scientific reports

OPEN

Role of epigenetics in the etiology of hypospadias through penile foreskin DNA methylation alterations

Martin Kaefer[1], Richard Rink[1], Rosalia Misseri[1], Paul Winchester[2], Cathy Proctor[2], Millissia Ben Maamar[3], Daniel Beck[3], Eric Nilsson[3] & Michael K. Skinner[3]✉

Abnormal penile foreskin development in hypospadias is the most frequent genital malformation in male children, which has increased dramatically in recent decades. A number of environmental factors have been shown to be associated with hypospadias development. The current study investigated the role of epigenetics in the etiology of hypospadias and compared mild (distal), moderate (mid shaft), and severe (proximal) hypospadias. Penile foreskin samples were collected from hypospadias and non-hypospadias individuals to identify alterations in DNA methylation associated with hypospadias. Dramatic numbers of differential DNA methylation regions (DMRs) were observed in the mild hypospadias, with reduced numbers in moderate and low numbers in severe hypospadias. Atresia (cell loss) of the principal foreskin fibroblast is suspected to be a component of the disease etiology. A genome-wide (>95%) epigenetic analysis was used and the genomic features of the DMRs identified. The DMR associated genes identified a number of novel hypospadias associated genes and pathways, as well as genes and networks known to be involved in hypospadias etiology. Observations demonstrate altered DNA methylation sites in penile foreskin is a component of hypospadias etiology. In addition, a potential role of environmental epigenetics and epigenetic inheritance in hypospadias disease etiology is suggested.

Abbreviations

DMRs	Differential DNA methylation regions
DES	Diethylstilbestrol
MeDIP	Methylated DNA immunoprecipitation
NGS	Next generation sequencing
FDR	False discovery rate
PCA	Principal component analysis
Supplemental Table S6	Gene symbols and names

Hypospadias involves abnormal penile and urethral development in newborn males. This is one of the most frequent genital malformations in male children, which has increased dramatically in past decades[1–3]. The prevalence of hypospadias in the U.S. is approximately 6.8 per 1000 male births, which has increased from 6.1 per 1000 in 1997, for an 11.5% increase over 15 years[4]. This supports a potential environmental impact and role of epigenetic inheritance in hypospadias disease etiology[3,5,6]. The severity of hypospadias is classified as mild (distal penile opening near the tip of the penis) (50% of cases), moderate (mid shaft penile opening midway up the penis) (30%), and severe (proximal penile opening at or below the scrotum) (20%). Clinical factors associated with increased risk of hypospadias include: having an affected family member, maternal hypertension and preeclampsia, small for dates at birth, prematurity, and multiple gestation[7]. Although urethral formation occurs between the 8th and 14th week of pregnancy, the developmental origins or etiology of hypospadias is largely unknown[1–3]. The assumption is genetic mutations in critical genes will be the cause of the disease[7–10]. However,

[1]Department of Pediatric Urology, Riley Hospital for Children, Indiana University School of Medicine, Indiana University, Indianapolis, IN 46202-5201, USA. [2]Department of Pediatrics, St. Franciscan Hospital, School of Medicine, Indiana University, Indianapolis, IN 46202-5201, USA. [3]Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA. ✉email: skinner@wsu.edu

genome-wide association studies (GWAS) suggest less than 10% of the hypospadias patients have correlated genetic mutations[11]. Heritability estimates for hypospadias range between 57 and 77%[12–14]. Observations suggest a potential role of epigenetic inheritance[15] of hypospadias and environmental impacts are now thought to be critical[8]. Since environmental factors such as toxicants or pharmaceuticals (diethylstilbestrol, DES) cannot directly alter DNA sequence, environmental epigenetics is thought to be the primary molecular mechanism involved[6]. Abnormal androgen production or action in later fetal development is associated with hypospadias[1]. It has been postulated that changes in concentrations of sex hormones during the crucial period of penile fetal development (weeks 8–14) through androgenesis and environmental exposures (i.e., endocrine disruptors) are major factors. Environmental factors such as diet and chemical exposures of the mother and fetus are also felt to likely be involved in the etiology of hypospadias[5,15–17], and are known to act through epigenetics[6].

Penile development is dependent on normal fibroblast maturation and eventual foreskin formation[1]. Abnormal foreskin development is one of the primary factors in hypospadias. This penile tissue is androgen responsive and thus would be vulnerable to endocrine disruptor exposures (i.e., anti-androgenic fungicide vinclozolin) throughout development. Male sexual differentiation regulates testosterone production in the testis and actions on the genital target tissues through the androgen receptor[1,2], which includes foreskin development in male penile development. In addition to androgen action on the androgen receptor, Hox genes are also involved in male fibroblast development[9,10]. Environmental exposures and factors have also been shown to have dramatic impacts on the penile developmental process, independent of classic genetic mutations[5,15–17]. A number of reviews and literature have suggested a potential role for environmental epigenetics in the etiology of hypospadias[2,8,18,19], but few direct experimental studies have been performed.

Epigenetics is defined as "molecular factors and processes around DNA that regulate genome activity, independent of DNA sequence, and are mitotically stable"[6,20]. The known molecular factors and processes involved are DNA methylation, histone modifications, chromatin structure, non-coding RNA, and RNA methylation[6]. All these factors can regulate gene expression directly, independent of DNA sequence. In contrast to DNA sequence that generally cannot be modified directly by environmental exposures, epigenetic mechanisms have evolved to respond to environmental factors to regulate gene expression and phenotypic variation[6]. Epigenetics operates in all areas of biology from cell and developmental biology to evolutionary biology[20,21]. Epigenetics is the precursor to any gene expression event and integrates with genetics to regulate all biological processes[6,20,21], including reproductive development in the male and penile formation. Since epigenetics mediates environmental impacts on biology, epigenetic alterations that are associated with the etiology of hypospadias were investigated.

An additional element to environmental impacts on epigenetic regulation of biology[6], is that when the epigenetic alterations become programmed into the germline (sperm or egg) they can be transmitted to the embryo after fertilization to influence the health of the offspring and subsequent generations. This phenomenon is termed environmentally induced epigenetic transgenerational inheritance and is a non-genetic form of inheritance that can be transmitted for hundreds of generations[6,22]. The phenomenon has been shown in all organisms from plants to humans[6,20]. The transgenerational occurrence of hypospadias in humans following diethylstilbestrol (DES) exposure has been suggested to be the result of epigenetic transgenerational inheritance[2,18,19,23]. Although no direct epigenetic analyses have been reported, the DES-induced transmission and inheritance of hypospadias has been observed[8,24–26]. Therefore, the increase in hypospadias frequency in the population over the past decades through environmental impacts may not simply be direct fetal exposure, but due in part to epigenetic transgenerational inheritance from ancestral exposures to environmental contaminants[8]. Other studies have linked severity of hypospadias with maternal exposure to environmental toxicants during pregnancy[27,28].

The current study was designed to collect foreskin fibroblasts from mild (distal), moderate (mid shaft), and severe (proximal) hypospadias male patients and compare with foreskin from non-disease control population patients. The epigenetics (i.e., DNA methylation) in the foreskin cell populations were assessed and compared to identify differential DNA methylation regions (DMRs) associated with hypospadias disease etiology. The current study utilized a genome-wide DNA methylation analyses that potentially assesses greater than 95% of the genome[29]. An initial study on hypospadias previously published used an array based analysis of CpG islands that involves less than 1% of the genome, but did show differences in DNA methylation between control and hypospadias patient tissues[2]. A number of studies have demonstrated in environmentally exposed populations the presence of increased hypospadias frequency and suggested epigenetic effects, but generally did not directly assess epigenetic alterations[5,15–18,30–32]. The most common exposure correlated with hypospadias occurrence was diethylstilbestrol (DES). Various endocrine disruptors including 17α estradiol, progesterone, anti-androgenic pesticide vinclozolin, and the antihistamine loratadine at physiological levels cause hypospadias in animal models[33–37]. More recent studies have assessed epigenetic alterations at specific genomic sites[2,19,23], but not at a genome-wide level. Therefore, the current study reports genome-wide alterations in DNA methylation in foreskin tissues from mild, moderate, and severe hypospadias patients to provide new insights into the etiology of hypospadias.

