Available online at www.sciencedirect.com
ScienceDirect
Molecular Genetics and Metabolism 92 (2007) 122–130
Molecular Genetics and Metabolism
www.elsevier.com/locate/ymgme

Identification of the molecular defects in Spanish and Argentinian mucopolysaccharidosis VI (Maroteaux–Lamy syndrome) patients, including 9 novel mutations

Elena Garrido [a,b,c], Amparo Chabás [b,d], Maria Josep Coll [b,d], Mariana Blanco [e], Carmen Domínguez [b,f], Daniel Grinberg [a,b,c], Lluïsa Vilageliu [a,b,c], Bru Cormand [a,b,c,*]

[a] *Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, edifici annex, 3ª planta, E-08028 Barcelona, Catalonia, Spain*
[b] *CIBERER, Instituto de Salud Carlos III, Barcelona, Catalonia, Spain*
[c] *Institut de Biomedicina de la Universitat de Barcelona (IBUB), Catalonia, Spain*
[d] *Institut de Bioquímica Clínica, Hospital Clínic, Corporació Sanitària Clínic, Barcelona, Catalonia, Spain*
[e] *Laboratorio de Neuroquímica Dr. N. A. Chamoles, Buenos Aires, Argentina*
[f] *Centre d'Investigacions en Bioquímica i Biologia Molecular, Hospital Vall d'Hebron, Barcelona, Catalonia, Spain*

Received 21 March 2007; received in revised form 7 June 2007; accepted 7 June 2007
Available online 20 July 2007

Abstract

Maroteaux–Lamy syndrome, or mucopolysaccharidosis VI (MPS VI), is an autosomal recessive lysosomal storage disorder caused by a deficiency of *N*-acetylgalactosamine-4-sulfatase or arylsulfatase B (ARSB). We aimed to analyze the spectrum of mutations responsible for the disorder in Spanish and Argentinian patients, not previously studied. We identified all the ARSB mutant alleles, nine of them novel, in 12 Spanish and 4 Argentinian patients. The new changes were as follows: six missense mutations: c.245T>G [p.L82R], c.413A>G [p.Y138C], c.719C>T [p.S240F], c.922G>A [p.G308R], c.1340G>T [p.C447F] and c.1415T>C [p.L472P]; one nonsense mutation: c.966G>A [p.W322X]; and two intronic changes involving splice sites: c.1142+2T>A, in the donor splice site of intron 5, which promotes skipping of exon 5, and c.1143−1G>C, which disrupts the acceptor site of intron 5, resulting in skipping of exon 6. We also report 10 previously described mutations as well as several non-pathogenic polymorphisms. Haplotype analysis indicated a common origin for most of the mutations found more than once. Most of the patients were compound heterozygotes, whereas only four of them were homozygous. These observations confirm the broad allelic heterogeneity of the disease, with 19 different mutations in 16 patients. However, the two most frequent mutations, c.1143−1G>C and c.1143−8T>G, present in both populations, accounted for one-third of the mutant alleles in this group of patients.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Maroteaux–Lamy syndrome; Mucopolysaccharidosis VI; MPS VI; *N*-Acetylgalactosamine-4-sulfatase; Arylsulfatase B; *ARSB* gene; Mutational analysis; Haplotype analysis

Mucopolysaccharidosis VI (MPS VI) or Maroteaux–Lamy syndrome (MIM# 253200) is a rare lysosomal, autosomal recessive storage disorder caused by a deficiency of *N*-acetylgalactosamine-4-sulfatase (EC 3.1.6.12) or arylsulfatase B (ARSB), a lysosomal enzyme involved in the degradation of dermatan sulfate. Intralysosomal accumulation of this glycosaminoglycan causes progressive damage mainly in skeletal tissues. Patients usually present with dwarfism, rough facies, dysostosis multiplex and diverse skeletal defects, as well as aortic valve dysfunction, hepatosplenomegaly and corneal clouding, without mental involvement, although psychomotor retardation has been reported [1]. The clinical severity of the disease ranges from relatively attenuated to severe forms and at least three distinct ages at onset have been differentiated: infantile (severe

* Corresponding author. Address: Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, edifici annex, 3ª planta, E-08028 Barcelona, Catalonia, Spain. Fax: +34 93 4034420.
E-mail address: bcormand@ub.edu (B. Cormand).

1096-7192/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2007.06.002

form, with onset before 2 years of age and rapid progression), juvenile (intermediate form, with onset in late childhood) and adult (mild form, with onset after the second decade and slower progression).

Characterization of the *ARSB* cDNA [2] allowed the identification of the first mutations in MPS VI patients [3]. The *ARSB* gene, located on chromosome 5q13–q14 [4], contains 8 exons and spans over 209 kb (genome.ucsc.edu) [5]. The crystal structure of the encoded enzyme, 4-sulfatase has been resolved [6]. The nascent polypeptide has 533 amino acids and the mature form of the protein, with a molecular weight of 57 kDa, comprises two 7 and 8 kDa peptides that are disulfide-linked to a 43-kDa species [7].

To date, more than 70 mutations have been identified in the *ARSB* gene, most of which are missense changes. These mutations have been reported in several populations: Russia [8–10], Australia [3,11,12], USA [8,13–16], France [8,15,17,18], Germany [8,17], Portugal [15], Italy [19,20], China [21–23], Brazil and Chile [15,24,25].

Here, we present the identification of all the disease-causing mutations in 16 Spanish and Argentinian MPS VI patients by direct sequencing of the coding region and the splice sites of the *ARSB* gene. Most of the patients were compound heterozygotes, with only four individuals bearing homozygous mutations.

Materials and methods

Subjects

We studied 12 Spanish (ML1–ML10 and ML15–ML16) and 4 Argentinian (ML11–ML14) unrelated MPS VI patients, whose parents were not consanguineous. The Spanish patients were diagnosed at the *Institut de Bioquímica Clínica* (Barcelona), except for patient ML16, who was diagnosed at the *Hospital de la Vall d'Hebron* (Barcelona). The Argentinian patients were diagnosed at the *Laboratorio de Neuroquímica Dr. N. A. Chamoles* (Buenos Aires). Diagnosis was based on identification of dermatan sulfate in urine and the enzymatic measurement of ARSB activity. When parental samples were available, their carrier status was confirmed.

