

Supplemental Information

# **A SMN-Dependent U12 Splicing Event Essential for Motor Circuit Function**

Francesco Lotti, Wendy L. Imlach, Luciano Saieva, Erin S. Beck, Le T. Hao,  
Darrick K. Li, Wei Jiao, George Z. Mentis, Christine E. Beattie, Brian D.  
McCabe, and Livio Pellizzoni

## Inventory of Supplemental Information

### Supplemental Data

**Figure S1**, related to Figure 1. Characterization of NIH3T3 cell lines with regulated knockdown of endogenous mouse SMN.

**Figure S2**, related to Figure 2. Specificity of the effects of SMN deficiency in NIH3T3 cells.

**Figure S3**, related to Figure 3. Expression and splicing analysis of putative U12 intron-containing genes in *Drosophila*.

**Figure S4**, related to Figure 4. Efficiency of RNAi-mediated knockdown of SMN target genes in *Drosophila*.

**Figure S5**, related to Figure 5. Stasimon is an evolutionarily conserved protein that is highly expressed in the CNS.

**Figure S6**, related to Figure 6. Specificity of Stasimon effects on motor axon development in zebrafish embryos.

**Figure S7**, related to Figure 7. Effects of SMN deficiency on Stasimon U12 splicing and mRNA expression of in SMA mice.

**Table S1**, related to Figure 1. U12 intron-containing mouse genes analyzed in this study.

**Table S2**, related to Figure 3. *Drosophila* genes with bioinformatically predicted U12 introns analyzed in this study.

**Table S3**. Primers and probes used in this study.

### Supplemental Experimental Procedures

### Supplemental References

**Figure S1, related to Figure 1. Characterization of NIH3T3 cell lines with regulated knockdown of endogenous mouse SMN.**

(A) Schematic representation of the lentiviral vectors used to establish NIH3T3 cell lines with inducible RNAi knockdown of endogenous mouse SMN (NIH3T3-Smn<sub>RNAi</sub>) as well as cell lines with inducible RNAi knockdown of endogenous mouse SMN and constitutive expression of RNAi-resistant human SMN (NIH3T3-SMN/Smn<sub>RNAi</sub>). See also Supplemental Experimental Procedures. (B) RT-qPCR analysis of mouse SMN mRNA expression in wild-type NIH3T3 (Control), NIH3T3-Smn<sub>RNAi</sub> and NIH3T3-SMN/Smn<sub>RNAi</sub> cells cultured without (blue bars) or with (red bars) Dox for 5 days. RNA levels in Dox-treated cells are expressed relative to those in untreated cells. Data are represented as mean and SEM. Dox treatment has no effect on expression of SMN mRNA in wild-type NIH3T3 cells while causes knockdown of SMN mRNA in NIH3T3-Smn<sub>RNAi</sub> and NIH3T3-SMN/Smn<sub>RNAi</sub> cells to a similar extent. (C) SMN deficiency decreases U1 snRNP assembly in NIH3T3 cells. Extracts from wild-type NIH3T3 (Control), NIH3T3-Smn<sub>RNAi</sub> and NIH3T3-SMN/Smn<sub>RNAi</sub> cells cultured without (-) or with (+) Dox for 5 days were prepared using three independent biological replicates per group and analyzed in parallel. *In vitro* snRNP assembly experiments were carried out with radioactive U1 snRNA and extracts from NIH3T3 cells followed by immunoprecipitation with anti-SmB antibodies, electrophoresis on denaturing polyacrylamide gels and autoradiography. In SMN-deficient NIH3T3 cells, the reduction in U1 snRNP assembly and the degree of its correction by human SMN is proportional to the levels of SMN expression relative to the uninduced state (see Figure 1A). (D) Quantification of SMN-dependent U1 snRNP assembly defects in NIH3T3 cells. The amounts of immunoprecipitated U1 snRNA from the experiment in (C) were quantified and RNA levels in Dox-treated cells expressed relative to those in untreated cells. Data are represented as mean and SEM. (E) SMN deficiency decreases snRNP assembly of all snRNAs of the Sm class. *In vitro* snRNP

assembly experiments were carried out with each of the indicated snRNAs and extracts from NIH3T3-Smn<sub>RNAi</sub> cells cultured without or with Dox for 5 days as in (C). The amount of each snRNA immunoprecipitated from snRNP assembly reactions with extracts from Dox-treated cells were quantified and expressed relative to those from untreated cells. Data are represented as mean and SEM. (F) Expression and splicing analysis of U12 intron-containing mouse genes in NIH3T3 cells. RT-PCR analysis of the indicated U12 intron-containing genes in wild-type NIH3T3 (Control) and NIH3T3-Smn<sub>RNAi</sub> cells cultured without (-) or with (+) Dox for 5 days. Genes and exons monitored by PCR are indicated on the left. Schematics of spliced and intron-containing mRNAs are shown on the right. Red lines highlight the position of U12 introns. The -RT lanes correspond to RT-PCR reactions lacking reverse transcriptase. (G) Expression and splicing analysis of U2 intron-containing mouse genes in NIH3T3 cells. RT-PCR analysis of the indicated genes was carried out as in (F). The *Drosophila* homologs of *Mgat1* (CG13431), *Sf3a1* (CG16941), *Znf830* (CG11839) and *Zrsr2* (CG3294) are U12 intron-containing genes, three of which (CG13431, CG16941 and CG11839) are affected by SMN deficiency in *smn* mutant third-instar larvae.

**Figure S2, related to Figure 2. Specificity of the effects of SMN deficiency in NIH3T3 cells.**

(A) Analysis of U2 intron splicing in mouse SMN target genes with U12 introns. RT-PCR analysis of U2 splicing in the indicated genes was carried out in wild-type NIH3T3 (Control) and NIH3T3-Smn<sub>RNAi</sub> cells cultured without (-) or with (+) Dox for 5 days. Genes and exons monitored by PCR are indicated on the left. Schematics of spliced and intron-containing mRNAs are shown on the right. Gapdh was used for RNA calibration across samples. The -RT lanes correspond to RT-PCR reactions lacking reverse transcriptase.

(B) Dox has no effects on proliferation of NIH3T3 cells. Equal numbers of wild-type NIH3T3 cells were plated and cultured either with or without Dox for the indicated number of days. Cell number was determined at each time point.

(C) RNAi-resistant human SMN corrects cell proliferation defects caused by depletion of mouse SMN in NIH3T3 cells. Equal numbers of NIH3T3-SMN/Smn<sub>RNAi</sub> cells were plated and cultured either with or without Dox for the indicated number of days. Cell number was determined at each time point.

(D) Serum deprivation decreases proliferation of NIH3T3 cells. Equal numbers of wild-type NIH3T3 cells were plated and cultured in the presence of either 10% FBS or 2% FBS for the indicated number of days. Cell number was determined at each time point.

(E) Serum deprivation does not decrease SMN levels in NIH3T3 cells. Western blot analysis of wild-type NIH3T3 cells cultured in the presence of 2% FBS for the indicated number of days.

