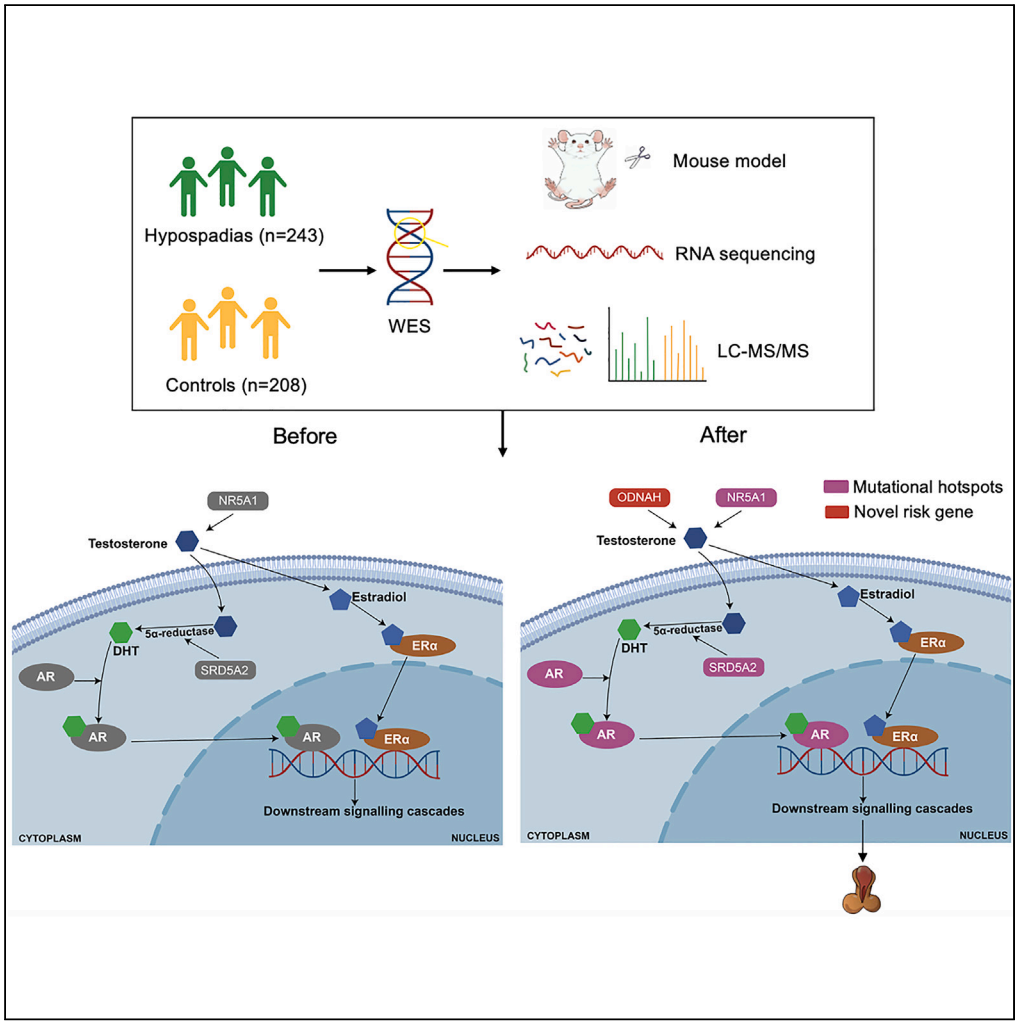
iScience

CellPress
OPEN ACCESS

Article

Whole-exome sequencing study of hypospadias



Zhongzhong Chen, Yunping Lei, Richard H. Finnell, ..., Yaping Wang, Hua Xie, Fang Chen

zhongzhongchen@gmail.com (Z.C.)
drxiehua@163.com (H.X.)
chenfang01@sjtu.edu.cn (F.C.)

Highlights

NR5A1, *SRD5A2*, and *AR* were identified as mutational hotspots in severe hypospadias

Mutations in outer dynein arm heavy chain genes associated with hypospadias

Ciliary genes may affect urethral development by influencing androgen levels

Chen et al., iScience 26, 106663
May 19, 2023
https://doi.org/10.1016/j.isci.2023.106663

iScience

CellPress
OPEN ACCESS

Article

Whole-exome sequencing study of hypospadias

Zhongzhong Chen,[1,2,7,8,*] Yunping Lei,[3,7] Richard H. Finnell,[3,4,7] Yu Ding,[1,7] Zhixi Su,[5] Yaping Wang,[1] Hua Xie,[1,*] and Fang Chen[1,6,*]

SUMMARY

Hypospadias results from the impaired urethral development, which is influenced by androgens, but its genetic etiology is still unknown. Through whole exome sequencing analysis, we identified *NR5A1*, *SRD5A2*, and *AR* as mutational hotspots in the etiology of severe hypospadias, as these genes are related to androgen signaling. Additionally, rare damaging variants in cilia-related outer dynein arm heavy chain (*ODNAH*) genes (*DNAH5*, *DNAH8*, *DNAH9*, *DNAH11*, and *DNAH17*) ($p = 8.5 \times 10^{-47}$) were significantly enriched in hypospadias cases. The *Dnah8* KO mice exhibited significantly decreased testosterone levels, which had an impact on urethral development and disrupted steroid biosynthesis. Combined with trios data, transcriptomic, and phenotypical and proteomic characterization of a mouse model, our work links ciliary genes with hypospadias. Overall, a panel of *ODNAH* genes with rare damaging variants was identified in 24% of hypospadias patients, providing significant insights into the underlying pathogenesis of hypospadias as well as genetic counseling.

INTRODUCTION

Hypospadias is one of the most common male congenital disorders. It occurs in approximately 2 per 1,000 pregnancies which included live births, stillbirths, and elective terminations of pregnancy.[1] The prevalence has been increasing during the period of 1980–2010.[1] The androgen-dependent signals have fundamental roles in masculinization and are crucial for the developmental process of the external genitalia,[2] which has lifelong effects on the urinary function, fertility, and sexual health.[3] The development of hypospadias is influenced by androgens, specifically testosterone (T), and dihydrotestosterone (DHT).[2] Testosterone is the major androgen[4] and can be converted to DHT. Even with the recent advances in explaining the genetic regulation of masculinization using genomic sequencing,[2] the genetic etiology of hypospadias remains far from fully understood.[5]

Hypospadias is considered to be a complex congenital disorder caused by multiple genetic and environmental interacting factors.[6] In studies of hypospadias in human cohorts, the majority of genetic investigations have either focused common variant analysis such as genome-wide association studies (GWASs) or narrowly focused on known key genes in hypospadias cohorts.[7,8] To date, hypospadias risk-associated variants have been reported in populations of European descent using GWASs, explaining approximately 9.4% of the observed genetic variance.[9,10] Recently, Kalfa and colleagues attempted to address the "missing heritability" in hypospadias using a targeted next-generation sequencing (NGS) panel including 336 genes.[11] While their findings are informative, the "missing heritability" question was not completely addressed.[5] Previous investigations demonstrated that missense and loss-of-function (LoF) variants are more likely than other variant types to adversely affect protein function; therefore, they are considered more likely to be causative than other types of genetic variants.[12] Rare damaging variants (LoF and deleterious missense variants), which dramatically alter protein sequences, may explain the unaccounted "missing heritability" not accounted for by GWAS analyses, and have been effectively used in identifying causal genes in birth defect diseases, such as neural tube defects.[13,14] These data suggest that the rare damaging variants may contribute to the etiology of hypospadias in humans; however, there is a paucity of such data that has been published on human hypospadias to date. A more extensive unbiased genetic approach supported by strong statistical evidence is needed to better understand the genetic etiology of hypospadias.[5]

Unlike most previous hypospadias studies that focused primarily on common variant associations using GWASs[9,10] or candidate genes resequencing using NGS method,[11] whole-exome sequencing

[1]Department of Urology, Shanghai Children's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200062, China

[2]Urogenital Development Research Center, Shanghai Children's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200062, China

[3]Center for Precision Environmental Health, Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

[4]Departments of Molecular and Human Genetics and Medicine, Baylor College of Medicine, Houston, TX 77030, USA

[5]School of Life Sciences, Fudan University, Shanghai 200438, China

[6]Clinical Research Center For Hypospadias Pediatric College, Shanghai Jiao Tong University School of Medicine, Shanghai 200062, China

[7]These authors contributed equally

[8]Lead contact

*Correspondence: zhongzhongchen@gmail.com (Z.C.), drxiehua@163.com (H.X.), chenfang01@sjtu.edu.cn (F.C.)

https://doi.org/10.1016/j.isci.2023.106663

(WES) or whole-genome sequencing (WGS) which include all coding genes may impact the number of detectable variants and is more likely to study the "missing heritability" of hypospadias, especially for the discovery of novel causative genes. In this study, we performed WES and RNA sequencing in a large hypospadias cohort. Mice models and label-free quantitative proteomics were used to evaluate the genetic effects of novel identified genes to hypospadias. To date, no published studies had applied WES to discern the genetic landscape of a large cohort of hypospadias patients, and this approach was undertaken in the belief that it would provide significant insights into the genetic etiology of hypospadias.

RESULTS

Mutations in AR, NR5A1, and SRD5A2 genes function as mutational hotspots in severe hypospadias

We performed WES on 191 patients with severe hypospadias (Figure 1A) (Table S1) with a mean depth of 142X and achieved a 99.9% breadth of coverage of target nucleotides (Figure S1). Population sequencing efforts have identified that each hypospadias participant contains on average approximately 123 rare damaging variants (Figure S1), which is close to the number of the protein-truncating variants reported in the UK biobank array.[15] Our initial analysis focused on comparing the burden of rare damaging variants in 191 Han Chinese individuals with severe hypospadias to 208 normal Han Chinese controls in 336 candidate gene set from study by the Kalfa team[11] (Figure 1A). The results demonstrated that rare damaging variants in six genes *CNTNAP3*, *HLA-DRB1*, *NR5A1*, *SRD5A2*, *DMD*, and *AR* were significantly increased in hypospadias cases as compared to controls ($p < 0.01$) (Figure 1B). However, no gene was associated with hypospadias after multiple-test correction ($q > 0.05$). We further investigated the burden of rare damaging variants in 30 previously reported known hypospadias risk-associated gene set.[8,16] We identified that 27.2% (52/191) of hypospadias patients carried at least one rare damaging variant in 21 genes, including: *NR5A1*, *SRD5A2*, *AR*, *GLI2*, *HSD3B2*, *GLI3*, *AKR1C3*, *ESR2*, *FGFR2*, *HSD3B1*, *MAP3K1*, *BMP4*, *BMP7*, *CHD7*, *DGKK*, *ESR1*, *FGF8*, *HOXA4*, *SHH*, *SP7*, *WT1* (Figure 1C). All 21 previously reported hypospadias risk-associated genes were included in burden analysis. Results showed that mutations in *AR*, *NR5A1*, and *SRD5A2* were significantly increased in severe hypospadias compared with controls ($p < 0.01$, binomial test) (Figure 1C)(Table 1), indicating that they may function as mutational hotspots in the etiology of severe hypospadias. Furthermore, rare damaging mutations of *NR5A1*, *SRD5A2*, and *AR* genes occurs in approximately 14% of all severe hypospadias cases, whereas only one rare damaging mutation was found in these three genes in 208 Chinese controls. Additionally, all of the variants were classified as pathogenic or likely pathogenic (Table 2) according to the ACMG/AMP guidelines.

