Published online 8 March 2022

NAR Genomics and Bioinformatics, 2022, Vol. 4, No. 1 **1**
https://doi.org/10.1093/nargab/lqac016

A comprehensive atlas of fetal splicing patterns in the brain of adult myotonic dystrophy type 1 patients

Max J.F. Degener [1,†], Remco T.P. van Cruchten [1,†], Brittney A. Otero [2], Eric T. Wang [2], Derick G. Wansink [3] and Peter A.C. 't Hoen [1,*]

[1]Centre for Molecular and Biomolecular Informatics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands, [2]Department of Molecular Genetics and Microbiology, Center for NeuroGenetics, Genetics Institute, University of Florida, FL 32610-0266 Gainesville, FL, USA and [3]Department of Cell Biology, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands

Received October 10, 2021; Revised January 28, 2022; Editorial Decision February 08, 2022; Accepted February 13, 2022

ABSTRACT

In patients with myotonic dystrophy type 1 (DM1), dysregulation of RNA-binding proteins like MBNL and CELF1 leads to alternative splicing of exons and is thought to induce a return to fetal splicing patterns in adult tissues, including the central nervous system (CNS). To comprehensively evaluate this, we created an atlas of developmentally regulated splicing patterns in the frontal cortex of healthy individuals and DM1 patients, by combining RNA-seq data from BrainSpan, GTEx and DM1 patients. Thirty-four splice events displayed an inclusion pattern in DM1 patients that is typical for the fetal situation in healthy individuals. The regulation of DM1-relevant splicing patterns could partly be explained by changes in mRNA expression of the splice regulators *MBNL1*, *MBNL2* and *CELF1*. On the contrary, interindividual differences in splicing patterns between healthy adults could not be explained by differential expression of these splice regulators. Our findings lend transcriptome-wide evidence to the previously noted shift to fetal splicing patterns in the adult DM1 brain as a consequence of an imbalance in antagonistic *MBNL* and *CELF1* activities. Our atlas serves as a solid foundation for further study and understanding of the cognitive phenotype in patients.

INTRODUCTION

Myotonic dystrophy type 1 (DM1; OMIM #160900), also known as Steinert's disease or dystrophia myotonica, is an autosomal dominant neuromuscular disease with a highly variable, multisystemic clinical presentation, affecting skeletal and smooth muscles, the central nervous system (CNS), the heart and several other organs. DM1 is the most common form of adult onset muscular dystrophy with an estimated global prevalence of 1:8000 (1). The clinical phenotype of DM1 is most notably defined by myotonia, a delayed relaxation of skeletal muscles following contraction. Additionally, affected individuals show a variable combination of progressive weakness of distal muscle groups, insulin resistance, cardiac arrhythmia, cataract, fatigue, cognitive impairment and changes in personality and behavior.

The cause of DM1 is a (CTG)n trinucleotide repeat expansion in the 3′-untranslated region of the DM1 protein kinase (*DMPK*) gene, located on the long arm of chromosome 19 (2). The central mechanism of DM1 pathophysiology involves the formation of RNA hairpin structures by the (CUG)n repeat in *DMPK* transcripts (3,4). Proteins of the muscleblind like splicing regulator (MBNL) family are recruited by these hairpin structures, bind to repetitive 'YGCY' motifs of the (CUG)n expansion and form RNA foci (5–7). Consequently, MBNL proteins, trapped in nuclear RNA foci, are depleted from their normal RNA targets (8). Since MBNL proteins also regulate their own splicing, the depletion of MBNL proteins leads to a further loss of MBNL function due to the formation of other splice variants (9). Furthermore, the presence of expanded *DMPK* transcripts increases levels of hyperphosphorylated, stable CUG-binding protein and ETR3-like factor 1 (CELF1) (10).

RNA-binding proteins (RBPs) like CELF1, MBNL1 and MBNL2 function as trans-acting regulators of alternative splicing (11,12). These RBPs regulate alternative splicing by binding to short RNA motifs in the pre-mRNA. Depending on the binding sites and the RBPs involved, splicing of the target RNA elements can be promoted or suppressed (13). Although the functional impact of many individual splice events remains ill understood, the coordination of these events, driven by binding of RBPs, determines the

*To whom correspondence should be addressed. Tel: +31 24 361 93 90; Email: peter-bram.thoen@radboudumc.nl
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

development of tissues, particularly of the heart, skeletal muscle and brain (14). During heart development in mice and chickens, MBNL1 and CELF1 mediate a highly conserved transition from fetal to adult splicing patterns (15). Similarly, loss of *Mbnl1* and *Mbnl2* in the brain of adult knockout (KO) mice shifts the splicing profile to an earlier stage of CNS development and causes mis-splicing of several important developmentally regulated exons (16,17).

A cornerstone of DM1 pathophysiology is the aberrant alternative splicing of many pre-mRNA products and the preferential expression of fetal transcript variants in the adult state. Numerous studies revealed mis-splicing in DM1 patient tissue and animal models by means of RT-PCR, microarray and RNA-Sequencing (e.g. (17–24)). Mis-splicing in DM1 is not limited to one type of tissue, but affects virtually all systems associated with the disease, in particular skeletal and cardiac muscles as well as the CNS (19–20,22). Importantly, abnormal splicing of certain genes has been linked to characteristic DM1 symptoms, albeit with variable levels of evidence (17,19,25–27). The alternative splicing of a number of these phenotype-linked genes is regulated by MBNL1 or MBNL2 (*DMD*, *BIN1*, *SCN5A*), CELF1 (*CLCN1*, *RYR1*, *INSR*, *TNNT2*) or both (*CACNA1S*, *PKM*, *GRIN1*, *MAPT*). Therefore, the loss-of-function of MBNL1/2, due to their sequestration in foci, and the gain-of-function of CELF1, due to its hyperphosphorylation, could explain the dysregulation of alternative splicing in the disease state (7,28). MBNL1/2 and CELF1 regulate splicing in an antagonistic fashion by either promoting adult (MBNL1/2) or fetal (CELF1) splicing patterns (29,30). Consequently, the altered activity of MBNL1/2 and CELF1 will induce the expression of fetal transcript variants in adult DM1 tissues (8) and DM1 mouse models (24).

Here, we set out to create an atlas of all developmentally regulated splicing patterns in the human frontal cortex and studied to what extent they were affected in the brains of DM1 patients. For this, we analyzed publicly available RNA-Seq data from healthy individuals, generated by the BrainSpan and GTEx projects (31,32), together with recently published data from the brains of DM1 patients and controls (23). In contrast to previous more targeted approaches, this now allowed us to make a comprehensive inventory of splicing events demonstrating a transition toward the fetal state in the DM1 frontal cortex. Moreover, we assessed to what extent this could be explained by the RNA expression levels of *CELF1* and *MBNL1/2*.

MATERIALS AND METHODS

Datasets

The BrainSpan Atlas of the Developing and Adult Human Brain is a resource for studying the spatial and temporal development of the human brain (31). The RNA-Seq dataset consists of multiple CNS regions and age groups (from 8 weeks before birth to 40 years of age). The BrainSpan consortium only included samples that did not have any confounding pathological factors (e.g. cerebrovascular incidents and tumors), did not show chromosomal abnormalities or malformations and lesions, had not been subjected to drug or alcohol abuse and showed a RNA integrity number (RIN) of at least 5. Our selection consisted of 139 samples from 39 individuals (23 males, 16 females) in which four subregions of the frontal cortex were represented (i.e. dorsolateral: $N = 35$, ventrolateral: $N = 36$, medial: $N = 35$ and orbital: $N = 33$) (Figure 1A). Since we found no differences in splicing of our exons of interest between these subregions, these data were pooled in further analyses (Supplementary Figure S1 and Supplementary Table S1). The RNA-Seq was paired-end with a read length of 76 bp. Access to FASTQ files and comprehensive phenotype data was requested from the database of Genotypes and Phenotypes (dbGaP; study accession: phs000731.v2.p1). FASTQ files were downloaded from the sequence read archive (SRA).

The Otero *et al*. (2021) dataset consists of frontal cortex samples from 21 DM1 adult onset patients (9 males, 12 females) and 8 controls (4 males, 4 females) (Figure 1B, (23)). The samples were retrieved from frozen brain tissue of adult individuals. The age of the donors ranged from 39 to 77 years (median = 59). Controls were labeled as 'unaffected'. The RNA-Seq was paired-end with a read length of 75 bp. More details can be found in the associated publication (23). The raw RNA-Seq data were downloaded from GEO (study accession: GSE157428).

The GTEx project provides RNA-Seq data from a variety of tissues from healthy donors (32). We focused on the human frontal cortex and included only samples with RIN above 7, coming from donors with non-disease related death (Hardy Scale below 3), in which autolysis of the extracted tissue was not visible upon inspection by a pathologist. The final selection consisted of 54 samples from 54 donors (42 males, 12 females). The age of the donors ranged from 23 to 70 (median = 60.5). The RNA-Seq was paired-end with a read length of 76 bp. For the selected samples, raw gene-level counts were downloaded from the public GTEx portal (release V8). Access to BAM files and comprehensive sample and phenotype data was requested from dbGaP (study accession: phs000424.v8.p2). Following approval, selected BAM files and metadata were downloaded from the corresponding repository on the NHGRI AnVIL (Genomic Data Science Analysis, Visualization, and Informatics Labspace) platform.