Results

The objectives of the current study were to identify the DNA methylation alterations in the foreskin of newborn male children with and without hypospadias, and to compare epigenetic alterations between mild (distal), moderate (mid shaft), and severe (proximal) cases of hypospadias. The foreskin sample collections were found histologically to be greater than 90% pure fibroblast populations in the control and hypospadias tissue collected. Vascular tissue (< 5%) was the primary contaminant. Samples consisted of foreskin tissue collected during a Hypospadias surgical repair (cases) or a circumcision procedure (controls). Samples were processed and stored at − 20 °C following surgery. Prior to sample collection, approvals to conduct the study were obtained from Indiana University (IU) Institutional Review Board (IRB # 1503167555). The legal guardian of a study participant was consented, and then samples collected from Riley Hospital for Children IU Health and Franciscan Health,

Indianapolis IN. The severity was categorized as mild (distal), moderate (mid shaft), and severe (proximal) based on the urethral meatus location at the time of the procedure, as identified by the Urologist. Hypospadias samples and circumcision foreskin tissues were analyzed to assess epigenetic alterations in each group specifically for comparison. The samples were obtained from non-Hispanic white Caucasian males less than 36 months of age between 2016 and 2018 with the male karyotype being confirmed with the sequence data, Supplemental Table S1. The control (non-disease) group (n = 15) had a mean ± SD age 13 ± 5.8 months, the mild hypospadias (n = 17) had a mean age 11 ± 4.2 months, the moderate hypospadias (n = 13) had a mean age 10 ± 3.8 months, and the severe hypospadias (n = 6) had a mean age 10 ± 3.2 months, Supplemental Table S1. There was no statistical difference in the mean age between the disease and control groups (p < 0.292). The sample identification, hospital, age, sex, and disease category are presented, Supplemental Table S1. The frozen samples were shipped on ice to Washington State University and stored at − 80 °C until DNA preparation and analysis.

DNA was isolated from small < 5 mm fragments of minced foreskin samples and DNA isolated and analyzed with a methylated DNA immunoprecipitation (MeDIP) procedure to obtain methylated DNA for subsequent sequencing (Seq) for an MeDIP-Seq protocol[38], as described in the Methods. This procedure can provide a genome-wide analysis of greater than 95% of the genome[29]. Differential DNA methylation regions (DMRs) were identified by comparing the control and hypospadias case samples from each individual newborn male child in the separate groups. The DMRs were identified for each group and presented in Fig. 1. The control group was compared with each of the mild (distal), moderate (mid shaft), and severe (proximal) hypospadias groups to identify the DMRs. The DMRs at various p value statistical thresholds are presented, and p < 1e−05 (i.e., p < 0.00001) was used for all subsequent analyses. The number of DMR with adjacent DMR 1 kb windows are shown at a significance level of p < 1e−05. The majority of DMR for each group involved a single 1 kb window, with some DMRs with two adjacent 1 kb windows or more, Fig. 1. The mild hypospadias had 2725 DMRs at p < 1e−05 edgeR and a false discovery rate (FDR) p value less than 0.01, Fig. 1A. The moderate hypospadias had 95 DMRs at p < 1e−05 edgeR and a false discovery rate (FDR) p value less than 0.1, Fig. 1B. The severe hypospadias had 22 DMRs at p < 1e−05 edgeR and a false discovery rate (FDR) p value less than 0.1 for 15 DMR (Fig. 1C). A combination analysis of all mild, moderate, and severe hypospadias had 734 DMRs at p < 1e−05 edgeR and FDR p value less than 0.05 (Fig. 1D). The lists of DMRs for each group and genomic features are presented in Supplemental Table S2 for mild, Supplemental Table S3 for moderate, Supplemental Table S4 for severe, and Supplemental Table S5 for all hypospadias. These tables present for each group the DMR name, chromosomal locations, start position, DMR length in bp, log-fold change (for positive increased DNA methylation or for negative decreased DNA methylation) for each DMR, gene associations and functional categories for the associated genes, Supplemental Tables S2–S5. Approximately half of the DMRs had an increase in DNA methylation and the rest a decrease in DNA methylation. The chromosomal locations of the DMRs (red arrowheads) for each group are presented in Fig. 2. The DMRs are present on most chromosomes, with groups of DMRs throughout the genome. The black boxes indicate clusters of DMRs at particular regions (Fig. 2). No obvious common regions are observed between the groups. The genomic features of the DMRs were investigated. The CpG density of the DMRs for each group is predominantly between 1 and 3 CpG/100 bp, Supplemental Figure S1. These regions with low CpG density are considered CpG deserts[39], which represents the majority (> 90%) of the genome[29]. The sizes of the DMRs for each group are predominantly 1 or 2 kb (Supplemental Figure S1).

An overlap of the DMRs for each group demonstrated the majority of DMRs were distinct for each group except the All group, at p < 1e−05 (i.e., p < 0.00001) (Fig. 3A). As expected, the All hypospadias group had good overlap with the mild (distal) hypospadias and other groups (Fig. 3A). Further analysis of potential overlap used an extended overlap analysis with a comparison of the p < 1e−05 DMRs with the other groups at p < 0.05 threshold. The extended overlap demonstrated a much higher overlap with the All hypospadias group having greater than 90% overlap with the other groups and the mild hypospadias having 92% with moderate, 82% with severe and 77% with all (Fig. 3B). The severe (proximal) hypospadias had the lowest level of overlap with 25% with mild and 32% with all hypospadias. Therefore, a high DMR overlap between the hypospadias stages was observed with a reduced statistical threshold. A principal component analysis (PCA) of the DMR sets within each group demonstrated for all comparisons a good separation between the hypospadias and control group DMRs (Fig. 4). The DMR PCA demonstrate the control (non-disease) samples are distinct from the hypospadias samples at DMR sites.

The final analysis investigated the DMR associated genes for each of the hypospadias groups. The DMRs within 10 kb of a gene to include proximal and distal promoter regions were identified. The DMR gene association does not confirm a regulatory site for gene expression, and many distal DMRs not identified as associated with genes may also influence gene expression through ncRNA mediated mechanisms. The DMR associated genes listed in Supplemental Tables S2–S5 were analyzed for gene functional category (Fig. 5A). The transcription, signaling, and transport categories were prominent in each comparison group. The DMR associated gene groups were analyzed for pathway associations with ≥ 3 genes in the pathway for each comparison group (Fig. 5B). The mild hypospadias had the highest number of pathway associations, due to having the highest number of DMRs, followed by the all combined hypospadias, with metabolism, pathways in cancer, and neurodegeneration pathways being predominant. The moderate and severe hypospadias had too few DMR and associated genes for this analysis. The DMR associated genes in the mild hypospadias group were further analyzed with a gene network analysis with a focus on genes previously shown to be involved in hypospadias (Fig. 6). The moderate and severe hypospadias groups had too few genes for this analysis, but both did show links with CYP11A1 and CYP1A1 involved in steroidogenesis of androgen. The identified DMR associated genes were correlated with major diseases, including hypospadias, and cellular processes, using the Pathway Studio (Elsevier Inc.) analysis described in the Methods. Hypospadias, severe hypospadias, posterior hypospadias, congenital malformation, androgen deficiency, and feminization were diseases over-represented in the mild group of DMR associated genes. Skin development, genital development, fibroblast differentiation, and keratinocyte proliferation are cellular processes

DMR Identifications

(A) Mild Hypospadias DMR Biomarkers

p-value	All				Multiple Window			
0.001	19255				964			
1e-04	6922				286			
1e-05	**2725**				**90**			
1e-06	1145				28			
1e-07	530				16			

Number of Significant Windows	1	2	3	4	5	6	8	9	13
Number of DMR	2635	75	3	6	1	1	1	2	1

(B) Moderate Hypospadias DMR Biomarkers

p-value	All	Multiple Window
0.001	1658	30
1e-04	346	5
1e-05	**95**	**3**
1e-06	37	1
1e-07	18	1

Number of Significant Windows	1	3	5
Number of DMR	92	2	1

(C) Severe Hypospadias DMR Biomarkers

p-value	All	Multiple Window
0.001	583	3
1e-04	83	0
1e-05	**22**	**0**
1e-06	6	0
1e-07	4	0

Number of Significant Windows	1
Number of DMR	22

(D) Mild, Moderate, and Severe Hypospadias DMR Biomarkers

p-value	All				Multiple Window		
0.001	9556				210		
1e-04	2374				61		
1e-05	**734**				**30**		
1e-06	313				16		
1e-07	160				13		

Number of Significant Windows	1	2	3	4	5	7	8
Number of DMR	704	19	3	3	3	1	1

Figure 1. Hypospadias DMR. DMR identification and numbers. The number of DMRs found using different p value cutoff thresholds. The All Window column shows all DMRs. The Multiple Window column shows the number of DMRs containing at least two significant windows (1000 bp each). The number of DMRs with the number of significant windows (1000 bp per window) at a p value threshold of p < 1e−05 for DMR is presented. (**A**) Mild (distal) hypospadias DMRs. (**B**) Moderate (mid shaft) hypospadias DMRs. (**C**) Severe (proximal) hypospadias DMRs. (**D**) Combined mild, moderate, and severe hypospadias DMRs.

