Figure S1

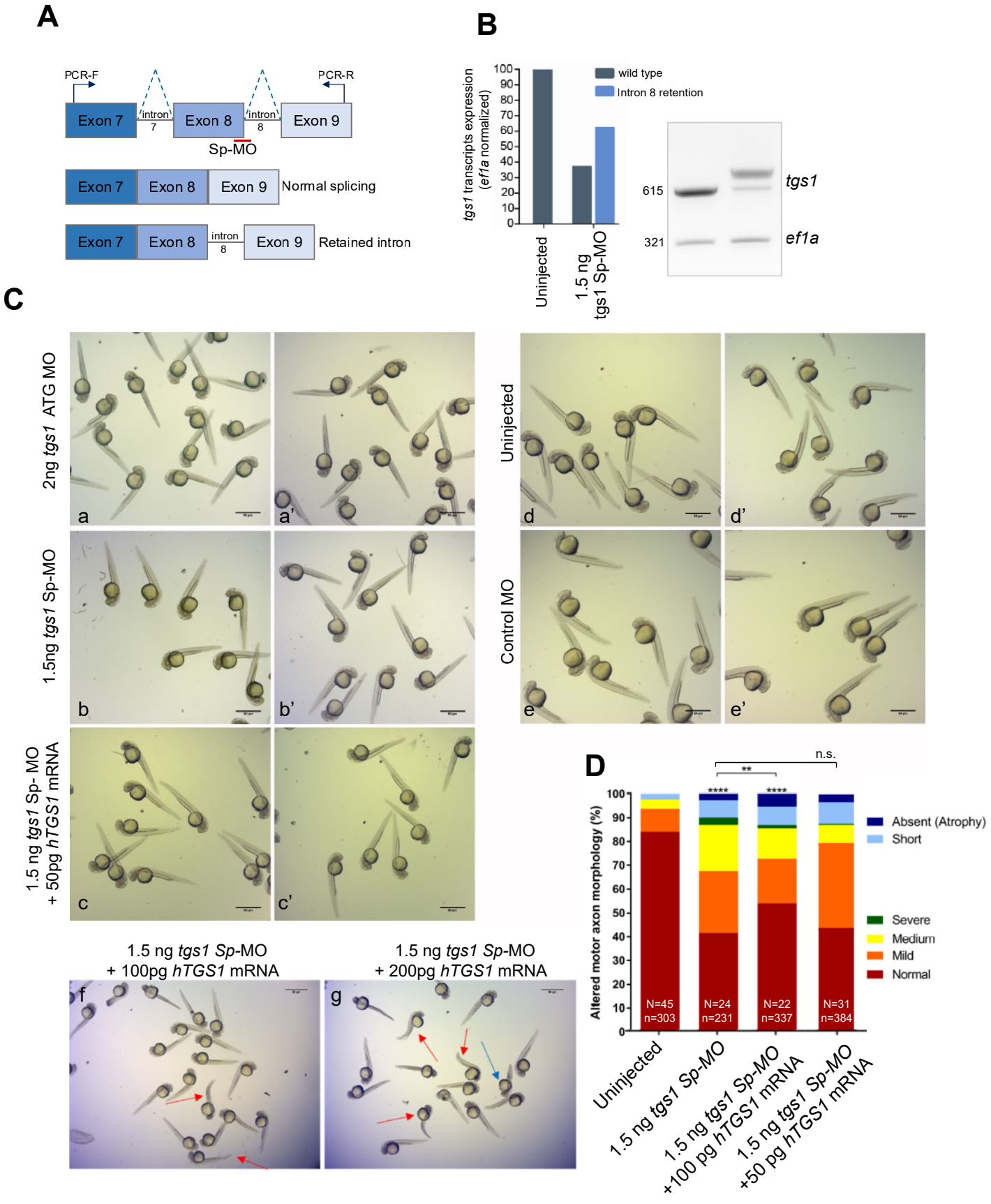

A
PCR-F
PCR-R
Exon 7
intron 7
Exon 8
intron 8
Exon 9
Sp-MO
Exon 7
Exon 8
Exon 9
Normal splicing
Exon 7
Exon 8
intron 8
Exon 9
Retained intron
B
tgs1 transcripts expression (ef1a normalized)
wild type
Intron 8 retention
Uninjected
1.5 ng tgs1 Sp-MO
615
321
tgs1
ef1a
C
2ng tgs1 ATG MO
1.5ng tgs1 Sp-MO
1.5 ng tgs1 Sp- MO + 50pg hTGS1 mRNA
Uninjected
Control MO
a
a'
b
b'
c
c'
d
d'
e
e'
1.5 ng tgs1 Sp-MO + 100pg hTGS1 mRNA
1.5 ng tgs1 Sp-MO + 200pg hTGS1 mRNA
f
g
D
Altered motor axon morphology (%)
n.s.
**
****
****
Absent (Atrophy)
Short
Severe
Medium
Mild
Normal
N=45 n=303
N=24 n=231
N=22 n=337
N=31 n=384
Uninjected
1.5 ng tgs1 Sp-MO
1.5 ng tgs1 Sp-MO +100 pg hTGS1 mRNA
1.5 ng tgs1 Sp-MO +50 pg hTGS1 mRNA