Significant excess of mutations were identified in ODNAH genes in severe hypospadias

The omnigenic theory proposed that genetic mutations near genes that are expressed in relevant tissues contribute substantially to heritability.[17] Given that the development of hypospadias is affected by androgen signaling, and the major androgen is testosterone, which is produced by testes, we proposed that defects of genes specifically expressed in testis might be closely associated with hypospadias.[5] To better understand the genetic contribution of genes specifically and highly expressed in the testis tissue, we carried out genetic burden tests to investigate the role of rare damaging variants in 950 testis-specific highly expressed (tissue enriched) genes[18] and their orthologs for further exploring the risk of severe hypospadias. Significant excess of rare damaging variants was identified in 9 genes (*IQCM*, *DNAH8*, SPATA31A6, *CCDC168*, *DNAH17*, SPATA31A3, *GOLGA8T*, *WDR87*, and *DNHD1*) ($p < 0.01$) (Figure S2). Although no gene passed the Bonferroni multiple-test correction, enrichment of rare damaging variants across two testis-specific genes (*DNAH8* and *DNAH17*) were initially identified from outer dynein arm heavy chain (*ODNAH*) gene. It suggests that *ODNAH* gene family may potentially contribute to the genetic etiology of hypospadias. We then showed that the *ODNAH* gene set (*DNAH5*, *DNAH8*, *DNAH9*, *DNAH11*, and *DNAH17*) demonstrated a significant increase in the number of rare damaging variants after the Bonferroni multiple-test correction. Twenty-seven rare damaging variants in *DNAH8*, *DNAH9*, and *DNAH17* were selected in the available severe hypospadias cases, of which 96.4% (26/27) were confirmed by Sanger sequencing (Figure 2A). All 26 variants were identified as heterozygous (Figure 2A) and very rare or novel from gnomAD database (Table 3). After removing the false positive mutation, a significant excess of rare damaging variants was identified in *ODNAH* genes in patients with severe hypospadias compared to 208 Han Chinese controls ($p = 4.8 \times 10^{-17}$, binomial test) (Table S2) (Figure 1D). When compared to 2,504 controls from 1KGP, we found a higher burden of rare damaging variants in *ODNAH* genes (odds

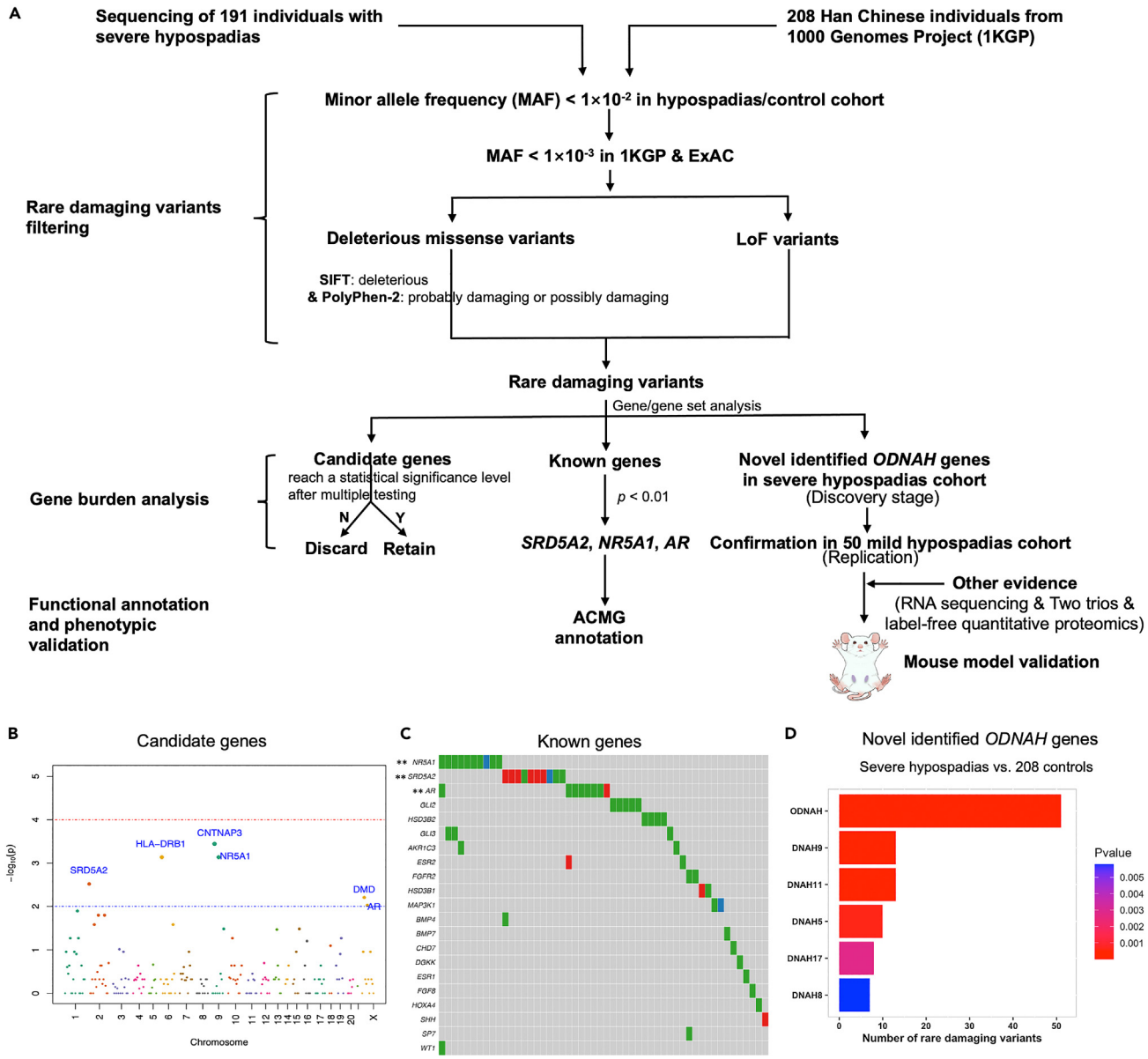


Figure 1. High-impact rare genetic variants in severe hypospadias

(A) Analytic workflow of rare damaging variants in candidate genes, known genes and novel causative genes in hypospadias.

(B) Manhattan plot of genetic burden in 336 known and candidate genes[11] in severe hypospadias. The horizontal dashed red and blue line indicates the preset threshold of $p = 1 \times 10^{-4}$ and $p = 1 \times 10^{-2}$, separately.

(C) Rare damaging variants in known risk-associated genes in patients with severe hypospadias. The asterisks indicate genes with significant excess of rare damaging variants (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

(D) Rare damaging variants were enriched in *ODNAH* genes for 191 individuals with severe hypospadias compared to 208 Han Chinese controls from 1000 Genomes Project (1KGP). Significance of association was displayed with p values.

ratio, 22.2, 95% CI, 13.5 to 37.0; $p = 1.2 \times 10^{-33}$) (Table S2). Among all 51 detected *ODNAH* rare damaging variants, 19.6% (10/51) of the variants were located in the dynein heavy domain (Figure 2B). Apart from *ODNAH* genes, the genetic burden in the inner dynein arm heavy chain (*IDNAH*) genes from the *DNAH* family (TF316836) was further evaluated (Table S3). Only *DNAH1* and *DNAH2* showed significant enrichments of rare damaging variants ($p < 0.01$, binomial test). Overall, rare damaging variants of the *ODNAH* genes identified in 22.5% of (43/191) sporadic severe hypospadias cases, are believed to contribute to the genetic etiology of hypospadias.

Table 1. Burden analysis of rare damaging variants in *SRD5A2*, *NR5A1* and *AR* by severe hypospadias vs. control comparison

	Number of rare damaging variants[a]		
Genes	Severe hypospadias (N = 191)	Han Chinese controls (N = 208)	P^b
SRD5A2	10	0	6.3×10^{-4}
NR5A1	10	0	6.3×10^{-4}
AR	9	1	9.0×10^{-3}

[a]Rare damaging variants (LoF and D-mis) with MAF<0.1% in 1KGP and ExAC databases.
[b]*P*-value based on binomial test by severe hypospadias vs. control.

Excess of ODNAH variants were confirmed in patients with mild hypospadias

To further evaluate our initial observation of significant differences between severe and mild hypospadias cases, we performed a replication study using 50 mild hypospadias patients. Rare damaging variants of *ODNAH* genes were found in 30% (15/50) of samples from patients with mild hypospadias. Overall, a significant excess of rare damaging variants was found in *ODNAH* genes in mild hypospadias patients ($p = 2.0 \times 10^{-11}$, binomial test) and the combined result yielded statistical significance with $p = 1.7 \times 10^{-18}$ (binomial test) compared to 208 Han Chinese controls (Table S2). To strengthen the genetic evidence of *ODNAH* genes, we also investigated rare damaging variants in *ODNAH* genes in 104 Japanese people in Tokyo (JPT) from 1 KG. Rare damaging variants in *ODNAH* genes were identified in 1.9% (2/104) of JPT, significantly ($p = 3.2 \times 10^{-6}$, binomial test) less than rare damaging variants compared to 50 mild hypospadias patients (15/50). When compared 241 hypospadias cases to 2,504 controls from 1KGP, we found a higher burden of rare damaging variants in *ODNAH* genes (odds ratio, 22.8, 95% CI, 14.3 to 37.1; $p = 8.5 \times 10^{-47}$) (Figure 3A) (Table 4). There was no excess of rare damaging variants in severe hypospadias cases compared with mild hypospadias cases (odds ratio, 0.89, 95% CI, 0.45 to 1.85; p = 0.65) (Table S4) (Figure 3B).

Dysregulated expression and dynamic network in AR and ODNAH mutated hypospadias

To better understand how *AR* and *ODNAH* genes variants contribute to severe hypospadias, 12 hypospadiac prepuce samples and six non-hypospadiac prepuce samples (controls) were analyzed using RNA-seq. The gene expression of *AR* in *AR*-mutated cases, *DNAH8* in *DNAH8*-mutated cases and *DNAH17* in *DNAH17*-mutated cases was significantly differentially expressed from the controls (Figure 4A). Unsupervized hierarchical cluster analysis revealed that 15 out of 30 previously reported hypospadias risk-associated genes were differentially expressed between hypospadias and control samples (Figure 4A). The network propagation analysis indicated that rare damaging variants in *AR* may directly alter the expression of *GLI1*, *GLI2*, *ATF3*, and *CHD7*, and indirectly affect the expression of *DNAH17* through protein-protein interactions (PPIs) (Figure 4B). Compared to controls, *AR*, *DNAH17*, *ATF3*, *CHD7*, and the homeobox genes (*HOXA4* and *HOXB6*) were differentially expressed in patients with *DNAH8* mutations (Figure 4C), whereas 13 genes might be indirectly affected by mutations in *DNAH17* (Figure 4D). Additionally, *AR*, *ATF3*, *CHD7*, and *DNAH17*, which act as the crucial nodes with the highest connectivity in these three subnetworks as shown in the Venn diagram of Figure 4A, might be closely associated with the development of hypospadias.

Dnah8-knockdown mice decreased testosterone production, impact the urethral development and dysregulated steroid biosynthesis

Given the fact that rare damaging variants in *ODNAH* genes were significantly enriched in both severe and mild hypospadias, we further evaluated the *in vivo* functional and pathogenetic effect of *ODNAH* genes on testosterone level and hypospadias phenotype through investigating the Dnah8-knockout ($Dnah8^{-/-}$) mouse generated by CRISPR/Cas9 technology. We designed two sgRNAs targeting exon 4 on *Dnah8* to completely excise exon 4 (Figure 5A). Deletion of the 439 bp target sequence was verified by Sanger sequencing (Figure S3). Mouse genotypes were confirmed by PCR (Figure 5B). We investigated the expression of *Dnah8* in the testicles of adult mice, including WT, HET, and HO mice (Figure S4). Immunohistochemistry staining revealed that *Dnah8* (brown) was expressed in the center of seminiferous tubule, with a particularly strong localization in the cytoplasm of round spermatids and flagella of elongated spermatids in mouse testes. Compared to the WT group, the expression of Dnah8 on the flagella of mature sperm in

Table 2. Rare damaging variants in *SRD5A2*, *NR5A1* and *AR* featuring severe hypospadias

