NOVA1 promotes SMN2 exon 7 splicing by binding the UCAC motif and increases SMN protein expression

Abstract

Spinal muscular atrophy (SMA) is a rare hereditary neuromuscular disease with a high lethality rate in infants. Variants in the homologous genes survival of motor neuron (SMN1) and SMN2 have been reported to be SMA pathogenic factors. Previous studies showed that a high inclusion rate of SMN2 exon 7 increased SMN expression, which in turn reduced the severity of SMA. The inclusion rate of SMN2 exon 7 was higher in neural tissues than in non-neural tissues. Neuro-oncological ventral antigen (NOVA) is a splicing factor that is specifically and highly expressed in neurons. It plays a key role in nervous system development and in the induction of nervous system diseases. However, it remains unclear whether this splicing factor affects SMA. In this study, we analyzed the inclusion level of SMN2 exon 7 in different tissues in a mouse model of SMA (genotype smn1/smn1) and littermate controls (genotype smn1/+). We found that inclusion level of SMN2 exon 7 was high in the brain and spinal cord tissue, and that NOVA1 was also highly expressed in nervous system tissues. In addition, SMN2 exon 7 and NOVA1 were expressed synchronously in the central nervous system. We further investigated the effects of NOVA1 on disease and found that the number of neurons in the anterior horn of spinal cord decreased in the mouse model of SMA during postnatal days 1–7, and that NOVA1 expression levels in motor neurons decreased simultaneously as spinal muscular atrophy developed. We also found that in vitro expression of NOVA1 increased the inclusion of SMN2 exon 7 and expression of the SMN2 protein in the U87MG cell line, whereas the opposite was observed when NOVA1 was knocked down. Finally, point mutation and RNA pull-down showed that the UCAC motif in SMN2 exon 7 plays a critical role in NOVA1 binding and promoting the inclusion of exon 7. Moreover, CA was more essential for the inclusion of exon 7 than the order of Y residues in the motif. Collectively, these findings indicate that NOVA1 interacts with the UCAC motif in exon 7 of SMN2, thereby enhancing inclusion of exon 7 in SMN2, which in turn increases expression of the SMN protein.

Key Words: exon 7 inclusion; motor neuron; neuro-oncological ventral antigen 1; SMN2 splicing; spinal cord; spinal muscular atrophy; splicing factors; UCAC motif

Introduction

Spinal muscular atrophy (SMA) is a rare hereditary neuromuscular disease that is primarily characterized by degeneration of alpha motor neurons in the anterior horn of the spinal cord and affects innervation of skeletal muscles, leading to muscle weakness, muscle atrophy, paralysis, or even death (Arnold and Fischbeck, 2018). SMA is caused by deletion of the survival of motor neuron 1 (SMN1) gene on chromosome 5 or by point mutations in SMN2 leading to a lack of expression of functional full-length SMN (SMN FL) (Brzustowicz et al., 1990; Lefebvre et al., 1995). SMN is a housekeeping protein that is part of the SMN-Gemin multiprotein complex (Borg et al., 2015). It is involved in assembly of small nuclear ribonucleoproteins and regulation of precursor mRNA (pre-mRNA) splicing, and also participates in a variety of physiological processes, including stress response, axon transport, cytoskeletal dynamics, mitochondrial and bioenergy pathways, and ubiquitin pathways (Chaytow et al., 2018). Thus, SMN is an important molecule that is involved in multiple activities that are essential for human life.

In addition to SMN1, humans have a second, homologous, gene called SMN2. The key difference between them is that the sixth nucleotide in exon 7 of the SMN1 gene is C, while SMN2 contains a T at the same location. Only 10% of the transcripts produced from SMN2 are full-length and generate a stable SMN protein, while the rest are missing exon 7 (SMN2Δ7) and generate a truncated protein that has no biological function and is extremely unstable (Anderton et al., 2013). Hence, variation in the sixth nucleotide is the key factor that influences the inclusion of exon 7 (Lefebvre et al., 1995). Clinical studies have shown that the severity of SMA is negatively correlated with the amount of functional SMN protein, as well as SMN2 copy number (Lefebvre et al., 1997; Wirth, 2000). The amount of SMN FL protein generated by SMN2 is not sufficient to compensate for defects in SMN1; therefore, promoting the...
inclusion of SMN2 exon 7 can effectively ameliorate the severity of SMA and improve patient survival (Hua et al., 2011; Wu et al., 2017; Son and Yokota, 2018). SMN2 transcripts are thus increasingly important for the treatment of SMA.

Alternative splicing is an important mechanism that regulates protein diversity and is ubiquitous in the occurrence and development of many diseases. Two types of splicing factors are involved in the regulation of pre-mRNA splicing. A typical type I serine-rich (SR) protein that promotes pre-mRNA splicing contains an RNA recognition motif and an RS region which encodes a series of serine and arginine (Long and Caceres, 2000; Ankö, 2014). Type II splicing factors, such as the heterogeneous nuclear ribonucleoproteins (hnRNPs) proteins, which contain KH domains, inhibit pre-RNA splicing (Martínez-Contreras et al., 2007). Previous studies have shown that the NOVA family, neuro-oncological against (NOVA), which is specifically expressed in neurons, promotes or inhibits splicing depending on which site the protein binds to on the pre-mRNA (Ule et al., 2006; Leggere et al., 2007). The NOVA family proteins play important roles in the regulation of SMN2 splicing (Hua et al., 2008; Wee et al., 2014). For example, Hua et al. (2010) showed that blocking HNRNP A1/A2 binding to the intron splicing silencer located in SMN2 intron 7 significantly increased the inclusion of SMN2 exon 7 and up-regulated the expression of SMN FL. Moreover, NOVA plays a key role in nervous system development, as well as the development of neurological diseases (Buckanovich et al., 1993). NOVA1 regulates alternative splicing by binding the YCAY motif in the target gene mRNA precursor via its KH domain (Hsu et al., 1994). The presence of NOVA1 was found to eliminate the regulatory effect of NOVA1 on alternative splicing (Ludow et al., 2018), and the UCAC sequence was found in exon 7 of SMN2. However, the relationship between the NOVA family and SMA remains unclear. Here, we scanned for splicing factors that regulate SMN2 exon 7 inclusion in SMA mice, which harbor two copies of SMN2 (Hsieh-Li et al., 2000), and investigated the mechanism by which they promote SMN2 exon 7 inclusion.

