Published in final edited form as:

Am J Med Genet A. 2021 October ; 185(10): 3028–3041. doi:10.1002/ajmg.a.62439.

Exome sequencing of child-parent trios with bladder exstrophy: Findings in 26 children

Georgia Pitsava[1], **Marcia L. Feldkamp**[2], **Nathan Pankratz**[3], **John Lane**[3], **Denise M. Kay**[4], **Kristin M. Conway**[5], **Gary M. Shaw**[6], **Jennita Reefhuis**[7], **Mary M. Jenkins**[7], **Lynn M. Almli**[7], **Andrew F. Olshan**[8], **Faith Pangilinan**[9], **Lawrence C. Brody**[9], **Robert J. Sicko**[4], **Charlotte A. Hobbs**[10], **Mike Bamshad**[11], **Daniel McGoldrick**[12], **Deborah A. Nickerson**[12], **Richard H. Finnell**[13], **James Mullikin**[14], **Paul A. Romitti**[5,#], **James L. Mills**[1,#], **University of Washington Center for Mendelian Genomics, NISC Comparative Sequencing Program and the National Birth Defects Prevention Study**

[1]Division of Intramural Population Health Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland

[2]Division of Medical Genetics, Department of Pediatrics, 295 Chipeta Way, Suite 2S010, University of Utah School of Medicine, Salt Lake City, Utah

[3]Department of Laboratory Medicine and Pathology, University of Minnesota Medical School, Minneapolis, Minnesota

[4]Division of Genetics, Wadsworth Center, New York State Department of Health, Albany, New York

[5]Department of Epidemiology, College of Public Health, The University of Iowa, Iowa City, Iowa

[6]Department of Pediatrics, Stanford University School of Medicine, Stanford, California

[7]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, Georgia

[8]Department of Epidemiology, Gillings School of Global Public Health, Chapel Hill, North Carolina

[9]Gene and Environment Interaction Section, National Human Genome Research Institute, Bethesda, Maryland

Corresponding Author James L. Mills, MD, MS, 6710B Building, Room 3117, NICHD, NIH, Bethesda, MD, 20892, Phone: 301-496-5394, millsj@exchange.nih.gov.

[#] Co-senior authors

AUTHOR CONTRIBUTIONS

James L. Mills and Paul A. Romitti supervised the study. Lynn M. Almli., Kristin M. Conway, Marcia L. Feldkamp, Richard H. Finnell, Charlotte A. Hobbs, Mary M. Jenkins, Andrew F. Olshan, Jennita Reefhuis, Paul A. Romitti, Garry M. Shaw, NBDPS, contributed to the NBDPS study design and data and specimen acquisition. Mike Bamshad, Lawrence C. Brody, Denise M. Kay, Daniel McGoldrick, James Mullikin, Deborah A. Nickerson, Faith Pangilinan, Robert J. Sicko, NISC Comparative Sequencing Program, and UWCMG conducted experimental studies. John Lane and Nathan Pankratz analyzed the data, and Mike Bamshad, Lawrence C. Brody, Marcia L. Feldkamp, Denise M. Kay, John Lane, Daniel McGoldrick, James L. Mills, Deborah A. Nickerson, Nathan Pankratz, Georgia Pitsava, Paul A. Romitti, Robert J. Sicko and UWCMG interpreted the data. Georgia Pitsava created the initial draft of the manuscript and reviewed the literature and Marcia L. Feldkamp, Denise M. Kay, John Lane, James L. Mills, Nathan Pankratz, Faith Pangilinan, and Paul A. Romitti critically revised the manuscript. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

Dr. Finnell was formerly in a leadership position with TeratOmic Consulting LLC, a now dissolved consulting firm.

DATA AVAILABILITY

The data are from a CDC-funded multi-site study and not publicly available at this time due to privacy or ethical restrictions.

[10]Rady Children's Institute for Genomic Medicine, San Diego, California

[11]Department of Pediatrics, University of Washington, Seattle, Washington

[12]Department of Genome Sciences, University of Washington, Seattle, Washington

[13]Center for Precision Environmental Health, Baylor College of Medicine, Houston, Texas

[14]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland

Abstract

Bladder exstrophy (BE) is a rare, lower ventral midline defect with the bladder and part of the urethra exposed. The etiology of BE is unknown but thought to be influenced by genetic variation with more recent studies suggesting a role for rare variants. As such, we conducted paired-end exome sequencing in 26 child/mother/father trios. Three children had rare (allele frequency ≤0.0001 in several public databases) inherited variants in *TSPAN4*, one with a loss-of-function variant and two with missense variants. Two children had loss-of-function variants in *TUBE1*. Four children had rare missense or nonsense variants (one per child) in *WNT3*, *CRKL*, *MYH9*, or *LZTR1,* genes previously associated with BE. We detected 17 *de novo* missense variants in 13 children and three *de novo* loss-of-function variants *(AKR1C2*, *PRRX1*, *PPM1D)* in three children (one per child). We also detected rare compound heterozygous loss-of-function variants in *PLCH2* and *CLEC4M* and rare inherited missense or loss-of-function variants in additional genes applying autosomal recessive (three genes) and X-linked recessive inheritance models (13 genes). Variants in two genes identified may implicate disruption in cell migration (*TUBE1*) and adhesion (*TSPAN4*) processes, mechanisms proposed for BE, and provide additional evidence for rare variants in the development of this defect.

Keywords

TUBE1 ; *TSPAN4* ; birth defects; bladder exstrophy; genetics; ventral body wall closure

1 INTRODUCTION

Bladder exstrophy (BE), a severe lower midline ventral defect, includes a spectrum of urological abnormalities in which all or part of the abdominal wall fails to close, resulting in the bladder and part of the urethra being exposed. Additional features of BE include pelvic bone malformations, widened pubic bones, cleft clitoris and shortened vagina in female infants, and epispadias in male infants (Ebert, Reutter, Ludwig, & Rosch, 2009; Reinfeldt Engberg, Mantel, Fossum, & Nordenskjold, 2016). The estimated prevalence of BE varies from 0.5–4.6 per 100,000 live births (Siffel et al., 2011). In most (Higgins, 1962; ICBDMS, 1987; Lattimer & Smith, 1966; Nelson, Dunn, & Wei, 2005; Reinfeldt Engberg et al., 2016; Reutter et al., 2011; Shapiro, Lepor, & Jeffs, 1984; Yang et al., 1994), but not all studies (Nelson et al., 2005; Yang et al., 1994), male predominance has been reported, ranging from 1.5–6.0:1.

The etiology of BE remains largely unknown; however, recent findings for familial cases (Froster, Heinritz, Bennek, Horn, & Faber, 2004; Kajbafzadeh, Tajik, Payabvash, Farzan,

& Solhpour, 2006; Reutter, Hoischen, et al., 2007; Reutter, Shapiro, & Gruen, 2003), from analysis (Messelink, Aronson, Knuist, Heij, & Vos, 1994; Reutter et al., 2003; Shapiro et al., 1984) of twin studies (Reutter, Qi, et al., 2007), and of reports of BE in individuals with trisomy 21 continue to support an underlying genetic susceptibility (Ebert et al., 2009; Reutter, Betz, Ludwig, & Boemers, 2006; Reutter et al., 2009). More recently, this susceptibility has been supported by findings from array-based approaches, including common variants identified through genome-wide association studies (Draaken et al., 2015; Reutter et al., 2014; Zhang, Knapp, Kause, Reutter, & Ludwig, 2017) and rare variants identified through copy number analyses (Beaman et al., 2019; Draaken et al., 2015; Draaken et al., 2013; Draaken et al., 2010; Jorgez et al., 2014; Lundin et al., 2010; Soderhall et al., 2014; von Lowtzow et al., 2016). Building on investigations for rare variants for BE, we conducted exome sequencing on a modestly sized, but well-characterized, population-based sample of children with BE and their parents. Our approach provides additional data on rare variants that may contribute to the etiology of BE and provides sequencing data for use in future studies.

2 MATERIALS AND METHODS

2.1 National Birth Defects Prevention Study

The National Birth Defects Prevention Study (NBDPS) was a population-based, case-control study designed to investigate risk factors for over 30 major structural birth defects. The study methods and population have been detailed previously (Reefhuis et al., 2015). In brief, 10 sites (Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, Utah) participated in the NBDPS. The birth defect surveillance program at each site ascertained children with NBDPS-eligible defects among pregnancies with estimated dates of delivery during October 1997 through December 2011.

2.1.1 Child classification—All children with BE, with or without epispadias, diagnosed by physical examination were ascertained. Children with epispadias without BE and those where BE was a component of the OEIS complex (omphalocele, exstrophy of the cloaca, imperforate anus and spine abnormalities) were excluded. A board-certified clinical geneticist at each NBDPS site reviewed clinical information abstracted from medical records to verify eligibility (Rasmussen et al., 2003). Consistency across centers was established by a clinical geneticist who performed the final classification of each child diagnosed with BE. Children were classified as isolated (no additional major birth defect) or multiple (one or more additional major birth defects in an unrelated organ system). Children with known chromosomal or monogenic etiologies were excluded.

2.1.2 Interview data and specimen collection—Case mothers completed a computer-assisted telephone interview six weeks to two years after their estimated date of delivery. The interview asked about pre-pregnancy and pregnancy exposures and behaviors and sociodemographic information. Following the interview, mothers were asked to collect buccal cell specimens from themselves, their child (if living), and the child's father (if available) (Reefhuis et al., 2015). Mothers who participated in the NBDPS with a previous child, those who could not complete the interview in English or Spanish, or those who were

incarcerated or otherwise did not have custody of their child at the time of recruitment were excluded.