RNA-Seq data processing

To allow comparison across datasets, we applied the same RNA-Seq pre-processing and expression quantification steps on FASTA files from BrainSpan and Otero *et al*. (2021) as specified for GTEx V8 (https://gtexportal.org/home/documentationPage). FASTQ files were aligned to the hg38 genome with STAR 2.7.0f (33). The STAR index was built with the genome sequence (GRCh38.p10) in FASTA format and the comprehensive gene annotation, both retrieved from GENCODE release v26. Samtools 1.9 was used to sort and index the resulting BAM files (34). Gene-level counts were calculated by RNASeQC 2.3.5 (35) and normalized to log2 counts per million (CPM) with the trimmed mean of M-values method, as implemented by the edgeR 3.28.0 package (36). RNASeQC 2.3.5 required a col-

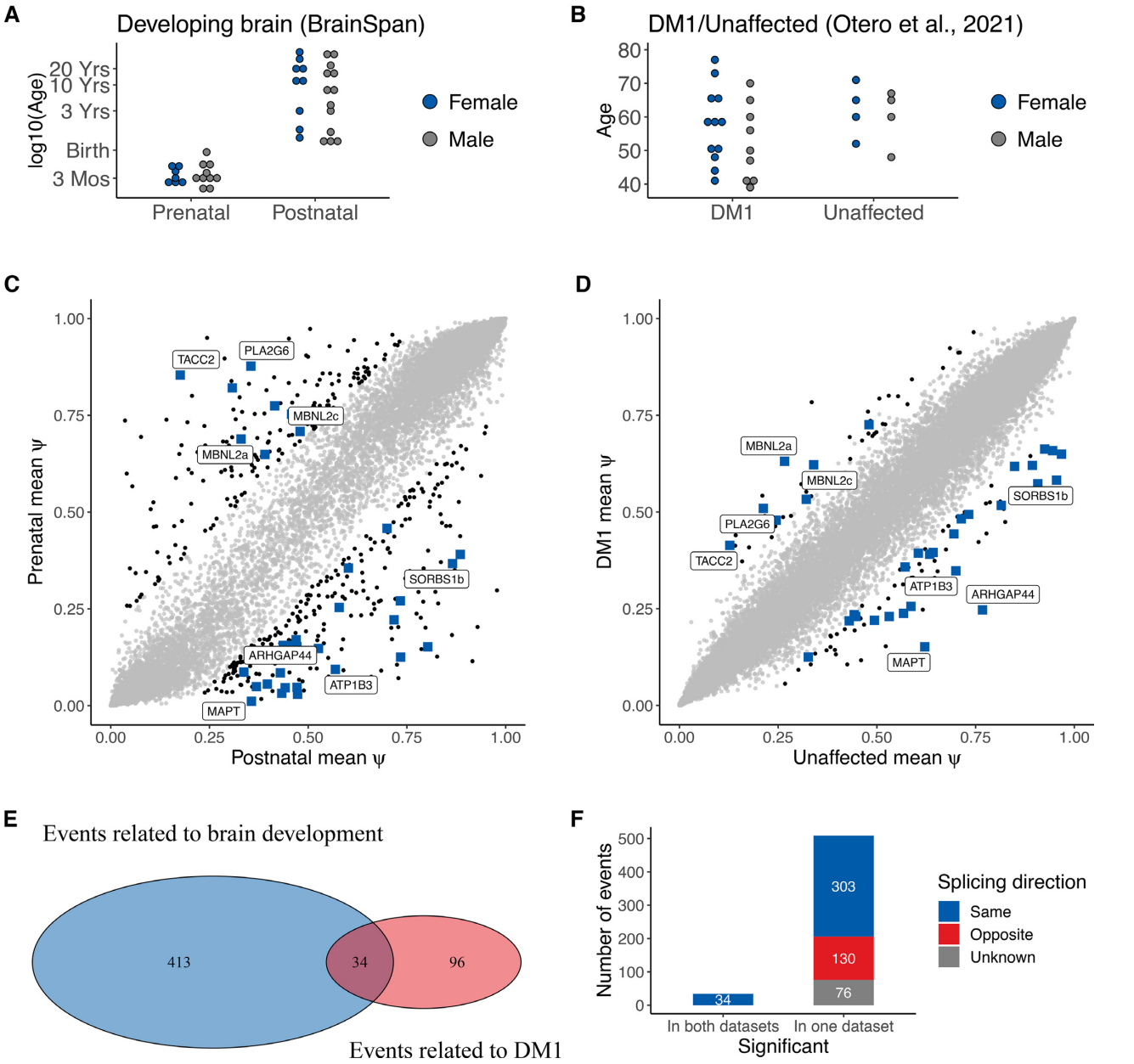

Figure 1. A quarter of the aberrant splice events in adult DM1 brains are associated with brain development. (**A**) Overview of the developmental stage of 39 healthy donors in the BrainSpan dataset. In total, the donors provided 139 RNA-Seq samples from four subregions of the frontal cortex. (**B**) Overview of the DM1 dataset from Otero *et al.* (2021) containing RNA-Seq samples from the frontal cortices of 21 DM1 adult patients and 8 unaffected controls. (**C**) Scatter plot of mean Ψ in prenatal (*y*-axis) and postnatal (*x*-axis) samples for all measured splice events in the developmental dataset. The 447 splice events with significant differences between groups ($|\Delta\Psi| > 0.2$, $P < 0.01$ by rank-sum test) are in black. Events shaped as a blue-colored square were significantly different in both datasets and those with the largest mean $\Delta\Psi$ in the DM1 dataset are labeled by their name (see also Figure 2). (**D**) Scatter plot of mean Ψ in DM1 (*y*-axis) and healthy (*x*-axis) samples for all splice events measured in the DM1 dataset. The 130 splice events with significant differences between groups ($|\Delta\Psi| > 0.2$, $P < 0.01$ by rank-sum test) are in black. Further labeling also as in (C). (**E**) Venn diagram of splice events that featured a marked and significantly different change ($|\Delta\Psi| > 0.2$, $P < 0.01$ by rank-sum test) between prenatal and postnatal samples (i.e. related to brain development), and between DM1 and unaffected samples (i.e. related to DM1). (**F**) Histogram of splice events significantly altered in both the BrainSpan and the DM1 datasets (left bar) of in just one of the datasets (right bar). The splicing direction indicates whether exon inclusion increased or decreased in the same or opposite direction when comparing the DM1 and prenatal group to the unaffected and postnatal group. The direction of change for some events is marked as 'unknown' due to missing data in one of the datasets. Here, events were considered significant when $P < 0.01$ by rank-sum test in both datasets. The 34 high-confidence events all show the same direction of change.

lapsed annotation file which lists only a single transcript per gene (35). The *collapse_annotation.py* script from GTEx was used to collapse the comprehensive GENCODE v26 gene annotation. Pathway analysis with the May 2021 release of Wikipathways (37) was performed using ClueGO v2.5.8 (38) with GO term fusion and the Bonferroni right-sided post test.

Calculation of percent spliced-in

To estimate exon inclusion for exon skipping events, percent spliced-in (PSI, Ψ) values were computed with MISO 0.5.4 (39). MISO estimates the Ψ for a given splice event by using Bayesian inference on a combination of inclusion reads, exclusion reads, junction-spanning reads and upstream/downstream exonic reads. Splice events which were not detected in each sample within a dataset were excluded from that dataset. The function *gff_make_annotation* of rnaseqlib 0.1 (40) was used to generate MISO annotation files of the most recent genome build (hg38) based on the comprehensive GENCODE v26 annotation in the UCSC genePred format.

Calculation of correlations and partial correlations

Ordinal relationships between estimates of exon inclusion and RNA expression were assessed by computing the Spearman's rank correlation. The function *corr.test* of the psych 1.8.12 R package was used to compute the correlation coefficients, to test for significant correlations and to control the false discovery rate (FDR) at 5% with the Benjamini–Hochberg procedure (41,42). Correlations were assessed (1) between Ψ estimates of selected splice events and age of the sample donor and (2) between Ψ estimates and RNA expression of selected RBPs (i.e. *CELF1*, *MBNL1*, *MBNL2*) and (3) between Ψ estimates of all high-confidence splice event pairs (i.e. pairwise cross-correlations). To account for the influence of other variables on the pairwise correlations between splice events, partial correlations were computed with the ppcor 1.1 R package (43). A partial correlation represents the association of two variables (e.g. inclusion of exon A and exon B) while removing the effect of one or more other variables (e.g. RNA expression of gene A and/or gene B). The age of the sample donor, the disease state (DM1 or unaffected) and the RNA expression level of the RBPs were separately taken into account for the partial correlation analysis. The resulting controlled partial correlations were compared to the zero-order pairwise correlations.

Data availability/sequence data resources

No new sequence data were generated. Data access for the BrainSpan Atlas of the Developing and Adult Human Brain and GTEx project must be requested from dbGaP (BrainSpan study accession: phs000731.v2.p1; GTEx study accession: phs000424.v8.p2). Raw gene-level counts for GTEx can be downloaded from the public GTEx portal (release V8, https://gtexportal.org/home/). The RNA-Seq data of Otero *et al*. (2021) are available on GEO (study accession: GSE157428).

Data availability/novel programs, software, algorithms

The R scripts that were used to analyze the processed RNA-Seq data are available in the GitHub repository (https://github.com/cmbi/BrainDM1).

Web sites/data base referencing

- STAR 2.7.0f (https://github.com/alexdobin/STAR/releases/tag/2.7.0f)
- Samtools 1.9 (http://www.htslib.org/doc/1.9/samtools.html)
- RNASeqQC 2.3.5 (https://github.com/getzlab/rnaseqc/releases/tag/v2.3.5)
- MISO 0.5.4 (https://pypi.python.org/pypi/misopy/0.5.4)
- rnaseqlib 0.1 (https://github.com/yarden/rnaseqlib)
- ClueGO 2.5.8 (https://apps.cytoscape.org/download/cluego/2.5.8)
- edgeR 3.28.0 (https://bioconductor.org/packages/release/bioc/html/edgeR.html)
- psych 1.8.12 (https://cran.r-project.org/src/contrib/Archive/psych/psych_1.8.12.tar.gz)
- ppcor 1.1 (https://cran.r-project.org/web/packages/ppcor/index.html).
- *collapse_annotation.py* script (https://github.com/broadinstitute/gtex-pipeline/tree/master/gene_model)
- GENCODE release v26 annotation (https://www.gencodegenes.org/human/release_26.html)
- GENCODE v26 annotation in UCSC genePred format (http://hgdownload.soe.ucsc.edu/goldenPath/hg38/database/wgEncodeGencodeCompV26.txt.gz)

Statistical analyses

P values were corrected for multiple comparisons by controlling the FDR at 5% with the Benjamini–Hochberg procedure. Significant group differences in exon inclusion or gene expression were assessed by the nonparametric Wilcoxon rank-sum test. For the transcriptome-wide comparison, *P* values were not FDR-corrected and splice events were considered significant when $P < 0.01$ and $|\Delta\Psi| > 0.2$ between groups. Based on these cutoffs and random shuffling of samples, we estimated an FDR of below 5%. A significant overlap between sets of splice events was assessed by the Fisher's exact test. Spearman's rank correlation was chosen in favor of Pearson correlation since the Ψs for the majority of splice events were skewed and not normally distributed, as assessed by the Shapiro–Wilk test ($P < 0.05$).

