DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Short Communication

The MBNL/CELF Splicing Factors Regulate Cytosolic Sulfotransferase 4A1 protein expression During Cell Differentiation.

Misgana Idris, Neville J. Butcher and Rodney F. Minchin

Laboratory for Molecular and Cellular Pharmacology, School of Biomedical Sciences, University of Queensland, Brisbane, Queensland, Australia, 4072

Running Title: Role of MBNL and CELF in SULT4A1 splicing

Corresponding author: Prof Rod Minchin, School of Biomedical Sciences, University of Queensland, Brisbane, QLD, Australia 4072. Email: r.minchin@uq.edu.au

Telephone: 61-7-3346-1620

Number of text pages: 17

Number of tables: 0

Number of figures: 4

Number of references: 31

Number of words – Abstract: 240

Number of words – Introduction: 472

Number of words – Results and Discussion: 1613

Abbreviations: SULT4A1, sulfotransferase 4A1; WT, wild-type; hiPSC, human induced pluripotent stem cells, 6p pseudo-exon 6; RA, retinoic acid; RIPA, radioimmunoprecipitation assay buffer.

DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

ABSTRACT

Sulfotransferase 4A1 (SULT4A1) is a sulfotransferase-like protein that is highly conserved between species. In human tissues, there are 2 transcripts, one that produces a full length protein and one that produces an unstable truncated protein. The second transcript, which includes a pseudo-exon between exons 6 and 7 (6p), is widely expressed while the first is more restricted. Differentiation of neuronal cells results in the removal of the pseudo-exon and subsequent SULT4A1 protein expression. Recent studies with SULT4A1 knockout mice showed that the protein is essential for normal development and that its absence leads to a severe neurological phenotype. Here, the regulation of SULT4A1 6p splicing was investigated during neuronal differentiation using SH-SY5Y cells, human induced pluripotent stem cells and mouse embryonic tissue. In all 3 models, pseudo-exon 6p was removed during differentiation, resulting in stable SULT4A1 protein expression. Using a minigene splicing assay, a region upstream of pseudo-exon 6p was identified that is essential for correct splicing of SULT4A1 mRNA. Within this region, there were binding motifs for 4 RNA processing factors (MBNL-1, MBNL-2, CELF-1 and CELF-2). Time-dependent changes in SULT4A1 protein and MBNL/CELF protein during differentiation supported their role in correctly splicing the SULT4A1 mRNA. Furthermore, ectopic expression of each factor produced efficient splicing in the minigene assay as well as correct splicing of the endogenous SULT4A1 mRNA. These results show that SULT4A1 mRNA is a target for MBNL/CELF-dependent splicing, which may be essential in producing stable, functional SULT4A1.

Introduction

Sulfotransferase 4A1 (SULT4A1) is a cytosolic sulfotransferase expressed primarily in neuronal cells. It is highly conserved between species but its biological function remains elusive. SULT4A1 polymorphisms have been associated with neurological diseases such as schizophrenia (Brennan and Condra, 2005; Meltzer et al., 2008) and there has been conflicting

evidence suggesting SULT4A1 may be a biomarker for response to antipsychotic drugs such as olanzapine (Ramsey et al., 2014; Wang et al., 2014). Two recent studies have highlighted the physiological importance of SULT4A1. Firstly, Falany's group successfully deleted the gene in a murine model and observed a severe phenotype that included tremors, ataxia, abnormal gait and absence seizures (Garcia et al., 2018). This important work suggests that SULT4A1 is fundamental to normal neuronal development. Secondly, two reports have independently shown that SULT4A1 protein expression is dependent on alternative splicing of its transcript (Sidharthan et al., 2014; Hashiguchi et al., 2018). While the SULT4A1 protein is mostly confined to neuronal tissue, its mRNA is more widely expressed (Falany et al., 2000; Sidharthan et al., 2014). Two different transcripts have been reported, one that includes a pseudo-exon between exons 6 and 7 (6p), termed variant transcript, and one where 6p has been removed, termed wild-type (WT) transcript. The variant transcript introduces a premature stop codon resulting in an unstable, truncated protein (Sidharthan et al., 2014). In human tissues, SULT4A1 protein expression correlates well with the WT transcript consistent with the notion that the variant transcript does not produce measurable protein levels (Sidharthan et al., 2014). Thus, alternative splicing appears to be an important mechanism for regulating SULT4A1 expression.

In both neuroblastoma SH-SY5Y and SK-N-MC cells, the switch from the variant to the WT transcript is seen during cellular differentiation (Sidharthan et al., 2014). This alternative splicing is also observed over time in cultured neuron-glia mixed cells (Hashiguchi et al., 2018). Alternative splicing of RNA results in multiple protein isoforms from a single gene. Of the ~20,000 protein-coding genes in human cells, more than 90% generate multiple transcripts, providing greater diversity to the human genome (Scotti and Swanson, 2015). Splicing is regulated by RNA processing factors that recognise explicit RNA binding motifs and direct the splicing machinery to specific splice sites. These factors are expressed in a cell-specific and a temporal-dependent manner to regulate protein function. During embryonic

development, alternative splicing is central to cell reprogramming. Thus, it is not surprising that errors in splicing are associated with many human diseases such as Duchenne muscular dystrophy, amyotrophic lateral sclerosis, dilated cardiomyopathy and frontotemporal dementia (Scotti and Swanson, 2015).

In the present study, the regulation of SULT4A1 splicing was investigated to identify potential splicing factors responsible for the maturation of SULT4A1 mRNA. For this, several models of cell differentiation were used including retinoic acid (RA)-treated neuroblastoma cells, differentiated human induced pluripotent stem cells (hiPSCs) and mouse embryonic tissue.

Materials and Methods

Cell culture and differentiation. SH-SY5Y cells were grown in advanced DMEM/F12 medium containing 10% fetal calf serum, 4 mM glutamine and 1% penicillin/streptomycin in a humidified atmosphere of 95% air, 5% CO_2 at 37ºC. The cells were differentiated in the same medium with 10 μM *trans*-retinoic acid (Sigma, St. Louis, MI, USA) as previously described where differentiation was confirmed morphologically and by the up-regulation of NeuN expression (Encinas et al., 2000; Sidharthan et al., 2014). Human induced-pluripotent stem cells (C11; WiCell, Madison, WI, USA) were cultured in mTeSR/1 (STEMCELL Technologies, VIC, Australia) /Matrigel (BD Biosciences, NSW, Australia) as previously described (Ovchinnikov et al., 2015) and differentiated with 10 μM *trans*-retinoic acid for up to 28 days. The neuronal marker NeuN increased with time, as determined by Western blot (Supplementary Fig 1A). Differentiation of the hiPSCs was also evident from the morphological changes where the cells formed spheroids with prominent dendritic outgrowths (Supplementary Fig 1B).

Cell transfections. For minigene splicing assays, SH-SY5Y cells were seeded in 6-well plates at 0.5×10^6 cells/well and allowed to adhere overnight. Cells were then transiently transfected with 2 μg of the minigene plasmids using Lipofectamine2000 (Thermo Fisher

Scientific, VIC, Australia). After 4 h of transfection, cells were treated with 10 μM retinoic acid for 5 days before processing for downstream assays. For experiments with co-expression of splicing factors and minigene constructs, cells were co-transfected with 1μg of minigene plasmid and 3μg of FLAG-tagged MBNL or CELF expressing plasmids (kind gifts from Thomas Cooper, Baylor College). To create non-clonal SH-SY5Y cell lines that stably express either MBNL or CELF proteins, FLAG-tagged MBNL or CELF plasmids were first linearized by digestion with the endonuclease enzyme PciI, gel purified and transfected into SH-SY5Y cells using Lipofectamine2000. The transfected cells were then subcultured four times under antibiotic selection before the stable expression of the splicing factor proteins was confirmed by Western blot.

