Integrated phenotypic and genotypic approach defines *EBF3*-related genotype-phenotype relationships

Cole A. Deisseroth, BS[1,2,14,16], Aarushi Nayak[3], Nathan D. Bliss[4], Vanesa Lerma, BS, MA[2,5], Ashley W. LeMaire, PhD[6], Vinodh Narayanan, MD[7], Christopher Balak, BS[8], Ginevra Zanni, MD, PhD[9], Enza Maria Valente, MD, PhD[10,11], Enrico Bertini, MD[9], Paul J. Benke, MD[12], Michael F. Wangler, MD[2,13,16], Hsiao-Tuan Chao, MD, PhD[1,2,5,13,14,15,17]

[1]Medical Scientist Training Program, Baylor College of Medicine, Houston, TX, 77030, USA

[2]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, 77030, USA

[3]Case Western Reserve University, Cleveland, OH, 44106, USA

[4]Department of Biomedical Engineering, Texas A&M University, College Station, TX, 77843, USA

[5]Department of Pediatrics, Division of Neurology and Developmental Neuroscience, Baylor College of Medicine, Houston, TX, 77030, USA

[6]Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, 77030, USA

[7]Center for Rare Childhood Disorders, Translational Genomics Research Institute, Phoenix, AZ, 85004, USA

[8]Biomedical Sciences Graduate Program, University of California at San Diego, San Diego, CA, 92093, USA

[9]Department of Neurosciences, Unit of Neuromuscular and Neurodegenerative Disorders, Bambino Gesu Children's Research Hospital IRCCS, Rome, 00146, Italy

[10]Department of Molecular Medicine, University of Pavia, Pavia, 27100, Italy

[11]Neurogenetics Research Centre, IRCCS Mondino Foundation, Pavia, 27100, Italy

[12]Joe DiMaggio Children's Hospital, Hollywood, FL, 33021, USA

[13]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030, USA

[14]Department of Neuroscience, Baylor College of Medicine, Houston, TX, 77030, USA

[15]McNair Medical Institute, The Robert and Janice McNair Foundation, Houston, TX, 77030, USA

[16]Co-corresponding authors: Cole.Deisseroth@bcm.edu (C.A.D.), mw147467@bcm.edu (M.F.W.)

[17]Corresponding author: hc140077@bcm.edu (H.T.C.), 832-826-0461, mailing address: 1250 Moursund St, Suite 925, Room 930.01, Houston, TX, 77030

Abstract

Purpose: Hypotonia, Ataxia, and Delayed Development syndrome is a neurodevelopmental disorder caused by heterozygous *Early B-Cell Factor 3* (*EBF3*) loss-of-function variants. Identified in 2016, the full spectrum of clinical findings and the relationship between the *EBF3* genotype and clinical outcomes has not been determined beyond its namesake features.

Methods: We combined a phenotypic assessment of 33 individuals molecularly diagnosed with *EBF3* pathogenic variants with a meta-analysis of 34 previously reported individuals. The combined 62 unique individuals enabled comparative cross-sectional phenotype and genotype analysis in the largest cohort to date of affected individuals. Cardinal distinguishing features were identified that facilitate phenotypic stratification for clinical diagnosis. We developed assessment scales to ascertain individuals at risk for pathogenic *EBF3* variants, stratify the clinical severity, and connect variant-specific molecular phenotypes to clinical outcomes.

Results: Our findings show that a specific class of *EBF3* variants affecting the evolutionarily conserved Zinc Finger (ZNF) motif, which is critical for stabilizing the protein interaction with the DNA target sequence, is associated with an increased risk of persistent motor and language impairments.

Conclusion: These findings highlight the impact of combining variant-specific molecular phenotypes with comprehensive clinical data to predict neurodevelopmental outcomes and potentially guide personalized decisions for therapeutic interventions.

Introduction

Early B-Cell Factor 3 (*EBF3*; OMIM*607407; HGNC:19087) encodes a Collier/Olf/EBF (COE) transcription factor located on chromosome 10q26. COE transcription factors control gene expression via DNA methylation, nucleosome remodeling, and active chromatin modifications[1–3], and regulate nervous system, immune system, bone, and muscle development[4–7]. Loss-of-function manipulations of *EBF3* homologs in worms (*unc-3*, *CeO/E*), flies (*knot, collier*), frogs (*Xcoe2*), and mice (*Ebf3*) are deleterious to survival and impair neurodevelopment[8]. In mice, *Ebf3* has been shown to regulate neuronal proliferation, specification, and migration[6,9–11]. Cellular studies revealed that *EBF3* binds to the promoters of many genes associated with neurodevelopmental disorders[12], suggesting that EBF3 dysfunction could result in a wide phenotypic spectrum.

Over 30 years ago, genomic deletions containing *EBF3* were identified in the 10q26-deletion syndrome, which is characterized by delayed development, hypotonia, ataxia, growth problems, neuropsychiatric comorbidities, and genitourinary abnormalities (OMIM #609625)[13–17]. However, the principal genetic alterations mediating pathogenesis in 10q26-deletion syndrome was unknown until recently. In 2016, *de novo EBF3* variants and a heterozygous *Ebf3* single gene deletion were discovered to cause the Hypotonia, Ataxia, and Delayed Development Syndrome (HADDS, OMIM#617330), which is characterized by the namesake features and impaired expressive language, neuropsychiatric comorbidities, and genitourinary abnormalities[12,18–20]. The strong overlap in features between HADDS and 10q26-deletion syndrome suggests that *EBF3* haploinsufficiency may be the principal pathogenic factor in 10q26-deletion syndrome.

Although the discovery of a monogenic disorder associated with pathogenic *EBF3* variants is fairly recent, there is increasing evidence that *EBF3*-related disorders may be more common than previously recognized because of a wide range of clinical severity[20–25]. This is consistent with a prior statistical model of *de novo* variants for autism spectrum disorder (ASD) and developmental and intellectual disability (DD/ID) that identified *EBF3* as one of ~1,000 genes with significant enrichment of *de novo* loss-of-function variants in affected individuals[26]. Therefore, expanding our understanding of the HADDS phenotypic spectrum and co-morbid conditions will advance diagnosis, facilitate prognosis, and improve neurodevelopmental outcomes for individuals with HADDS, 10q26-deletion syndrome, and other prevalent neurodevelopmental disorders with similar features.

A common challenge faced with newly recognized or relatively rare neurodevelopmental disorders is that incomplete knowledge about distinguishing features limits the stratification of affected individuals for traditional phenotypically driven diagnostics. To expand our understanding of the HADDS phenotypic spectrum and identify individuals at risk for pathogenic *EBF3* variants, we conducted a comprehensive analysis of 33 individuals with HADDS and correlated the findings with a literature meta-analysis for a combined assessment of 62 unique individuals. Our findings expand our understanding of the HADDS phenotypic spectrum and reveal that *EBF3* variant type and location are associated with quantifiable clinical severity.

MATERIALS AND METHODS

BCM investigators systematically evaluated 22 individuals in person at Texas Children's Hospital and 11 through medical record reviews for a total of 33 individuals. We compared our cohort to a meta-analysis of nine studies containing a total of 34 individuals[12,18–25]. Five of the individuals enrolled in our study were previously reported. Probands 7, 15, 19, 25, and 26 from our study (**Table S1**) are the same as meta-analysis probands M-11, M-1, M-25, M-31, and M-32, respectively (**Table S2**). We assessed a total of 62 unique individuals and compared demographics, clinical features, and molecular findings in our cohort to the meta-analysis (**Tables S1-S5**). Variant nomenclature is according to gene transcript (GenBank:NM_001005463.2, GRCh37). Odds ratios (OR) = AD/BC, where A is the number of probands from our cohort with the feature, B is the number of probands from our cohort without the feature, C is the number of probands from the meta-analysis with the feature, and D is the number of probands from the meta-analysis without the feature. 95% confidence intervals (CI) were calculated for a sample size *n* and proportion *p*:

$$CI = p \pm 1.96\sqrt{\frac{p(1-p)}{n}}$$

For height, weight, and occipito-frontal circumference (OFC) measurements, we calculated LMS (lambda-mu-sigma)-based Z-scores[27], where for a given L (degree of skewness), M (median), and S (variation coefficient) values, and an experimental value X, the z-score is = $[(X/M)^L - 1]/(LS)$. L, M, and S values for height and weight based on sex and age (birth to 20 years) are from the Centers for Disease Control and Prevention (CDC)[28]. For OFC, the L, M, and S values were taken from the World Health Organization (WHO, birth to five years)[29]. Delayed development was determined by failure to achieve developmental milestones within the CDC's

standard age range (**Tables S4, S5**)[30].

