Meeting report on the NIDDK/AUA Workshop on Congenital Anomalies of External Genitalia: challenges and opportunities for translational research

H. Scott Stadler b,***,1, Craig A. Peters b,c,++,1, Renea M. Sturm d,1, Linda A. Baker b, Carolyn J.M. Best e, Victoria Y. Bird f,g, Frank Geller h, Deborah K. Hoshizaki i, Thomas B. Knudsen j, Jenna M. Norton k, Rodrigo L.P. Romao k, Martin J. Cohn l,*,1

Summary

Congenital anomalies of the external genitalia (CAEG) are a prevalent and serious public health concern with lifelong impacts on the urinary function, sexual health, fertility, tumor development, and psychosocial well-being of affected individuals. Complications of treatment are frequent, and data reflecting long-term outcomes in adulthood are limited. To identify a path forward to improve treatments and realize the possibility of preventing CAEG, the National Institute of Diabetes and Digestive and Kidney Diseases and the American Urological Association convened researchers from a range of disciplines to coordinate research efforts to fully understand the different etiologies of these common conditions, subsequent variation in clinical phenotypes, and best practices for long term surgical success. Meeting participants concluded that a central data hub for clinical evaluations, including collection of DNA samples from patients and their parents, and short interviews to determine familial penetrance (small pedigrees), would accelerate research in this field. Such a centralized data hub will advance efforts to develop detailed multidimensional phenotyping and will enable access to genome sequence analyses and associated metadata to define the genetic bases for these conditions. Inclusion of tissue samples and integration of clinical studies with basic research using human cells and animal models will advance efforts to identify the developmental mechanisms that are disrupted during development and will add cellular and molecular granularity to phenotyping CAEG. While the discussion focuses heavily on hypospadias, this can be seen as a potential template for other conditions in the realm of CAEG, including cryptorchidism or the exstrophy–epispadias complex. Taken together with long-term clinical follow-up, these data could inform surgical choices and improve likelihood for long-term success.

Introduction

Congenital anomalies of the external genitalia (CAEG) are a prevalent and serious public health concern with lifelong impacts on the urinary function, sexual health, fertility, tumor development, and psychosocial well-being of affected individuals. CAEG include a broad set of conditions that present in a wide spectrum, ranging from mild variants of normal to extreme anomalies that interfere with function (Table 1). Although many conditions require surgical management, complications are frequent and high-quality epidemiological data reflecting long-term outcomes in adulthood are limited. Improving treatments and realizing the possibility of preventing CAEG will require the coordinated efforts of diverse research disciplines to fully understand the different etiologies of these common conditions, subsequent variation in clinical phenotypes, and best practices for long term surgical success. The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the American Urological Association (AUA) convened researchers from a range of disciplines, including pediatric urology, developmental biology, genetics & genomics, computational systems biology, toxicology, endocrinology, and epidemiology at AUA...
Headquarters in Linthicum, Maryland in May 2018 to discuss the best way forward to address this public health concern. Meeting participants discussed research needs relating to: 1) defining the molecular determinants of normal external genital development to improve understanding of CAEG etiologies, 2) mapping molecular pathways to the cellular behaviors that drive genital development, 3) systems modeling to unravel complex tissue interactions that coordinate morphogenesis 4), improving the quality of epidemiological reporting in the era of “Big Data,” and 5) new approaches to clinical phenotyping. The workshop resulted in development of a general framework for advancing research and clinical care to improve outcomes for people with CAEG, which we describe in this article (See Box 1). The goal was to provide a broad and inclusive framework from varied perspectives, including basic scientists, epidemiologist, clinicians, environmental scientists and others. This framework is intended to serve as a template for organizing current and future knowledge to better develop understanding of these conditions in order to facilitate possible prevention and clinical care. There was broad support for future workshops that target specific needs, such as development of more granular clinical phenotyping protocols that incorporate anatomic, molecular, and genetic data.

Defining the molecular determinants of normal external genital development to better understand the causes of CAEG

During embryogenesis, external genital development involves a complex and dynamic series of cellular events, including differential cell proliferation, apoptosis, migration, and sorting [9–14]. Mechanistically, these cellular processes are controlled by tissue-specific growth factor signaling and sex hormones, which regulate the activity of target genes that orchestrate development of the genitalia in a sex-specific manner [14–18]. At present, our understanding of how these molecular signals are integrated to control the growth and differentiation of the primitive genital anlagen is incomplete, as the full complement of the signaling factors necessary for genital development remains to be elucidated. Equally important is the lack of information regarding how these molecular signals are regulated (or perturbed) by genomic or environmental factors, which, when identified, will provide mechanistic insights into the molecular pathology of CAEG.

The external genitalia arise as paired genital swellings located on each side of the embryonic cloaca [19]. During development (around embryonic day 28 in human and 10.5 in mice) these genital swellings fuse to form the genital tubercle, the structure from which the external genitalia is derived. Genital tubercle development involves some of the same processes as embryonic limb and digit development, including the specification and patterning of tissues along three spatial planes or axes, namely proximodistal (base-to-tip), anteroposterior (cranial side, which will become the dorsal aspect of the phallus, to caudal side, which will become ventral), and mediolateral (midline to left and right sides) [20,21]. To achieve this differential patterning, the endodermal urethral plate epithelium grows out from the cloaca with the genital swellings and functions as the critical organizing center of the genital tubercle. Signaling interactions between the urethral plate and adjacent mesenchyme are mediated by secreted proteins - including SHH, BMP4, BMP7, WNT5a, and FGF10 - that control many of the cellular processes required for genital development [10,13,14,22,23].

