

Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in *MECP2*

Samin A. Sajan, PhD^{1,2,9}, Shalini N. Jhangiani, MS³, Donna M. Muzny, MS³, Richard A. Gibbs, PhD^{3,4}, James R. Lupski, MD, PhD³⁻⁵, Daniel G. Glaze, MD¹, Walter E. Kaufmann, MD⁶, Steven A. Skinner, MD⁷, Fran Annese, MSW⁷, Michael J. Friez, PhD⁷, Jane Lane, RN⁸, Alan K. Percy, MD⁸ and Jeffrey L. Neul, MD, PhD^{1,2,4,9}

Purpose: Rett syndrome (RTT) is a neurodevelopmental disorder caused primarily by de novo mutations in *MECP2* and sometimes in *CDKL5* and *FOXG1*. However, some RTT patients lack mutations in these genes.

Methods: Twenty-two RTT patients without apparent *MECP2*, *CDKL5*, and *FOXG1* mutations were subjected to both whole-exome sequencing and single-nucleotide polymorphism array–based copy-number variant (CNV) analyses.

Results: Three patients had *MECP2* mutations initially missed by clinical testing. Of the remaining 19, 17 (89.5%) had 29 other likely pathogenic intragenic mutations and/or CNVs (10 patients had 2 or more). Interestingly, 13 patients had mutations in a gene/region previously reported in other neurodevelopmental disorders (NDDs),

thereby providing a potential diagnostic yield of 68.4%. These mutations were significantly enriched in chromatin regulators (corrected $P = 0.0068$) and moderately enriched in postsynaptic cell membrane molecules (corrected $P = 0.076$), implicating glutamate receptor signaling.

Conclusion: The genetic etiology of RTT without *MECP2*, *CDKL5*, and *FOXG1* mutations is heterogeneous, overlaps with other NDDs, and complicated by a high mutation burden. Dysregulation of chromatin structure and abnormal excitatory synaptic signaling may form two common pathological bases of RTT.

Genet Med advance online publication 12 May 2016

Key Words: chromatin regulation; CNV; exome sequencing; glutamate signaling; Rett syndrome

INTRODUCTION

Rett syndrome (RTT; MIM 312750) is a neurodevelopmental disorder that mostly affects girls and is caused primarily by de novo mutations in the *Methyl-CpG-binding Protein 2 (MECP2)* gene on the X chromosome.¹ The disorder, which affects approximately 1 out of 10,000 live female births,² is characterized by apparently normal early development in the first 6–18 months of life, followed by psychomotor regression involving loss of speech and hand use and development of gait problems and characteristic repetitive hand stereotypies.³ RTT cases that satisfy all the revised diagnostic criteria of the disease are classified as typical RTT, and almost 97% of these patients carry de novo mutations in *MECP2*.^{3,4} Cases that satisfy some but not all of the diagnostic criteria are classified as atypical RTT, which are further divided based on overall severity or profile of symptoms. Up to 86% of atypical cases with mild symptoms, including the “preserved speech” variant of RTT, can be accounted for by mutations in *MECP2*.^{4,5} Some atypical cases with early onset of seizures before regression (“early seizure” variants) are due to de novo mutations in *CDKL5*, whereas those that regress earlier and have gross early

abnormal development (“congenital” variants of RTT) are caused by mutations in *FOXG1*.^{6,7} However, mutations in the latter two genes account for a substantially smaller proportion of atypical RTT cases when compared to mutations in *MECP2*.

The primary RTT gene *MECP2* codes for a methyl-CpG binding protein that binds to chromatin and both activates and represses gene transcription, as demonstrated by studies of gene expression changes in brains of knockout mice and of those overexpressing *MECP2*, in which reciprocal changes in expression were observed for many genes.⁸ Attempts have been made to show that *MECP2*, *CDKL5*, and *FOXG1* share some common pathways.⁹ For instance, MeCP2 can regulate the expression of *CDKL5*, whose protein product can, in turn, phosphorylate MeCP2. Some similarity has also been suggested between *MECP2* and *FOXG1* based on their overlapping domains of expression in the brain.⁶ Despite these observations, it remains unclear which specific biological functions or pathways may be affected in RTT. More recently, mutations in a few additional genes have been found in a few cases of RTT-like disorders. These genes include *MEF2C*,¹⁰ *WDR45* (ref. 11), and *STXBP1* (refs. 12,13). Other genes that have

¹Section of Child Neurology and Developmental Neuroscience, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA; ²Jan and Dan Duncan Neurological Research Institute, Texas Children’s Hospital, Houston, Texas, USA; ³Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA; ⁴Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA; ⁵Department of Pediatrics, Baylor College of Medicine and Texas Children’s Hospital, Houston, Texas, USA; ⁶Department of Neurology, Boston Children’s Hospital, Boston, Massachusetts, USA; ⁷Greenwood Genetic Center, Greenwood, South Carolina, USA; ⁸Department of Pediatrics, University of Alabama at Birmingham, Birmingham, Alabama, USA. ⁹Current address: Department of Neurosciences, University of California San Diego, San Diego, California, USA. Correspondence: Jeffrey L. Neul (jneul@ucsd.edu)