RESULTS

Splicing patterns in the frontal cortex of adult DM1 patients resemble those of developing brains

DM1-related aberrant splicing has been associated with the expression of fetal transcript variants (17,26,44–47). We performed a comprehensive evaluation of the developmental regulation of DM1-related exon skipping events by comparing exon inclusion in frontal cortex samples from the human BrainSpan dataset (Figure 1A) and from a recent dataset including 21 DM1 patients and 8 controls (Figure 1B) (23,31). PSI (Ψ) values were calculated for all splice

Table 1. Overview of the 34 DM1-relevant developmental splice events in the frontal cortex.

Event	Skipped exon coordinates	Length	Inclusion in DM1
MBNL1	chr3:152446704-152446757	54	Increased
ADD1	chr4:2926642-2926675	34	Increased
TACC2	chr10:122245065-122245073	9	Increased
MBNL2a	chr13:97356796-97356849	54	Increased
MBNL2b	chr13:97366459-97366553	95	Increased
MBNL2c	chr13:97366459-97366553	95	Increased
TCF3	chr19:1615285-1615485	201	Increased
PLA2G6	chr22:38128269-38128430	162	Increased
SOS1	chr2:38989270-38989314	45	Decreased
ATP1B3	chr3:141902136-141902227	92	Decreased
LRRFIP2	chr3:37091467-37091538	72	Decreased
PALLD	chr4:168925233-168925278	46	Decreased
PACRGL	chr4:20702141-20702221	81	Decreased
SEPT11	chr4:77036773-77036837	65	Decreased
GABRG2	chr5:162151730-162151753	24	Decreased
NRCAM	chr7:108166921-108167073	153	Decreased
GOLGA2	chr9:128272785-128272865	81	Decreased
SORBS1a	chr10:95351209-95351376	168	Decreased
SORBS1b	chr10:95351216-95351376	161	Decreased
SORBS1c	chr10:95375973-95376056	84	Decreased
NUMA1	chr11:72012401-72012442	42	Decreased
CAMKK2	chr12:121244573-121244615	43	Decreased
DNM1L	chr12:32705825-32705863	39	Decreased
PPHLN1	chr12:42384940-42384996	57	Decreased
DCLK1	chr13:35788212-35788285	74	Decreased
PACS2	chr14:105385685-105385717	33	Decreased
TJP1	chr15:29719777-29720016	240	Decreased
KIFC3	chr16:57759125-57759153	29	Decreased
MTSS1L	chr16:70679315-70679323	9	Decreased
ARHGAP44	chr17:12973302-12973319	18	Decreased
MAPT	chr17:45971859-45971945	87	Decreased
CSNK1D	chr17:82245976-82246039	64	Decreased
DLGAP1	chr18:3656084-3656113	30	Decreased
DMD	chrX:31126642-31126673	32	Decreased

All events showed a significant difference ($|\Delta\Psi| > 0.2$, $P < 0.01$ by rank-sum test) in frontal cortex samples between DM1 patients and unaffected controls and between prenatal and postnatal healthy individuals. Genomic coordinates are based on the hg38 genome build. Note that the coordinates for MBNL2b and MBNL2c are duplicated because these events share the same skipped exon but varying up- and downstream exons (see Supplementary Table S4).

events and reflect the inclusion rate of a cassette exon that is either included or skipped within the boundaries of its flanking upstream and downstream exons.

In the developing brain dataset, we identified 447 events with significant differences ($|\Delta\Psi| > 0.2$, $P < 0.01$ by rank-sum test) between prenatal and postnatal frontal cortex (Figure 1C and Supplementary Table S2). One hundred thirty events displayed significant differences ($|\Delta\Psi| > 0.2$, $P < 0.01$ by rank-sum test) between DM1 and healthy frontal cortex samples (Figure 1D; Supplementary Table S3; (23)). A significant overlap of 34 exons ($P < 2.2$ e-16, Fisher's exact test) was found between the sets of splice events in the developing brains and the DM1 brains (Figure 1E; Table 1; Supplementary Table S4, Supplementary Figure S2 (48)). The absolute $\Delta\Psi$ for DM1 versus control was on average larger in this subset of 34 events compared to all 130 events related to DM1 ($P < 0.01$ by rank-sum test, Supplementary Figure S3). These 34 events were present in 30 different genes and involved 32 unique cassette exons. When there were multiple events in the same gene, these events were labeled with a/b/c etc. (see Supplementary Table S4 for the exact nomenclature). Two cassette exons were present in multiple events differing by a variable 5′ end (SORBS1a, SORBS1b) or the upstream exons (MBNL2b, MBNL2c). Compared to the background of all differentially spliced genes in the DM1 dataset, these 30 were enriched (Bonferroni adj. $P = 0.025$) for genes in the Wikipathway 'ectoderm differentiation', represented by *DMD*, *KIFC3*, *NUMA1* and *TCF3* (37). Out of all splice events that changed significantly in one or both datasets, 337 events, including the overlap of 34 events, changed toward the prenatal inclusion pattern in the DM1 brains, whereas only 130 events changed in a direction that resembles the postnatal state (Figure 1F). Together this confirms, on a transcriptome-wide scale, previous observations that DM1 brains reflect embryonic/fetal splicing patterns. The list of 34 splice events contains many of the previously described aberrantly spliced exons in DM1 and known MBNL targets, such as *MBNL1* exon 5, *MBNL2* exon 5, *DMD* exon 78 and *MAPT* exon 3.

Closer inspection of the developmental regulation of splicing for these 34 splice events demonstrated gradual changes from the fetal through the perinatal to the postnatal period. In Figure 2, we display the splice events with the largest mean difference in Ψ between the DM1 frontal cortex and that of controls. For all splice events with an increase of Ψ during development, Ψ values were lower in DM1 patients than in controls, and vice versa. The strongest increases in Ψ throughout development were observed in *ARHGAP44*, *ATP1B3*, *MAPT* and *SORBS1*

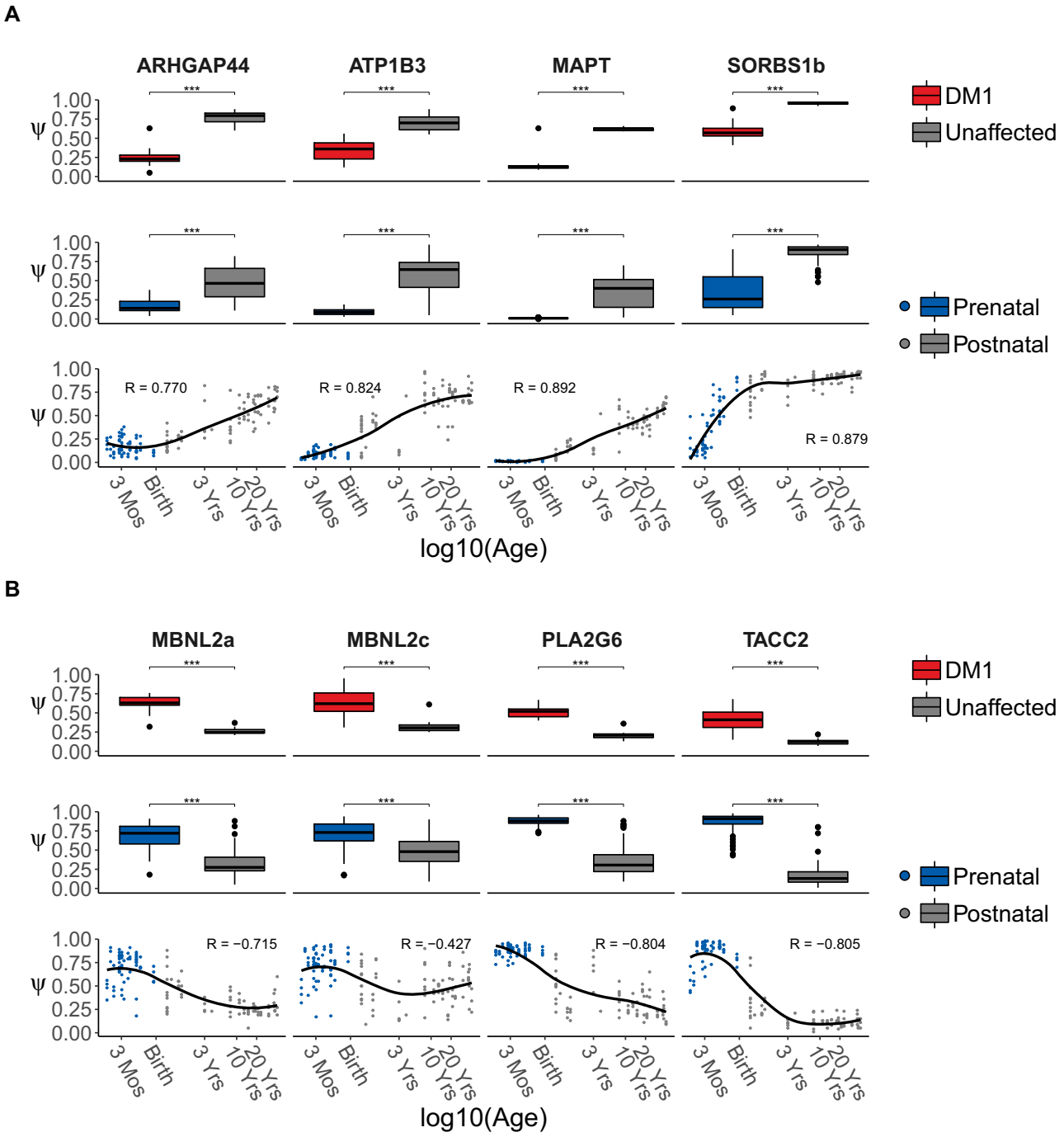

Figure 2. Splice events with strongest up- and downregulation in DM1 show gradual change through development. Splice events with the largest decrease (**A**) or increase (**B**) in Ψ in DM1 patients compared to unaffected adults. The top rows in panels (A) and (B) show boxplots of Ψs for events selected for the largest difference in mean Ψ between frontal cortex samples from DM1 patients (in red) and unaffected adults (in gray). The boxplots in the middle rows are based on frontal cortex samples from healthy prenatal (in blue) and postnatal donors (in gray). All samples collected before and after birth were pooled to form the prenatal and postnatal groups, respectively. A significant difference between groups was assessed by the rank-sum test. *P*-values were FDR-corrected with the Benjamini–Hochberg procedure (ns: $P > 0.05$, ***: $P < 0.001$). The bottom row shows scatterplots of Ψ for the selected events in the healthy, developing human frontal cortex. Post-conceptional age was recorded in days and plotted on a log10 scale. Arbitrary *x*-axis labels were chosen to highlight specific timepoints during development (Mos = Months, Yrs = Years). The regression curve was estimated by the LOESS method.