Genomic DNA isolation

Genomic DNA was obtained from fibroblasts (individuals ML1–ML5) or from whole blood (the rest of the patients, and one or both parents in most cases) using the *Wizard® Genomic DNA Purification Kit* (Promega, Madison, WI, USA). The genotype of patient ML7 was inferred from parental results. Patients ML9a and ML9b, and ML13a and ML13b were affected sibs and mutation analysis was carried out in both sibs and also in their parents. Informed consent was obtained from all participants and the protocols were approved by the Ethical Committee of the University of Barcelona, in accordance with the Helsinki Declaration.

Mutation analysis

The coding region and all the exon–intron boundaries of the *ARSB* gene were amplified by PCR in 8 fragments. The intronic primers were designed from the genomic sequence of *ARSB* using the Primer3 software (frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primer sequences are available from the authors upon request. The PCRs consisted of 100 ng of template DNA, 1 U of *Taq* DNA Polymerase (Promega, Madison, WI, USA), 1.75 mM $MgCl_2$ in the recommended buffer, 100 μM of dNTPs and 20 pmol of each primer in a final volume of 50 μl. For the amplification of exon 1, 10% DMSO was added. The reactions were subjected to 10 s at 98 °C, and 35 cycles of denaturation at 96 °C for 15 s, annealing at 53 °C for 40 s and elongation at 72 °C for 40 s, with a final extension step at 72 °C for 5 min. The PCRs were performed in a GeneAmp® PCR System 2700 thermocycler (PE Applied Biosystems, Foster City, CA, USA). The PCR products were column-purified with the *GFX® PCR DNA and Gel Band Purification Kit* (Amersham Pharmacia Biotech, Amersham, UK), sequenced with the *ABI Prism BigDye® Terminator Cycle Sequencing v.3.1* kit (Applied Biosystems, Foster City, CA, USA), purified with Sephadex G-50 (Amersham Biosciences, Molsheim, France) and run in an *ABI PRISM 3700 DNA analyzer* (Applied Biosystems, Foster City, CA, USA). Sequence chromatograms were analyzed with the Seq-Man™ II software (DNASTAR Inc., Madison, WI, USA).

All mutations, except c.899_1142del, were confirmed by overnight digestion of the corresponding PCR product with 5–10 U of the restriction enzyme indicated (Table 1), followed by electrophoresis on a 2% agarose gel or 8–12% polyacrylamide gel and ethidium bromide staining. In some cases we used the same primers as in the mutational screening, whereas in other cases we designed mismatched primers to generate the desired restriction site (Table 1). Fifty control samples were screened for the mutations identified in the patients, following the same restriction protocol.

The cDNA sequence numbering follows Genbank Accession No. M32373.1, with nucleotide 560, the adenine of the ATG start codon, corresponding to position +1. The protein numbering is based on the reference sequence NP_000037.2. Mutation nomenclature follows HGVS guidelines (http://www.hgvs.org).

RT-PCR analysis

Total RNA was obtained from cultured fibroblasts of patients ML1, ML2, ML3, ML4, ML5 and the parents of ML9, using the *Qiaschredder® Kit* and the *RNeasy® MiniKit* (QIAGEN, Hilden, Germany). Single-stranded cDNA was synthesized by RT-PCR as follows: 3 μl of RNA was incubated with 2 μl of oligo(dT) (0.5 μg/μl) (Amersham Pharmacia Biotech, Amersham, UK) in a final volume of 15 μl at 70 °C during 5 min. cDNA synthesis was performed with 200 U of M-MLV/H(−) Point reverse transcriptase (Promega, Madison, WI, USA), 10% of the recommended buffer (5×) and 2.5 mM dNTPs in a final volume of 25 μl for 1 h at 42 °C, and the reaction was stopped by incubation at 70 °C for 10 min. Amplification of a fragment spanning exons 4–8 of the *ARSB* cDNA was performed using the exonic primers 5′-CAGAG AAGCCTCTGTTTCTC-3′ (forward) and 5′-CTTTGTGACGATGTGA GGA-3′ (reverse). Amplification of β-actin cDNA was performed using commercially available primers (*Human β-Actin Control Amplimer Set*, BD Biosciences Clontech, Palo Alto, CA, USA). The *Discovery Series Quantity One 1-D Analysis Software* (http://www.bio-rad.com, Hercules, CA, USA) was used to quantify the relative intensity of gel bands.

Haplotype analysis

Haplotype analysis, including 7 polymorphic sites along the *ARSB* gene, was performed by sequencing genomic PCR products of patients bearing mutations present more than once in our series (c.1143−8T>G, c.1143−1G>C, c.427delG and p.C447F). The polymorphisms selected were representative of five distinct haplotypic blocks, as defined by the SNP Browser 3.5 software (PE Applied Biosystems, Foster City, CA, USA) using a Linkage Disequilibrium Unit (LDU) threshold of 0.3: rs3733895 (block 1, intron 1), rs6870443 (block 2, intron 3), rs1065757 and rs17220759 (block 3, exon 5), rs25414 and rs25413 (block 4, exon 6) and rs3822473 (block 5, intron 7). Two of the polymorphic sites (rs1065757 or p.V358M and rs25414 or p.S384N) were genotyped in 50 healthy Spanish individuals to determine their allelic frequencies by PCR amplification of genomic DNA and restriction analysis. The PCR conditions were the same as described in the "Mutation analysis" section. For p.V358M, we used the primers 5′-GGGAATTTAGGGTGGG

Table 1
Primers and enzymes used for PCR and restriction analysis of the *ARSB* mutant alleles identified in the present study

