

## ARTICLE

# A rare variant (c.863G > T) in exon 7 of *SMN1* disrupts mRNA splicing and is responsible for spinal muscular atrophy

Yu-jin Qu<sup>1</sup>, Jin-li Bai<sup>1</sup>, Yan-yan Cao<sup>1</sup>, Wen-hui Zhang<sup>1</sup>, Hong Wang<sup>1</sup>, Yu-wei Jin<sup>1</sup> and Fang Song<sup>\*,1</sup>

Proximal spinal muscular atrophy (SMA) is an autosomal recessive neuromuscular disorder caused by deletion or mutation of *SMN1* (survival motor neuron 1). *SMN* exon 7 splicing is regulated by a number of exonic and intronic regulatory sequences and the trans-factors that bind them. Variants located in or near these regulated regions should be evaluated to determine their effect on splicing. We identified the rare variant c.863G > T (r.835\_\*3del, p.Gly279Glufs\*5) in exon 7 of *SMN1* in three patients affected with type I or type II SMA. Most of the *SMN1* transcripts exhibited complete loss of exon 7 *in vivo*. The *ex vivo* splicing assay demonstrated that the variant disrupts inclusion of exon 7 (~85%) in the *SMN1* mRNA; replacement with various bases yielded a variety of splicing effects in *SMN1* and *SMN2* pre-mRNA. The c.863G > T (r.835\_\*3del, p.Gly279Glufs\*5) variant is located in a region that includes binding sites for multiple splicing factors including Tra2β1. Thus, the variant disrupts Tra2β1 binding, but does not affect binding of hnRNP A1. These findings demonstrate how rare variants influence pre-mRNA splicing of *SMN* and reveal the functional influence of c.863G > T (r.835\_\*3del, p.Gly279Glufs\*5) variant in patients with SMA. *European Journal of Human Genetics* (2016) 24, 864–870; doi:10.1038/ejhg.2015.213; published online 30 September 2015

## INTRODUCTION

Proximal spinal muscular atrophy (SMA) is a common neuromuscular disorder caused by loss of α-motor neurons in the spinal cord due to homozygous deletion or mutation of the survival motor neuron 1 gene (*SMN1*). SMA is the most frequent genetic cause of infantile death with an incidence of ~1 in 6000–10000 live births and a carrier frequency of 1 in 50.<sup>1–3</sup> The *SMN1* gene is located on chromosome 5q13<sup>4</sup> and produces full-length (FL) *SMN* mRNA. An almost identical gene, named *SMN2*, expresses abundant levels of transcript lacking exon 7 due to a C-to-T nucleotide substitution at position 6 within exon 7; the transcript encodes an SMNΔ7 protein, which disrupts SMN oligomerization, thus enhancing degradation of the SMN monomer.<sup>5</sup> Childhood SMA is divided into types (I–III) based on the age of onset and achieved maximum motor abilities. SMA severity is inversely correlated with *SMN2* gene copy number.<sup>6–9</sup>

Alternative pre-mRNA splicing is an important mechanism of gene regulation. Accurate exon identification requires classical splicing signals and various cis-elements such as exonic splicing enhancers (ESEs) and silencers (ESSs) (reviewed in ref. 10–12). Multiple exonic cis-elements within the *SMN* gene and its regulators mediate splicing of *SMN* exon 7, which includes several negative and positive cis-elements<sup>13</sup> such as Exinct, SF2/ASF-ESE, hnRNP A1-ESS, Tra2β1-ESE, and the 3'Cluster.

Variants that alter exonic splicing regulatory elements usually affect the functions in human genetic disease.<sup>14</sup> In this paper, we identified a rare variant (c.863G > T, r.835\_\*3del, p.Gly279Glufs\*5) in exon 7 of the *SMN1* gene in three patients with different phenotypes. This variant lies adjacent to the Tra2β1 enhancer element. We demonstrated the c.863G > T (r.835\_\*3del, p.Gly279Glufs\*5) variant

interferes with the correct splicing of exon 7 due to disruption of Tra2β1. Our results suggest the variant abolishes *SMN1* exon 7 inclusion *in vivo* and in an *ex vivo* minigene splicing assay, thus affecting the function of SMN.

## MATERIALS AND METHODS

### Patients and controls

Patient 1 (SMA I, ID: sm08120) and patient 2 (SMA II, ID:sm08117) were reported in a previous study.<sup>15</sup> Patient 3 (ID:sm14002) was a male newborn infant diagnosed with SMA I. In the late stages of pregnancy, the fetal movement decreased significantly. When the little boy was born, he showed faint cry, profound hypotonia, symmetrical flaccid paralysis; and poor suckling and swallowing abilities after birth. The infant died at age 20 days due to respiratory failure caused by severe pneumonia. He was the second affected patient of this family, the first affected patient was his elder sister with similar severe phenotype. She was born normally, showed abnormal symptom at 2 months old with hypotonia and symmetrical flaccid paralysis, and died at 3 months old with severe pneumonia and respiratory failure. Informed consent was obtained from all three families to participate in this study but RNA would not be obtained from patient 3 because he died prior to genetic diagnosis. In addition, only patient 1 agreed to the establishment of an *ex vivo* skin fibroblast cell line. The skin fibroblast cell lines from two healthy children (N1 and N2) with two copies of *SMN1* gene and two copies of *SMN2* gene were also established as the controls. This study was approved by the Capital Institute of Pediatric Ethics Committee.

### Sequencing and copy number determination

The *SMN* gene from exon 7 to exon 8 was PCR-amplified from genomic DNA with primers R111 and 541C1120.<sup>4</sup> The PCR products were subcloned and 8–10 *SMN1* clones were sequenced on an ABI 3730 automatic sequencer (Applied Biosystems, Foster, CA, USA) as described.<sup>16</sup> *SMN1*, *SMN2*, *NAIP*,

<sup>1</sup>Department of Medical Genetics, Capital Institute of Pediatrics, Beijing, P.R. China

\*Correspondence: Professor F Song, Department of Medical Genetics, Capital Institute of Pediatrics, No 2, Ya Bao Road, ChaoYang District, Beijing 100020, P.R. China. Tel: +86 10 85695541; Fax: +86 10 85628367; E-mail: songf\_558@263.net or ychsi1234@263.net

Received 16 March 2015; revised 10 August 2015; accepted 25 August 2015; published online 30 September 2015

*P44*, and *H4F5* gene copy number was determined by multiplex ligation-dependent probe amplification (MLPA) with a SALSA MLPA kit (P021-A1; MRC-Holland, Amsterdam, The Netherlands). The sequencing reads were aligned with the *SMN1* genomic DNA reference sequence (NG\_008691.1) and mRNA reference sequences (NM\_000344.3). And the variant was submitted to the Leiden Muscular Dystrophy *SMN1* Mutation Database ([www.LOVD.nl/SMN1](http://www.LOVD.nl/SMN1) (database ID SMN1\_00057)).

