

**Title:** Disruption of RFX family transcription factors causes autism, attention deficit/hyperactivity disorder, intellectual disability, and dysregulated behavior

**Authors:**

Holly K. Harris,<sup>1,2,43</sup> Tojo Nakayama,<sup>3,4,43</sup> Jenny Lai,<sup>3,42</sup> Boxun Zhao,<sup>3,4</sup> Nikoleta Argyrou,<sup>3,4</sup> Cynthia S. Gubbels,<sup>3,4</sup> Aubrie Soucy,<sup>3,4</sup> Casie A. Genetti,<sup>3,4</sup> Lance H. Rodan,<sup>3,4,5</sup> George E. Tiller,<sup>6</sup> Gaetan Lesca,<sup>7</sup> Karen W. Gripp,<sup>8</sup> Reza Asadollahi,<sup>9</sup> Ada Hamosh,<sup>10</sup> Carolyn D. Applegate,<sup>10</sup> Peter D. Turnpenny,<sup>11</sup> Marleen E.H. Simon,<sup>12</sup> Catharina (Nienke) M.L. Volker-Touw,<sup>12</sup> Koen L.I. van Gassen,<sup>12</sup> Ellen van Binsbergen,<sup>12</sup> Rolph Pfundt,<sup>13</sup> Thatjana Gardeitchik,<sup>13</sup> Bert B.A. de Vries,<sup>13</sup> Ladonna L. Imken,<sup>14</sup> Catherine Buchanan,<sup>14</sup> Marcia Willing,<sup>15</sup> Tomi L. Toler,<sup>15</sup> Emily Fassi,<sup>15</sup> Laura Baker,<sup>8</sup> Fleur Vansenne,<sup>16</sup> Xiadong Wang,<sup>17</sup> Julian L. Ambrus Jr.,<sup>18</sup> Madeleine Fannemel,<sup>19</sup> Jennifer E. Posey,<sup>20</sup> Emanuele Agolini,<sup>21</sup> Antonio Novelli,<sup>21</sup> Anita Rauch,<sup>9</sup> Paranchai Boonsawat,<sup>9</sup> Christina R. Fagerberg,<sup>22</sup> Martin J. Larsen,<sup>22</sup> Maria Kibaek,<sup>22</sup> Audrey Labalme,<sup>7</sup> Alice Poisson,<sup>7</sup> Katelyn K. Payne,<sup>23</sup> Laurence E. Walsh,<sup>23,24</sup> Kimberly Aldinger,<sup>25</sup> Jorune Balciuniene,<sup>26</sup> Cara Skraban,<sup>26</sup> Christopher Gray,<sup>26</sup> Jill Murrell,<sup>26</sup> Caleb P. Bupp,<sup>27</sup> Giulia Pascolini,<sup>28</sup> Paola Grammatico,<sup>28</sup> Martin Broly,<sup>29</sup> Sébastien Küry,<sup>29</sup> Mathilde Nizon,<sup>29</sup> Iqra Ghulam Rasool,<sup>30,31</sup> Muhammad Yasir Zahoor,<sup>30</sup> Cornelia Kraus,<sup>31</sup> André Reis,<sup>31</sup> Muhammad Iqbal,<sup>32</sup> Kevin Uguen,<sup>33,34</sup> Severine Audebert-Bellanger,<sup>33</sup> Claude Ferec,<sup>33,34</sup> Sylvia Redon,<sup>33,34</sup> Janice Baker,<sup>35</sup> Yunhong Wu,<sup>36</sup> Guiseppa Zampino,<sup>37</sup> Steffan Syrbe,<sup>38</sup> Ines Brosse,<sup>38</sup> Rami Abou Jamra,<sup>39</sup> William B. Dobyns,<sup>40</sup> Lilian L. Cohen,<sup>41</sup> Pankaj B. Agrawal,<sup>3,4</sup> Alan Beggs,<sup>3,4</sup> Timothy W. Yu\*<sup>3,4</sup>

**Affiliations:**

- 1) Division of Developmental Medicine, Department of Medicine, Boston Children's Hospital, Boston, MA, 02115, USA
- 2) Department of Pediatrics, Baylor College of Medicine and Meyer Center for Developmental Pediatrics, Texas Children's Hospital, Houston, TX, 77054, USA
- 3) Division of Genetics and Genomics, Department of Pediatrics, Boston Children's Hospital, Boston, MA, 02115, USA
- 4) The Manton Center for Orphan Disease Research, Boston Children's Hospital, Boston, MA, 02115, USA
- 5) Department of Neurology, Boston Children's Hospital, Boston, MA 02215, USA
- 6) Department of Genetics, Kaiser Permanente, Los Angeles, CA, 90027, USA
- 7) Department of Medical Genetics, Lyon University Hospital, 69500 Bron, France
- 8) Division of Medical Genetics, Nemours/A.I. DuPont Hospital for Children, Wilmington, DE, 19803, USA
- 9) Institute of Medical Genetics, University of Zurich, 8952 Schlieren-Zurich, Switzerland
- 10) Department of Genetic Medicine, Johns Hopkins University, Baltimore, MD, 21287, USA
- 11) Peninsula Clinical Genetics, Royal Devon and Exeter NHS Foundation Trust, Exeter EX1 2ED, UK
- 12) Department of Genetics, University Medical Centre Utrecht, Utrecht, The Netherlands
- 13) Department of Human Genetics, Radboud University Medical Centre, 6500 HB Nijmegen, The Netherlands
- 14) Dell Children's Medical Group, Department of Clinical and Metabolic Genetics, Austin, TX, 78723, USA
- 15) Division of Genetics and Genomic Medicine, Washington University School of Medicine in St. Louis, St. Louis, MO, 63110, USA
- 16) Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands
- 17) CIPHERGENE, Beijing, China
- 18) Division of Allergy, Immunology, and Rheumatology, SUNY at Buffalo School of Medicine, Buffalo, NY, 14203, USA
- 19) Department of Medical Genetics, Oslo University Hospital, 0372 Oslo, Norway
- 20) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030, USA

- 21) *Laboratory of Medical Genetics, Bambino Gesù Children's Hospital, 00146, Rome, Italy*
- 22) *Department of Clinical Genetics, Odense University Hospital, 5000 Odense, Denmark*
- 23) *Department of Neurology, Indiana University Health Neuroscience Center, Indianapolis, IN, 46202, USA*
- 24) *Department of Medical and Molecular Genetics, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, 46202, USA*
- 25) *Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA, 98101, USA*
- 26) *Division of Genomic Diagnostics, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA*
- 27) *Spectrum Health Helen DeVos Children's Hospital, Grand Rapids, MI 49503, USA*
- 28) *Laboratory of Medical Genetics, Department of Molecular Medicine, Sapienza University, San Camillo-Forlanini Hospital, 00185 Roma RM, Italy*
- 29) *CHU Nantes, Service de Génétique Médicale, Nantes, France; L'institut du thorax, INSERM, CNRS, UNIV Nantes, CHU Nantes, Nantes, France*
- 30) *Institute of Biochemistry & Biotechnology, University of Veterinary & Animal Sciences, 54000 Lahore, Pakistan*
- 31) *Institute of Human Genetics, University of Erlangen-Nuremberg, 91054 Erlangen, Germany*
- 32) *Department of Biochemistry and Biotechnology, The Islamia University of Bahawalpur, Punjab, Pakistan*
- 33) *Department of Medical Genetics, Brest University Hospital, 29200 Brest, France*
- 34) *Univ Brest, Inserm, EFS, UMR 1078, GGB, F-29200 Brest, France*
- 35) *Department of Genomic Medicine, Children's Minnesota, Minneapolis, MN 55404*
- 36) *Shanxi Children's Hospital, Taiyuan, China*
- 37) *Center for Rare Disease and Congenital Defects, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Università Cattolica del Sacro Cuore, 00168, Rome, Italy*
- 38) *Division of Pediatric Epileptology, Center for Pediatric and Adolescent Medicine, University Hospital Heidelberg, Heidelberg, Germany*
- 39) *Institute of Human Genetics, University of Leipzig Medical Center, Germany*
- 40) *Departments of Pediatrics and Genetics, University of Minnesota, Minneapolis, MN, 55455, USA*
- 41) *Division of Medical Genetics, Weill Cornell Medical College, New York, NY, 10021, USA*
- 42) *Program in Neuroscience, Harvard University, Boston, MA, 02115, USA*
- 43) *These authors contributed equally to this work*

**\*Correspondence to:**

Timothy W. Yu  
Division of Genetics and Genomics  
Boston Children's Hospital  
300 Longwood Avenue, Mailstop BCH3150  
Boston, MA 02115  
[timothy.yu@childrens.harvard.edu](mailto:timothy.yu@childrens.harvard.edu)

## ABSTRACT

**Purpose:** We describe a novel neurobehavioral syndrome of autism spectrum disorder, intellectual disability, and attention deficit/hyperactivity disorder associated with *de novo* or inherited deleterious variants in members of the *RFX* family of genes. *RFX* genes are evolutionarily conserved transcription factors that act as master regulators of central nervous system development and ciliogenesis.

**Methods:** We assembled a cohort of 36 individuals (from 31 unrelated families) with *de novo* mutations in *RFX3*, *RFX4*, and *RFX7*. We describe their common clinical phenotypes and present bioinformatic analyses of expression patterns and downstream targets of these genes as they relate to other neurodevelopmental risk genes.

**Results:** These individuals share neurobehavioral features including autism spectrum disorder (ASD), intellectual disability, and attention-deficit/hyperactivity disorder (ADHD); other frequent features include hypersensitivity to sensory stimuli and sleep problems. *RFX3*, *RFX4*, and *RFX7* are strongly expressed in developing and adult human brain, and X-box binding motifs as well as *RFX* ChIP-seq peaks are enriched in the cis-regulatory regions of known ASD risk genes.

**Conclusion:** These results establish deleterious variation in *RFX3*, *RFX4*, and *RFX7* as important causes of monogenic intellectual disability, ADHD and ASD, and position these genes as potentially critical transcriptional regulators of neurobiological pathways associated with neurodevelopmental disease pathogenesis.

