Supporting Information

Zhang et al. 10.1073/pnas.1319280110

SI Materials and Methods

RNA-seq Data Alignment. RNA-seq data were aligned to reference genome (University of California at Santa Cruz, UCSC, mm9) using Tophat (v1.3.1) by default parameters (1). Uniquely mapped and properly paired alignments were kept for further analysis. Reads per exon were grouped using RefSeq annotation and thus RPKM (reads per kilobase of exon model per million mapped reads) can be calculated (2). Genes with RPKM > 2 were considered as expressed genes.

Differential Gene-Expression Analysis. Differential gene-expression sequence analysis (3) was performed to detect differentially expressed genes. The input count data for each gene was computed by BEDTools (4). Replicates were considered when estimating dispersion for each gene and then negative binomial test was conducted to test differentially expressed genes.

Alternative Splicing Events Analysis. Mixture of isoforms (MISO) (5) (v0.4.1 with default parameters) was applied to the aligned RNA-seq data to identify differentially spliced isoforms between WT and spinal muscular atrophy (SMA) samples, which estimates expression at alternative splice-event level by computing Ψ (percent spliced isoform, PSI), and calculates Bayes factors to measure the differential expression. The alternative splicing events were filtered by the criteria modified from MISO manual: (*i*) at least one inclusion read, (*ii*) one exclusion read, such that (*iii*) the sum of inclusion and exclusion reads is at least 10 and (*iv*) the $|\Delta\Psi|$ is at least 0.2, and (*v*) the Bayes factor is at least 10, and *i*–*v* are true in one of the samples. Same alternative splicing events are required to be identified in both SMA samples.

Two-Exon–Skipped and Aberrant Splicing-Event Analysis. We developed custom pipelines to identify two-exon–skipping and aberrant splicing events by analyzing junction reads outputted from Tophat. For two-exon–skipped events, the coordinates were obtained from AceView database. Junction reads were then extracted accordingly and normalized to total number of mapped reads. We defined the inner junction as the "inclusion junction" across two skipping exons and the outer junction as the "exclusion junction" connected to the other two exons. We constructed a two-by-two contingency table using the inner junctions and the outer junctions in WT and SMA samples. *P* values were generated by Fisher's exact test using R (www.r-project.org) and $P < 0.01$ was set as threshold. Similarly, for the aberrant splicing events, we identified aberrant junctions by requiring one splicing site within the canonical site, whereas the other site matched the annotated splicing site. The number of canonical junctions and aberrant junctions in WT and SMA samples were then normalized and used to compute *P* values. For a splicing event to be considered as aberrant, we required at least 20 reads (in SMA) that support each aberrant junction (not annotated in RefSeq, UCSC, AceView, or Ensembl databases), twofold or more of these reads in SMA compared with WT, and $P < 0.01$.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Pathway Analyses of Dysregulated Genes. We used DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/) for all of the analysis. For the genes affected in SMA motor neurons (MNs) or white matter (WM), a background corresponding to all genes expressed in the MNs or WM was used.

RT-PCR and Quantitative RT-PCR. For all of the semiquantitative RT-PCR and quantitative RT-PCR (RT-qPCR) experiments analyzing mRNAs, we used the synthesized cDNA samples by Ovation RNA-seq System V2. Ten nanograms of cDNA were used in each PCR using Platinum Taq DNA polymerase (Invitrogen). PCR products were resolved in 4.5% Nusieve 3:1 agarose (Lonza) gels and visualized in a Molecular Imager Gel Doc XR system (Bio-Rad). Sequences of primers used in the RT-PCR experiments are listed in Table S1. RT-qPCR experiments were performed on a 7900HT Fast Real-time PCR system (ABI), following the manufacturer's recommended conditions. mRNA levels were measure using SYBR Green and Taqman Gene Expression Assays (ABI) (Tables S1 and S2), and snRNA levels were measured following procedures previously described (6).

1. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
2. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621–628.
3. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
4. Quinlan AR, Hall IM (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.
5. Katz Y, Wang ET, Airoldi EM, Burge CB (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* 7(12):1009–1015.
6. Zhang Z, et al. (2008) SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* 133(4):585–600.