### Cell culture

Patient 1 fibroblast primary cell culture was established by the tissue explant method. Fibroblast and HEK293 cells were grown in Dulbecco's modified Eagle's medium (Gibco, New York, NY, USA) with 10~15% fetal bovine serum (Gibco, Newcastle, NSW, Australia), 100 U/ml penicillin and 100 µg/ml streptomycin at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere.

### Transcription of *SMN1*

Total RNA was isolated from the peripheral blood and/or skin fibroblast cells of patients and controls. First-strand cDNA synthesis was performed with 0.5 µg total RNA, random primers, and M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) according to manufacturer's protocols. Specific PCR primers (SMN575<sup>17</sup> and 541C1120<sup>4</sup>) were used to amplify *SMN* exons 1–8 (including *SMN1* and *SMN2*) with LA Taq polymerase (TAKARA, Kyoto, Japan). The *SMN* transcript-products were subcloned into pGEM-T Easy (Promega, Madison, WI, USA); *SMN1* subclones were screened by restriction digestion (*DraI* and *DdeI*) and sequenced.

Real-time PCR was used to quantify *SMN* transcripts as described by Tiziano *et al*<sup>18</sup>. Primers SMN\_mgb-F and SMN\_mgb-R were used to amplify the *SMN1* and *SMN2* genes; the specific probes for *SMN1* and *SMN2* are listed in Table 1. GAPDH served as the internal control (primers GAPDH\_abs-F and GAPDH\_abs-R; probe GAPDH-MGB).<sup>18</sup> Each 20-µl reaction contained 2×GoldStar TaqMan (KANGWEI, Pekin, China), 20 ng cDNA, 0.4 µl each primer (10 pmol/µl), and 4 pmol of the *SMN1*, *SMN2*, or GAPDH probe; reactions were performed on a 7500 Real-Time PCR System (Applied Biosystems). Samples were assayed in duplicate and repeated at least three times; data analysis was performed in SDS version 1.4 (Applied Biosystems).

### Plasmid construction

*SMN* minigenes (pEasy-M2-*SMN1*, pEasy-M2-*SMN2*, and pEasy-M2-*SMN1E7/SMN2E8*) were constructed as follows. A 5.8-kb segment of the human *SMN1*, hybrid *SMN1E7/SMN2E8*, and *SMN2* (exon 6, intron 6, 54 nt of exon 7, intron 7, and exon 8) were amplified with primers hSMNE6F and

hSMNE8R (Table 1) from human genomic DNA template. The amplified products were cloned into expression vector pEasy-M2 (TRANS, Pekin, China) and sequenced to ensure that no base substitutions had been acquired during cloning. The resulting plasmids were pEasy-M2-*SMN1*, pEasy-M2-*SMN1E7/SMN2E8*, and pEasy-M2-*SMN2*. Short minigenes were constructed by amplifying the three wild-type plasmids with overlapping primers (SMN16F and SMN16R). Plasmids were constructed by PCR-based site-directed mutagenesis with fly-*pfu* DNA polymerase (TRANS) with the three short mini-plasmids as templates. We constructed one splicing-positive plasmid c.853-1G>A and three mutant plasmids (863A, 863C, 863T) with the primers described in Table 1. These short minigenes consisted of exon 6, part of intron 6, 54 nt of exon 7, intron 7, and exon 8.

### Transfection and *ex vivo* splicing analysis

For transfection, 1 × 10<sup>5</sup> HEK293 cells were seeded in each well of a six-well plate in DMEM with 10% FBS. On the second day, 2 µg of each minigene plasmid was mixed with 10 µl TransLipid Transfection Reagent (TRANS) for 20 min and transfected in 1 ml serum-/antibiotic-free medium. After 4–6 h, the medium was replaced with 2 ml antibiotic-free DMEM with 10% FBS. Cells were harvested after 24 h and total RNA was extracted. Reverse transcription was performed in a total volume of 40 µl from 5 µg total RNA using random primers and M-MLV Reverse Transcriptase (Invitrogen). To ensure amplification of plasmid-derived transcripts, we amplified the cDNA from the transfected minigenes with plasmid-specific forward primer pEasy-M2F (Table 1) and the FAM-labeled *SMN*-specific reverse primer SMNE8R (Table 1) under the following cycle conditions: 94 °C for 5 min; 28 cycles of 90 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min; 72 °C for 8 min. Amplified products were detected by 2% agarose gel electrophoresis and ethidium bromide staining. To calculate the splicing efficiency, a semi-quantitative analysis was performed with a plasmid-specific primer and a FAM-labeled *SMN*-specific primer; the PCR products were run on the ABI 3730 automatic sequencing system (Applied Biosystems) and the raw data were analyzed with Gene marker version 1.75. Splicing efficiency was calculated by dividing the FL transcripts by the sum of the FL and truncated transcripts. These assays were repeated at least three times.

### RNA pull-down

RNA-binding assays were performed with an RNA–protein binding assay kit (Thermo Scientific, Rockford, IL, USA). Biotin-labeled RNA SMNr.863 g (5'-biotin-AAAGAAGGAAGGUGCUCACAUUC) and SMN r.863 u (5'-biotin-AAAGAAGGAAUGUGCUCACAUUC) were synthesized by Invitrogen. The same biotin-labeled level of RNAs were added to an equivalent volume of beads and incubated for 60 min at 4 °C with rotation. After washing, HeLa nuclear extract was added to the RNA-bead mixture and incubated for 2 h at 4 °C with rotation. Bound complexes were washed three times with buffer and eluted in SDS-PAGE loading buffer. Bound proteins were analyzed by SDS-PAGE and immunoblotting with anti-Tra2β1, anti-hnRNP A1, anti-SF2/ASF antibodies (Abcam, Cambridge, UK). The experiments were repeated at least three times.

## RESULTS

### The rare *SMN1*.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant was associated with lower levels of FL *SMN1* transcript in the peripheral blood of two patients and skin fibroblasts of patient 1

Clinical characteristics are described in Table 2. The sequences of *SMN1* exon 7-8 genomic DNA from all three patients revealed a base substitution of G to T at position c.863 (position +29 in *SMN1* exon 7). All three patients carried only one *SMN1* copy. Patients 1 and 2 carried three copies of *SMN2*, indicating that one *SMN1* gene was converted to *SMN2*. The remaining modifier genes (*NAIP*, *P44*, and *H4F5*) were present in the normal two copies. Patient 3 carried two *SMN2* genes, zero copies of *NAIP*, and one copy of *P44* and *H4F5* (Table 2).