Mutation	Enzyme	Primers (position)[b]	Fragment length (bp)
238delG c.237-243delGGTGCTC p.L82R[a] p.R95Q	BsiHKAI (−) BsiHKAI (−) MvaI (−) BstUI (−)	F: 5′-TTCCTCATTCTATCAGCGGTACAAG-3′ (E1) R: 5′-GAGAAGCCGCCGGGACCCATAACT-3′ (I1)	522
p.Y138C[a] c.427delG p.R160Q p.R160X	RsaI (+) BslI (+) StyI (+) DdeI (−)	F: 5′-GAAGGCCATTTTATCTGCTTGT-3′ (I1) R: 5′-TGATTGCACTTGGGTGTGTT-3′ (I2)	337
p.S240F[a]	HinfI (+)(mismatch)	F: 5′-TCTCTACCTTGCTCTCGAGT-3′ (E4) R: 5′-GCTAACCGCTCCAATTTGTC-3′ (I4)	289
p.G308R[a]	StyI (+)(mismatch)	F: 5′-ACGGAGGGCAGACTTTGCCA-3′ (E5) R: 5′-TCAGGCTGCTCTTGGAGTTT-3′ (I5)	296
p.P313A p.R315Q c.1142+2T>A[a]	ApaI (+) TaqI (−) KpnI (−)	F: 5′-GGGAATTTAGGGTGGGAAAA-3′ (I4) R: 5′-TCAGGCTGCTCTTGGAGTTT-3′ (I5)	444
p.W322X[a]	DdeI (+)(mismatch)	F: 5′-GGAAGAAAATGGAGCCTCTG-3′ (E5) R: 5′-TCAGGCTGCTCTTGGAGTTT-3′ (I5)	209
c.1143−1G>C[a]	DdeI (+)(mismatch)	F: 5′-TGTGGAGACCTCCAAATTCA-3′ (I5) R: 5′-GGGGATGGGCTTCCCTCA-3′ (I5/E6)	158
c.1143−8T>G	BstEII (+)(mismatch)	F: 5′-TTCAAAGGGTCCCAGAATCA-3′ (I5) R: 5′-TGGGCTTCCTTCACTGGTAA-3′ (I5/E6)	285
p.C447F[a]	HindIII (+)(mismatch)	F: 5′-TTGTTTTCTTTGCTAAGCT-3′ (I7/E8) R: 5′-AAAAGGCCTGAGGTCCAACT-3′ (E8)	337
p.L472P[a]	MvaI (+)	F: 5′-ATGTTTCCACACCCACAACC-3′ (I7) R: 5′-AAAAGGCCTGAGGTCCAACT-3′ (E8)	430

(+) Creates a restriction site, (−) abolishes a restriction site. Mutation nomenclature follows HGVS recommendations (http://www.hgvs.org). cDNA sequence numbering is according to GenBank Accession No. M32373.1, with nucleotide 560 as the A of the ATG start codon. Protein numbering is based on the reference sequence NP_000037.2.

[a] Novel mutation.

[b] E, exon; I, intron.

AAAA-3′ (forward) and 5′-TCAGGCTGCTCTTGGAGTTT-3′ (reverse), and the PCR products were digested with BspHI and electrophoresed on a 2% agarose gel. For p.S384N, we used a forward mismatched primer, 5′-TTGTTTTCCAGTGAAGCAA-3′, and the reverse primer 5′-GTTCAATGGCTTAAAGGCTA-3′. The PCR products were digested with BstEII and electrophoresed on 12% polyacrylamide.

Results

Clinical and biochemical findings

The clinical presentation of all patients ranged from intermediate to severe (Table 2). We found no case with the mild form of the disease. Enzyme activity measured in leukocytes, cultured skin fibroblasts or dried blood spots on paper filter ranged from about 1 to 23% of that of mean normal values (Table 2).

Analysis of the ARSB gene in MPS VI patients

The entire open reading frame of the *ARSB* gene and all the exon–intron boundaries were analyzed in 16 patients by direct sequencing of genomic PCR products. As a result, all 32 mutant alleles were identified (Table 2). In total, 19 distinct mutations were found, nine of which were novel (Fig. 1). The new mutant alleles were six missense mutations (p.L82R, p.Y138C, p.S240F, p.G308R, p.C447F and p.L472P), one nonsense mutation (p.W322X) and two intronic changes affecting splice sites (c.1143−1G>C and c.1142+2T>A). Screening of 100 chromosomes from healthy Spanish individuals failed to detect these mutations. In addition, we found several previously described mutations in this series of patients: c.238delG [11], c.237-243delGGTGCTC [9], p.R95Q [12], c.427delG [14], p.R160Q and p.R160X [9], p.P313A [26], p.R315Q [20], c.1143−8T>G [25] and a genomic deletion including exon 5, the limits of which have not been defined [17]. Fig. 1 shows the position along the gene of all the mutations found in our study.

The most prevalent mutation was c.1143−1G>C, accounting for 21.9% (7/32) of the mutant alleles, followed by c.1143−8T>G which represents 12.5% (4/32). All other mutant alleles were found only once or twice. Overall, 25% of the patients bore homozygous mutations (ML2, ML5, ML8 and ML13), whereas the rest were compound heterozygotes.

Table 2
Genotypes, clinical data and biochemical analysis of the patients included in the study