Whole mouse fetal samples. Female CD1 mice were selected on gestational days 8.5, 10.5, 12.5, 14.5 and 18.5 of pregnancy to be euthanized for analysis of fetal tissues. All procedures were reviewed by and received the approval of the University of Queensland Animal Ethics Office (approval numbers numbers AE03917 and AE02544). Total RNA was isolated using Trizol (Thermo Fisher Scientific, VIC, Australia) and converted to cDNA. PCR for SULT4A1 was performed as described below. Protein samples for E8.5, E12.5 and E14.5 fetal tissues were prepared in radioimmunoprecipitation assay buffer (RIPA; 150 mM NaCl, 1% Triton X-100, 0.1% sodium dodecyl sulfate, 0.5% sodium deoxycholate, 50 mM Tris, pH 8.0,) by brief (3 × 5 sec) sonication on ice. Protein samples (50 μg) were electrophoresed and Western blotted for SULT4A1 as described below.

RT-PCR. Total RNA was extracted using a RNeasy mini kit (Qiagen, VIC, Australia) and 2 μg was reverse transcribed to cDNA using SuperScript II (Thermo Fisher Scientific, VIC, Australia). PCRs were performed to study the pattern of expression of transcripts of SULT4A1, minigenes and various neuronal splicing factors, and β-actin was used as control. All primers used are listed in Supplementary Table 1 along with the PCR conditions for

amplification. The PCR amplicons were then separated on 2% agarose gels containing Hyragreen (1:20,000; ACTgene, Piscataway, NJ, USA) for visualisation.

Minigene cloning and splicing assay. Genomic DNA was isolated from SH-SY5Y cells using Wizard Genomic DNA Purification kit (Promega, NSW, Australia) and the alternative pseudo-exon 6p with its 623 bp 5'-upstream and 821 bp 3'-downstream flanking intronic sequences was amplified using AccuPrime Pfx DNA polymerase (Thermo Fisher Scientific, VIC, Australia) and forward and reverse primers with engineered *Kpn*I and *Hin*dIII restriction sites, respectively (Supplementary Table 2). The PCR conditions used for amplification were: initial denaturation at 95^{o}C for 5 min, followed by 35 cycles of melting at 95^{o}C for 15 s, annealing at 58^{o}C for 30 s and elongation at 68^{o}C for 1.5 min. The PCR product was electrophoresed on a 0.8% agarose gel, purified using PureLink Quick Gel Extraction kit (Thermo Fisher Scientific, VIC, Australia), and then double digested with *Kpn*I and *Hin*dIII high fidelity enzymes in Cutsmart buffer (New England Biolabs, Ipswich, MA, USA) for 2 h. The PCR product was first cloned into p-Bluescript and the *Nde*I restriction site in the intronic region of the insert upstream of pseudo-exon 6p was removed by site directed mutagenesis using GENEART Site-Directed Mutagenesis system (Thermo Fisher Scientific, VIC, Australia) as per the manufacturer's instructions. The insert was then re-amplified from p-Bluescript using forward and reverse primers with engineered *Nde*I restriction sites and the amplified product was digested with *Nde*I (New England Biolabs, Ipswich, MA, USA) and recloned into the same site in pTBNde minigene plasmid, a gift from Franco Pagani (Addgene plasmid #14125) (Pagani et al., 2003). A series of deletion minigene constructs were generated by amplifying the different segments of the parent insert with AccuPrime Pfx DNA polymerase using primers listed in Supplementary Table 2. The sequence of each minigene insert was verified by DNA sequencing prior to use in the minigene splicing assay. Minigene constructs were transiently transfected into SH-SY5Y

cells using Lipofectamine2000 (Thermo Fisher Scientific, VIC, Australia) as described above and the splicing patterns examined by RT-PCR.

Western blot analysis. Cells were washed twice with cold PBS and then lysed in radioimmunoprecipitation buffer (RIPA). The lysates were centrifuged at 14,000 × g for 10 min at 4°C, supernatants were recovered and protein was quantified by the Bradford assay (Bradford, 1976). Cytosolic and nuclear protein fractions were prepared with a cell fractionation kit (Abcam, VIC, Australia). All protein samples (10 μg) were electrophoresed on 12% SDS-PAGE gels, transferred to nitrocellulose membranes, and blocked in 5% skim milk in PBS-T for 1 h. Membranes were incubated overnight at 4°C with primary antibody and then with horseradish peroxidase conjugated-secondary antibody for 1 h. Primary antibodies used in this study were: anti-SULT4A1 (12578-1-AP, ProteinTech Group, Chicago, IL, USA), anti-MBNL-1 (ab108519, Abcam), anti-CELF-1 (ab9549, Abcam), anti-MBNL-2 (sc-136167, Santa Cruz Biotechnology, Dallas, TX, USA), anti-CELF-2 (sc-47731, Santa Cruz Biotechnology), anti β-actin (#3700, Cell Signaling Technology, Danvers, MA, USA), anti-α-tubulin (#3873, Cell Signaling Technology), anti-histone H4 (#2592, Cell Signaling Technology) and anti-FLAG-M2 (A8592, Sigma).

Data analysis. Data were analysed either with Student's *t*-test or one-way ANOVA using GraphPad Prism 7 (GraphPad Software, San Diego, CA) and presented as a mean with standard error of the mean. Densitometry was performed with ImageJ software to quantify the level of expression of proteins normalised to loading controls.

Identification of potential binding domains for trans-acting factors within the nucleotide sequences encompassing the alternative pseudo-exon 6p and its surrounding intronic regions was performed using SFmap (Akerman et al., 2009; Paz et al., 2010) in conjunction with SpliceAid 2 (Piva et al., 2012).

Results and Discussion

Initially, the expression of SULT4A1 was determined in SH-SY5Y cells and human pluripotent stem cells (hiPSC) differentiated into neuronal-like cells. In the SH-SY5Y cells, differentiation was confirmed as described elsewhere (Sidharthan et al., 2014). cDNA from each cell-line was amplified using primers located in exon 6 and 7 to give an amplicon of 397 bp when 6p was present or 270 bp when it was absent (Fig. 1A). SULT4A1 mRNA splicing changed over time from the variant to the WT transcript (Fig. 1A, upper panel). At the same time, SULT4A1 protein expression increased (Fig 1A, lower panel). By day 12, nearly all of the SULT4A1 transcript was WT. In the hiPSC differentiated over 28 days, neuronal-like cells were evident as shown by their morphological characteristics and increased expression of the neuron-specific marker NeuN (Supplementary Fig. 1). Similar to that seen in the SH-SY5Y cells, SULT4A1 transcript splicing changed during differentiation (Fig 1B, upper panel). At day 0, almost all of the transcript included exon 6p whereas, by day 14, there was a marked shift to the WT transcript. This was accompanied by an increase in SULT4A1 protein expression. By day 28, the WT transcript was dominant and protein expression was maximal (Fig. 1B, lower panel). To assess SULT4A1 mRNA splicing *in vivo*, whole embryonic mouse tissue from E8.5 to E18.5 was analysed (Fig 1C). At E8.5, only variant SULT4A1 transcript was present whereas, by E18.5, only the WT transcript was evident. SULT4A1 protein expression correlated with the appearance of the WT transcript. These results show that SULT4A1 splicing in both cells and embryonic tissue changes in association with cell differentiation. As there is little or no protein expression from the variant transcript, differential splicing of pseudo-exon 6p may be an important mechanism for the regulation of SULT4A1 protein expression at the post-transcriptional level. Moreover, the similar results seen in the human and mouse cells, along with the data by Hashiguchi et al (Hashiguchi et al., 2018), indicate that the change in SULT4A1 splicing may be consistent across species.

