Variants in *ZFX* are associated with an X-linked neurodevelopmental disorder with recurrent facial gestalt

Authors

James L. Shepherdson, Katie Hutchison, Dilan Wellalage Don, ..., Peggy J. Farnham, Cheol-Hee Kim, Marwan Shinawi

Correspondence

zebrakim@cnu.ac.kr (C.-H.K.), mshinawi@wustl.edu (M.S.)

We describe a cohort of 18 individuals with germline variants in *ZFX*, which encodes a transcription factor not previously associated with a human disease. In addition to identifying recurrent clinical features, molecular characterization of variants in cultured cells, *in silico* modeling, and a zebrafish model suggest potential modes of pathogenicity.

Shepherdson et al., 2024, The American Journal of Human Genetics *111*, 487–508
March 7, 2024

https://doi.org/10.1016/j.ajhg.2024.01.007

Variants in *ZFX* are associated with an X-linked neurodevelopmental disorder with recurrent facial gestalt

James L. Shepherdson,[1,46] Katie Hutchison,[2,46] Dilan Wellalage Don,[3,46] George McGillivray,[4,5] Tae-Ik Choi,[3] Carolyn A. Allan,[6] David J. Amor,[5,7] Siddharth Banka,[8,9] Donald G. Basel,[10] Laura D. Buch,[11] Deanna Alexis Carere,[12] Renée Carroll,[13] Jill Clayton-Smith,[14] Ali Crawford,[15] Morten Dunø,[16] Laurence Faivre,[17,18] Christopher P. Gilfillan,[19,20] Nina B. Gold,[21,22] Karen W. Gripp,[23] Emma Hobson,[24] Alexander M. Holtz,[25] A. Micheil Innes,[26] Bertrand Isidor,[27,28]

(Author list continued on next page)

Summary

Pathogenic variants in multiple genes on the X chromosome have been implicated in syndromic and non-syndromic intellectual disability disorders. *ZFX* on Xp22.11 encodes a transcription factor that has been linked to diverse processes including oncogenesis and development, but germline variants have not been characterized in association with disease. Here, we present clinical and molecular characterization of 18 individuals with germline *ZFX* variants. Exome or genome sequencing revealed 11 variants in 18 subjects (14 males and 4 females) from 16 unrelated families. Four missense variants were identified in 11 subjects, with seven truncation variants in the remaining individuals. Clinical findings included developmental delay/intellectual disability, behavioral abnormalities, hypotonia, and congenital anomalies. Overlapping and recurrent facial features were identified in all subjects, including thickening and medial broadening of eyebrows, variations in the shape of the face, external eye abnormalities, smooth and/or long philtrum, and ear abnormalities. Hyperparathyroidism was found in four families with missense variants, and enrichment of different tumor types was observed. In molecular studies, DNA-binding domain variants elicited differential expression of a small set of target genes relative to wild-type ZFX in cultured cells, suggesting a gain or loss of transcriptional activity. Additionally, a zebrafish model of *ZFX* loss displayed an altered behavioral phenotype, providing additional evidence for the functional significance of *ZFX*. Our clinical and experimental data support that variants in *ZFX* are associated with an X-linked intellectual disability syndrome characterized by a recurrent facial gestalt, neurocognitive and behavioral abnormalities, and an increased risk for congenital anomalies and hyperparathyroidism.

Introduction

Syndromic and non-syndromic X-linked intellectual disability has been associated with more than 160 genes on the X chromosome.[1–4] X-linked intellectual disability partially accounts for an excess of intellectual disability in males compared to females, and both dominant and recessive X-linked intellectual disability have the potential to cause disease in females as a result of skewed X-inactivation.[4,5] The higher frequency of intellectual disability-associated genes on the X chromosome as compared to the autosomes is supported by an enrichment of brain-expressed transcripts on the X chromosome.[4] Several of these genes are involved in transcriptional regulation, including the transcription factors (TFs) *ZNF711* (MIM: 314990) and *ARX* (MIM: 300382) and the chromatin modifier *MECP2* (MIM: 300005), demonstrating the potential for defects in transcriptional control to cause intellectual disability with concomitant pleomorphic phenotypes.[6]

ZFX (MIM: 314980), located in Xp22.11, encodes a conserved C2H2 zinc-finger TF that has not previously been reported in association with X-linked intellectual disability. ZFX is thought to primarily function by binding to CpG island promoter regions and activating the expression of many target genes.[7] Structurally, ZFX is composed of an amino-terminal transactivation domain followed by

[1]Medical Scientist Training Program, Washington University School of Medicine, St. Louis, MO, USA; [2]Department of Biochemistry and Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [3]Department of Biology, Chungnam National University, Daejeon 34134, Korea; [4]Victorian Clinical Genetics Services, Parkville, VIC 3052, Australia; [5]Murdoch Children's Research Institute, Parkville, VIC 3052, Australia; [6]Hudson Institute of Medical Research, Monash University, and Department of Endocrinology, Monash Health, Melbourne, Australia; [7]Department of Paediatrics, The University of Melbourne, Parkville 3052, VIC, Australia; [8]Division of Evolution, Infection and Genomics, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK; [9]Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Health Innovation Manchester, Manchester, UK; [10]Division of Genetics, Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI, USA; [11]Greenwood Genetic Center, Greenwood, SC, USA; [12]GeneDx, Gaithersburg, MD 20877, USA; [13]Adelaide Medical School and Robinson Research Institute, The University of Adelaide, Adelaide, SA, Australia; [14]Manchester Centre for Genomic Medicine, Manchester University NHS Foundation Trust, Manchester, UK; [15]Medical Genomics Research, Illumina Inc, San Diego, CA, USA; [16]Department of Clinical Genetics, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark; [17]Centre de Référence Anomalies du Développement et Syndromes Malformatifs, FHU

(Affiliations continued on next page)

Adam Jackson,[8,9] Panagiotis Katsonis,[29] Leila Amel Riazat Kesh,[24] Genomics England Research Consortium, Sébastien Küry,[27,28] François Lecoquierre,[30] Paul Lockhart,[5,7] Julien Maraval,[17,18] Naomichi Matsumoto,[31] Julie McCarrier,[10] Josephine McCarthy,[20] Noriko Miyake,[31,32] Lip Hen Moey,[33] Andrea H. Németh,[34,35] Elsebet Østergaard,[16,36] Rushina Patel,[37] Kate Pope,[5] Jennifer E. Posey,[29] Rhonda E. Schnur,[12] Marie Shaw,[13] Elliot Stolerman,[11] Julie P. Taylor,[15] Erin Wadman,[23] Emma Wakeling,[38] Susan M. White,[4,5,7] Lawrence C. Wong,[39] James R. Lupski,[29,40,41,42] Olivier Lichtarge,[29] Mark A. Corbett,[13] Jozef Gecz,[13,43] Charles M. Nicolet,[2] Peggy J. Farnham,[2,45] Cheol-Hee Kim,[3,45,*] and Marwan Shinawi[44,45,*]

13 zinc-finger domains, of which the final three are necessary and sufficient for recruitment to promoter regions.[8,9]

Initially thought to play a role in sex determination,[10–12] ZFX has more recently been investigated for its role in stem cell self-renewal and oncogenesis. ZFX is necessary for self-renewal of murine embryonic stem cells and the maintenance of murine adult hematopoietic stem cells.[13] In mouse models, loss of *Zfx* was shown to impair the development of basal cell carcinoma, medulloblastoma, acute myeloid leukemia, and acute T-lymphoblastic leukemia, and it has been hypothesized that ZFX may promote metaplastic transformation.[9,14,15] Male and female $Zfx^{-/-}$ mice displayed impaired embryonic growth and reduced adult body size and germ cell number.[16] Partial neonatal lethality was observed, primarily in males, with only 10% of hemizygous knockout (KO) males surviving to weaning.[16]

ZFX is part of a gene family that includes *ZFY* (MIM: 490000), located in Yp11.2 and demonstrating 96% overall sequence similarity to *ZFX*, and *ZNF711* (MIM: 314990), located in Xq21.1 and demonstrating 87% sequence similarity in the crucial DNA-contacting zinc-finger domains.[9] Of note, variants in *ZNF711* have previously been reported in individuals with X-linked intellectual disability (intellectual developmental disorder, X-linked 97; MIM: 300803).[17,18] Genome-wide binding patterns of ZFX, measured by chromatin immunoprecipitation and sequencing (ChIP-seq), show considerable overlap with binding of ZFY and ZNF711.[9] Cultured HEK293 cells lacking *ZFX*, *ZFY*, and *ZNF711* demonstrated impaired proliferation and transcriptome alterations relative to wild-type (WT) cells.[9]

Recurrent somatic mutations in the final zinc finger of ZFX, resulting in substitution of arginine 786 with glutamine (p.Arg786Gln [c.2357G>A]) or leucine (p.Arg786Leu [c.2357G>T]), were identified in 6 of 130 samples in a sporadic parathyroid adenoma sequencing study.[19] In the Catalog of Somatic Mutations in Cancer (COSMIC, https://cancer.sanger.ac.uk/cosmic), an additional 8 tumors with p.Arg786Gln missense variants are reported (COSMIC: COSV58447078): four endometrioid carcinomas, one melanoma, one acute lymphoblastic leukemia, one sarcoma, and one squamous cell carcinoma.

Germline variants in *ZFX*, however, remain largely uncharacterized. The NIH-sponsored ClinVar database of human genetic variation contains no nonsense or frameshift variants in the coding sequence of *ZFX*, and nine reported missense variants each with a single submitter (accessed October 6, 2023).[20] The p.Arg786Gln variant, in particular, has not been reported in the human germline (dbSNP: rs748417793).

Variants in TFs like *ZFX* have the potential to cause broad and pleomorphic functional effects through diverse mechanisms. Truncated protein variants that lack a DNA-binding domain or trigger protein surveillance mechanisms such as nonsense-mediated mRNA decay (NMD) can be predicted to cause disease through protein insufficiency/loss-of-function mechanisms.[21] Single-amino-acid variants in TF DNA binding domains have been shown to both abrogate and alter DNA binding specificity with the potential for both loss- and gain-of-function mechanisms, as well as cause dominant-negative interactions by interfering with DNA binding by other TFs.[22,23]

TRANSLAD, Hôpital d'Enfants, Dijon, France; [18]INSERM UMR1231, Equipe GAD, Université de Bourgogne-Franche Comté, 21000 Dijon, France; [19]Eastern Health Clinical School, Monash University, Melbourne, VIC, Australia; [20]Department of Endocrinology, Eastern Health, Box Hill Hospital, Melbourne, VIC, Australia; [21]Harvard Medical School, Boston, MA, USA; [22]Division of Medical Genetics and Metabolism, Massachusetts General Hospital, Boston, MA, USA; [23]Division of Medical Genetics, Nemours Children's Hospital, Wilmington, DE, USA; [24]Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust, Department of Clinical Genetics, Chapel Allerton Hospital, Leeds, UK; [25]Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA, USA; [26]Departments of Medical Genetics and Pediatrics and Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada; [27]Nantes Université, CHU Nantes, Service de Génétique Médicale, 44000 Nantes, France; [28]Nantes Université, CHU Nantes, CNRS, INSERM, l'institut du Thorax, 44000 Nantes, France; [29]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; [30]Univ Rouen Normandie, Inserm U1245 and CHU Rouen, Department of Genetics and Reference Center for Developmental Disorders, 76000 Rouen, France; [31]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [32]Department of Human Genetics, Research Institute, National Center for Global Health and Medicine, Tokyo 162-8655, Japan; [33]Department of Genetics, Penang General Hospital, George Town, Penang, Malaysia; [34]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK; [35]Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Foundation Trust, Oxford, UK; [36]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; [37]Medical Genetics, Kaiser Permanente Oakland Medical Center, Oakland, CA, USA; [38]North East Thames Regional Genetic Service, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK; [39]Medical Genetics, Kaiser Permanente Downey Medical Center, Downey, CA, USA; [40]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; [41]Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; [42]Texas Children's Hospital, Houston, TX, USA; [43]South Australian Health and Medical Research Institute, Adelaide, SA, Australia; [44]Division of Genetics and Genomic Medicine, Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA
[45]Senior authors
[46]These authors contributed equally
*Correspondence: zebrakim@cnu.ac.kr (C.-H.K.), mshinawi@wustl.edu (M.S.)
https://doi.org/10.1016/j.ajhg.2024.01.007.

Here, we present detailed clinical and molecular data on 18 individuals with germline variants in *ZFX*, in two subgroups: 7 with truncating (frameshift or nonsense) variants predicted to result in loss of function via NMD of the resulting transcript or truncation of the DNA-binding zinc fingers, and 11 with missense variants in and around the two most carboxy-terminal zinc fingers responsible for DNA binding specificity, including two individuals with a germline p.Arg786Gln variant. The great majority of individuals exhibited developmental delay/intellectual disability, hypotonia, and overlapping facial features. We also observed increased incidence of congenital anomalies and hyperparathyroidism. We used *in silico*, *in vitro*, and *in vivo* approaches to functionally characterize the observed variants. For a subset of missense variants mapping within the DNA-contacting zinc fingers, we characterized the effects on genome-wide DNA binding by ChIP-seq and observed transcriptional perturbation in variants by RNA-seq. We also modeled the group of variants with truncations causing predicted loss of function in a zebrafish *ZFX* deletion model and characterized behavioral alterations consistent with an intellectual disability phenotype.

Material and methods

Probands

Through a collaborative effort involving clinicians and researchers from multiple institutions, we identified 18 individuals, including 3 individuals from one family, with missense and truncating predicted deleterious variants in *ZFX*. The connection between all collaborators was facilitated using GeneMatcher and virtual meetings.[24] Several individuals were recruited through consortia and evaluated by their local medical geneticists, including the 100,000 Genomes Project, DECIPHER, and the Deciphering Developmental Disorders Study (details in Table S1).[25] Exome sequencing (ES) or genome sequencing (GS) for other subjects was ordered as part of their clinical diagnostic workup for developmental delay/intellectual disability, congenital anomalies, and/or dysmorphic facial features by local specialists, typically clinical geneticists or neurologists. The probands or their guardians signed informed consent for publication, which was approved by the Institutional Review Board of their respective institutions.

Sequencing

All individuals had either ES or GS performed under a mixture of both clinical and research protocols. For specific sequencing details, see Table S1 ("Genetic Testing") and the supplemental notes. Trio whole-genome sequencing was performed by Complete Genomics on a single individual from family 6 (proband 6B). Segregation of the coding c.2321A>G (p.Tyr774Cys) variant in *ZFX* was determined through gDNA extracted from blood of consented family members by Sanger sequencing (see supplemental notes; Figure 3). For proband 12, no causative variants were identified using panels, but subsequent panel-agnostic re-analysis using a previously described pipeline identified a *ZFX* variant (for details, see supplemental notes).[26–28] The sequencing methodology and variant interpretation criteria were based on the local protocols of each laboratory. All variants were classified by the performing laboratories as variants of uncertain significance (VUSs) as is required for a candidate gene. In light of the evidence presented in this study (although ACMG criteria are not applicable for a novel disease classification), we putatively reclassified the variants according to the applicable ACMG standards and guidelines.[29–31] Variant annotation was based on *ZFX* transcript GenBank: NM_003410.4. Sanger sequencing was used to confirm positive sequencing findings, except for probands 3 and 9. All ES and GS studies were performed as trios.

Evolutionary Action (EA)

We evaluated the impact of missense variants on protein function using the EA method.[32] The EA scores come from solving a formal equation that states that the impact of each variant equals the importance of the variant residue times the magnitude of the change. The importance of each residue is calculated according to the Evolutionary Trace scores[33,34] and the magnitude of change is calculated according to substitution odds. Objective assessments of the Critical Assessment of Genome Interpretation (CAGI) community have showed that EA performed consistently well among the state-of-the-art.[35,36]

For the evolutionary analysis of *ZFX*, we used 184 homologous sequences, obtained with BLASTp search[37] using the NM_003410 sequence and the databases "non-redundant NCBI," "UniRef100," and "UniRef90."[38] The sequences included orthologs from distant species, such as zebra finch (*Taeniopygia*) and pufferfish, and we aligned them using MUSCLE.[39] The ET scores were color-mapped on the 3D structure using PyMOL v.2.5.2 and the PyETV plugin.[40] The 3D structure of ZFX was generated using AlphaFold[41] and the sequence with Uniprot: P17010.

X-inactivation studies

For family 6, X-inactivation studies were performed following previously established protocols.[42] Variable number tandem repeat loci adjacent to *FMR1*, *AR*, and *RP2*, which are subject to DNA methylation on the inactive X, were amplified by PCR from genomic DNA extracted from blood. The relative quantities of amplicons from gDNA digested with *Hpa*II or untreated were compared by fragment analysis. Females with skewed X-inactivation will show greater than 90% loss of one allele where sizes are informative in the *Hpa*II digested sample. Primers used in the assay were as follows: *FMR1*, forward 5′-GCTCAGCTCCGT TTCGGTTTCACTTCCGGT-3′ and reverse 5′-[HEX]AGCCCCGCA CTTCCACCACCAGCTCCTCCA-3′; *AR*, forward 5′-TCCAGAATCT GTTCCAGAGCGTGC-3′ and reverse 5′-[FAM]GCTGTGAAGGTTG CTGTTCCTCAT-3′; *RP2*, forward 5′-[NED]TGACATAGCGAGACC CTGTG-3′ and reverse 5′-TGGTGGGTTCTCTAGCTGG-3′.

For proband 9, an X-inactivation study was performed by the Greenwood Diagnostic Laboratory (Greenwood, SC) using an androgen receptor X-inactivation assay.[43]

In vitro molecular characterization

Plasmid constructs

The plasmids described in this study were generated from the expression vector *ZFX* (GenBank: NM_003410) Human Tagged Open Reading Frame (ORF) Clone from Origene (Cat #RC214045). The parent vector for this construct is pCMV6-Entry (Origene #PS100001), in which expression is driven from the cytomegalovirus (CMV) promoter (this plasmid is also used as a negative control in relevant experiments). The same strategy, using two primers, was used to introduce each of the four single-nucleotide variants found in the cohort into the *ZFX* coding region. One primer contains the variant near the middle of the sequence, while the other primer amplifies in the other direction and abuts, but does not contain, the variant. For p.Arg786Gln and p.Arg764Trp,

two primer pairs were used (denoted X01-3′ and X01-5′ or X02_3′ and X02_5′, respectively); see Table S2 for a list of primers used for cloning. For p.Tyr774Cys and p.Thr771Met, different primers were used to incorporate the variants (X04-5′ and X05_5′, respectively), but the same primer was used to amplify in the other direction (X05_04_3′). Complementary nucleotides were added to the ends of the primers so that plasmid assembly could occur via Gibson assembly and not rely on blunt-end ligation. Following amplification using the reverse-oriented primers to generate the entire vector sequence and verification by gel electrophoresis, fragments were purified by Ampure magnetic beads (Beckman Coulter #A63881). Assembly was carried out using Gibson Assembly mix (NEB #2611L), then a portion of each reaction was transformed into chemically competent CopyCutter *E. coli* (Lucigen #C400CH10). Plasmids were isolated from transformants, and the presence of the introduced variants was verified by Sanger sequencing (Azenta Life Sciences service facility). Large-scale plasmid preparations were subsequently isolated by different column-based protocols, but all utilized the induction of plasmid copy number in the Lucigen cells (Lucigen #C400CH10) that is required for these toxic plasmids.

Cell culture

The cell line used in this study, designated DKO, was derived from (female) HEK293T cells (which lack a Y chromosome and thus *ZFY*) in which the endogenous *ZFX* and *ZNF711* genes were inactivated through CRISPR-mediated deletions.[9] Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (Thermo Fisher #10437036) plus 1% penicillin and 1% streptomycin at 37°C with 5% CO_2. Cell lines were authenticated via the short tandem repeat (STR) method and validated to be mycoplasma free using a universal mycoplasma detection kit (ATCC #30-1012K).

Transactivation assays

To test the activity of the WT and variant ZFX proteins, a transient transfection assay using the *ZFX* expression vectors was performed using the DKO cell line. Transfections for RNA preparation were carried out using Lipofectamine 3000 (Thermo Fisher #L3000015), according to the manufacturer's instructions. Much larger cell numbers were required for ChIP assays; the same transfection protocol was used but scaled appropriately. All transfections were performed in triplicate wells.

RNA preparation and RNA-seq

To recover RNA for expression analyses, cells were lysed 24 h after transfection in triplicate wells using TRI Reagent (Zymo #R2050-1-200), and RNA was recovered by isopropanol precipitation. RNA integrity was confirmed using a Bioanalyzer (Agilent) using an RNA integrity number (RIN) of 9.0 as a quality cutoff for library preparation. Samples were submitted to Novogene for their standard RNA-seq service (150 bp paired-end reads with a total minimum output of 6 Gb). Efficacy of transactivation prior to library preparation and sequencing was monitored by quantitative reverse-transcription PCR (RT-qPCR) of known responsive target genes, as described.[9] All RNA-seq assays were performed in triplicate (see Table S3A).

Chromatin preparation, immunoprecipitation, and ChIP-seq

To obtain chromatin for ChIP-seq assays, cells were split 1:2 24 h after transfection and allowed an extra day of growth. After this additional 24 h, chromatin crosslinking and chromatin immunoprecipitation was carried out using established lab protocols.[7,9] Briefly, cells were crosslinked in 1% formaldehyde for 10 min before quenching with 125 mM (final) glycine. Following washing in phosphate-buffered saline (PBS), crosslinked cells were swollen in cell lysis buffer (5 mM PIPES [pH 8.0], 85 mM KCl, 1% NP-40), then lysed by sonication in nuclei lysis buffer (0.1% sodium dodecyl sulfate [SDS], 50 mM Tris-HCl [pH 8], 10 mM ethylenediaminetetraacetic acid [EDTA]). All buffers included cOmplete Protease Inhibitor EDTA-free tablets (Roche #11836153001). Lysis volumes were based on cell pellet volumes. Sonication was carried out with a Diagenode Bioruptor Pico using conditions to generate fragments 400–800 bp in length. Chromatin and antibody amounts vary depending on the experiment; for the anti-FLAG ChIP-seq experiments, 100 μg soluble chromatin was diluted 1:5 in radioimmunoprecipitation assay buffer (RIPA, 150 mM NaCl, 50 mM Tris-HCl [pH 8.0], 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) and incubated overnight with 12.5 μL of anti-FLAG antibody (Cell Signaling Technology #5419S). Magnetic Protein A/G beads (Pierce #88803) were added for an additional 2 h of incubation. Bead-immunocomplexes were then washed twice with RIPA (150 mM NaCl, 50 mM Tris-HCl [pH 8.0], 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS), followed by washing three times with IP Wash Buffer II (100 mM Tris-Cl [pH 9.0], 500 mM LiCl, 1% NP-40, 1% sodium deoxycholate). Elution was performed in 150 μL of elution buffer (50 mM $NaHCO_3$, 1% SDS), then ChIP samples and inputs (10 μL of precleared chromatin lysis plus 140 μL elution buffer) were reverse crosslinked overnight at 65°C. DNA was purified using Qiagen MinElute PCR clean up columns (cat #28006) and quantified spectrophotometrically. ChIP-seq libraries were prepared using the KAPA HyperPrep kit (Roche #KK8503) following the manufacturer's protocol. Final cycle amplification numbers varied depending on input DNA and ChIP parameters. Libraries were quality checked by qPCR for target enrichment and visually by Bioanalyzer, then submitted to Novogene and sequenced on a Novaseq 6000, providing a minimum of 10 Gb of data in paired-end 150-bp format. All ChIP-seq assays were performed in duplicate (see Table S3B). Datasets from the two independent experiments were merged for downstream analyses.

Data analysis and visualization

RNA-seq. All RNA-seq analyses were carried out using Partek Flow modules (v.10.0.21.1103 at project initiation), as provided through the Norris Medical Library Bioinformatics Service. Following quality trimming, fastq files provided by Novogene were aligned using STAR v.2.7.8a with the default alignment parameters provided. Transcript and gene counts were generated using hg38 GENCODE Genes release 36. Gene counts were then analyzed for differential expression by DESeq2(R) 3.5. Default parameters were used except that fit type was set to "local" and not "parametric."

ChIP-seq. All ChIP-seq data were aligned to hg38 using Bowtie2 software (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). Only reads that aligned to a unique position in the genome with no more than two sequence mismatches were retained for further analysis. Duplicate reads that mapped to the same exact location in the genome were counted only once to reduce clonal amplification effects. Normalization was done across samples using an equal number of uniquely mapped reads. Biological replicates of ChIP-seq datasets for WT and each mutant were performed. ChIP-seq peaks were called using the narrowPeak setting in MACS2 (https://github.com/taoliu/MACS), using the ENCODE pipeline and Irreproducible Discovery Rate (IDR) method to determine reproducible peaks. Peaks were annotated using HOMER v.4.11 (http://homer.ucsd.edu/homer/), and average tag density plots were generated using CEAS software (https://liulab-dfci.github.io/software/).

Proportional Venn diagrams were generated using the web tool DeepVenn.[44]

In vivo characterization in zebrafish

Zebrafish husbandry

Animal experiments were conducted in accordance with protocols approved by the Animal Ethics Committee of Chungnam National University (202012A-CNU-170). Adult fish were reared under standard conditions with a 14 h/10 h light/dark cycle. We obtained embryos by natural mating of heterozygous adult zebrafish, and embryos were reared in egg water at 28.5°C. WT and KO zebrafish were obtained from the Zebrafish Center for Disease Modeling. All fish used in behavioral tests were fully grown, sexually mature, 3- to 12-month-old fish, ranging in size between 3.0 and 3.5 cm standard length.