(F) Effect of serum deprivation on SMN-dependent U12 splicing events. RT-qPCR analysis of U12 intron retention in Clcn7, Parp1, Tspan31 and Tmem41b mRNAs. Red bars show RNA levels in NIH3T3-Smn<sub>RNAi</sub> cells cultured for 3 days with Dox relative to untreated cells (dotted line). Blue bars show RNA levels in wild-type NIH3T3 cells cultured for 3 days with 2% FBS relative to cells grown with 10% FBS (dotted line). Data are represented as mean and SEM.

(G) Effect of serum deprivation on SMN-dependent U12 splicing events. RT-qPCR analysis of the levels of aberrantly

spliced *Tmem41b* and exon-skipped *Clcn7* mRNAs. Red bars show RNA levels in NIH3T3-*Smn*<sub>RNAi</sub> cells cultured for 5 days with Dox relative to untreated cells (dotted line). Blue bars show RNA levels in wild-type NIH3T3 cells cultured for 5 days with 2% FBS relative to cells grown with 10% FBS (dotted line). Data are represented as mean and SEM.

**Figure S3, related to Figure 3. Expression and splicing analysis of putative U12 intron-containing genes in *Drosophila*.**

(A) RT-PCR analysis of bioinformatically-predicted U12 intron-containing genes whose expression is not affected in *Drosophila smn* mutants. Equal amounts of total RNA from control, *smn*<sup>73A0</sup> and *U6atac*<sup>K01105</sup> third-instar *Drosophila* larvae were used. Genes and exons monitored by PCR are indicated on the left. Schematics of spliced and intron-containing mRNAs are shown on the right. Red lines highlight the position of U12 introns. The –RT lanes correspond to RT-PCR reactions lacking reverse transcriptase.

(B) RT-PCR analysis of *Drosophila* genes that contain only U2 introns. *CG7939 (RpL32)* was used for RNA calibration across samples in both (A) and (B). The mouse homologs of *CG8594 (Clcn7)*, *CG8545 (Nol1)*, *CG40411 (Parp1)*, *CG6335 (Harsl)* and *CG8454 (Vps16)* are U12 intron-containing genes affected by SMN deficiency in NIH3T3 cells.

(C) RT-qPCR analysis of U12 intron-containing genes whose mRNA expression is affected in both *U6atac*<sup>K01105</sup> and *smn*<sup>73A0</sup> mutants using total RNA from control and *smn*<sup>X7</sup> third-instar *Drosophila* larvae. The mRNA levels in *smn*<sup>X7</sup> mutants were expressed relative to control larvae. Data are represented as mean and SEM.

**Figure S4, related to Figure 4. Efficiency of RNAi-mediated knockdown of SMN target genes in *Drosophila*.**

RT-qPCR analysis of mRNA expression levels was carried out using total RNA from control and RNAi larvae. For all genes except *CG16941*, the effects of expression of UAS-RNAi constructs by a ubiquitous driver (*da-Gal4*) were analyzed in whole third-instar *Drosophila* larvae. Due to embryonic lethality of ubiquitous *CG16941* RNAi, mRNA knockdown of this gene was analyzed in muscle tissue following expression of the UAS-RNAi construct with a muscle-specific driver (*G14-Gal4*). The mRNA levels in RNAi larvae were expressed relative to those in the corresponding controls (without Gal4 expression). Data are represented as mean and SEM.



**Figure S5, related to Figure 5. Stasimon is an evolutionarily conserved protein that is highly expressed in the CNS.**

(A) Alignment of Stasimon protein sequences highlights strong evolutionary conservation across species. The amino acid sequences of human STASIMON/TMEM41b (GenBank™ NP\_055827) and orthologs from *P. troglodytes* (Ensembl ID ENSPTRG00000003346), *C. familiaris* (GenBank™ XP\_851421), *R. norvegicus* (GenBank™ NP\_001012358), *M. musculus* (GenBank™ NP\_705745), *G. gallus* (GenBank™ NP\_001008469), *X. tropicalis* (GenBank™ NP\_001016955), *D. rerio* (GenBank™ NP\_001073456), *C. elegans* (GenBank™ NP\_001073456) and *D. melanogaster* (GenBank™ NP\_573225) were aligned using ClustalW2 and BOXSHADE 3.21. Black boxes highlight identical amino acids; grey boxes highlight similar amino acids. (B) Schematic representation of Stasimon protein structure showing the position of the predicted six transmembrane domains and SNARE-associated domain. (C) *In situ* hybridization in *Drosophila* embryos shows ubiquitous expression of *Stasimon* mRNA with remarkably high levels in both brain and ventral cord. (D) *In situ* hybridization in the spinal cord of wild type mice at P4 shows widespread neuronal expression of *Stasimon* mRNA with remarkably high levels in motor neurons and DRG neurons. Image from the Allen Brain Atlas (<http://mousespinal.brain-map.org>) (Lein et al., 2007).

**Figure S6, related to Figure 6. Specificity of Stasimon effects on motor axon development in zebrafish embryos.**

(A) Schematic representation of a portion of the zebrafish *stasimon* pre-mRNA that includes the splice junction between intron 1 and exon 2 targeted by *stas* MO. Splice sites (GU and AG) and branch point adenosine (A) are indicated (top panel). RT-qPCR analysis of Stasimon mRNA levels in zebrafish *Tg(mnx1:GFP)* embryos injected with Control MO and *stas* MO normalized to  $\beta$  actin mRNA (bottom panel). Data are represented as mean and SEM. (B) Western blot analysis of hnRNP Q protein levels in *Tg(mnx1:GFP)* embryos injected with Control MO and *hnrnp* Q MO. (C) Representative lateral view of motor axons in zebrafish *Tg(mnx1:GFP)* morphants injected with *hnrnp* Q MO. (D) Quantification of the effects of hnRNP Q deficiency on motor axon development in zebrafish. Motor axons were scored in *Tg(mnx1:GFP)* embryos injected with Control MO or *hnrnp* Q MO. Embryos were classified as severe, moderate, mild or no defects based on the severity of motor axons defects, as previously described (Carrel et al., 2006), and the percentage for each group is shown. Data are represented as mean and SEM. (E) Quantification of the effects of Stasimon overexpression on normal motor axon development in zebrafish. Motor axons were scored in *Tg(mnx1:GFP)* embryos injected with either control MO or STAS RNA and embryos were classified as in (D). Data are represented as mean and SEM. (F) Representative lateral view of motor axons in zebrafish *Tg(mnx1:GFP)* morphants co-injected with *smn* MO and Bcl2 RNA. (G) Quantification of Bcl2 effects on SMN-dependent motor axons defects in zebrafish. Motor axons were scored in *Tg(mnx1:GFP)* embryos injected with *smn* MO or *smn* MO together with Bcl2 RNA and embryos were classified as in (D). Data are represented as mean and SEM. (H) Western blot analysis of Smn protein levels in *Tg(mnx1:GFP)* embryos injected with Control MO as well as *smn* MO either in the presence or in the absence of co-injected STAS RNA and Bcl2 RNA.

