Published in final edited form as:

Genet Med. 2021 June ; 23(6): 1028–1040. doi:10.1038/s41436-021-01114-z.

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

A full list of authors and affiliations appears at the end of the article.

Abstract

Purpose: We describe a novel neurobehavioral phenotype of autism spectrum disorder, intellectual disability, and/or 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 38 individuals (from 33 unrelated families) with *de novo* variants in *RFX3*, *RFX4*, and *RFX7*. We describe their common clinical phenotypes and present

[*]**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.

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, conducted variant analyses, 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, AB, and LLC contributed clinical case information and/or analyzed exome data. PBA and AB supported research subject enrollment.

ETHICS DECLARATION

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 Gesu 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, Universita Cattolica del Sacro Cuore, University Hospital Heidelberg, University of Leipzig Medical Center, University of Washington, Oslo University Hospital, 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. Consent for the publication of full-face photographs was obtained from all appropriate individuals in Fig. 1.

DECLARATION OF INTERESTS

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to timothy.yu@childrens.harvard.edu.

SUPPLEMENTAL DATA

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

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/or 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 a likely role of deleterious variation in *RFX3, RFX4*, and *RFX7* in cases of monogenic intellectual disability, ADHD and ASD, and position these genes as potentially critical transcriptional regulators of neurobiological pathways associated with neurodevelopmental disease pathogenesis.

INTRODUCTION

Autism spectrum disorder (ASD), marked by deficits in social communication and the presence of restricted interests and repetitive behavior, is highly heritable and genetically heterogeneous, with *de novo* loss-of-function variants as known contributors to ASD risk.[1] ASD is often comorbid with other neurodevelopmental diagnoses, including attention-deficit/hyperactivity disorder (ADHD). Emerging evidence also points to a role of *de novo* loss-of-function variants in ADHD.[2]

RFX3 is a member of the regulatory factor X (*RFX*) gene family which encodes transcription factors with a highly-conserved DNA binding domain. *RFX3* is expressed in several tissues including developing and adult brain, and other *RFX* family members (*RFX1, 4, 5*, and *7*) are also highly expressed in brain tissue, with expression patterns of *RFX1, 3, 4* and *7* clustering tightly.[3]

We report a series of 38 individuals from 33 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,[4] but to date *RFX4* and *RFX7* have not been previously associated with human disease. Analysis of case clinical data reveals common features including intellectual disability (ID), ASD, and/or ADHD, delineating a novel neurobehavioral phenotype associated with *RFX* haploinsufficiency.

MATERIALS AND METHODS

Case Ascertainment and Data Collection

We obtained phenotypic data from 15 unrelated individuals with loss-of-function variants in *RFX3*, 4 unrelated individuals with loss-of-function variants in *RFX4*, and 14 unrelated individuals with loss-of-function variants in *RFX7*. Individual case summaries for all individuals are provided (Supplemental Data). 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 and contributed photographs are 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), Autism Diagnostic Observation Schedule (ADOS), or Autism Diagnostic Interview, Revised (ADI-R). 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 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.[5] Predictions of the functional effects for all variants were assessed using MutationTaster, SIFT, PolyPhen2, PROVEAN, LRT, and MutationAssessor, and the total number of algorithms out of six with a deleterious prediction is referred to as the Nonsynonymous Damaging score (NsynD) as previously described.[6]

Cell transfection and culture

Human *RFX3* (NM_134428.2; Human *RFX3* cDNA) was cloned into V5-tagged mammalian expression vectors using the Gateway cloning system (Thermo Fisher Scientific). Point mutations were introduced with the QuikChange Lighting 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 using Lipofectamine 3000 (Thermo Fisher Scientific). The transfected cells were cultured for 48 hours before harvesting.

WEB RESOURCES
Ensembl Variant Effect Predictor, https://uswest.ensembl.org/info/docs/tools/vep/index.html
Online Mendelian Inheritance in Man (OMIM), https://www.omim.org/
GeneDx ClinVar submission page, https://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/
Ambry Genetics, https://ambrygen.com/clinician/our-scientific-excellence

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 six times. One-way ANOVA by repeated measures was employed. Multiple comparison correction was performed by using Dunnett statistical testing.

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,[7] and additional ChIP-seq data for RFX3_K562 were obtained from RegulomeDB[8] (Table S6). Functional binding genes (1 kb upstream/downstream of TSS [transcriptional start site]) were annotated using ChIPseeker.[9] ASD risk gene lists included 102 TADA genes from Satterstrom *et al.* and 253 ASD/ID genes from Coe *et al.* (Table S7).[4,10] 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).[11] The SYSCILIA Gold Standard (SCGSv.1) was used as a gold standard of known ciliary genes in human.[12] Customized KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed using clusterProfiler[13] to determine the enrichment for KEGG pathways, ciliary genes, 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.[14] For all analyses, motif models were obtained from the JASPAR 2020 database.[15] 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. Motifs were classified as enriched based on fold-enrichment > 1.5 over randomly selected background sequences with matched GC% content, and q-value (Benjamini) <0.01. Enhancer sequences associated with genes of interest were obtained from Enhancer Atlas 2.0.[16] ASD risk gene lists were obtained as noted previously to include 102 TADA genes and 124 ASD/ID genes reaching exome-wide significance.[4,10]

RESULTS

Case series of individuals with *de novo* or inherited *RFX3* variants

We identified and obtained clinical information from 18 individuals bearing loss-of-function variants in *RFX3* via GeneMatcher.[17] Genotypic information is provided in Table 1, clinical phenotypes are summarized in Tables 2 and S1, and predicted variant impacts 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 1A).

There were thirteen males and five females, with no sex-based differences in severity of phenotype. All individuals had neurodevelopmental delays, with formally recorded clinical diagnoses of ASD (72%) and ID of varying severity (borderline to moderate) or global developmental delay in young children (78%) and ADHD (56%) (Table 2 and 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 (Figure 1B). 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, uncal asymmetry, partially empty sella, or prominent ventricles. One individual had mild thinning of the corpus callosum (Table 2, individual RFX3–10). 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.