**KEYWORDS:** autism, intellectual disability, attention-deficit/hyperactivity disorder, transcription factor, neurodevelopment, genetic disorders

## INTRODUCTION

Autism spectrum disorder (ASD), marked by deficits in social communication and the presence of restricted interests and repetitive behavior,<sup>1</sup> is highly heritable and genetically heterogeneous,<sup>2</sup> with *de novo* loss-of-function variants known to be an important contributor to ASD risk.<sup>3-5</sup> ASD is often comorbid with other neurodevelopmental diagnoses, including attention-deficit/hyperactivity disorder (ADHD), a behavioral diagnosis that captures a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning.<sup>1</sup> Emerging evidence also points to a role of *de novo* loss-of-function variants in ADHD,<sup>6-8</sup> and recent analyses of genome-wide association studies from over 20,000 individuals with ADHD suggests a role for common variation at 12 significant risk loci, implicating genes involved in neurotransmitter regulation and neuronal plasticity.<sup>9</sup>

*RFX3* is a member of the regulatory factor X (*RFX*) gene family which encodes transcription factors with a highly-conserved DNA binding domain. Located on chromosome 9p24.2, the gene is comprised of 307,708 base pairs and encodes a transcribed protein that is 749 amino acids in length. *RFX3* is expressed in several tissues including developing and adult brain. Other *RFX* family members (*RFX1*, 4, 5, and 7) are also highly expressed in brain tissue<sup>10-13</sup>, with expression patterns of *RFX1*, 3, 4 and 7 clustering especially tightly.<sup>13</sup>

We report a series of 36 individuals from 31 families with deleterious, mostly *de novo* variants in three brain-expressed members of the *RFX* family: *RFX3*, *RFX4*, or *RFX7*. *RFX3* was among 102 genes recently identified as statistically enriched for *de novo* variants in a large-scale analysis of trio exome data from individuals with ASD,<sup>14</sup> but to date *RFX4* and *RFX7* have not been previously associated with human disease. Analysis of their clinical presentations reveals common features including intellectual disability (ID), ASD, and/or ADHD, delineating a novel neurobehavioral phenotype associated with *RFX* haploinsufficiency.

## MATERIALS AND METHODS

## **Ethics Statement**

This series was compiled via an international collaborative effort involving Boston Children's Hospital, Kaiser Permanente, Lyon University Hospital, Nemours/A.I. DuPont Hospital for Children, University of Zurich, Johns Hopkins University, Peninsula Clinical Genetics at Royal Devon and Exeter NHS Foundation Trust, University Medical Centre Utrecht, Radboud University Medical Centre, Dell Children's Medical Group, Washington University School of Medicine in St. Louis, University Medical Center Groningen, Ciphergene, SUNY at Buffalo School of Medicine, Oslo University Hospital, Baylor College of Medicine, Bambino Gesù Children's Hospital, Odense University Hospital, Indiana University Health Neuroscience Center, Seattle Children's Research Institute, Children's Hospital of Philadelphia, Spectrum Health Helen DeVos Children's Hospital, Sapienza University and San Camillo-Forlanini Hospital, CHU Nantes et Service de Génétique Médicale, University of Erlangen-Nuremberg, The Islamia University of Bahawalpur, Brest University Hospital, Children's Minnesota, Università Cattolica del Sacro Cuore, University Hospital Heidelberg, University of Leipzig Medical Center, University of Washington, and Weill Cornell Medical College. Collaboration was facilitated by the online genetics/genomics resource GeneMatcher. Affected individuals were clinically assessed by at least one clinical geneticist from one of the participating centers. De-identified clinical data from collaborating institutions (collected with local IRB approval or deemed exempt from IRB review as per local institutional policy) was shared for analysis and publication under a study protocol approved by the Boston Children's Hospital IRB.

## **Case Ascertainment and Data Collection**

We obtained phenotypic data from 15 unrelated individuals with loss-of-function variants in *RFX3*, 3 unrelated individuals with loss-of-function variants in *RFX4*, and 13 unrelated individuals with loss-of-function variants in *RFX7*. De-identified individual case summaries are available upon request. Variants arose *de novo* with the exception of four related individuals from the same nuclear family with the same heterozygous loss-of-function variant in *RFX3*, and three other related cases in *RFX4* (homozygous for an

inherited missense variant). Pedigree information is shown in Figure 1. Diagnoses of ASD were reported in the medical record, but not uniformly evaluated by standardized measures such as the Diagnostic and Statistical Manual, Fourth or Fifth Edition (DSM-IV and 5),<sup>1,15</sup> Autism Diagnostic Observation Schedule (ADOS),<sup>16</sup> or Autism Diagnostic Interview, Revised (ADI-R).<sup>17</sup> Similarly, ID and ADHD diagnoses were accepted per clinician report and not always accompanied by standardized cognitive or behavioral testing measures.

### **Exome Sequencing**

Individuals included underwent exome sequencing on a clinical or research basis. Seven of the individuals were sequenced through GeneDx using genomic DNA from the proband or proband plus parents, captured using either the Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA) or the IDT xGen Exome Research Panel v1.0, and sequenced on an Illumina system with 100bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19, and variants were analyzed and interpreted as previously described<sup>18</sup> using variant classification criteria publicly available on the GeneDx ClinVar submission page (see Web Resources). Two cases of *RFX7* were sequenced through Ambry Genetics whose gene and variant classification process are available on the AmbryGenetics web page. The remainder of the individual's exome sequencing was performed through the clinicians' institutions or an external laboratory or research program (see Acknowledgments).

### **Variant Analyses**

Variant genomic coordinates are reported in relation to the Human Dec. 2013 (GRCh38/hg38) Assembly. The reference mRNA and protein sequences used are *RFX3* NM\_134428.2, NP\_602304.1; *RFX4* NM\_213594.2, NP\_998759.1; and *RFX7* NM\_022841.5, NP\_073752.5. The variant databases gnomAD v2.1.1 and v3 were examined for the presence of each variant.<sup>19</sup> Predicted pathogenicities for all variants were assessed by MutationTaster<sup>20</sup> (Schwarz et al., 2010). For missense variants, pathogenicity predictions were obtained from six algorithms: MutationTaster,<sup>20</sup> SIFT,<sup>21</sup> PolyPhen2,<sup>22</sup> PROVEAN,<sup>23</sup>

LRT,<sup>24</sup> and MutationAssessor,<sup>25</sup> and the total number of algorithms out of six with a deleterious prediction is referred to as the Nonsynonymous Damaging score (NsynD). In addition, the regional missense constraint metric MPC score<sup>26</sup> and CADD scaled score<sup>27</sup> are reported in Table S2 and Table S3. MPC  $\geq 2$  is considered highly deleterious.

### **Cell transfection and culture**

Human *RFX3* (NM\_134428.2; Human *RFX3* cDNA) was cloned into V5-or Myc- tagged mammalian expression vectors using the Gateway cloning system (Thermo Fisher Scientific). Point mutations were introduced with the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) to incorporate variants from affected individuals. To quantify the expression level of exogenous *RFX3*, equal amounts of tagged-*RFX3* expression vectors were transfected into Hela cells or HEK293T using Lipofectamine 3000 (Thermo Fisher Scientific). The transfected cells were cultured for 48 hours before harvesting. Cell extracts were analyzed by immunoblotting, using antibodies raised against RFX3 (HPA035689, Sigma-Aldrich), V5 (R960-25, Thermo Fisher Scientific), or beta actin (ab6276, Abcam). Blots were scanned on a Li-Cor Odyssey imager (Li-Cor). Signal intensities were quantified using Image Studio Lite (Li-Cor). Each immunoblot analysis was replicated three times.

### **KEGG pathway and ASD gene set over-representation analysis**

ChIP-seq and eCLIP-seq narrowPeak bed files for RFX family members, CREBBP, EP300, FMR1, FXR1, and FXR2 were obtained from the ENCODE portal,<sup>28</sup> and additional ChIP-seq data for *RFX3\_K562* were obtained from RegulomeDB<sup>29</sup> (Table S6). Functional binding genes (1 kb upstream/downstream of TSS [transcriptional start site]) were annotated using ChIPseeker.<sup>30</sup> ASD risk gene lists included 102 TADA genes from Satterstrom *et al.* and 253 ASD/ID genes from Coe *et al.* (Table S7).<sup>14,31</sup> Differentially expressed genes (DEGs) in ASD brains were extracted from Velmeshev *et al.* excluding endothelial DEGs that could originate from vascular cells in the brain (Table S7).<sup>32</sup> Customized KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed using

clusterProfiler<sup>33</sup> to determine the enrichment for KEGG pathways, ASD risk gene sets, and ASD DEGs. Multiple testing correction was performed using Benjamini-Hochberg correction (Table S8). Annotations, statistical analyses, and plots were implemented in R.

### **Motif analysis**

For motif occurrence analysis, FIMO was used to scan promoter sequences for individual occurrences of RFX motifs.<sup>34</sup> For all analyses, motif models were obtained from the JASPAR 2020 database.<sup>35</sup> Motifs searched for included RFX3 (MA0798.1), RFX4 (MA0799.1), and RFX7 (MA1554.1). Promoter sequences were defined as -1000 base pairs and +500 base pairs relative to the transcription start site. Motif occurrences were classified as significant based on a reporting threshold of p-value <0.00001 and q-value (Benjamini) <0.10. For motif enrichment analysis, we used the HOMER findMotifs.pl and findMotifsGenome.pl scripts.<sup>36</sup> Motifs were classified as enriched based on fold-enrichment > 1.5 and q-value (Benjamini) <0.01. Enhancer sequences associated with genes of interest were obtained from Enhancer Atlas 2.0.<sup>37</sup> ASD risk gene lists were obtained as noted previously to include 102 TADA genes and 124 ASD/ID genes reaching exome-wide significance.<sup>14,31</sup>

## **RESULTS**

**Index Case:** A teenaged male presented for genetic evaluation of ASD, ADHD, and borderline ID. His prenatal course was uncomplicated. Developmental delays in language, social communication, and motor functioning were first observed at 18 months of age. He had a history of sleep disturbance that improved by adolescence. He had notable deterioration of his behavioral functioning in adolescence, requiring placement in an intensive therapeutic program. His behavioral profile was marked by mood lability, anxiety, impulsivity, aggression, noise hypersensitivity and self-soothing motor stereotypies (e.g. rocking). An electroencephalogram (EEG) was obtained for a clinical concern for absence seizures but was normal. Medical history was otherwise unremarkable. Growth parameters were normal including head circumference. Physical examination was notable for subtle dysmorphisms including a high arched



palate and down-turned corners of the mouth, mild hypotonia and generalized joint hyperlaxity. Strength, reflexes, and sensation were normal.

Exome sequencing revealed a heterozygous in-frame deletion of three nucleotides in the coding region of *RFX3* (c.584\_586delGAA; p.(Glu195del)). This variant is absent from dbSNP and gnomAD and parental sequencing revealed the deletion to be *de novo*. It is predicted to lie within the conserved DNA binding domain of *RFX3* and removes an amino acid (glutamate-195) that is evolutionarily conserved in all 9/9 species examined down to zebrafish. *RFX3* is under significant loss-of-function constraint in databases of human population variation (pLI score 1.0 in ExAC, gnomAD).