HADDS Developmental and Behavioral Scale (HADDS-DBS)

The HADDS-DBS assesses for seven major and 14 minor features (**Table 1**). Major features include decreased pain sensitivity (HP:0007328), weak/absent cry (HP:0001612), motor delay (HP:0001270), hypotonia (HP:0001252), speech delay (HP:0000750), ataxia (HP:0001251), and hypomimia (HP:0000338). Minor features include poor eye contact, strabismus, repetitive scratching in any location, repetitive scratching on back or neck, attraction to lights, aversion to crowds, aversion to loud noises, constipation, frontal bossing, recurrent urinary tract infections, feeding difficulties, cerebellar anomalies, gait abnormalities, and short stature. Based on the number of diagnostic features identified, the individuals are stratified into four tiers (**Table S6)**. Higher-numbered tiers indicates a higher risk for a pathogenic *EBF3* variant.

Positive (PPV) and negative (NPV) predictive values were calculated by combining the HADDS-DBS tier assignments with the molecularly confirmed genetic diagnoses. True positives (TP) are individuals with HADDS and a Tier-3 or -4 assignment. False positives (FP) are individuals with non-*EBF3* related genetic syndromes and a Tier-3 or -4 assignment. True negatives (TN) are individuals with non-*EBF3* related genetic syndromes and a Tier-1 or -2 assignment. False negatives (FN) are individuals with HADDS and a Tier-1 or -2 assignment. Sensitivity = TP/(TP+FN). Specificity = TN/(TN+FP). PPV = TP/(TP+FP). NPV = TN/(TN+FN).

For each diagnostic tier, OR = AD/BC, where A is the number of probands from our cohort at

the specified tier, B is the number of probands from our cohort not at the specified tier, C is the number of probands from the meta-analysis at the specified tier, and D is the number of probands from the meta-analysis not at the specified tier. Statistically significant differences at the 0.05 level were determined by one-way ANOVA with Tukey's post-hoc analysis.

HADDS Developmental Delay Severity Scale (HADDS-DDSS)

The four-tier HADDS-DDSS assesses for delayed development and neuroanatomical findings (**Table S7**). Higher-numbered tiers indicates increased severity. Statistically significant differences at the 0.05 level were determined by one-way ANOVA with Tukey's post-hoc analysis. OR for each severity tier were calculated using the same formula as the HADDS-DBS.

RESULTS

General features, demographics, and inheritance in HADDS

Demographics, growth parameters, genotype, and phenotype findings in our cohort were evaluated with a comparative literature meta-analysis (**Fig.1A-D, Table S3**). Growth parameters reveal below-average height and weight for individuals with HADDS (**Table S3**). However, head circumference both at birth and recent assessment remained within average parameters. Gender distribution was equivalent between our cohort and the meta-analysis (OR=1.19) (**Fig.1A**). More than 90% of affected individuals were identified at <18 years old (**Fig.1B**). Self-reported ethnicity in our study included European (76%), African (3%), East Asian (3%), Middle Eastern (6%), and multiethnic (12%) descent. Ethnicity was reported in <45% of individuals in the meta-analysis, but there was overall a similar distribution (**Fig.1C**).

The inheritance pattern for the majority of *EBF3* pathogenic variants was *de novo* in 79% of our study and 88% of the meta-analysis (OR=0.50, **Fig.1D**). Mendelian inheritance of a pathogenic *EBF3* variant was previously reported in one affected mother and son pair[23], and inheritance of a pathogenic *EBF3* variant from an unaffected parent with mosaicism was reported in an unaffected mother and two affected children[12]. In our study, we identified Mendelian inheritance in four affected individuals from three unrelated families (OR=4.55) and mosaicism in the parents of two affected and unrelated individuals (OR=1.03) (**Fig.1D**).

Pathogenic *EBF3* variants and genomic deletions

We identified 15 individuals with noncoding *EBF3* variants and genomic deletions (**Fig.1E**) and 47 individuals with coding variants (**Fig.1F**). The majority of *EBF3* variants cluster within the N-terminal DNA binding domain with 27% of pathogenic variants located within five amino acids of the zinc finger motif (ZNF, amino acid 157-170, OR=2.2), a critical region for stabilizing the interaction with the DNA target[31]. A prior *in vitro* study of the paralogous EBF1 showed that disrupting conserved residues in the ZNF, including p.Arg163, abolished DNA-binding activity[31]. Intriguingly, we identified the most frequent recurrent *EBF3* variants affected the p.Arg163 triplet in eight unrelated individuals with p.Arg163Gln (c.488G>A) in three, p.Arg163Pro (c.488G>C) in two, p.Arg163Trp (c.487C>T) in two, and p.Arg163Leu (c.488G>T) in one individual. Other recurrent variants affected p.Arg209 in six, p.Arg206 in five, p.Cys161 in three, and p.His157 in three individuals. The recurrent variants are likely due to the position of the affected nucleotides in CpG-dinucleotide islands, which are mutational hotspots underlying over one-third of *de novo* missense variants associated with human diseases[32,33]. Other potential mechanisms include the selfish spermatogonial selection process[34],

the error-prone replication hypothesis[35], or a selection bias for similar phenotypes.

Common neurologic features in HADDS

Prevalent neurologic features identified in the cohort include delayed development, motor incoordination, perturbed expressive language, altered sensory processing, neuropsychiatric comorbidities, and cerebellar alterations. Consistent between our cohort and the meta-analysis, we found that >95% of affected individuals had delayed motor development and >80% had delayed expressive language development (OR=2.14, **Fig.2A**). Intriguingly, we found discordance between expressive and receptive language development with 66% (CI[16, 56]) of individuals diagnosed with HADDS able to follow one-step verbal commands on time and 57% (CI[15, 59]) able to follow two-step verbal commands on time. This finding suggests that individuals with HADDS have a greater prevalence of delayed expressive language development, potentially related to motor incoordination and speech apraxia due to cerebellar dysfunction[36,37], but maintain age appropriate receptive language development. Another common language finding was dysarthria (21%, CI[7, 35]), which was identified in a similar frequency in the meta-analysis (OR=1.04, **Fig.2B**). We also identified speech apraxia in 21% (CI[7, 35], OR=4.31, **Fig.2B**) and weak or absent cry in 97% (CI[91, 103]) of affected individuals in our study, a higher prevalence than previously reported (OR 512, **Fig.2B**).

We identified a high prevalence of hypotonia (100%) and ataxia (78%, CI[64, 92], OR=2.6) in our cohort (**Fig.2C**). The comprehensive phenotypic analysis also revealed a high frequency of motor incoordination (63%, CI[47, 80], OR=8.17), hypomimia (57%, CI[40, 74], OR=7.87), and hypertonia (39%, CI[22, 56], OR=21.45), but these features were reported in less than 20% of

affected individuals in the meta-analysis (**Fig.2C**). Additional neurologic findings include an increased risk of seizures (39%, CI[22, 56], OR=21.45), febrile seizures (28%, CI[13, 43], OR=3.875), and abnormalities on electroencephalogram (EEG) recordings (18%, CI[5, 31] OR=1.04) (**Fig.2D**). Intriguingly, we found that altered sensory processing is a cardinal feature of HADDS. All individuals in our study had chronic decreased pain response, but a smaller prevalence was identified in the meta-analysis (18%, CI[5, 30]), potentially due to incomplete ascertainment (**Fig.2E**). Other sensory findings in our cohort include repetitive scratching behaviors (51%, CI[34, 69]), hyperacusis (48%, CI[31, 66]), and discomfort with loud noises (45%, CI [28, 62]) (**Fig.2E**), all of which were not previously reported.

Neuropsychiatric comorbidities are commonly observed in affected individuals with HADDS. We found that these features occur at a higher frequency than previously reported. Prevalent neuropsychiatric comorbidities in HADDS include autistic features (52%, CI [34, 69], OR=17.6), formal autism diagnosis (27%, CI [12, 42], OR=8.82), attention deficit and hyperactivity disorder (ADHD, 18%, CI [5, 31], OR=1.03), and emotional lability (18%, CI [5, 31], OR=3.56) (**Fig.2F**). Common autistic features identified in our study include stereotypies (63%, CI [47, 80], OR=11.8), poor eye contact (39%, CI [23, 56]), attraction to lights (24%, CI [10, 39]), and aversion to crowds (21%, CI [7, 35]) (**Fig.2G**). In contrast, the meta-analysis identified autistic features, autism, or autism spectrum disorder in 9-18% of individuals (**Fig.2F**).