Materials and Methods

Animals

SMN2 mice (genotype smn−/−; SMN2tm1tm1) (RRID: SCR_005570) and littermate control mice (genotype smn−/−; SMN2tm1tm1) on postnatal day 1 (P1), P4, and P7 were used in this study. The background strains were FVB inbred mice. The parental mice were donated by the Hua lab at the Institute of Neuroscience of Chinese Academy of Sciences (Shanghai, China). The spinal cords from P1, P4, and P7 mice were collected for RNA or protein extraction. For the mutation experiments, 1 μg of pCI-SMN1 and pCI-SMN2, pEGFP-SMN1, and pEGFP-SMN2 according to a previous method described by the Hua lab (Hua et al., 2010), was used for all transfections. To knock down NOVA1-sense: 5′-AGA UCA UUG UUC AGU ACG A-3′, antisense: 5′-UCC AAG UCA ACA AUA CAC-3′. The negative control sequences were: NOVA1-sense: 5′-UUC UCC GGA GUC AGC UTT-3′, antisense: 5′-AGA UCA GCU CAG ATT-3′. siRNAs were designed using online software (http://biodevextra.cea.fr/DSIR/DSIR.html), synthesized by Shanghai GenePharma, and transfected using Lipofectamine 2000 (Life Technologies).

Point mutation

The PCR reaction mixture used to construct the SMN2 mutation mini-genes was similar to that used to construct the NOVA1 overexpression plasmid. The reaction conditions were 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 210 seconds, for 30 cycles total. The annealing temperature was adjusted according to the primer 1m value or determined by gradient PCR. The mutation primers are listed in Additional Table 1. The PCR products were treated with Dpn I to remove dsDNA and then used to detect the products after PCR amplification (usually 25 μL of the reaction mixture). One microliter of DpnI was added to the remaining PCR product, which was then allowed to digest in a 37°C water bath for 4 hours. After digestion, 8 μl of the product was transformed, and a single colony was selected for the plasmid was extracted and sequenced to verify whether the mutation was successful.

Reverse transcription-polymerase chain reaction

Total RNA was extracted from spinal cord tissues at P4 or from U87MG and HEK-293T cells using TRIzol reagent according to the manufacturer’s instructions (Vazyme, Nanjing, China; Cat# R4001-01). In brief, 1 μg of each RNA sample was used as 20 μl reaction for first-strand cDNA synthesis with oligo(dT)18 and M-MLV reverse transcriptase. PCR was carried out in a 25-μl reaction mixture. The PCR products were separated on 2% agarose gels, followed by signal quantitation using ImageJ software (National Institutes of Health, Bethesda, MD, USA). Exon 7 inclusion was calculated using the following formula: exon 7 inclusion ratio (%) = [FL (exon 7 inclusion)] / (FL + Δ7 exon 7 exclusion)] × 100 (Hua et al., 2007).

Quantitative polymerase chain reaction

Total RNA was extracted from SMA mice or from U87MG cells. The 10-μL qPCR reaction mixture contained 5 μL 2x ChamQ Universal SYBR qPCR Master Mix (Cat# Q711, Vazyme), 0.5 μL forward primer, 0.5 μL reverse primer, 1 μL cDNA, and 3 μL ddH2O. The quantitative polymerase chain reaction (qPCR) conditions were 45 cycles of 95°C for 10 seconds, 60°C for 20 seconds, and 72°C for 20 seconds. The sequences of the primers used for qPCR are shown in Additional Table 2, and the mRNA levels are shown as heat maps.

Western blot assay

Protein samples from the spinal cord at P4 or from U87MG cells were separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and electroblotted onto polyvinylidene fluoride membranes (Millipore, FFP39), followed by blocking with 5% skim milk (Cat# A60669, Sangong, Shanghai, China). Then the membranes were incubated with primary antibodies overnight at 4°C and secondary antibodies for 2 hours at room temperature. The primary antibodies used were as follows: mouse anti-SMN1 antibody (1:500, BD Biosciences, Franklin Lakes, NJ, USA; Cat# 610644, RRID: AB_397973), rabbit anti-NOVA1 antibody (Cat AP-PCR-13042, Sangong, Shanghai, China), mouse anti-AKT antibody (1:1000, Santa Cruz Biotechnology, Santa Cruz, CA, USA; Cat sc-72778), secondary rabbit anti-mouse (1:1000, Cat# D110087) and goat anti-rabbit antibodies (1:1000, Cat# D110088). The membranes were subsequently washed with TBST and then incubated with the Tanon-5200Multi Gel Imaging System (Tanon Science & Technology, Shanghai, China). The scanned images were imported into ImageJ software (Schneider et al., 2012). The signals were normalized to actin.

