NUDT21-spanning CNVs lead to neuropsychiatric disease and altered MeCP2 abundance via alternative polyadenylation

Vincenzo A Gennarino[1,2†], Callison E Alcott[2,3,12†], Chun-An Chen[1,2], Arindam Chaudhury[4,5], Madelyn A Gillentine[1,2], Jill A Rosenfeld[1], Sumit Parikh[6], James W Wheless[7], Elizabeth R Roeder[1,8], Dafne DG Horovitz[9], Erin K Roney[1], Janice L Smith[1], Sau W Cheung[1], Wei Li[10], Joel R Neilson[4,5], Christian P Schaaf[1,2]*, Huda Y Zoghbi[1,2,3,8,11]*

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, United States; [2]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, United States; [3]Program in Developmental Biology, Baylor College of Medicine, Houston, United States; [4]Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, United States; [5]Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, United States; [6]Center for Child Neurology, Cleveland Clinic Children's Hospital, Cleveland, United States; [7]Department of Pediatric Neurology, Neuroscience Institute and Tuberous Sclerosis Clinic, Le Bonheur Children's Hospital, University of Tennessee Health Science Center, Memphis, United States; [8]Department of Pediatrics, Baylor College of Medicine, Houston, United States; [9]Depto de Genetica Medica, Instituto Nacional de Saude da Mulher, da Criança e do Adolescente Fernandes Figueira, Rio de Janeiro, Brazil; [10]Division of Biostatistics, Dan L Duncan Cancer Center, Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, United States; [11]Howard Hughes Medical Institute, Baylor College of Medicine, Houston, United States; [12]Medical Scientist Training Program, Baylor College of Medicine, Houston, United States

*For correspondence: hzoghbi@bcm.edu (HYZ); schaaf@bcm.edu (CPS)

†These authors contributed equally to this work

Abstract The brain is sensitive to the dose of MeCP2 such that small fluctuations in protein quantity lead to neuropsychiatric disease. Despite the importance of MeCP2 levels to brain function, little is known about its regulation. In this study, we report eleven individuals with neuropsychiatric disease and copy-number variations spanning *NUDT21*, which encodes a subunit of pre-mRNA cleavage factor Im. Investigations of *MECP2* mRNA and protein abundance in patient-derived lymphoblastoid cells from one *NUDT21* deletion and three duplication cases show that *NUDT21* regulates MeCP2 protein quantity. Elevated *NUDT21* increases usage of the distal polyadenylation site in the *MECP2* 3′ UTR, resulting in an enrichment of inefficiently translated long mRNA isoforms. Furthermore, normalization of *NUDT21* via siRNA-mediated knockdown in duplication patient lymphoblasts restores MeCP2 to normal levels. Ultimately, we identify *NUDT21* as a novel candidate for intellectual disability and neuropsychiatric disease, and elucidate a mechanism of pathogenesis by MeCP2 dysregulation via altered alternative polyadenylation.
DOI: 10.7554/eLife.10782.001

eLife digest The X-chromosome carries a number of genes that are involved in a child's intellectual development. One of these genes encodes a protein called MeCP2, which is important for brain function after birth. Mutations in the *MECP2* gene cause a disorder known as Rett syndrome. At around 18 months of age, affected children begin to lose the cognitive and motor skills that they had previously acquired. Individuals with extra copies of this gene also show cognitive impairments. For both diseases, individuals with levels of the MeCP2 protein that are the most different from those found in healthy individuals also show the most severe symptoms.

To produce the protein that is encoded by a particular gene, enzymes inside the cell must first make a copy of that gene using a molecule called messenger ribonucleic acid (or mRNA). This mRNA is then used as a template to assemble the protein itself. In the case of *MECP2*, two different mRNA templates are produced: a long version and a short version. A gene called *NUDT21* makes a protein that regulates whether the long or short version of *MECP2* mRNA is made.

Gennarino, Alcott et al. have now discovered that people with too many, or too few, copies of the *NUDT21* gene have intellectual disabilities and altered levels of MeCP2 protein. Specifically, individuals with extra copies of *NUDT21*—and thus higher levels of the corresponding protein—produce more of the long *MECP2* mRNA. The production of proteins from this long mRNA is less efficient than from the short mRNA; therefore, these individuals have lower levels of MeCP2 protein. The opposite is true for individuals who lack a copy of the *NUDT21* gene.

To confirm these data, Gennarino, Alcott et al. grew cells in the laboratory from patients with extra copies of the *NUDT21* gene and found that reducing the production of its protein returned the levels of the MeCP2 protein back to normal. These findings show that alterations in the *NUDT21* gene cause changes in the level of MeCP2 protein in cells and leads to neuropsychiatric diseases.
DOI: 10.7554/eLife.10782.002

Main text

Methyl CpG-binding protein 2 (MeCP2) binds methylated cytosines (*Lewis et al., 1992*, *Guo et al., 2014*) and affects the expression of thousands of genes (*Chahrour et al., 2008*). *MECP2* loss-of-function is the predominant cause of Rett syndrome, a postnatal neurological disorder that typically manifests around 18 months of age with developmental regression (*Amir et al., 1999*; *Ravn et al., 2005*; *Pan et al., 2006*). *MECP2* mutations additionally cause other neurological disorders, such as non-specific autism (*Carney et al., 2003*), Angelman-like syndrome (*Imessaoudene et al., 2001*), and intellectual disability (*Orrico et al., 2000*). And in the absence of *MECP2* mutations, decreased MeCP2 expression has been detected in the brains of patients with autism, Angelman syndrome, and Prader–Willi syndrome (*Samaco et al., 2004*; *Nagarajan et al., 2006*).

Duplications within Xq28 involving *MECP2* account for 1% of X-linked intellectual disability (*del Gaudio et al., 2006*; *Friez et al., 2006*; *Lugtenberg et al., 2006*; *Meins et al., 2005*; *Van Esch et al., 2005*; *Cox et al., 2003*) and triplications encompassing *MECP2* lead to a more severe phenotype (*del Gaudio et al., 2006*). Mouse studies have established that *MECP2* alone within the duplicated and triplicated region is sufficient to cause all the neurological phenotypes of the duplication and triplication syndromes (*Collins et al., 2004*; *Samaco et al., 2012*). Notably, even small changes in MeCP2 protein level lead to neurocognitive deficits and behavioral abnormalities, and the severity of the phenotype correlates with the level of MeCP2 (*Chao and Zoghbi, 2012*).

MECP2 is distinctive for its long 3′ UTR of about 8.5 kb that contains two predominant poly-adenylation (p(A)) sites: a proximal p(A) site located just after the last exon, and a distal p(A) site approximately 8 kb from the final exon, which result in short and long messenger ribonucleic acid (mRNA) isoforms, respectively (*Coy et al., 1999*; *Shahbazian et al., 2002*). 3′ UTRs are important for fine-tuning transcript and protein levels because they contain binding sites for regulatory molecules such as RNA-binding proteins and microRNAs (miRNAs), (*Bartel, 2009*, *Gennarino et al., 2012*; *Gennarino et al., 2015*) which induce degradation of the mRNA or inhibit its translation. Thus, long isoforms allow for greater regulation due to an increased number of regulatory elements. The long 3′ UTR of *MECP2* harbors more than 50 putative miRNA-binding sites, including the miRNAs known to bind *MECP2*—miR-483-5p, miR-132 and miR-155—and reduce MeCP2 protein abundance (*Klein et al., 2007*; *Kuhn et al., 2010*; *Han et al., 2013*). The proximal p(A) site of *MECP2* is increasingly used throughout postnatal development, resulting in more

short mRNA isoforms and a concomitant increase in protein (*Balmer et al., 2003*), which may be driven by the loss of regulatory binding sites in its 3′ UTR. These observations highlight the complex and precise regulation of MeCP2 in the human brain.