Patient	Origin	Genotype (cDNA)	Predicted effect (protein)	Age at onset/ diagnosis	Clinical signs	Clinical classification[d]	ARSB activity (% of control)[e]	GAG[f]
ML1	Spain	c.[937C>G] + [899_1142del][a]	p.[P313A] + [D300VfsX193]	?/4y	HC, GR, CC, CFF	Intermediate	17 (L), 8 (F)	ND
ML2	Spain	c.[1143−8T>G][b] + [1143−8T>G][b]	p.[S381RfsX12] + [S381RfsX12]	Birth/2y	H, PMR, CFF, SK, CC, IH	Intermediate-severe	12 (F)	22.15
ML3	Spain	c.[479G>A] + [1143−1G>C][b]	p.[R160Q] + [S381RfsX12]	4y/5y	H, SK, JS	Intermediate	19 (F)	32.77
ML4	Spain	c.[427delG] + [966G>A]	p.[V143SfsX41] + [W322X]	Birth/14y	CCO, SK, CFF, PMR, HC, GR, HS, IH	Intermediate-severe	23 (F)	50.8
ML5	Spain	c.[245T>G] + [245T>G]	p.[L82R] + [L82R]	2y/3y	HS, CFF, SK, IH, JS, GR	Intermediate	12 (L), 8 (F)	10.39
ML6	Spain	c.[1415T>C] + [1143−1G>C][b]	p.[L472P] + [S381RfsX12]	?/7y	MC, slight VD, CFF, SK	Intermediate	12 (L)	ND
ML7	Spain	c.[237-243delGGTGCTC] + [1143−8T>G][b]	p.[V80WfsX32] + [S381RfsX12]	4y/11y	SK, PMR, H, CCO, MC	Severe	13 (L)	18.32
ML8	Spain	c.[478C>T] + [478C>T]	p.[R160X] + [R160X]	Birth/9y	VD, CFF, CC, slight SK, CCO	Intermediate	8 (L)	52
ML9a/b	Spain	c.[238delG] + [1142+2T>A][c]	p.[V80CfsX34] + [D300VfsX193]	?/19m	GR, CFF	Intermediate	8/4 (L), 16/17 (F)	16.69/14.24
ML10	Spain	c.[1340G>T] + [427delG]	p.[C447F] + [V143SfsX41]	Birth/8m	PMR, CFF, severe VD, IH, facial paralysis, JS, CCO	Severe	<1 (L)	ND
ML11	Argentina	c.[413A>G] + [1143−8T>G][b]	p.[Y138C] + [S381RfsX12]	3m/7y	MC, CFF, IH, PMR, SK, HS, RR, UH, JS	Severe	8 (L)	ND
ML12	Argentina	c.[284G>A] + [1143−1G>C][b]	p.[R95Q] + [S381RfsX12]	2y/6y	HS, CFF, CCO, VD, RR, SK	Intermediate	10 (L)	ND
ML13a/b	Argentina	c.[1143−1G>C][b] + [1143−1G>C][b]	p.[S381RfsX12] + [S381RfsX12]	1m/1y; birth/4y	ML13a: CFF, HS, SK, GR, CC, UH, JS, RR, recurrent otitis; ML13b: H, CFF, UH, RR	Severe	7/7 (L)	ND
ML14	Argentina	c.[944G>A] + [922G>A]	p.[R315Q] + [G308R]	1y/3y7m	CFF, SK, H, VD, JS, severe visual impairment, RR	Intermediate-severe	5 (L)	ND
ML15	Spain	c.[1340G>T] + [1143−1G>C][b]	p.[C447F] + [S381RfsX12]	4m/5y	HS, CFF, SK, RR	Intermediate	1 (F)	122.08
ML16	Spain	c.[719C>T] + [1143−1G>C][b]	p.[S240F] + [S381RfsX12]	?/20m	CFF, SK, HS, MC, JS, CC	Intermediate	13.5 (L), 10 (F)	ND

CC, corneal clouding; CCO, cardiac complications; CFF, coarse facial features; GR, growth retardation; H, hepatomegaly; HC, hydrocephalus; HS, hepatosplenomegaly; IH, impaired hearing; MC, macrocephaly; JS, joint stiffness; PMR, psychomotor retardation; RR, recurring rhinitis; SK, skeletal abnormalities; UH, umbilical hernia; VD, vascular disease; ND, not determined. Mutation nomenclature follows HGVS recommendations (http://www.hgvs.org). cDNA sequence numbering is according to GenBank Accession No. M32373.1, with nucleotide 560 as the A of the ATG start codon. Protein numbering is based on the reference sequence NP_000037.2.

[a] Genomic deletion of exon 5.

[b] Exon 6 skipping.

[c] Exon 5 skipping.

[d] Based on the extent of organ involvement, the presence of psychomotor retardation and the age at death.

[e] ARSB activity in the patients' leukocytes (L) or skin fibroblasts (F) referred to the mean value of several controls analyzed in the same experiment. Spanish patients: ARSB activity measured in L (normal values: 53–149 nmol/h/mg) or F (normal values: 210–650 nmol/h/mg). Argentinian patients: ARSB activity measured in dried blood spots on filter paper: μmol/h/l (normal values: 5.7–28.8 μmol/h/l).

[f] Urine glycosaminoglycans (GAG) concentration, expressed as mg GAG/mmol creatinine. Normal values: 8.01 ± 3.67 (0–1 years); 4.15 ± 2.07 (2–4 years); 3.02 ± 1.6 (4–10 years); 2.76 ± 0.99 (10–15 years); 1.48 ± 0.87 (15–20 years).

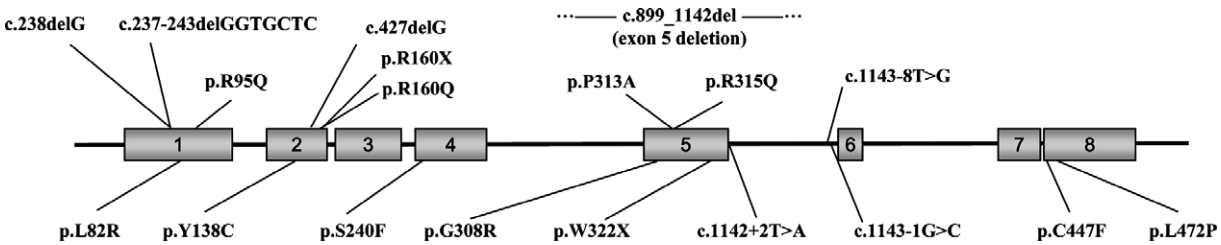

Fig. 1. Mutations identified in the *ARSB* gene in Spanish and Argentinian MPS VI patients. Gray boxes correspond to exons. Exons and introns are drawn to different scales. The positions of the mutations described in this study are shown: top, mutations previously described; below, mutations described here for the first time.

Mutations affecting RNA integrity or stability

Several of the mutations identified may affect the integrity of the *ARSB* mRNA transcript. Thus, skin fibroblasts from patients ML1, ML2, ML3, ML4 and ML5, and from the parents of patient ML9 were used for RNA experiments. ML1 showed a genomic deletion that included exon 5, ML2 (c.1143−8T>G) and ML3 (c.1143−1G>C) carried intronic nucleotide substitutions that altered the intron 5 acceptor splice site, either at the 100% conserved AG dinucleotide or at the pyrimidine track, and ML9 (c.1142+2T>A) had an intronic change that affected the donor splice site GT consensus of the same intron. RT-PCR and sequence analysis demonstrated the absence of exon 5 in *ARSB* RNA from patient ML1 (Fig. 2, faint lower band of 565 bp, indicated by an arrow). A similar band was observed for the father of patient ML9 but, in this case, it was generated by skipping of exon 5, caused by a splicing mutation. In patients ML2 and ML3, alterations in the splice sites of c.1143−8T>G and c.1143−1G>C produced the skipping of exon 6. Thus, a band of 738 bp was observed in these two patients (Fig. 2). The removal of either exon 5 or 6 produced a frameshift, leading to truncated proteins of 491 and 391 aa, respectively. Patient ML5, homozygous for a missense mutation (p.L82R), was used as a control for RNA integrity.