Submitted 14 October 2015; accepted 24 February 2016; advance online publication 12 May 2016. doi:10.1038/gim.2016.42

been found to be mutated in a few RTT patients but have been associated primarily with non-RTT neurodevelopmental disabilities are *IQSEC2* (ref. 13), *SCN8A*,¹³ and *SMC1A*,¹⁴ suggesting that they might impact some shared biological pathways important to brain development and/or maintenance of proper brain function.

In this study we hypothesized that genes other than *MECP2*, *CDKL5*, and *FOXG1* could contribute to RTT. We used genomic approaches to identify some of the genetic causes of both typical and atypical RTT patients who lack mutations in *MECP2*, *CDKL5*, and *FOXG1*, anticipating that at least some of the cases will be due to mutations in genes already implicated in other neurodevelopmental disorders involving epilepsy, intellectual disability, and autism spectrum disorder (ASD) owing to their phenotypic overlap with RTT. We carried out a combination of exome sequencing and high-density single-nucleotide polymorphism (SNP) array-based copy-number variant (CNV) analyses of a total of 22 RTT patients lacking mutations in the above three genes.

MATERIALS AND METHODS

Patient cohort and clinical diagnosis

Written informed consent was obtained from all parents for participation in this study, which was approved by the Institutional Review Board of the Baylor College of Medicine. All participants were enrolled in the Rett Syndrome Natural History Study (U54HD061222, ClinicalTrials.gov NCT00299312), and enrollment in this study required either a clinical diagnosis of RTT or a pathogenic mutation in *MECP2*. Diagnosis of either typical or atypical RTT was made by expert clinicians (J.L.N., D.G.G., W.E.K., S.A.S., A.K.P.) following the recently revised diagnostic criteria.³ The requirements for a diagnosis of typical RTT are evidence of a period of regression followed by stabilization, loss of acquired hand skills, loss of acquired spoken language, gait abnormalities, and stereotyped hand movements. These are considered the main criteria for diagnosis, and the presence of these features in the study participants is described in detail in **Supplementary Table S1** online. The diagnosis of atypical RTT requires the period of regression followed by stabilization and 2 of the 4 remaining main criteria in addition to 5 of 11 supportive criteria. These are also outlined in **Supplementary Table S1** online.

Genomic DNA was extracted from peripheral blood at Baylor Molecular Genetics Diagnostic Laboratory according to standard Clinical Laboratory Improvement Amendments-approved methods. Clinical efforts to arrive at a molecular diagnosis included Sanger sequencing of coding regions of known genes (*MECP2*, *CDKL5*, *FOXG1*) and assessing structural variations through a combination of methods, including multiplex ligation-dependent probe amplification, Southern blotting, and bacterial artificial chromosome or oligonucleotide array comparative genomic hybridization.

SNP genotyping and CNV analysis

Genome-wide CNV analysis was performed by genotyping probands on the Illumina Omni 2.5m SNP array using standard procedures in the Laboratory for Translational Genomics at Baylor College of Medicine. Penn CNV

(<http://penncnv.openbioinformatics.org/en/latest/user-guide/download/>) was used to identify CNVs from arrays that had a call rate >99%, SD of log *R* ratio <0.3, and a GC wave factor between -0.04 and +0.04. All samples satisfied these criteria. Two samples (cases 102000 and 101329) resulted in more than 800 CNV calls each and were removed. CNVs from the remaining samples were filtered to retain those that were at least 30kb long with 10 or more SNPs and a confidence score of at least 10 and that impacted at least one exon of at least one protein-coding gene.

Exome sequencing and variant identification

Genomic DNA of probands and parents was processed for paired-end whole-exome sequencing on an Illumina HiSeq 2000 (Illumina, San Diego, CA) at the Baylor-Hopkins Center for Mendelian Genomics. Exome capture was achieved using either the Baylor College of Medicine-developed Human Genome Sequencing Center (HGSC) Core reagent or NimbleGen's VCRome 2.1 reagent (Roche NimbleGen, Madison, WI).¹⁵ More than 6 Gb of uniquely aligned sequence was produced per individual, with at least 85% of bases covered by $\geq 20\times$ and overall average coverage of $87\times$. Alignments were made using a Burrows-Wheeler Aligner (BWA v0.5.9, <https://github.com/lh3/bwa>) for the hg19 reference human genomes and duplicates flagged by Picard v1.98 (<http://broadinstitute.github.io/picard/>). Variants were identified by following the best practice work flow of the Genome Analysis Toolkit (GATK v2.5-2, <https://github.com/broadinstitute/gatk>) and annotated using ANNOVAR (v2014Sept09, <http://annovar.openbioinformatics.org/en/latest/user-guide/download/>).