### **Case series of individuals with *RFX3* pathogenic variants**

We identified and obtained clinical information from 18 individuals inclusive of our index case via GeneMatcher,<sup>38</sup> all bearing loss-of-function variants in *RFX3*. Genotypic information is provided in Table 1, clinical phenotypes are summarized in Tables 2 and S1, and variant pathogenicity predictions are summarized in Supplemental Table S2. A total of 15 distinct variants were identified: two frameshift variants, two canonical splice donor variants, eight missense variants, one in-frame deletion, one 42 kb deletion removing the last two exons of *RFX3*, and one 227 kb deletion involving only *RFX3*. In one family, an affected parent transmitted a frameshift variant to three affected children; all other variants were *de novo* and novel (Figure 1).

Phenotypic data from these 18 individuals (thirteen males and five females) are presented in Table 2. There were no sex-based differences in severity of phenotype. All individuals had neurodevelopmental delays, with formally recorded clinical diagnoses of ASD (72%) and ID (61%) of varying severity (borderline to moderate; described as global developmental delay in young children), and ADHD (56%) (Table S1). Many showed a distinct behavioral pattern marked by easy excitability/overstimulation, hypersensitivity to sensory (particularly auditory) stimuli, anxiety, emotional dysregulation and/or aggression (13/15 [87%] with specific behavioral information provided). Three individuals were reported to have seizures (17%). Some individuals had sleep difficulties (44%) including

limited total duration of sleep, frequent awakenings, or early morning awakenings. Subtle non-specific and non-recurrent dysmorphisms were commonly reported (61%), including broad nasal bridge, high, arched palate, and hand and foot abnormalities (tapered fingers, widely spaced toes), but no consistent recognizable features were shared by all individuals. Both macrocephaly (six individuals) and microcephaly (two individuals) were reported (8/11 individuals [73%] with a head circumference measurement or percentile provided). Magnetic resonance imaging (MRI) of the brain was available for eight individuals, with reports of non-specific findings in four, including white matter changes, uncus asymmetry, partially empty sella, or prominent ventricles. One individual had mild thinning of the corpus callosum (Table 2). Five of seven individuals (71%) who were past the onset of puberty (ages 12-30 years) had reports of behavioral and/or cognitive worsening at the time of puberty/adolescence. Three had increased aggression specifically noted. Three were described as having manic and/or psychotic symptoms, specifically two described as having hallucinations (one requiring psychiatric hospitalization) and another described as having conversations with imaginary friends. Three were reported to have had decline in cognition, one in adolescence and another around 28 years of age.

### ***RFX3* pathogenic variants lead to reduced *RFX3* levels**

The types of *RFX3* variants encountered strongly supported a loss-of-function model in which mutations disrupt or destabilize expression of *RFX3* protein. We tested this hypothesis by engineering a subset of variants into *RFX3* cDNAs for expression analysis in HeLa and HEK293 cells (Figure 2B). As expected, the frameshift variant (p.Trp569ValfsX6, predicted to cause nonsense-mediated decay of the mutant mRNA) led to essentially complete loss of *RFX3* protein expression. Missense variants p.(Asp611Tyr), p.(Leu443Ile), p.(Phe383Ser), and the stop-gain variant p.(Glu195del) also led to significant decreases in *RFX3* levels, suggesting that *RFX3* pathogenic variants lead to haploinsufficiency through reduced protein levels (Figure 2B).

### **Pathogenic variants in additional RFX family genes cause related neurodevelopmental phenotypes**

Additional individuals were ascertained who harbored loss-of-function variants in other closely related genes of the *RFX* family. Thirteen individuals bearing *de novo* loss-of-function variants in *RFX7* were identified (Tables 1 and 2), including three frameshift variants, five stop gain variants, one in-frame deletion, and two missense variants (Table 1, Table S2). Again, more males were identified than females (eight males, four females) without differences in phenotype based on sex. All individuals had language delay, and most had ID/global developmental delay (85%) (Tables 2 and 3). While formal diagnoses of ASD (31%) and/or ADHD (31%) were less consistent, autistic features and/or significant behavioral challenges akin to those seen in *RFX3* individuals were reported in the majority of cases, including excitability/overstimulation, sensitivity to sensory (particularly auditory) stimuli, a high pain threshold, emotional dysregulation, aggression, and anxiety (8/8, 100% of those with specific behavioral information provided). Abnormal head size (five individuals with microcephaly and three with macrocephaly) was noted in 8/10 (80%) that provided head circumference measurements. In 4/9 patients (44%) who had neuroimaging, MRI abnormalities were observed (Dandy-Walker malformation, cerebellar tonsillar herniation, an abnormality of the basal ganglia, and a fourth case with limited information but an “abnormal brain MRI” noted). Subtle clinical dysmorphisms were reported in 92% including abnormalities of the hands and feet such as widely spaced toes, syndactyly, or long tapered fingers (50%) (Table 2). Again, no consistent dysmorphisms were evident across individuals.

Five individuals with probable loss-of-function *RFX4* variants were also identified (Tables 1 and 2). Two were individuals who harbored *de novo* *RFX4* variants, including an in-frame deletion (*RFX4* p.(Tyr639\_Ser643del)), and a predicted pathogenic missense variant (*RFX4* p.(Arg79Ser)) (Table 1, Supplemental Table S2). Three additional related individuals (siblings) were homozygous for a missense variant (p.Thr247Met) altering a well-conserved threonine residue. Of these five individuals, three were female and two were male. Most were noted to have ID (80%) and ASD (80%). All individuals were normocephalic. Neuroimaging was performed in one and demonstrated asymmetric volume loss. Seizures were described in two individuals (40%). No consistent dysmorphisms were evident.

### ***RFX3*, *RFX4*, and *RFX7* variant analyses**

In total, 31 distinct variants in *RFX* family members (15 *RFX3*, 3 *RFX4*, and 13 *RFX7* were identified (Table 1). Excluding related individuals, each case involved a novel variant (e.g., there were no recurrent variants). *RFX3*, *RFX4*, and *RFX7* each exhibit intolerance to loss-of-function variation in human population databases (gnomAD, pLI scores = 1.00). All variants were absent from gnomAD except for *RFX7* p.Pro964\_Thr965del, which is detected at a very low frequency in gnomAD v2.1.1 (AF 0.00007677) leading us to formally classify it as a variant of uncertain significance (VUS).

Nineteen protein truncating variants and 12 missense variants were found. We assessed the pathogenicity of the missense variants using six algorithms (MutationTaster, SIFT, PolyPhen2, PROVEAN, LRT, and MutationAssessor) and 11/12 missense variants were predicted to be damaging by at least four algorithms (NsynD score  $\geq 4$ ) and the twelfth was predicted damaging by three algorithms (*RFX4* p.(Thr247Met), NsynD3) (Table S2). All missense variants affect highly conserved amino acids (PhastCons vertebrate, mammalian, and primate scores ranging from 0.99-1.00) (Table S2). *RFX* transcription factors are defined by a conserved, specialized winged-helix type DNA binding domain (DBD) that recognize the X-box motif. In addition to the DBD, *RFX3* and *RFX4* have three known domains that are associated with dimerization (DD).<sup>13</sup> All variants identified in *RFX3* are located in either the DBD or one of the dimerization domains (15/15, 100%) (Figure 2A, Table 1). *RFX4* and *RFX7* variants did not exhibit clustering to specific functional domains.

We examined 35 additional reported variants in *RFX3*, *RFX4*, and *RFX7* from prior studies of *de novo* or inherited variants in ASD and neuropsychiatric conditions (Table S3, Figure S1).<sup>4,14,39-42</sup> Missense variants from the literature tended to be of milder predicted pathogenicity than those reported here (Figure 2C). Sixteen were in *RFX3*, including five *de novo* variants (four protein truncating and one missense variant predicted damaging by all six algorithms, NsynD6, supportive of likely pathogenicity), seven inherited variants (four CNVs, one frameshift variant (p.(Pro408fs)) and three missense variants (p.(Thr151Ala), p.(Ala101Thr), p.(Arg615His); NsynD scores 3-6), and four copy number variants (all microdeletions) were reported for which parental inheritance was not established (Table S3). Among

previously reported *RFX7* variants, one was a *de novo* frameshift and one was an inherited frameshift variant. There were also six reported inherited missense variants (6/6 with NsynD >4), and two *de novo* missense variants that are likely benign. Finally, there were nine previously reported *RFX4* variants, only one of which was *de novo* (a missense variant lacking strong evidence of pathogenicity), and eight inherited missense variants of varying predicted pathogenicity (NsynD scores 3-6).

### ***RFX* expression is enriched in human brain**

*RFX3*, *RFX4* and *RFX7* have been reported to have relatively high expression in human fetal cortex.<sup>43</sup> To determine whether specific cell types are affected by *RFX* haploinsufficiency, we examined single-cell transcriptomes from developing and adult human cortex (Figure 3A-F, Figure S2A-B).<sup>32,44</sup> In developing human cortex, *RFX3* and *RFX7* exhibited the strongest brain expression, with *RFX3* most highly expressed in maturing excitatory upper enriched neurons, *RFX4* most highly expressed in outer radial glia, and *RFX7* most highly expressed in interneurons from the medial ganglionic eminence (Figure 3A-C). We also examined *RFX* expression patterns in the adult human cortex (Figure 3D-F, Figure S2A-B).<sup>32,44</sup> Again, *RFX3* and *RFX7* exhibited the highest expression. *RFX3* was most highly expressed in glutamatergic layer 2/3 neurons, followed by astrocytes. *RFX7* was expressed in both inhibitory and excitatory neurons. *RFX4* expression was much lower overall, but highest in astrocytes (Figure 3F). These patterns are similar to prior reports from developing and adult mouse neocortex, suggesting that developmental and cell-specific regulation of *RFX* expression is highly conserved during mammalian cortical development.<sup>45</sup> These expression profiles suggest that *RFX* pathogenic variants may lead to our observed neurodevelopmental phenotypes by altering early developmental cell fates or by impacting the function of upper-layer cortical neurons, astrocytes, and interneurons.

### ***RFX* binding motifs are present in ASD risk gene cis-regulatory regions**

Dysregulated gene expression, especially in upper layer cortical neurons, has been implicated in ASD pathogenesis.<sup>32,46</sup> Given the expression of *RFX3* in layer 2/3 neurons and the autistic features of

individuals reported here, we considered whether *RFX* family genes might be important transcriptional regulators of ASD risk genes. RFX family transcription factors bind to a characteristic consensus motif called an X-box (GTHNYY AT RRNAAC)<sup>47</sup> with individual family members having additional specificity for particular subsequences within this consensus. We therefore performed *RFX3*, *4*, and *7* motif enrichment analysis in upstream regulatory sequences of 187 ASD risk genes (the union of 102 TADA genes from Satterstrom et al., 2020 and 124 genes meeting exome-wide significance from Coe et al, 2019)<sup>14,31</sup> and an additional set of 447 genes identified (via single cell transcriptome analyses) to be upregulated in ASD brains.<sup>32</sup> We found enrichment of X-box motifs (q-value <0.05) in human ESC-neuron specific enhancers for ASD risk genes (Table S5A). As a group, *RFX3* and *RFX4* motifs were particularly enriched (q-value <0.005), while the *RFX7* motif was not (q-value 0.48). X-box, *RFX3* and *RFX4* motifs were similarly enriched in the enhancer regions of genes upregulated in ASD brains (Table S5B).<sup>32</sup> Enrichment of *RFX* motifs in promoter regions of ASD risk genes and DEGs did not emerge (data not shown). Last, we analyzed available RFX ChIP-seq data from the ENCODE project (Table S6) to determine enrichment for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, ASD risk gene sets, and ASD DEGs (Table S7). RFX functional binding genes from most ENCODE cell lines were significantly enriched in ASD risk genes and DEGs after multiple testing correction (p.adjust < 0.05; Benjamini-Hochberg's correction; Figure 3G, Figure S4, Table S8). Across cell lines, there was a positive correlation between enrichment in ASD genes and RFX expression levels in that cell type (Figure S5, Table S9), indicating that higher RFX expression levels may be required to engage ASD relevant targets.