Pre-messenger RNA cleavage factor Im (CFIm) regulates 3′ UTR length by mediating alternative poly-adenylation (APA) (*Millevoi and Vagner, 2010*). It binds pre-messenger mRNA as a hetero-tetramer that consists of a *NUDT21*-encoded CFIm25 dimer and any two of the paralogs CFIm59 or CFIm68 (or its splice variant, CFIm72) (*Kim et al., 2010*; *Yang et al., 2011*). In vitro, CFIm25 or CFIm68 knockdown leads to a transcriptome-wide increase in proximal APA site usage, whereas knockdown of other p(A) proteins, such as CFIm59, CPSF, CSTF, or CFIIm, has minimal effects on APA (*Gruber et al., 2012*; *Masamha et al., 2014*). Several models have been proposed to explain the role of CFIm25 in APA, but the exact mechanism remains unknown. A total of 59 genes have increased proximal p(A) site usage when comparing low- to high-*NUDT21* expressing glioblastoma tumors, and of these, 24 show the same effect as that following CFIm25 knockdown in HeLa cells. *MECP2* is the most affected, and its protein levels increase accordingly (*Masamha et al., 2014*). Moreover, Clip-seq data show enriched CFIm25 binding of the *MECP2* 3′ UTR near its p(A) sites (*Gruber et al., 2012*). Based on this, we hypothesized that CFIm25 regulates MeCP2 in humans by changing the ratio of short to long mRNA isoforms and that *NUDT21* gain- or loss-of-function will be correlated with neuropsychiatric disease.

To support our hypothesis, we sought patients with copy-number variation (CNV) of *NUDT21*. We identified eight individuals with duplications and three with deletions spanning *NUDT21*, all of whom had undergone clinical chromosome microarray testing. Of these, five consented to provide a detailed medical history (*Figure 1* and *Table 1*), and four provided blood samples for molecular studies (three with *NUDT21*-spanning duplications and one with a deletion).

Lymphoblastoid cell lines are commonly used in patient-specific molecular and functional studies of neurological disease (*Sie et al., 2009*). We, therefore, established immortalized lymphoblastoid cell lines of the four individuals and tested them for relative changes in MeCP2 protein and mRNA levels, and APA site usage. As hypothesized, in the duplication patients, we found ∼60% more CFIm25 and ∼50% less MeCP2 protein (*Figure 2A*), while in the deletion patient, we found ∼50% less CFIm25 and ∼50% more MeCP2 protein (*Figure 2B*), when compared to 13 age-matched control subjects. In order to prove that *NUDT21* alone can regulate MeCP2 levels, we performed an siRNA-mediated knockdown of *NUDT21* and assessed for changes in MeCP2 protein abundance. We found that *NUDT21* knockdown increases MeCP2 in patient-derived lymphoblastoid cells, and that reducing it to wild-type levels in the *NUDT21*-duplication patients rescues MeCP2 protein levels to that of the healthy controls (*Figure 2C*). To our knowledge, this is the first time that patient-derived lymphoblastoid cells have been transfected or transduced with siRNAs and that the functional consequence of a candidate gene within a larger CNV has been validated in vitro.

We then set out to understand how *NUDT21* CNVs lead to altered MeCP2 protein levels by investigating their effects on *MECP2* mRNA. Quantitative RT-PCR showed that total *MECP2* mRNA was elevated in both the duplication cases and the deletion patient (*Figure 3A*). However, using primer pairs that detected only the long 3′ UTR isoform of *MECP2*, we found that duplication patients have increased long *MECP2* (*Figure 3A*), indicating that elevated CFIm25 increases distal APA site usage. Conversely, the deletion patient showed a decrease in the long *MECP2* isoform, further indicating that *NUDT21* levels correlate with *MECP2* 3′ UTR length (*Figure 3A*). Northern blot analysis of RNA from the lymphoblastoid cells confirmed the switching between long and short *MECP2* isoforms for both duplication and deletion patients (*Figure 3—figure supplement 1*). Our analysis of the *MECP2* mRNA levels in *NUDT21*-spanning duplication and deletion subjects shows that CFIm25 regulates the ratio of short to long mRNA isoforms (*Figure 3A*).

To clarify how an increase in *MECP2* mRNA may still lead to decreased MeCP2 protein in the *NUDT21*-duplication patients, we performed a polyribosome fractionation assay to assess translation efficiency of total and long *MECP2* mRNA (*Figure 3—figure supplement 2*). We found that the *NUDT21*-duplication patients translate *MECP2* less efficiently than healthy controls, with a ∼50% reduction in total *MECP2* mRNA in the polyribosome fraction and a dramatic enrichment of the long isoform in the non-polyribosome fraction (*Figure 3B*). These data support the hypothesis that *NUDT21*-duplication patients have decreased MeCP2 protein despite increased mRNA because *NUDT21* promotes distal APA site usage and the elevated RNA pool consists largely of the inefficiently translated long *MECP2* isoform (*Figure 3C*).
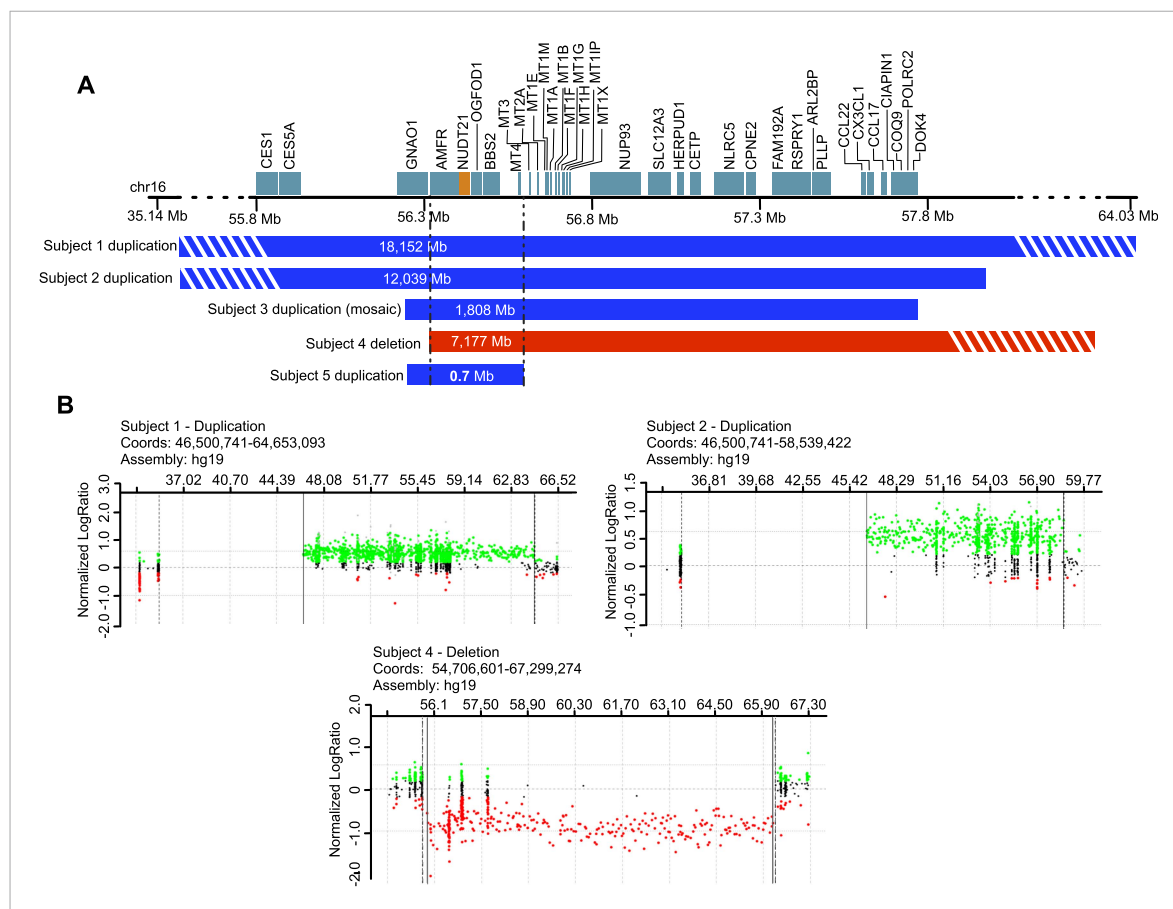
Figure 1. Subjects with *NUDT21*-spanning copy-number variations (CNVs). (**A**) Five intrachromosomal rearrangements of chromosome 16q including the *NUDT21* gene, identified by clinical array comparative genomic hybridization. Duplications shown in blue, deletion in red. Del, deletion; Dup, duplication; Mb, megabases. The striped bars indicate that the copy variant is not drawn to scale. (**B**) Array plots of oligonucleotide arrays on subjects 1, 2, and 4. The array plots of subject 3 and 5 were unavailable due to the closure of the Signature Genomics microarray laboratory in June 2014. Black dots indicate probes with normal copy-number, green dots indicate copy-number gain, and red dots indicate copy-number loss. Solid and dotted lines respectively define the minimum and maximum expected boundaries of the CNVs.