Variants were filtered to select only those whose inheritance appeared to be consistent with dominant or recessive models of disease (de novo, homozygous, compound heterozygous). Because RTT results in a clinically obvious and severe phenotype, it is extremely unlikely to be caused by variants present in control populations or in populations with other nonneurodevelopmental diseases, even at low frequencies. Thus, for de novo variants, we prioritized only those not found in the dbSNP138, 1000 Genomes, ESP6500, and ExAC databases. For compound heterozygous variants, the frequency of each individual variant had to be less than 0.005 (with no homozygotes reported for either variant) so as to be consistent with a reasonable combined incidence of typical and atypical RTT cases not caused by mutations in *MECP2*, *CDKL5*, and *FOXG1* of approximately 0.000025, which is 25% of the total incidence of RTT of 1 out of 10,000. The total read-depth cutoff was set at 10; for heterozygous variants, at least two reads had to carry the variant. Additionally, the proportion of reads with the heterozygous variant had to be 15–85%. Missense variants were prioritized based on their predicted deleteriousness as determined by 12 tools (SIFT, Polyphen2_HDIV, Polyphen2_HVAR, LRT, MutationTaster, MutationAssessor, FATHMM, RadialSVM, LR, and VEST3, and conservation scores from GERP++_RS and CADD, <http://annovar.openbioinformatics.org/en/latest/user-guide/download/#-for-filter-based-annotation>). The following additional criteria were used to select likely pathogenic variants from RTT patients for whom DNA

samples of one or both parents were unavailable: occurrence in genes previously reported to have de novo mutations in epileptic encephalopathies,^{16,17} ASD,^{18–22} intellectual disability,²³ unexplained developmental delays,^{24,25} and observation of a nervous system phenotype in mice (phenotype code MP:0003631 from Mouse Genome Informatics (<http://www.informatics.jax.org/phenotypes.shtml>)).

Sanger validation of candidate variants from exome data

Standard polymerase chain reaction (PCR) was used to amplify products with 300–800 bp for Sanger sequencing. Briefly, 20–30 ng of genomic DNA template and KAPA HiFi Hotstart DNA polymerase (KAPA Biosystems, Woburn, MA) were used for amplification in a 30- μ l reaction per the manufacturer's instructions. All forward and reverse primers were designed to have M13F-41 (GGTTTCCAGTCACGAC) and M13R-27 (GGAAACAGCTATGACCATG) universal sequences at their 5' ends, respectively. PCR products were cleaned with a clean-up kit (Qiagen, Valencia, CA, or Bioneer, Alameda, CA) and sequenced at SeqWright, LoneStar Sequencing (both Houston, TX), or Eton Bioscience (San Diego, CA).

RESULTS

Overview of genetic findings

Of the 22 patients examined, 11 had a clinical diagnosis of typical RTT and 11 had atypical RTT (**Supplementary Table S1** online), as defined by the consensus criteria, which are outlined in **Supplementary Table S1**.³ Notably, all patients showed regression followed by stabilization, specifically either a loss of hand skills or spoken language, gait abnormalities, and characteristic repetitive hand stereotypies. Exomes of both unaffected parents of six typical and seven atypical RTT patients were also sequenced. All variants considered to be likely pathogenic are presented in **Supplementary Table S4**. This table also lists all de novo mutations identified from exome analysis, regardless of whether they were considered likely pathogenic. All CNVs and exome variants selected for Sanger validation per case are listed in **Supplementary Table S2**. The Sanger sequence of one mosaic de novo mutation is presented in **Supplementary Figure S1**. The intensity and B-allele frequency plots of CNVs are provided as **Supplementary Figures S2–S10**.

Three patients were found to have causative *MECP2* mutations that were initially missed during clinical testing. One, a 5-bp frameshift deletion (p.E50fs) in the third exon of *MECP2* not present in the unaffected mother, was eventually detected in the clinic upon resequencing. The second was a de novo 17-bp frameshift duplication c.41_57dup17 (p.R20fs) initially undetected by clinical sequencing because this exon was not routinely sequenced. However, a revised sequencing report was able to detect this mutation. Our exome sequence data could not detect this mutation because of the high GC content of the first exon of *MECP2*, a molecular feature that can decrease capture efficiency in the hybridization-based capture step of exome sequencing. Considering this, we Sanger-sequenced the first exon of

MECP2 in all the remaining patients and found one de novo mutation (M1V) in the initiation codon. This exact mutation has been reported in a typical RTT patient and is expected to abolish the normal translation of the MeCP2_e1 transcript, which is the more abundant isoform in the nervous system.²⁶