Figure S2

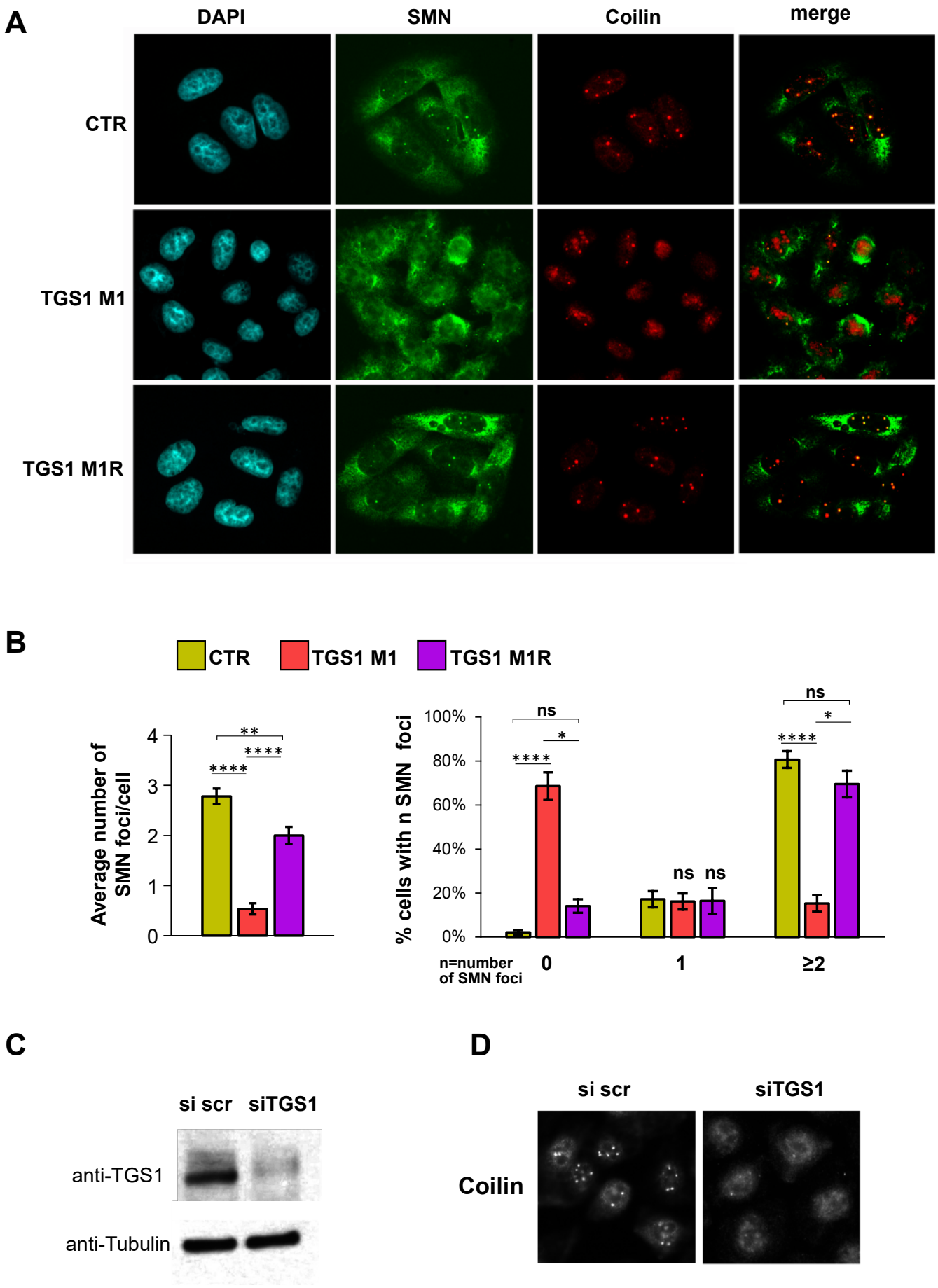
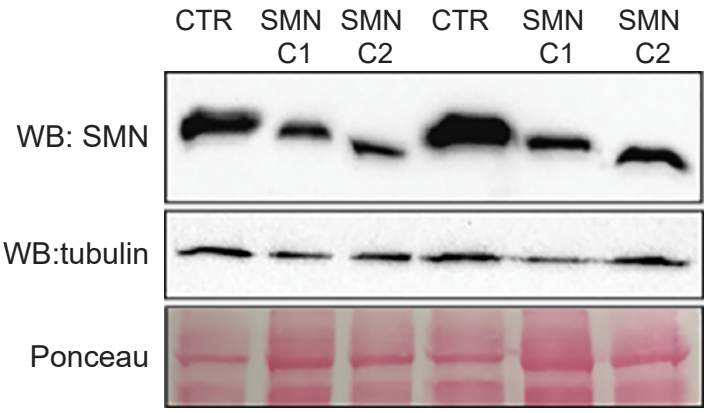

A
DAPI
SMN
Coilin
merge
CTR
TGS1 M1
TGS1 M1R
B
CTR
TGS1 M1
TGS1 M1R
Average number of SMN foci/cell
% cells with n SMN foci
n=number of SMN foci
0
1
≥2
ns
C
si scr
siTGS1
anti-TGS1
anti-Tubulin
D
si scr
siTGS1
Coilin

