Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Birth Defects Res. Author manuscript; available in PMC 2023 April 01.

Published in final edited form as:

Birth Defects Res. 2022 April ; 114(7): 215–227. doi:10.1002/bdr2.1987.

Exome sequencing identifies variants in infants with sacral agenesis

Georgia Pitsava[1], **Marcia L. Feldkamp**[2], **Nathan Pankratz**[3], **John Lane**[3], **Denise M. Kay**[4], **Kristin M. Conway**[5], **Charlotte Hobbs**[6], **Gary M. Shaw**[7], **Jennita Reefhuis**[8], **Mary M. Jenkins**[8], **Lynn M. Almli**[8], **Cynthia Moore**[8], **Martha Werler**[9,10], **Marilyn L. Browne**[11,12], **Chris Cunniff**[13], **Andrew F. Olshan**[14], **Faith Pangilinan**[15], **Lawrence C. Brody**[15], **Robert J. Sicko**[4], **Richard H. Finnell**[16], **Michael J. Bamshad**[17], **Daniel McGoldrick**[18], **Deborah A. Nickerson**[18], **James C. Mullikin**[19], **Paul A. Romitti**[5,#], **James L. Mills**[1,#],

UW 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, USA

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

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

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

Corresponding author: James L. Mills, MD, MS, 6710B Rockledge Drive Room 3117, NICHD, NIH, Bethesda, MD, 20892, millsj@exchange.nih.gov.

[#]Co-senior authors

Author Contributions

Study design: Brody, Browne, Feldkamp, Hobbs, Jenkins, Moore, Olshan, Reefhuis, Romitti, Shaw, Werler, National Birth Defects Prevention Study

Data and specimen acquisition: Almli, Browne, Conway, Feldkamp, Finnell, Hobbs, Jenkins, Moore, Olshan, Reefhuis, Keegan, Romitti, Shaw, Werler, National Birth Defects Prevention Study

Experimental studies: Bamshad, Brody, Kay, McGoldrick, Mullikin, Nickerson, Pangilinan, Sicko, NISC Comparative Sequencing Program, University of Washington Center for Mendelian Genomics,

Statistical analysis: Lane, Pankratz

Data interpretation: Bamshad, Brody, Feldkamp, Kay, Lane, McGoldrick, Mills, Nickerson, Pankratz, Pitsava, Romitti, Sicko, University of Washington Center for Mendelian Genomics

Initial draft of manuscript: Pitsava

Critical revision of manuscript: Feldkamp, Kay, Lane, Mills, Pankratz, Pangilinan, Romitti

Review of manuscript: Almli, Bamshad, Brody, Conway, Cunniff, Finnell, Hobbs, Jenkins, McGoldrick, Mullikin, Nickerson, Olshan, Reefhuis, Keegan, Browne, Canfield, Shaw, Sicko

Study supervision: Mills, Romitti

Following analyses by the UW-CMG analytic group, the results were additionally reviewed by Dr. Nathan Pankratz.

Compliance with ethical standards

Conflicts of interest None declared, except that Dr. Finnell was formerly in a leadership position with TeratOmic Consulting LLC, a now dissolved consulting firm.

Informed Consent

Signed informed consent was obtained for all participants of the study.

Ethics approval The National Birth Defects Prevention Study (NBDPS) protocol was approved by the institutional review board at each NBDPS site.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

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

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

[7]Department of Pediatrics, Stanford University School of Medicine, Stanford, California, USA

[8]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA, USA

[9]Slone Epidemiology Center at Boston University, Boston, MA

[10]Department of Epidemiology, School of Public Health, Boston University, Boston, MA

[11]New York State Department of Health, Birth Defects Registry, Albany, New York, USA

[12]Department of Epidemiology and Biostatistics, University at Albany School of Public Health, Rensselaer, New York, USA

[13]Department of Pediatrics, Weill Cornell Medical College, New York, NY, USA

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

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

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

[17]Department of Pediatrics, University of Washington, Seattle, Washington, USA

[18]Department of Genome Sciences, University of Washington, Seattle, Washington, USA

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

Abstract

Background—Sacral agenesis (SA) consists of partial or complete absence of the caudal end of the spine and often presents with additional birth defects. Several studies have examined gene variants for syndromic forms of SA, but only one has examined exomes of children with non-syndromic SA.

Methods—Using buccal cell specimens from families of children with non-syndromic SA, exomes of 28 child-parent trios (eight with and 20 without a maternal diagnosis of pregestational diabetes) and two child-father duos (neither with diagnosis of maternal pregestational diabetes) were exome sequenced.

Results—Three children had heterozygous missense variants in *ID1* (Inhibitor of DNA Binding 1), with CADD scores >20 (top 1% of deleterious variants in the genome); two children inherited the variant from their fathers and one from the child's mother. Rare missense variants were also detected in *PDZD2* (PDZ Domain Containing 2; N=1) and *SPTBN5* (Spectrin Beta, Non-erythrocytic 5; N=2), two genes previously suggested to be associated with SA etiology. Examination of variants with autosomal recessive and X-linked recessive inheritance identified five and two missense variants, respectively. Compound heterozygous variants were identified in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

several genes. In addition, 12 *de novo* variants were identified, all in different genes in different children.

Conclusions—To our knowledge, this is the first study reporting a possible association between *ID1* and non-syndromic SA. Although maternal pregestational diabetes has been strongly associated with SA, the missense variants in *ID1* identified in two of three children were paternally inherited. These findings add to the knowledge of gene variants associated with non-syndromic SA and provide data for future studies.

Keywords

variant; *ID1* ; sacral agenesis; birth defects; congenital abnormality

1 | INTRODUCTION

Sacral agenesis (SA) is a rare (0.01–0.05 per 1,000 live births [(Caird, Hall, Bloom, Park, & Farley, 2007)]) birth defect characterized by varying degrees of agenesis of the lower spinal column (Sonek et al., 1990). Respiratory, genitourinary, central nervous system, and gastrointestinal defects have been reported with SA (Boulas, 2009; Pang, 1993; Sarnat, Case, & Graviss, 1976). Few factors, genetic or environmental, have been reported that contribute to the development of SA. Maternal diabetes mellitus during organogenesis is by far the most important. Maternal smoking has also been reported to be associated with increased risk of SA in the National Birth Defects Prevention Study (NBDPS) (Nalbandyan et al., 2019).