From the exome data of the remaining 19 patients, we selected 78 variants for Sanger-based confirmation, of which 13 (16.7%) were loss-of-function (nonsense, splice, and frameshift insertions or deletions), 4 were in-frame insertions or deletions, 1 was a stop-loss mutation, and 60 were missense mutations. From these, a total of 15 de novo mutations were confirmed in 11 trios, giving a rate of 1.36 such mutations per trio. One de novo mutation was apparently mosaic. Three (25%) de novo mutations were loss-of-function. One de novo deletion CNV was also identified. At least one likely pathogenic mutation was found in 17 of the 19 patients (89.5%), with 13 having mutations previously associated with other neurodevelopmental disorders, thereby providing a potential molecular diagnostic yield of 68.4%. This suggests that severe neurodevelopmental disorders are more likely than not caused by genetic defects due to new mutations.

An increased mutation burden potentially contributes to RTT phenotype

Ten of 19 patients (52.6%) had more than one likely pathogenic mutation identified from exome sequence data, CNV analysis, or both. This is a high proportion of cases with multiple likely causal variants, suggesting that a high burden of mutation may contribute to the final disease phenotype in these cases. Even though not all individual de novo mutations were considered to contribute to disease, we noted that 4 of the 11 patients whose mutations were part of complete trios carried two or more such mutations from exome sequence data. We therefore determined the overall rate of such protein-altering de novo mutations in RTT patients and compared it with the same rate reported in controls.²⁷ Because 15 confirmed de novo mutations were identified out of a total of 695,695,712 high-quality bases sequenced at a depth of at least 10 \times , the rate was 1.36 de novo mutations per trio, or 2.16×10^{-8} per base per generation. Although this rate is higher, it is not statistically significantly different from the reported²⁷ control rate of 1.47×10^{-8} (binomial $P = 0.15$), which probably reflects the small sample size of 11 trios. However, when the two patients with confirmed de novo mutations in *MECP2* (one of whom also had five additional de novo mutations, all of which are listed in **Supplementary Table S4** online) were included, the observed rate of such mutations was 1.70 per trio (22 de novo mutations in 13 trios with 829,661,092 high-quality bases sequenced at a depth of at least 10 \times) or 2.57×10^{-8} per base per generation, which is significantly higher than the reported rate in controls (binomial $P = 0.009$). Hence, a high burden of de novo mutations may be a feature of RTT in general, which, when combined with CNVs, results in an increased overall mutation burden that contributes to RTT.

Enrichment of chromatin regulators and glutamate receptor signaling molecules in genes with likely pathogenic mutations

We asked whether the genes with likely pathogenic mutations in our patients were significantly enriched for those that code for proteins with common biological functions. We compiled a list of 46 genes (Table 1) comprising those with likely pathogenic intragenic mutations identified from exome sequencing as well as select genes impacted by CNVs in our patients with intragenic de novo mutations that had been reported in at least one patient in large-scale exome sequencing studies of ASD, intellectual disability, epilepsy, and other developmental disorders. Using the DAVID functional annotation tool (<https://david.ncifcrf.gov/>), we found that there was highly significant enrichment of the term “chromatin regulator” (uncorrected $P = 0.00011$; Benjamini-Hochberg corrected $P = 0.0068$) of the Protein Information Resource database. The six genes within this term were *ACTL6B*, *BRD1*, *CHD4*, *HDAC1*, *SMARCB1*, and *TRRAP*. There was also a moderate enrichment of the term “postsynaptic cell membrane” (uncorrected $P = 0.002$; Benjamini-Hochberg corrected $P = 0.076$). The four genes within this term were *GABRB2*, *GRIN2A*, *GRIN2B*, and *SHANK3*, with the last three being members of the glutamate receptor signaling pathway.

We next asked whether the protein products of genes listed in Table 1, when analyzed together with the known RTT genes *MECP2*, *CDKL5*, and *FOXG1*, physically interact with one another, colocalize, or participate in the same step of a given pathway. GeneMania (<http://www.GeneMANIA.org/>) identified 23 of the 46 genes (50%) that interacted with one another (or with other genes reported to have mutations in neurodevelopmental disorders and/or with genes showing an expression change in *MECP2* mutant model system) either directly or indirectly in at least one of these three ways (Figure 1). To ascertain that the enrichment discovered from these cases is not spurious, a similar analysis using 65 genes with de novo loss-of-function and missense mutations predicted to be deleterious observed in control individuals in several studies did not yield significant results (Supplementary Table S3 and Figure S11 online).^{18,20,27,28} Functional annotation using DAVID showed an enrichment of “phosphoprotein,” comprising 41 genes and a corrected $P = 0.0093$, which was not considered a highly specific enrichment. These analyses support the contention that many genes mutated in our RTT patients share some features with the other known RTT genes, further implicating them as having a role in this disease.