**Figure S7, related to Figure 7. Effects of SMN deficiency on *Stasimon* U12 splicing and mRNA expression in SMA mice.**

(A) Schematic of the portion of *Stasimon* pre-mRNA from exon 3 to exon 5 including the splice sites and introns are shown at the top. Green lines indicate the aberrant splicing event due to activation of a cryptic 5' splice site in exon 3 that is caused by SMN deficiency. Schematics of intron-containing and aberrantly spliced *Stasimon* mRNAs are shown at the bottom. Arrows represent the primers used for RT-qPCR. (B) RT-qPCR analysis of the aberrantly spliced *Stasimon* mRNA in the brain and kidney from control and SMA mice at the indicated post-natal days. Data are represented as mean and SEM. (C) RT-qPCR analysis of *Stasimon* U12 intron retention in the brain and kidney from control and SMA mice at the indicated post-natal days. Data are represented as mean and SEM. (D) RT-qPCR analysis of *Stasimon* mRNA expression in the brain, kidney, spinal cord and L1 DRG from control and SMA mice at the indicated post-natal days. Data are represented as mean and SEM. (E) Representative epifluorescence and bright-field images of the ventral spinal cord before and after LCM of CTb-488 labeled motor neurons. (F) Representative epifluorescence and bright-field images of the DRG before and after LCM of CTb-488 labeled proprioceptive neurons. (G) RT-qPCR analysis of ChAT and Parvalbumin expression validating the identity of the collected CTb-488<sup>+</sup> LCM neurons in the spinal cord as motor neurons and in the DRG as proprioceptive neurons. Data are represented as mean and SEM. The levels of *Stasimon* mRNA in LCM iliopsoas motor neurons relative to LCM iliopsoas proprioceptive neurons at P6 were similar.

## Supplemental Experimental Procedures

### Lentiviral constructs and viral production

With the exception of commercially available pLenti6/TR (Invitrogen), all other lentiviral constructs were generated by standard cloning techniques using the pRRLSIN.cPPT.PGK-GFP.WPRE vector (Addgene plasmid 12252) as a backbone (Dull et al., 1998; Zufferey et al., 1998). Schematic representations of these constructs are shown in Figure S1A. The pLenti6/TR construct constitutively expresses the tetracycline-dependent repressor (TetR) protein under the control of the CMV promoter as well as the blasticidin resistance gene from the SV40 promoter. The pLenti.pur/Smn<sub>RNAi</sub> construct expresses an shRNA targeting mouse SMN mRNA (5'-GAAGAAUGCCACAACUCCC-3') under the control of a tetracycline-regulated H1<sub>TO</sub> promoter as well as the puromycin resistance gene from the PGK promoter. The pLenti.hyg/SMN constitutively expresses an RNAi-resistant, epitope-tagged human SMN (Flag and Strep fused in tandem at the amino-terminus) under the control of the PGK promoter as well as the hygromycin resistance gene from the SV40 promoter. Viral stocks pseudotyped with the vesicular stomatitis G protein (VSV-G) were prepared by transient co-transfection of 293T cells using the ViraPower™ Lentiviral Packaging Mix (Invitrogen) following manufacturer's instructions.

### NIH3T3 cell lines and tissue culture

The NIH3T3 cell lines used in this study were generated through lentiviral transduction. 48 hours after lentiviral transduction carried out as described previously (Dull et al., 1998), NIH3T3 cells were split at a 1 to 5 ratio and grown in medium containing the appropriate antibiotic at the following final concentrations: 5 µg/ml Blasticidin-S hydrochloride (Invitrogen), 5 µg/ml Puromycin (Sigma), and 250 µg/ml Hygromycin B (Invitrogen). NIH3T3-Smn<sub>RNAi</sub> cells were generated by transduction of wild-

type NIH3T3 cells with pLenti6/TR and pLenti.pur/Smn<sub>RNAi</sub> followed by antibiotic selection and cloning by limiting dilution. In these cells, TetR binding to the H1<sub>TO</sub> promoter represses shRNA transcription in normal conditions while addition of the tetracycline analogue doxycycline to the culture medium triggers shRNA expression and RNAi-mediated knockdown of endogenous mouse SMN (Figures 1A and S1B). To control for potentially non-specific effects of shRNA expression, NIH3T3-SMN/Smn<sub>RNAi</sub> cells were generated by transduction of NIH3T3-Smn<sub>RNAi</sub> cells with pLenti.hyg/SMN followed by antibiotic selection. These cells express an epitope-tagged human SMN isoform that is resistant to RNAi (Figures 1A and S1B) and expected to correct defects caused specifically by depletion of endogenous mouse SMN in NIH3T3-Smn<sub>RNAi</sub> cells.

Mouse NIH3T3 fibroblasts were grown in DMEM with high glucose (Invitrogen) containing 10% of FBS (HyClone), 2 mM glutamine (Gibco), and 1% penicillin and streptomycin (Gibco). RNAi was induced by addition to the growth medium of Doxycycline HCl (Fisher Scientific) at the final concentration of 100 ng/ml. For serum deprivation experiments, wild-type NIH3T3 cells were cultured in the presence of 2% FBS. Cell number was determined with an automatic digital cell counter (ADAM, Digital Bio).

### **Antibodies and Western blot**

The following monoclonal antibodies were used in this study: anti-SMN clone 8 (BD Transduction Laboratories), *Drosophila*-specific anti-SMN (Chang et al., 2008), anti-SmB 18F6 (Carissimi et al., 2006), anti- $\beta$ -actin (Sigma), anti-Tubulin DM 1A (Sigma), anti-FLAG (Sigma). For zebrafish experiments, an anti-hnRNP Q rabbit polyclonal antibody (Abcam), an anti-SMN mouse monoclonal antibody (2E6, gift from Dr. Glenn Morris) and an anti- $\beta$  actin mouse monoclonal antibody (Santa Cruz) were used. For immunohistochemistry of mouse tissue, an anti-ChAT goat polyclonal antibody

(Millipore) and an anti-parvalbumin chicken polyclonal antibody (Covance) were used.

Total protein extracts for Western blot analysis were prepared by homogenization of *Drosophila* third-instar larvae, zebrafish embryos or NIH3T3 cells in SDS sample buffer (2% SDS, 10% glycerol, 5%  $\beta$ -mercaptoethanol, 60 mM Tris-HCl pH 6.8, bromophenol blue) followed by brief sonication and boiling. Protein concentration was measured using RC DC Protein Assay (Bio-Rad). All protein samples were analyzed by SDS/PAGE on 12% polyacrylamide gels followed by transfer onto nitrocellulose membrane and Western blot. SMN levels were determined relative to signal intensity in the serial dilution followed by normalization to Tubulin using ImageJ.

### **SMA mice and laser capture microdissection**

All animal procedures were approved by the Institutional Laboratory Animal Care and Use Committee of Columbia University. FVB.Cg-Tg(*SMN2\**delta7)4299Ahmb Tg(*SMN2*)89Ahmb *Smn1tm1Msd/J* (JAX Stock No:005025) mice were interbred to obtain SMA- $\Delta$ 7 (*Smn*<sup>-/-</sup>; *SMN2*<sup>+/+</sup>; *SMN $\Delta$ 7*<sup>+/+</sup>) mice (Le et al., 2005). Genotyping of SMA mice was carried out as previously described (Gabanella et al., 2007).

