Molecular Therapy
Methods & Clinical Development
Original Article

A Generic Assay to Detect Aberrant *ARSB* Splicing and mRNA Degradation for the Molecular Diagnosis of MPS VI

Mike Broeders,[1,2,3] Kasper Smits,[1,2,3] Busra Goynuk,[1,2,3] Esmee Oussoren,[4,5] Hannerieke J.M.P. van den Hout,[4,5] Atze J. Bergsma,[1,2,3] Ans T. van der Ploeg,[4,5] and W.W.M. Pim Pijnappel[1,2,3]

[1]Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; [2]Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; [3]Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, the Netherlands; [4]Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; [5]Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, the Netherlands

Identification and characterization of disease-associated variants in monogenic disorders is an important aspect of diagnosis, genetic counseling, prediction of disease severity, and development of therapy. However, the effects of disease-associated variants on pre-mRNA splicing and mRNA degradation are difficult to predict and often missed. Here we present a generic assay for unbiased identification and quantification of arylsulfatase B (*ARSB*) mRNA for molecular diagnosis of patients with mucopolysaccharidosis VI (MPS VI). We found that healthy control individuals have inefficient *ARSB* splicing because of natural skipping of exon 5 and inclusion of two pseudoexons in introns 5 and 6. Analyses of 12 MPS VI patients with 10 different genotypes resulted in identification of a 151-bp intron inclusion caused by the c.1142+2T>C variant and detection of low *ARSB* expression from alleles with the c.629A>G variant. A special case showed skipping of exon 4 and low *ARSB* expression. Although no disease-associated DNA variant could be identified in this patient, the molecular diagnosis could be made based on RNA. These results highlight the relevance of RNA-based analyses to establish a molecular diagnosis of MPS VI. We speculate that inefficient natural splicing of *ARSB* may be a target for therapy based on promotion of canonical splicing.

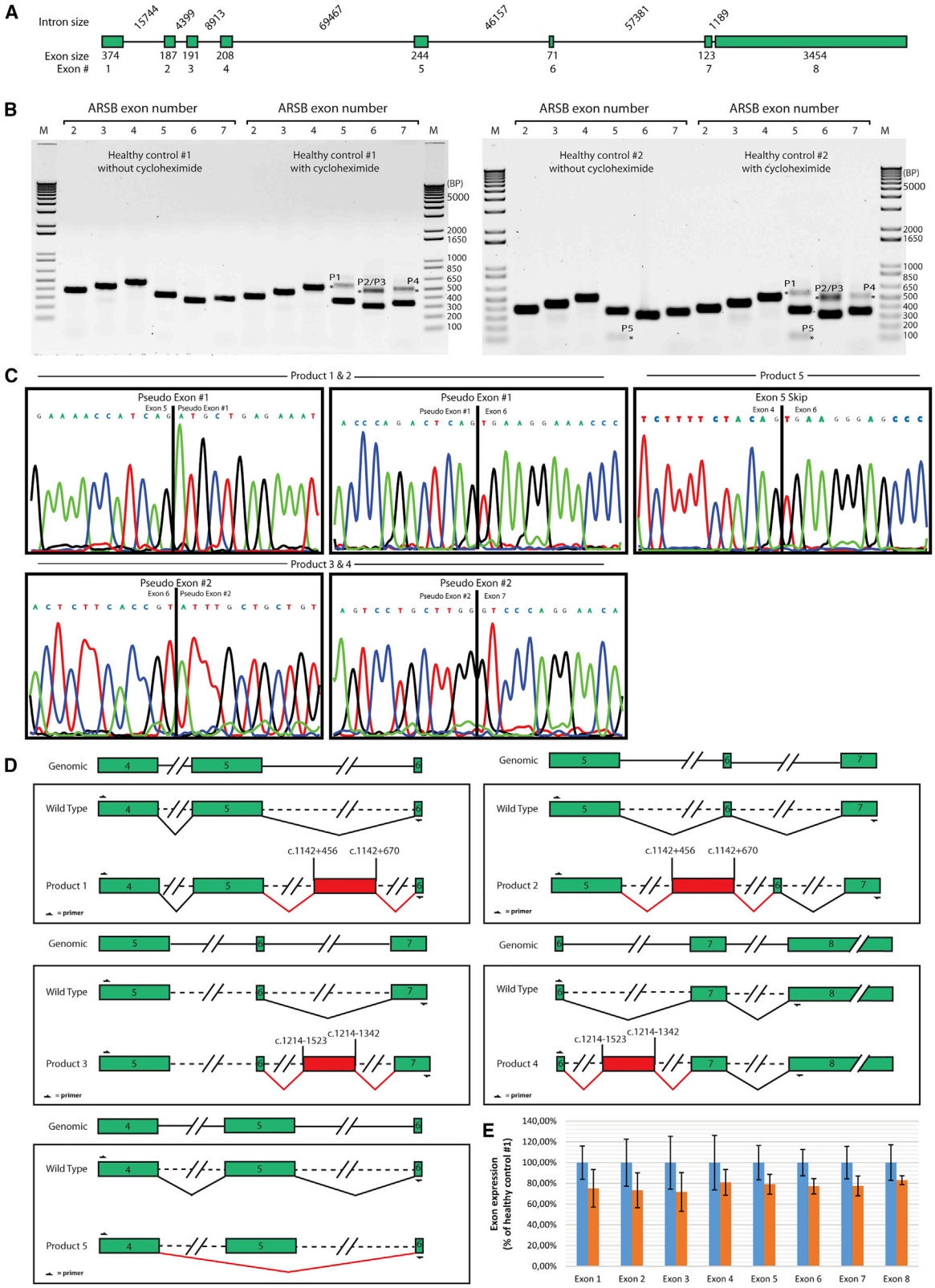
INTRODUCTION

Mucopolysaccharidosis VI (MPS VI; MIM: 253200) is a lysosomal storage disease caused by disease-associated variants in the *arylsulfatase B* (*ARSB*) gene and has an autosomal recessive inheritance. ARSB enzyme deficiency results in failure to degrade glycosaminoglycans (GAGs) within lysosomes, resulting in their accumulation and lysosomal pathology in multiple tissues. The most prominently affected tissues include cartilage, bone, cornea, heart valves, and visceral organs.[1–6] Patients with MPS VI can be diagnosed with a slowly or rapidly progressive form of the disease, although a broad clinical spectrum exists.[5–8] Enzyme replacement therapy (ERT) for MPS VI, where the human recombinant ARSB enzyme is administered weekly via intravenous infusion, has been available since 2005. ERT improves body growth and symptoms in visceral organs and extends life expectancy but has limited efficacy in connective tissues.[8–12] The molecular diagnosis of MPS VI is based on ARSB enzyme deficiency and the presence of disease-associated variants in the *ARSB* gene.[6] Current diagnostic assays for identification of DNA variants are based on sequence analysis of exons and short stretches of adjacent introns. However, disease-associated variants outside of these regions are missed in such analyses. In addition, the effects of variants on pre-mRNA splicing and/or mRNA expression remain unknown.

To date, over 200 unique variants in the *ARSB* gene have been reported,[13] of which 37% are classified as likely pathogenic and 16% as pathogenic, and insufficient data are available for classification of ~45% of variants. Currently ~5% of the reported *ARSB* variants have been reported to affect splicing, but it is likely that this is an underestimation based on the reported frequency of splicing variants in human disease.[14,15] This amounts to up to ~20% of disease-associated variants in general, implying that up to an additional ~30 of the 200 known *ARSB* variants for which insufficient information is available or that are currently classified as missense might have an undocumented effect on splicing, which should be investigated further.[14,16,17] The effect of variants on splicing are difficult to predict *in silico*.[18,19] Even when a strong effect on splicing is predicted, the outcome of aberrant splicing is unclear.[20] For instance, loss of a splice site may result in exon skipping or inclusion, intron retention, and/or utilization of a cryptic splice site. Depending on the number of nucleotides in the skipped or included sequence, the reading frame may remain intact or become disrupted. The latter can result in mRNA degradation via the nonsense-mediated decay (NMD) pathway and can remain unnoticed unless this pathway is inhibited experimentally;

Received 7 July 2020; accepted 11 September 2020;
https://doi.org/10.1016/j.omtm.2020.09.004.

Correspondence: W.W.M. Pim Pijnappel, Erasmus MC University Medical Center, Wytemaweg 80, 3015 CN Rotterdam, the Netherlands.
E-mail: w.pijnappel@erasmusmc.nl



(legend on next page)

for example, using cycloheximide (CHX). Identification of disease-associated variants and understanding their mechanism of action is becoming increasingly important to confirm the diagnosis, for genetic counseling, and for development of novel therapies.[20,21] Newborn screening has been shown to be feasible for MPS VI.[22–24] It will be of great importance for newborn screening programs to have a full understanding of *ARSB* disease-associated variants, including their putative effects on RNA processing and stability.

Previously, we developed a generic approach for identification and characterization of variants that affect *GAA* splicing in Pompe disease and applied this assay to identify multiple novel splicing events caused by disease-associated variants that were known or that were missed by standard diagnostics.[25–27] Here we tailored this assay for detection of *ARSB* splicing in MPS VI and extended it to enable detection of *ARSB* transcripts that undergo mRNA decay. The *ARSB* splicing assay yielded novel information at the RNA level for 4 patients with three different genotypes of 12 MPS VI patients in the Netherlands, including an RNA-based molecular diagnosis of MPS VI in a patient lacking *ARSB* disease-associated DNA variants. In addition, it revealed inefficient canonical splicing of exons 5 and 6 in all analyzed patients and control individuals, providing a putative target for a generic strategy for MPS VI based on improving canonical splicing.