Figure S3

A



SMN CRISPR clone C1	p. [Ser230del];[Pro226del; Ser230del]
SMN CRISPR clone C2	p.[C231fs*8];[Pro226del; Leu228fs*12]

B

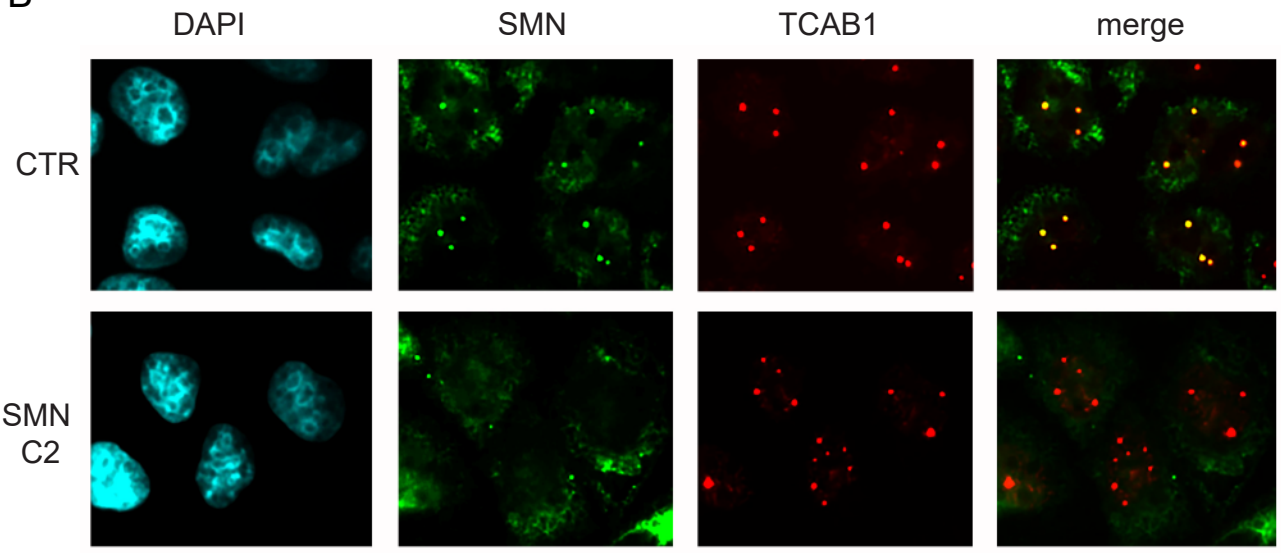
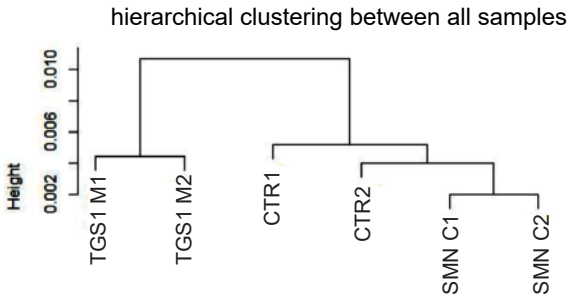
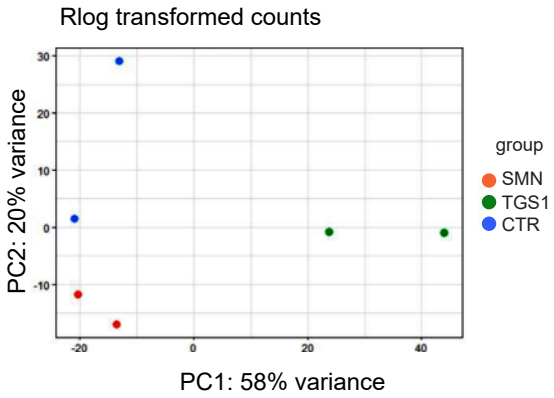


Figure S4

A



C



B

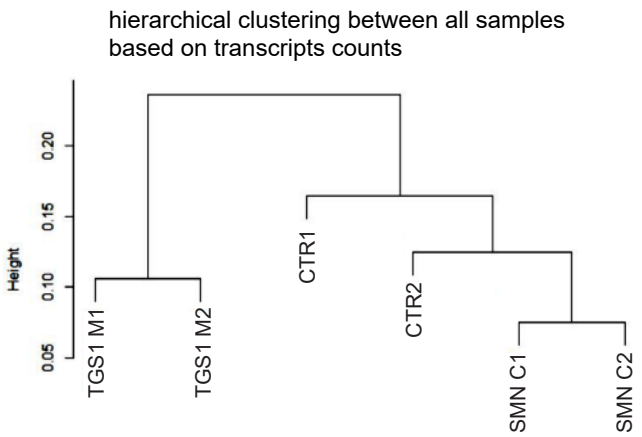
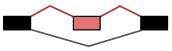
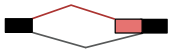
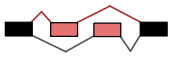
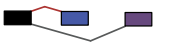