Sample	Sex	Age (month)	Phenotype	Gene	Chr.	Position[a]	Minor/major allele	Protein change	SIFT[b]	PP2[c]	ACMG/AMP
S156	M	38	Severe	*NR5A1*	9	124482790	A/G	p.L452F	D	D	Likely pathogenic
S72	M	58	Severe	*NR5A1*	9	124482915	A/G	p.P410L	D	D	Likely pathogenic
S21	M	28	Severe	*NR5A1*	9	124493095	T/C	p.D309N	D	D	Likely pathogenic
S149	M	29	Severe	*NR5A1*	9	124493095	T/C	p.D309N	D	D	Likely pathogenic
S50	M	12	Severe	*NR5A1*	9	124500119	A/G	p.R281C	D	D	Likely pathogenic
S48	M	25	Severe	*NR5A1*	9	124500239	A/G	p.R241W	D	D	Likely pathogenic
S47	M	23	Severe	*NR5A1*	9	124500621	G/GA	p.R114Gfs*182	NA	NA	Pathogenic
S44	M	20	Severe	*NR5A1*	9	124500709	T/C	p.R84H	D	P	Likely pathogenic
S5	M	36	Severe	*NR5A1*	9	124503139	T/G	p.R62S	D	D	Likely pathogenic
S102	M	16	Severe	*NR5A1*	9	124503183	C/T	p.Y47C	D	D	Likely pathogenic
S8	M	9	Severe	*AR*	X	67545318	T/C	p.Q58*	NA	NA	Pathogenic
S107	M	20	Severe	*AR*	X	67545325	T/A	p.Q60L	D	P	Likely pathogenic
S50	M	12	Severe	*AR*	X	67545325	T/A	p.Q60L	D	P	Likely pathogenic
S191	M	22	Severe	*AR*	X	67546343	A/G	p.W399*	NA	NA	Pathogenic
S137	M	18	Severe	*AR*	X	67643285	T/C	p.P549L	D	D	Likely pathogenic
S95	M	25	Severe	*AR*	X	67686057	G/A	p.K606E	D	D	Likely pathogenic
S45	M	16	Severe	*AR*	X	67711620	A/C	p.L702I	D	D	Likely pathogenic
S1	M	24	Severe	*AR*	X	67722899	A/G	p.R841H	D	D	Likely pathogenic
S107	M	20	Severe	*AR*	X	67723804	T/C	p.S909F	D	D	Likely pathogenic
S182	M	53	Severe	*SRD5A2*	2	31526224	T/C	p.R246Q	D	D	Likely pathogenic
S184	M	26	Severe	*SRD5A2*	2	31526224	T/C	p.R246Q	D	D	Likely pathogenic
S172	M	35	Severe	*SRD5A2*	2	31529348	G/GA	p.F219Sfs*60	NA	NA	Pathogenic
S97	M	34	Severe	*SRD5A2*	2	31529419	T/C	p.G196S	D	D	Likely pathogenic
S143	M	18	Severe	*SRD5A2*	2	31533629	T/C	p.W140*	NA	NA	Pathogenic
S167	M	19	Severe	*SRD5A2*	2	31533640	T/G	p.Y136*	NA	NA	Pathogenic
S151	M	42	Severe	*SRD5A2*	2	31580690	A/G	p.Q71*	NA	NA	Pathogenic
S3	M	19	Severe	*SRD5A2*	2	31580885	A/G	p.Q6*	NA	NA	Pathogenic
S4	M	12	Severe	*SRD5A2*	2	31580885	A/G	p.Q6*	NA	NA	Pathogenic
S14	M	19	Severe	*SRD5A2*	2	31580885	A/G	p.Q6*	NA	NA	Pathogenic

NA, not available; ACMG/AMP, the guidelines of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.
[a]Positions are given in bp from GRCh38.
[b]SIFT predictions: D, deleterious.
[c]PolyPhen-2 (PP2) predictions: B, benign; D, probably damaging; P, possibly damaging.

the seminiferous tubules of the HET group was significantly decreased. In the HO group, no obvious expression was detected. Sexually mature WT, HET, and HO male mice (8–10 weeks old) were individually caged with sexually mature (8–10 weeks old) WT C57BL/6 female mice (one male with three females) for 2 months, and copulatory plugs were checked every morning. Normal mounting and copulatory plugs were all observed in three groups. However, male mice in the HO group failed to produce offspring over 2 months of breeding, whereas WT and HET group males routinely produced offspring. HO mice lead to male infertile. Since androgen signaling errors are responsible for the development of hypospadias[2] and urethra closure occurs during embryogenesis, we initially detected the concentration of testosterone, which is the major androgen.[4] Compared to the wildtype (*Dnah8*$^{+/+}$) mice at E16.5, intratesticular testosterone levels decreased 30% ($p < 0.05$) in heterozygous (*Dnah8*$^{+/-}$) mice and 59% ($p < 0.01$) in homozygous (*Dnah8*$^{-/-}$) mice (Figure 5C). We further investigated the expression of *Dnah8* in fetal testis in WT, HET and HO mice. Compared to WT mice, the immunofluorescence staining results demonstrated that Dnah8 was significantly decreased in both HET mice and HO mice (Figure 5D). Next, examination of

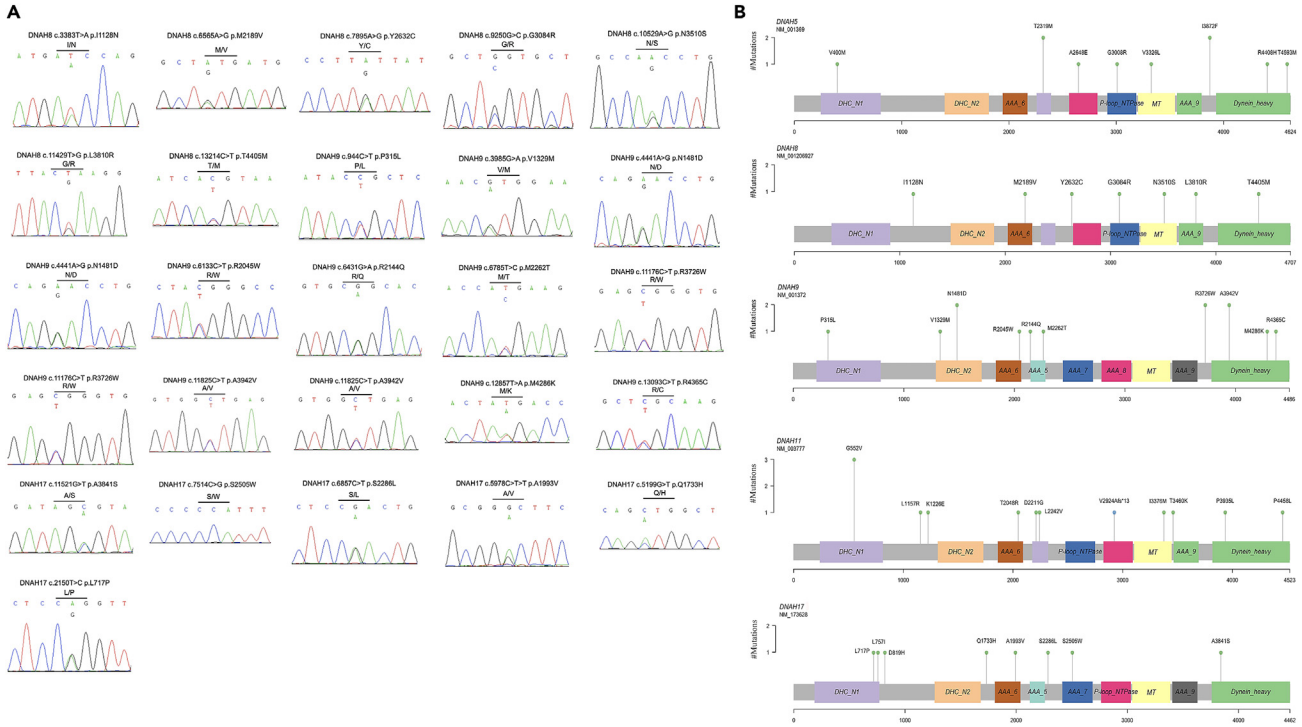


Figure 2. Sanger confirmation and mapping of rare damaging mutations in the *DNAH5*, *DNAH8*, *DNAH9*, *DNAH11*, and *DNAH17* genes

(A) 26 selected rare damaging variants in *DNAH8*, *DNAH9*, and *DNAH17* that were confirmed by Sanger sequencing in severe hypospadias cases.

(B) Mapping of rare damaging mutations in the *DNAH5*, *DNAH8*, *DNAH9*, *DNAH11*, and *DNAH17* genes in severe hypospadias cases.

external genitalia revealed abnormal penile morphology with varying degrees in HO mice at E16.5 (Figure 5E). Abnormal penile morphology was detected in 26.7% (n = 12/45) of the HET and HO mice. All of these mice showed ectopic opening of the urethral duct at the base of the penis. In contrast, no abnormal penile morphology was observed in WT mice (n = 0/35). This difference was statistically significant ($p = 0.0075$, Fisher test). In the WT group, an obvious urethral opening was observed, and the urethral plate epithelium on both sides had fused in the midline to form a closed urethral seam. In contrast, the HET group showed no obvious urethral opening, and the urethral plate epithelium on both sides did not fuse at the midline. The HO group exhibited an obvious abnormal dehiscent urethral opening.

Testosterone is a steroid hormone produced in men mainly by Leydig cells in the testes. Since all rare damaging variants in *DNAH8* genes were heterozygous mutations identified in hypospadias patients, combined with transcriptomic characterization, we further investigated the possible mechanism by which *Dnah8*$^{+/-}$ mice are implicated in hypospadias. Specifically, we used label-free quantitative mass spectrometry to analyze the protein levels in the testes of both HET and WT mice at E16.5. A total of 708 differentially expressed proteins were detected ($p < 0.05$) in the testes of HET mice compared with WT mice. As the genetic related genes of hypospadias affect its development mainly through biological processes,[7] we annotated the differentially expressed proteins and conducted a GO biological process by David database.[19] Multiple significant biological process ($q < 0.05$) were found to be involved in the differentially expressed proteins between WT and HET (Figure 5F) (Table S5), including steroid biosynthesis, protein transport, translocation, transport, ATP synthesis, and sterol biosynthesis. Steroid biosynthesis is the top significantly enriched biological process after Bonferroni correction ($q = 2.3 \times 10^{-4}$). Meanwhile, to localize Dnah8 presence in fetal testis, we performed the immunofluorescence staining of 3β-hydroxysteroid dehydrogenase (3β-HSD), a fetal Leydig cell specific marker and steroid biosynthesis-related protein, on E16.5 testis sections. Results showed that fetal Leydig cells were present in WT mice but fewer in number at E16.5 in both HET mice and HO mice (Figure 5G).

Table 3. 26 rare damaging variants in *DNAH8*, *DNAH9* and *DNAH17* featuring severe hypospadias

Sample	Sex	Age (month)	Gene	Chr.	Position[a]	Minor/major allele	Protein change	SIFT[b]	PP2[c]	MAF in ExAC[d]	MAF in gnomAD[e]
S34	M	20	*DNAH17*	17	78444611	A/C	p.A3841S	D	P	NA[f]	NA[f]
S83	M	14	*DNAH17*	17	78485003	C/G	p.S2505W	D	D	NA[f]	NA[f]
S190	M	15	*DNAH17*	17	78486468	A/G	p.S2286L	D	P	4.1E-04	3.3E-04
S144	M	12	*DNAH17*	17	78495023	A/G	p.A1993V	D	D	NA[f]	NA[f]
S96	M	15	*DNAH17*	17	78501865	A/C	p.Q1733H	D	D	NA[f]	NA[f]
S169	M	42	*DNAH17*	17	78558136	G/A	p.L717P	D	D	8.2E-06	4.1E-06
S168	M	58	*DNAH8*	6	38815517	A/T	p.I1128N	D	D	NA[f]	NA[f]
S34	M	20	*DNAH8*	6	38866657	G/A	p.M2189V	D	D	9.1E-05	1.2E-04
S47	M	23	*DNAH8*	6	38882946	G/A	p.Y2632C	D	D	5.8E-05	4.6E-05
S57	M	30	*DNAH8*	6	38906309	C/G	p.G3084R	D	D	NA[f]	NA[f]
S152	M	37	*DNAH8*	6	38921373	G/A	p.N3510S	D	D	1.2E-04	1.2E-04
S102	M	16	*DNAH8*	6	38931965	G/T	p.L3810R	D	D	4.9E-05	5.1E-05
S29	M	23	*DNAH8*	6	38990172	T/C	p.T4405M	D	P	4.0E-04	3.6E-04
S143	M	18	*DNAH9*	17	11617450	T/C	p.P315L	D	D	7.4E-05	4.5E-05
S136	M	34	*DNAH9*	17	11689807	A/G	p.V1329M	D	P	9.1E-05	7.7E-05
S157	M	26	*DNAH9*	17	11690263	G/A	p.N1481D	D	D	1.5E-04	1.3E-04
S143	M	18	*DNAH9*	17	11690263	G/A	p.N1481D	D	D	1.5E-04	1.3E-04
S31	M	18	*DNAH9*	17	11744818	T/C	p.R2045W	D	D	2.5E-05	2.0E-05
S70	M	27	*DNAH9*	17	11747587	A/G	p.R2144Q	D	P	1.3E-04	1.0E-04
S113	M	17	*DNAH9*	17	11756614	C/T	p.M2262T	D	D	NA[f]	NA[f]
S21	M	28	*DNAH9*	17	11891840	T/C	p.R3726W	D	D	4.3E-04	3.8E-04
S100	M	11	*DNAH9*	17	11891840	T/C	p.R3726W	D	D	4.3E-04	3.8E-04
S25	M	12	*DNAH9*	17	11923889	T/C	p.A3942V	D	D	4.5E-04	5.6E-04
S143	M	18	*DNAH9*	17	11923889	T/C	p.A3942V	D	D	4.5E-04	5.6E-04
S24	M	22	*DNAH9*	17	11961880	A/T	p.M4286K	D	P	2.6E-04	2.2E-04
S92	M	10	*DNAH9*	17	11962116	T/C	p.R4365C	D	D	1.6E-05	8.1E-06