Nissl staining

Nissl staining was carried out with a kit according to the manufacturer’s instructions (Beyotime, Shanghai China, Cat# CO117). The spinal cord tissue sections were dehydrated at P1, P4, and P7, and clarified with 95% formaldehyde at 4°C overnight, washed in 0.01 M phosphate-buffered saline (PBS), and embedded in paraffin blocks. Then, 4-μm-thick paraffin sections were cut for Nissl staining. In brief, the paraffin sections were dehydrated, rehydrated, and then rehydrated with water, dehydrated with 95% ethanol for 2 minutes, 70% ethanol for 3 minutes, and left to air-dry. Subsequently, tissue sections were placed in Nissl staining solution for 3–10 minutes at room temperature, washed with pure water twice for a few seconds, dehydrated twice with 95% ethanol for 2 minutes, cleared twice with xylene for 5 minutes, sealed with neutral gum, and observed under a microscope (OLYMPUS BX51, Tokyo, Japan).

Research Article
Immunofluorescence

Spinal cord tissue from control and SMA mice at P1, P4, and P7 was fixed with 4% formaldehyde at 4°C overnight, dehydrated with 30% sugar solution, embedded in tissue freezing medium (O.C.T.), frozen, then cut into 6-μm-thick sections using a freezing microtome, rehydrated, and stained for immunofluorescence. In brief, the frozen sections were placed in an oven at 60°C, dried for 30 minutes, washed twice with PBS for 2 minutes, and were incubated with 0.3% Triton X-100 permeabilization solution at room temperature for 30 minutes. Subsequently, these sections were washed with 0.1 M PBS and blocked with 5% bovine serum albumin (BSA) blocking solution for 30 minutes at room temperature. After the blocking solution was discarded, the sections were incubated with the primary antibodies (anti-Nova1/2, Novus Biologicals, Cat# NBP1-89785; anti-ChAT, Abcam, Cat# ab190702) and chicken anti-cholesterol acyltransferase (Chat; 1:50; Abcam, Cat# ab34419) diluted in 0.1 M PBS for 2 hours to the tissue sections. Then, the tissue sections were incubated in a humid box at 4°C overnight. The next day, the tissue sections were washed with 0.1 M PBS for 5 minutes. After clowwise addition of the secondary antibodies Cy3-goat anti-rabbit (IgG (1:1000; Sangon, Shanghai, China, Cat# D110062) and fluorescein isothiocyanate-donkey anti-chicken IgY (1:1000; Sangon, Cat# D110201) diluted in 0.1 M PBS, as well as DAPI (1:2000; Sangong, Cat# E607303), the tissue sections were incubated for 2 hours in a humid box at room temperature in the dark. Then, tissue sections were washed three times with 0.1 M PBS for 5 minutes, then NOVA1 and CHAT co-localization was observed using a fluorescent mounting tablet (OLYMPUS BX51).

RNA pull-down

The biotin-labeled RNA probes WT (5'-AGG UGC UCA CAU UCC U-3') and Mut (5'-AGG UGC UCA CAU UCC U-3') were purchased from Shanghai GenePharma for use in the RNA pull-down experiment. After washing streptavidin agarose beads (Sigma, Cat# S1638-1ML) with streptavidin washing buffer three times, biotin-labeled RNA was added to the precleaved beads, and the mixture was incubated at 4°C for 1–4 hours on a 100 × g rotator. Then, the beads were precipitated and washed using RNA-streptavidin interaction buffer. Next, 100 μL of nucleoprotein extracted from U87MG cells was mixed with the beads at a protein concentration of 3–5 μg/μL. The tubes were maintained at 4°C overnight on a 100 × g rotator. After centrifugation, the supernatant was discarded. An appropriate volume of protein loading buffer was added, and the mixture was then placed in a water bath at 90°C for 10 minutes. After cooling on ice, the samples were centrifuged at 10,000 × g at 4°C for 10 minutes, and the supernatant was removed for western blot analysis.

Statistical analysis

No statistical methods were used to predetermine sample sizes; however, our sample sizes were similar to those reported previously (Mao et al., 2021). The software program SPSS 16.0 (SPSS, Chicago, IL, USA) was used for statistical analysis. Data are presented as mean ± standard deviation (SD). The statistical significance of the differences between groups was analyzed using the Student’s t-test. P < 0.05 indicates that the difference was statistically significant.

Results

The inclusion of SMN2 exon 7 is high in neural tissues

To investigate the differences in the inclusion of SMN2 exon 7 in different tissues, we evaluated eight tissues—heart, liver, spleen, lung, kidney, brain, spinal cord, and hindlimb muscle in control and SMA mice at P4 (Figure 1A and B). The results showed that SMN2 splicing resulting in inclusion of SMN2 exon 7 was different between tissues. A high rate of exon 7 inclusion was observed in the brain and spinal cord, and a low rate in muscle in control and SMA mice (P < 0.001) and the mixture was then placed in a water bath at 90°C for 10 minutes. After cooling on ice, the samples were centrifuged at 10,000 × g at 4°C for 10 minutes, and the supernatant was removed for western blot analysis.