DOI: 10.7554/eLife.10782.003

The CFIm25 subunit of pre-messenger RNA CFIm represents the first identified post-transcriptional protein regulator of MeCP2 and provides important insights into MeCP2 regulation during normal development and disease.

The presented cases of individuals with *NUDT21* CNVs and subsequent MeCP2 level changes suggest *NUDT21* as a novel candidate gene for intellectual disability and neuropsychiatric disease. Their CNVs, however, are non-recurrent, and all affect several genes in addition to *NUDT21,* so that the actual effect of the *NUDT21* dosage change alone on phenotype cannot be determined at this point. CNVs affecting only *NUDT21* or point mutation cases of *NUDT21* would provide important insight. However, we found no *NUDT21*-spanning CNVs in 30,466 healthy controls (*Pinto et al., 2007*; *Simon-Sanchez et al., 2007*; *Zogopoulos et al., 2007*; *International Schizophrenia C, 2008*; *Jakobsson et al., 2008*; *Itsara et al., 2009*; *Kirov et al., 2009*; *Shaikh et al., 2009*; *Conrad et al., 2010*; *International HapMap et al., 2010*; *Vogler et al., 2010*; *Banerjee et al., 2011*; *Campbell et al., 2011*; *Cooper et al., 2011*; *Genomes Project et al., 2012*), while among our group of 87,200 patients who underwent clinical chromosome microarray analysis, mostly for neurodevelopmental disorders, we identified 11 *NUDT21*-spanning CNVs (p = 0.025, one-tailed Chi square test without Yates correction), suggesting that CNVs involving *NUDT21* are indeed pathogenic. Moreover, of the four other genes common to all the patients' CNVs, only one is associated with neurological disease, *BBS2*, but it is autosomal recessive and is characterized by distinctive features not seen in the patients we studied.

Table 1. Molecular and clinical characteristics of four individuals with *NUDT21* duplication and one individual with *NUDT21* deletion

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
Sex	F	M	M	F	F
Age	10 years	5 years	15 years	13 years	8 years
Deletion/duplication	Duplication	Duplication	Duplication	Deletion	Duplication
Coordinates	chr16: 46,500,741-64,653,093	chr16: 46,500,741-58,539,422	chr16: 55,725,264-57,533,101	chr16: 56,344,856-63,521,523	chr16: 55,919,145-56,619,283
Size	18.15 Mb	12.04 Mb	1.81 Mb	7.18 Mb	0.70 Mb
Zygosity	Heterozygous	Heterozygous	Mosaic marker chromosome (present in 45% of cells)	Heterozygous	Heterozygous
Inheritance	De novo	De novo	De novo	Unknown	Not maternal*
Additional CNVs	52 kb duplication of chr6: 152, 716, 341-152, 768, 421 (maternal)	None	None	144 kb duplication of chr17: 427, 284-571, 275 (inheritance unknown)	None
Developmental delay/intellectual disability	Intellectual disability (clinical impression: moderate)	Significant developmental delay	Intellectual disability	Intellectual disability (full scale IQ 53)	Intellectual disability
Developmental regression	Yes, at 2 years of age	No	Yes	No	No
Autism spectrum disorder	Yes	No	Yes	No	Yes
Epilepsy	Single, isolated seizure at 5 years of age	No	No	Intractable, symptomatic partial onset seizures	No
ADHD	Yes	Yes	Unknown	Yes	No

F, female; M, male; chr, chromosome; MB, megabases; CNVs, copy number variants; kb, kilobases; ADHD, attention deficit hyperactivity disorder.
*Father not available.
DOI: 10.7554/eLife.10782.004

The clinical presentation of the three individuals with increased *NUDT21* copy number and decreased MeCP2 protein levels is not that of classic Rett syndrome, which is not surprising since individuals with Rett lack the protein in 50% of their cells, whereas the duplication of *NUDT21* causes a partial decrease in all cells. However, these individuals suffer from intellectual disability and autism spectrum disorder, and two of the three individuals manifested significant developmental regression, which is a rare clinical phenomenon and considered a hallmark of Rett syndrome. The identification of additional patients will help us better define the phenotypic overlap and differences between *NUDT21*-duplication and Rett syndrome, and for *NUDT21* deletion or loss-of-function and *MECP2* duplication syndrome.

The clinical and molecular data presented provide insight into the post-transcriptional regulation of *MECP2*, and identify *NUDT21* as a novel candidate for intellectual disability and neuropsychiatric disease. Further, we developed a new way of validating the role of individual genes within pathogenic CNVs by effectively transfecting patient-derived lymphoblastoid cells. Ultimately, this study provides an example of how the complexity of a CNV goes well beyond the affected chromosomal domain and the genes affected within. The better we understand gene networks, gene-protein, and protein–protein interactions, the better we will be positioned to identify the molecular bases of various neuropsychiatric disorders and design treatment strategies for the affected individuals.