For the most part, the developmental processes regulated by these growth factors during genital development have been investigated through the use of genetically engineered mice, which determined that the formation of the urethral tube proceeds by central canalization of the urethral plate [9,24]. While anatomical differences clearly exist between the human and mouse penis, including the presence of bone (os penis or baculum) and distal fibrocartilage (occasionally referred to as the male urogenital mating protuberance) in mice, it is likely that the fundamental developmental processes and molecular signals mediating genital outgrowth and urethral tube formation, two sites commonly affected in CAEG, are generally conserved. However, comparative studies of urethral tube formation in mouse and human have identified potential differences in the genes of lumen formation; human urethral tube development appears to involve a combination of fusion and canalization processes [25,26], whereas in mice the urethral lumen seems to arise by cavitation of the bilaminar urethral plate [9,24,27]. The extent to which the underlying molecular mechanisms differ in mouse and human urethrae will require further investigation. Genomic analyses of individuals affected by CAEG have identified mutations and copy number variants in several of the developmental genes identified as essential for mouse genital development, including EPHB2, EFNB2, EYA1, FGFR2, and HOXA13, highlighting the translational
<table>
<thead>
<tr>
<th>Virilized or Masculinized Phenotype</th>
<th>Nonvirilized or Feminized Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prepuce</strong></td>
<td><strong>Clitoris</strong></td>
</tr>
<tr>
<td>Structure</td>
<td>Structure / size</td>
</tr>
<tr>
<td>Hypospadias-associated vs isolated Phimosis</td>
<td>Enlarged (CAH, androgen exposure)</td>
</tr>
<tr>
<td>Phimosis</td>
<td>Diminutive (CAIS)</td>
</tr>
<tr>
<td><strong>Penis</strong></td>
<td><strong>Vestibule/ Urethra</strong></td>
</tr>
<tr>
<td>Structure</td>
<td>Structure</td>
</tr>
<tr>
<td>Congenital megaprepuce</td>
<td>Labial fusion</td>
</tr>
<tr>
<td>Penoscrotal webbing</td>
<td>Urethral prolapse</td>
</tr>
<tr>
<td>Micropenis</td>
<td>Urethral polyps (tumor excluded)</td>
</tr>
<tr>
<td>Alignment</td>
<td>Paraurethral or Gartner’s cysts</td>
</tr>
<tr>
<td>Torsion</td>
<td>Ectopic ureterocele, ectopic urethral insertion</td>
</tr>
<tr>
<td>Chordee with hypospadias</td>
<td>See DSD: Ambiguous genitalia</td>
</tr>
<tr>
<td>Congenital chordee</td>
<td>Pure urogenital sinus anomaly</td>
</tr>
<tr>
<td><strong>External urethra</strong></td>
<td>Persistent cloacal anomalies</td>
</tr>
<tr>
<td>Structure</td>
<td><strong>Vagina</strong></td>
</tr>
<tr>
<td>Megalourethra (prune belly syndrome)</td>
<td>Structure</td>
</tr>
<tr>
<td>Anterior urethral valves/diverticula</td>
<td>Transverse or longitudinal vaginal septum</td>
</tr>
<tr>
<td>Meatal location</td>
<td>Vaginal atresia or agenesis (MRKH)</td>
</tr>
<tr>
<td>Hypospadias</td>
<td>Urogenital sinus/cloacal anomalies, aplasia, see DSD</td>
</tr>
<tr>
<td>Epispadias (with vs without exstrophy)</td>
<td>Female epispiadias</td>
</tr>
<tr>
<td>Fistula (with vs without anorectal malformation)</td>
<td>Hymenal anomaly (skin tag, imperforate)</td>
</tr>
<tr>
<td><strong>Number</strong></td>
<td><strong>Cervix/uterus</strong></td>
</tr>
<tr>
<td>Aphallia</td>
<td>Structure</td>
</tr>
<tr>
<td>Diphallia</td>
<td>Uterine anomalies of lateral or vertical fusion</td>
</tr>
<tr>
<td>Number</td>
<td>Cervical stenosis, absence or lack of patency</td>
</tr>
<tr>
<td><strong>Scrotal</strong></td>
<td>Embryonic cyst</td>
</tr>
<tr>
<td>Location/ structure</td>
<td>Congenital fistula</td>
</tr>
<tr>
<td>Penoscrotal transposition</td>
<td>Uterine or cervical agenesis/aplasia/hypoplasia</td>
</tr>
<tr>
<td>Bifid scrotum</td>
<td><strong>Number</strong></td>
</tr>
<tr>
<td>Ectopic scrotum</td>
<td><strong>Gonadal / Fallopian tubes</strong></td>
</tr>
<tr>
<td>Scrotal hypoplasia</td>
<td>Structure</td>
</tr>
<tr>
<td><strong>Number</strong></td>
<td>Duplication with or without insertion anomaly</td>
</tr>
<tr>
<td><strong>External lesions</strong></td>
<td>See DSD: includes streak gonad</td>
</tr>
<tr>
<td>Structure</td>
<td>Absence/torsion</td>
</tr>
<tr>
<td>Vascular malformations</td>
<td>Accessory ovary</td>
</tr>
<tr>
<td>Penile or urethral mass/cyst</td>
<td>Fallopian tube cyst</td>
</tr>
<tr>
<td><strong>Testis</strong></td>
<td>Embryonic cyst of the broad ligament</td>
</tr>
<tr>
<td>Location</td>
<td><strong>Number</strong></td>
</tr>
<tr>
<td>Undescended testis</td>
<td>Urogenital sinus/cloacal anomalies, (continued on next page)</td>
</tr>
<tr>
<td>Retractile testis</td>
<td><strong>Duplication anomalies of</strong></td>
</tr>
<tr>
<td>Splenogonadal fusion</td>
<td><strong>Number</strong></td>
</tr>
<tr>
<td>Transverse testicular ectopia</td>
<td><strong>Number</strong></td>
</tr>
</tbody>
</table>
relevance of developmental genetic analysis of CAEG in mice [11, 28–31].

