Mutations in *EBF3* Disturb Transcriptional Profiles and Cause Intellectual Disability, Ataxia, and Facial Dysmorphism

Frederike Leonie Harms,[1,23] Katta M. Girisha,[2,23] Andrew A. Hardigan,[3,4,23] Fanny Kortüm,[1] Anju Shukla,[2] Malik Alawi,[5,6,7] Ashwin Dalal,[8] Lauren Brady,[9] Mark Tarnopolsky,[9] Lynne M. Bird,[10,11] Sophia Ceulemans,[11] Martina Bebin,[12] Kevin M. Bowling,[3] Susan M. Hiatt,[3] Edward J. Lose,[13] Michelle Primiano,[14] Wendy K. Chung,[14] Jane Juusola,[15] Zeynep C. Akdemir,[16] Matthew Bainbridge,[17] Wu-Lin Charng,[16] Margaret Drummond-Borg,[18] Mohammad K. Eldomery,[16] Ayman W. El-Hattab,[19] Mohammed A.M. Saleh,[20] Stéphane Bézieau,[21] Benjamin Cogné,[21] Bertrand Isidor,[21,22] Sébastien Küry,[21] James R. Lupski,[16] Richard M. Myers,[3] Gregory M. Cooper,[3,*] and Kerstin Kutsche[1,*]

From a GeneMatcher-enabled international collaboration, we identified ten individuals affected by intellectual disability, speech delay, ataxia, and facial dysmorphism and carrying a deleterious *EBF3* variant detected by whole-exome sequencing. One 9-bp duplication and one splice-site, five missense, and two nonsense variants in *EBF3* were found; the mutations occurred de novo in eight individuals, and the missense variant c.625C>T (p.Arg209Trp) was inherited by two affected siblings from their healthy mother, who is mosaic. EBF3 belongs to the early B cell factor family (also known as Olf, COE, or O/E) and is a transcription factor involved in neuronal differentiation and maturation. Structural assessment predicted that the five amino acid substitutions have damaging effects on DNA binding of EBF3. Transient expression of EBF3 mutant proteins in HEK293T cells revealed mislocalization of all but one mutant in the cytoplasm, as well as nuclear localization. By transactivation assays, all EBF3 mutants showed significantly reduced or no ability to activate transcription of the reporter gene *CDKN1A*, and in situ subcellular fractionation experiments demonstrated that EBF3 mutant proteins were less tightly associated with chromatin. Finally, in RNA-seq and ChIP-seq experiments, EBF3 acted as a transcriptional regulator, and mutant EBF3 had reduced genome-wide DNA binding and gene-regulatory activity. Our findings demonstrate that variants disrupting *EBF3*-mediated transcriptional regulation cause intellectual disability and developmental delay and are present in ~0.1% of individuals with unexplained neurodevelopmental disorders.

Intellectual disability (ID) is a common phenotype with extreme clinical and genetic heterogeneity. Widespread application of whole-genome and whole-exome sequencing (WES) has greatly increased the identification of genetic causes of non-syndromic and syndromic forms of ID.[1,2] WES together with the freely accessible tool GeneMatcher, which brings together clinicians and researchers with an interest in the same gene, can improve the identification of genes in which mutations cause disease.[3]

We investigated a family with three healthy and two affected children, who both presented with global developmental delay, febrile seizures, and gait instability with frequent falls. Both probands and one healthy sibling were subjected to WES as described previously.[4,5] We initially hypothesized a Mendelian recessive trait. By employing our internal pipeline to prioritize potentially disease-causing mutations,[4,5] we did not identify any rare, potentially pathogenic biallelic variants in the affected siblings (data not shown). WES data were then analyzed for heterozygous variants (non-synonymous, frameshift, and intronic variants at exon-intron boundaries ranging from −10 to +10) absent in dbSNP138, the 1000 Genomes Browser, the NHLBI Exome Sequencing Project Exome Variant Server (EVS), and the Exome Aggregation Consortium (ExAC) Browser,[6] shared by both affected subjects, and absent in the healthy sibling. Exonic variants and intronic alterations at exon-intron boundaries ranging from −10 to +10, which were clinically associated and

[1]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [2]Department of Medical Genetics, Kasturba Medical College, Manipal University, 576104 Manipal, India; [3]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA; [4]Department of Genetics, University of Alabama at Birmingham, Birmingham, AL 35294, USA; [5]Bioinformatics Service Facility, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [6]Center for Bioinformatics, University of Hamburg, 20246 Hamburg, Germany; [7]Virus Genomics, Heinrich Pette Institute, Leibniz Institute for Experimental Virology, 20246 Hamburg, Germany; [8]Diagnostics Division, Centre for DNA Fingerprinting and Diagnostics, 500001 Hyderabad, Telangana, India; [9]Department of Pediatrics, McMaster University Medical Center, Hamilton, ON L8N 3Z5, Canada; [10]Department of Pediatrics, University of California, San Diego, San Diego, CA 92123, USA; [11]Division of Genetics/Dysmorphology, Rady Children's Hospital San Diego, San Diego, CA 92123, USA; [12]Department of Neurology, University of Alabama at Birmingham, Birmingham, AL35294, USA; [13]Department of Genetics, University of Alabama at Birmingham, Birmingham, AL 35294, USA; [14]Department of Pediatrics and Medicine, Columbia University, New York, NY 10032, USA; [15]GeneDx, Gaithersburg, MD 20877, USA; [16]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; [17]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA; [18]Cook Children's Genetic Clinic, Fort Worth, TX 76102, USA; [19]Division of Clinical Genetics and Metabolic Disorders, Department of Pediatrics, Tawam Hospital, 15258 Al-Ain, United Arab Emirates; [20]Section of Medical Genetics, Children's Hospital, King Fahad Medical City, 11564 Riyadh, Saudi Arabia; [21]Service de Génétique Médicale, Centre Hospitalier Universitaire Nantes, 44093 Nantes Cedex 1, France; [22]INSERM UMR-S 957, 44035 Nantes, France
[23]These authors contributed equally to this work
*Correspondence: gcooper@hudsonalpha.org (G.M.C.), kkutsche@uke.de (K.K.)
http://dx.doi.org/10.1016/j.ajhg.2016.11.012.

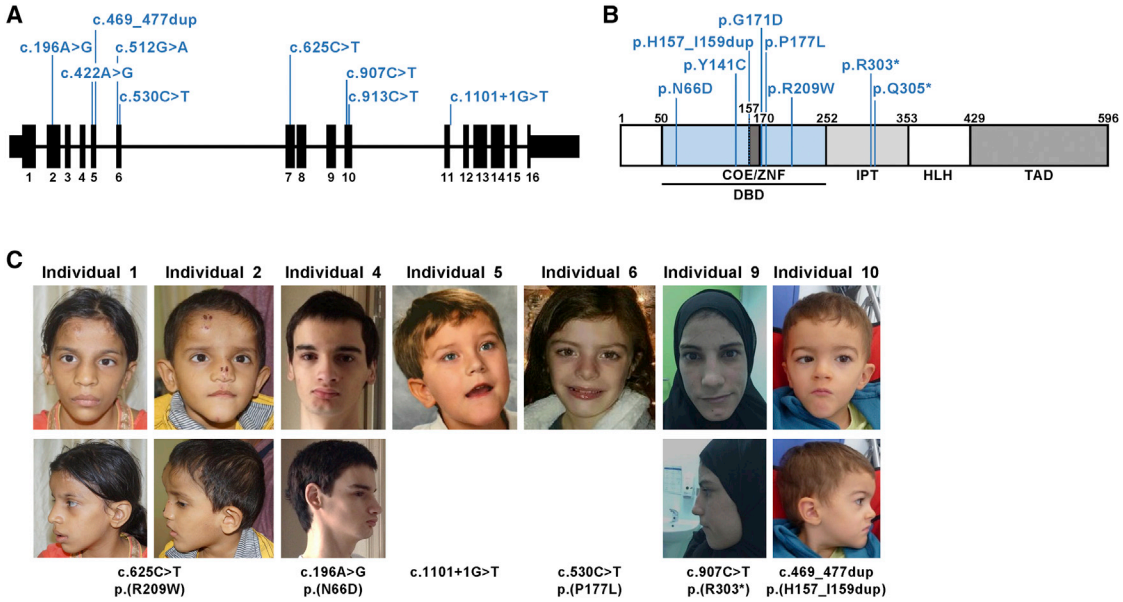

Figure 1. *EBF3* Mutations Identified in Ten Individuals with ID

(A) Schematic representation of the exon-intron structure of *EBF3*. Black bars represent exons, and black lines represent introns. Mutations identified in the ID-affected individuals are indicated above the exon-intron structure.