that had correlations with the hypospadias-disease-connected DMR associated genes (Fig. 6). Previously identified hypospadias associated genes are present, including GLI3, CYP11A1, CYP1A1, EGF, TGFBR3, FGFR2, RYR1, INSL3, ADAT3, ARNT2, SNAP29, CYPIA2, DGKK[9,10]. In addition to the previous known hypospadias associated disease categories, a large number of hypospadias DMR associated genes identified were correlated with the cellular processes presented (Fig. 6). The DMR associated genes and names in Fig. 6 are presented in Supplemental Table S6. These DMR associated genes have not been directly associated to hypospadias previously, but require consideration due to the current study observations.
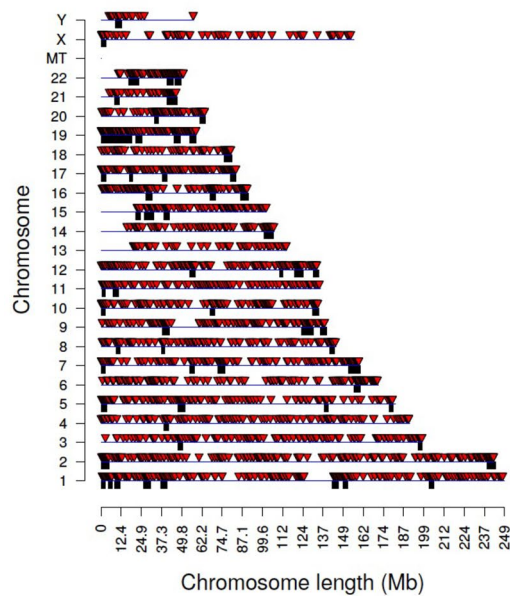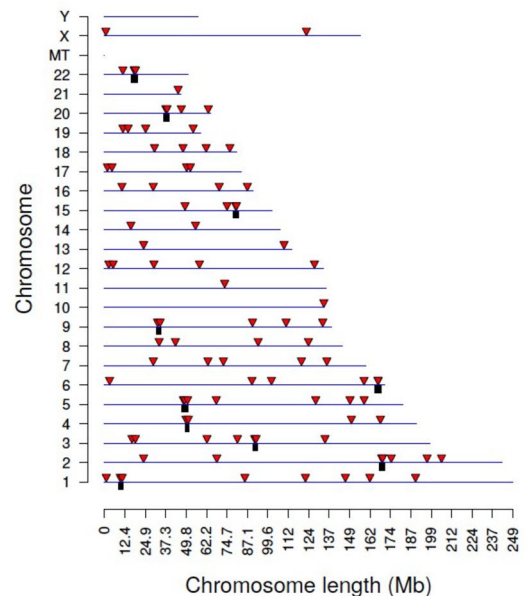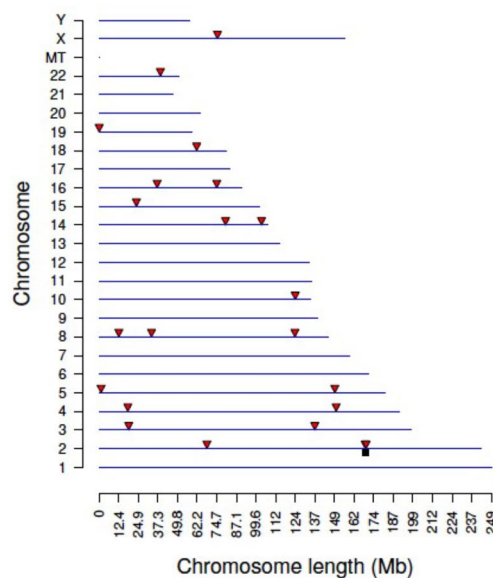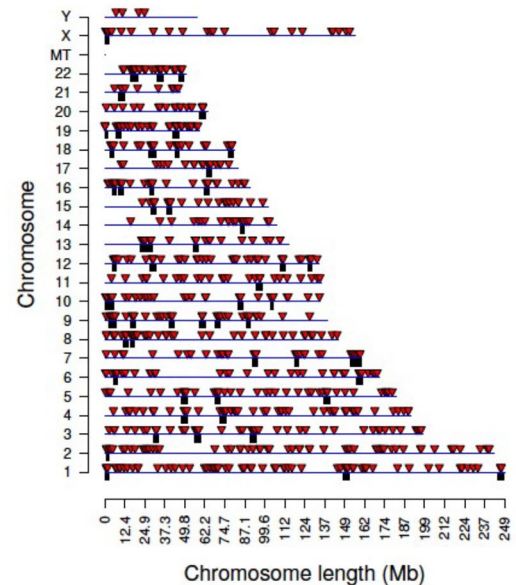
Chromosomal Locations

Figure 2. DMR chromosomal locations. The DMR locations on the individual chromosomes is represented with a red arrowhead and a cluster of DMRs with a black box. All DMRs containing at least one significant window at a p value threshold of 1e−05 for DMR are shown. (**A**) Mild (distal) hypospadias. (**B**) Moderate (mid shaft) hypospadias. (**C**) Severe (proximal) hypospadias. (**D**) Combined mild, moderate, and severe hypospadias.

Discussion

Hypospadias is the most frequent malformation of the penis and the second most common birth defect in male newborns after cryptorchidism[40], which has increased in frequency over the past decades by 11.5%[4]. This corresponds with an increase in all chronic disease, which now impacts over 80% of the human population worldwide[41,42]. As with many diseases, including hypospadias, environmental exposures are thought to be the primary source for this increase in chronic disease[41]. Although genetics and genetic mutations have been speculated to be the primary molecular basis for hypospadias[9,10], as with most other disease, environmental factors seldom directly induce genetic mutations, so fall short of explaining the molecular basis of disease etiology.

Hypospadias DMR Overlap

(A) Overlap p<1e-05
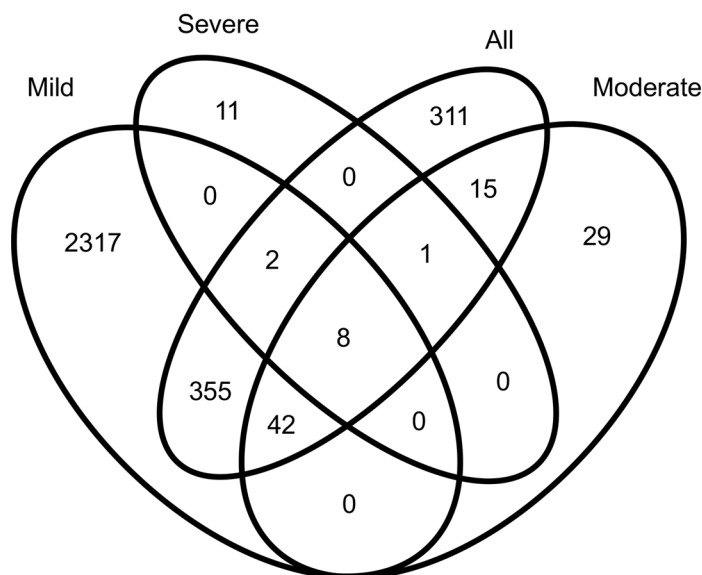
(B) Extended Overlap

p<1e-05 〱 p<0.05	Mild vs. Control	Moderate vs. Control	Severe vs. Control	All vs. Control
Mild vs. Control	2725 (100%)	1360 (50%)	668 (25%)	2689 (99%)
Moderate vs. Control	87 (92%)	95 (100%)	41 (43%)	94 (99%)
Severe vs. Control	18 (82%)	15 (68%)	22 (100%)	20 (91%)
All vs. Control	565 (77%)	464 (63%)	238 (32%)	734 (100%)

Figure 3. Hypospadias group DMR overlap. (**A**) Venn diagram. (**B**) Extended overlap p < 1e−05 compared to p < 0.05 comparison. The number of DMRs overlapped and percentage presented.