What is lacking is a molecular understanding of how these signals are regulated and integrated to facilitate the growth and closure of the external genitalia. Indeed, most CAEG have unknown etiologies with a presumed mix of monogenic and/or multifactorial processes involving both genetic and environmental factors (Table 2). Recent advances in genomic sequencing, such as whole exome sequencing, RNA-seq, single cell RNA-seq, ATAC-seq, ChIP-seq, and ChIRP-seq, will be instrumental in interrogating patient samples for the genetic and epigenetic causes underlying CAEG. Fundamental research is required to address knowledge gaps in our understanding of the earliest stages of genital development, including the mechanisms that determine the positioning, outgrowth, and sexual differentiation of the external genital

<table>
<thead>
<tr>
<th>Virilized or Masculinized Phenotype</th>
<th>Nonvirilized or Feminized Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anomalies of the vas deferens, epididymis, processus vaginalis</strong></td>
<td><strong>Structure</strong></td>
</tr>
<tr>
<td>Perinatal torsion Polorchidism See DSD: Disorders of gonadal development</td>
<td>clitoris, urethra, vagina, uterus (with vs without bladder duplication) Hernia/hydrocele Epididymal cyst Congenital absence of the vas deferens Anomalous vasal insertion</td>
</tr>
</tbody>
</table>

Table 1 (continued)

<table>
<thead>
<tr>
<th>Virilized or Masculinized Phenotype</th>
<th>Nonvirilized or Feminized Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Differences of Sexual Development</strong></td>
<td></td>
</tr>
<tr>
<td>46,XX DSD</td>
<td>46,XY DSD</td>
</tr>
<tr>
<td><strong>Disorders of Gonadal Development</strong></td>
<td><strong>Disorders of Gonadal Development</strong></td>
</tr>
<tr>
<td>Gonadal dysgenesis Testicular DSD (e.g. SRY+, dup SOX9) Ovotesticular DSD</td>
<td>Complete gonadal dysgenesis Partial gonadal dysgenesis Bilateral vanishing testis/ testicular regression syndrome Ovotesticular DSD</td>
</tr>
<tr>
<td><strong>Androgen Excess</strong></td>
<td><strong>Disorders of Androgen Synthesis or Action</strong></td>
</tr>
<tr>
<td>Fetal (congenital adrenal hyperplasia) Fetoplacental (aromatase deficiency) Maternal (ovarian, adrenal tumors, exogenous)</td>
<td>Androgen biosynthesis defect (17, 20-Lyase deficiency) Defect in androgen activity (complete or partial androgen insensitivity syndrome) LH receptor defects (Leydig cell agenesis/ unresponsiveness) Disorders of AMH and AMH receptor (persistent Müllerian duct syndrome) Disorders of testosterone metabolism by peripheral tissues (5α-reductase deficiency) Turner syndrome Ovotesticular DSD/mixed gonadal dysgenesis</td>
</tr>
<tr>
<td><strong>Other 46,XX DSD</strong></td>
<td></td>
</tr>
<tr>
<td>Cloacal extrophy Mayer–Rokitansky–Küster–Hauser syndrome</td>
<td></td>
</tr>
<tr>
<td><strong>47,XXY DSD</strong></td>
<td>Klinefelter syndrome</td>
</tr>
<tr>
<td><strong>Disorders of Gonadal Differentiation</strong></td>
<td><strong>Syndromes of gonadal dysgenesis</strong></td>
</tr>
<tr>
<td>Seminiferous tubule dysgenesis Klinefelter Syndrome 46, XX Male</td>
<td>Turner syndrome Pure gonadal dysgenesis Mixed gonadal dysgenesis Partial gonadal dysgenesis</td>
</tr>
<tr>
<td>45,X DSD 45,X/46,XY, 46,XX/46,XY DSD Unclassified</td>
<td></td>
</tr>
</tbody>
</table>

anlagen as well as the mechanisms by which environmental factors disrupt these processes. Critical morphogenetic processes needing further study include epithelial–mesenchymal transition, cell polarity, cell adhesion versus repulsive cues, cell death, mechanotransduction, and the roles of extracellular matrix, vascular endothelial cells, neuronal cells, and neural crest cells. The roles of epigenetic factors, long noncoding RNAs, and miRNAs have just begun to be interrogated in genital tubercle development [32,33], but are generally important in embryogenesis [34,35]. Thus, the quest to identify the pivotal molecular mediators needed for human embryonic external genital morphogenesis should continue, as understanding their identities and functions will be required for insight into the epigenetic, genetic, and environmental causes of human CAEG.