Zebrafish generation

In zebrafish, *zfx* on chromosome 24 is the ortholog of human *ZFX* (Figure S2); we aligned the human ZFX (from transcript GenBank: NM_003410.4) and zebrafish zfx (Ensembl: ENSDART00000110652.4) protein sequences via the EMBOSS Needle tool and found 65.5% similarity and 53.6% identity. To understand the *in vivo* role of ZFX, we established a KO zebrafish model utilizing the CRISPR-Cas9 system.[45] CRISPR single guide (sg) RNAs targeting *zfx* were identified using CRISPRScan (https://www.crisprscan.org/), and oligonucleotides were selected (Table S7). *In vitro* transcription was carried out using 150–200 ng of template and the MaxiScript T7 Kit (Ambion #AM1312). RNA was precipitated with isopropanol. Cas9 expression vector (Addgene #46757) was linearized with XbaI (NEB #R0145) and purified with an agarose gel DNA extraction kit (ELPIS #EBD-1009). Cas9 mRNA was transcribed with the mMESSAGE mMACHINE T3 Kit (Ambion #AM1348) and then purified by lithium chloride precipitation following the manufacturer's protocol. One-cell-stage zebrafish embryos were injected with 300 ng/μL Cas9 mRNA and 150 ng/μL sgRNA. For genotyping F0, PCR products (20 μL) were re-annealed in a thermal cycler under the following conditions: 95°C for 2 min, 95°C–85°C at 2°C/s, 85°C–25°C at 0.1°C/s, then kept at 4°C. Part (16 μL) of the re-annealed mixture was incubated with 0.2 μL of T7 endonuclease I, 2 μL of NEB Buffer 2, and 1.8 μL of nuclease-free water at 37°C for 40 min. See Table S7 for genotyping primers.

Human ZFX variant over-expression in zebrafish

Human WT *ZFX* cDNA (GenBank: NM_003410.4) was subcloned into the pCS2+ expression vector. Missense variants were introduced by site-directed mutagenesis using WT *ZFX* pCS2+ vector as template. *ZFX* mRNA was transcribed with the mMESSAGE mMACHINE SP6 Transcription Kit (Ambion #AM1340) and then purified by lithium chloride following the manufacturer's protocol. For over-expression, three different concentrations (100, 200, and 400 pg) of mRNAs were microinjected into 1- to 2-cell-stage zebrafish embryos.

Quantitative real-time PCR

Total RNA was isolated from brains of adult *zfx* KO and WT zebrafish via easy-Blue Total RNA isolation kit (iNtRON Biotechnology #17061). One microgram of RNA was converted to cDNA by SuperScript III First-Strand Synthesis System (Invitrogen #18080-51). RT-qPCR was conducted in triplicate using TB Green Premix Ex Taq II master mix (Takara #RR82LR) and a Thermal Cycler Dice Real Time System III (Takara, Japan). β-actin was used as a reference gene, and relative gene expression levels were calculated by the $2^{-}\Delta^{Ct}$ method.

Behavioral tests

Novel tank assay. A novel tank test was performed as previously described in Kim et al.[46] to evaluate swimming activity and anxiety-like behavior. Each male WT (n = 8) or KO (n = 17) zebrafish was introduced to a tank (24 × 15 × 15 cm) filled with system water (up to 10 cm) and recorded with a video camera (Sony HDR-CX190) for 10 min. The videos were analyzed using EthoVision XT software. To examine exploratory activity, the tank was divided into three equally sized zones (top, middle, and bottom) as shown in Figure 6A. The percentage of average time spent in each zone and distance moved in first 5 min were analyzed.

Scototaxis test. To assess anti-anxiety behavior, we evaluated scototaxis (dark/light preference) as previously described in Maximino et al.[47] For this, we modified the experimental setup as shown in Figure 6C by dividing the light zone into gray and bright zones using an external light source underneath the tank. Each zebrafish was placed in the tank followed by a 5-min habituation period. Swimming activity was recorded from above for 10 min, and the percentage of average time spent in each region was analyzed for WT (n = 8) and KO (n = 10) zebrafish.

Adult startle tap test. To elicit an unconditioned startle response to a stimulus, a tap startle test was performed as described in Eddins et al.[48] with adult zebrafish. We used only three tap stimuli, as the WT fish showed greater adaptation. Each fish was placed in a round container (95 mm in diameter and 40 mm in height) filled with system water to the height of 25 mm. The round arena was then placed in the observation chamber under bright light conditions (mimicking daylight). Fish were left for a 10-min acclimatization time, allowing for stable swimming bouts. After the acclimatization period, a sequence of three taps (stimulus intensity level at 5) were evoked at 1-min intervals (Figure 6F). The average velocity for 5 s before and after each tap was analyzed using DanioVision. The post-tap velocity data of each WT (n = 11) and KO (n = 17) zebrafish were normalized for the analysis.

Statistical analysis

All results were expressed as mean ± SEM. Statistical analysis was performed using GraphPad Prism software (v.8.01 for Windows, GraphPad Software Inc., USA). In all experiments, comparisons between WT and KO fish were done using a one-tailed Student's t test. The effect size estimations were calculated using a web application (https://www.estimationstats.com/).[49] The estimates were performed using 5,000 bootstraps, and the confidence interval was bias-corrected and accelerated. Group differences are normalized to Hedges' g. Statistical significance is shown as follows: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Results

We identified and recruited 18 individuals from 16 unrelated families for this study. Probands 6A, 6B, and 6C are relatives from the same family. Fourteen individuals were assigned male at birth (78%), while four were assigned female, with one self-identifying as male (referred to by their assigned sex throughout this manuscript). The mean age of subjects in our cohort was 16 years (range: 8–34 years). A full description of demographic, phenotypic, and genetic data is presented in Table S1, and a summary of the key findings in our cohort is found in Table 1 with additional details provided in the clinical findings section below.

Table 1. Summary of demographic and clinical features in individuals with *ZFX* variants

Demographic and clinical features	Frequency[a] (%)		
	Missense	Frameshift	Total
Sex	males: 8/11 (73)	males: 6/7 (86)	males: 14/18 (78)
Mean Age (±SD)	16.04 (±7.95)	12.24 (±5.7)	14.52 (±7.2)
De novo inheritance	6/11 (55)	4/7 (57)	10/18 (56)
Development and neurobehavioral			
Developmental delay/intellectual disability	10/11 (91)	7/7 (100)	17/18 (94)
Gross motor delay	9/11 (82)	7/7 (100)	16/18 (89)
Fine motor delay	9/11 (82)	7/7 (100)	16/18 (89)
Speech delay	10/11 (91)	6/7 (86)	16/18 (89)
ADHD	5/11 (45)	1/7 (14)	6/18 (33)
Autism	5/11 (45)	1/7 (14)	6/18 (33)
Autistic traits[b]	1/11 (9)	2/7 (29)	3/18 (17)
Other behavioral abnormalities	9/11 (82)	1/7 (14)	10/18 (56)
Therapeutic services or special education	10/11 (91)	7/7 (100)	17/18 (94)
Neurological			
Hypotonia	9/11 (82)	4/7 (57)	13/18 (72)
Abnormal brain MRI	4/7 (57)	6/6 (100)	10/13 (77)
Sleep problems	6/11 (55)	2/7 (29)	8/18 (44)
Epilepsy	1/11 (9)	1/7 (14)	2/18 (11)
Growth parameters			
Height $\leq$ 10%ile	4/11 (36)	4/7 (57)	7/18 (39)
OFC[c] > 90%ile	4/10 (40)	2/6 (33)	6/16 (38)
Vision and hearing			
Vision/eye abnormalities	7/11 (64)	6/7 (86)	13/18 (72)
Hearing loss	7/11 (64)	2/7 (29)	9/18 (50)
Dysmorphic features			
Broad/thick eyebrows ± synophrys	9/11 (82)	6/7 (86)	15/18 (83)
External eye findings	9/11 (82)	6/7 (86)	15/18 (83)
Other face abnormalities	10/11 (91)	5/7 (71)	15/18 (83)
Smooth ± long philtrum	8/11 (72)	6/7 (83)	14/18 (78)
Ear abnormalities	10/11 (91)	4/7 (57)	14/18 (78)
Thin upper lip	10/11 (91)	3/7 (43)	13/18 (72)
Nose abnormalities	6/11(55)	6/7 (86)	12/18 (67)
Forehead abnormalities	6/11 (55)	4/7 (57)	10/18 (56)
Macroglossia	7/11 (64)	1/7 (14)	8/18 (44)
Miscellaneous			
Musculoskeletal findings	10/11 (91)	7/7 (100)	17/18 (94)
Inguinal hernia	9/11 (82)	4/7 (57)	13/18 (72)
Umbilical hernia	8/11 (72)	1/7 (14)	9/18 (50)
Genitourinary congenital anomalies (males)	7/8 (88)	3/6 (50)	10/14 (71)
Abnormal echocardiography	4/9 (44)	4/5 (80)	8/14 (57)

(*Continued on next page*)

Table 1. ***Continued***

Demographic and clinical features	Frequency[a] (%)		
Gastrointestinal abnormalities	10/11 (91)	2/7 (29)	12/18 (67)
Abnormal renal ultrasound	5/10 (50)	4/7 (57)	9/17 (53)
Hyperparathyroidism	3/7 (43)	0/5 (0)	3/12 (25)

[a]Denominator can vary based on data availability
[b]Autistic traits but no formal diagnosis with autism
[c]OFC, occipital frontal circumference

Variant characterization

Four hemizygous or heterozygous *ZFX* variants were identified in 8 male and 3 female subjects: c.2290C>T (p.Arg764Trp), c.2312C>T (p.Thr771Met), c.2321A>G (p.Tyr774Cys), and c.2357G>A (p.Arg786Gln) (see Tables 2 and S1). In addition, 7 unrelated subjects (6 males and 1 female) were found to have truncating variants delineated in Table 2. None of these *ZFX* variants have been observed in the Genome Aggregation Database (gnomAD, last accessed on October 6, 2023).[50]

Based on *in silico* analysis, the truncation variants lead to the creation of nonsense codons at or shortly downstream of the genetic change and are predicted to cause loss of function either through NMD or truncation of the essential DNA-binding zinc fingers (ZF11–ZF13). The results of *in silico* variant effect prediction tools for the missense variants are shown in Table 2. Three missense variants were maternally inherited in 5 subjects (including the 3 related subjects), and 6 subjects had *de novo* missense variants, while the truncating variants were *de novo* in 4 subjects and maternally inherited in 3 subjects. All missense variants were located within or between zinc fingers 12 and 13 in the ZFX protein (Figure 1).

We analyzed the *ZFX* variants using the EA method (Figure 1B).[32] Briefly, EA uses sequence homology and phylogenetic distances to prioritize the variants in each protein and score them on a scale from 100 (pathogenic) to 0 (benign). The missense sequence variants p.Arg764Trp, p.Thr771Met, p.Tyr774Cys, and p.Arg786Gln had EA scores of 56.77, 72.08, 72.47, and 54.03, respectively. These EA scores indicate intermediate to high impact on *ZFX* function, and they are larger than the average EA score that corresponds to missense variants due to random nucleotide changes (EA score of 42.27).

X-Inactivation

X-inactivation studies on gDNA extracted from blood were available for all informative females in family 6 (discussed below and in Figure 3C) and proband 9. Extreme skewing was detected in both carrier females in family 6 (p.Tyr774Cys) and in proband 9 (p.Arg786Gln), while random inactivation was observed in a non-carrier female in family 6. Early studies of *ZFX* reported it to escape X-inactivation, and recent single-cell RNA sequencing experiments have supported this claim.[11,55]

Clinical findings

Prenatal findings and growth

In this cohort, nine individuals (50%) were born via Cesarean section and two via induced vaginal delivery. The indications for these types of deliveries included fetal distress, pre-eclampsia, pregnancy-induced hypertension, fetal growth restriction, and oligohydramnios. Six of the participants (33%) were born prematurely (33–36 weeks of gestation). Furthermore, 4 subjects (22%) were born small for gestational age (defined as birth weight $<10^{th}$ percentile for gestational age) and 5 subjects (28%) had birth weight or length equal or above the 90^{th} percentile, which along with additional clinical and physical findings prompted molecular testing for Beckwith-Wiedemann syndrome in 4 of these subjects.

Postnatal growth parameters showed a trend for shorter stature in 44% of subjects; 4/7 (57%) subjects with truncating variants as compared to 4 individuals (36%) with missense variants had height equal to or below the 10^{th} percentile for age. Proband 14 receives growth hormone treatment, and his height on last assessment was within normal range, and proband 15 had borderline short stature with height at the 11^{th} percentile. In addition, large head size (occipitofrontal circumference $>90^{th}$ percentile) was found in 38% of probands with available head-circumference measurements.

General findings

Nine subjects (50%) were found to have conductive or sensorineural hearing loss. Six subjects from the missense variant group required placement of ear tubes for recurrent ear infections and/or chronic serous otitis media, one subject from the truncation group was reported to have hyperacusis and chronic middle ear effusion, and another required ear tube placement.

Seven of the 11 subjects (64%) with missense variants and 6 probands with truncation variants (86%) had eye or vision abnormalities. These abnormalities included refractive errors, strabismus, astigmatism, nystagmus, optic nerve hypoplasia, and retinal detachment.

Neurodevelopmental and behavioral findings

All male probands in our cohort had developmental delay, intellectual disability, or learning disability; their intellectual disability was with variable severity ranging between borderline (low normal) to moderate. Two of the four females in this cohort were described as very intelligent, one had normal cognition (with relative weaknesses in

Table 2. Table of *ZFX* variants

Proband(s)	DNA (GenBank: NC_000023.11)	cDNA (GenBank: NM_003410.4)	Protein (GenBank: NP_003401.2)	CADD score	REVEL score	NMD predicted?	Putative applicable ACMG criteria[a]	Consequence	Domain
1, 2	g.24211248C>T	c.(2290C>T)	p.Arg764Trp	25	0.3	–	PM2, PS2[(2)], PP3	missense	ZF12
3, 4, 5	g.24211270C>T	c.(2312C>T)	p.Thr771Met	24.6	0.46	–	PM2, PS2, PP3	missense	ZF12-ZF13 linker
6A–6C, 7	g.24211279A>G	c.(2321A>G)	p.Tyr774Cys	27.3	0.57	–	PM2, PS2[(7)], PP3	missense	ZF12-ZF13 linker
8, 9	g.24211315G>A	c.(2357G>A)	p.Arg786Gln	24.6	0.34	–	PM2, PS2[(9)], PP3	missense	ZF13
10	g.24207442A>AT	c.(768dup)	p.Lys257*	29.3	–	yes	PVS1, PM2	truncation	acidic domain
11	g.24210271G>GT	c.(1319dup)	p.Leu440Phefs*21	32	–	no	PVS1_Strong, PS2, PM2	truncation	ZF1
12	g.24209008G>GGA	c.(1205_1206dup)	p.Arg403Glufs*12	28.2	–	no	PVS1_Strong, PS2, PM2	truncation	acidic domain
13	g.24210953CAT>C	c.(1996_1997del)	p.Met666Valfs*2	32	–	no	PVS1_Strong, PS2, PM2	truncation	ZF9
14	g.24179650G>GT	c.(529dup)	p.Ser177Phefs*12	27	–	yes	PVS1, PS2, PM2	truncation	acidic domain
15	g.24179543ATG>A	c.(423_424del)	p.Ser142*	25	–	yes	PVS1, PM2	truncation	acidic domain
16	g.24179237CTG>C	c.(115_116del)	p.Val39Phefs*14	26.2	–	yes	PVS1, PM2	truncation	N-terminal region

CADD[51] and REVEL[52] scores were calculated for the missense variants, and putative applicable American College of Medical Genetics (ACMG) criteria are listed.[29–31]

[a]For criteria that only apply to a subset of probands in each row, the applicable individuals(s) are indicated by superscript. PP3 was applied on the basis of CADD score. According to current ACMG criteria, PVS1 is applied for a gene where loss-of-function is a known mechanism of disease. However, PVS1 was used to upgrade the classification of the truncating variants based on evidence provided in this study.

visuo-motor coordination and executive functioning skills related to planning and decision-making), and one had borderline cognitive abilities. All subjects with truncating variants (including one female proband) and all male subjects as well as one female subject with missense variants had gross and fine motor delay. Of note, the other two female subjects with missense variants had normal developmental motor skills. In addition, with the exception of one female subject (proband 4), all probands with missense variants and six subjects with truncation variants had speech delay. All subjects, except one female proband (proband 4) with a missense variant, required therapeutic services, including physical, occupation, and speech therapies, with or without special education at school.

Neurobehavioral problems were also prevalent. A total of 50% of the probands were either diagnosed with autism spectrum disorder (ASD) (6/18) or reported to have autistic traits (3/18), and approximately 33% (6/18) of individuals had a diagnosis of attention-deficit/hyperactivity disorder (ADHD). Three individuals (17%) had ADHD as well as ASD. Other behavioral problems, such as anger and tantrums, were more common among individuals with missense variants than in those with truncating variants (82% vs. 14%; $p = 0.006$). Six participants (55%) with missense variants and two subjects (29%) with truncating variants had sleep problems or difficulty.

Hypotonia was found in 13/18 (72%) of the subjects, which may have played a role in the motor delay that was observed in these individuals. However, the hypotonia in two of the individuals was mild and resolved in late infancy/early childhood. Epilepsy was reported in two subjects only.

Abnormal but nonspecific brain magnetic resonance imaging findings, including cerebral atrophy, arachnoid cysts, delayed myelination, corpus callosum abnormalities, cerebellar atrophy, or pituitary abnormalities, were found in 10 out of 13 subjects (77%) who underwent brain imaging as part of their workup.

Facial features

Most cohort members exhibited recurrent facial features with an overlapping facial gestalt in a subset of individuals (Figures 2A, 2B, and 3A). Several individuals were described as having coarse facial features. The most prominent feature found in 15/18 individuals was broadening and/or thickening of eyebrows especially on the medial side with or without synophrys. Fifteen individuals (83%) had differences in the shape of the face: 7 individuals had pointed chins, 5 had long faces, and 2 individuals had midface hypoplasia. External eye findings were seen in 15/18 individuals, including 6 individuals with epicanthal folds and downslanting palpebral fissures in 8 individuals. Smooth and, in a subset of individuals, long philtrum was found in 14/18 individuals and thin upper lip in 13/18 individuals; the latter was mostly among individuals with a missense variant. Sixty-four percent (7/11) of individuals with missense variants had macroglossia, 3 of whom required tongue reduction surgery. Only one individual with a truncation variant had macroglossia ($p = 0.04$), and another had tongue protrusion. Forehead abnormalities were reported in 10/18 and included broad forehead, frontal bossing, and metopic

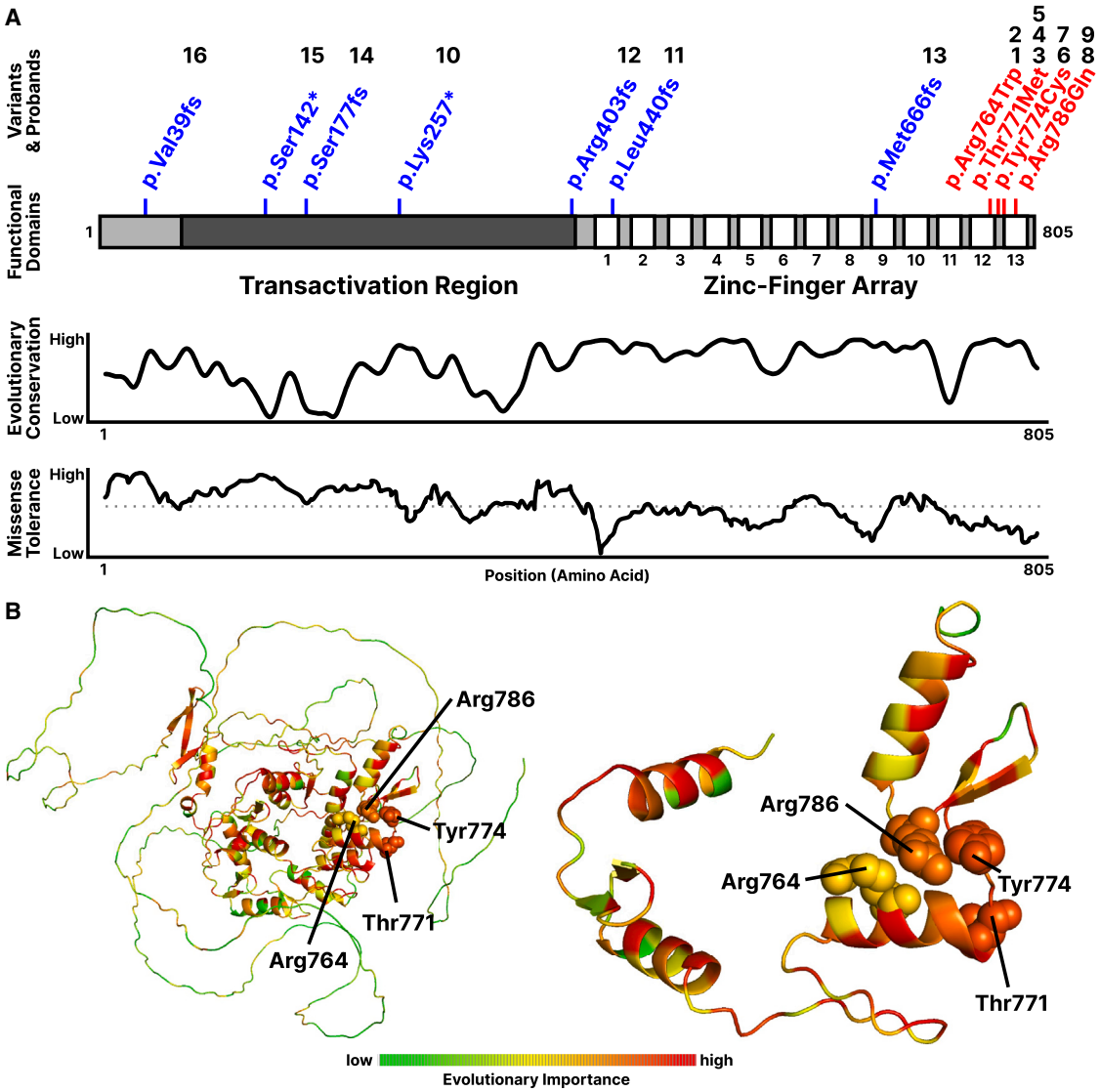


Figure 1. Cohort variants and existing characterization of ZFX

(A) Distribution of cohort variants throughout the *ZFX* coding sequence, with missense variants in red and truncating variants in blue (numbers indicate the corresponding probands). Evolutionary sequence conservation shown for *ZFX* across 19 vertebrate species.[53] Missense tolerance ratios (MTRs) are calculated from gnomAD v2.0 exomes; dotted line indicates 50th percentile for MTR.[54].

(B) Evolutionary Action scores superimposed on a predicted ZFX structure, with residues colored proportional to evolutionary importance.

ridging. Sixty-seven percent of individuals (12/18) had nasal anomalies: 5 had depressed nasal bridges, 8 had a bulbous/wide nasal tip, and 5 had hanging/prominent columella. Seventy-eight percent of individuals (14/18) had ear anomalies including abnormalities in size, helices, low set, posteriorly rotated, and prominent ears. A small subset of individuals had blond hair, but this trait was not well documented in other family members and we do not consider it at this time as one of the physical characteristics of this condition.

Skeletal abnormalities were reported in all individuals except one. The most prominent skeletal findings were found in the hands (15 individuals) and feet (10 individuals) (Figure 2C). Hand findings included the shape of distal phalanges, abnormal shape and length of fingers, abnormal creases in hands, restriction or hypermobility in interphalangeal joints, and deep-seated fingernails. Feet findings included deep-seated toenails, abnormal or deep creases in soles, hallux valgus deformity, proximal insertion of a toe (1 individual) and polydactyly (1 individual). Joint hypermobility was found in 6 individuals with missense variants, one of whom met the diagnostic clinical criteria for Ehler Danlos syndrome-hypermobility type (proband 9). Three individuals with missense variants and 2 individuals with truncating variants had restriction in elbow extension. Pectus excavatum or carinatum was reported in 4 individuals. Spine anomalies were found in 6 individuals including 5 with scoliosis (all with missense variants) and sacral dysgenesis (1).