RESULTS

Healthy Control Individuals

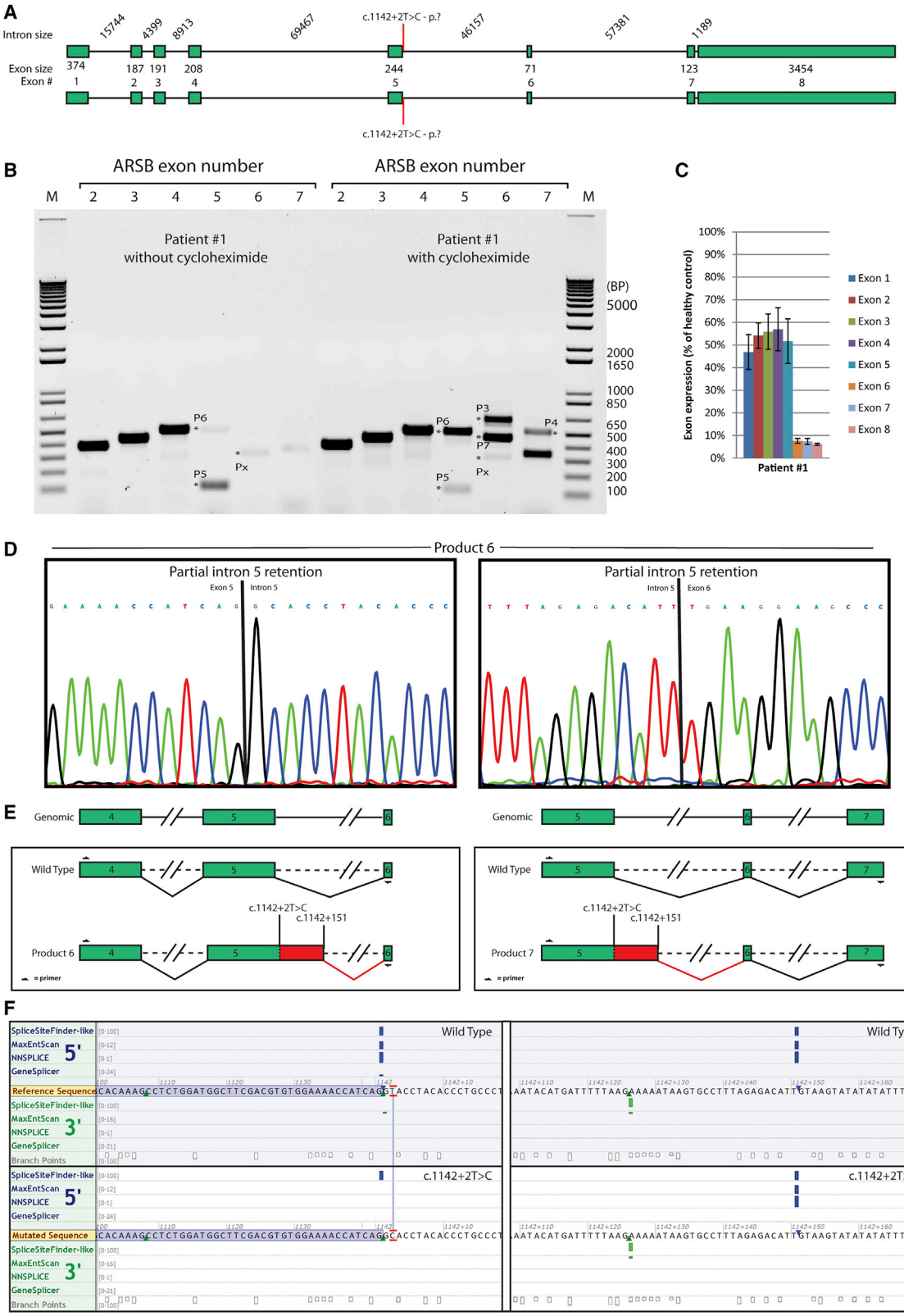
The generic splicing assay consists of flanking-exon RT-PCR (using primers in exons flanking the exon of interest) and exon-internal RT-qPCR (using primers within the exon of interest) of all exons, followed by Sanger sequence analysis of splicing products, as decribed.[25] We tailored this assay to *ARSB* and included treatment with CHX to inhibit NMD to detect aberrant splicing products that changed the reading frame and caused a premature termination codon. Primary fibroblasts of two healthy control individuals were used to validate the approach. Flanking-exon RT-PCR analysis of all but the first and last exons showed that all canonical splicing products were detected at the expected sizes in healthy control individuals 1 and 2 (Figures 1A and 1B). Inhibition of NMD by treatment with CHX resulted in identification of additional aberrant products in healthy control individual 1 (products 1–4). Healthy control individual 2 showed identical products, and, in addition, showed a lowly expressed product (product 5) that was insensitive to NMD inhibition. Repeated analyses showed that product 5 can be easily missed in RT-PCR analysis, likely because of its low abundance, and that it could also be detected in control individual 1 in some cases (data not shown). Sequence analysis showed that products 1 and 2 contained an inclusion of a pseudoexon of 214 nt from c.1142+456 to c.1142+670 in intron 5. Analysis of products 3 and 4 showed inclusion of a pseudoexon of 181 nt from c.1214–1523 to c.1214–1342 in intron 6. Product 5 was the result of a perfect skip of exon 5, resulting in loss of 244 nt (Figures 1C and 1D). All of these products resulted in an out-of-frame product that was predicted to undergo NMD. In agreement, products 1–4 were only detected after inhibition of NMD; however, product 5 was detected irrespective of CHX treatment, suggesting escape from NMD. RT-qPCR analysis showed quantification of all *ARSB* exons in healthy control individuals 1 and 2. The expression levels of all exons were similar in both healthy control individuals, with slightly lower levels in control individual 2 compared control individual 1 (Figure 1E). In summary, these results established the generic splicing assay for *ARSB* and showed inefficient *ARSB* splicing in healthy control individuals.

Patient 1

Patient #1 is homozygous for c.1142+2T>C, which is located near the splice donor of exon 5 (Figure 2A; Table 1). This variant has been published before by our center.[9] No splicing analysis has been reported by us or others, but Garrido et al.[28,29] reported that the similar variant c.1142+2T>A induces skipping of exon 5. *In silico* splicing prediction of both variants indicated loss of the canonical splice donor of exon 5 (Figure 2F, left panel; data not shown). Flanking-exon RT-PCR analysis of all exons showed the presence of canonical splicing products for exons 2–4 but very low expression of canonical splicing products for exons 5–7 (Figure 2B). After inhibition of NMD, several splicing products for exons 5–7 were detected, suggesting that the 3′ part of the *ARSB* mRNA was mostly degraded under normal growth conditions (see below for identification of these products). Quantification of *ARSB* mRNA expression using exon-internal RT-qPCR analysis of cells grown in the absence of CHX showed that expression of exons 6–8 was below 10% of the levels in healthy control individuals. Exons 1–5 showed expression of ~50% of that of healthy control individuals, suggesting that this part of the mRNA partially escaped mRNA degradation. Flanking-exon RT-PCR analysis further showed two aberrant splice products for exon 5: one with a low molecular weight (MW) (product 5) and one with a higher MW than expected (product 6). Expression of exon 5 products in the absence of CHX was low in flanking-exon RT-PCR because of low expression of exon 6, in which the reverse primer was located. Sequence analysis indicated that product 5 was the result of complete skipping of exon 5. However, this exon 5 skip cannot be attributed to the c.1142+2T>C variant because it also occurs in healthy control individuals (Figure 1), and there was no indication that the level of exon 5 skipping was

Figure 1. Splicing Analysis of Two Healthy Control Individuals

(A) Schematic overview of the *ARSB* gene with intron size, exon size, and exon number. (B) Flanking-exon RT-PCR analysis of two healthy control individuals. Exon numbers are indicated above the lanes. PCR products were separated by electrophoresis on a 1% agarose gel. Asterisks indicate alternative splicing products detected in healthy individuals. Numbers besides the bands refer to the products analyzed in further detail. (C) Sequence analysis of the aberrant splicing products shown in (B). (D) Cartoons of splicing events detected in healthy control individuals. Exons are indicated as green boxes. Introns are depicted as lines. A broken line indicates that the intron is longer than suggested in this drawing. Canonical splicing is indicated in black and alternative splicing in red. (E) Exon-internal RT-qPCR analysis of healthy control individuals 1 and 2. Values are means relative to healthy control individual 1. *GAPDH* was used for normalization. Error bars indicate SD (n = 3).



(legend on next page)

increased compared with the healthy control individuals. Product 6 was identified after NMD inhibition. Sequence analysis of this product showed partial retention of intron 5 (Figures 2D and 2E, left panels). This was confirmed using flanking-exon RT-PCR analysis of exon 6, resulting in product 7, which was also larger than expected and showed the same partial inclusion of intron 5 (Figures 2D and 2E, right panels). Inclusion of the first 151 nt of intron 5 causes a shift in the reading frame, likely causing mRNA degradation starting at exon 6. Interestingly, partial retention of intron 5 seemed to prevent utilization of the cryptic splice acceptor seen in healthy control individuals in intron 5 at c.1142+456 (as judged from the absence of a band corresponding to product 1 (P1) and P2, as seen in Figure 1). Px likely represented the canonical splicing product for exon 6, based on size. Topo cloning followed by sequencing did not result in its identification, likely because of its very low abundance. Similar to healthy control individuals, inclusion of the pseudoexon in intron 6 was detected in this patient, as seen from the higher-MW P3 and P4 for flanking-exon RT-PCR analysis of exons 6 and 7, respectively (Figure 2B). Splicing prediction indicated the presence of a strong cryptic splice donor at position c.1142+151, which is the new donor that was indeed used in P6 and P7 and resulted in intron retention (Figure 2F, right panel). Taken together, these results indicate that c.1142+2T>C has the following effects: (1) it causes partial retention of 151 nt of intron 5, which leads to a frameshift and degradation of the downstream part of the mRNA; (2) it prevents utilization of the pseudoexon in intron 5; and (3) it has negligible effects on exon 5 skipping.

Patients 2–9

We applied the splicing assay to 8 additional patients with missense variants to assess whether these variants could also affect splicing. Patients 2 and 3 were homozygous for c.454C>T, located in exon 2. Patient 4 was heterozygous for c.629A>G and c.937C>G, which are located in exons 3 and 5, respectively. Patient 5–8 were homozygous for a disease-associated variant located in exon 5; patient 5 carried c.971G>T, patients 6 and 8 carried c.937C>T, and patient 7 carried c.995T>G. Patient 9 was homozygous for c.903C>G in exon 5 and c.1151G>A in exon 6. All of these missense variants have a predicted or demonstrated deleterious effect on ARSB enzyme activity.[9] For all of these variants, *in silico* prediction indicated no effect on splicing (data not shown). Flanking-exon RT-PCR analysis of cells grown in the absence or presence of CHX indicated normal *ARSB* splicing in all cases (Figure S1). All analyzed patients showed inclusion of pseudoexons in introns 5 and 6, similar to healthy control individuals. We conclude that the missense variants that were present in patients 2–9 did not induce detectable aberrant *ARSB* splicing.

Patients 10 and 11

Patients 10 and 11 are siblings and compound heterozygous for the c.629A>G missense variant located in exon 3 and the c.979C>T nonsense variant located in exon 5 (Figure 3A). The c.979C>T variant results in a premature translation termination codon with predicted mRNA degradation via the NMD pathway. *In silico* splicing prediction of both variants indicated no effects on splicing (data not shown). Both variants have been classified previously as disease associated.[9,30] Flanking-exon RT-PCR analysis showed all canonical splicing products in both patients, although the levels were severely reduced compared with healthy controls. The expression levels of all exons increased upon CHX treatment, indicating degradation by NMD. The same non-canonical products (i.e., products with the pseudoexons in introns 5 and 6) as in healthy controls were detected after NMD inhibition (Figures 3B and 3C). The exon-internal RT-qPCR assay confirmed the low levels of ARSB mRNA expression, which were approximately 20% of the values of healthy control individuals for all exons in both patients. These results indicate that these patients have low *ARSB* mRNA expression. Although a 50% reduction in mRNA expression can be explained by NMD of mRNA that is expressed from the c.979C>T allele, the cause of the additional 30% reduction in mRNA expression from the c.629A>G allele is unclear. Further research is needed to investigate whether this additional reduction in mRNA expression is caused by c.629A>G itself or by another, still unidentified *ARSB* sequence variant.