**Mapping molecular pathways to the cellular behaviors that drive genitourinary development**

Only the minority of hypospadias cases have an identifiable monogenic or chromosomal abnormality, although recent discoveries of variation in gene copy number [28,36] highlight the importance of distinguishing normal genetic variation from variants associated with phenotypic disorders. Challenges in identifying causes of external genital anomalies might indicate either non-genetic underpinnings, such as environmental factors, or technical limitations of screening methods. Whether primary causes are genetic, environmental, or a combination of the two, abnormal external genital development results from disruptions in the cellular and molecular processes that regulate genital morphogenesis. Indeed, animal model studies are beginning to show that mutations in developmental control genes and exposures to endocrine disrupting chemicals (EDC) during critical developmental periods can result in mis-regulation of the same genetic pathways and cellular processes [37–39]. Thus, two major challenges are (a) identifying the molecular and cellular mechanisms that control normal outgrowth, tissue patterning, and urethral tube formation, and (b) understanding how these processes are disrupted in individuals with CAEG. Knowing which genes orchestrate these processes during normal development will allow investigators to determine the range and scope of mutations and allelic differences that contribute to the commonality of these defects (see Box 1).

**Potential antiandrogenic and estrogenic actions of endocrine disruptors**

Experimental models demonstrate that several classes of chemicals ubiquitously present in the environment possess anti-androgenic and estrogenic activity in a dose-dependent, additive manner [40]. These environmental EDCs include industrial products like phthalates, bisphenol A, and PCBs, agricultural biocides such as DDT, hexachlorobenzene, and vinclozolin, as well as pharmaceutical and personal care products. Environmental EDCs have been suggested as a potential cause for an increase in the prevalence of male reproductive disorders, such as hypospadias, cryptorchidism, infertility and testicular cancer, referred to as the testicular dysgenesis syndrome hypothesis [41].

Evidence in humans is limited to pregnancy cohort studies focused on anthropometric measurements of the external genitalia (e.g., anogenital distance, penile length/width) [42–44], and small case–control studies concentrated on specific classes of compounds and occurrence of hypospadias and cryptorchidism in the offspring [45]. A recent systematic review concludes that limited epidemiological evidence exists for a small increase in the risk of male reproductive disorders after exposure to persistent environmental EDCs (e.g., pesticides) but not for rapidly metabolized chemicals (e.g. phthalates). However, the authors also conclude that future, higher-quality epidemiological studies could change the weight of the evidence in either direction [46]. Potential antiandrogenic and estrogenic action of EDCs deserves continued attention from the research community.

**Human genetic studies on hypospadias**

Epidemiological evidence for a genetic component in hypospadias was observed in a familial aggregation study of isolated hypospadias, with recurrence risk ratios of 50.8, 11.6 and 3.27 for male twin pairs and the first- and second-degree relatives of a case, respectively [47]; in additional heritability analyses on twins of known zygosity, the model with additive genetics and unique environment provided the best fit with a heritability estimate of 77% (95% CI: 57%–90%).

Human genetic studies long focused on rare mutations in candidate genes. These candidate genes were selected from multiple pathways, including SHH, WNT, BMP, FGF, androgen and estrogen signaling, and ephrin receptor and homeobox genes [48]. Identified variants are usually rare, and moderate sample sizes and sparse follow-up in independent cohorts made verifying specific hypospadias risk variants difficult.

Improvements in microarray technologies enabled more comprehensive approaches investigating common variation over the whole genome. A genome-wide association study (GWAS) with 1000 surgery-confirmed isolated cases of hypospadias and more than 5000 controls identified and replicated 18 loci at genome-wide significance, explaining 9% of hypospadias risk [29]. The identified regions were enriched for genes with key roles in embryonic development (e.g., HOXA4, IRX5, IRX6 and EYA1). Skin and the urogenital system—two physiological systems with an obvious link to hypospadias—yielded significant results in pathway analysis. Findings for the musculoskeletal system require further investigation. Analyzing the heritability explained by all the genetic information captured in the genome-wide SNP data yielded a 57% estimate. Thus, extending the GWAS by SNP-array genotyping of an additional large group of hypospadias cases has potential to identify more associated loci and further elucidate underlying mechanisms. Because the projects generating these data did not specifically target hypospadias, implementing functional data such as DNA methylation in preputial tissue...
Table 2  Diagnosis, description, incidence, and developmental/genetic causes of hypospadias.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Genital organ</th>
<th>Anatomic Variance</th>
<th>Epidemiology</th>
<th>Developmental cause(s)</th>
<th>Genetic Associations/Molecular Mediators</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypospadias</strong></td>
<td>• Preputial structure (deficient, hooded)</td>
<td>Distal (85%) vs mid to proximal forms based on location of meatus</td>
<td>0.8% of male births</td>
<td>Abnormal urethrogenesis with complex etiology involving both environmental and genetic factors</td>
<td>A minority of cases are secondary to an identified monogenic or chromosomal abnormality.</td>
</tr>
<tr>
<td></td>
<td>• Chordee (ventral penile curvature)</td>
<td>Varied classification schemes</td>
<td>• Most common male external genital birth defect</td>
<td>Computational modeling of adverse outcome pathways: urethral tube closure dependent upon molecular induction of mesenchymal proliferation and epithelial morphogenesis, linked to androgen signaling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Urethral structure and meatal location (ectopic, proximal)</td>
<td>• Forme fruste</td>
<td>• 10–15% familial</td>
<td>Data needed at cellular level regarding how molecular cues are translated into morphological anomalies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Severe hypospadias</td>
<td>• Standard hypospadias</td>
<td>• Recurrence risk ratios for male twin pairs, first- and second-degree relatives of a case: 50.8, 11.6 and 3.3 respectively</td>
<td>• Temporal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Hypospadias variations</td>
<td>• Hypospadias variations</td>
<td>• ~10% of cases explained by identified genetic variants</td>
<td>• Tissue-specific (cell–cell, cell–matrix interactions)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Potential variants:</td>
<td></td>
<td></td>
<td>• Biomechanical forces</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Megameatus intact prepuce</td>
<td></td>
<td>&gt;280 OMIM citations (<a href="http://www.omim.gov">www.omim.gov</a>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Chordee</td>
<td></td>
<td>Environmental/Parental Factors:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• DSD associated defects (e.g. micropenis, ambiguous genitalia, cryptorchidism or testicular dysgenesis)</td>
<td></td>
<td>• Chemical: pharmaceutical and environmental endocrine-disrupting chemicals, biocides</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Penoscrotal webbing or transposition, bifid scrotum</td>
<td></td>
<td>• Parental:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Placental insufficiency, maternal age, valproic acid use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• examples of factors that may increase risk</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1) Penile development/midline body patterning
- MID1; Xq22: Midline 1; X-linked Opitz Syndrome
- 22q11.2: Opitz G/BBB Autosomal Dominant Syndrome
- HOXD13; 2q31-q32: Homeobox gene D13
- SHH; 7q36: Sonic hedgehog
- BMP4; 14q22-q23: Bone morphogenetic factor 4
- BMP7; 20q13: Bone morphogenetic factor 7
- FGF8; 10q24: Fibroblast growth factor 8
- FGFR2; 10q26: Fibroblast growth factor receptor 2
- FOXA1/A2; 14q21.1/20p11.21 Forkhead Box A1/A2
- GLI3; 7p14.1: GLI family zinc finger 3