[a]Positions are given in bp from GRCh38.
[b]SIFT predictions: D, deleterious.
[c]PolyPhen-2 (PP2) predictions: B, benign; D, probably damaging; P, possibly damaging.
[d]MAF from Exome Aggregation Consortium (ExAC) database.
[e]MAF from Genome Aggregation Database (gnomAD) database.
[f]Not available.

ODNAH genes cause hypospadias by impairing steroid biosynthesis and may be inherited through unaffected mothers

Based on the genomic data obtained and our mouse model, we proposed that *ODNAH* genes are likely to cause hypospadias by adversely affecting the steroid biosynthesis and androgen signaling, as well as altering the expression of *AR*, *NR5A1*, and *SRD5A2* (Figure 6A). Furthermore, *ODNAH* genes are expressed at higher average mRNA levels in testis tissue compared to other 26 tissues based on the Human Protein Atlas (www.proteinatlas.org), indicating that they may contribute to the abnormal morphogenesis (Figure 6B). To examine the transmission patterns that are caused by either inherited or *de novo* mutations, we recruited samples from two family based trios with mutations of *ODNAH* genes to decipher the genetic predisposition. Among these two trios, one ultra-rare (MAF = 1.6×10^{-5} in the ExAC database) variant (*DNAH8*, NM_001206927.1:c.4690C>T, p.Pro1564Ser or p.P1564S) was identified in *DNAH8* (Figure 6C). Another ultra-rare (MAF = 8.3×10^{-6} in the ExAC database) variant (*DNAH11*, NM_001277115.1:c.9721G>A, p.Ala3241Thr or p.P3241T) was identified in *DNAH11* (Figure 6C). Both are damaging variants and located in a highly conserved region and were inherited from the mothers. This strongly suggested an association with the observed hypospadias.

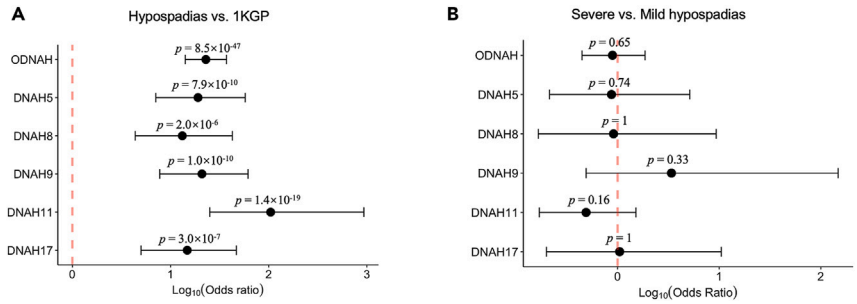


Figure 3. Burden of rare damaging variants in *ODNAH* genes between hypospadias cohort and 1KGP, as well as different type of hypospadias cohort

(A) Rare damaging variants were enriched in *ODNAH* genes for 241 individuals with hypospadias compared to 2,504 controls from 1000 Genomes Project (1KGP).

(B) No rare damaging variants were enriched in *ODNAH* genes for 191 individuals with severe hypospadias compared to 50 individuals with mild hypospadias. Significance of association was displayed with p values; errors bars indicated 95% confidence intervals (CIs) of the corresponding $\log_{10}$ (Odds ratios).

DISCUSSION

Hypospadias is a male-specific congenital disorder. Testosterone is the major androgen, which is responsible for the development of hypospadias. This study represents the first large-scale WES investigation to explore the genetic contribution of rare damaging variants and their potential impact on the etiology of hypospadias. Results indicated that WES has clear technical benefits over limited gene panels as it includes all coding genes and can discover novel causative genes in hypospadias. In addition to confirming that 27% of severe hypospadias patients were found to carry rare damaging variants in the previously reported hypospadias risk-associated genes, this study provides significant findings in several respects.

First, among the previously reported genes leading to the risk of hypospadias, we demonstrated that three testosterone signaling associated genes (Figure 6): nuclear receptor subfamily 5 group A member 1 (*NR5A1*), steroid reductase 5α type 2 (*SRD5A2*), and androgen receptor (*AR*) genes function as mutational hotspots in the etiology of severe hypospadias. In testis Leydig cells, *NR5A1* activates the transcription of several genes involved in steroidogenesis, which leads to the synthesis of testosterone required for the development of external genitalia in a male embryo.[20] *SRD5A2* is known to convert testosterone to DHT, whereas *AR* mediates the androgenic effects of testosterone and DHT. With respect to hypospadias, genetic factors are involved in the testis determination, gonadal steroid synthesis (testosterone and DHT), hormonal biosynthesis, or the transduction mediated by AR.[21] Given the fact that hypospadias is affected by multiple genes and pathways,[8,16] it should be no surprise that the genetic contribution of key genes is

Table 4. Burden analysis of rare damaging variants in *ODNAH* genes by hypospadias vs. control comparison

	Number of rare damaging variants[a]					
Genes	Severe[b], N = 191	Mild[c], N = 50	All[d], N = 241	1KGP Controls, N = 2,504	OR(95% CI)[e]	*P*[f]
ODNAH	51	15	66	30	22.8 (14.3–37.1)	8.5×10^{-47}
DNAH5	10	3	13	7	19.3 (7.1–57.5)	7.9×10^{-10}
DNAH8	7	2	9	7	13.3 (4.4–42.6)	2.0×10^{-6}
DNAH9	13	1	14	7	20.7 (7.7–61.3)	1.0×10^{-10}
DNAH11	13	7	20	2	103.6 (24.9–934.9)	1.4×10^{-19}
DNAH17	8	2	10	7	14.8(5.0–46.3)	3.0×10^{-7}

[a]Rare damaging variants (LoF and D-mis) with MAF<0.1% in 1KGP and ExAC databases.
[b]Severe hypospadias.
[c]Mild hypospadias.
[d]Severe and mild hypospadias.
[e]ORs and 95% CIs were estimated by hypospadias vs. 2,504 controls from 1000 Genomes Project (1KGP).
[f]*P*-value based on binomial test by 241 individuals with hypospadias vs. 2,504 controls from 1KGP.

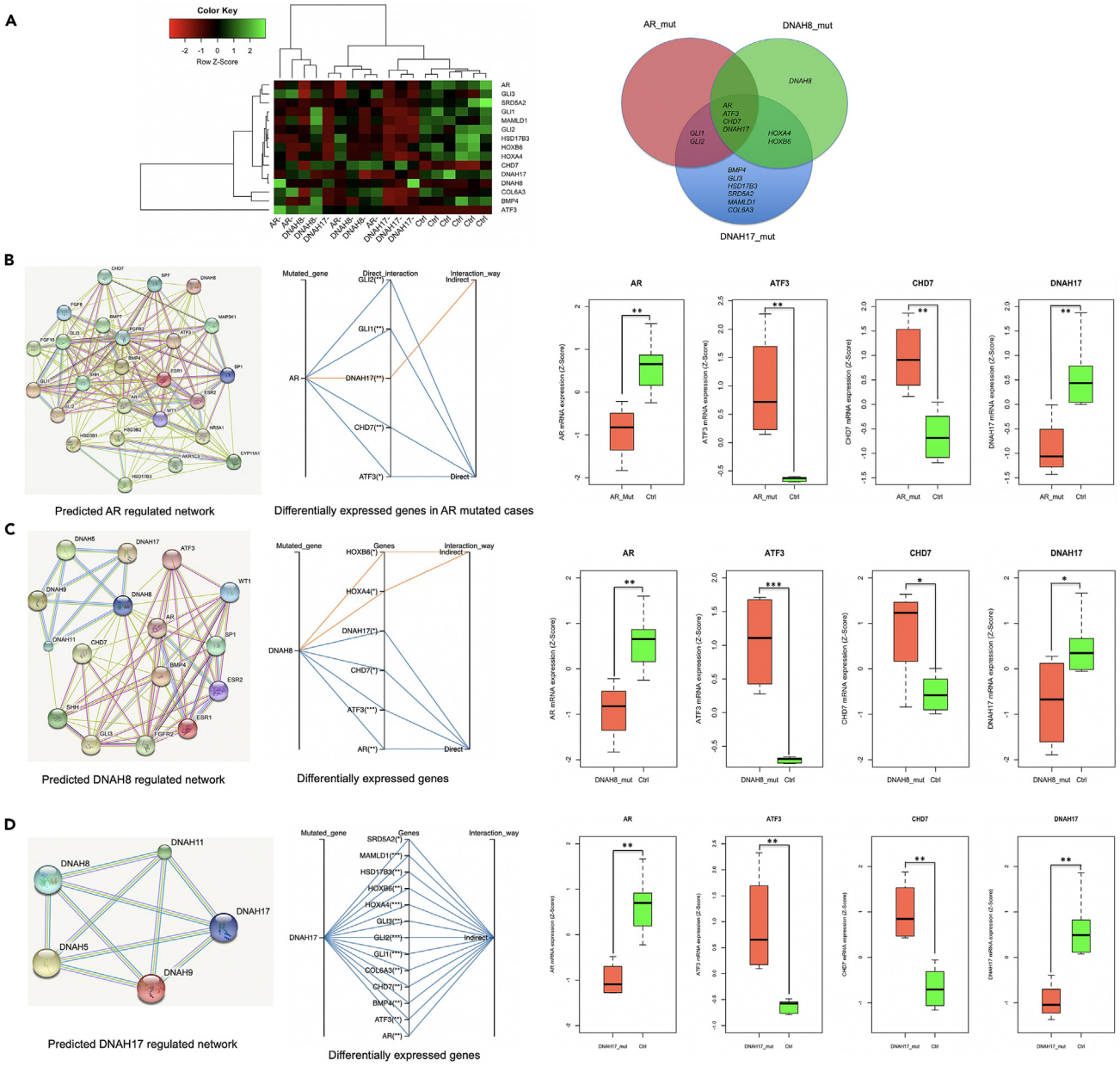


Figure 4. Differential mRNA expression of known hypospadias risk genes in *AR*, *DNAH8*, and *DNAH17* mutated severe hypospadias cases compared with controls

(A) Hierarchical clustering and Venn diagram of 15 differentially expressed genes in *AR*, *DNAH8*, and *DNAH17* mutated patients compared with controls.

(B) Predicted protein-protein interactions (PPIs) network of AR, differentially expressed genes directly or indirectly interacted with AR in *AR* mutated cases based on PPIs, and expression level of four co-regulated genes in *AR* mutated cases.

(C) Predicted protein-protein interactions (PPIs) network of DNAH8, differentially expressed genes directly or indirectly interacted with DNAH8 in *DNAH8* mutated cases based on PPIs, and expression level of four co-regulated genes in *DNAH8* mutated cases.