Hypospadias has been shown to be impacted by a number of environmental toxicants and factors, such as DES, that increase the frequency of the disease in the human population[5,15–17,25,26,30–32]. Factors such as DES do not have the capacity to impact DNA sequence (i.e., genetic mutations), but have been shown to promote epigenetic alterations[43,44]. Previous reports have speculated on the role of epigenetics in hypospadias, but few studies[2] have directly assessed epigenetic alterations in tissues related to hypospadias. As with many pathologies, various degrees of hypospadias [i.e., mild (distal), moderate (mid shaft), severe (proximal)] reflect the magnitude of abnormal or disrupted development of the tissue. Little molecular information has been reported on the etiology of these different stages of the disease. Clinical treatments of the various severities of hypospadias are different, and the abnormal tissue development distinct[45,46]. The current study was designed to investigate the role of environmental epigenetics in hypospadias etiology, and to compare the mild (distal), moderate (mid shaft), and severe (proximal) hypospadias with regards to epigenetic alterations.
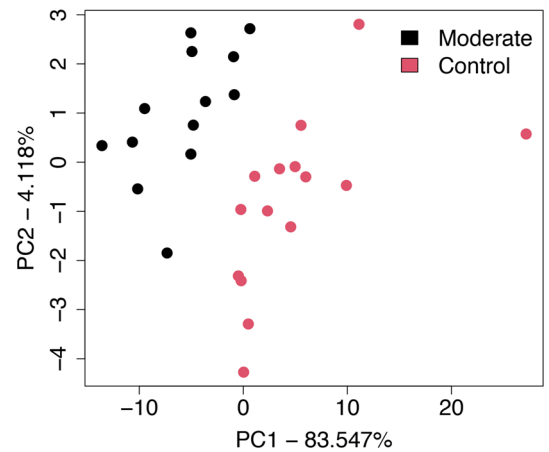
Abnormal penile development can result in hypospadias and involves the androgen responsive development of fibroblasts and foreskin[47]. Therefore, hypospadias etiology involves molecular alterations in the fibroblast of the foreskin to, in part, generate the hypospadias. Although genetic mutations have been shown to be involved, genome-wide association studies (GWAS) suggest less than ten percent of hypospadias patients have associated genetic mutations[11]. Transcriptome analyses of non-diseased and hypospadias foreskin have been reported and support the role of altered gene expression having a role in hypospadias[48,49]. The question is what regulates the altered transcriptome during the development of hypospadias. Epigenetics is now known to be the primary regulator of gene expression events, so is the precursor molecular process that promotes the gene transcription. The current study identified DNA methylation alterations between hypospadias and non-disease penile fibro-blast and foreskin samples (Fig. 1). Interestingly, the mild (distal) hypospadias cases had the highest number of DMRs, followed by fewer DMR in the moderate hypospadias penile fibroblast, and a dramatically reduced number of DMRs in the severe (proximal) hypospadias. This contrasts with the intuitive expectation that the severe hypospadias would have the most dramatic impacts. In considering the current observations, several
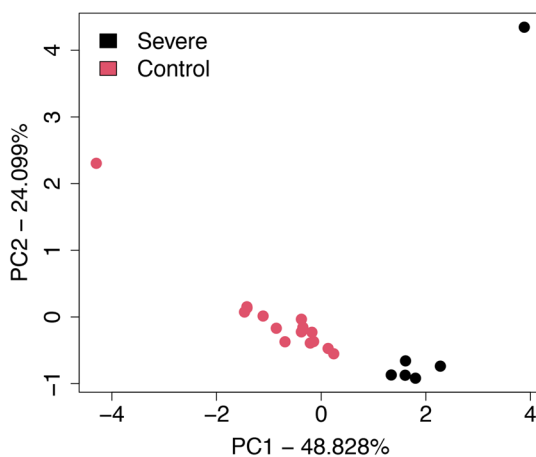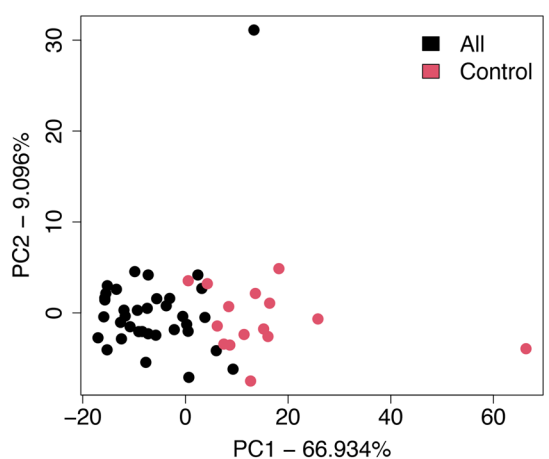
Figure 4. DMR principal component analysis (PCA). The first two principal components used. The underlying data is the RPKM read depth for DMR genomic windows. (**A**) Mild (distal) hypospadias, (**B**) Moderate hypospadias (moderate), (**C**) Severe (proximal) hypospadias, and (**D**) Combined mild, moderate, and severe hypospadias.

potential explanations can be made. The first is that the foreskin in cases of mild hypospadias contain cells that have been impacted by the developmental etiology of disease, so those cells are still present. In contrast, with the severe hypospadias those affected fibroblast populations are removed (i.e., agenesis), and those that remain are similar to the mature fibroblasts in the non-disease foreskin. Therefore, a transient transmission of the fibroblasts leads to a dramatic alteration in the number of DMRs and transcriptome for the mild hypospadias group, but the progression to severe removes this cell population, and the remaining fibroblasts appear similar to the surrounding shaft fibroblasts. The moderate (mid shaft) hypospadias cells are intermediate with some of the epigenetic altered fibroblasts remaining. This hypothesis, developed from the current study observations, is supported by the overlap of the DMRs between the mild, moderate, and severe hypospadias tissue. This suggests a loss (atresia) of foreskin fibroblasts with large numbers of DMRs in the severe (proximal) hypospadias compared to the mild hypospadias. An alternate interpretation is that the sample size for the severe hypospadias was less, so a lower statistical significance was observed. However, the combined mild, moderate, and severe hypospadias foreskin provided a significant number of DMRs that showed a very high percentage overlap in DMRs between the individual stages (Fig. 4). Therefore, the suggestion that severe (proximal) hypospadias may have a reduced number of the affected cell populations is a viable hypothesis proposed for further investigations.

The genomic features of the foreskin hypospadias fibroblasts were investigated. The chromosomal locations of the DMRs were throughout the genome with no major location over-represented. The DMRs were in low CpG density deserts[39] and 1 or 2 kb in size. There was generally a good separation of principal components in the PCA analysis between the hypospadias and non-disease DMRs. This demonstrates the hypospadias DMR

DMR Gene Associations

(A) DMR Associated Gene Functional Category
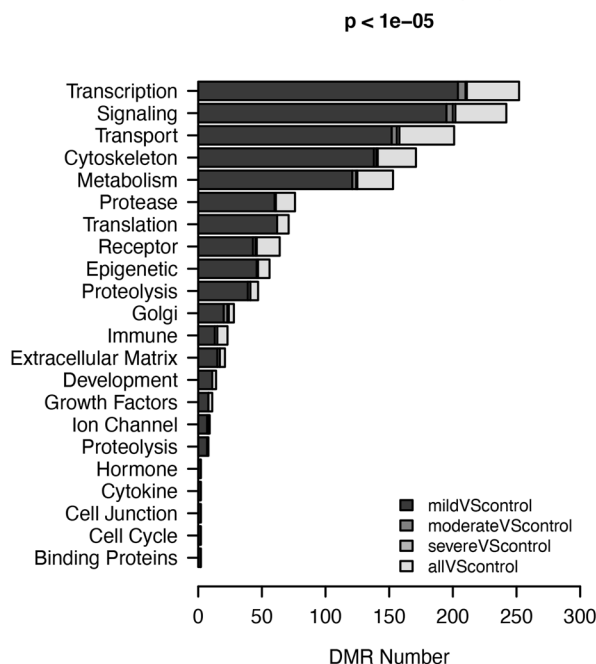
(B) DMR Associated Gene Pathways

Mild
hsa01100 Metabolic pathways (128)
hsa05200 Pathways in cancer (53)
hsa05022 Pathways of neurodegeneration (48)
hsa05014 Amyotrophic lateral sclerosis (35)
hsa04151 PI3K-Akt signaling pathway (35)
hsa05168 Herpes simplex virus 1 infection (35)
hsa05010 Alzheimer disease (33)
hsa04144 Endocytosis (31)
hsa05165 Human papillomavirus infection (31)
hsa05016 Huntington disease (29)
hsa04080 Neuroactive ligand-receptor (29)
hsa04024 cAMP signaling pathway (28)
hsa04020 Calcium signaling pathway (27)
hsa04010 MAPK signaling pathway (27)
hsa05017 Spinocerebellar ataxia (26)

Moderate
hsa05207 Carcinogenesis receptor (3)
hsa04080 Neuroactive ligand-receptor (3)