Maternal pregestational diabetes is a known risk factor for SA, with a relative risk of approximately 200 in infants of diabetic mothers compared to infants of non-diabetic mothers, much higher than any other birth defect (Aslan, Yanik, Celikaslan, Yildirim, & Ceylan, 2001; Boulas, 2009; Mills, 1982; Mills, Baker, & Goldman, 1979; Tinker et al., 2020; Wender-Ozegowska et al., 2005). Insulin-dependent diabetes mellitus has been reported in 15–25% of mothers of children affected with SA (Lynch, Wang, Strachan, Burn, & Lindsay, 2000), with poor metabolic control being a key risk factor in diabetic women (Mills, 2010). Human or animal evidence for additional environmental risk factors is limited.

Several studies have reported on pathogenic variants in individuals with syndromic forms of SA (Belloni et al., 2000; Cretolle et al., 2006; Hagan et al., 2000; Kibar et al., 2007; Lynch et al., 2000; Nowaczyk et al., 2000; Postma et al., 2014; Ross et al., 1998). To date, only a small pilot whole exome sequencing and copy number variant study conducted on four child-parent trios has reported on variants for non-syndromic SA (Porsch et al., 2016). Advancing this work, the current study conducted whole exome sequencing on 28 child-parent trios and two child-father duos with non-syndromic SA. This study was designed to add to our limited knowledge of gene variants possibly associated with non-syndromic SA and to contribute to future studies.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2 | METHODS

2.1 | Study Sample

The National Birth Defects Prevention Study (NBDPS) was a multisite, population-based case-control study that investigated risk factors for more than 30 major structural birth defects (Reefhuis et al., 2015). The birth defect surveillance program at each NBDPS site (Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah) ascertained children with NBDPS-eligible defects among pregnancies with estimated dates of delivery during the period October 1, 1997 through December 31, 2011. Over 6.5 million live births were included in the NBDPS. The study protocol was approved by the institutional review board at each NBDPS site.

Children with SA diagnosed by physical and imaging examinations were included in the NBDPS. Clinical information abstracted from medical records was reviewed by a board-certified clinical geneticist at each site to verify eligibility. One clinical geneticist performed a final classification of each child with SA to ensure consistency in classification across the NBDPS sites (Rasmussen et al., 2003). Only children diagnosed with SA and classified as isolated or multiple (i.e. non-syndromic) were eligible for the NBDPS; those with known chromosomal or monogenic etiologies were excluded. Isolated SA was defined as sacral defects with or without additional axial skeletal defects as well as with or without sequence or secondary defects, such as lower limb deformations or hypoplasia and/or tethered spinal cord. Children with additional major birth defects considered unrelated to SA, including appendicular skeletal defects, were classified as multiple. Due to informed consent requirements, we listed the defects diagnosed in aggregate.

Data collection for the NBDPS has been described in detail elsewhere (Reefhuis et al., 2015). Information about the pregnancy, sociodemographic information and exposures was collected by maternal interview. Each mother was also asked to collect buccal cell specimens from herself, the child's father (if available), and their child (if living). The telephone interview was completed by 109 (63.0%) of 173 mothers; six pregnancies ended in fetal loss, leaving 103 mothers of liveborn children. Of these, 60 (58.2%) mothers, 43 (41.7%) fathers, and 53 (51.5%) children with SA provided buccal cell specimens: 39 child-parent trios, 14 mother-child duos, four mother-father duos, and three mothers only. The 39 child-parent trios were evaluated for sequencing; five of these trios were excluded because of poor DNA quality, one failed quality control, three had child specimens with very low DNA, and two had maternal specimens with very low DNA, leaving 28 trios and two child-father duos for analysis.

2.2 | Specimen processing and sequencing

Buccal cell specimens with adequate DNA (≥200 ng, assessed by quantitative real-time PCR targeting the RNaseP gene) were sent to the National Institutes of Health Intramural Sequencing Center at the National Human Genome Research Institute. DNA from 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. For exome enrichment and coverage, libraries were prepared using the Nimblegen

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SeqCap EZ Exome + UTR Library (v3.0) kit and sequenced on an Illumina HiSeq 2500 instrument (Jenkins et al., 2019).

Regarding exome sequencing, 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, five to six 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. The pipeline included 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).

2.4 | Annotation

All variants were annotated for functional impact (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, including dbSNP (Sherry et al., 2001) version 151, 1000 Genomes (Genomes Project et al., 2015), National Heart Lung and Blood Institute (NHLBI) GO Exome Sequencing Project 6500 exomes (NHLBI-ESP), Exome Aggregation Consortium (ExAC), and Genome Aggregation Database (gnomAD) (Karczewski et al., 2019).

2.5 | Specimen quality control

Reported child and parent sex and relatedness were verified using Peddy v0.4.2 (Pedersen & Quinlan, 2017). Sample quality was assessed with FastQC v0.11.2 (FastQC, Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), VerifyBamID's freemix score (Jun et al., 2012), and an internal quality control pipeline (Jun et al., 2012). All specimens had sufficient coverage, with >75% of targets covered at 20X.

2.6 | Filtering and prioritization

The genotyping pipeline identified 268,346 unique variants in the study sample, with 224,142 (83.5%) of the variants deemed high quality (not in a GATK tranche, depth >10, genotype quality [GQ] >50). Of these 224,142 high quality variants, 38,129 affected an amino acid (missense or nonsense variants) and were deemed putatively functional. Genes in which loss-of-function variants were found in more than one child with SA and those in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

which loss-of-function and missense variants were found in more than two children were selected for further review. We further prioritized genes based on their gene damage index (GDI) (Itan et al., 2015), which is intended to estimate how well a gene tolerates genetic alterations based on the frequency and expected severity of cumulative missense variants in a population. Genes with lower GDI show less evidence of such genetic variation, indicating genetic alterations may be more likely to have severe consequences. Of the 38,129 variants, 530 met the primary filtering criteria of maximum AF ≤0.0001 and in a gene with a GDI <50%. Variants were also prioritized using secondary filtering criteria, which were the same as the primary filtering criteria but less restrictive, with regard to rarity in the control population (AF ≤0.001) and GDI (<75th percentile); 2,099 variants met the secondary filtering criteria. Furthermore, all putatively functional (missense or loss-of-function) rare variants (AF ≤0.001) in genes previously associated with SA 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 are summarized in Supplemental Table 1.