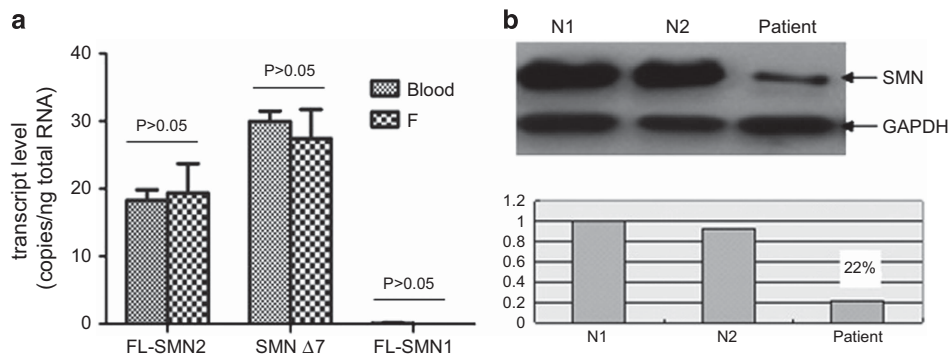
We used real-time PCR to measure the effect of the *SMN1*c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant on levels

**Table 1** Oligonucleotide sequences

Primer	Sequence (5'–3')
SMN_mgb-F	TGGTACATGAGTG GCTATCATACTG
SMN_mgb-R	GTGAGCACC TTCCTCTTTTT
SMN1 probe	FAM-ATG GGTTCAGAA-MGB-NFQ
SMN2 probe	FAM-ATGGGTTTTAGAA- MGB -NFQ
hSMNE6F	CGATCTCGAGATAATCCCCACCACTC
hSMNE8R	GCTACCC GGGCACATACGCCACATACA
SMN16F	AT TATATGGTAG GTAATCACTC AGCATCTT
SMN16R	GAGTGATTACCTACCATATAATAGCCAGTATGA
c.835-1A-F	TTCCTTTATT TTCCTTACAA GGTTCAGAC AAA
c.835-1A-R	TTGTAAGGAAAATAAAGGAAGTTAAAAAATA
c.863A-F	CAAAATCAAAAAGAAGGAAAGTGCTCACATTCC
c.863A-R	TTTCCTCTTTTTGATTTTGCTGAAACCCCTGT
c.863T-F	CAAAATCAAAAAGAAGGAATGTGCTCACATTCC
c.863T-R	ATTCTCTTTTTGATTTTGCTGAAACCCCTGT
c.863C-F	CAAAATCAAAAAGAAGGAAAGTGCTCACATTCC
c.863C-R	GTTCTCTTTTTGATTTTGCTGAAACCCCTGT
pEasy-M2F	TCTAGAGGATCGCCCTTCGA
SMNE8R-FAM	FAM-GTGGTGCATTTAGTGCTGT

**Table 2 SMA patient characteristics associated with c.863G>T(r.835\_\*3del, p.Gly279Gluufs\*5)**

Case No	Phenotype	Gender	Age	Age at				Progression	Gene copy numbers				
				Age onset	diagnosis	Maximum motor function			<i>SMN1</i>	<i>SMN2</i>	<i>NAIP</i>	<i>p44</i>	<i>H4F5</i>
1	Type Ib	F	6 years 11 months	5 months	18 months	Sit with supported, not stand	Alive, severe scoliosis, limbs muscle atrophy	1	3	2	2	2	
2	Type II	F	8 years 4 months	17 months	2 years 10 months	Sit and stand unsupported, not walk	Alive, can sit not stand, tongue-trembling, hands-trembling	1	3	2	2	2	
3	Type I	M	Died at 20 days old	0 day	30 days	Profound hypotonia, cannot sit and stand	Died in respiratory failure caused by severe pneumonia	1	2	0	1	1	



**Figure 1** *SMN* transcript and SMN protein levels in patient 1. (a) The levels of *FL-SMN1*, *FL-SMN2*, and *SMNΔ7* transcript did not differ between blood cells and skin fibroblasts; *FL-SMN1* transcript levels were near the limit of detection. (b) Semiquantitative detection of SMN protein in skin fibroblasts of patient 1. SMN protein expression was very low, only 22% of control levels (two normal individuals carrying two copies of *SMN1* gene and two *SMN2*).

of *FL-SMN1* (*FL-SMN1*) transcript in the blood of patients 1 and 2. Our results showed that the amount of *FL-SMN1* was significantly reduced, consistent with prior report.<sup>16</sup> Sequencing of the *SMN1* clones also showed that the *SMN1* transcript was missing exon 7. In comparison with normal controls with the same *SMN2* copy numbers, the amount of *FL-SMN1* transcript was very low and SMN protein levels were significantly reduced (only 22% of control levels) in the skin fibroblasts of patient 1 (Figure 1). Our results also showed no significant difference in the levels of *SMN* transcripts (*FL-SMN1*, *FL-SMN2*, and *SMNΔ7*) between the peripheral blood and skin fibroblasts of patient 1 (Figure 1). Thus, the c.863G>T (r.835\_\*3del, p.Gly279Gluufs\*5) variant was associated with reduced inclusion of exon 7 and lower expression of *FL-SMN1* mRNA and SMN protein.