Consistent with the neurodevelopmental impairments, neuroimaging studies in our cohort revealed altered cerebellar structure in 33% (CI [17, 49]) of individuals, compared to 18% (CI [9, 38]) in the meta-analysis (OR=1.63, **Fig.2H**). Cerebellar vermian hypoplasia was observed in

30% (CI [15, 46]) of individuals in our study, compared to 18% (CI [5, 30]) in the meta-analysis (OR=2.03), and is the most common neuroanatomical finding. Less frequent findings in our cohort include prominent ventricles (9%, CI [-1, 19]), abnormal corpus callosum (one individual), delayed myelination (one individual), hydrocephalus (one individual), and partial rhomboencephalosynapsis (one individual) (**Fig.2H**).

Non-neurological findings in HADDS

Other features commonly seen in HADDS include craniofacial, gastrointestinal, ophthalmologic, genitourinary, and musculoskeletal involvement. In our study, 84% (CI[73, 97]) of individuals had at least one craniofacial feature (OR=0.75). Recurrent features included deep-set eyes or prominent forehead (48%, CI[31, 66]), ear abnormalities (36%, CI[20, 53]), mouth anomalies (39%, CI[23, 56]), straight eyebrows (21%, CI[7, 35]), tall forehead (21%, CI[7, 35]), long face (18%, CI[5, 31]), pointed chin (18%, CI[5, 31]), myopathic facies (12%, CI[1, 23]), triangular facies (9%, CI[-1, 19]), anteverted nares (9%, C [-1, 19]), and short, broad chin (6%, CI[-2, 14]) (**Fig.3A**). Ophthalmologic findings also occur frequently in HADDS with strabismus reported in 87% (CI[77, 99]) of individuals in our cohort compared to 61% (CI[45, 78]) in the meta-analysis (OR=4.49) and nystagmus reported in 6% (CI[2, 14], OR=1.03) of individuals in both our cohort and the meta-analysis (**Fig.3B**).

In our cohort, we identified a high frequency of constipation (84%, CI[73, 97], OR=42), feeding difficulties (78%, CI[64, 92], OR=12) and gastroesophageal reflux (57%, CI[41, 74], OR=10), compared to a meta-analysis frequency of 12-21% for these features (**Fig.3C**). Genitourinary complications are frequently observed in both males and females. In our cohort, we found that

recurrent urinary tract infections (UTIs) occurred in 48% (CI[31, 66]) of individuals compared to 24% (CI[9, 38]) in the meta-analysis (OR=3.06), with a higher prevalence in females (**Fig.3D**). Anatomical findings in our study included abnormal kidney or bladder anatomy (33%, CI[17, 49], OR=2.9), vesicoureteric reflux (18%, CI[5, 31], OR=0.62), abnormal external genitalia (9%, CI[-1, 19], OR=0.47), and cryptorchidism (6%, CI[-2, 14], OR=0.48) (**Fig.3D**), which had a similar frequency in the meta-analysis. Female reproductive organ anomalies were seen in 6% of the meta-analysis (**Table S2**) but were not observed in our cohort (**Table S1**). The most prevalent musculoskeletal finding was short stature, which we identified in 33% (CI[17, 49]) of individuals in our cohort and was similarly reported in 41% (CI[24, 58] of the meta-analysis (OR=0.71). Other musculoskeletal dysmorphisms occurred in less than 10% of individuals in our cohort and in the meta-analysis (OR≤3). Dysmorphisms included clinodactyly (9%, CI[-1,19]), pectus excavatum (6%, CI[-2,14]), and low muscle mass (6%, CI[-2,14]) (**Fig.3E**). Short digits were also reported in two individuals in the literature (**Table S2**).

Diagnostic utility of the HADDS developmental and behavioral scale

To aid clinical recognition and expedited diagnosis of HADDS, we developed a diagnostic assessment scale (HADDS-DBS) to identify individuals "at risk" for pathogenic *EBF3* variants and prioritize *EBF3*-specific gene testing. We designed the HADDS-DBS to address the challenges of assessing children in various age groups and developed the scale for use by caregivers and healthcare providers. The HADDS-DBS can be used to screen for *EBF3* pathogenic variant "risk" in children as young as 15-months old and takes the form of a symptom questionnaire comprised of 21 clinical features (**Table 1**). Based on the number of major and minor clinical features identified in the questionnaire, the HADDS-DBS assigns one of four tiers

based on increasing risk of carrying a pathogenic *EBF3* variant with recommendations for molecular testing (**Table S6**). When applied to our cohort, the HADDS-DBS identified 32 individuals as Tier-4 (97%), and one individual as Tier-2 (3%) (**Fig.4A, Table S8**). When applied to the 34 individuals in the meta-analysis, the HADDS-DBS identified three individuals as Tier-4 (8%, OR=331), seven individuals as Tier-3 (21%), 12 individuals as Tier-2 (35%, OR=0.06), and 12 individuals as Tier-1 (35%) (**Fig.4A**, **Table S9**). These differences in phenotypic stratification between our cohort and the meta-analysis reflects the variability in phenotypic ascertainment across different clinical sites and highlights the value of comprehensive clinical phenotyping (**Fig.4A**, $p<0.001$, F=84.17).

To determine the specificity of the HADDS-DBS, we applied the diagnostic scale to three neurodevelopmental disorders with similar clinical features (hypotonia, delayed development, ataxia, and variable cognitive impairment), but do not have *EBF3* gene alterations. Three studies were selected that described eight individuals with Harel-Yoon syndrome (HYS, OMIM#617183) caused by *ATAD3A* variants[38], 20 individuals with Mental Retardation and Microcephaly with Pontine and Cerebellar Hypoplasia (MICPCH, OMIM#300749) caused by *CASK* variants[39], and 15 individuals with Takenouchi-Kosaki syndrome (TKS, OMIM#616737) caused by *CDC42* variants[40]. Of the eight individuals with HYS[38], two were classified Tier-1 (25%) and six as Tier-2 (75%). For the 20 individuals with MICPCH[39], 10 were classified as Tier-1 (50%) and 10 as Tier-2 (50%). Finally, of the 15 individuals with TKS[40], 14 were classified as Tier-1 (93%) and one as Tier-2 (7%). Statistical comparisons with one-way ANOVA reveal a significant difference for the phenotypic stratification of the HADDS meta-analysis compared to individuals with either HYS, MICPCH, or TKS (**Fig.4A**, $p<0.05$, F=84.17)

with sensitivity=0.29, specificity=1.0, PPV=1.0, and NPV=0.64. In contrast, comparison of our HADDS cohort to individuals with either HYS, MICPCH, or TKS reveal sensitivity=0.97, specificity=1.0, PPV=1.0, and NPV=0.97, which is consistent with development of the HADDS-DBS based on findings from our cohort.

Correlating *EBF3* variant type with severity of motor and language delays

We used the HADDS-DDSS to identify molecular phenotype groups with differential risk of motor and language impairments (**Table S7**). Of the combined 32 individuals diagnosed with HADDS who were more than five-years old, we classified nine individuals as Tier-1, 16 as Tier-2, three as Tier-3, and four as Tier-4. We examined the HADDS-DDSS classification relative to the *EBF3* variant type (noncoding, gene deletion, or coding) and the location of coding variants relative to the ZNF. For the five individuals with *EBF3* noncoding variants or gene deletion, four were classified as Tier-2 and one was classified as Tier-4 severity (**Fig.4B**). For the 18 individuals with *EBF3* coding variants located outside of the ZNF, nine were classified as Tier-1, seven as Tier-2, one as Tier-3, and one as Tier-4. Finally, for the nine individuals with *EBF3* coding variants located in the ZNF region, the HADDS-DDSS classified five individuals as Tier-2, two as Tier-3, and two as Tier-4 (**Fig.4B**). A significant correlation was identified between increased clinical severity and coding variants within five amino acids of the ZNF (**Fig.4B**, $p<0.05$, F=4.57). Intriguingly, >25% of the individuals with HADDS have *EBF3* variants located within the ZNF, which may reflect a selection bias for consistent phenotypes or increased clinical testing due to more pronounced phenotypic features.