Severe
None

All
hsa01100 Metabolic pathways (31)
hsa05168 Herpes simplex virus 1 infection (15)
hsa05200 Pathways in cancer (14)
hsa04080 Neuroactive ligand-receptor (10)
hsa05022 Pathways of neurodegeneration –
          multiple diseases (9)
hsa04390 Hippo signaling pathway (9)
hsa04144 Endocytosis (8)
hsa05016 Huntington disease (8)
hsa04724 Glutamatergic synapse (7)
hsa04151 PI3K-Akt signaling pathway (7)
hsa05165 Human papillomavirus infection (7)
hsa04713 Circadian entrainment (6)
hsa05014 Amyotrophic lateral sclerosis (6)
hsa04024 cAMP signaling pathway (6)
hsa04060 Cytokine-cytokine receptor (6)

Figure 5. DMR gene associations. (**A**) the DMR associated gene functional categories are presented for mild, moderate, severe and combined all groups as indicated by the inset. The gene category and DMR number listed. (**B**) DMR associated gene pathways (KEGG) with KEGG identifier and number of pathway genes in brackets.

genomic features are distinct from those in the non-disease samples. The final analysis investigated the DMR gene associations and the disease relationships (Figs. 5, 6). Although the DMRs cannot be definitively correlated with gene expression, the genomic locations of the DMRs can be associated with genes in a manner similar to GWAS genetic mutations. The DMRs within 10 kb of a gene were associated with a gene to include proximal and distal promoter gene considerations. The DMR-associated gene categories were predominant in transcription and signaling, which are larger gene families in the genome and involved in molecular control of genome activity and cellular function. The associated gene pathways suggest metabolism and signaling. A number of the DMR-associated genes also demonstrate correlations with hypospadias disease and cellular processes previously identified (Fig. 6). Therefore, the experimental design of the current study and epigenetic analysis was partially
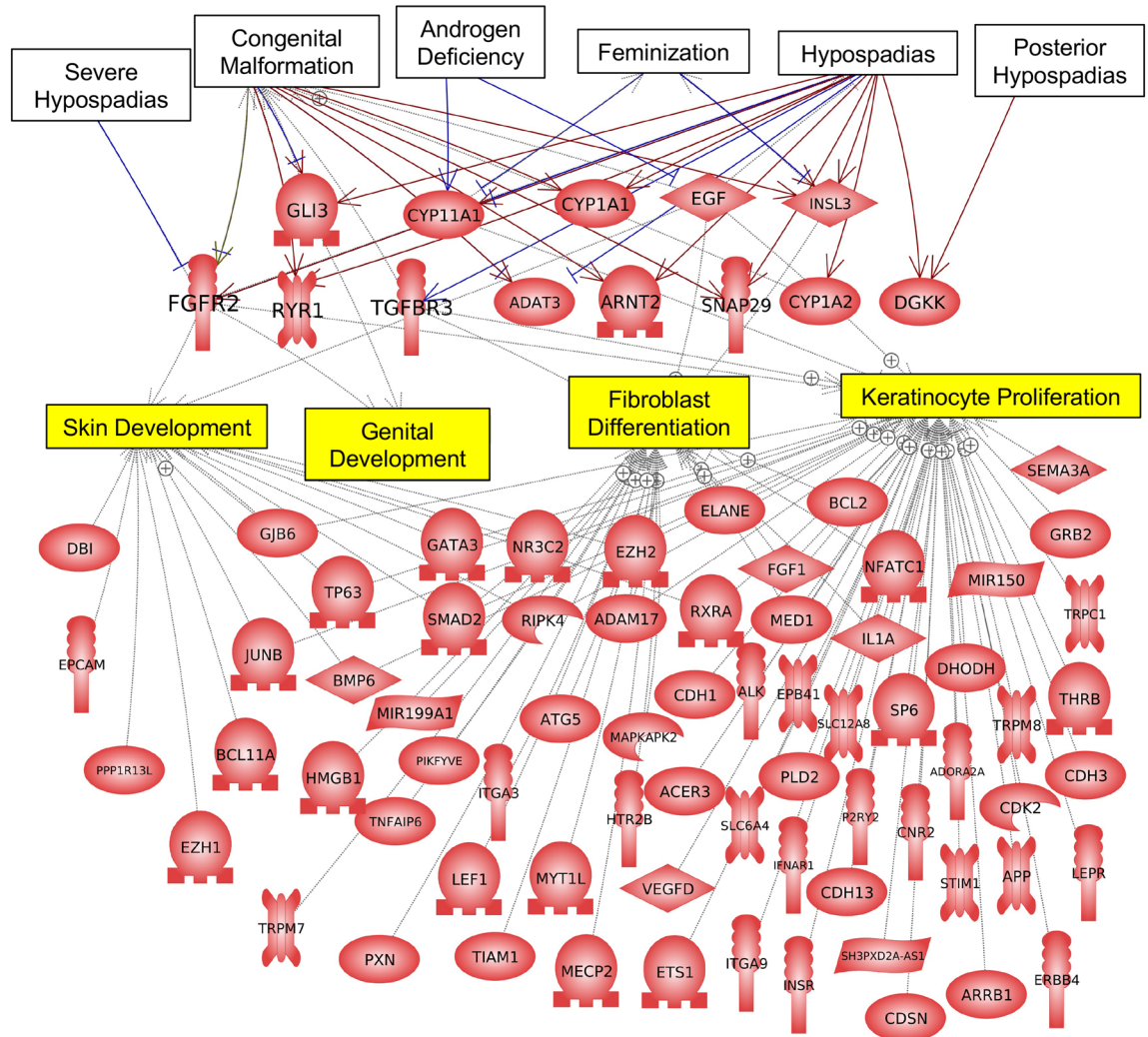
Figure 6. DMR associated gene networks. The hypospadias disease categories (white box) and related categories presented are correlated with DMR associated genes. The disease associated cellular processes (yellow box) are correlated with the known hypospadias associated genes, as well as DMR associated genes correlated to the cellular processes, but not known to previously be associated with hypospadias (lower set genes). Gene symbols and correlated names are presented in Supplemental Table S6.

validated through the GWAS and transcriptome identification of previously known hypospadias associated genes and processes. Novel observations identify a large number (i.e., 66) of new DMR-associated genes and three non-coding RNA potentially associated in hypospadias, which included DBI, EPCAM, PPP1R13L, EZH1, BCL11A, JUNB, GJB6, TP63, BMP6, HMGB1, TRPM7, GATA3, SMAD2, MIR199A1, PIKFYVE, TNFAIP6, PXN, NR3C2, RIPK4, ATG5, ITGA3, LEF1, TIAM1, MYT1L, EZH2, ADAM17, MAPKAPK2, HTR2B, VEGFD, MECP2, ETS1, ACER3, CDH1, RXRA, ELANE, ALK, SLC6A4, ITGA9, INSR, IFNAR1, PLD2, EPB41, MED1, FGF1, BCL2, NFATC1, IL1A, SLC12A8, P2RY2, CDH13, SH3PXD2A-AS1, CDSN, CNR2, SP6, SEMA3A, GRB2, MIR150, TRPC1, DHODH, TRPM8, THRB, ADORA2A, CDK2, CDH3, STIM1, APP, LEPR, ARRB1, ERBB4 (Fig. 6). The gene symbols and correlated gene names are presented in Supplemental Table S6. Novel hypospadias associated pathways and cellular processes were also identified (Figs. 5, 6). Therefore, in addition to the insights into the presence and role of epigenetics in hypospadias etiology, a large number of potential novel associated genes and cellular processes are also provided and need to be considered in hypospadias development.

The current study demonstrates a role for epigenetics in hypospadias etiology. Direct environmental exposures, such as DES, have been shown to alter epigenetics and critical gene expression events to generate the etiology of disease, both in the individual exposed and in subsequent generations[6]. Epigenetics has evolved to be responsive to the environment and allow phenotypic variation and adaptation to facilitate natural selection

and evolution[21]. When environmental exposures are at critical developmental periods, they can reprogram the normal cellular differentiation to promote a susceptibility for disease later in life, such as hypospadias in the male fetus and newborn. In addition, direct exposure impacts on epigenetics in the germline (sperm or egg) can transmit epigenetic and transcriptome alterations to the subsequent generation embryonic stem cell that then alters the epigenome and transcriptome of all derived somatic cells in the individual[6]. This results in increased disease susceptibility later in life, or in the case of hypospadias, fetal penile development and newborn hypospadias. This generational phenomenon is termed "environmentally induced epigenetic transgenerational inheritance" and is a non-genetic form of inheritance that impacts the offspring phenotypes and health for generations to come[6]. In plants, worms, and flies, the phenomenon has been transmitted hundreds of generations, and in mammals minimally five generations[6]. This environmentally induced epigenetic transgenerational inheritance phenomenon now needs to be incorporated into studies of disease etiology and provides an explanation for the dramatic increase in chronic disease incidence in the human population[6]. Although epigenetic changes linked to aging have been observed[50,51], this is not relevant to the current study as fetal development and early life periods like hypospadias will not have such aging impacts. In considering the etiology of hypospadias, the direct and transgenerational actions of the environment on hypospadias is required. Previous DES studies have suggested the potential epigenetic transgenerational inheritance of hypospadias[25,26]. Therefore, the parental and ancestral impacts on hypospadias etiology need to be investigated. Recently, transgenerational epigenetics of parental or ancestral exposures have been shown to be involved in parental transmitted autism and male infertility etiology[52,53]. Similar studies with hypospadias are required in the future and will provide further insights into environmentally induced hypospadias etiology.