#### Minigene construction and analysis

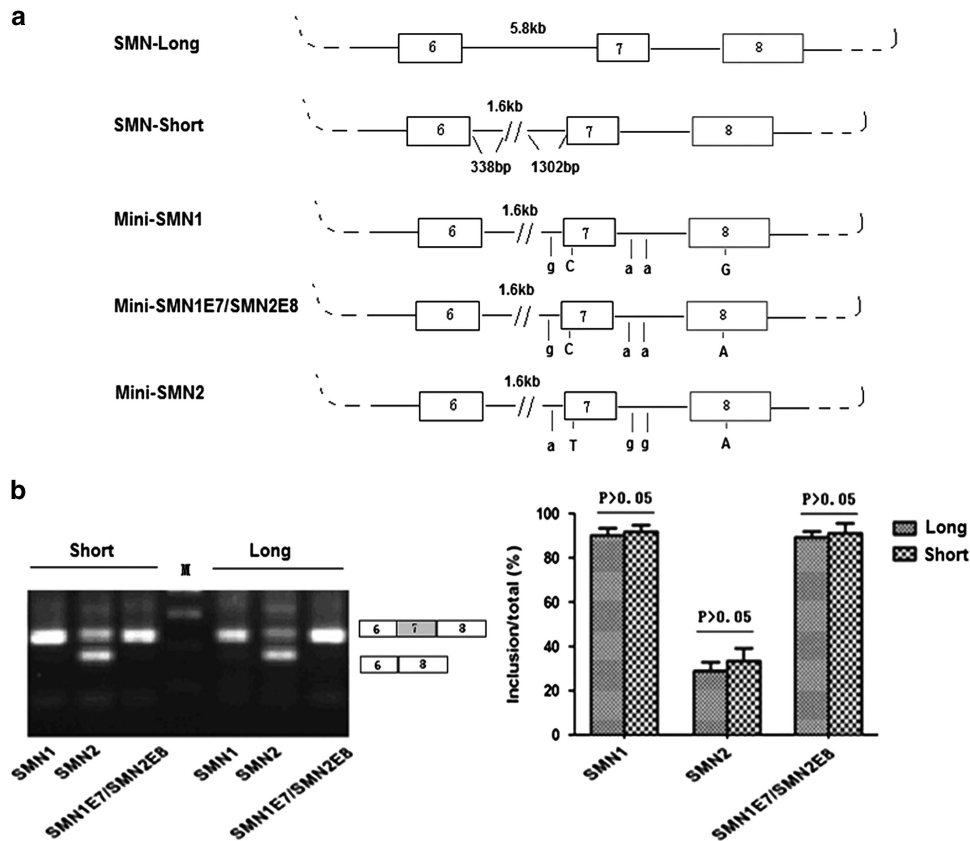
Two types of minigenes were constructed (Figure 2a). The *SMN*-Long construct contains exon 6, intron 6, exon 7, intron 7, and exon 8; *SMN*-Short was missing ~4.2 kb of intron 6. For each plasmid, we constructed three minigenes: *SMN1*, *SMN2*, and a hybrid *SMN1E7/SMN2E8* gene. To compare the splicing effects of these minigenes, we measured the percentage of *FL-SMN* transcript by RT-PCR after transfection of HEK293 cells. Over 90% of the long and short *SMN1* and *SMN1E7/SMN2E8* transcripts retained exon 7; in contrast, ~70% of the long and short *SMN2* transcripts excluded exon 7 (Figure 2b), although there was no significant splicing difference between long and short plasmid types ( $P>0.05$ ). Thus, shortening the otherwise long intron 6 had no effect on splicing and most of the intronic 6 sequence is not essential for exon 7 inclusion. In follow-up studies, mutant minigenes were therefore constructed from the short plasmids.

#### The c.863G>T (r.835\_\*3del, p.Gly279Gluufs\*5) variant inhibits inclusion of exon 7 in *SMN1*, *SMN2*, and hybrid *SMN1E7/SMN2E8* genes

To determine whether the c.863G>T (r.835\_\*3del, p.Gly279Gluufs\*5) variant is the sole cause of exon 7 exclusion, we performed *ex vivo* splicing assays using minigenes that mimic the natural context of the *SMN1*, *SMN2*, and hybrid *SMN1E7/SMN2E8* genes (Figure 3). To evaluate the splicing effect of the mutant minigene, we constructed a positive control splicing minigene that includes the classical splicing site c.835-1G>A in the *SMN1*-Short plasmid; this construct yielded 100% exclusion of exon 7. The c.863G>T (r.835\_\*3del, p.Gly279Gluufs\*5) *SMN1* minigene variant reduced *SMN1* exon 7 inclusion from 91.7% to 14.8% ( $t=23.45$ ,  $P=0.000$ , Figure 3a). These data are consistent with the results observed in patient blood and skin fibroblasts. Similar results were observed for the *SMN2* and hybrid *SMN1E7/SMN2E8* minigenes. In the *SMN2* gene, the c.863G>T (r.835\_\*3del, p.Gly279Gluufs\*5) variant disrupted *SMN2* exon 7 inclusion from 33.3% to 0.0% ( $t=7.56$ ,  $P=0.002$ , Figure 3b). In the hybrid *SMN1E7/SMN2E8* gene, this variation caused a significant reduction of exon 7 inclusion from 91.1% to 15.5% ( $t=23.45$ ,  $P=0.000$ , Figure 3c).

#### The splicing effect of various c.863 substitutions

We constructed two additional substitutions at c.863 in *SMN1*, *SMN2*, and the hybrid *SMN1E7/SMN2E8* genes and observed their varying influence on exon 7 splicing. In the *SMN1* and *SMN1E7/SMN2E8* genes (Figure 3), the c.863G>A variant reduced exon 7 inclusion from 91.7% to 81.3% ( $t=3.3$ ,  $P=0.016$ ) and from 91.1% to 80.8% ( $t=3.02$ ,  $P=0.057$ ), respectively. The results for the *SMN2* gene were



**Figure 2** Deletion of 4.2 kb of intron 6 has no significant effect on *SMN* exon 7 inclusion/exclusion. (a) Schematic diagram of the *SMN* (*SMN1*, *SMN2*, and *SMN1E7/SMN2E8*) minigenes. Interconversion of *SMN1* and *SMN2* often occurs to form two hybrid genes (*SMN2E7/SMN1E8* and *SMN1E7/SMN2E8*) in normal individuals. The *SMN2E7/SMN1E8* hybrid gene leads to exon 7 skipping as in the *SMN2* gene, where the exon 7 +6 site is a T. The *SMN1E7/SMN2E8* hybrid gene retains exon 7 and does not produce the SMA phenotype. We constructed c.863 variants in the *SMN1E7/SMN2E8* hybrid minigene to determine whether its splicing was consistent with that of variation in *SMN1*. The 4.2 kb central sequence of intron 6 was deleted from the *SMN*-Short minigene and the five variant nucleotides in *SMN1* and *SMN2* are shown in the *SMN*-Short minigenes. (b) Splicing effects were analyzed by qualitative and semi-quantitative RT-PCR; no significant difference was found between the long- and short-type plasmids ( $P > 0.05$ ).

similar, as the variant decreased exon 7 inclusion from 33.3% to 19.3% ( $t = 3.06$ ,  $P = 0.038$ ). The c.863G>C variant promoted exon 7 inclusion in all three genes: in *SMN1* and *SMN1E7/SMN2E8*, exon 7 inclusion increased from 91% to 95% ( $P > 0.05$ ); in *SMN2*, the increase was from 33.3% to 54.9% ( $t = -4.48$ ,  $P = 0.011$ ).