DISCUSSION

Despite increased information available regarding *EBF3*-related HADDS, the need remains for a clear understanding of the phenotypic variability and how the *EBF3* genotype affects specific clinical findings. Our comprehensive phenotyping in a large cohort of individuals molecularly diagnosed with pathogenic *EBF3* variants allow for an expanded understanding of the variation in clinical outcomes and improved clinical management (**Fig.S1**). Neuroanatomical studies reveal that the cerebellum is the most common structure perturbed in individuals with *EBF3* pathogenic variants. Extensive anatomical and functional studies revealed critical roles for the cerebellum in motor and non-motor functions, including language, cognition, and social behavior[37]. Thus, many of the features observed in individuals with HADDS may be consistent with cerebellar dysfunction. To further understand the relationship between the *EBF3* molecular phenotype and clinical outcome, we designed two assessment scales that determine an individual's risk for having a pathogenic *EBF3* variant (HADDS-DBS) and their degree of functional impairments (HADDS-DDSS). Our findings validated the clinical utility of the scales, revealed critical insights into the *EBF3* genotype-phenotype map, and demonstrated that molecular phenotypes correlate with quantitative clinical findings (**Fig.4**).

The ability to broadly predict the risk of a pathogenic *EBF3* variant advances clinical diagnosis, facilitates prognosis, and aids in resolving *EBF3* variants of uncertain significance (VUS). Moreover, we were interested in determining whether the *EBF3* molecular phenotype could provide insights into the variation of clinical outcomes. By integrating the HADDS-DDSS assessment with the *EBF3* molecular phenotype, we found that missense variants within the ZNF define a subgroup of individuals at risk for persistent motor and language impairments. This

finding may be related to a stronger dominant-negative effect resulting from *EBF3* missense variants located within the ZNF. Haploinsufficient non-coding *EBF3* variants resulting in nonsense mediated decay or *EBF3* gene deletion appears to correlate with intermediate clinical outcomes. However, this interpretation is limited by the small sample size. Furthermore, the variation in clinical outcomes associated with *EBF3* variants raises the possibility that subtle impairments in EBF3 function or protein levels may result in a limited subset of HADDS features or manifest as a non-syndromic ASD or developmental delay. This hypothesis is supported by prior work identifying *EBF3* as one of ~1,000 genes with significant enrichment of *de novo* loss-of-function variants in individuals with ASD and DD/ID[26].

By using the largest set of clinically annotated *EBF3* variants examined to date, we strengthen prior findings by showing that HADDS-associated *EBF3* variants are related to a wide spectrum of clinical outcomes, identified Mendelian transmission of potentially hypomorphic variants, and expanded our understanding of cardinal diagnostic features. Furthermore, we found that the distribution of pathogenic *EBF3* variants remains predominantly in the DNA-binding domain, suggesting that variants in the C-terminal domains are less damaging to EBF3 function. Of note, we found that *EBF3* variants affecting the ZNF are associated with an increased risk of persistent motor and/or language impairments. These differential risk profiles may account for the increased frequency and severity of *de novo* missense variants in the ZNF and surrounding nucleotides, whereas milder variants resulting in subtle impairments in EBF3 function may not be identified primarily on the basis of a clinical HADDS diagnosis.

Despite advances in our recognition of *EBF3*-related HADDS, the neurobiological basis of

differential risk profiles based on *EBF3* variant type and location remains to be elucidated. As EBF3 functions through homo- or heterodimerization, the missense variants may result in a dominant-negative effect that alters the function of the wildtype EBF3 protein and reduces the overall amount of functional EBF3 dimers. This would result in increased perturbations to the EBF3-dependent transcriptional regulation in the developing nervous system. In contrast, *EBF3* single gene deletion or non-coding variants may preserve functional EBF3 dimers but result in an overall reduction of EBF3 available to regulate developmental processes. Ideally, a comprehensive analysis would include the effect of variation on EBF3's protein-protein interactions, subcellular localizations, regulation of target gene expression profiles, and nuclear function. The increasing recognition of individuals at risk for pathogenic *EBF3* variants and new high-throughput assays might make such datasets available for expanded analysis of the neurobiological consequences of the molecular phenotype.

In conclusion, we show that pathogenic *EBF3* variants cause a broad spectrum of phenotypic severity. This expands our understanding of defining clinical features in HADDS. Through the integration of the HADDS-DBS and HADDS-DDSS assessments with the molecular phenotype, we found that a subset of individuals with *EBF3* missense variants within the ZNF have a higher risk of functional impairments. Furthermore, the phenotypic overlap between HADDS with ASD and DD/ID suggest that subtle EBF3 dysfunction may result in non-syndromic features and that HADDS may share common pathogenic mechanisms with ASD and DD/ID. Longitudinal follow-up of these individuals in conjunction with new prospective recruitment efforts will be needed to comprehensively delineate the *EBF3* genotype-phenotype map, which may lead to promising avenues for therapeutic interventions.

ACKNOWLEDGEMENTS

We thank the families for participation in this study, Keri Ramsey with clinical assistance, and Mingshan Xue, Sahana Murthy, Harim Delgado-Seo, Maimuna Paul, and Joshi Stephen for manuscript feedback. There is no targeted funding for this study. H.T.C. is partially supported from NIH-1DP5OD026428. C.D. is supported by the BCM Medical Scientist Training Program. N.D.B. was supported by the Autism Science Foundation. V.L. is supported by The Gordon and Mary Cain Foundation and Annie and Bob Graham. EB and GZ are members of the European Reference Network for Rare Neurological Diseases-Project #739510. EB, EMV, and GZ are supported by the Ministry of Health (Ricerca Corrente 2020, Ricerca Finalizzata NET-2013-02356160).

DATA AVAILABILITY

De-identified data regarding molecular and clinical findings are available upon request.

AUTHOR INFORMATION

Conceptualization: C.D., H.T.C.; Data curation: C.D., H.T.C., A.N., N.D.B, V.L., A.W.L., V.N., C.B., G.Z., E.M.V., E.B., P.J.B., M.F.W., H.T.C.; Formal analysis: C.D., H.T.C.; Funding acquisition: H.T.C.; Investigation: H.T.C.; Methodology: C.D., H.T.C.; Project administration: H.T.C., V.L.; Resources: H.T.C.; Software: H.T.C.; Supervision: H.T.C.; Validation: C.D., H.T.C.; Visualization: C.D., A.N., N.D.B., H.T.C.; Writing—original draft: C.D., H.T.C.; Writing—review & editing: C.D., H.T.C., A.N., N.D.B, V.L., A.W.L., V.N., C.B., G.Z., E.M.V., E.B., P.J.B., M.F.W., H.T.C.

ETHICS DECLARATION

Individuals were enrolled in a study approved by the Institutional Review Board of Baylor College of Medicine (BCM), following its ethical standards. Written and informed consent was obtained from all individuals before the study was conducted. All procedures followed national standards. All data were de-identified.

REFERENCES

1. Decker T, Pasca di Magliano M, McManus S, et al. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity*. 2009;30(4):508-520. doi:10.1016/j.immuni.2009.01.012

2. Maier H, Ostraat R, Gao H, et al. Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. *Nat Immunol*. 2004;5(10):1069-1077. doi:10.1038/ni1119

3. Treiber T, Mandel EM, Pott S, et al. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription- independent poising of chromatin. *Immunity*. 2010;32(5):714-725. doi:10.1016/j.immuni.2010.04.013

4. Green YS, Vetter ML. EBF proteins participate in transcriptional regulation of Xenopus muscle development. *Dev Biol*. 2011;358(1):240-250. doi:10.1016/j.ydbio.2011.07.034

5. Kuriki M, Sato F, Arai HN, et al. Transient and lineage-restricted requirement of Ebf3 for sternum ossification. *Development*. 2020;147(9). doi:10.1242/dev.186239

6. Wang SS, Lewcock JW, Feinstein P, Mombaerts P, Reed RR. Genetic disruptions of O/E2 and O/E3 genes reveal involvement in olfactory receptor neuron projection. *Development*. 2004;131(6):1377-1388. doi:10.1242/dev.01009

7. Sugiyama T, Omatsu Y, Nagasawa T. Niches for hematopoietic stem cells and immune cell progenitors. *Int Immunol*. 2019;31(1):5-11. doi:10.1093/intimm/dxy058

8. Liberg D, Sigvardsson M, Akerblad P. The EBF/Olf/Collier family of transcription factors: regulators of differentiation in cells originating from all three embryonal germ layers. *Mol Cell Biol*. 2002;22(24):8389-8397. doi:10.1128/mcb.22.24.8389-8397.2002

9. Chiara F, Badaloni A, Croci L, et al. Early B-cell factors 2 and 3 (EBF2/3) regulate early migration of Cajal-Retzius cells from the cortical hem. *Dev Biol*. 2012;365(1):277-289. doi:10.1016/j.ydbio.2012.02.034

10. Fulp CT, Cho G, Marsh ED, Nasrallah IM, Labosky PA, Golden JA. Identification of Arx transcriptional targets in the developing basal forebrain. *Hum Mol Genet*. 2008;17(23):3740-3760. doi:10.1093/hmg/ddn271

11. Garel S, Garcia-Dominguez M, Charnay P. Control of the migratory pathway of facial branchiomotor neurones. *Development*. 2000;127(24):5297-5307.

