMF1-9F: GTTGTGGGGAAGCACTGTCT	SUMF1-9R: ATGCACCACACCATAAAGCA

denaturation at 94 °C for 2 min, 35 cycles of denaturation (94 °C for 40 s) and annealing (54–56 °C for 30 s), and a final extension step (72 °C for 5 min). PCR fragments were purified with the *GFX PCR DNA and Gel Band purification Kit* (Amersham Pharmacia Biotech, Amersham, UK). Sequencing was performed using the *BigDye Terminator Cycle Sequencing v3.1* kit (PE Applied Biosystems, Foster City, CA, USA) in an *ABI PRISM 3700 DNA analyzer* (PE Applied Biosystems, Foster City, CA, USA).

Minigene constructions

Human genomic DNA was amplified from normal and mutant (IVS7 + 5G>T) *SUMF1* exon 7 to generate a fragment that contained exon 7 along with 327 bp of the 5′ and 405 bp of the 3′ intronic flanking sequences using the forward oligonucleotide, 5′-CATGTACATCCTC TCCCTACTCACCCAAT-3′; and the reverse one, 5′-C ATGTACACCCATCAGAACTCAGGCAAC-3′. Both oligonucleotides contain a *Bsr*GI restriction enzyme site at their 5′ ends, used to clone the product into a previously produced minigene plasmid. This plasmid, pGLB1 (used in our group for the study of splicing defects in GM-1 gangliosidosis patients) was based in a pcDNA3.1 vector into which a DNA fragment containing exons 7, 8, and 9 and introns 7 and 8 of the *GLB1* gene was cloned between the *Bam*HI and *Apa*I sites of the vector. Intron 7 of *GLB1* gene has a *Bsr*GI restriction site, which was used to clone *SUMF1* amplified fragments from the wild-type and the mutant alleles. The hybrid plasmids were named pGLB1-SUMF1wt and pGLB1-SUMF1mut.

Splicing assay

The splicing assay was performed by transfecting 1 μg of each minigene plasmid with 4 μl of Lipofectamine 2000 Reagent (Life technologies, Basel, Switzerland) into 2.5×10^5 HeLa cells grown in 6-well plates. Total RNA was isolated from cultured HeLa cells 24 h after transfection using the *QIAshredder* (QIAGEN, Hilden, Germany) and the *RNeasy Mini Kit* (QIAGEN, Hilden, Germany). RNA integrity was verified by 1% agarose gel electrophoresis, and its concentration was determined by OD 260 nm. Finally, total RNA was stored at −80 °C.

To obtain total cDNA from the samples, the reverse transcription reaction was carried out in a total volume of 25 μl containing 1 μg of total RNA, 1 μg of oligo-dT, 200 U of the *M-MLV Reverse Transcript RNAse H-*(Promega, Madison, WI, USA), 10 mM of each dNTP, 5 μl of M-MLV RT 5× Reaction Buffer (Promega, Madison, WI, USA), and nuclease-free water. The mixture was incubated at 42 °C for 1 h and then heated to 70 °C for 10 min to inactivate the reaction. The final cDNA was stored at −20 °C.

To amplify specifically those transcripts expressed from the plasmid, a PCR using specific plasmid primers was performed. Primers were T7-F 5′-AATACGACTC ACTATAGGGA-3′ and SP6-R 5′-CATTTAGGTGA CACTA-3′.

Results and discussion

For patient MSD1, leukocyte arylsulfatase A and B were not detectable, and arylsulfatase C was 4% of mean control values. Glycosaminoglycans (GAG) excretion in urine was increased: 47.1 mg/mmol creatinine (control values: 5.0 ± 2.7; range: 1.8–12.4) with high levels of heparan sulfate (HS), dermatan sulfate (DS), and keratan sulfate (KS). Urinary levels of sulfatide were also elevated: 16.7 nmols/mg creatinine (age-matched control values: 1.1 ± 0.8; range: 0.2–2.8).