Expression of splicing factors in control and SMA mice

Exon inclusion or exclusion is determined by splicing events that are regulated by splicing factors. Classic splicing factors, such as HNRNP (Martinez-Contreras et al., 2007), SR (Long and Caceres, 2009), and NOVA (Ule et al., 2003, 2006) family members, have been confirmed to regulate a variety of splicing events. To clarify the relationship between the expression of splicing factors in SMA mice and the differential splicing of SMN2 among different tissues, we studied the expression of various splicing factors in neural and non-neural tissues from SMA or control mice at postnatal day 4 using qPCR and heat maps. The results showed that the expression of 20 members of the HNRNP, SR, and NOVA families differed among tissues, while Nova was highly expressed in nerve tissues in both control mice (Figure 2A) and SMA mice (Figure 2B).

Nov2 is down-regulated with SMA progression

To investigate the relationship between the neuro-specific expression of NOVA family members and the progression of SMA disease, SMN and NOVA1 mRNA and protein levels were detected in the spinal cord of SMA mice at P1, P4, and P7. In control mice, SMN FL mRNA (P1 vs. P7, P < 0.01, P4 vs. P7, P < 0.05) and protein (P1 vs. P7, P < 0.05, P4 vs. P7, P < 0.05) levels were significantly reduced from P1 to P7 (Figure 2A–C). Meanwhile, NOVA1 mRNA expression levels (P1 vs. P4, P < 0.05, P1 vs. P7, P < 0.01) were significantly increased on P4 and P7 compared with those on P1, and NOVA1 protein levels (P1 vs. P4, P < 0.05, P1 vs. P7, P < 0.05) were significantly higher on P4 and P7 than P1 (Figure 3D–F). In SMA mice, SMN FL mRNA (P1 vs. P7, P < 0.01, P4 vs. P7, P < 0.05) and protein (P1 vs. P4, P < 0.05, P1 vs. P7, P < 0.05) levels decreased significantly from P1 to P7 (Figure 3G–I). Meanwhile, NOVA1 mRNA levels were significantly decreased on P4 and P7 compared with those on P1 (P1 vs. P4, P < 0.01, P1 vs. P7, P < 0.05), and NOVA1 protein levels were significantly lower on P7 than on P1 and P4 (P1 vs. P4, P < 0.01, P1 vs. P7, P < 0.001, P4 vs. P7, P < 0.05; Figure 3J–L).
wild-type and mutated mini-genes were then transferred into U87MG cells. Bundles of hollow Nissl-positive plaques were found in the SMA anterior horn regions, unlike in control mice. We also found that the motor neuron marker ChAT partially co-localized with NOVA1 in the spinal anterior horn at the time points studied (Figure 4B). Taken together, NOVA1 expression was correlated with splicing of the SMA-related gene SMN2, SMN protein expression, and several motor neurons in the spinal cord anterior horn in SMA mice, indicating that NOVA1 may be involved in the disease process in SMA mice.

NOVA1 regulates the inclusion of SMN2 exon 7 in vitro
To investigate the mechanism by which NOVA1 regulates SMN2 splicing, that is, inclusion of SMN2 exon 7, we transfected siRNA or over expressed plasmid of NOVA1 into U87MG cells. NOVA1 mRNA (P < 0.01; Figure 5A) and protein (P < 0.05; Figure 5B and C) expression levels were significantly decreased after siRNA-mediated knock down. The inclusion of SMN2 exon 7 was significantly reduced when NOVA1 expression was knocked down (P < 0.05; Figure 5D and E). Meanwhile, the SMN protein level was also decreased (P < 0.01; Figure 5F and G). In contrast, NOVA1 mRNA (P < 0.001; Figure 5H) and protein (P < 0.01; Figure 5I and J) expression levels were up-regulated after plasmid transfection. NOVA1 expression was correlated with the rate of SMN2 exon 7 inclusion (P < 0.05; Figure 5K and L), and SMN protein expression increased (P < 0.05; Figure 5M and N). NOVA1 expression did not change after SMN knockdown (Additional Figure 2). In short, NOVA1 significantly promoted the inclusion of SMN2 exon 7 and increased the expression of SMN protein.

NOVA1 binds to the UCAC motif in exon 7 of SMN2
Sequence analysis identified the presence of a single UCAC motif, which is the consensus sequence for the YCAY motif, in exon 7 of SMN2 (34 to 37 bp) (Figure 6A). To confirm that NOVA1 targets the UCAC sequence to regulate inclusion of SMN2 exon 7, we constructed an SMN2 mini-gene in which the UCAC motif was deleted and a version in which it was mutated to UAAC and investigated the effect of these changes on SMN2 splicing (Additional Figure 3). The SMN2 wild-type and mutated mini-genes were then transferred into HEK293T cells. The results showed that deletion of the UCAC sequence significantly reduced SMN2 exon 7 inclusion (P < 0.05), and that mutation of the UCAC motif to UAAC reduced its inclusion even more significantly (P < 0.001; Figure 6B and C).

To confirm that NOVA1 targeted binding the UCAC sequence to regulate SMN2 splicing, RNA pull-down experiments were performed. NOVA1 was detected bound to SMN2 WT (UCAC) mRNA, but not to SMN2 MUT (UAAC) mRNA (Figure 6D), indicating that the UCAC sequence (34–37 bp) in exon 7 of SMN2 is the binding site for NOVA1.

Next, we tested the effect of mutation in the core “CA” sequence to the following sequences: UGAC, UCAC, UACG, UGCG, and UCAC (Figure 6E). We found that mutation of the UCAC sequence to UCAU (P < 0.001), CCAC (P < 0.05), UGAC (P < 0.01), UCAC (P < 0.001), UGCG (P < 0.001), or UUGC (P < 0.01) reduced the rate of SMN2 exon 7 inclusion (Figure 6F–H). However, mutation to CCAC (P < 0.01) or UGCG (P < 0.01) significantly up-regulated inclusion of exon 7 (Figure 6F–H), indicating that the “CA” sequence has a significant impact on inclusion.