Figure S5

A

UP ↑ DOWN ↓		number of events		SMN		TGS1	
		Annotated	Novel	Annotated	Novel	Annotated	Novel
	Intron retention	8,648	3,735	↑ 56 (0.65%)	34 (0.91%)	199 (2.30%)	82 (2.20%)
				↓ 44 (0.51%)	45 (1.20%)	100 (1.16%)	88 (2.36%)
	Exon skipping	50,665	972	↑ 143 (0.28%)	6 (0.62%)	328 (0.65%)	22 (2.26%)
				↓ 112 (0.22%)	5 (0.51%)	261 (0.52%)	11 (1.13%)
	Alternative 5' SS	19,132	868	↑ 80 (0.42%)	5 (0.58%)	182 (0.95%)	11 (1.27%)
				↓ 91 (0.48%)	6 (0.69%)	146 (0.76%)	19 (2.19%)
	Alternative 3' SS	21,248	919	↑ 85 (0.40%)	2 (0.22%)	251 (1.18%)	21 (2.29%)
				↓ 100 (0.47%)	7 (0.76%)	135 (0.64%)	13 (1.41%)
	Mutually exclusive exons	8,272	82	↑ 23 (0.28%)	0 (0.00%)	48 (0.58%)	0 (0.00%)
				↓ 39 (0.47%)	0 (0.00%)	66 (0.80%)	0 (0.00%)
	Alternative first exon	120,379	311	↑ 98 (0.08%)	4 (1.29%)	158 (0.13%)	9 (2.89%)
				↓ 92 (0.08%)	3 (0.96%)	207 (0.17%)	6 (1.93%)
	Alternative last exon	43,694	92	↑ 80 (0.18%)	0 (0.00%)	121 (0.28%)	0 (0.00%)
				↓ 68 (0.16%)	0 (0.00%)	98 (0.22%)	3 (3.26%)
Total number of events		279,017		1228 (0.44%)		2585 (0.93%)	