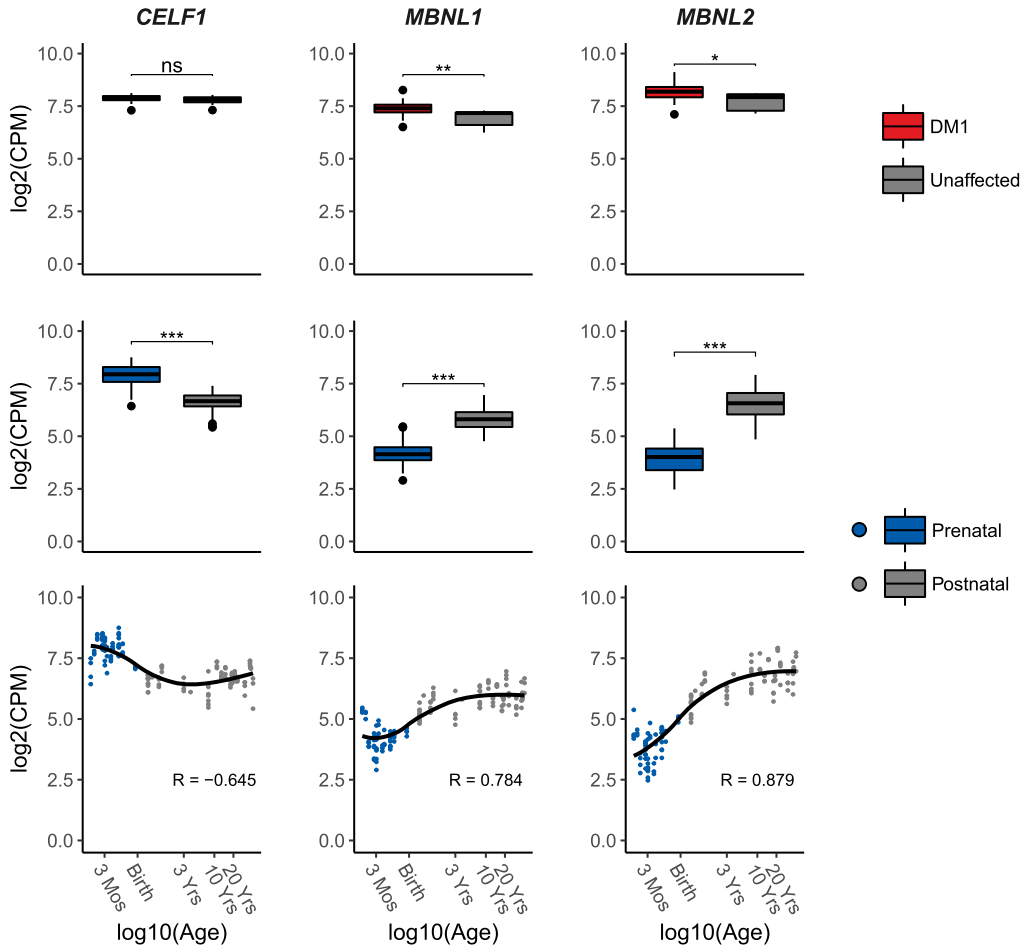

Figure 3. Differences in RNA expression of DM1-relevant splicing factors throughout human brain development and between DM1 patients and unaffected adults. The top row shows boxplots of RNA expression (log2 of CPM) of *CELF1, MBNL1 and MBNL2* in frontal cortex samples from DM1 patients (in red) and unaffected adults (in gray). The boxplots in the middle row are based on frontal cortex samples from healthy prenatal (in blue) and postnatal donors (in gray). Significant differences between groups were assessed by the rank-sum test. P-values were FDR-corrected with the Benjamini–Hochberg procedure (ns: $P > 0.05$, *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$). The bottom row shows scatterplots of RNA expression (log2 CPM) of *CELF1, MBNL1 and MBNL2* in the healthy, developing frontal cortex. Post-conceptional age was plotted on a log10 scale. *x*-axis labels were chosen to highlight arbitrary timepoints during development (Mos = Months, Yrs = Years). The regression curve was estimated by the LOESS method.

transcripts, whilst the strongest decreases in Ψ were observed in MBNL2, *PLA2G6* and *TACC2*. For most of these genes, pronounced changes in splicing patterns were observed during the perinatal period, but the timing and the extent to which these patterns changed from the perinatal period to the adult stage differed considerably between exons. For an overview of Ψs of these exons across sex and CNS subregions, see Supplemental Figures S4 and S5, respectively.

Expression of *MBNL* and *CELF1* in the developing and DM1 brain

Since RBPs are known to regulate alternative splicing during human tissue development, we inspected the RNA expression levels of RBPs that have been linked to DM1 pathophysiology, i.e. those of the MBNL and CELF families, throughout development and in the disease situation (7,14). For *CELF1*, as well as *CELF2, -3, -5* and *-6*, we observed a significant (FDR < 0.05) downregulation in prenatal and postnatal samples (Figure 3 and Supplementary Figure S6; log2 fold change for CELF1 ($\log FC_{CELF1}$) = -1.23). On the contrary, for *MBNL1* and *MBNL2*, as well as *CELF4*, expression increased gradually and significantly (FDR < 0.05; $\log FC_{MBNL1} = 1.60$; $\log FC_{MBNL2} = 2.66$) during development, whereas no change was observed for *MBNL3* (Supplementary Figure S7). The RNA expression levels of *CELF1*, *-3* and *-6* did not differ between DM1 cases and controls, while a slight upregulation of *MBNL1*, *-2* and *-3* was found in DM1 ($\log FC_{CELF1} = -0.09$; $\log FC_{MBNL1} = -0.50$; $\log FC_{MBNL2} = -0.44$, $\log FC_{MBNL3} = -0.60$). For CELF2, -4 and -5 we noticed a subtle downregulation in DM1 (Supplementary Figure S6). Since *MBNL1/2* and *CELF1* are best characterized in the context of DM1, and most of their family member featured similar or only modestly different effects, we focused on *CELF1* and *MBNL1/2* for further analyses. Lastly, expression levels of *MBNL1/2* and *CELF1* slightly differed between sexes during development but not when comparing DM1 cases to unaffected controls (Supplementary Figure S8).