(D) Predicted protein-protein interactions (PPIs) network of DNAH17, differentially expressed genes directly or indirectly interacted with DNAH8 in *DNAH17* mutated cases based on PPIs, and expression level of four co-regulated genes in *DNAH17* mutated cases. The asterisk indicates a statistical difference (*p < 0.05, **p < 0.01, ***p < 0.001) by Student's *t* test.

not well defined. Our study suggests that mutations in the *AR* and *NR5A1* genes were enriched in severe hypospadias. Alterations in the *AR* gene led to the androgen insensitivity syndrome (AIS), including both the complete (CAIS) and the partial form (PAIS). More *AR* mutations are found in individuals with CAIS than in individuals with PAIS.[22] This study demonstrated that rare damaging mutations in *AR* are more strongly

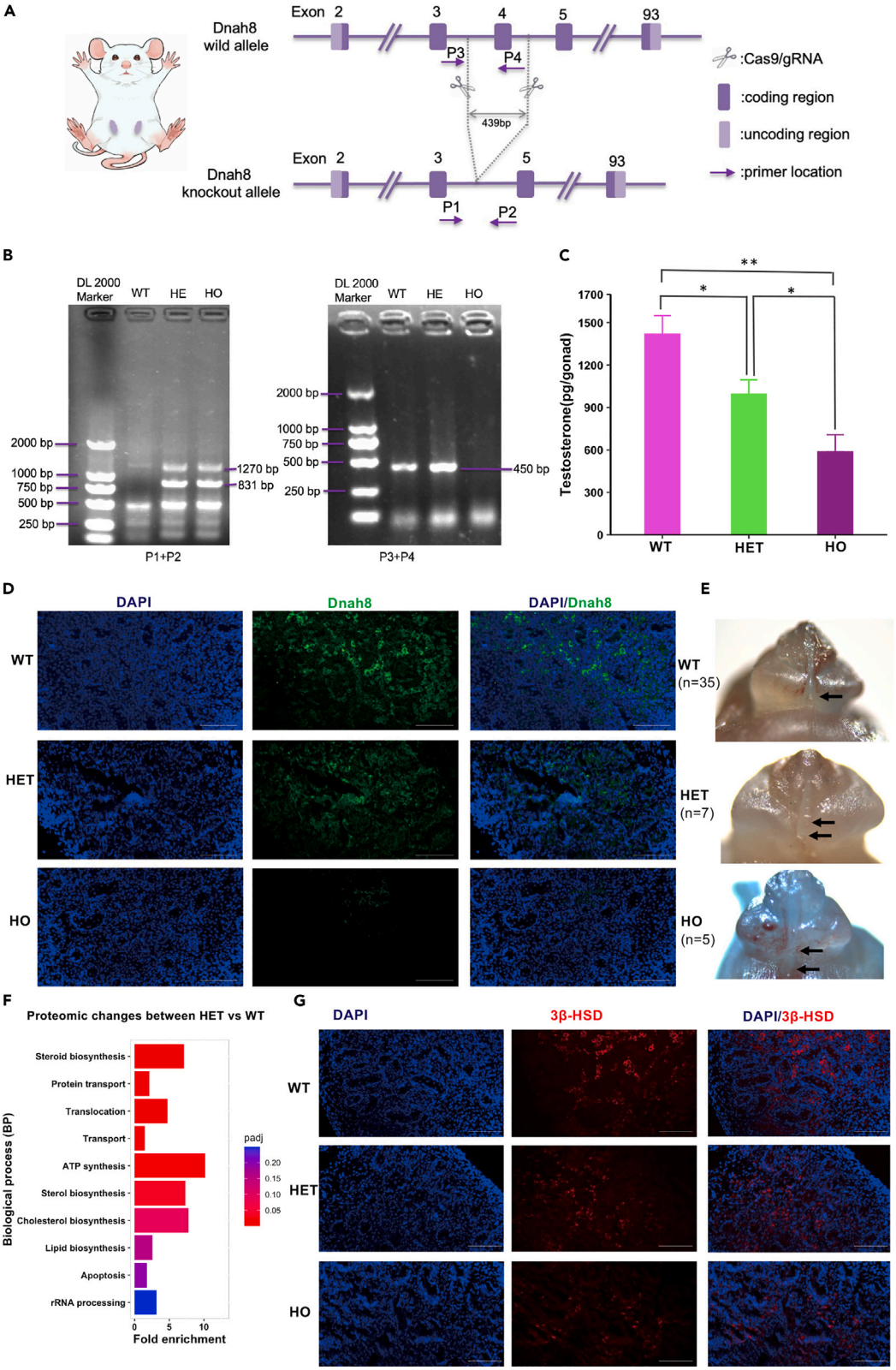

A
Dnah8
wild allele
Exon 2 3 4 5 93
P3
P4
439bp
:Cas9/gRNA
:coding region
:uncoding region
:primer location
Dnah8
knockout allele
Exon 2 3 5 93
P1
P2
B
DL 2000 Marker
WT
HE
HO
2000 bp
1000 bp
750 bp
500 bp
250 bp
1270 bp
831 bp
P1+P2
450 bp
P3+P4
C
Testosterone(pg/gonad)
1700
1500
1200
900
600
300
0
**
*
*
WT
HET
HO
D
DAPI
Dnah8
DAPI/Dnah8
WT
HET
HO
E
WT
(n=35)
HET
(n=7)
HO
(n=5)
F
Proteomic changes between HET vs WT
Steroid biosynthesis
Protein transport
Translocation
Transport
ATP synthesis
Sterol biosynthesis
Cholesterol biosynthesis
Lipid biosynthesis
Apoptosis
rRNA processing
Biological process (BP)
padj
0.20
0.15
0.10
0.05
0
5
10
Fold enrichment
G
DAPI
3β-HSD
DAPI/3β-HSD
WT
HET
HO

Figure 5. Generation of Dnah8-KO mice by CRISPR-Cas9, testosterone levels, immunofluorescence and proteomic analysis

(A) Schematic diagram of the CRISPR/Cas9 targeting region for Dnah8.

(B) PCR genotyping of Dnah8 wildtype (WT, *Dnah8*$^{+/+}$) (453 bp), heterozygous (HET, *Dnah8*$^{+/-}$) (453 bp and 831 bp), and homozygous (HO, *Dnah8*$^{-/-}$) (831 bp) mice.

(C) Intratesticular testosterone levels as measured by ELISA in WT, HET, and HO mice at E16.5, n = 3/group. Results are represented as mean ± SEM, the asterisk indicates a statistical difference (*$p < 0.05$, **$p < 0.01$).

(D) Immunofluorescence of Dnah8 (green) in testes obtained from WT, HET, and HO with E16.5 mice. Nuclei are stained with DAPI (blue). Scale bar: 100 μm.

(E) Urogenital tracts were evaluated for Dnah8-mice at E16.5. Single arrow marks the normal urethral opening; multiple arrows show ectopic urethral opening.

(F) Label-free quantitative mass spectrometry-based proteomics demonstrated that *Dnah8* knocdown mice significantly dysregulated the biological process of steroid biosynthesis.

(G) Immunofluorescence of 3β-HSD (Leydig cell specific marker) in testes obtained from WT, HET, and HO with E16.5 mice. Nuclei are stained with DAPI (red). Scale bars = 100 μm.

associated with severe hypospadias than with mild hypospadias. The *Nr5a1*$^{-/-}$ male mice implicated androgen deficiency with lowered testosterone,[23] and LoF mutations in *NR5A1*, tend to be associated with severe hypospadias.[24] Between 5% and 20% of disorders of sex development patients with *NR5A1* mutations without CAIS, suggesting *NR5A1* deficiency might act in a dose-dependent manner.[25] Since the treatment paradigm of CAIS and PAIS may differ, *AR* and *NR5A1* mutations are crucial for optimal urologic patients' management. Furthermore, several studies have been reported in which patients diagnosed with PAIS carried mutations in *SRD5A2*.[26] Inactivating mutations in *SRD5A2* can lead to a broad spectrum of masculinizing defects, including hypospadias of varying degrees.[27] Even though both *AR* and *SRD5A2* mutations impaired the conversion of testosterone to DHT, the management of such patients is different. While hormone replacement therapy is an effective option for *SRD5A2* deficiency, resistance to AR-directed therapies is still a challenge. Taken together, mutations in *SRD5A2, AR,* and *NR5A1* are crucial for understanding the genetic heterogeneity of hypospadias. Due to the possible treatment variability of this disorder, sequence analysis of *SRD5A2, AR,* and *NR5A1* should be considered in order to better inform the diagnosis and treatment of hypospadias patients. Although other candidate genes *CNTNAP3, HLA-DRB1,* and *DMD* were moderately enriched in rare damaging variants, further research is required to understand their mediated mechanisms in hypospadias.

Another interesting discovery arising from our study is the significant enrichment of rare damaging variants of *ODNAH* genes in 24% of our hypospadias cohort. *DNAH* genes encode axonemal dynein heavy chain. Axonemal dyneins form the inner and outer chains, and through the ATPase activity of their heavy chains, play an important role in the beating of both cilia and sperm.[28] In contrast to *IDNAH* genes or inner dynein arms (IDAs), the significant excess of rare damaging variants in *ODNAH* genes is indicative of their being a major factor in hypospadias causation. Hypospadias has lifelong effects on the urinary function and fertility,[3] which were mediated by testosterone signaling related genes, such as *NR5A1*.[2] In the absence of testosterone, spermatogenesis rarely proceeds beyond meiosis.[29] Among the five *ODNAH* genes or outer dynein arms (ODAs), mutations in *DNAH17*,[30] and *DNAH8*[31] have recently been shown to result in isolated male infertility. Although rare damaging variants in *ODNAH* genes were reported to cause infertility, their roles in testosterone production are unknown. *ODNAH* genes, which are mainly expressed in human testis and are therefore sex-limited, are expected to undergo differential selection. Deleterious mutations that impact either viability before reproduction or fertility are not expected to be transmitted to subsequent generations.[32] Consistent with the inheritance of male-infertility-causative mutations,[33] this study suggests that hypospadias-causative mutations may be inherited through unaffected mothers, who should therefore be a central part of future genomic studies. Ongoing efforts are required to recruit additional hypospadias patients for evaluation of this inheritance model.

Furthermore, our study demonstrated that the knockout of *Dnah8* mice significantly decreases testosterone levels, dysregulated steroid biosynthesis, and may complement known ciliary function. Our research revealed the importance of *Dnah8* in the production of testosterone required for the formation of the male urethra. Binding of testosterone to AR activates AR signals,[34] which regulated various hypospadias risk-associated genes.[16] In addition to AR signaling, ESR1 signaling remains important for the development of male urethra. Imbalance of AR and ESR1 signaling may result in the development of hypospadias.[35] The dysregulation of *AR* and the estrogen-responsive gene *ATF3* in *DNAH8,* and *DNAH17* mutated hypospadias provided further support that such a link exists (Figure 4). Together with the hub role of AR in the PPI network comprised of proteins encoded by previously reported hypospadias risk-associated genes,[16]