Taken together, these findings show that the UCAC sequence is the key sequence required for NOVA1 to specifically bind to SMN2 and regulate splicing. Therefore, more functional, stable SMN2 FL is expressed in neural tissues, which express high levels of NOVA1, than in tissues that express less NOVA1 (Figure 7A). In contrast, in non-neural tissues that express lower levels of NOVA1, more nonfunctional SMN2/7 mRNA is produced, leading to the production of nonfunctional and unstable SMN protein that is quickly degraded (Figure 7B).
NOVA1 is expressed at high levels in neural tissue and regulates SMN2 without exon 7. After knocking down NOVA1, the expression of SMN2 exon 7 was assessed by qPCR and western blot. Relative mRNA expression was normalized to the siNC group. The efficiency of NOVA1 knockdown in U87MG cells was determined by qPCR and western blot. Western blot detection of NOVA1 protein after RNA pull-down. Schematic diagram of UCAC mutation cassette. Determination of inclusion rate of exon 7 in the context of UCAC deletion or UAAC mutation. Inclusion rate of exon 7 of SMN2 in U87MG cells. The pathological changes that occur during this period indicates that gene polymorphism may play a significant role in regulating SMA from the perspective of splicing. SMA with symptoms similar to human SMA was established by Hung Li and coworkers. These mice carry two copies of the SMN2 gene (genotype smn<sup>−/−</sup>, SMN2<sup>2tg/0</sup>) and develop typical symptoms of SMA 4 days after birth. This indicates that gene polymorphism may play a significant role in regulating the pathological changes that occur during this period. SMA, which is caused by a deficit in functional SMN protein synthesis due to SMN1 mutation, is a rare disease that seriously compromises quality of life. Comparison of the splicing of SMN1 and its homologue SMN2 has shown that promoting the inclusion of SMN2 exon 7 can effectively reduce the severity of SMA. NOVA1 regulates SMN2 without exon 7, resulting in little NOVA1 binding to the UCAC motif and the production of much SMN2 Δ7 mRNA, which was translated into unstable, nonfunctional SMN protein that was degraded quickly. NOVA1: Neuro-oncological ventral antigen 1; SMN2: survival of motor neuron 2; U87MG: a human glioma cell line.

Figure 5 | NOVA1 regulates SMN2 exon 7 inclusion in U87MG cells.

(A–C) The efficiency of NOVA1 knockdown in U87MG cells was determined by qPCR (A) and western blot (B, C). Relative mRNA expression was normalized to the siNC group. (D, E) Inclusion rate of exon 7 of SMN2 after knocking down NOVA1. (F, G) SMN protein expression level after NOVA1 knockdown. (H–J) The level of NOVA1 overexpression in U87MG cells was assessed by qPCR (H) and western blot (I, J). (K, L) Inclusion rate of exon 7 after NOVA1 overexpression. (M, N) SMN protein expression levels after NOVA1 overexpression. Data are expressed as mean ± SD (n = 3 independent assays). *P < 0.05, **P < 0.01, ***P < 0.001 (Student’s t-test). NOVA1: Neuro-oncological ventral antigen 1; qPCR: quantitative polymerase chain reaction; SMN2: survival of motor neuron 2; U87MG: a human glioma cell line.

Figure 6 | The UCAC motif is the NOVA1 binding site and plays an important role in SMN2 splicing.

(A) Schematic diagram of the SMN2 mini-gene with the NOVA1 binding site located in exon 7. (B, C) Inclusion of SMN2 exon 7 in the context of UCAC deletion or UAAC mutation. (D) Western blot detection of NOVA1 protein after RNA pull-down. (E) Schematic diagram of UCAC mutation cassette. (F) Detection of SMN2 splicing after mutation of the CA part of the UCAC sequence. (H) Quantitation of the data shown in F and G. Data are expressed as mean ± SD (n = 3 independent assays). *P < 0.05, **P < 0.01, ***P < 0.001, vs. UCAC (Student’s t-test). FL: Full-length; NOVA1: neuro-oncological ventral antigen 1; SMN2: survival of motor neuron 2; Δ7: without exon 7.

Figure 7 | NOVA1 is expressed at high levels in neural tissue and regulates SMN2 splicing by binding the UCAC motif to promote SMN2 FL mRNA and protein expression.

(A) NOVA1, an RNA-binding protein, is highly expressed in neural tissue, where it promotes the inclusion of SMN2 exon 7 by binding the UCAC motif, resulting in the production of high levels of SMN2 FL in the nucleus that are then translated into stable, functional SMN protein in the cytoplasm by the ribosomes. (B) In contrast, NOVA1 was expressed at low levels in non-neural tissue, resulting in little NOVA1 binding to the UCAC motif and the production of much SMN2 Δ7 mRNA, which was translated into unstable, nonfunctional SMN protein that was degraded quickly. NOVA1: Neuro-oncological ventral antigen 1; SMN2: survival of motor neuron 2; SMN2 FL: SMN 2 full-length; SMN2 Δ7- SMN 2 without exon 7.
However, different pathological changes take place in different tissues and organs of SMA mice. The tissue specificity of SMN2 splicing was revealed by a previous study of a mouse model of type III SMA (Chen et al., 2015), suggesting the importance of SMN2 alternative splicing in SMA mice.