For patient 2, plasma iduronosulfate sulfatase and leukocyte arylsulfatase A, B, and C were 9.7, 3.5, 2.7, and 11.3% of mean control values, respectively. GAG excretion in urine was increased: 32.2 mg/mmol creatinine (age-matched control values: 8.0 ± 3.7; range: 2.2–16.5) with high levels of heparan sulfate and dermatan sulfate.

In the case of patient MSD3, six sulfatases were analyzed. Activity of arylsulfatase A, arylsulfatase B, galactose-6-sulfatase, heparan sulfatase, *N*-acetyl glucosamine-6 sulfatase, and iduronate sulfatase was 5.8, 7.4, 2.5, 3.3, 4.4, and 5.1% of mean control values, respectively. All were measured in leukocytes except for the last one that was measured in dried blood spots on filter

paper, according to the method described in Chamoles et al. [9]. Thin layer chromatography of GAG showed presence of abnormally high levels of DS, HS, and KS in urine of patient 3 compared to control samples (not shown).

We amplified and sequenced the nine exons of the *SUMF1* gene from the three MSD patients and found all six mutant alleles. Each patient was homozygous for a different mutation (Table 2).

The first patient (MSD1) was homozygous for a c.463T>C substitution in exon 3, with serine 155 replaced by a proline (p.S155P). The change was confirmed by *Hin*fI restriction analysis. This mutation, which lies in the N-terminal subdomain, was reported by Cosma et al. [6,7] in two patients (patient 1 and 4) who were homozygous for this allele. There is some discrepancy regarding the clinical phenotype between the two publications by Cosma et al. [6]: while in the first one the phenotypes of the two patients were described as moderate and mild-moderate, in the discussion of the second work [7], homozygosity for S155P was considered to cause a severe phenotype. The phenotype for our patient MSD1 could be considered moderate, consistent with the first description of the two patients with the same genotype in Cosma et al. [6]. Regarding particular clinical features, there are differences between the two patients in Cosma et al. [7]. Patient MSD1, described here, shares mental retardation with both of them, seizures, hypotonia, and lack of visceromegaly with patient 1, and ichthyosis with patient 4. According to Dierks et al. [8], mutation p.S155P would cause the loss of three hydrogen bonds: the Ser-OH interaction with the side chain of His208, with the side chain of Ser357, and with the main chain NH of Phe156.

The second patient (MSD2) was homozygous for a c.1033C>T substitution in exon 9, with arginine 345 replaced by a cysteine (p.R345C). This mutation was also previously reported by Cosma et al. [6,7] in two homozygous and two compound heterozygous patients. The phenotype of the homozygotes p.R345C/p.R345C was considered to be mild-moderate [6], consistent with that of the MSD2 patient described here. This mutation, which lies in the C-terminal subdomain, was shown to have an apparently mild effect in functional analyses [7]. According to Dierks et al. [8], mutation p.R345C would cause the loss of four hydrogen bonds. The Arg345 side chain forms two hydrogen bonds with the side chain of Glu113 and two with the Tyr342 main chain carbonyl group, which are all lost. Cys345 may also interfere with the correct Cys235/Cys346 disulfide formation due to its proximity to Cys346.

The third patient (MSD3) was homozygous for a novel mutation, IVS7 + 5 G>T (Fig. 2A), which affects the splice donor site of intron 7. We analyzed the

Table 2
Genotypes for *SUMF1* mutations and polymorphisms in the MSD patients

Patient	Mutations	188G>A	1116C>T	1135AA>GT	1185C>T
MSD1	p.S155P/p.S155P	G/G	T/T	AA/AA	T/T
MSD2	p.R345C/p.R345C	G/G	C/C	GT/GT	C/C
MSD3	IVS7 + 5G>T/ IVS7 + 5G>T	G/G	C/C	GT/GT	C/C