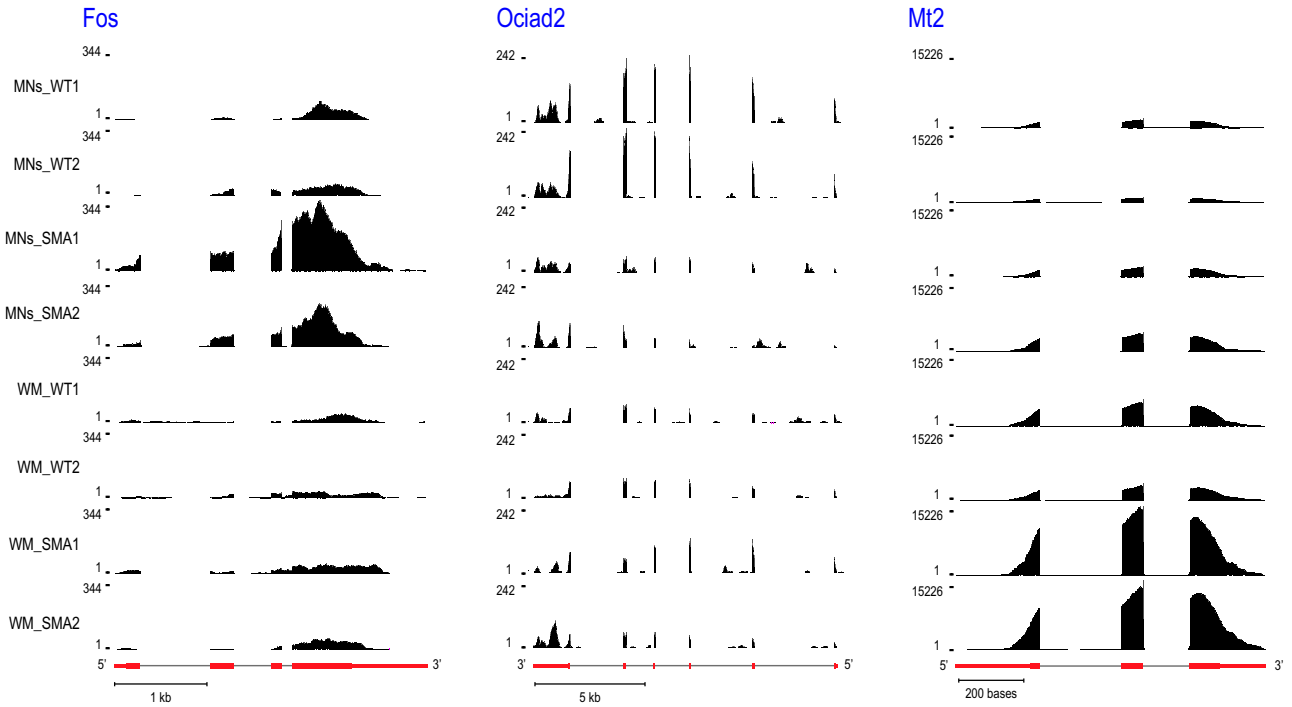

Fig. S1. Genome browser visuals of genes affected by differential expression in cell type-specific manner. UCSC Genome Browser view of Fos, Ociad2, and Mt2 displaying RNA-seq mapped reads profile, showing MN-specific up-regulation of Fos and down-regulation of Ociad2, and WM-specific up-regulation of Mt2 mRNA levels.

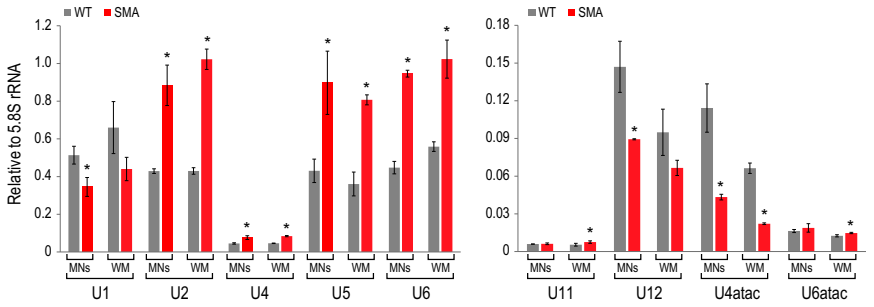

Fig. S2. snRNA levels are affected in SMA MNs and WM at postnatal day (P)3. Total RNA from MNs ($n = 3$) and WM ($n = 3$) of P3 SMA mice were purified after laser-capture microdissection (LCM) procedures. snRNA levels in each sample were measured by RT-qPCR. RNA input was normalized to 5.8S rRNA level. Error bars indicate SD (*$P < 0.05$, Student t test).