(B) Domain structure of EBF3, including the positions of the identified amino acid alterations. Amino acid numbers are given. Abbreviations are as follows: DBD, DNA-binding domain with an atypical zinc finger (ZNF; COE motif); IPT, Ig-like/plexins/transcription factors; HLH, helix-loop-helix motif; and TAD, transactivation domain.

(C) Photographs of seven individuals show subtle yet distinct facial dysmorphism. All show a long face, tall forehead, high nasal bridge, deep philtrum, straight eyebrows, strabismus, short and broad chin, and mildly dysmorphic ears. Consent for the publication of photographs was obtained for the seven subjects.

unknown in public databases, were retained. This analysis identified 16 variants (Table S1). We used objective metrics from the ExAC Browser to prioritize genes intolerant to functional variation (pLI $\geq$ 0.9 and high *Z* scores);[6] we identified five genes with strong selection against various classes of variants for segregation analysis in the family (Table S1). Four variants were inherited from a healthy parent and/or were present in two healthy siblings (Table S2). The missense variant c.625C>T (p.Arg209Trp) in *EBF3* (MIM: 607407; GenBank: NM_001005463.2) was confirmed in both affected siblings and was absent in the father and all healthy siblings (Figure S1 and Table S2). In leukocyte-derived DNA from the mother, the Sanger sequence profile showed a very low signal for the mutated base (thymine) superimposed on the wild-type sequence (cytosine), suggesting that she had mosaicism for the *EBF3* variant (Figure S1).[7] By cloning the mutation-bearing *EBF3* amplicon and then Sanger sequencing the colony PCR products, we confirmed the mother to be a mosaic carrier (18% and 4% of leukocytes and buccal cells, respectively, were heterozygous for the *EBF3* variant; Figure S1); parenthetically, maternal mosaics are at greater recurrence risk.[7] Given that *EBF3* is relatively intolerant to functional genetic variation (Table S1)[8,9] and the variant p.Arg209Trp was computationally predicted to be deleterious (Table S3), we next submitted *EBF3* to GeneMatcher and were matched with eight other research groups.

In addition to the above family, eight unrelated affected individuals with variants in *EBF3* were identified through WES[10–14] by groups that independently submitted to GeneMatcher. In addition to c.625C>T (p.Arg209Trp), we found missense variants c.196A>G (p.Asn66Asp) in subject 4, c.422A>G (p.Tyr141Cys) in subject 7, c.512G>A (p.Gly171Asp) in subject 8, and c.530C>T (p.Pro177Leu) in subject 6; 9-bp duplication c.469_477dup (p.His157_Ile159dup) in subject 10; nonsense variants c.907C>T (p.Arg303*) in subject 9 and c.913C>T (p.Gln305*) in subject 3; and splice-site mutation c.1101+1G>T in subject 5 (Figures 1A and 1B and Table S3). All variants were computationally predicted to affect protein function (Table S3) and were absent in the 1000 Genomes Browser, EVS, and ExAC Browser. All eight additional variants were confirmed to have arisen de novo (Table 1). The in-frame duplication and the five amino acid substitutions affect highly conserved residues and are invariant among *EBF* family paralogs (Figure S2). We note that the samples from all individuals in this study were collected, sequenced, and analyzed with the participants' consent and under the appropriate supervisory committees to protect human subjects at each research site contributing to this study.

Clinical features consistent among all individuals with a *EBF3* mutation (10/10) were ID, speech delay, and motor developmental delay. Ataxia was reported in 6/8, and seizures were reported in 2/9. Brain imaging revealed cerebellar vermian hypoplasia in 2/8 (Table 1). Facial dysmorphism was mild, and commonly seen features included a long face, tall forehead, high nasal bridge, deep philtrum,

Table 1. Clinical Features of Subjects with Mutations in *EBF3*

	Family 1			Family 2	Family 3	Family 4	Family 5	Family 6	Family 7	Family 8	Family 9
	Subject 1	Subject 2	Mother	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	Subject 9	Subject 10
Mutation in *EBF3*[a]	c.625C>T (p.Arg209Trp)	c.625C>T (p.Arg209Trp)	c.[625C=/>T]	c.913C>T (p.Gln305*)	c.196A>G (p.Asn66Asp)	c.1101+1G>T	c.530C>T (p.Pro177Leu)	c.422A>G (p.Tyr141Cys)	c.512G>A (p.Gly171Asp)	c.907C>T (p.Arg303*)	c.469_477dup (p.His157_Ile159dup)
Origin of mutation	inherited	inherited	de novo (mosaicism)	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo
Sex	female	male	female	male	male	male	female	male	female	female	male
Age at assessment	9 years, 3 months	3 years, 4 months	33 years	5 years, 9 months	16 years, 6 months	4 years, 6 months	2 years, 7 months	1 year, 11 months	13 years	25 years	3 years, 5 months
Birth weight in grams (SD)	2,500 (−2)	1,780 (−3.5)	NA	2,495 (−2)	3,500 (−0.2)	4,054 (+0.7)	3,033 (−0.7)	2,891 (−1)	3,068 (−0.6)	2,900 (−1)	3,190 (−0.7)
Birth length in cm (SD)	NA	NA	NA	45.7 (−2)	50.8 (+0.3)	NA	NA	48.2 (−0.7)	49.5 (0)	NA	49 (−0.25)
OFC at birth in cm (SD)	NA	NA	NA	33.5 (−1.3)	NA	NA	NA	35.5 (−0.25)	NA	NA	36 (+1)
Weight in kg (SD)	22 (−2)	11 (−2)	NA	21 (+0.5)	45.8 (−2.2)	25.67 (+3.4)	13.3 (0)	10.6 (−1.5)	42.2 (−1)	55 (−0.5)	12.4 (−1)
Length in cm (SD)	119 (−3)	91 (−2)	NA	108 (−1.2)	168.6 (−0.8)	110.49 (+1)	90.2 (−0.2)	NA	NA	167 (+0.8)	95 (−1)
OFC in cm (SD)	50 (−3)	48.3 (−2.5)	NA	51.5 (−1)	57 (+0.7)	NA	49.75 (+1)	49.5 (+0.7)	51.9 (−1.7)	NA	51 (+0.5)
Neurological Abnormalities											
ID	+	+	−	+	+	+	+	+	+	+	+
Motor developmental delay	+	+	−	+	+	+	+	+	+	+	+
Speech delay	+	+	−	+	+	+	+	+	+	+	+
Ataxia	+	+	−	+	+	NA	+	−	wide-based bent knee, dystonic gait	−	NA
Seizures	+	+	−	−	−	−	−	−	NA	−	−
Tone	normal	normal	normal	mild hypotonia in early childhood, normal now	normal	truncal hypotonia	truncal hypotonia	hypotonia	hypotonia as infant, dystonia now	normal	truncal hypotonia
Brain MRI	normal	ND	ND	Normal	CVH	normal	normal	normal	normal	ND	CVH
Craniofacial Abnormalities											
Long face	+	+	+	−	+	NA	NA	−	+	+	−
Deep philtrum	+	+	+	−	+	NA	+	−	+	−	−