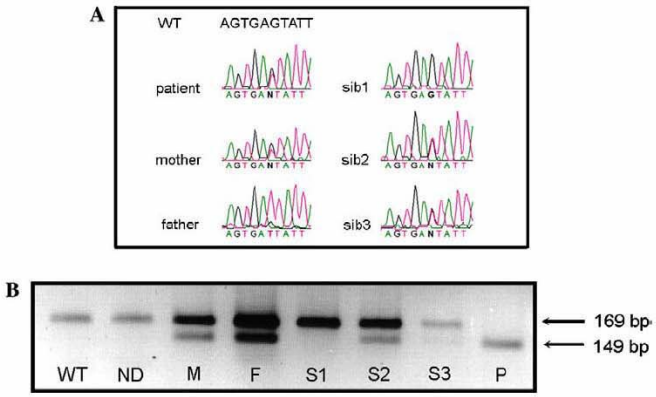


Fig. 2. Identification of mutation IVS7 + 5G>T in patient MSD3 and her family. (A) Chromatograms showing the G to T change in homozygosity in the patient, and in heterozygosity in both parents and siblings 2 and 3. (B) A reverse mismatch primer, in intron 7, was designed to amplify a 169 bp fragment including the position of the mutated nucleotide, when used together with a forward primer in intron 6. A *Vsp*I site is created when the IVS7 + 5G>T mutation is present. The wild-type allele produces a 169 bp-PCR fragment since it has no *Vsp*I site. Thus, the result is similar to the undigested sample. The 149 bp band corresponds to the mutant allele. WT, wild-type; ND, non-digested sample; M, mother; F, father; S, sibling; P, patient.

patient's family, and found the mutation, in heterozygosis, in both parents and in two of the three siblings. This mutation was confirmed by mismatch PCR and *Vsp*I restriction analysis (Fig. 2B). The mutation was absent in 182 control chromosomes.

Unfortunately, the lack of sample from this patient and from the family precluded mRNA analysis that could have shown that the change affects the splicing process. However, the change is very likely to be the molecular cause of the disease, since there are many examples of pathogenic mutations due to a nucleotide substitution at the +5 position of the splicing donor site. They include the substitution of the G (consensus) nucleotide by any of the other three. Just to name a few examples, we could mention the following: a G to A substitution in the IVS1 + 5G>A mutation of the β-globin gene causing β-thalassemia [10]; a G to C substitution in type-1 neurofibromatosis [11]; or a G to T substitution (like the one reported here) in factor-VII deficiency [12] or in β-thalassemia [13].

To confirm the hypothesis that the change produces aberrant splicing, we performed a minigene assay, using the pGLB1 plasmid in which *SUMF1* exon 7 and flanking regions, from the wild-type (pGLB1-SUMF1wt) or IVS7 + 5T (pGLB1-SUMF1mut) allele, was cloned (see Subjects and methods). Fig. 3A shows the results of this minigene assay. Control lanes (2) and (3) show the result of RT-PCR amplifications using the T7 and SP6 primers in cells transfected either with the intact pcDNA3.1 plasmid (2) or with the pGLB1 plasmid (3). Expected bands of 154 and 300 bp, respectively, were obtained. When plasmids pGLB1-SUMF1wt and pGLB1-SUMF1mut were assayed, different results were obtained. The wild-type allele band (425 bp) corresponds to a mature transcript in which *SUMF1* exon 7 was included (Fig. 3B: 4, grey box). However, the mutant allele gave rise to a longer PCR fragment (514 bp). Sequencing analysis showed that 89 bp of intron 7 were included in the mature mRNA of the mutant allele (Fig. 3B: 5, black box), indicating that the mutation prompts the splicing machinery to choose a cryptic donor site inside intron 7 (Fig. 3C). This cryptic site has a MaxEntScan score [14] of 3.64, intermediate between the *SUMF1* intron 7 wild-type donor site (8.29) and the mutant counterpart (−2.69). Although, we cannot rule out that in vivo a different cryptic site might be used (since only part of the *SUMF1* intron 7 was included in the minigene construct), the minigene experiment proved that the IVS7 + 5 G>T substitution affects normal splicing, which means it can be considered a pathogenic mutation.