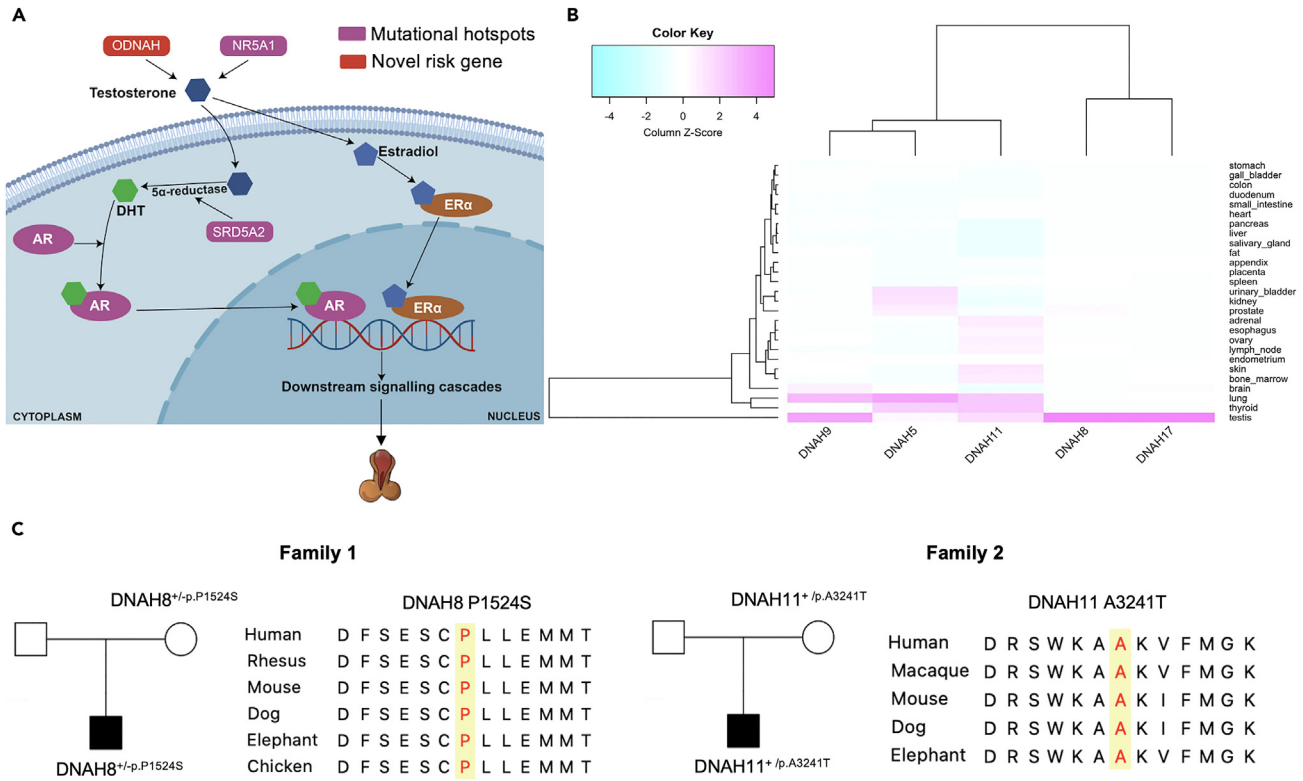


Figure 6. Potential mechanism, tissue expression and pedigree analysis of *ODNAH* genes in the development of hypospadias

(A) *ODNAH*, *NR5A1*, *SRD5A2*, *AR* genes increase risk of hypospadias by affecting the synthesis or androgenic effects of testosterone.

(B) *DNAH5*, *DNAH8*, *DNAH9*, *DNAH11*, and *DNAH17* are highly expressed in testis compared with other tissues.

(C) Two pedigrees with rare damaging mutations of *DNAH8* and *DNAH11* inherited from mothers are shown. The pedigree plot with shapes for male (square), female (circle). The shapes that are shaded indicate hypospadias case.

we posit that *ODNAH* genes are likely to cause hypospadias by inactivating testosterone, which may disrupt the AR signals. As androgen is a major steroid hormones, the significantly dysregulated steroid biosynthesis (Figure 5F) supported this hypothesis. Cilia play crucial roles for the proper development and function of the male reproductive organs, including the testes and the quality of sperm.[36] Disruptions in ciliary function can lead to various developmental and reproductive disorders, including infertility[31] and congenital abnormalities.[37] Bardet-Biedl Syndrome (BBS) is a rare autosomal recessive ciliopathy that affects ciliary function and is associated with various complications, including hypospadias.[37] Although the exact relationship between BBS and hypospadias is not fully understood, it is thought to be related to the role of cilia in urethral development.[38] Although ciliary function can be affected by androgen levels,[39] it is unclear whether cilia affect androgens and how they do so. Our current study demonstrated that ciliary gene may affect androgen levels, which in turn affect urethral development, providing novel insights into the existing pathogenesis of hypospadias.

Finally, this study provides a more complete understanding of the genetic architecture of hypospadias. Defining the genetic architecture of a complex disease is central to the clinical goals of human genetics, which is aid in disease screening, diagnosis, prognosis, and therapy.[40] Hypospadias are a heterogeneous category of male birth defects and are considered to be multifactorial inheritance. However, the genetic architecture underlying hypospadias remains incompletely understood. Recently, an "omnigenic" model of inheritance was proposed that many complex traits are directly affected by a relatively small number of core genes and are largely driven by peripheral genes with no direct role in disease.[17] According to the omnigenic theory, genetic mutations near genes that are expressed in relevant tissues contribute substantially to heritability, while genes which are specifically expressed in other tissues contribute little or nothing.[17] Thus, we proposed that rare damaging mutations of core genes such as testis specific or highly expressed genes would lead to hypospadias. As expected, our study demonstrated that rare damaging

variants is heavily concentrated in *ODNAH* genes. Together with the mouse model, the results provide convincing evidence that *ODNAH* genes tend to have strong effects as the role of core genes. The severity of hypospadias can vary widely, ranging from mild to severe cases, and can be considered a phenotypic continuum. Although the goal of this study is to hunt for core genes that drive hypospadias, it should be noted that the accumulation of extremely rare damaging variants might yielded more severe subtype of hypospadias. Compared to mild hypospadias cases, there was no excess of rare damaging variants identified in severe hypospadias cases (Table S4). In addition, we found that each hypospadias participant contains on average approximately 120 rare damaging variants (Figure S1). Therefore, other rare damaging variants or environmental factors might perturb genetic systems to produce phenotypes. Our recent study showed that triple compound rare damaging variants (two variants from *SLC25A5* and one variant from *AR*) in the same patient rather than single mutation yielded severe hypospadias.[35] Additionally, this study showed that compound rare damaging variants of *DNAH17* p.A3841S and *DNAH8* p.M2189V were found in the same patient with severe hypospadias (Table 3). The omnigenic model also extends the idea of "widespread pleiotropy" on different traits,[17] which were demonstrated in hypospadias. For example, genetic variants in *DNAH8* are associated with male infertility[31] and hypospadias. *DNAH11* variants are significantly associated with congenital heart diseases (CHDs).[41] The recent study showed that boys born with hypospadias increased cardiovascular risk,[42] indicating that *DNAH11* might be important to understand the pleiotropy between CHDs and hypospadias. In brief, this study improves our understanding of genetic architecture of hypospadias, especially the genetic etiology of core genes in hypospadias.

In conclusion, this study demonstrated the importance of utilizing WES in exploring rare damaging variants in hypospadias in an effort to explain the "missing heritability" unidentified in previously published GWASs and targeted genes NGS panel analyses of hypospadias cohorts. Our studies demonstrated that mutations in *NR5A1*, *SRD5A2,* and *AR* genes are crucial for establishing an optimal treatment program for hypospadias patients. In addition, this manuscript now provides unequivocal evidence that mutations in the *ODNAH* genes (*ODNAH5*, *ODNAH8*, *ODNAH9*, *ODNAH11*, *ODNAH17*) contribute to the risk of hypospadias by affecting steroid biosynthesis and the synthesis of testosterone (Figure 6A). This study also complements known ciliary function and will be informative for the genetic counseling of families dealing with hypospadias.

Limitations of the study

There are three limitations to our study. (1) Although this is the largest WES analysis in hypospadias cohort, the sample size for WES is relatively modest compared with many WES. Therefore, the statistical power is still limited for specific gene and gene sets. (2) The severity of hypospadias can vary widely and be thought of as a phenotypic continuum. In some cases, some individuals with milder forms of hypospadias may not be diagnosed. (3) Although *DNAH8* deficiency provides insights into the understanding of the development of hypospadias through androgen reduction, the more functional consequences of other *ODNAH* genes are warranted to better understanding the validity of their roles in the development of hypospadias. Additionally, the cause of male infertility, whether it is due to hypospadias or defects in sperm flagella, remains unclear and warrants further research.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects and samples
  - Generation of Dnah8-knockout mouse model
- METHOD DETAILS
  - Rare variants and functional annotation
  - RNA sequencing and differential gene expression analysis
  - Genotyping analysis
  - Testosterone measurement

- Immunofluorescence staining
- Label-free quantitative mass spectrometry-based proteomics and LC-MS/MS analysis

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106663.

ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (81870459, 81970572). We are grateful to Shiqian Chen and Fuying Lan for providing pictures used in this manuscript.

AUTHOR CONTRIBUTIONS

Z.C., Y.L., and F.C. directed and designed the study. Y.W. recruited study subjects. Z.C., X.Z., and Y.L. performed bioinformatics analysis and function annotation. Z.C., Y.D., and Y.L. conducted the experiments. Z.C., Y.L., R.F., H.X., and F.C. prepared the manuscript; all authors reviewed the manuscript.

DECLARATION OF INTERESTS

Dr. Richard Finnell formerly held a leadership position in the now dissolved TeratOmic Consulting LLC. He also receives travel funds to participate in editorial board meetings of the journal Reproductive and Developmental Medicine, which is published by the Red Hospital of Fudan University. All other authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: December 21, 2022
Revised: March 1, 2023
Accepted: April 7, 2023
Published: April 12, 2023

REFERENCES

1. Yu, X., Nassar, N., Mastroiacovo, P., Canfield, M., Groisman, B., Bermejo-Sánchez, E., Ritvanen, A., Kiuru-Kuhlefelt, S., Benavides, A., Sipek, A., et al. (2019). Hypospadias prevalence and trends in international birth defect surveillance systems, 1980-2010. Eur. Urol. *76*, 482–490. https://doi.org/10.1016/j.eururo.2019.06.027.

2. Matsushita, S., Suzuki, K., Murashima, A., Kajioka, D., Acebedo, A.R., Miyagawa, S., Haraguchi, R., Ogino, Y., and Yamada, G. (2018). Regulation of masculinization: androgen signalling for external genitalia development. Nat. Rev. Urol. *15*, 358–368. https://doi.org/10.1038/s41585-018-0008-y.

3. Stadler, H.S., Peters, C.A., Sturm, R.M., Baker, L.A., Best, C.J.M., Bird, V.Y., Geller, F., Hoshizaki, D.K., Knudsen, T.B., Norton, J.M., et al. (2020). Meeting report on the NIDDK/AUA workshop on congenital anomalies of external genitalia: challenges and opportunities for translational research. J. Pediatr. Urol. *16*, 791–804. https://doi.org/10.1016/j.jpurol.2020.09.012.

4. Bardin, C.W., and Catterall, J.F. (1981). Testosterone: a major determinant of extragenital sexual dimorphism. Science *211*, 1285–1294. https://doi.org/10.1126/science.7010603.

5. Chen, Z., Xie, H., and Chen, F. (2021). Re: vuthy Ea, Anne Bergougnoux, Pascal Philibert, et al. How Far Should We Explore Hypospadias? Next-generation Sequencing Applied to a Large Cohort of Hypospadiac Patients. Eur Urol 2021;79:507-515. Eur. Urol. *80*, e10–e11. https://doi.org/10.1016/j.eururo.2021.03.030.

6. Carmichael, S.L., Shaw, G.M., and Lammer, E.J. (2012). Environmental and genetic contributors to hypospadias: a review of the epidemiologic evidence. Birth Defects Res. A Clin. Mol. Teratol. *94*, 499–510. https://doi.org/10.1002/bdra.23021.

7. Chen, Z.-Z., Gao, Y.-Q., Xie, H., Huang, Y.-C., Chen, F., and Lei, Y.-P. (2022). Transcription factors dysregulated in three complex birth defects datasets. Reprod. Dev. Med. *6*, 79–85. https://doi.org/10.1097/rd9.0000000000000018.

8. Bouty, A., Ayers, K.L., Pask, A., Heloury, Y., and Sinclair, A.H. (2015). The genetic and environmental factors underlying hypospadias. Sex. Dev. *9*, 239–259. https://doi.org/10.1159/000441988.

9. Geller, F., Feenstra, B., Carstensen, L., Pers, T.H., van Rooij, I.A.L.M., Körberg, I.B., Choudhry, S., Karjalainen, J.M., Schnack, T.H., Hollegaard, M.V., et al. (2014). Genome-wide association analyses identify variants in developmental genes associated with hypospadias. Nat. Genet. *46*, 957–963. https://doi.org/10.1038/ng.3063.