Mutations generating premature termination codons (PTC) located at a minimum of about 50 bp upstream from the last exon–exon junction of the mRNA are predicted to induce the degradation of the transcript by nonsense-mediated mRNA decay (NMD) [27]. Our RT-PCR results indicate that mutations c.899_1142del (patient ML1),

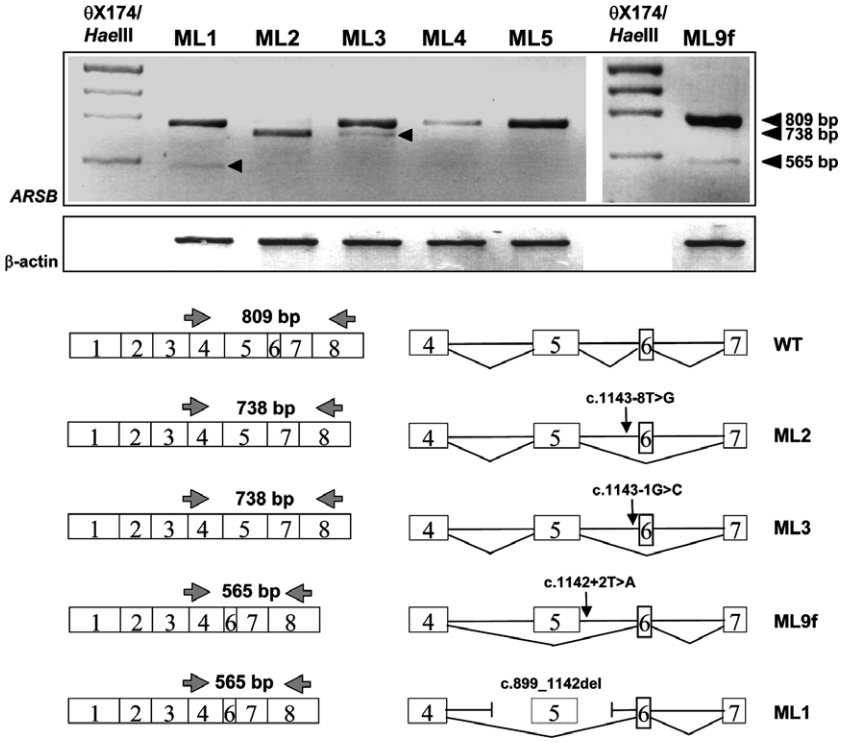

Fig. 2. RT-PCR experiments on skin fibroblasts RNA from Maroteaux–Lamy patients. The amplified fragment, 809 bp in length, spans exon 4–8 of the *ARSB* cDNA. The absence of exon 6 (71 bp) in patients ML2 and ML3 results in a lower 738-bp band. Patient ML1 and the father of patient ML9 (ML9f) display a 565-bp band, which corresponds to the lack of exon 5 (244 bp). Patient ML4 shows a fainter band of normal size, and ML5 shows a band of normal size and intensity.

Table 3
Haplotype analysis of patients bearing mutations present more than once in our series (c.1143−8T>G, c.427delG, c.1143−1G>C and p.C447F)

		Polymorphisms, allelic frequencies and haplotypes[a]						
		c.313-26T>C rs3733895 (I1)	c.691-22T>C rs6870443 (I3)	p.V358M rs1065757 (E5)	p.V376M rs17220759 (E5)	p.S384N rs25414 (E6)	p.P397P rs25413 (E6)	c.1337-32C>G rs3822473 (I7)
Mutation	**Patient and origin[b]**	T: 0.608 C: 0.392 Num. chr: 120	T: 0.833 C: 0.167 Num. chr: 120	G:0.568 A: 0.432 Num. chr: 118	G:0.833 A: 0.167 Num. chr: 48	G:0.933 A:0.067 Num. chr:120	G:0.839 A:0.161 Num. chr:118	C:0.967 G:0.033 Num. chr:120
c.1143-8T>G	ML2 (Sp)	T	C	G	G	G	A	C
	ML2 (Sp)	T	C	G	G	G	A	C
	ML7 (Sp)	T/C	T/C	G/A	G	G	G/A	C
	ML11 (A)	T	C	G	G	G	A	C
c.427delG	ML4 (Sp)	C	T	G	A	G	G	C
	ML10 (Sp)	C	T	G	A	G	G	C
c.1143-1G>C	ML3 (Sp)	T	T/C	G/A	G	G	G	C
	ML6 (Sp)	T	T	A	G	G	G	C
	ML12 (A)	T	T	A	G	G	G	C
	ML13 (A)	T	T	A	G	G	G	C
	ML13 (A)	T	T	A	G	G	G	C
	ML15 (Sp)	T	T	A	G	G	G	C
	ML16 (Sp)	T	T	A	G	G	G	C
p.C447F	ML10 (Sp)	T/C	T	A	G	G	G	C
	ML15 (Sp)	T	T	A	G	G	G	C

All the nucleotides depicted are in the coding strand of the gene.
Gray cells indicate shared genotypes and white cells denote phase ambiguities.
Allele frequency data are from Caucasoid individuals of the Central European (CEU) panel of the HapMap database (www.hapmap.org) except for rs17220759, from the Perlegen AFD European panel (genome.perlegen.com/browser).
[a] E: Exon; I: Intron.
[b] Sp: Spain; A: Argentina.

c.1142+2T>A (ML9), c.1143−1G>C (ML3), c.427delG (ML4) and c.966C>A (p.W322X) (ML4) triggered the NMD process. In patient ML1 and in the father of patient ML9, the ratio between normal *ARSB* RNA and the RNA lacking exon 5 was 2.3:1 and 18.5:1, respectively, whereas in patient ML3 the ratio [normal RNA]:[RNA without exon 6] was 3.6:1. Patient ML4 displayed a single fainter band of normal size in the RT-PCR experiment, which was validated through comparison with the β-actin transcript: the ratio [*ARSB* RNA]:[β-actin RNA] was 1:5 in ML4 whereas in ML5, used as a control, it was 1:0.8.

Haplotype analysis

We performed haplotype analysis in order to test whether the mutations appearing more than once in our series of patients (c.1143−8T>G, c.427delG, c.1143−1G>C and p.C447F) had a single origin or were recurrent events. Seven polymorphisms were selected along the *ARSB* gene and genotyped in the patients (Table 3). The four chromosomes bearing mutation c.1143−8T>G (two from the same patient, ML2), the two c.427delG alleles, the seven c.1143−1G>C alleles (two from patient ML13) and the p.C447F alleles shared common haplotypes when all the variants analyzed were considered. When phases were unknown, the results were consistent with the presence of a shared haplotype.