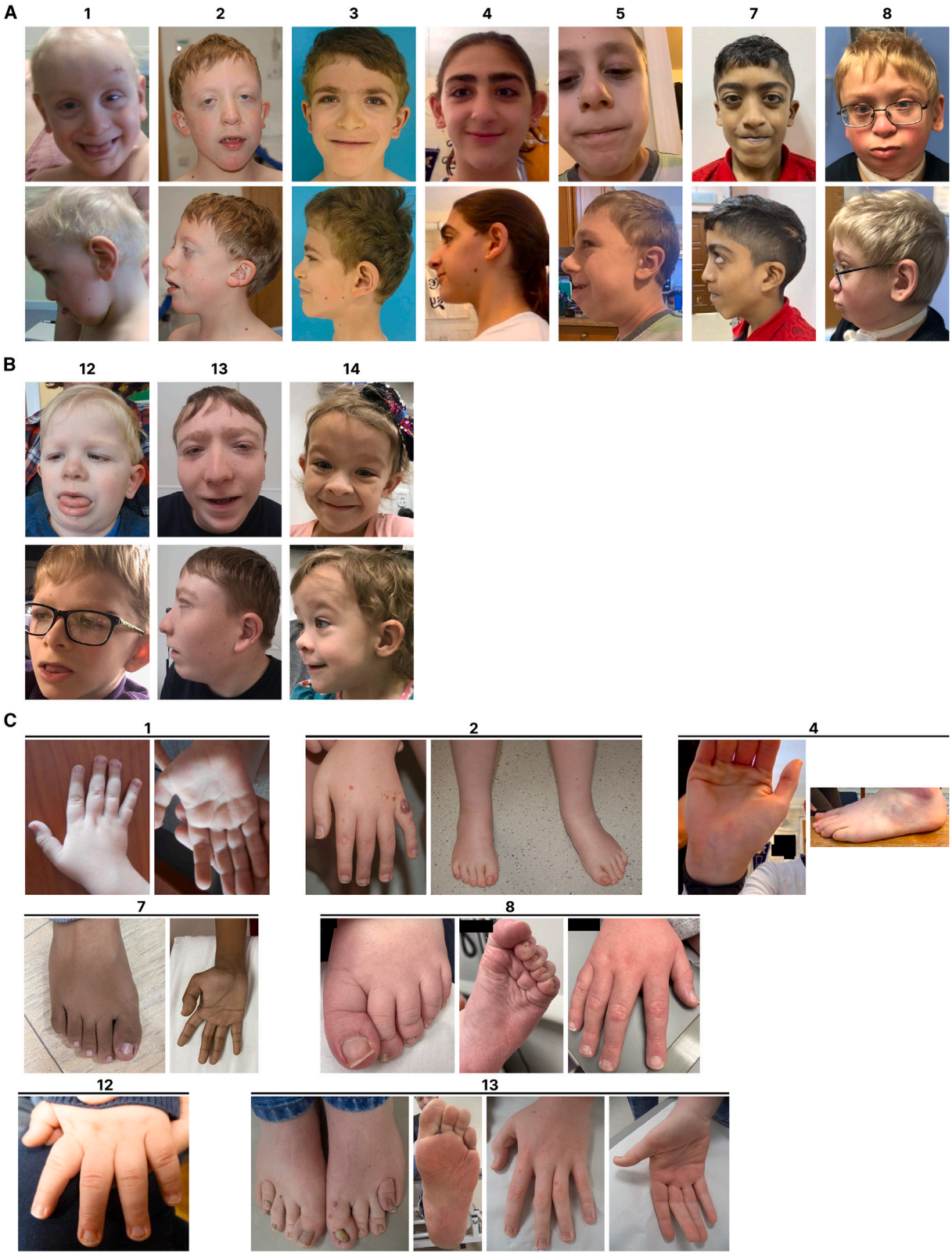
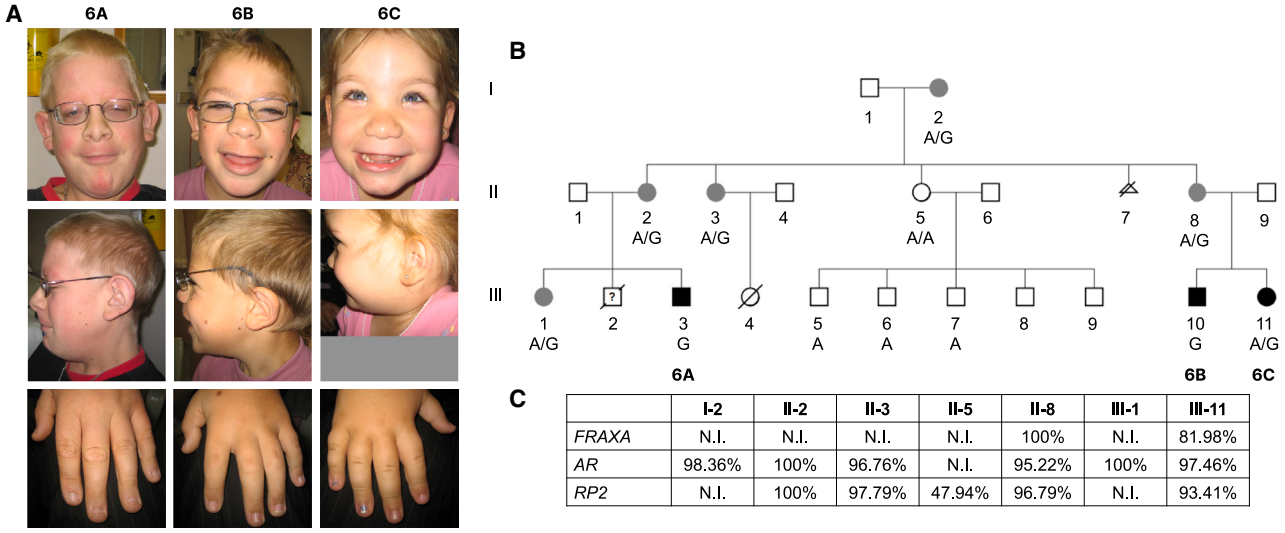


Figure 2. Facial features of the presented individuals with *ZFX* variants
(A) Subjects with missense variants in the *ZFX* DNA binding domain.
(B) Subjects with truncating *ZFX* variants.
(C) Extremities of the indicated subjects. See supplemental information and Table S1 for additional details; images not available for all individuals.



	I-2	II-2	II-3	II-5	II-8	III-1	III-11
FRAXA	N.I.	N.I.	N.I.	N.I.	100%	N.I.	81.98%
AR	98.36%	100%	96.76%	N.I.	95.22%	100%	97.46%
RP2	N.I.	100%	97.79%	47.94%	96.79%	N.I.	93.41%

Figure 3. Characterization of a family with an inherited *ZFX* variant
(A) Facial features and extremities of probands 6A–6C (see text and supplemental information for additional details).
(B) Three-generation pedigree of probands 6A–6C and family members. Dark black circles and squares indicate affected individuals; gray circles indicate carrier females diagnosed with hyperparathyroidism (except for III-1). A is the wild-type *ZFX* allele; G is the variant *ZFX* allele at same position (GRCh38 chrX: 24229396A>G, c.(2438A>G), p.Tyr774Cys).
(C) Results of X-inactivation studies showing skewing in all carrier females and random inactivation in a noncarrier female (II-5).

Other less common findings included skull abnormalities (5/18), high arched palate (4/18), micro/retrognathia (3/17), wide mouth (3/18), and short neck (2/18).

Congenital anomalies

Fifty-seven percent of individuals (8/14) who underwent echocardiography had abnormal findings. Four individuals with missense variants had variable structural cardiac abnormalities, while 4 individuals with truncation variants had atrial/ventricular septal defects and/or patent ductus arteriosus that spontaneously closed. Of note, probands 2 and 7 had a dilated aortic root, and proband 7 required a beta blocker to stop the progression.

A renal ultrasound was performed for all but one individual, and an abnormality was documented in 9 individuals (53%), but the majority of the findings were mild or transient, such as hydronephrosis or pelvicaliectasis. Interestingly, one individual had renal calculi and two had nephrocalcinosis.

Genitourinary anomalies were found in 69% of males, 7/8 males with missense variants (hypospadias in 6, cryptorchidism/retractable testes in 4) and in 3/6 males with truncating variants (hypospadias in 2, cryptorchidism in 1).

Seventy-eight percent of individuals had inguinal and/or umbilical hernia. Nine of 11 individuals with missense variants had inguinal hernias, seven of whom also had umbilical hernia, and one individual had epigastric, hiatal, and umbilical hernia. Four of 7 of the subjects with truncating variants had inguinal hernia, one of whom also had umbilical hernia.

Other rare congenital anomalies among individuals with *ZFX* variants included palate, dental, and upper airway anomalies. Cleft palate was found in two individuals with missense variants. Six subjects (33% of individuals) had dental anomalies: four had supernumerary teeth (one with delayed eruption), one had retained baby teeth, and one had crowded teeth. One individual was tracheostomy dependent, and another had mild tracheal stenosis and laryngomalacia requiring continuous positive airway pressure in infancy and arytenoid release surgery at 3 months of age.

Miscellaneous findings

Gastrointestinal abnormalities were observed in 12 individuals (67%), more common among individuals with missense variants ($p = 0.008$), and included feeding difficulties, gastroesophageal reflux, and constipation; two individuals required gastrostomy tube placement, and two additional individuals needed naso/orogastric tubes for a short period of time. Three individuals with missense variants and one individual with a truncation variant had pyloric stenosis.

Because of signs of overgrowth, umbilical hernia, macroglossia, and coarse facial features, four individuals underwent molecular testing for Beckwith-Wiedemann syndrome. Six individuals had molecular testing for other overgrowth syndromes (3/6 Simpson-Golabi-Behmel/*GPC3*, 2/6 Sotos/*NSD1*, 2/6 Costello), and three individuals had biochemical testing for mucopolysaccharidosis. All tests did not identify a genetic variant or biochemical evidence of the relevant disorder.

Hyperparathyroidism

Early in this study, an endocrine workup of proband 8 revealed inappropriately normal to high-normal levels of parathyroid hormone (PTH), which progressed to persistent PTH elevation, causing symptomatic hypercalcemia and eventually necessitating parathyroidectomy at 13 years of age. The mother of proband 8, who is a carrier

for the same *ZFX* variant, also underwent parathyroidectomy for hyperparathyroidism. The histopathology in proband 8 and his mother was consistent with parathyroid hyperplasia.

Intriguingly, the same *ZFX* variant (c.2357G>A [p.Arg786Gln]) that was found in this family has been previously reported as a somatic variant in sporadic parathyroid adenoma.[19] Based on this information, all collaborators were asked to share results of calcium and PTH levels if already drawn or order them in all probands. This approach led to the discovery of additional individuals with hyperparathyroidism. Proband 9, who carries the same variant (p.Arg786Gln) as proband 8, was diagnosed at 12 years of age with "familial hypocalciuric hypercalcemia" and at 15 years of age with parathyroid adenoma requiring parathyroidectomy. Proband 4, who carries the missense variant p.Thr771Met, exhibited biochemical abnormalities consistent with hyperparathyroidism (hypercalcemia and inappropriately normal PTH) and was evaluated for familial hypocalciuric hypercalcemia, but she is currently asymptomatic. Furthermore, four females with the p.Tyr774Cys variant from a three-generation family (including the mothers of probands 6A–6C; see Figure 3B and supplemental notes) were diagnosed with hyperparathyroidism and needed parathyroidectomy; the histology in one proband was consistent with parathyroid hyperplasia. None of the female relatives of probands 5A–5C and 7 with hyperparathyroidism exhibit the other clinical features seen in probands in this cohort. Notably, hypercalcemia secondary to hyperparathyroidism was observed only among individuals with missense variants. As previously discussed, X-inactivation studies were available for a subset of the individuals (for family 6, see Figure 3C); skewed X-inactivation may be a possible explanation for the observed phenotypes, but further sequencing studies are required.

Additional types of tumor and vascular anomalies were enriched in our cohort. Two individuals (probands 1 and 9) had multiple hemangiomas; proband 9 also had lymphatic malformations and was diagnosed at age 13 with thyroid papillary carcinoma. Proband 1 had hepatic angiomatosis, and proband 9 had hepatic adenoma. The mother of proband 8 was diagnosed with sarcoma. Proband 13 underwent a resection of benign suprarenal ganglioneuroma at 11 years of age. One of the female carriers in family 6 (III-1 in Figure 3C) had metastatic colorectal adenocarcinoma. Four individuals with missense variants had melanocytic nevi.

Missense ZFX variants demonstrate altered target gene expression

To explore the biochemical activities of several of the variant ZFX proteins present in the affected individuals, we individually introduced the p.Arg764Trp (probands 1 and 2), p.Thr771Met (probands 3, 4, and 5), p.Tyr774Cys (probands 6A–6C and 7), and p.Arg786Gln (probands 8 and 9) single-nucleotide changes into an expression vector containing a WT copy of a FLAG-tagged ZFX protein (Table S2). These four variants are located within the region of ZFX that we have previously shown to be required for DNA binding[9]; Arg786 and Arg764 are located within the finger loops of zinc finger 13 and zinc finger 12, respectively, whereas Thr771 and Tyr774 are located within the linker between zinc fingers 12 and 13 (Figure 1). The p.Arg786Gln, p.Arg764Trp, and p.Thr771Met variants have also been identified in The Cancer Genome Atlas database (TCGA, https://www.cancer.gov/ccg/research/genome-sequencing/tcga), indicating a possible role of these variants in carcinogenesis as well as in development. To evaluate the DNA binding and transactivation activities of the ZFX proteins harboring the individual amino acid changes, we used a cell line (DKO), derived from HEK293T cells, in which all endogenous *ZFX* family member protein expression has been eliminated through CRISPR-mediated deletion.[9] The use of these cells allows a direct comparison of the binding pattern and activity of the transfected ZFX variant proteins to the activity of the transfected WT *ZFX*, without interference from endogenous proteins; see Table S3 for details concerning all genomic datasets used in this study.

The DKO cell line was transfected with the different *ZFX* expression constructs or the parental vector as a control; 24 h after transfection, the cells were harvested, and the genomic DNA binding patterns for the variant ZFX proteins were analyzed and compared to the activity of WT ZFX. Genome-wide DNA binding profiles of the WT and variant ZFX proteins were determined by performing ChIP-seq assays using a FLAG antibody. Examination of the binding patterns revealed that, in general, the variant proteins had a very similar DNA binding pattern as did WT ZFX (Figure 4A); see also Figure S1B for heatmaps showing global comparisons of the binding patterns of the different ZFX proteins. However, we note that the variant ZFX proteins did appear to have slightly lower peaks than WT ZFX; this was especially noticeable for p.Thr771Met. These differences in peak height do not correlate with the expression levels of the variant proteins, as compared to the levels of WT ZFX; for example, p.Thr771Met has the most-reduced binding activity but is expressed at the same level as WT ZFX (Figure S1). As ZFX has previously been shown to be primarily localized to proximal promoter regions and the functional binding is confined to the sites located at +240 in the promoter region,[7,9] the reproducible promoter peaks (ranked by peak score) located within ± 2 kb of known promoters were selected for WT ZFX and the mutant ZFX proteins for further analyses (peak files are provided in Table S4). Pairwise comparisons between the peak sets from the WT and variant ZFX proteins indicate a very high degree of overlap (Figure 4B), with all pairwise comparisons showing greater than 80% identity. Although, taken overall, the same promoter regions were bound by the WT and variant proteins, it was possible that

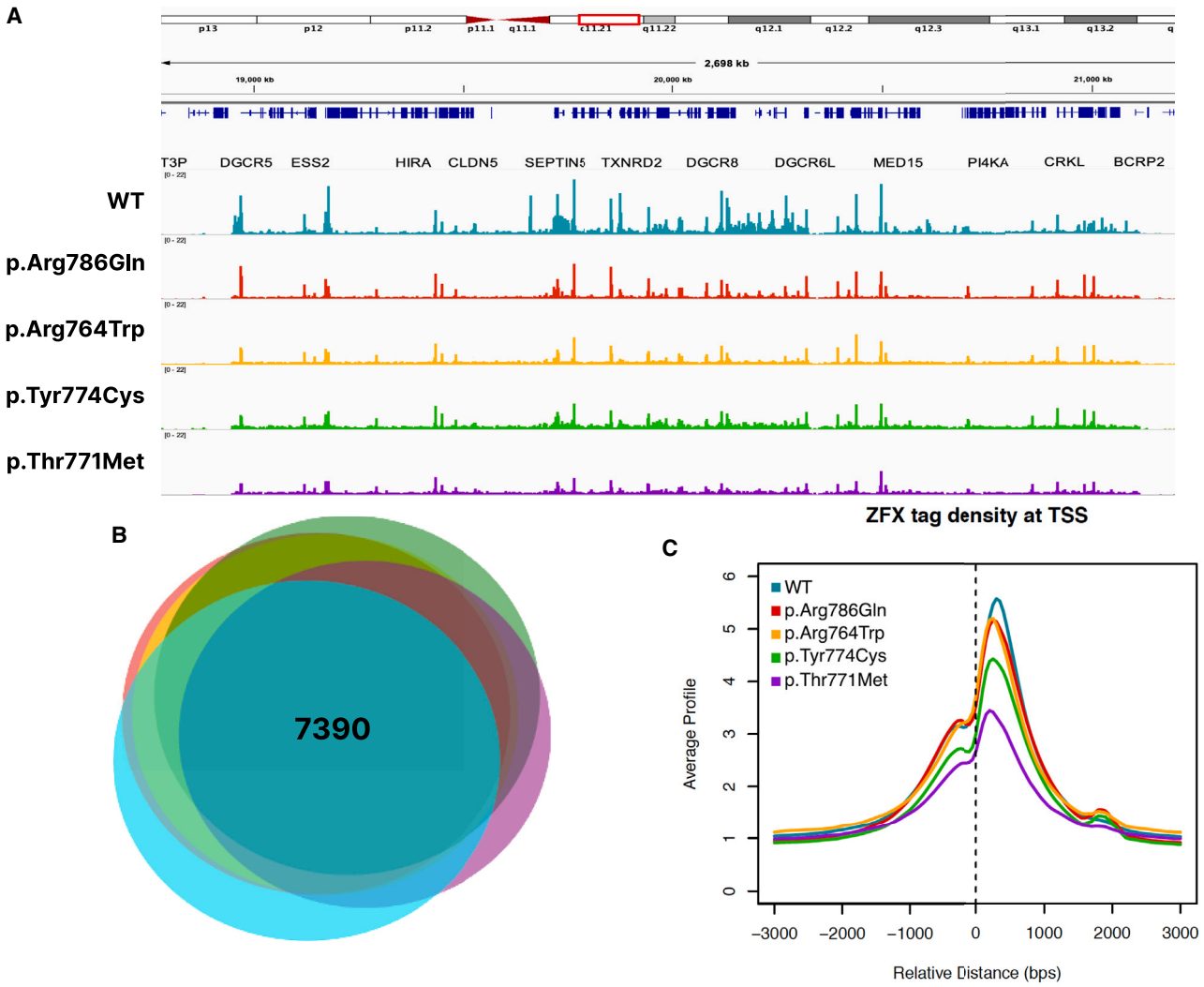


Figure 4. Characterization of wild-type and missense ZFX DNA binding by ChIP-seq
(A) Shown is a browser track displaying the genomic binding patterns of WT and variant ZFX proteins in a representative 2.4k-kb region of chromosome 22.
(B) Proportional Venn diagram of overlaps among the top 12,000 called peaks within 2 kb of the transcription start site (TSS) for WT and variant ZFX proteins. The number of peaks common to all ZFXs is indicated.
(C) Shown are the DNA binding profiles of WT and variant ZFX proteins from −2 kb to +2 kb from the TSS.

the variants bound to different positions within the same promoters as did WT ZFX. To investigate this possibility, a tag density analysis of the promoter peaks for WT and variant ZFXs was performed (Figure 4C). We have previously shown that endogenous WT ZFX predominantly binds downstream of the transcription start site at position +240.[9] Therefore, as expected, the transfected WT ZFX also displays this same pattern; we also found that each variant ZFX also has the same downstream binding pattern. Taken together, these data indicate that although the four missense variants are within the region of the ZFX protein required for DNA binding, these variants do not cause any major changes in genomic DNA binding patterns. However, we note that heatmaps of the ChIP-seq data suggest that the mutants have enhanced binding, as compared to WT ZFX, at a small set of promoters (Figure S1B).

Although the missense variants had only subtle effects on DNA binding, it remained possible that they could affect the transcriptional activation properties of ZFX. Therefore, we next transfected DKO cells with the WT or variant *ZFX* expression plasmids or parental vector as a control and prepared RNA for RNA-seq. Genes with expression changes caused by transfection of the WT or variant ZFX proteins were identified using DEseq2 (R). Changes in the transcriptome caused by each *ZFX* construct, as compared to the parental vector, are visualized by volcano plots (Figure 5A), and lists of all genes used in these plots are provided in Table S5. These experiments clearly demonstrate that the four variant ZFX proteins are active in the transactivation assay and the majority of the differentially regulated genes are upregulated. These results are consistent with classification of ZFX as a transcriptional activator.[7,9] Importantly, they also clearly demonstrate

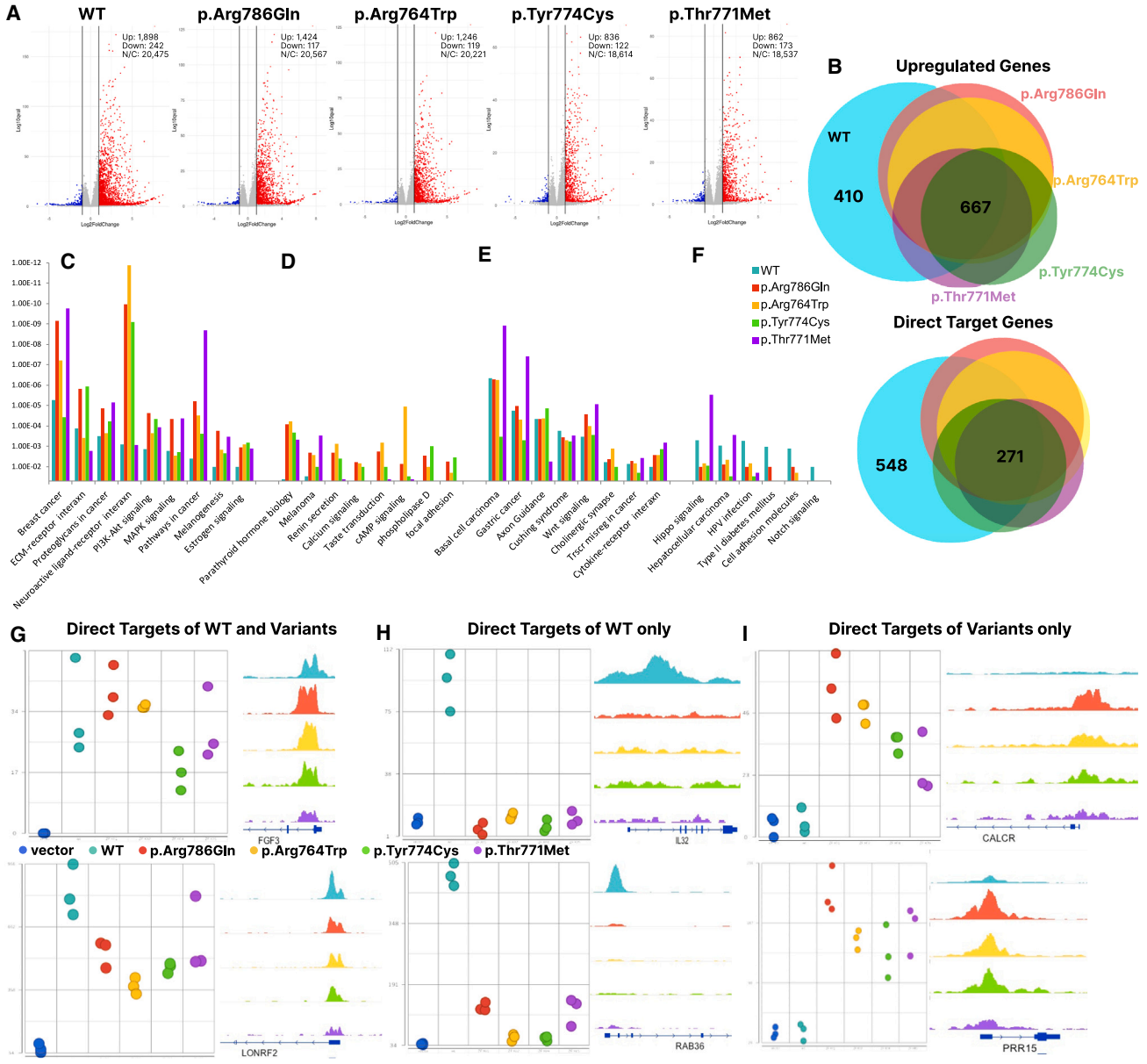


Figure 5. Characterization of differential expression in the context of missense *ZFX* variants

(A) Gene expression changes following transfection of WT or variant ZFX proteins into DKO cells. Volcano plots show gene expression changes in cells transfected with plasmids expressing the indicated ZFX compared to transfection with the vector alone. RNAs with increased expression are shown by red dots, RNAs with decreased expression are shown by blue dots; cut-offs used were a 2-fold change in expression and a q value < 0.05. The numbers of RNA with increased (Up), decreased (Down), and no change (N/C) in expression are displayed.

(B) Top: Overlap analysis of all genes activated by WT and variant ZFX proteins. The numbers of genes induced by all five proteins (667) and only by WT ZFX (410) are indicated. Bottom: Overlap analysis of direct targets activated by WT and variant ZFX proteins. The numbers of direct targets common to all five proteins (271) and direct targets unique to WT ZFX (548) are indicated.

(C–F) KEGG pathway analysis using the direct targets of WT and variant ZFX proteins. The top significant pathways are organized into four groups: (C) pathways more highly enriched in the set of genes regulated by the variant ZFXs than in the set of genes regulated by WT ZFX, (D) pathways only identified in sets of genes regulated by the variant ZFX proteins, (E) pathways enriched in both variant and WT gene sets, and (F) pathways enriched more in the sets of genes regulated by WT ZFX. Significance is plotted on the y axis.

(G–I) Examples of differences in expression and binding patterns of direct targets for WT ZFX vs. variant ZFX proteins; left panels show the expression level of the gene (values on the y axis represent normalized read counts mapping to all gene transcripts) in cells transfected with the WT and mutant ZFX proteins, whereas right panels show the ChIP-seq signals (browser shots) of the different ZFX WT and mutant constructs at the promoter of that gene. (G) Direct targets of both WT and variant ZFX, (H) direct targets of WT ZFX only, and (I) direct targets only of variant ZFX proteins.

that the amino acid changes do not create a complete loss-of-function ZFX protein. However, there are fewer upregulated genes upon transfection of p.Tyr774Cys and p.Thr771Met, suggesting that the ZFX linker variants may have slightly lower transactivation ability than the WT ZFX or the ZFX finger variants; this may be due to the slightly lower occupancy of these TFs on the promoter regions, as shown by the tag density plots in Figure 4C.