Variants in additional RFX family genes are associated with similar neurodevelopmental phenotypes

Additional individuals were ascertained who harbored loss-of-function variants in other closely related genes of the *RFX* family. Fourteen individuals bearing *de novo* loss-of-function variants in *RFX7* were identified (Tables 1 and 2), including four frameshift variants, five stop gain variants, one in-frame deletion, and two missense variants (Table 1, Table S2). Slightly more males were identified than females (eight males, six females) without differences in phenotype based on sex. All individuals had language delay, and most had ID/global developmental delay (93%) (Table 2, Supplemental Table 1). While formal diagnoses of ASD (36%) and/or ADHD (29%) 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 7/11

(64%) that provided head circumference measurements. In 5/11 patients (45%) 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 86% 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 (Figure 1D)

Six individuals with probable loss-of-function *RFX4* variants were also identified (Tables 1 and 2). Three were individuals who harbored *de novo RFX4* variants, including an in-frame deletion (RFX4 p.(Tyr639_Ser643del)), and two predicted damaging missense variants (RFX4 p.(Arg79Ser) and p.(Thr362Ala)) (Table 1, Supplemental Table S2). We also report a pedigree in which three additional related individuals (siblings) were homozygous for a missense variant in *RFX4* (p.Thr247Met) altering a well-conserved threonine residue. Parents of these siblings were first cousins, each heterozygous for the mutation and without any known neurobehavioral phenotype, and a heterozygous sibling was similarly reported as neurotypical; this pedigree therefore raises the possibility that the *RFX4* phenotype may be associated with both monoallelic and biallelic inheritance as has been described for several other genetic conditions.[18] Of these six individuals, three were female and three were male. All were noted to have ID or global developmental delay (100%) and most had documented ASD (83%). Four individuals were normocephalic, and one was microcephalic. Neuroimaging was performed in two and demonstrated asymmetric volume loss in one individual and absent pituitary gland in another individual with hypopituitarism. The latter individual also presented with cleft lip and palate. Seizures were described in two individuals (33%). No consistent dysmorphisms were evident (Figure 1C).

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

In total, 33 distinct variants in *RFX* family members (15 *RFX3*, 4 *RFX4*, and 14 *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) (see Supplemental RFX7 Case Descriptions Individual 14 for further details). The fact that the majority of variants identified are predicted to cause outright protein truncation or gene deletion (20 out of 33) strongly supports a loss-of-function / haploinsufficiency model. Of the thirteen missense variants, eleven were predicted to be damaging by at least four of six algorithms (NsynD score >=4) and two missense variants were predicted damaging by at least 2 algorithms (*RFX4* p.(Thr247Met) and p.(Thr362Ala)) (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).[3] *RFX4* and *RFX7* variants did not exhibit clustering to specific functional domains, but all of the non-truncating (missense or in-frame deletion) variants identified in *RFX3* were found to be located in the DBD or one of the dimerization domains (Figure 2A, Table 1). We engineered five of the non-truncating variants – p.(Glu195del), p.(Leu241Trp), p.(Phe383Ser), p.(Leu443Ile), and p.(Asp611Tyr) – into a V5-*RFX3* heterologous expression vector for protein stability analyses in HeLa cells (Figure S1). The majority of these variants resulted in significant decreases in detectable RFX3 levels, consistent with a destabilizing impact on protein expression. Two missense variants, p.(Leu241Trp) and p.(Leu443Ile) (residing in the DNA binding domain or the second extended protein dimerization domain, respectively) did not appear to impact protein stability, raising the possibility that they might disrupt more specific functional interactions of RFX3 to be investigated more thoroughly in the future.

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 S2).[4,19–22] Missense variants from the literature tended to be of milder predicted deleteriousness than those reported here (Figure 2B). 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 deleteriousness), 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 deleteriousness (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.[23] 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 S3A–B).[11,24] 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 S3A–B).[11,24] 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 expression profiles suggest that *RFX* deleterious 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.[11,25] 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)[26] 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) [4,10] and an additional set of 447 genes identified to be upregulated in ASD brains.[11] 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).[11] 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, ciliary genes, 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 S5, Table S8). Across cell lines, there was a positive correlation between enrichment in ASD genes and RFX expression levels in that cell type (Figure S6, 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 *MYT1L* (Figure S4B, D),[14] and *RFX3* ENCODE ChIP-seq data from HepG2 cells confirmed RFX3 binding peaks in the promoters of AP2S1, KDM6B, and NONO (Figure S4E–G). Notably, *de novo* loss-of-function variants in *KDM6B* (MIM 611577) 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* (MIM 300084) and *MYT1L* (MIM 613084) are a cause of X-linked and autosomal dominant intellectual disability, respectively (Table S4). 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 phenotype 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. Like ID/DD and ASD more generally,[27] individuals with RFX variants also exhibited a male bias.

This report complements accumulating statistical genetic evidence for *RFX3* as an ASD risk gene,[4,21] 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.[20,28] *RFX3* also lies within the region of the chromosome 9p deletion syndrome (OMIM#158170), associated with developmental delay, ID, and ASD, although the size of the deletions in this syndrome make *RFX3* unlikely to be the sole contributor.

This report also implicates both *RFX4* and *RFX7* as causes of human neurodevelopmental disorders. Individuals with *RFX4* or *RFX7* variants were somewhat more severely affected than those with *RFX3* variants, with *RFX7* less likely to be associated with ASD or ADHD, but showing almost uniform diagnoses of language delay and ID (92%). There were fewer individuals 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 primary cilia,[29–31] and biallelic *Rfx3* knockout in mice results in situs inversus, hydrocephalus, and deficits in corpus callosum formation.[31–33] 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. On the other hand, the individuals described in this report lack systemic features of ciliopathies, suggesting the alternative hypothesis that *RFX* haploinsufficiency may directly dysregulate the expression of ASD risk genes while leaving ciliary genes intact (Figure 3H). Future work, which may include analyzing cilia morphology and function in cellular or animal models of *RFX* haploinsufficiency, or characterizing transcriptome-wide effects of *RFX* gene disruption, may prove helpful in distinguishing these hypotheses.