Materials and methods

Human subjects

Of approximately 52,000 patients referred to the Baylor College of Medicine (BCM) Medical Genetics Laboratory (MGL) for clinical array comparative genomic hybridization (aCGH) analysis between April 2007 and February 2013, six probands with copy number variants affecting *NUDT21* were identified: five duplication cases and one deletion. We also screened the database of Signature Genomics
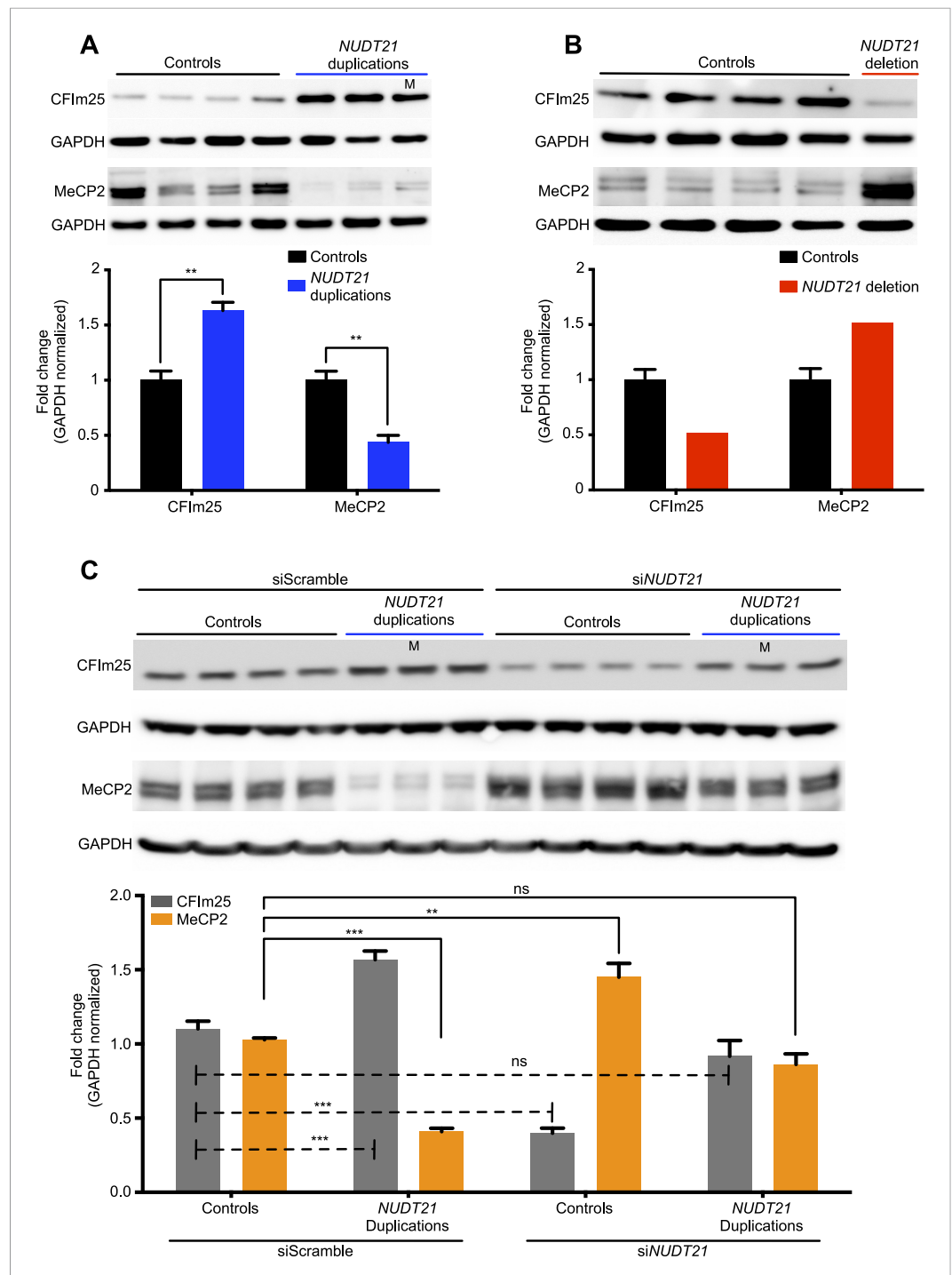
Figure 2. CFIm25 regulates MeCP2 protein levels in patient-derived lymphoblastoid cells with *NUDT21* CNVs. (**A**) Representative western blot picture for three duplication patients compared to four age-matched controls showing the increase of CFIm25 and decrease of MeCP2 protein levels. (**B**) Representative western blot picture for one deletion patient compared to four age-matched controls showing the decrease of CFIm25 and increase of MeCP2 protein levels. Quantification of protein levels for both CFIm25 and MeCP2 from three duplication patients and one deletion patient compared to a total of 13 age-matched controls are shown below the corresponding western blot. Data represent mean ± SEM from a total of six technical replicates. Data were normalized to GAPDH protein levels. (**C**) Western blot and its relative quantification showing that knockdown of *NUDT21* by siRNA-*NUDT21* nucleofection increases MeCP2 in control and duplication subjects, and normalizing CFIm25 in duplication patients rescue MeCP2 to control levels. Data represent mean ± SEM

Figure 2. continued on next page

Figure 2. Continued

from four age-matched control and three duplication cases. Data were normalized to GAPDH protein levels. **$p < 0.01$, ***$p < 0.001$. M, mosaic patient.

The following figure supplement is available for figure 2:

Figure supplement 1. siGLO nucleofection showing patient-derived lymphoblastoid cells can be transfected with small RNA.

Laboratories, based on approximately 35,200 samples submitted for clinical aCGH between November 2007 and February 2013, and identified three duplication cases and two deletions. The search was limited to copy-number variants <20 megabases. The clinical labs shared our contact information with the referring providers of all eleven cases, and we were subsequently contacted by five families who expressed interest in participating in this research study. Their copy-number variants had been detected by clinical aCGH on the following platforms: CMA-HR + SNP V9.1.1, Baylor College of Medicine (subject 1); Oligo V8.1.1, Baylor College of Medicine (subject 2); SignatureChipOS v2.0, Signature Genomics (subject 3); and Oligo V6.5, Baylor College of Medicine (subject 4). Following informed consent, approved by the Institutional Review Board for Human Subject Research at Baylor College of Medicine, we performed a comprehensive chart review of medical records and neuropsychological testing. A venous blood sample was provided by the probands in order to establish immortalized lymphoblastoid cell lines. All individuals with *NUDT21* copy-number variants were enrolled in a research study approved by the Institutional Review Board of Baylor College of Medicine (H-25466). The consent form specifically allows for sharing of medical information and physical exam findings. All individuals whose lymphoblastoid cell lines were enrolled as controls had been enrolled in a research study approved by the Institutional Review Board of Baylor College of Medicine (H-25531) as unaffected sibling controls.

Lymphoblastoid cell culture

Venous blood of the recruited probands was drawn into ACD solution A tubes. Buffy coat was prepared, lymphocytes were pelleted, and transformed with Epstein–Barr virus and cyclosporin A following standard procedures. Cell lines were grown in RPMI 1640 medium (Invitrogen, Carlsbad, CA, United States) supplemented with 10% fetal bovine serum (Atlanta Biological, Flowery Branch, GA, United States) and 1% penicillin/streptomycin. Cell cultures were maintained at 37°C in a humidified incubator supplemented with 5% CO_2. Medium was renewed every 2 to 3 days to maintain the cell density between 1×10^5 and 2×10^6 cells/ml.