Finally, single gene analyses showed enrichment of *RFX3* and *RFX4* motifs in the promoters of five ASD-associated genes (FIMO p-value <0.0001, q-value <0.1): *AP2S1*, *KDM6B*, *ANK2*, *NONO*, and *MYTIL* (Figure S3B, D),<sup>34</sup> and *RFX3* ENCODE ChIP-seq data from HepG2 cells confirmed RFX3 binding peaks in the promoters of *AP2S1*, *KDM6B*, and *NONO* (Figure S3E-G). Notably, *de novo* loss-of-function variants in *KDM6B* cause a neurodevelopmental syndrome that has phenotypic overlap with *RFX3* haploinsufficiency as described in this report, namely mild global delays, delayed speech, hypotonia, and features of ASD and ADHD, while loss-of-function variants in *NONO* and *MYTIL* are a

cause of X-linked and autosomal dominant intellectual disability, respectively (Table S4).<sup>48–55</sup> These cases support the model that *RFX* members may be transcriptional activators of a subset of ASD risk genes via actions at both enhancer and promoter sites.

## DISCUSSION

Our results delineate a novel human neurobehavioral syndrome including ASD, ID and/or ADHD due to deleterious variants in *RFX* family transcription factors. While presence of neuroimaging findings, seizures, and dysmorphisms varied between different *RFX* family members, the behavioral phenotypes of individuals with *RFX3*, *RFX4*, and *RFX7* were strikingly similar, and often included sensory hypersensitivity and impulsivity.

This report complements accumulating statistical genetic evidence for *RFX3* as an ASD risk gene,<sup>14,41</sup> and extends these findings to the closely related *RFX* family members *RFX4* and *RFX7*. Two-thirds of individuals with *RFX3* variants in our series carried an ASD diagnosis, half had ADHD, and just over half of individuals had ID. Several individuals with *RFX3* variants also exhibited post-pubertal cognitive or behavioral regression sometimes accompanied by psychosis. *RFX3* CNVs have been previously reported in schizophrenia<sup>40,56</sup>

In contrast to individuals with *RFX3*, individuals with *RFX7* variants showed almost uniform diagnoses of language delay and ID (92%). Individuals with *RFX7* variants were less likely to have ASD or ADHD diagnoses. Fewer individuals were identified with *RFX4* variants, but those identified had high rates of ASD and ID.

*RFX* family members have been previously known for their biological roles in cilia development. The *RFX3* transcription factor activates core components necessary for development and maintenance of both motile and immotile cilia,<sup>57–59</sup> and biallelic *Rfx3* knockout in mice results in situs inversus, hydrocephalus, and deficits in corpus callosum formation.<sup>59–61</sup> This raises the question of whether the neurodevelopmental phenotypes reported here may be mechanistically related to cilia development – e.g., a hypomorphic human ciliopathy. The majority of described genetic ciliopathies are recessive, and

therefore not due to haploinsufficiency, but some (e.g. Meckel syndrome, Joubert syndrome, Bardet-Biedl syndrome, oral-facial-digital syndrome type I) may be associated with neurodevelopmental abnormalities and/or brain malformations. Conversely, the individuals described in this report lack systemic features of ciliopathies.

In a competing model, the neurodevelopmental consequences of *RFX* haploinsufficiency may reflect non-ciliary roles of RFX transcription factors in brain development (Figure 3H). Finally, while not formally included in the case series presented here, we make note of an *RFX4* case that was ascertained after an early pregnancy termination and noted to have bilateral ventriculomegaly on autopsy. The fetus was determined to be a compound heterozygote for missense variants in *RFX4* that are of uncertain significance (maternally inherited c.55C>T, p.(Arg19Trp) and paternally inherited c.1519A>T, p.(Ile507Phe)), raising the possibility that biallelic mutations in *RFX* genes may lead to more severe phenotypes than those described in this series.

Enriched expression of *RFX3* in upper cortical layer neurons places this gene in cells that are involved in communication between regions of the cortex important for higher cognition and social behavior,<sup>62</sup> raising the possibility that haploinsufficiency may disrupt either the developmental specification, synaptic connectivity, or electrophysiological function of this set of neurons. Projection neurons in this layer have been implicated in ASD by analyses of co-expression networks of autism genes,<sup>63,64</sup> and superficial cortical neurons exhibit the strongest amount of differential gene expression in ASD brains compared to controls.<sup>32,46</sup> Sun and colleagues in fact showed strong enrichment of RFX motifs in differentially acetylated peaks upregulated in ASD brains compared to controls.<sup>46</sup> Future studies aimed at understanding the downstream targets of RFX family members in human brain may shed new light on pathways important to the molecular pathogenesis of ASD, ADHD, and ID.



## **SUPPLEMENTAL DATA**

Supplemental Data include detailed clinical descriptions for each affected individual, five figures, and nine tables.

## **ACKNOWLEDGEMENTS**

We thank the patients and their families for their participation and inclusion in this work. We also thank all the genetic counselors working with our patients and families. We thank the DDD study for providing exome sequencing results. We thank Billie Lianoglou, MS, CGC (Fetal Treatment Center, University of California at San Francisco Fetal Treatment Center, San Francisco, CA, 94158, USA) for her contributions of an RFX4 variant of unclear significance (see supplementary material). We thank GeneDx for their contributions as a collaborator via GeneMatcher, and Ambry Genetics for their contributions including facilitation of clinician communication. JL was supported by award T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. BZ was supported by the Manton Center Pilot Project Award and Rare Disease Research Fellowship. JEP and JRL were supported in part by the National Human Genome Research Institute (NHGRI) and National Heart Lung and Blood Institute (NHBLI) to the Baylor-Hopkins Center for Mendelian Genomics (BHCMG, UM1 HG006542). KU, SAB, CF, and SR were supported by the French Ministry of Health and the Health Regional Agency from Bretagne, Pays de la Loire and Centre Val de Loire (HUGODIMS 2, 2017). The DDD study is supported by the Wellcome Trust and the UK Department of Health Innovation Challenge Fund [HICF-1009-003] and the Wellcome Trust Sanger Institute [grant no. WT098051] (Nature 2015;519:223-8). TWY was supported by grant nos. NIH/NIMH R01MH113761, NICHD/NHGRI/NIH U19HD077671 and NIH/NICHD U24HD0938487, and by a SFARI Pilot Research Award.

## **AUTHOR CONTRIBUTIONS**

TWY, HKH, and TN conceptualized the paper, collected case information from collaborators, conducted and managed all functional studies, drafted the initial manuscript, and edited and revised the manuscript. TWY provided research study oversight. JL designed and performed *RFX* motif analysis, analyzed published brain transcriptome and single-cell RNA-sequencing data, analyzed *RFX3*, 4, and 7 variants and their predicted pathogenicities, cataloged *RFX* variants in previously published studies and analyzed their pathogenicities, contributed to the manuscript, and edited and revised the manuscript. BZ collected ChIP-seq data and performed over-representation analysis, contributed to the manuscript, and edited and revised the manuscript. NA and CSG conducted functional studies of the impact of *RFX3* variants on protein stability. AS performed *RFX* variant analyses. CAG assisted in research enrollment. LR, RP, TG, BBAdV, MEHS, KLIvG, EvB, C(N)MLV, AH, CDA, LLI, CB, MW, EF, TLT, KWG, LB, FV, PR, XW, JLA, MF, GET, JEP, JRL, EA, AN, RA, ARa, PB, CRF, MJL, MK, GL, AL, AP, KKP, LEW, KA, JB, CS, JM, CPB, GP, PG, MB, SK, MN, IGR, MYZ, CK, ARe, MI, KU, SA, CF, SR, MI, PDT, JB, YW, GZ, SS, IB, RAJ, WBD, and LLC contributed clinical case information and/or analyzed exome data. PBA and AB supported research subject enrollment.

## DECLARATION OF INTERESTS

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to [timothy.yu@childrens.harvard.edu](mailto:timothy.yu@childrens.harvard.edu).

## REFERENCES

1. American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.).
2. Geschwind, D.H. (2011). Genetics of autism spectrum disorders. *Trends Cogn Sci* *15*, 409–416.
3. Iossifov, I., O’Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* *515*, 216–221.
4. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* *515*, 209–215.
5. Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., Di Lullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* *485*, 237–241.
6. Banaschewski, T., Becker, K., Scherag, S., Franke, B., and Coghill, D. (2010). Molecular genetics of attention-deficit/hyperactivity disorder: An overview. *Eur. Child Adolesc. Psychiatry* *19*, 237–257.
7. Faraone, S. V., Perlis, R.H., Doyle, A.E., Smoller, J.W., Goralnick, J.J., Holmgren, M.A., and Sklar, P. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* *57*, 1313–1323.
8. Satterstrom, K.F., Kosmicki, J.A., Wang, J., Breen, M.S., Rubeis, S., An, J.-Y., Peng, M., Collins, R., Grove, J., Klei, L., et al. (2019). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *BioRxiv* 484113.
9. Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., Baldursson, G., Belliveau, R., Bybjerg-Grauholm, J., Bækvad-Hansen, M., et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* *51*, 63–75.