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,[34] 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,[35,36] and superficial cortical neurons exhibit the strongest amount of differential gene expression in ASD brains compared to controls.[11,25] Sun and colleagues in fact showed strong enrichment of RFX motifs in differentially acetylated peaks upregulated in ASD brains compared to controls.[25] 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Holly K. Harris, MD[1,2,46], Tojo Nakayama, MD, PhD[3,4,46], Jenny Lai, BA[3,45,46], Boxun Zhao, PhD[3,4], Nikoleta Argyrou, MSc[3,4], Cynthia S. Gubbels, MD, PhD[3,4], Aubrie Soucy, BA[3,4], Casie A. Genetti, MS, CGC[3,4], Victoria Suslovitch, MS, CGC[3,4], Lance H. Rodan, MD, FRCP[3,4,5], George E. Tiller, MD, PhD[6], Gaetan Lesca, MD, PhD[7], Karen W. Gripp, MD[8], Reza Asadollahi, MD, PhD[9], Ada Hamosh, MD, MPH[10], Carolyn D. Applegate, MS, CGC[10], Peter D. Turnpenny, MD, BSc[11], Marleen E.H. Simon, MD[12], Catharina M.L. Volker-Touw, MD, PhD[12], Koen L.I. van Gassen, PhD[12], Ellen van Binsbergen, PhD[12], Rolph Pfundt, PhD[13], Thatjana Gardeitchik, MD[13], Bert B.A. de Vries, MD, PhD[13], LaDonna L. Immken, MD[14], Catherine Buchanan, MS, CGC[14], Marcia Willing, MD, PhD[15], Tomi L. Toler, MS, CGC[15], Emily Fassi, MS, CGC[15], Laura Baker, MS, CGC[8], Fleur Vansenne, MD, PhD[16], Xiadong Wang, PhD[17], Julian L. Ambrus Jr., MD[18], Madeleine Fannemel, MD, MHA[19], Jennifer E. Posey, MD, PhD[20], Emanuele Agolini, PhD[21], Antonio Novelli, PhD[21], Anita Rauch, MD[9], Paranchai Boonsawat, PhD[9], Christina R. Fagerberg, MD, PhD[22], Martin J. Larsen, PhD[22], Maria Kibaek, MD[22], Audrey Labalme, PhD[7], Alice Poisson, MD[7], Katelyn K. Payne, RN, CGC[23], Laurence E. Walsh, MD[23,24], Kimberly Aldinger, PhD[25], Jorune Balciuniene, PhD[26], Cara Skraban, MD[26], Christopher Gray, MS, CGC[26], Jill Murrell, PhD[26], Caleb P. Bupp, MD, FACMG[27], Giulia Pascolini, MD, PhD[28], Paola Grammatico, PhD[28], Martin Broly, MD[29], Sébastien Küry, PhD, DVM[29], Mathilde Nizon, MD[29], Iqra Ghulam Rasool, MPhil, PhD[30,31], Muhammad Yasir Zahoor, DVM, PhD[30], Cornelia Kraus, PhD[31], André Reis, MD[31], Muhammad Iqbal, PhD[32], Kevin Uguen, MD[33,34], Severine Audebert-Bellanger, MD[33], Claude Ferec, MD, PhD[33,34], Sylvia Redon, MD[33,34], Janice Baker, MS, CGC[35], Yunhong Wu, MD[36], Guiseppe Zampino, MD[37], Steffan Syrbe, MD[38], Ines Brosse, MD[38], Rami Abou Jamra, MD[39], William B. Dobyns, MD[40], Lilian L. Cohen, MD, MPH[41], Anne Blomhoff, MD, PhD[42], Cyril Mignot, MD[43,44], Boris Keren, MD, PhD[43], Thomas Courtin, MD[43], Pankaj B. Agrawal, MD, MMSc[3,4], Alan H. Beggs, PhD[3,4], Timothy W. Yu, MD, PhD[3,4,*]

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 Gesu 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, Universita 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]Department 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]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

[43]APHP.Sorbonne Université, Département de Génétique, Groupe Hospitalier Pitié-Salpêtrière, Paris, France

[44]Centre de Référence Déficiences Intellectuelles de Causes Rares, Paris, France

[45]Program in Neuroscience, Harvard University, Boston, MA, 02115, USA

[46]These authors contributed equally to this work

ACKNOWLEDGEMENTS

We thank Billie Lianoglou, MS, CGC (Fetal Treatment Center, University of California at San Francisco Fetal Treatment Center) for her contributions of an *RFX4* variant of unclear significance (see Supplementary). We thank the genetic counselors at GeneDx (Kirsty McWalter, MS, CGC and Erin Torti, MS, CGC) and Ambry Genetics (Meghan Towne, MS, CGC, LGC, Zoe Powis, MS, CGC, and Deepali Shinde, PhD) 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) and 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.

PERMISSIONS

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for Figure S3C were obtained from the GTEx Portal, dbGaP accession number phs000424.v8.p2 on 05/01/2020.

DATA AND CODE AVAILABILITY

All data referred to in this manuscript is either provided in the main text/ supplementary material, or appropriately referenced (where derived from pre-existing, publicly accessible datasets).

REFERENCES

1. Sanders SJ, Murtha MT, Gupta AR, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature. 2012;485(7397):237–241. http://www.ncbi.nlm.nih.gov/RefSeq/; Accessed April 24, 2020. [PubMed: 22495306]
2. Banaschewski T, Becker K, Scherag S, Franke B, Coghill D. Molecular genetics of attention-deficit/hyperactivity disorder: An overview. Eur Child Adolesc Psychiatry. 2010;19:237–257. [PubMed: 20145962]
3. Sugiaman-Trapman D, Vitezic M, Jouhilahti E-M, et al. Characterization of the human RFX transcription factor family by regulatory and target gene analysis. BMC Genomics. 2018;19(1):181. doi:10.1186/s12864-018-4564-6 [PubMed: 29510665]

4. Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020;180(3):568–584. doi:10.1016/j.cell.2019.12.036 [PubMed: 31981491]
5. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434–443. doi:10.1038/s41586-020-2308-7 [PubMed: 32461654]
6. Doan RN, Lim ET, Rubeis S, et al. Recessive gene disruptions in autism spectrum disorder. Nat Genet. doi:10.1038/s41588-019-0433-8
7. Davis CA, Hitz BC, Sloan CA, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. 2018;46(D1):D794–D801. doi:10.1093/nar/gkx1081 [PubMed: 29126249]
8. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790–1797. doi:10.1101/gr.137323.112 [PubMed: 22955989]
9. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31(14):2382–2383. doi:10.1093/bioinformatics/btv145 [PubMed: 25765347]
10. Coe BP, Stessman HAF, Sulovari A, et al. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat Genet. 2019;51(1):106–116. doi:10.1038/s41588-018-0288-4 [PubMed: 30559488]
11. Velmeshev D, Schirmer L, Jung D, et al. Single-cell genomics identifies cell type-specific molecular changes in autism. Science. 2019;364(6441):685–689. doi:10.1126/science.aav8130 [PubMed: 31097668]
12. van Dam T, Wheway G, Slaats GG, Group S, Huynen MA, Giles RH. The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. Cilia. 2013:7. doi:10.1186/2046-2530-2-7
13. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–287. doi:10.1089/omi.2011.0118 [PubMed: 22455463]
14. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. Bioinformatics. 2011;27(7):1017–1018. doi:10.1093/bioinformatics/btr064 [PubMed: 21330290]
15. Fornes O, Castro-Mondragon JA, Khan A, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2020;48(D1):D87–D92. doi:10.1093/nar/gkz1001 [PubMed: 31701148]
16. Gao T, Qian J. EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. Nucleic Acids Res. 2020;48(D1):D58–D64. doi:10.1093/nar/gkz980 [PubMed: 31740966]
17. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. Hum Mutat. 2015;36(10):928–930. doi:10.1002/humu.22844 [PubMed: 26220891]
18. Harel T, Yesil G, Bayram Y, et al. Monoallelic and Biallelic Variants in EMC1 Identified in Individuals with Global Developmental Delay, Hypotonia, Scoliosis, and Cerebellar Atrophy. Am J Hum Genet. 2016;98(3):562–570. doi:10.1016/j.ajhg.2016.01.011 [PubMed: 26942288]
19. Sahoo T, Theisen A, Rosenfeld J a, et al. Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. Genet Med. 2011;13(10):868–880. doi:10.1097/GIM.0b013e3182217a06 [PubMed: 21792059]
20. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science. 2008;320(5875):539–543. doi:10.1126/science.1155174 [PubMed: 18369103]
21. Li J, Wang L, Guo H, et al. Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. Mol Psychiatry. 2017;22(9):1282–1290. doi:10.1038/mp.2017.140 [PubMed: 28831199]
22. Krumm N, Turner TN, Baker C, et al. Excess of rare, inherited truncating mutations in autism. Nat Genet. 2015;47(6):582–588. doi:10.1038/ng.3303 [PubMed: 25961944]

23. Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011;478(7370):483–489. doi:10.1038/nature10523 [PubMed: 22031440]
24. Polioudakis D, de la Torre-Ubieta L, Langerman J, et al. A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. Neuron. 2019;103(5):785–801. doi:10.1016/j.neuron.2019.06.011 [PubMed: 31303374]
25. Sun W, Poschmann J, Cruz-Herrera del Rosario R, et al. Histone Acetylome-wide Association Study of Autism Spectrum Disorder. Cell. 2016;167(5):1385–1397. doi:10.1016/j.cell.2016.10.031 [PubMed: 27863250]
26. Efimenko E, Bubb K, Mak HY, et al. Analysis of xbx genes in C. elegans. Development. 2005;132(8):1923–1934. doi:10.1242/dev.01775 [PubMed: 15790967]
27. Polyak A, Rosenfeld JA, Girirajan S. An assessment of sex bias in neurodevelopmental disorders. Genome Med. 2015;7(1). doi:10.1186/s13073-015-0216-5
28. Sahoo T, Theisen A, Rosenfeld JA, et al. 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. 2011;13(10):868–880. doi:10.1097/GIM.0b013e3182217a06
29. El Zein L, Ait-Lounis A, Morle L, et al. RFX3 governs growth and beating efficiency of motile cilia in mouse and controls the expression of genes involved in human ciliopathies. J Cell Sci. 2009;122(17):3180–3189. doi:10.1242/jcs.048348 [PubMed: 19671664]
30. Choksi SP, Lauter G, Swoboda P, Roy S. Switching on cilia: transcriptional networks regulating ciliogenesis. Development. 2014;141(7):1427–1441. doi:10.1242/dev.074666 [PubMed: 24644260]
31. Bonnafe E, Touka M, AitLounis A, et al. The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. Mol Cell Biol. 2004;24(10):4417–4427. doi:10.1128/MCB.24.10.4417-4427.2004 [PubMed: 15121860]
32. Benadiba C, Magnani D, Niquille M, et al. The ciliogenic transcription factor RFX3 regulates early midline distribution of guidepost neurons required for corpus callosum development. PLoS Genet. 2012;8(3):e1002606. doi:10.1371/journal.pgen.1002606
33. Baas D, Meiniel A, Benadiba C, et al. A deficiency in RFX3 causes hydrocephalus associated with abnormal differentiation of ependymal cells. Eur J Neurosci. 2006;24(4):1020–1030. doi:10.1111/j.1460-9568.2006.05002.x [PubMed: 16930429]
34. Sorensen SA, Bernard A, Menon V, et al. Correlated gene expression and target specificity demonstrate excitatory projection neuron diversity. Cereb Cortex. 2015;25(2):433–449. doi:10.1093/cercor/bht243 [PubMed: 24014670]
35. Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell. 2013;155(5):997. doi:10.1016/j.cell.2013.10.020 [PubMed: 24267886]
36. Parikshak NN, Luo R, Zhang A, et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell. 2013;155(5):1008. doi:10.1016/j.cell.2013.10.031 [PubMed: 24267887]
37. Durand B, Sperisen P, Emery P, et al. RFXAP, a novel subunit of the RFX DNA binding complex is mutated in MHC class II deficiency. EMBO J. 1997;16(5):1045–1055. doi:10.1093/emboj/16.5.1045 [PubMed: 9118943]
38. Reith W, Siegrist CA, Durand B, Barras E, Mach B. Function of major histocompatibility complex class II promoters requires cooperative binding between factors RFX and NF-Y. Proc Natl Acad Sci U S A. 1994;91(2):554–558. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC42987/. [PubMed: 8290561]
39. Piasecki BP, Burghoorn J, Swoboda P. Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals. Proc Natl Acad Sci. 2010;107(29):12969–12974. doi:10.1073/pnas.0914241107
40. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012;489(7416):391–399. doi:10.1038/nature11405 [PubMed: 22996553]