Lymphoblastoid nucleofection

Lymphoblastoid cells were nucleotransfected with 300 nM of either Ambion small interfering RNA si*NUDT21* (s21771) or siScramble (4390843) using Amaxa 4D-Nucleofector and P3 Primary Cell 4D-Nucleofector X Kit (Lonza, Cologne, Germany, cat# V4XP-3024). Cells (2×10^7) were centrifuged at 200×g for 10 min at room temperature, then resuspended in 100 μl P3 Primary Cell 4D-Nucleofector Solution with the added supplement. In microfuge tubes, 300 nM siRNA was mixed with the suspended cells and transferred to Nucleocuvette Vessels prior to nucleofection. Cells were nucleofected using the program FI-115 (efficiency). Post-nucleofection, cells were resuspended with 0.5-ml pre-warmed media (RPMI, 10% FBS, and 1% antibiotic–antimycotic) and transferred to T25 flasks. After 6 hr, medium was removed and replaced with fresh media. Cells were incubated at 37°C with 5% CO_2 for 48 hr prior to protein extraction and western blot analysis. The nucleofection protocol was optimized by transfecting 2×10^7 lymphoblastoid cells with 2 μM of siGLO Green transfection indicator (Dharmacon, Lafayette, CO, United States) or Control (empty) either using the program FI-115 (efficiency) or EO-115 (functionality). Cells were later incubated at 37°C with 5% CO_2 and collected at different time points from 12 to 72 hr for cytometric analysis using a BD LSR FORTESSA (BD Biosciences, San Jose, CA, United States) according to manufacture's instructions. We determined that the efficiency protocol was preferable because the nucleofection rate was higher, and the survival rate was acceptable. We extracted protein at 48 hr post-nucleofection because the survival was satisfactory, and it was the latest time point at which nucleofected small RNA was nearly 100%, which would allow the siRNA the most time to act (*Figure 2—figure supplement 1*).
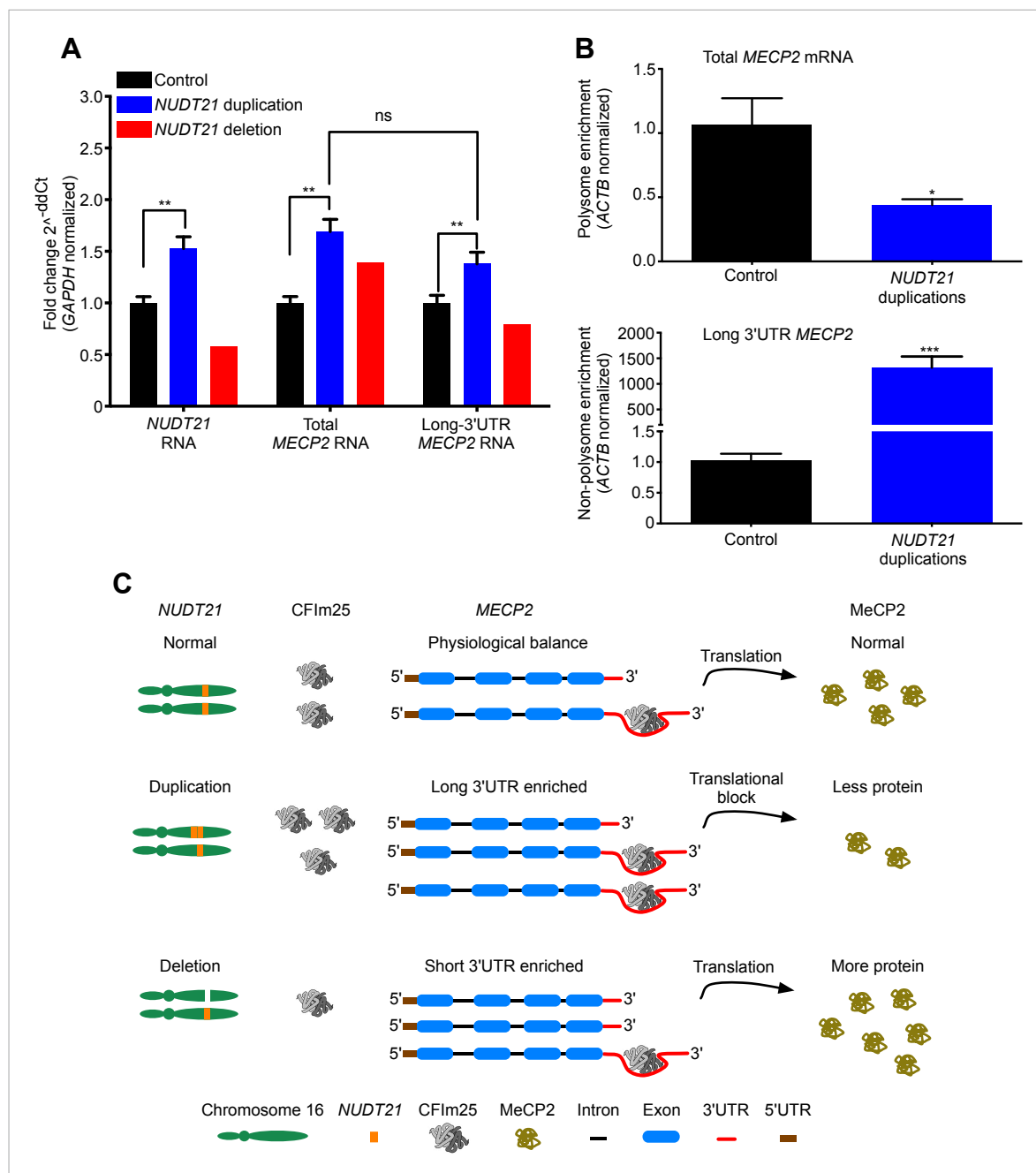
Figure 3. *NUDT21* mRNA levels correlate with inefficiently translated long *MECP2* mRNA. (**A**) RNA quantification by quantitative RT-polymerase chain reaction (qRT-PCR) from lymphoblastoid cells of *NUDT21* duplication and deletion patients. The bar graph shows the total mRNA fold change of *NUDT21*, total *MECP2*, and long *MECP2* for the three duplication patients and one deletion patient compared to 13 age-matched controls. Data represent mean ± SEM from five independent experiments. Data were normalized to *GAPDH* mRNA levels. (**B**) Relative polyribosomal and non-poyribosomal enrichment of total and long *MECP2* mRNA isoforms of *NUDT21* duplication patients compared to age-matched controls. Data represent mean ± SEM from a total of three control and duplication cases. Data were normalized to *ACTB* mRNA levels. (**C**) Proposed model showing that duplication and deletion patients have more or less CFIm25, respectively leading to a relative increase in long and short *MECP2* 3′ UTR isoforms. In both cases, there is an accumulation of mRNA: in the deletion patient, this leads to more MeCP2 protein, but in the duplication patients, it results in less MeCP2 protein due to a translational block from the CFIm25-mediated increase in long *MECP2* isoforms and putative binding of miRNAs or RNA-binding proteins to the 3′ UTR. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
DOI: 10.7554/eLife.10782.007

The following figure supplements are available for figure 3:

Figure supplement 1. Northern blot assay from patient-derived lymphoblastoid cells.
DOI: 10.7554/eLife.10782.008

Figure supplement 2. Polyribosome fractionation traces of control and duplication subjects.
DOI: 10.7554/eLife.10782.009