Patient 12

Patient 12 was diagnosed with MPS VI based on deficient ARSB enzymatic activity and clinical symptoms; however, no disease-associated variants in the *ARSB* gene were found at the DNA level by standard diagnostic analysis (Figure 4A). Flanking-exon RT-PCR showed that all canonical splicing products were present at very low levels (Figure 4B). This was likely due to degradation by NMD because expression of all exons was elevated following inhibition of NMD using CHX. A very low level of expression for the canonical exon 4 product was detected (top band in RT-PCR of exon 4; Figure 4). Exon-internal RT-qPCR analysis confirmed the low *ARSB* expression, with levels of all exons that were below 10% compared with the levels of healthy control individual 1 (Figure 4C). Besides all canonical splicing products, an aberrant product (P8) with a low MW was detected using flanking-exon RT-PCR analysis of exon 4. This product was the result of a perfect skip of exon 4, as shown by sequence analysis (Figure 4D). The size of exon 4 is 208 nt, and skipping of this exon results in a frameshift, likely causing NMD. This was indeed observed because the exon 4-skipped mRNA was more abundant following inhibition of NMD (Figure 4E). In addition, we observed that the pseudoexon in intron 5 (P1 and P2 in Figure 1)

Figure 2. Splicing Analysis of Patient 1

(A) A schematic overview of the *ARSB* gene, with the disease-associated variant indicated by a red solid line. (B) Flanking-exon RT-PCR analysis of patient 1. (C) Exon-internal RT-qPCR analysis of patient 1. Values are means relative to healthy control individual 1. *GAPDH* was used for normalization. Error bars indicate SD (n = 3). (D) Sequence analysis of the aberrant splicing products shown in (B). (E) Cartoons of aberrant splicing events detected in patient 1. (F) Alamut Visual splicing prediction for five algorithms of variant c.1142+2T>C compared with the wild type. The left panel shows the canonical splice donor, and the right panel shows the splice donor used in P6 and P7. Larger bars indicate higher scores in the algorithms.

Table 1. Characteristics of Included Patients

Patient	ΔDNA (cDNA HGVS Nomenclature)	Location	ΔProtein (HGVS Nomenclature)	ΔDNA (cDNA HGVS Nomenclature)	Location	ΔProtein (HGVS Nomenclature)	Type of Disease Progression	ARSB Activity in Fibroblast (nmol/h*mg)	Age at Diagnosis (Years)	Ethnicity
1	c.1142+2T>C	intron 5	p.?	c.1142+2T>C	intron 5	p.?	rapidly	84.8	2.9	Pakistani
2	c.454C>T	exon 2	p.(R152W)	c.454 C>T	exon 2	p.(R152W)	slowly	61.9	0.7	Turkish
3	c.454C>T	exon 2	p.(R152W)	c.454C>T	exon 2	p.(R152W)	slowly	79.9	7.5	Turkish
4	c.629A>G	exon 3	p.(Y210C)	c.937C>G	exon 5	p.(P313A)	slowly	50.2	10.3	Dutch
5	c.971G>T	exon 5	p.(G324V)	c.971G>T	exon 5	p.(G324V)	rapidly	57.6	1.4	Guinea
6	c.937C>T	exon 5	p.(P313S)	c.937C>T	exon 5	p.(P313S)	rapidly	41.8	6	Caribbean
7	c.995T>G	exon 5	p.(V332G)	c.995T>G	exon 5	p.(V332G)	rapidly	69.4	2.7	Moroccan
8	c.937C>G	exon 5	p.(P313A)	c.937C>G	exon 5	p.(P313A)	rapidly	?	4.6	Dutch
9	c.[903C>G;1151G>A]	exons 5 & 6	p.[(N301K);(S384N)]	c.[903C>G;1151G>A]	exons 5 & 6	p.[(N301K);(S384N)]	rapidly	85.7	1.9	Turkish
10	c.629A>G	exon 3	p.(Y210C)	c.979C>T	exon 5	p.(R327X)	slowly	40.7	6.4	Dutch
11	c.629A>G	exon 3	p.(Y210C)	c.979C>T	exon 5	p.(R327X)	slowly	82.3	5.9	Dutch
12		unknown			unknown		rapidly	36.3	3.1	Dutch/Zimbabwe

ARSB activity range in fibroblasts: healthy individuals, 320–1,080 nmol/h*mg; patients, 19–105 nmol/h*mg.

seen in all individuals was lacking, likely because of the exon 4 skip, but that the pseudoexon in intron 6 (P3 and P4 in Figure 1) remained present.

To identify the genomic DNA variant that might be involved, we performed sequence analysis of all exons and their boundaries (up to 500 nt into the introns), but this failed to result in identification of any disease-associated variant. Because *ARSB* has very large intron sizes (8.9 kb for intron 3 and 69.4 kb for intron 4), it could be a candidate for recursive splicing (RS). The predicted consensus sequence AGGTRAGW used by Sibley et al.[31] was present eight times at several locations in intron 4, but targeted sequencing of these regions did not reveal the presence of any *ARSB* DNA sequence variant, and RNA sequencing (RNA-seq) analysis did not reveal RS in *ARSB* (data not shown). Heterozygous SNPs were detected at the DNA level in intron 2 and exon 5 and at the RNA level at a 50/50 ratio in exon 5, arguing against a large gene deletion or a promoter variant.

To examine the genomic organization, a SNP array of the patient as well as of the parents was performed. This showed two maternal gains of 484 kb and 225 kb on the X chromosome and two paternal gains on chromosome 5 and 17 of 38 kb and 300 kb, respectively (Figure 4F). The X chromosome gains and the gain on chromosome 17 did not contain any genes known to be involved in splicing that could explain the results of the splicing assay. The gain on chromosome 5 was located in 5q14.1, which is also the location of the *ARSB* gene, and consisted of exons 2–4 of the *ARSB* gene. Although the exact location of this gain could not be determined, it is possible that it would disrupt one allele of patient 12 and cause skipping of exon 4. Nevertheless, it is unclear how this putative mono-allelic disruption could explain the results of the splicing assay because the results of the flanking-exon RT-PCR and the exon-internal RT-qPCR suggested degradation of both alleles. To examine this further, we analyzed the parents. Both parents showed expression levels of *ARSB* that were ~50% of levels in healthy control individuals. The father contributed the allele with the gain in 5q14.1, which might have disrupted ARSB expression to a large extent and might have caused skipping of exon 4. It is unclear why exon 4 skipping was minimally present in the father and so abundant in the patient. The mother likely contributed an allele with low *ARSB* expression, but it remained unclear what the underlying DNA variant was. Irrespective of the lack of an *ARSB* genotype for this patient, the results emphasize that analysis at the RNA level using the generic splicing assay enabled molecular diagnosis of MPS VI, consisting of low *ARSB* mRNA expression and skipping of exon 4.

DISCUSSION

Aberrant ARSB Splicing in Healthy Control Individuals

ARSB is a very large gene (total size, 208 kb) with exceptionally large introns; intron 1 is large at 15.7 kb, but introns 4, 5, and 6 are huge at 69.4, 46.1, and 57.3 kb, respectively. In comparison, the average size of a human protein-coding gene is 67 kb, and the average intron size is

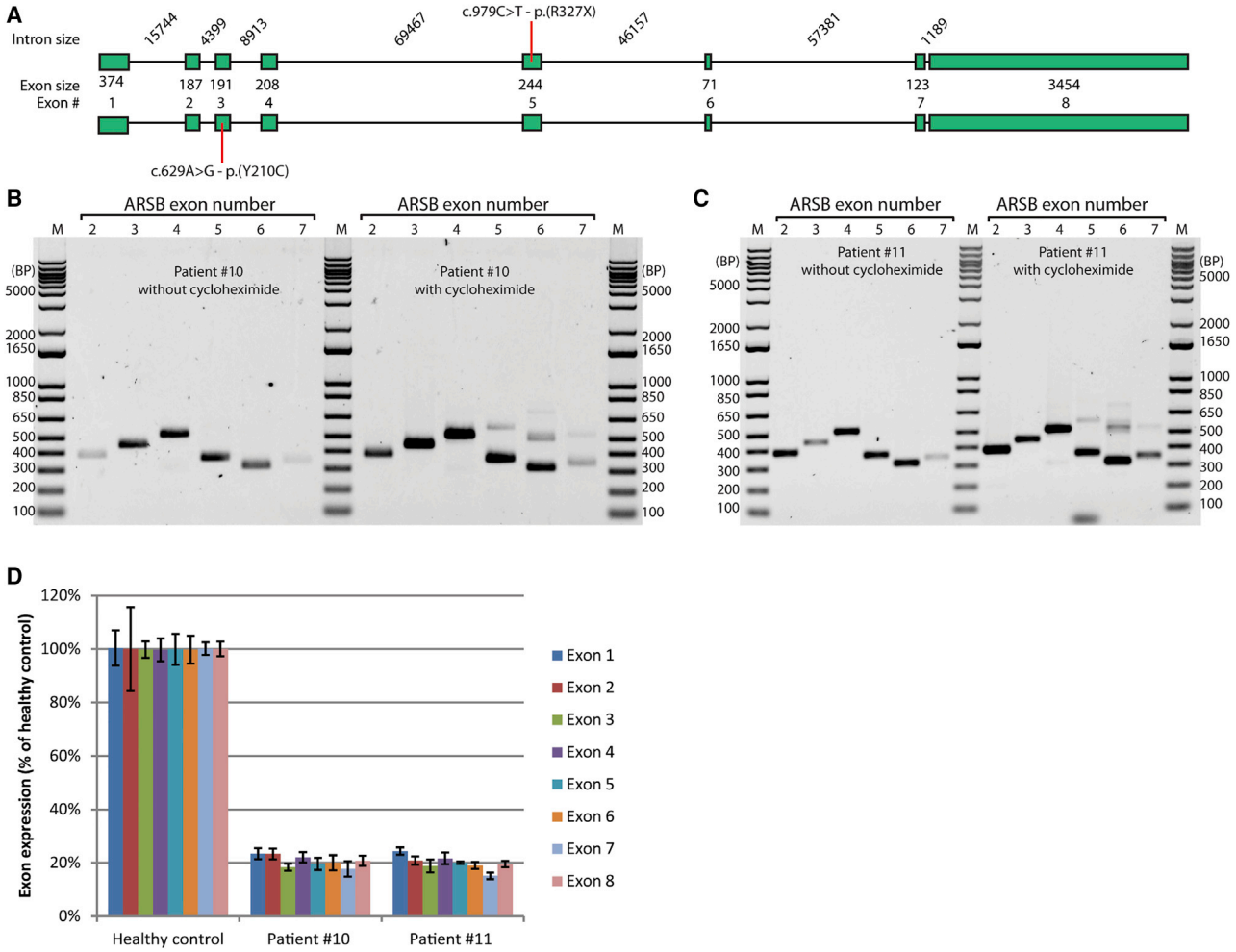


Figure 3. Splicing Analyses of Patients 10 and 11

(A) Schematic overview of the *ARSB* gene, with the disease-associated variant indicated by a red solid line. (B) Flanking-exon RT-PCR analysis of patient 10. (C) Flanking-exon RT-PCR analysis of patient 11. (D) Exon-internal RT-qPCR analysis of patients 10 and 11. Values are means relative to healthy control individual 1. *GAPDH* was used for normalization. Error bars indicate SD (n = 3).