A proportional Venn diagram of the overlaps of the sets of genes with increased expression (Figure 5B, top) shows that the linker variants (p.Thr771Met and p.Tyr774Cys) and the zinc finger variants (p.Arg786Gln and p.Arg764Trp) cause changes in expression of many of the same genes as regulated by WT ZFX. However, some genes whose expression is increased by WT ZFX are not regulated by the variant ZFX proteins, and the variant proteins also have their own distinct sets of regulated genes. We note that changes in the transcriptome mediated by the WT and variant ZFX proteins are a combination of altered expression of direct target genes (i.e., genes having a promoter bound by ZFX and that show upregulation of expression in the transactivation assay) and of genes in downstream signaling pathways controlled by the direct target genes. Although analysis of the entire set of altered genes can provide insight into the biological consequences of the variants, analysis of the set of altered direct target genes can provide insight into the mechanisms by which the pathogenic variants affect ZFX transcriptional activity. Therefore, we next examined the effects of the missense variants on the regulation of direct target genes. For these analyses, we combined the DNA binding profiles of the WT and variant ZFX proteins with the RNA expression profiles, selecting genes that show a greater than 2-fold increase in expression and that have a ZFX peak (taken from the top-ranked 12,000 promoter peaks) at their promoter region, to generate lists of direct target genes for each ZFX protein. From this analysis, we identified 1,300 direct targets for WT ZFX and 994, 879, 665, and 619 direct target genes for p.Arg786Gln, p.Arg764Trp, p.Tyr774Cys, and p.Thr771Met, respectively; lists of the direct target genes for WT ZFX and the variant ZFX proteins are provided in Table S6. To compare the sets of direct targets for the WT and ZFX variant proteins, a proportional Venn diagram was created (Figure 5B, bottom). Again, we found that the WT ZFX, the linker variant (p.Thr771Met and p.Tyr774Cys), and the finger variants (p.Arg786Gln and p.Arg764Trp) can regulate distinct sets of direct target genes, in addition to the 271 direct target genes activated by all ZFX proteins. Both sets of overlaps reveal that the transcriptional profiles of cells transfected with finger variants are more similar to each other than to the transcriptional profiles of cells transfected with the linker variants (and vice versa).

The fact that the finger variants p.Arg786Gln and p.Arg764Trp and the linker variants p.Thr771Met and p.Tyr774Cys increase the expression of some genes that are not affected by WT ZFX and also fail to increase the expression of some genes that are affected by WT ZFX suggests that the pathogenic variants may change the cellular phenotype. To assess the possible biological consequences of the variant-specific and WT-specific targets, we performed the Partek gene pathway-KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis. The pathways enriched in sets of genes regulated by WT and variant ZFX proteins were organized into four groups: pathways more highly enriched in the sets of genes regulated by the variant ZFXs than in the set of genes regulated by WT ZFX, pathways only identified in sets of genes regulated by the variant ZFX proteins, pathways identified in both variant and WT gene sets, and pathways enriched more in the sets of genes regulated by WT ZFX (Figures 5C–5F). These data, taken together with the gene expression and DNA binding data, suggest that altered activities of the variant ZFX proteins may not only cause a defect in certain phenotypes through an inability to activate certain genes but may also result in the acquisition of new cellular phenotypes in cells harboring the variants. The substantial enrichment of the neuroactive ligand physiology pathway in the sets of genes transfected with the variant ZFX proteins is of particular interest given the neurodevelopmental phenotypes observed in this cohort. Examples of genes that are regulated by both WT and variant ZFX, only WT ZFX, and only variant ZFX are shown in Figures 5G–5I; of interest is *CALCR*, which falls within the neuroactive ligand physiology pathway and is only activated by the variant ZFX proteins.

Loss-of-function ZFX zebrafish model supports neurocognitive phenotype

To model truncating loss-of-function *ZFX* variants *in vivo*, we generated a *zfx* KO zebrafish model and characterized neurocognitive behavioral changes in adult zebrafish. To generate *zfx* zebrafish, we injected Cas9 mRNA and CRISPR guide RNAs targeting *zfx* exon 2, screened for mosaic F0 founders, and established a mutant allele containing a 25-bp deletion in *zfx*, which is predicted to cause a frameshift with premature termination (Figure S2). Additionally, we sought to analyze the spatiotemporal mRNA expression of *zfx* in WT zebrafish larvae using a previously reported protocol.[45] Whole-mount *in situ* hybridization indicated *zfx* expression in the brain regions of 24-hpf (hours postfertilization) and 48-hpf larvae (Figure S2C). In order to quantitate the mRNA expression level in KO organisms, real-time qPCR was conducted using RNA obtained from adult zebrafish brains. *zfx* mRNA expression was significantly reduced in *zfx* KO zebrafish compared to WT siblings ($p = 0.008$, Figure S2D).

KO zebrafish displayed no significant differences in terms of gross morphology compared to WT zebrafish during early developmental stages (Figure S3). Additionally, of note, over-expression of two *ZFX* missense variants (p.Arg764Trp and p.Arg786Gln) by microinjection of mRNA in WT zebrafish did not yield any developmental morphological phenotypes (Figure S4). However, the lack

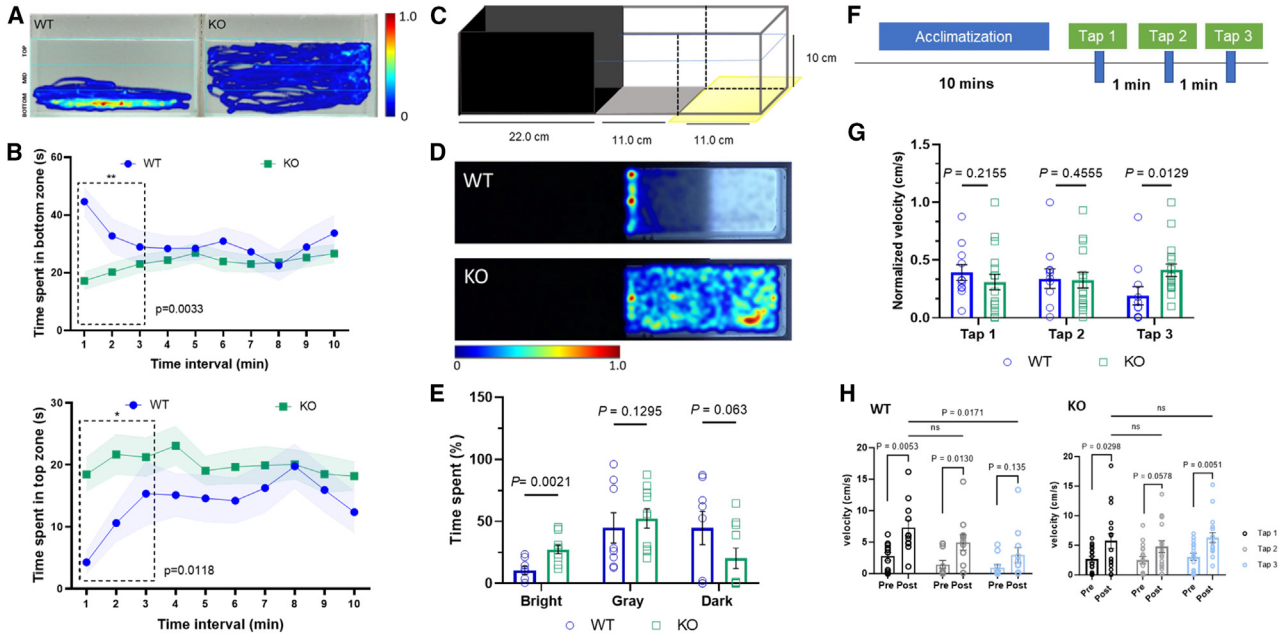


Figure 6. Behavioral characterization of *zfx* knockout zebrafish

(A) Representative heatmap image for novel tank assay during the first 3 min. The lines indicate the boundaries of three vertically different zones (top, middle, and bottom).

(B) Quantification of average time spent in novel tank assay. Upper graph shows WT spent significantly increased time in bottom zone ($t = 2.983$, $df = 23$, $p = 0.0033$, $n = 8$); lower graph shows KO fish spent higher time in the top zone ($t = 2.424$, $df = 23$, $p = 0.0118$, $n = 17$).

(C) Test apparatus for the scototaxis assay modified with gray and bright zone. The external light source is placed underneath the tank to evaluate the specific preference of adult zebrafish.

(D) Representative heatmap image for scototaxis assay.

(E) Quantification of percentage of average time spent in bright, gray, and dark zones. KO zebrafish spent significantly increased time in the bright zone ($t = 3.516$, $df = 16$, $p = 0.0021$, $n = 10$) compared to WT siblings ($n = 8$).

(F) Behavioral paradigm showing the 10-min acclimatization in the round container followed by three tap stimuli with 1-min intervals.

(G) Quantification of normalized velocity (cm/s) post-tap. WT ($n = 11$) and KO ($n = 17$) indicated similar response for both the first and second tap stimuli. For the third tap, KO zebrafish displayed significantly increased response ($t = 2.378$, $df = 26$, $p = 0.0129$).

(H) Quantification of velocity change (pre- and post-tap). WT fish showed a trend of decreasing response to the tap stimuli, whereas KO fish showed significantly increased response to the third tap ($t = 2.907$, $df = 16$, $p = 0.0051$).

of gross morphological changes does not preclude the potential for more subtle alterations in adult zebrafish behavior.

To evaluate behavioral changes in adult *zfx* KO zebrafish relative to WT, we performed a battery of behavioral tests, including a novel tank test, social interaction assay, mirror biting test, scototaxis assay, and adult startle tap test. The novel tank diving test is used to assay swimming activity and evaluate anxiety (equivalent to the rodent open field test).[56] When placed in a novel environment, zebrafish typically display a behavior called "novel tank diving" where they rapidly swim in the water column and explore their surroundings by diving and swimming along the bottom of the tank. This behavior is thought to be a response to an unfamiliar environment and can be used as a measure of anxiety-like behavior.[57] As expected, sibling WT fish spent significantly more time in the bottom zone. The time spent in the top zone was significantly increased in *zfx* KO zebrafish compared to WT fish during the first 3 min (Figures 6A and 6B), suggesting reduced anxiety or an altered anxiolytic response. There was no significant difference in distance moved by WT and *zfx* KO, indicating normal motor functions of *zfx* KO zebrafish. Social interaction and mirror biting tests showed no significant difference in behavior between WT and *zfx* KO fish.

We next performed a scototaxis test to assay dark/gray/light preference. WT zebrafish prefer to spend more time in darker conditions.[47] However, *zfx* KO zebrafish spent significantly increased time in the bright zone compared to WT fish (Figures 6C–6E). In the gray zone, both WT and *zfx* KO zebrafish spent equal time ($t = 0.5461$, $df = 16$, $p = 0.1295$). Although not statistically significant, WT fish spent more time in the dark zone compared to *zfx* KO zebrafish ($t = 1.604$, $df = 16$, $p = 0.063$). These results are consistent with the novel tank test and suggest decreased baseline anxiety or an altered anxiolytic response.

To examine the sensorimotor response and habituation of adult zebrafish, we also performed a startle tap test. In a startle tap test, the motor startle response is measured to determine an initial startle response and evaluate habituation with repeated testing.[48] Typically, zebrafish can adapt to the startle response over time when exposed to repeated taps in the same environment.[58,59] No significant

differences were observed in responses for both the first and second tap stimuli between WT and *zfx* KO fish. However, for the third stimulus, *zfx* KO fish demonstrated a significantly increased response compared to WT fish (Figures 6F–6H). While WT fish showed a trend of significantly decreasing response by the third tap stimulus, *zfx* zebrafish showed the opposite response, indicating a lack of habituation to the external tap stimuli. Taken together, these experiments support that *zfx* KO zebrafish exhibit altered behavioral activity compared to WT fish, suggesting that loss of *zfx* function may result in neurocognitive abnormalities.

Discussion

Here, we present a large cohort with an overlapping syndromic intellectual disability phenotype linked by variants in the TF *ZFX*. These individuals share overlapping and recurrent facial features, neurocognitive abnormalities, behavioral problems, and congenital anomalies. A subset of individuals with missense variants in the DNA-binding domain demonstrate hyperparathyroidism. Molecular characterization of genome-wide binding profiles of the missense *ZFX* variants by ChIP-seq and transcriptional dysregulation by RNA-seq shows altered transcriptional activation of a small set of gene targets compared to WT ZFX. Furthermore, a zebrafish model of *ZFX* loss demonstrated abnormal anxiolytic and habituation behaviors in several tests, including a novel tank, scototaxis, and adult tap test—supporting altered neurocognitive behavior in association with truncating *ZFX* variants. Taken together, our data suggest that the individuals in this cohort represent a complex X-linked intellectual disability syndrome.

Individuals in this study cohort displayed considerable phenotypic variability, prompting clinicians to order non-targeted molecular studies, including ES and chromosomal microarray analysis, to evaluate the probands. The main indications for ordering such studies included developmental delay, intellectual disability, multiple congenital anomalies, and dysmorphic features, all of which are associated with remarkable locus heterogeneity (for example, there are numerous genes associated with syndromic neurodevelopmental disorders). Nevertheless, targeted testing for Fragile X syndrome, which is considered a first-tier testing candidate for individuals with developmental delay/intellectual disability, was ordered for several probands. In addition, four subjects underwent molecular testing for Beckwith-Wiedemann syndrome because of overgrowth, umbilical hernia, and macroglossia, and six additional individuals underwent molecular testing for other overgrowth syndromes including Simpson-Golabi-Behmel, Sotos, and Costello syndromes. Mucopolysaccharidosis has been also considered on the differential diagnosis of three probands based on coarse facial features, some overgrowth, and hernias. Hyperparathyroidism is an emerging phenotype associated with certain *ZFX* variants and, if present in individuals with syndromic neurodevelopmental disorder (NDD), should probably prompt consideration of this diagnostic entity.

Eight individuals had additional genetic findings of varying clinical significance (Table S1, supplemental notes), and a subset of them may have contributed to or modified the individuals' phenotypes. Proband 9 was found to have a paternally inherited variant in *KCNH2* (MIM: 152427), monoallelic variants in which are associated with long QT syndrome-2 (MIM: 613688). This finding is consistent with the proband's personal and family history of prolonged QT syndrome, and it is not related to the phenotype associated with the *ZFX* variant. Proband 12 carries a pathogenic *OCA2* (MIM: 611409) variant and a heterozygous missense variant in *MC1R* (MIM: 155555). The relevance of these variants to the proband's phenotype is unclear. His mother is clinically affected with oculocutaneous albinism and carries 2 pathogenic variants in *OCA2*, but he does not display classical features of this condition, and his nystagmus is thought to be secondary to optic nerve hypoplasia. Proband 14 had a paternally inherited nonsense variant in *ROBO1* (MIM: 602430), which can explain the proband's pituitary hypoplasia and growth hormone deficiency (MIM: 620303; combined or isolated pituitary hormone deficiency-8). This is a typical example of a blended phenotype in the same individual. ES of proband 15 also revealed a VUS in *TFAP2A* (MIM: 107580), heterozygous pathogenic variants in which are associated with branchiooculofacial syndrome (MIM: 113620). The proband does have hearing loss, which has been described in this condition, but the absence of the other syndromic cardinal features made it seem unlikely that the *TFAP2A* variant is the cause of this proband's phenotype. Four additional probands had small copy number variants of unclear clinical significance, but we cannot completely exclude a potential modifying effect.

Multiple lines of evidence support the pathogenicity of the observed *ZFX* variants, including shared and recurrent phenotypes in affected individuals, structural and *in silico* variant analysis, lack of these variants in the general population based on publicly available databases, and functional molecular data demonstrating altered transcriptional activity in missense variants and altered behavior in a zebrafish loss-of-function model. It is important to note that *ZFX* is a highly intolerant gene for loss-of-function variants, as demonstrated by high probability of loss-of-function intolerance score (pLI = 1.0), low loss-of-function observed/expected upper bound fraction (LOEUF = 0.16),[60] and low haploinsufficiency score (HI index = 7.61%).[61]

Although we cannot yet estimate the frequency of *ZFX*-related disorder, *ZFX* variants appear to be very uncommon. GeneDx confirmed *de novo* missense or frameshift variants in *ZFX* for four affected individuals and maternally inherited missense variants in 12 affected hemizygotes out of 125,778 individuals undergoing clinical

ES/GS for an indication of NDD. Of the *de novo* variants, there was one hemizygous missense variant, one heterozygous missense variant, and two heterozygous frameshift variants. Because of the design and scope of this study, we were unable to calculate the enrichment of *ZFX* variants in NDD. Nevertheless, none of the variants reported in this study were found in gnomAD (last accessed: October 6, 2023). *ZFX* is not only intolerant for loss-of-function variants (based on pLI, LOEUF, and HI index) but also highly constrained for missense variants. In fact, the missense constraint *Z* score for ZFX, which calculates the deviation of observed from predicted missense variants, is +3.15, with higher scores indicating that the transcript is more constrained and less tolerant of missense variation.[62]

Variants in TFs, by nature of their complex cellular functions, can result in pleiotropic phenotypes. Most fundamentally, TFs can be reduced to two functional domains: (1) a structured DNA-binding domain that interacts with a specific DNA motif and (2) a structured or unstructured transcriptional effector domain responsible for activation (the recruitment of transcriptional machinery to a target gene) and/or repression (the binding of corepressors and chromatin remodelers). Variants in TFs have the potential to impair or alter DNA binding specificity, inhibit transcriptional effector function, result in loss of TF expression entirely, or cause a gain of transcriptional function. In this *ZFX* cohort, we observe both predicted loss-of-function truncating variants and DNA-binding domain missense variants that exhibit the potential for gain-of-function behavior. This complexity is consistent with the observed pleiotropy of the individuals in this cohort, which may reflect variations in the particular gene targets of the different *ZFX* variants.

It is of particular interest that the missense variants cluster near the penultimate and ultimate zinc fingers of ZFX, which we have shown to be critical for genomic DNA binding activity.[9] Although several variants fall within the zinc fingers, the variant proteins are not significantly impaired in their recruitment to promoter regions. Importantly, we show that the missense variants generally bind to the same promoters and activate the same genes as does WT ZFX, and thus they are not inactivating. However, it is also clear that some genes are (1) direct targets of WT ZFX but not the variant ZFX proteins and (2) direct targets of one or more of the variant ZFX proteins but not of the WT ZFX (Figures 5G–5I). Taken together, our studies suggest that the missense variants may result in both gain and loss of transcriptional function at a small set of genes.

In particular, the hyperparathyroidism observed across multiple individuals with three different nearby *ZFX* DNA-binding domain variants supports the possible role of altered binding affinity or activation in the pathogenesis of this phenotype. Hyperparathyroidism was also observed in carrier females of the p.Arg786Gln and p.Tyr774Cys variants who were otherwise phenotypically unaffected, suggesting that the parathyroid phenotype could result from an altered protein activity leading to mis-activation of downstream non-target genes. Our molecular studies of the missense variants demonstrating differential transcriptional activation (e.g., binding to and activation of *CALCR*, which encodes the calcitonin receptor, by missense ZFX proteins but not WT ZFX) support this hypothesis, but further molecular studies of these specific variants will be needed to investigate this link.

The prevalence of affected male individuals supports an X-linked recessive inheritance pattern, but the manifesting heterozygote females, including probands 4, 9, and 14 and members of the 6A–6C family, complicate this classification. Skewed mosaic X-inactivation in females, resulting from either random chance or selection during development, can vary the presence and severity of X-linked intellectual disability phenotypes.[63,64] X-inactivation studies were only available for a subset of the cohort members, but for the 6A–6C family, several affected females displayed substantial X-inactivation skewing (see Figure 3C), and additional X-inactivation studies for proband 9 also demonstrated substantial skewing. It is important to note that only proband 6C from family 6 had clinical findings similar to the remainder of the cohort, whereas other females mainly had hyperparathyroidism. One possible explanation is variation of X-inactivation in different tissues, but the contribution of other genetic and nongenetic modifiers cannot be excluded. The consequence of this in their presented phenotype, and indeed the potential for disease in carrier females, will necessitate additional characterization.

The specific molecular pathomechanisms by which *ZFX* variants cause disease will require further studies, but this work reflects a comprehensive effort to collate a phenotypic profile of *ZFX* disease. The specific role of *ZFX* DNA-binding domain variants in the pathogenesis of neurocognitive phenotypes and hyperparathyroidism, in particular, is ripe for further characterization. Additionally, further investigation of the gene targets and transcriptional regulatory network of *ZFX* may yield substantial insights into molecular disease mechanisms.

A primary limitation of this study is the heterogeneity of clinical assessment of the participating individuals, with evaluations by different professionals resulting in the potential for interobserver bias. For instance, not all participants had a comprehensive and systematic neurocognitive evaluation, which restricts the potential for comparisons across some neurocognitive traits. Furthermore, due to lack of availability of biological specimens, molecular studies were not performed on all truncating variants, although their predicted loss-of-function phenotypes were modeled in zebrafish and assessed based on in silico models. Additionally, the small cohort size and the fact that the majority of the *ZFX* variants were private to each proband or family made the assessment of genotype-phenotype correlation difficult. Furthermore, because of the original design and scope of this study, we could not perform a formal statistical test of enrichment for

ZFX variants in affected individuals as compared to controls. In the *in vitro* studies, co-deletion of partner TFs *ZFY* and/or *ZNF711* was performed, although in this cohort these genes are unaffected. However, there is precedent for variants in only a single copy of a redundant X/Y gene pair to cause disease, such as the gene pair *NLGN4X* (MIM: 300427) and *NLGN4Y* (MIM: 400028), of which *NLGN4X* specifically is associated with X-linked intellectual disability (intellectual developmental disorder, X-linked, MIM: 300495). The need to delete homologous genes in models may just be a consequence of assay sensitivity.

In summary, both truncation and missense variants in the TF *ZFX* are associated with a complex X-linked intellectual disability syndrome with characteristic developmental and behavioral abnormalities. Further characterization of variants is needed to reveal the precise pathomechanism, particularly with regards to the recurrent hyperparathyroidism observed in a subset of affected individuals.

Data and code availability

All raw and some processed files used in this study are available at the Gene Expression Omnibus. The accession number for the data reported in this paper is GEO: GSE218691.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2024.01.007.

Acknowledgments

We thank the cohort members and their families for participating in this study. For additional acknowledgments, including detailed funding information, please see the supplemental acknowledgments.

Author contributions

Investigation: J.L.S., K.H., D.W.D., G.M., T.-I.C., C.A.A., D.J.A., S.B., D.G.B., L.D.B., D.A.C., R.C., J.C.-S., A.C., M.D., L.F., C.P.G., N.B.G., K.W.G., E.H., A.M.H., A.M.I., B.I., A.J., P.K., L.A.R.K., S.K., F.L., P.L., J. Maraval, N. Matsumoto, J. McCarrier, J. McCarthy, N. Miyake, L.H.M., A.H.N., E.Ø., R.P., K.P., J.E.P., R.E.S., M. Shaw, E.S., J.P.T., E. Wadman, E. Wakeling, S.M.W., L.C.W., J.R.L., O.L., M.A.C., J.G., C.M.N., P.J.F., C.-H.K., M. Shinawi; writing—original draft: J.L.S., D.W.D., P.K., C.M.N., M. Shinawi; writing—review & editing: J.L.S., K.H., D.W.D., G.M., C.A.A., D.J.A., S.B., D.G.B., L.D.B., D.A.C., J.C.-S., A.C., M.D., L.F., C.P.G., N.B.G., K.W.G., E.H., A.M.H., A.M.I., B.I., A.J., P.K., L.A.R.K., S.K., F.L., P.L., J. Maraval, N. Matsumoto, J. McCarrier, J. McCarthy, N. Miyake, L.H.M., A.H.N., E.Ø., R.P., K.P., J.E.P., R.E.S., E.S., J.P.T., E. Wadman, E. Wakeling, S.M.W., L.C.W., J.R.L., O.L., M.A.C., J.G., C.M.N., P.J.F., C.-H.K., M. Shinawi; resources: G.M., C.A.A., D.J.A., S.B., D.G.B., L.D.B., D.A.C., R.C., J.C.-S., M.D., L.F., C.P.G., N.B.G., K.W.G., E.H., A.M.H., A.M.I., B.I., A.J., L.A.R.K., G.E.R.C., S.K., F.L., P.L., J. Maraval, N. Matsumoto, J. McCarrier, J. McCarthy, N. Miyake, L.H.M., A.H.N., E.Ø., R.P., K.P., R.E.S., M. Shaw, E.S., E. Wadman, E. Wakeling, S.M.W., L.C.W., M.A.C., P.J.F., C.-H.K., M. Shinawi; supervision: O.L., J.G., P.J.F., C.-H.K., M. Shinawi; funding acquisition: J.G., P.J.F., C.-H.K., M. Shinawi; conceptualization: M. Shinawi; project administration: M. Shinawi.

Declaration of interests

D.A.C. and R.E.S. are employees of GeneDx, LLC. A.C. and J.P.T. are employees and shareholders of Illumina, Inc. L.D.B. performs advisory board, consulting, and speaking arrangement work unrelated to the present study for Sanofi S.A., Horizon Therapeutics, Amicus Therapeutics, and Chiesi Farmaceutici S.p.A. N.B.G. has received personal fees from Pfizer Inc. and RCG Consulting for work unrelated to the present study. J.R.L. has stock ownership in 23andMe and is a paid consultant to Genome International.