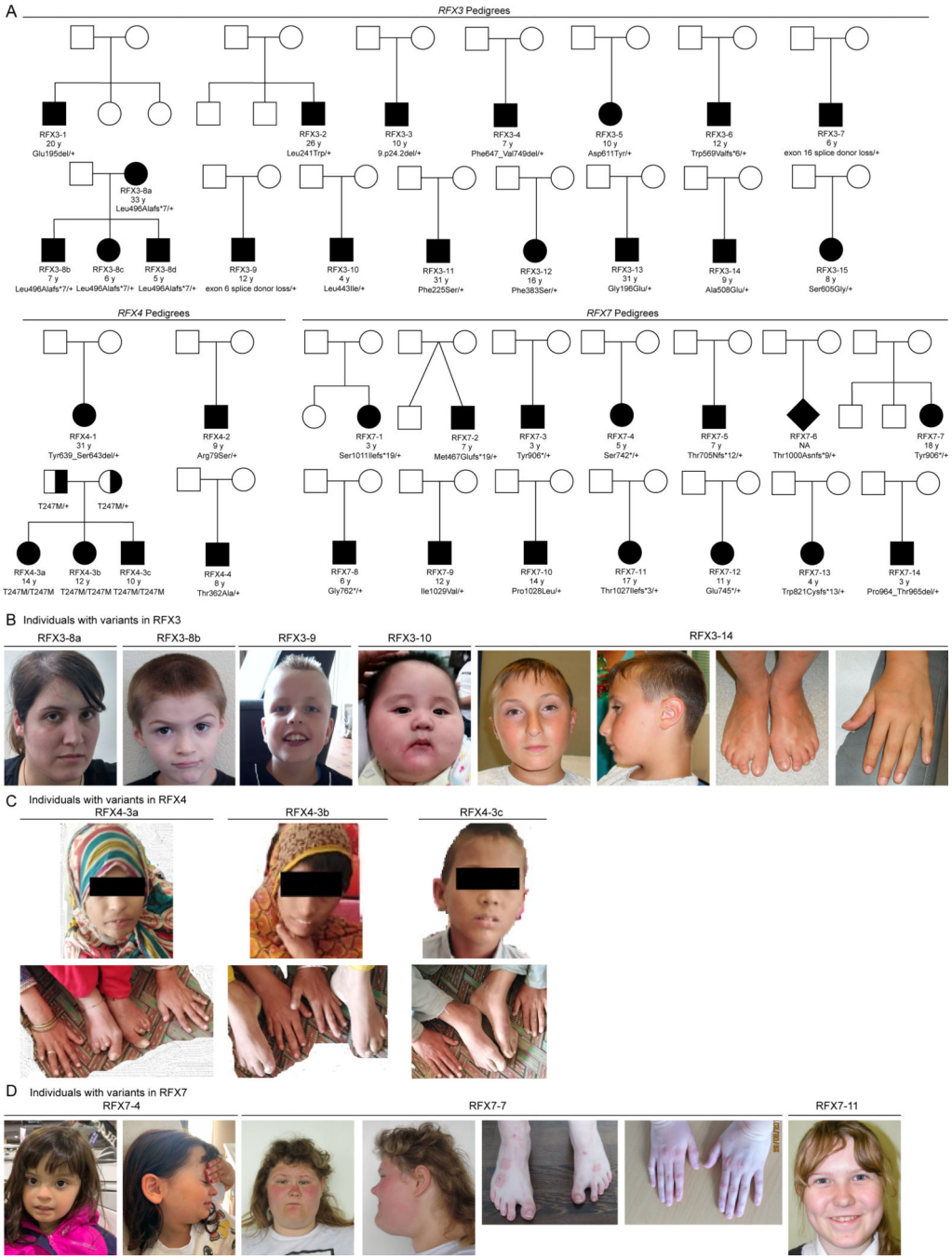

Figure 1. Pedigrees of reported individuals with *RFX3*, *RFX4*, and *RFX7* variants.

Pedigrees and clinical photographs of individuals with variants in *RFX3, RFX4*, and *RFX7*. (A) *RFX3, RFX4*, and *RFX7* case pedigrees. All pedigrees show de novo origin of variants except for RFX3-8a-d: a 33 year-old affected mother carrying the variant p.(Leu496Alafs*7) with transmission to three children, and pedigree RFX4-3a-c: three affected children homozygous for p.(Thr247Met).

(B) Individuals with *RFX3* variants.

(C) Individuals with *RFX4* variants.

(D) Individuals with *RFX7* variants.