Discussion

We identified all the mutant *ARSB* alleles in 16 Spanish and Argentinian Maroteaux–Lamy unrelated patients. Most of the patients (12/16) were compound heterozygotes, while only four of them were homozygotes, thereby confirming the broad allelic heterogeneity of the disease. Two of the mutations (c.1143−1G>C and c.1143−8T>G) were present in patients from both countries. Some mutations previously reported to be relatively frequent in other populations, such as p.Y120C and p.H393P, present in 4 chromosomes each [12], were not found in our sample of patients. A total of 9 alleles correspond to novel mutations, six of which produce amino acid changes, one involving a nonsense alteration, and two affecting splice sites.

The most prevalent mutations in this group of patients were the intronic splice site changes c.1143−1G>C and c.1143−8T>G, which altogether accounted for one-third of the total number of mutant alleles. Stop codon mutations are a frequent type of pathogenic change in MPS VI and have been reported to account for 16% of the total mutations described for the disease [28]. We identified two distinct nonsense changes out of 19 mutations (10.5%), which were present in 3 mutant alleles out of 32.

Of the 7 new mutations involving *ARSB* coding regions described herein (p.L82R, p.Y138C, p.S240F, p.G308R, p.W322X, p.C447F and p.L472P), p.W322X caused the

most evident structural modification on 4-sulfatase, as it resulted in a truncated protein that lacks the 212 carboxy-terminal residues. Mutations p.L82R, p.Y138C, p.S240F, p.G308R, p.C447F and p.L472P did not involve any amino acid of the conserved catalytic site of the sulfatase family, defined by the consensus sequences CTPSR (residues 91–95 in ARSB) and GKWHLG (residues 144–149). Moreover, neither are these mutations located in a conserved position within the human sulfatase family nor in the ARSB enzyme of different species. However, several mutations out of these conserved regions have been reported to be important for ER, Golgi and lysosome trafficking, glycosylation, proteolysis, enzyme stability and substrate binding [29]. Several observations indicate that the novel missense changes that we identified are disease-causing mutations. First, they were not found in a screening of 100 control chromosomes from healthy Spanish individuals, neither were they in any SNP database. This would indicate that they are not neutral polymorphisms. Second, these changes, together with a second mutated allele, were the only ones found in each patient after an extensive mutation analysis of all the coding and periexonic regions of the *ARSB* gene. Third, all the amino acid substitutions caused considerable charge, polarity or structural changes in the residues involved, although protein secondary structure prediction using the PSIPRED software did not reveal substantial alterations (bioinf.cs.ucl.ac.uk/psipred) [30]. Finally, the pathogenicity of several of the changes identified is supported by the findings of other studies. In this sense, mutation p.C447F was located close to a potential glycosylation site that may affect affinity for the mannose-6-phosphate receptor and the recaptation of the enzyme by distinct cell types [31]. In any case, biochemical characterization of the mutant proteins, currently under way in our lab, is needed to confirm the involvement of these amino acid substitutions in the disease phenotype.

Two other novel mutations identified, c.1142+2T>A (patient ML9) and c.1143−1G>C (ML3, ML6, ML12, ML13, ML15 and ML16) and a previously described change, c.1143−8T>G (ML2, ML7 and ML11), affected the donor and acceptor splice sites of intron 5. Mutations c.1143−1G>C and c.1143−8T>G altered the canonical acceptor site and the pyrimidine track, respectively, and mRNA analysis demonstrated exon 6 skipping, which resulted in frameshift and premature termination of the encoded protein. The c.1143−8T>G mutation was previously described [25] but its effect on splicing was not analyzed. The splicing score predicted for the c.1143−8T>G mutation (http://www.fruitfly.org/seq_tools/splice.html) [32] is reduced from 0.81 (T allele) to 0.57 (G allele), which was consistent with our experimental results.

Mutation c.1142+2T>A altered the 100% conserved GT dinucleotide in the donor splice site of intron 5, causing skipping of exon 5, protein frameshift and premature termination. Mutation c.899_1142del showed the same effect as that of c.1142+2T>A, although generated by a different mechanism: mutation c.899_1142del corresponded to a genomic deletion starting in intron 4 and ending in intron 5. In both cases, the mRNA lacked exon 5.

In addition, these three splice mutations and the genomic deletion may cause mRNA degradation by NMD [27]. This could also happen for nonsense (p.R160X and p.W322X) and frameshift (c.238delG, c.237-243delGGTGCTC and c.427delG) mutations. For some of the patients bearing these mutations, RT-PCR results indicated that NMD had taken place.

It was very difficult to establish genotype–phenotype correlations. On the one hand, the high allelic heterogeneity observed in our series of patients, which is the rule in this disease, prevented the identification of unrelated individuals with the same genotype. On the other hand, patients such as ML8, who was homozygous for a null allele (p.R160X), was diagnosed at 8 years of age, while patients bearing missense mutations in one allele, such as ML10 (p.C447F) or ML11 (p.Y138C), or in both, such as ML14 (p.R315Q;p.G308R), were diagnosed at an earlier age.

Here, we studied several non-pathogenic polymorphisms, which had been previously described or were present in SNP databases: c.313−26T>C (rs3733895), c.691−22T>C (rs6870443), c.1072G>A (rs1065757 or p.V358M) [13], c.1126G>A (rs17220759 or p.V376M) [3], c. 1151G>A (rs25414 or p.S384N), c.1191G>A (rs25413 or p.P397P) [33] and c.1337−32C>G (rs3822473). We did not find any of several polymorphic variants in our series of patients that had been previously described in other Caucasoid populations (p.I114I, p.L124L and p.G324G) [20].