Western blot

Lymphoblastoid cell suspension cell cultures were collected at 6×10^6 confluence and processed for protein extraction. Cell pellets were lysed with modified RIPA buffer (180 mM NaCl, 0.5% NP-40, 0.5% Triton X-100, 2% SDS, and complete protease inhibitor cocktail (Roche, France)) by pipetting them up and down with a p1000 tip and then placed at room temperature rotisserie shaker for 10 min followed by boiling for 5 min. Then the lysates were repeatedly passed through a 27-G needle with a syringe to reduce viscosity due to DNA breakage, followed by centrifugation at 13,000 rpm at room temperature for 15 min. Proteins were quantified by Pierce BCA Protein Assay Kit (Thermo Scientific, Waltham, MA, United States) and resolved by high resolution NuPAGE 4–12% Bis-Tris Gel (Life Technologies, Waltham, MA, United States) according to the manufacturer's instruction. Antibodies: Mouse α-NUDT21 (1:500, Santa Cruz Biotechnology, Dallas, TX, United States, sc-81109); rabbit α-serum MeCP2 N-terminus (1:5000, Zoghbi Lab, #0535); mouse α-GAPDH (1:10,000, Millipore, Billerica, MA, United States, CB1001-500UG).

RNA extraction and quantitative real-time PCR

Lymphoblastoid cell suspension cell cultures were collected at 6×10^6 confluence and processed for RNA extraction. Total RNA was obtained using the miRNeasy kit (Qiagen, Netherlands) according to the manufacturer's instructions. RNA was quantified using the NanoDrop 1000 (Thermo Fisher, Waltham, MA, United States). Quality of RNA was assessed by gel electrophoresis. cDNA was synthesized using Quantitect Reverse Transcription kit (Qiagen, Netherlands) starting from 1 μg of DNase-treated RNA. Quantitative RT-polymerase chain reaction (qRT-PCR) experiments were performed using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, United States) with PerfeCta SYBR Green FastMix, ROX (Quanta Biosciences, Gaithersburg, MD, United States). Real-time PCR results were analyzed using the comparative Ct method normalized against the housekeeping gene *GAPDH* (*Vandesompele et al., 2002*). The range of expression levels was determined by calculating the standard deviation of the ΔCt (*Pfaffl, 2001*). To ensure efficacy of the genomic DNA elimination, we ran negative control samples in the qRT-PCR that did not have reverse transcriptase (-RT) in the cDNA synthesis reaction.

Northern blot assay

Total RNA was obtained from lymphoblastoid cell suspension cell cultures as mentioned above. Total RNA (15 μg) was separated on 1.2% formaldehyde agarose gel and transferred to Hybond-N+ nylon membrane (Amersham, United Kingdom), followed by UV crosslinking with Stratalinker 2400 (Agilent Technologies, Santa Clara, CA, United States). *MECP2* cDNA probe (1431 bp), which contains a majority of exon1 and exon2, was synthesized by random primed DNA labeling kit (Roche, France) according to the manufacturer's instructions. The blots were hybridized using ULTRAhyb hybridization buffer (Applied Biosystems, Waltham, MA, United States) according to the manufacturer's instructions. In brief, the blots were hybridized with radiolabeled probes (10^6 cpm/ml) in ULTRAhyb hybridization buffer for 2 hr at 68°C. The blots were then washed with 2× SSC, 0.05% SDS for 20 min at room temperature, followed by washing with 0.1× SSC, 0.1% SDS at 50°C for 10 min. The blots were exposed for 72 hr for analysis. For loading control, the blots were then stripped in boiling 0.2% SDS and hybridized with *GAPDH* probe.

Polysomal profiling

1×10^6 lymphoblastoid cells were treated with either 100 μg/ml cycloheximide (Sigma–Aldrich, St. Louis, MO, United States) for 60 min at 37°C before being washed with cold PBS containing 100 μg/ml Cycloheximide or Puromycin. The cells were then lysed in polysome lysis buffer (10 mM Tris-Cl, pH 7.4, 5 mM MgCl$_2$, 100 mM KCl, 1% (vol/vol) Triton X-100, 0.5% (wt/vol) deoxycholate, 1000 U/ml RNasin, 2 mM DTT and 100 μg/ml Cycloheximide in DEPC-treated water), incubated on ice for 10 min and centrifuged at 16,000×*g* at 4°C to obtain the clarified post-nuclear lysate. Post-nuclear clarified lysates (50 OD units) were layered on top of 10–50% sucrose gradients and centrifuged at 28,000 rpm (100,000×*g*) for 4 hr at 4°C using an SW41 rotor (Beckman, Fullerton, CA, United States). Sucrose gradients were fractionated using a BR-184 tube piercer and a syringe pump (Brandel, Gaithersburg, MD, United States) fitted with a UA-6 UV detector (Teledyne ISCO, Lincoln, NE, United States). Digital data were collected during fractionation using the DI-158U USB data acquisition device (DATAQ Instruments, Akron, OH, United States). The digital data were processed using the Peak Chart Data Acquisition Software and UV absorption at 254 nm was plotted vs time (from top to bottom or fraction number of sucrose gradients) (*Figure 3—figure supplement 2*). TRIzol LS reagent

(Life Technologies, Carlsbad, CA, United States) was used to extract RNA from the various polysome fraction and total lysate aliquots as per the manufacturer's recommendations.

Primers

In order to unambiguously distinguish spliced cDNA from genomic DNA contamination, specific exon primers were designed to amplify across introns of the genes tested. The primers for all target genes tested were designed with Primer3 v.0.4.0 (*Koressaar and Remm, 2007*). Primer sequences:

qRT-PCR

Human *GAPDH*: Forward, 5′-CGACCACTTTGTCAAGCTCA-3′; Reverse, 5′-TTACTCCTTGGAGGCCATGT-3′

Human *MECP2*: Forward, 5′-GATCAATCCCCAGGGAAAAGC-3′; Reverse, 5′-CCTCTCCCAGTTA CCGTGAAG-3′

Human long 3′ UTR: Forward, 5′-CATCCCGTGCTTTTAAGGAA-3′; Reverse, 5′-CATTCTCCCCAGC TCTTCAG-3′

Human short 3′ UTR: Forward, 5′-TCTGACAAAGCTTCCCGATT-3′; Reverse, 5′-CCAACTACTCC CACCCTGAA-3′

Northern blot

Human *MECP2*: Forward, 5′-GAAGAAAAGTCAGAAGACCAGGACCT-3′; Reverse, 5′-GCTAACTCTC TCGGTCACGGG-3′

Human *GAPDH*: Forward, 5′-ACCACAGTCCATGCCATCAC-3′; Reverse, 5′-CACCACCCTGTTGC TGTAGCC-3′

Acknowledgements

Additional information

Competing interests

HYZ: Senior editor, *eLife*. The other authors declare that no competing interests exist.

Funding

Author contributions

HYZ, Conception and design, Analysis and interpretation of data, Drafting or revising the article; VAG, Design the experiments and Performed the experiments, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; CEA, Design the experiments and Performed the experiments, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; C-AC, MAG, Contributed to molecular work and lymphoblast cultures, Acquisition of data; AC, Performed the polyribosome experiments, Acquisition of data; JAR, EKR, JLS, SWC, SP, JWW, ERR, DDGH, Examined the patients and provided clinical data; WL, Shared CFIm25 data prior to publication, Contributed unpublished essential data; JRN, Analyzed the polyribosome experiments; CPS, Serves as communicating author for all clinical data, Conception and design, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs

Jill A Rosenfeld, http://orcid.org/0000-0001-5664-7987

Ethics

Human subjects: All individuals with NUDT21 copy number variants were enrolled in a research study approved by the Institutional Review Board of Baylor College of Medicine (H-25466). The consent form specifically allows for sharing of medical information and physical exam findings. All individuals whose lymphoblastoid cell lines were enrolled as controls had been enrolled in a research study approved by the Institutional Review Board of Baylor College of Medicine (H-25531) as unaffected sibling controls.