Received: August 6, 2023
Accepted: January 17, 2024
Published: February 6, 2024

References

1. Tejada, M.I., and Ibarluzea, N. (2020). Non-syndromic X Linked Intellectual Disability: Current Knowledge in Light of the Recent Advances in Molecular and Functional Studies. Clin. Genet. *97*, 677–687. https://doi.org/10.1111/cge.13698.
2. Stevenson, R.E., Schwartz, C.E., and Rogers, R.C. (2013). Malformations among the X-linked Intellectual Disability Syndromes. Am. J. Med. Genet. *161*, 2741–2749. https://doi.org/10.1002/ajmg.a.36179.
3. Schwartz, C.E., Louie, R.J., Toutain, A., Skinner, C., Friez, M.J., and Stevenson, R.E. (2023). X-Linked Intellectual Disability Update 2022. Am. J. Med. Genet. *191*, 144–159. https://doi.org/10.1002/ajmg.a.63008.
4. Lubs, H.A., Stevenson, R.E., and Schwartz, C.E. (2012). Fragile X and X-Linked Intellectual Disability: Four Decades of Discovery. Am. J. Hum. Genet. *90*, 579–590. https://doi.org/10.1016/j.ajhg.2012.02.018.
5. Fieremans, N., Van Esch, H., Holvoet, M., Van Goethem, G., Devriendt, K., Rosello, M., Mayo, S., Martinez, F., Jhangiani, S., Muzny, D.M., et al. (2016). Identification of Intellectual Disability Genes in Female Patients with a Skewed X-Inactivation Pattern. Hum. Mutat. *37*, 804–811. https://doi.org/10.1002/humu.23012.
6. Kaufman, L., Ayub, M., and Vincent, J.B. (2010). The Genetic Basis of Non-Syndromic Intellectual Disability: A Review. J. Neurodev. Disord. *2*, 182–209. https://doi.org/10.1007/s11689-010-9055-2.
7. Rhie, S.K., Yao, L., Luo, Z., Witt, H., Schreiner, S., Guo, Y., Perez, A.A., and Farnham, P.J. (2018). ZFX Acts as a Transcriptional Activator in Multiple Types of Human Tumors by Binding Downstream from Transcription Start Sites at the Majority of CpG Island Promoters. Genome Res. *28*, 310–320. https://doi.org/10.1101/gr.228809.117.
8. Mardon, G., Luoh, S.W., Simpson, E.M., Gill, G., Brown, L.G., and Page, D.C. (1990). Mouse Zfx Protein Is Similar to Zfy-2: Each Contains an Acidic Activating Domain and 13 Zinc Fingers. Mol. Cell Biol. *10*, 681–688. https://doi.org/10.1128/mcb.10.2.681-688.1990.
9. Ni, W., Perez, A.A., Schreiner, S., Nicolet, C.M., and Farnham, P. (2020). Characterization of the ZFX Family of Transcription

Factors That Bind Downstream of the Start Site of CpG Island Promoters. Nucleic Acids Res. *48*, 5986–6000. https://doi.org/10.1093/nar/gkaa384.

10. Page, D.C., Mosher, R., Simpson, E.M., Fisher, E.M., Mardon, G., Pollack, J., McGillivray, B., de la Chapelle, A., and Brown, L.G. (1987). The Sex-Determining Region of the Human Y Chromosome Encodes a Finger Protein. Cell *51*, 1091–1104. https://doi.org/10.1016/0092-8674(87)90595-2.
11. Schneider-Gädicke, A., Beer-Romero, P., Brown, L.G., Nussbaum, R., and Page, D.C. (1989a). *ZFX* Has a Gene Structure Similar to *ZFY*, the Putative Human Sex Determinant, and Escapes X Inactivation. Cell *57*, 1247–1258. https://doi.org/10.1016/0092-8674(89)90061-5.
12. Schneider-Gädicke, A., Beer-Romero, P., Brown, L.G., Mardon, G., Luoh, S.-W., and Page, D.C. (1989b). Putative Transcription Activator with Alternative Isoforms Encoded by Human *ZFX* Gene. Nature *342*, 708–711. https://doi.org/10.1038/342708a0.
13. Galan-Caridad, J.M., Harel, S., Arenzana, T.L., Hou, Z.E., Doetsch, F.K., Mirny, L.A., and Reizis, B. (2007). Zfx Controls the Self-Renewal of Embryonic and Hematopoietic Stem Cells. Cell *129*, 345–357. https://doi.org/10.1016/j.cell.2007.03.014.
14. Palmer, C.J., Galan-Caridad, J.M., Weisberg, S.P., Lei, L., Esquilin, J.M., Croft, G.F., Wainwright, B., Canoll, P., Owens, D.M., and Reizis, B. (2014). Zfx Facilitates Tumorigenesis Caused by Activation of the Hedgehog Pathway. Cancer Res. *74*, 5914–5924. https://doi.org/10.1158/0008-5472.can-14-0834.
15. Weisberg, S.P., Smith-Raska, M.R., Esquilin, J.M., Zhang, J., Arenzana, T.L., Lau, C.M., Churchill, M., Pan, H., Klinakis, A., Dixon, J.E., et al. (2014). ZFX Controls Propagation and Prevents Differentiation of Acute T-Lymphoblastic and Myeloid Leukemia. Cell Rep. *6*, 528–540. https://doi.org/10.1016/j.celrep.2014.01.007.
16. Luoh, S.W., Bain, P.A., Polakiewicz, R.D., Goodheart, M.L., Gardner, H., Jaenisch, R., and Page, D.C. (1997). *Zfx* Mutation Results in Small Animal Size and Reduced Germ Cell Number in Male and Female Mice. Development *124*, 2275–2284. https://doi.org/10.1242/dev.124.11.2275.
17. van der Werf, I.M., Van Dijck, A., Reyniers, E., Helsmoortel, C., Kumar, A.A., Kalscheuer, V.M., de Brouwer, A.P., Kleefstra, T., van Bokhoven, H., Mortier, G., et al. (2017). Mutations in Two Large Pedigrees Highlight the Role of *ZNF711* in X-linked Intellectual Disability. Gene *605*, 92–98. https://doi.org/10.1016/j.gene.2016.12.013.
18. Tarpey, P.S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy, C., O'Meara, S., Latimer, C., Dicks, E., Menzies, A., et al. (2009). A Systematic, Large-Scale Resequencing Screen of X-chromosome Coding Exons in Mental Retardation. Nat. Genet. *41*, 535–543. https://doi.org/10.1038/ng.367.
19. Soong, C.-P., and Arnold, A. (2014). Recurrent ZFX Mutations in Human Sporadic Parathyroid Adenomas. Oncoscience *1*, 360–366. https://doi.org/10.18632/oncoscience.38.
20. Landrum, M.J., Lee, J.M., Benson, M., Brown, G.R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., et al. (2018). ClinVar: Improving Access to Variant Interpretations and Supporting Evidence. Nucleic Acids Res. *46*, D1062–D1067. https://doi.org/10.1093/nar/gkx1153.
21. Angelozzi, M., Karvande, A., Molin, A.N., Ritter, A.L., Leonard, J.M.M., Savatt, J.M., Douglass, K., Myers, S.M., Grippa, M., Tolchin, D., et al. (2022). Consolidation of the Clinical and Genetic Definition of a *SOX4*- Related Neurodevelopmental Syndrome. J. Med. Genet. *59*, 1058–1068. https://doi.org/10.1136/jmedgenet-2021-108375.
22. Sharma, R., Sahoo, S.S., Honda, M., Granger, S.L., Goodings, C., Sanchez, L., Künstner, A., Busch, H., Beier, F., Pruett-Miller, S.M., et al. (2022). Gain-of-Function Mutations in RPA1 Cause a Syndrome with Short Telomeres and Somatic Genetic Rescue. Blood *139*, 1039–1051. https://doi.org/10.1182/blood.2021011980.
23. Hoshino, A., Boutboul, D., Zhang, Y., Kuehn, H.S., Hadjadj, J., Özdemir, N., Celkan, T., Walz, C., Picard, C., Lenoir, C., et al. (2022). Gain-of-Function *IKZF1* Variants in Humans Cause Immune Dysregulation Associated with Abnormal T/B Cell Late Differentiation. Sci. Immunol. *7*, eabi7160. https://doi.org/10.1126/sciimmunol.abi7160.
24. 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. https://doi.org/10.1002/humu.22844.
25. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Vooren, S.V., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am. J. Hum. Genet. *84*, 524–533. https://doi.org/10.1016/j.ajhg.2009.03.010.
26. Faundes, V., Newman, W.G., Bernardini, L., Canham, N., Clayton-Smith, J., Dallapiccola, B., Davies, S.J., Demos, M.K., Goldman, A., Gill, H., et al. (2018). Clinical Assessment of the Utility of Sequencing and Evaluation as a Service (CAUSES) Study, Deciphering Developmental Disorders (DDD) Study, S. Banka, Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders. Am. J. Hum. Genet. *102*, 175–187. https://doi.org/10.1016/j.ajhg.2017.11.013.
27. Jackson, A., Banka, S., Stewart, H., Robinson, H., Lovell, S., and Clayton-Smith, J. (2021). Recurrent KCNT2 missense variants affecting p.Arg190 result in a recognizable phenotype. Am. J. Med. Genet. *185*, 3083–3091. https://doi.org/10.1002/ajmg.a.62370.
28. Pagnamenta, A.T., Jackson, A., Perveen, R., Beaman, G., Petts, G., Gupta, A., Hyder, Z., Chung, B.H.-Y., Kan, A.S.-Y., Cheung, K.W., et al. (2022). Biallelic TMEM260 variants cause truncus arteriosus, with or without renal defects. Clin. Genet. *101*, 127–133. https://doi.org/10.1111/cge.14071.
29. ACMG Laboratory Quality Assurance Committee, S. Richards, Aziz, N., S. Bale, D.B., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., E. Spector, K.V., and Rehm, H.L. (2015). Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. *17*, 405–423. https://doi.org/10.1038/gim.2015.30.
30. Abou Tayoun, A.N., Pesaran, T., DiStefano, M.T., Oza, A., Rehm, H.L., Biesecker, L.G., and Harrison, S.M. (2018). ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI), Recommendations for Interpreting the Loss of Function PVS1 ACMG/AMP Variant Criterion. Hum. Mutat. *39*, 1517–1524. https://doi.org/10.1002/humu.23626.
31. Pejaver, V., Byrne, A.B., Feng, B.-J., Pagel, K.A., Mooney, S.D., Karchin, R., O'Donnell-Luria, A., Harrison, S.M., Tavtigian, S.V., Greenblatt, M.S., et al. (2022). ClinGen Sequence Variant Interpretation Working Group, Calibration of Computational Tools for Missense Variant Pathogenicity Classification and ClinGen Recommendations for PP3/BP4 Criteria. Am. J.

Hum. Genet. *109*, 2163–2177. https://doi.org/10.1016/j.ajhg.2022.10.013.
32. Katsonis, P., and Lichtarge, O. (2014). A Formal Perturbation Equation between Genotype and Phenotype Determines the Evolutionary Action of Protein-Coding Variations on Fitness. Genome Res. *24*, 2050–2058. https://doi.org/10.1101/gr.176214.114.
33. Lichtarge, O., Bourne, H.R., and Cohen, F.E. (1996). An Evolutionary Trace Method Defines Binding Surfaces Common to Protein Families. J. Mol. Biol. *257*, 342–358. https://doi.org/10.1006/jmbi.1996.0167.
34. Mihalek, I., Res, I., and Lichtarge, O. (2004). A Family of Evolution-Entropy Hybrid Methods for Ranking Protein Residues by Importance. J. Mol. Biol. *336*, 1265–1282. https://doi.org/10.1016/j.jmb.2003.12.078.
35. Katsonis, P., and Lichtarge, O. (2019). CAGI5: Objective Performance Assessments of Predictions Based on the Evolutionary Action Equation. Hum. Mutat. *40*, 1436–1454. https://doi.org/10.1002/humu.23873.
36. Katsonis, P., and Lichtarge, O. (2017). Objective Assessment of the Evolutionary Action Equation for the Fitness Effect of Missense Mutations across CAGI-blinded Contests. Hum. Mutat. *38*, 1072–1084. https://doi.org/10.1002/humu.23266.
37. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic Local Alignment Search Tool. J. Mol. Biol. *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
38. Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R., and Wu, C.H. (2007). UniRef: Comprehensive and Non-Redundant UniProt Reference Clusters. Bioinformatics *23*, 1282–1288. https://doi.org/10.1093/bioinformatics/btm098.
39. Edgar, R.C. (2004). MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. Nucleic Acids Res. *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340.
40. Lua, R.C., and Lichtarge, O. (2010). PyETV: A PyMOL Evolutionary Trace Viewer to Analyze Functional Site Predictions in Protein Complexes. Bioinformatics *26*, 2981–2982. https://doi.org/10.1093/bioinformatics/btq566.
41. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
42. Thouin, M.M., Giron, J.M., and Hoffman, E.P. (2002). Detection of Nonrandom X Chromosome Inactivation. Curr. Protoc. Hum. Genet. *35*. https://doi.org/10.1002/0471142905.hg0907s35.
43. Allen, R.C., Zoghbi, H.Y., Moseley, A.B., Rosenblatt, H.M., and Belmont, J.W. (1992). Methylation of HpaII and HhaI Sites near the Polymorphic CAG Repeat in the Human Androgen-Receptor Gene Correlates with X Chromosome Inactivation. Am. J. Hum. Genet. *51*, 1229–1239.
44. T. Hulsen, DeepVenn – a Web Application for the Creation of Area-Proportional Venn Diagrams Using the Deep Learning Framework Tensorflow.Js,Preprint at arXiv. https://doi.org/10.48550/ARXIV.2210.04597.
45. Lee, Y.-R., Khan, K., Armfield-Uhas, K., Srikanth, S., Thompson, N.A., Pardo, M., Yu, L., Norris, J.W., Peng, Y., Gripp, K.W., et al. (2020). Mutations in FAM50A Suggest That Armfield XLID Syndrome Is a Spliceosomopathy. Nat. Commun. *11*, 3698. https://doi.org/10.1038/s41467-020-17452-6.
46. Kim, O.-H., Cho, H.-J., Han, E., Hong, T.I., Ariyasiri, K., Choi, J.-H., Hwang, K.-S., Jeong, Y.-M., Yang, S.-Y., Yu, K., et al. (2017). Zebrafish Knockout of Down Syndrome Gene, DYRK1A, Shows Social Impairments Relevant to Autism. Mol. Autism. *8*, 50. https://doi.org/10.1186/s13229-017-0168-2.
47. Maximino, C., Marques de Brito, T., Dias, C.A. G.d.M., Gouveia, A., and Morato, S. (2010). Scototaxis as Anxiety-like Behavior in Fish. Nat. Protoc. *5*, 209–216. https://doi.org/10.1038/nprot.2009.225.
48. Eddins, D., Cerutti, D., Williams, P., Linney, E., and Levin, E.D. (2010). Zebrafish Provide a Sensitive Model of Persisting Neurobehavioral Effects of Developmental Chlorpyrifos Exposure: Comparison with Nicotine and Pilocarpine Effects and Relationship to Dopamine Deficits. Neurotoxicol. Teratol. *32*, 99–108. https://doi.org/10.1016/j.ntt.2009.02.005.
49. Ho, J., Tumkaya, T., Aryal, S., Choi, H., and Claridge-Chang, A. (2019). Moving beyond P Values: Data Analysis with Estimation Graphics. Nat. Methods *16*, 565–566. https://doi.org/10.1038/s41592-019-0470-3.
50. Chen, S., Francioli, L.C., Goodrich, J.K., Collins, R.L., Kanai, M., Wang, Q., Alföldi, J., Watts, N.A., Vittal, C., Gauthier, L.D., et al. (2024). A genomic mutational constraint map using variation in 76,156 human genomes. Nature *625*, 92–100. https://doi.org/10.1038/s41586-023-06045-0.
51. Rentzsch, P., Schubach, M., Shendure, J., and Kircher, M. (2021). CADD-Splice—Improving Genome-Wide Variant Effect Prediction Using Deep Learning-Derived Splice Scores. Genome Med. *13*, 31. https://doi.org/10.1186/s13073-021-00835-9.
52. Ioannidis, N.M., Rothstein, J.H., Pejaver, V., Middha, S., McDonnell, S.K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., et al. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am. J. Hum. Genet. *99*, 877–885. https://doi.org/10.1016/j.ajhg.2016.08.016.
53. Chang, K.T., Guo, J., di Ronza, A., and Sardiello, M. (2018). Aminode: Identification of Evolutionary Constraints in the Human Proteome. Sci. Rep. *8*, 1357. https://doi.org/10.1038/s41598-018-19744-w.
54. Silk, M., Petrovski, S., Ascher, D.B., and MTR-Viewer. (2019). Identifying Regions within Genes under Purifying Selection. Nucleic Acids Res. *47*, W121–W126. https://doi.org/10.1093/nar/gkz457.
55. Wainer Katsir, K., and Linial, M. (2019). Human Genes Escaping X-inactivation Revealed by Single Cell Expression Data. BMC Genom. *20*, 201. https://doi.org/10.1186/s12864-019-5507-6.
56. Kysil, E.V., Meshalkina, D.A., Frick, E.E., Echevarria, D.J., Rosemberg, D.B., Maximino, C., Lima, M.G., Abreu, M.S., Giacomini, A.C., Barcellos, L.J.G., et al. (2017). Comparative Analyses of Zebrafish Anxiety-Like Behavior Using Conflict-Based Novelty Tests. Zebrafish *14*, 197–208. https://doi.org/10.1089/zeb.2016.1415.
57. Blaser, R., and Gerlai, R. (2006). Behavioral Phenotyping in Zebrafish: Comparison of Three Behavioral Quantification Methods. Behav. Res. Methods *38*, 456–469. https://doi.org/10.3758/BF03192800.
58. Wong, K., Elegante, M., Bartels, B., Elkhayat, S., Tien, D., Roy, S., Goodspeed, J., Suciu, C., Tan, J., Grimes, C., et al. (2010). Analyzing Habituation Responses to Novelty in Zebrafish (Danio Rerio). Behav. Brain Res. *208*, 450–457. https://doi.org/10.1016/j.bbr.2009.12.023.

59. Pittman, J.T., and Lott, C.S. (2014). Startle Response Memory and Hippocampal Changes in Adult Zebrafish Pharmacologically-Induced to Exhibit Anxiety/Depression-like Behaviors. Physiol. Behav. *123*, 174–179. https://doi.org/10.1016/j.physbeh.2013.10.023.
60. 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. (2016). Exome Aggregation Consortium, Analysis of Protein-Coding Genetic Variation in 60,706 Humans. Nature *536*, 285–291. https://doi.org/10.1038/nature19057.
61. Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and Predicting Haploinsufficiency in the Human Genome. PLoS Genet. *6*, e1001154. https://doi.org/10.1371/journal.pgen.1001154.
62. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The Mutational Constraint Spectrum Quantified from Variation in 141,456 Humans. Nature *581*, 434–443. https://doi.org/10.1038/s41586-020-2308-7.
63. Plenge, R.M., Stevenson, R.A., Lubs, H.A., Schwartz, C.E., and Willard, H.F. (2002). Skewed X-Chromosome Inactivation Is a Common Feature of X-Linked Mental Retardation Disorders. Am. J. Hum. Genet. *71*, 168–173. https://doi.org/10.1086/341123.
64. Wongkittichote, P., Wegner, D.J., and Shinawi, M.S. (2021). Novel Exon-Skipping Variant Disrupting the Basic Domain of HCFC1 Causes Intellectual Disability without Metabolic Abnormalities in Both Male and Female Patients. J. Hum. Genet. *66*, 717–724. https://doi.org/10.1038/s10038-020-00892-9.

The American Journal of Human Genetics, Volume *111*

Supplemental information

Variants in *ZFX* are associated with an X-linked neurodevelopmental disorder with recurrent facial gestalt

James L. Shepherdson, Katie Hutchison, Dilan Wellalage Don, George McGillivray, Tae-Ik Choi, Carolyn A. Allan, David J. Amor, Siddharth Banka, Donald G. Basel, Laura D. Buch, Deanna Alexis Carere, Renée Carroll, Jill Clayton-Smith, Ali Crawford, Morten Dunø, Laurence Faivre, Christopher P. Gilfillan, Nina B. Gold, Karen W. Gripp, Emma Hobson, Alexander M. Holtz, A. Micheil Innes, Bertrand Isidor, Adam Jackson, Panagiotis Katsonis, Leila Amel Riazat Kesh, Genomics England Research Consortium, Sébastien Küry, François Lecoquierre, Paul Lockhart, Julien Maraval, Naomichi Matsumoto, Julie McCarrier, Josephine McCarthy, Noriko Miyake, Lip Hen Moey, Andrea H. Németh, Elsebet Østergaard, Rushina Patel, Kate Pope, Jennifer E. Posey, Rhonda E. Schnur, Marie Shaw, Elliot Stolerman, Julie P. Taylor, Erin Wadman, Emma Wakeling, Susan M. White, Lawrence C. Wong, James R. Lupski, Olivier Lichtarge, Mark A. Corbett, Jozef Gecz, Charles M. Nicolet, Peggy J. Farnham, Cheol-Hee Kim, and Marwan Shinawi

SUPPLEMENTARY CASE REPORTS

Patient 1

The proband is the second child born to non-consanguineous parents. He was born via C section secondary to severe preeclampsia at 32 4/7 weeks of gestation. Pregnancy was complicated by fetal growth restriction, short femur length and left prenatal pyeliectasis. Family history was unremarkable. His older sister is healthy.

At birth, Apgar score was 9/10/10. Growth parameters were at 5th percentile for weight (1650g) and length (41.5cm), and 10th percentile for OFC (30cm). He exhibited a transient neonatal respiratory distress and central hypotonia. Facials features at birth included bulging eyes, coarse facial traits, pointed chin, smooth philtrum, thin upper lip, deep plantar creases, squared finger tips and deep-seated nails.

His physical examination and additional workup revealed hypospadias (operated at 2 years old), bilateral inguinal hernias and horseshoe kidney. He had hepatic angiomatosis and cutaneous haemangiomas. Evaluation of the heart revealed a repolarization abnormality with normal cardiac architecture and normal heart function.

His evaluation by audiology revealed moderate hearing loss and his ophthalmological examination showed hyperopia and astigmatism. Initial MRI and skeletal X-rays were normal.

During first weeks of life, his main medical issues were nephrocalcinosis secondary to hyperoxaluria and hypoglycaemia secondary to cortisol deficiency. He also presented feeding difficulties, persistent hypotonia and joint hypermobility.

Thought the clinical follow-up, he exhibited a global developmental delay compatible with mild intellectual disability after testing. He acquired walking at 3.5 years old with persistent fine motor difficulties. First words were said after 2 years old with 2-word phrases at age 3 years. He did not acquire complete sentences, leading to communication through isolated words and sign language.

He went to specialized school and exhibited initial difficulties during social interaction and intolerance to frustration. The proband still continues receiving speech therapy, psychomotricity and physical therapy. He is depended for some daily life activities (personal hygiene, self-dressing).

Hi height gain slowed around 8 years of age without endocrinal etiology. Last height was around -1.7 SD at 11 years old and weight and OFC remained in the norms.

Several medical conditions appeared, including digestive issues (esophagitis with gastric ulcer, constipation, dysphagia), respiratory issues (asthma), recurrent otitis (ear tube placement) and cutaneous lipomas. Hyperoxaluria remained while cortisol deficiency resolved.

At 8 years of age, an MRI revealed new cerebral features with moderate cerebellum atrophy, choroid plexus cyst and enlarged frontal sub-arachnoid space.

Several non-conclusive genetic investigations were performed including an array-CGH, metabolic workup, and sequencing of HRAS and other genes in the RAS signalling pathway.

Finally, a trio exome found a missense VUS, c.2290C>T p.(Arg764Trp), in the ZFX gene. It is inherited from his asymptomatic mother and asymptomatic maternal grandmother with no other male carrier.

Patient 2

This patient was born at 42 weeks to non-consanguineous parents by emergency caesarean section for fetal distress. He weighed 2.68kg (2nd centile) with relative macrocephaly (50th centile). He was hypotonic and required intubation and ventilation.

At 5 months, he was seen in the genetics clinic. He had a large anterior fontanelle, a high prominent forehead and dysplastic, and low set, posteriorly rotated ears. He had long, down slanting, palpebral fissures, with ectropion of the temporal aspect of his lower lids. He had blocked nasolacrimal ducts and a tear duct repair. He had a bulbous nasal tip and a prominent pointed chin. His hair was fine and slow growing. He had hyper-extensibility of his finger joints, a single palmar crease and small nails. A vascular lesion on the dorsum of his little finger was noted. Over time, he developed red-brown, discrete macules and papules scantily distributed on his face, trunk and limbs. His skin was soft, pale and translucent. He had seborrheic eczema of his scalp for several years. He had delayed tooth eruption and possible extra teeth. He had bilateral undescended testes with bilateral inguinal hernias and an umbilical hernia. He was noted to have scoliosis in his teenage years.

He had delayed development. He sat late but before 1 year and was walking at 2 years, 1 month of age. He couldn't hold a pen at 5 years, 4 months. He had no sounds or words at 2 years, 10 months. His first words were at 3.5 years. At 5 years, 4 months, he had nasal speech which was difficult to understand. He had input from the speech and language therapy and physiotherapy teams at various ages. He was social at 5 years and 4 months and played with other children. At 12 years of age, he had an IQ of 48. He had full time support in school. At age 17, he has new diagnosis of autism. A re-assessment of IQ suggests this is in the normal range (report awaited). He is in college and completing foundation level 2.

He had a cranial ultrasound that showed an abnormality in his ventricles. His brain MRI revealed an unusually large choroid plexus. He has severe sensorineural deafness in his left ear. He had an abdominal ultrasound scan which was normal. He initially had a normal echocardiogram but developed aortic dilatation at 6 years of age. This has remained stable on atenolol. An array CGH in 2014 showed a - VOUS 7q21.13-7q21.13 0.42Mb loss. A buccal smear karyotype was normal – 46XY with no 12p tetrosomy. In 2014, he had normal 15 gene Rasopathy panel testing. He had normal FBN1, SKI, TGFBR1 and TGFBR2 sequencing. The DDD study and 100,000 genomes project initially found no pathogenic results.

He has the X:24229365-24229365 C>T missense variant which was picked up from the DDD study and parental testing confirmed that the variant was *de novo*.