Mothers of 75 of 107 (70.1%) eligible children with BE completed the telephone interview; one pregnancy ended in fetal loss, leaving mothers of 74 children. Of these, 39 mothers (52.7%), 39 children (52.7%), and 37 fathers (50.0%) provided buccal cell specimens; these specimens included 37 case child-parent trios, one child-mother dyad, one mother only, and one child only. Only the 37 child-parent trios were considered for sequencing; one of these trios was from a family with a child diagnosed with OEIS and, therefore, was excluded from analysis, leaving 36 trios.

As described elsewhere (Jenkins et al., 2019), two different types of cytobrushes were used to collect specimens, “wet brushes” (cytobrushes packaged in closed plastic tubes preventing air drying [Cyto-Pak Cytosoft Brushes CP-5B, Medical Packaging Corporation, Camarillo, CA]) and “dry brushes” (cytobrushes packaged in open paper-backed peel pouches [Cytology Brush Pack CYB-1, Medical Packaging Corporation]). A pilot study determined the DNA quantity and quality from wet brushes to be sub-optimal for exome sequencing (Jenkins et al., 2019). Because seven of the 36 case child-parent trios were wet brushes, they were not sequenced. Of the remaining 29 dry brush case child-parent trios, three were excluded due to inadequate DNA, leaving 26 trios (78 specimens) for exome sequencing. Signed informed consent was obtained for all participants providing buccal cell specimens. The study protocol for the NBDPS was approved by the institutional review board at each NBDPS site.

2.2 Specimen processing and sequencing

Buccal specimens that had adequate DNA amounts (≥200 ng, assessed by quantitative real-time polymerase chain rection (PCR) targeting the RNaseP gene) were sent to the National Institutes of Health Intramural Sequencing Center (NISC) at the National Human Genome Research Institute. Each specimen was quantified using a Qubit assay to measure double-stranded DNA, and DNA molecular weight was assessed by running a small portion of each specimen on an agarose gel.

2.2.1 Exome enrichment and coverage—The exome sequencing capture kit used at NISC was a standard, commercially available kit, the NimbleGen SeqCap EZ Exome+UTR Library (Version 3.0) and covered 96 Mb (Roche NimbleGen, 2013). As described (Jenkins et al., 2019), the DNA was sheared mechanically, and targeted fragments were captured by probe hybridization and amplified before sequencing.

2.2.2 Exome sequencing—NISC generated read lengths of 126 bases on an Illumina HiSeq 2500 instrument. Paired-end reads generated approximately 250 bp of sequence from each fragment in the library. A total of 38 million paired-end 126 bp reads were targeted and as many as 48 libraries were pooled and sequenced across as many lanes as needed to achieve the targeted number of reads (938 million read pairs or 76 million reads pre-library); thus, 5–6 libraries were run per lane. Image analysis and base calling were performed using Illumina Genome Analyzer Pipeline software (versions 1.18.64.0) with default parameters.

2.3 Alignment and genotype calling

FASTQ files were processed using a Genome Analysis Toolkit (GATK) v3.7 based pipeline, including BWA-MEM v0.7.17 for alternate contig aware alignment to the hg38 reference genome (GRCh38_full_analysis_set_plus_decoy_hla.fa), Picard Tools v2.6.0 to mark duplicates (Picard, retrieved from http://broadinstitute.github.io/picard/), and GATK for indel realignment, base quality recalibration, genotyping (HaplotypeCaller), variant quality score recalibration, and to split multiallelic sites (Li & Durbin, 2010; McKenna et al., 2010). Following processing, the median number of on-target reads per specimen was 68 million (69% of total reads overlapping a target, 79% of reads within 126 base pairs of a target), resulting in a median target coverage of 61X (95% of targets covered at 20X).

2.4 Annotation

All variants were annotated for functional impact (putative amino acid changes and predicted deleteriousness, e.g. the Combined Annotation Dependent Depletion [CADD] score (Rentzsch, Witten, Cooper, Shendure, & Kircher, 2019) using SnpEff v4.3r (Cingolani et al., 2012) and ANNOVAR v2018Apr16 (Wang, Li, & Hakonarson, 2010). ANNOVAR was also used to annotate the presence and allele frequency (AF; including ancestry-specific frequencies) of each variant in several public databases, dbSNP (Sherry et al., 2001) version 151, 1000 Genomes (Genomes Project et al., 2015), NHLBI GO Exome Sequencing Project 6500 exomes (NHLBI-ESP), Exome Aggregation Consortium and Genome Aggregation Database (Konrad J. Karczewski et al., 2019).

2.5 Specimen quality control

Reported sex and relatedness for each specimen were verified using Peddy v0.4.2 (Pedersen & Quinlan, 2017). Specimen quality was assessed with FastQC v0.11.2 (FastQC, Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), VerifyBamID's freemix score, and an internal QC pipeline (Jun et al., 2012). All specimens had sufficient target coverage (>75% of targets covered at 20X). Two specimens (one mother and one child from separate trios) were flagged as potentially contaminated (freemix >0.1). The two trios with potential contamination were not analyzed for *de novo* variation, and the child was flagged as lower quality in screens for inherited variation.

2.6 Identifying CNVs, duplications or deletions in the autosome

ExomeDepth (Plagnol et al., 2012) was used to call autosomal copy number variants (CNVs, duplications or deletions in the genome) using the targets in the NimbleGen SeqCap EZ Exome+UTR Library (Version 3.0) to define exon boundaries. ExomeDepth calls were imported into Genvisis (http://www.genvisis.org), which was used to merge calls within each sample (if the gap between two calls of the same copy number was less than 20% of the size of the larger call), and filter for large events (>100kb and spanning more than 10 targets) with an ExomeDepth Bayes factor quality score >10. CNVs were then annotated with AnnotSV v3.0.7 (Geoffroy et al., 2018) to determine the allele frequency in external databases and the predicted pathogenicity (AnnotSV uses a pathogenicity score based on recommendations from the American College of Medical Genetics and Clinical Genome Resource (Riggs et al., 2020).

Annotations from AnnotSV were used to remove CNVs completely contained in a benign region (e.g., the same copy number is present in healthy populations such as gnomAD v2.1 at >1% allele frequency or homozygous deletions are present in multiple healthy individuals) and retain only CNVs that were predicted to be “Likely pathogenic” or “Pathogenic”. Additionally, CNVs present in a population of 68 internal controls were removed. The remaining rare and putatively functional calls were manually reviewed using the Genvisis Trailer visualization module and plots produced by ExomeDepth and excluded any that appeared to be artifact (e.g., due to poor mapping quality, highly variable coverage in controls).

2.7 Filtering and prioritization

Our genotyping pipeline identified 255,289 unique variants. Of these 215,871 (84.6%) were deemed high quality (not in a GATK tranche, depth ≥10, genotype quality [GQ] ≥50). Of the 215,871 high quality variants, 35,852 affected an amino acid (missense or nonsense variants) and were deemed putatively functional. We focused on genes in which loss-of-function variants were found in more than one individual and loss-of-function and missense variants that were found in more than two individuals. Novel and known gene candidates were prioritized using the gene damaging index (GDI), focusing on genes with mutational burdens in the lowest 50th percentile (Itan et al., 2015). Of the 35,852 variants, 469 met our primary criteria (maximum AF across all control populations ≤0.0001 and in a gene with a GDI <50). We further prioritized variants that were high quality using secondary criteria, which were the same as our primary criteria but less restrictive in terms of rarity in the control population (AF ≤0.001) and GDI (<75th percentile), leaving 1,778 variants. All putatively functional (missense or loss-of-function) rare variants (AF ≤0.001) in genes previously associated with BE were manually reviewed using the Integrative Genomics Viewer (IGV) v2.4.13 (Robinson et al., 2011) (Supplemental Table 1), regardless of the quality.

Filtering criteria for compound heterozygous, homozygous, and variants in X-linked genes are summarized in Supplemental Table 1. The individual who was the product of a consanguineous union was confirmed by identifying large runs of homozygosity using bcftools (using the “roh” command) (Narasimhan et al., 2016) and was considered in a separate analysis for the autosomal recessive model (Supplemental Table 4).

For calling *de novo* variants, we used the GATK genotype refinement pipeline (retrieved from https://gatkforums.broadinstitute.org/gatk/discussion/4723/genotype-refinement-workflow) to incorporate pedigree and population priors into the GQ scores and to flag potential *de novo* variants. *De novo* variants flagged by GATK were further filtered to require GQ ≥40 and total depth ≥ 8 in all members of the trio, alternate allele depth ≥5 in the child, and AF ≤0.001 in external databases (Supplemental Table 1). *De novo* variants passing all filters were manually reviewed with IGV.

2.8 Sanger sequencing

Sanger sequencing was used to validate a subset of variants detected by exome sequencing in child and parent specimens. Variants were selected for validation if they were in genes

that either met our primary filters or were previously reported in the literature. In addition, variants were sequenced if they were *de novo* but were flagged by GATK as having an increased likelihood of being a false positive (i.e., in a tranche). DNA was amplified in 20 μl reactions using Platinum II Taq Hot-Start Master Mix (Thermo Fisher Scientific, Waltham, MA), and 0.2 μM each primer (M13-tailed; Thermo Fisher Scientific). PCR products were cleaned-up using ExoSAP-IT (USB Corporation; Cleveland, Ohio) and sequenced using BigDye Terminator v.3.1 Cycle Sequencing chemistry kits (Thermo Fisher Scientific) on an ABI 3130xl or 3730 DNA Analyzer (Thermo Fisher Scientific). Sequence chromatograms were analyzed using FinchTV v.1.4.0 (Geospiza; Seattle, WA) and the BLAST-Like Alignment Tool (BLAT) (Kent, 2002).