The p.V376M substitution is considered a non-pathogenic variant, according to several expression studies [3,14,15], and has a population frequency of 0.167 as calculated from the Perlegen AFD European Panel (genome.perlegen.com/browser) [34]. The p.V358M polymorphism is a frequent amino acid substitution among Caucasoids (38% of 220 alleles) [13] and was present in 32% of the 100 alleles of healthy Spanish individuals screened in our study. This substitution was described recently and it was shown to cause a reduction in ARSB activity when accompanying a disease-causing mutation in the same chromosome [14,15]. A third putative non-synonymous polymorphic variant, p.S384N, was considered in a previous report [10, article in Russian], as quoted in a subsequent publication [28], a severe MPS VI mutation, as it was found at homozygosity in a patient with a rapidly progressing phenotype. However, we have found this change in one Spanish and in one Argentinian patient who also bore two other mutations: c.238delG and c.1142+2T>A (ML9) and p.R315Q and G308R (ML14). In patient ML9, p.S384N was in the same allele as c.1142+2T>A, a mutation that affects the intron 5 donor splice site causing skipping of exon 5 and premature termination of the protein. In patient ML14, the analysis of the mother indicated that the p.S384N change was most likely in phase with

p.R315Q, as in one previously reported patient [15], with genotype p.[L72R] + [R315Q;S384N]. The mutation p.R315Q has been identified by several authors at homozygosity in patients with intermediate to severe phenotypes [20,25] and had undetectable activity after *in vitro* expression studies [15]. Also, a patient was reported with the missense changes p.S384N, p.R152W and p.L360P, although the phase was not established in this case [10]. The change p.S384N is described as a SNP in the databases (rs25414), with a minor allele frequency (MAF) of 0.067 in the Central European (CEU) panel of the HapMap Project (http://www.hapmap.org) [35]. In the Spanish control population that we studied, the frequency of the least common allele, an asparagine residue, was 0.07 (over a total of 100 chromosomes). From these data we would expect a frequency of about 5:1000 for the p.[S384N] + [S384N] genotype in the general population, which would be in contradiction with the low incidence of the disorder, about 1:300,000 live births or less [36,37]. All these observations indicate that this residue might be a non-pathogenic polymorphism rather than a pathogenic mutation, although it may act as a modifier allele, as is the case of p.V358M, when it accompanies another mutation in the same chromosome.

Our haplotype analyses for all the mutations found more than once in Spanish and Argentinian MPS VI patients showed that mutations c.1143−8T>G, c.427delG, c.1143−1G>C and p.C447F were always in the context of the same haplotype, clearly supporting a common origin for the alleles bearing each of these four changes.

Acknowledgments

The authors thank the following physicians for sending patient samples: Dr. Gracia , Dr. Nieto, Dr. González Venero (Madrid); Dr. Arroyo (Cáceres); Dr. Prieto, Dr. García (València); Dr. Colomer, Dr. Pineda (Barcelona); Dr. Martí (Canarias); Dr. Caimari (Mallorca); Dr. Vela (Sevilla); Dr. del Toro (Barcelona). We are also grateful to R. Rycroft for revising the English. Financial support was provided by CICYT (SAF2003-00386, SAF2006-12276), FIS (Redes Temáticas, REDEMETH G03/054 and PI051343) and AGAUR (2005SGR00848).

References

[1] E.F. Neufeld, J. Muenzer, The mucopolysaccharidoses, in: C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle (Eds.), The Metabolic and Molecular Basis of Inherited Disease, eighth ed., McGraw-Hill, New York, 2001, pp. 2465–2494.

[2] C. Peters, B. Schmidt, W. Rommerskirch, K. Rupp, M. Zuhlsdorf, M. Vingron, H.E. Meyer, R. Pohlmann, K. von Figura, Phylogenetic conservation of arylsulfatases. cDNA cloning and expression of human arylsulfatase B, J. Biol. Chem. 265 (1990) 3374–3381.

[3] G. Wicker, V. Prill, D. Brooks, G. Gibson, J.J. Hopwood, K. von Figura, C. Peters, Mucopolysaccharidosis VI (Maroteaux–Lamy syndrome). An intermediate clinical phenotype caused by substitution of valine for glycine at position 137 of arylsulfatase, Br. J. Biol. Chem. 266 (1991) 21386–21391.

[4] T. Litjens, E.G. Baker, K.R. Beckmann, C.P. Morris, J.J. Hopwood, D.F. Callen, Chromosomal localization of ARSB, the gene for human *N*-acetylgalactosamine-4-sulphatase, Hum. Genet. 821 (1989) 67–68.

[5] S. Modaressi, K. Rupp, K. von Figura, C. Peters, Structure of the human arylsulfatase B gene, Biol. Chem. Hoppe Seyler 374 (1993) 327–335.

[6] C.S. Bond, P.R. Clements, S.J. Ashby, C.A. Collyer, S.J. Harrop, J.J. Hopwood, J.M. Guss, Structure of a human lysosomal sulfatase, Structure 5 (1997) 277–289.

[7] T. Kobayashi, K. Honke, T. Jin, S. Gasa, T. Miyazaki, A. Makita, Components and proteolytic processing sites of arylsulfatase B from human placenta, Biochim. Biophys. Acta 1159 (1992) 243–247.

[8] D. Isbrandt, G. Arlt, D.A. Brooks, J.J. Hopwood, K. von Figura, C. Peters, Mucopolysaccharidosis VI Maroteaux–Lamy syndrome: six unique arylsulfatase B gene alleles causing variable disease phenotypes, Am. J. Hum. Genet. 454 (1994) 454–463.

[9] E. Voskoboeva, D. Isbrandt, K. von Figura, X. Krasnopolskaya, C. Peters, Four novel mutant alleles of the arylsulfatase B gene in two patients with intermediate form of mucopolysaccharidosis VI (Maroteaux–Lamy syndrome), Hum. Genet. 93 (1994) 259–264.

[10] E. Voskoboeva, K.D. Krasnopol'skaia, K. Peters, K. von Figura, Identification of mutations in the arylsulfatase B gene in Russian mucopolysaccharidosis type VI patients, Genetika 36 (2000) 837–843.

[11] T. Litjens, C.P. Morris, E.F. Robertson, C. Peters, K. von Figura, J.J. Hopwood, An *N*-acetylgalactosamine-4-sulfatase mutation delta G238 results in a severe Maroteaux–Lamy phenotype, Hum. Mutat. 1 (1992) 397–402.

[12] T. Litjens, D.A. Brooks, C. Peters, G.J. Gibson, J.J. Hopwood, Identification, expression, and biochemical characterization of *N*-acetylgalactosamine-4-sulfatase mutations and relationship with clinical phenotype in MPS-VI patients, Am. J. Hum. Genet. 58 (1996) 1127–1134.