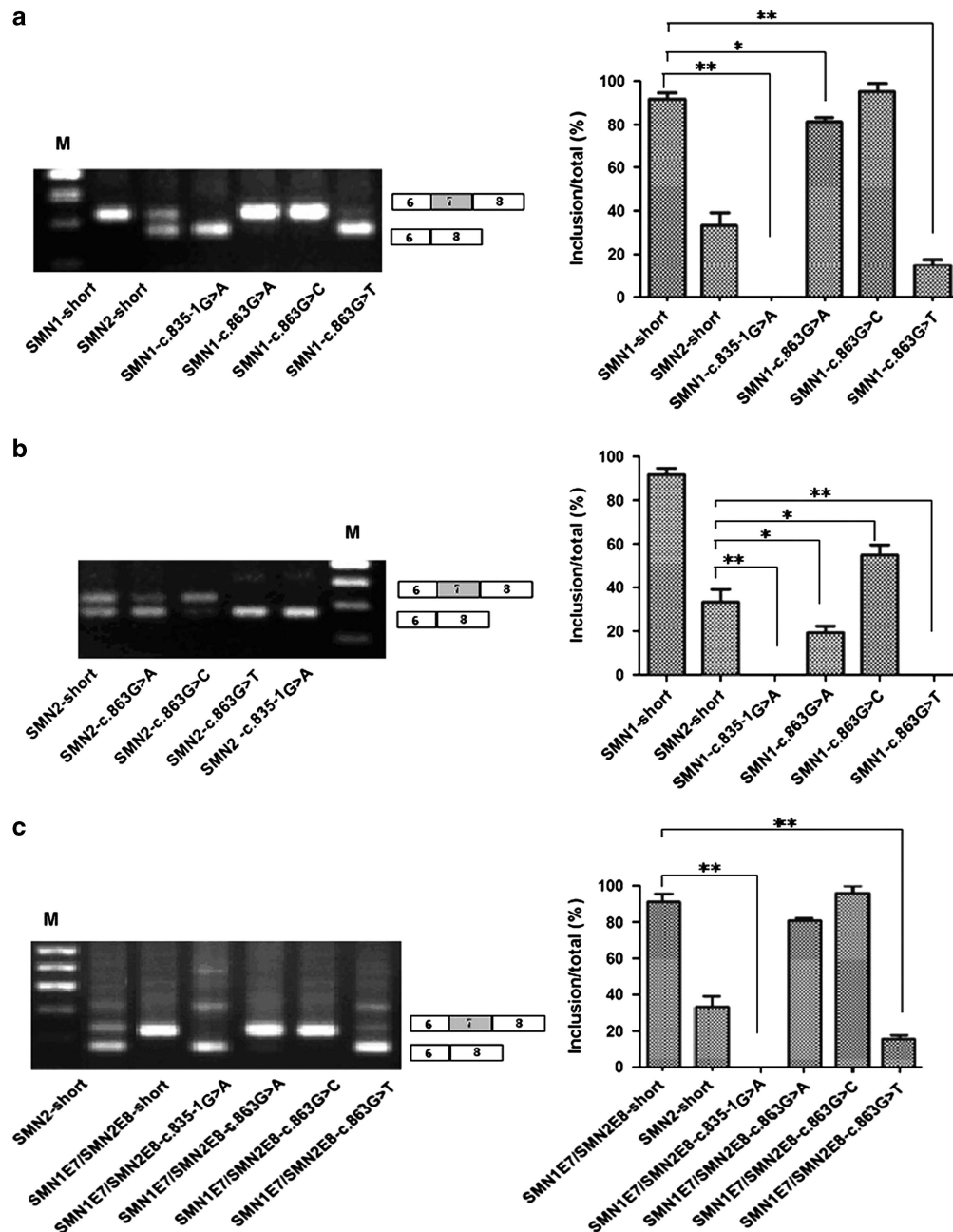
### The c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variation disrupts Tra2β1 binding

Variants at c.863, particularly c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5), may influence splicing of *SMN* exon 7. We then identified the splicing factors that bind the central region of exon 7, which includes the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) site (position +29). Pull-down assays were performed with synthetic RNAs corresponding to positions +19 to 41 of exon 7 (Figure 4a). Biotin-labeled RNAs (including wild-type r.863 g and mutant-type r.863 u) were incubated with HeLa nuclear extract and analyzed by western blotting (Figure 4b) for the key regulators (hnRNP A1, Tra2β1, and SF2/ASF<sup>19–21</sup>) of *SMN* exon 7 splicing. The c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant disrupted binding of Tra2β1, the ESE-binding factor that recognizes the center of exon 7 (+19–+27). The splicing enhancer protein SF2/ASF did not bind to wild-type or mutant RNAs, whereas hnRNP A1 interacted with both, suggesting the presence of a binding element for this regulator within exon 7. The c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant did not disrupt hnRNP

A1-binding, however, indicating that +29 variant site is not the key nucleotide in the hnRNP A1-binding element. Our results suggest the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant promotes exon 7 skipping by disrupting the interaction of splicing enhancer Tra2β1.

### DISCUSSION

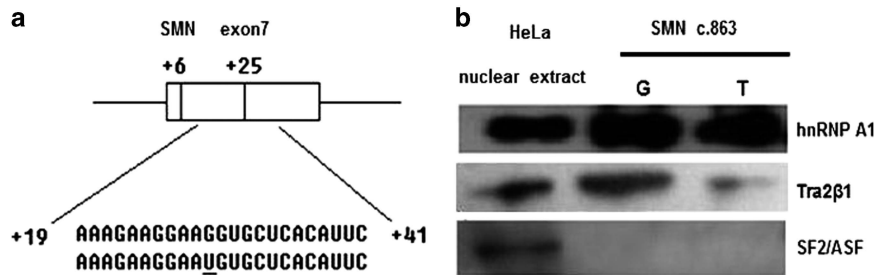
Here, we identified a rare variant (c.863G>T, r.835\_\*3del, p.Gly279Glufs\*5) in center exon 7 of the *SMN1* gene in three patients and this variant abolished *SMN1* exon 7 inclusion both *in vivo* and *ex vivo* assays. Our results demonstrated that this c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant does not affect SMN protein function via amino-acid substitution; instead, it disrupts the pre-mRNA splicing of *SMN1* by interfering with Tra2β1 binding. Sterne et al<sup>22</sup> used bioinformatics and biochemical methods to show that ~25% of missense, nonsense, and silent variants result in aberrant splicing of pre-mRNA. These changes were traditionally considered in the context of amino-acid substitutions rather than pre-mRNA splicing, although we now know such variants are common causes of human genetic disease. Studies of an SMA phenotype modifier, the *SMN2* gene, demonstrated that a single base substitution in exon 7+6 causes abnormal splicing of *SMN2* pre-mRNA, resulting in skipping of this exon and reduced expression of the functional SMN protein.<sup>5</sup> Our results present clear evidence that the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant in the central portion of *SMN1* exon 7