Combined observations identify a better understanding of the etiology of hypospadias on a molecular level so as to integrate environmental impacts and clarify the severity of hypospadias progression. The following components of hypospadias etiology need to be considered: (1) Environmental factors are critical in the induction of hypospadias and will act through epigenetic alterations in the penile tissues during development; (2) although genetic factors will be important, environmental factors generally cannot directly alter DNA sequence and require a precursor epigenetic modification to influence the transcriptome; (3) the differential DNA methylation regions (DMRs) that develop alter gene expression events that associate with hypospadias etiology; (4) DMR associated genes and cellular processes previously associated with hypospadias etiology were identified, as well as an expanded number of potential novel DMR associated genes were identified that need to be integrated into hypospadias etiology; and (5) The progression of hypospadias from mild (distal), moderate (mid shaft), and severe (proximal) appears to involve a reduction in detected DNA methylation alterations that suggest the more severe forms may have lost (i.e., agenesis) the foreskin fibroblast and epithelial cells involved in the mild hypospadias stage. More thorough epigenetic analysis of hypospadias at various developmental stages, in response to known environmental exposures, and generational impacts will allow further advances in understanding the etiology of hypospadias.

Methods

Clinical sample collection and analysis.
Indiana University (IU) Health Hospitals Riley Hospital for Children and Franciscan Health, Indianapolis, Indiana, USA provided samples for the current study. Informed consent and HIPAA authorization were obtained from the legal guardians of all minor age participants prior to the clinical sample collection. The study protocol was approved by the Indiana University Institutional Review Board (IRB) #1503167555. All research was performed in accordance with relevant guidelines and regulations. Following IRB approval foreskin samples were prospectively collected from 51 non-Hispanic white males less than 20 months of age who were undergoing a surgical repair for hypospadias (cases) or simple circumcision (controls). Of the 51 participants, 36 (70.6%) were cases and 15 (29.4%) were controls. Controls were mean age matched to cases with no statistical difference. The mean ± SD age for cases was 11 ± 3.8 months and for controls 13 ± 5.8 months was not significant (p < 0.068). Hypospadias cases were categorized by levels of severity by the Urologist in accordance with the medical and surgical procedure codes: n = 17 (47.2%) mild (distal), n = 13 (36.1%) moderate (mid shaft), n = 6 (16.7%) severe (proximal). The demographic data for these subjects is presented in Supplemental S1. Foreskin tissues were after collection stored at − 20 °C until shipment and then stored at − 80 °C.

DNA preparation.
Frozen human foreskin tissue samples were stored at − 80 °C and thawed for analysis. Genomic DNA from foreskin tissue samples was prepared as follows: the foreskin tissue was suspended in 750 µl of cell lysis solution and 3.5 µl of Proteinase K (20 mg/ml). This suspension was incubated at 55 °C for 3 h, then vortexed and centrifuged briefly. The lysis solution was then transferred to a new 1.5 µl microcentrifuge tube. The microcentrifuge tube with the foreskin tissue was centrifuged again to retain any remaining solution which was combined with the transferred lysis solution. The foreskin tissue had 300 µl of protein precipitation solution (Promega, A795A, Madison, WI) added to the lysis solution. The sample was incubated on ice for 15 min, then centrifuged at 4 °C for 30 min. The supernatant was transferred to a fresh 2 ml microcentrifuge tube and 1000 µl ice cold isopropanol was added along with 2 µl glycoblue. This suspension was mixed thoroughly and incubated at − 20 °C overnight. The suspension was then centrifuged at 4 °C for 20 min, the supernatant was discarded, and the pellet was washed with 75% ethanol, then air-dried and resuspended in 100 µl H_2O. DNA concentration was measured using the Nanodrop (Thermo Fisher, Waltham, MA).

Methylated DNA immunoprecipitation (MeDIP).
Methylated DNA Immunoprecipitation (MeDIP) with genomic DNA was performed as follows: individual DNA samples (2–4 µg of total DNA) were diluted to 130 µl with 1× Tris–EDTA (TE, 10 mM Tris, 1 mM EDTA) and sonicated with the Covaris M220 using the

300 bp setting. Fragment size was verified on a 2% E-gel agarose gel. The sonicated DNA was transferred from the Covaris tube to a 1.7 ml microfuge tube, and the volume was measured. The sonicated DNA was then diluted with TE buffer (10 mM Tris HCl, pH7.5; 1 mM EDTA) to 400 µl, heat-denatured for 10 min at 95 °C, then immediately cooled on ice for 10 min. Then 100 µl of 5× IP buffer and 5 µg of antibody (monoclonal mouse anti 5-methyl cytidine; Diagenode #C15200006) were added to the denatured sonicated DNA. The DNA-antibody mixture was incubated overnight on a rotator at 4 °C. The following day magnetic beads (Dynabeads M-280 Sheep anti-Mouse IgG; 11201D) were pre-washed as follows: the beads were resuspended in the vial, then the appropriate volume (50 µl per sample) was transferred to a microfuge tube. The same volume of Washing Buffer (at least 1 ml 1× PBS with 0.1% BSA and 2 mM EDTA) was added and the bead sample was resuspended. The tube was then placed into a magnetic rack for 1–2 min and the supernatant was discarded. The tube was removed from the magnetic rack and the beads were washed once. The washed beads were resuspended in the same volume of 1X IP buffer (50 mM sodium phosphate ph 7.0, 700 mM NaCl, 0.25% Triton X-100) as the initial volume of beads. 50 µl of beads were added to the 500 µl of DNA-antibody mixture from the overnight incubation, then incubated for 2 h on a rotator at 4 °C. After the incubation, the bead-antibody-DNA complex was washed three times with 1X IP buffer as follows: the tube was placed into a magnetic rack for 1–2 min and the supernatant was discarded, then the magnetic bead antibody pellet was washed with 1X IP buffer 3 times. The washed bead antibody DNA pellet was then resuspended in 250 µl digestion buffer with 3.5 µl Proteinase K (20 mg/ml). The sample was incubated for 2–3 h on a rotator at 55 °C, then 250 µl of buffered phenol–chloroform-isoamylalcohol solution was added to the sample, and the tube was vortexed for 30 s and then centrifuged at 14,000×g for 5 min at room temperature. The aqueous supernatant was carefully removed and transferred to a fresh microfuge tube. Then 250 µl chloroform were added to the supernatant from the previous step, vortexed for 30 s and centrifuged at 14,000×g for 5 min at room temperature. The aqueous supernatant was removed and transferred to a fresh microfuge tube. To the supernatant 2 µl of glycoblue (20 mg/ml), 20 µl of 5 M NaCl and 500 µl ethanol were added and mixed well, then precipitated in − 20 °C freezer for 1 h to overnight. The precipitate was centrifuged at 14,000×g for 20 min at 4 °C and the supernatant was removed, while not disturbing the pellet. The pellet was washed with 500 µl cold 70% ethanol in − 20 °C freezer for 15 min then centrifuged again at 14,000×g for 5 min at 4 °C and the supernatant was discarded. The tube was spun again briefly to collect residual ethanol to the bottom of the tube and as much liquid as possible was removed with gel loading tip. The pellet was air-dried at room temperature until it looked dry (about 5 min) then resuspended in 20 µl H_2O or TE buffer. DNA concentration was measured in Qubit (Life Technologies) with ssDNA kit (Molecular Probes Q10212).

MeDIP-Seq analysis. The MeDIP DNA samples (50 ng of each) were used to create libraries for next generation sequencing (NGS) using the NEBNext Ultra RNA Library Prep Kit for Illumina (San Diego, CA) starting at step 1.4 of the manufacturer's protocol to generate double stranded DNA. After this step the manufacturer's protocol was followed. Each sample received a separate index primer. NGS was performed at WSU Spokane Genomics Core using the Illumina HiSeq 2500 with a PE50 application, with a read size of approximately 50 bp and approximately 10–32 million reads per sample, and 10–11 sample libraries each were run in one lane.