(*Continued on next page*)

Table 1. ***Continued***

	Family 1			Family 2	Family 3	Family 4	Family 5	Family 6	Family 7	Family 8	Family 9
	Subject 1	**Subject 2**	**Mother**	**Subject 3**	**Subject 4**	**Subject 5**	**Subject 6**	**Subject 7**	**Subject 8**	**Subject 9**	**Subject 10**
Tall forehead	+	+	+	−	+	NA	+	−	+	+	−
High nasal bridge	+	+	+	−	+	NA	NA	−	+	+	−
Straight eyebrows	+	+	+	−	+	NA	NA	−	−	−	+
Strabismus	+	+	−	+	+	+	+	+	+	−	+
Ears	low set, posteriorly rotated, small ear lobes	low set, posteriorly rotated	low set, posteriorly rotated	−	small ear lobes	NA	NA	−	normal	NA	low set
Short and broad chin	+	+	+	−	+	NA	NA	−	prominent chin	broad chin	−
Other	thick vermillion of upper and lower lips	−	−	small mouth, short philtrum, micrognathia	thick vermillion of upper and lower lips	flat nasal bridge	relative macrocephaly, upward-slanting palpebral fissures, long eyelashes, straight hair	hypertelorism	facial asymmetry, submucous cleft palate	upward-slanting palpebral fissures, broad nasal bridge, smooth philtrum, thin upper lip, chin dimple	thin lower lip
Other Findings											
Additional clinical anomalies	−	−	−	syndactyly of the second and third toes, inguinal hernia (repaired)	dysarthria, pectus excavatum, thin and scooped nails, orchiopexy for undescended testicles, very mild hypospadias, strabismus surgery, intoeing due to femoral anteversion, attention deficit disorder, gastroesophageal reflux, possible eosinophilic esophagitis	−	−	congenital heart disease (atrial septal defect)	complete (right) and partial (left) syndactyly of the second and third toes, limited facial expression, neurogenic bladder (vesicostomy), feeding difficulties, swallowing of semisolid food only, mild scoliosis, dysplastic right kidney, bilateral vesicoureteric reflux, megacolon with intractable constipation, severe hip and knee contractures	−	bilateral talipes equinovarus, phimosis, recurrent lower urinary-tract infections, constipation

Abbreviations are as follows: +, present; −, absent; CVH, cerebellar vermian hypoplasia; ID, intellectual disability; MRI, magnetic resonance imaging; NA, not available; ND, not determined; OFC, occipital frontal circumference; and SD: standard deviation.
[a]According to mRNA reference sequence GenBank: NM_001005463.2.

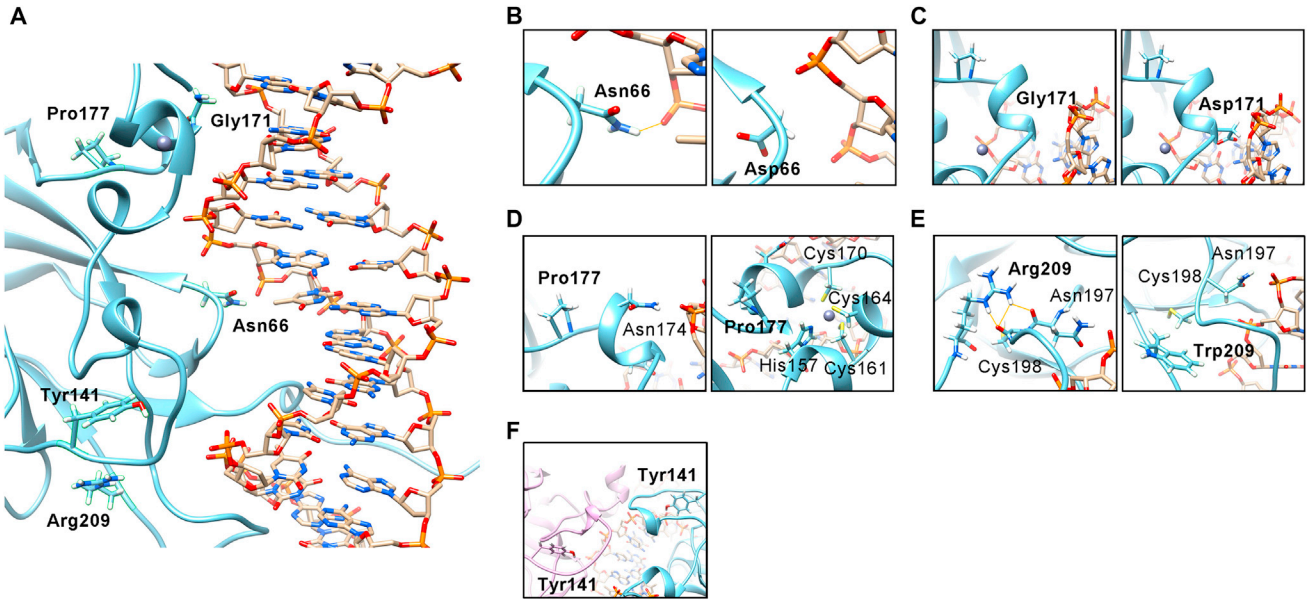

Figure 2. Structural Impact of *EBF3* Missense Mutations
(A) Model of the DBD of an EBF3 monomer (cyan ribbon; affected residues are shown as sticks) bound to DNA (sticks). Hydrogen bonds are represented by yellow lines. Major interactions of affected residues are shown in (B)–(F).
(B) Asn66 forms a hydrogen bond with a DNA phosphate group (left), which is disrupted by the substitution (right). In addition, the negative charge of aspartate at position 66 most likely leads to electrostatic repulsion of the phosphate group.
(C) Gly171 is part of the protein-DNA interface (left). Substitution of Gly171 with the negatively charged asparagine could lead to an electrostatic repulsion of the DNA backbone (right).
(D) Pro177 is closely localized to Asn174, which forms a hydrogen bond with the DNA (left), and to the zinc finger (right). Replacement of Pro177 with leucine could lead to a conformational change altering the position of Asn174 and possibly of the zinc finger, reducing the DNA-binding capacity of EBF3. His157 and Cys161, Cys164, and Cys170 are invariant residues coordinating Zn^{2+} (right). In-frame duplication of the amino acids His157, Glu158, and Ile159 could cause a conformational change of the zinc finger, reducing the DNA-binding capability of EBF3.
(E) Arg209 forms hydrogen bonds with the backbone of Cys198 and Asn197, the latter of which forms a hydrogen bond with the DNA (left). Substitution at Arg209 leads to disruption of these hydrogen bonds, probably affecting the positioning of Asn197 (right).
(F) Tyr141 is localized within a loop that is not directly involved in DNA binding but rather in EBF3 dimer formation (one EBF3 monomer is indicated by a pink ribbon, and the other is marked by a blue ribbon). Alteration of Tyr141 could lead to a conformational change at the dimer interface, probably resulting in reduced stability of the EBF3 dimer and interfering with its ability to bind to DNA.

straight eyebrows, strabismus, and short and broad chin (Figure 1C).