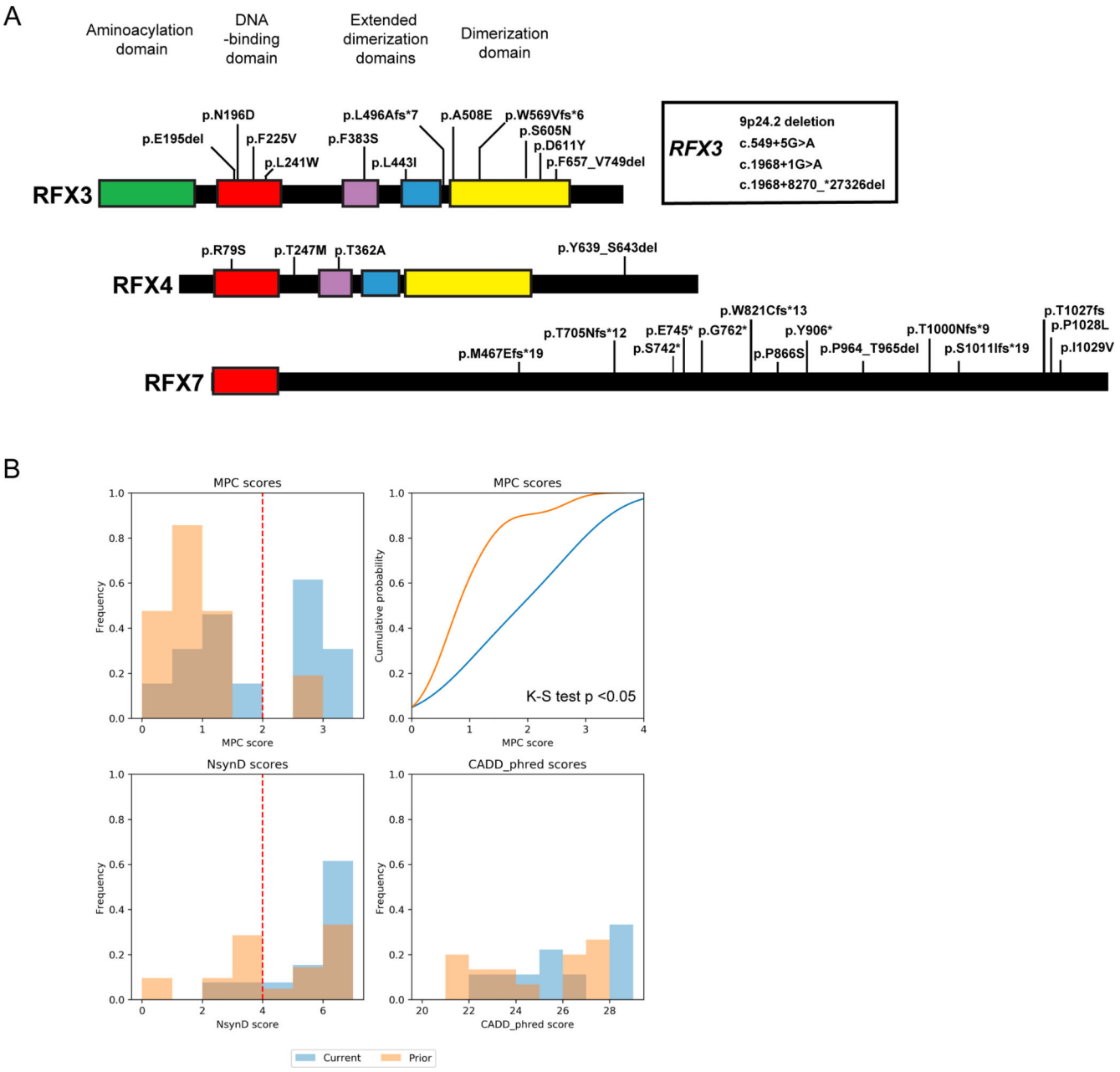

Figure 2. Distribution and predicted deleteriousness of RFX variants.

(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).

(C) Missense variant deleteriousness 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 <0.05 (p-value=0.015). MPC, Missense badness, PolyPhen-2, and Constraint. NsynD, Nonsynonymous Damaging score. CADD, Combined Annotation Dependent Depletion.

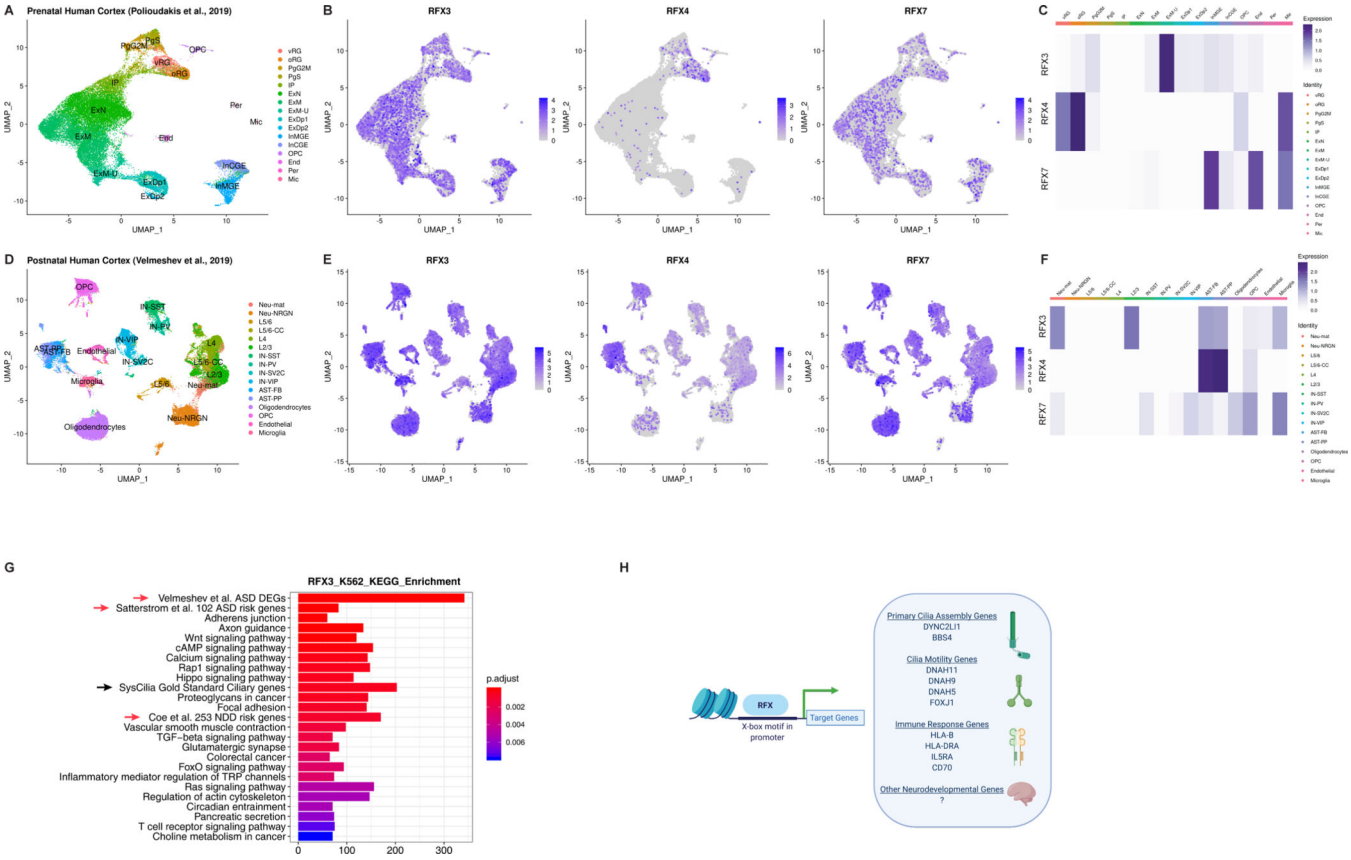

Figure 3. RFX3, RFX4, and RFX7 expression patterns in human cortex and haploinsufficiency gene dosage model.

(A) Transcriptomic cell types in the prenatal human cortex identified by single-cell RNA-sequencing.[24]

(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.[11]

(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, ciliary genes, 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 Piasecki, Durand, Reith, Sugiaman-Trapman.[3,37–39] 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, SVC2 expressing interneurons. IN-VIP, VIP interneurons. AST-FB, fibrous astrocytes. AST-PP, protoplasmic astrocytes. OPC, oligodendrocyte precursor cells.

Table 1.