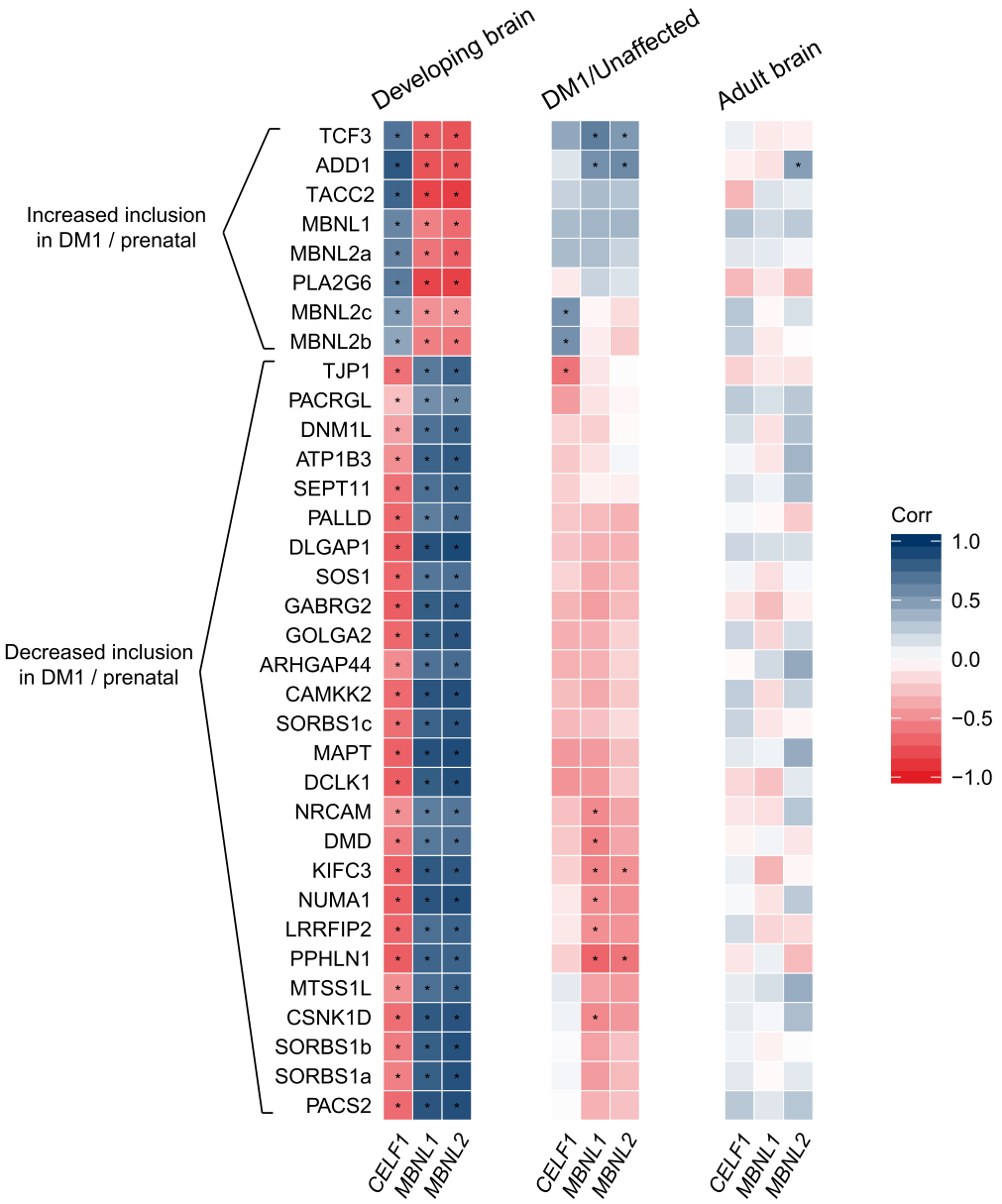

Figure 4. Correlation of alternative splicing and RNA expression of *CELF1*, *MBNL1* and *MBNL2*. Correlations were computed between Ψ of high-confidence exon skipping events and RNA expression levels of *CELF1, MBNL1* and *MBNL2* within the same samples from the healthy, developing frontal cortex (left), from the frontal cortex of DM1 patients and unaffected adults (middle) and from the healthy, adult frontal cortex (right). The splice events are hierarchically clustered based on the average distance between the correlations for the DM1/unaffected samples. The color scale on the right reflects the value of the Spearman's rank correlation coefficient. Asterisks indicate a significant correlation (per study FDR-corrected $P < 0.05$).

DM1-relevant developmental splice events are associated with expression of *CELF1* and *MBNL1/2* mRNA

Next, we assessed the potential effects of the balance between *CELF1* and *MBNL* mRNA expression on the splicing patterns of the set of 34 DM1-relevant developmental splice events. To achieve this, we calculated correlations between *CELF1* and *MBNL* expression on the one hand, and the Ψ of these splice events on the other hand. In the developing brain, we observed that all 8 splice events with increased inclusion in the prenatal state and in DM1 patients were strongly positively correlated with the expression of *CELF1* and strongly negatively correlated with the expression of *MBNL1* and *MBNL2* (Figure 4 and Supplementary Figure S9, examples in Supplementary Figure S10A). Conversely, the 26 events with increased inclusion in the postnatal state and decreased inclusion in DM1 patients were strongly negatively correlated with the expression of *CELF1* and strongly positively regulated with the expression of *MBNL1* and *MBNL2*. Such correlations were much weaker within the sets of prenatal or postnatal samples, indicating that changes during development are driv-

ing these correlations (Supplementary Figure S11). Collectively, these findings indicate a high level of coordination within this set of splice events during brain development, and suggest an association with the balance between the activity of the CELF1 and MBNL splicing factors.

In the dataset of DM1 and control adult frontal cortex samples, we did not observe similarly strong correlations between the inclusion rates of the 34 DM1-relevant developmental splice events and the expression levels of *CELF1* and *MBNL1/2* (Figure 4 and Supplementary Figure S9, examples in Supplementary Figure S10B). Given the low overall variation of *CELF1* and *MBNL1/2* between DM1 patients, recall Figure 3, this was not unexpected. However, upon closer inspection of these correlations, we can appreciate that the set of 8 splice events with increased inclusion in the prenatal state and in DM1 patients were positively correlated with the RNA expression of *CELF1*, *MBNL1* and *MBNL2*. Again, the 26 events with increased inclusion in the postnatal state and decreased inclusion in DM1 showed the opposite pattern. The correlations between *MBNL1/2* expression and splicing in the DM1 and unaffected brain showed opposite trends compared to the correlations observed in the developing brain. Notably, for the set of 96 splice events, which were found to be relevant to DM1 but not significantly developmentally regulated (Figure 1E), the same trends were present (Supplementary Figure S12). These trends were not observed when analyzing the DM1 and unaffected brain separately (Supplementary Figure S13).

Since the DM1 dataset contained only 8 control individuals, we performed a similar analysis in the much larger set of GTEx frontal cortex samples from healthy adult donors, which has a higher statistical power ($N = 54$, Figure 4). In this dataset only splicing of the exon in *ADD1*, a known MBNL2 target (17), correlated significantly to *MBNL2* expression. We found no further significant correlations, indicating that the degree of splicing regulation during brain development is much larger than the interindividual differences in splicing regulation in the adult brain.

Prenatal-like splicing patterns in the DM1 brain are not driven by variation in mRNA expression of *CELF1* and *MBNL1/2*

We finally set out to investigate to what extent the coordinated splicing differences observed in the developing and the DM1 frontal cortex can be explained by variation in sample donor age and splicing factor RNA expression. For this analysis, we calculated pairwise correlations between the exon inclusion levels of all 34 high-confidence splice events. A pairwise correlation of two splice events captures the degree to which the two exons are both included or excluded (i.e. strong positive correlation) or are mutually exclusive (i.e. strong negative correlation). Next, we controlled the pairwise correlations for variation in age, splicing factor expression and disease state by calculating the partial correlations and compared the distribution of partial correlations to the zero-order (uncorrected) correlations (Figure 5). A difference between the partial and zero-order correlation suggests that the zero-order correlation can be partially explained by one of the factors that are controlled for in the partial correlation. In the context of the DM1 frontal cortex, we also considered the disease state (DM1 or unaffected).

As expected, controlling for age in the developing brain cohort, that features age as its main contrast, resulted in a strong decrease in pairwise correlations. Since the expression of *CELF1* and *MBNL1/2* was strongly related to the age of the subject, as shown in Figure 3, these factors could not be assessed independently and showed the same effect on the pairwise correlations. Nonetheless, the data support the notion that the coordinated changes in splicing during development are linked to a shift in the balance between *CELF1* and *MBNL1/2* expression. In the DM1 cohort, pairwise correlations were strongly decreased after controlling for disease state but remained similar after controlling for age or splicing factor expression. Thus, we conclude that coordination of splicing in the DM1 frontal cortex is driven by differences in controls vs DM1 patients. In the healthy adult brain dataset pairwise correlations were generally low and were unaltered by controlling for the factors mentioned above.

DISCUSSION

In this work, we offer a comprehensive overview of DM1 splice abnormalities in the frontal cortex that are related to development. We show that the majority of dysregulated splice events in frontal cortices in DM1 patients are changed toward prenatal splice variants, and identify 34 events with high confidence and a large shift in Ψ. This transcriptome-wide analysis provides insight in how DM1-pathophysiology reflects a reversed developmental situation.

Various studies in humans and model organisms have shown that brain development goes hand-in-hand with changes in alternative splicing (14,49). This was also observed in our analysis of the BrainSpan dataset, where we identified around 450 splice events that were different between prenatal and postnatal human frontal cortex. Of the 130 splice events that were dysregulated in the DM1 frontal cortex, 34 overlapped with the ∼450 developmentally regulated splice events, a significant enrichment. We focus on these 34 events because we assume that splice events that are developmentally regulated are more likely to have a physiological effect. Among these were known DM1-related splice events such as those in *ADD1*, *CAMKK2*, *DMD*, *MAPT*, *MBNL1*, *MBNL2* and *SORBS1* (21,50), as well as several novel events. The splice events in *GABRG2*, *NRCAM*, *ARHGAP44*, *DLGAP1* and *CSNK1D* were also found misspliced ($|\Delta\Psi| \geq 0.2$, FDR ≤ 0.05) by Goodwin *et al*. in a previous RNA-Seq study on the DM1 frontal cortex, adding confidence to these exons for their relevance for DM1 (20).

Among the 34 events, altered splicing of *GABRG2*, *TCF3*, *NRCAM*, *ARHGAP44*, *DLGAP*, *SEPT11*, *DCLK1* and *CSNK1D* may be most relevant for the DM1 brain phenotypes given their known functions in the CNS. Of these, only the events in *SEPT11*, *DCLK1* and *CSNK1D* lead to a change in reading frame (exon length is not divisible by three), potentially leading to mRNA degradation or dysfunctional proteins. Still, e.g. for *DCLK1* the identified splice event results in a different C-terminus of the protein,

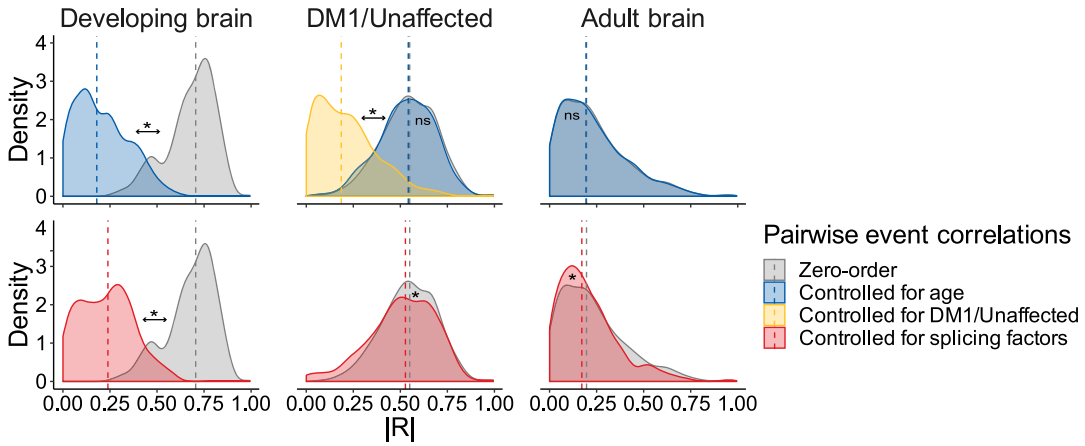

Figure 5. Pairwise correlations of splice events are partially explained by age differences and splicing factor RNA expression in the developing frontal cortex but not in the DM1 or unaffected adult frontal cortex. Pairwise (cross-) correlations were computed between the Ψ of all 34 high-confidence splice events. The density plots show the distribution of the absolute Spearman's rank correlation coefficients for frontal cortex samples from the healthy, developing brain (left), from DM1 patients and unaffected adults (middle) and from the healthy, adult brain (right). Zero-order (i.e. uncorrected) pairwise correlations (in gray) were separately controlled for the following variables: age of the sample donor (in blue), disease state (DM1 or unaffected; in yellow) or RNA expression of *CELF1, MBNL1* and *MBNL2* (in red). Differences between zero-order distributions and their corrected counterparts were assessed by the rank-sum test. *P*-values were FDR-corrected with the Benjamini–Hochberg procedure (ns: $P > 0.05$, *: $P < 0.05$). The dashed vertical line indicates the median correlation.

differentiating between DLCK1-alpha and -beta, of which it is known that alpha form is dominant in the prenatal brain (51). Of note, *DCLK1* (isoform-) expression is associated with cognitive abilities (52) and anxiety (53).