3 RESULTS

Among the 26 live born children with BE, 18 (69.2%) were males, 23 (88.4%) were full term, and all were singleton births. Among case mothers, age at delivery was nearly equally divided between 20–29 and 30–39 years (50.0% and 46.2%, respectively), and most were non-Hispanic white (73.1%) (Table 1a). Several children had additional malformations (Table 1b).

Using our primary criteria for variant prioritization, two genes emerged. Five children inherited heterozygous variants, two in the *TUBE1* gene and three in the *TSPAN4* gene (Table 2a). The variants identified in *TUBE1* (c.1093A>T, p.R365X; c.673dupG, p.V225fs) were loss-of-function variants, with each predicted to lead to absence of the protein product due to nonsense mediated decay. Of the three children with putatively functional variants in *TSPAN4*, one had a loss-of-function variant (c.411_432+3del, p.A137fs), whereas the other two had missense variants (c.491G>A, p.R164Q; c.698C>G, p.A233G) (Table 2a). One of these missense variants (c.491G>A, p.R164Q) lies within the tetraspanin extracellular EC2 domain, as annotated in the InterPro database (Mitchell et al., 2019). Each of the five variants detected in *TUBE1* and *TSPAN4* have mappability scores of 1 (Derrien et al., 2012). All children with *TUBE1* and *TSPAN4* variants were of non-Hispanic white race/ethnicity. The GDI percentile of *TUBE1* is 49.1% suggesting that nearly half of all human genes are less tolerant to damaging variation than *TUBE1*. Other bioinformatic algorithms reveal conflicting evidence regarding the relevance of loss-of-function variants in *TUBE1*. Two algorithms suggest that heterozygous loss-of-function variants in this gene could lead to haploinsufficiency. The gnomAD loss-of-function observed/expected (o/e) constraint score was below 1 (0.31 [90% CI: 0.17 – 0.58]) (K. J. Karczewski et al., 2020; Lek et al., 2016) indicating that fewer than expected loss-of-function *TUBE1* variants were found in gnomAD (Konrad J. Karczewski et al., 2019; Lek et al., 2016). In addition, the percentile of the predicted haploinsufficiency score (Huang, Lee, Marcotte, & Hurles, 2010) of *TUBE1* was low (14.86%, range: 0 – 100; with 0 being more likely to cause haploinsufficiency) indicating that heterozygous loss-of-function variants are more likely to lead to a severe phenotype. However, gnomAD's probability of loss of function intolerance (pLI) was only 0.03 (ranges from 0 to 1; 0 being more likely to be tolerant to loss-of-function variation), which contradicts the two scores above. *TSPAN4* has a GDI percentile of 40.3%, a gnomAD loss-of-function o/e constraint score of 0.87 (0.56 – 1.41), a gnomAD missense o/e constraint score of 1.09 (0.96 – 1.24), haploinsufficiency score of 71.26%, and a gnomad

pLI of 0. Therefore, loss-of-function and potentially damaging missense variants in *TSPAN4* are predicted to be well-tolerated.

Table 2b lists inherited variants near or in genes previously reported to be associated with BE in the literature, *WNT3*, *CRKL*, *LZTR1*, *MYH9*, *ISL1*, *TP63*, and *SLC20A1*. We identified four variants – one per child and one each in four of these genes (*WNT3*, *CRKL*, *LZTR1, MYH9*). Missense variants were identified in three children: c.286G>C, p.D96H in *WNT3*, c.76A>G, p.T26A in *CRKL* and c.4958A>G, p.E1653G in *MYH9* (Table 2b); one child had a nonsense variant (c.628C>T, p.R210X) in the *LZTR1* gene.

Our investigation of variants using autosomal recessive or X-linked recessive models (Tables 3 and 4 respectively) identified 3 missense and 13 loss-of-function variants. We also identified 17 *de novo* missense and 3 *de novo* loss-of-function variants (Table 5). Two of the *de novo* variants (in *AKR1C2* and *USP18*) were flagged as being potentially false positive by GATK due to mapping differences between reads with and without alternate alleles; however, both were confirmed by Sanger sequencing. All 20 *de novo* variants passed manual review in IGV, including these two variants. Screenshots of all variants reviewed in IGV can be found in the Supplemental Table 1. The loss-of-function variants were found one each in *AKR1C2*, *PRRX1*, and *PPM1D*; this last variant occurred in a child classified clinically as ‘multiple’ (congenital anomalies of the skeletal system in addition to bladder exstrophy).

We detected inherited compound heterozygous loss-of-function variants in *PLCH2* and *CLEC4M* that met our primary screening criteria (Supplemental Table 2). Other compound heterozygous variants of interest identified are listed in Supplemental Table 2. Variants in additional genes that met our secondary inclusion criteria are listed in Supplemental Table 3. Of note, we identified one missense variant in *CD151* in a consanguineous family (Supplemental Table 4). *CD151* is a member of the tetraspanin family, which creates complexes with integrins and other molecules and is involved in cell adhesion. Many of the children in our study had more than one variant (see case child numbers in tables), which may suggest an oligogenic model of inheritance for BE.

We identified a total of seven rare autosomal CNVs with length greater than 100kb in six children (three deletions and four duplications). All seven were predicted to be pathogenic by AnnotSV. None of the regions with CNVs were identified in more than one child. Six of the CNVs were inherited and one was a *de novo* deletion in child 10 (child “10” had an additional paternally inherited duplication). A paternally inherited duplication in child “26” was found within 22q11.21 (chr22:17400001–21700000 on the hg38 reference genome) that has previously been associated with bladder exstrophy (among the genes overlapped by the duplication in child “26” are *CRKL*, *THAP7*, and *LZTR1*) (Figure 1). Additional information for all seven CNVs (e.g. the breakpoints from ExomeDepth and annotations from AnnotSV), the images used to manually review call quality and screenshots from the UCSC Genome Browser (Kent et al., 2002) (http://genome.ucsc.edu) can be found in supplementary file 2.

A subset of 13 variants was selected for confirmation via Sanger sequencing. This subset included 1) the five variants in *TUBE1* and *TSPAN4* (Table 2a) selected using our primary criteria for candidate variants; 2) the five variants in *WNT3*, *MYH9*, *LZTR1*, *THAP7*, and *CRKL* (Table 2b) selected using our criteria for known genes; 3) an additional variant in *MYH9* (rs768181629) that we suspected to be an artifact and was not described in the manuscript but was still nominated for validation because it is in a gene with prior evidence of association; and 4) two *de novo* variants in *AKR1C2* and *USP18* that were flagged by GATK. All five variants in *TUBE1* and *TSPAN4*, plus variants in *MYH9* (rs780486984), *WNT3* and *LZTR1* were validated in probands and the parent from which the variant was inherited. The *CRKL* variant was validated in the proband, but parent DNA was not available for validation studies. *THAP7* and *MYH9* rs768181629 variants were the only two that did not validate; both were lower quality *de novo* variants with low alternate allele counts. The *de novo* variants in *AKR1C2* and *USP18* were confirmed in probands and were absent from both parents.

4 DISCUSSION

There have been numerous postulated mechanisms for the pathogenesis of BE; cell migration and cell adhesion defects are two of the theories that have been proposed (Martinez-Frias, Bermejo, Rodriguez-Pinilla, & Frias, 2001; Sadler, 2010; Sadler & Feldkamp, 2008). We identified variants in *TUBE1*, which is implicated in cell migration (Chang, Giddings, Winey, & Stearns, 2003; Garcin & Straube, 2019), and *TSPAN4*, which is involved in both cell migration (Berditchevski, 2001) and cell adhesion (Jiang et al., 2019; Levy, Todd, & Maecker, 1998). In two children, we identified heterozygous predicted loss-of-function variants in *TUBE1*, including a stop-gain mutation and an insertion leading to a frameshift mutation. The protein epsilon tubulin 1, encoded by *TUBE1*, is localized to the centrosome (Chang & Stearns, 2000). Directed cell migration requires polarization of cells via the microtubule-organizing center (MTOC) (centrosome) and its disruption results in loss of cell polarity (Cheng et al., 2019). *TUBE1* is a gene crucial for the organization of the microtubules during centriole duplication. Members of this superfamily are essential in maintaining the microtubule cytoskeleton (Chang et al., 2003). One theory is that BE is caused by failure of one or both lateral body wall folds to move ventrally to meet in the midline resulting in internal organs, such as the bladder, protruding through the defective region (Duhamel, 1963; Feldkamp, Carey, & Sadler, 2007; Sadler & Feldkamp, 2008). It is known that dorsal movement of the neural folds to form the neural tube is driven in part by alterations in the shaping and position of the cells. (Colas & Schoenwolf, 2001; Lee & Nagele, 1985; Sadler, 2005); we hypothesize that the same may potentially be true for the ventral abdominal wall (Lee & Nagele, 1985; Sadler & Feldkamp, 2008). The hypothesis that *TUBE1* variants could be possibly be associated with defective ventral body wall closure is also supported by the finding that the cadherin receptor gene *CELSR3,* a planar cell polarity core gene which, like TUBE1 is related to centriole positioning (Carvajal-Gonzalez, Mulero-Navarro, & Mlodzik, 2016; Su et al., 2016), has been suggested as a potential BE-associated gene (Reutter et al., 2016).