**Figure 3** *Ex vivo* splicing assay of *SMN1*, *SMN2*, and hybrid *SMN1E7/SMN2E8* minigenes. The mutant minigenes were constructed by PCR-based site-directed mutagenesis and transfected into the HEK293 cells. The picture on the left shows the products of RT-PCR on ethidium bromide-stained agarose gel. The quantitative results are shown on the right. Error bars indicate the SD; *P*-values were calculated by Student's *t*-test (\*\**P*<0.01, \**P*<0.05). The c.863T minigenes showed increased exon 7 skipping; the c.863A minigenes showed a slight increase of exon 7 exclusion; the c.863C minigenes showed increased inclusion of exon 7. (a) *SMN1* minigenes; (b) *SMN2* minigenes; (c) hybrid *SMN1E7/SMN2E8* minigenes.

interferes with correct splicing of this exon, thereby causing the absence of SMN protein and resulting in a severe SMA phenotype. Low-level FL-*SMN1* transcript inclusion of exon 7 was confirmed by *in vivo* studies and sequencing in peripheral blood cells and skin fibroblasts from affected patients. SMN protein levels were significantly reduced in skin fibroblasts from patient 1. Our *ex vivo* splicing assay confirmed that the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant promoted exon 7 skipping in SMN pre-mRNA (*SMN1*, *SMN2*, and hybrid *SMN1E7/SMN2E8*), retaining only a small amount of FL transcript. We thus confirmed that the c.863G>T (r.835\_\*3del,

p.Gly279Glufs\*5) variant, which increases exon 7 skipping, causes incorrect splicing of *SMN1* and is not a missense change in the traditional sense. The change effect of c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant in *SMN1* is similar to that in *SMN2* (c.840C>T). Thus, the amount of functional SMN protein is sharply reduced *in vivo* owing to rapid degradation of the truncated SMN protein.<sup>23</sup> Consequently, motor neurons in the spinal cord anterior horn undergo degeneration and necrosis, resulting in a severe SMA phenotype. This has enhanced our understanding of coding region variants that lead to changes in pre-mRNA splicing.



**Figure 4** RNA pull-down assay for splicing factors bound to the central region of *SMN* exon 7. (a) Schematic of the synthetic RNA sequences and the pertinent positions in exon 7 of *SMN*; the underlined base indicates the r.863G>u(c.863G>T) variant. (b) Western blot analysis. Synthetic RNAs were labeled with biotin and immobilized on agarose beads; proteins were detected with antibodies specific for hnRNP A1, Tra2β1, and SF2/ASF.

Over the past decade, extensive studies have been conducted on exon 7 splicing of *SMN1* and *SMN2* and several splicing-related elements and their corresponding splicing factors have been identified.<sup>10,13,19–33</sup> There are overlapping enhancer and silencer sequences within exon 7. For example, Cartegni and Krainer<sup>20</sup> identified an SF2/ASF-binding ESE at position +6 of *SMN1* exon 7. Kashima *et al*<sup>27</sup> found that a C>T substitution at position +6 of *SMN2* exon 7 destroys the SF2/ASF-binding ESE and produces an hnRNP A1-binding ESS. Several studies have shown that a Tra2β1-binding ESE (AAAGAAGGA) is present in the central portion of exon 7 at positions +19–+27.<sup>19,25,34</sup> A study of the c.859G>C variant (at position +25 of exon 7) showed that the central portion of exon 7 (positions +16–+39) also bind the splicing-inhibitor hnRNP A1 and the splicing activator SF2/ASF.<sup>35</sup> Our RNA pull-down assays showed that SF2/ASF does not bind exon 7 at positions +19–+41, consistent with predictions generated in ESEfinder 2.0. The findings of Vezain *et al*<sup>35</sup> led us to suggest that the important nucleotides for SF2/ASF binding may lie within positions +16–+19 of exon 7, not +19–+41. In addition, the central portion of exon 7 is bound by the hnRNP A1 inhibitor but does not require c.863G. The c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant interferes with binding of the splicing activator Tra2β1. A sharp decline in binding of Tra2β1 to the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) site suggests c.863G might be involved in the Tra2β1-ESE sequence, although this finding remains to be verified. Our findings and those of Vezain *et al*<sup>35</sup> suggest the central region of exon 7 binds the Tra2β1 activator and the hnRNP A1 inhibitor. Two recent studies have also shown binding of two splicing factors (hnRNP M and PSF) at sites that partially overlap with the Tra2β1-ESE sequence.<sup>32–33</sup> The evidence suggests sequence overlap between ESE and ESS elements, and between binding sequences for various splicing activators. These components mutually constrain each other to maintain homeostasis, loss of which leads to aberrant splicing. Variants located in or near regulated regions must be considered in the context of splicing effects.

We also showed that exon 7 inclusion in *ex vivo* splicing assays can be improved or disrupted by variants at c.863 site: the highest level of FL transcript was produced by a G to C substitution, followed by the wild-type base G and A. The base T at this position led to a significant reduction in the level of FL transcript. The splicing effects of these four different bases at c.863 in three minigene constructs (*SMN1*, *SMN1E7/SMN2E8*, and *SMN2*) were similar. Multi-species sequence comparisons suggest c.863G is highly conserved; only an A occurs at this site in mice. *In vitro* splicing data showed that the c.863G>T variant led to exon 7 skipping in most *SMN* transcripts. Our patients exhibited severe phenotypes of SMA type I or II and significantly reduced levels of FL-*SMN1* transcript. We conclude that the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant affects function. It is interesting to note that

a G to C substitution at c.863 resulted in an increase of FL-*SMN* transcripts. Our findings and those of Vezain *et al*<sup>34</sup> indicated that the splicing-inhibitor hnRNP A1 can bind the central region of exon 7. The optimal binding sequence for hnRNP A1 is TAGGGA/T. There are two similar sequences (four bases matched the optimal hnRNP A1-binding sequence) in the central region of *SMN* exon 7: GAAGGA (at positions +22–+27) and GAAGGT (at positions +26–+31). The c.859G>C variant (at position +25 of exon 7) in *SMN2*, which was first reported in SMA by Feldkötter *et al*,<sup>6</sup> leads to increased inclusion of *SMN* exon 7 by creating a new binding site for SF2/ASF<sup>36</sup> or interfering with hnRNP A1 bindings.<sup>35</sup> However, increased expression of FL *SMN2* rescues SMA phenotypic severity. We found that c.859G>C variant (+25G to C) reduced the hnRNP A1 sequence match to three bases (GAACGA). Variant of +29 G to C alters the sequence to GAACGT, similar to the G>C substitution at position +25. This also led to increased exon 7 inclusion, so we suggest a substitution of C at +29 might disrupt binding of splicing-inhibitor hnRNP A1.