12. Harms FL, Girisha KM, Hardigan AA, et al. Mutations in EBF3 Disturb Transcriptional Profiles and Cause Intellectual Disability, Ataxia, and Facial Dysmorphism. *Am J Hum Genet*. 2017;100(1):117-127. doi:10.1016/j.ajhg.2016.11.012

13. Lewandowski RC, Kukolich MK, Sears JW, Mankinen CB. Partial deletion 10q. *Hum Genet*. 1978;42(3):339-343. doi:10.1007/BF00291317

14. Shapiro SD, Hansen KL, Pasztor LM, et al. Deletions of the long arm of chromosome 10. *Am J Med Genet*. 1985;20(1):181-196. doi:10.1002/ajmg.1320200122

15. Mehta L, Duckett DP, Young ID. Behaviour disorder in monosomy 10qter. *J Med Genet*. 1987;24(3):185-186. doi:10.1136/jmg.24.3.185

16. Tanabe S, Akiba T, Katoh M, Satoh T. Terminal deletion of chromosome 10q: clinical features and literature review. *Pediatr Int*. 1999;41(5):565-567. doi:10.1046/j.1442-200x.1999.01105.x

17. Irving M, Hanson H, Turnpenny P, et al. Deletion of the distal long arm of chromosome 10; is there a characteristic phenotype? A report of 15 de novo and familial cases. *Am J Med Genet A*. 2003;123A(2):153-163. doi:10.1002/ajmg.a.20220

18. Chao H-T, Davids M, Burke E, et al. A Syndromic Neurodevelopmental Disorder Caused by De Novo Variants in EBF3. *Am J Hum Genet*. 2017;100(1):128-137. doi:10.1016/j.ajhg.2016.11.018

19. Sleven H, Welsh SJ, Yu J, et al. De Novo Mutations in EBF3 Cause a Neurodevelopmental Syndrome. *Am J Hum Genet*. 2017;100(1):138-150. doi:10.1016/j.ajhg.2016.11.020

20. Lopes F, Soares G, Gonçalves-Rocha M, Pinto-Basto J, Maciel P. Whole Gene Deletion of EBF3 Supporting Haploinsufficiency of This Gene as a Mechanism of Neurodevelopmental Disease. *Front Genet*. 2017;8:143. doi:10.3389/fgene.2017.00143

21. Blackburn PR, Barnett SS, Zimmermann MT, et al. Novel de novo variant in EBF3 is likely to impact DNA binding in a patient with a neurodevelopmental disorder and expanded phenotypes: patient report, in silico functional assessment, and review of published cases. *Cold Spring Harb Mol Case Stud*. 2017;3(3):a001743. doi:10.1101/mcs.a001743

22. Tanaka AJ, Cho MT, Willaert R, et al. De novo variants in EBF3 are associated with hypotonia, developmental delay, intellectual disability, and autism. *Cold Spring Harb Mol Case Stud*. 2017;3(6). doi:10.1101/mcs.a002097

23. Beecroft SJ, Olive M, Quereda LG, et al. Cylindrical spirals in two families: Clinical and genetic investigations. *Neuromuscul Disord*. Published online December 25, 2019. doi:10.1016/j.nmd.2019.12.006

24. Harkness JR, Beaman GM, Teik KW, et al. Early B-cell Factor 3-Related Genetic Disease Can Mimic Urofacial Syndrome. *Kidney Int Rep*. 2020;5(10):1823-1827. doi:10.1016/j.ekir.2020.07.001

25. Hildebrand MS, Jackson VE, Scerri TS, et al. Severe childhood speech disorder. *Neurology*. 2020;94(20):e2148. doi:10.1212/WNL.0000000000009441

26. Samocha KE, Robinson EB, Sanders SJ, et al. A framework for the interpretation of de novo mutation in human disease. *Nat Genet*. 2014;46(9):944-950. doi:10.1038/ng.3050

27. Flegal KM, Cole TJ. *Construction of LMS Parameters for the Centers for Disease Control and Prevention 2000 Growth Charts*. US Department of Health and Human Services, Centers for Disease Control and …; 2013.

28. Kuczmarski RJ. *2000 CDC Growth Charts for the United States: Methods and Development*. Department of Health and Human Services, Centers for Disease Control and …; 2002.

29. World Health Organization. *WHO Child Growth Standards: Methods and Development*. World Health Organization; 2007.

30. Milestone Checklist. Accessed March 24, 2020. https://www.cdc.gov/ncbddd/actearly/pdf/checklists/Checklists-with-Tips_Reader_508.pdf

31. Hagman J, Gutch MJ, Lin H, Grosschedl R. EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains. *EMBO J*. 1995;14(12):2907-2916.

32. Magewu AN, Jones PA. Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. *Mol Cell Biol*. 1994;14(6):4225-4232. doi:10.1128/mcb.14.6.4225

33. Mancini D, Singh S, Ainsworth P, Rodenhiser D. Constitutively methylated CpG dinucleotides as mutation hot spots in the retinoblastoma gene (RB1). *Am J Hum Genet*. 1997;61(1):80-87. doi:10.1086/513898

34. Goriely A, McGrath JJ, Hultman CM, Wilkie AOM, Malaspina D. "Selfish spermatogonial selection": a novel mechanism for the association between advanced paternal age and neurodevelopmental disorders. *Am J Psychiatry*. 2013;170(6):599-608. doi:10.1176/appi.ajp.2013.12101352

35. Dhokarh D, Abyzov A. Elevated variant density around SV breakpoints in germline lineage lends support to error-prone replication hypothesis. *Genome Res*. 2016;26(7):874-881. doi:10.1101/gr.205484.116

36. Bolceková E, Mojzeš M, Van Tran Q, et al. Cognitive impairment in cerebellar lesions: a logit model based on neuropsychological testing. *Cerebellum Ataxias*. 2017;4:13. doi:10.1186/s40673-017-0071-9

37. Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. *Annu Rev Neurosci*. 2009;32:413-434. doi:10.1146/annurev.neuro.31.060407.125606

38. Harel T, Yoon WH, Garone C, et al. Recurrent De Novo and Biallelic Variation of ATAD3A, Encoding a Mitochondrial Membrane Protein, Results in Distinct Neurological Syndromes. *Am J Hum Genet*. 2016;99(4):831-845. doi:10.1016/j.ajhg.2016.08.007

39. Moog U, Kutsche K, Kortüm F, et al. Phenotypic spectrum associated with CASK loss-of-function mutations. *J Med Genet*. 2011;48(11):741-751. doi:10.1136/jmedgenet-2011-100218

40. Martinelli S, Krumbach OHF, Pantaleoni F, et al. Functional Dysregulation of CDC42 Causes Diverse Developmental Phenotypes. *Am J Hum Genet*. 2018;102(2):309-320. doi:10.1016/j.ajhg.2017.12.015

TABLE AND FIGURE LEGENDS

Table 1: HADDS-Diagnostic Behavioral Scale (HADDS-DBS) questionnaire

The HADDS-DBS scale is comprised of seven major and 14 minor features and designed for use by parents and clinicians. Seven of the clinical features are considered "major" due to their high prevalence among the patients in the current cohort. These major features include decreased pain sensitivity (HP:0007328), weak/absent cry (HP:0001612), motor delay (HP:0001270), hypotonia (HP:0001252), speech delay (HP:0000750), ataxia (HP:0001251), and hypomimia (HP:0000338). The remaining 14 clinical features are considered "minor" and include poor eye contact, strabismus, repetitive itching affecting any location, repetitive itching specifically on the back of the neck, attraction to lights, aversion to crowds, aversion to loud noises, constipation, frontal bossing, recurrent urinary tract infections, feeding difficulties, cerebellar anomalies, gait abnormalities, and short stature.