10. van der Zanden, L.F.M., van Rooij, I.A., Feitz, W.F.J., Knight, J., Donders, A.R.T., Renkema, K.Y., Bongers, E.M., Vermeulen, S.H., Kiemeney, L.A., Veltman, J.A., et al. (2011). Common variants in DGKK are strongly associated with risk of hypospadias. Nat. Genet. *43*, 48–50. https://doi.org/10.1038/ng.721.

11. Ea, V., Bergougnoux, A., Philibert, P., Servant-Fauconnet, N., Faure, A., Breaud, J., Gaspari, L., Sultan, C., Paris, F., and Kalfa, N. (2021). How far should we explore hypospadias? Next-Generation sequencing applied to a large cohort of hypospadiac patients. Eur. Urol. *79*, 507–515. https://doi.org/10.1016/j.eururo.2020.12.036.

12. Sveinbjornsson, G., Albrechtsen, A., Zink, F., Gudjonsson, S.A., Oddson, A., Másson, G., Holm, H., Kong, A., Thorsteinsdottir, U., Sulem, P., et al. (2016). Weighting sequence variants based on their annotation increases power of whole-genome association studies. Nat. Genet. *48*, 314–317. https://doi.org/10.1038/ng.3507.

13. Chen, Z., Kuang, L., Finnell, R.H., and Wang, H. (2018). Genetic and functional analysis of SHROOM1-4 in a Chinese neural tube defect cohort. Hum. Genet. *137*, 195–202. https://doi.org/10.1007/s00439-017-1864-x.

14. Chen, Z., Lei, Y., Zheng, Y., Aguiar-Pulido, V., Ross, M.E., Peng, R., Jin, L., Zhang, T., Finnell, R.H., and Wang, H. (2018). Threshold for neural tube defect risk by accumulated singleton loss-of-function variants. Cell Res. *28*, 1039–1041. https://doi.org/10.1038/s41422-018-0061-3.

15. DeBoever, C., Tanigawa, Y., Lindholm, M.E., McInnes, G., Lavertu, A., Ingelsson, E., Chang, C., Ashley, E.A., Bustamante, C.D., Daly, M.J., and Rivas, M.A. (2018). Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. Nat. Commun. *9*, 1612. https://doi.org/10.1038/s41467-018-03910-9.

16. Chen, Z., Lin, X., Lei, Y., Chen, H., Finnell, R.H., Wang, Y., Xu, J., Lu, D., Xie, H., and Chen, F. (2019). Genome-wide association study in Chinese cohort identifies one novel hypospadias risk associated locus at 12q13.13. BMC Med. Genom. *12*, 196. https://doi.org/10.1186/s12920-019-0642-0.

17. Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An expanded view of complex traits: from polygenic to omnigenic. Cell *169*, 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038.

18. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419. https://doi.org/10.1126/science.1260419.

19. Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., and Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. *50*, W216–W221. https://doi.org/10.1093/nar/gkac194.

20. Ozisik, G., Achermann, J.C., Meeks, J.J., and Jameson, J.L. (2003). SF1 in the development of the adrenal gland and gonads. Horm. Res. *59* (*Suppl 1*), 94–98. https://doi.org/10.1159/000067831.

21. Kalfa, N., Gaspari, L., Ollivier, M., Philibert, P., Bergougnoux, A., Paris, F., and Sultan, C. (2019). Molecular genetics of hypospadias and cryptorchidism recent developments. Clin. Genet. *95*, 122–131. https://doi.org/10.1111/cge.13432.

22. Gottlieb, B., Beitel, L.K., Nadarajah, A., Paliouras, M., and Trifiro, M. (2012). The androgen receptor gene mutations database: 2012 update. Hum. Mutat. *33*, 887–894. https://doi.org/10.1002/humu.22046.

23. Karpova, T., Ravichandiran, K., Insisienmay, L., Rice, D., Agbor, V., and Heckert, L.L. (2015). Steroidogenic factor 1 differentially regulates fetal and adult leydig cell development in male mice. Biol. Reprod. *93*, 83. https://doi.org/10.1095/biolreprod.115.131193.

24. Köhler, B., Lin, L., Mazen, I., Cetindag, C., Biebermann, H., Akkurt, I., Rossi, R., Hiort, O., Grüters, A., and Achermann, J.C. (2009). The spectrum of phenotypes associated with mutations in steroidogenic factor 1 (SF-1, NR5A1, Ad4BP) includes severe penoscrotal hypospadias in 46,XY males without adrenal insufficiency. Eur. J. Endocrinol. *161*, 237–242. https://doi.org/10.1530/EJE-09-0067.

25. Achermann, J.C., Ozisik, G., Ito, M., Orun, U.A., Harmanci, K., Gurakan, B., and Jameson, J.L. (2002). Gonadal determination and adrenal development are regulated by the orphan nuclear receptor steroidogenic factor-1, in a dose-dependent manner. J. Clin. Endocrinol. Metab. *87*, 1829–1833. https://doi.org/10.1210/jcem.87.4.8376.

26. Avendaño, A., Paradisi, I., Cammarata-Scalisi, F., and Callea, M. (2018). 5-alpha-Reductase type 2 deficiency: is there a genotype-phenotype correlation? A review. Hormones (Basel) *17*, 197–204. https://doi.org/10.1007/s42000-018-0013-9.

27. Chávez, B., Ramos, L., Gómez, R., and Vilchis, F. (2014). 46,XY disorder of sexual development resulting from a novel monoallelic mutation (p.Ser31Phe) in the steroid 5alpha-reductase type-2 (SRD5A2) gene. Mol. Genet. Genom. Med. *2*, 292–296. https://doi.org/10.1002/mgg3.76.

28. King, S.M. (2016). Axonemal dynein arms. Cold Spring Harbor Perspect. Biol. *8*, a028100. https://doi.org/10.1101/cshperspect.a028100.

29. Walker, W.H. (2011). Testosterone signaling and the regulation of spermatogenesis. Spermatogenesis *1*, 116–120. https://doi.org/10.4161/spmg.1.2.16956.

30. Whitfield, M., Thomas, L., Bequignon, E., Schmitt, A., Stouvenel, L., Montantin, G., Tissier, S., Duquesnoy, P., Copin, B., Chantot, S., et al. (2019). Mutations in DNAH17, encoding a sperm-specific axonemal outer dynein arm heavy chain, cause isolated male infertility due to asthenozoospermia. Am. J. Hum. Genet. *105*, 198–212. https://doi.org/10.1016/j.ajhg.2019.04.015.

31. Liu, C., Miyata, H., Gao, Y., Sha, Y., Tang, S., Xu, Z., Whitfield, M., Patrat, C., Wu, H., Dulioust, E., et al. (2020). Bi-Allelic DNAH8 variants lead to multiple morphological abnormalities of the sperm flagella and primary male infertility. Am. J. Hum. Genet. *107*, 330–341. https://doi.org/10.1016/j.ajhg.2020.06.004.

32. McClellan, J., and King, M.C. (2010). Genetic heterogeneity in human disease. Cell *141*, 210–217. https://doi.org/10.1016/j.cell.2010.03.032.

33. Gershoni, M., and Pietrokovski, S. (2014). Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. Nat. Commun. *5*, 4438. https://doi.org/10.1038/ncomms5438.

34. Sakkiah, S., Wang, T., Zou, W., Wang, Y., Pan, B., Tong, W., and Hong, H. (2017). Endocrine disrupting chemicals mediated through binding androgen receptor are associated with diabetes mellitus. Int. J. Environ. Res. Publ. Health *15*, 25. https://doi.org/10.3390/ijerph15010025.

35. Chen, Z., Lin, X., Wang, Y., Xie, H., and Chen, F. (2020). Dysregulated expression of androgen metabolism genes and genetic analysis in hypospadias. Mol. Genet. Genom. Med. *8*, e1346. https://doi.org/10.1002/mgg3.1346.

36. Rosenfeld, C.S. (2019). Male reproductive tract cilia beat to a different drummer. Proc. Natl. Acad. Sci. USA *116*, 3361–3363. https://doi.org/10.1073/pnas.1900112116.

37. Moore, S.J., Green, J.S., Fan, Y., Bhogal, A.K., Dicks, E., Fernandez, B.A., Stefanelli, M., Murphy, C., Cramer, B.C., Dean, J.C.S., et al. (2005). Clinical and genetic epidemiology of Bardet-Biedl syndrome in Newfoundland: a 22-year prospective, population-based, cohort study. Am. J. Med. Genet. *132A*, 352–360. https://doi.org/10.1002/ajmg.a.30406.

38. Quinlan, R.J., Tobin, J.L., and Beales, P.L. (2008). Modeling ciliopathies: primary cilia in development and disease. Curr. Top. Dev. Biol. *84*, 249–310. https://doi.org/10.1016/S0070-2153(08)00605-4.

39. Dragićević, S. (2010). Evaluation of health-related quality of life in patients with prostate cancer after treatment with radical retropubic prostatectomy and permanent prostate brachytherapy. Urol. Int. *85*, 173–179. https://doi.org/10.1159/000318187.

40. Timpson, N.J., Greenwood, C.M.T., Soranzo, N., Lawson, D.J., and Richards, J.B. (2018). Genetic architecture: the shape of the genetic contribution to human traits and disease. Nat. Rev. Genet. *19*, 110–124. https://doi.org/10.1038/nrg.2017.101.

41. Liu, S., Chen, W., Zhan, Y., Li, S., Ma, X., Ma, D., Sheng, W., and Huang, G. (2019). DNAH11 variants and its association with congenital heart disease and heterotaxy syndrome. Sci. Rep. *9*, 6683. https://doi.org/10.1038/s41598-019-43109-6.

42. Lucas-Herald, A.K., Montezano, A.C., Alves-Lopes, R., Haddow, L., Alimussina, M., O'Toole, S., Flett, M., Lee, B., Amjad, S.B., Steven, M., et al. (2022). Vascular dysfunction and increased cardiovascular risk in hypospadias. Eur. Heart J. *43*, 1832–1845. https://doi.org/10.1093/eurheartj/ehac112.

43. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74. https://doi.org/10.1038/nature15393.

44. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. *4*, 1073–1081. https://doi.org/10.1038/nprot.2009.86.

45. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*,

248–249. https://doi.org/10.1038/nmeth0410-248.

46. McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. Bioinformatics *26*, 2069–2070. https://doi.org/10.1093/bioinformatics/btq330.

47. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of medical genetics and genomics and the association for molecular Pathology. Genet. Med. *17*, 405–424. https://doi.org/10.1038/gim.2015.30.

48. Jiang, H., Lei, R., Ding, S.W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinf. *15*, 182. https://doi.org/10.1186/1471-2105-15-182.

49. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

50. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295. https://doi.org/10.1038/nbt.3122.

51. Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M.R., Armean, I.M., Bennett, R., Bhai, J., Billis, K., Boddu, S., et al. (2019). Ensembl 2019. Nucleic Acids Res. *47*, D745–D751. https://doi.org/10.1093/nar/gky1113.

52. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80. https://doi.org/10.1186/gb-2004-5-10-r80.

53. Chen, Z., Zhang, T., Lin, J., Yan, Z., Wang, Y., Zheng, W., and Weng, K.C. (2014). GeneSense: a new approach for human gene annotation integrated with protein-protein interaction networks. Sci. Rep. *4*, 4474. https://doi.org/10.1038/srep04474.

54. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372. https://doi.org/10.1038/nbt.1511.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-DNAH8 Antibody	Abcam	Cat# ab121989; RRID:AB_11127702
Anti-HSD3B1 Antibody	Abcam	Cat# ab55268; RRID:AB_942015
Anti-Rabbit IgG H&L (goat), Alexa Fluor® 488	Abcam	Cat# ab150077; RRID:AB_2630356
Anti-mouse IgG H&L (goat), Cy3	Servicebio	Cat# GB21301; RRID:AB_2923552
Biological samples		
Blood samples and foreskin specimen from children with hypospadias	This study	N/A
Chemicals, peptides, and recombinant proteins		
4,6-diamidino-2-phenylindole (DAPI)	Servicebio	G1012
Animal Genomic DNA Quick Extraction Kit for PCR Analysis	Beyotime	D0065S
TaKaRa Ex Taq®	Takara	RR001A
Water-DEPC Treated Water	Sangon Biotech	B501005
4% Paraformaldehyde Fix Solution	Beyotime	P0099
1X PBS Buffer	Sangon Biotech	B540626
Critical commercial assays		
Testosterone ELISA kit	Beyotime	PT872
Deposited data		
1KGP data	1000 Genome Project	www.internationalgenome.org
ExAC data	ExAC	http://exac.broadinstitute.org
RNA expression data for different tissues	Human Protein Atlas	www.proteinatlas.org (BioProject: PRJEB4337)
GnomAD data	GnomAD	https://gnomad.broadinstitute.org
Experimental models: Organisms/strains		
Mouse: C57BL/6 DNAH8$^{+/+}$	Shanghai Model Organisms Center	N/A
Mouse: C57BL/6 DNAH8$^{-/+}$	Shanghai Model Organisms Center	This study
Mouse: C57BL/6 DNAH8$^{-/-}$	Shanghai Model Organisms Center	This study
Oligonucleotides		
Genotyping primer: DNAH8 P1 Forward AAGGGCTGCTGGCAGAATAC	This study	N/A
Genotyping primer: DNAH8 P2 Reverse ATCCTGATGTCCCCGAGTGT	This study	N/A
Genotyping primer: DNAH8 P3 Forward CAGAGGTTTCTCCGTGGAGTC	This study	N/A
Genotyping primer: DNAH8 P4 Reverse TTCAGCGTCTTACAGCCACC	This study	N/A
Software and algorithms		
PolyPhen-2	PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/
SIFT	SIFT	http://sift.bii.a-star.edu.sg/
TreeFam database	TreeFam	http://www.treefam.org/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
STRING	STRING	https://www.string-db.org
R 4.1.0	R studio	https://www.r-project.org/
GraphPad Prism 7	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
ImageJ	ImageJ	https://imagej.nih.gov/ij/

RESOURCE AVAILABILITY

Lead contact

Further information and any related requests should be directed to and will be fulfilled by the lead contact, Dr. Zhongzhong Chen (zhongzhongchen@gmail.com).

Materials availability

All reagents generated in this study being made available upon request to the lead contact.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects and samples

A total of 243 hypospadias patients with mean age 2.3 $\pm$ 0.1 years, including 241 sporadic patients and two trios (an affected proband with both parents), were enrolled from 2011 to 2019 at Shanghai Children's Hospital, Shanghai Jiaotong University. All patients were non-consanguineous Chinese and having pathologically diagnosed hypospadias recruited from the Department of Urology at Shanghai Children's Hospital. According to the abnormal location of the urethral opening, the patients were divided into three categories: mild (glandular), moderate (penile), or severe (in the scrotum or perineum). Of the 241 sporadic patients ultimately enrolled, there were 79.3% severe cases and 20.7% mild cases (Table S1). We performed WES on genomic DNA extracted from blood samples of all 241 sporadic patients and two trios. For the RNA sequencing analysis, 18 outer layers of prepuces of children who underwent consecutive circumcision either because of phimosis (controls; n = 6) or because of hypospadias repair (*AR* mutated severe hypospadias, n = 4; *DNAH8* mutated severe hypospadias, n = 4; *DNAH17* mutated severe hypospadias, n = 4), were included.

The study was conducted with the approval of the Ethics Committee of the Shanghai Children's Hospital in China (approval #: 2020R018-E01). Each patient was informed of the purpose of the study and written consent forms were obtained from all participants or their parent/legal guardian.

Generation of Dnah8-knockout mouse model

A mouse model of the *Dnah8* knockout was generated by using CRISPR-Cas9 technology. Briefly, Cas9 mRNA was *in vitro* transcribed with mMESSAGE mMACHINE T7 Ultra Kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. Two sgRNAs targeted to delete exons 4 were *in vitro* transcribed using the MEGAshortscript Kit (ThermoFisher, USA). One sgRNA targeted to intron 3 of gene *Dnah8* was 5'- AGACTTAAAGACCTTCCCGAAGG -3'; the other sgRNA targeted to intron 4 of gene *Dnah8* was 5'- CATTATCTTAAGAACACTCT GGG-3'. *In vitro*-transcribed Cas9 mRNA and sgRNAs were injected into zygotes of C57BL/6J mouse and transferred to pseudopregnant recipients. The F0 mice were validated by PCR and sequencing using primer pairs: F-5'-AAGGGCTGCTGGCAGAATAC-3'; R-5'-ATCCTGATGTCCCCGAGTGT-3'. The positive F0 mice were chosen and crossed with C57BL/6J mice to obtain F1 heterozygous *Dnah8* knockout mice. The genotype of F1 mice was identified by PCR and confirmed by DNA sequencing. Male and female F1 heterozygous mice were intercrossed to produce the

homozygous *Dnah8* knockout mice. All animal handling and experimental procedures were carried out in accordance with the guidelines stipulated and approved by the Ethics Committee of the Shanghai Jiao Tong University Affiliated Sixth People's Hospital.

METHOD DETAILS

Rare variants and functional annotation

We systematically investigated the mutation patterns of rare damaging variants in a Chinese cohort composed of 191 severe hypospadias cases and 208 controls of the Han Chinese population (CHS + CHB)from the 1KGP[43] (1000 Genomes Project) (www.internationalgenome.org). Then we confirmed our results in 50 mild hypospadias cases and two trios. We performed the WES analysis and functional annotation using the method of Chen.[14] The missense variants that were predicted to be probably deleterious by both Sorting Intolerant From Tolerant (SIFT)[44] and Polymorphism Phenotyping version 2 (PolyPhen-2),[45] were annotated as probably deleterious missense variants (D-mis). Variants with a minor allele frequency (MAF) > 1% in the hypospadias or controls cohort were excluded. Rare probably damaging variants [LoF (loss of function, including splice acceptor/donor, stop gain/lost, initiator codon and frameshift indels) and D-mis] were further selected that had a MAF < 0.1% in the 1KGP and ExAC databases (http://exac.broadinstitute.org). The candidate rare damaging variants in 30 previously reported hypospadias risk associated genes[8,16] were interrogated to demonstrate their genetic contribution to the etiology of hypospadias. *ODNAH* gene family (*DNAH5*, *DNAH8*, *DNAH9*, *DNAH11* and *DNAH17*) (TreeFam accession number: TF316836) members were identified using the EMBL-EBI database (http://www.treefam.org/family/TF316836). Twenty-seven identified variants in *DNAH8*, *DNAH9* and *DNAH17* were selected to be further validated by Sanger sequencing. These rare variants were also annotated with MAFs from gnomAD database (v170228) (https://gnomad.broadinstitute.org) via VEP.[46] Additionally, candidate rare damaging variants for known genes were classified according to the guidelines of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP).[47] Of note, PP3 is used to support a deleterious effect on the gene or gene product based on multiple lines of computational evidence. PM2 is applied when the variant is absent from controls or if the MAF in ExAC database < 0.1%. PP2 is applied when missense variant in a gene with a low rate of benign missense variation. PP4 is applied when patient's phenotype or family history highly specific for a disease with a single genetic etiology. If the variants are located in a mutational hot spot or well-established functional domain, then PM1 will be applied. When the variants are ''non-frameshift deletions/insertions'' in a non-repeat region or stop-loss variants, PM4 will be applied. PVS1 is applied when there is a null variant in a gene where loss of function (LoF) is a known mechanism of disease.

RNA sequencing and differential gene expression analysis

RNA-Seq fastq files were generated from the Illumina NovaSeq 6000 sequencer. Skewer software[48] was used to dynamically remove the 3' ends, linker sequences, and low mass fragments of raw reads. Quality assessment was carried out by FastQC tool (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed clean reads were then mapped to the human reference genome (GRCh38) using STAR.[49] The transcriptome was assembled via StringTie software[50] based on the Ensembl database[51] annotation. The FPKM (Fragments Per Kilobase of transcript per Million mapped reads) method was applied to assess differentially expressed genes. Function heatmap.2 was utilized to create the graphical display of the dendrogram.[52] GeneSense[53] and STRING (https://www.string-db.org) were used to identify protein-protein interactions (PPIs). RNA expression data for 27 different tissues that were obtained from 95 individuals as part of the Human Protein Atlas[18] (www.proteinatlas.org) (BioProject: PRJEB4337).

Genotyping analysis

Genotyping for the Dnah8 gene was performed using standard PCR conditions. Forward (F) and reverse (R) primers were: mutant-P1(F): 5'-AAGGGCTGCTGGCAGAATAC-3', mutant-P2(R): 5'-ATCCTGATG TCCCCGAGTGT-3' , wildtype-P3(F): 5'-CAGAGGTTTCTCCGTGGAGTC-3', wildtype-P4(F): 5'-TTCAGC GTCTTACAGCCACC-3'. Expected band sizes were 453 bp for wildtype and 831 bp for knockout ($Dnah8^{-/-}$) animals, respectively, and heterozygous mice showed both bands.

Testosterone measurement

The intratesticular samples were collected at each experimental time point for the quantitative measurement of testosterone levels. Testosterone levels in the medium were determined using a mouse

testosterone (T) enzyme-linked immunosorbent assay (ELISA) (Beyotime Institute of Biotechnology, Jiangsu, China) according to the manufacturer's instructions. Independent measurements of testosterone levels in each group (wildtype mice, *Dnah8*$^{+/-}$ mice and *Dnah8*$^{-/-}$ mice) were repeated three times at E16.5.

Immunofluorescence staining

Using the embedded specimen previously described from which 6μm thick sections were cut using a paraffin-microtome (RM2235; Leica Biosystems). Sections were mounted on slides treated with Polylysine, bake at 60 °C for 8 hours. After xylene dewaxing and gradient alcohol rehydration, citrate buffer (pH6.0, 0.01 mol/L) was used for antigen repair by microwave repair. These sections were washed with PBS three times and blocked with 5% bovine serum albumin having 0.1% Triton X-100. After blocking, the testicular sections were immersed in primary antibody for 12 hours at 4°C and then incubated with a secondary antibody at 37°C for 1-2 hours.Anti-DNAH8 (1:100, ab121989, Sigma-Aldrich) and anti-HSD3B1(1:200, ab55268, Sigma-Aldrich) were used as primary antibodies. IgG H&L (Alexa Fluor 488, ab150077, Abcam) and IgG H&L (Cy3, GB21301, Servicebio) were used as secondary antibodies. The nuclei were stained simultaneously with 4′,6-diamidino-2-phenylindole (DAPI). All the images were captured using a I8 microscope (Leica Microsystems).

Label-free quantitative mass spectrometry-based proteomics and LC-MS/MS analysis

Fetal male mice were obtained from 16.5 day pregnant mice, and bilateral testicular tissues were obtained by surgical operation. SDT (4% SDS, 100mM Tris HCl, pH7.6) buffer was used for the cleavage and protein extraction testis obtained from WT, and HET with E16.5 mice. BCA protein assay kit (Bio Rad, USA) was used to quantify the amount of protein. 20 μg protein of each sample was mixed with 5X loaded buffer solution and boiled for 5 minutes, then protein was separated in SDS-PAGE gel (constant voltage 180V, 45 minutes). Protein digestion was performed by trypsin according to the Filtration Assisted Sample Preparation (FASP) procedure. For each HE and WT sample, three technical replicates were analyzed.

LC-MS/MS analysis, protein identification and quantification, and bioinformatics analysis are completed by Applied Protein Technology (APL, Shanghai) company. Use MaxQuant 1.6.14 software to combine and search the original MS data of each sample using Label-free quantitative (LFQ) method for identification and quantitative analysis. *P* value is calculated by significance B algorithm.[54] Differentially expressed proteins were annotated by David database.[19]

QUANTIFICATION AND STATISTICAL ANALYSIS

Binomial testing and Fisher's exact test were performed for burden analysis of rare coding variants by comparisons of hypospadiac patients and controls. Two-tailed Student's t tests were conducted to determine the differentially expressed genes. Testicular testosterone levels are shown as the mean ± standard error of the mean (SEM) and corresponding statistical analyses were performed using one tailed t-test. Differences were considered to be significant at $P < 0.05$ and are represented by *, those at $P \leq 0.01$ are indicated by **, and those at $P \leq 0.001$ are represented by ***. All statistical analyses were performed by R software (http://www.R-project.org).