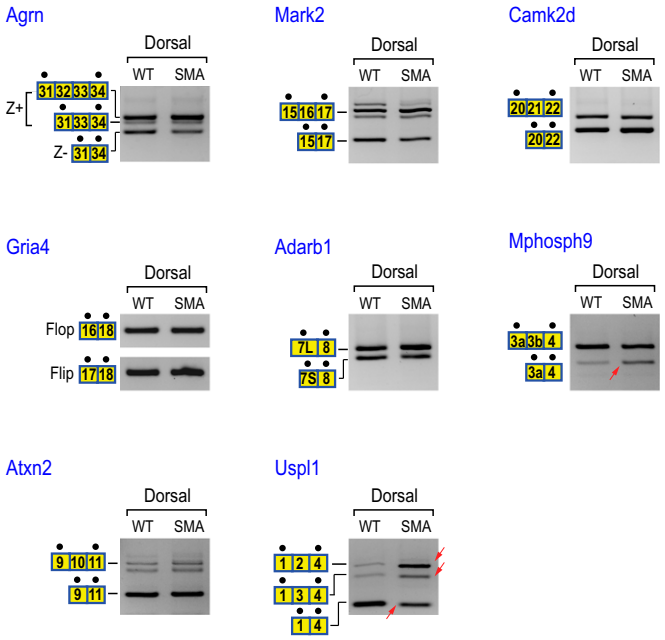

Fig. S3. Splicing events detected in SMA MNs and WM are cell-type specific. RT-PCR reactions were performed as described in Fig. 3*A*, using total RNA purified from the dorsal region (Dorsal) of a WT and an SMA mouse littermate at P1. Most of the splicing changes identified in MNs or WM are not detected in Dorsal, except Mphosph9 and Uspl1.

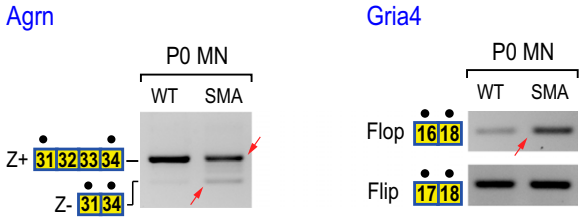

Fig. S4. Splicing abnormalities in Agrn and Gria4 are also found in P0 SMA MNs. RT-PCR reactions were performed as described in Fig. 3*A*, using total RNA purified from the MNs of a WT and an SMA mouse littermate. Splicing changes in Agrn and Gria4 similar to those observed at P1 were detected at P0.

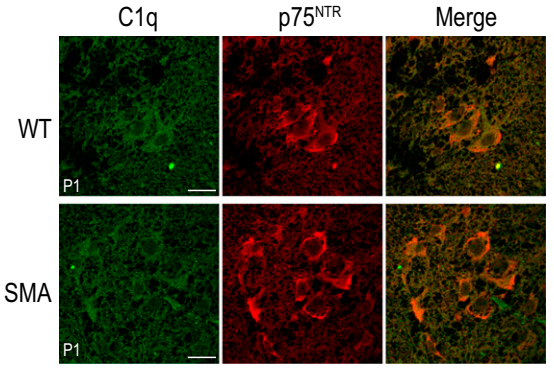

Fig. S5. Immunofluorescence staining of C1q did not detect protein level difference between WT and SMA MNs at P1. Immunofluorescence staining of C1q protein complex was performed as described in Fig. 3*B*; similar levels of weak C1q staining was detected in P1 WT and SMA MNs. (Scale bars, 25 μm.)