2) Testis determination
- NR5A1; 9q33.3: Steroidogenic factor-1
- SRY; Yp11.3: Sex-determining region of Y chromosome
- WT-1; 11p13: Wilms’ tumor –1
- NROB1; Xp21.3-p21.2: DAX deleted in azoosperma
- LHCGR; 2p21: Luteinizing hormone receptor

3) Androgen production/signaling
- AR; Xq11.2-q12: Androgen receptor
- SRD5A2; 2p23: 5α-reductase, type II
- HSD17B3; 9q22: 17β-hydroxysteroid dehydrogenase type 3
- HSD3B2; 1p13.1: 3β-hydroxysteroid dehydrogenase type 2
- ESR1; 6q25.1: Estrogen receptor 1
- ESR2; 14q23.2: Estrogen receptor 2
- MAMLD1; Xq28: Mastermind-like domain containing 1
- ATF3; 1q32.3: Activating transcription factor 3

4) Potential GWAS-identified:
- DGKK; Xp11.22: diacylglycerol kinase kappa
- EYA1; 8q13.3: EYA transcriptional coactivator and phosphatase 1
- HOXA4; 7p15.2: Homeobox gene A4
- IRX5; 16q12.2/IRX6; 16q12.2: Iroquois homeobox genes
- SP1/SP7; 12q13.13: Sp1/7 transcription factor

5) Key susceptibility loci to environmental exposures: the interaction between genetic and endocrine signals

6) Potential role of epigenetic mechanisms, long noncoding RNAs, miRNAs, unfolded protein response (autophagy-apoptosis), metabolome
or fetal murine gene expression during urethral development could further improve pathway analyses.

Re-thinking genetics

Beyond searching for mutations in coding and regulatory regions of developmental control genes, we must define key susceptibility loci that influence individual environmental exposure response. Differential sensitivity to anti-androgenic or estrogenic chemicals, for example, may have a genetic basis. Much of external genital development occurs before sexual differentiation of the gonads and is virtually identical in males and females. After gonadal differentiation and onset of steroidogenesis, systemically circulating sex hormones direct the genital tubercle down masculinizing or feminizing pathways, which involve strikingly different morphogenetic processes. Understanding how these pathways affect expression of genes that govern genital morphogenesis in target tissues will be essential in determining how chemicals may disrupt or mimic sex hormones, leading to CAEG. Interactions between genetic and endocrine signals in the external genitalia are largely unknown at present. Because the incidence of hypospadias appears to have outpaced the mutation rates for most protein coding genes, characterizing epigenetic mechanisms capable of influencing the expression/function of loci involved in penile development and urethral closure will be critical. Candidate epigenetic mechanisms to consider for potential roles in penile development include DNA methylation, modifications affecting chromatin condensation (such as histone acetylation/methylation), and the expression of non-coding RNAs (micro RNAs and long non-coding RNAs) that can regulate the expression of multiple protein-coding genes. Traditionally, perturbations in these epigenetic mechanisms have not been examined in clinical genetic evaluations and are likely to be identified as contributing factors in the pathology of CAEG due to their capacity to regulate the expression of susceptibility loci. Establishing the relevant linkages between genetic variation, epigenetic influences, and environmental factors will be essential for a true understanding of many of the CAEG conditions. Moreover, identification of mechanisms of these interactions will permit a functional understanding of the causes of CAEG and has the potential to generate markers of specific processes and lead to therapeutic interventions.

Focus on the cell: connecting molecules to morphological anomalies

The paucity of data at the cellular level has hampered efforts to connect genotype to phenotype in the external genitalia. Determining CAEG etiologies will require understanding of how molecular cues, whether intrinsic or environmental, are translated into morphological anomalies. Areas of exploration include the relationship of gene expression to cellular behavior in a temporal and tissue-specific manner during external genital development; the roles of cell–cell and cell–matrix communication in urethral tube development and how these processes influence cell and tissue mechanics, integrity, and morphogenesis. Understanding the causes of failed urethral tube development will require detailed investigations into the consequences of gene expression patterns on cell polarity, shape, migration, apoptosis, and proliferation. These questions present an opportunity to develop experimental tests that will define the actual mechanisms that regulate cellular processes during human genital development.