In addition, in the search for mutations, we detected two new polymorphisms in exon 9 of the *SUMF1* gene: c.1135-1136 AA>GT and c.1185C>T. Frequencies were established by the analysis of 22 control chromosomes. For the c.1135-1136 AA>GT polymorphism, frequencies were 45.5% for AA and 45.5% for GT. A third allele, AT,

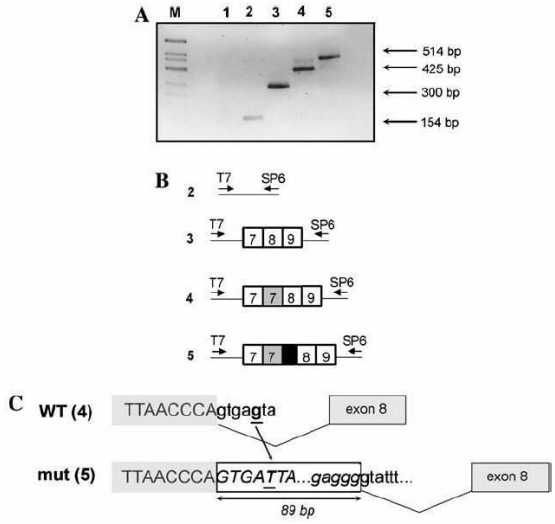


Fig. 3. Minigene assay for mutation IVS7 + 5G>T. (A) RT-PCR amplification on RNA extracted from HeLa cells untransfected (1) or transfected with different plasmids (2–5): intact pcDNA3.1 (2), pGLB1 plasmid (3), pGLB1-SUMF1wt (4), and pGLB1-SUMF1mut (5). M: molecular weight marker. (B) Diagram of the different amplification products. Numbers correspond to lanes in A. Sites for primers T7 and SP6 are indicated. White boxes: GLB1 exons. Grey box: exon 7 of the SUMF1 gene. Black box: part of SUMF1 intron 7 included in the mature RNA as a consequence of the IVS7 + 5G>T change. (C) Detailed sequence corresponding to the 3′ end of *SUMF1* exon 7 (grey) and 5′ end of the following intron of constructs bearing the wt (4) or the mutant (5) allele. The G to T change is indicated in bold and underlined. Capital letters indicate exonic sequence. Boxed sequence in italics corresponds to the extended *SUMF1* exon 7 (black box in B).

was found only twice (9%); and for the c.1185C>T polymorphism, C was at 54.5% and T, 45.5%.

We then performed a haplotype analysis with these two new polymorphisms and two previously described by Dierks et al. [5] (c.188G>A and c.1116C>T). Two haplotypes, G-T-AA-T and G-C-GT-C, were found (Table 2). The first haplotype was present in the alleles bearing the p.S155P mutation and the second one in the alleles bearing either the p.R345C or the IVS7 + 5 G>T mutation. Mutations p.S155P and p.R345C were found several times by Cosma et al. [6,7] in 4 and 6 alleles, respectively, out of the 40 mutant alleles analyzed. It would be interesting to perform a haplotype analysis of these patients to find out if exists a single origin for each of these mutations. In this regard, Cosma et al. [6] found the p.R345C mutation only in patients from Sicily and a haplotype analysis performed in some of them suggested a founder effect. However, haplotype data from the Sicilian patients are not available.

In conclusion, we increased by three the small number of MSD patients for whom a mutational analysis has been performed. We identified six mutant alleles and reported a new mutation. This mutation (IVS7 + 5G>T) affects position +5 of the consensus donor splice site,

which leads to an alteration of the normal splicing of the gene, as shown by a minigene assay approach.

Acknowledgments

We are grateful to Dr. J.M. Pérez (Hospital Casa de Maternitat, Barcelona) for the clinical description of patient MSD2. We are also grateful to R. Rycroft for revising the English. The technical assistance of J. Jarque and H. Sellés is gratefully acknowledged. A.D.F. is a recipient of a fellowship from the Spanish Ministerio de Educación y Cultura and R.S. is a recipient of a fellowship from the Spanish Ministerio de Ciencia y Tecnologia. The research was supported by CICYT (SAF 2000-0200 and SAF 2003-00386) and FIS (Redes Temáticas, G03/054 REDEMETH).