Regarding *CSNK1D* (Casein kinase I isoform delta), exon 9 inclusion prevents further downstream translation of exon 10 by introducing a stop codon (54). Interestingly, this transcript variant alters the circadian rhythm and sleep alterations are a prevalent DM1 phenotype (55,56). Ongoing efforts aim to develop small molecules for modulating CSNK1D activity but target specificity remains a major challenge (54).

GABRG2 (Gamma-Aminobutyric Acid Type A Receptor Subunit Gamma 2) has been associated with a plethora of neuro(developmental) disorders, mostly related to epilepsy (57–59). *GABRG2* is known to be differentially spliced during development, where the short variant is expressed in early development (60). This alternatively spliced (micro-) exon is 24 nucleotides long and codes for an 8-amino acid stretch that alters GABA[A]R complex composition and modulates its activity (61). MS-based evidence has been gathered that, at least in rats, both isoforms are translated into protein (62). In schizophrenic patients the ratio of short (γ2S) versus long (γ2L) GABRG2 isoforms is also altered, although in a way opposite to the DM1-situation presented in the data from Otero *et al.*, with a relative increase of the short isoform (63). Additionally, *GABRG2* variants have been observed in individuals affected by Dravet syndrome, autism spectrum disorder, developmental delay and intellectual disability (58,64). Of major interest is that GABRG2 is a known drug target, and that the ratio of different isoforms can be altered via drugs (65–67).

Although not identified as dysregulated in the Goodwin study (20), *TCF3* (Transcription Factor 3, or E2-alpha [E2A]) is a hit with a possibly important impact. TCF3 functions as a transcriptional regulator in neuronal differentiation, with an impact on many other genes (68). Interestingly, TCF3 can regulate IL-6 signaling, which is disturbed in DM1 (69,70). Since there are differences in splicing aberrations between brain regions in DM1, our findings may not always be representative for other brain regions than the frontal cortex (71).

We analyzed the expression of the *CELF* and *MBNL* mRNAs throughout development and in the frontal cortex of adult DM1 patients and controls. It is important to note that *CELF* and *MBNL* family members are strongly post-transcriptionally and post-translationally regulated (72–75), meaning that extrapolations of mRNA level to protein expression and activity should be interpreted with caution (9,30,76). In the future, it would be of high interest to carry out an analysis as presented here on proteomics data, which are currently scarce.

In the literature, MBNL2 is generally considered the brain-related member of the MBNL family, but it is also known that MBNL1 can compensate for loss of MBNL2 function, or even plays an important role in the brain development itself (17,77–78). We noted that *MBNL1*, but not *MBNL3*, is also expressed in the human brain throughout development at levels comparable to those of *MBNL2*. This is confirmed via the Allen brain map and by recent single cell analyses (79,80).

Consistent with data from the mouse heart (15), we observed a gradual increase in *MBNL1/2* mRNA expression and a decrease of *CELF1* expression with age in the human brain. The balance in MBNL1/2 and CELF activity during development is likely explained by these different transcription levels, suggesting that regulation of *MBNL* expression during development is mainly transcriptional, while post-transcriptional regulation of *MBNL* expression may dominate in adults (72–75). This is especially relevant because many exons are sensitive to specific levels of MBNL1/2 activity (81). In turn, alterations in this balance

are a likely cause of the regulation of a proportion of the developmentally regulated splicing events, as was demonstrated in the brains of Mbnl2 knock-out mice and muscles of Mbnl1 knock-out mice. In addition, there are exons in both *MBNL1* (exon 5) and *MBNL2* (exon 5 and 8) that are alternatively spliced throughout development (17,82). These are likely part of an autoregulatory mechanism controlling the localization and activity of these splicing factors (9). *MBNL1* exon 1 is also involved in MBNL1 autoregulation (9), but we found no differences in splicing in any of our analyses (data not shown).

In DM1 patients, we observed a switch to fetal splice isoforms, but this was not accompanied by a decrease but by a slight increase in the *MBNL1/2* mRNA expression, while the *CELF1* mRNA expression was unchanged. This confirms the basis of the altered balance between MBNL and CELF in DM1 patients: on the one hand MBNL loss-of-function is mainly caused by the differential splicing of *MBNL1/2* and/or the entrapment of MBNL1/2 proteins in foci which is insufficiently compensated by an increase in *MBNL1/2* transcription levels, and on the other hand the increased activity of CELF1/-2 is due to increased protein levels and/or their phosphorylation level (10,83). In summary, MBNL1/2 and CELF1 seem transcriptionally regulated during development, but are heavily post-transcriptionally regulated in adult brains and muscles, e.g. through altered splicing and altered cellular localization (MBNL1/2) and post-translational modification (CELF1) (9,84,85).

In conclusion, we connected developmentally regulated splicing events with those dysregulated in DM1. This provides insights into the DM1 disease mechanism and helps to prioritize splice events for further investigation.

DATA AVAILABILITY

All software used in the RNA-Seq data processing can be acquired via the following links:

- STAR 2.7.0f (https://github.com/alexdobin/STAR/releases/tag/2.7.0f)
- Samtools 1.9 (http://www.htslib.org/doc/1.9/samtools.html)
- RNASeqQC 2.3.5 (https://github.com/getzlab/rnaseqc/releases/tag/v2.3.5)
- MISO 0.5.4 (https://pypi.python.org/pypi/misopy/0.5.4)
- rnaseqlib 0.1 (https://github.com/yarden/rnaseqlib)
- ClueGO 2.5.8 (https://apps.cytoscape.org/download/cluego/2.5.8)
- edgeR 3.28.0 (https://bioconductor.org/packages/release/bioc/html/edgeR.html)
- psych 1.8.12 (https://cran.r-project.org/src/contrib/Archive/psych/psych_1.8.12.tar.gz)
- ppcor 1.1 (https://cran.r-project.org/web/packages/ppcor/index.html).

The *collapse_annotation.py* script can be downloaded from the GTEx GitHub repository (https://github.com/broadinstitute/gtex-pipeline/tree/master/gene_model). GENCODE release v26 annotation files are available under https://www.gencodegenes.org/human/release_26.html. The comprehensive GENCODE v26 annotation in the UCSC genePred format can be retrieved via http://hgdownload.soe.ucsc.edu/goldenPath/hg38/database/wgEncodeGencodeCompV26.txt.gz.

The R scripts that were used to create all figures and tables are available in the GitHub repository (https://github.com/cmbi/BrainDM1).

Data access for the BrainSpan Atlas of the Developing and Adult Human Brain and GTEx project must be requested from dbGaP (BrainSpan study accession: phs000731.v2.p1; GTEx study accession: phs000424.v8.p2). Raw gene-level counts for GTEx can be downloaded from the public GTEx portal (release V8, https://gtexportal.org/home/). The RNA-Seq data of Otero *et al*. (2021) is available on GEO (study accession: GSE157428).

SUPPLEMENTARY DATA

Supplementary Data are available at NARGAB Online.

ACKNOWLEDGEMENTS

The BrainSpan datasets used for the analysis described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000731.v2.p1. Submission of the data, phs000731.v2.p1, to dbGaP was provided by Dr. Nenad Sestan. Collection of the data and analysis was supported by grants from the National Institutes of Health (MH089929, MH081896, and MH090047). Additional support was provided by the Kavli Foundation, a James S. McDonnell Foundation Scholar Award, NARSAD, and the Foster-Davis Foundation.
The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health (commonfund.nih.gov/GTEx). Additional funds were provided by the NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. Donors were enrolled at Biospecimen Source Sites funded by NCI\Leidos Biomedical Research, Inc. subcontracts to the National Disease Research Interchange (10XS170), Roswell Park Cancer Institute (10XS171), and Science Care, Inc. (X10S172). The Laboratory, Data Analysis, and Coordinating Center (LDACC) was funded through a contract (HHSN268201000029C) to the The Broad Institute, Inc. Biorepository operations were funded through a Leidos Biomedical Research, Inc. subcontract to Van Andel Research Institute (10ST1035). Additional data repository and project management were provided by Leidos Biomedical Research, Inc. (HHSN261200800001E). The Brain Bank was supported supplements to University of Miami grant DA006227. Statistical Methods development grants were made to the University of Geneva (MH090941 & MH101814), the University of Chicago (MH090951, MH090937, MH101825, & MH101820), the University of North Carolina - Chapel Hill (MH090936), North Carolina State University (MH101819), Harvard University (MH090948), Stanford University (MH101782), Washington University (MH101810), and to the University of Pennsylvania (MH101822). The datasets used for the analyses described in this manuscript were obtained

from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000424.v8.p2.

FUNDING

This work was supported by an E-rare3 - JTC2018 grant to the ReCognitION consortium.
Conflict of interest statement. None declared.