10. Feng, C., Li, J., and Zuo, Z. (2011). Expression of the transcription factor regulatory factor X1 in the mouse brain. *Folia Histochem. Cytobiol.* 49, 344–351.
11. Aftab, S., Semenc, L., Chu, J., and Chen, N. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evol. Biol.* 8, 226.
12. Ma, K., Zheng, S., and Zuo, Z. (2006). The transcription factor regulatory factor X1 increases the expression of neuronal glutamate transporter type 3. *J. Biol. Chem.* 281, 21250–21255.
13. Sugiaman-Trapman, D., Vitezic, M., Jouhilahti, E.-M., Mathelier, A., Lauter, G., Misra, S., Daub, C.O., Kere, J., and Swoboda, P. (2018). Characterization of the human RFX transcription factor family by regulatory and target gene analysis. *BMC Genomics* 19, 181.
14. Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L., et al. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* 180, 568–584.
15. American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.).
16. Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., and Bishop, S. (2012). Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1-4.
17. Lord, C., Rutter, M., and Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* 24, 659–685.
18. Retterer, K., Juusola, J., Cho, M.T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K.G., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. *Genet. Med.* 18, 696–704.
19. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L.,

- Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443.
20. Schwarz, J.M., Rödelberger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7, 575–576.
21. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4, 1073–1081.
22. Adzhubei, I., Jordan, D.M., and Sunyaev, S.R. (2013). Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. *Curr. Protoc. Hum. Genet.* 07, Unit7.20.
23. Choi, Y., and Chan, A.P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 31, 2745–2747.
24. Chun, S., and Fay, J.C. (2009). Identification of deleterious mutations within three human genomes. *Genome Res.* 19, 1553–1561.
25. Reva, B., Antipin, Y., and Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* 39, e118–e118.
26. Samocha, K.E., Kosmicki, J.A., Karczewski, K.J., O’Donnell-Luria, A.H., Pierce-Hoffman, E., MacArthur, D.G., Neale, B.M., and Daly, M.J. (2017). Regional missense constraint improves variant deleteriousness prediction. *BioRxiv* 148353.
27. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 47, D886–D894.
28. Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., et al. (2018). The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* 46, D794–D801.
29. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J.,

Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 22, 1790–1797.

30. Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31, 2382–2383.

31. Coe, B.P., Stessman, H.A.F., Sulovari, A., Geisheker, M.R., Bakken, T.E., Lake, A.M., Dougherty, J.D., Lein, E.S., Hormozdiari, F., Bernier, R.A., et al. (2019). Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nat. Genet.* 51, 106–116.

32. Velmeshev, D., Schirmer, L., Jung, D., Haeussler, M., Perez, Y., Mayer, S., Bhaduri, A., Goyal, N., Rowitch, D.H., and Kriegstein, A.R. (2019). Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* 364, 685–689.

33. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287.

34. Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018.

35. Fornes, O., Castro-Mondragon, J.A., Khan, A., van der Lee, R., Zhang, X., Richmond, P.A., Modi, B.P., Correard, S., Gheorghe, M., Baranašić, D., et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 48, D87–D92.

36. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* 38, 576–589.

37. Gao, T., and Qian, J. (2020). EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. *Nucleic Acids Res.* 48, D58–D64.

38. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: A Matching Tool for

Connecting Investigators with an Interest in the Same Gene. *Hum. Mutat.* 36, 928–930.

39. Sahoo, T., Theisen, A., Rosenfeld, J. a, Lamb, A.N., Ravnan, J.B., Schultz, R. a, Torchia, B.S., Neill, N., Casci, I., Bejjani, B. a, et al. (2011). Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genet. Med.* 13, 868–880.

40. Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A., et al. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320, 539–543.

41. Li, J., Wang, L., Guo, H., Shi, L., Zhang, K., Tang, M., Hu, S., Dong, S., Liu, Y., Wang, T., et al. (2017). Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. *Mol. Psychiatry* 22, 1282–1290.

42. Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.-X., et al. (2015). Excess of rare, inherited truncating mutations in autism. *Nat. Genet.* 47, 582–588.

43. Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature* 478, 483–489.

44. Polioudakis, D., de la Torre-Ubieta, L., Langerman, J., Elkins, A.G., Shi, X., Stein, J.L., Vuong, C.K., Nichterwitz, S., Gevorgian, M., Opland, C.K., et al. (2019). A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron* 103, 785–801.

45. Loo, L., Simon, J.M., Xing, L., McCoy, E.S., Niehaus, J.K., Guo, J., Anton, E.S., and Zylka, M.J. (2019). Single-cell transcriptomic analysis of mouse neocortical development. *Nat. Commun.* 10, 1–11.

46. Sun, W., Poschmann, J., Cruz-Herrera del Rosario, R., Parikshak, N.N., Hajan, H.S., Kumar, V., Ramasamy, R., Belgard, T.G., Elanggovan, B., Wong, C.C.Y., et al. (2016). Histone Acetylome-wide

Association Study of Autism Spectrum Disorder. *Cell* 167, 1385–1397.

47. Efimenko, E., Bubb, K., Mak, H.Y., Holzman, T., Leroux, M.R., Ruvkun, G., Thomas, J.H., and Swoboda, P. (2005). Analysis of *xbx* genes in *C. elegans*. *Development* 132, 1923–1934.

48. Stolerman, E.S., Francisco, E., Stallworth, J.L., Jones, J.R., Monaghan, K.G., Keller-Ramey, J., Person, R., Wentzensen, I.M., McWalter, K., Keren, B., et al. (2019). Genetic variants in the *KDM6B* gene are associated with neurodevelopmental delays and dysmorphic features. *Am. J. Med. Genet. Part A* 179, 1276–1286.

49. Mircescu, D., Langouët, M., Rio, M., Moutton, S., Siquier-Pernet, K., Bole-Feysot, C., Cagnard, N., Nitschke, P., Gaspar, L., Žnidarič, M., et al. (2015). Mutations in *NONO* lead to syndromic intellectual disability and inhibitory synaptic defects. *Nat. Neurosci.* 18, 1731–1736.

50. Reinstein, E., Tzur, S., Cohen, R., Bormans, C., and Behar, D.M. (2016). Intellectual disability and non-compaction cardiomyopathy with a de novo *NONO* mutation identified by exome sequencing. *Eur. J. Hum. Genet.* 24, 1635–1638.

51. Scott, D.A., Hernandez-Garcia, A., Azamian, M.S., Jordan, V.K., Kim, B.J., Starkovich, M., Zhang, J., Wong, L.-J., Darilek, S.A., Breman, A.M., et al. (2017). Congenital heart defects and left ventricular non-compaction in males with loss-of-function variants in *NONO*. *J. Med. Genet.* 54, 47–53.

52. de Ligt, J., Willemsen, M.H., van Bon, B.W.M., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C., et al. (2012). Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N. Engl. J. Med.* 367, 1921–1929.

53. De Rocker, N., Vergult, S., Koolen, D., Jacobs, E., Hoischen, A., Zeesman, S., Bang, B., Béna, F., Bockaert, N., Bongers, E.M., et al. (2015). Refinement of the critical 2p25.3 deletion region: the role of *MYT1L* in intellectual disability and obesity. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 17, 460–466.

54. Blanchet, P., Bebin, M., Bruet, S., Cooper, G.M., Thompson, M.L., Duban-Bedu, B., Gerard, B.,



Piton, A., Suckno, S., Deshpande, C., et al. (2017). MYT1L mutations cause intellectual disability and variable obesity by dysregulating gene expression and development of the neuroendocrine hypothalamus. *PLoS Genet.* *13*, e1006957.

55. Loid, P., Mäkitie, R., Costantini, A., Viljakainen, H., Pekkinen, M., and Mäkitie, O. (2018). A novel MYT1L mutation in a patient with severe early-onset obesity and intellectual disability. *Am. J. Med. Genet. Part A* *176*, 1972–1975.

56. Sahoo, T., Theisen, A., Rosenfeld, J.A., Lamb, A.N., Ravnán, J.B., Schultz, R.A., Torchia, B.S., Neill, N., Casci, I., Bejjani, B.A., et al. (2011). Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genet. Med. Off. J. Am. Coll. Med. Genet.* *13*, 868–880.

57. El Zein, L., Ait-Lounis, A., Morle, L., Thomas, J., Chhin, B., Spassky, N., Reith, W., and Durand, B. (2009). RFX3 governs growth and beating efficiency of motile cilia in mouse and controls the expression of genes involved in human ciliopathies. *J. Cell Sci.* *122*, 3180–3189.

58. Choksi, S.P., Lauter, G., Swoboda, P., and Roy, S. (2014). Switching on cilia: transcriptional networks regulating ciliogenesis. *Development* *141*, 1427–1441.

59. Bonnafé, E., Touka, M., AitLounis, A., Baas, D., Barras, E., Ucla, C., Moreau, A., Flamant, F., Dubruille, R., Couble, P., et al. (2004). The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. *Mol. Cell. Biol.* *24*, 4417–4427.

60. Benadiba, C., Magnani, D., Niquille, M., Morlé, L., Valloton, D., Nawabi, H., Ait-Lounis, A., Otsmane, B., Reith, W., Theil, T., et al. (2012). The ciliogenic transcription factor RFX3 regulates early midline distribution of guidepost neurons required for corpus callosum development. *PLoS Genet.* *8*, e1002606.

61. Baas, D., Meiniel, A., Benadiba, C., Bonnafé, E., Meiniel, O., Reith, W., and Durand, B. (2006). A deficiency in RFX3 causes hydrocephalus associated with abnormal differentiation of ependymal cells.

Eur. J. Neurosci. 24, 1020–1030.

62. Sorensen, S.A., Bernard, A., Menon, V., Royall, J.J., Glattfelder, K.J., Desta, T., Hirokawa, K., Mortrud, M., Miller, J.A., Zeng, H., et al. (2015). Correlated gene expression and target specificity demonstrate excitatory projection neuron diversity. *Cereb. Cortex* 25, 433–449.

63. Willsey, A.J., Sanders, S.J., Li, M., Dong, S., Tebbenkamp, A.T., Muhle, R.A., Reilly, S.K., Lin, L., Fertuzinhos, S., Miller, J.A., et al. (2013). XCoexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* 155, 997.

64. Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). XIntegrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* 155, 1008.

65. Durand, B., Sperisen, P., Emery, P., Barras, E., Zufferey, M., Mach, B., and Reith, W. (1997). RFXAP, a novel subunit of the RFX DNA binding complex is mutated in MHC class II deficiency. *EMBO J.* 16, 1045–1055.

66. Reith, W., Siegrist, C.A., Durand, B., Barras, E., and Mach, B. (1994). Function of major histocompatibility complex class II promoters requires cooperative binding between factors RFX and NF-Y. *Proc. Natl. Acad. Sci. U. S. A.* 91, 554–558.

67. 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.

## FIGURE TITLES AND LEGENDS

### Figure 1. Individuals with Pathogenic Variants in *RFX* Members.

*RFX3*, *RFX4*, and *RFX7* case pedigrees. All mutations were de novo except for p.(Leu496Alafs\*7) which was found in an affected mother and three children, and p.(Thr247Met), found in the homozygous state in three affected children.

### Figure 2. Characterization of *RFX* Variants.

(A) Mapping of selected *RFX* variants to domains. Whole gene deletion and intronic variants are not illustrated. *RFX3* (NP\_602304.1), *RFX4* (NP\_998759.1), *RFX7* (NP\_073752.5).

(B) Immunoblot analysis of *RFX3* abundance in Hela cells overexpressed with *RFX3* mutants.

Quantification of *RFX3* abundance was performed. Values represent the average of three independent experiments. Error bars represent standard deviation. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ .

(C) Missense variant pathogenicity scores for the currently reported variants (current) and prior reported variants (prior) in *RFX3*, *4*, and *7*. The distribution of MPC scores for missense variants reported in this study is significantly different from that of prior reported missense variants, Kolmogorov-Smirnov (K-S) test  $p$ -value  $< 0.01$ . MPC, Missense badness, PolyPhen-2, and Constraint. NsynD, Nonsynonymous Damaging score. CADD, Combined Annotation Dependent Depletion.

### Figure 3. *RFX3*, *RFX4*, and *RFX7* expression patterns in human cortex and haploinsufficiency gene dosage model.

(A) Transcriptomic cell types in the prenatal human cortex identified by single-cell RNA-sequencing.<sup>44</sup>

(B) *RFX3*, *4*, and *7* expression patterns in single cells of the prenatal human cortex.

(C) Heatmap of *RFX3*, *4*, and *7* expression levels among cell types in the prenatal human cortex.

(D) Transcriptomic cell types in the postnatal human cortex identified by single-cell RNA-sequencing.<sup>32</sup>

- (E) *RFX3*, *4*, and *7* expression patterns in single cells of the postnatal human cortex.
- (F) Heatmap of *RFX3*, *4*, and *7* expression levels among cell types in the postnatal human cortex.
- (G) The enrichment of KEGG pathways, ASD risk gene sets, and ASD differentially expressed genes (DEGs) among *RFX3* ChIP-seq binding targets. Pathways and ASD gene sets are ranked by their statistical significance (p.adjust values, Benjamini-Hochberg's correction). Red arrows indicate ASD risk gene sets and ASD DEGs. X-axis shows the number of genes bound by *RFX* in their promoter regions.
- (H) Binding of *RFX* family transcription factors bind to X-box motif in promoter regions of ciliary and immunologic genes. Target gene lists obtained from Durand, Reith, Sugiaman-Trapman.<sup>13,65,66</sup> Model of *RFX* gene dose-dependent regulation of genes. In tissues with higher expression of *RFX* genes, ASD genes are activated. Lower levels of *RFX* genes are sufficient to activate ciliary genes.
- vRG, ventricular radial glia. oRG, outer radial glia. PgG2M, cycling progenitors G2/M phase. PgS, cycling progenitors S phase. IP, intermediate progenitors. ExN, migrating excitatory. ExM, maturing excitatory. ExM-U, maturing excitatory upper enriched. ExDp1, excitatory deep layer 1. ExDp2, excitatory deep layer 2. InMGE, interneuron MGE. InCGE, interneuron CGE. OPC, oligodendrocyte precursor cells. End, endothelial. Per, pericyte. Mic, microglia. Neu-mat, immature neurons. Neu-NRGN, NRGN expressing neurons. L5/6, layer 5/6 excitatory neurons. L5/6-CC, layer 5/6 excitatory cortico-cortical projection neurons. L4, layer 4 excitatory neurons. L2/3, layer 2/3 excitatory neurons. IN-SST, somatostatin interneurons. IN-PV, parvalbumin interneurons. IN-SV2C, SV2C expressing interneurons. IN-VIP, VIP interneurons. AST-FB, fibrous astrocytes. AST-PP, protoplasmic astrocytes. OPC, oligodendrocyte precursor cells.

## SUPPLEMENTAL FIGURE TITLES AND LEGENDS

### Figure S1. Pathogenicity scores of missense variants in *RFX3*, *RFX4*, and *RFX7*.

(A) The *RFX3* missense variants reported in this study have a higher frequency of pathogenic predictions than prior reported missense variants.

(B) *De novo* missense variants are more likely pathogenic than inherited missense variants.

(C) *RFX3* missense variants are more likely damaging than *RFX4* or *7* missense variants.

(D) Prior reported missense variants in *RFX3*, *4*, and *7* do not show clear associations between predicted pathogenicity and inheritance.

MPC, Missense badness, PolyPhen-2, and Constraint. NsynD, Nonsynonymous Damaging score.

CADD, Combined Annotation Dependent Depletion.

**Figure S2. *RFX3*, *RFX4*, and *RFX7* expression patterns in human cortex.**

(A) *RFX3*, *4*, and *7* expression patterns in single cells of adult human cortex (Allen Human Brain Atlas).<sup>67</sup> Cell types identified in human cortex shown in t-SNE plot with taxonomy of clusters. Expression color-scale units are Log<sub>2</sub>(CPM+1). Data from Allen Brain Map single-nucleus RNA-sequencing of human cortex (Image credit: Allen Institute).

(B) Heatmap visualizing *RFX3*, *4*, and *7* expression among adult human cortical cell types, along with canonical cell type markers. (Image credit: Allen Institute).

(C) *RFX3*, *RFX4*, and *RFX7* expression across all human tissues. (Data Source: GTEx Analysis Release V8 dbGaP Accession phs000424.v8.p2).

**Figure S3. Presence of *RFX3* and *RFX4* binding motifs in ASD-associated gene promoters.**

(A) *RFX3* binding motif MA0798.1 (JASPAR 2020).

(B) ASD risk genes with significant *RFX3* binding motif occurrences.

(C) *RFX4* binding motif MA0799.1 (JASPAR 2020).

(D) ASD risk genes with significant *RFX4* binding motif occurrences.

(E) – (G) RFX3 ChIP-seq binding peaks are located in promoter region of ASD-associated genes (E) AP2S1, (F) KDM6B, (G) NONO.

All motif occurrences were identified using FIMO, q-values <0.10. RFX3 in HepG2 ChIP-seq binding peaks obtained from ENCODE GSM2534235.

**Figure S4. Customized KEGG enrichment analysis of RFX functional binding sites.**

The enrichment of KEGG pathways, ASD risk gene sets, and ASD differentially expressed genes (DEGs) for different RFX binding profiles. (A–C) representative barplots of the enriched customized KEGG pathways were shown using RFX1 (A), and RFX5 (B-C) ChIP-seq data. Pathways and ASD gene sets are ranked by their statistical significance (p.adjust values, Benjamini-Hochberg's correction). Red arrows indicate ASD risk gene sets and ASD DEGs. X-axis shows the number of genes bound by RFX in their promoter regions. Data used for Figure S4D were obtained from the GTEx Portal, dbGaP accession number [phs000424.v8.p2](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108084) on 05/01/2020.

**Figure S5. Correlation of enrichment significance in ASD gene sets with *RFX* expression level.**

*RFX* expression levels in different cell lines were correlated with the enrichment significances [ $-\log_{10}(\text{p.adjust})$ ] for (A) ASD 514 DEGs; (B) TADA 102 ASD risk genes; (C) 253 ASD/ID risk genes. Red dotted line indicates the P-value = 0.05. X-axis shows the GTEx expression of *RFX* in different cell lines derived from a variety of human tissues.

## TABLE TITLES AND LEGENDS

### **Table 1. Molecular findings in individuals with ASD, ADHD, and/or ID and variants in *RFX3*, *RFX4*, or *RFX7*.**

Molecular characterization of *RFX3*, *RFX4*, and *RFX7* variants reported in this study. Chromosome structure is described according to the Human Dec. 2013 (GRCh38/hg38) Assembly. RefSeq identifiers: *RFX3* NM\_134428.2, NP\_602304.1; *RFX4* NM\_213594.2, NP\_998759.1; *RFX7* NM\_022841.5, NP\_073752.5. Protein domains were obtained from Sugiaman-Trapman et al., 2018.<sup>13</sup> Italics indicates individual has a variant of uncertain significance.

### **Table 2. Clinical features of individuals with pathogenic variants in *RFX3*, *RFX4*, or *RFX7*.**

\*ASD = We recognize the heterogeneity in ASD diagnoses. For our table, individuals were considered to have ASD if documented in the clinical note as having ASD diagnosed via formal measure, according to clinical expertise, or documented as having clear "autistic features," a designation we considered equivalent to a diagnosis of ASD for purposes of this report. Granular description of social communication or restrictive and repetitive behavior data to determine DSM-5 diagnosis was not uniformly available. NA, not available. ID, intellectual disability. GDD, global developmental delay. ADHD, attention deficit hyperactivity disorder. MRI, magnetic resonance imaging. CT, computerized tomography. PVL, periventricular leukomalacia. Italics indicates individual has a variant of uncertain significance.

## SUPPLEMENTAL TABLE TITLES AND LEGENDS

### **Table S1. Summary of clinical features associated with pathogenic variants in *RFX3*, *RFX4* or *RFX7*.**

Clinical features summarized by the number of positive reports out of the total number of individuals per gene, with the exception of "Micro or Macrocephaly" and "Neuroimaging findings," where the fraction is

the number of positive reports out of the total number of reports available per gene. Proportions are shown in parentheses.

**Table S2. Predicted pathogenicity of *RFX* variants.**

NA, not applicable. Italics indicates individual has a variant of uncertain significance.

**Table S3. *RFX* variants referenced in prior published studies of neurodevelopmental delay.**

RefSeq identifiers: RFX3 NM\_134428.2, NP\_602304.1; RFX4 NM\_213594.2, NP\_998759.1; RFX7 NM\_022841.5, NP\_073752.5. gDNA coordinates are provided as originally reported by each reference (Hg19: Krumm et al., 2015, Li et al., 2017, Sahoo et al., 2011, Walsh et al., 2008. Hg38: De Rubis et al., 2014, Tabet et al., 2015). AD, activating domain. DBD, DNA binding domain. DD, dimerization domain. NA, not applicable.

**Table S4. Phenotypes associated with ASD risk genes potentially regulated by *RFX* members.**

De novo loss of function variants in *KDM6B*, *NONO*, and *MYT1L* are associated with NDD/ID phenotypes. These phenotypes are similar to those seen in the individuals with *RFX* variants in this present study. AD, autosomal dominant. XL, X-linked.

**Table S5. Enrichment of *RFX* binding motifs in enhancer regions associated with ASD risk genes and ASD upregulated genes in human embryonic stem cell (ESC)-derived neurons.**

A) *RFX* binding motifs enriched in human ESC-neuron enhancer regions associated with ASD risk genes, q-value <0.05. B) *RFX* binding motifs enriched in human ESC-neuron enhancer regions associated with ASD upregulated genes, q-value <0.05. Lack of enrichment for *RFX* binding motifs in promoters of ASD upregulated genes. Fold-enrichment reported as the % motifs identified in target sequences/% motifs identified in background sequences. All motif enrichment analyses were performed using HOMER tool.