B

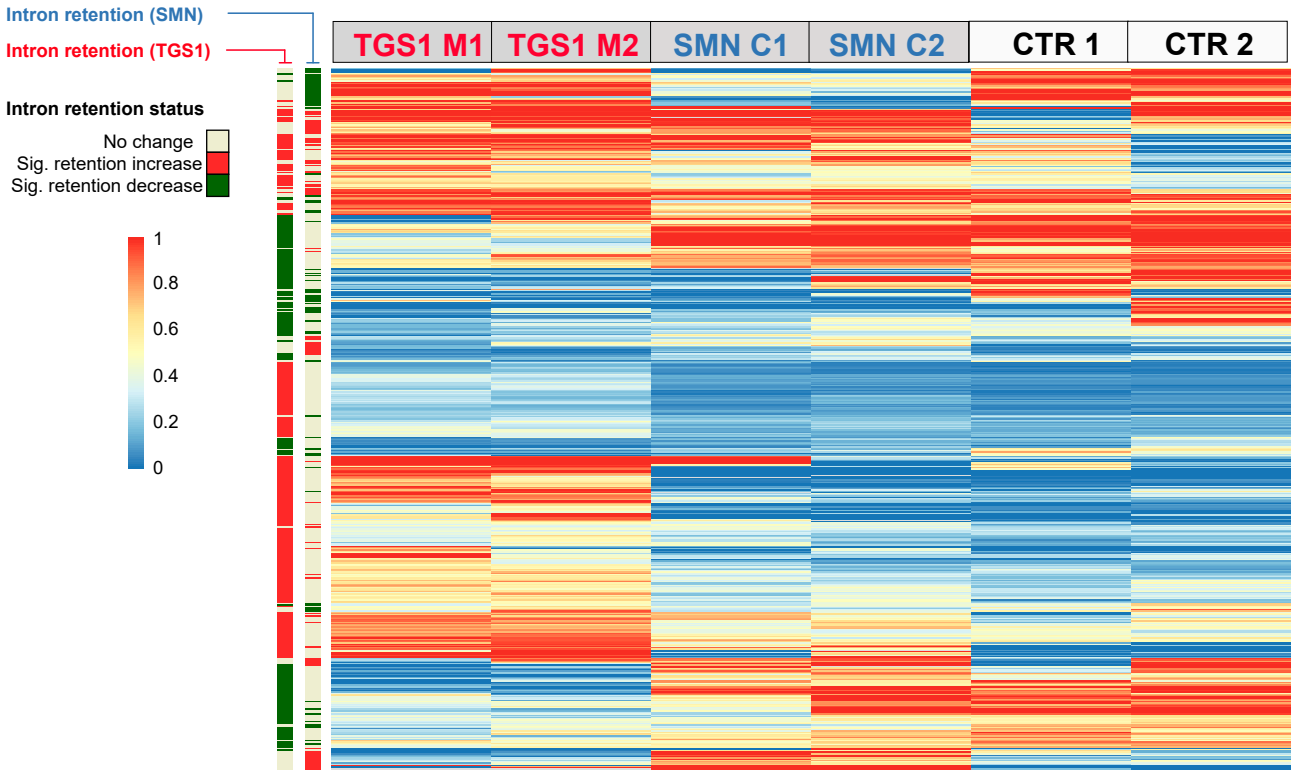


Figure S6

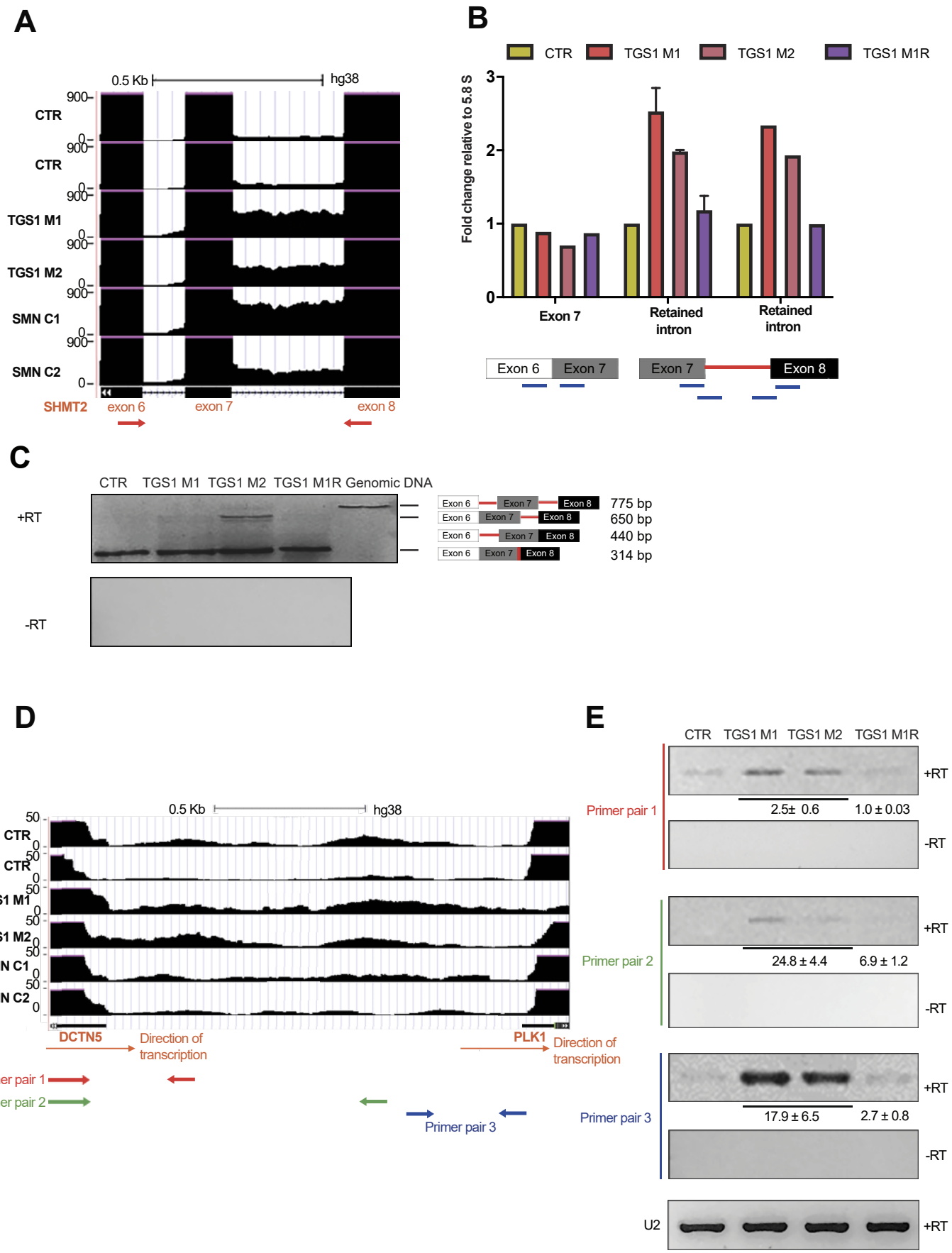
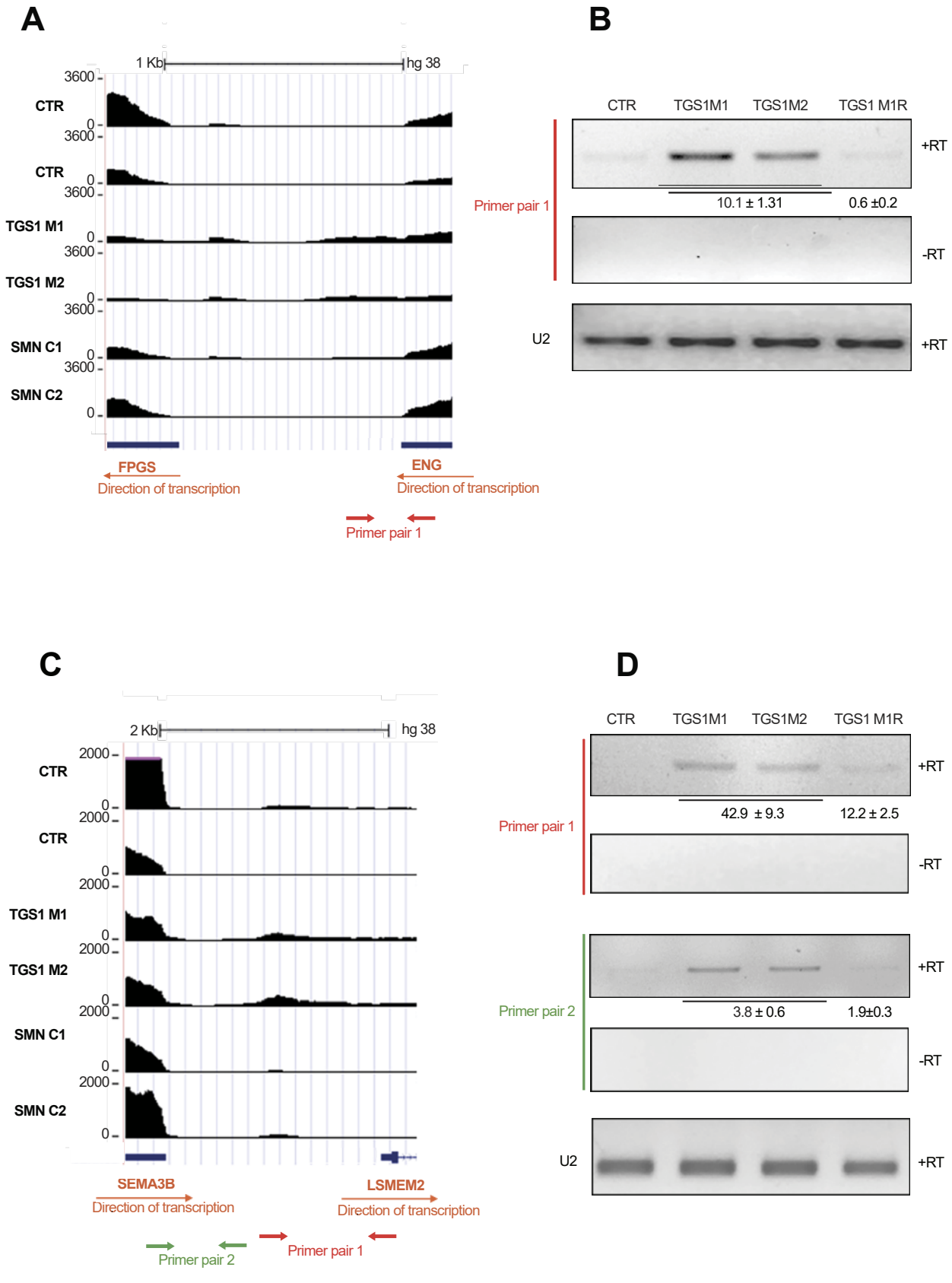