Experimental models and approaches

Animal models

Some of these challenges cannot be addressed directly in human cells or tissues and require animal models. The mouse remains a system of choice for numerous reasons, including its well-developed genetic and genomic infrastructure, advantages as a developmental system, and amenability to genetic manipulation. External genitalia of mice and humans have considerable differences in their anatomy and development, including urethral tube formation, which must be considered when extrapolating between species. Nonetheless, molecular genetic studies suggest that the developmental mechanisms are largely conserved between mice and humans.
Although rodent models will remain essential tools in studies of external genital development, additional models are needed. Historical thinking suggested that phallicus of different vertebrates were related only by common functions with no shared evolutionary history, raising questions about the wisdom of studying genital development in an organism in which the penis could be wholly unrelated to that of humans. However, recent work indicates the amniote penis evolved once, and the external genitalia of mammals, birds, and reptiles are homologous, with similarities due to common ancestry [19,49,50]. This discovery opens possibilities for use of novel, comparative models of genital development with potential advantages over current models, exemplified in recent use of bird and reptile embryos to fate map the progenitor cells of the genital tubercle [51,52], discovery of similarities between guinea pig and human urethral tube formation [53], and use of marsupials to study genital morphogenesis during their extended period of development in the pouch [54,55]. The “optimal” animal model depends on the question under study; therefore, the field must be open to use of non-traditional models that could enable new experimental approaches and insights applicable to humans.

**In vitro models**

*In vitro* models using human induced pluripotent stem cells or genital tissues will enable identification, confirmation, and integration of molecular and cellular processes required for genital development. These investigations may be enhanced by the use of three-dimensional organoids, which have revolutionized studies of brain, kidney, and heart development [56–58]. Research in external genital development has not leveraged organoid systems, which represent an untapped resource for modeling, testing, and perturbing the processes hypothesized to underlie CAEG. Additionally, by procuring human tissues from clinical settings, organ culture may enable identification of cellular processes mediating human genital development and closure (i.e., matrix deposition, cell migration, epithelial–mesenchymal transition, apoptosis). Clinical samples will also facilitate development of cell lines to establish the genetic and epigenetic hierarchies functioning in the external genitalia and provide high-throughput avenues to assess susceptibility of these hierarchies to various environmental factors. Finally, such *in vitro* approaches may allow determination of the genetic and environmental inputs necessary to engineer new tissues, providing clinically relevant alternatives to repair or replace tissues affected by congenital malformation, injury, or disease.

**Systems modeling of complex tissue interactions in silico**

Understanding how tissues and organs arise during embryogenesis is a central question in developmental biology and a challenge for predicting developmental toxicity. While alterations in key signaling pathways (e.g., SHH, WNT, TGFβ, EPH receptor tyrosine kinases, retinoic acid) can disrupt genitourinary development, the critical developmental processes causal in CAEG are poorly understood. To address this knowledge-gap, it is essential that approaches complementary to genetic and toxicological screens be used. At the forefront of these approaches is systems modeling, which allows complex defects such as CAEG to be computationally interrogated. This approach provides a mechanism to assess adverse outcomes pathways in a high throughput manner, identifying the roles of genetic mutations or putative teratogens in the pathology of CAEG [59]. Indeed, recent applications of systems modeling to kidney development identified a Turing mechanism by which GDNF, RET, and WNT11 mediate the complex tissue dynamics controlling branching morphogenesis [60,61]. Systems modeling will also facilitate analysis of how disruption of signaling events functioning simultaneously during the genital development causes CAEG, as many of these signaling events are often integrated during formation of the external genitalia [39,62,63].

For hypospadias, computational models successfully recapitulated sexually dimorphic development of the genital tubercle controlled by SHH, FGF10, and androgen pathways through modulation of stochastic cell behaviors, including differential adhesion, motility, proliferation, and apoptosis. Proper urethral tube closure in this model was shown to depend quantitatively on SHH- and FGF10-induced effects on mesenchymal proliferation and epithelial apoptosis. Moreover, the model also linked these cellular processes to androgen signaling. In the absence of androgen, genital tubercle development was feminized and with partial androgen deficiency, the model resolved with incomplete urethral tube closure, providing an *in-silico* platform for probabilistic prediction of hypospadias risk across combinations of minor perturbations to the system [64,65].

Finally, in a human cell-based “biomimetic system” designed to assess the influence of environmental factors on controlled tissue events [66], organoid survival was dependent on signaling through EGF, IGF, HGF, and FGF pathways, and organoid fusion was disrupted by inhibition of BMP signaling. Concordance between the effects of EGF, FGF, and BMP inhibitors on palatal organoid fusion and epithelial cell migration *in vitro* suggested critical dependence on epithelial morphogenesis and, potentially, an innate feedback mechanism. These findings demonstrate promising utility of integrative modeling and adverse outcome pathways to decode the “toxicological blueprint of active substances” that interact with the developing embryo [67].

**Distinguishing pattern from process: imaging developmental dynamics**

Understanding the morphogenetic processes and tissue movements involved in human and mouse urethral development has been confounded by descriptions that conflate anatomical patterns with developmental processes [26,68]. Recent advances in imaging technology enable real-time, three-dimensional imaging of genital morphogenesis [69], providing opportunities to directly observe processes previously inferred from static, two-dimensional images (e.g., histological sections) [25,26]. When applied to mouse models that allow single cell labeling and lineage tracing, such live imaging approaches can reveal the spatiotemporal processes involved in urethral tube formation.