[13] W.D. Jin, C.E. Jackson, R.J. Desnick, E.H. Schuchman, Mucopolysaccharidosis type VI: identification of three mutations in the arylsulfatase B gene of patients with the severe and mild phenotypes provides molecular evidence for genetic heterogeneity, Am. J. Hum. Genet. 50 (1992) 795–800.

[14] L. Karageorgos, P. Harmatz, J. Simon, A. Pollard, P.R. Clements, D.A. Brooks, J.J. Hopwood, Mutational analysis of mucopolysaccharidosis type VI patients undergoing a trial of enzyme replacement therapy, Hum. Mutat. 23 (2004) 229–233.

[15] L. Karageorgos, D.A. Brooks, P. Harmatz, D. Ketteridge, A. Pollard, E.L. Melville, E. Parkinson-Lawrence, P.R. Clements, J.J. Hopwood, Mutational analysis of mucopolysaccharidosis type VI patients undergoing a phase II trail of enzyme replacement therapy, Mol. Genet. Metab. 90 (2007) 164–170.

[16] C.M. Simonaro, E.H. Schuchman, *N*-Acetylgalactosamine-4-sulfatase: identification of four new mutations within the conserved sulfatase region causing mucopolysaccharidosis type VI, Biochim. Biophys. Acta 1272 (1995) 129–132.

[17] G. Arlt, D.A. Brooks, D. Isbrandt, J.J. Hopwood, J. Bielicki, T.M. Bradford, C.A. Bindloss-Petherbridge, K. von Figura, C. Peters, Juvenile form of mucopolysaccharidosis VI Maroteaux–Lamy syndrome. A C-terminal extension causes instability but increases catalytic efficiency of arylsulfatase B, J. Biol. Chem. 269 (1994) 9638–9643.

[18] D. Isbrandt, J.J. Hopwood, K. von Figura, C. Peters, Two novel frameshift mutations causing premature stop codons in a patient with the severe form of Maroteaux–Lamy syndrome, Hum. Mutat. 7 (1996) 361–363.

[19] G.R. Villani, N. Balzano, P. Di Natale, Two novel mutations of the arylsulfatase B gene in two Italian patients with severe form of mucopolysaccharidosis, Hum. Mutat. 11 (1998) 410, Mutation in brief 127.

[20] G.R. Villani, N. Balzano, D. Vitale, M. Saviano, V. Pavone, P. Di Natale, Maroteaux–Lamy syndrome: five novel mutations and their structural localization, Biochim. Biophys. Acta 1453 (1999) 185–192.

[21] J.Y. Wu, C.F. Yang, C.C. Lee, J.G. Chang, F.J. Tsai, A novel mutation (Q239R) identified in a Taiwan Chinese patient with type VI

mucopolysaccharidosis (Maroteaux–Lamy syndrome), Hum. Mutat. 15 (2000) 389–390.
[22] C.F. Yang, J.Y. Wu, S.P. Lin, F.J. Tsai, Mucopolysaccharidosis type VI: report of two Taiwanese patients and identification of one novel mutation, J. Formos. Med. Assoc. 100 (2001) 820–823.
[23] W. Dou, X. Gu, C. Peng, J. Zheng, J. Chen, W. Zhang, S. Huang, H.Z. Sheng, Two novel mutations of the arylsulfatase B gene in a Chinese MPS VI child undergoing bone marrow transplantation therapy, Clin. Chim. Acta 374 (2006) 171–172.
[24] M.F. Petry, T. Dieter, M. Burin, R. Giugliani, S. Leistner, Identification of a novel mutation in the ARSB gene that is frequent among Brazilian MPS VI patients, Genet. Test. 7 (2003) 347–349.
[25] M.F. Petry, K. Nonemacher, J.C. Sebben, I.V. Schwartz, A.C. Azevedo, M.G. Burin, A.R. De Rezende, C.A. Kim, R. Giugliani, S. Leistner-Segal, Mucopolysaccharidosis type VI: identification of novel mutations on the arylsulphatase B gene in South American patients, J. Inherit. Metab. Dis. 286 (2005) 1027–1034.
[26] D.A. Brooks, G.J. Gibson, L. Karageorgos, L.K. Hein, E.F. Robertson, J.J. Hopwood, An index case for the attenuated end of the mucopolysaccharidosis type VI clinical spectrum, Mol. Genet. Metab. 853 (2005) 236–238.
[27] L.E. Maquat, Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics, Nat. Rev. Mol. Cell. Biol. 5 (2004) 89–99.
[28] T. Litjens, J.J. Hopwood, Mucopolysaccharidosis type VI: structural and clinical implications of mutations in *N*-acetylgalactosamine-4-sulfatase, Hum. Mutat. 18 (2001) 282–295.
[29] S. Esposito, N. Balzano, A. Daniele, G.R. Villani, K. Perkins, B. Weber, J.J. Hopwood, P. Di Natale, Heparan *N*-sulfatase gene: two novel mutations and transient expression of 15 defects, Biochim. Biophys. Acta 1501 (2000) 1–11.
[30] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (2000) 404–405.
[31] M. Fuller, J.J. Hopwood, D.S. Anson, Receptor mediated binding of two glycosylation forms of *N*-acetylgalactosamine-4-sulphatase, Biochim. Biophys. Acta 1406 (1998) 283–290.
[32] M.G. Reese, F.H. Eeckman, D. Kulp, D. Haussler, Improved splice site detection in genie, J. Comp. Biol. 4 (1997) 311–323.
[33] E.H. Schuchman, C.E. Jackson, R.J. Desnick, Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA, and regions of amino acid identity with arylsulfatases A and C, Genomics 6 (1990) 149–158.
[34] D.A. Hinds, L.L. Stuve, G.B. Nilsen, E. Halperin, E. Eskin, D.G. Ballinger, K.A. Frazer, D.R. Cox, Whole-genome patterns of common DNA variation in three human populations, Science 307 (2005) 1072–1079.
[35] G.A. Thorisson, A.V. Smith, L. Krishnan, L.D. Stein, The international HapMap project web site, Genome Res. 15 (2005) 1592–1593.
[36] P.J. Meikle, J.J. Hopwood, A.E. Clague, W.F. Carey, Prevalence of lysosomal storage disorders, JAMA 281 (1999) 249–254.
[37] J. Nelson, J. Crowhurst, B. Carey, L. Greed, Incidence of the mucopolysaccharidoses in Western Australia, Am. J. Med. Genet. A 123 (2003) 310–313.