DISCUSSION

It is well known that *MeCP2*, the product of the primary RTT gene, has the capacity to alter chromatin structure. Notable lines of evidence include abnormal organization of heterochromatin during neural differentiation of a *Mecp2*-deficient mouse embryonic stem cell line, the inability of *MeCP2* containing many RTT-causing missense mutations to cause heterochromatin to cluster, and the requirement of

Table 1 Forty-six genes with de novo and likely pathogenic mutations contributing to RTT identified from either exome sequencing or CNV analysis used for enrichment testing of biological functions

Gene	Source
<i>ACTL6B</i>	Exome
<i>AUTS2</i>	CNV
<i>BRAF</i>	Exome
<i>BRD1</i>	CNV
<i>CELSR1</i>	CNV
<i>CHD4</i>	Exome
<i>CHL1</i>	CNV
<i>CNTN6</i>	CNV
<i>FAM151A</i>	Exome
<i>FAT3</i>	Exome
<i>GABRB2</i>	Exome
<i>GRIN2A</i>	Exome
<i>GRIN2B</i>	Exome
<i>HDAC1</i>	Exome
<i>HERC2</i>	CNV
<i>IMPDH2</i>	Exome
<i>IQGAP3</i>	Exome
<i>IQSEC2</i>	Exome
<i>KCNJ10</i>	Exome
<i>LAMB2</i>	Exome
<i>LRRC40</i>	Exome
<i>NAGA</i>	CNV
<i>NCOR2</i>	Exome
<i>NDNL2</i>	CNV
<i>OTUD7A</i>	CNV
<i>PLXNB2</i>	CNV
<i>PPP6R2</i>	CNV
<i>SBF1</i>	CNV
<i>SCO2</i>	CNV
<i>SCUBE1</i>	CNV
<i>SHANK3</i>	CNV
<i>SLC2A1</i>	CNV
<i>SLC39A13</i>	Exome
<i>SLC6A1</i>	CNV
<i>SMARCB1</i>	CNV
<i>SMC1A</i>	Exome
<i>SPECC1L</i>	CNV
<i>STXBP1</i>	Exome
<i>TCF20</i>	CNV
<i>TCF4</i>	Exome
<i>TRPM1</i>	CNV
<i>TRRAP</i>	Exome
<i>TUBGCP6</i>	CNV
<i>WDR45</i>	Exome
<i>WNT7B</i>	CNV
<i>ZNF536</i>	Exome

Genes from CNVs were selected if they had intragenic *de novo* mutations reported previously by large-scale exome sequencing studies of ASD, intellectual disability, epilepsy, or developmental delays (see text for appropriate references). All 46 genes served as input to determine enriched terms using the DAVID functional annotation tool as well as to generate the interaction network in Figure 1, except that in the latter case three RTT genes were also included (*MECP2*, *CDKL5*, and *FOXG1*).

CNV, copy-number variant; RTT, Rett syndrome.