To identify possible intronic sequences that regulate SULT4A1 mRNA splicing, a minigene splicing assay was used (Pagani et al., 2003). A genomic fragment containing

pseudo-exon 6p with flanking intronic sequences (623 bp 5'-upstream and 821 bp 3'-downstream) was cloned into the minigene construct (Supplementary Fig. 2A). Differentiated SH-SY5Y cells were transiently transfected with the minigene and mRNA was isolated for analysis by RT-PCR using forward and reverse primers located upstream and downstream of the insert. This allowed for identification of mRNA both with and without pseudo-exon 6p (Supplementary Fig. 2B). Next, deletion mutants of the minigene were constructed by sequentially removing sequences in the upstream and downstream regions of pseudo-exon 6p. Each minigene was transiently transfected into SH-SY5Y cells with and without RA-induced differentiation, and mRNA was isolated to evaluate splicing. Endogenous SULT4A1 mRNA was also measured as a positive control (Fig. 2A, right panels). The full-length minigene showed differentiation-dependent splicing of 6p (Fig 2A). The ratio of variant to WT transcript is shown to the right of each gel and, for the full-length minigene, this was 0.69. Deletion of the upstream sequence to 208 bp (5' Δ 406) showed splicing of 6p similar to that seen with the full-length construct (ratio = 0.73). However, when the upstream sequence was deleted to only 176 bp, very little splicing of the minigene was observed, although endogenous SULT4A1 mRNA was efficiently spliced in the same cells. Similar deletions in the downstream sequence had little effect on splicing (Fig. 2A, 3' Δ 528 and 3' Δ 220). These results suggested that a region essentail for the splicing of 6p is located between 176 and 208 bp upstream of the pseudo-exon. This 32 bp sequence is shown in Fig. 2B (in red). Analysis of this region using SpliceAid 2 and SFmap databases identified potential *cis*-elements for 2 RNA processing factors - MBNL and CELF (Fig 2B). Both interact with similar sequences comprising CUG/CUU motifs (Lambert et al., 2014).

To determine whether the MBNL and CELF RNA processing factors could splice the SULT4A1 minigene, each family member (MBNL1, MBNL2, CELF1, CELF2) was ectopically expressed with the 5' Δ 338 minigene into undifferentiated SH-SY5Y cells. All 4 RNA processing factors were able to splice the 5' Δ 338 minigene (Fig. 2C, upper panel). Next,

DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

SH-SY5Y cells were stably transfected with a construct for each RNA processing factor and the splicing of endogenous SULT4A1 transcript was investigated (Fig. 2D). Again, all 4 factors were able to remove pseudo-exon 6p from the endogenous SULT4A1 transcript (Fig. 2D, upper panel). Taken together, these results indicate that either of the MBNL and CELF RNA processing factors can promoter SULT4A1 6p splicing.

To identify which of the RNA processing factors are expressed during cell differentiation, SH-SY5Y cells were treated with retinoic acid over 10 days and MBNL/CELF mRNA was estimated by RT-PCR (Fig 3A). SULT4A1 splicing with time was evident in these experiments (uppermost panel). Moreover, all 4 factors were expressed in the untreated cells (day 0) as well as during differentiation. MBNL-2 mRNA is, itself, subject to exon splicing, which can regulate both its activity as well as its subcellular localization (Sznajder et al., 2016). There are 3 exons in the C-terminus containing 54, 36 and 95 bases, respectively, which are differentially spliced, and these are evident in the RT-PCR products following differentiation of the SH-SY5Y cells (Fig. 3A).

Since mRNA levels do not always correlate with protein expression, the levels of the CELF and MBNL proteins were quantified in the differentiated SH-SY5Y cells (Fig. 3B). Western blots of whole cell extracts showed increased SULT4A1 expression with RA treatment. By contrast, there was a significant decrease in CELF-1 and MBNL-1 protein expression. There was no change in CELF-2 when normalized to tubulin levels whereas MBNL-2 protein increased significantly.

To determine changes in MBNL and CELF expression in another model of differentiation, each of the RNA processing factors was measured in the human pluripotent stem cells during differentiation. Similar to that seen in the SH-SY5Y cells, MBNL-2 protein increased with time (Fig. 4). However, changes in the expression of the other RNA processing factors were quite different. Both CELF-2 and MBNL-1 proteins increased until day 14

following which they decreased back towards levels seen in the undifferentiated cells. There was little change in the expression of CELF-1.

In the present study, the MBNL/CELF family of RNA processing factors were shown to splice the SULT4A1 variant transcript. Each of these proteins binds to similar RNA sequences to execute their splicing effects (Kalsotra et al., 2008; Wang et al., 2012; Konieczny et al., 2014). However, they are not always co-ordinately expressed. For example, in the developing heart, CELF-1 expression is down-regulated and MBNL-1 is up-regulated during differentiation (Kalsotra et al., 2010). Moreover, there are complex, and often competing, functions for the MBNL and CELF proteins (Kalsotra et al., 2008; Wang et al., 2015). For example, MBNL-1 can behave as a repressor of MBNL-2 activity by inhibiting MBNL-2 splicing. In MBNL-1 null mice, there is an up-regulation of MBNL-2 in muscle and heart tissue, with its enrichment in the nuclear compartment (Lee et al., 2013). In the present study, differentiation of both the SH-SY5Y cells and the hiPSCs resulted in a decrease in MBNL-1 expression and an increase in MBNL-2 expression. Moreover, the decrease in MBNL-1 was accompanied by alternative splicing of MBNL-2 (Fig 4). These results support the reciprocal nature of their respective functions. It is noteworthy that variant SULT4A1 mRNA is wide-spread in human tissues but splicing of exon 6p appears to be controlled in a tissue- and time-specific manner. This may reflect the complexity of the MBNL/CELF splicing machinery which also shows tissue- and time-specific expression in the body.

There is accumulating evidence suggesting MBNL proteins have a dominant role in executing a switch from the fetal to adult pattern of mRNA splicing with their levels increasing during development and differentiation (Han et al., 2013; Konieczny et al., 2014). By contrast, CELF-1 has an opposing role, orchestrating fetal splicing pattern with its level decreasing with development (Lee and Cooper, 2009; Poulos et al., 2011). MBNL proteins show increased expression in differentiated adult tissue, where MBNL-1 is responsible for adult type splicing in heart and skeletal muscle tissues. MBNL-2 is the dominant protein in brain and differentiated

neurons (Charizanis et al., 2012; Wang et al., 2012). In the present study, SULT4A1 protein was absent early in mouse embryogenesis but was highly expressed by E18.5. This suggests that SULT4A1 is part of the alternative splicing network in the developing nervous system (Raj and Blencowe, 2015), which is supported by the severe neurological phenotype seen in SULT4A1 null mice (Garcia et al., 2018).

While an *in vivo* function for SULT4A1 is currently lacking, the recently reported mouse knockout animals indicates it is critical for normal development (Garcia et al., 2018). Thus, genetic alterations in the gene may be associated with developmental abnormalities. The human SULT4A1 gene has a very low incidence of mutations (SNP frequency < 1:700), which supports a vital role in development (Minchin et al., 2008). The current study highlights the importance of splicing for correct expression of the SULT4A1 protein. Thus, dysregulation of the RNA processing factors responsible for SULT4A1 splicing may generate a phenotype similar to that described in the mouse knockout model. While functionally important genetic polymorphisms in the MBNL genes have not been widely reported, sequestration by transcripts with expanded CUG or CCUG repeats has been noted in the onset of myotonic dystrophy (Miller et al., 2000). It will be important in the future to determine whether SULT4A1 splicing is altered in pathological conditions where the activities of these splicing factors change.

DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Acknowledgements

The authors acknowledge assistance from Dr James Hudson, University of Queensland, with the differentiation of the hiPSC's and Dr Paul Dawson, Mater Research Institute, for mouse embryonic tissue and cDNA. The authors thank Dr Thomas Cooper, Baylor College of Medicine, for the kind gift of the MBNL and CELF constructs and Dr Franco Pagani for the original minigene construct.

DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Authorship contributions.

Participated in research design: Idris, Butcher, Minchin.

Conducted experiments: Idris, Butcher.

Performed data analysis: Idris, Minchin.

Wrote or contributed to the writing of the manuscript: Idris, Butcher, Minchin.

References

Akerman M, David-Eden H, Pinter RY, and Mandel-Gutfreund Y (2009) A computational approach for genome-wide mapping of splicing factor binding sites. *Genome Biology* **10:**R30-R30.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72:**248-254.

Brennan MD and Condra J (2005) Transmission disequilibrium suggests a role for the sulfotransferase-4A1 gene in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet* **139B:**69-72.

Charizanis K, Lee KY, Batra R, Goodwin M, Zhang C, Yuan Y, Shiue L, Cline M, Scotti MM, Xia G, Kumar A, Ashizawa T, Clark HB, Kimura T, Takahashi MP, Fujimura H, Jinnai K, Yoshikawa H, Gomes-Pereira M, Gourdon G, Sakai N, Nishino S, Foster TC, Ares M, Jr., Darnell RB, and Swanson MS (2012) Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron* **75:**437-450.

Encinas M, Iglesias M, Liu Y, Wang H, Muhaisen A, Ceña V, Gallego C, and Comella JX (2000) Sequential Treatment of SH-SY5Y Cells with Retinoic Acid and Brain-Derived Neurotrophic Factor Gives Rise to Fully Differentiated, Neurotrophic Factor-Dependent, Human Neuron-Like Cells. *J Neurochem* **75:**991-1003.

Falany CN, Xie X, Wang J, Ferrer J, and Falany JL (2000) Molecular cloning and expression of novel sulphotransferase-like cDNAs from human and rat brain. *Biochem J* **346 Pt 3:**857-864.

Garcia PL, Hossain MI, Andrabi SA, and Falany CN (2018) Generation and Characterization of SULT4A1 Mutant Mouse Models. *Drug Metab Dispos* **46:**41-45.

Han H, Irimia M, Ross PJ, Sung HK, Alipanahi B, David L, Golipour A, Gabut M, Michael IP, Nachman EN, Wang E, Trcka D, Thompson T, O'Hanlon D, Slobodeniuc V, Barbosa-Morais NL, Burge CB, Moffat J, Frey BJ, Nagy A, Ellis J, Wrana JL, and Blencowe BJ (2013) MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* **498:**241-245.

Hashiguchi T, Shindo S, Chen SH, Hong JS, and Negishi M (2018) Sulfotransferase 4A1 Increases Its Expression in Mouse Neurons as They Mature. *Drug Metab Dispos* **46:**860-864.

Kalsotra A, Wang K, Li P-F, and Cooper TA (2010) MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development. *Genes & Development* **24:**653-658.

Kalsotra A, Xiao X, Ward AJ, Castle JC, Johnson JM, Burge CB, and Cooper TA (2008) A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc Natl Acad Sci U S A* **105:**20333-20338.

Konieczny P, Stepniak-Konieczna E, and Sobczak K (2014) MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res* **42:**10873-10887.

Lambert N, Robertson A, Jangi M, McGeary S, Sharp Phillip A, and Burge Christopher B (2014) RNA Bind-n-Seq: Quantitative Assessment of the Sequence and Structural Binding Specificity of RNA Binding Proteins. *Molecular Cell* **54:**887-900.

Lee JE and Cooper TA (2009) Pathogenic mechanisms of myotonic dystrophy. *Biochem Soc Trans* **37:**1281-1286.

Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, Warren SA, Chamberlain CM, Finn D, Hong H, Ashraf H, Kasahara H, Ranum LP, and Swanson MS (2013) Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol Med* **5:**1887-1900.

Meltzer HY, Brennan MD, Woodward ND, and Jayathilake K (2008) Association of Sult4A1 SNPs with psychopathology and cognition in patients with schizophrenia or schizoaffective disorder. *Schizophr Res* **106:**258-264.

Miller JW, Urbinati CR, Teng-umnuay P, Stenberg MG, Byrne BJ, Thornton CA, and Swanson MS (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *The EMBO Journal* **19:**4439-4448.

Minchin RF, Lewis A, Mitchell D, Kadlubar FF, and McManus ME (2008) Sulfotransferase 4A1. *The International Journal of Biochemistry & Cell Biology* **40:**2686-2691.

Ovchinnikov DA, Hidalgo A, Yang SK, Zhang X, Hudson J, Mazzone SB, Chen C, Cooper-White JJ, and Wolvetang EJ (2015) Isolation of contractile cardiomyocytes from human pluripotent stem-cell-derived cardiomyogenic cultures using a human NCX1-EGFP reporter. *Stem Cells Dev* **24:**11-20.

Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, and Baralle FE (2003) New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum Mol Genet* **12:**1111-1120.

Paz I, Akerman M, Dror I, Kosti I, and Mandel-Gutfreund Y (2010) SFmap: a web server for motif analysis and prediction of splicing factor binding sites. *Nucleic Acids Res* **38:**W281-285.

Piva F, Giulietti M, Burini AB, and Principato G (2012) SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. *Human Mutation* **33:**81-85.

Poulos MG, Batra R, Charizanis K, and Swanson MS (2011) Developments in RNA splicing and disease. *Cold Spring Harb Perspect Biol* **3:**a000778.

Raj B and Blencowe BJ (2015) Alternative Splicing in the Mammalian Nervous System: Recent Insights into Mechanisms and Functional Roles. *Neuron* **87:**14-27.

Ramsey TL, Liu Q, and Brennan MD (2014) Replication of SULT4A1-1 as a pharmacogenetic marker of olanzapine response and evidence of lower weight gain in the high response group. *Pharmacogenomics* **15:**933-939.

Scotti MM and Swanson MS (2015) RNA mis-splicing in disease. *Nature Reviews Genetics* **17:**19.

Sidharthan NP, Butcher NJ, Mitchell DJ, and Minchin RF (2014) Expression of the orphan cytosolic sulfotransferase SULT4A1 and its major splice variant in human tissues and cells: dimerization, degradation and polyubiquitination. *PLoS One* **9:**e101520.

Sznajder LJ, Michalak M, Taylor K, Cywoniuk P, Kabza M, Wojtkowiak-Szlachcic A, Matloka M, Konieczny P, and Sobczak K (2016) Mechanistic determinants of MBNL activity. *Nucleic Acids Res* **44:**10326-10342.

Wang D, Li Q, Favis R, Jadwin A, Chung H, Fu DJ, Savitz A, Gopal S, and Cohen N (2014) SULT4A1 haplotype: conflicting results on its role as a biomarker of antipsychotic response. *Pharmacogenomics* **15:**1557-1564.

Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, Lecuyer E, and Burge CB (2012) Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150:**710-724.