A
0.5 Kb
hg38
CTR
CTR
TGS1 M1
TGS1 M2
SMN C1
SMN C2
SHMT2
exon 6
exon 7
exon 8
B
CTR
TGS1 M1
TGS1 M2
TGS1 M1R
Fold change relative to 5.8 S
Exon 7
Retained intron
Retained intron
Exon 6
Exon 7
Exon 7
Exon 8
C
CTR
TGS1 M1
TGS1 M2
TGS1 M1R
Genomic DNA
+RT
-RT
775 bp
650 bp
440 bp
314 bp
D
0.5 Kb
hg38
CTR
CTR
TGS1 M1
TGS1 M2
SMN C1
SMN C2
DCTN5
Direction of transcription
PLK1
Direction of transcription
Primer pair 1
Primer pair 2
Primer pair 3
E
CTR
TGS1 M1
TGS1 M2
TGS1 M1R
Primer pair 1
+RT
-RT
2.5± 0.6
1.0 ± 0.03
Primer pair 2
+RT
-RT
24.8 ± 4.4
6.9 ± 1.2
Primer pair 3
+RT
-RT
17.9 ± 6.5
2.7 ± 0.8
U2
+RT

Figure S7


A
1 Kb
hg 38
3600
CTR
0
3600
CTR
0
3600
TGS1 M1
0
3600
TGS1 M2
0
3600
SMN C1
0
3600
SMN C2
0
FPGS
Direction of transcription
ENG
Direction of transcription
Primer pair 1
B
CTR
TGS1M1
TGS1M2
TGS1 M1R
Primer pair 1
+RT
10.1 ± 1.31
0.6 ±0.2
-RT
U2
+RT
C
2 Kb
hg 38
2000
CTR
0
2000
CTR
0
2000
TGS1 M1
0
2000
TGS1 M2
0
2000
SMN C1
0
2000
SMN C2
0
SEMA3B
Direction of transcription
LSMEM2
Direction of transcription
Primer pair 2
Primer pair 1
D
CTR
TGS1M1
TGS1M2
TGS1 M1R
Primer pair 1
+RT
42.9 ± 9.3
12.2 ± 2.5
-RT
Primer pair 2
+RT
3.8 ± 0.6
1.9±0.3
-RT
U2
+RT

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Injection of a *tgs1* Sp-MO induces intron 8 retention and causes CaP-MN defects

(A) Diagram showing *tgs1* gene exon/introns structure. The location of the *tgs1* Sp-MO at the intron-exon boundary is indicated with a red line. Dotted line indicates normal splicing of the introns 7 and 8. Arrows: primers used for RT-PCR to verify that the *tgs1* Sp-MO injection induces accumulation of mis-spliced tgs1 transcripts retaining intron 8.