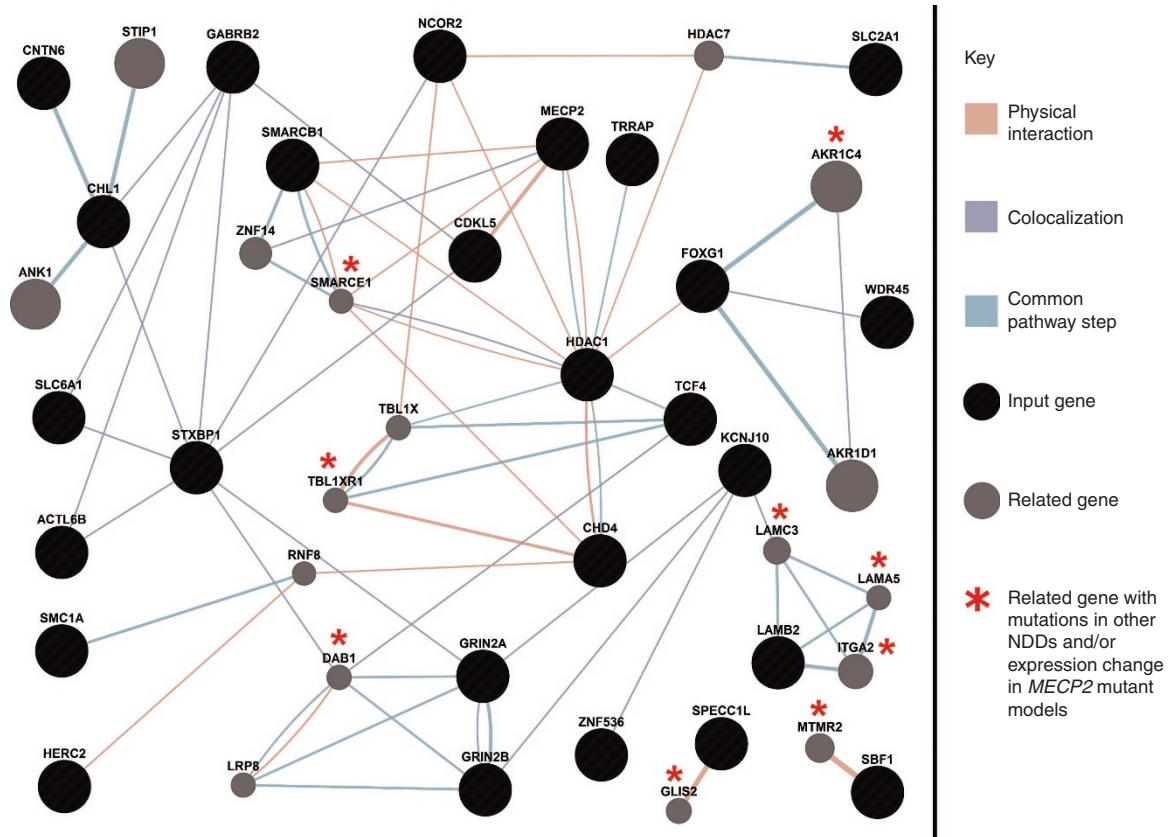


Figure 1 An interaction network of genes with likely pathogenic mutations contributing to RTT in our patients. Black circles represent input genes; gray circles represent genes highly related to the input genes chosen by the network-building algorithm to maximize connectivity. The network was generated using an input list of 46 genes with likely pathogenic mutations listed in [Table 1](#) as well as the 3 known RTT genes: *MECP2*, *CDKL5*, and *FOXG1*. Of the 46 genes, 23 were found to interact among one another, either directly or indirectly through at least one of three ways: physical interactions (orange lines), colocalization of protein products (light blue lines), and participating in the same step of a given pathway (light green lines). Asterisks indicate genes related to input genes that have been reported to either carry de novo mutations in at least one patient with other NDDs (*TBL1XR1*, *MTMR2*, *AKR1C4*) or whose expression has been reported to be significantly altered in a *MECP2* mutant model system (*DAB1*, *ITGA2*, *LAMA5*), or both (*GLIS2*, *LAMC3*, *SMARCE1*). Network weighting was assigned based on query genes to maximize connectivity among input genes, and, at most, 20 related genes and 10 related attributes were allowed to be incorporated in the network. NDD, neurodevelopmental disorder.

MeCP2 binding to chromatin to form loops that bring distal regions into close proximity for the proper transcription of specific genes.⁸ Because RTT is commonly considered part of ASD, it is not surprising that other studies have also uncovered mutations in chromatin regulators in ASD.²⁹ Given that we observed a significant enrichment of mutations in genes coding for chromatin regulators despite our small sample size compared to those of ASD studies, it is possible that dysregulation of normal chromatin architecture plays a more important role in the etiology of RTT, which is more severe than ASD. Nevertheless, our results demonstrate the presence of a shared biological function that is disrupted more often in both RTT and ASD and raises the possibility of further research into discovering overlapping treatment options for these two related, yet distinct, disorders.

Dysregulation of neuronal excitation is one factor that leads to RTT because many patients have seizures. Interestingly, previous studies utilizing *Mecp2* mutant mice have implicated both the glutamate and GABA signaling pathways. For instance,

MeCP2-deficient hippocampal glutamatergic neurons exhibit a significant reduction in synaptic response, whereas those that overexpress *Mecp2* display a higher response.³⁰ Additionally, ablating *Mecp2* function in cortical excitatory neurons but not inhibitory forebrain neurons leads to spontaneous seizures in mice.³¹ The glutamate signaling pathway is also dysregulated in other neurodevelopmental disorders related to RTT, such as ASD³² and intellectual disability.²³ In light of this, our results reinforce the role of abnormal glutamatergic signaling in RTT and, given its importance in other disorders, warrant further research to explore the possibility of treatment options that modulate this neuronal pathway as is being done for ASD.³² Our exome and CNV data did not reveal likely pathogenic variants in many genes from the GABA signaling pathway, possibly because of the small size of our patient cohort.