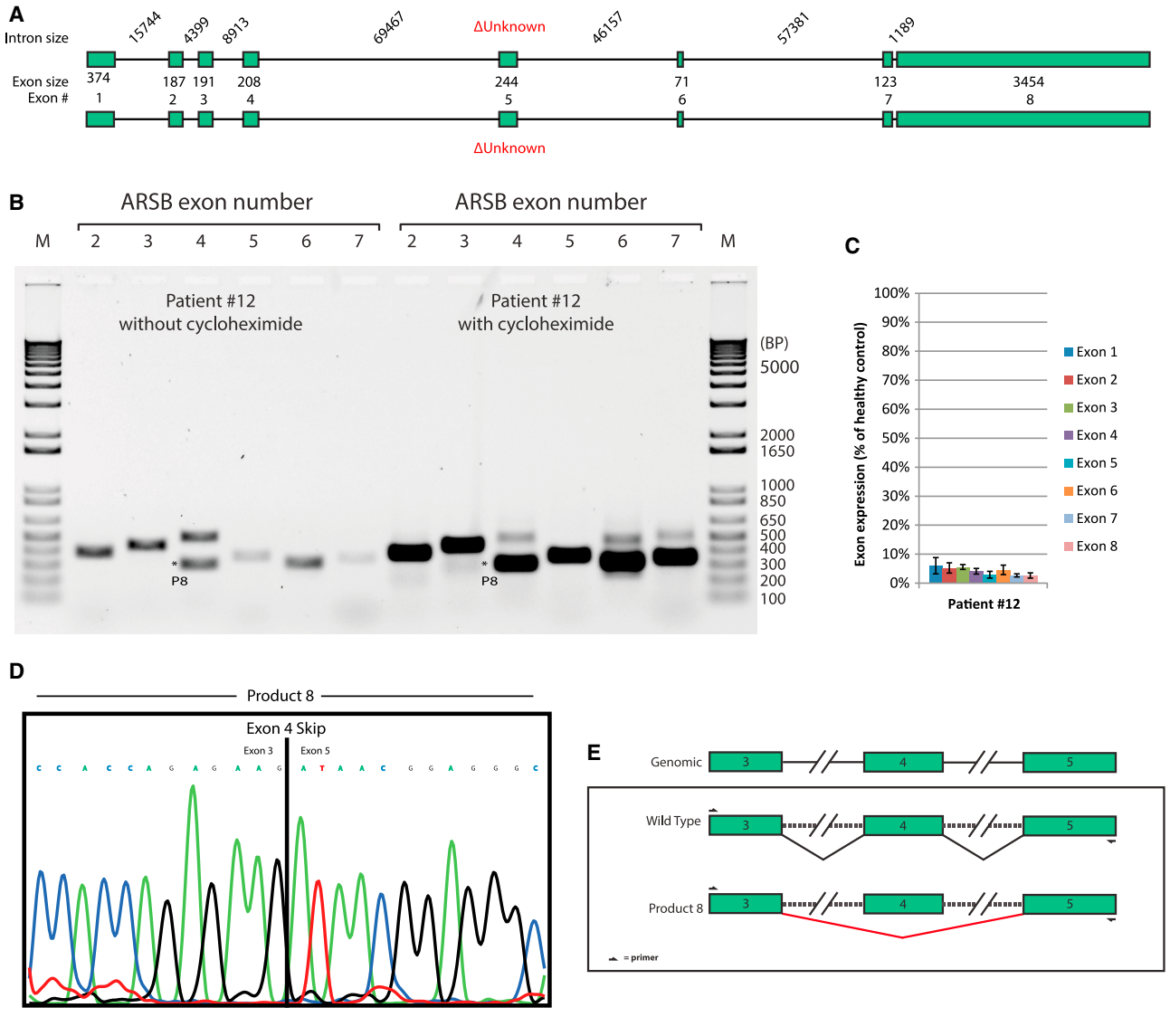
~6 kb.[32] This poses several challenges for production of ARSB protein. First, the pre-mRNA transcript is long, so takes more time than average for RNA polymerase II to generate a full-length transcript. Second, correct pre-mRNA splicing is a challenge because long introns contain, on average, more competitive cryptic splice sites, and long introns impose physical challenges for correct lariat formation during splicing.

Indeed, we found that splicing of *ARSB* is rather inefficient; exon 5 was skipped, and pseudoexons in introns 5 and 6 were utilized in a subset of transcripts. All of these aberrant splicing events resulted in an out-of-frame transcript. Skipping of exon 5 was present at low levels, produced an out-of-frame mRNA, but escaped NMD, as is sometimes the case.[33] Detection of exon 5 skipping in healthy control individuals should prevent misinterpretation of the effects of disease-associated *ARSB* variants in MPS VI.

Inclusion of pseudoexons in introns 5 and 6 did result in NMD and could only be detected after CHX treatment. Their inclusion was not obvious from *in silico* predictions (Figures S2 and S3). For the pseudoexon in intron 5, the splice acceptor of exon 6 had a relatively weak predicted strength, but the predicted strength of the cryptic splice acceptor at c.1142+456 was much weaker. Interestingly, the splice donor of this pseudoexon at c.1142+670 could not be predicted by any of the programs used. Cryptic splices sites of the pseudoexon in intron 6 were predicted *in silico* and were moderately strong. This highlights the need for experimental testing of splicing outcomes.

Relevance of Aberrant Natural Splicing

Although, at first sight, pseudoexon inclusion or exon skipping in healthy control individuals seem to be undesired forms of aberrant splicing, these are actually known to be among the mechanisms to regulate gene expression in various organisms ranging from plants



F

Chr	Band	Position	Length (Kb)	# probes		Patient #12	Mother	Father
5	q14.1	chr5:78234241-78272540	38	19	GAIN	Yes	No	Yes
17	p13.3	chr17:2819869-3119767	300	117	GAIN	Yes	No	Yes
X	p22.33	chrx:970843-1455218	484	74	GAIN	Yes	Yes	No
X	p22.33	chrx:1477190-1702312	225	76	GAIN	Yes	Yes	No

Figure 4. Splicing Analysis of Patient 12

(A) Schematic overview of the *ARSB* gene; no disease-associated variants were identified for this patient. (B) Flanking-exon RT-PCR analysis of patient 12. (C) Exon-internal RT-qPCR analysis of patient 12. Values are means relative to healthy control 1. *GAPDH* was used for normalization. Error bars indicate SD (n = 3). (D) Sequence analysis of the aberrant splicing products shown in (B). (E) Cartoon of the aberrant splicing event detected in patient 12. (F) SNP array results for patient 12 and parents.

to humans. For example, it has been shown that many genes involved in cell cycle regulation, splicing, development, tissue homeostasis, and immune regulation are subject to alternative splicing to produce aberrant transcripts in healthy individuals.[20,34–36] This may reflect the need to rapidly regulate RNA expression levels during dynamic cellular processes. In *ARSB*, natural aberrant splicing might be useful to provide a level of regulation for turnover of GAGs in cartilage via ARSB-mediated degradation, which is known to be highly dynamic in response to mechanical stress.[37] Natural aberrant splicing can also play a role in human disease. Natural cryptic splice sites can overtake

canonical splicing when a canonical splice site is weakened; for example, as seen in β-thalassemia[38–41] and in Pompe disease, caused by the common c.-32-13T>G (IVS1) variant.[42] However, in the case of MPS VI, we did not observe simple competition between inclusion of a canonical versus a pseudoexon. For example, in patient 1, the c.1142+2T>C variant weakened the splice donor of exon 5, resulting in retention of the most 5′ part of intron 5 and inhibition rather than promotion of utilization of the pseudoexon in intron 5. Cryptic splice sites and pseudoexons also offer the opportunity to design antisense oligonucleotides (AONs) that block utilization of cryptic splice sites and promote canonical splicing in human disease.[20,21,43] We and others have shown recently that it is feasible for the lysosomal storage disease Pompe disease to restore *GAA* expression after aberrant splicing caused by several *GAA* variants, including c.1552-3C>G, c.1256A>T, c.2190-345A>G,[26] and the common IVS1 variant.[42,44–46] However, AONs that blocked the cryptic splice sites of the pseudoexons in *ARSB* introns 5 and 6 failed to promote canonical *ARSB* splicing (Figure S4), suggesting a more complex underlying mechanism for *ARSB* splicing. We speculate that the large sizes of introns 5 and 6 offer too many alternative options for splicing to allow a simple competition model between utilization of canonical splice sites and the cryptic splice sites of the pseudoexons.

Unexpected Aberrant Splicing and mRNA Expression in MPS VI Patients

We found that the c.1142+2T>C variant in patient 2 caused intron retention of the most 5′ part of intron 5 by utilizing a predicted cryptic splice acceptor site (patient 1; Figure 2). A similar variant, c.1142+2T>A, has been described in the literature. In that case, only cells from the father that carried c.1142+2T>A were analyzed in the absence of CHX. It was concluded that this variant causes skipping of exon 5. However, we found that exon 5 skipping also occurs in healthy control individuals, opening the possibility that c.1142+2T>A was not the cause of the exon 5 skip and that this variant may also cause intron 5 retention. This would require further analysis of cells grown in the presence of CHX because the intron retention product can be easily missed in normally growing cells because of NMD.

Patient 1 in the present study showed NMD of only the 3′ part of the mRNA (Figure 2). At present we do not have an explanation for this observation. Aberrant splicing around exon 5 shows variable levels of NMD; exon 5 skipping occurs in healthy and diseased individuals, causes a frameshift, but does not induce NMD, whereas a premature termination codon in exon 5 does induce NMD (caused by c.979C> T in patients 10 and 11; Figure 3). Although exons 1–5 encode the ARSB signal peptide and the sulfatase domains, exons 6–8 encode the C terminus of the ARSB protein, for which no conserved domain is presently known.[47,48] Future experiments are needed to elucidate the mechanisms underlying regulation of NMD around exon 5.

Patients 10 and 11 contained the c.629A>G variant on the second allele. *ARSB* mRNA expression was 20% rather than the expected 50% for all exons (the other allele underwent NMD), indicating that the c.629A>G allele was associated with low expression. We speculate that this missense was linked to another variant that was responsible for the low mRNA levels of this allele (for example, a promoter variant), which needs further analysis. Nevertheless, this result highlights the value of quantitative analysis of *ARSB* mRNA expression to detect effects on mRNA expression levels.