EBF3 encodes early B cell factor 3, which is one of four members of the EBF transcription factor family (also known as Olf, COE, or O/E). All EBFs consist of an N-terminal DNA-binding domain (DBD), an Ig-like/plexins/transcription factors (IPT) domain with yet unknown function, a helix-loop-helix (HLH) domain, which is critical for homo- and heterodimer formation, and a C-terminal transactivation domain (TAD) (Figure 1B).[15] EBF1 has been discovered as a key factor for B cell differentiation[16] and olfactory nerve signaling.[17] However, expression of *ebf1*, *ebf2*, and *ebf3* in early post-mitotic neurons during embryogenesis suggests a role in neuronal differentiation and maturation.[18] Ebf3 acts downstream of the proneural transcription factor neuroD in late neural differentiation in *Xenopus*[19] and is transcriptionally repressed by ARX,[20] variants in which cause a spectrum of developmental disorders ranging from ID to brain-malformation syndromes.[21] Silencing, genomic deletion, and somatic point mutations in *EBF3* exist in diverse types of cancer,[22] and EBF3-mediated induction of cell-cycle arrest and apoptosis suggests that EBF3 acts as a tumor suppressor by regulating the expression of specific target genes and controlling a potential anti-neoplastic pathway.[15] All germline *EBF3* missense variants affect the DBD (Figure 1B and Figure S2). Interaction between the DBD and DNA is dependent on a zinc-coordination motif, the zinc knuckle (COE motif), located between His157 and Cys170 (Figure 1B and Figure S2). Thus, the p.His157_Ile159dup variant most likely affects DNA binding, as has been shown for the EBF3 p.His157Asn mutant.[23,24]

We explored the structural impact of the five missense mutations by using a homology model for the DNA-bound configuration of the DBD of EBF3 with the DNA duplex containing the EBF1 consensus sequence (Figure 2A).[25] The three-dimensional structure of the DBD of wild-type and mutant EBF3 (amino acids 50–251) was obtained by means of homology modeling with the web-based service SWISS-MODEL.[26] The crystallographic structure of EBF1 bound to DNA (PDB: 3MLP) at 2.8-Å resolution was used as a template.[25] Molecular graphics were developed with UCSF Chimera.[27] Two of the five alterations, p.Asn66Asp and p.Gly171Asp, were predicted to directly affect DNA

binding (Figures 2B and 2C). Pro177 is closely localized to the zinc finger (His157, Cys161, Cys164, and Cys170) and to Asn174, which forms a hydrogen bond with DNA. Replacement of Pro177 with leucine causes a conformational change, probably affecting the correct positioning of the zinc knuckle and destabilizing the protein-DNA complex (Figure 2D). Arg209 does not directly interact with DNA but rather forms hydrogen bonds with the backbone of Cys198 and Asn197, the latter of which directly interacts with DNA. The p.Arg209Trp change is expected to alter the positioning of Asn197 and binding affinity for DNA (Figure 2E). Tyr141 is localized within a loop probably involved in EBF3 dimer formation. Alteration of this residue could lead to a conformational change at the dimer interface, resulting in reduced stability of the EBF3 dimer and interfering with its ability to interact with DNA (Figure 2F).

To analyze the functional effect of the *EBF3* mutations, we transiently and efficiently expressed FLAG-tagged wild-type and mutant proteins in HEK293T cells. By immunocytochemistry followed by epifluorescence microscopy, we confirmed exclusive nuclear localization of wild-type EBF3 ($EBF3^{WT}$) (Figure 3A). In contrast, most mutant proteins showed both nuclear and cytoplasmic distribution (Figure 3A). We next performed a reporter gene assay (Luciferase Reporter Assay System, Promega) to assess whether EBF3 mutants still mediate transcriptional activation of the target gene *CDKN1A* (p21).[23] First, we analyzed dose-dependent activation of $EBF3^{WT}$ on the pGL2-p21 promoter-Luc construct and identified a maximum activation with 6–8 μg DNA of $EBF3^{WT}$ construct (Figure S3). In contrast to p53, for which *CDKN1A* is a prototypical target gene, and $EBF3^{WT}$, EBF3 mutants had significantly reduced or no ability to activate transcription of the reporter gene (Figure 3B). Activation of the *CDKN1A* promoter was not significantly reduced when $EBF3^{WT}$ was co-expressed with mutant p.Gly171Asp, p.Pro177Leu, or p.Arg209Trp (Figure S4). In contrast, co-expression of $EBF3^{WT}$ and each of the mutants p.Asn66Asp, p.Tyr141Cys, p.His157_Ile159dup, p.Arg303*, and p.Gln305* caused a significant reduction in reporter activity by 40%–50% (Figure S4), suggesting a potential dominant-negative impact of these variants on the wild-type allele. The dominant-negative effect was also observed when increasing amounts of the EBF3 mutant p.Asn66Asp or p.Gln305* were expressed in the presence of a fixed amount of $EBF3^{WT}$, demonstrating dose-dependent reduction in reporter gene activity (Figure S4). However, the observed nonsense variants are predicted to undergo nonsense-mediated mRNA decay in vivo, and our data also suggest that the truncated protein p.Gln305*, if produced in vivo, would only partially localize to the nucleus. Thus, it remains unclear to what extent pathogenesis results from dominant-negative or loss-of-function mechanisms.

To study the interaction between EBF3 mutants and chromatin, we performed in situ subcellular fractionation.[30] We treated transiently transfected HEK293T cells expressing either $EBF3^{WT}$ or one of the mutants (p.Asn66Asp, p.Tyr141Cys, p.Pro177Leu, p.Arg209Trp, or p.Gln305*) to extract cytoplasmic proteins and then selectively extracted non-tightly chromatin-bound proteins and free protein aggregates within the nucleus. Detection of FLAG-tagged EBF3 proteins by immunoblotting demonstrated that all mutants were present in both the cytoplasmic and nuclear fractions (Figure 3C). In contrast, $EBF3^{WT}$ was absent in the fraction containing cytoplasmic proteins and only barely detectable in the nuclear fraction, indicating that it is tightly associated with chromatin (Figure 3C).

To further test the hypothesis that the *EBF3* variants affect EBF3 DNA binding and gene regulation, we transfected SK-N-SH cells with constructs encoding either $EBF3^{WT}$ or p.Pro177Leu EBF3 ($EBF3^{P177L}$) (Figure S5), stably selected the cells for integration and expression, and then performed RNA sequencing (RNA-seq). The SK-N-SH cell line is derived from a neuroblastoma that has little to no endogenous *EBF3* expression, and we used it to minimize any influence of endogenous EBF3 on our functional assessments. Overexpression of $EBF3^{WT}$ or $EBF3^{P177L}$ resulted in 679 or 377 transcripts, respectively, expressed differentially between transfected and untransfected cells (false-discovery rate [FDR] < 0.05; Figures 4A and 4B and Table S4). Gene Ontology (GO) terms were identified with the online tool g:Profiler[34] with all GO term annotation categories. Significantly enriched GO terms associated with $EBF3^{WT}$ expression included various neuron- and signaling-related pathways, supporting a role for EBF3 in neurodevelopment. In contrast, $EBF3^{P177L}$ did not yield as significant an enrichment in neuron-related pathways (Figure S6). Overall, analyses of the SK-N-SH $EBF3^{WT}$ and $EBF3^{P177L}$ transcriptomes indicated that EBF3 targets a wide variety of genes and that the p.Pro177Leu substitution alters EBF3-mediated gene regulation.