REFERENCES

1. Meola,G. (2013) Clinical aspects, molecular pathomechanisms and management of myotonic dystrophies. *Acta Myol.*, **32**, 154–165.
2. Mahadevan,M., Tsilfidis,C., Sabourin,L., Shutler,G., Amemiya,C., Jansen,G., Neville,C., Narang,M., Barceló,J., O'Hoy,K. *et al.* (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science*, **255**, 1253–1255.
3. Napierała,M. and Krzyzosiak,W.J. (1997) CUG repeats present in myotonin kinase RNA form metastable 'slippery' hairpins. *J. Biol. Chem.*, **272**, 31079–31085.
4. van Cruchten,R.T.P., Wieringa,B. and Wansink,D.G. (2019) Expanded CUG repeats in DMPK transcripts adopt diverse hairpin conformations without influencing the structure of the flanking sequences. *RNA*, **25**, 481–495.
5. Goers,E.S., Purcell,J., Voelker,R.B., Gates,D.P. and Berglund,J.A. (2010) MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.*, **38**, 2467–2484.
6. Smith,K.P., Byron,M., Johnson,C., Xing,Y. and Lawrence,J.B. (2007) Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type i is blocked at entry into SC-35 domains. *J. Cell Biol.*, **178**, 951–964.
7. Sznajder,Ł.J. and Swanson,M.S. (2019) Short tandem repeat expansions and RNA-mediated pathogenesis in myotonic dystrophy. *Int. J. Mol. Sci.*, **20**, 3365.
8. Goodwin,M. and Swanson,M.S. (2014) RNA-binding protein misregulation in microsatellite expansion disorders. *Adv. Exp. Med. Biol.*, **825**, 353–388.
9. Konieczny,P., Stepniak-Konieczna,E. and Sobczak,K. (2018) MBNL expression in autoregulatory feedback loops. *RNA Biol.*, **15**, 1–8.
10. Kuyumcu-Martinez,N.M., Wang,G.S. and Cooper,T.A. (2007) Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-Mediated hyperphosphorylation. *Mol. Cell*, **28**, 68–78.
11. Dasgupta,T. and Ladd,A.N. (2012) The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *Wiley Interdiscip. Rev. RNA*, **3**, 104–121.
12. 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.
13. Raj,B. and Blencowe,B.J. (2015) Alternative splicing in the mammalian nervous system: recent insights into mechanisms and functional roles. *Neuron*, **87**, 14–27.
14. Baralle,F.E. and Giudice,J. (2017) Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.*, **18**, 437–451.
15. Kalsotra,A., Xiao,X., Ward,A.J., Castle,J.C., Johnson,J.M., Burge,C.B. and Cooper,T.A. (2008) A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc. Natl. Acad. Sci. USA*, **105**, 20333–20338.
16. Weyn-Vanhentenryck,S.M., Feng,H., Ustianenko,D., Duffié,R., Yan,Q., Jacko,M., Martinez,J.C., Goodwin,M., Zhang,X., Hengst,U. *et al.* (2018) Precise temporal regulation of alternative splicing during neural development. *Nat. Commun.*, **9**, 2189.
17. Charizanis,K., Lee,K.Y., Batra,R., Goodwin,M., Zhang,C., Yuan,Y., Shiue,L., Cline,M., Scotti,M.M., Xia,G. *et al.* (2012) Muscleblind-like 2-Mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron*, **75**, 437–450.
18. Du,H., Cline,M.S., Osborne,R.J., Tuttle,D.L., Clark,T.A., Donohue,J.P., Hall,M.P., Shiue,L., Swanson,M.S., Thornton,C.A. *et al.* (2010) Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.*, **17**, 187–193.
19. Freyermuth,F., Rau,F., Kokunai,Y., Linke,T., Sellier,C., Nakamori,M., Kino,Y., Arandel,L., Jollet,A., Thibault,C. *et al.* (2016) Splicing misregulation of SCN5A contributes to cardiac-conduction delay and heart arrhythmia in myotonic dystrophy. *Nat. Commun.*, **7**, 11067.
20. Goodwin,M., Mohan,A., Batra,R., Lee,K.Y., Charizanis,K., Fernández Gómez,F.J., Eddarkaoui,S., Sergeant,N., Buée,L., Kimura,T. *et al.* (2015) MBNL sequestration by toxic RNAs and RNA misprocessing in the myotonic dystrophy brain. *Cell Rep.*, **12**, 1159–1168.
21. Nakamori,M., Sobczak,K., Puwanant,A., Welle,S., Eichinger,K., Pandya,S., Dekdebrun,J., Heatwole,C.R., McDermott,M.P., Chen,T. *et al.* (2013) Splicing biomarkers of disease severity in myotonic dystrophy. *Ann. Neurol.*, **74**, 862–872.
22. Wang,E.T., Treacy,D., Eichinger,K., Struck,A., Estabrook,J., Wang,T.T., Bhatt,K., Westbrook,T., Sedehizadeh,S., Ward,A. *et al.* (2018) Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Hum. Mol. Genet.*, **28**, 1312–1321.
23. Otero,B.A., Poukalov,K., Hildebrandt,R.P., Thornton,C.A., Jinnai,K., Fujimura,H., Kimura,T., Hagerman,K.A., Sampson,J.B., Day,J.W. *et al.* (2021) Transcriptome alterations in myotonic dystrophy frontal cortex. *Cell Rep.*, **34**, 108634.
24. Tanner,M.K., Tang,Z. and Thornton,C.A. (2021) Targeted splice sequencing reveals RNA toxicity and therapeutic response in myotonic dystrophy. *Nucleic Acids Res.*, **49**, 2240–2254.
25. Charlet-B.,N., Savkur,R.S., Singh,G., Philips,A.V., Grice,E.A. and Cooper,T.A. (2002) Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cell*, **10**, 45–53.
26. Jiang,H., Mankodi,A., Swanson,M.S., Moxley,R.T. and Thornton,C.A. (2004) Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.*, **13**, 3079–3088.
27. Savkur,R.S., Philips,A.V. and Cooper,T.A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat. Genet.*, **29**, 40–47.
28. Brinegar,A.E. and Cooper,T.A. (2016) Roles for RNA-binding proteins in development and disease. *Brain Res.*, **1647**, 1–8.
29. Ho,T.H., Charlet-B,N., Poulos,M.G., Singh,G., Swanson,M.S. and Cooper,T.A. (2004) Muscleblind proteins regulate alternative splicing. *EMBO J.*, **23**, 3103–3112.
30. Wang,E.T., Ward,A.J., Cherone,J.M., Giudice,J., Wang,T.T., Treacy,D.J., Lambert,N.J., Freese,P., Saxena,T., Cooper,T.A. *et al.* (2015) Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.*, **25**, 858–871.
31. Miller,J.A., Ding,S.L., Sunkin,S.M., Smith,K.A., Ng,L., Szafer,A., Ebbert,A., Riley,Z.L., Royall,J.J., Aiona,K. *et al.* (2014) Transcriptional landscape of the prenatal human brain. *Nature*, **508**, 199–206.
32. Ardlie,K.G., DeLuca,D.S., Segrè,A.V., Sullivan,T.J., Young,T.R., Gelfand,E.T., Trowbridge,C.A., Maller,J.B., Tukiainen,T., Lek,M. *et al.* (2015) The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (80-.)*, **348**, 648–660.
33. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
34. Li,H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**, 2987–2993.
35. DeLuca,D.S., Levin,J.Z., Sivachenko,A., Fennell,T., Nazaire,M.-D., Williams,C., Reich,M., Winckler,W. and Getz,G. (2012) RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics*, **28**, 1530–1532.
36. Robinson,M.D. and Oshlack,A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.*, **11**, R25.
37. Martens,M., Ammar,A., Riutta,A., Waagmeester,A., Slenter,D.N., Hanspers,K., Miller,R.A., Digles,D., Lopes,E.N., Ehrhart,F. *et al.* (2021) WikiPathways: connecting communities. *Nucleic Acids Res.*, **49**, D613–D621.
38. Bindea,G., Mlecnik,B., Hackl,H., Charoentong,P., Tosolini,M., Kirilovsky,A., Fridman,W.-H., Pagès,F., Trajanoski,Z. and Galon,J. (2009) ClueGO: a cytoscape plug-in to decipher functionally grouped