Molecular bioinformatics and statistics. Basic read quality was verified using information produced by the FastQC program[54]. Reads were filtered and trimmed to remove low quality base pairs using Trimmomatic[55]. The reads for each sample were mapped to the GRCh38 human genome using Bowtie2[56] with default parameter options. The mapped read files were then converted to sorted BAM files using SAMtools[57]. To identify DMR, the reference genome was broken into 1000 bp windows. Duplicate reads were considered PCR artifacts and removed. The MEDIPS R package[58] was used to calculate differential coverage between control and exposure sample groups using TMM normalization. Only genomic windows with an average read depth across all samples of at least 10 reads were analyzed for differential coverage. The edgeR p value (the exact p value for differential expression based on the negative binomial mode)[59] was used to determine the relative difference between the two groups for each genomic window. FDR adjusted p values were also calculated. Windows with an edgeR p value less than 10^{-5} (i.e., p < 1e−05 or p < 0.00001) were considered DMRs. The DMR edges were extended until no genomic window with an edgeR p value less than 0.1 remained within 1000 bp of the DMR. CpG density and other information was then calculated for the DMR based on the reference genome. DMR were annotated using the NCBI provided annotations. The genes that overlapped with DMR were then input into the KEGG pathway search[60,61] to identify associated pathways. The DMR associated genes were then sorted into functional groups by reducing Panther[62] protein classifications into more general categories. All MeDIP-Seq genomic data obtained in the current study have been deposited in the NCBI public GEO database (GEO #: GSE200681). Lists of DMR associated genes were analyzed for functional relationships using the KEGG database (https://www.genome.jp/kegg/pathway.html) and Pathway Studio software (version 12.2.1.2: database of functional relationships and pathways of mammalian proteins; Elsevier).

Ethics approval and consent to participate. Indiana University (IU) Health Hospitals Riley Hospital for Children and Franciscan Health, Indianapolis, Indiana, USA provided samples for the current study. Informed consent and HIPAA authorization were obtained from the legal guardians of all minor age participants and approved prior to the clinical sample collection. The study protocol was approved by the Indiana University Institutional Review Board (IRB) #1503167555. All research was performed in accordance with relevant guidelines and regulations.

Data availability

References

1. Blaschko, S. D., Cunha, G. R. & Baskin, L. S. Molecular mechanisms of external genitalia development. *Differentiation* **84**, 261–268. https://doi.org/10.1016/j.diff.2012.06.003 (2012).
2. Choudhry, S. *et al.* Genome-wide DNA methylation profiling of CpG islands in hypospadias. *J. Urol.* **188**, 1450–1455. https://doi.org/10.1016/j.juro.2012.03.047 (2012).
3. Kaefer, M. Hypospadias-nature and nurture. *J. Urol.* **200**, 714–715. https://doi.org/10.1016/j.juro.2018.07.038 (2018).
4. Chen, M. J., Karaviti, L. P., Roth, D. R. & Schlomer, B. J. Birth prevalence of hypospadias and hypospadias risk factors in newborn males in the United States from 1997 to 2012. *J. Pediatr. Urol.* **14**(425), e421-e425 e427. https://doi.org/10.1016/j.jpurol.2018.08.024 (2018).
5. Skakkebaek, N. E. *et al.* Male reproductive disorders and fertility trends: Influences of environment and genetic susceptibility. *Physiol. Rev.* **96**, 55–97. https://doi.org/10.1152/physrev.00017.2015 (2016).
6. Nilsson, E., Sadler-Riggleman, I. & Skinner, M. K. Environmentally induced epigenetic transgenerational inheritance of disease. *Environ. Epigenet.* **4**(1–13), dvy016. https://doi.org/10.1093/eep/dvy016 (2018).
7. Thorup, J., Nordenskjold, A. & Hutson, J. M. Genetic and environmental origins of hypospadias. *Curr. Opin. Endocrinol. Diabetes Obes.* **21**, 227–232. https://doi.org/10.1097/MED.0000000000000063 (2014).
8. Bouty, A., Ayers, K. L., Pask, A., Heloury, Y. & Sinclair, A. H. The genetic and environmental factors underlying hypospadias. *Sexual Dev. Genet. Mol. Biol. Evol. Endocrinol. Embryol. Pathol. Sex Determin. Differ.* **9**, 239–259. https://doi.org/10.1159/000441988 (2015).
9. Pineyro-Ruiz, C., Serrano, H., Perez-Brayfield, M. R. & Jorge, J. C. New frontiers on the molecular underpinnings of hypospadias according to severity. *Arab J. Urol.* **18**, 257–266. https://doi.org/10.1080/2090598X.2020.1760589 (2020).
10. Joodi, M. *et al.* The genetic factors contributing to hypospadias and their clinical utility in its diagnosis. *J. Cell. Physiol.* **234**, 5519–5523. https://doi.org/10.1002/jcp.27350 (2019).
11. Geller, F. *et al.* Genome-wide association analyses identify variants in developmental genes associated with hypospadias. *Nat. Genet.* **46**, 957–963. https://doi.org/10.1038/ng.3063 (2014).
12. Stoll, C., Alembik, Y., Roth, M. P. & Dott, B. Genetic and environmental factors in hypospadias. *J. Med. Genet.* **27**, 559–563. https://doi.org/10.1136/jmg.27.9.559 (1990).
13. Schnack, T. H., Poulsen, G., Myrup, C., Wohlfahrt, J. & Melbye, M. Familial coaggregation of cryptorchidism, hypospadias, and testicular germ cell cancer: A nationwide cohort study. *J. Natl. Cancer Inst.* **102**, 187–192. https://doi.org/10.1093/jnci/djp457 (2010).
14. Schnack, T. H. *et al.* Familial aggregation of hypospadias: A cohort study. *Am. J. Epidemiol.* **167**, 251–256. https://doi.org/10.1093/aje/kwm317 (2008).
15. Marrocco, G. *et al.* Environmental, parental and gestational factors that influence the occurrence of hypospadias in male patients. *J. Pediatr. Urol.* **11**, 12–19. https://doi.org/10.1016/j.jpurol.2014.10.003 (2015).
16. Xing, J. S. & Bai, Z. M. Is testicular dysgenesis syndrome a genetic, endocrine, or environmental disease, or an unexplained reproductive disorder?. *Life Sci.* **194**, 120–129. https://doi.org/10.1016/j.lfs.2017.11.039 (2018).
17. Spade, D. J. *et al.* Differential response to abiraterone acetate and di-n-butyl phthalate in an androgen-sensitive human fetal testis xenograft bioassay. *Toxicol. Sci.* **138**, 148–160. https://doi.org/10.1093/toxsci/kft266 (2014).
18. Rice, W. R., Friberg, U. & Gavrilets, S. Sexually antagonistic epigenetic marks that canalize sexually dimorphic development. *Mol. Ecol.* **25**, 1812–1822. https://doi.org/10.1111/mec.13490 (2016).
19. Vottero, A. *et al.* Evidence for epigenetic abnormalities of the androgen receptor gene in foreskin from children with hypospadias. *J. Clin. Endocrinol. Metab.* **96**, E1953-1962. https://doi.org/10.1210/jc.2011-0511 (2011).
20. BenMaamar, M., Nilsson, E. E. & Skinner, M. K. Epigenetic transgenerational inheritance, gametogenesis and germline development. *Biol. Reprod.* **105**, 570–592. https://doi.org/10.1093/biolre/ioab085 (2021).
21. Skinner, M. K. Environmental epigenetics and a unified theory of the molecular aspects of evolution: A neo-Lamarckian concept that facilitates neo-Darwinian evolution. *Genome Biol. Evol.* **7**, 1296–1302. https://doi.org/10.1093/gbe/evv073 (2015).
22. Anway, M. D., Cupp, A. S., Uzumcu, M. & Skinner, M. K. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308**, 1466–1469. https://doi.org/10.1126/science.1108190 (2005).
23. Aiba, T. *et al.* Exploring disease-specific methylated CpGs in human male genital abnormalities by using methylated-site display-amplified fragment length polymorphism (MSD-AFLP). *J. Reprod. Dev.* **65**, 491–497. https://doi.org/10.1262/jrd.2019-069 (2019).
24. Skinner, M. K. Endocrine disruptors in 2015: Epigenetic transgenerational inheritance. *Nat. Rev. Endocrinol.* **12**, 68–70. https://doi.org/10.1038/nrendo.2015.206 (2016).
25. Fenichel, P., Brucker-Davis, F. & Chevalier, N. The history of Distilbene(R) (Diethylstilbestrol) told to grandchildren—the transgenerational effect. *Ann. Endocrinol. (Paris)* **76**, 253–259. https://doi.org/10.1016/j.ando.2015.03.008 (2015).
26. Diethylstilbestrol (DES): Also harms the third generation. *Prescrire Int.* **25**, 294–298 (2016).
27. Yang, Q. *et al.* Literature on the aetiology of hypospadias in the last 5 years: Molecular mechanism and environmental factors. *Andrologia* **46**, 583–591. https://doi.org/10.1111/and.12125 (2014).
28. Kitley, W. *et al.* in *Pediatric Urology Fall Congress* (Indiana University).
29. Beck, D., BenMaamar, M. & Skinner, M. K. Genome-wide CpG density and DNA methylation analysis method (MeDIP, RRBS, and WGBS) comparisons. *Epigenetics* **17**, 518–530. https://doi.org/10.1080/15592294.2021.1924970 (2022).
30. Wautier, A. *et al.* Genital tract and reproductive characteristics in daughters of women and men prenatally exposed to diethylstilbestrol (DES). *Therapie* **75**, 439–448. https://doi.org/10.1016/j.therap.2019.10.004 (2020).
31. Tournaire, M. *et al.* Birth defects in children of men exposed in utero to diethylstilbestrol (DES). *Therapie* **73**, 399–407. https://doi.org/10.1016/j.therap.2018.02.007 (2018).
32. Tournaire, M. *et al.* Adverse health effects in children of women exposed in utero to diethylstilbestrol (DES). *Therapie* **71**, 395–404. https://doi.org/10.1016/j.therap.2016.01.006 (2016).
33. Liu, B. *et al.* Estradiol upregulates activating transcription factor 3, a candidate gene in the etiology of hypospadias. *Pediatr. Dev. Pathol.* **10**, 446–454. https://doi.org/10.2350/06-04-0079.1 (2007).
34. Carmichael, S. L. *et al.* Maternal progestin intake and risk of hypospadias. *Arch. Pediatr. Adolesc. Med.* **159**, 957–962. https://doi.org/10.1001/archpedi.159.10.957 (2005).
35. Buckley, J., Willingham, E., Agras, K. & Baskin, L. S. Embryonic exposure to the fungicide vinclozolin causes virilization of females and alteration of progesterone receptor expression in vivo: An experimental study in mice. *Environ. Health Glob. Access Sci. Source* **5**, 4. https://doi.org/10.1186/1476-069X-5-4 (2006).