To determine whether the p.Pro177Leu substitution affects EBF3 binding across the genome, we performed chromatin immunoprecipitation coupled with massively parallel sequencing (ChIP-seq[38]) in both $EBF3^{WT}$- and $EBF3^{P177L}$-transfected SK-N-SH cells. $EBF3^{WT}$ and $EBF3^{P177L}$ genome-wide binding sites were identified by established methods.[39] We identified 21,046 binding sites for $EBF3^{WT}$ and 4,193 binding sites for $EBF3^{P177L}$, 4,081 of which were shared (Figure 4C), i.e., $EBF3^{P177L}$ binds to a minor subset of $EBF3^{WT}$ binding sites. MEME-Suite motif analysis[40] with HOCOMO (v.10) identified an enrichment of the canonical EBF zinc-binding motif in called peaks from assays of both wild-type and mutant proteins (Figure 4C). We then investigated whether differentially expressed genes were closer to either $EBF3^{WT}$ or $EBF3^{P177L}$ binding sites. Consistent with the greater number of binding sites in $EBF3^{WT}$, genes were more likely to be closer to $EBF3^{WT}$ binding sites than $EBF3^{P177L}$ binding sites, but upregulated differentially expressed genes in both $EBF3^{WT}$ and $EBF3^{P177L}$ experiments were closer to binding sites than were non-significant genes (Figure 4D), indicating that $EBF3^{P177L}$ retains some regulatory function. To determine whether $EBF3^{WT}$ proximal binding sites have

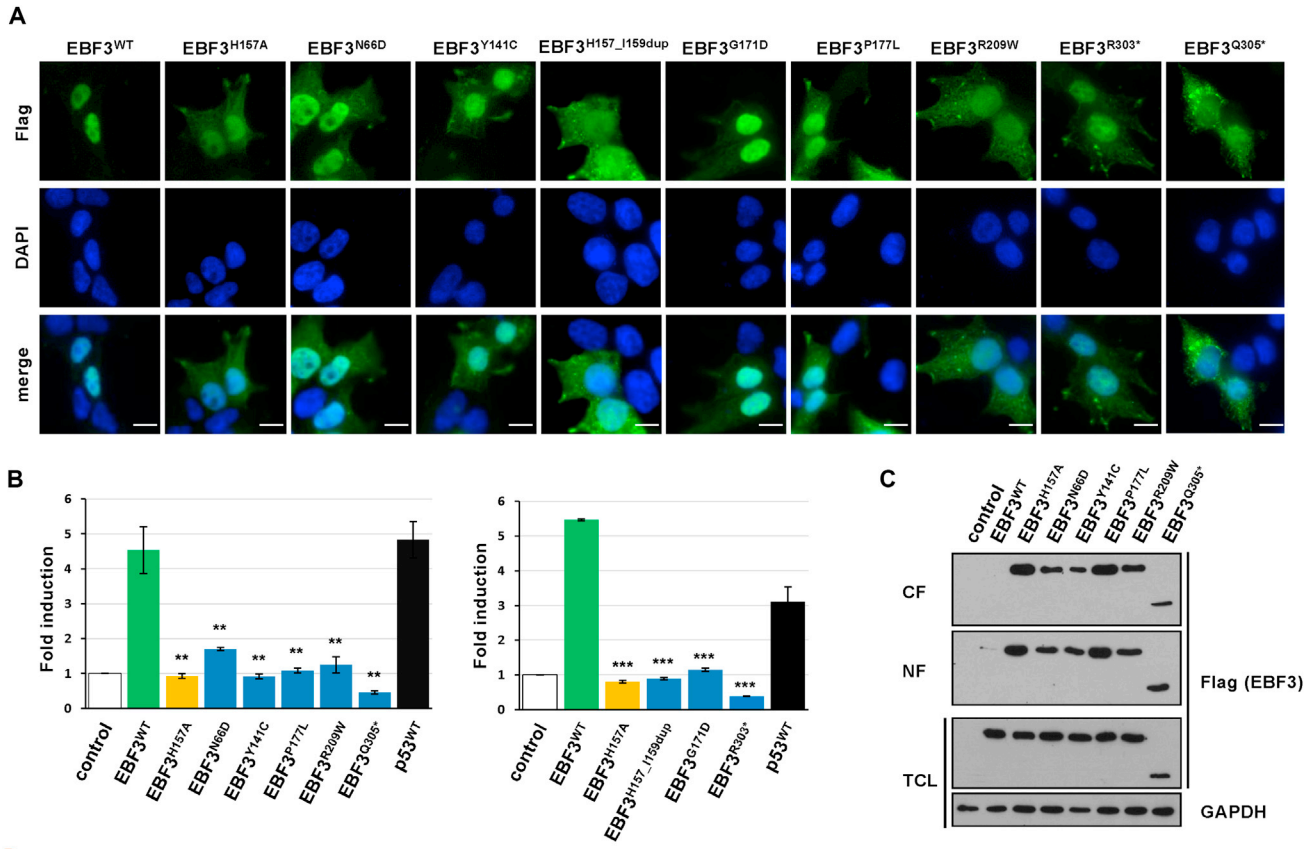

Figure 3. EBF3 Mutants Show Impaired DNA Binding and Altered Subcellular Localization

(A) HEK293T cells were transiently transfected with EBF3 expression constructs, fixed, treated with permeabilization-blocking solution, and incubated in mouse monoclonal anti-FLAG M2 antibody solution (1:200 dilution; clone F-3165, Sigma-Aldrich). After washing, cells were incubated with Alexa Fluor 488 coupled to goat anti-mouse IgG (1:1,000 dilution; ThermoFisher) and embedded in mounting solution (ProLong Diamond Antifade Mountant with DAPI, ThermoFisher). Cells were analyzed with the Olympus IX-81 epifluorescence microscope. Wild-type EBF3 ($EBF3^{WT}$, green) was exclusively localized in the nucleus (blue), whereas the DNA-binding-deficient mutant p.His157Ala ($EBF3^{H157A}$) and the disease-associated mutants p.Asn66Asp ($EBF3^{N66D}$), p.Tyr141Cys ($EBF3^{Y141C}$), p.Gly171Asp ($EBF3^{G171D}$), p.His157_Ile159dup ($EBF3^{H157\_I159dup}$), p.Pro177Leu ($EBF3^{P177L}$), p.Arg209Trp ($EBF3^{R209W}$), and p.Arg303* ($EBF3^{R303*}$) were also located in the cytoplasm. Representative images are shown. Error bars represent 10 μm.

(B) EBF3 mutants show impaired activation of luciferase reporter expression under the control of the *CDKN1A* (p21) promoter. HEK293T cells were transiently transfected with the expression construct(s) of interest, together with pGL2-p21 (CDKN1A) promoter-Luc[28] and pREN in a 1:1:3 ratio of pREN(2 μg):pGL2-p21 promoter-Luc(2 μg):pFLAG-CMV4-EBF3 or pFLAG-CMV4-$p53^{WT}$(6 μg). pGL2-p21 promoter-Luc encodes *Photinus* luciferase, and pREN is a derivate of the pFiRe-basic[29] encoding *Renilla* luciferase. Wild-type p53 was used as an internal control. Dual luciferase assays were done with the extracts of transfected cells 48 hr after transfection. Data were normalized to the activity of *Renilla* luciferase, and basal promotor activity for transfection with pFLAG-CMV4-cassetteA (control vector) was considered to be 1. Compared to the empty vector (control; white bar), expression of $EBF3^{WT}$ (green bar) and p53 (black bar) led to 4- to 5-fold elevated promoter activity. The DNA-binding-deficient $EBF3^{H157A}$ mutant (yellow bar) and all disease-associated EBF3 mutants (blue bars) showed strongly reduced or no activation of the luciferase reporter. The normalized luciferase activity (mean ± SD) of three independent experiments is depicted as the fold induction relative to that of cells transfected with a control vector. All comparisons are in reference to $EBF3^{WT}$, and p values were calculated with the two-sided Student's t test (**$p < 0.005$, ***$p < 0.0005$).