The GATK genotype refinement pipeline (Retrieved from https://gatkforums.broadinstitute.org/gatk/discussion/4723/genotype-refinement-workflow) was used to incorporate pedigree and population priors into the GQ scores and to flag potential *de novo* variants. The d*e novo* variants flagged by GATK were further filtered to have GQ >40 and minimum total depth greater than eight in all members of the trio, a minimum alternate allele depth of ≥five in the child, and allele frequency <0.001 in external databases. *De novo* variants passing all filters were manually reviewed with IGV.

2.7 | 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 a) were in genes that met the primary filtering criteria; b) had quality concerns following manual review in IGV and were either *de novo* or in genes previously reported in the literature to be associated with SA; or c) were compound heterozygous and the inheritance could not be fully determined from the sequencing data. 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 (Kent, 2002).

3 | RESULTS

Among the 30 liveborn children with non-syndromic SA included in the study, 17 (56.7%) were female, 21 (70%) were full term, and 29 (96.7%) were singleton births. Among mothers, age at delivery most often was in the 20–29 years range (70%), and the majority were non-Hispanic White (73.3%) (Table 1). Other birth defects diagnosed among the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

children with SA are presented in Table 2. Four cases were classified as isolated birth defects and 26 as multiple.

Exome sequencing was performed on eight trios with a maternal diagnosis of pregestational diabetes and 20 trios and two child-father duos without a maternal diagnosis of pregestational diabetes. The median number of on-target reads per sample was 62 million (66% of total reads overlapping a target, 76% of reads within 126 base pairs of a target), resulting in a median target coverage of 52X. All specimens had sufficient target coverage at 20X (median percent of targets covered at >20X for all specimens was 93%, range 74% to 98%) and did not show signs of contamination with other human DNA samples (maximum freemix was 0.04).

Three children with SA, all non-Hispanic White, had rare heterozygous variants in *ID1* (Inhibitor of DNA Binding 1). Two of the children, both of non-diabetic mothers, inherited the variant from their fathers, and the third child inherited the variant from a diabetic mother. *ID1* was the only gene that met the primary filtering criteria of ≥3 high quality functional variants that were very rare (AF<0.0001) and were in a gene with a GDI<50th percentile (Table 3). The three children with SA harbored different missense variants (p.S36A, p.S39G, and p.P54L). Two of these variants (p.S36A and p.P54L) are present in the gnomAD database, with allele frequencies of 0.000027 and 0.000032, respectively, while p.S39G is not. One, p.P54L, lies within the Myc-type, basic helix-loop-helix (bHLH) domain; thus, it might disrupt the DNA-binding ability of ID1. None of these variants are present in ClinVar. All three variants are predicted by CADD to be within the top 1% of deleterious variants in the genome. The phenotypes among these three children differed. One had isolated SA and the other two were classified as multiple, with additional defects that included cardiac and genitourinary defects (one child) and gastrointestinal and central nervous system defects (one child).

A total of 45 rare, inherited compound heterozygous variants were detected across 22 genes that met our filtering criteria (Table 4, Supplemental Table 1). Additionally, investigation of variants inherited in autosomal recessive and X-linked recessive models (Table 5) identified missense variants in seven genes (one variant in each gene). One of the variants identified in *BRS3* (Bombesin Receptor Subtype 3) is involved in the regulation of glucose metabolism (Feng et al., 2011; Gonzalez, Moreno, & Jensen, 2015). There were no instances that the same gene had variants meeting filtering criteria in more than one child.

Next, variants in 34 genes previously reported to be associated with SA in the literature were investigated as a stand-alone hypothesis (Supplemental Table 2). Heterozygous missense variants were identified in three children with SA: p.R2992Q and p.S583L in *SPTBN5* (Spectrin Beta, Non-erythrocytic 5*)* and p.P1033R in *PDZD2* (PDZ Domain Containing 2). However, the GDI scores are in the 99th percentile for *SPTBN5* and the 85th percentile for *PDZD2*, which suggests that these missense variants may be well tolerated.

In addition, *de novo* variants were investigated, revealing 11 *de novo* missense and one *de novo* loss-of-function variant in 11 different genes, but no variant was present in more than one child (Supplemental Table 3). Lastly, filtering criteria were relaxed to see if there were

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

any genes that met secondary filtering criteria (Supplemental Table 4). In total, across all genes examined, 24 of the 30 children with non-syndromic SA had variants identified in more than one gene.

Sanger sequencing was performed to validate seven variants in five genes, three in *ID1*, and one each in *ABLIM2* (Actin Binding LIM Protein Family Member 2), *PTF1A* (Pancreas Associated Transcription Factor 1a), *HOXA13* (Homeobox A13) and *DYNC1H1* (Dynein Cytoplasmic 1 Heavy Chain 1). These variants were selected because they met the primary filtering criteria (*ID1*) or they had quality concerns in manual review (*de novo* and previously reported genes-*PTF1A, HOXA13, DYNC1H1*) or, in the case of compound heterozygotes, their inheritance could not be determined entirely from the sequencing data (*ABLIM2*). The variants identified for *ID1* that met the primary filtering criteria (Table 3) were confirmed in all six individuals (three children, two fathers, one mother). For the compound heterozygous event in *ABLIM2* (Table 4), one variant was confirmed in the father, and Sanger sequencing confirmed that the second allele seen in the child was maternally inherited (the mother only had sufficient DNA for Sanger sequencing but not whole exome sequencing). The variants identified in *PTF1A* and *HOXA13* (Supplemental Table 2) that met the filtering criteria for known genes did not validate. The single *de novo* variant in *DYNC1H1* (Supplemental Table 3) was confirmed to be present in the child and absent in the parents.