Patient 3

The proband is the first male child to non-consanguineous parents. The pregnancy was complicated by polyhydramnios. He postnatally exhibited feeding difficulties and prolonged neonatal jaundice. He had pyloric stenosis, which required surgery at 6 months of age.

The proband had macroglossia, which required surgery, hypospadias and bilateral cryptorchidism, umbilical hernia and three cutaneous hemangiomas. He exhibited global developmental delay and learning difficulties. His independent walking was at 3 years of age. He also exhibited delays in speech and fine motor skills. He is able to read and write

but with difficulties. He also exhibited recurrent respiratory infections and had mild hypogammaglobulinemia.

At last evaluation at the age of 10yand 6m, his weight was 35kg/+1SD, height was 147 +1.5SD, and OFC was 56 cm (+1.5SD). His main dysmorphic features were large mouth, everted lower lip, thin upper lip, deep plantar creases, and deep-seated nails. Bilateral hallux valgus and supernumerary tooth were also noted.

Because of his clinical course and findings, Simpson-Golabi Behmel and Costello syndromes were clinically considered on the differential diagnosis.

Several non-conclusive genetic investigations were performed including an array-CGH, metabolic workup, and sequencing of HRAS and other genes in the RAS signaling pathway. His trio exome sequencing revealed a de novo missense, c.2312C>T p.(Thr771Met) in the ZFX gene.

Patient 4

This female proband was born at 39w5d gestational age by uncomplicated vaginal delivery. Bilateral inguinal hernias were identified and repaired in infancy. Macroglossia was noted in early childhood prompting a genetics evaluation where an exam revealed additional findings including large hands and feet, prominent lashes, and macrotia. Her workup revealed normal Beckwith-Wiedemann methylation studies, a normal chromosome microarray, and normal urine glycosaminoglycans. Other issues in childhood included recurrent otitis media requiring bilateral tympanostomy tubes and a small, narrow palate that required a palate expander. She also has atypical scoliosis with grade 2 spondylolisthesis at L5 relative to S1 and L5 bilateral spondylolysis.

This individual developed multiple endocrinopathies during adolescence including persistent hypercalcemia. Comprehensive evaluation revealed inappropriately normal PTH level and urine calcium excretion leading to a working diagnosis of familial hypocalciuric hypercalcemia. She also has a history of amenorrhea with evidence of hypogonadotropic hypogonadism. Her first menses was at the age of 15 years and she had a total of 11 cycles before developing amenorrhea. Labs have shown persistently low LH, borderline low FSH, and low estradiol levels with normal testosterone, DHEA-S, prolactin, and thyroid studies. Brain MRI did not reveal causative pituitary findings.

This individual met her developmental milestones on time and did not require any special educational services. She is an excellent student and is attending college. She does have neuropsychiatric challenges including generalized anxiety disorder, ADHD, and autism spectrum disorder. She re-established care with genetics during early adulthood due to her history of hypercalcemia. Exam at the time revealed thick eye brows, down-slanting palpebral fissures, and a prominent chin. Fragile X repeat testing was sent due to her amenorrhea and this returned normal. Trio exome sequencing was sent that was non-diagnostic, but did reveal a heterozygous, *de novo*, candidate variant of uncertain significance in the *ZFX* gene [c.2312 C>T; p.(T771M)].

Patient 5

Proband ZFX05 was born at 41 4/7 weeks gestational age to a healthy non-consanguineous couple with no known family history except for a bicuspid aortic valve in the father. Pregnancy was complicated by decreased fetal movements around the time of his due date. Respiratory effort at birth was weak and there was no heartbeat detected. He received chest compressions, was intubated, and underwent a cooling protocol. He was noted to have macroglossia, organomegaly, and hypo- then hyperglycemia. A clinical diagnosis of Beckwith-Wiedemann was made. Cardiac imaging showed a bicuspid aortic

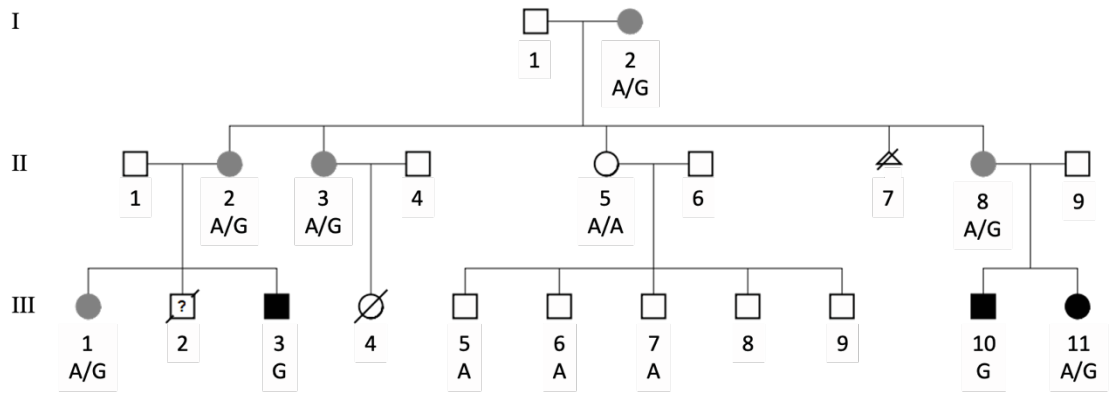
valve, coarctation of the aorta and PDA. He had repair of the PDA and coarctation of the aorta within the first 3 weeks of life and then repair of a pyloric stenosis and inguinal and umbilical hernias. He had a 9 weeks NICU stay. At 9 months of life he had a tongue reduction and then required re-repair of an inguinal hernia and hypospadias repair. He had many episodes of otitis media during childhood and required several PE tubes, an adenoidectomy, and repeat tongue reduction.

Developmental milestones were all delayed. He sat at 1 year of age, walked at 2 years of age, and had his first expressive words at 5 years of age. He was diagnosed with autism and has always been in special classes in school and continues to receive speech, physical, and occupational therapies.

Work-up included array CGH showing a paternally inherited 10q21.3 duplication of uncertain significance and normal Beckwith-Wiedemann methylation studies. Urine organic acids, plasma amino acids, a peroxismal disorders panel, and carnitine levels were normal or inconclusive.

At 10 years of age a whole exome sequencing trio was done showing a de novo ZFX variant of uncertain significance.

Family 6



Pedigree of Family ZFX-6: Dark black circles and squares- affected individuals. Light black circles- carrier females diagnosed with hyperparathyroidism (except for III-1). A is the wild type ZFX allele chrX-24229396 A>G (c.2438A>G:p.(Tyr774Cys)); G is the mutated ZFX allele at same position.

Patient 6A (III:3 in the pedigree): This is a male proband who was born at 39 weeks of gestation via Caesarean section for symptoms of pre-eclampsia. The pregnancy was otherwise uncomplicated. This was his mother's third pregnancy. Amniocentesis was performed and results of karyotype (46, XY) and alpha-fetoprotein were normal. He was macrosomic at birth with a weight of 4.125kg, a length of 53cm, and a head circumference of 37.5cm all above the 90thcentile for gestational age. He was noted after delivery to have dysmorphic and "coarse features" including thick lips, macroglossia and large hands and feet. He was normoglycemic. He had palpable liver and spleen 2cm below the costal margin. A small umbilical hernia, a large right inguinal hernia, undescended testes and hypospadias were also found. Because of the constellation of findings, Beckwith-Wiedemann syndrome was suspected clinically. At 11 years of age, a diagnosis of Simpson-Golabi-Behmel syndrome (SGBS) was made based on an X-linked inheritance pattern in the family,

diaphragmatic hernia in deceased sibling III:2, and III:3's history of neonatal overgrowth, coarse facial features with thick lips and macroglossia, hypotonia and intellectual disability. Proband III:3 had a right inguinal herniorrhaphy with orchidopexy at 2 weeks of age. At 12 months of age, he underwent a reduction surgery for his macroglossia and rerouting of his sublingual salivary ducts for excessive drooling. His hypospadias and umbilical hernia were repaired surgically at 4 years of age. He required multiple sets of tympanostomy drainage tubes for chronic serous otitis media. He also had an adenoidectomy and tonsillectomy. Ultrasound surveillance was recommended for Beckwith-Wiedemann related tumours but these were unremarkable. Developmentally, Proband III:3 had global developmental delay presenting with hypotonia and delayed motor milestones in infancy. He received physiotherapy in infancy and early intervention with physiotherapy, occupational therapy and speech therapy in early childhood. His neuropsychological evaluation at 6 years of age confirmed mild intellectual disability and his cognitive abilities were consistent with those of a three-year-old child. He attended a mainstream primary school with support from an integration aide. His social and self-help skills as well as his behaviours were judged to be satisfactory. He was noted to have mild coordination problems (not specified).

On physical examination at 6 years and 7 months, his head circumference was at the 80%ile, weight at the 50%ile, and height at the 5%ile. His build was described as "bulky", he had a prominent tongue despite a reduction surgery and an "erythematous nevus" over his eyebrows, outer cheeks and chin. He had small longitudinal pits in the skin behind his left ear lobe and helix. He had prominence of the right posterior thoracic wall but scoliosis was not confirmed. He did not have any evidence of hemihypertrophy. Subsequently, a mild thoracic scoliosis was diagnosed radiologically 8y8m with associated short stature (height < 3%ile). Scoliosis surgery was not required.

In addition to scoliosis, III:3's skeletal phenotype included flexion contractures of both elbows (20degrees) and both knees (10 degrees). He had broad hands with mild limitation of extension of the MCP and PIP joints of the fingers of both hands. His DIP joints in the fingers of both hands could hyperextend. He had bilateral hallux valgus deformity, very short 4th metatarsal, and his 4th toe crossed over his third toe bilaterally.

Proband III:3's ectodermal features included abnormal anterior and posterior head hair whorls, sparse eyebrows with lateral thinning and general hypertrichosis. As a child, he had soft skin generally, keratosis pilaris on his trunk and upper arms. His childhood facial findings included naevus simplex over his medial forehead, both upper eyelids and his chin. As a young man, he had seborrheic dermatitis of the forehead and eyebrows. Proband III:3's fingernails and toenails were small in childhood. As a young adult, his fingernails were hyperconvex and angled back with a more acute nail-bed: nail angle.

Proband III:3 had five supernumerary maxillary teeth adjacent to his central upper incisors and medial to his upper canine's preventing the normal eruption of his canines and his lateral upper incisors. These were removed surgically prior to orthodontic treatment. Additional concerns were a crossbite, thickened gums, dental caries and poor oral hygiene. The risk of caries and oral hygiene issues was exacerbated by a dry mouth following the redirection of III:3's salivary ducts.

Proband III:3 was referred for an ophthalmology assessment at 7 years for significant myopia requiring glasses. His corrected visual acuity was 6/12 on the right and 6/24 on the left. Complete absence of the sphincter papillate bilaterally was noted. Proband III:3 had bilateral retinal detachment as a teenager. He had multiple surgeries but vision could only be saved in one eye.

Proband III:3 developed epilepsy that was more difficult to control in his early adult years. His epilepsy presented with a 2-minute non-febrile generalized convulsion at 9 years of age. He was not hypoglycaemic at the time. He had additional generalized tonic colonic seizures in childhood and was treated with carbamazepine. His epilepsy was well controlled until his early 20's when he had more frequent seizures. EEG at 22 years of age for frequent partial seizures showed frequent bilateral interictal activity that was more prominent in the posterior temporal regions. A single generalized seizure was recorded from sleep. Levetiracetam, Lamotrigine and lacosamide treatments were not effective. He was reviewed by a neurologist at 23 years of age and was treated with clonazepam and carbamazepine. By 25 years of age, his seizure control had improved with infrequent dyscognitive and convulsive seizures on a combination of carbamazepine, phenytoin and clobazam.

A striking feature of III:3's phenotype was his long-standing hypothalamic hypogonadism with secondary osteopenia. An initial clinical finding was an absence of body hair persisting into adulthood. At 18 years of age, despite absent body hair, his penis and testicles were noted to be "normal". He did not have gynaecomastia. His bone age was 14 years at a chronological age of 18 years and 3 months. He suffered an ankle fracture at age 16 years following a fall. He had fracture of his left upper humerus following another fall at 21 years of age and a different fracture of his left mid-shaft of his humerus during a seizure shortly after the first humerus fracture healed. His bone density showed a T-score of -3.9 at both his lumbar spine and hip and he was commenced on Strontium ranelate by his GP. He was formally assessed by an Endocrinologist at 22 years of age. He had a "generally hypogonadal" appearance but without gynaecomastia. His left testis was 12.5 mL and his right testis 10 mL in volume. Both were soft in texture. Penile length was 5 cm. Despite his testicular volume, his androgen profile was consistent with hypogonadotropic hypogonadism with baseline serum testosterone levels of 2.2 nmol/L and 2.3 nmol/L and an LH of 2 IU/L and FSH of 4 IU/L. His SHBG was elevated (126 nmol/L, his calculated free testosterone was 16pmol/L and his oestradiol level was not detectable. Cortisol, prolactin, thyroid function, parathyroid hormone level and serum calcium, phosphate, magnesium and vitamin D levels were all normal. Screening for other secondary causes of osteopenia was negative. Mild thrombocytopaenia was thought to be caused by anti-epileptic medications and did not require treatment. Proband III:3's androgen replacement was commenced slowly, initially using testogel, subsequently Axiron transdermal testosterone solution and more recently intramuscular reandron. At 33 years of age III:3's testosterone level was 18.2 nmol/L on Testosterone undecanoate 1000mg IM every 16 weeks.

Proband III:3 had a friendly and happy predisposition as a child. His behaviour in his teenage years was repeatedly noted to be challenging and of concern to his parents. Repeat educational psychology testing placed his full-scale IQ at 60. He was enrolled in a special education secondary school for his schooling for additional support. As an adult in his mid-twenties, Proband III:3's behaviour is much less challenging and no longer a concern for his family. Currently, he attends a supported work program one day a week and music and theatre groups 2 days per week. He lives with his parents and is not in a relationship currently.

Male baby III:2's clinical status is uncertain and DNA was not available for genetic testing due to neonatal demise shortly after delivery in 1986. He presented antenatally with maternal polyhydramnios at 30 weeks' gestation. Ultrasound identified hydrops fetalis and what was thought to be an intrathoracic "growth". Antenatal banded karyotype following amniocentesis was normal (46, XY). No further genetic testing was undertaken at the time. He was delivered at 35 weeks' gestation by Caesarian section. He died 25 minutes after

birth secondary to respiratory insufficiency. A post mortem examination was performed and revealed a large right congenital diaphragmatic hernia with liver herniation. There was cardiac distortion and presumed cardiac compromise with hydrops fetalis, and pulmonary hypoplasia bilaterally, more severe on the right side. At autopsy, syndromic features were not noted nor discussed with the family and it is possible that II:2 had incidental isolated right congenital diaphragmatic hernia. However, it was partly on the basis of III:2's CDH that a possible diagnosis of Simpson-Golabi Behmel syndrome was considered for III:3 and III:10.

Female III:1 died at 38 years and 8 months of age from metastatic colorectal carcinoma that had been diagnosed at 37 years and 2 months of age. She had no symptoms or signs of hyperparathyroidism and her pertinent biochemical studies were normal. She had no other medical problems and had normal childhood development and cognitive abilities.

Female child III:4 had agenesis of the corpus callosum, a possible malformation of cortical development, global developmental delay, epilepsy and severe visual impairment. Her problems were judged not to be related to the problems in her cousins III:2 and III:3. DNA was not available for retrospective ZFX gene testing.

At 34-week gestation, a prenatal ultrasound scan identified a fetal brain anomaly with increased intracerebral fluid filled spaces thought to be dilated ventricles with fetal hydrocephalus. However, this was not confirmed on postnatal imaging. A CT scan at one week of age identified agenesis of the corpus callosum with a high-riding and dilated third ventricle and colpocephaly with dilated posterior horns of the lateral ventricles. Delayed sulcation and reduced grey-white matter differentiation raised the possibility of a malformation of cortical development. MRI was not performed. As a newborn III:4 had a relatively small head with a narrow forehead. She fed normally and was not noted have any abnormal neurological findings. She needed phototherapy for unconjugated hyperbilirubinemia. She was seen by an ophthalmologist at 10 weeks of age for reduced visual acuity. She was not reacting to light and her pupils constricted sluggishly. Her eyes were both normal in structure with the exception of optic disc pallor and her left optic disc being smaller than her right. These changes were not thought to be significant enough to diagnose optic nerve hypoplasia. She was reviewed by a clinical geneticist at the same time. She was feeding and sleeping well. Her head circumference was 39cm (25^{th} -50^{th} centile) with open fontanelles and she had good head control. She had mild dysmorphic facial features including mild hypotelorism. She had pectus carinatum with "some curving of the dorsal spine". A clinical diagnosis of holoprosencephaly was made and her parents were counselled that III:4 was at the mild end of the holoprosencephaly spectrum. III:4 was fostered at approximately 6 months of age and there are no medical records of her findings thereafter. The family reports that she subsequently struggled with epilepsy and was not able to sit independently, walk or talk at 18 months of age. The family did not have significant contact with III:4's adoptive parents after this time.

Female II:2 (Mother of Patient 6A) had regular health and development in early childhood. She was diagnosed with scoliosis at 9 years of age and managed with a Milwaukee brace for three years. The angle of her scoliosis curve exceeded 45 degrees despite bracing and she had surgery with Harrington rod fixation at age 12 years. II:2 had cognitive abilities in the normal range (not formally tested) and attended mainstream schooling. She did not complete her final year of schooling but married at a young age and had her daughter, III:1, at the age of 18. She and her partner had two more male children over 5 years, including III:2 who died shortly after birth due to CDH and the family proband III:3. II:2 stayed home to raise her children and to manage III:3's needs. She reports that back pain was a chronic issue and she had secondary problems with right hip pain as a result of her spinal curve. II:2 joined the workforce at 43 years of age as a supermarket cashier. She had increased back pain as a result and needed time off work intermittently in her 40's. Ongoing symptoms resulted in an extensive two stage anterior and posterior spinal fusion using rib grafts at 50 years of age. She experienced post-operative back pain, pain and paresthesia related to the

harvesting of her ribs, unexplained right sided abdominal, and right foot paresthesia in the absence of bladder symptoms. She experienced restricted movement of her spine but also limited tolerance for sitting and reduced effort tolerance. She was not able to work and received a disability pension. At 51 years of age, II:2 saw a gastro-enterologist for persistent abdominal pains and constipation in the setting of opioid analgesia requirements. Endoscopy and laparoscopy were negative. Blood tests at the time identified hypercalcemia and she was referred to an endocrinologist with bone aches, lethargy, short term memory loss and hyperparathyroidism. II:2 was assessed for parathyroidectomy. Pre-operative blood biochemistry showed calcium 2.64 mmol/L (RR 2.15-2.65), corrected calcium 2.70 mmol/L (RR 2.15-2.65), phosphate 0.8mmol/L (RR 0.8-1.4), ALP 149 U/L (RR 30-115), and otherwise unremarkable LFTs. PTH was 15.6pmol/L (RR 1.5-7.6). A 24hr urine collection showed a urinary calcium excretion of 12.5 mmol/24hr (RR 2.0-7.5). Pre-operative ultrasound neck imaging identified a 12mm left parathyroid lesion inferior to the left thyroid lobe. A parathyroid Sestamibi scan noted retention in left and right inferior poles of the thyroid gland. Ultrasound images of the thyroid gland revealed bulkiness of both thyroid lobes with multiple nodules in the right lobe, left lobe and isthmus, and small cysts in the left lobe totaling 16 in number. The largest two nodules were in the left isthmus and 19 and 16mm in size, respectively. Fine needle aspirate from the largest left nodule and the largest right nodule (11mm) showed thyroid tissue with benign changes.
Bilateral neck exploration and parathyroidectomy was performed.
The left inferior parathyroid was markedly enlarged (640mg), contained packed sheets of chief cells. The right superior parathyroid was moderately enlarged (290mg) with predominantly shaved chief cells. The right inferior parathyroid was mildly enlarged. Atypia or malignancy was not detected. Histology was judged to be in keeping with parathyroid gland hyperplasia. The left superior parathyroid gland was not identified. Postoperative biochemistry 3weeks after surgery showed persistent hypercalcemia (corrected calcium 2.97 mmol/L) and increased parathyroid hormone (9.8pmol/L). Repeat localization studies placed the 4th parathyroid gland at/ within the left lower pole of the thyroid gland. The gland was removed approximately one year after the initial surgery. It was moderately enlarged (252g) with predominantly chief and clear cells seen on histology. Blood biochemistry had normalized 2 weeks later. Osteopenia/ osteoporosis was diagnosed 5 months post-surgery on DEXA scan with left femoral neck BMD 0.698 g/cm2 (T score -1.4, in the osteopenic range) and left forearm BMB 0.494 g/cm2 (T score -3.3, in the osteoporotic range). She had a fractured toe with significant trauma. She was treated with combined vitamin D3 and calcium tablets and magnesium but not bisphosphonates. In addition to calcium and vitamin D supplementation, II:2's current regular medications at 57 years of age include levodopa / carbidopa for restless legs, a statin for hypercholesterolemia, Gabapentin for nerve pain, paracetamol and turmeric for musculoskeletal pain, and vitamin C, CoQ10 and zinc supplements for wellbeing. II:2 has hyperconvex and ridged finger-nails and small toes nails similar to her son III:3's. She has an interpupillary distance of 5.8cm with a head circumference of 55cm and is not considered to be dysmorphic.

Female II:3 is a 56-year-old woman who was diagnosed with hyperparathyroidism when she was 50 years old. He initial symptoms included lethargy but she had no constipation or kidney stones. She underwent parathyroidectomy of 2 parathyroid glands.
She fractured her right femur 10 years ago after a fall from standing height. Her other medical issues include osteoarthritis, depression, fibromyalgia, and osteoporosis.

Female II-8 is the mother of Probands ZFX06_B and ZFX06_C. She is 48 years old and was diagnosed with hyperparathyroidism at age 43 when she presented with lethargy. She underwent parathyroidectomy of 3 ½ glands. Her past medical history is unremarkable. She had no developmental or cognitive concerns. She currently receives no medications.

Male Patient 6B (III:10 in the pedigree) was delivered vaginally following face presentation at 39 weeks of gestation. Mild polyhydramnios was noted late in the pregnancy but routine ultrasound scans were otherwise unremarkable. He was normally grown with a birth weight of 3,315g (50%ile), a length of 51cm (50-90%ile) and a head circumference of 36.5cm (90%ile). Facial swelling settled over 24 hours but his macroglossia persisted. He was treated with triple phototherapy at 36 hours for neonatal jaundice (unconjugated bilirubin 315umol/L, conjugated 5umol/l). He and his mother were O positive and direct Coombs was negative. A left inguinal hernia was noted. He was transferred to a tertiary neonatal unit on day of life 3 with the combination of the hernia, jaundice, breathing difficulties and dysmorphic facial features. On admission, he was treated with oxygen, antibiotics and phototherapy. Echocardiogram showed a PDA, a PFO, persistent pulmonary hypertension and poor left ventricle contraction that was treated with oxygen and dobutamine for 6 days and captopril for 10 days. Capillary glucose was monitored and hypoglycemia was not recorded. Jaundice resolved on day 6. He was assessed by a clinical geneticist and given a diagnosis of Beckwith-Wiedemann syndrome. A renal ultrasound scan showed large but otherwise normal kidneys (R5.6cm/Lt5.9cm). His macroglossia caused feeding difficulties that improved with a Habermann teat feeder.

His banded karyotype was normal. He was diagnosed with pyloric stenosis at four weeks of age after he had vomiting with increasing frequency for 72 hours. He had a pyloromyotomy and bilateral inguinal herniotomies.

III:10 was evaluated in genetics at 4 years of age. He had mild global developmental delay. He had walked at 21 months of age and could run and climb stairs at 4 years. His first words were at 18 months and he was speaking in full sentences by four years of age. As a young child, III:10 attended an early childhood intervention program for speech and occupational therapy. Bilateral chronic serous otitis media had been treated with ear ventilation tubes. He had mild atopic eczema and infrequent episodes of wheezing. He attended a regular kindergarten. His height was on the 75th centile, his weight was 2 kg above the 97th centile and his head circumference was just above the 98th centile. A diagnosis of Simpson-Golabi- Behmel was considered. However, GPC3 sequencing was normal.

He was seen again at 8 years of age with dysmorphic features and mild global developmental delay. He had a number of morphological features in common with his maternal male cousin III:3 and his sister III:11. His speech was not clear as a result of his macroglossia. He could feed himself, color and paint. He struggled with scissors and his drawings were immature. He attended a mainstream primary school with an integration aide for support. Formal testing recorded his full-scale IQ as 67. His growth followed his previously noted centiles. His medical problems included persistent macroglossia, bilateral supernumerary upper jaw teeth (above the premolar/ molar tooth junction), myopia without retinal detachment, enlarged adenoids and tonsils treated surgically, scoliosis, and an admission for pneumonia.

III:10 attended a special education school for his secondary schooling. He had delayed puberty onset (Tanner stage 2 pubic hair and 5-6ml testicle volume at 14y6mo) with a growth spurt, concordant bone and chronological ages and some early blood testosterone (0.9 nmol/L) and LH (1.0 IU/L) level changes as assessed by an endocrinologist pre-scoliosis surgery. When reviewed at 15y3mo of age, puberty had progressed with 12 ml testicle volumes and Tanner stage 3 pubic hair development and he was discharged from endocrinology follow up. As age 16 years, he had a staged anterior thoracolumbar release and posterior instrumentation and fusion procedure for progressive scoliosis. Two subsequent lumbar CSF leaks were treated with blood patches. Chronic rhinosinusitis was treated surgically by endoscopic sinus washout, nasal septoplasty, inferior turbinoplasty, nasal polypectomy and partial adenoidectomy. Glaucoma was diagnosed and treated medically. As a young adult he currently has a mixed weekly program including part time work at a disability service in a supported environment two days a week, a day of performing art and dance, a day volunteering in an animal shelter and a day for therapies (physiotherapy and speech therapy) and medical appointments.