Given the clinical features shared by RTT with other neurodevelopmental disorders, such as ASD, Pitt-Hopkins syndrome, and Cornelia de Lange Syndrome, it is not surprising that most of our patients had likely pathogenic variants

in genes and CNVs that had previously been reported in patients with other disorders. However, what was surprising was that 10 of the 19 patients (52.6%) carried two or more likely pathogenic mutations, including a combination of intragenic variants and CNVs, suggesting the importance of increased mutation burden in causing disease. Even though 4 of the 11 RTT probands who were part of complete trios who lacked mutations in the three known RTT genes carried at least two de novo mutations, the overall burden of such mutations was not significantly different from the reported rate in control trios. Interestingly, including just two additional trios who had de novo *MECP2* mutations revealed that this rate was significantly higher. Thus, it is possible that there is an increased burden of de novo intragenic mutations in RTT in general, and it will be interesting to assess these, as well as CNVs, in a larger cohort of patients, particularly those who also harbor causal *MECP2* mutations, and perform detailed genotype–phenotype correlations.

Similar increases in mutation burden have been observed in other neurologic disorders such as Charcot–Marie–Tooth disease, a peripheral neuropathy in which a high burden of rare variants has been shown to contribute to variable expressivity, possibly by destabilizing different pathways and protein networks that could in turn modulate the phenotype.³³ This underscores the importance of using both exome sequencing and CNV analyses to identify a specific combination of likely causal variants that could help explain the variability of phenotypes in individual patients who otherwise meet the overall diagnostic criteria of a particular disorder.

Our study shows that the genetic etiology of RTT patients without mutations in *MECP2*, *CDKL5*, and *FOXG1* is heterogeneous because we did not find any recurrent pathogenic variants. Although our cohort of RTT patients is, to date, the largest reported one of its kind subjected to both exome sequencing and CNV analysis, recurrence will undoubtedly be observed as larger cohorts are analyzed. A particular focus on genes involved in chromatin remodeling and glutamate signaling in additional patients could help identify recurrence and/or novel RTT genes because these pathways were overrepresented by the genes found mutated in our cohort. We also note the usefulness of many large-scale exome sequencing studies of ASD, intellectual disability, epileptic encephalopathy, and unexplained development delay, as these have revealed de novo mutations in many genes in single patients. Smaller-scale studies with more phenotypic information can potentially bolster the evidence supporting the involvement of many of these genes in neurological disease because studies of these smaller cohorts may also identify single patients with deleterious de novo mutations in those same genes. The challenge will be to compare in detail the clinical phenotypes of the respective patients from disparate studies and cohorts, keeping in mind that any phenotypic differences may not necessarily exclude the gene in question as being causal because there could be additional variants elsewhere that modify the phenotype.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

We thank all the subjects and their families for participating in this study and the local coordinators who helped facilitate sample collection and processing. This work was funded by grants from the US National Institutes of Health (U54HG006542 and U54HD061222), the International Rett Syndrome Foundation, the Rett Syndrome Research Trust and the Cynthia and Anthony Petrello Scholar fund at the Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Eunice Kennedy Shriver Child Health and Human Development Institute.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999;23:185–188.
- Laurvick CL, de Klerk N, Bower C, et al. Rett syndrome in Australia: a review of the epidemiology. *J Pediatr* 2006;148:347–352.
- Neul JL, Kaufmann WE, Glaze DG, et al.; RettSearch Consortium. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol* 2010;68:944–950.
- Neul JL, Lane JB, Lee HS, et al. Developmental delay in Rett syndrome: data from the natural history study. *J Neurodev Disord*. 2014;6:20.
- Percy AK, Lane JB, Childers J, et al. Rett syndrome: North American database. *J Child Neurol* 2007;22:1338–1341.
- Ariani F, Hayek G, Rondinella D, et al. *FOXG1* is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet* 2008;83:89–93.
- Bahi-Buisson N, Nectoux J, Rosas-Vargas H, et al. Key clinical features to identify girls with *CDKL5* mutations. *Brain* 2008;131:2647–2661.
- Lombardi LM, Baker SA, Zoghbi HY. *MECP2* disorders: from the clinic to mice and back. *J Clin Invest* 2015;125:2914–2923.
- Guerrini R, Parrini E. Epilepsy in Rett syndrome, and *CDKL5*- and *FOXG1*-gene-related encephalopathies. *Epilepsia* 2012;53:2067–2078.
- Lambert L, Bienvenu T, Allou L, et al. *MEF2C* mutations are a rare cause of Rett or severe Rett-like encephalopathies. *Clin Genet* 2012;82:499–501.
- Ohba C, Nabatame S, Iijima Y, et al. De novo *WDR45* mutation in a patient showing clinically Rett syndrome with childhood iron deposition in brain. *J Hum Genet* 2014;59:292–295.
- Romaniello R, Saettini F, Panzeri E, Arrigoni F, Bassi MT, Borgatti R. A de-novo *STXBP1* gene mutation in a patient showing the Rett syndrome phenotype. *Neuroreport* 2015;26:254–257.
- Olson HE, Tambunan D, LaCoursiere C, et al. Mutations in epilepsy and intellectual disability genes in patients with features of Rett syndrome. *Am J Med Genet A* 2015;167:2017–2025.
- Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature* 2014;511:344–347.
- Lupski JR, Gonzaga-Jauregui C, Yang Y, et al. Exome sequencing resolves apparent incidental findings and reveals further complexity of *SH3TC2* variant alleles causing Charcot-Marie-Tooth neuropathy. *Genome Med* 2013;5:57.
- Epi KC, Allen AS, et al.; Epi4K Consortium and Epilepsy Phenome/Genome Project. De novo mutations in epileptic encephalopathies. *Nature* 2013;501:217–221.
- EuroEPINOMICS-RES Consortium; Epilepsy Phenome/Genome Project; Epi4K Consortium. De novo mutations in synaptic transmission genes including *DNM1* cause epileptic encephalopathies. *Am J Hum Genet* 2014;95:360–370.
- Iossifov I, O'Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 2014;515:216–221.