Disrupted fusion of the lateral body folds, once they meet in the midline, is another theory to explain failure of ventral body wall closure. Functional variants in the *TSPAN4*

gene might contribute to BE because of disrupted fusion. Two out of the three *TSPAN4* variants were missense variants, and the other a loss-of-function variant. *TSPAN4* belongs to the tetraspanin superfamily. Tetraspanins are known to play a role in cell adhesion and motility (Jiang et al., 2019; Levy et al., 1998) and form complexes with adhesion receptors from the integrin family, another mediator of cell adhesion (Berditchevski & Odintsova, 1999; Maecker, Todd, & Levy, 1997). In attached cells, tetraspanins remain concentrated within adhesion structures that morphologically are similar to Rac-dependent peripheral focal complexes, which stabilize adhesive interactions of the cell front (Berditchevski & Odintsova, 1999; Nakamura, Iwamoto, & Mekada, 1995; Nobes & Hall, 1995). Tetraspanins may also be involved in either post-adhesion signaling or affect integrin-mediated cell attachment directly (Shaw et al., 1995; Yanez-Mo et al., 1998). In other tissues that fuse during embryogenesis, such as closure of the lip and palate and the neural tube, adhesion is initiated by cell surface molecules, such as glycoproteins, and more permanent adhesion between the two folds takes place through new cell-to-cell contact (Burk, Sadler, & Langman, 1979; Colas & Schoenwolf, 2001; Greene & Pratt, 1976; Sadler, 2005). Although the exact process by which the ventral wall fuses is unknown, our data suggest that the function of *TSPAN4* in cell adhesion could potentially be involved in fusion of the lateral folds.

Tetraspanins are also involved in cell migration. The human TSPAN4-like membrane protein (*TM4SF1*), is crucial for nanopodia formation. Nanopodia have been shown to play an important role in human endothelial cell polarization and migration, and their formation is dependent on *TM4SF1* expression, as demonstrated by their dose-dependent increase or reduction in various cell lines (e.g. fibroblasts, melanocytes) (Zukauskas et al., 2011). In the *Xenopus* model, *Tspan4* has been shown to be expressed during gastrulation in the dorsal and ventral mesoderm (Kashef, Diana, Oelgeschlager, & Nazarenko, 2013); in humans, lateral body folds are formed from the parietal layer of lateral plate mesoderm and overlying ectoderm (Sadler & Feldkamp, 2008). We hypothesize that *TSPAN4* may play a role in both cell adhesion and cell migration.

The gnomAD loss-of-function o/e constraint score suggests that healthy individuals do have loss-of-function variants in *TUBE1* but that they are observed less often than expected in healthy individuals. In addition, a haploinsufficiency score (Huang et al., 2010) for this gene was low, suggesting that heterozygous loss-of-function variants in *TUBE1* are more likely to lead to a severe phenotype. Variants in *TSPAN4* are predicted to be better tolerated. It is possible that the inherited variants in these two genes in children in our cohort may have occurred by chance. Therefore, functional follow-up would be needed to establish a causal relationship with bladder exstrophy for either gene. In addition, no pathogenic deletions or duplications of either gene for BE were noted in ClinVar, ClinGen, OMIM, dbVar, or DECIPHER.

We identified and confirmed variants in *WNT3, CRKL, LZTR1,* and *MYH9* – four of the eight genes previously reported to be associated with BE – in four patients. Each gene was reported in association with BE as part of a deletion or insertion in the gene region, or as a point mutation. The association of BE with a variant near *WNT3* was suggested by previous copy number variant (CNV) analysis and supported by mouse expression data and

genome-wide association studies (Reutter et al., 2014). *WNT3* variants have been associated with other defects (Niemann et al., 2004), whereas *WNT* is related to persistent cloaca and *p63* that is a *WNT* suppressor has been associated with human bladder exstrophy-epispadias complex (Ching et al., 2010; Katoh et al., 2016; Mahfuz, Darling, Wilkins, White, & Cheng, 2013; Qi et al., 2011).

Duplications of the 22q11.21 region have been associated with classic BE (Beaman et al., 2019; Draaken et al., 2014; Draaken et al., 2010; Lundin et al., 2010). Genes included in the duplication region – *Crkl,* and *Lztr1* – have been suggested as candidate genes for BE in mouse studies (Draaken et al., 2014). We identified rare sequence variants in each of these two genes, including a rare nonsense variant in *LZTR1* also seen in the autosomal recessive form of Noonan syndrome (Johnston et al., 2018). If a duplication of a large region like this increases the risk of disease, it likely has a gain-of-function effect. The missense variant we have identified in *CRKL* could also have a gain-of-function effect, but the loss-of-function allele (stop gain mutation in the middle of the gene) in *LZTR1* would be expected to have the opposite effect.

We confirmed a *MYH9* missense variant in one child and father. *MYH9* encodes a non-muscle myosin heavy chain IIA essential for cell adhesion in cell migration during embryonic development. Interestingly, another *de novo* missense variant in this gene was previously reported as the causative gene in an individual with BE and congenital macrothrombocytopenia. (Utsch et al., 2006).

Lastly, we identified a *de novo* stop-gain mutation in exon 6 of *PPM1D*. Stop-gain variants in the last exons of this gene (exons 5 or 6) have been reported to cause an intellectual disability syndrome along with facial dysmorphology, short stature, small hands and/or feet and hyperlordosis (Jansen et al., 2017). The child sequenced in our study reportedly had craniofacial dysmorphology and skeletal anomalies that were not reported by Jansen et al. To the best of our knowledge, this is the first case with a *PPM1D* mutation reported to have BE.

A primary strength of our analysis is the use of a trio-based design to discern between inherited versus *de novo* variants. Other strengths include that children were systematically ascertained for the NBDPS from population-based surveillance programs, minimizing missed cases within each site's catchment area, and medical record information for each child was reviewed by clinical geneticists, producing a well-characterized study sample. In addition, we expanded investigation of rare variants by applying exome sequencing. Limitations of our study are the modest sample size with only a subset of eligible children providing DNA and limited racial/ethnic diversity among families that provided DNA. Nonetheless, these families were from a population-based sample and well-characterized. Another limitation was the use of exome sequencing rather than whole genome sequencing, which prevented us from studying non-coding regions; however, these exome data identified genes that could potentially be candidates for disease association and contribute to understanding the genetic etiology of BE. Lastly, we did not screen additional case populations for the variants detected because they were all rare, and therefore, even if they are causal, the likelihood of detecting them in additional populations was low. Functional

studies were not conducted because none of the variants were recurrent. Despite these limitations and the potential for chance findings, our study suggests new hypotheses and our data can contribute to more large-scale efforts to characterize genetic susceptibility of BE.

Additional considerations regarding our findings included that the heterozygous variants we found in *TUBE1* and *TSPAN4* were each inherited from one of the unaffected parents. Some of the alleles in these two genes are present in public databases in a higher frequency than the disease prevalence. As such, we hypothesize that incomplete penetrance or other genes may be involved in BE etiology. Also, some children carried multiple variants. For example, five variants were seen in child 1 and seven variants were detected in child 8; therefore, although they were potential candidates based on initial screening, a number of the candidate variants are likely to be rare variants not associated with BE.

In conclusion, our study identified variants in several genes that may potentially contribute to BE. Our results underscore the potential role of genetic factors in the pathogenesis of BE and provide important clues for future investigations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the NBDPS participants, scientists, staff, and Genetics Collaborative Working Group. We also thank Dr. Marilyn Browne at the Birth Defects Research Section, New York State Department of Health, and the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at https://gnomad.broadinstitute.org/about. This work was carried out in part using computing resources at the University of Minnesota Supercomputing Institute. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the National Institutes of Health, or the California Department of Public Health.

Statement of Financial Support

This work was funded by the Intramural Research Program of the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development (Contract numbers HHSN275201100001I and HHSN27500005) and by the Division of Intramural Research of the National Human Genome Research Institute, National Institutes of Health. This work was also supported through Centers for Disease Control and Prevention (CDC) cooperative agreements under PA #96043, PA #02081, FOA #DD09-001, FOA #DD13-003, and NOFO #DD18-001 to the Centers for Birth Defects Research and Prevention participating in the National Birth Defects Prevention Study (NBDPS) and/or the Birth Defects Study To Evaluate Pregnancy exposureS (BD-STEPS), and the Iowa Center for Birth Defects Research and Prevention U01 DD001035 and U01 DD001223 (PAR). The University of Washington Center for Mendelian Genomics (UW-CMG) was funded by NHGRI and NHLBI grants UM1 HG006493 and U24 HG008956.