In this study, we detected the inclusion of SMN2 exon 7 in eight different tissues from SMA mice at P4. Interestingly, we found that SMN2 exon 7 is present in different tissues, and the inclusion pattern varies in different tissues of SMA and control mice. On the basis of the hypothesis that different expression levels of certain splicing factors in neural and non-neural tissues were responsible for the variable inclusion rates, we analyzed the expression of three types of classic splicing factors, HNRNP SR, and the NOVA family, in different tissues. Enhanced expression of NOVA1 was detected in neural tissues in both control and SMA mice (Graus et al., 1993). NOVA2 is localized in the neocortex and hippocampus (Yang et al., 1998). In this study, we found that NOVA1 expression is inversely correlated with the inclusion of SMN2 exon 7 and is significantly down-regulated as the disease progresses in SMA mice. Immunofluorescence experiments also confirmed that NOVA1 was expressed in the anterior horn motor neurons of the spinal cord of SMA mice, whereas it was undetectable in their inclusion of SMN2 exon 7. The NOVA-deficient mice created by Jensen and others (Jensen et al., 2000; Ruggiu et al., 2009) using gene-editing technology exhibit motor dysfunction and die within 10 days of birth because of motor neuron apoptosis. The symptoms observed in these mice were similar to those seen in our SMA mice, suggesting a causal association between NOVA expression and SMA severity. Here, we showed for the first time that NOVA1 is highly expressed in neural tissues and enhances SMN translation by promoting the inclusion of SMN2 exon 7.

Decreased SMN synthesis and loss of SMN function dominate the development of SMA. As a ubiquitously expressed protein, SMN protein is involved in various biological processes, including pre-mRNA splicing, apoptosis, axon transport, cytoskeleton dynamics, mitochondria, and bioenergy pathways, as well as other activities (Schrank et al., 1997; Paushkin et al., 2002; Fallini et al., 2012). The current study showed that SMN protein expression was notably down-regulated in parallel with SMA progression, which is consistent with the results reported by Groen et al. (2018). These findings confirm that temporal and spatial expression of SMN affects the severity of SMA. However, we also found that, while SMN is translated from SMN2 FL mRNA, its expression level was dramatically decreased with increased inclusion of exon 7, which may not be able to produce normal levels of SMN. Therefore, we established a positive correlation between SMN2 exon 7 inclusion and SMN protein expression. We found that both NOVA1 and SMN were simultaneously down-regulated in SMA mice. Similarly, NOVA1 and SMN were both down-regulated in the NOVA1-deficient mice. Protein expression in vitro. Related studies have shown that SMN can interact with splicing factors, such as SMN and HNRNP in axons, axon terminals, and the cytoplasm of mouse motor neurons (Dombert et al., 2014). Thus, it was speculated that NOVA1 may directly bind to SMN to regulate its expression, thereby affecting SMA severity; this putative regulation mechanism needs further study (Tadesse et al., 2008; Singh et al., 2017). Interestingly, separate mutation of the outer pyrimidine residues or the core “CA” sequence results in a reduced reduction in the rate of SMN2 exon 7 inclusion, and mutation of the core sequence had the greatest impact. Therefore, the “CA” sequence in the YCAY motif is critical for the interaction with NOVA1. In addition, mutation of UCAC to CCAC or UCGC greatly promoted the inclusion of exon 7. The NOVA-dependent mutations may have created new RNA-binding sites. Further research is needed to verify this hypothesis.

Potential treatments for SMA that are currently being investigated include soybean butyrate (Chang et al., 2001) and induced pluripotent stem cells (Corti et al., 2012). Although these methods are somewhat promising, their use is limited by minimal efficacy or by safety concerns. In recent years, gene therapy and antisense oligonucleotides have become popular products for the treatment of SMA. Spinraza is the first antisense oligonucleotide drug developed by Hua et al. (2011), to treat SMA in children and adults (Hoy, 2017). In addition, Zolgensma, the first gene therapy drug used to treat SMA in children younger than 2 years, was also approved by the Food and Drug Administration (FDA) for clinical treatment in May 2019 (Hoy, 2019). However, multiple systems have been reported to be affected in SMA (Farrar et al., 2017), and, while these drugs are effective in one or more systems, their limitations and side effects cannot be ignored. The small nuclear RNA splicing, relatively high efficiency and a short half-life in the circulation. In contrast, overexpression vectors can be persistently expressed, have larger molecular weights, and potentially risk causing mutations in host genes if they cross the blood-brain barrier. At the same time, they cannot completely and effectively cure SMA, creating a greater economic burden on patients. Thus, the development of more effective SMA treatments is urgently needed.

In conclusion, SMA development correlates with down-regulation of NOVA1 expression, which exhibits the same inter-tissue differences as SMN2 splicing, and NOVA1 binds to the UCAC sequence located at 3A–3B of SMN2 exon 7 to regulate SMN2 splicing. These findings provide new perspectives on correcting SMN gene splicing and treating SMA.

Acknowledgments: We thank Professor Yi-Min Hua from Institute of Neuroscience in Soochow University for providing SMA mice and guidance in constructing SMN2 mini-gene.

Author contributions: Study design and manuscript revision: LCW, JJS; experiment implementation: LLD, JJS, LCW, ZHC; data analysis: LCW, JJS; experiment design and manuscript draft: LLD, LCW. All authors read and approved the final manuscript.

Conflicts of interest: We hereby declare that there is no conflict of interest between the authors of this manuscript.