In addition to initial clinical diagnoses of familial Beckwith-Wiedemann syndrome and Simpson-Golabi-Behmel syndrome, a diagnosis of FG syndrome had been considered at 8 years of age but ruled out based on careful review of his clinical course and facial features.

Patient 6C (III:11 in the pedigree) is an 18-year-old female who completed secondary school and who works as a waitress. Fetal macrosomia and polyhydramnios were noted antenatally. She was delivered at 38 weeks of gestation by emergency Caesarian section following maternal antepartum hemorrhage due to abruptio placentae and an abnormal cardiotocograph. She required resuscitation (initial Apgar scores of 2^{1}, 5^{5} and 5^{10}) and 48 hours of ventilation for blood aspiration with right lung atelectasis. She was noted to have coarse facial features, macroglossia and macrosomia and clinically diagnosed with Simpson-Golabi-Behmel syndrome in a manifesting female with reference to the same clinical diagnosis in her older male sibling. She had neonatal jaundice that did not require treatment (peak bilirubin of 249/8 (umol/l) on day 5 of life). She had cyanotic episodes observed at 2 weeks of age that were attributed to her macroglossia causing intermittent upper airway obstruction documented on polysomnography. Her obstruction improved when sleeping prone and she did not require surgery. GERD was diagnosed in the neonatal period by pH study and managed with thickened feeds. Her auditory brainstem response was normal for both ears. She attended mainstream schooling and had a part time integration aide during primary school. She struggled with mathematics. She attended mainstream secondary school without an aide and completed her senior education certificate in a technical stream. As a young adult, she has a learner driving license and is aiming for her driver's license. She lives with her parents.

Female I:2 is a 76-year-old woman who was diagnosed with hyperparathyroidism and underwent right inferior minimally invasive parathyroidectomy when she was 73 years old. Her medical history is otherwise unremarkable.

Methods: Whole genome sequencing (Complete Genomics) was performed on a single individual from family 6 (Patient 6B). Single nucleotide and structural variants were filtered against public databases (gnomAD, minor allele frequency (MAF) < 0.0002; Complete Genomics 69 genomes MAF = 0) and based on an X-linked inheritance pattern. The coding c.2321A>G variant was segregated through gDNA extracted from blood of consented family members by Sanger sequencing.

Patient 7

This is a male patient who was born at term via spontaneous vaginal delivery to non-consanguineous parent. He has 2 healthy older siblings. His birth weight was 2.75kg. The pregnancy was complicated with gestational diabetes mellitus, requiring insulin treatment, and polyhydramnios, which was detected at 23 weeks of gestation. Labor was complicated by shoulder dystocia with low Apgar scores. He was admitted to neonatal ward as he had poor sucking, cleft palate, hypotonia, coarse and dysmorphic features. He had poor feeding since birth, requiring orogastric tube feeding, anti-reflux formula milk and anti-reflux medication. Due to feeding difficulty, a magnetic resonance imaging (MRI) of brain and spine was performed on day 6 of life, and it revealed subacute left parietal and posterior fossa subdural haemorrhage, right parieto-occipital subarachnoid haemorrhage, and bilateral choroid plexus cysts. This was treated conservatively.
He had stridor and required non-invasive ventilation since 1 week of age. Airway assessment revealed laryngotracheomalacia and mid-tracheal stenosis. He underwent aryepiglottoplasty surgery and bilateral inguinal hernia repair at 3 months of age. He was discharged home with non-invasive ventilation at 4 months old. Besides the history of

bilateral middle ear effusion, his hearing and eye assessment were reported to be normal. He stopped using non-invasive ventilation and orogastric tube feeding at one year old. He was doing well without frequent hospitalisation after that.
He had global developmental delay, for which he required speech therapy and occupational therapy since infancy. At his last assessment at age 9 years and 10 months, he has mild intellectual disability. He is attending special education classes. He can communicate in short sentences and interacts well with friends in school. He understands a lot and able to identify many objects. He is independent for his daily activities. His weight and height were at the 75th centile and head circumference was at the 50th centile. Physical examination showed prominent forehead, hypertelorism, downslanting palpebral fissures, thin upper lip, prominent jaw, low set posteriorly rotated ears, pectus carinatum, reducible umbilical hernia, lax finger joints, hallux valgus bilaterally. He has a normal parathyroid hormone level, echocardiogram, kidney scan, chromosomal microarray and metabolic workup.
Trio whole exome sequencing was performed using NovaSeq was performed at Takara Bio Inc. (Kusatsu, Japan).

Patient 8

Patient ZFX08 is a 13-year-old male with global developmental delay, dysmorphic features, hearing loss, congenital heart defects, Pierre-Robin sequence (cleft palate repaired), feeding difficulties (G tube dependent), chromic respiratory failure (tracheostomy-dependent), and hyperparathyroidism complicated by hypercalcemia requiring parathyroidectomy at age 13 years.

He was at 35 weeks of gestation via C-section with low Apgar scores and pregnancy was complicated by oligohydramnios, fetal growth restriction 2 vessel cord. His birth weight was 1.96 kg, and his birth length was 48 cm. He did not pass the hearing newborn screening. The patient was in the neonatal intensive care unit for 4 months because of respiratory distress and feeding problems. He underwent tracheostomy at 3 weeks of age. He exhibited feeding difficulties poor weight gain during the first few months of life. He was found postnatally to have aortic coarctation, which required 2 repair procedures and he remained with mildly dilated ascending aorta. His subsequent echocardiography revealed mildly redundant mitral and tricuspid valves. He required hearing aids for his conductive hearing loss. His sleep study showed obstructive events. He was diagnosed with gastroesophageal reflux and feeding difficulties, for which he had G-tube placement and Nissen fundoplication. He had 2 corrective eye surgeries for strabismus and continues exhibiting myopia and astigmatism and was diagnosed with bilateral posterior polar cataracts. He is status post cataract extraction and intraocular lens and Yag in both eyes.

Patient ZFX08 developed hypercalcemia secondary to primary hyperparathyroidism. He required 2 ½ gland parathyroidectomy at age 13 years and pathology was remarkable for one hypercellular parathyroid tissue with no adenoma. He was found to have bilateral non-obstructing renal calculi but his renal function studies are normal.

He exhibits recurrent tracheobronchitis and asthma. He failed decannulation but currently is not mechanically ventilated and doesn't require oxygen supplementation.

He underwent umbilical hernia reduction and inguinal hernia repair and is followed for chronic mild thrombocytopenia of unknown etiology. His thrombocytopenia panel was negative.

Patient ZFX08 started exhibiting significant motor and language delays since 1 year of age. He started sitting unsupported around 3 years, crawling 3 years and walking at 4 years but continues requiring some support with walking. He started using utensils at 4-1/2 years. He

is nonverbal and currently exhibits global developmental delay, autism and significant behavioral concerns, specifically aggression and agitation. He is not potty trained yet. He attends special education classes at 7th grade. He receives physical, occupation and speech therapies as well as ABA for autism

Family history is significant for a brother who died at 4 months of age due to multiple complications of extreme prematurity (24 weeks gestation) including liver failure, NEC, sepsis and respiratory failure. No genetic testing was performed for him. The mother had a history of delayed speech and needed speech therapy. She underwent 3 surgeries for parathyroid glands due to hyperparathyroidism. Histology showed parathyroid hyperplasia. She had a history of chronic anemia due to bleeding gastric ulcers. She underwent an excision of a benign tumor on the back of her knee (fibroma). The proband's father had a history of bipolar and ADHD. The parents are of Northern European ancestry.

On most recent physical examination, he was small for his age: his weight was 27.2 kg (<1 %ile; Z= -2.94), height 127 cm (<1 %ile; Z= -3.61) and OFC was 52.3 cm (12 %ile; Z= -1.15). Tracheostomy and G tube were in place. He had dysmorphic features including relative macrocephaly, broad forehead with metopic ridging, low anterior hairline, smooth and long philtrum, thin upper lip, micrognathia, thickened helices with asymmetric ears that were relatively large. Pectus excavatum was noted. He had squared finger tips, symmetric abnormal creases in feet, deep seated toenails, hallux valgus deformity. Small phallus and small scrotum sac with retractable testes. He had truncal hypotonia and stiffness in both elbows.

Results of initial exome sequencing were negative. However, exome reanalysis revealed that the patient is hemizygous for the p.R786Q (c.2357G>A) variant in the X-linked gene ZFX.

Patient 9

Proband ZFX09 is an 18-year-old biological female who is now a transgendered male and was the third-born child of non-consanguineous parents. He was born at 38-week gestation by induced vaginal delivery due to pregnancy-induced hypertension. There was a nuchal cord, but Apgars quickly returned to normal. At two months of age, he had an echocardiogram due to possible murmur and the study was essentially normal. Prominent coronary arteries without obstruction or aneurysm were noted but repeat study at 13 months of age was normal with normal coronary arteries noted.

At 3.5 months of age, the patient had herniorrhaphy for an umbilical and two epigastric hernias. He had seven hemangiomas of infancy on skin, scalp and liver. A right scapular hemangioma and a right parietooccipital scalp dermoid were concurrently excised at the time of the herniorrhaphies.

The patient has a history of longstanding eustachian tube dysfunction and was first seen in ENT at 12 months of age due to a history of recurrent episodes of otitis media and bilateral middle ear effusions, at which time he underwent bilateral tympanotomies and insertion of middle ear vent tubes. At 15 months of age, the patient was re-evaluated in ENT due to snoring at night and restless sleeping. Vent tubes were replaced and a partial selective adenoidectomy was performed at 16 months of age. The adenoids were 4+ and almost totally obstructing the posterior choanae. The patient's breathing initially improved and there was one episode of tonsillitis. The patient was seen again in ENT and the impression was dysphagia, for which speech therapy consultation was recommended.

The patient achieved developmental milestones as follows: sat alone at 6 months, crawled at 9 months, walked with support at 11 months, walked alone at 15 months. Speech/language milestones were difficult to evaluate. The patient had avoidance of meat and other difficult to chew and swallow textures due to choking, which could be severe enough to result in vomiting. He had a tonsillar voice quality which impacted speech acceptability and intelligibility, with an occasional slight wet sound to phonation. Speech was essentially unintelligible to the unfamiliar listener on first production. Speech output was reduced. ENT impression was that the patient manifested a significant feeding/swallowing problem related to oropharyngeal airway restriction secondary to significantly enlarged palatine tonsils and possibly adenoids.

A sleep study performed at 22 months of age showed evidence of sleep-disordered breathing characterized mainly by hypopneas associated with moderate hypoxemia. There was elevated pCO2 suggestive of possible obstructive hypoventilation. The patient underwent an adenotonsillectomy for upper airway obstruction as well as dysphagia associated with tonsillar hypertrophy.

At 2y3m of age, the patient was seen by dermatology for a lymphatic malformation of the left axilla. At 2y6m of age, the patient had intermittent rectal prolapse and constipation. A solitary rectal polyp was removed during colonoscopy. At 4y10m of age the patient was seen by Genetics due to history of multiple hemangiomas and lymphangiomas. Clinical exam noted long and down slanting palpebral fissures, upturned nose, high arched palate, normal uvula, small chin, and hypoplastic fifth toenails bilaterally. Chromosome analysis was 46,XX.

The patient was noted to have prolonged bleeding with cuts and nosebleeds, and testing by Haematology was negative. Concern for possible Ehlers-Danlos syndrome was raised. The patient was noted to have a history of hypermobility and low tone.

The patient underwent reduction of lingual tonsillar hypertrophy and expansion pharyngoplasty for low lying palate and had another abnormal sleep study, showing evidence of residual obstructive sleep apnea. He followed up in Genetics at age 5y10m, at which time clinical exam noted prominent nose, slightly cupped ears, pectus excavatum and flaring of ribs inferiorly, hypermobility of finger, knees, and thumbs, skin showed some hyperextensibility. The geneticist felt that the patient's clinical features were consistent with Ehlers-Danlos syndrome type 3 (hypermobility type), and the history of vascular malformation suggested the possibility Klippel-Trenaunay-Weber syndrome but insurance did not authorize testing. DNA microarray was recommended and came back normal.

At age 7y, the patient was seen in the ER for acute onset of left-sided chest pain described as sharp and constant, and chest wall tenderness on physical exam led to consideration of costochondritis. The patient was seen repeatedly in the ER for chest pain.

He had septal cautery for epistaxis as well as re-excision of recurrent lymphatic malformation from left axilla and excision of melanocytic nevi from temporal scalp, periumbilical area, and posterior scalp.

The patient underwent neuropsychology evaluation at age 10y and his general intellectual abilities were average. He demonstrated relative weaknesses in visuo-motor coordination and executive functioning skills related to planning and decision-making. The weaknesses noted were not at a level consistent with a diagnosis, and the deficits described were likely related to difficulties in learning math. Re-evaluation occurred at age 15y, and the patient was felt to be at risk for a diagnosis of a specific learning disorder in math.

The patient underwent total thyroidectomy at age 13y. Pathology showed a focus of papillary thyroid cancer classified by ATA pediatric thyroid cancer as low risk. He underwent remnant ablation and there was no cervical lymphadenopathy. At age 15y the patient had surgical removal of a neck mass and histology was consistent with a parathyroid adenoma. He required a short course of calcitriol and calcium carbonate and has since maintained eucalcemia without any medications. His calcium baseline is now normal whereas he previously had slight hypercalcemia before the parathyroid adenoma was removed.

Other medical concerns include congenital QT syndrome due to a parentally inherited *KCNH2* variant, bilateral mixed conductive and sensorineural hearing loss, cyclic vomiting syndrome, gastroparesis, gastroesophageal reflux, chronic abdominal pain, chronic headaches, migraine, idiopathic intracranial hypertension, proteinuria and microscopic hematuria, POTS, euthyroid autoimmune thyroiditis, familial hypocalciuric hypercalcemia, generalized anxiety disorder and major depressive disorder. He also had a year long history of fatigue, nausea, dizziness, shortness of breath, and rapid weight gain. A core biopsy was performed of a lesion on the left 10th rib and showed benign/reactive fibro-osseous lesion with no evidence of malignancy. He has had a longstanding history of generalized body and bone pain.

The patient underwent WGS trio analysis in 2018 was found to have a *de novo* variant in ZFX (R786Q) through research analysis of WGS raw data. A paternally inherited KCNH2 c.526C>T (p.Arg176Trp) was also identified in this patient, consistent with a family history of prolonged QT interval.

Genome sequencing for patient 8 and parental samples was conducted through Illumina (Medical Genomics Research, Illumina Inc) using the Illumina NovoSeq at an average of 40 x coverage and evaluated single nucleotide variants (SNVs), small indels, copy number variants and mitochondrial DNA SNVs.

Patient 10

Proband ZFX10 is a 10-year-old male who was the first-born child of his non-consanguineous parents. His mother was 20 years old at the time of conception. There were no teratogen exposures. During the pregnancy, nonspecific findings on ultrasound were found including a 2-vessel cord and enlarged cisterna magna. Owing to poor fetal growth, there was an induction of delivery at 35 weeks gestational age. Birth weight was 1860 grams, reflecting fetal growth restriction, and Apgar scores were 6 and 9, at 1 and 5 minutes, respectively. At birth, other features were noted including bilateral post axial polydactyly of the feet and hypospadias. No specific syndrome was apparent at the time on physical exam.

His early childhood was complicated by a vein of Galen thrombosis and bilateral subdural bleeds after a minor trauma. He was also diagnosed with congenital nystagmus in infancy. He has mesocardia and a history of a ventricular septal defect, now closed. He also has bilateral hydronephrosis. He has a history of mild to moderate developmental delay, and now probable mild intellectual disability with ADHD. However, on formal assessment at 11 years of age, he does not have evidence of autism. In follow up, his height is below the 3rd percentile with head circumference at the 55th percentile. He has a distinctive facial appearance characterized by turricephaly, a long face, eyebrows that are sparse laterally and full medially with marked synophrys. He has short and upslanting palpebral fissures with remnant epicanthal folds, a long nose, prominent columella and relatively indistinct philtrum. After birth, he had normal chromosomal microarray. Over the course of his childhood, other

diagnostic investigations included normal 7- dehydrocholesterol and a ciliopathy sequencing panel (Blueprint Genetics) both nondiagnostic. Calcium and parathyroid hormone levels are normal in the patient and his mother. The family was then enrolled for clinical trio exome sequencing (Blueprint Genetics) with the detection of a novel maternally inherited frameshift variant in *ZFX*.

Patient 11

Proband ZFX11 was born at 37 weeks' gestation by Caesarean section. He spent 30 days in the NICU following delivery with APGARs of 0 at 1 minute and 5 at 5 minutes. He had an extensive workup in the NICU with concern for sepsis, was noted to have feeding difficulties, and required a g-tube while admitted. While admitted, he was observed to have a mild/moderate central renal collecting structure prominence of questionable significance and an increasing bilateral hydronephrosis with moderate right and moderately left pyelocaliectasis. His g-tube was removed at discharge.

He was noted to have mild developmental delays early on. Fine motor weakness, gross motor delays, and history of speech delay are all noted in this patient's history. He required surgical intervention for bilateral inguinal hernias (2010) and for superior oblique palsy of the right eye (2019). He was also noted to have left sided strabismic amblyopia. He has also had ventilation tubes inserted in bilateral ears and required removal of both his tonsils and his adenoids.

Over the course of his genetics workup, he was evaluated by several different geneticists and had multiple genetic tests performed. A chromosomal microarray and a next generation sequencing analysis of 23 genes associated with the RAS pathway were completed in 2015, both of which were normal.

Physical examinations revealed multiple distinctive features, though nonspecific for a known genetic syndrome. Features noted on his physical examination included asymmetry of the face, hypertrichosis to the forehead, a prominent forehead, thick hair, a low posterior hairline, synophrys, thick eyebrows, low-set posteriorly-rotated cupped ears (left>right), leftward nasal deviation, high palate, retained baby teeth, a short and asymmetric neck with prominent right posterior musculature, a grade II-III/VI systolic murmur best heard over the left sternal border, jagged nail edges, and slightly decreased extension to his bilateral elbows.

He was evaluated through the school system in 2015 and was noted to have a level of intellectual functioning within the average range, although with varied abilities, and his academic achievement was observed to be adequate aside from math, with his adaptive functioning in the average to above average range. Following evaluation by Cardiology in 2016, he was found to have a small hemodynamically insignificant muscular ventricular septal defect. In 2018 he was evaluated through Paediatric Psychology and was noted to have a full-scale IQ of 76 in the 5th percentile and a 95% confidence interval of 71-83, denoted as in the borderline to low average range. He was worked up for a palpable lump of the right popliteal region in 2019 with bilateral superficial varicosities without evidence of thrombophlebitis found.

In 2020 he had whole exome sequencing trio performed through the Greenwood Genetic Center, which identified a de novo variant in *ZFX*. Following this, PTH and Calcium levels were tested and were normal.

Patient 12

This 8-year-old boy was born at 38+5 weeks by vaginal delivery following an unremarkable pregnancy. He was the first child to non-consanguineous parents. His mother has a clinical diagnosis of oculocutaneous albinism and Autistic Spectrum Disorder (Pedigree Figure). His father has no medical problems. He has a younger brother who is now 6 years of age and has nystagmus with abnormal visual evoked potentials but his development has been normal.

Born at 2.58kg (-1.70 SD), he struggled to feed and required nasogastric feeding support for the first week of life. He otherwise made good progress. At the age of 7 months, he was noted to have nystagmus and small nails and was hence referred to genetics.

He sat at 11 months, started commando crawling at 12 months but did not walk independently until 23 months. He babbled by 12 months but did not say single words until two and a half years. He started speaking in sentences by 4 years, although his pronunciation was poor due to his protruding tongue.

He has several autistic traits, including intolerance to change and lack of understanding of social norms, although does not have a formal diagnosis. He has sensory processing difficulties and is under dental review as he has a tendency to push his tongue against his bottom teeth, which has resulted in wearing down of the secondary roots. He attends a specialist school for the visually impaired and his English and Maths skills are currently at Key Stage 1 (5-7 years) whilst all other subjects are at Key Stage 2 level (7-11 years). He has recently developed tightness in his quadriceps, which is worse when he has been sat for long periods of time. He has regular physiotherapy but this does not limit his ability to walk or exercise.

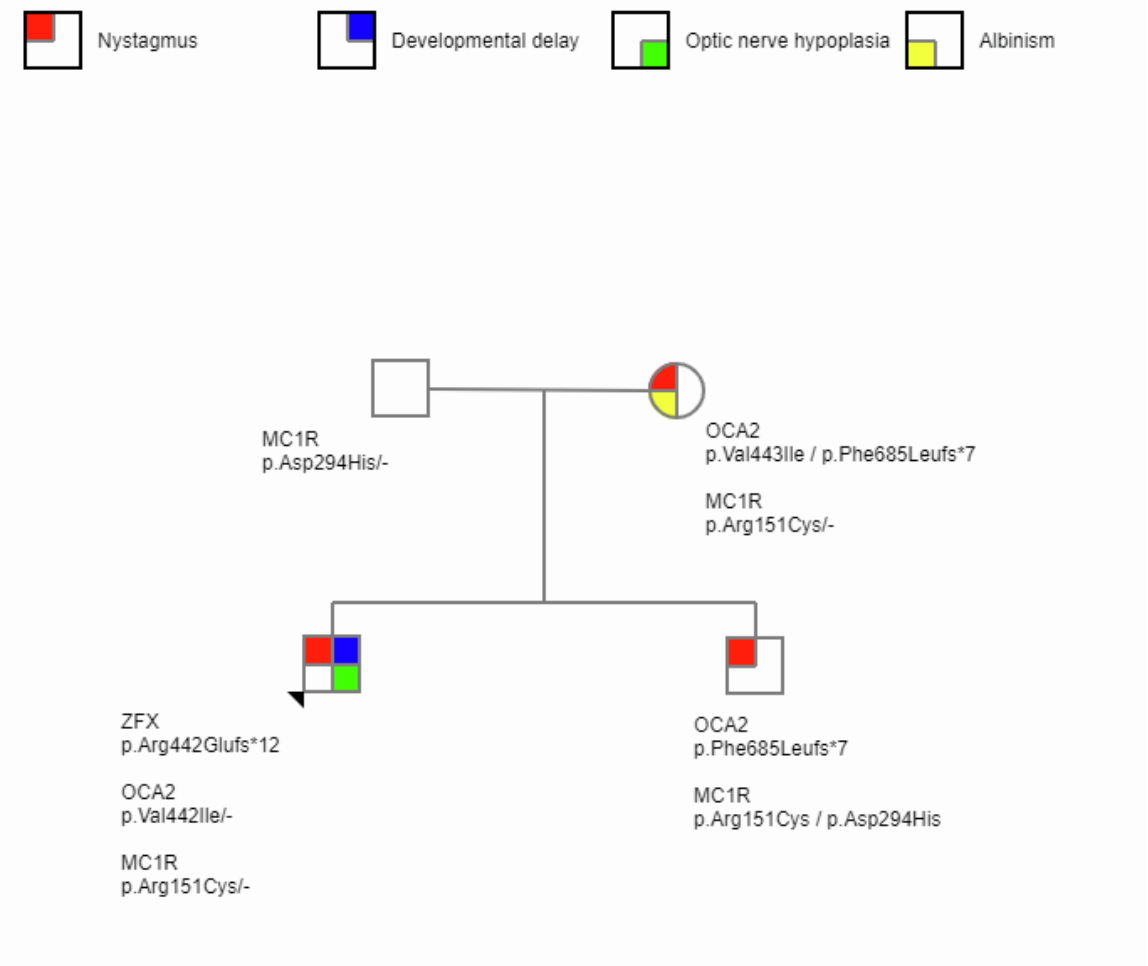
On physical examination at 7 years of age, hypotonia, small deep-set nails, tongue protrusion, long philtrum and deep palmar creases were noted. Appearance of blepharophimosis. He had normal anterior segments of both eyes and clear corneas. Well pigmented irides and well pigmented retinae but small grey optic discs, indicating optic nerve hypoplasia.

An MRI brain aged 1 year showed thin optic tracts and chiasm, a relatively small pituitary, a retrocerebellar arachnoid cyst and periventricular leukomalacia. An ultrasound of the renal tract was normal at this time also. His arginine stimulation test for growth hormone was low at age 5 years and he was commenced on growth hormone injections.

Karyotype and chromosomal microarray were unremarkable. He was recruited to the 100 000 Genomes Project where no causative variants were identified using panels: Infantile nystagmus v1.3, Ocular and oculocutaneous albinism v1.21, Optic neuropathy v2.1, Retinal disorders v2.6, IUGR and IGF abnormalities v1.30 and Intellectual disability v3.3.

Subsequent panel agnostic re-analysis identified a *de novo* frameshift variant in *ZFX* (NM_001330327.1:c.1322_1323dupGA, p.Arg442Glufs*12). This variant is absent from population databases, occurs in exon 10/10 and is likely to escape nonsense-mediated decay, although this is predicted to truncate the C2H2 zinc fingers 2-13. C2H2 fingers 11-13 have recent been shown to be essential for DNA-binding of ZFX and for its subsequent function as a transcriptional activator.

We also identified likely pathogenic and pathogenic variants in both *OCA2* and *MC1R* in the mother and brother of the proband, which explain their ocular albinism, although the proband is unaffected in this respect (see pedigree).



Patient 13

The patient is a 24-year-old man. He is the younger of two sons born to healthy, unrelated Caucasian parents. There is no significant family history of intellectual disability.