References

Beaman GM, Woolf AS, Cervellione RM, Keene D, Mushtaq I, Urquhart JE, . . . Newman WG (2019). 22q11.2 duplications in a UK cohort with bladder exstrophy-epispadias complex. Am J Med Genet A, 179(3), 404–409. doi:10.1002/ajmg.a.61032 [PubMed: 30628148]

Berditchevski F. (2001). Complexes of tetraspanins with integrins: more than meets the eye. J Cell Sci, 114(Pt 23), 4143–4151. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11739647 [PubMed: 11739647]

Berditchevski F, & Odintsova E. (1999). Characterization of integrin-tetraspanin adhesion complexes: role of tetraspanins in integrin signaling. J Cell Biol, 146(2), 477–492. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10427099 [PubMed: 10427099]

Burk D, Sadler TW, & Langman J. (1979). Distribution of surface coat material on nasal folds of mouse embryos as demonstrated by concanavalin A binding. Anat Rec, 193(2), 185–196. doi:10.1002/ar.1091930202 [PubMed: 426293]

Carvajal-Gonzalez JM, Mulero-Navarro S, & Mlodzik M. (2016). Centriole positioning in epithelial cells and its intimate relationship with planar cell polarity. Bioessays, 38(12), 1234–1245. doi:10.1002/bies.201600154 [PubMed: 27774671]

Chang P, Giddings TH Jr., Winey M, & Stearns T. (2003). Epsilon-tubulin is required for centriole duplication and microtubule organization. Nat Cell Biol, 5(1), 71–76. doi:10.1038/ncb900 [PubMed: 12510196]

Chang P, & Stearns T. (2000). Delta-tubulin and epsilon-tubulin: two new human centrosomal tubulins reveal new aspects of centrosome structure and function. Nat Cell Biol, 2(1), 30–35. doi:10.1038/71350 [PubMed: 10620804]

Cheng HW, Hsiao CT, Chen YQ, Huang CM, Chan SI, Chiou A, & Kuo JC (2019). Centrosome guides spatial activation of Rac to control cell polarization and directed cell migration. Life Sci Alliance, 2(1). doi:10.26508/lsa.201800135

Ching BJ, Wittler L, Proske J, Yagnik G, Qi L, Draaken M, . . . Boyadjiev SA (2010). p63 (TP73L) a key player in embryonic urogenital development with significant dysregulation in human bladder exstrophy tissue. Int J Mol Med, 26(6), 861–867. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21042780 [PubMed: 21042780]

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, . . . Ruden DM (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin), 6(2), 80–92. doi:10.4161/fly.19695 [PubMed: 22728672]

Colas JF, & Schoenwolf GC (2001). Towards a cellular and molecular understanding of neurulation. Dev Dyn, 221(2), 117–145. doi:10.1002/dvdy.1144 [PubMed: 11376482]

Derrien T, Estelle J, Marco Sola S, Knowles DG, Raineri E, Guigo R, & Ribeca P. (2012). Fast computation and applications of genome mappability. PLoS One, 7(1), e30377. doi:10.1371/journal.pone.0030377

Draaken M, Baudisch F, Timmermann B, Kuhl H, Kerick M, Proske J, . . . Reutter H. (2014). Classic bladder exstrophy: Frequent 22q11.21 duplications and definition of a 414 kb phenocritical region. Birth Defects Res A Clin Mol Teratol, 100(6), 512–517. doi:10.1002/bdra.23249 [PubMed: 24764164]

Draaken M, Knapp M, Pennimpede T, Schmidt JM, Ebert AK, Rosch W, . . . Reutter H. (2015). Genome-wide association study and meta-analysis identify ISL1 as genome-wide significant susceptibility gene for bladder exstrophy. PLoS Genet, 11(3), e1005024. doi:10.1371/journal.pgen.1005024

Draaken M, Mughal SS, Pennimpede T, Wolter S, Wittler L, Ebert AK, . . . Ludwig M. (2013). Isolated bladder exstrophy associated with a de novo 0.9 Mb microduplication on chromosome 19p13.12. Birth Defects Res A Clin Mol Teratol, 97(3), 133–139. doi:10.1002/bdra.23112 [PubMed: 23359465]

Draaken M, Reutter H, Schramm C, Bartels E, Boemers TM, Ebert AK, . . . Ludwig M. (2010). Microduplications at 22q11.21 are associated with non-syndromic classic bladder exstrophy. Eur J Med Genet, 53(2), 55–60. doi:10.1016/j.ejmg.2009.12.005 [PubMed: 20060941]

Duhamel B. (1963). Embryology of Exomphalos and Allied Malformations. Arch Dis Child, 38(198), 142–147. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21032411 [PubMed: 21032411]

Ebert AK, Reutter H, Ludwig M, & Rosch WH (2009). The exstrophy-epispadias complex. Orphanet J Rare Dis, 4, 23. doi:10.1186/1750-1172-4-23 [PubMed: 19878548]

Feldkamp ML, Carey JC, & Sadler TW (2007). Development of gastroschisis: review of hypotheses, a novel hypothesis, and implications for research. Am J Med Genet A, 143A(7), 639–652. doi:10.1002/ajmg.a.31578 [PubMed: 17230493]

Froster UG, Heinritz W, Bennek J, Horn LC, & Faber R. (2004). Another case of autosomal dominant exstrophy of the bladder. Prenat Diagn, 24(5), 375–377. doi:10.1002/pd.879 [PubMed: 15164413]

Garcin C, & Straube A. (2019). Microtubules in cell migration. Essays Biochem, 63(5), 509–520. doi:10.1042/EBC20190016 [PubMed: 31358621]

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, . . . Abecasis GR (2015). A global reference for human genetic variation. Nature, 526(7571), 68–74. doi:10.1038/nature15393 [PubMed: 26432245]

Geoffroy V, Herenger Y, Kress A, Stoetzel C, Piton A, Dollfus H, & Muller J. (2018). AnnotSV: an integrated tool for structural variations annotation. Bioinformatics, 34(20), 3572–3574. doi:10.1093/bioinformatics/bty304 [PubMed: 29669011]

Greene RM, & Pratt RM (1976). Developmental aspects of secondary palate formation. J Embryol Exp Morphol, 36(2), 225–245. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1033980 [PubMed: 1033980]

Higgins CC (1962). Exstrophy of the bladder: report of 158 cases. Am Surg, 28, 99–102. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/13907242 [PubMed: 13907242]

Huang N, Lee I, Marcotte EM, & Hurles ME (2010). Characterising and predicting haploinsufficiency in the human genome. PLoS Genet, 6(10), e1001154. doi:10.1371/journal.pgen.1001154

ICBDMS. (1987). Epidemiology of bladder exstrophy and epispadias: a communication from the International Clearinghouse for Birth Defects Monitoring Systems. Teratology, 36(2), 221–227. doi:10.1002/tera.1420360210 [PubMed: 3424208]

Itan Y, Shang L, Boisson B, Patin E, Bolze A, Moncada-Velez M, . . . Casanova JL (2015). The human gene damage index as a gene-level approach to prioritizing exome variants. Proc Natl Acad Sci U S A, 112(44), 13615–13620. doi:10.1073/pnas.1518646112 [PubMed: 26483451]

Jansen S, Geuer S, Pfundt R, Brough R, Ghongane P, Herkert JC, . . . de Vries BB (2017). De Novo Truncating Mutations in the Last and Penultimate Exons of PPM1D Cause an Intellectual Disability Syndrome. Am J Hum Genet, 100(4), 650–658. doi:10.1016/j.ajhg.2017.02.005 [PubMed: 28343630]

Jenkins MM, Almli LM, Pangilinan F, Chong JX, Blue EE, Shapira SK, . . . National Birth Defects Prevention, S. (2019). Exome sequencing of family trios from the National Birth Defects Prevention Study: Tapping into a rich resource of genetic and environmental data. Birth Defects Res. doi:10.1002/bdr2.1554

Jiang D, Jiang Z, Lu D, Wang X, Liang H, Zhang J, . . . Yu L. (2019). Migrasomes provide regional cues for organ morphogenesis during zebrafish gastrulation. Nat Cell Biol, 21(8), 966–977. doi:10.1038/s41556-019-0358-6 [PubMed: 31371827]

Johnston JJ, van der Smagt JJ, Rosenfeld JA, Pagnamenta AT, Alswaid A, Baker EH, . . . Biesecker LG (2018). Autosomal recessive Noonan syndrome associated with biallelic LZTR1 variants. Genet Med, 20(10), 1175–1185. doi:10.1038/gim.2017.249 [PubMed: 29469822]

Jorgez CJ, Rosenfeld JA, Wilken NR, Vangapandu HV, Sahin A, Pham D, . . . Lamb DJ (2014). Genitourinary defects associated with genomic deletions in 2p15 encompassing OTX1. PLoS One, 9(9), e107028. doi:10.1371/journal.pone.0107028

Jun G, Flickinger M, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, . . . Kang HM (2012). Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. Am J Hum Genet, 91(5), 839–848. doi:10.1016/j.ajhg.2012.09.004 [PubMed: 23103226]

Kajbafzadeh AM, Tajik P, Payabvash S, Farzan S, & Solhpour AR (2006). Bladder exstrophy and epispadias complex in sibling: case report and review of literature. Pediatr Surg Int, 22(9), 767–770. doi:10.1007/s00383-006-1741-5 [PubMed: 16896811]

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, . . . MacArthur DG (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature, 581(7809), 434–443. doi:10.1038/s41586-020-2308-7 [PubMed: 32461654]

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, . . . MacArthur DG (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv, 531210. doi:10.1101/531210

Kashef J, Diana T, Oelgeschlager M, & Nazarenko I. (2013). Expression of the tetraspanin family members Tspan3, Tspan4, Tspan5 and Tspan7 during Xenopus laevis embryonic development. Gene Expr Patterns, 13(1–2), 1–11. doi:10.1016/j.gep.2012.08.001 [PubMed: 22940433]

Katoh I, Fukunishi N, Fujimuro M, Kasai H, Moriishi K, Hata R, & Kurata S. (2016). Repression of Wnt/beta-catenin response elements by p63 (TP63). Cell Cycle, 15(5), 699–710. doi:10.1080/15384101.2016.1148837 [PubMed: 26890356]