Spinal cord and DRG from control and SMA mice injected in the iliopsoas muscle with CTb-488 were embedded in OCT, cryo-sectioned (10  $\mu$ m) and adhered to UV treated PEN membrane slides (Leica). Tissue sections were immediately fixed in ethanol for 15 seconds prior to laser capture microdissection (LCM). LCM was carried out with the Leica DM6000B under 40X magnification to isolate CTb-488<sup>+</sup> motor neurons from the L1-2 spinal cord segment and large (>20 $\mu$ m) CTb-488<sup>+</sup> proprioceptive neurons from L1-2 DRG of control and SMA mice. Immunohistochemistry and confocal microscopy of CTb-labeled tissues was performed to validate the identity of CTb-488<sup>+</sup> neurons in the

spinal cord as motor neurons (ChAT<sup>+</sup>) and in the DRG as proprioceptive neurons (parvalbumin<sup>+</sup>) as previously reported (Arber et al., 2000; Mentis et al., 2011).

### **Immunohistochemistry and confocal microscopy**

For immunohistochemical analysis of mouse tissue, the spinal cord and the associated DRG (L1 and L2 spinal segments) were dissected under *in vitro* conditions and immersion fixed with 4% paraformaldehyde as previously described (Mentis et al., 2011). The spinal cord and DRG were harvested under cold (~16°C), oxygenated (95% O<sub>2</sub>, 5% CO<sub>2</sub>) artificial cerebrospinal fluid (128.35 mM NaCl, 4 mM KCl, 0.58 mM NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, 21 mM NaHCO<sub>3</sub>, 30 mM D-Glucose, 0.1 mM CaCl<sub>2</sub>·H<sub>2</sub>O, and 2 mM MgSO<sub>4</sub>·7H<sub>2</sub>O). The tissue was immersion-fixed in 4% paraformaldehyde for 4 hours and then washed in 0.01 M phosphate buffer saline (PBS). Tissues were embedded in warm 5% Agar and serial transverse sections (70-80 μm) were cut on a Vibratome. The sections were blocked with 10% normal donkey serum in PBS-T (0.01 M PBS with 0.1% Triton X-100, pH 7.4) and incubated overnight at room temperature with a goat anti-ChAT polyclonal antibody (1:100) and a chicken anti-parvalbumin antibody (1:16,000). Secondary antibody incubations were performed for 3 hours with species-specific antisera coupled to either Rhodamine-RX or Cy5 (Jackson laboratories) diluted at 1:50 in PBS-T. After secondary antibody incubations, the sections were washed in PBS and mounted on slides with an anti-fading solution made of Glycerol:PBS (3:7). Sections were imaged using an SP5 Leica confocal microscope.

### **RNA analysis**

Total RNA from *Drosophila* third-instar larvae, zebrafish embryos, mouse tissues and NIH3T3 cells was purified using TRIzol reagent (Invitrogen) and treated with RNase-free DNase I (Ambion) to remove DNA contamination. RNA from LCM neurons was

purified using the Absolutely RNA Nanoprep Kit (Agilent) and linear amplification was performed with the MessageAmp II aRNA Amplification Kit (Ambion), according to the manufacturer's instructions. A mixture of oligo-dT primers and random hexamers was used to generate cDNA using Advantage RT-for-PCR kit (Clontech) and 1 µg of total RNA following manufacturer's instructions. For semi-quantitative RT-PCR analysis, 2.5% of the cDNA was used and PCR reactions were analyzed by electrophoresis on 1.5% agarose gels followed by staining with GelRed™ (Biotium). Data represent PCR reactions within the linear range of amplification as determined for each primer pair independently and in every set of experiments by sample collection every two cycles. The identity of all the PCR products was confirmed by DNA sequencing. For quantitative RT-qPCR analysis, 1% of the cDNA was used and each measurement was carried out in triplicates in a standard 3-step qPCR reaction with a Mastercycler ep Realplex<sup>4</sup> (Eppendorf) PCR system and *Power SYBR® Green PCR Master Mix* (ABI). RT-qPCR data from NIH3T3 cells and *Drosophila* larvae were normalized to *Gapdh* and *RpL32* (CG7939) mRNAs, respectively. For Northern blot analysis, total RNA from *Drosophila* third-instar larvae (2 µg) was analyzed by electrophoresis on 8% polyacrylamide/8M urea denaturing gel and transferred to a Hybond+ membrane (GE Healthcare). Radioactive antisense RNA probes against *Drosophila* snRNAs and 5.8S rRNA were transcribed *in vitro* from DNA oligonucleotide templates. Quantification was carried out using a Typhoon PhosphorImager (Molecular Dynamics). The list of primers and probes is shown in Table S3.

### ***In vitro* snRNP assembly and analysis of snRNP levels**

NIH3T3 cell extracts were prepared by homogenization in ice-cold reconstitution buffer (20 mM Hepes-KOH pH 7.9, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 5% glycerol) containing 0.01% NP40 as previously described (Gabanella et al., 2007).



Radioactive snRNAs were generated by run-off transcription with T7 polymerase from template DNA in the presence of [ $\alpha^{32}$ -P] UTP (3000 Ci/mmol) and m7G cap analogue (New England Biolabs), and then purified from denaturing polyacrylamide gels. *In vitro* snRNP assembly experiments with radioactive snRNAs and NIH3T3 cell extracts (25  $\mu$ g) followed by immunoprecipitation with anti-SmB antibodies were carried out according to established procedures (Gabanella et al., 2007; Gabanella et al., 2005; Pellizzoni et al., 2002; Workman et al., 2009). Quantification of immunoprecipitated snRNAs was carried out using a Typhoon PhosphorImager (Molecular Dynamics). For analysis of endogenous snRNP levels, NIH3T3 cell extracts (200  $\mu$ g) were immunoprecipitated with anti-SmB antibodies in RSB-500 buffer (500 mM NaCl, 10 mM Tris-HCl pH 7.4, 2.5 mM MgCl<sub>2</sub>) containing 0.1% NP40, EDTA-free protease inhibitor cocktail (Roche) and phosphatase inhibitors (50 mM NaF, 0.2 mM Na<sub>3</sub>VO<sub>4</sub>) for 2 h at 4°C (Gabanella et al., 2007). After extensive washing with the same buffer, bound RNAs were recovered by proteinase K treatment, phenol/chloroform extraction and ethanol precipitation. The levels of snRNAs were measured using real-time RT-qPCR following the procedure previously described (Workman et al., 2009). The primers used are listed in Table S3.

### ***Drosophila* genetics and analysis**

*smn* mutants: we used an outcrossed stock of the SMN missense mutation *smn*<sup>73A0</sup> (Chan et al., 2003) (Gift from Dr. Greg Matera, UNC) and Df(3L)*smn*<sup>X7</sup>, a small deletion that completely removes the SMN coding region without perturbing nearby genes (Chang et al., 2008). *U6atac* mutant: we used the P-element insertion P(LacW)K01105 which has previously been validated to disrupt U6atac (Otake et al., 2002). SMN target gene screen: SMN target genes were inhibited by driving transgenic RNAi (Dietzl et al., 2007) expression in all tissues using da-Gal4; UAS-Dcr2 (Dietzl et al.,

2007; Perrin et al., 2003); pan-neuronally using C155-Gal4; UAS-Dcr2 (Lin and Goodman, 1994) and in muscles using G14-Gal4 (Aberle et al., 2002). *stas* mutants: the *stas* P-element insertion mutant P{EPgy2}CG8408[EY04008] (Bellen et al., 2004) was obtained from the Bloomington *Drosophila* stock center. UAS-Stas: Stasimon cDNA was amplified from LD12309 (DGRC) and cloned into pBID-UASC (Wang et al., 2012). The construct was inserted at the attP40 landing site on chromosome 2L by phiC31 transgenesis (Groth et al., 2004). Stasimon was expressed in all neurons using nsyb-Gal4 (Bushey et al., 2009), cholinergic neurons using Cha-Gal4 (Salvaterra and Kitamoto, 2001) and glutamatergic motor neurons using OK371-Gal4 (Mahr and Aberle, 2006).