Figure 1: Demographics and molecular findings in *EBF3*-related HADDS

Comparison of demographic and *EBF3* variant findings from 33 individuals in this study and 34 individuals in the meta-analysis. A total of 62 unique individuals were evaluated. **(A)** Gender distribution was equivalent between our cohort and the meta-analysis. **(B)** The molecularly confirmed diagnosis of *EBF3*-related HADDS occurred predominantly in the pediatric population, with more than 90% of affected individuals identified at less than 18 years old. **(C)** Ethnicity distribution in our study compared to meta-analysis. UNK = Unknown, EUR = European, AFR = African, EAS = East Asian, SAS = South Asian, MEA = Middle Eastern, MUL = Multiethnic. **(D)** Inheritance pattern for the majority of identified *EBF3* pathogenic variants was *de novo*. **(E, F)** There are 15 individuals with noncoding *EBF3* variants and

genomic deletions (**E**) and 47 individuals with coding variants (**F**). Published variants are shown in magenta, variants unique to this study are in purple, and variants shared between our cohort and the literature are in black. DBD is the DNA binding domain, ZNF is the zinc finger motif, IPT/TIG stands for immunoglobin-like, plexins, transcription factors/transcription factor immunoglobin, fold structure, and HLH is the helix-loop-helix region.

Figure 2: Frequency of neurological features in *EBF3*-related HADDS

From the combined analysis of 62 unique individuals, we identified distinct common neurologic features associated with pathogenic *EBF3* variants. Findings from our cohort are shown in solid bars and findings from the meta-analysis are shown in the outlined bars. Data shown as percent of individuals. Sample sizes are shown above each bar. Females are shown in red and males in blue. **(A)** Percent of individuals achieving language and gross motor developmental milestones on time (green) or delayed (purple) reveal prevalent motor and language delays. **(B)** Other language findings in our cohort include speech delay, weak or absent crying during infancy, speech apraxia, and dysarthria. **(C)** Frequent motor findings observed in affected individuals include hypotonia, motor delay, ataxia, motor incoordination, hypomimia, and hypertonia. **(D)** A risk of seizures, febrile seizures, and abnormal EEG findings are observed in our cohort. **(E)** Sensory alterations are prevalent in HADDS with frequent findings of decreased pain response, repetitive itching behaviors, hyperacusis, and aversion to loud noise. **(F)** Prevalent neuropsychiatric comorbidities in HADDS include autistic features, autism, ADHD, and emotional lability. **(G)** Common autistic features identified in our study include stereotypies, poor eye contact, attraction to bright lights, and aversion to crowded environments. **(H)** Neuroanatomical findings observed in our HADDS cohort include cerebellar-vermian hypoplasia

(CVH), prominent ventricles, abnormal corpus callosum, delayed myelination, hydrocephalus, and partial rhomboencephalosynapsis (RES).

Figure 3: Frequency of non-neurological features in *EBF3*-related HADDS

Other features commonly seen in individuals with HADDS include craniofacial, gastrointestinal, ophthalmologic, genitourinary, and musculoskeletal involvement. Findings from the current study are shown in solid bars and findings from the meta-analysis are shown in the outlined bars. Data is shown as a percentage of individuals. Sample sizes are shown above each bar. Females are shown in red and males in blue. **(A)** Craniofacial dysmorphisms are frequently observed with 84% of affected individuals in our cohort found to have at least one craniofacial finding. **(B)** Strabismus is the most common ophthalmology finding in individuals with HADDS. Nystagmus was reported in a subset of individuals. **(C)** Constipation, feeding difficulties, and gastroesophageal reflux disease (GERD) are common gastrointestinal (GI) findings identified in affected individuals. **(D)** Genitourinary complications are frequently observed in both genders with recurrent urinary tract infections (UTIs) and bladder control problems as the most prevalent findings. Other genitourinary findings include vesicoureteral reflux (VUR), abnormal bladder and/or kidney anatomy, abnormal external genitalia, and cryptorchidism **(E)** Musculoskeletal findings include short stature, clinodactyly, pectus excavatum, and low muscle mass.

Figure 4: Diagnostic assessment scale and genotype-phenotype mapping in *EBF3*-related HADDS

The HADDS-DBS is a symptom questionnaire designed to identify individuals at risk for pathogenic *EBF3* variants. **(A)** The HADDS-DBS diagnostic scale was applied to individuals

with molecularly confirmed HADDS enrolled in our study and previously reported in the literature. The majority of affected individuals enrolled in our study are classified as Tier-4, whereas the majority of affected individuals in the meta-analysis are classified as Tiers-2 and -3. When applied to three non-*EBF3* related disorders (OMIM#617183, #300749, and #616737), the HADDS-DBS primarily classified the reported individuals as Tier-1. Data is shown as a percentage of each cohort classified in Tiers-1 through -4. Sample sizes are shown above each bar. Tier-1 is shown in black, Tier-2 is shown in purple, Tier-3 is shown in blue, and Tier-4 is shown in red. Statistically significant differences were determined by one-way ANOVA with Tukey's post-hoc analysis. Significance shown as * $p<0.05$ and *** $p<0.001$. **(B)** The HADDS-DDSS is a symptom assessment scale designed to determine the severity of neurological limitations. The scale is based on readily ascertained developmental milestones (age of first word and age of independent ambulation) and the presence or absence of MRI brain findings. **(i)** The HADDS-DDSS tier assignment is examined relative to the *EBF3* variant type and location relative to the ZNF motif for genotype-phenotype mapping. A significant correlation was identified between symptom severity and coding variants within the ZNF motif (amino acids 157-170) and the flanking five amino acids (total range 152-175). Data is shown as the percentage of each cohort classified in Tiers-1 through -4. Sample sizes are shown above each bar. Tier-1 is in black, Tier-2 is in purple, Tier-3 is in blue, and Tier-4 is in red. Statistically significant differences were determined by one-way ANOVA with Tukey's post-hoc analysis. Significance shown as * $p<0.05$. **(ii)** Illustration shows the distribution of *EBF3* missense variants within five amino acids flanking the ZNF motif.

Figure 1

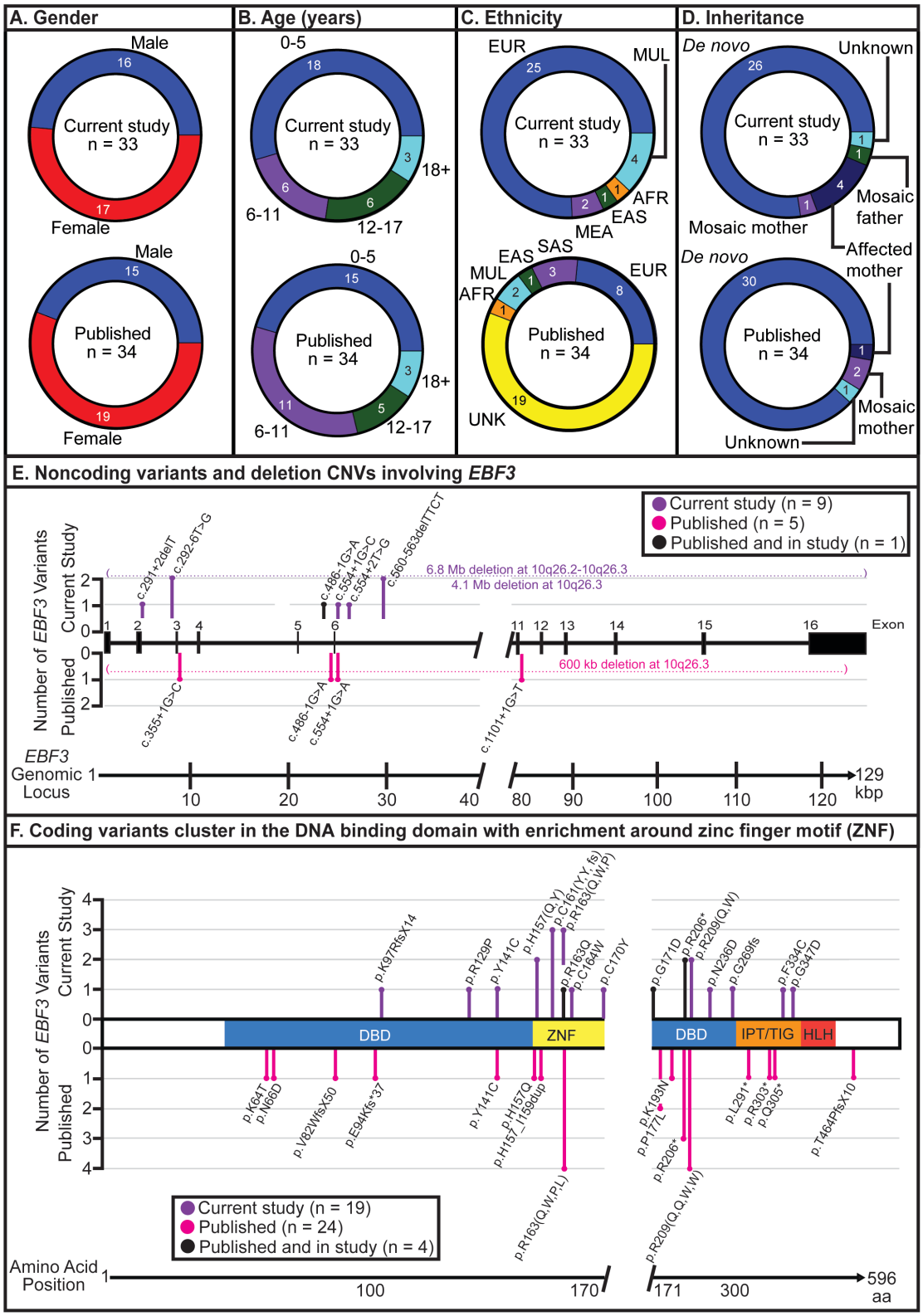
A. Gender
Male
Female
Current study n = 33
Published n = 34
B. Age (years)
0-5
6-11
12-17
18+
C. Ethnicity
EUR
MUL
AFR
EAS
MEA
SAS
UNK
D. Inheritance
De novo
Unknown
Mosaic father
Mosaic mother
Affected mother
Mosaic mother
E. Noncoding variants and deletion CNVs involving EBF3
Current study (n = 9)
Published (n = 5)
Published and in study (n = 1)
Number of EBF3 Variants
Current Study
Published
6.8 Mb deletion at 10q26.2-10q26.3
4.1 Mb deletion at 10q26.3
600 kb deletion at 10q26.3
Exon
EBF3 Genomic Locus
kbp
F. Coding variants cluster in the DNA binding domain with enrichment around zinc finger motif (ZNF)
DBD
ZNF
IPT/TIG
HLH
Current study (n = 19)
Published (n = 24)
Published and in study (n = 4)
Amino Acid Position
aa