4 | DISCUSSION

This study identified several rare gene variants in children with non-syndromic SA, including three variants in *ID1,* a gene that has not been previously associated with SA. *ID1* is a member of a four-protein family (*ID1-ID4*) of helix-loop-helix (HLH) transcriptional regulatory proteins (Billestrup, 2011; Norton, 2000) that play a role in cell growth and differentiation (OMIM#600349). These genes participate in crucial developmental processes, including neurogenesis, myogenesis, and sex determination (Jones, 2004), as well as cardiogenesis (Cunningham et al., 2017; Fraidenraich & Benezra, 2006). In mice, *Id1* has been shown to be related to calvarial but not skeletal bone development (Maeda, Tsuji, Nifuji, & Noda, 2004) and to act as a negative regulator of insulin secretion, possibly contributing to beta-cell dysfunction and glucose intolerance in type 2 diabetes (Akerfeldt & Laybutt, 2011). Given the strong association between pregestational diabetes and SA, it is noteworthy that other HLH factors, neurogenin3 and neuroD, are critical in beta-cell development and function and, reportedly, a mutation in either of these two genes can lead to diabetes (Jorgensen et al., 2007), providing biological plausibility for the observed association between *ID1* and SA.

We investigated how likely *ID1* is to tolerate damaging variation. The number of amino acids coded for by *ID1* is 155, which is quite a bit fewer than 375, the median number for homo sapiens (Karczewski et al., 2019). Based on size alone, smaller genes would be less likely to be mutated than larger genes. The GDI percentile of *ID1* is 47.0% suggesting that nearly half of all human genes are less tolerant to damaging variation than *ID1*. The gnomAD missense observed/expected (oe) constraint score of 1.59 (1.39 – 1.81) (Karczewski et al., 2019) indicates that, given the properties of *ID1* such as size, more

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

missense variants were found in *ID1* than expected in the gnomAD database. Therefore, it is possible that the three children with missense variants found in our cohort may be by chance, and functional follow up would be needed to help establish a causal relationship.

Two of the three children with an *ID1* variant in the current study inherited the variant from their fathers, and the third inherited the variant from a diabetic mother. These findings suggest that *ID1* may act by two mechanisms: indirectly by a diabetogenic effect on the mother in conjunction with poor glycemic control or directly to disrupt embryonic development when inherited from the father. If associated with SA, these *ID1* variants may have reduced penetrance or interact with other inherited or non-inherited factors, because each was inherited from a parent without SA, and two variants were present in seven heterozygous individuals in gnomAD who were reported to have SA.

One small pilot study of SA performed whole exome sequencing and copy number variant analysis using DNA from four child-parent trios; one of the children had a diabetic mother (Porsch et al., 2016). Of the 14 genes reported in the pilot study, our study identified different variants in only two of the genes, *PDZD2* and *SPTBN5*. *PDZD2* is considered to be a diabetes-related gene, as it is involved in fetal pancreatic progenitor cell proliferation. The pilot whole exome sequencing study identified a homozygous variant in *PDZD2* (Porsch et al., 2016), whereas the variant identified in our study was heterozygous. The pilot study also identified a single *de novo* mutation (p.E25K, predicted to be benign) in *SPTBN5* (Porsch et al., 2016), whereas variants identified in *SPTBN5* in our study were inherited. No other study has linked *SPTBN*5 with any disease. Even though variants in these genes were found in non-syndromic children with SA, they are of unclear significance due to their high GDIs (99.7% and 84.9%, respectively). Additionally, different *CYP26A1* variants than those identified in our study have been reported in a human study of SA but were not considered to be a major risk factor (De Marco et al., 2006).

Several mutated genes in mouse studies, including *Acd* (Keegan et al., 2005), *Pcsk5* (Szumska et al., 2008), *Wnt-3a* (Greco et al., 1996), *Cdx2* (Savory et al., 2009), *Cdx4*, *Hoxc13*, and *Hoxb13* (Young et al., 2009) have been shown to cause a phenotype similar to SA; however, variants in these genes have not been identified in humans. Our study did not find any variants in genes reported in previous mouse studies (genes are reviewed in Supplemental Table 2); however, we identified a variant in the *PTF1A* gene, which has been linked to phenotypes similar to SA in mouse studies (Lugani et al., 2013; Semba et al., 2013; Vlangos et al., 2013).

De novo mutations are good candidates for genetic etiology of SA. We examined the published literature for the 12 *de novo* variants that we identified (Supplementary Table 3) to determine whether any of the corresponding genes had reported functions related to the pathogenesis of SA. Among the 12 genes, four had reported effects on embryonic development, but none of these effects were related to vertebral or spinal cord embryogenesis.

Strengths of this study include the NBDPS study design that required that children with non-syndromic SA be ascertained from population-based surveillance programs and that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

each child was clinically well-characterized, with review of medical record data by clinical geneticists. It also expands the number of non-syndromic SA cases investigated to date. The use of the child-parent trio design for exome sequencing allowed for distinguishing between *de novo* and inherited variants. Limitations of the current study include the modest sample size and that only a subset of eligible children provided DNA. Also, the study sample had limited diversity in race/ethnicity. Additionally, exome rather than genome sequencing was performed, precluding studying non-coding regions. The heterogeneity of the phenotypes in the children may have complicated the effort to identify genes. The heterozygous variants found in *ID1* were inherited from unaffected parents; therefore, these variants may have incomplete penetrance, which could not be demonstrated. Additional study samples were not screened for the variants detected because the variants were rare, and even if causal, the likelihood of detecting the variants in additional populations was low. Lastly, functional studies were not conducted because none of the variants were recurrent. Regardless of these limitations, the current study offers data and insights into mechanisms underlying the genetic etiology of non-syndromic SA in humans that could contribute to future investigations.

In conclusion, exome sequencing of 28 child-parent trios and two child-father duos identified three children with non-syndromic SA, each inheriting a different rare variant in the *ID1* gene. Additional rare variants identified using the established primary and secondary filtering criteria were not identified in more than one child in the study sample. These rare variants may also contribute to SA etiology, but very large sample sizes would be required to replicate these associations. Many children had variants identified in more than one gene. Together, these findings suggest no single, common cause of non-syndromic SA in protein-coding regions using our filtering criteria. Future sequencing studies for non-syndromic SA require collaboration among multiple study populations, including populations with increased racial/ethnic diversity, to examine both rare variants reported here and additional variants. Data from the current study can provide a starting point for this collaboration.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the National Birth Defects Prevention Study participants, scientists, staff, and Genetics Collaborative Working Group. We also thank the Genome Aggregation Database 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 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

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Study and/or the Birth Defects Study To Evaluate Pregnancy exposureS, and the Iowa Center for Birth Defects Research and Prevention U01 DD001035 and U01 DD001223. The University of Washington Center for Mendelian Genomics was funded by the National Human Genome Research Institute and the National Heart, Lung and Blood Institute grants UM1 HG006493 and U24 HG008956.