(C) EBF3 mutants are not tightly bound to chromatin. 24 hr after transfection of HEK293T cells with EBF3 expression constructs or pFLAG-CMV4-cassetteA (control vector), in situ subcellular fractionation was performed.[30] Cells were incubated with CSK buffer containing 0.1% Triton-X. The cytoplasmic extracts were removed, and proteins were precipitated. Cells were subsequently treated with CSK buffer supplemented with 0.5% Triton-X. Nuclear extracts were removed, and proteins were precipitated. Total cell lysate (TCL), cytoplasmic fraction (CF), and nuclear fraction (NF) were analyzed by SDS-PAGE and immunoblotting with a mouse monoclonal anti-FLAG M2 peroxidase conjugate (1:50,000 dilution; Sigma-Aldrich). For control of equal loading, TCL was analyzed with mouse anti-GAPDH antibody (1:10,000 dilution; Abcam). The mutant EBF3 proteins were present in both the CF and NF. In marked contrast, $EBF3^{WT}$ was present in only minimal amounts in the NF, demonstrating exclusive nuclear localization and strong chromatin binding. Data represent four independent experiments.

some functional transcriptional consequence that is altered by the p.Pro177Leu substitution, we compared log_2 fold changes for $EBF3^{WT}$ and $EBF3^{P177L}$ significantly differentially expressed genes with a transcription start site (TSS) within 50 kb of shared $EBF3^{WT}$ and $EBF3^{P177L}$ binding sites ($n = 146$). These genes had significantly ($p = 0.0177$) smaller log_2 fold changes in $EBF3^{P177L}$ samples than in $EBF3^{WT}$ samples, as quantified on the basis

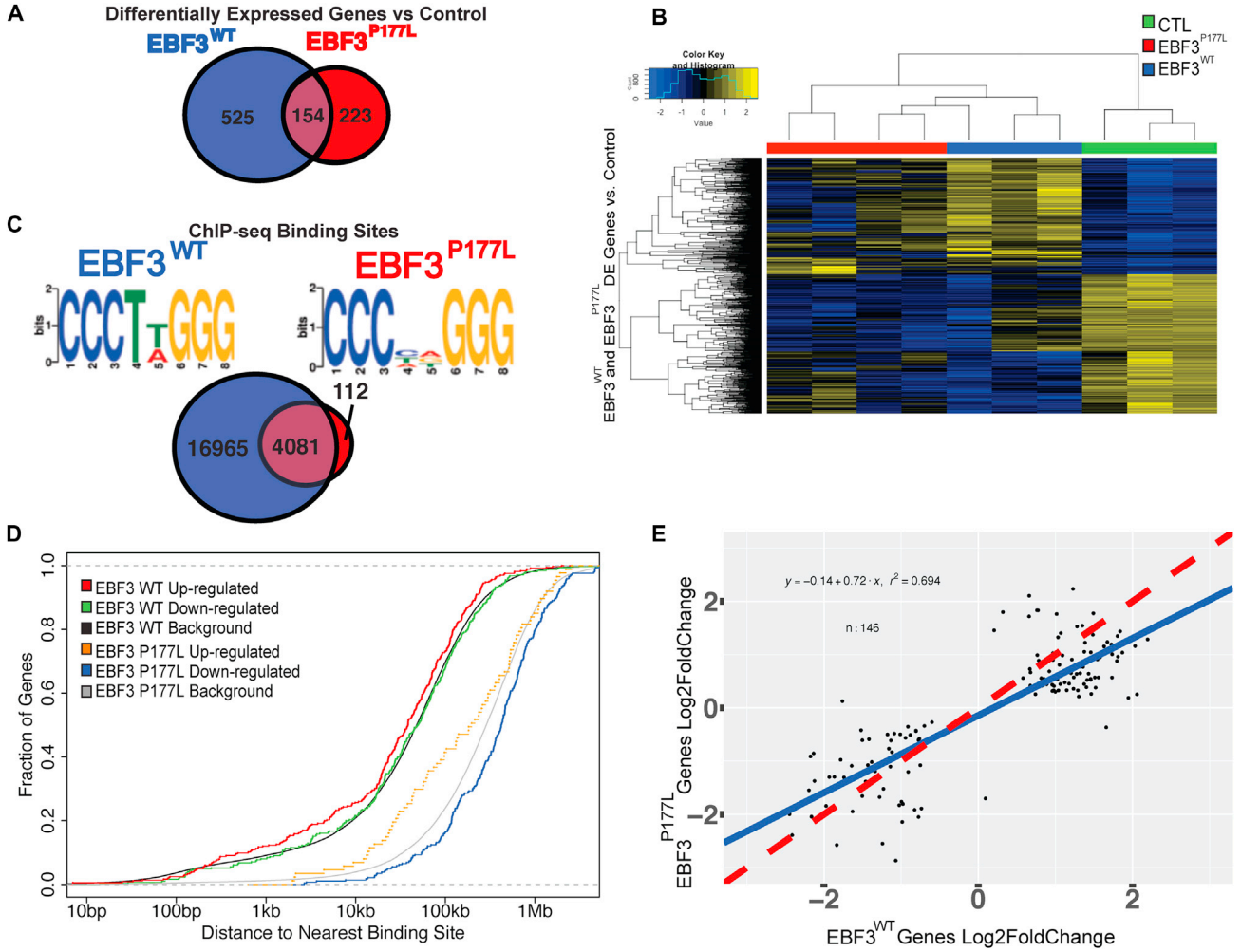

Figure 4. EBF3^{P177L} Overexpression Shows Less Transcriptome Alteration and Whole-Genome EBF3 Occupancy Than EBF3^{WT}

(A) Expression of EBF3^{P177L} has less transcriptome alteration than EBF3^{WT}. We generated two SK-N-SH (ATCC HTB-11) EBF3^{WT}- and two EBF3^{P177L}-expressing stable cell-line seed stocks, each of which represented a pool of two transfections (1×10^6 SK-N-SH cells transfected with 5 μg of respective overexpression construct). Replicates from these samples were grown and maintained independently under selection (along with SK-N-SH controls), and then RNA was isolated and used to make sequencing libraries. RNA-seq libraries were prepared with the Nextera DNA Library Sample Prep Kit according to established protocols.[31] Libraries were sequenced on an Illumina HiSeq 2500. All statistical analyses were performed in R (v.3.2.1). RNA-seq reads were processed with aRNApipe.[32] To perform differential gene-expression analysis, we used the R DESeq2 package[33] (v.1.8.2) with default Wald-test hypothesis testing with an adjusted p value (FDR) cutoff of 0.05. GO term enrichment was performed with the online tool g:Profiler[34] with all GO term annotation categories. The Venn diagram depicts genes identified as significantly differentially expressed between SK-N-SH cells (control) and EBF3^{WT} and EBF3^{P177L} cells; 154 were shared between EBF3^{WT} and EBF3^{P177L} cells.

(B) A heatmap of DESeq2 variance-stabilized RNA-seq expression values compares SK-N-SH control (CTL), EBF3^{WT}, and EBF3^{P177L} samples for genes determined to be significantly different between either EBF3^{WT} or EBF3^{P177L} cells and control SK-N-SH cells.

(C) EBF3^{P177L} reduces genome-wide EBF3 binding sites determined by ChIP-seq (bottom). Available anti-EBF3 antibodies have been found to have some degree of cross-reactivity with other EBF family members, which limits the interpretability of specific family member binding sites.[35] Therefore, we performed ChIP by using a high-affinity monoclonal anti-FLAG M2 antibody (Sigma) that targets the C-terminal 3X-FLAG epitope of the *EBF3* cDNA constructs. Libraries were sequenced on an Illumina NextSeq. ChIP-seq reads were aligned with the Burrows-Wheeler Aligner to the UCSC Genome Browser (hg19), and peaks were identified for each replicate with MACS2.1.0[36] with an -mfold cutoff of [10, 30]. We merged replicate overlapping peaks with BEDTools[37] to generate the final peak lists used in downstream analyses. The most significant motifs with centrally enriched distribution for EBF3^{WT} and EBF3^{P177L} were identified with MEME-Suite (top).