Here, we report three patient carrying the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant, two with type I and one with type II SMA. Patients 1 and 3 were type I, but Patient 3 had a more severe phenotype. Patient 3 suffered disease onset at late-stage of pregnancy and had a flaccid body at birth, with death occurring at age 20 days due to lung infection and respiratory failure. MLPA results suggested that different modifier gene backgrounds likely cause phenotypic differences between patients. Patient 3 had only two copies of *SMN2*, zero copies of *NAIP*, and one copy each of *P44* and *H4F5*. In Patient 3, therefore, full deletion of *SMN1* occurred on one chromosome in the presence of a larger deletion that included the *NAIP*, *P44*, and *H4F5* genes; the other chromosome carried the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant in *SMN1* and a deletion of *NAIP*. Patient 1 carried three copies of *SMN2* and normal complements of *NAIP*, *P44*, and *H4F*, suggesting conversion of one copy of *SMN1* to *SMN2*. Previous studies showed that expansion of the *SMN1* deletion to *NAIP*, *P44*, and *H4F5* is always associated with a severe phenotype, including early age of onset and early death.<sup>37–39</sup> Conversion of *SMN1* to *SMN2*<sup>40,41</sup> results in increased *SMN2* copy number and is inversely correlated with phenotype severity. Our recent report,<sup>9</sup> revealed that the median age onset is 1 month in SMA patients with *SMN1-SMN2-NAIP* genotype of 0-2-0, whereas in patients with 0-3-2 genotype were 8 months. The median survival time of 0-2-0 patients was only 6 months, whereas the lifespan of patients with 0-3-2 is almost not affected. Based on the c.863G>T functional effect, the genotype of patient 3 is similar to 0-2-0, whereas other two patients are similar to 0-3-2. In addition, the *PLS3* gene at position Xq23 is protective in female SMA patients.<sup>42–45</sup> Our recent research<sup>45</sup> revealed that the *PLS3* gene may have age- and gender-specific role (female protective) in the clinical severity in our SMA patients. So the male

patient 3 exhibited a more serious phenotype than the female patients, a phenotypic difference that may partly be attributable to *PLS3*. Patient 2 had a moderate phenotype with a later age of onset and better motor function than patient 1, but both patients had same *SMN1* genotype and *SMN2-NAIP-P44-H4F5* copy number, and no significant difference in FL-*SMN* transcript levels. We suggest other factors such as profilin Ila<sup>46,47</sup> and zinc finger protein<sup>48</sup> may also influence SMA phenotype.

In summary, our study clearly demonstrates that the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant can lead to exon 7 skipping in a majority of *SMN1* transcripts. Patients with this variant showed more severe SMA phenotype. Our results also indicate that the c.863G>T (r.835\_\*3del, p.Gly279Glufs\*5) variant results in incorrect splicing of exon 7 by interfering with the binding of splicing activator Tra $\beta$ 1, but does not promote binding of the splicing-inhibitor hnRNP A1. Our results suggest overlap of the ESE and ESS signals in the central portion of exon 7. Therefore, variants in this region require comprehensive consideration and integrated analysis.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### ACKNOWLEDGEMENTS

The authors are grateful to the patients and their families for their participation and cooperation in this study. This research was supported by the National Natural Science Foundation of China (Project no. 81100933 and 81470056), the Capital Health Research and Development of Special (Project no. 2011-1008-03), and the Research Foundation of Capital Institute of Pediatrics (Project no. Fangxiang-2014-01).