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

Gene	Individual	Inheritance	gDNA (GRCh38)	cDNA	Protein	Category	Domain
RFX3	RFX3–1	*de novo*	chr9:g.3293222_3293224del	c.584_586del	p.(Glu195del)	Inframe deletion	DBD
	RFX3–2	*de novo*	chr9:g.3293086A>C	c.722T>G	p.(Leu241Trp)	Missense	DBD
	RFX3–3	*de novo*	chr9:g.9p24.2del	NA	(gene deletion)	Deletion	(all)
	RFX3–4	*de novo*	chr9:g.3197716_3239762del	c.1968+8270_*27326del	(exon 17 and exon 18 deleted) p.Phe647_Val749del	Deletion	DD
	RFX3–5	*de novo*	chr9:g.3248169C>A	c.1831 G>T	p.(Asp611Tyr)	Missense	DD
	RFX3–6	*de novo*	chr9:g.3257101dupG	c.1704dup	p.(Trp569Valfs*6)	Frameshift	DD
	RFX3–7	*de novo*	chr9:g.3248031C>T	c.1968+1 G>A	exon 16 splice donor loss	Splicing	DD
	RFX3–8a-d	inherited	chr9:g.3263053_3263054del	c.1486_1487del	p.(Leu496Alafs*7)	Frameshift	DD
	RFX3–9	*de novo*	chr9:g.3301541C>T	c.549+5G>A	exon 6 splice donor loss	Splicing	DBD
	RFX3–10	*de novo*	chr9:g.3270401G>T	c.1327C>A	p.(Leu443Ile)	Missense	EDD
	RFX3–11	*de novo*	chr9:g.3293134A>G	c.674T>C	p.(Phe225Ser)	Missense	DBD
	RFX3–12	*de novo*	chr9:g.3271057A>G	c.1148T>C	p.(Phe383Ser)	Missense	EDD
	RFX3–13	*de novo*	chr9:g.3293221C>T	c.587G>A	p.(Gly196Glu)	Missense	DBD
	RFX3–14	*de novo*	chr9:g.3263017G>T	c.1523C>A	p.(Ala508Glu)	Missense	DD
	RFX3–15	*de novo*	chr9:g.3256992T>C	c.1813A>G	p.(Ser605Gly)	Missense	DD
RFX4	RFX4–1	*de novo*	chr12:g.106750773_106750787del	c.1915_1929del	p.(Tyr639_Ser643del)	Inframe deletion	NA
	RFX4–2	*de novo*	chr12:g.106654271C>A	c.235C>A	p.(Arg79Ser)	Missense	DBD
	RFX4–3a-c	recessive (homozygous)	chr12:g.106696353C>T	c.740C>T	p.(Thr247Met)	Missense	NA
	RFX4–4	*de novo*	chr12:g.106715490A>G	c.1084A>G	p.(Thr362Ala)	Missense	DD
RFX7	RFX7–1	*de novo*	ch r15:g.56094696del	c.3032del	p.(Ser10111lefs*19)	Frameshift	NA
	RFX7–2	unknown (adopted)	chr15:g.56096328_56096329del	c.1399_1400del	p.(Met467Glufs*19)	Frameshift	NA
	RFX7–3	*de novo*	chr15:g.56095010G>T	c.2718C>A	p.(Tyr906*)	Stop Gain	NA
	RFX7–4	*de novo*	chr15:g.56095503G>C	c.2225C>G	p.(Ser742*)	Stop Gain	NA
	RFX7–5	*de novo*	chr15:g.56095615dupT	c.2113dup	p.(Thr705Asnfs*12)	Frameshift	NA
	RFX7–6	*de novo*	chr15:g.56094730dupT	c.2998dup	p.(Thr1000Asnfs*9)	Frameshift	NA
	RFX7–7	*de novo*	chr15:g.56095010G>C	c.2718C>G	p.(Tyr906*)	Stop Gain	NA
	RFX7–8	*de novo*	chr15:g.56095444C>A	c.2284G>T	p.(Gly762*)	Stop Gain	NA

Gene	Individual	Inheritance	gDNA (GRCh38)	cDNA	Protein	Category	Domain
	RFX7–9	*de novo*	chr15:g.56094643T>C	c.3085A>G	p.(Ile1029Val)	Missense	NA
	RFX7–10	*de novo*	chr15:g.56094645G>A	c.3083C>T	p.(Pro1028Leu)	Missense	NA
	RFX7–11	*de novo*	ch r15:g.56094648del	c.3080del	p.(Thr1027Ilefs*3)	Frameshift	NA
	RFX7–12	*de novo*	chr15:g.56095495C>A	c.2233G>T	p.(Glu745*)	Stop Gain	NA
	RFX7–13	*de novo*	chr15:g.56095266_56095269dup	c.2459_2462dup	p.(Trp821Cysfs*13)	Frameshift	NA
	RFX7–14	*de novo*	*chr15:g. 56094864_56094869del*	*c.2859_2864del*	*p.(Pro964_Thr965del)*	*Inframe deletion*	*NA*

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.[3] Italics indicates individual has a variant of uncertain significance.

Genet Med. Author manuscript; available in PMC 2022 September 14.

Table 2.

nical features of individuals with variants in *RFX3*, *RFX4*, or *RFX7*

dividual	Age	Sex	Presentation	Variant	Languag e delay	Motor delay	ASD*	ID	ADHD	Behavioral Profile	Sleep issues	Seizures	Hypotonia; Other Neurologic Findings	Dysmorphism	Micro or macrocephaly	Neuroimaging Findings	Other Medical or Neuropsychiatric Features
FX3																	
FX3–1	20 years	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
FX3–2	26 years	M	ASD, ID, ADHD	missense	yes	yes	yes	yes	yes	sensory seeking behavior, aggression, biting, pica, self-injury, sensory	yes	no	yes	yes	no	normal/ nonspecific findings	hypogonadism, strabismus, bipolar, behavioral and mild cognitive
FX3–3	10 years	M	ASD, ADHD	deletion	no	yes	yes	no	yes	sensory hypersensitivity	yes	yes	NA	yes	macrocephaly	NA	strabismus
FX3–4	7 years	M	ADHD, anxiety	deletion	yes	yes	yes	yes	yes	NA	NA	no	NA	yes	mild macrocephaly	normal/ nonspecific findings	myopia
FX3–5	10 years	F	GDD	missense	no	yes	yes	no	yes	oppositional behavior, aggression	NA	no	NA	yes	mild macrocephaly	NA	NA
FX3–6	12 years	M	ASD, GDD, ADHD	frameshift	no	no	yes	no	yes	anxiety, aggression, emotional dysregulation	NA	no	yes	no	macrocephaly	PVL	anxiety, mild cognitive and behavioral decline at puberty
FX3–7	6 years	M	ASD, ADHD	splicin g	yes	no	no	yes	NA	aggression, sensory-seeking behavior, elopement, and impulsivity	yes	no	yes	no	no	partially empty sella	myopia
FX3–8a	33 years	F	ASD, ID	frameshift	yes	NA	yes	yes	NA	NA	NA	NA	NA	yes	NA	NA	NA
FX3–8b	7 years	M	ASD, ID, ADHD	frameshift	yes	NA	yes	yes	yes	aggression, biting	NA	no	NA	no	NA	NA	NA
FX3–8c	6 years	F	GDD	frameshift	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