(D) Significantly upregulated genes are closer to EBF3^{WT} and EBF3^{P177L} binding sites. A cumulative distribution function (CDF) plot of EBF3^{WT} and EBF3^{P177L} shows the distance from the GRCh37 TSS to the nearest EBF3^{WT} and EBF3^{P177L} binding sites for genes identified as upregulated or downregulated (FDR < 0.05, $\log_2$ fold change > 0 and < 0, respectively) in comparison to background.

(E) EBF3^{WT} or EBF3^{P177L} significantly differentially expressed genes with a TSS within 50 kb of shared ChIP-seq binding sites exhibit a relatively greater absolute $\log_2$ fold change in EBF3^{WT} than the same genes in EBF3^{P177L} expression data. Linear regression of $\log_2$ fold changes for these genes ($n = 146$) exhibits a downward-skewed slope of 0.72 (blue line) in comparison to the null expectation of perfect correspondence (slope = 1, dashed red line), indicating comparatively reduced alteration of expression for significant EBF3^{WT} genes by EBF3^{P177L}.

of a random sampling of 10,000 146-gene sets matched for total count distributions from the EBF3^{WT} and EBF3^{P177L} DESeq2 results (Figure 4E and Figure S7). Thus, even when binding occurs, the p.Pro177Leu substitution appears to lead to reduced transcriptional alteration. Together, these findings are consistent with the notion that EBF3 acts as a proximal regulator of transcription at *cis*-regulatory sequences and support the hypothesis that EBF3^{P177L} has reduced binding and regulatory function as a result of partial disruption of the DBD.

Finally, we assessed the significance of the observed enrichment of *EBF3* mutations within affected individuals. Although many submissions to GeneMatcher result from assessments of individual families (precluding meaningful rate estimation), we identified three cohorts that allow quantification of mutation enrichment. Two individuals described here were found in a Clinical Sequencing Exploratory Research study[41] on a series of 375 unrelated families. According to the Samocha et al.[9] estimated probabilities for observing a de novo protein-altering mutation within *EBF3* ($10^{-4.856}$ per chromosome), the p value for observing two such individuals in this cohort is ~1.9×10^{-4}. Additionally, five de novo protein-altering variants (two missense, one frameshift, and two splice) in *EBF3* have recently been reported in a series of 4,293 families with individuals with developmental disorders ($p = 3.8 \times 10^{-6}$), although neither of these variants nor *EBF3* was discussed or concluded to be pathogenic.[42] Both missense variants in this cohort are located in the DBD of EBF3, and one (p.Pro177Leu) is identical to a variant in this study, indicating mutational recurrence. Finally, a meta-analysis of 6,206 trios with autism, schizophrenia, ID, and/or epilepsy identified two individuals (one with autism and one with ID) with de novo disruptions of *EBF3* ($p = 0.042$).[43] The lower rate of *EBF3* mutation in this last cohort is likely to be at least partly due to the inclusion of 1,063 trios with schizophrenia. The combined signal of enrichment of de novo *EBF3* mutations among individuals with neurodevelopmental disorders is highly significant (nine de novo *EBF3* mutations in 10,874 individuals, $p = 9.4 \times 10^{-9}$).

In conclusion, our results show that de novo mutations disrupting the regulatory functions of the conserved neurodevelopmental transcription-factor-encoding gene *EBF3* lead to ID, ataxia, and facial dysmorphism. *EBF3* mutations are estimated to affect ~1 in 1,000 individuals with otherwise unexplained neurodevelopmental disorders. Our study furthermore underscores the importance of data sharing and collaborative human genetics, for which tools such as GeneMatcher[3] promote the assembly of unique sample collections that enable the discovery and analysis of genetic contributions to disease.

Accession Numbers

WES data from the two trios sequenced at HudsonAlpha have been submitted to dbGAP and are available under study accession number dbGAP: phs001089.v1.p1 and individual IDs SRX1716411, SRX1716420, SRX1716424, SRX1716279, SRX1716280, and SRX1716281. The RNA-seq and ChIP-seq data described in this paper have been submitted to GEO and are available under accession number GEO: GSE90682.

Supplemental Data

Supplemental Data include seven figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.11.012.

Acknowledgments

We thank the families who contributed to this study. We thank the HudsonAlpha Clinical Sequencing Exploratory Research team (Shirley Simmons, Kelly East, Whitley Kelley, Candice Finnila, David Gray, Michelle Amaral, and Michelle Thompson); Ryne Ramaker and Brian Roberts for help with data analysis; Verena Kolbe for technical assistance; Hans-Jürgen Kreienkamp for help with structural analysis and critical reading; Stefan Kindler and Claudia Schob for help with the luciferase assay; and the Microscopy Imaging Facility of the University Medical Center Hamburg-Eppendorf for technical support. We thank Martin Walsh for plasmid pGL2-p21 promoter-Luc (Addgene plasmid 33021), Stefan Kindler for pREN, and Kristoffer Rieken for LeGO-iG2-Puro+-p53. This work was supported by grants from the NIH (UM1HG007301 to HudsonAlpha, 1R21NS094047-01 to K.M.G. as co-principal investigator, HG006542 to the Baylor-Hopkins Center for Mendelian Genomics, and 5T32GM008361-21 to the University of Alabama at Birmingham), the Simons Foundation (337701 to W.K.C.), and the Deutsche Forschungsgemeinschaft (KO 4576/1-1 and KO 4576/1-2 to F.K. and KU 1240/10-1 to K.K.). W.-L.C. was supported by the Cancer Prevention & Research Institute of Texas training program (RP140102). J.J. is an employee of GeneDx. J.R.L. has stock ownership in 23andMe, is a paid consultant for Regeneron Pharmaceuticals, has stock options in Lasergen Inc., is a member of the scientific advisory board of Baylor Genetics, and is a co-inventor on multiple patents related to molecular diagnostics. The Baylor College of Medicine derives revenue from the chromosomal microarray analysis and clinical exome sequencing offered at Baylor Genetics. M.B. is the founder of Codified Genomics LLC.

Received: August 10, 2016
Accepted: November 11, 2016
Published: December 22, 2016

Web Resources

1000 Genomes Project, http://www.1000genomes.org/
BioRxiv, http://biorxiv.org/
Clustal Omega, http://www.ebi.ac.uk/Tools/msa/clustalo/
COSMIC, http://cancer.sanger.ac.uk/cosmic
dbNSFP, https://sites.google.com/site/jpopgen/dbNSFP
dbSNP, http://www.ncbi.nlm.nih.gov/SNP/
ExAC Browser, http://exac.broadinstitute.org/
GeneMatcher, http://www.genematcher.org/
Genic Intolerance, http://genic-intolerance.org/
NCBI Gene, http://www.ncbi.nlm.nih.gov/gene/
NCBI HomoloGene, http://www.ncbi.nlm.nih.gov/homologene/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/

OMIM, http://www.omim.org/
RefSeq, https://www.ncbi.nlm.nih.gov/refseq/
Protein Data Bank (PDB), http://www.rcsb.org/pdb/home/home.do
SWISS-MODEL, http://swissmodel.expasy.org/
UCSC Genome Browser, http://genome.ucsc.edu
UCSF Chimera, http://www.cgl.ucsf.edu/chimera/