- Pearn J: Incidence, prevalence, and gene frequency studies of chronic childhood spinal muscular atrophy. *J Med Genet* 1978; **15**: 409–413.
- Pearn J: Classification of spinal muscular atrophies. *Lancet* 1980; **1**: 919–922.
- McAndrew PE, Parsons DW, Simard LR *et al*: Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMN1 and SMN2 gene copy number. *Am J Hum Genet* 1997; **60**: 1411–1422.
- Lefebvre S, Bürglen L, Reboullet S *et al*: Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 1995; **80**: 155–165.
- Lorson CL, Hahnen E, Androphy EJ, Wirth B: A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci USA* 1999; **96**: 6307–6011.
- Feldkötter M, Schwarzer V, Wirth R, Wienker TF, Wirth B: Quantitative analyses of *SMN1* and *SMN2* based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet* 2002; **70**: 358–368.
- Mailman MD, Heinz JW, Papp AC *et al*: Molecular analysis of spinal muscular atrophy and modification of the phenotype by *SMN2*. *Genet Med* 2002; **4**: 20–26.
- Cuscó I, Barceló MJ, Rojas-García R *et al*: *SMN2* copy number predicts acute or chronic spinal muscular atrophy but does not account for intrafamilial variability in siblings. *J Neurol* 2006; **253**: 21–25.
- Qu YJ, Ge XS, Bai JL *et al*: Association of copy numbers of survival motor neuron gene 2 and neuronal apoptosis inhibitory protein gene with the natural history in a Chinese spinal muscular atrophy cohort. *J Child Neurol* 2015; **30**: 429–436.
- Cartegni L, Chew SL, Krainer AR: Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; **3**: 285–298.
- Pagani F, Baralle FE: Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet* 2004; **5**: 389–396.
- Havens MA, Duelli DM, Hastings ML: Targeting RNA splicing for disease therapy. *Wiley Interdiscip Rev RNA* 2013; **4**: 247–266.
- Singh RN: Evolving concepts on human SMN pre-mRNA splicing. *RNA Bio* 2007; **4**: 7–10.
- Cooper TA, Wan L, Dreyfuss G: RNA and disease. *Cell* 2009; **36**: 777–793.
- Qu YJ, Song F, Yang YL, Jin YW, Bai JL: Compound heterozygous mutation in two unrelated cases of Chinese spinal muscular atrophy patients. *Chin Med J (Engl)* 2011; **124**: 385–389.
- Qu YJ, Ge XS, Bai JL *et al*: Subtle mutations in the *SMN1* gene in Chinese patients with SMA: p.Arg288Met mutation causing *SMN1* transcript exclusion of exon7. *BMC Med Genet* 2012; **13**: 86.
- Sun Y, Grimmmler M, Schwarzer V, Schoenen F, Fischer U, Wirth B: Molecular and functional analysis of intragenic *SMN1* mutations in patients with spinal muscular atrophy. *Hum Mutat* 2005; **25**: 64–71.
- Tiziano FD, Pinto AM, Fiori S *et al*: SMN transcript levels in leukocytes of SMA patients determined by absolute real-time PCR. *Eur J Hum Genet* 2010; **18**: 52–58.
- Hofmann Y, Lorson CL, Stamm S, Androphy EJ, Wirth B: Htra2-beta 1 stimulates an exonic splicing enhancer and can restore full-length SMN expression to survival motor neuron 2 (*SMN2*). *Proc Natl Acad Sci USA* 2000; **97**: 9618–9623.
- Cartegni L, Krainer AR: Disruption of an SF2/ASF-dependent exonic splicing enhancer in *SMN2* causes spinal muscular atrophy in the absence of *SMN1*. *Nat Genet* 2002; **30**: 377–384.
- Kashima T, Manley JL: A negative element in *SMN2* exon 7 inhibits splicing in spinal muscular atrophy. *Nat Genet* 2003; **34**: 460–463.
- Sterne WT, Howard J, Mort M, Cooper DN, Sanford JR: Loss of exon identity is a common mechanism of human inherited disease. *Genome Res* 2011; **21**: 1563–1571.
- Lorson CL, Strasswimmer J, Yao JM *et al*: SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nat Genet* 1998; **16**: 63–66.
- Lorson CL, Androphy EJ: An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. *Hum Mol Genet* 2000; **9**: 259–265.
- Hofmann Y, Wirth B: hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1. *Hum Mol Genet* 2002; **11**: 2037–2049.
- Fairbrother WG, Yeh RF, Sharp PA, Burge CB: Predictive identification of exonic splicing enhancers in human genes. *Science* 2002; **297**: 1007–1013.
- Kashima T, Rao N, David CJ, Manley JL: hnRNP A1 functions with specificity in repression of *SMN2* exon 7 splicing. *Hum Mol Genet* 2007; **16**: 3149–3159.
- Cartegni L, Hastings ML, Calarco JA, de Stanchina E, Krainer AR: Determinants of exon 7 splicing in the spinal muscular atrophy genes, *SMN1* and *SMN2*. *Am J Hum Genet* 2006; **78**: 63–77.
- Singh NN, Androphy E J, Singh R N: In vivo selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes. *RNA* 2004; **10**: 1291–1305.
- Singh NN, Androphy E J, Singh R N: An extended inhibitory context causes skipping of exon 7 of *SMN2* in spinal muscular atrophy. *Biochem Biophys Res Commun* 2004; **315**: 381–388.
- Pedrotti S, Bielli P, Paronetto MP *et al*: The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in *SMN2* alternative splicing in spinal muscular atrophy. *EMBO J* 2010; **29**: 1235–1247.
- Cho S, Moon H, Loh TJ *et al*: hnRNP M facilitates exon 7 inclusion of *SMN2* pre-mRNA in spinal muscular atrophy by targeting an enhancer on exon 7. *Biochim Biophys Acta* 2014; **1839**: 306–315.
- Cho S, Moon H, Loh TJ *et al*: PSF contacts exon 7 of *SMN2* pre-mRNA to promote exon 7 inclusion. *Biochim Biophys Acta* 2014; **1839**: 517–525.
- Young PJ, DiDonato CJ, Hu D, Kothary R, Androphy EJ, Lorson CL: SRp30c-dependent stimulation of survival motor neuron (SMN) exon 7 inclusion is facilitated by a direct interaction with hTra2 beta 1. *Hum Mol Genet* 2002; **11**: 577–587.
- Veizan M, Saugier-Verber P, Goïna E *et al*: A rare *SMN2* variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. *Hum Mutat* 2010; **31**: E1110–E1125.
- Prior TW, Krainer AR, Hua Y *et al*: A positive modifier of spinal muscular atrophy in the *SMN2* gene. *Am J Hum Genet* 2009; **85**: 408–413.
- Tsai CH, Jong YJ, Hu CJ *et al*: Molecular analysis of SMN, *NAIP* and *P44* genes of SMA patients and their families. *J Neurol Sci* 2001; **190**: 35–40.
- Amara A, Adala L, Ben Charfeddine I *et al*: Correlation of *SMN2*, *NAIP*, *p44*, *H4F5* and *Occludin* genes copy number with spinal muscular atrophy phenotype in Tunisian patients. *Eur J Paediatr Neurol* 2012; **16**: 167–174.
- He J, Zhang QJ, Lin QF *et al*: Molecular analysis of *SMN1*, *SMN2*, *NAIP*, *GTF2H2*, and *H4F5* genes in 157 Chinese patients with spinal muscular atrophy. *Gene* 2013; **518**: 325–329.
- Hahnen E, Schönling J, Rudnik-Schöneborn S, Zerres K, Wirth B: Hybrid survival motor neuron genes in patients with autosomal recessive spinal muscular atrophy: new insights into molecular mechanisms responsible for the disease. *Am J Hum Genet* 1996; **59**: 1057–1065.
- Van der Steege G, Grootsholten PM, Cobben JM *et al*: Apparent gene conversions involving the SMN gene in the region of the spinal muscular atrophy locus on chromosome 5. *Am J Hum Genet* 1996; **59**: 834–838.
- Oprea GE, Kröber S, McWhorter ML *et al*: Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. *Science* 2008; **320**: 524–527.
- Stratigopoulos G, Lanzano P, Deng L *et al*: Association of plastin 3 expression with disease severity in spinal muscular atrophy only in postpubertal females. *Arch Neurol* 2010; **67**: 1252–1256.
- Bernal S, Also-Rallo E, Martínez-Hernández R *et al*: Plastin 3 expression in discordant spinal muscular atrophy (SMA) siblings. *Neuromuscul Disord* 2011; **21**: 413–419.
- Yanyan C, Yujin Q, Jinli B, Yuwei J, Hong W, Fang S: Correlation of *PLS3* expression with disease severity in children with spinal muscular atrophy. *J Hum Genet* 2014; **59**: 24–27.
- Bowerman M, Shafey D, Kothary R: Smn depletion alters profilin II expression and leads to upregulation of the RhoA/ROCK pathway and defects in neuronal integrity. *J Mol Neurosci* 2007; **32**: 120–131.
- Bowerman M, Anderson CL, Beauvais A, Boyl PP, Witke W, Kothary R: SMN, profilin IIa and plastin 3: a link between the deregulation of actin dynamics and SMA pathogenesis. *Mol Cell Neurosci* 2009; **42**: 66–74.
- hmad S, Wang Y, Shaik GM, Burghes AH, Gangwani L: The zinc finger protein ZPR1 is a potential modifier of spinal muscular atrophy. *Hum Mol Genet* 2012; **21**: 2745–2758.