**Table S6. ChIP-seq and eCLIP datasets from the ENCODE project.**

**Table S7. Customized ASD risk and ASD DEG gene sets.**

**Table S8. Enrichment statistics of customized KEGG analysis.**

**Table S9. Correlation between RFX expression level and enrichment for ASD gene sets.**

**Table 1. Molecular findings in individuals with ASD, ADHD, and/or ID and variants in *RFX3*, *RFX4*, or *RFX7***

Gene	Inheritance	gDNA (GRCh38)	cDNA	Protein	Category	Domain
<i>RFX3</i>	<i>de novo</i>	chr9:g.3293222_3293224del	c.584_586del	p.Glu195del	Inframe deletion	DBD
	<i>de novo</i>	chr9:g.3293086A>C	c.722T>G	p.Leu241Trp	Missense	DBD
	<i>de novo</i>	chr9:g.9p24.2del	NA	(gene deletion)	Deletion	(all)
	<i>de novo</i>	chr9:g.3197716_3239762del	c.1968+8270_*27326del	(exon 17 and exon 18 deleted) p.Phe647_Val749del	Deletion	DD
	<i>de novo</i>	chr9:g.3248169C>A	c.1831G>T	p.Asp611Tyr	Missense	DD
	<i>de novo</i>	chr9:g.3257101dupG	c.1704dupG	p.Trp569Valfs*6	Frameshift	DD
	<i>de novo</i>	chr9:g.3248031C>T	c.1968+1G>A	exon 16 splice donor loss	Splicing	DD
	inherited	chr9:g.3263053_3263054del	c.1486_1487del	p.Leu496Alafs*7	Frameshift	DD
	<i>de novo</i>	chr9:g.3301541C>T	c.549+5G>A	exon 6 splice donor loss	Splicing	DBD
	<i>de novo</i>	chr9:g.3270401G>T	c.1327C>A	p.Leu443Ile	Missense	EDD
	<i>de novo</i>	chr9:g.3293134A>G	c.674T>C	p.Phe225Ser	Missense	DBD
	<i>de novo</i>	chr9:g.3271057A>G	c.1148T>C	p.Phe383Ser	Missense	EDD
	<i>de novo</i>	chr9:g.3293221C>T	c.587G>A	p.Gly196Glu	Missense	DBD
	<i>de novo</i>	chr9:g.3263017G>T	c.1523C>A	p.Ala508Glu	Missense	DD
<i>de novo</i>	chr9:g.3256992T>C	c.1813A>G	p.Ser605Gly	Missense	DD	
<i>RFX4</i>	<i>de novo</i>	chr12:g.106750773_106750787del	c.1915_1929del	p.Tyr639_Ser643del	Inframe deletion	NA
	<i>de novo</i>	chr12:g.106654271C>A	c.235C>A	p.Arg79Ser	Missense	DBD
	recessive (homozygous)	chr12:g.106696353C>T	c.740C>T	p.Thr247Met	Missense	NA
<i>RFX7</i>	<i>de novo</i>	chr15:g.56094696del	c.3032del	p.Ser1011Ilefs*19	Frameshift	NA
	unknown (adopted)	chr15:g.56096328_56096329del	c.1399_1400del	p.Met467Glu fs*19	Frameshift	NA
	<i>de novo</i>	chr15:g.56095010G>T	c.2718C>A	p.Tyr906*	Stop Gain	NA
	<i>de novo</i>	chr15:g.56095503G>C	c.2225C>G	p.Ser742*	Stop Gain	NA
	<i>de novo</i>	chr15:g.56095615dupA	c.2113dupT	p.Thr705Nfs*12	Frameshift	NA
	<i>de novo</i>	chr15:g.56094730dupT	c.2998dupA	p.Thr1000AsnfsTer9	Frameshift	NA
	<i>de novo</i>	chr15:g.56095010G>C	c.2718C>G	p.Tyr906*	Stop Gain	NA
	<i>de novo</i>	chr15:g.56095444C>A	c.2284G>T	p.Gly762*	Stop Gain	NA
	<i>de novo</i>	chr15:g.56094643T>C	c.3085A>G	p.Ile1029Val	Missense	NA
	<i>de novo</i>	chr15:g.56094645G>A	c.3083C>T	p.Pro1028Leu	Missense	NA
	<i>de novo</i>	chr15:g.56094648del	c.3080del	p.Thr1027fs	Frameshift	NA
	<i>de novo</i>	chr15:g.56095495C>A	c.2233G>T	p.Glu745*	Stop Gain	NA
	<i>de novo</i>	chr15:g.56094864_56094869del	c.2859_2864del	p.Pro964_Thr965del	Inframe deletion	NA

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

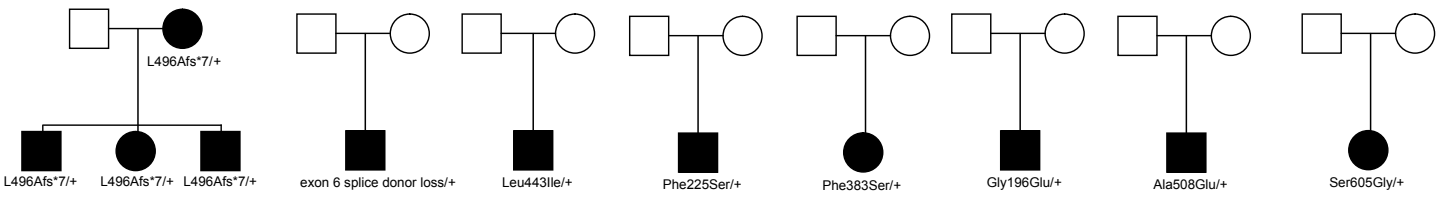
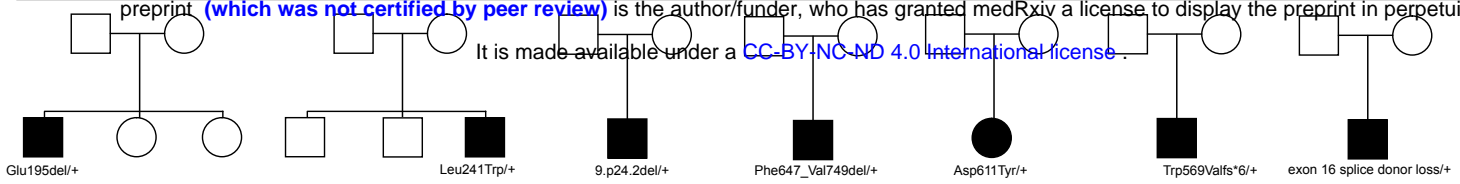
Table 2. Clinical features of individuals with pathogenic variants in *RFX3*, *RFX4*, or *RFX7*

Age (years)	Sex	Presentation	Variant	Language delay	Motor delay	ASD*	ID	ADHD	Behavioral Profile	Sleep issues	Seizures	Hypotonia or Neurologic Findings	Dysmorphism	Micro or macrocephaly	Neuroimaging (MRI unless specified)	Other Medical or Neuropsychiatric
<b><i>RFX3</i></b>																
20-24	M	ASD, ID, ADHD	inframe deletion	yes	yes	yes	yes	yes	mood swings, anxiety, aggression, sensory hypersensitivity, rocking	yes	no	yes	yes	No	NA	behavioral decline in adolescence
0-4	M	GDD	missense	NA	yes	NA	NA	NA	NA	NA	yes	yes	yes	Microcephaly	thin corpus callosum	NA
5-9	M	ADHD, anxiety	deletion	yes	yes	yes	yes	yes	NA	NA	no	NA	yes	Mid macrocephaly	normal nonspecific findings	myopia
5-9	M	ASD, ADHD	splicing	yes	no	no	yes	NA	aggression, sensory-seeking behavior, elopement, and impulsivity	yes	no	yes	no	No	partially empty sella	myopia
5-9	M	ASD, ID, ADHD	frameshift	yes	NA	yes	yes	yes	aggression, biting	NA	no	NA	no	NA	NA	NA
5-9	F	GDD	frameshift	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
5-9	M	GDD	frameshift	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
5-9	M	GDD	missense	yes	yes	no	yes	no	No	no	no	no	yes	NA	NA	bilateral conductive hearing loss
5-9	F	ASD, ID, ADHD	missense	yes	yes	yes	yes	yes	aggression, anxiety, impulsivity	yes	no	no	yes	mid macrocephaly	NA	asthma
10-14	M	ASD, ADHD	deletion	no	yes	yes	no	yes	sensory hypersensitivity	yes	yes	NA	yes	Macrocephaly	NA	strabismus
10-14	F	GDD	missense	no	yes	yes	No	yes	oppositional behavior, aggression, anxiety	NA	no	NA	yes	Mid macrocephaly	NA	NA
10-14	M	ASD, GDD, ADHD	frameshift	no	no	yes	no	yes	aggression, anxiety, emotional dysregulation	NA	no	yes	no	macrocephaly	PVL	anxiety, mild cognitive and behavioral decline at puberty
10-15	M	ASD, ID	splicing	yes	NA	yes	yes	no	sensory hypersensitivity	yes	no	yes	yes	Mid macrocephaly	normal/nonspecific findings	NA
15-19	F	ID, ASD	missense	yes	no	yes	yes	yes	impulsivity, mood swings	no	no	NA	yes	Macrocephaly	normal	hallucinations, mania, behavioral decline
25-29	M	ASD, ID, ADHD	missense	yes	yes	yes	yes	yes	sensory seeking behavior, aggression, biting, pica, self-injury, sensory hypersensitivity	yes	no	yes	yes	No	normal/nonspecific findings	Hypogonadism, strabismus, bipolar, behavioral and mild cognitive decline in adolescence
30-35	M	ASD	missense	NA	NA	yes	NA	yes	NA	yes	NA	NA	NA	NA	NA	Chrohn's disease, anxiety, depression
30-35	M	ASD	missense	no	no	yes	yes	no	aggression, elopement	yes	yes	no	no	NA	uncal asymmetry	behavioral decline, hallucinations
30-35	F	ASD, ID	frameshift	yes	NA	yes	yes	NA	NA	NA	NA	NA	yes	NA	NA	NA
<b><i>RFX4</i></b>																
5-9	M	ID ASD, epilepsy	missense	NA	NA	NA	NA	NA	NA	NA	yes	NA	NA	NA	NA	NA
10-14	F	ID, ASD, behavior problems	missense	yes	yes	yes	yes	NA	yes, behavior challenges	NA	No	NA	no	no	NA	skeletal abnormalities
10-14	F	ID, ASD, behavior problems	missense	yes	yes	yes	yes	NA	yes, behavior challenges	NA	No	NA	no	no	NA	skeletal abnormalities
10-14	M	ID, ASD, behavior problems	missense	yes	yes	yes	yes	NA	yes, behavior challenges	NA	No	NA	no	no	NA	none
30-35	F	ASD, ID, epilepsy	inframe deletion	NA	NA	yes	yes	NA	hand flapping, hand wringing, inappropriate laughter	NA	Yes, generalized intractable	NA	NA	NA	asymmetric volume loss	NA
<b><i>RFX7</i></b>																
0-4	F	ID	frameshift	yes	yes	no	yes	no	NA	no	no	no	yes	microcephaly	normal	NA
5-9	M	ID, ASD	frameshift	yes	yes	yes	yes	no	sensory seeking behavior, sensory hypersensitivity, attention seeking behavior	no	NA	no	yes	macrocephaly	NA	underdeveloped scrotum
0-4	M	GDD, ASD	stop-gain	yes	no	yes	yes	yes	low frustration tolerance, hair pulling, high pain threshold	NA	no	NA	yes	NA	NA	eczema
5-9	F	GDD, ASD	stop-gain	yes	yes	yes	yes	no	sensory hypersensitivity and sensory seeking, excitable, high pain threshold	no	no	yes	yes	no	normal	hydronephrosis, constipation
5-9	M	ASD, GDD	frameshift	yes	yes	yes	yes	no	excitable, laughs easily, sensory hypersensitivity, aggressive when younger	yes	no	yes	no	macrocephaly	normal	none
0-4	M	abnormal brain MRI	inframe deletion	yes	NA	NA	NA	NA	NA	NA	NA	NA	yes	no	cerebellar vermis hypoplasia with marked uprotation, cystic enlargement of 4th ventricle	asymptomatic atrial septal and ventricular septal defects
5-9	M	ID	stop-gain	yes	yes	no	NA	NA	sensory hypersensitivity	no	no	yes	yes	macrocephaly	NA	neuroblastoma
10-14	M	GDD	missense	yes	yes	no	yes	yes	NA	no	febrile seizure x 1	hyperreflexia	no	microcephaly	normal	hypocalcemia
10-14	M	ID	missense	yes	yes	no	yes	no	no	yes	epilepsy with myoclonic seizures	dystonic movements, mixed hypohypertonia	yes	microcephaly	abnormality of the basal ganglia, delayed CNS myelination	optic nerve hypoplasia, cataract, mild hearing loss, laryngomalacia, recurrent bronchitis/pneumonia, cryptorchidism
10-14	F	GDD	stop-gain	yes	yes	NA	NA	NA	potential anxiety	NA	no	yes	yes	NA	abnormal myelination but no structural abnormality	loose anagen hair syndrome
15-19	F	ID, ADHD	stop-gain	yes	yes	no	yes	yes	sensory hypersensitivity, aggression, anxiety, hair pulling, skin picking, nail biting	no	no	no	yes	microcephaly	normal (CT)	obesity, mild scoliosis
15-19	F	ID	frameshift	yes	no	no	yes	no	short attention span, excessive fears/phobias, fixated interests, sensory hypersensitivity, poor social interactions	no	yes, absence during day and grand mal seizures in sleep but none since 14	no	long fingers	no	normal	none
NA	NA	ID, ASD	frameshift	yes	yes	yes	yes	yes	mood swings, anxiety, aggression, self-injury	yes	no	no	yes	NA	cerebellar tonsillar herniation, abnormal 4th ventricle	polyphagia