(B) Representative agarose gel showing *tgs1* mRNA products as measured by semi-quantitative RT-PCR in uninjected (control) and *tgs1* Sp-MO injected embryos. As compared with control fish, injection of the *tgs1* Sp-MO reduced the amount of correctly spliced transcripts (represented by the wildtype amplicon of 615 bp) and induced an intron 8-retained splice transcript of 750 bp. This transcript contains a stop codon upstream the sequence encoding the Tgs1 catalytic domain and is therefore predicted to produce a non-functional protein. Densitometric analysis of *tgs1* transcripts normalized against the *ef1a* housekeeping gene shows that *tgs1 Sp*-MO reduces WT *tgs1* transcripts by approximately 65%.

(C) Morphology of zebrafish larvae after *tgs1* MOs and *TGS1* mRNA injection. Bright field images of fish larvae injected with *tgs1* MOs (alone or in combination with an RNA encoding FLAG-tagged human TGS1 (*hTGS1*) and control groups (control MO and uninjected). Gross morphological defects are not observed upon injection of 2 ng *tgs1* ATG MO (panels a, a'), 1.5 ng *tgs1* Sp-MO (alone or in combination with 50 pg *TGS1* mRNA; panels b, b', c, c'), and 1.5 ng control-non targeting – MO (panels d, d') or in uninjected animals (e, e'). In contrast, injection of 1.5 ng *tgs1* Sp-MO in combination with 100 pg or 200 pg h*TGS1* RNA resulted in cardiac edema, tail bending and body axis malformations (arrows in panels f, g). Scale bar: 50 μm.

(D) Based on overall appearance, CaP-MNs were classified as described in Fig. 4B. Zebrafish larvae were injected with 1.5 ng of *tgs1 Sp*-MO alone or in combination with 100 pg or 50 pg h*TGS1* RNA. Results are presented in percentages from 3 independent

experiments (n = axons analysed; N= animals tested. ****, p<0.0001 Chi-square test); n.s. not significant. The capacity of 100 pg h*TGS1* RNA to rescue the neuronal defects induced by injection of *tgs1 Sp*-MO cannot be unequivocally assessed, as larvae injected with this dosage of h*TGS1* RNA display developmental abnormalities (shown in image f of panel C).

Figure S2. *TGS1* knockdown disrupts Cajal body organization in HeLa and UMUC3 cells

(A) Cajal body organization in *TGS1*-proficient (CTR), *TGS1 M1* and*TGS1* M1 expressing *FLAG-TGS1* (*TGS1 M1R*) HeLa cells . Cells were stained with anti-SMN and anti-Coilin antibodies and with DAPI (DNA). SMN accumulates in the CBs and colocalizes with the CB marker Coilin in 99% of control cells. Depletion of TGS1 leads to a reduction in the coilin-rich CBs, loss of SMN nuclear localization, and ectopic SMN granules in the cytoplasm. These defects are ameliorated by stable expression of a *FLAG-TGS1* transgene (M1R).

(B) Quantification of the number of SMN/Coilin positive foci in the nuclei of cells with the indicated genotypes. At least 200 cells counted per sample. Error bars, S.E.M. *, **, **** p < 0.05, 0.01 and 0.0001, respectively; ns, not significant, two-way ANOVA).
(C) Western blotting showing TGS1 abundance in UMUC3 cells treated with *TGS1* siRNA or non-targeting dsRNA (si scr) for 6 days. Tubulin is a loading control.
(**D)** UMUC3 cells treated with *TGS1* siRNA or non-targeting dsRNA (si scr) stained with anti-Coilin antibodies and DAPI. Note that *TGS1* knockdown disrupts CB organization.

Figure S3. Characterization of CRISPR-induced *SMN* mutant Hela cells
(A) SMN expression in two independent CRISPR-derived clones of HeLa cells (C1, C2) revealed by Western blotting (WB). CTR, parental cell line; tubulin is a loading control. Note that the *SMN* frameshift mutations in the C2 mutant clone result in a truncated protein.

(B) Cajal body organization in *SMN*-proficient (CTR) and *SMN C2* mutant HeLa cells. Cells were stained with anti-SMN and anti-TCAB1 antibodies, and with DAPI (DNA). SMN accumulates in the CBs and colocalizes with the CB marker TCAB1 in 99% of control cells. The *SMN* mutation present in the C2 mutant clone abrogates nuclear localization of SMN, while the TCAB1-rich foci were unaffected.

Figure S4. Hierarchical clustering and PCA analysis of RNA samples from *CTR*, *TGS1* and *SMN* mutant cells

(A, B) Hierarchical clustering and principal component analysis based on gene expression (A, C) and transcript expression (B) show similarity among samples within the same group.

Figure S5. Patterns of alternative splicing in *TGS1* and *SMN* mutant cells

(A) Number of significantly changed alternative splicing (AS) events across 7 major AS types. Annotated AS events are extracted from the GENCODE annotation and novel AS events are obtained from the assembled transcriptome by Scallop(50). The column "number of events" indicates the total number of alternative splicing events in each category and "*SMN*" and "*TGS1*" columns show the number of differentially spliced events in each mutant category, where the first and second rows in each category show the number of events with increased and decreased alternative isoforms, respectively.