Figure 2

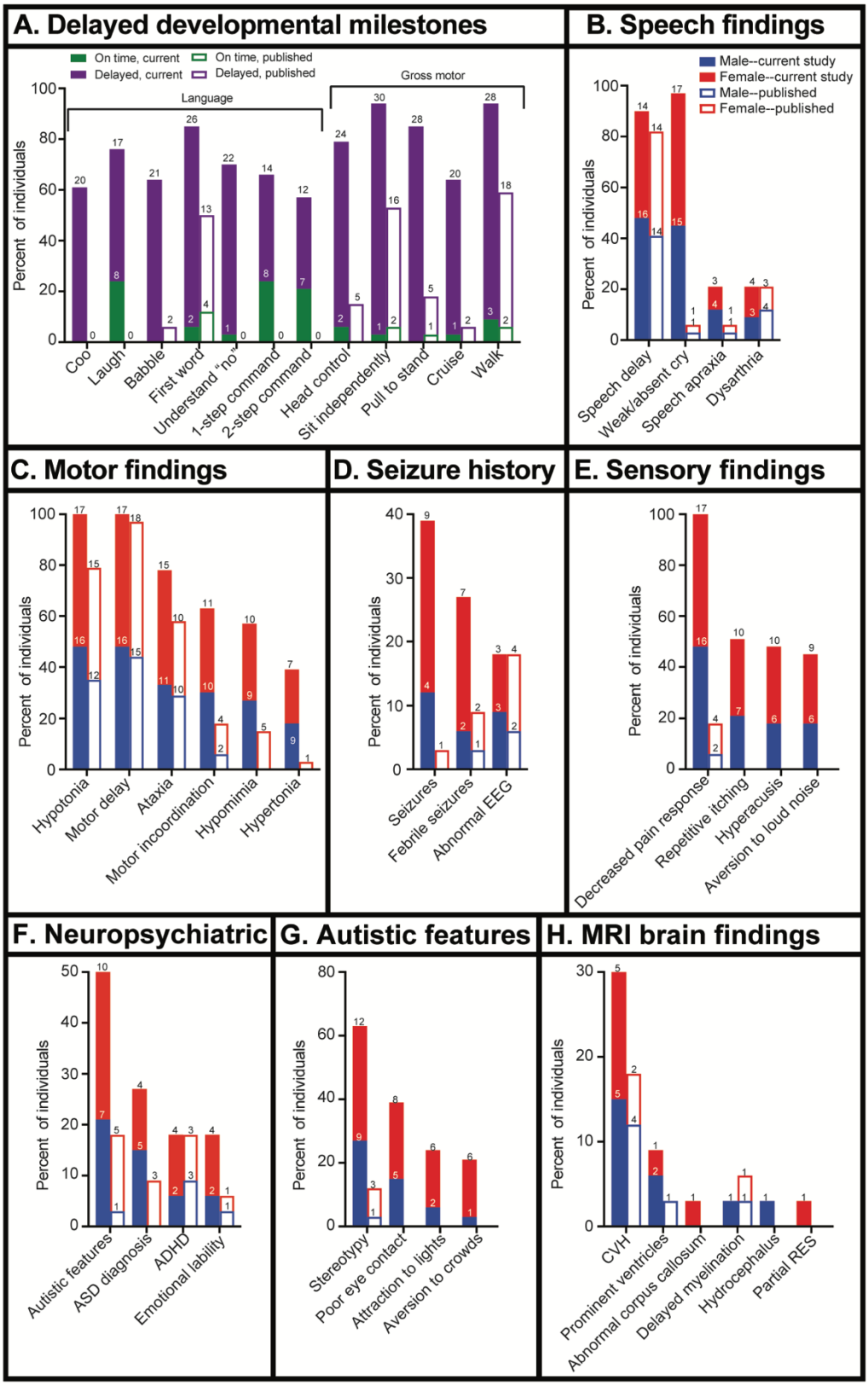
A. Delayed developmental milestones
On time, current
Delayed, current
On time, published
Delayed, published
Language
Gross motor
Percent of individuals
Coo
Laugh
Babble
First word
Understand "no"
1-step command
2-step command
Head control
Sit independently
Pull to stand
Cruise
Walk
B. Speech findings
Male--current study
Female--current study
Male--published
Female--published
Speech delay
Weak/absent cry
Speech apraxia
Dysarthria
C. Motor findings
Hypotonia
Motor delay
Ataxia
Motor incoordination
Hypomimia
Hypertonia
D. Seizure history
Seizures
Febrile seizures
Abnormal EEG
E. Sensory findings
Decreased pain response
Repetitive itching
Hyperacusis
Aversion to loud noise
F. Neuropsychiatric
Autistic features
ASD diagnosis
ADHD
Emotional lability
G. Autistic features
Stereotypy
Poor eye contact
Attraction to lights
Aversion to crowds
H. MRI brain findings
CVH
Prominent ventricles
Abnormal corpus callosum
Delayed myelination
Hydrocephalus
Partial RES