*Drosophila* NMJ electrophysiology was performed as previously described (Imlach and McCabe, 2009). Rhythmic motor activity was recorded from muscle 6 of abdominal segment A1 using a semi-intact preparation where the brain and peripheral nerves remained intact (Imlach et al., 2012). For analysis of locomotion, 60-second video recordings of the locomotor paths were made with a digital video camera (Sentech STC-620CC), recorded with Final Cut Express 4.0 (Apple), converted with QuickTime 7.6.4 (Apple) and analyzed using DIAS 3.4.2 (Soll Technologies) (Imlach et al., 2012). For analysis of muscle size, dissected third-instar larvae were fixed and stained with Alexa Fluor-labelled phalloidin, and the area of muscle 6 in hemisegment A3 was measured (Imlach et al., 2012).

### ***In situ* hybridization**

*In situ* hybridization of *Drosophila* embryos was performed as previously described (Kosman et al., 2004), with the following modifications. After incubation in anti-Dig-AP Fab fragments antibody (1:500, Roche), *in situ* experiments were developed for 10-30 minutes using the BCIP/NBT Alkaline Phosphatase Substrate Kit IV (Vector

Labs). Antisense and sense Stasimon RNA probes were transcribed *in vitro* from PCR products amplified from the BDGP cDNA clone LD12309 (Rubin et al., 2000). The open reading frame of *Drosophila* Stasimon was amplified using the following forward 5'-GATAATACGACTCACTATAGGGAGAGCTCGAAATTAACCCTCACTA-3' and reverse 5'-GCAGATCTGATATCATCGCCACT-3' primers for the sense probe, as a negative control, an antisense probe was amplified using the following forward 5'-TGGCGGCCGCTCTAGAACTAG-3' and reverse 5'-GCTCGAAATTAACCCTCACTA-3' primers.

### **Zebrafish motor axon outgrowth**

Zebrafish and embryos were maintained at ~28.5°C and staged by hours or days post-fertilization (Westerfield, 1995). Transgenic *Tg(mnx1:0.6hsp70:GFP)os26* zebrafish embryos that express GFP in ventrally projecting motor axons (Dalgin et al., 2011) and are referred to as *Tg(mnx1:GFP)* were injected with ~4.5 ng of antisense morpholino oligonucleotides (MO) as previously described (McWhorter et al., 2003). The sequence of all the MOs used in this study is shown in Table S3. Control, *smn* and *tdp43* MOs were previously described (Kabashi et al., 2011; McWhorter et al., 2003). A reticule was used to measure the bolus volume injected into each embryo. For over-expression of human *Stasimon* (*STAS*), the open reading frame (Accession number NM\_015012) cloned in pcDNA3.1 was amplified with forward primer 5'-ATAGGATCCATATGGCGAAAGGCAGAGTC-3' and reverse primer 5'-AGCTCGAGAGCTTACTCAAATTTCTGCTTTAG-3' and subcloned into the pCS2+ vector using BamHI and XhoI sites. The Bcl2 (Accession number NM\_001030253.2) cDNA construct in pCS2+ was a kind gift from Dr. Thomas Look (Langenau et al., 2005). Plasmid DNA was linearized with NotI and capped RNA was generated using the Sp6

mMESSAGE mMACHINE kit (Ambion, Austin, TX) following the protocol of the manufacturer. One-two cell stage *Tg(mnx1:GFP)* embryos were injected with ~4.5 ng of the indicated MOs with or without 200-300 pg of synthetic *STAS* or *Bcl2* RNAs using an MPPI-2 Pressure Injector (Applied Scientific Instrumentation, Eugene, OR). Injections of MO and RNA were performed according to previous protocols (Carrel et al., 2006). To visualize motor axons in GFP transgenic animals, live *Tg(mnx1:GFP)* zebrafish were anesthetized with tricaine, mounted on glass coverslips for observation using a Leica TCS-SL scanning confocal microscope. Ten motor axons from each side of the embryo were scored (total of 20 per embryo) at 28 hours post-fertilization and used to classify the embryo as severe, moderate, mild, or no defects according to previously described criteria based on number and type of motor axon abnormalities (Carrel et al., 2006). At least 3 independent sets of injections were performed per condition and 18-25 embryos (360-500 motor axons) scored in each experiment. The exact percentage of severe, moderate, mild, or no defects is dependent on the batch of *smn* MO and thus may be different in this report than in previously published reports (Carrel et al., 2006; McWhorter et al., 2003; Oprea et al., 2008).

### **Statistical analysis**

Statistical analysis was carried out using two-tailed unpaired Student's t-test and the Prism 5 (GraphPad) software. Data are represented as mean and standard error of the mean (SEM) from independent experiments and P values are indicated as follows: \* =  $p < 0.05$ ; \*\* =  $p < 0.01$ ; \*\*\* =  $p < 0.001$ . Statistical analysis of RNA expression in mouse tissues was carried out using one-way ANOVA followed by the Student-Newman-Keuls post hoc test and the asterisk (\*) indicates  $p < 0.05$ . For zebrafish motor axon characterization, the distribution of larval classifications (severe, moderate, mild, and no

defects) was analyzed by comparing median values of each group with a two-tailed Mann-Whitney nonparametric rank test as previously described (Carrel et al., 2006).

## Supplemental References

Aberle, H., Haghghi, A.P., Fetter, R.D., McCabe, B.D., Magalhaes, T.R., and Goodman, C.S. (2002). wishful thinking encodes a BMP type II receptor that regulates synaptic growth in *Drosophila*. *Neuron* 33, 545-558.

Arber, S., Ladle, D.R., Lin, J.H., Frank, E., and Jessell, T.M. (2000). ETS gene *Er81* controls the formation of functional connections between group Ia sensory afferents and motor neurons. *Cell* 101, 485-498.

Bellen, H.J., Levis, R.W., Liao, G., He, Y., Carlson, J.W., Tsang, G., Evans-Holm, M., Hiesinger, P.R., Schulze, K.L., Rubin, G.M., *et al.* (2004). The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* 167, 761-781.

Bushey, D., Tononi, G., and Cirelli, C. (2009). The *Drosophila* fragile X mental retardation gene regulates sleep need. *J Neurosci* 29, 1948-1961.

Carissimi, C., Saieva, L., Gabanella, F., and Pellizzoni, L. (2006). *Gemin8* is required for the architecture and function of the survival motor neuron complex. *J Biol Chem* 281, 37009-37016.

Carrel, T.L., McWhorter, M.L., Workman, E., Zhang, H., Wolstencroft, E.C., Lorson, C., Bassell, G.J., Burghes, A.H., and Beattie, C.E. (2006). Survival motor neuron function in motor axons is independent of functions required for small nuclear ribonucleoprotein biogenesis. *J Neurosci* 26, 11014-11022.