Wang ET, Ward AJ, Cherone JM, Giudice J, Wang TT, Treacy DJ, Lambert NJ, Freese P, Saxena T, Cooper TA, and Burge CB (2015) Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res* **25:**858-871.

Footnote:

This work was supported by a grant from the National Health and Medical Research Council of Australia [Grant # 1099135].

Figure legends

Fig. 1. SULT4A1 transcript and protein expression during cell differentiation and mouse embryonic development. (A) SH-SY5Y cells were cultured for up to 12 days in DMEM/F12 medium in the presence of 10 μM retinoic acid. At each time point, cells were harvested for mRNA and protein determination. RT-PCR was performed using primers located in exon 6 and exon 7. The variant transcript, which contained pseudo-exon 6p, produced a PCR product of 397 bp while the WT transcript produced a PCR product of 270 bp. RT-PCR results are shown in the upper 2 images while SULT4A1 protein is shown in the lower 2 images. (B) SULT4A1 mRNA and protein in hiPSCs during RA-induced differentiation. (C) Expression of SULT4A1 mRNA and protein in mouse embryonic tissue from day E8.5 to E18.5. β-Actin was used as a house-keeping gene for RT-PCR and α-tubulin as a loading control for Western blots. Molecular weight markers are shown on the right of each Western blot.

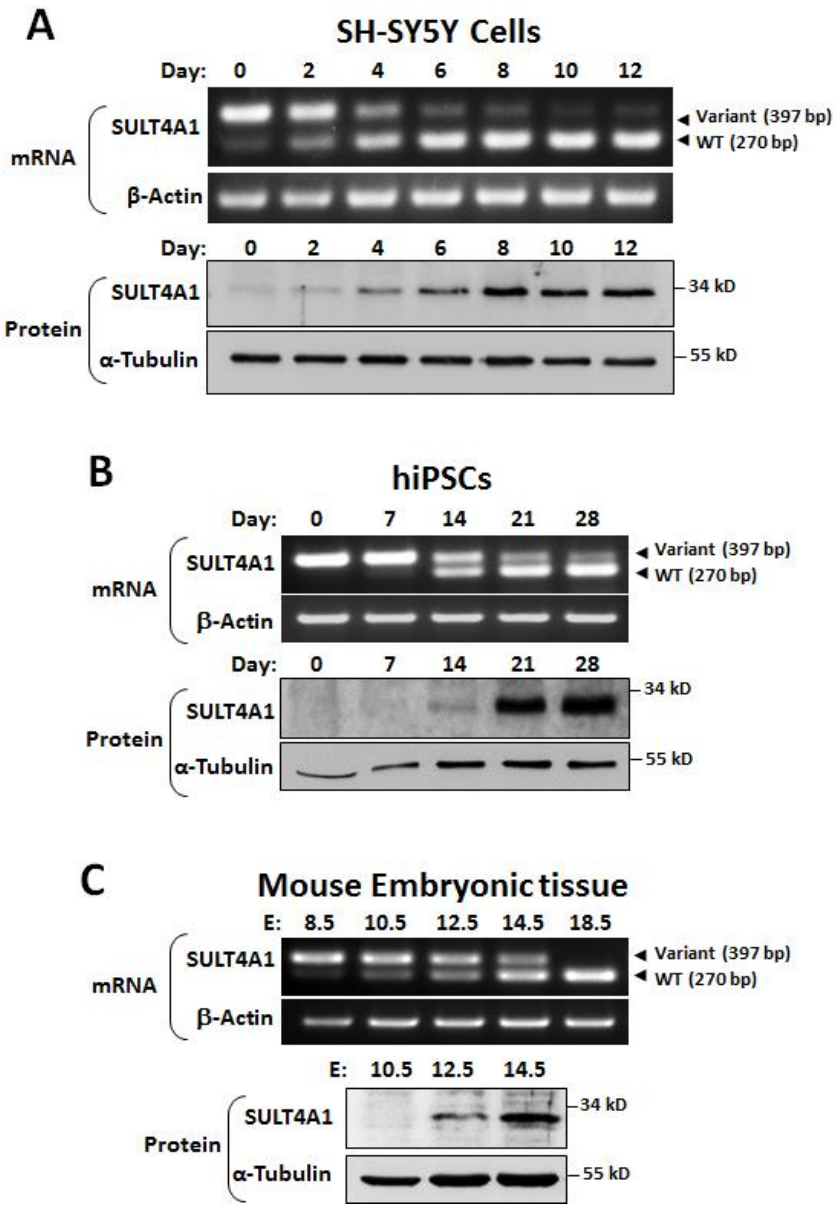
Fig. 2. Minigene splicing assay of SULT4A1 pseudo-exon 6p in SH-SY5Y cells differentiated with retinoic acid (RA). (A) Deletion minigene assays: 10 minigene plasmids with sequential deletion of intronic segments of the ~1.6 kb full length minigene were constructed and transfection into SH-SY5Y cells with and without RA treatment. Switching of the endogenous SULT4A1 transcript was used as a positive control (far right panels). The nomenclature of the deletion minigenes is based on the number of base pairs deleted either from the 5' or 3' ends of the full length minigene. RA untreated (-) and treated (+) lanes are labelled and the ratio of variant to WT PCR product (V/WT) is shown next to the RT-PCR gels. Results are representative of duplicate experiments. (B) Sequence of the SULT4A1 gene upstream of pseudo-exon 6p (grey arrow) showing the 32 bases sequence between 5' Δ 406 and 5' Δ 438 (red). Putative binding motifs for RNA processing factors are shown in the boxes with blue = MBNL and green = CELF. Numbers refer to the number of bases upstream of the 6p pseudo-exon. (C) Effect of ectopic expression of the MBNL and CELF genes on splicing of the 5' Δ 338 minigene construct in undifferentiated SH-SY5Y cells. Cells were transfected with each RNA splicing factor and the minigene. Upper panel shows the PCR products of the minigene transcript with the variant and WT indicated. Expression of each protein is shown as a FLAG Western blot below. (D) Effect of MBNL and CELF proteins on the splicing of endogenous SULT4A1 in undifferentiated SH-SY-5Y cells. Upper panel shows the variant and WT PCR products. Expression of each protein is shown as a FLAG Western blot below. Results are representative of duplicate experiments. EV = empty vector. Molecular weight markers are shown on the right of each Western blot.

Fig. 3. Expression of SULT4A1, MBNL and CELF transcripts in SH-SY5Y cells during RA-induced differentiation. (A) SH-SY5Y cells were treated with 10 μM RA for 10 days. RNA was isolated at the indicated times and amplified by RT-PCR. For MBNL2, splice variants were detected. (B). Western blots of protein expression after 10 days of RA treatment. Quantification of protein levels is shown alongside each Western blot. Data are mean with standard deviations (n = 3). Asterisk indicates significant differences (Student's t-test, $p < 0.05$)

Fig. 4. Expression of MBNL and CELF proteins in hiPSC during neuronal differentiation with 10 μM RA. Proteins were identified by Western blots and quantified by densitometry as the ratio of protein to α-tubulin (MBNL proteins) or β-actin (CELF proteins). Molecular weight markers are shown on the right of each Western blot. For MBNL-1 and 2, closed symbols are the upper band and open symbols are the lower band.