Kent WJ (2002). BLAT--the BLAST-like alignment tool. Genome Res, 12(4), 656–664. doi:10.1101/gr.229202 [PubMed: 11932250]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, & Haussler D. (2002). The human genome browser at UCSC. Genome Res, 12(6), 996–1006. doi:10.1101/gr.229102 [PubMed: 12045153]

Lattimer JK, & Smith MJ (1966). Exstrophy closure: a followup on 70 cases. J Urol, 95(3), 356–359. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/5906001 [PubMed: 5906001]

Lee HY, & Nagele RG (1985). Studies on the mechanisms of neurulation in the chick: interrelationship of contractile proteins, microfilaments, and the shape of neuroepithelial cells. J Exp Zool, 235(2), 205–215. doi:10.1002/jez.1402350207 [PubMed: 3903030]

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, . . . Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536(7616), 285–291. doi:10.1038/nature19057 [PubMed: 27535533]

Levy S, Todd SC, & Maecker HT (1998). CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. Annu Rev Immunol, 16, 89–109. doi:10.1146/annurev.immunol.16.1.89 [PubMed: 9597125]

Li H, & Durbin R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics, 26(5), 589–595. doi:10.1093/bioinformatics/btp698 [PubMed: 20080505]

Lundin J, Soderhall C, Lunden L, Hammarsjo A, White I, Schoumans J, . . . Nordenskjold A. (2010). 22q11.2 microduplication in two patients with bladder exstrophy and hearing impairment. Eur J Med Genet, 53(2), 61–65. doi:10.1016/j.ejmg.2009.11.004 [PubMed: 20045748]

Maecker HT, Todd SC, & Levy S. (1997). The tetraspanin superfamily: molecular facilitators. FASEB J, 11(6), 428–442. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9194523 [PubMed: 9194523]

Mahfuz I, Darling T, Wilkins S, White S, & Cheng W. (2013). New insights into the pathogenesis of bladder exstrophy-epispadias complex. J Pediatr Urol, 9(6 Pt B), 996–1005. doi:10.1016/j.jpurol.2013.05.001 [PubMed: 23743131]

Martinez-Frias ML, Bermejo E, Rodriguez-Pinilla E, & Frias JL (2001). Exstrophy of the cloaca and exstrophy of the bladder: two different expressions of a primary developmental field defect. Am J Med Genet, 99(4), 261–269. doi:10.1002/ajmg.1210 [PubMed: 11251990]

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, . . . DePristo MA (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res, 20(9), 1297–1303. doi:10.1101/gr.107524.110 [PubMed: 20644199]

Messelink EJ, Aronson DC, Knuist M, Heij HA, & Vos A. (1994). Four cases of bladder exstrophy in two families. J Med Genet, 31(6), 490–492. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8071977 [PubMed: 8071977]

Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, . . . Finn RD (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res, 47(D1), D351–D360. doi:10.1093/nar/gky1100 [PubMed: 30398656]

Nakamura K, Iwamoto R, & Mekada E. (1995). Membrane-anchored heparin-binding EGF-like growth factor (HB-EGF) and diphtheria toxin receptor-associated protein (DRAP27)/CD9 form a complex with integrin alpha 3 beta 1 at cell-cell contact sites. J Cell Biol, 129(6), 1691–1705. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7790364 [PubMed: 7790364]

Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, & Durbin R. (2016). BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics, 32(11), 1749–1751. doi:10.1093/bioinformatics/btw044 [PubMed: 26826718]

Nelson CP, Dunn RL, & Wei JT (2005). Contemporary epidemiology of bladder exstrophy in the United States. J Urol, 173(5), 1728–1731. doi:10.1097/01.ju.0000154821.21521.9b [PubMed: 15821570]

NHLBI-ESP. Retrieved from https://esp.gs.washington.edu/drupal/

Niemann S, Zhao C, Pascu F, Stahl U, Aulepp U, Niswander L, . . . Muller U. (2004). Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family. Am J Hum Genet, 74(3), 558–563. doi:10.1086/382196 [PubMed: 14872406]

Nobes CD, & Hall A. (1995). Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. Cell, 81(1), 53–62. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7536630 [PubMed: 7536630]

Pedersen BS, & Quinlan AR (2017). Who's Who? Detecting and Resolving Sample Anomalies in Human DNA Sequencing Studies with Peddy. Am J Hum Genet, 100(3), 406–413. doi:10.1016/j.ajhg.2017.01.017 [PubMed: 28190455]

Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, . . . Nejentsev S. (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics, 28(21), 2747–2754. doi:10.1093/bioinformatics/bts526 [PubMed: 22942019]

Qi L, Chen K, Hur DJ, Yagnik G, Lakshmanan Y, Kotch LE, . . . Boyadjiev SA (2011). Genome-wide expression profiling of urinary bladder implicates desmosomal and cytoskeletal dysregulation in the bladder exstrophy-epispadias complex. Int J Mol Med, 27(6), 755–765. doi:10.3892/ijmm.2011.654 [PubMed: 21431277]

Rasmussen SA, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA, & National Birth Defects Prevention, S. (2003). Guidelines for case classification for the National Birth Defects Prevention Study. Birth Defects Res A Clin Mol Teratol, 67(3), 193–201. doi:10.1002/bdra.10012 [PubMed: 12797461]

Reefhuis J, Gilboa SM, Anderka M, Browne ML, Feldkamp ML, Hobbs CA, . . . National Birth Defects Prevention, S. (2015). The National Birth Defects Prevention Study: A review of the methods. Birth Defects Res A Clin Mol Teratol, 103(8), 656–669. doi:10.1002/bdra.23384 [PubMed: 26033852]

Reinfeldt Engberg G, Mantel A, Fossum M, & Nordenskjold A. (2016). Maternal and fetal risk factors for bladder exstrophy: A nationwide Swedish case-control study. J Pediatr Urol, 12(5), 304 e301–304 e307. doi:10.1016/j.jpurol.2016.05.035

Rentzsch P, Witten D, Cooper GM, Shendure J, & Kircher M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res, 47(D1), D886–D894. doi:10.1093/nar/gky1016 [PubMed: 30371827]

Reutter H, Betz RC, Ludwig M, & Boemers TM (2006). MTHFR 677 TT genotype in a mother and her child with Down syndrome, atrioventricular canal and exstrophy of the bladder: implications of a mutual genetic risk factor? Eur J Pediatr, 165(8), 566–568. doi:10.1007/s00431-006-0116-1 [PubMed: 16602006]

Reutter H, Bokenkamp A, Ebert AK, Rosch W, Boemers TM, Nothen MM, & Ludwig M. (2009). Possible association of Down syndrome and exstrophy-epispadias complex: report of two new cases and review of the literature. Eur J Pediatr, 168(7), 881–883. doi:10.1007/s00431-008-0852-5 [PubMed: 18923839]

Reutter H, Boyadjiev SA, Gambhir L, Ebert AK, Rosch WH, Stein R, . . . Jenetzky E. (2011). Phenotype severity in the bladder exstrophy-epispadias complex: analysis of genetic and nongenetic contributing factors in 441 families from North America and Europe. J Pediatr, 159(5), 825–831 e821. doi:10.1016/j.jpeds.2011.04.042 [PubMed: 21679965]

Reutter H, Draaken M, Pennimpede T, Wittler L, Brockschmidt FF, Ebert AK, . . . Mattheisen M. (2014). Genome-wide association study and mouse expression data identify a highly conserved 32 kb intergenic region between WNT3 and WNT9b as possible susceptibility locus for isolated classic exstrophy of the bladder. Hum Mol Genet, 23(20), 5536–5544. doi:10.1093/hmg/ddu259 [PubMed: 24852367]

Reutter H, Hoischen A, Ludwig M, Stein R, Radlwimmer B, Engels H, . . . Weber RG (2007). Genome-wide analysis for micro-aberrations in familial exstrophy of the bladder

using array-based comparative genomic hybridization. BJU Int, 100(3), 646–650. doi:10.1111/j.1464-410X.2007.07086.x [PubMed: 17669146]

Reutter H, Keppler-Noreuil K, C EK, Thiele H, Yamada G, & Ludwig M. (2016). Genetics of Bladder-Exstrophy-Epispadias Complex (BEEC): Systematic Elucidation of Mendelian and Multifactorial Phenotypes. Curr Genomics, 17(1), 4–13. doi:10.2174/1389202916666151014221806 [PubMed: 27013921]

Reutter H, Qi L, Gearhart JP, Boemers T, Ebert AK, Rosch W, . . . Boyadjiev SA (2007). Concordance analyses of twins with bladder exstrophy-epispadias complex suggest genetic etiology. Am J Med Genet A, 143A(22), 2751–2756. doi:10.1002/ajmg.a.31975 [PubMed: 17937426]

Reutter H, Shapiro E, & Gruen JR (2003). Seven new cases of familial isolated bladder exstrophy and epispadias complex (BEEC) and review of the literature. Am J Med Genet A, 120A(2), 215–221. doi:10.1002/ajmg.a.20057 [PubMed: 12833402]

Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, . . . Martin CL (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med, 22(2), 245–257. doi:10.1038/s41436-019-0686-8 [PubMed: 31690835]

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, & Mesirov JP (2011). Integrative genomics viewer. Nat Biotechnol, 29(1), 24–26. doi:10.1038/nbt.1754 [PubMed: 21221095]

Sadler TW (2005). Embryology of neural tube development. Am J Med Genet C Semin Med Genet, 135C(1), 2–8. doi:10.1002/ajmg.c.30049 [PubMed: 15806586]

Sadler TW (2010). The embryologic origin of ventral body wall defects. Semin Pediatr Surg, 19(3), 209–214. doi:10.1053/j.sempedsurg.2010.03.006 [PubMed: 20610194]