Chan, Y.B., Miguel-Aliaga, I., Franks, C., Thomas, N., Trulzsch, B., Sattelle, D.B., Davies, K.E., and van den Heuvel, M. (2003). Neuromuscular defects in a *Drosophila* survival motor neuron gene mutant. *Hum Mol Genet* 12, 1367-1376.

Chang, H.C., Dimlich, D.N., Yokokura, T., Mukherjee, A., Kankel, M.W., Sen, A., Sridhar, V., Fulga, T.A., Hart, A.C., Van Vactor, D., *et al.* (2008). Modeling spinal muscular atrophy in *Drosophila*. *PLoS One* 3, e3209.

Dalgin, G., Ward, A.B., Hao le, T., Beattie, C.E., Nechiporuk, A., and Prince, V.E. (2011). Zebrafish *mnx1* controls cell fate choice in the developing endocrine pancreas. *Development* 138, 4597-4608.

Dietzl, G., Chen, D., Schnorrer, F., Su, K.C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., *et al.* (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448, 151-156.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *J Virol* 72, 8463-8471.

Gabanella, F., Butchbach, M.E., Saieva, L., Carissimi, C., Burghes, A.H., and Pellizzoni, L. (2007). Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. *PLoS One* 2, e921.

- Gabanella, F., Carissimi, C., Usiello, A., and Pellizzoni, L. (2005). The activity of the spinal muscular atrophy protein is regulated during development and cellular differentiation. *Hum Molecular Genet* 14, 3629-3642.
- Groth, A.C., Fish, M., Nusse, R., and Calos, M.P. (2004). Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166, 1775-1782.
- Imlach, W., and McCabe, B.D. (2009). Electrophysiological methods for recording synaptic potentials from the NMJ of *Drosophila* larvae. *J Vis Exp.* 10, 1109.
- Imlach, L.W., Beck, E.S., Choi, B.J., Lotti, F., Pellizzoni, L., and McCabe, B.D. (2012). SMN is required for sensory-motor circuit function in *Drosophila*. *Cell*.
- Kabashi, E., Bercier, V., Lissouba, A., Liao, M., Brustein, E., Rouleau, G.A., and Drapeau, P. (2011). FUS and TARDBP but not SOD1 interact in genetic models of amyotrophic lateral sclerosis. *PLoS Genet* 7, e1002214.
- Kosman, D., Mizutani, C.M., Lemons, D., Cox, W.G., McGinnis, W., and Bier, E. (2004). Multiplex detection of RNA expression in *Drosophila* embryos. *Science* 305, 846.
- Langenau, D.M., Jette, C., Berghmans, S., Palomero, T., Kanki, J.P., Kutok, J.L., and Look, A.T. (2005). Suppression of apoptosis by bcl-2 overexpression in lymphoid cells of transgenic zebrafish. *Blood* 105, 3278-3285.
- Le, T.T., Pham, L.T., Butchbach, M.E., Zhang, H.L., Monani, U.R., Covert, D.D., Gavrilina, T.O., Xing, L., Bassell, G.J., and Burghes, A.H. (2005). SMNDelta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Hum Mol Genet* 14, 845-857.
- Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., *et al.* (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168-176.
- Lin, D.M., and Goodman, C.S. (1994). Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. *Neuron* 13, 507-523.
- Mahr, A., and Aberle, H. (2006). The expression pattern of the *Drosophila* vesicular glutamate transporter: a marker protein for motoneurons and glutamatergic centers in the brain. *Gene Expr Patterns* 6, 299-309.
- McWhorter, M.L., Monani, U.R., Burghes, A.H., and Beattie, C.E. (2003). Knockdown of the survival motor neuron (Smn) protein in zebrafish causes defects in motor axon outgrowth and pathfinding. *J Cell Biol* 162, 919-931.
- Mentis, G.Z., Blivis, D., Liu, W., Drobac, E., Crowder, M.E., Kong, L., Alvarez, F.J., Sumner, C.J., and O'Donovan, M.J. (2011). Early functional impairment of sensory-motor connectivity in a mouse model of spinal muscular atrophy. *Neuron* 69, 453-467.

Oprea, G.E., Krober, S., McWhorter, M.L., Rossoll, W., Muller, S., Krawczak, M., Bassell, G.J., Beattie, C.E., and Wirth, B. (2008). Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. *Science* 320, 524-527.

Otake, L.R., Scamborova, P., Hashimoto, C., and Steitz, J.A. (2002). The divergent U12-type spliceosome is required for pre-mRNA splicing and is essential for development in *Drosophila*. *Mol Cell* 9, 439-446.

Pellizzoni, L., Yong, J., and Dreyfuss, G. (2002). Essential role for the SMN complex in the specificity of snRNP assembly. *Science* 298, 1775-1779.

Perrin, L., Bloyer, S., Ferraz, C., Agrawal, N., Sinha, P., and Dura, J.M. (2003). The leucine zipper motif of the *Drosophila* AF10 homologue can inhibit PRE-mediated repression: implications for leukemogenic activity of human MLL-AF10 fusions. *Mol Cell Biol* 23, 119-130.

Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., *et al.* (2000). Comparative genomics of the eukaryotes. *Science* 287, 2204-2215.

Salvaterra, P.M., and Kitamoto, T. (2001). *Drosophila* cholinergic neurons and processes visualized with Gal4/UAS-GFP. *Brain Res Gene Expr Patterns* 1, 73-82.