A Complex Case: Molecular Diagnosis at the mRNA Level

Patient 12 was deficient in ARSB enzyme activity, but no DNA variants could be identified by standard diagnostics. We found a number of aberrations, including low mRNA expression of all exons, skipping of exon 4, and gain of *ARSB* exons 2–4 in 5q14.1, which is the location of the *ARSB* gene. This helped to establish the molecular diagnosis at the mRNA level, but we could not identify the DNA variants that were responsible. Below we discuss possible scenarios for this patient.

The father contains gain of *ARSB* exons 2–4 in 5q14.1 on one allele. He has ∼50% *ARSB* mRNA expression from both alleles (data not shown) and shows very low but detectable skipping of exon 4 (Figure S5). We hypothesize that the gain of *ARSB* 2–4 landed in the ARSB gene to induce exon 4 skipping and disrupt *ARSB* expression via NMD. However, the father had only minor levels of exon 4 skipping compared with patient 12. We hypothesize that patient 12 may contain a *de novo* variant in *cis* or in *trans* that results in worsening of exon 4 skipping and NMD.

The mother also has low *ARSB* mRNA expression of ∼50% of the levels of healthy control individual 1 levels (data not shown). No disease-associated *ARSB* variant was detected in the mother. A heterozygous SNP was detected in exon 5 in mRNA from patient 12 at a 50% ratio in the absence and presence of CHX (data not shown), indicating that both alleles were expressed. This argued against a simple scenario in which the maternal allele underwent NMD. The underlying molecular mechanism remained obscure and may suggest a *de novo* variant in the maternal allele that caused low mRNA expression in patient 12.

Conclusions

In conclusion, of 12 MPS VI patients who were present in the Netherlands at the time of analysis, we identified novel aberrant splicing/mRNA expression events in 4 patients with three different genotypes. This highlights the need to implement systematic diagnostic analyses at the mRNA level to establish a molecular diagnosis and to understand the pathogenic mechanism of *ARSB* variants that can cause MPS VI. For future diagnostic implementation of the generic splicing assay, we anticipate that next-generation sequencing will be useful to make this assay suitable for diagnosis of a broad spectrum of human diseases within a single run.

MATERIALS AND METHODS

Patients and Healthy Control Individuals

Patients were diagnosed with MPS VI at the Center for Lysosomal and Metabolic Diseases of Erasmus MC, Rotterdam, the Netherlands. Diagnosis was based on clinical symptoms, ARSB enzyme deficiency in leukocytes and/or fibroblasts, and *ARSB* variant analysis. Analysis

was performed on patient material obtained with informed consent. For healthy control individuals, analysis was performed on material for an unrelated disease obtained with informed consent. The study protocol was approved by the Medical Ethical Committee at Erasmus Medical Center, Rotterdam.

Nomenclature

The nomenclature of variants and splice sites is according to Human Genome Variation Society (HGVS) standards (http://www.hgvs.org/mutnomen/). RefSeq NM_000046.5 was used as a *ARSB* reference transcript and NC_000005.10 as the genomic sequence.[49]

Splicing Prediction

Splicing prediction was carried out as described before.[25] Alamut Visual version 2.6.1 was used to predict 5′ and 3′ splice junctions using five different algorithms (see the description of algorithms at http://www.interactive-biosoftware.com/doc/alamutvisual/2.6/splicing.html).

Cell Culture

Primary human fibroblasts were cultured in high-glucose DMEM (Lonza), 10% fetal bovine serum (Thermo Fisher Scientific), and 100 U/mL penicillin/streptomycin/glutamine (Lonza). Cells were grown in the presence of 5% CO_2. Cells were passaged at 80% to 90% confluency with TrypLE (Gibco). All cell lines were routinely tested for mycoplasma infection using the MycoAlert Mycoplasma Detection Kit (Lonza) and were negative. To inhibit NMD, 20 μg/ml CHX (Sigma) was added to the medium 48 h prior to RNA isolation. Phosphorodiamidate morpholino oligo (PMO) AON design and treatment at 20 μM final concentration was performed as described previously.[26,42,44]

RNA Isolation and cDNA Preparation

RNA was isolated using the RNAeasy Miniprep Kit with on-column DNase treatment (QIAGEN) using guidelines provided by the manufacturer. 400 ng of RNA was used for generation of cDNA using the iScript cDNA Synthesis Kit (Bio-Rad). The cDNA solution was diluted 10 times before use in RT-PCR or RT-qPCR.

Flanking-Exon RT-PCR Analysis and Exon-Internal RT-qPCR Analysis

Flanking-exon RT-PCR was performed with FastStart Taq Polymerase (Roche) on a Bio-Rad S1000 thermal cycler. Exon-internal RT-qPCR was carried out using iTaq SYBR Green Supermix (Bio-Rad) on a cfx96rts cycler (Bio-Rad). All samples were measured in triplicate and normalized against Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*). All primer sets used for RT-qPCR showed high efficiency and specificity based on melting-curve analysis and standard curve measurements. Primers are shown in Table S1.

Sequencing

Direct sequencing of products identified in flanking-exon RT-PCR was performed using the Big Dye Terminator Kit v.3.1 (Applied Biosystems). To obtain pure DNA samples, PCR products visible on the gel were stabbed with a 20-μL pipette tip, and DNA on the tip was resuspended in 10 μL H_2O. A 1-μL aliquot was subsequently used in a new PCR (as described above). Excess primers and dinucleotide triphosphates (dNTPs) were removed using FastAP thermosensitive alkaline phosphatase (Thermo Fisher Scientific) according to the manufacturer's protocol. Alternatively, PCR products were cloned into a pCR2.1-TOPO (Thermo Fisher Scientific) cloning vector according to the manufacturer's protocol, and then M13 primers were used for sequencing (Table S1). Samples were purified with Sephadex G-50 (GE Healthcare), and the sequence was determined on an AB3130 genetic analyzer (Applied Biosystems).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtm.2020.09.004.

AUTHOR CONTRIBUTIONS

M.B. and W.W.M.P.P. conceived and designed the study and drafted the manuscript. M.B., K.S., and B.G. performed the experiments. W.W.M.P.P supervised the study. M.B., E.O., H.J.M.P.v.d.H., A.J.B., A.T.v.d.P., and W.W.M.P.P. were involved in data interpretation and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

This work has received funding from Zeldzame Ziekten Fonds/WE Foundation, Metakids (project number 2018-082), and Stofwisselkracht.

REFERENCES

1. Azevedo, A.C., Schwartz, I.V., Kalakun, L., Brustolin, S., Burin, M.G., Beheregaray, A.P., Leistner, S., Giugliani, C., Rosa, M., Barrios, P., et al. (2004). Clinical and biochemical study of 28 patients with mucopolysaccharidosis type VI. Clin. Genet. *66*, 208–213.

2. Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., Childs, B., Kinzler, K.W., and Vogelstein, B. (2001). The Metabolic Bases of Inherited Disease. (McGraw-Hill).

3. Spranger, J.W., Koch, F., McKusick, V.A., Natzschka, J., Wiedemann, H.R., and Zellweger, H. (1970). Mucopolysaccharidosis VI (Maroteaux-Lamy's disease). Helv. Paediatr. Acta *25*, 337–362.

4. Stumpf, D.A., Austin, J.H., Crocker, A.C., and LaFrance, M. (1973). Mucopolysaccharidosis type VI (Maroteaux-Lamy syndrome). I. Sulfatase B deficiency in tissues. Am. J. Dis. Child. *126*, 747–755.

5. Valayannopoulos, V., Nicely, H., Harmatz, P., and Turbeville, S. (2010). Mucopolysaccharidosis VI. Orphanet J. Rare Dis. *5*, 5.

6. Giugliani, R., Harmatz, P., and Wraith, J.E. (2007). Management guidelines for mucopolysaccharidosis VI. Pediatrics *120*, 405–418.

7. Swiedler, S.J., Beck, M., Bajbouj, M., Giugliani, R., Schwartz, I., Harmatz, P., Wraith, J.E., Roberts, J., Ketteridge, D., Hopwood, J.J., et al. (2005). Threshold effect of urinary glycosaminoglycans and the walk test as indicators of disease progression in a survey of subjects with Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). Am. J. Med. Genet. A. *134A*, 144–150.

8. Brands, M.M., Oussoren, E., Ruijter, G.J., Vollebregt, A.A., van den Hout, H.M., Joosten, K.F., Hop, W.C., Plug, I., and van der Ploeg, A.T. (2013). Up to five years experience with 11 mucopolysaccharidosis type VI patients. Mol. Genet. Metab. *109*, 70–76.

9. Brands, M.M., Hoogeveen-Westerveld, M., Kroos, M.A., Nobel, W., Ruijter, G.J., Özkan, L., Plug, I., Grinberg, D., Vilageliu, L., Halley, D.J., et al. (2013). Mucopolysaccharidosis type VI phenotypes-genotypes and antibody response to galsulfase. Orphanet J. Rare Dis. *8*, 51.

10. Decker, C., Yu, Z.F., Giugliani, R., Schwartz, I.V., Guffon, N., Teles, E.L., Miranda, M.C., Wraith, J.E., Beck, M., Arash, L., et al. (2010). Enzyme replacement therapy for mucopolysaccharidosis VI: Growth and pubertal development in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. J. Pediatr. Rehabil. Med. *3*, 89–100.

11. Harmatz, P., Giugliani, R., Schwartz, I.V., Guffon, N., Teles, E.L., Miranda, M.C., Wraith, J.E., Beck, M., Arash, L., Scarpa, M., et al.; MPS VI Study Group (2008). Long-term follow-up of endurance and safety outcomes during enzyme replacement therapy for mucopolysaccharidosis VI: Final results of three clinical studies of recombinant human N-acetylgalactosamine 4-sulfatase. Mol. Genet. Metab. *94*, 469–475.

12. Harmatz, P., Yu, Z.F., Giugliani, R., Schwartz, I.V., Guffon, N., Teles, E.L., Miranda, M.C., Wraith, J.E., Beck, M., Arash, L., et al. (2010). Enzyme replacement therapy for mucopolysaccharidosis VI: evaluation of long-term pulmonary function in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. J. Inherit. Metab. Dis. *33*, 51–60.

13. Tomanin, R., Karageorgos, L., Zanetti, A., Al-Sayed, M., Bailey, M., Miller, N., Sakuraba, H., and Hopwood, J.J. (2018). Mucopolysaccharidosis type VI (MPS VI) and molecular analysis: Review and classification of published variants in the ARSB gene. Hum. Mutat. *39*, 1788–1802.

14. Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D.N., and Sanford, J.R. (2011). Loss of exon identity is a common mechanism of human inherited disease. Genome Res. *21*, 1563–1571.