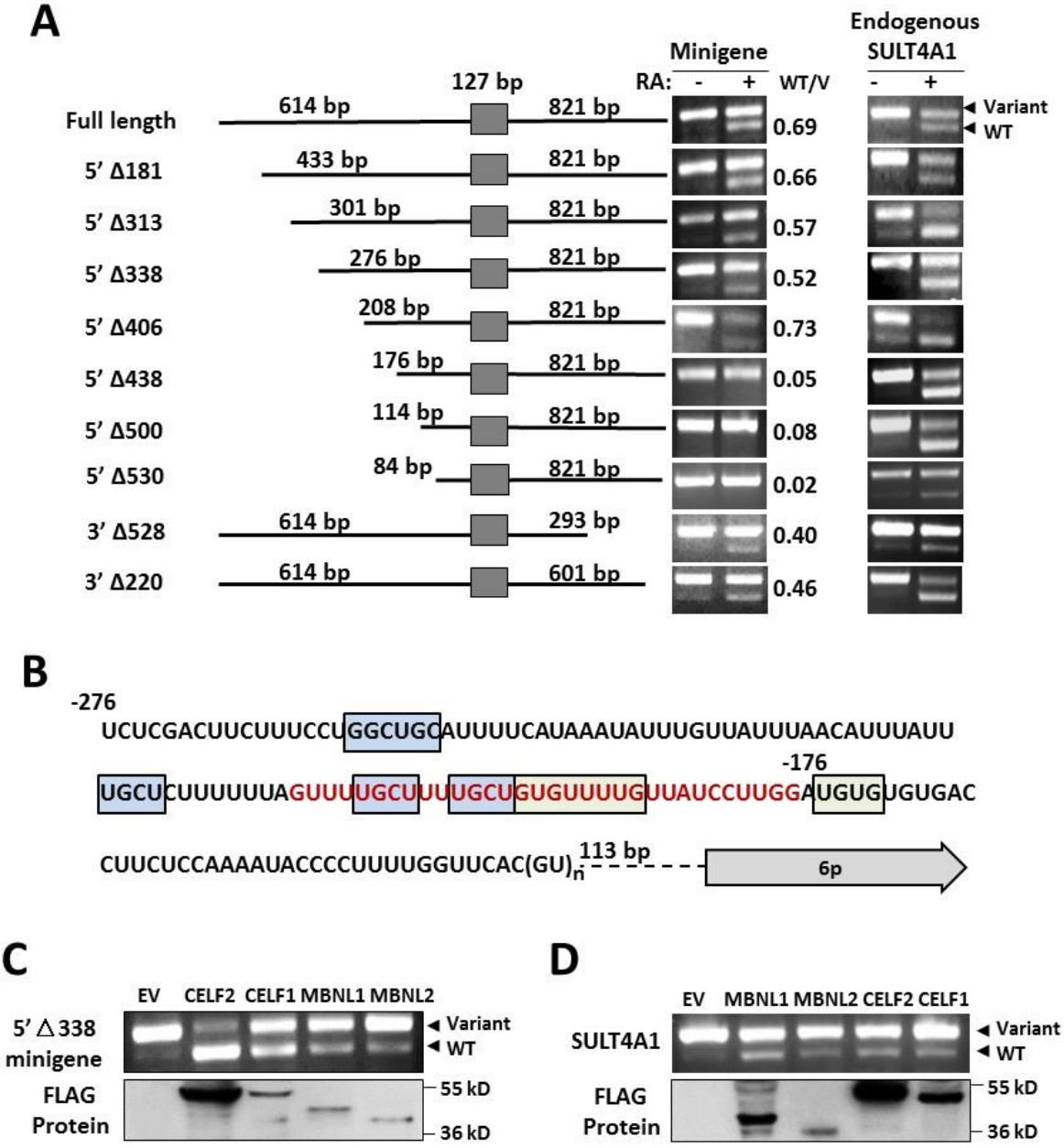
DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 1



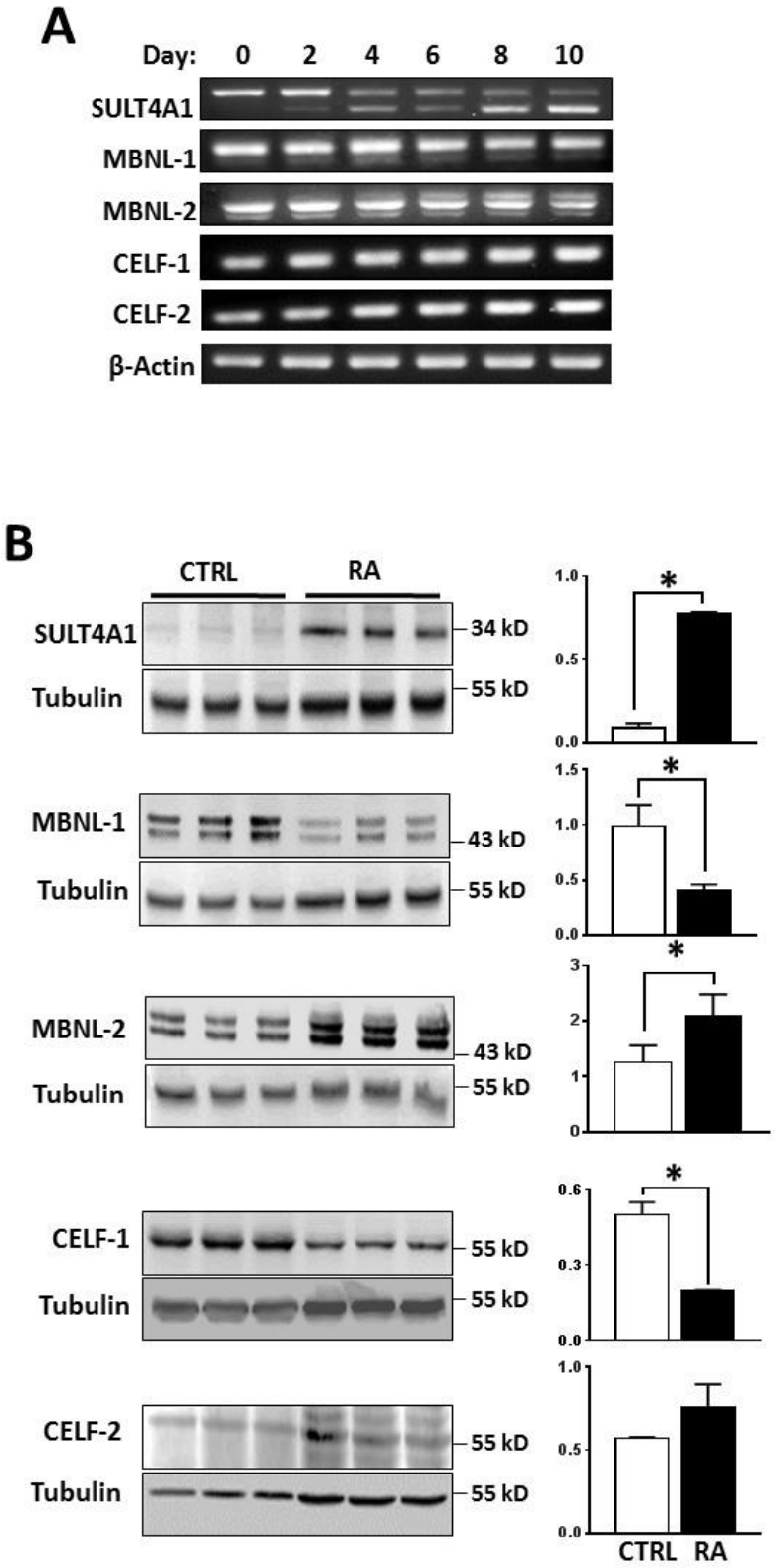
DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 2