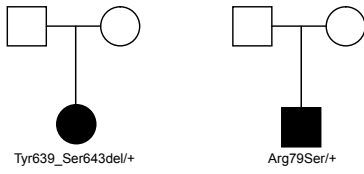
Figure 1. Individuals with pathogenic variants in RFX members.

medRxiv preprint doi: <https://doi.org/10.1101/2020.09.09.20187304>; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

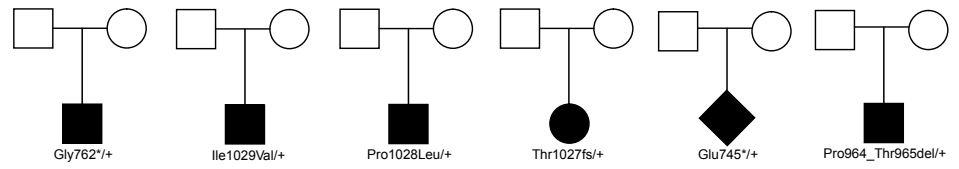
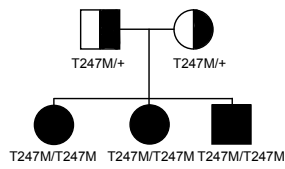
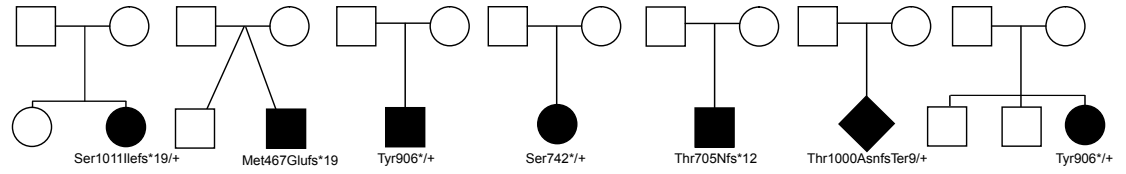
It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



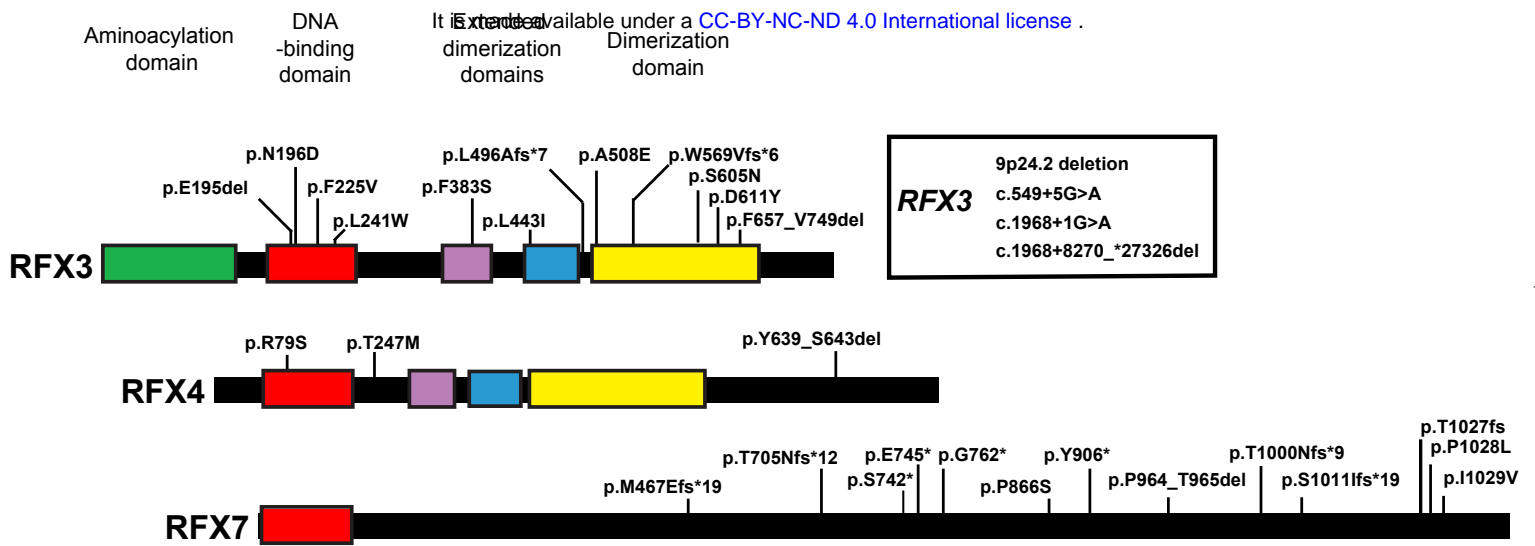
*RFX4* Pedigrees



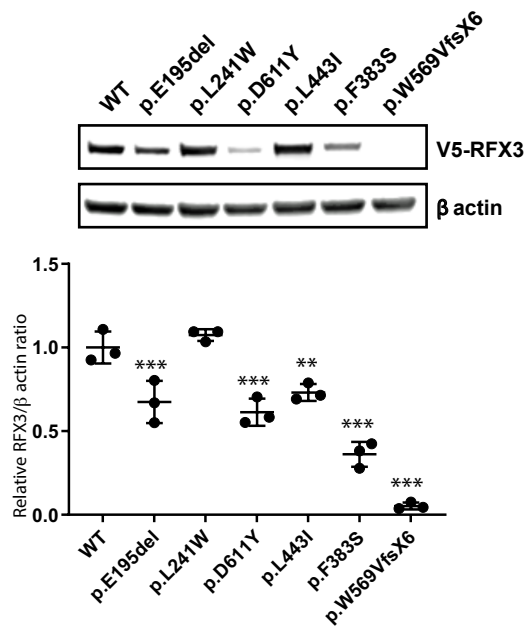
*RFX7* Pedigrees



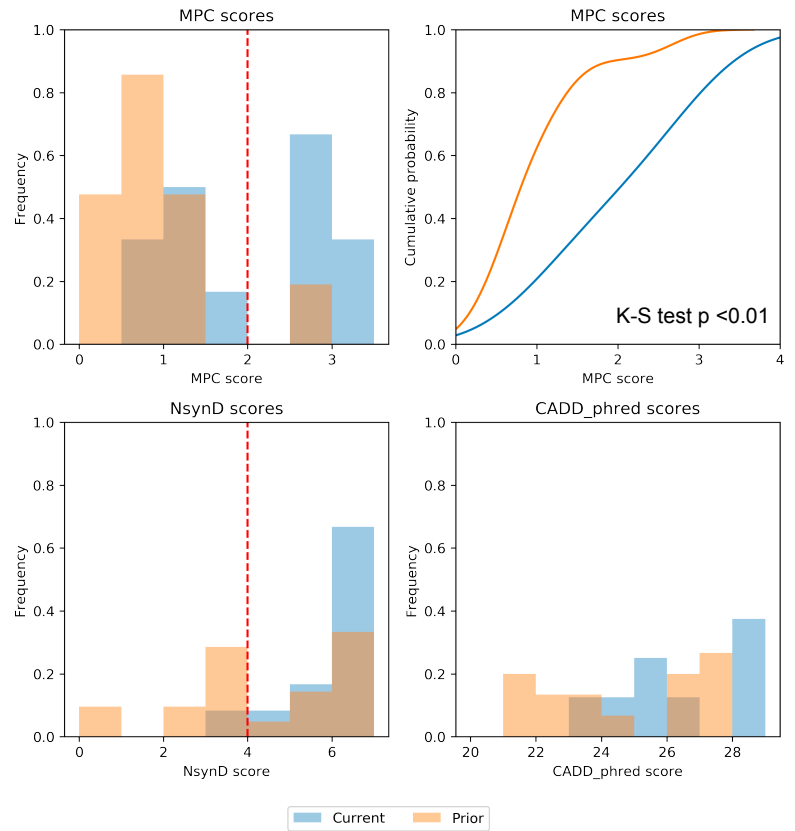
A



B



C



# Figure 3: RFX3, RFX4 and RFX7 expression patterns in human cortex