15. Stenson, P.D., Mort, M., Ball, E.V., Evans, K., Hayden, M., Heywood, S., Hussain, M., Phillips, A.D., and Cooper, D.N. (2017). The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum. Genet. *136*, 665–677.

16. Lim, K.H., Ferraris, L., Filloux, M.E., Raphael, B.J., and Fairbrother, W.G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. Proc. Natl. Acad. Sci. USA *108*, 11093–11098.

17. Soukarieh, O., Gaildrat, P., Hamieh, M., Drouet, A., Baert-Desurmont, S., Frébourg, T., Tosi, M., and Martins, A. (2016). Exonic Splicing Mutations Are More Prevalent than Currently Estimated and Can Be Predicted by Using In Silico Tools. PLoS Genet. *12*, e1005756.

18. Moles-Fernández, A., Duran-Lozano, L., Montalban, G., Bonache, S., López-Perolio, I., Menéndez, M., Santamariña, M., Behar, R., Blanco, A., Carrasco, E., et al. (2018). Computational Tools for Splicing Defect Prediction in Breast/Ovarian Cancer Genes: How Efficient Are They at Predicting RNA Alterations? Front. Genet. *9*, 366.

19. Schleit, J., Bailey, S.S., Tran, T., Chen, D., Stowers, S., Schwarze, U., and Byers, P.H. (2015). Molecular Outcome, Prediction, and Clinical Consequences of Splice Variants in COL1A1, Which Encodes the proα1(I) Chains of Type I Procollagen. Hum. Mutat. *36*, 728–739.

20. Bergsma, A.J., van der Wal, E., Broeders, M., van der Ploeg, A.T., and Pim Pijnappel, W.W.M. (2018). Alternative Splicing in Genetic Diseases: Improved Diagnosis and Novel Treatment Options. Int. Rev. Cell Mol. Biol. *335*, 85–141.

21. Kuijper, E.C., Bergsma, A.J., Pijnappel, W.W.M.P., and Aartsma-Rus, A. (2020). Opportunities and challenges for antisense oligonucleotide therapies. J. Inherit. Metab. Dis. Published online May 11, 2020. https://doi.org/10.1002/jimd.12251.

22. Arunkumar, N., Langan, T.J., Stapleton, M., Kubaski, F., Mason, R.W., Singh, R., Kobayashi, H., Yamaguchi, S., Suzuki, Y., Orii, K., et al. (2020). Newborn screening of mucopolysaccharidoses: past, present, and future. J. Hum. Genet. *65*, 557–567.

23. Scott, C.R., Elliott, S., Hong, X., Huang, J.Y., Kumar, A.B., Yi, F., Pendem, N., Chennamaneni, N.K., and Gelb, M.H. (2020). Newborn Screening for Mucopolysaccharidoses: Results of a Pilot Study with 100 000 Dried Blood Spots. J. Pediatr. *216*, 204–207.

24. Kubaski, F., de Oliveira Poswar, F., Michelin-Tirelli, K., Burin, M.G., Rojas-Málaga, D., Brusius-Facchin, A.C., Leistner-Segal, S., and Giugliani, R. (2020). Diagnosis of Mucopolysaccharidoses. Diagnostics (Basel) *10*, E172.

25. Bergsma, A.J., Kroos, M., Hoogeveen-Westerveld, M., Halley, D., van der Ploeg, A.T., and Pijnappel, W.W. (2015). Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach. Hum. Mutat. *36*, 57–68.

26. Bergsma, A.J., In 't Groen, S.L., Verheijen, F.W., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2016). From Cryptic Toward Canonical Pre-mRNA Splicing in Pompe Disease: a Pipeline for the Development of Antisense Oligonucleotides. Mol. Ther. Nucleic Acids *5*, e361.

27. In 't Groen, S.L.M., de Faria, D.O.S., Iuliano, A., van den Hout, J.M.P., Douben, H., Dijkhuizen, T., Cassiman, D., Witters, P., Barba Romero, M.Á., de Klein, A., et al. (2020). Novel *GAA* Variants and Mosaicism in Pompe Disease Identified by Extended Analyses of Patients with an Incomplete DNA Diagnosis. Mol. Ther. Methods Clin. Dev. *17*, 337–348.

28. Garrido, E., Chabás, A., Coll, M.J., Blanco, M., Domínguez, C., Grinberg, D., Vilageliu, L., and Cormand, B. (2007). Identification of the molecular defects in Spanish and Argentinian mucopolysaccharidosis VI (Maroteaux-Lamy syndrome) patients, including 9 novel mutations. Mol. Genet. Metab. *92*, 122–130.

29. Garrido, E., Cormand, B., Hopwood, J.J., Chabás, A., Grinberg, D., and Vilageliu, L. (2008). Maroteaux-Lamy syndrome: functional characterization of pathogenic mutations and polymorphisms in the arylsulfatase B gene. Mol. Genet. Metab. *94*, 305–312.

30. Litjens, T., Brooks, D.A., Peters, C., Gibson, G.J., and Hopwood, J.J. (1996). 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*, 1127–1134.

31. Sibley, C.R., Emmett, W., Blazquez, L., Faro, A., Haberman, N., Briese, M., Trabzuni, D., Ryten, M., Weale, M.E., Hardy, J., et al. (2015). Recursive splicing in long vertebrate genes. Nature *521*, 371–375.

32. Piovesan, A., Caracausi, M., Antonaros, F., Pelleri, M.C., and Vitale, L. (2016). GeneBase 1.1: a tool to summarize data from NCBI gene datasets and its application to an update of human gene statistics. Database (Oxford) *2016*, baw153.

33. Dyle, M.C., Kolakada, D., Cortazar, M.A., and Jagannathan, S. (2020). How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. Wiley Interdiscip. Rev. RNA *11*, e1560.

34. Baralle, F.E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. Nat. Rev. Mol. Cell Biol. *18*, 437–451.

35. Lee, Y., and Rio, D.C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu. Rev. Biochem. *84*, 291–323.

36. Dhir, A., and Buratti, E. (2010). Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies. FEBS J. *277*, 841–855.

37. Vynios, D.H. (2014). Metabolism of cartilage proteoglycans in health and disease. BioMed Res. Int. *2014*, 452315.

38. Dobkin, C., and Bank, A. (1983). A nucleotide change in IVS 2 of a beta-thalassemia gene leads to a cryptic splice not at the site of the mutation. Prog. Clin. Biol. Res. *134*, 127–128.

39. Dobkin, C., and Bank, A. (1985). Reversibility of IVS 2 missplicing in a mutant human beta-globin gene. J. Biol. Chem. *260*, 16332–16337.

40. Wong, C., Antonarakis, S.E., Goff, S.C., Orkin, S.H., Forget, B.G., Nathan, D.G., Giardina, P.J., and Kazazian, H.H., Jr. (1989). Beta-thalassemia due to two novel nucleotide substitutions in consensus acceptor splice sequences of the beta-globin gene. Blood *73*, 914–918.

41. Treisman, R., Orkin, S.H., and Maniatis, T. (1983). Specific transcription and RNA splicing defects in five cloned beta-thalassaemia genes. Nature *302*, 591–596.

42. van der Wal, E., Bergsma, A.J., van Gestel, T.J.M., In 't Groen, S.L.M., Zaehres, H., Araúzo-Bravo, M.J., Schöler, H.R., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2017). GAA Deficiency in Pompe Disease Is Alleviated by Exon Inclusion in iPSC-Derived Skeletal Muscle Cells. Mol. Ther. Nucleic Acids *7*, 101–115.

43. Bennett, C.F. (2019). Therapeutic Antisense Oligonucleotides Are Coming of Age. Annu. Rev. Med. *70*, 307–321.

44. van der Wal, E., Bergsma, A.J., Pijnenburg, J.M., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2017). Antisense Oligonucleotides Promote Exon Inclusion and Correct the Common c.-32-13T>G GAA Splicing Variant in Pompe Disease. Mol. Ther. Nucleic Acids *7*, 90–100.

45. Aung-Htut, M.T., Ham, K.A., Tchan, M., Johnsen, R., Schnell, F.J., Fletcher, S., and Wilton, S.D. (2020). Splice modulating antisense oligonucleotides restore some acid-alpha-glucosidase activity in cells derived from patients with late-onset Pompe disease. Sci. Rep. *10*, 6702.

46. Goina, E., Peruzzo, P., Bembi, B., Dardis, A., and Buratti, E. (2017). Glycogen Reduction in Myotubes of Late-Onset Pompe Disease Patients Using Antisense Technology. Mol. Ther. *25*, 2117–2128.

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

48. Pfam. Protein: ARSB_HUMAN (P15848). http://pfam.xfam.org/protein/P15848.

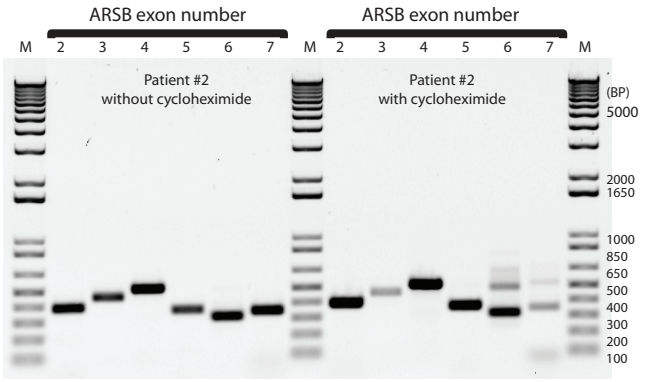
49. den Dunnen, J.T., Dalgleish, R., Maglott, D.R., Hart, R.K., Greenblatt, M.S., McGowan-Jordan, J., Roux, A.F., Smith, T., Antonarakis, S.E., and Taschner, P.E. (2016). HGVS Recommendations for the Description of Sequence Variants: 2016 Update. Hum. Mutat. *37*, 564–569.