Sadler TW, & Feldkamp ML (2008). The embryology of body wall closure: relevance to gastroschisis and other ventral body wall defects. Am J Med Genet C Semin Med Genet, 148C(3), 180–185. doi:10.1002/ajmg.c.30176 [PubMed: 18655098]

Shapiro E, Lepor H, & Jeffs RD (1984). The inheritance of the exstrophy-epispadias complex. J Urol, 132(2), 308–310. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6737583 [PubMed: 6737583]

Shaw AR, Domanska A, Mak A, Gilchrist A, Dobler K, Visser L, . . . Willett BJ (1995). Ectopic expression of human and feline CD9 in a human B cell line confers beta 1 integrin-dependent motility on fibronectin and laminin substrates and enhanced tyrosine phosphorylation. J Biol Chem, 270(41), 24092–24099. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7592610 [PubMed: 7592610]

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, & Sirotkin K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res, 29(1), 308–311. doi:10.1093/nar/29.1.308 [PubMed: 11125122]

Siffel C, Correa A, Amar E, Bakker MK, Bermejo-Sanchez E, Bianca S, . . . Olney RS (2011). Bladder exstrophy: an epidemiologic study from the International Clearinghouse for Birth Defects Surveillance and Research, and an overview of the literature. Am J Med Genet C Semin Med Genet, 157C(4), 321–332. doi:10.1002/ajmg.c.30316 [PubMed: 22002949]

Soderhall C, Lundin J, Lagerstedt-Robinson K, Grigelioniene G, Lackgren G, Kockum CC, & Nordenskjold A. (2014). A case with bladder exstrophy and unbalanced X chromosome rearrangement. Eur J Pediatr Surg, 24(4), 353–359. doi:10.1055/s-0033-1349056 [PubMed: 23801353]

Su L, Zhang Z, Gan L, Jiang Q, Xiao P, Zou J, . . . Jiang H. (2016). Deregulation of the planar cell polarity genes CELSR3 and FZD3 in Hirschsprung disease. Exp Mol Pathol, 101(2), 241–248. doi:10.1016/j.yexmp.2016.09.003 [PubMed: 27619161]

Utsch B, DiFeo A, Kujat A, Karle S, Schuster V, Lenk H, . . . Trobs RB (2006). Bladder exstrophy and Epstein type congenital macrothrombocytopenia: evidence for a common cause? Am J Med Genet A, 140(20), 2251–2253. doi:10.1002/ajmg.a.31454 [PubMed: 16969870]

von Lowtzow C, Hofmann A, Zhang R, Marsch F, Ebert AK, Rosch W, . . . Draaken M. (2016). CNV analysis in 169 patients with bladder exstrophy-epispadias complex. BMC Med Genet, 17(1), 35. doi:10.1186/s12881-016-0299-x [PubMed: 27138190]

Wang K, Li M, & Hakonarson H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res, 38(16), e164. doi:10.1093/nar/gkq603 [PubMed: 20601685]

Yanez-Mo M, Alfranca A, Cabanas C, Marazuela M, Tejedor R, Ursa MA, . . . Sanchez-Madrid F. (1998). Regulation of endothelial cell motility by complexes of tetraspan molecules CD81/TAPA-1 and CD151/PETA-3 with alpha3 beta1 integrin localized at endothelial lateral junctions. J Cell Biol, 141(3), 791–804. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9566977 [PubMed: 9566977]

Yang P, Khoury MJ, Stewart WF, Beaty TH, Chee E, Beatty JC, . . . Gordis L. (1994). Comparative epidemiology of selected midline congenital abnormalities. Genet Epidemiol, 11(2), 141–154. doi:10.1002/gepi.1370110205 [PubMed: 8013895]

Zhang R, Knapp M, Kause F, Reutter H, & Ludwig M. (2017). Role of the LF-SINE-Derived Distal ISL1 Enhancer in Patients with Classic Bladder Exstrophy. J Pediatr Genet, 6(3), 169–173. doi:10.1055/s-0037-1602387 [PubMed: 28794909]

Zukauskas A, Merley A, Li D, Ang LH, Sciuto TE, Salman S, . . . Jaminet SC (2011). TM4SF1: a tetraspanin-like protein necessary for nanopodia formation and endothelial cell migration. Angiogenesis, 14(3), 345–354. doi:10.1007/s10456-011-9218-0 [PubMed: 21626280]

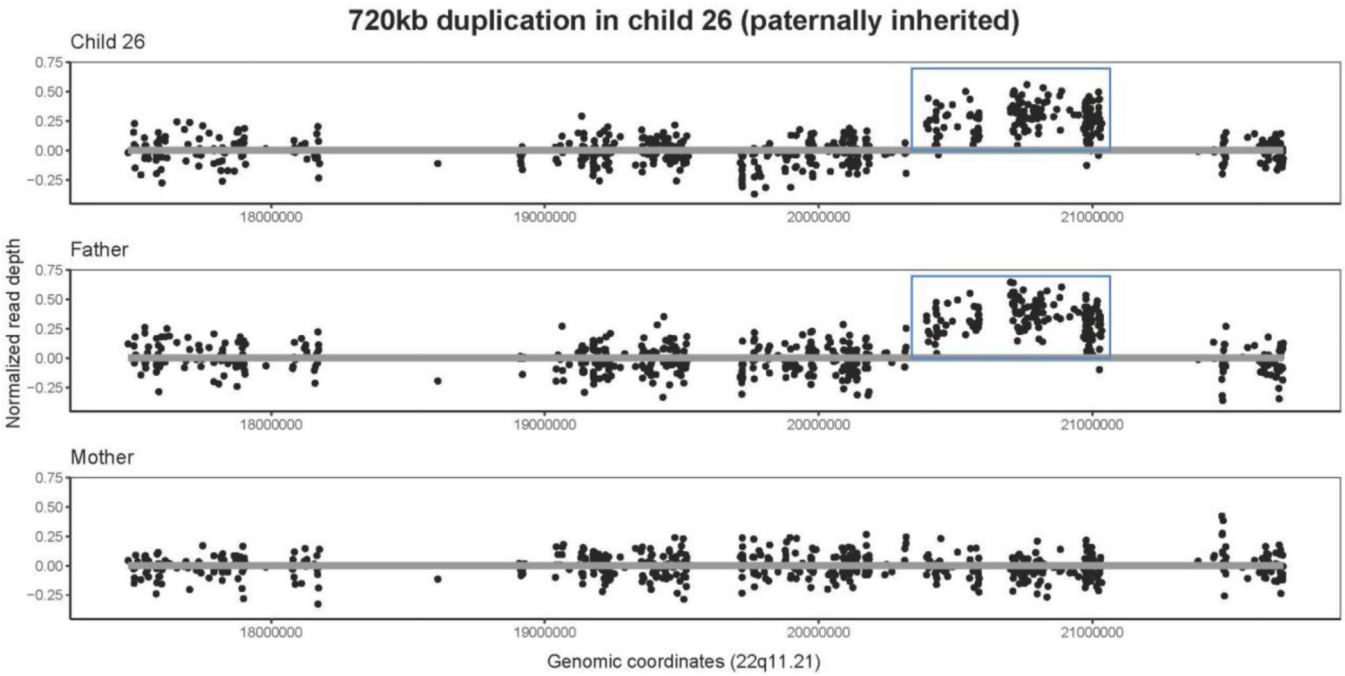

Figure 1.

Population normalized read depths generated by Genvisis (http://www.genvisis.org) for each target in the NimbleGen SeqCap EZ Exome+UTR Library (Version 3.0) within the 22q11.21 region.

The rectangle depicts the paternally inherited 720kb duplication in child 26 that was detected by ExomeDepth

Table 1a.

Selected child and maternal characteristics of sequenced children with bladder exstrophy, National Birth Defects Prevention Study, 1997–2011

Characteristic	Sequenced children with BE N (%)
Child	
Sex	
Male	18 (69.2)
Female	8 (30.8)
Gestational age (weeks)	
Preterm (<37)	3 (11.6)
Term (37–45)	23 (88.4)
Plurality	
Singleton	26 (100)
Maternal	
Age at delivery (years)	
20–29	13 (50.0)
30–39	12 (46.2)
≥40	1 (3.8)
Race/Ethnicity[a]	
Non-Hispanic white	19 (73.1)
Hispanic	3 (11.5)
Asian/Pacific Islander	1 (3.9)
Other	3 (11.5)

BE bladder exstrophy

[a] self-reported

Table 1b.

Additional malformations seen in bladder exstrophy cases.

Malformation	Number of children (%)
Abnormal genitalia[a]	13 (50.0)
Epispadias	12 (46.1)
Anteriorly displaced anus	9 (34.6)
Heart defects[b]	7 (26.9)
Pelvic diastasis	5 (19.2)
Inguinal hernia	4 (15.4)
Skeletal dysmorphology[c]	4 (15.4)
Congenital hip dislocation	3 (11.5)
Craniofacial dysmorphology	3 (11.5)
Dorsal chordee	2 (7.7)
Displaced umbilicus	2 (7.7)
Other:	
Undescended testicles	1 (3.8)
Omphalocele	1 (3.8)
Patent urachus	1 (3.8)
Urethral atresia	1 (3.8)
Absent kidney	1 (3.8)

[a]Includes: bifid clitoris, bifid penis, hypospadias, micropenis, small bifid phallus, short urethra

[b]Includes: absent pulmonary valve, atrial septal defect, dysplastic pulmonary valve, patent foramen ovale, patent ductus arteriosus, pulmonary artery stenosis, Tetralogy of Fallot, tricuspid regurgitation, transposition of great vessels, ventricular septal defect

[c]Includes: fused ribs, fused vertebrae, hypoplastic toenails, small anterior fontanelle, thumb hypoplasia

Table 2.