REFERENCES

Akerfeldt MC, & Laybutt DR (2011). Inhibition of Id1 augments insulin secretion and protects against high-fat diet-induced glucose intolerance. Diabetes, 60(10), 2506–2514. doi:10.2337/db11-0083 [PubMed: 21940780]

Aslan H, Yanik H, Celikaslan N, Yildirim G, & Ceylan Y. (2001). Prenatal diagnosis of Caudal Regression Syndrome : a case report. BMC Pregnancy Childbirth, 1(1), 8. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11782287 [PubMed: 11782287]

Belloni E, Martucciello G, Verderio D, Ponti E, Seri M, Jasonni V, . . . Scherer SW (2000). Involvement of the HLXB9 homeobox gene in Currarino syndrome. Am J Hum Genet, 66(1), 312–319. doi:10.1086/302723 [PubMed: 10631160]

Billestrup N. (2011). ID'ing a novel inhibitor of beta-cell function, Id1. Diabetes, 60(10), 2455–2456. doi:10.2337/db11-1084 [PubMed: 21949000]

Boulas MM (2009). Recognition of caudal regression syndrome. Adv Neonatal Care, 9(2), 61–69; quiz 70–61. doi:10.1097/ANC.0b013e31819de44f [PubMed: 19363325]

Caird MS, Hall JM, Bloom DA, Park JM, & Farley FA (2007). Outcome study of children, adolescents, and adults with sacral agenesis. J Pediatr Orthop, 27(6), 682–685. doi:10.1097/01.bpo.0000190841.69412.eb [PubMed: 17717471]

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]

Cretolle C, Zerah M, Jaubert F, Sarnacki S, Revillon Y, Lyonnet S, & Nihoul-Fekete C. (2006). New clinical and therapeutic perspectives in Currarino syndrome (study of 29 cases). J Pediatr Surg, 41(1), 126–131; discussion 126–131. doi:10.1016/j.jpedsurg.2005.10.053 [PubMed: 16410121]

Cunningham TJ, Yu MS, McKeithan WL, Spiering S, Carrette F, Huang CT, . . . Colas AR, (2017). Id genes are essential for early heart formation. Genes Dev, 31(13), 1325–1338. doi:10.1101/gad.300400.117 [PubMed: 28794185]

De Marco P, Merello E, Mascelli S, Raso A, Santamaria A, Ottaviano C, . . . Capra V. (2006). Mutational screening of the CYP26A1 gene in patients with caudal regression syndrome. Birth Defects Res A Clin Mol Teratol, 76(2), 86–95. doi:10.1002/bdra.20225 [PubMed: 16463413]

Feng Y, Guan XM, Li J, Metzger JM, Zhu Y, Juhl K, . . . Zhou YP (2011). Bombesin receptor subtype-3 (BRS-3) regulates glucose-stimulated insulin secretion in pancreatic islets across multiple species. Endocrinology, 152(11), 4106–4115. doi:10.1210/en.2011-1440 [PubMed: 21878513]

Fraidenraich D, & Benezra R. (2006). Embryonic stem cells prevent developmental cardiac defects in mice. Nat Clin Pract Cardiovasc Med, 3 Suppl 1, S14–17. doi:10.1038/ncpcardio0402 [PubMed: 16501623]

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]

Gonzalez N, Moreno P, & Jensen RT (2015). Bombesin receptor subtype 3 as a potential target for obesity and diabetes. Expert Opin Ther Targets, 19(9), 1153–1170. doi:10.1517/14728222.2015.1056154 [PubMed: 26066663]

Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, & Camper SA (1996). Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. Genes Dev, 10(3), 313–324. doi:10.1101/gad.10.3.313 [PubMed: 8595882]

Hagan DM, Ross AJ, Strachan T, Lynch SA, Ruiz-Perez V, Wang YM, . . . Homfray T. (2000). Mutation analysis and embryonic expression of the HLXB9 Currarino syndrome gene. Am J

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Hum Genet, 66(5), 1504–1515. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10749657 [PubMed: 10749657]

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]

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

Jones S. (2004). An overview of the basic helix-loop-helix proteins. Genome Biol, 5(6), 226. doi:10.1186/gb-2004-5-6-226 [PubMed: 15186484]

Jorgensen MC, Ahnfelt-Ronne J, Hald J, Madsen OD, Serup P, & Hecksher-Sorensen J. (2007). An illustrated review of early pancreas development in the mouse. Endocr Rev, 28(6), 685–705. doi:10.1210/er.2007-0016 [PubMed: 17881611]

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]

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

Keegan CE, Hutz JE, Else T, Adamska M, Shah SP, Kent AE, . . . Hammer GD (2005). Urogenital and caudal dysgenesis in adrenocortical dysplasia (acd) mice is caused by a splicing mutation in a novel telomeric regulator. Hum Mol Genet, 14(1), 113–123. doi:10.1093/hmg/ddi011 [PubMed: 15537664]

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

Kibar Z, Torban E, McDearmid JR, Reynolds A, Berghout J, Mathieu M, . . . Gros P. (2007). Mutations in VANGL1 associated with neural-tube defects. N Engl J Med, 356(14), 1432–1437. doi:10.1056/NEJMoa060651 [PubMed: 17409324]

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]

Lugani F, Arora R, Papeta N, Patel A, Zheng Z, Sterken R, . . . Gharavi AG (2013). A retrotransposon insertion in the 5' regulatory domain of Ptf1a results in ectopic gene expression and multiple congenital defects in Danforth's short tail mouse. PLoS Genet, 9(2), e1003206. doi:10.1371/journal.pgen.1003206

Lynch SA, Wang Y, Strachan T, Burn J, & Lindsay S. (2000). Autosomal dominant sacral agenesis: Currarino syndrome. J Med Genet, 37(8), 561–566. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10922380 [PubMed: 10922380]

Maeda Y, Tsuji K, Nifuji A, & Noda M. (2004). Inhibitory helix-loop-helix transcription factors Id1/Id3 promote bone formation in vivo. J Cell Biochem, 93(2), 337–344. doi:10.1002/jcb.20154 [PubMed: 15368360]

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]