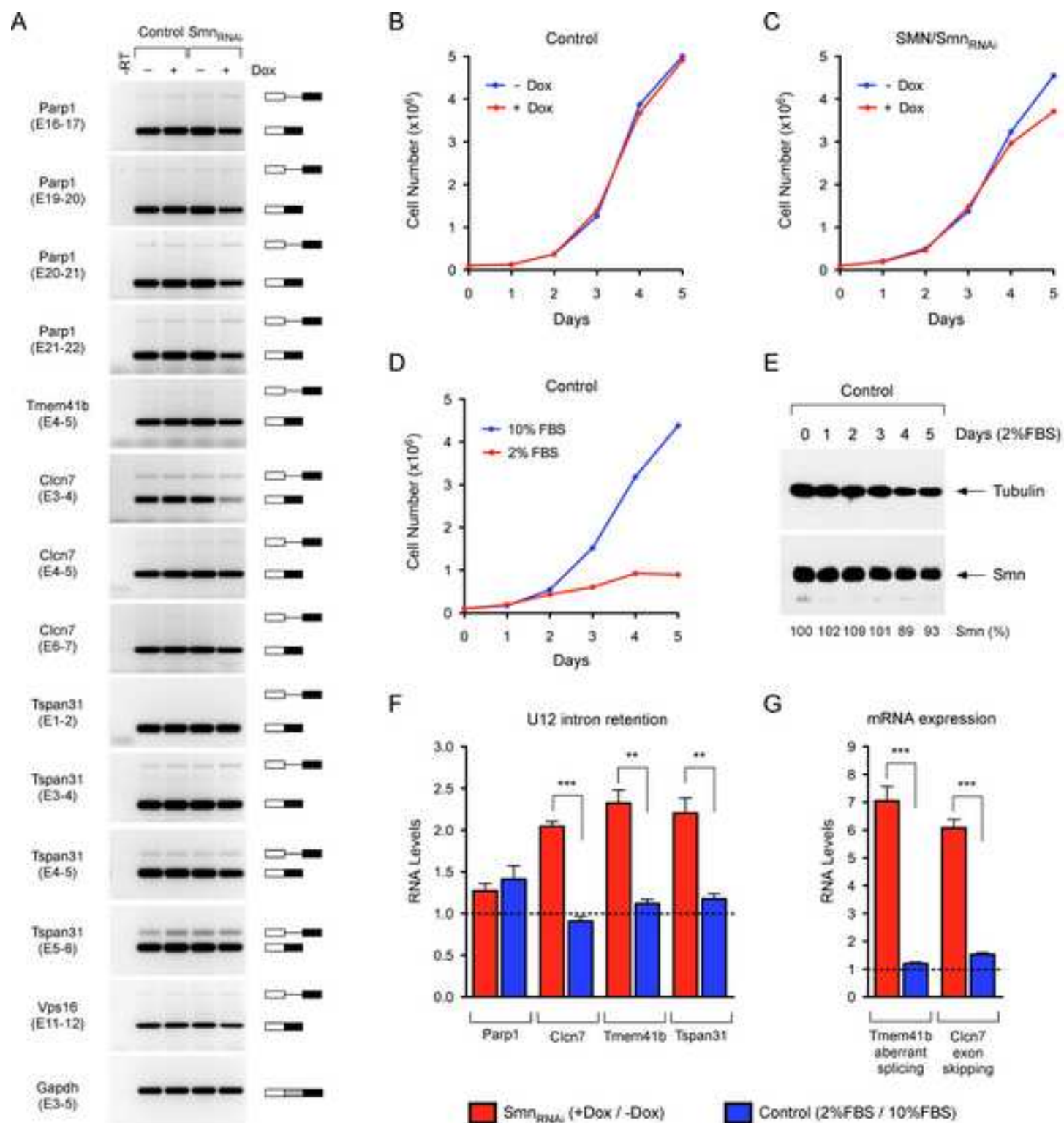
Wang, J.W., Beck, E.S., and McCabe, B.D. (2012). A modular toolset for recombination transgenesis and neurogenetic analysis of *Drosophila*. *PloS One* 7, e42102.

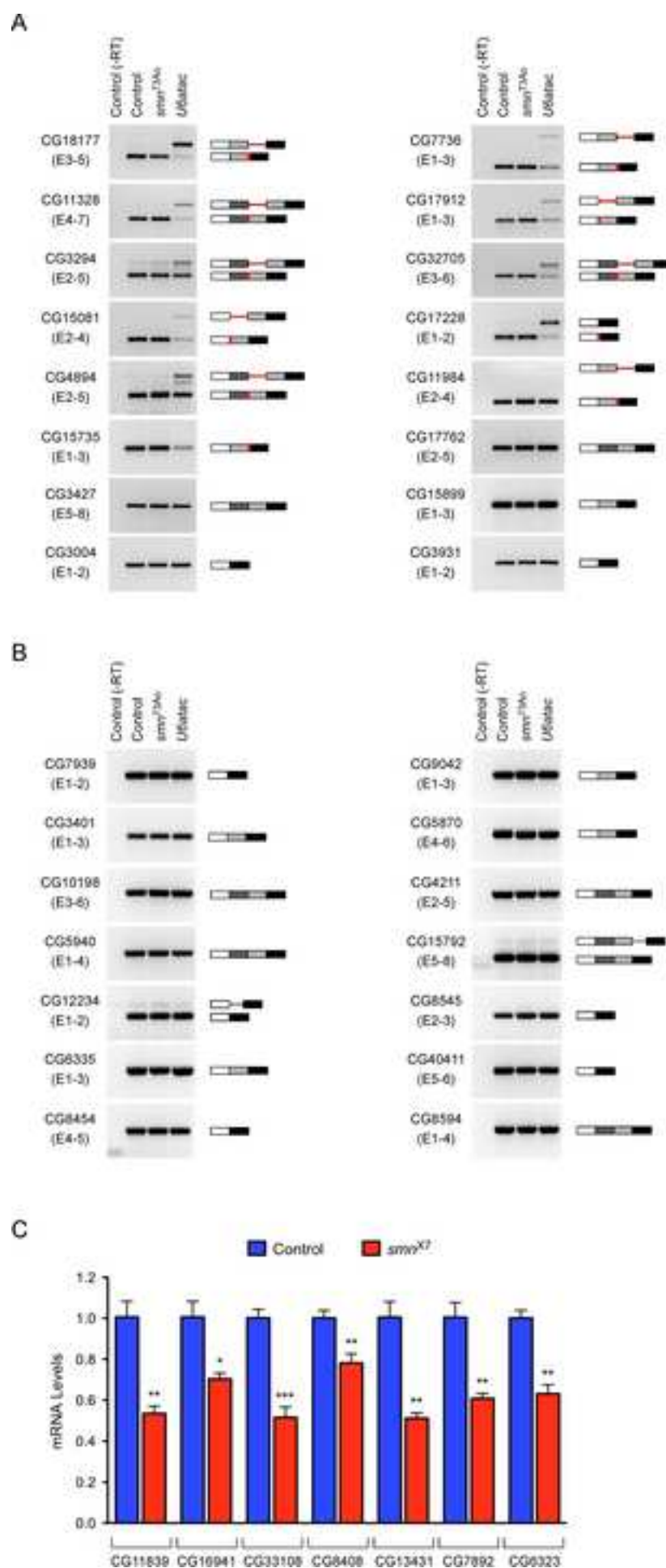
Workman, E., Saieva, L., Carrel, T.L., Crawford, T.O., Liu, D., Lutz, C., Beattie, C.E., Pellizzoni, L., and Burghes, A.H. (2009). A SMN missense mutation complements SMN2 restoring snRNPs and rescuing SMA mice. *Hum Mol Genet* 18, 2215-2229.