At 12-week gestation increased nuchal thickening and mild pericardial effusion was noted on ultrasound scan. CVS testing showed normal karyotype. Pericardial effusion persisted on antenatal scans at 25 and 28 weeks. At 32 weeks his mother suffered a large antepartum haemorrhage. He was born by Caesarean Section at 33 week-gestation due to maternal pre-eclampsia, with a birth weight of 2.01kg (50th centile).

He had prolonged neonatal jaundice and feeding difficulties, requiring admission to special care baby unit (SCBU) for 4 months. He underwent bilateral inguinal hernia repair at 2 days of age and repair of pyloric stenosis at 6 weeks. There was initial unconjugated hyperbilirubinaemia, followed by cholestasis. A liver biopsy showed evidence of giant cell hepatitis, which gradually resolved throughout early childhood.

At 11 years, a routine follow-up abdominal ultrasound showed a large suprarenal mass. This was found to be a benign suprarenal ganglioneuroma, for which he underwent complete surgical excision.

His early developmental milestones were all delayed, with particular delay in expressive language. He was late to sit and walked independently at 25 months. He has intellectual disability and attended Special School. At 12 years of age his verbal and non-verbal IQ were

in the 3-4 year age range. He is currently at college studying life skills and enjoys playing basketball and football.

He has required surgery for hypospadias. He had a large patent ductus arteriosus and small atrial septal defect on echocardiogram which closed spontaneously. He also suffered from recurrent respiratory infections and was found to have mild IgG and IgM deficiency.

Ophthalmological review was reported as showing partial ocular albinism and astigmatism. He also has very dry eyes for which he uses regular eye drops. He has bilateral moderate conductive hearing loss for which he had three sets of grommets in the past and has been prescribed hearing aids.

He remains under the care of the adult immunology service but no longer has recurrent chest infections. He takes Solifenacin for night-time enuresis. There is no history of seizures and he is otherwise generally healthy.

His final height is 173 cm (28^{th} centile) and head circumference 59.4 cm (93^{rd} centile), with current weight 69.3kg (60^{th} centile). He has distinctive facial features with long palpebral fissures, long eyelashes, thick medial eyebrows, up-turned ear lobes, a thin upper lip, smooth philtrum, and micrognathia. He has limited mouth opening. His teeth are crowded with a pointed incisor on the left. His speech is nasal in tone. He has mild ulnar deviation of his 3^{rd} and 4^{th} fingers in the left hand and fetal pads on his fingertips on both sides. His 4^{th} toes are proximally placed. He also has hallux valgus and thick soles. There is slight restriction of elbow extension bilaterally.

Previous genetic investigations included array CGH, *NSD1* gene sequencing and 11p15 methylation analysis, all of which were normal. His brain MRI scan at 3 years was reported as showing delayed myelination for age.

A recent spinal X-ray showed no evidence of scoliosis but several other bony changes. Both clavicles are angulated with an apparent pseudoarthrosis forming with the coracoid processes of the scapula. There are 12 pairs of ribs with angulated slender posterior elements. In the thoracic spine the vertebral body endplates appear irregular. The femoral necks appear short with irregular ossification of the ischium.

Patient 14

Proband ZFX14 was an 8-year-old girl at her last evaluation. She initially presented for a medical genetic evaluation at age 3 years for growth hormone deficiency, pituitary stalk interruption syndrome, developmental dysplasia of the hip, and sacral dysgenesis with deviation of rectum and anus. Prenatal history was remarkable for abnormal ultrasounds identifying congenital heart defect and macrocephaly. Neonatal history was remarkable for hypoglycemia and NG feeds. Postnatally, she followed with cardiology and had a VSD that closed spontaneously. Head and neck imaging revealed hypoplasia of the facial bones and temporomandibular joints, as well as cervical spine stenosis. On physical exam she was noted to have short stature, mild pectus excavatum, hypotonia, macrocephaly, dolichocephaly, prominent forehead, sparse hair, mid-face hypoplasia, epicanthal folds, low and depressed nasal bridge, wide and upturned short nasal tip, long and flattened philtrum, unilateral preauricular tag, spatulate fingers and scooped toenails. She rolled over at age 10-11 months and walked at 26 months. She had hip surgery and spica cast at 12 months. She required PT and OT. Her motor skills improved after starting growth hormone therapy at age 3 years. Her first words were at 10 months, and she was putting 2-3 words together at 13 months. At age 8 years she was doing well in school without additional support or therapies. Her height was in the 25^{th}-50^{th} centile for age. Previous genetic testing revealed 46, XX

chromosomes and a Xp11.23 de novo duplication on SNP array. She also had a Growth Hormone Receptor multigene panel that was non diagnostic. Given non diagnostic testing, trio exome sequencing was recommended which identified a paternally inherited variant of uncertain significance in *ROBO1* (c.928C>T; p. R130X) and a *de novo* variant of uncertain significance in *ZFX* (c.529dupT; p. S177Ffs*12).

Patient 15

Proband ZFX15 is a now 13-year-old male of White British origin. His older sister who is well. Pregnancy was complicated by hyperemesis and pre-eclampsia. A risk of Down syndrome was identified by maternal serum biochemical screening. He was born via emergency C section following secondary to poor fetal movements at 36 weeks. He was transferred to the NICU due to respiratory distress. He required NG feeding. MRI showed small left intraventricular haemorrhage. He had relative macrocephaly at birth and Sotos syndrome was queried but *NSD1* testing was normal. Other early features noted included bilateral inguinal hernia, seizures (later resolved), hypospadias, wide palpebral fissures, prominent forehead. He was discharged when he was 3 weeks of age. He was noted to have developmental delay; he was able to walk at 2 years, but not running until 4 years of age. Examination then revealed a high arched palate in addition to the other previously noted features noted. He was recruited into the DDD study (PMID: 25529582). However, no pathogenic variants were identified. He was then recruited to the 100,000 Genomes study (PMID: 34758253) and the genome sequencing identified the two variants: ZFX (subject of this paper and TFAP2A (Branchiooculofacial syndrome). The TFAP2A was not thought to cause patient's phenotype and it is likely not relevant.

The proband developed thoracolumbar scoliosis around 10 years of age. He is wearing a brace but the scoliosis continues progressing and a spinal fusion planned. He also exhibited vision deterioration (vision was 6/12 iat age 7 yo and now it is less than 6/60) around 10 years of age. Electrophysiology revealed no evidence of significant retinal pathway dysfunction but cortical visual evoked potentials were consistent with significant bilateral optic nerve or posterior visual pathway dysfunction. His retinal photo with optos showed macula changes affecting outer retinal layers and pigment epithelium (but not clear if that is the reason for his visual problem or if it is related to his previous diagnosis of papilledema); these changes remained stable between images in 10/2020 and optos images from 3/2022 and 9/2022. Currently he is registered as severely sight impaired and his vision is further down (6/76 RE and 6/150 left eye). Bilateral optic atrophy was also noted. He also had deterioration in hearing at age 12 when he found to have high frequency sensorineural hearing loss - recently had cochlear implants. He continues exhibiting occasional "blank spells.". He used to sleep very well but recently he started to wake in the night and needs to sleep with parents or becomes very distressed

Patient 16

Proband ZFX16 is a 15-year-old boy, the third child of healthy non-consanguineous parents. During pregnancy, right-sided hydronephrosis was found by ultrasound. The delivery was uncomplicated. He had surgery for pelviourethral obstruction at 4 years of age. His initial development regarding contact, including eye contact was described as normal. Crawling and walking were delayed (not specified), and he still has slight motor problems. His language development was also delayed and he used non-verbal language. He spoke in short sentences by age 3 years. At age 5½ years, he started receiving physiotherapy due to delayed motor skills and at age 14 he started receiving occupational therapy. He is described as having problems with social interaction, concentration and organization. He

was diagnosed with infantile autism at age 14 years. At age 15 years, he goes to normal school, but his cognitive level is below for what is expected for his age.

His physical exam at 14 years of age showed dysmorphic features including thin upper lip, long, almost flat philtrum, downslanting palpebral fissures, long and big nose, big nostrils, large ears and slightly large tongue.

SUPPLEMENTARY FIGURES

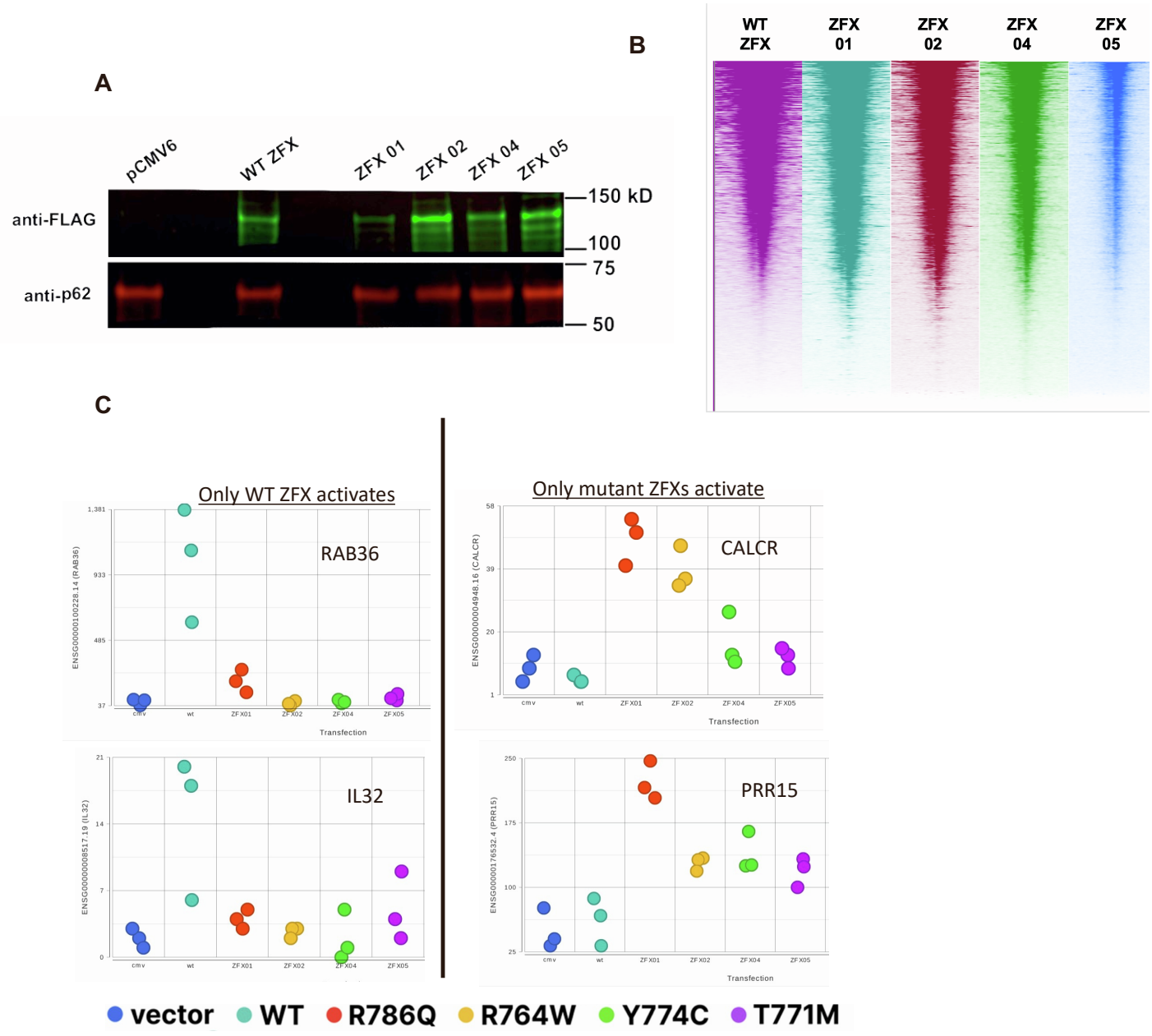
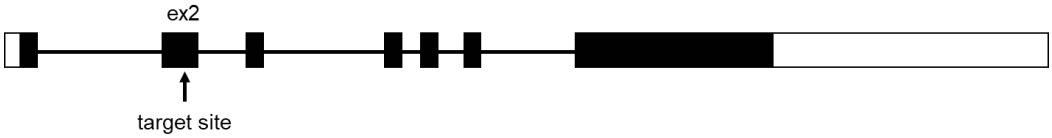


Figure S1: (A) Western blot analysis of WT and mutant ZFX proteins in DKO cells. Equal amounts of nuclear lysate (~ 40 ug) were loaded into each well of a 4-15% gel (Bio-Rad CAT#4561085) After transfer, the membrane was probed with anti-FLAG (green) and anti-p62 (red) antibodies. WT ZFX or mutant ZFX proteins are indicated with the green arrow; the nuclear-localized protein p62 loading control is indicated with the red arrow. (B) Shown are heat maps of ChIP-seq data for wt ZFX and ZFX missense mutants at promoters of all protein-coding genes; all missense ZFX datasets are ranked according to the wt ZFX-bound promoters. (C) Shown are activation levels for two direct targets of wt ZFX only (left) and 2 direct targets of the variant ZFX proteins only (right) in cells transfected with the wt and mutant ZFX proteins; values on the Y-Axis represent normalized read counts for all transcripts of the indicated gene. Key: R786Q = ZFX01; R764W= ZFX02; Y774C= ZFX04; and T771M =ZFX05.

A

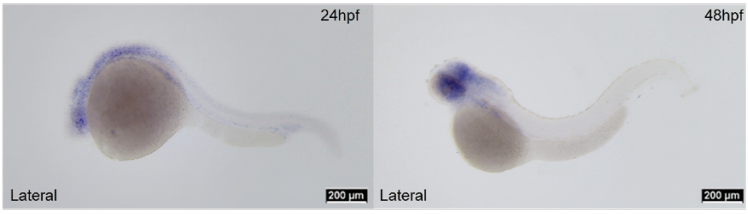


B

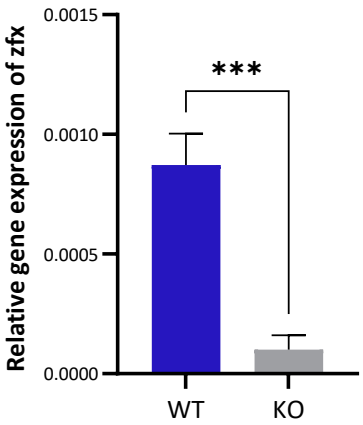
```
WT        5’- GCTCTGGAGGATGAAGGGCTGCAGGTGGATGTGGTGACCGACGCTCAGGTACAAGAGGAT- 3’
-25 bp        GCTCTGGAGGATGAAGG-------------------------GCTCAGGTACAAGAGGAT

WT         5’-GTGACCGACG----------------------CTCAGGTACAAGAGGATCCAGACACCTGC -3’
+23 bp        GTGACCGACGGTACAAGACGGTACTCAGGTACAAGAAGATCCAGACACCTGCGGACACCTGC
```

C



D



E

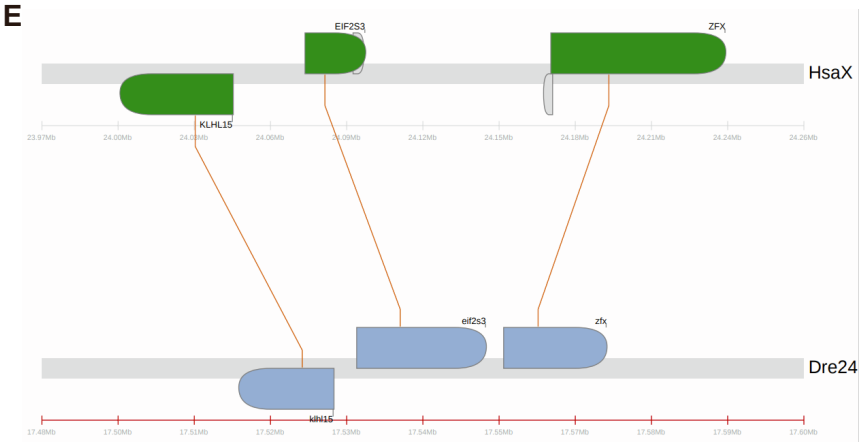


Figure S2: (A) Generation of zfx mutant zebrafish. Schematic showing the genomic structure of zebrafish zfx and CRISPR-Cas9 target site in exon 2. Black boxes, coding regions; white box, untranslated region; black lines, introns. (B) Targeted sequence for WT and 25-bp deletion and 23-bp insertion mutants. (C) Whole mount *in situ*

hybridization of *zfx* in wildtype zebrafish larvae. (D) Whole brain extracts measured by qPCR reveals *zfx* mRNA expression level was significant low in KO compered to WT (p=0.008)(E) Visualization of synteny map of ZFX surrounding region comparing human and zebrafish reference genomes obtained by Synteny Database (http://syntenydb.uoregon.edu/synteny_db/). hpf = hours post fertilization

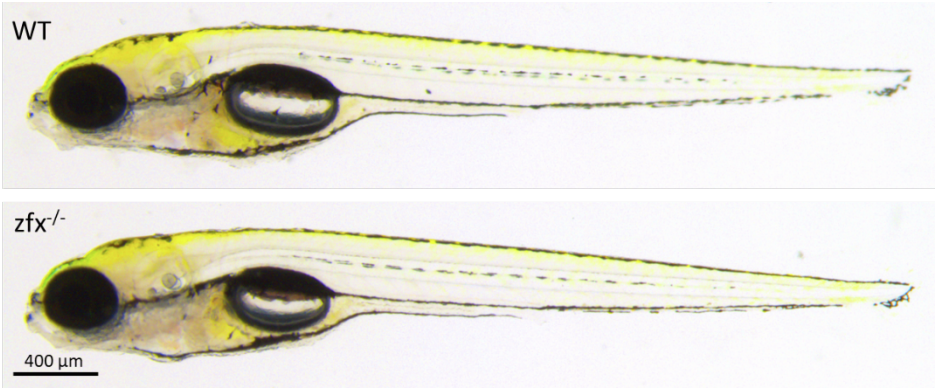


Figure S3: Representative bright field images of WT and *zfx* mutant zebrafish at 5 dpf. *zfx* homozygous mutants show relatively normal development at 5 dpf.

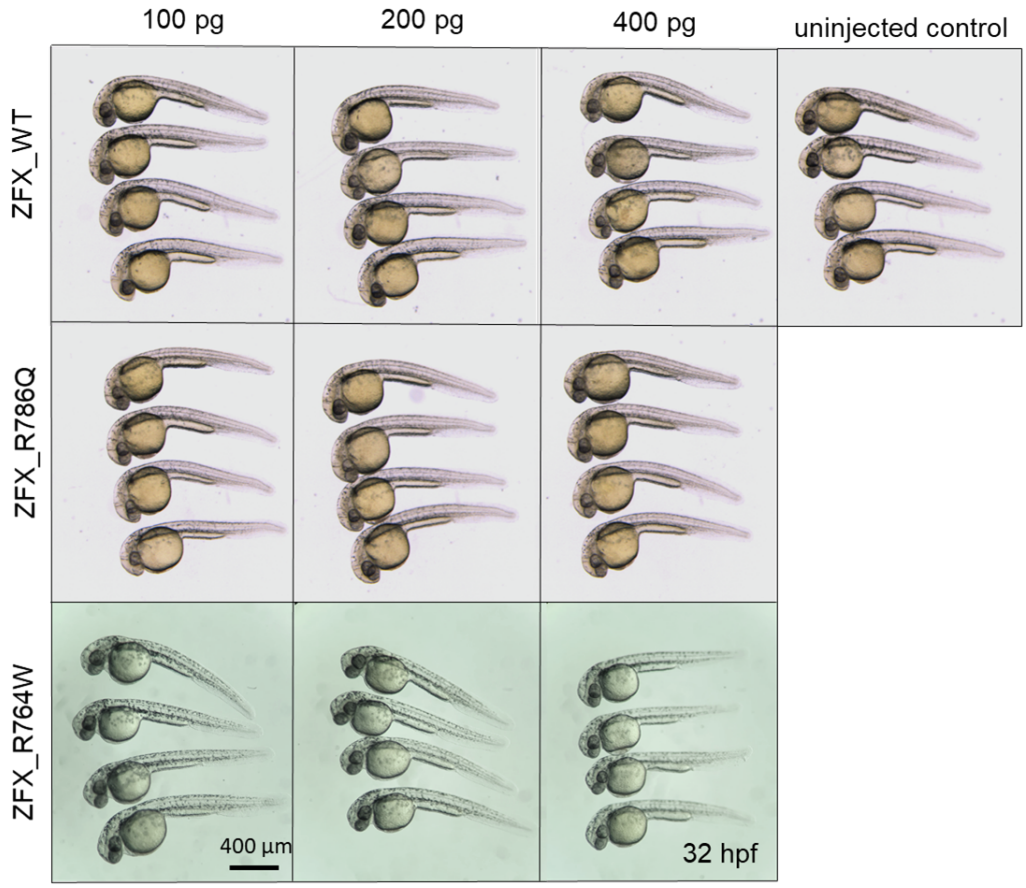


Figure S4. Functional analysis of human ZFX variants using *zfx* mutant zebrafish model. Overexpression experiments by microinjection of mRNAs. Three different concentration (100, 200 and 400 pg) of mRNAs, human wild-type or variant forms of ZFX (R764W and R786Q) were injected.

SUPPLEMENTARY TABLES

Table S1: Clinical, molecular, and demographic summaries for cohort participants.

See Supplementary_Table_S1.xlsx

Table S2: Primers used for creating mutant ZFX expression constructs. The red nucleotide indicates the site of the mutation. ZFX01: p.Arg786Gln; ZFX02: p.Arg764Trp; ZFX04: p.Tyr774Cys; ZFX05: p.Thr771M

Primer name	Primer sequence (5' to 3')
ZFX01_3	GGAAGCCTTTCTTGCAGTACTCACACCGGTGAGG
ZFX01_5	GTACTGCAAGAAAGGCTTCCAAAGACCTTCAGAAAAGAACCAG
ZFX02_3	TAAAGCCTGAGGCATCTGTAGTGCTATACTCACAGTACTCACACTG
ZFX02_5	GCACTACAGATGCCTCAGGCTTTAAACGGCACGTTATTTCCATTCACAC
ZFX05_04_3	GTGAATGGAAATAACGTGCCGTTTAAAGCCTGAGGCATCTGTAGTGC
ZFX04_5	ACGGCACGTTATTTCCATTCACACGAAAGACTGTCCTCACCGGTGTG
ZFX05_5	ACGGCACGTTATTTCCATTCACATGAAAGACTATCCTCACCGGTGTG

Table S3: RNA-seq and ChIP-seq data used in this study.

See Supplementary_Table_S3.xlsx

Table S4: ChIP peaks called for the indicated ZFX genotypes.

See Supplementary_Table_S4.xlsx

Table S5: Gene-level differential expression values computed for the indicated genotype comparisons.

See Supplementary_Table_S5.xlsx

Table S6: Identified direct target genes of the indicated ZFX mutants.

See Supplementary_Table_S6.xlsx

Table S7: Primer sequences used for generation of zfx CRISPR knockout zebrafish model.

Name	Information	Sequence	Purpose
zfx-FP	Forward	5'-GATGATGAAGGGCTTGGCAC-3'	Genotyping
zfx-RP	Reverse	5'-cgctacaatctcactcacAGG-3'	Genotyping
Oligo		5'-taatacgactcactataGGATGTGGTGACCGACGCTCgttttagagctagaa-3'	sgRNA

SUPPLEMENTARY ACKNOWLEDGEMENTS

The authors are grateful to Yibu Chen and members of the Norris Medical Library Bioinformatic Core for assistance. Software and computing resources used are funded by the USC Office of Research and the Norris Medical Library. The authors are also grateful to to Zexun Wu and Suhn Rhie for bioinformatic assistance with the ChIP-seq pipeline, and to Emily Hsu and Nathan Zemke for help with data visualizations. We also appreciate the many helpful suggestions and comments provided by members of the Farnham lab during the course of these studies.

The authors wish to acknowledge A/Prof Nicola Kilpatrick, a pediatric dentist from Melbourne who brought the dental phenotype of the three affected children in Family 6 to the clinical geneticist's attention, triggering further investigation of the family.

This work was funded in part by the National Institutes of Health (1R01GM133450, P30CA014089, R35NS105078). J.L.S. was supported by F30EY033640 and T32GM007200. C.-H.K. was supported by grants from the National Research Foundation of Korea (2018M3A9B8021980, 2020R1A5A8017671, 2021R1A2C1008506). This work was funded in part by NHGRI UM1 HG011758 to the Baylor College of Medicine Genomic Research to Elucidate the Genetics of Rare disease (BCM-GREGoR) program. A.J. and S.B. acknowledge the support of Solve-RD. The Solve-RD project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 779257.

This work was supported by the Japan Agency for Medical Research and Development (AMED) under grant numbers JP23ek0109674, JP23ek0109549, JP23ek0109617, JP23ek0109648 (N. Matsumoto); JSPS KAKENHI under grant numbers JP22H03047 (N.Miyake); and the Takeda Science Foundation (N. Matsumoto).

Part of this study makes use of data generated by the DECIPHER community: A full list of centers who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. Funding for the DECIPHER project was provided by Wellcome Grant# WT223718/Z/21/Z. The DECIPHER consortium, which carried out the original analysis and collection of the data, bears no responsibility for the further analysis or interpretation of the data.

The DDD study presents independent research commissioned by the Health Innovation Challenge Fund# HICF-1009-003, a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute (Grant# WT098051). The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome or the Department of Health. The DDD study has UK Research Ethics Committee approval (Cambridge South REC# 10/H0305/83, Republic of Ireland REC# GEN/284/12). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. See Nature PMID: 25533962 or www.ddduk.org/access.html for full acknowledgement.

The results here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Part of this research was made possible through access to the data and findings generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care).

The 100,000 Genomes Project is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support. The study was supported in part by the NIHR Manchester Biomedical Research Centre (Grant# NIHR203308).