gene ontology and pathway annotation networks. *Bioinformatics*, **25**, 1091–1093.
39. Katz,Y., Wang,E.T., Airoldi,E.M. and Burge,C.B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods*, **7**, 1009–1015.
40. Katz,Y. (2022) Rnaseqlib 0.1. https://rnaseqlib.readthedocs.io/en/clip/, (02 March 2022, date last accessed).
41. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B*, **57**, 289–300.
42. Revelle,W. (2015) Package 'psych' - Procedures for psychological, psychometric and personality research. R Package.
43. Kim,S. (2015) ppcor: an r package for a fast calculation to Semi-partial correlation coefficients. *Commun. Stat. Appl. Methods*, **22**, 665–674.
44. Dhaenens,C.M., Tran,H., Frandemiche,M.L., Carpentier,C., Schraen-Maschke,S., Sistiaga,A., Goicoechea,M., Eddarkaoui,S., Van Brussels,E., Obriot,H. *et al.* (2011) Mis-splicing of tau exon 10 in myotonic dystrophy type 1 is reproduced by overexpression of CELF2 but not by MBNL1 silencing. *Biochim. Biophys. Acta - Mol. Basis Dis.*, **1812**, 732–742.
45. Furuta,M., Kimura,T., Nakamori,M., Matsumura,T., Fujimura,H., Jinnai,K., Takahashi,M.P., Mochizuki,H. and Yoshikawa,H. (2018) Macroscopic and microscopic diversity of missplicing in the central nervous system of patients with myotonic dystrophy type 1. *Neuroreport*, **29**, 235–240.
46. Leroy,O., Wang,J., Maurage,C.A., Parent,M., Cooper,T., Buée,L., Sergeant,N., Andreadis,A. and Caillet-Boudin,M.L. (2006) Brain-specific change in alternative splicing of tau exon 6 in myotonic dystrophy type 1. *Biochim. Biophys. Acta - Mol. Basis Dis.*, **1762**, 460–467.
47. Suenaga,K., Lee,K.-Y., Nakamori,M., Tatsumi,Y., Takahashi,M.P., Fujimura,H., Jinnai,K., Yoshikawa,H., Du,H., Ares,M. *et al.* (2012) Muscleblind-Like 1 knockout mice reveal novel splicing defects in the myotonic dystrophy brain. *PLoS One*, **7**, e33218.
48. Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
49. Su,C.-H., D,D. and Tarn,W.-Y. (2018) Alternative splicing in neurogenesis and brain development. *Front. Mol. Biosci.*, **5**, 12.
50. Wang,E.T., Treacy,D., Eichinger,K., Struck,A., Estabrook,J., Wang,T.T., Bhatt,K., Westbrook,T., Sedehizadeh,S., Ward,A. *et al.* (2018) Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Hum. Mol. Genet.*, **28**, 1312–1321.
51. Burgess,H.A. and Reiner,O. (2002) Alternative splice variants of doublecortin-like kinase are differentially expressed and have different kinase activities. *J. Biol. Chem.*, **277**, 17696–17705.
52. Le Hellard,S., Håvik,B., Espeseth,T., Breilid,H., Løvlie,R., Luciano,M., Gow,A.J., Harris,S.E., Starr,J.M., Wibrand,K. *et al.* (2009) Variants in Doublecortin- and Calmodulin kinase like 1, a gene up-regulated by BDNF, are associated with memory and general cognitive abilities. *PLoS One*, **4**, e7534.
53. Schenk,G.J., Veldhuisen,B., Wedemeier,O., McGown,C.C., Schouten,T.G., Oitzl,M., de Kloet,E.R. and Vreugdenhil,E. (2010) Over-expression of δC-DCLK-short in mouse brain results in a more anxious behavioral phenotype. *Physiol. Behav.*, **101**, 541–548.
54. Xu,P., Ianes,C., Gärtner,F., Liu,C., Burster,T., Bakulev,V., Rachidi,N., Knippschild,U. and Bischof,J. (2019) Structure, regulation, and (patho-)physiological functions of the stress-induced protein kinase CK1 delta (CSNK1D). *Gene*, **715**, 144005.
55. Fustin,J.-M., Kojima,R., Itoh,K., Chang,H.-Y., Ye,S., Zhuang,B., Oji,A., Gibo,S., Narasimamurthy,R., Virshup,D. *et al.* (2018) Two Ck1δ transcripts regulated by m6A methylation code for two antagonistic kinases in the control of the circadian clock. *Proc. Natl. Acad. Sci.*, **115**, 5980–5985.
56. Laberge,L., Gagnon,C. and Dauvilliers,Y. (2013) Daytime sleepiness and myotonic dystrophy. *Curr. Neurol. Neurosci. Rep.*, **13**, 340.
57. Kang,J.Q. and MacDonald,R.L. (2016) Molecular pathogenic basis for GABRG2Mutations associatedwith a spectrum of epilepsy syndromes, from generalized absence epilepsy to dravet syndrome. *JAMA Neurol.*, **73**, 1009–1016.
58. Wang,T., Hoekzema,K., Vecchio,D., Wu,H., Sulovari,A., Coe,B.P., Gillentine,M.A., Wilfert,A.B., Perez-Jurado,L.A., Kvarnung,M. *et al.* (2020) Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. *Nat. Commun.*, **11**, 4932.
59. Zhang,C.-Q., Catron,M.A., Ding,L., Hanna,C.M., Gallagher,M.J., Macdonald,R.L. and Zhou,C. (2021) Impaired state-dependent potentiation of GABAergic synaptic currents triggers seizures in a genetic generalized epilepsy model. *Cereb. Cortex*, **31**, 768–784.
60. Whiting,P., McKernan,R.M. and Iversen,L.L. (1990) Another mechanism for creating diversity in γ-aminobutyrate type a receptors: RNA splicing directs expression of two forms of γ2 subunit, one of which contains a protein kinase c phosphorylation site. *Proc. Natl. Acad. Sci. USA*, **87**, 9966–9970.
61. Meier,J. and Grantyn,R. (2004) Preferential accumulation of GABAA receptor γ2L, not γ2S, cytoplasmic loops at rat spinal cord inhibitory synapses. *J. Physiol.*, **559**, 355–365.
62. Chen,Z.W., Fuchs,K., Sieghart,W., Townsend,R.R. and Evers,A.S. (2012) Deep amino acid sequencing of native brain GABAA receptors using high-resolution mass spectrometry. *Mol. Cell. Proteomics*, **11**, M111.011445.
63. Huntsman,M.M., Tran,B.Van, Potkin,S.G., Bunney,W.E. and Jones,E.G. (1998) Altered ratios of alternatively spliced long and short γ2 subunit mRNAs of the γ-amino butyrate type a receptor in prefrontal cortex of schizophrenics. *Proc. Natl. Acad. Sci. USA*, **95**, 15066–15071.
64. Steel,D., Symonds,J.D., Zuberi,S.M. and Brunklaus,A. (2017) Dravet syndrome and its mimics: beyond SCN1A. *Epilepsia*, **58**, 1807–1816.
65. Tyndale,R.F., Bhave,S.V., Hoffmann,E., Hoffman,P.L., Tabakoff,B., Tobin,A.J. and Olsen,R.W. (1997) Pentobarbital decreases the γ-Aminobutyric acidareceptor subunit gamma-2 long/short mRNA ratio by a mechanism distinct from receptor occupation. *J. Pharmacol. Exp. Ther.*, **283**, 350–357.
66. Eom,W., Lee,J.M., Park,J., Choi,K., Jung,S.-J. and Kim,H.-S. (2011) The effects of midazolam and sevoflurane on the GABA $_A$ receptors with alternatively spliced variants of the γ2 subunit. *Korean J. Anesthesiol.*, **60**, 109-118.
67. Benkwitz,C., Banks,M.I. and Pearce,R.A. (2004) Influence of GABAA receptor γ2 splice variants on receptor kinetics and isoflurane modulation. *Anesthesiology*, **101**, 924–936.
68. Yi,S., Huang,X., Zhou,S., Zhou,Y., Anderson,M.K., Zúñiga-Pflücker,J.C., Luan,Q. and Li,Y. (2020) E2A regulates neural ectoderm fate specification in human embryonic stem cells. *Development*, **147**, dev190298.
69. Robinson,K.F., Narasipura,S.D., Wallace,J., Ritz,E.M. and Al-Harthi,L. (2020) β-Catenin and TCFs/LEF signaling discordantly regulate IL-6 expression in astrocytes. *Cell Commun. Signal.*, **18**, 93.
70. Nakamori,M., Hamanaka,K., Thomas,J.D., Wang,E.T., Hayashi,Y.K., Takahashi,M.P., Swanson,M.S., Nishino,I. and Mochizuki,H. (2017) Aberrant myokine signaling in congenital myotonic dystrophy. *Cell Rep.*, **21**, 1240–1252.
71. Nishi,M., Kimura,T., Igeta,M., Furuta,M., Suenaga,K., Matsumura,T., Fujimura,H., Jinnai,K. and Yoshikawa,H. (2020) Differences in splicing defects between the grey and white matter in myotonic dystrophy type 1 patients. *PLoS One*, **15**, e0224912.
72. Gates,D.P., Coonrod,L.A. and Berglund,J.A. (2011) Autoregulated splicing of muscleblind-like 1 (MBNL1) pre-mRNA. *J. Biol. Chem.*, **286**, 34224–34233.
73. Konieczny,P., Stepniak-Konieczna,E., Taylor,K., Sznajder,L.J. and Sobczak,K. (2017) Autoregulation of MBNL1 function by exon 1 exclusion from MBNL1 transcript. *Nucleic Acids Res.*, **45**, 1760–1775.
74. Kajdasz,A., Niewiadomska,D., Sekrecki,M. and Sobczak,K. (2022) Distribution of alternative untranslated regions within the mRNA of the CELF1 splicing factor affects its expression. *Sci. Rep.*, **12**, 190.
75. Poulos,M.G., Batra,R., Li,M., Yuan,Y., Zhang,C., Darnell,R.B. and Swanson,M.S. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. *Hum. Mol. Genet.*, **22**, 3547–3558.
76. Kalsotra,A., Wang,K., Li,P.-F. and Cooper,T.A. (2010) MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development. *Genes Dev.*, **24**, 653–658.
77. Wang,P.Y., Chang,K.T., Lin,Y.M., Kuo,T.Y. and Wang,G.S. (2018) Ubiquitination of MBNL1 is required for its cytoplasmic localization and function in promoting neurite outgrowth. *Cell Rep.*, **22**, 2294–2306.
78. Lee,K.Y., Chang,H.C., Seah,C. and Lee,L.J. (2019) Deprivation of muscleblind-like proteins causes deficits in cortical neuron

distribution and morphological changes in dendritic spines and postsynaptic densities. *Front. Neuroanat.*, **13**, 75.
79. Hawrylycz,M.J., Lein,E.S., Guillozet-Bongaarts,A.L., Shen,E.H., Ng,L., Miller,J.A., van de Lagemaat,L.N., Smith,K.A., Ebbert,A., Riley,Z.L. *et al.* (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, **489**, 391–399.
80. Eze,U.C., Bhaduri,A., Haeussler,M., Nowakowski,T.J. and Kriegstein,A.R. (2021) Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat. Neurosci.*, **24**, 584–594.
81. Wagner,S.D., Struck,A.J., Gupta,R., Farnsworth,D.R., Mahady,A.E., Eichinger,K., Thornton,C.A., Wang,E.T. and Berglund,J.A. (2016) Dose-dependent regulation of alternative splicing by MBNL proteins reveals biomarkers for myotonic dystrophy. *PLos Genet.*, **12**, e1006316.
82. Thomas,J.D., Sznajder,Ł.J., Bardhi,O., Aslam,F.N., Anastasiadis,Z.P., Scotti,M.M., Nishino,I., Nakamori,M., Wang,E.T. and Swanson,M.S. (2017) Disrupted prenatal RNA processing and myogenesis in congenital myotonic dystrophy. *Genes Dev.*, **31**, 1122–1133.
83. Hernández-Hernández,O., Guiraud-Dogan,C., Sicot,G., Huguet,A., Luilier,S., Steidl,E., Saenger,S., Marciniak,E., Obriot,H., Chevarin,C. *et al.* (2013) Myotonic dystrophy CTG expansion affects synaptic vesicle proteins, neurotransmission and mouse behaviour. *Brain*, **136**, 957–970.
84. Iakova,P., Wang,G.L., Timchenko,L., Michalak,M., Pereira-Smith,O.M., Smith,J.R. and Timchenko,N.A. (2004) Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate. *EMBO J.*, **23**, 406–417.
85. Philips,A.V., Timchenko,L.T. and Cooper,T.A. (1998) Disruption of splicing regulated by a CUG-Binding protein in myotonic dystrophy. *Science*, **280**, 737–741.