Author statement: This paper has been posted as a preprint on Research Square with doi: https://doi.org/10.21203/rs.3.rs-798194/v1, which is available from: https://assets.researchsquare.com/files/rs-798194/v1/4f3e7178-3170-4df5-8894-2e46bf019299.pdf?c=1637245333.

Availability of data and materials: All data generated or analyzed during this study are included in this published article and its supplementary information files.

Open access statement: This is an open access journal, and articles are distributed under the terms of the Creative Commons AttributionNonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Additional files: Additional Table 2: Primer sequence used to construct the SMN2 mini-gene with UCAC sequence deletion or mutation on SMN2 exon 7. Additional Table 2: Primer sequence used in this experiment. Additional Figure 1: PGCT7-NOVA1 sequence by forward CMV promoter and reverse primer. Additional Figure 2: Western blot detection of NOVA1 expressions after SMN knockdown. Additional Figure 3: Diagrams of sequencing results before and after UCAC sequence deletion or mutation.

References
Anderton RS, Meloni BP, Mastaglia FL, Boulos S (2013) Spinal muscular atrophy and the antiapoptotic role of survival of motor neuron (SMN) protein. Mol Neurobiol 47:821–832.
Ankó ML (2014) Regulation of gene expression programmes by serine-arginine rich splicing factors. Semin Cell Dev Biol 32:11–21.
Aoyagi Y, Kobayashi H, Tanaka K, Ozawa T, Nitta H, Tsuji S (1999) A de novo splice donor site mutation causes in-frame deletion of 14 amino acids in the proteolipid protein in Pelzuesa-Mezrbacher disease. Ann Neurol 46:112–115.
Arnold ES, Fischbeck KH (2018) Spinal muscular atrophy. Hand Clin Neuro 148:591–601.
Borg RM, Bordonne R, Vassallo N, Cauchi RJ (2015) Genetic interactions between the members of the SMN-gemin complex in drosophila. PLoS One 10:e0130974.
Brzustowicz LM, Lehrer T, Castilla LH, Penchaszadeh GK, Wilhelmsen KC, Daniels R, Davies KE, Leppert M, Ziter F, Wood D, Dubovitz V, Zerres K, Hausmanowa-Petrusewicz I, Ott J, Munsat TL, Gilliam TC (1990) Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome q91:1–2:3:3. Nature 344:540–541.
Buckanovich RJ, Posner JB, Darnell RB (1993) Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. Neuro 11:657–672.
Chang JG, Hsieh-Li HM, Jing JY, Wang NM, Tsai CH, Li H (2001) Treatment of spinal muscular atrophy by sodium butyrate. Proc Natl Acad Sci U S A 98:9808–9813.
Chaytov H, Huang YT, Gillingwater TH, Fuller KME (2018) The role of survival motor neuron protein (SMN) protein in homeostasis. Cell Mol Life Sci 75:3877–3894.
Chen YC, Chang JG, Jing JY, Liu TY, Yoo CY (2015) High expression level of Tra2-β1 is responsible for increased SMN2 exon 7 inclusion in the testis of SMA mice. PLoS One 10:e0120721.
Corti S, Nizzardo M, Simone C, Falcone M, Nardini M, Ronchi D, Donadoni C, Salani S, Riboldi G, Magri F, Menozzi G, Bonaglia C, Rizzo F, Bresolin N, Comi GP (2012) Genetic correction of human induced pluripotent stem cells from patients with spinal muscular atrophy. Sci Transl Med 4:165ra152.
Additional Table 1 Primer sequence used to construct the SMN2 mini-gene with UCAC sequence deletion or mutation on SMN2 exon 7

| Name         | Primer sequence (5’-3’)                                           |
|--------------|-------------------------------------------------------------------|
| DelUCAC-F    | TCAAAAAGAAGGAAGGTGCATTCCCTTAATAAGG                                |
| UAAC-F       | TCAAAAAGAAGGAAGGTGCTAACATCCTTAATAAGG                              |
| UCAU-F       | TCAAAAAGAAGGAAGGTGCTCATATTCCTTAATAAGG                              |
| CCAU-F       | TCAAAAAGAAGGAAGGTCCCATATTCCTTAATAAGG                              |
| CCAC-F       | TCAAAAAGAAGGAAGGTCATATTCCTTAATAAGG                                |
| UGAC-F       | TCAAAAAGAAGGAAGGTGCTGACATTCCTTAATAAGG                              |
| UUAC-F       | TCAAAAAGAAGGAAGGTCATTCCCTTAATAAGG                                |
| UAGC-F       | TCAAAAAGAAGGAAGGTGCTAGCATTCCTTAATAAGG                              |
| UUGC-F       | TCAAAAAGAAGGAAGGTCGCTGCCATTCCTTAATAAGG                             |
| UUGC-F       | TCAAAAAGAAGGAAGGTGCTGCCATTCCTTAATAAGG                             |
| UUGC-F       | TCAAAAAGAAGGAAGGTGCTGCCATTCCTTAATAAGG                             |
| UUGC-F       | TCAAAAAGAAGGAAGGTGCTGCCATTCCTTAATAAGG                             |
| mutant primers-R | GCACCTTTCCTTTTTTGATTTTGCTCTAAAACCT                              |