References

[1] J.J. Hopwood, A. Balabio, Multiple sulfatase deficiency and the nature of the sulfatase family, in: 8th ed, C.R. Scriver, A.L. Beaudet, W.S. Sly (Eds.), The Metabolic and Molecular Bases of Inherited Disease, vol. III, McGraw-Hill, New York, 2001, pp. 3725–3732.

[2] H. Kihara, Genetic heterogeneity in metachromatic leukodystrophy, Am. J. Hum. Genet. 34 (1982) 171–181.

[3] W. Rommerskirch, K. von Figura, Multiple sulfatase deficiency: catalytically inactive sulfatases are expressed from retrovirally introduced sulfatase cDNAs, Proc. Natl. Acad. Sci. USA 89 (1992) 2561–2565.

[4] B. Schmidt, T. Selmer, A. Ingendoh, K. von Figura, A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency, Cell 82 (1995) 271–278.

[5] T. Dierks, B. Schmidt, L.V. Borissenko, J. Peng, A. Preusser, M. Mariappan, K. von Figura, Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme, Cell 113 (2003) 435–444.

[6] M.P. Cosma, S. Pepe, I. Annunziata, R.F. Newbold, M. Grompe, G. Parenti, A. Ballabio, The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases, Cell 113 (2003) 445–456.

[7] M.P. Cosma, S. Pepe, G. Parenti, C. Settembre, I. Annunziata, R. Wade-Martins, C. Di Domenico, P. Di Natale, A. Mankad, B. Cox, G. Uziel, G.M. Mancini, E. Zammarchi, M.A. Donati, W.J. Kleijer, M. Filocamo, R. Carrozzo, M. Carella, A. Ballabio, Molecular and functional analysis of SUMF1 mutations in multiple sulfatase deficiency, Hum. Mutat. 23 (2004) 576–581.

[8] T. Dierks, A. Dickmanns, A. Preusser-Kunze, B. Schmidt, M. Mariappan, K. von Figura, R. Ficner, M.G. Rudolph, Molecular basis for multiple sulfatase deficiency and mechanism for formylglycine generation of the human formylglycine-generating enzyme, Cell 121 (2005) 541–552.

[9] N.A. Chamoles, M. Blanco, D. Gaggioli, Diagnosis of alpha-L-iduronidase deficiency in dried blood spots on filter paper: the possibility of newborn diagnosis, Clin. Chem. 47 (2001) 2192.

[10] S. Danckwardt, G. Neu-Yilik, R. Thermann, U. Frede, M.W. Hentze, A.E. Kulozik, Abnormally spliced beta-globin mRNAs: a single point mutation generates transcripts sensitive and insensitive to nonsense-mediated mRNA decay, Blood 99 (2002) 1811–1816.

[11] M. Baralle, D. Baralle, L. De Conti, C. Mattocks, J. Whittaker, A. Knezevich, C. Ffrench-Constant, F.E. Baralle, Identification of a mutation that perturbs NF1 agene splicing using genomic DNA samples and a minigene assay, J. Med. Genet. 40 (2003) 220–222.

[12] F.H. Herrmann, K. Wulff, K. Auberger, V. Aumann, F. Bergmann, K. Bergmann, E. Bratanoff, D. Franke, M. Grundeis, W. Kreuz, H. Lenk, H. Losonczy, B. Maak, G. Marx, C. Mauz-Korholz, H. Pollmann, M. Serban, A. Sutor, G. Syrbe, G. Vogel, N. Weinstock, E. Wenzel, K. Wolf, Molecular biology and clinical manifestation of hereditary factor VII deficiency, Semin. Thromb. Hemost. 26 (2000) 393–400.

[13] M.D. Bashyam, L. Bashyam, G.R. Savithri, M. Gopikrishna, V. Sangal, A.R. Devi, Molecular genetic analyses of beta-thalassemia in South India reveals rare mutations in the beta-globin gene, J. Hum. Genet. 49 (2004) 408–413.

[14] G. Yeo, C.B. Burge, Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals, J. Comput. Biol. 11 (2004) 377–394.