**Developmental biomechanics**

Direct investigation of biomechanical forces speculated to act during urethral tube development are needed. For example, the ventral aspect of the urethral plate is
believed to be under tension during lumen formation and may be susceptible to rupture when epithelial mechanical integrity is compromised \[37\]. Such investigation may reveal the relationship between epithelial maturation and lumen formation and commonalities in tubulogenesis across organ systems. Application of engineering principles to genital developmental systems research will enhance understanding of genetic and environmental effects on the cellular processes that drive morphogenesis.

**Improving the quality of epidemiological reporting in the era of “Big Data”**

Recent epidemiological studies have yielded conflicting results regarding prevalence of CAEG. Conflicts may result from inconsistent recognition and classification of CAEG, due to the lack of a standardized classification system. A 30-year, population-based Danish study found hypospadias rates doubled \[70\], with the final prevalence similar to the rate seen in Nova Scotia, Canada around the same time \[71\]. However, prevalence in Nova Scotia remained stable over the same time period \[71\]. A recent meta-analysis of studies around the globe, covering over 90 million births, reported the mean prevalence of hypospadias (per 10,000 births) as 19.9 in Europe, 34.2 in North America, 5.2 in South America, 0.6—69 in Asia, 5.9 in Africa, and 17.1—34.8 in Australia \[72\]. The authors noted the extreme heterogeneity of published datasets, confounding factors, such as the absence of standardized definitions of hypospadias, conflicting results of studies from the same regions, and a complete absence of data from some regions \[72\]. Multiple reports suggest an increasing prevalence, although we lack clarity as to how increased numbers of registries, better data capture and reporting, or over reporting of mild cases may affect prevalence estimates. Inconsistencies across global hypospadias prevalence data suggest misclassification might be an issue. Hence, accurate epidemiology will require improved quality of CAEG recognition and reporting, including diagnosis, comorbidities and systematic phenotypic descriptions (see section entitled “New approaches to clinical phenotyping” below for further details).

Data sources analyzed for reporting CAEG in recent decades have been inconsistent, challenging determination of the true CAEG prevalence. Needs include: (i) registry standardization, including consistent reporting of long-term data, reduced site variation, and multi- (versus single-) institution reporting; (ii) persistent data collection over time and across geography; (iii) consistent case discovery efforts; and (iv) a global standard that aligns the full range of phenotypes with classification codes (e.g., International Classification of Diseases, 10th Revision [ICD-10]).

Strategies are in place to systematically mine existing electronic health record (EHR) databases through graphical interfaces of integrated data repository systems (e.g., Informatics for Integrating Biology and the Bedside [I2B2]). Such state and national databases will aid in identifying regional prevalence patterns and variables correlated to specific CAEG using existing cohorts identified by ICD-10 or other coding systems. Zip codes available in EHRs can facilitate geospatial analysis and mapping of CAEG to identify potential hot spots and target interventions. Machine learning algorithms \[73\] can leverage EHR data for prediction models, and may integrate genetic variables from databases with curated clinical and genetic information (e.g., BioVU). Such models may predict surgical outcomes (e.g., complication or repeat surgery risk after correction of hypospadias) based on phenotypic, genetic and epigenetic relationships. Large databases may facilitate identification of maternal exposures before and during pregnancy associated with CAEG based on demographic, geographic and temporal information. Thus, by using high performance statistical software to cluster large EHR datasets, phenotypic variations (VP) associated with CAEG could be defined as a summation of variances, including genomic (VG), environmental (VE) and genome—environmental interaction (VGE), such that $VP = VG + VE + VGE$.

![Clinical Registry](image)

**Figure 1** A model for the structure of a clinical registry that will utilize standardized criteria for phenotyping, tissue banking, analysis of tissues, integration of clinical outcomes, and data synthesis to determine risk and improve patient outcomes.
New approaches to clinical phenotyping

Enhancing understanding of CAEG will rely on robust phenotyping with precise and objective characterization of individual genital patterns, their developmental phenotype, that of their family, and genetic subtyping through individual and family genotyping and pedigree analysis. Robust description may include tissue, cellular and molecular phenotyping as well (See Table 3). Phenotypes must be based on reproducible categories that consistently apply, thus allowing for the transfer of knowledge over time. Description of genital anatomy requires objective measurable parameters that do not presuppose any particular pattern or relationship. CAEG are characterized by a wide variation in severity, yet phenotypic overlap complicates pattern recognition, as distinct characteristics seem associated with various patterns. Viewing CAEG as a mosaic of elements that contribute to the ultimate individual phenotype may improve categorization and better reflect developmental patterns.

Multiple potential methods may enable accurate and objective genital phenotyping, including standardized photographs, templated measurements, or ultrasonography. More complex methods, including physical molding and 3-D surface imaging, may prove valuable. Hypospadias description, for example, begins with meatal location, degree of ventral curvature, penoscrotal transposition, and bifid scrotum, but will need to include further parameters. Descriptive templates exist [74], but some may inappropriate categorize presumed yet unproven relationships. A single, consensus-based descriptive template—enhanced by visual imaging and recording—may enable objective comparison and categorization of patterns.