DelUCAC-F: The forward primer of deletion of bases UCAC; SMN2: survival of motor neuron 2.
| Name        | Forword (5’-3’)                           | Reverse (5’-3’)                           |
|-------------|-------------------------------------------|-------------------------------------------|
| mGapdh      | CCGTAGACAAATATGGTGAGGT                  | CGTAGGTTGAGTCTAACCAG                  |
| mHnrnpk     | AGGAAGAAACCTCCCCCAACCC                  | TGCCCAATTCCACATCTCATC                  |
| mHnrnpl     | AGTGGAGCTGACCTTGGTGGAGGA                | AGCAGTGCTCTCACAATCTACCAG                |
| mHnrnp          | AGTCAGCCAGTAAAGAAAGG                  | AGGCACTGCTGCAGCATATGATG                 |
| mHnrnu          | AGTCCTCACAGCCCCATGTGTT                | AGCAGTGCTGCTGCAGCATATGATG                 |
| mHnrnpd        | ATGGTGTTCCACAGAGCAAGGA                | TGGTAGTTCCACAGAGCAAGGA                |
| mHnrnpf        | TGCCATTGGTTCGACGCAAGA                | AGGCGCTCCCTGTCGTT                   |
| mHnrnph2       | ACCTTTTGTAGTGGAATGGCAGTC                | AGAGCTCCAGTGATGTTACCT                   |
| mHnrnpl        | AGCCGTTGAGGGAAGAGAAAGA                | ACCCTGTTAGAGTTCTTCTCAAAGG               |
| mHnrnpu        | TCGGGTGAGTTTGAGGAGTCTGAAG              | ACCACAGAGATGCTCCTG                   |
| mSrsf10          | AGTGCTCTGTCTGGTGCAGAAGCC                | ACCATAAGCGACAAATTCGG                   |
| mSrsf1          | TTGCCTTCGTTGAGGTCTGAG                  | CGGTACCCGTCGAGTCTAG                   |
| mSrsf3          | ATTTGAAAGTCCCAGAGATGC                 | TCTTTTACACGCCAGATG                   |
| mSrsf6          | TGCCGTTGAGCTGACTCAGAAG                 | AGGTATCCCGACTGTTACCC                   |
| mSrsf2          | AGCCACACCAAGTCCTCCAGAGAAG              | ATGGACGGATGAGCTGAGTTG                  |
| mSrsf4          | ATCTGTGAGCCGTCAAGAAGCCG                | ACCACAGAGATGCTCCTG                   |
| mSrsf5          | ACTCAAGAGCGAGTCTGACTGAGG               | TCTCTTCGCTCGTCCAGG                   |
| mSrsf7          | ATTCAGATGCTAGAGCCGGGAGAAGC             | ACTGTGGCATAATCCGGAG                   |
| mSrsf9          | AGATCGAGCTCAAGAAGCCGG                 | ACAGTGCCCATATCCGG                     |
| mNova1         | ACTGGAGCCACTATCCTGAGT                  | ATCCATGACGCTGGTCCAGG                  |
| mNova2         | AGAGAGAGAGGCACTACAGAG                 | AGACAGAGGACACCATACAG                   |
| hGAPDH         | GAAGGTGGAGGCAGCTGGAGGT                 | TGGAGATGCTGGTAGGATG                   |
| hSMN1/2        | ACCACCCCCATCTATCTACTATG                | GAAGTGGAGACACTTCCTTCTT                 |
| hNOVA1        | GCCAATCTTCCCCAAGCTACCA                | TGCTCCCTACAGCTACCA                   |

*kGAPDH*: Human glyceraldehyde-3-phosphate dehydrogenase; *hNOVA1*: human neuro- oncological ventral antigen 1; *hSMN1/2*: human survival of motor neuron 1/2; *mGapdh*: mouse glyceraldehyde-3-phosphate dehydrogenase; *mHnrnpd*: mouse heterogeneous nuclear ribonucleoprotein D; *mHnrnpf*: mouse heterogeneous nuclear ribonucleoprotein F; *mHnrnph2*: mouse heterogeneous nuclear ribonucleoprotein H2; *mHnrnph3*: mouse heterogeneous nuclear ribonucleoprotein H3; *mHnrnph*: mouse heterogeneous nuclear ribonucleoprotein K; *mHnrnp*: mouse heterogeneous nuclear ribonucleoprotein L; *mHnrnpi*: mouse heterogeneous nuclear ribonucleoprotein I; *mSrsf*: mouse serine and arginine rich splicing factor.
Additional Figure 1 PCGT7-NOVA1 sequence by forward CMV primer and reverse primer.
CMV: Cytomegaviyns; NOVA1: neuro-oncoological ventral antigen 1.
Additional Figure 2 Western blot detection of NOVA1 expressions after SMN knockdown.
Knockdown efficiency of SMN1/2 was checked by western blot in U87MG cells. Protein level of NOVA1 did not obviously change after SMN1/2 inhibition. The experiment was repeated three times. NOVA1: Neuro-oncological ventral antigen 1; SMN: survival of motor neuron; U87MG: a human glioma cell line.
Additional Figure 3 Diagrams of sequencing results before and after UCAC sequence deletion or mutation. (A) Sequencing map of partial sequence before UCAC sequence mutation. (B) Sequencing map of partial sequence after mutation of UCAC to UAAC. (C) Sequencing diagram of partial sequence after deletion of UCAC sequence, and the red arrow indicates the deletion of UCAC point. (D-F) Sequencing map of partial sequence after mutation of pyrimidine combination at both ends of UCAC. (G-L) Sequencing map of partial sequence after further mutation of UCAC core sequence "CA", in which the red boxes showed the mutation sequence of the UCAC.