References

1. de Ligt, J., Willemsen, M.H., van Bon, B.W., 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.
2. Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature *511*, 344–347.
3. 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.
4. Girisha, K.M., Kortüm, F., Shah, H., Alawi, M., Dalal, A., Bhavani, G.S., and Kutsche, K. (2016). A novel multiple joint dislocation syndrome associated with a homozygous nonsense variant in the EXOC6B gene. Eur. J. Hum. Genet. *24*, 1206–1210.
5. Kortüm, F., Caputo, V., Bauer, C.K., Stella, L., Ciolfi, A., Alawi, M., Bocchinfuso, G., Flex, E., Paolacci, S., Dentici, M.L., et al. (2015). Mutations in KCNH1 and ATP6V1B2 cause Zimmermann-Laband syndrome. Nat. Genet. *47*, 661–667.
6. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
7. Campbell, I.M., Shaw, C.A., Stankiewicz, P., and Lupski, J.R. (2015). Somatic mosaicism: implications for disease and transmission genetics. Trends Genet. *31*, 382–392.
8. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.
9. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nat. Genet. *46*, 944–950.
10. Tanaka, A.J., Cho, M.T., Retterer, K., Jones, J.R., Nowak, C., Douglas, J., Jiang, Y.-H., McConkie-Rosell, A., Schaefer, G.B., Kaylor, J., et al. (2016). De novo pathogenic variants in CHAMP1 are associated with global developmental delay, intellectual disability, and dysmorphic facial features. Cold Spring Harb Mol Case Stud *2*, a000661.
11. Bainbridge, M.N., Wang, M., Wu, Y., Newsham, I., Muzny, D.M., Jefferies, J.L., Albert, T.J., Burgess, D.L., and Gibbs, R.A. (2011). Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol. *12*, R68.
12. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N. Engl. J. Med. *369*, 1502–1511.
13. Yang, Y., Muzny, D.M., Xia, F., Niu, Z., Person, R., Ding, Y., Ward, P., Braxton, A., Wang, M., Buhay, C., et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. JAMA *312*, 1870–1879.
14. Isidor, B., Küry, S., Rosenfeld, J.A., Besnard, T., Schmitt, S., Joss, S., Davies, S.J., Lebel, R.R., Henderson, A., Schaaf, C.P., et al. (2016). De Novo Truncating Mutations in the Kinetochore-Microtubules Attachment Gene CHAMP1 Cause Syndromic Intellectual Disability. Hum. Mutat. *37*, 354–358.
15. Liao, D. (2009). Emerging roles of the EBF family of transcription factors in tumor suppression. Mol. Cancer Res. *7*, 1893–1901.
16. Boller, S., and Grosschedl, R. (2014). The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. Immunol. Rev. *261*, 102–115.
17. Wang, M.M., and Reed, R.R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature *364*, 121–126.
18. Garel, S., Marín, F., Mattéi, M.G., Vesque, C., Vincent, A., and Charnay, P. (1997). Family of Ebf/Olf-1-related genes potentially involved in neuronal differentiation and regional specification in the central nervous system. Dev. Dyn. *210*, 191–205.
19. Pozzoli, O., Bosetti, A., Croci, L., Consalez, G.G., and Vetter, M.L. (2001). Xebf3 is a regulator of neuronal differentiation during primary neurogenesis in Xenopus. Dev. Biol. *233*, 495–512.
20. Quillé, M.L., Carat, S., Quéméner-Redon, S., Hirchaud, E., Baron, D., Benech, C., Guihot, J., Placet, M., Mignen, O., Férec, C., et al. (2011). High-throughput analysis of promoter occupancy reveals new targets for Arx, a gene mutated in mental retardation and interneuronopathies. PLoS ONE *6*, e25181.
21. Olivetti, P.R., and Noebels, J.L. (2012). Interneuron, interrupted: molecular pathogenesis of ARX mutations and X-linked infantile spasms. Curr. Opin. Neurobiol. *22*, 859–865.
22. Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., et al. (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. *43*, D805–D811.
23. Zhao, L.Y., Niu, Y., Santiago, A., Liu, J., Albert, S.H., Robertson, K.D., and Liao, D. (2006). An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis. Cancer Res. *66*, 9445–9452.
24. Fields, S., Ternyak, K., Gao, H., Ostraat, R., Akerlund, J., and Hagman, J. (2008). The 'zinc knuckle' motif of Early B cell Factor is required for transcriptional activation of B cell-specific genes. Mol. Immunol. *45*, 3786–3796.
25. Treiber, N., Treiber, T., Zocher, G., Treiber, N., Treiber, T., Zocher, G., and Grosschedl, R. (2010). Structure of an Ebf1 : DNA complex reveals unusual DNA recognition and structural homology with Rel proteins. Genes Dev. *24*, 2270–2275.
26. Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. *31*, 3381–3385.
27. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF

Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.
28. Nishio, H., and Walsh, M.J. (2004). CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription. Proc. Natl. Acad. Sci. USA *101*, 11257–11262.
29. Schütt, J., Falley, K., Richter, D., Kreienkamp, H.J., and Kindler, S. (2009). Fragile X mental retardation protein regulates the levels of scaffold proteins and glutamate receptors in postsynaptic densities. J. Biol. Chem. *284*, 25479–25487.
30. Sawasdichai, A., Chen, H.-T., Abdul Hamid, N., Jayaraman, P.-S., and Gaston, K. (2010). In situ subcellular fractionation of adherent and non-adherent mammalian cells. J. Vis. Exp. *6*, 1–5.
31. Gertz, J., Varley, K.E., Davis, N.S., Baas, B.J., Goryshin, I.Y., Vaidyanathan, R., Kuersten, S., and Myers, R.M. (2012). Transposase mediated construction of RNA-seq libraries. Genome Res. *22*, 134–141.
32. Alonso, A., Lasseigne, B.N., Williams, K., Nielsen, J., Ramaker, R.C., Hardigan, A.A., Johnston, B., Roberts, B.S., Cooper, S.J., Marsal, S., et al. (2016). aRNApipe: A balanced, efficient and distributed pipeline for processing RNA-seq data in high performance computing environments. bioRxiv. http://dx.doi.org/10.1101/060277.
33. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.
34. Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. *44* (W1), W83–W89.
35. Rajakumari, S., Wu, J., Ishibashi, J., Lim, H.W., Giang, A.H., Won, K.J., Reed, R.R., and Seale, P. (2013). EBF2 determines and maintains brown adipocyte identity. Cell Metab. *17*, 562–574.
36. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.
37. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.
38. Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497–1502.
39. Reddy, T.E., Pauli, F., Sprouse, R.O., Neff, N.F., Newberry, K.M., Garabedian, M.J., and Myers, R.M. (2009). Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. Genome Res. *19*, 2163–2171.
40. Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. *37*, W202–W208.
41. Green, R.C., Goddard, K.A.B., Jarvik, G.P., Amendola, L.M., Appelbaum, P.S., Berg, J.S., Bernhardt, B.A., Biesecker, L.G., Biswas, S., Blout, C.L., et al.; CSER Consortium (2016). Clinical Sequencing Exploratory Research Consortium: Accelerating Evidence-Based Practice of Genomic Medicine. Am. J. Hum. Genet. *98*, 1051–1066.
42. McRae, J.F., Clayton, S., Fitzgerald, T.W., Kaplanis, J., Prigmore, E., Rajan, D., Sifrim, A., Aitken, S., Akawi, N., Alvi, M., et al. (2016). Prevalence, phenotype and architecture of developmental disorders caused by de novo mutation. bioRxiv. http://dx.doi.org/10.1101/049056.
43. Lelieveld, S.H., Reijnders, M.R.F., Pfundt, R., Yntema, H.G., Kamsteeg, E.J., de Vries, P., de Vries, B.B.A., Willemsen, M.H., Kleefstra, T., Löhner, K., et al. (2016). Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat. Neurosci. *19*, 1194–1196.