Figure 3

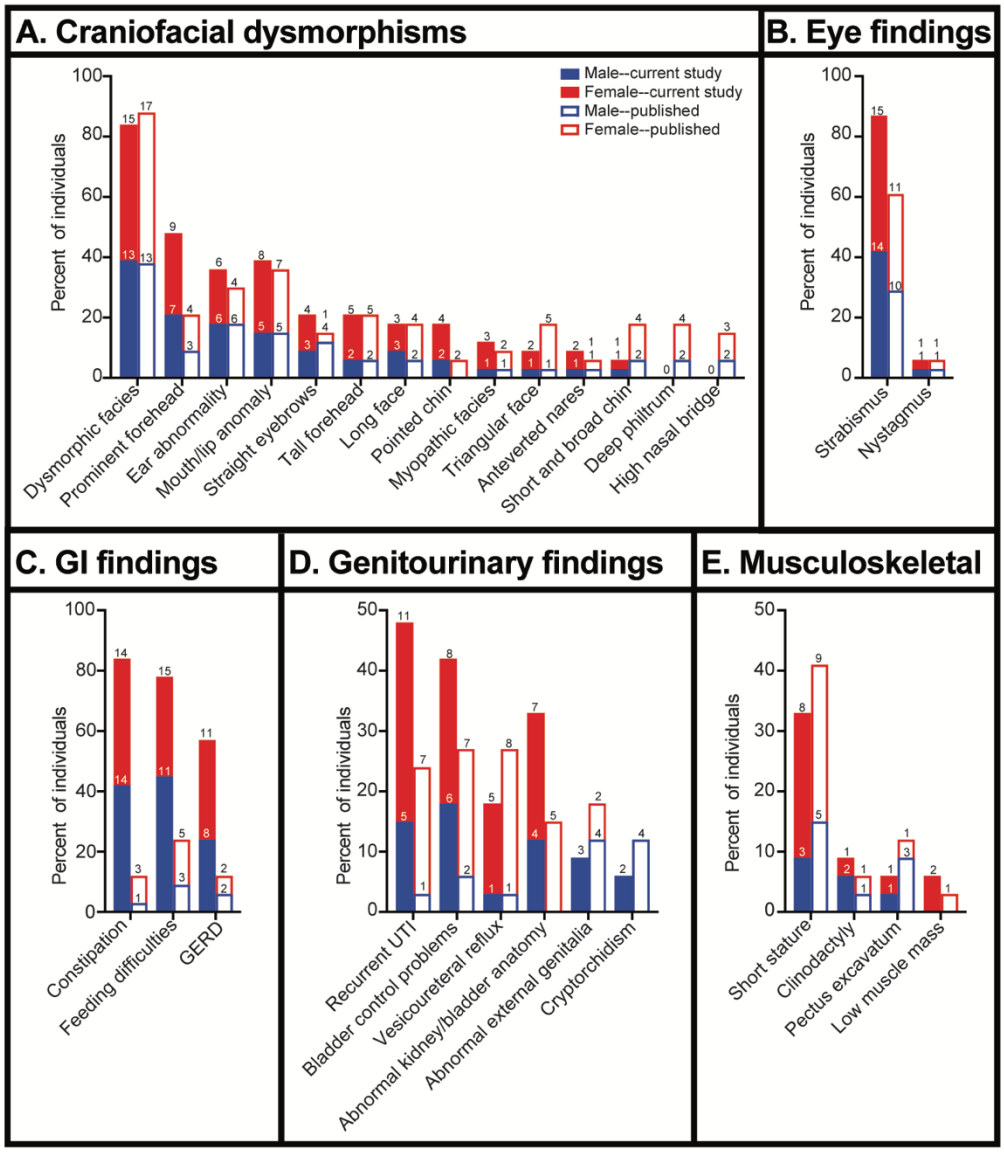
A. Craniofacial dysmorphisms
Male--current study
Female--current study
Male--published
Female--published
Percent of individuals
Dysmorphic facies
Prominent forehead
Ear abnormality
Mouth/lip anomaly
Straight eyebrows
Tall forehead
Long face
Pointed chin
Myopathic facies
Triangular face
Anteverted nares
Short and broad chin
Deep philtrum
High nasal bridge
B. Eye findings
Strabismus
Nystagmus
C. GI findings
Constipation
Feeding difficulties
GERD
D. Genitourinary findings
Recurrent UTI
Bladder control problems
Vesicoureteral reflux
Abnormal kidney/bladder anatomy
Abnormal external genitalia
Cryptorchidism
E. Musculoskeletal
Short stature
Clinodactyly
Pectus excavatum
Low muscle mass

Figure 4

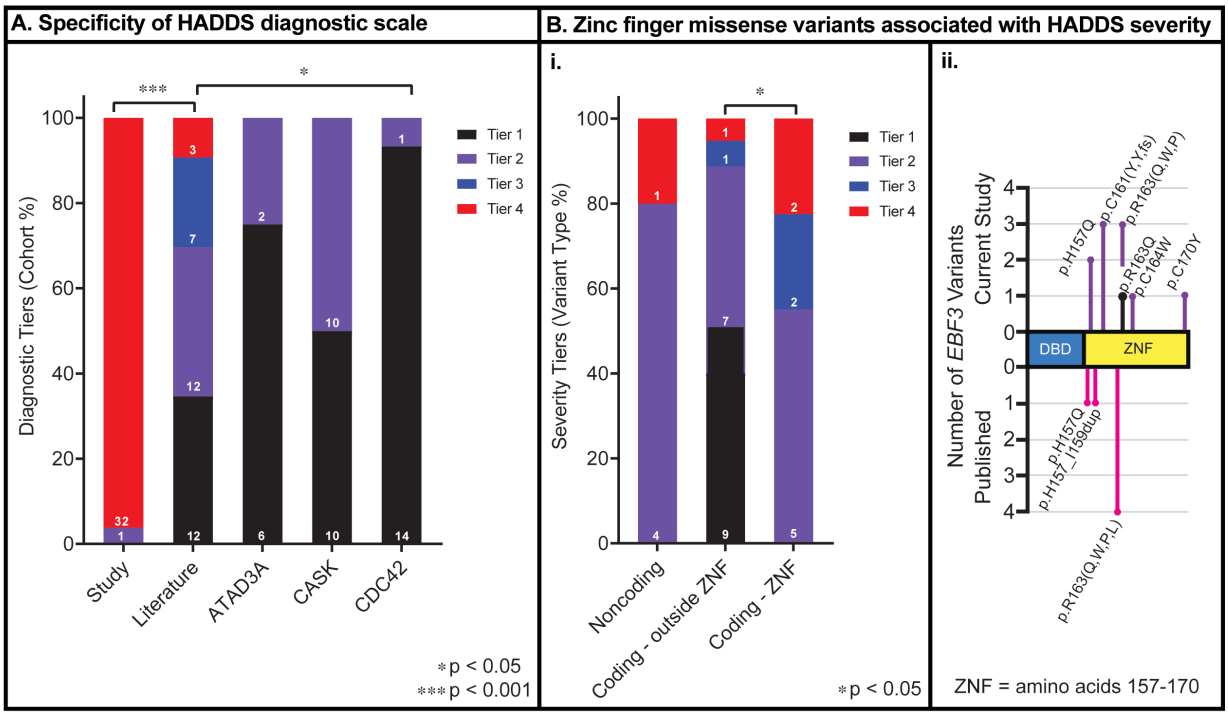
A. Specificity of HADDS diagnostic scale
B. Zinc finger missense variants associated with HADDS severity
i.
ii.
Diagnostic Tiers (Cohort %)
Severity Tiers (Variant Type %)
Tier 1
Tier 2
Tier 3
Tier 4
Study
Literature
ATAD3A
CASK
CDC42
Noncoding
Coding - outside ZNF
Coding - ZNF
*p < 0.05
***p < 0.001
Number of EBF3 Variants
Current Study
Published
DBD
ZNF
p.H157Q
p.C161(Y,Y,fs)
p.R163(Q,W,P)
p.R163Q
p.C164W
p.C170Y
p.H157_I159dup
p.R163(Q,W,P,L)
ZNF = amino acids 157-170

Table 1:

MAJOR FEATURE	Question (Yes or No)
Weak or absent cry	As a baby, does your child seem to rarely cry or cry weakly?
Decreased pain sensitivity	Does your child have a reduced response to painful stimuli, such as needle sticks or falls?
Hypomimia	Does your child have reduced facial expressions, such as difficulty with smiling?
Hypotonia	Does your child have low muscle tone or appear to be "floppy" or more "flexible" as a baby?
Ataxia	Does your child have trouble keeping their balance and coordinating their movements?
Speech delay	Is your child taking longer than 12-15 months to say their first word?
Motor delay	Is your child taking longer than expected to sit, stand, and/or walk without assistance?
MINOR FEATURE	**Question (Yes or No)**
Poor eye contact	Is your child less likely to look directly at another person's eyes?
Strabismus	Does your child's eyes have difficulty aligning with each other when looking at an object or person?
Repetitive scratching	Does your child exhibit repetitive scratching when tired or stressed and there is no clear evidence of another trigger (ie, insect bites, allergy)?
Repetitive stratching on back or neck	Does your child exhibit repetitive stratching on the back or neck?
Attraction to lights	Is your child drawn to bright lights?
Disturbed by loud noises	Does your child exhibit discomfort with loud noises?
Aversion to crowds	Does your child exhibit an aversion to crowded environments?
Constipation	Does your child have difficulty in emptying their bowels?
Frontal bossing	Does your child have a prominent forehead?
Recurrent urinary tract infections	Does your child have frequent urinary infections?
Difficulty feeding	As a baby, did you child have trouble nursing?
Cerebellar anomaly	Does your child have abnormal cerebellum findings on MRI brain imaging studies?
Unsteady, wide-based gait	Does your child have trouble keeping their balance while walking and walk with feet far apart?
Short stature	Is your child's height below expected based on parental heights?