dividual	Age	Sex	Presentation	Variant	Languag e delay	Motor delay	ASD*	ID	ADHD	Behavioral Profile	Sleep issues	Seizures	Hypotonia; Other Neurologic Findings	Dysmorphism	Micro or macrocephaly	Neuroimaging Findings	Other Medical or Neuropsychiatric Features
FX3–8d	5 years	M	GDD	frameshift	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
FX3–9	12 years	M	ASD, ID	splicing	yes	NA	yes	yes	no	sensory hypersensitivity	yes	no	yes	yes	mild macrocephaly	normal/ nonspecific findings	NA
FX3–10	4 years	M	GDD	missense	NA	yes	NA	NA	NA	NA	NA	yes	yes	yes	microcephaly	thin corpus callosum	NA
FX3–11	31 years	M	ASD	missense	NA	NA	yes	NA	yes	NA	yes	NA	NA	NA	NA	NA	Crohn's disease, anxiety, depression
FX3–12	16 years	F	ID, ASD	missense	yes	no	yes	yes	yes	impulsivity, mood swings	no	no	NA	yes	macrocephaly	normal	hallucinations, mania, behavioral decline
FX3–13	31 years	M	ASD	missense	no	no	yes	yes	no	aggression, elopement	yes	yes	no	no	NA	uncal asymmety	cognitive and behavioral decline, hallucinations
FX3–14	9 years	M	GDD	missense	yes	yes	no	yes	no	no	no	no	no	yes	NA	NA	bilateral conductive hearing loss
FX3–15	8 years	F	ASD, ID, ADHD	missense	yes	yes	yes	yes	yes	aggression, anxiety, impulsivity	yes	no	no	yes	mild macrocephaly	NA	asthma
FX4																	
FX4–1	31 years	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
FX4–2	9 years	M	ID ASD, epilepsy	missense	NA	NA	NA	NA	NA	NA	NA	yes	NA	NA	NA	NA	NA
FX4–3a	14 years	F	ID, ASD, behavior problems	missense	yes	yes	yes	yes	NA	yes, behavior challenges	NA	no	NA	no	no	NA	skeletal abnormalities
FX4–3b	12 years	F	ID, ASD, behavior problems	missense	yes	yes	yes	yes	NA	yes, behavior challenges	NA	no	NA	no	no	NA	skeletal abnormalities
FX4–3c	10 years	M	ID, ASD, behavior problems	missense	yes	yes	yes	yes	NA	yes, behavior challenges	NA	no	NA	no	no	NA	none

Genet Med. Author manuscript; available in PMC 2022 September 14.

dividual	Age	Sex	Presentation	Variant	Languag e delay	Motor delay	ASD*	ID	ADHD	Behavioral Profile	Sleep issues	Seizures	Hypotonia; Other Neurologic Findings	Dysmorphism	Micro or macrocephaly	Neuroimaging Findings	Other Medical or Neuropsychiatric Features
FX 4–4	8 years	M	GDD, ASD, behavior problems	missense	yes	yes	yes	yes	NA	sensory hypersensitivity, impulsivity, anxiety, mood swings	NA	no	NA	cleft lip and palate	microcephaly	absent pituitary	hypopituitarism
FX7																	
FX7–1	3 years	F	ID	frameshift	yes	yes	no	yes	no	NA	no	no	no	yes	microcephaly	normal	NA
FX7–2	7 years	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
FX7–3	3 years	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
FX7–4	5 years	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
FX7–5	7 years	M	ASD, GDD	frameshift	yes	yes	yes	yes	no	excitable, laughs easily, sensory hypersensitvty, aggressive when younger	yes	no	yes	no	macrocephaly	normal	none
FX7–6	NA	M	ID, ASD	frameshift	yes	yes	yes	yes	yes	mood swings, anxiety, aggression, selfinjury	yes	no	no	yes	NA	cerebellar tonsillar herniation, abnormal 4th ventricle	polyphagia
FX7–7	18 years	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

dividual	Age	Sex	Presentation	Variant	Languag e delay	Motor delay	ASD*	ID	ADHD	Behavioral Profile	Sleep issues	Seizures	Hypotonia; Other Neurologic Findings	Dysmorphism	Micro or macrocephaly	Neuroimaging Findings	Other Medical or Neuropsychiatric Features
FX7–8	6 years	M	ID	stop-gain	yes	yes	no	yes	NA	sensory hypersensitivity	no	no	yes	yes	macrocephaly	NA	neuroblastoma
FX7–9	12 years	M	GDD	missense	yes	yes	no	yes	yes	NA	no	febrile seizure x 1	hyperreflexia dystonic	no	microcephaly	normal	hypocalcemia optic nerve hypoplasia, cataract,
FX7–10	14 years	M	ID	missense	yes	yes	no	yes	no	no	yes	epilepsy with myoclonic seizures	movements, mixed hypo/ hypertonia	yes	microcephaly	abnormalty of the basal ganglia, delayed CNS myelination	mild hearing loss, laryngomalacia, recurrent bronchitis/ pneumonia, cryptorchidism
FX7–11	17 years	F	ID	frameshift	yes	no	no	yes	no	short attention span, excessive fears/phobais, fixated interests, sensory hypersensitivity, poor social interactions	no	yes; absence during day and grand mal seizures in sleep but none since 10 years	no	long fingers	no	normal	none
FX7–12	11 years	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
FX7–13	4 years	F	ID	frameshift	yes	yes	no	yes	no	stereotypies	yes	no	no	yes	no	subcortical hypersignal in the left temporal pole (cortical dysplasia or developmental venous anomaly)	ventricular septal defect, mild unilateral hearing loss
FX7–14	*3 years*	*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*

SD = 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, ording 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.