Mills JL (1982). Malformations in infants of diabetic mothers. Teratology, 25(3), 385–394. doi:10.1002/tera.1420250316 [PubMed: 7051398]

Mills JL (2010). Malformations in infants of diabetic mothers. Teratology 25:385–94. 1982. Birth Defects Res A Clin Mol Teratol, 88(10), 769–778. doi:10.1002/bdra.20757

Mills JL, Baker L, & Goldman AS (1979). Malformations in infants of diabetic mothers occur before the seventh gestational week. Implications for treatment. Diabetes, 28(4), 292–293. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/437367 [PubMed: 437367]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Nalbandyan M, Howley MM, Cunniff CM, Romitti PA, Browne ML, & National Birth Defects Prevention, S. (2019). Descriptive and risk factor analysis of nonsyndromic sacral agenesis: National Birth Defects Prevention Study, 1997–2011. Am J Med Genet A, 179(9), 1799–1814. doi:10.1002/ajmg.a.61290 [PubMed: 31294918]

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

Norton JD (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. J Cell Sci, 113 (Pt 22), 3897–3905. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11058077 [PubMed: 11058077]

Nowaczyk MJ, Huggins MJ, Tomkins DJ, Rossi E, Ramsay JA, Woulfe J, . . . Belloni E. (2000). Holoprosencephaly, sacral anomalies, and situs ambiguus in an infant with partial monosomy 7q/trisomy 2p and SHH and HLXB9 haploinsufficiency. Clin Genet, 57(5), 388–393. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10852374 [PubMed: 10852374]

Pang D. (1993). Sacral agenesis and caudal spinal cord malformations. Neurosurgery, 32(5), 755–778; discussion 778–759. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8492851 [PubMed: 8492851]

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]

Porsch RM, Merello E, De Marco P, Cheng G, Rodriguez L, So M, . . . Campbell DD (2016). Sacral agenesis: a pilot whole exome sequencing and copy number study. BMC Med Genet, 17(1), 98. doi:10.1186/s12881-016-0359-2 [PubMed: 28007035]

Postma AV, Alders M, Sylva M, Bilardo CM, Pajkrt E, van Rijn RR, . . . van Maarle MC (2014). Mutations in the T (brachyury) gene cause a novel syndrome consisting of sacral agenesis, abnormal ossification of the vertebral bodies and a persistent notochordal canal. J Med Genet, 51(2), 90–97. doi:10.1136/jmedgenet-2013-102001 [PubMed: 24253444]

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]

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]

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]

Ross AJ, Ruiz-Perez V, Wang Y, Hagan DM, Scherer S, Lynch SA, . . . Strachan T. (1998). A homeobox gene, HLXB9, is the major locus for dominantly inherited sacral agenesis. Nat Genet, 20(4), 358–361. doi:10.1038/3828 [PubMed: 9843207]

Sarnat HB, Case ME, & Graviss R. (1976). Sacral agenesis. Neurologic and neuropathologic features. Neurology, 26(12), 1124–1129. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1033479 [PubMed: 1033479]

Savory JG, Bouchard N, Pierre V, Rijli FM, De Repentigny Y, Kothary R, & Lohnes D. (2009). Cdx2 regulation of posterior development through non-Hox targets. Development, 136(24), 4099–4110. doi:10.1242/dev.041582 [PubMed: 19906845]

Semba K, Araki K, Matsumoto K, Suda H, Ando T, Sei A, . . . Yamamura K. (2013). Ectopic expression of Ptf1a induces spinal defects, urogenital defects, and anorectal malformations in Danforth's short tail mice. PLoS Genet, 9(2), e1003204. doi:10.1371/journal.pgen.1003204

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]

Sonek JD, Gabbe SG, Landon MB, Stempel LE, Foley MR, & Shubert-Moell K. (1990). Antenatal diagnosis of sacral agenesis syndrome in a pregnancy complicated by diabetes mellitus. Am J Obstet Gynecol, 162(3), 806–808. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2180307 [PubMed: 2180307]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Szumska D, Pieles G, Essalmani R, Bilski M, Mesnard D, Kaur K, . . . Bhattacharya S. (2008). VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase Pcsk5. Genes Dev, 22(11), 1465–1477. doi:10.1101/gad.479408 [PubMed: 18519639]

Tinker SC, Gilboa SM, Moore CA, Waller DK, Simeone RM, Kim SY, . . . National Birth Defects Prevention, S. (2020). Specific birth defects in pregnancies of women with diabetes: National Birth Defects Prevention Study, 1997–2011. Am J Obstet Gynecol, 222(2), 176 e171–176 e111. doi:10.1016/j.ajog.2019.08.028

Vlangos CN, Siuniak AN, Robinson D, Chinnaiyan AM, Lyons RH Jr., Cavalcoli JD, & Keegan CE (2013). Next-generation sequencing identifies the Danforth's short tail mouse mutation as a retrotransposon insertion affecting Ptf1a expression. PLoS Genet, 9(2), e1003205. doi:10.1371/journal.pgen.1003205

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]

Wender-Ozegowska E, Wroblewska K, Zawiejska A, Pietryga M, Szczapa J, & Biczysko R. (2005). Threshold values of maternal blood glucose in early diabetic pregnancy--prediction of fetal malformations. Acta Obstet Gynecol Scand, 84(1), 17–25. doi:10.1111/j.0001-6349.2005.00606.x [PubMed: 15603562]

Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, . . . Deschamps J. (2009). Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. Dev Cell, 17(4), 516–526. doi:10.1016/j.devcel.2009.08.010 [PubMed: 19853565]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Selected characteristics of children with sacral agenesis and birth mothers. National Birth Defects Prevention Study, 1997–2011.

Characteristic	N (%)
Child	
Sex	
Male	13 (43.3)
Female	17 (56.7)
Gestational age (weeks)	
Preterm (<37)	9 (30.0)
Term (37–45)	21 (70.0)
Plurality	
Multiple	1 (3.3)
Singleton	29 (96.7)
Mother	
Age at delivery (years)	
<20	3 (10.0)
20–29	21 (70.0)
30–39	5 (16.7)
≥40	1 (3.3)
Race/Ethnicity	
Non-Hispanic White	22 (73.3)
Non-Hispanic Black	1 (3.3)
Hispanic	7 (23.3)

Due to rounding, percentages may not total 100.0.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2.