(B) Heatmap of the Intron Retention (IR) events in *SMN* or *TGS1* mutant cells as annotated by two sidebars to the left (increased IR marked in Magenta and decreased in Green). Heatmap shows the PSI value of a given IR event across all six RNA-seq datasets.

Figure S6. Validation of aberrant transcript structures from *in silico* analyses in *TGS1* mutants

(A-C): Detection of intron 7 retention in the *SHMT2* gene.

(A) Diagram from USCS Genome browser showing the reads mapping to intron 7 in *TGS1* M1, *TGS1* M2 and control (CTR) cells. Transcripts containing this intron are also enriched in *SMN* mutant cells. See text for definition of HeLa cell genotypes.

(B) RTq-PCR comparative quantification of the retention level of intron 7 assayed using two different primer pairs (blue lines shown below the histograms) and the abundance of the adjacent exon 7. RTq-PCR was performed on cDNA samples from Control (CTR), *TGS1* mutant (*TGS1 M1, TGS1 M2*), and *TGS1 M1* rescued (*TGS1 M1R*, expressing TGS1-FLAG) HeLa cell lines. Bars represent means between three replicates, are relative to the parental cell line (CTR) and are normalized to 5.8S RNA.

(C) RT-qPCR analysis performed on cDNA samples from Control (CTR), TGS1 mutant (*TGS1 M1, TGS1 M2*), and *TGS1 M1* rescued (*TGS1 M1R*, expressing TGS1-FLAG) HeLa cell lines to detect the retention of intron 7. The positions of the primer pairs used for RT-PCR spanning exons 6 to 8 are reported below panel A (red arrows). In negative control samples (-RT), the reverse transcriptase was omitted from the reaction to rule out genomic DNA contamination. Genomic DNA amplified by PCR is a control. The numbers below the lanes show densitometric quantitation of the amplification products relative to the U2 amplification product (+/- SEM, three replicates).

(D-E). Validation of transcriptional events spanning the intergenic region between the *DCTN5* and *PLK1* genes.

(D) Diagram from USCS Genome browser shows reads mapping to the region of interest in the sequenced samples. The blue, red and green arrows indicate the he positions of the primer pairs designed to perform RT-PCR analysis shown in panel E.

(E) RT-PCR analysis of the genomic region indicated in panel D. U2 is a positive control.

Figure S7. Validation of transcriptional events spanning intergenic regions

(A, B) The densities of reads in the *SEMA3B-LSMEM2* region are reported in the diagrams from the UCSC Genome browser. The positions of the primer pairs designed

to perform RT-PCR analyses (red arrows) are indicated below the diagram (A). (B) RT-PCR analysis; densitometric quantitation of the amplification products is relative to the U2 amplification product as in Figure S6E. – RT indicates absence of reverse transcriptase; see text for definition of HeLa cell genotypes.

(C, D) The Densities of reads in the *FPGS-ENG* region are reported in the diagrams from the UCSC Genome browser. The positions of the primer pairs designed to perform RT-PCR analyses (red and green arrows) are indicated below the diagram (A). (B) RT-PCR analysis; densitometric quantitation of the amplification products is relative to the U2 amplification product as in Figure S6E.

SUPPLEMENTARY TABLES AVAILABLE AT:

https://figshare.com/s/13572f55424106336cac

SUPPLEMENTARY TABLES DESCRIPTION

Table S1. Differentially expressed transcripts in TGS1 and SMN HeLa mutant cells compared to control cells (pipeline 1)

List of the transcripts differentially expressed in TGS1 (sig_transcripts_WT_vs_TGS1) or SMN (sig_transcripts_WT_vs_SMN) mutant cells compared to control cells (WT), identified and quantified using Kallisto + Sleuth. Annotated and unannotated transcript isoforms are denoted by ENST- and gene- prefixes, respectively. Transcripts with Beta values less than -2 or greater than 2 and with q values < 0.05 were considered to be significantly differentially expressed (positive values: downregulated; negative values: upregulated)

Table S2. Classification of the differentially expressed, alternatively spliced transcripts.

Table S3. List of the transcripts differentially expressed in TGS1 or SMN mutant cells compared to control cells (pipeline 2).

List of the transcripts differentially expressed in TGS1 (TGS1 vs CTR filter) or SMN (SMN vs CTR filter) mutant cells compared to control cells (WT), identified and quantified using the combination of the Salmon, Scallop and Deseq2 softwares. Annotated and unannotated transcript isoforms are denoted by ENST- and gene- prefixes, respectively. Transcripts with $\log_2$ fold change values less than -2 or greater than 2 and with adjusted p values < 0.05 were considered to be significantly differentially expressed. Sheet *locations*: genomic coordinates for unannotated transcripts are reported.

Table S4. Summary of the total number of DE transcripts identified using two independent quantification pipelines

Table S5. Classification of novel transcripts

The unannotated transcripts were classified according to the length of their 5' or 3' UTR and for the presence of retained introns. For those transcripts carrying unspliced introns (labeled as TRUE), the genomic location of the retained sequence is indicated

Table S6. List of the oligonucleotides used in this study