19. O'Roak BJ, Deriziotis P, Lee C, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 2011;43:585–589.
20. O'Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 2012;485:246–250.
21. O'Roak BJ, Stessman HA, Boyle EA, et al. Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat Commun* 2014;5:5595.
22. Jiang YH, Yuen RK, Jin X, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am J Hum Genet* 2013;93:249–263.
23. Hamdan FF, Srour M, Capo-Chichi JM, et al. De novo mutations in moderate or severe intellectual disability. *PLoS Genet* 2014;10:e1004772.
24. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* 2014;385:1305–1314.
25. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* 2015;519:223–228.
26. Saunders CJ, Minassian BE, Chow EW, Zhao W, Vincent JB. Novel exon 1 mutations in *MECP2* implicate isoform MeCP2_e1 in classical Rett syndrome. *Am J Med Genet A* 2009;149A:1019–1023.
27. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 2012;380:1674–1682.
28. Sanders SJ, Murtha MT, Gupta AR, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 2012;485:237–241.
29. Lasalle JM. Autism genes keep turning up chromatin. *OA Autism* 2013;1:14.
30. Chao HT, Zoghbi HY, Rosenmund C. MeCP2 controls excitatory synaptic strength by regulating glutamatergic synapse number. *Neuron* 2007;56:58–65.
31. Zhang W, Peterson M, Beyer B, Frankel WN, Zhang ZW. Loss of MeCP2 from forebrain excitatory neurons leads to cortical hyperexcitation and seizures. *J Neurosci* 2014;34:2754–2763.
32. Rojas DC. The role of glutamate and its receptors in autism and the use of glutamate receptor antagonists in treatment. *J Neural Transm (Vienna)* 2014;121:891–905.
33. Gonzaga-Jauregui C, Harel T, Gambin T, et al.; Baylor-Hopkins Center for Mendelian Genomics. Exome sequence analysis suggests that genetic burden contributes to phenotypic variability and complex neuropathy. *Cell Rep* 2015;12:1169–1183.
34. Foti SB, Chou A, Moll AD, Roskams AJ. HDAC inhibitors dysregulate neural stem cell activity in the postnatal mouse brain. *Int J Dev Neurosci* 2013;31:434–447.
35. Carvill GL, McMahon JM, Schneider A, et al.; EuroEPINOMICS Rare Epilepsy Syndrome Myoclonic-Astatic Epilepsy & Dravet working group. Mutations in the GABA transporter SLC6A1 cause epilepsy with myoclonic-atonic seizures. *Am J Hum Genet* 2015;96:808–815.
36. Matejas V, Hinkes B, Alkandari F, et al. Mutations in the human laminin beta2 (*LAMB2*) gene and the associated phenotypic spectrum. *Hum Mutat* 2010;31:992–1002.
37. Sarkozy A, Carta C, Moretti S, et al. Germline BRAF mutations in Noonan, LEOPARD, and cardiofaciocutaneous syndromes: molecular diversity and associated phenotypic spectrum. *Hum Mutat* 2009;30:695–702.
38. Kashevarova AA, Nazarenko LP, Schultz-Pedersen S, et al. Single gene microdeletions and microduplication of 3p26.3 in three unrelated families: CNTN6 as a new candidate gene for intellectual disability. *Mol Cytogenet* 2014;7:97.
39. Coppinger J, McDonald-McGinn D, Zackai E, et al. Identification of familial and de novo microduplications of 22q11.21-q11.23 distal to the 22q11.21 microdeletion syndrome region. *Hum Mol Genet* 2009;18:1377–1383.
40. Esplin ED, Li B, Slavotinek A, et al. Nine patients with Xp22.31 microduplication, cognitive deficits, seizures, and talipes anomalies. *Am J Med Genet A* 2014;164A:2097–2103.
41. Srivastava S, Cohen J, Pevsner J, et al. A novel variant in *GABRB2* associated with intellectual disability and epilepsy. *Am J Med Genet A* 2014;164A:2914–2921.