Patterns of birth defects diagnosed in children with sacral agenesis. National Birth Defects Prevention Study, 1997–2011.

Case Description and Classification in NBDPS	Number of Children[a] (%)
Isolated [b]	4 (13.3)
Multiple [c,d]	26 (86.6)
Recognized phenotype[e]	8 (26.6)
Cardiac defects[f]	12 (40)
Craniofacial defects[g]	5 (16.6)
Gastrointestinal defects[h]	16 (53.3)
Genitourinary defects[i]	13 (43.3)
Limb/appendicular skeletal defects[j]	6 (20)
Distal axial skeletal defects[k]	10 (33.3)
Central nervous system defects[l]	6 (20)
Other musculoskeletal defects[m]	2 (6.6)
Respiratory defects[n]	1 (3.3)
Situs/heterotaxy defects[o]	2 (6.6)

[a] Cases can be counted in more than one category with exception of cases classified as isolated

[b] Isolated cases have sacral agenesis with/without additional axial skeletal defects and secondary defects (lower limb deformations or hypoplasia and/or tethered spinal cord)

[c] Multiple cases have sacral agenesis with/without additional axial skeletal defects and secondary defects (lower limb deformations or hypoplasia and/or tethered spinal cord) AND at least one additional defect considered unrelated to the primary defect, including appendicular skeletal defects

[d] Only defects considered to be major (i.e., have surgical, medical, or serious cosmetic importance) are included in the totals

[e] Includes 6 cases with VACTERL (vertebral defects, anal atresia, cardiac defects, tracheoesophageal fistula/ esophageal atresia, renal defects, and limb defects) and 2 cases with Goldenhar phenotypes

[f] Includes atrial, ventricular, and atrioventricular septal defects; aortic arch coarctation and hypoplasia; truncus arteriosus; pulmonary atresia; tetralogy of Fallot; and double outlet right ventricle

[g] Includes microtia, cleft palate, Pierre Robin anomaly, hemifacial microsomia, and macrostomia

[h] Includes imperforate anus, anal stenosis, tracheoesophageal fistula/esophageal atresia, malrotation, and cloacal variant

[i] Includes multicystic kidney, ureteropelvic junction obstruction, hydronephrosis, absent kidney, horseshoe kidney, crossed renal ectopia, duplicated collecting system, hypospadias, bicornuate/hypoplastic uterus, uterine didelphys, and absent vagina

[j] Includes radial aplasia, tibial aplasia, missing phalanges, and syndactyly

[k] Includes extra, missing and fused ribs, and cervical and thoracic vertebral defects (fusion, hemivertebrae, absent)

[l] Includes spinal cord syrinx, thin corpus callosum, cerebral hypoplasia, schizencephaly, thoracic meningomyelocele, lumbosacral lipomeningocele, intrathecal lipoma, hydrocephalus, and microphthalmia

[m] Includes congenital diaphragmatic hernia and craniosynostosis

[n] Includes tracheal stenosis and lung agenesis

[o] Includes situs ambiguus and complex situs inversus

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Variants in *ID1*, the gene that met primary filtering criteria[a] in infants with sacral agenesis. National Birth Defects Prevention Study, 1997–2011.

Child	Gene	Variant	Locus	[Allele1, Allele2]	GQ	AD	AF[b]	CADD	GDI
1	*ID1*	p.S36A[c]	20:31605493	[T, G]	99	[9, 18]	0.00004	24.1	47.0%
2	*ID1*	p.S39G[c]	20:31605502	[A, G]	99	[16, 19]	0.00002	25.9	47.0%
3	*ID1*	p.P54L[d]	20:31605548	[C, T]	99	[21, 16]	0.00004	29.0	47.0%

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

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

[b] Maximum allele frequency observed in any public database for any subpopulation.

[c] Confirmed in proband and father via Sanger sequencing.

[d] Confirmed in proband and mother via Sanger sequencing; child of diabetic mother

None of the variants were found in ClinVar.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 4.

Inherited compound heterozygous variants meeting filtering criteria[a], ordered by gene damage index, in infants with sacral agenesis. National Birth Defects Prevention Study, 1997–2011.