Table S1. Sequences of primers used in RT-PCR and SYBR green RT-qPCR experiments

Primer name	Sequence (5′–3′)
Gapdh LCM F	AAGGGCTCATGACCACAGTC
Gapdh LCM R	ACACATTGGGGGTAGGAACA
Chat LCM F	TGCAACACCTGGTACCTGAA
Chat LCM R	GCAGGGCTAGAGTTGACTGG
Gfap LCM F	TCCTTCCAAGGTTGTCCATC
Gfap LCM R	CCAATCAGCCTCAGAGAAGG
$p75^{NTR}$ LCM F	TGCAAAGCCTGCAACC
$p75^{NTR}$ LCM R	CATAGGAGCATCGGCACA
Agrn ex31 F	TGTCCTGGGGGCTTCTCTGG
Agrn ex34 R	CAACCTTTCCAATCCACAGCACC
Agrn ex33 R	CGGGAATCCAGAGTTTCGGG
Mark2 ex15 F	CACAAACCGAAGCAGGAAC
Mark2 ex17 R	TGCTTCGACTGGACACAC
Camk2d ex20 F	GCGGGATGCCAAAGAC
Camk2d ex22 R	ATTACACGTAGAACCTTCACAAC
Gria4 flop ex16 F	CGGGGGAGGTGACTC
Gria4 flip ex17 F	CCCAAGGACTCGGGAAG
Gria4 ex18 R	CTTCATTCTCTTCGCCTCTG
Dusp22 ex3 F	CAAGAGATGCAGAACAGTTGA
Dusp22 ex4 R	GGTTTTGAGATGGTGTGTCTG
Atxn2 ex9 F	TCAAGAGCTGCTTCTCACA
Atxn2 ex11 R	AGGAGCAGCTGCTTCAC
Uspl1 ex1 F	GGAGTTCGGGTCCACTG
Uspl1 ex4 R	CTTGCCTTTCGCTCTACAAG
Adarb1 ex7 F	GGTTCCGGCTGAAGGA
Adarb1 ex8 R	CCTCGCCAGACTCTATTTTC
Mphosph9 ex3a F	GGGCCACCTGTCATTCA
Mphosph9 ex4 R	TCTCAGAAAAGAAGCCCATTTG
Gria4 Flip qPCR F	GGTGAATGTGGACCCAAGGA
Gria4 Flip qPCR R	GCTACGTTGCTCAGGCTCAAG
Gria4 Flop qPCR F	GGAGGTGACTCCAAGGACAAGA
Gria4 Flop qPCR R	CCGCCAACCAGAATGTAGAAG

Table S2. Taqman gene expression assays used in the RT-qPCR experiments

Gene symbol	Taqman gene expression assay ID
Gapdh	4352932E
Etv1	Mm00514804_m1
Igf1	Mm00439560_m1
Igf2	Mm00439564_m1
C1qb	Mm01179619_m1

Dataset S1. RNA-seq reads mapping results summary

Dataset S1

Dataset S2. Differential expressed gene lists

Dataset S2

Differentially expressed genes identified by RNA-seq analysis are listed on separate work sheets, showing up-regulated and down-regulated genes in MNs or WM, and the genes affected in both cell types. Genes on each work sheet are ranked by fold-change.

Dataset S3. Splicing events identified in MNs

Dataset S3

Genes affected by splicing abnormalities in MNs are listed, representing seven categories as illustrated in Fig. 2*B*.

Dataset S4. Splicing events identified in WM

Dataset S4

Genes affected by splicing abnormalities in WM are listed, representing seven categories as illustrated in Fig. 2*B*.

Dataset S5. Overlapping between genes affected by differential expression and alternative splicing

Dataset S5

Genes affected by both differential expression (DE) and alternative splicing (AS) are listed on separate work sheets, showing overlap between DE and AS in MNs and WM on the first 4 work sheets. The fifth work sheet shows overlap of genes affected by AS in MNs and WM.

Dataset S6. Splicing events identified in genes containing U12 introns

Dataset S6

Genes affected by splicing are searched against U12DB, and genomic coordinates of the U12 introns and the regions affected by splicing are also listed. The genes affected in MNs and WM are listed on the first and second work sheet, respectively. None of the regions affected by splicing overlap with locations of the U12 introns.