36. Vilela, M. L. *et al.* Endocrine disruptors and hypospadias: Role of genistein and the fungicide vinclozolin. *Urology* **70**, 618–621. https://doi.org/10.1016/j.urology.2007.05.004 (2007).
37. Willingham, E., Agras, K., Vilela, M. & Baskin, L. S. Loratadine exerts estrogen-like effects and disrupts penile development in the mouse. *J. Urol.* **175**, 723–726. https://doi.org/10.1016/S0022-5347(05)00188-6 (2006).
38. Ben Maamar, M., Sadler-Riggleman, I., Beck, D. & Skinner, M. K. Genome-wide mapping of DNA methylation 5mC by Methylated DNA immunoprecipitation (MeDIP)—sequencing. *DNA Modif. Methods Mol. Biol.* **2198**, 301–310. https://doi.org/10.1007/978-1-0716-0876-0_23 (2021).
39. Skinner, M. K. & Guerrero-Bosagna, C. Role of CpG deserts in the epigenetic transgenerational inheritance of differential DNA methylation regions. *BMC Genom.* **15**, 692 (2014).
40. Sagodi, L., Kiss, A., Kiss-Toth, E. & Barkai, L. Prevalence and possible causes of hypospadias. *Orv. Hetil.* **155**, 978–985. https://doi.org/10.1556/OH.2014.29858 (2014).
41. Collaborators, G. B. D. R. F *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 2287–2323. https://doi.org/10.1016/S0140-6736(15)00128-2 (2015).
42. Rosenthal, M. D. *et al.* Chronic critical illness: Application of what we know. *Nutr. Clin. Pract.* **33**, 39–45. https://doi.org/10.1002/ncp.10024 (2018).
43. Efimova, O. A., Koltsova, A. S., Krapivin, M. I., Tikhonov, A. V. & Pendina, A. A. Environmental epigenetics and genome flexibility: Focus on 5-hydroxymethylcytosine. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms21093223 (2020).
44. Ho, S. M. *et al.* Environmental factors, epigenetics, and developmental origin of reproductive disorders. *Reprod. Toxicol.* **68**, 85–104. https://doi.org/10.1016/j.reprotox.2016.07.011 (2017).
45. Chan, K. H. *et al.* Provider perspectives on shared decision-making regarding hypospadias surgery. *J. Pediatr. Urol.* **16**, 307–315. https://doi.org/10.1016/j.jpurol.2020.03.015 (2020).
46. Leslie, J. A. *et al.* Corporeal grafting for severe hypospadias: A single institution experience with 3 techniques. *J. Urol.* **180**, 1749–1752. https://doi.org/10.1016/j.juro.2008.03.091 (2008) (**discussion 1752**).
47. Matsushita, S. *et al.* Regulation of masculinization: Androgen signalling for external genitalia development. *Nat. Rev. Urol.* **15**, 358–368. https://doi.org/10.1038/s41585-018-0008-y (2018).
48. Richard, M. A. *et al.* Altered mechanisms of genital development identified through integration of DNA methylation and genomic measures in hypospadias. *Sci. Rep.* **10**, 12715. https://doi.org/10.1038/s41598-020-69725-1 (2020).
49. Han, X. *et al.* Di (2-ethylhexyl) phthalate-induced hypospadias in SD rats is related with Mafb expression: A transcriptome profiling-based study. *Nan Fang Yi Ke Da Xue Xue Bao* **39**, 456–463. https://doi.org/10.12122/j.issn.1673-4254.2019.04.12 (2019).
50. Goel, N., Karir, P. & Garg, V. K. Role of DNA methylation in human age prediction. *Mech. Ageing Dev.* **166**, 33–41. https://doi.org/10.1016/j.mad.2017.08.012 (2017).
51. Karir, P., Goel, N. & Garg, V. K. Human age prediction using DNA methylation and regression methods. *Int. J. Inf. Technol.* **12**, 373–381. https://doi.org/10.1007/s41870-019-00390-y (2020).
52. Garrido, N. *et al.* Sperm DNA methylation epimutation biomarker for paternal offspring autism susceptibility. *Clin. Epigenet.* **13**, 1–13. https://doi.org/10.1186/s13148-020-00995-2 (2021).
53. Luján, S. *et al.* Sperm DNA methylation epimutation biomarkers for male infertility and FSH therapeutic responsiveness. *Sci. Rep.* **9**, 16786. https://doi.org/10.1038/s41598-019-52903-1 (2019).
54. Andrews, S. *FastQC: A quality control tool for high throughput sequence data.* https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
55. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 (2014).
56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. https://doi.org/10.1038/nmeth.1923 (2012).
57. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 (2009).
58. Lienhard, M., Grimm, C., Morkel, M., Herwig, R. & Chavez, L. MEDIPS: Genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics* **30**, 284–286. https://doi.org/10.1093/bioinformatics/btt650 (2014).
59. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140. https://doi.org/10.1093/bioinformatics/btp616 (2010).
60. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
61. Kanehisa, M. *et al.* Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Res.* **42**, D199-205. https://doi.org/10.1093/nar/gkt1076 (2014).
62. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–1566. https://doi.org/10.1038/nprot.2013.092 (2013).

Acknowledgements

Author contributions

M.K. patient's recruitment, clinical and sample collection oversight, data analysis, editing manuscript. R.R. patient's recruitment, clinical and sample collection oversight, data analysis, editing manuscript. R.M. patient's recruitment, clinical and sample collection oversight, data analysis, editing manuscript. P.W. patient's recruitment, clinical and sample collection oversight, data analysis, editing manuscript. C.P. patient's recruitment, clinical and sample collection oversight, data analysis, editing manuscript. M.B.M. molecular analysis, data analysis, editing manuscript. D.B. bioinformatics, data analysis, editing manuscript. E.N. sample processing, data analysis, editing manuscript. M.K.S. conceived, data analysis, funding acquisition, wrote and edited manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-27763-5.

Correspondence and requests for materials should be addressed to M.K.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.