OMTM, Volume *19*

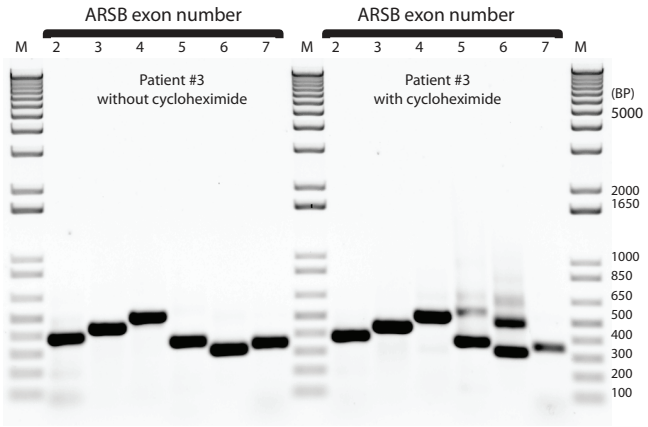
Supplemental Information

A Generic Assay to Detect Aberrant *ARSB* Splicing and mRNA Degradation for the Molecular Diagnosis of MPS VI

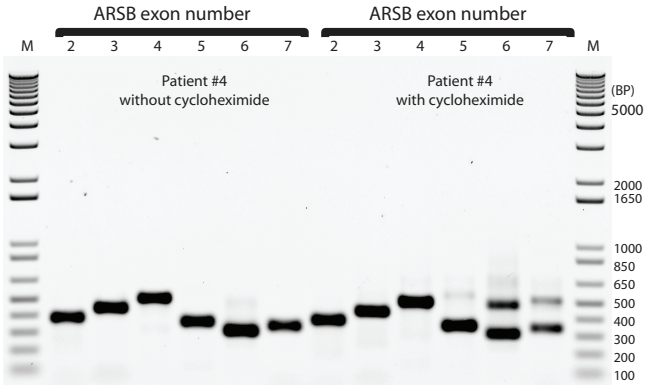
Mike Broeders, Kasper Smits, Busra Goynuk, Esmee Oussoren, Hannerieke J.M.P. van den Hout, Atze J. Bergsma, Ans T. van der Ploeg, and W.W.M. Pim Pijnappel



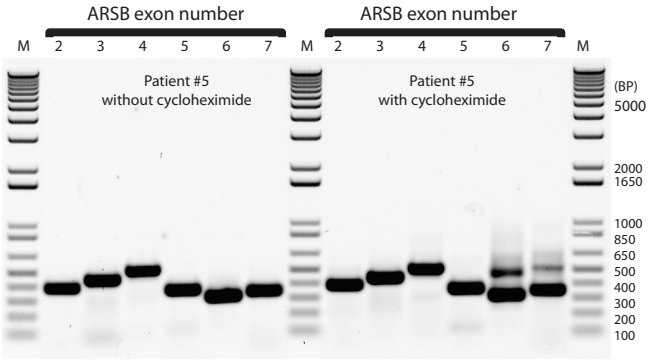
c.454C>T - Exon 2 - p.(R152W)
c.454C>T - Exon 2 - p.(R152W)



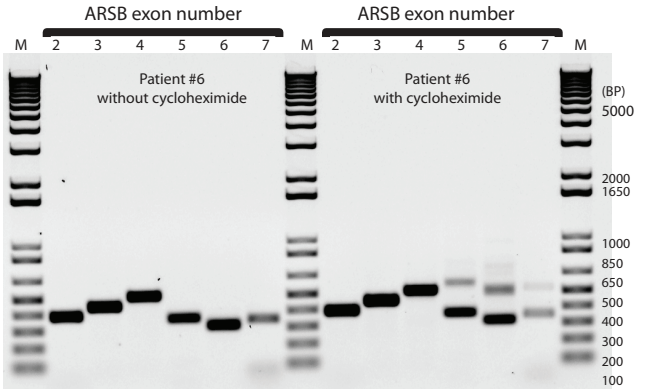
c.454C>T - Exon 2 - p.(R152W)
c.454C>T - Exon 2 - p.(R152W)



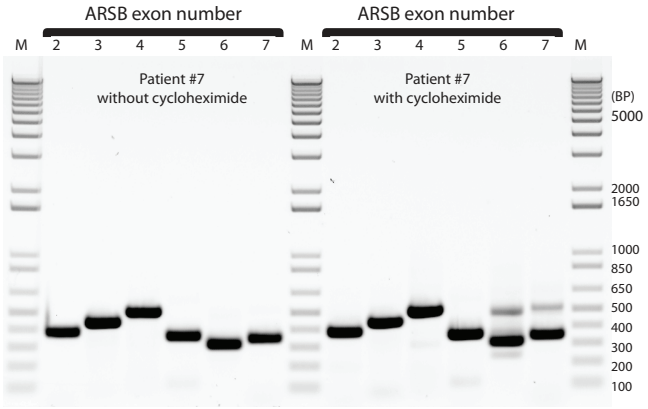
c.629A>G - Exon 3 - p.(Y210C)
c.937C>G - Exon 5 - p.(P313A)



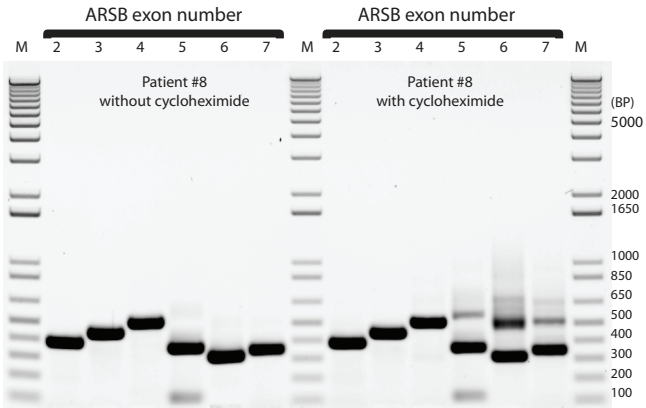
c.971G>T - Exon 5 - p.(G324V)
c.971G>T - Exon 5 - p.(G324V)



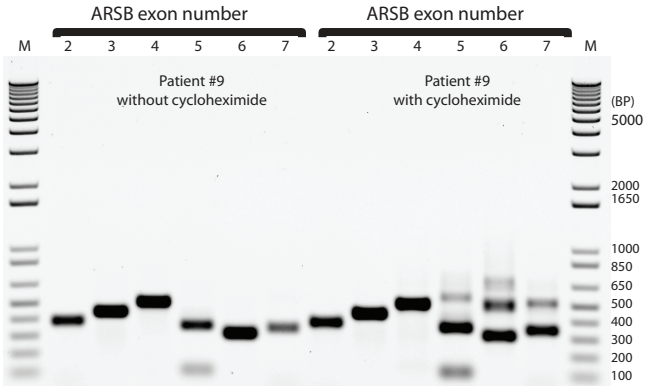
c.937C>T - Exon 5 - p.(P313S)
c.937C>T - Exon 5 - p.(P313S)



c.995T>G - Exon 5 - p.(V332G)
c.995T>G - Exon 5 - p.(V332G)



c.937C>G - Exon 5 - p.(P313A)
c.937C>G - Exon 5 - p.(P313A)



c.[903C>G;1151G>A] - Exon 5 & 6 - p.[(N301K);(S384N)]
c.[903C>G;1151G>A] - Exon 5 & 6 - p.[(N301K);(S384N)]

Figure S1. Analysis of patient #2 to #9 by flanking exon RT-PCR. The disease-associated variants and their location are indicated.

Figure S2. In silico prediction of splice donors and acceptors used by pseudo exons in ARSB intron 5 and intron 6.

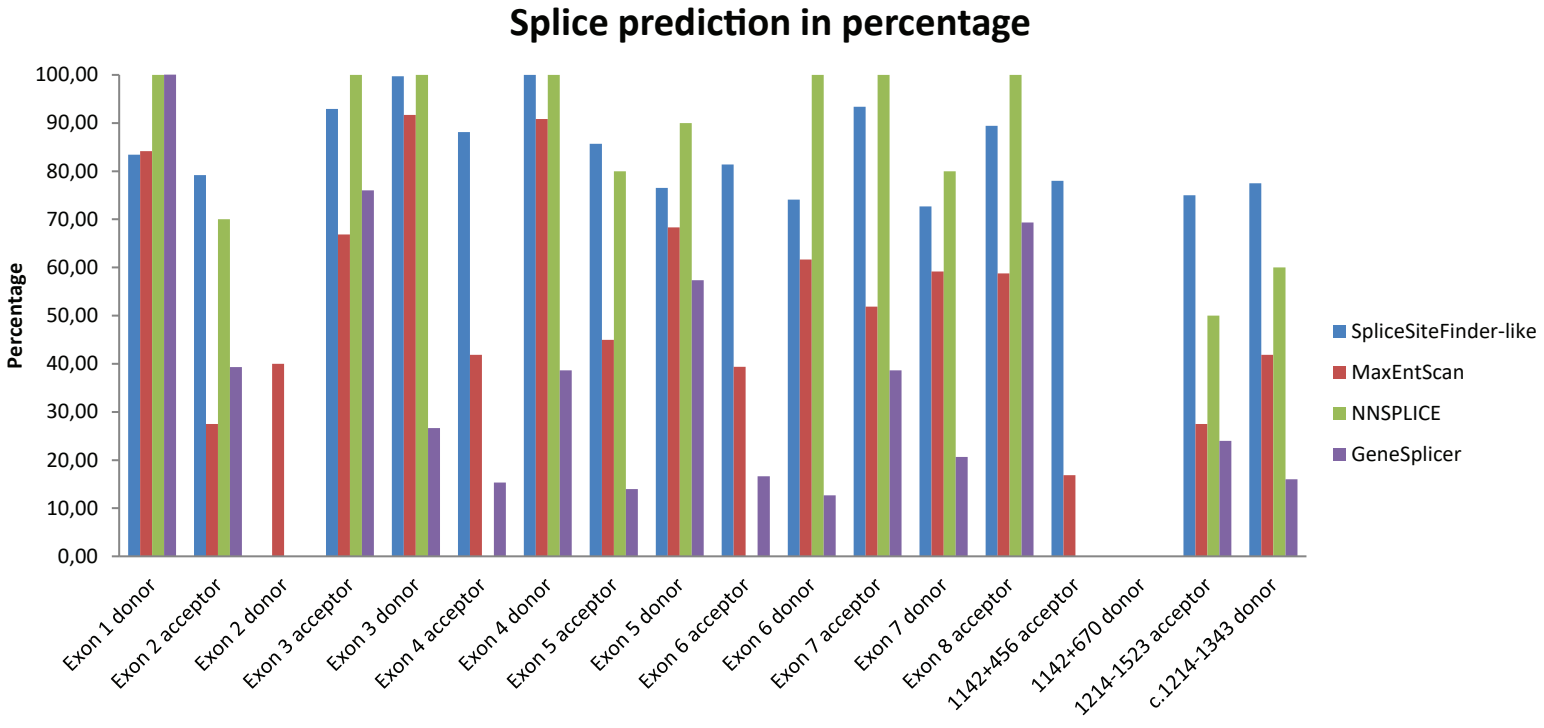


Figure S3. In silico splicing prediction of the strengths of canonical splice sites and cryptic splice sites utilized for pseudo exon inclusion.