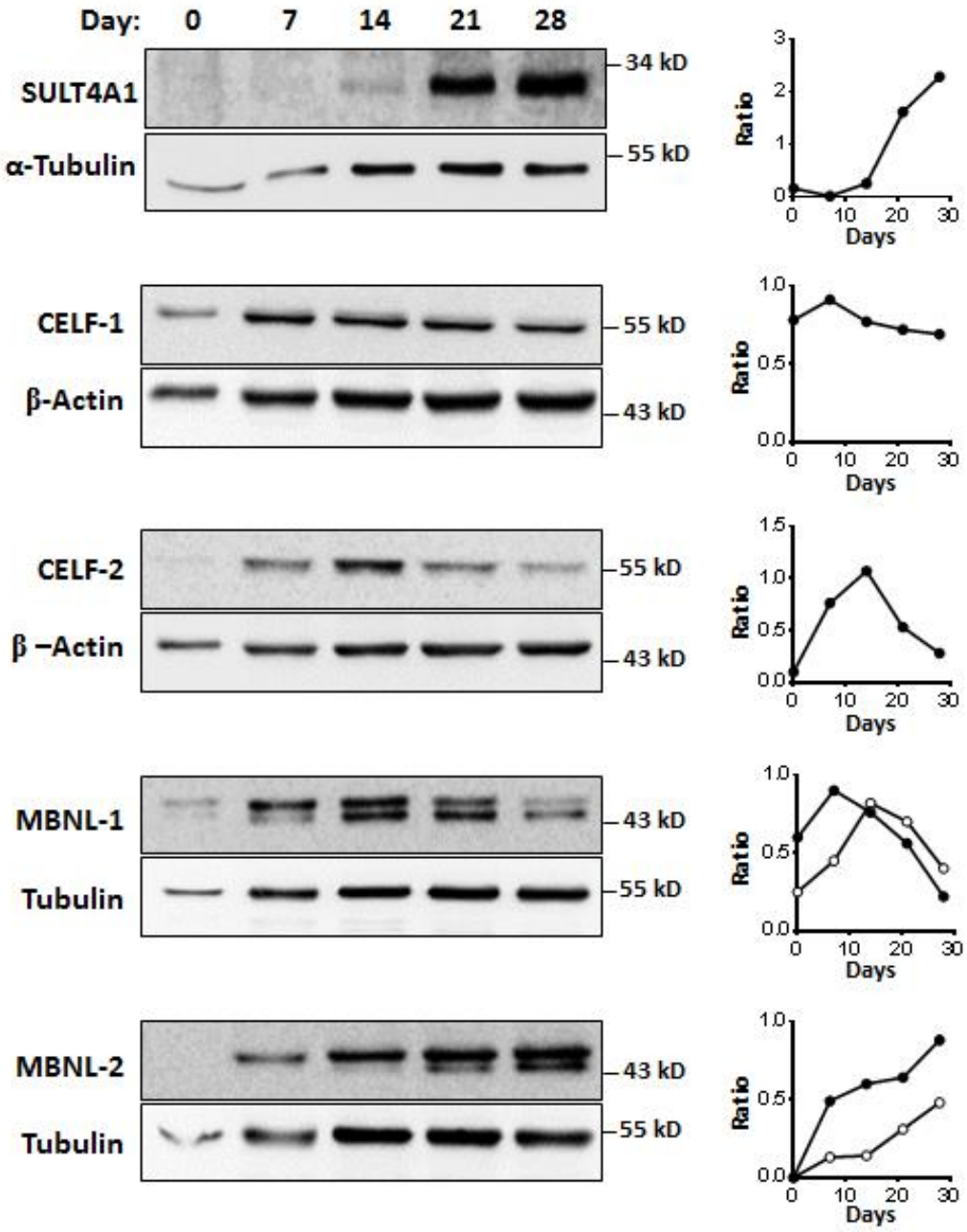
DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 3



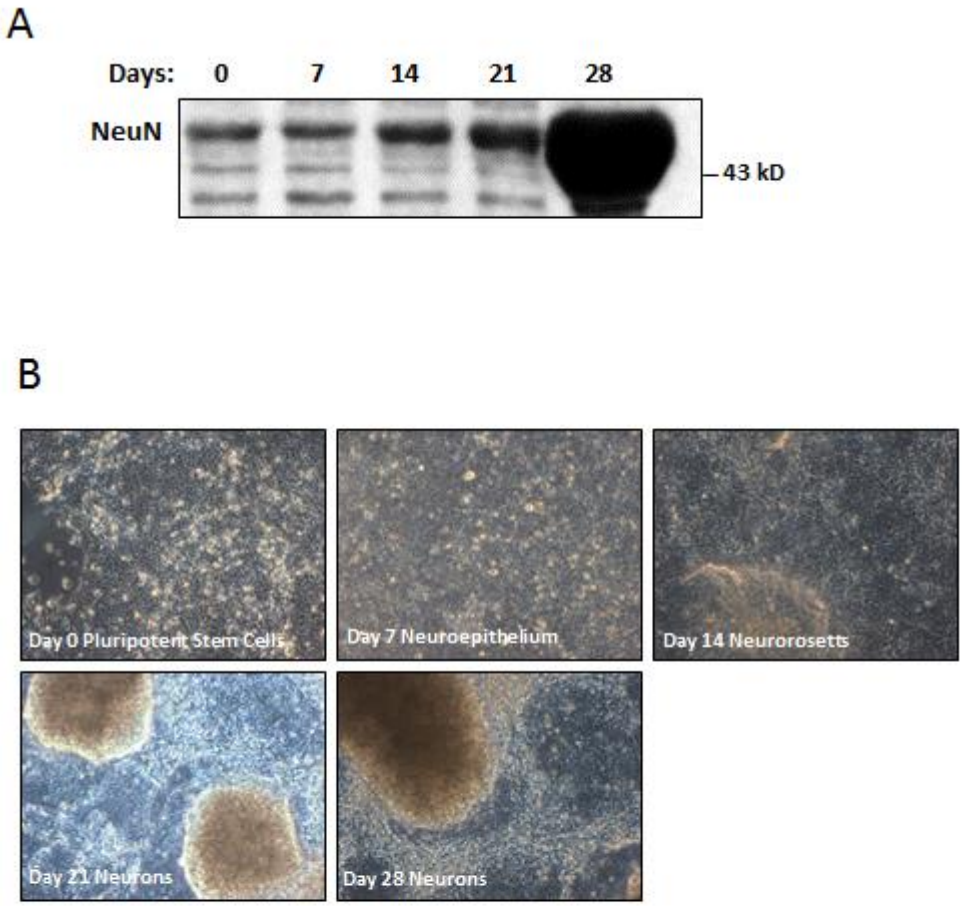
DMD Fast Forward. Published on January 3, 2019 as DOI: 10.1124/dmd.118.085290
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 4