Child	Gene	Variant	Locus	[Allele1, Allele2]	GQ	AD	AF[b]	CADD	GDI
11	*SFXN5*	p.T117M	2:73020246	[G, A]	99	[21, 22]	0.00010	25.3	15.9%
11	*SFXN5*	p.A13del	2:73071667	[AGCC, A]	51	[11, 2]	0.0096	.	15.9%
14	*ABLIM2*	p.D491N[c]	4:8008104	[C, T]	99	[30, 32]	0.0012	19.5	33.9%
14	*ABLIM2*	p.V49I	4:8106503	[C, T]	99	[19, 19]	0.00040	11.5	33.9%
9	*SLC22A11*	p.A2V	11:64556004	[C, T]	99	[37, 35]	0.0036	24.1	36.2%
9	*SLC22A11*	p.V543I	11:64571016	[G, A]	99	[31, 33]	0.0010	0.2	36.2%
7	*TMEM174*	p.N14Y	5:73173283	[A, T]	99	[34, 22]	0.00020	25.4	40.2%
7	*TMEM174*	p.G116R	5:73173589	[G, A]	99	[60, 62]	0.0064	22.7	40.2%
7	*SP100*	c.1207+2T>A[d]	2:230467208	[T, A]	99	[21, 15]	0.0038	2.0	40.3%
7	*SP100*	p.K739N	2:230515217	[G, C]	99	[28, 20]	0.00031	0.5	40.3%
7	*SP100*	p.E868K[d]	2:230515602	[G, A]	99	[35, 40]	0.0034	10.0	40.3%
29	*TMEM67*	p.P721S[e]	8:93799678	[C, T]	99	[25, 30]	0.0055	0.0	43.5%
29	*TMEM67*	p.A953G	8:93815398	[C, G]	99	[29, 14]	0.0000077	19.9	43.5%
29	*COL1A1*	p.P1179S	17:50186919	[G, A]	91	[7, 4]	0.000061	13.4	44.4%
29	*COL1A1*	p.G906S[e]	17:50189490	[C, T]	99	[11, 12]	0.0018	23.4	44.4%
2	*ZNF433*	p.N597D	19:12014973	[T, C]	99	[49, 50]	0.0098	11.8	47.0%
2	*ZNF433*	p.R375X	19:12015639	[G, A]	99	[42, 45]	0.0012	34.0	47.0%
19	*PODXL2*	p.G362R	3:127661112	[G, A]	99	[26, 23]	0.0010	19.8	49.1%
19	*PODXL2*	p.Q493K	3:127671485	[C, A]	99	[13, 25]	0.0010	21.9	49.1%
1	*APAF1*	p.E777K	12:98699432	[G, A]	99	[20, 21]	0.0060	22.8	55.3%
1	*APAF1*	p.S871L[e]	12:98706501	[C, T]	99	[13, 11]	0.0033	17.6	55.3%
15	*TRAF3IP3*	p.E285V	1:209775428	[A, T]	99	[31, 25]	0.00010	24.3	55.4%
15	*TRAF3IP3*	p.I378T	1:209777431	[T, C]	99	[21, 30]	0.0010	10.4	55.4%
6	*NRG2*	p.R760K	5:139848191	[C, T]	99	[5, 5]	0.0058	25.7	60.2%
6	*NRG2*	p.L221F	5:140042409	[G, A]	99	[17, 21]	0.00080	24.1	60.2%
4	*APC*	p.S966N	5:112838491	[G, A]	99	[55, 55]	0	25.7	61.8%
4	*APC*	p.R2505Q[e]	5:112843108	[G, A]	99	[47, 43]	0.0015	31.0	61.8%
1	*SLC12A4*	p.L842V	16:67946254	[G, C]	99	[33, 30]	0.00010	14.0	61.9%
1	*SLC12A4*	p.M587T	16:67948148	[A, G]	99	[14, 11]	0.000018	26.7	61.9%
8	*FRMD4A*	p.L1012V	10:13654432	[G, C]	99	[5, 8]	0.000067	23.5	62.2%
8	*FRMD4A*	p.T732A	10:13657395	[T, C]	99	[9, 13]	0.00060	15.7	62.2%
25	*ENPEP*	R437H	4:110513416	[G, A]	99	[35, 28]	0.0091	20.3	62.8%
25	*ENPEP*	p.R788L	4:110549748	[G, T]	99	[22, 26]	0.0010	27.7	62.8%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Child	Gene	Variant	Locus	[Allele1, Allele2]	GQ	AD	AF[b]	CADD	GDI
25	*SRCAP*	p.P1130S[e]	16:30721323	[C, T]	99	[26, 25]	0.0060	19.8	63.4%
25	*SRCAP*	p.A2774S	16:30738360	[G, T]	99	[27, 19]	0.000067	12.1	63.4%
24	*TULP4*	p.M701V	6:158501764	[A, G]	99	[16, 10]	0.00030	21.9	67.6%
24	*TULP4*	p.P779L	6:158501999	[C, T]	99	[13, 13]	0.00030	23.2	67.6%
26	*EEF1D*	p.G303S	8:143589325	[C, T]	99	[9, 14]	0.000033	22.4	67.8%
26	*EEF1D*	p.D66E	8:143590034	[G, T]	99	[19, 15]	0.00020	23.1	67.8%
7	*ITGA8*	p.P811T	10:15575536	[G, T]	99	[33, 36]	0.00020	23.5	68.1%
7	*ITGA8*	p.N501S	10:15613711	[T, C]	99	[12, 14]	0.0015	22.9	68.1%
21	*USP34*	p.A1750T	2:61280252	[C, T]	99	[29, 21]	0.0031	23.8	70.1%
21	*USP34*	p.A1410V	2:61296825	[G, A]	99	[13, 11]	0.0099	32.0	70.1%
21	*CCDC27*	p.G117S	1:3754148	[G, A]	99	[9, 6]	0.0015	0.7	74.7%
21	*CCDC27*	p.L639fs	1:3771467	[CT, C]	99	[16, 14]	0.0027	.	74.7%

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

[a] AD>10, GQ>50, AF≤0.01 in public databases, GDI<75th percentile, variants were inherited from different parents.

[b] Maximum allele frequency observed in any public database for any subpopulation.

[c] Variant confirmed in proband and mother via Sanger sequencing.

[d] Both of these variants were inherited from the father.

[e] Variants are present in ClinVar but predicted to be benign or likely benign or of uncertain significance; sacral agenesis was not one of the phenotypes reported.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 5.

Variants in recessive or X-linked recessive genes meeting the filtering criteria[a], ordered by gene damage index, in infants with sacral agenesis. National Birth Defects Prevention Study, 1997–2011.

Child	Gene	Variant	Locus	[Allele1, Allele2]	GQ	AD	Mode of inheritance	AF[b]	CADD	GDI
15	*SLC6A16*	p.L278V	19:49309695	[C, C]	99	[0, 46]	AR	0.0074	3.0	31.7%
10	*ZNF276*	p.R351W	16:89727323	[T, T]	99	[0, 49]	AR	0.0097	28.7	52.1%
28	*CLPB*	p.S223N[c]	11:72372993	[T, T]	63	[0, 21]	AR	0.0078	14.4	61.2%
3	*MISP*	p.R556H	19:758613	[A, A]	99	[0, 44]	AR	0.0099	27.4	70.2%
3	*PHKB*	p.N166S[c]	16:47515525	[G, G]	84	[0, 28]	AR	0.0030	23.4	72.8%
30	*CHD6*	p.N14D	20:41533564	[C, C]	99	[0, 78]	AR	0.0020	10.2	74.5%
23	*BRS3*	p.S380P	X:136492313	[C, -][d]	81	[0, 27]	XLR	0.000038	13.0	9.5%
2	*TENM1*	p.S1227I	X:124497031	[A, -][d]	99	[0, 34]	XLR	0	26.0	69.3%

AD allelic depth for the [Allele1, Allele2], *AF* allele frequency, *AR* autosomal recessive, *CADD* combined annotation dependent depletion score, *GDI* gene damage index, *GQ* genotype quality, *XLR* X-linked recessive.

[a] AD >10, GQ >50, AF ≤0.01 for AR and AF ≤0.0001 for XLR, GDI <75th percentile.

[b] Maximum allele frequency observed in any public database for any subpopulation.

[c] Variants are present in ClinVar but predicted to be benign or likely benign or of uncertain significance; sacral agenesis was not one of the phenotypes reported.

[d] Both children have only one allele because they are male and the genes are on the X chromosome.