Considerations for clinical implementation

Optimal phenotypic description must be both practical and efficient to be integrated into clinical care and must be implemented in a systematic manner to enable consistent assessment of individuals as they grow, develop or receive surgical intervention. Yet, an optimal tool must balance this consistency over time with the flexibility to adapt to evolving knowledge in the field. The method must account for the sensitive nature of CAEG phenotyping assessments, including risks and demands of patient confidentiality. Optimal patient phenotyping must also enable identification of associated concurrent conditions, facilitating identification of associated developmental patterns as clues to etiology.

Development of such a rigorous system requires multicenter participation. Phenotyping in CAEG may be informed by efforts of interdisciplinary multi-institutional consortia to standardize terminology, categorize and quantitate phenotypic features and catalog clinical phenotypes outside of CAEG via EHR-based data templates that fulfill clinical and research needs. This approach permits longitudinal follow up, access to clinical records, recording of biobanking, and potential for interoperability with other health IT systems, enabling cross-site collaboration. Resulting data must be shared and accessible while remaining secure. Such database systems will require broad coordination, oversight, and maintenance.

Standardizing surgical interventions and follow-up

It is premature to consider standardized surgical interventions because optimal interventions remain incompletely defined. A future definition of best practices may be developed based on robust outcomes analysis and linkage with mechanistic understanding of clinical patterns. Templated follow-up patterns will likely be most useful to evaluate interventions, and will include structural, functional and psychological parameters.

Inherent to defining outcomes is the challenge of maintaining follow-up in a highly mobile population. Providing value to patients and families in terms of knowledge, counseling and engagement in the development of new knowledge may enable improved monitoring, and in turn, better understanding of the underlying developmental pattern and clinical outcomes. Some hospital systems have research opt-out policies wherein patients may be contacted for research unless they specifically request no contact, which has increased research participation.

Potential to glean information from tissue sampling

Despite the delay between initiation of the condition during fetal development and tissue sampling long after development is completed, sampling of genital and somatic tissue may yield clues regarding underlying pathophysiology. Tissue sampling also may provide information about specific properties, such as wound healing, hormonal responsiveness, growth regulation, and fibrosis, which might influence surgical outcomes. Developing standardized protocols for tissue sampling during genital reconstruction may be of value. Identifying flexible preservation techniques that allow for current and future methodologies is essential. Since we cannot accurately predict all future methods, optimal methods might be those that minimize tissue effects.

Effective data-gathering, recording and assessment structure within a longitudinal timeframe is essential and will be most productive within the context of a collaborative network of investigators. A key task of this collaboration would be the development and implementation of standards of participation for all investigators. The broad view of this collaboration (Fig. 1) should involve iterative processes of change based on new knowledge. Machine learning strategies may facilitate integration of clinical, genetic, and biological parameters and provide relevant translational insights. These insights may feed back into the informational loop, furthering our understanding of the developmental, biological and clinical processes underlying each phenotype and associated outcomes, thereby enabling optimized selection of treatments to unique phenotypes.

Conclusions

At the conclusion of the 2018 CAEG Workshop held at the AUA Headquarters, the authors of this report were charged with writing a synthesis of the current state of research (clinical and basic science) and the key knowledge gaps in
CAEG (See Box 1). As indicated by this report, hypospadias was a major focus of the presentations and discussions at the workshop, and we propose that this roadmap for hypospadias research also serve as a model for investigations of other genitourinary anomalies. At the clinical level, a major challenge identified by workshop participants is the need for a standardized method to phenotype CAEG patients. Standardization of clinical phenotyping should also include the creation of a phenotype-associated biobank of DNA and possibly other samples (e.g., blood) from patients and both parents. This resource should be in a searchable format, allowing for future studies aimed at identifying key loci involved in discrete malformations as well as environmental factors that may contribute to defect etiology.

At the basic science level, a need was identified for additional studies aimed at defining the gene regulatory networks that control the developmental processes required for growth, generation of specific cell and tissue types, formation of a functional urethral tube, and sexual differentiation of the external genitalia. As the repertoire of single cell genomic tools expands, it is anticipated that new insights into molecular processes that regulate the growth and differentiation of the external genitalia will be gained. Included in these novel approaches will be studies aimed at defining the role of epigenetic factors—such as non-coding RNAs, DNA and/or RNA methylation, and exogenous (e.g., environmental) regulators of gene expression—in mediating the development of the external genitalia. Finally, at the cellular level, recent advances in live cell imaging should provide the necessary approaches to determine common and species-specific cellular and molecular processes necessary and/or sufficient for the growth and differentiation of the external genitalia.

By addressing these challenges in future clinical and basic science studies, more streamlined and effective mechanisms to define the molecular basis for CAEG should be realized. More importantly, the integration of clinical and basic science approaches to studying CAEG should provide the best roadmap towards reducing and eventually preventing CAEG in the global population, affording directed therapeutic approaches to patients and families, as well as providing new insights into mechanisms underlying patient susceptibility to these defects.

Disclaimers

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of their employers or funding agencies. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

The authors served as the writing group for the NIDDK/AUA Workshop on Congenital Anomalies of the External Genitalia and gratefully acknowledge the contributions of the following workshop participants: Andrew Adey (Oregon Health and Science University), Darius Bagli (University of Toronto), Larry Baskin (University of California - San Francisco), Spyros Darmanis (Chan Zuckerberg Biohub), L. Earl Gray (U.S. Environmental Protection Agency), Dorrie Lamb (Weill Cornell Medicine), Sean Li (Boston Children’s Hospital), Liang Ma (Washington University), Andrew Pask (University of Melbourne), Eric Vilain (Children’s National Medical Center and George Washington University), and Duncan Wilcox (University of Texas Southwestern Medical Center).

References

Congenital Anomalies of External Genitalia workshop