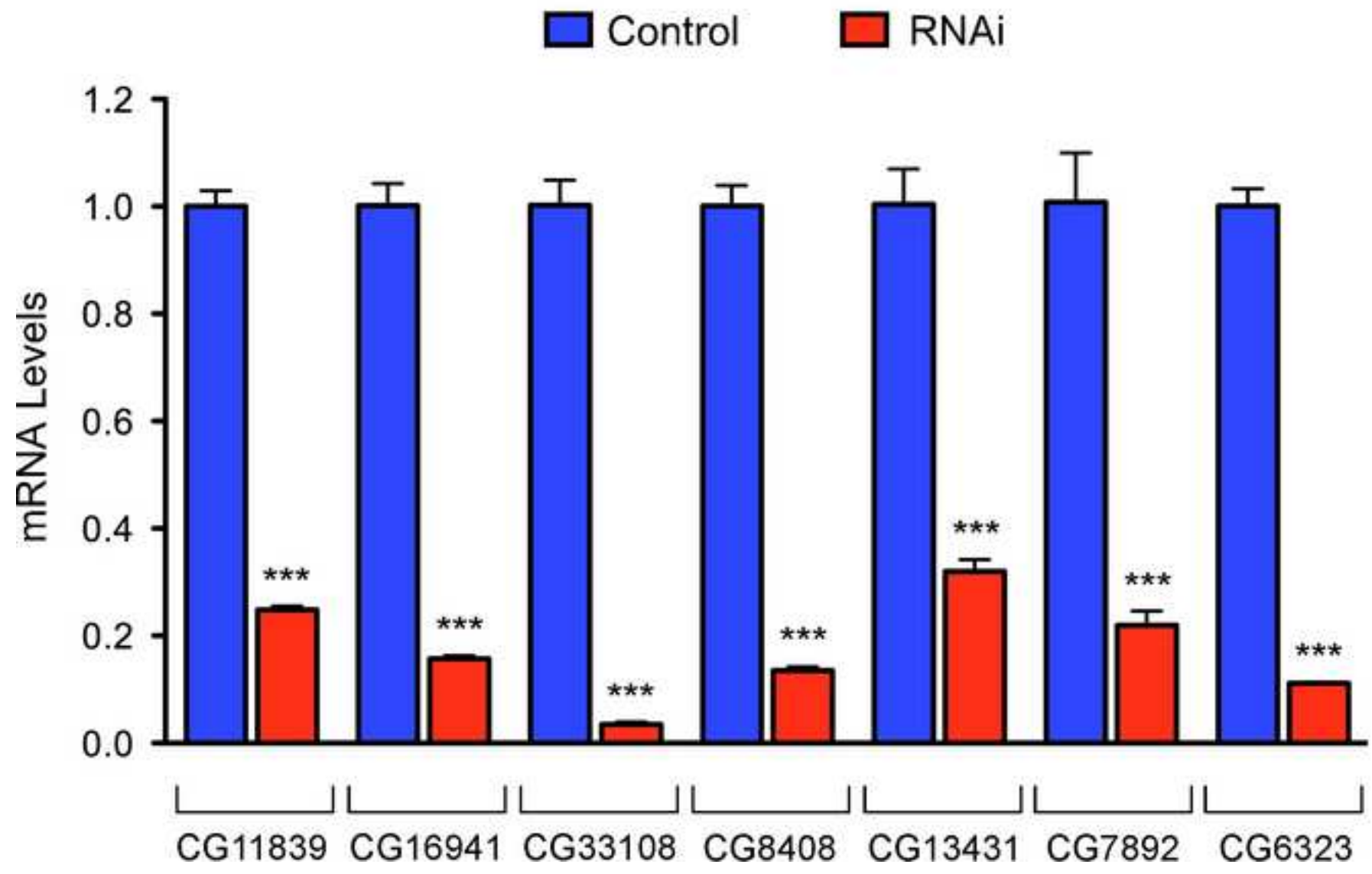
Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol* 72, 9873-9880.













A

<i>H. sapiens</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	SNV	IAQSAIR	52																																																								
<i>P. troglodytes</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	SNV	IAQSAIR	52																																																								
<i>C. familiaris</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	ACN	IAQSAIR	52																																																								
<i>H. norvegicus</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	PCV	IAQSAIR	52																																																								
<i>M. musculus</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	PCN	IAQSAIR	52																																																								
<i>G. gallus</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	AGQ	IAQSAIR	30																																																								
<i>X. tropicalis</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	FTY	TACSA	39																																																								
<i>D. rerio</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	PGQ	IAQSAIR	43																																																								
<i>C. elegans</i>	1	.....	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	S	Q	9																																																								
<i>D. melanogaster</i>	1	MSTCGV	IAISAD	EGIT	AKGRVARR	DLG--	AKHTT	PTD	AGAG	PL	LA	PGRR	QYER	.....	PKATEK	80																																																							
<i>H. sapiens</i>	53	LLILVSI	IPG	AA	VN	FL	YK	NP	PQ	LS	EE	ER	VN	HY	FD	MD	DA	KAL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	132																																				
<i>P. troglodytes</i>	53	LLILVSI	IPG	AA	VN	FL	YK	NP	PQ	LS	EE	ER	VN	HY	FD	MD	DA	KAL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	132																																				
<i>C. familiaris</i>	53	LLILVSI	IPG	AA	VN	FL	YK	NP	PQ	LS	EE	ER	VN	HY	FD	MD	DA	KAL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	132																																				
<i>H. norvegicus</i>	53	LLILVSI	IPG	AA	VN	FL	YK	NP	PQ	LS	EE	ER	VN	HY	FD	MD	DA	KAL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	132																																				
<i>M. musculus</i>	53	LLILVSI	IPG	AA	VN	FL	YK	NP	PQ	LS	EE	ER	VN	HY	FD	MD	DA	KAL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	132																																				
<i>G. gallus</i>	31	LLILVSI	IPG	AA	VN	FL	YK	NP	PQ	LS	EE	ER	VN	HY	FD	MD	DA	KAL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	110																																				
<i>X. tropicalis</i>	40	GRAT	K	FO	ET	RY	FS	CA	AF	CC	AA	SE	ER	VE	RE	DE	AK	AL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	110																																				
<i>D. rerio</i>	44	GRAT	K	FO	ET	RY	FS	CA	AF	CC	AA	SE	ER	VE	RE	DE	AK	AL	GE	VL	SK	YK	DT	FT	VQ	VL	VAT	FAT	IF	PL	QT	PA	IP	GS	123																																				
<i>C. elegans</i>	10	HPNE	V	L	A	T	T	V	S	I	P	S	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	89																																					
<i>D. melanogaster</i>	81	GGV	Y	A	G	I	Y	A	L	V	N	C	T	V	A	T	T	V	S	I	P	S	P	P	P	P	P	P	P	P	P	P	P	160																																					
<i>H. sapiens</i>	133	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	212				
<i>P. troglodytes</i>	133	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	212				
<i>C. familiaris</i>	133	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	212				
<i>H. norvegicus</i>	133	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	212				
<i>M. musculus</i>	133	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	212				
<i>G. gallus</i>	111	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	190				
<i>X. tropicalis</i>	130	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	199				
<i>D. rerio</i>	128	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	203				
<i>C. elegans</i>	90	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	169				
<i>D. melanogaster</i>	161	IFL	IL	SG	FL	Y	P	P	P	L	AL	FL	V	C	L	CG	GL	GA	S	F	C	Y	M	L	S	T	L	V	G	R	P	V	V	K	T	L	T	E	E	A	V	K	S	Q	V	E	R	R	R	R	L	I	M	T	I	I	F	L	R	I	T	F	F	L	P	M	240				
<i>H. sapiens</i>	213	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	291
<i>P. troglodytes</i>	213	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	291
<i>C. familiaris</i>	213	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	291
<i>H. norvegicus</i>	213	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	291
<i>M. musculus</i>	213	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	291
<i>G. gallus</i>	191	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	289
<i>X. tropicalis</i>	200	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	278
<i>D. rerio</i>	204	MF	I	N	I	S	P	V	I	N	V	L	K	Y	F	F	S	T	P	L	O	V	A	F	S	P	F	V	A	I	K	A	G	S	T	L	Y	Q	L	T	A	G	R	A	V	H	N	S	I	F	I	L	K	L	A	S	I	L	P	A	I	P	O	K	R	I	K	O	K	F	282
<i>C. elegans</i>	170	MF	I	N	I	S	P	V	I	N	V	L																																																											