Validated candidate bladder exstrophy-associated variants, National Birth Defects Prevention Study, 1997–2011

2a. Variants in genes that met primary filtering criteria[a]

Child	Gene	Variant	Locus (hg38)	AF[b]	CADD	GDI	Gene Function
1	*TUBE1*	p.R365X[c]	6:112072759	0	37.0	49.1%	Central role in organization of microtubules during centriole duplication
2	*TUBE1*	p.V225fs[c]	6:112076075	0	-	49.1%	Same as above
3	*TSPAN4*	p.A137fs[c]	11:865592	0.0000077	-	40.3%	Member of the tetraspanin family that plays a role in cell adhesion and motility
4	*TSPAN4*	p.R164Q[c]	11:865752	0.0001	23.1	40.3%	Same as above
5	*TSPAN4*	p.A233G[c]	11:866611	0.000071	21.9	40.3%	Same as above

AF allele frequency, *CADD* combined annotation dependent depletion score, *GDI* gene damaging index, *GQ* genotype quality

[a]Allelic depth ≥10, GQ ≥ 50, AF ≤0.0001, GDI <50th percentile, ≥2 loss-of-function variants or ≥3 loss-of-function or missense variants

[b]Maximum AF observed in public databases for any subpopulation

Transcripts used for annotation are provided in the Supplemental Table 1 with Integrative Genomics Viewer screenshots

[c]Variant confirmed in proband and parent via Sanger sequencing

2b.

Variants in genes previously reported to be associated with bladder exstrophy that met filtering criteria[a] and were validated by Sanger sequencing, ordered by GDI.

Child	Gene	Variant	Locus (hg38)	AF[b]	CADD	GDI	Gene Function (Associated Disease)
9	*WNT3*	p.D96H[c,e]	17:46773704	0.000067	21.6	4.4%	wingless-type MMTV integration site family member 3 (tetra-amelia syndrome)
6	*CRKL*	p.T26A[c,f]	22:20918010	0.000036	12.9	16.4%	CRK like proto-oncogene, adaptor protein, substrate of BCR-ABL tyrosine kinase
7	*LZTR1*	p.R210X[d,e]	22:20989659	0.00042	40.0	28.6%	leucine zipper-like transcription regulator 1 (Noonan syndrome 2, 10; susceptibility to schwannomatosis-2)
8	*MYH9*	p.E1653G[c,e]	22:36286821	0.000089	31.0	83.2%	myosin heavy chain 9 (autosomal dominant deafness 17; macrothrombocytopenia and granulocyte inclusions with or without nephritis or sensorineural hearing loss)

AF allele frequency, *CADD* combined annotation dependent depletion score, *GDI* gene damaging index

[a] AF ≤0.001, No quality filters

[b] Maximum AF observed in public databases for any subpopulation

[c] paternally inherited variant

[d] maternally inherited variant

[e] Variant confirmed in proband and parent via Sanger sequencing

[f] Variant confirmed in proband via Sanger sequencing. Parental DNA was not available for validation studies

Table 3.

Homozygous variants meeting filtering criteria[a], National Birth Defects Prevention Study, 1997–2011, Ordered by child and by GDI.

Child	Gene	Variant	Locus (hg38)	[Reference, Alternate]	GQ	AD	AF[b]	CADD	GDI
12	*TLE4*	p.S165T	9:79652696	[C, C]	99	[2, 97]	0.006	17.7	37.7%
12	*TMEM71*	p.S284L	8:132710946	[A, A]	99	[0, 49]	0.0036	2.3	39.6%
24	*GLMN*	p.A146T	1:92289110	[T, T]	99	[1, 105]	0.0075	18.6	55.0%

AD allelic depth for [Reference, Alternate], *AF* allele frequency, *CADD* combined annotation dependent depletion score, *GDI* gene damaging index, *GQ* genotype quality

[a]AD ≥10, GQ ≥50, AF ≤0.01, GDI <75th percentile

[b]Maximum AF observed in public databases for any subpopulation

Table 4.

Variants in X-linked genes (recessive inheritance model) meeting filtering criteria[a], National Birth Defects Prevention Study, 1997–2011, ordered by GDI.

Child[b]	Gene	Variant	Locus (hg38)	[Allele]	GQ	AD	AF[c]	CADD	GDI
25	*PCSK1N*	p.P178A	X:48831924	[C]	57	[0, 19]	0	10.7	1.9%
25	*FAM46D*	p.R260S	X:80443317	[A]	66	[0, 22]	0	26.5	8.2%
9	*MAGEB3*	p.S97P	X:30236213	[C]	99	[0, 49]	0.000086	2.0	9.0%
19	*ZFP92*	p.S58P	X:153420239	[C]	87	[0, 29]	0.000018	10.7	23.7%
4	*MED12*	p.A735T	X:71125123	[A]	66	[0, 22]	0.000082	16.9	31.2%
23	*FAAH2*	p.A27T	X:57286904	[A]	78	[0, 26]	0.00004	5.1	32.5%
15	*MTMR8*	p.L35V	X:64359449	[C]	99	[0, 72]	0.00009	12.2	33.2%
14	*GABRE*	c.1137+1G>A	X:151955367	[T]	78	[0, 26]	0	23.2	38.0%
15	*MAGED2*	p.A563T	X:54815548	[A]	99	[0, 44]	0.000071	9.0	44.2%
13	*MAGEE1*	p.S65R	X:76428125	[A]	99	[0, 61]	0.000095	14.9	51.0%
19	*ZCCHC18*	p.T129A	X:104114496	[G]	99	[0, 47]	0	9.0	55.3%
25	*SHROOM2*	p.S892A	X:9896582	[G]	99	[0, 50]	0	9.8	58.9%
14	*COL4A6*	p.S1359P	X:108163036	[G]	51	[0, 17]	0.000047	10.2	60.4%

AD allelic depth for [Allele], *AF* allele frequency, *CADD* combined annotation dependent depletion score, *GDI* gene damaging index, *GQ* genotype quality

[a] AD ≥10, GQ ≥50, AF ≤0.0001, GDI <75th percentile

[b] All children are male

[c] Maximum AF observed in public databases for any subpopulation

Table 5.

De novo missense and loss-of-function variants meeting filtering criteria[a], National Birth Defects Prevention Study, 1997–2011, ordered by GDI.

Child	Gene	Variant	Locus (hg38)	[Allele1, Allele2]	GQ	AD	AF[b]	CADD	GDI
12	*ACTB*	p.T351A	7:5527825	[T, C]	99	[30, 33]	0	23.5	1.4%
10	*ARHGAP36*	p.R441C	X:131086404	[C, T]	99	[63, 56]	0.0001	33.0	12.5%
5	*TUBGCP3*	p.T653I	13:112516568	[G, A]	99	[27, 23]	0	25.7	26.3%
4	*USP18*[c]	p.V132M	22:18161929	[G, A]	99	[23, 29]	0	23.9	31.9%
13	*c11orf84*	p.R327H	11:63826973	[G, A]	99	[24, 30]	0.0006	23.5	38.6%
14	*PPM1D*	p.E525X	17:60663307	[G, T]	99	[40, 42]	0.0002	44.0	40.6%
15	*PRRX1*	p.W141fs	1:170726223	[TG, T]	99	[30, 35]	0	-	45.3%
7	*AKR1C2*[c]	p.R96X	10:5000633	[G, A]	99	[92, 102]	0.0006	33.0	52.0%
15	*POLE2*	p.L225F	14:49664633	[T, G]	99	[58, 55]	0	26.2	53.2%
17	*MAP3K20*	p.G767R	2:173266646	[G, C]	99	[54, 54]	0	25.0	78.6%
16	*KMT2D*	p.R2922Q	12:49038591	[C, T]	99	[43, 35]	0.0001	24.3	78.8%
2	*FAHD1*	p.Y171C	16:1827741	[A, G]	99	[49, 23]	0.000036	23.7	79.7%
17	*BSN*	p.M3124T	3:49661216	[T, C]	99	[41, 34]	0.000018	9.1	80.4%
13	*HIPK1*	p.L74R	1:113940604	[T, G]	99	[25, 27]	0.0003	22.2	88.0%
18	*PKDREJ*	p.P847L	22:46260783	[G, A]	99	[76, 25]	0.00022	23.5	92.0%
18	*SLC22A24*	p.E295X	11:63118859	[C, A]	99	[53, 52]	0	2.6	92.1%
4	*KRT33A*	p.A373T	17:41346217	[C, T]	99	[32, 33]	0.0001	18.2	92.4%
7	*C7*	p.L283V	5:40947710	[C, G]	99	[78, 87]	0	27.4	94.0%
19	*ITPKB*	p.G279C	1:226736624	[C, A]	99	[34, 29]	0	15.2	97.1%
14	*TG*	p.S633F	8:132887270	[C, T]	99	[24, 31]	0	8.9	98.2%

AD allelic depth for [Allele1, Allele2], *AF* allele frequency, *CADD* combined annotation dependent depletion score, *GDI* gene damaging index, *GQ* genotype quality

[a] AD≥ 8, GQ ≥40, Max AF ≤0.001 in public databases

[b] Maximum AF observed in public databases for any subpopulation

[c] Validated *de novo* variants in probands. Variant not detected in either parent