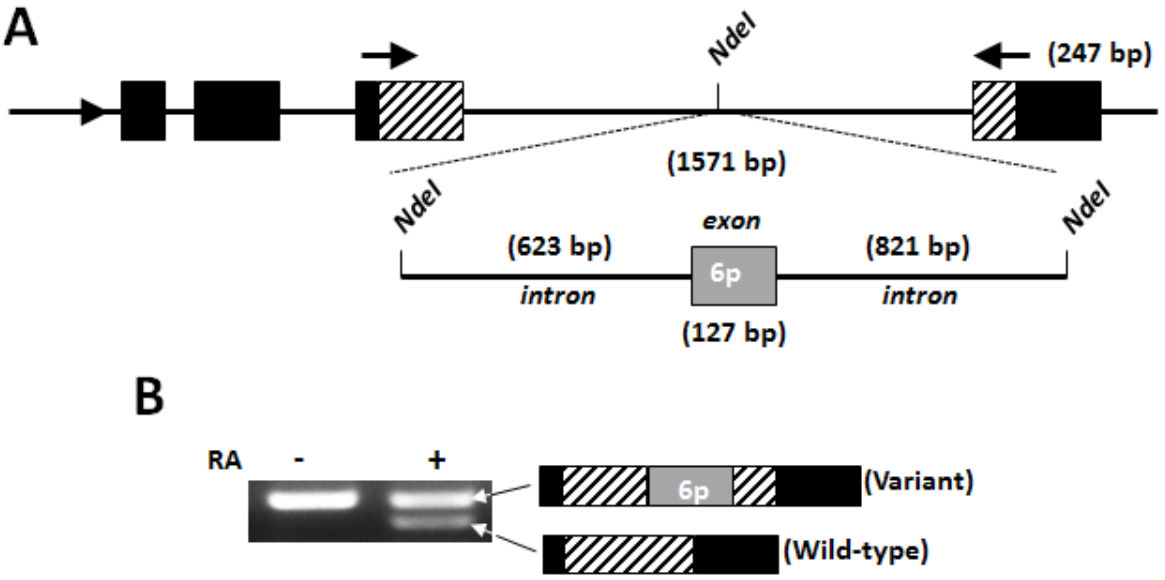


The MBNL/CELF Splicing Factors Regulate Cytosolic Sulfotransferase 4A1 protein expression During Cell Differentiation.

Supplementary Figures and Tables



Supplementary Figure 1. Differentiation of hiPSC over 28 days. A) Western blot of NeuN. B) Images of cells.



Supplementary Figure 2: Minigene splicing assay of SULT4A1 exon 6p. (A) Map of the minigene insert comprising ~1.6 kb genomic region encompassing the 127 bp long exon 6p, its 623 bp upstream and 821 bp downstream intronic regions, which was cloned into pTBNde minigene at a NdeI restriction site. (B) Splicing of minigene transcript after transfection into SH-SY5Y cells treated with RA shows two PCR products depending on whether the 127 bp exon 6p was spliced in or spliced out from the transcript. This result confirms that the minigene construct is spliced similarly to the endogenous SULT4A1 mRNA

Supplementary Table 1: Forward (FP) and reverse (RP) primers used for RT-PCR

Target	Primer	Sequence	PCR products (bp)	PCR conditions
SULT4A1	FP	5'-CTA CGG CTC CTG GTT TGA G-3'	270 & 397	Initial denaturation at 95 °C for 10 min, followed by **36** cycles of melting at 95 °C for 10 s, annealing at **58** °C for 10 s and elongation at 72 °C for **30** s.
	RP	5'-ATG GAG ACG GTG AAG ATG TC-3'		
Minigene	FP	5'-CAACTTCAAGCTCCTAAGCCACTGC-3'	247 & 374	Initial denaturation at 95 °C for 10 min, followed by **36** cycles of melting at 95 °C for 10 s, annealing at **55** °C for 10 s and elongation at 72 °C for **25** s
	RP	5'-TAGGATCCGGTCACCAGGAAGTTGGTTAAATCA-3'		
β- Actin	FP	5'-CCT CGC CTT TGC CGA TCC-3'	626	Initial denaturation at 95 °C for 10 min, followed by **33** cycles of melting at 95 °C for 10 s, annealing at **58** °C for 10 s and elongation at 72 °C for **45** s
	RP	5'-GGA TCT TCA TGA GGT AGT CAG TC-3'		
MBNL-2	FP	5'-ACTTGCAGGCCAAAATCAAA-3'	496, 532 & 586	Initial denaturation at 95 °C for 10 min, followed by **35** cycles of melting at 95 °C for 10 secs, annealing at **60** °C for 10 s and elongation at 72 °C for **40** s
	RP	5'-GTGGTCTGTCCAGAAGAACT-3'		
MBNL-1	FP	5'-GCCCAATACCAGGTCAACCA-3'	142 & 196	Initial denaturation at 95 °C for 10 minutes, followed by **33** cycles of melting at 95 °C for 10 secs, annealing at **54** °C for 10 secs and elongation at 72 °C for **25** secs
	RP	5'-TGTTAAAGACTGCGGTGGCA-3'		
CELF-2	FP	5'-CAGGGTGATG TTCTCTCCAT TT-3'	76	Initial denaturation at 95 °C for 10 min, followed by **33** cycles of melting at 95 °C for 10 s, annealing at **58** °C for 10 secs and elongation at 72 °C for **15** s
	RP	5'-GCCTCGACTCAGCCC ATC-3'		
CELF-1	FP	5'-CTGGA AGCCAGAAGGAAGGT-3'	81	Initial denaturation at 95 °C for 10 min, followed by **25** cycles of melting at 95 °C for 10 s, annealing at **58** °C for 10 s and elongation at 72 °C for **15** s
	RP	5'-GCAGG TCCTGATCACCAAAC-3'		

Supplementary Table 2: Forward (FP) and reverse (RP) primers used for minigene cloning

Minigene	Primer	Sequence	PCR conditions
KpnI-HindIII	FP	5'-TACTGGTACCTATGAACCTGATGGACACTTGG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 58 °C for 15 s and elongation at 72 °C for 2.0 min.
	RP	5'-TCATAAGCTTATGCAGATTTCAAGAAGCCTGC-3'	
SDM NdeI	FP	5'-ATACTGTGTGCCCATCTGCTCCTTCCCTTGG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 58 °C for 15 s and elongation at 72 °C for 2.0 min.
	RP	5'-CCAAGGGAAGGAGATGGGCACACAGTAT-3'	
Full-length	FP	5'-TACTCATATGAACCTGATGGACACTTGGTG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 58 °C for 15 s and elongation at 72 °C for 2.0 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
5′ Δ181	FP	5′-TACTCATATAACATAGGTCAGGCCTTCAG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 55 °C for 15 s and elongation at 72 °C for 1.5 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
5′ Δ313	FP	5′-TACTCATATGAGGAAGTGAACTGAACCAC-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 56 °C for 15 s and elongation at 72 °C for 1.5 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
5′ Δ338	FP	5′-TACTCATATGCTCTCGACTTCTTTCCTG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 56 °C for 15 s and elongation at 72 °C for 1.5 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
5′ Δ386	FP	5′-TACTCATATGCATTTATTTGCTCTTTTTTAG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 50 °C for 15 s and elongation at 72 °C for 1.5 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
5′ Δ406	FP	5′-TACTCATATGTTTTGCTTTTGCTGTG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 52 °C for 15 s and elongation at 72 °C for 1.5 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
5′ Δ438	FP	5′-TACTCATATGTGACCTTCTCC-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C

	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	for 15 s, annealing at 50 °C for 15 s and elongation at 72 °C for 1.5 min.
5′ Δ500	FP	5′-TACTCATATGTGAAGCAGCGTC-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 53 °C for 15 s and elongation at 72 °C for 1.0 min.
	RP	5'-TCATCATATGATGCAGATTTCAAGAAGCCTGC-3'	
3′ Δ528	FP	5'-TACTCATATGAACCTGATGGACACTTGGTG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 57 °C for 15 s and elongation at 72 °C for 1.0 min.
	RP	5′-TCATCATATGAATGTCCTTCACTAGCTCCG-3'	
3′ Δ220	FP	5'-TACTCATATGAACCTGATGGACACTTGGTG-3'	Initial denaturation at 95 °C for 5 min, followed by 40 cycles of melting at 95 °C for 15 s, annealing at 57 °C for 15 s and elongation at 72 °C for 1.5 min.
	RP	5′-TCATCATATGCTGAGCTGTGGGATAAACT-3'	