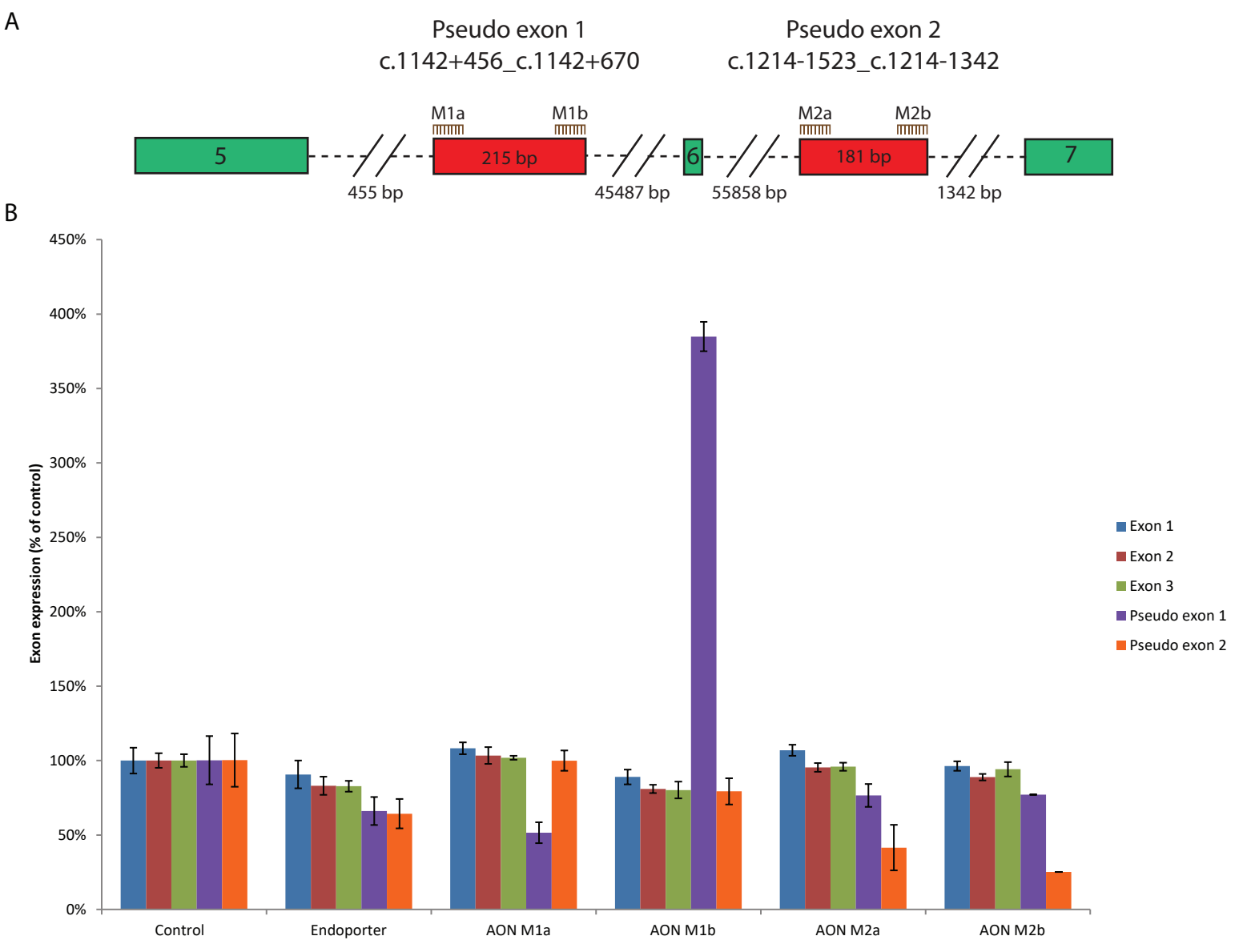


Figure S4. Antisense oligonucleotide (AON) treatment to block pseudo exon inclusion. **A:** Strategy for the AON treatment. PMO AONs (indicated in brown and shown in Tabel S1) were designed to block the splice acceptor and splice donor sites of pseudo exons 1 and 2 (indicated in red). **B:** Exon internal RT-qPCR analysis after AON treatment of healthy control fibroblasts. Treatment was for 24 hours. Although pseudo exon inclusion was decreased after AON treatment for pseudo exon 1 with M1a and for pseudo exon 2 with M2a and M2b, expression of canonical ARSB mRNA (measured for exons 1 to 3) did not increase. M1b promoted pseudo exon 1 inclusion by almost 4-fold for unknown reasons and decreased the overall expression slightly. GAPDH was used for normalization. Error bars indicate SD (n = 3).

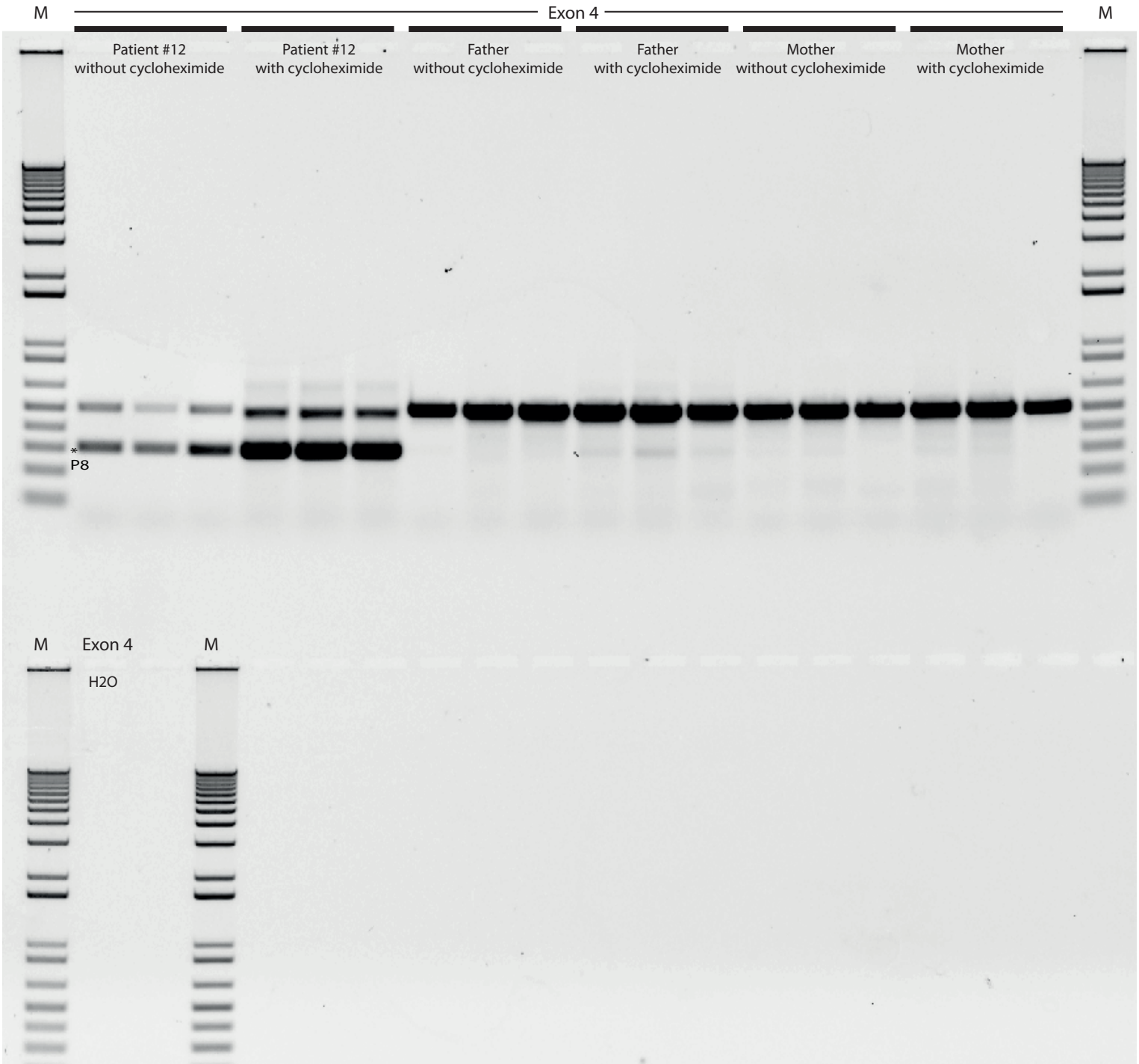


Figure S5. Splicing analysis of ARSB exon 4 of patient #12 and parents in triplicate.

Supplemental Table 1. Primers used for the RT-PCR, RT-qPCR and sequence reactions.

RT-qPCR		
Name	**Sequence (5'-3')**	**Product size (nt)**
GAPDH FW	ATGGGGAAGGTGAAGGTCG	70
GAPDH RV	TAAAAGCAGCCCTGGTGACC	
Exon 1 Fw	CAGACGACCTAGGCTGGAAC	101
Exon 1 Rv	GTAGTTGTCCAGGAGCACCC	
Exon 2 Fw	CCTGCCCCAGCTCCTAAAAG	108
Exon 2 Rv	GTATCAAATCCTCGGCGGGT	
Exon 3 Fw	ACGCTCTGAATGTCACACGA	129
Exon 3 Rv	GTGGATGGTTAGTTATGAGGGCT	
Exon 4 Fw	TCTCCAGTCTGTGCATGAGC	106
Exon 4 Rv	AAGGGACACCATTCCTGCAT	
Exon 5 Fw	GGCCCCTTCGAGGAAGAAAA	118
Exon 5 Rv	GAGATGTGGATGAGCTCCCG	
Exon 6 Fw	AAGCCCATCCCCCAGAATTG	64
Exon 6 Rv	CGGTGAAGAGTCCACGAAGT	
Exon 7 Fw	ATGGCTCCAGCAAAGGATGA	102
Exon 7 Rv	GCCCGTGAGGAGTTTCCAAT	
Exon 8 Fw	CACCAACCAAGACCCTCTGG	118
Exon 8 Rv	TGGTAGAACTGTAGGCGGGA	
RT-PCR		
Name	**Sequence (5'-3')**	**Product size (nt)**
Exon 1-3 Fw	GGGTGCTCCTGGACAACTAC	379
Exon 1-3 Rv	CCTGTTGCAACTTCTTCGCC	
Exon 2-4 Fw	ATGGCACCTGGGAATGTACC	444
Exon 2-4 Rv	GTGTTGTTCCAGAGCCCACT	
Exon 3-5 Fw	ACGCTCTGAATGTCACACGA	514
Exon 3-5 Rv	GTTGGCAGCCAGTCAGAGAT	
Exon 4-6 Fw	AAAAAGCAGTGGGCTCTGGA	361
Exon 4-6 Rv	CGGTGAAGAGTCCACGAAGT	
Exon 5-7 Fw	CAGAAGGGCGTGAAGAACCG	314
Exon 5-7 Rv	CCCGTGAGGAGTTTCCAATTTC	
Exon 6-8 Fw	ACTTCGTGGACTCTTCACCG	348
Exon 6-8 Rv	AGTACACGGGGACTGAGTGT	
M13 sequence primers (5'-3')		
M13 Rv	CAGGAAACAGCTATGAC	
M13 Fw	GTAAAACGACGGCCAG	
PMO AONs (5'-3')		
M1a	TTCTCAGCATCTAGAAGAAGGTATG	
M1b	ACGATTTCCAAATCTGAGTCTGGGT	
M2a	AGCAGCAAATCTTTAGCACCAAAGA	
M2b	CCTTATCCTCACCCAAGCAGGACTT	