References

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics* **23**:185–188. doi: 10.1038/13810.

Balmer D, Goldstine J, Rao YM, LaSalle JM. 2003. Elevated methyl-CpG-binding protein 2 expression is acquired during postnatal human brain development and is correlated with alternative polyadenylation. *Journal of Molecular Medicine* **81**:61–68.

Banerjee S, Oldridge D, Poptsova M, Hussain WM, Chakravarty D, Demichelis F. 2011. A computational framework discovers new copy number variants with functional importance. *PLOS ONE* **6**:e17539. doi: 10.1371/journal.pone.0017539.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136**(2):215–233. doi: 10.1016/j.cell.2009.01.002.

Carney RM, Wolpert CM, Ravan SA, Shahbazian M, Ashley-Koch A, Cuccaro ML, Vance JM, Pericak-Vance MA. 2003. Identification of MeCP2 mutations in a series of females with autistic disorder. *Pediatric Neurology* **28**:205–211. doi: 10.1016/S0887-8994(02)00624-0.

Campbell CD, Sampas N, Tsalenko A, Sudmant PH, Kidd JM, Malig M, Vu TH, Vives L, Tsang P, Bruhn L, Eichler EE. 2011. Population-genetic properties of differentiated human copy-number polymorphisms. *American Journal of Human Genetics* **88**:317–332. doi: 10.1016/j.ajhg.2011.02.004.

Chao HT, Zoghbi HY. 2012. MeCP2: only 100% will do. *Nature Neuroscience* **15**:176–177. doi: 10.1038/nn.3027.

Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY. 2008. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**(5880):1224–1229. doi: 10.1126/science.1153252.

Collins AL, Levenson JM, Vilaythong AP, Richman R, Armstrong DL, Noebels JL, David Sweatt J, Zoghbi HY. 2004. Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Human Molecular Genetics* **13**:2679–2689. doi: 10.1093/hmg/ddh282.

Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, Macarthur DG, Macdonald JR, Onyiah I, Pang AW, Robson S, Stirrups K, Valsesia A, Walter K, Wei J, Wellcome Trust Case Control Consortium, Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* **464**:704–712. doi: 10.1038/nature08516.

Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE. 2011. A copy number variation morbidity map of developmental delay. *Nature Genetics* **43**:838–846. doi: 10.1038/ng.909.

Cox JJ, Holden ST, Dee S, Burbridge JI, Raymond FL. 2003. Identification of a 650 kb duplication at the X chromosome breakpoint in a patient with 46,X,t(X;8)(q28;q12) and non-syndromic mental retardation. *Journal of Medical Genetics* **40**:169–174. doi: 10.1136/jmg.40.3.169.

Coy JF, Sedlacek Z, Bachner D, Delius H, Poustka A. 1999. A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3''-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. *Human Molecular Genetics* **8**:1253–1262. doi: 10.1093/hmg/8.7.1253.

del Gaudio D, Fang P, Scaglia F, Ward PA, Craigen WJ, Glaze DG, Neul JL, Patel A, Lee JA, Irons M, Berry SA, Pursley AA, Grebe TA, Freedenberg D, Martin RA, Hsich GE, Khera JR, Friedman NR, Zoghbi HY, Eng CM, Lupski JR, Beaudet AL, Cheung SW, Roa BB. 2006. Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genetics in Medicine* **8**:784–792. doi: 10.1097/01.gim.0000250502.28516.3c.

Friez MJ, Jones JR, Clarkson K, Lubs H, Abuelo D, Bier JA, Pai S, Simensen R, Williams C, Giampietro PF, Schwartz CE, Stevenson RE. 2006. Recurrent infections, hypotonia, and mental retardation caused by duplication of MECP2 and adjacent region in Xq28. *Pediatrics* **118**:e1687–95. doi: 10.1542/peds.2006-0395.

1000 Genomes Project Consortium; Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**:56–65. doi: 10.1038/nature11632.

Gennarino VA, D'Angelo G, Dharmalingam G, Fernandez S, Russolillo G, Sanges R, Mutarelli M, Belcastro V, Ballabio A, Verde P, Sardiello M, Banfi S. 2012. Identification of microRNA-regulated gene networks by expression analysis of target genes. *Genome Research* **22**(6):1163–1172. doi: 10.1101/gr.130435.111.

Gennarino VA, Singh RK, White JJ, De Maio A, Han K, Kim JY, Jafar-Nejad P, di Ronza A, Kang H, Sayegh LS, Cooper TA, Orr HT, Sillitoe RV, Zoghbi HY. 2015. Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels. *Cell* **160**(6):1087–1098. doi: 10.1016/j.cell.2015.02.012.

Gruber AR, Martin G, Keller W, Zavolan M. 2012. Cleavage factor Im is a key regulator of 3′ UTR length. *RNA Biology* **9**:1405–1412. doi: 10.4161/rna.22570.

Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, Zhu H, Chang Q, Gao Y, Ming GL, Song H. 2014. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature Neuroscience* **2**:215–22. doi: 10.1038/nn.3607.

Han K, Gennarino VA, Lee Y, Pang K, Hashimoto-Torii K, Choufani S, Raju CS, Oldham MC, Weksberg R, Rakic P, Liu Z, Zoghbi HY. 2013. Human-specific regulation of MeCP2 levels in fetal brains by microRNA miR-483-5p. *Genes & Development* **27**:485–490. doi: 10.1101/gad.207456.112.

Imessaoudene B, Bonnefont JP, Royer G, Cormier-Daire V, Lyonnet S, Lyon G, Munnich A, Amiel J. 2001. MECP2 mutation in non-fatal, non-progressive encephalopathy in a male. *Journal of Medical Genetics* **38**:171–174. doi: 10.1136/jmg.38.3.171.

International HapMap 3 Consortium; Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE, Altshuler DM, Gibbs RA, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Dermitzakis E, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Bonnen PE, Gibbs RA, Gonzaga-Jauregui C, Keinan A, Price AL, Yu F, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Schaffner SF, Zhang Q, Ghori MJ, McGinnis R, McLaren W, Pollack S, Price AL, Schaffner SF, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE. 2010. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**:52–58. doi: 10.1038/nature09298.

International Schizophrenia C. 2008. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**:237–241. doi: 10.1038/nature07239.

Itsara A, Cooper GM, Baker C, Girirajan S, Li J, Absher D, Krauss RM, Myers RM, Ridker PM, Chasman DI, Mefford H, Ying P, Nickerson DA, Eichler EE. 2009. Population analysis of large copy number variants and hotspots of human genetic disease. *American Journal of Human Genetics* **84**:148–161. doi: 10.1016/j.ajhg.2008.12.014.

Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung HC, Szpiech ZA, Degnan JH, Wang K, Guerreiro R, Bras JM, Schymick JC, Hernandez DG, Traynor BJ, Simon-Sanchez J, Matarin M, Britton A, van de Leemput J, Rafferty I, Bucan M, Cann HM, Hardy JA, Rosenberg NA, Singleton AB. 2008. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **451**:998–1003. doi: 10.1038/nature06742.

Kim S, Yamamoto J, Chen Y, Aida M, Wada T, Handa H, Yamaguchi Y. 2010. Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation. *Genes to Cells* **15**:1003–1013. doi: 10.1111/j.1365-2443.2010.01436.x.

Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, Holmans P, International Schizophrenia Consortium, Wellcome Trust Case Control Consortium, Craddock N, Owen MJ, O'Donovan MC. 2009. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Human Molecular Genetics* **18**:1497–1503. doi: 10.1093/hmg/ddp043.

Klein ME, Lioy DT, Ma L, Impey S, Mandel G, Goodman RH. 2007. Homeostatic regulation of MeCP2 expression by a CREB-induced microRNA. *Nature Neuroscience* **10**:1513–1514. doi: 10.1038/nn2010.

Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**:1289–1291. doi: 10.1093/bioinformatics/btm091.

Kuhn DE, Nuovo GJ, Terry AV Jr, Martin MM, Malana GE, Sansom SE, Pleister AP, Beck WD, Head E, Feldman DS, Elton TS. 2010. Chromosome 21-derived microRNAs provide an etiological basis for aberrant protein expression in human down syndrome brains. *The Journal of Biological Chemistry* **285**:1529–1543. doi: 10.1074/jbc.M109.033407.

Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, Bird A. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**(6):905–914.

Lugtenberg D, de Brouwer AP, Kleefstra T, Oudakker AR, Frints SG, Schrander-Stumpel CT, Fryns JP, Jensen LR, Chelly J, Moraine C, Turner G, Veltman JA, Hamel BC, de Vries BB, van Bokhoven H, Yntema HG. 2006. Chromosomal copy number changes in patients with non-syndromic X linked mental retardation detected by array CGH. *Journal of Medical Genetics* **43**:362–370. doi: 10.1136/jmg.2005.036178.

Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ. 2014. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **510**:412–416. doi: 10.1038/nature13261.

Meins M, Lehmann J, Gerresheim F, Herchenbach J, Hagedorn M, Hameister K, Epplen JT. 2005. Submicroscopic duplication in Xq28 causes increased expression of the MECP2 gene in a boy with severe mental retardation and features of Rett syndrome. *Journal of Medical Genetics* **42**:e12. doi: 10.1136/jmg.2004.023804.

Millevoi S, Vagner S. 2010. Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. *Nucleic Acids Research* **38**:2757–2774. doi: 10.1093/nar/gkp1176.

Nagarajan RP, Hogart AR, Gwye Y, Martin MR, LaSalle JM. 2006. Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. *Epigenetics* **1**:e1–e11.

Orrico A, Lam C, Galli L, Dotti MT, Hayek G, Tong SF, Poon PM, Zappella M, Federico A, Sorrentino V. 2000. MECP2 mutation in male patients with non-specific X-linked mental retardation. *FEBS Letters* **481**:285–288. doi: 10.1016/S0014-5793(00)01994-3.

Pan H, Li MR, Nelson P, Bao XH, Wu XR, Yu S. 2006. Large deletions of the MECP2 gene in Chinese patients with classical Rett syndrome. *Clinical Genetics* **70**:418–419. doi: 10.1111/j.1399-0004.2006.00694.x.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**:e45. doi: 10.1093/nar/29.9.e45.

Pinto D, Marshall C, Feuk L, Scherer SW. 2007. Copy-number variation in control population cohorts. *Human Molecular Genetics* **16**. doi: 10.1093/hmg/ddm241.

Ravn K, Nielsen JB, Skjeldal OH, Kerr A, Hulten M, Schwartz M. 2005. Large genomic rearrangements in MECP2. *Human Mutation* **25**:324. doi: 10.1002/humu.9320.

Samaco RC, Mandel-Brehm C, McGraw CM, Shaw CA, McGill BE, Zoghbi HY. 2012. Crh and Oprm1 mediate anxiety-related behavior and social approach in a mouse model of MECP2 duplication syndrome. *Nature Genetics* **44**:206–211. doi: 10.1038/ng.1066.

Samaco RC, Nagarajan RP, Braunschweig D, LaSalle JM. 2004. Multiple pathways regulate MeCP2 expression in normal brain development and exhibit defects in autism-spectrum disorders. *Human Molecular Genetics* **13**:629–639. doi: 10.1093/hmg/ddh063.

Shahbazian MD, Antalffy B, Armstrong DL, Zoghbi HY. 2002. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Human Molecular Genetics* **11**:115–124. doi: 10.1093/hmg/11.2.115.

Shaikh TH, Gai X, Perin JC, Glessner JT, Xie H, Murphy K, O'Hara R, Casalunovo T, Conlin LK, D'Arcy M, Frackelton EC, Geiger EA, Haldeman-Englert C, Imielinski M, Kim CE, Medne L, Annaiah K, Bradfield JP, Dabaghyan E, Eckert A, Onyiah CC, Ostapenko S, Otieno FG, Santa E, Shaner JL, Skraban R, Smith RM, Elia J, Goldmuntz E, Spinner NB, Zackai EH, Chiavacci RM, Grundmeier R, Rappaport EF, Grant SF, White PS, Hakonarson H. 2009. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Research* **19**:1682–1690. doi: 10.1101/gr.083501.108.

Sie L, Loong S, Tan EK. 2009. Utility of lymphoblastoid cell lines. *Journal of Neuroscience Research* **87**:1953–1959. doi: 10.1002/jnr.22000.

Simon-Sanchez J, Scholz S, Fung HC, Matarin M, Hernandez D, Gibbs JR, Britton A, de Vrieze FW, Peckham E, Gwinn-Hardy K, Crawley A, Keen JC, Nash J, Borgaonkar D, Hardy J, Singleton A. 2007. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity, and cell-line induced alterations in normal individuals. *Human Molecular Genetics* **16**:1–14. doi: 10.1093/hmg/ddl436.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3**.

Van Esch H, Bauters M, Ignatius J, Jansen M, Raynaud M, Hollanders K, Lugtenberg D, Bienvenu T, Jensen LR, Gecz J, Moraine C, Marynen P, Fryns JP, Froyen G. 2005. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *American Journal of Human Genetics* **77**:442–453. doi: 10.1086/444549.

Vogler C, Gschwind L, Röthlisberger B, Huber A, Filges I, Miny P, Auschra B, Stetak A, Demougin P, Vukojevic V, Kolassa IT, Elbert T, de Quervain DJ, Papassotiropoulos A. 2010. Microarray-based maps of copy-number variant regions in European and sub-Saharan populations. *PLOS ONE* **5**:e15246. doi: 10.1371/journal.pone.0015246.

Yang Q, Coseno M, Gilmartin GM, Doublie S. 2011. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure* **19**:368–377. doi: 10.1016/j.str.2010.12.021.

Zogopoulos G, Ha KC, Naqib F, Moore S, Kim H, Montpetit A, Robidoux F, Laflamme P, Cotterchio M, Greenwood C, Scherer SW, Zanke B, Hudson TJ, Bader GD, Gallinger S. 2007. Germ-line DNA copy number variation frequencies in a large North American population. *Human Genetics* **122**:345–353. doi: 10.1007/s00439-007-0404-5.