ORIGINAL ARTICLE

Male Endocrinology

Epididymis cell atlas in a patient with a sex development disorder and a novel *NR5A1* gene mutation

**Jian-Wu Shi[1],[*], Yi-Wen Zhou[2],[*], Yu-Fei Chen[1], Mei Ye[1], Feng Qiao[1], Jia-Wei Tian[1], Meng-Ya Zhang[1],
Hao-Cheng Lin[3], Gang-Cai Xie[1], Kin Lam Fok[4], Hui Jiang[3], Yang Liu[2], Hao Chen[1]**

This study aims to characterize the cell atlas of the epididymis derived from a 46,XY disorders of sex development (DSD) patient with a novel heterozygous mutation of the nuclear receptor subfamily 5 group A member 1 (*NR5A1*) gene. Next-generation sequencing found a heterozygous c.124C>G mutation in *NR5A1* that resulted in a p.Q42E missense mutation in the conserved DNA-binding domain of NR5A1. The patient demonstrated feminization of external genitalia and Tanner stage 1 breast development. The surgical procedure revealed a morphologically normal epididymis and vas deferens but a dysplastic testis. Microfluidic-based single-cell RNA sequencing (scRNA-seq) analysis found that the fibroblast cells were significantly increased (approximately 46.5%), whereas the number of main epididymal epithelial cells (approximately 9.2%), such as principal cells and basal cells, was dramatically decreased. Bioinformatics analysis of cell–cell communications and gene regulatory networks at the single-cell level inferred that epididymal epithelial cell loss and fibroblast occupation are associated with the epithelial-to-mesenchymal transition (EMT) process. The present study provides a cell atlas of the epididymis of a patient with 46,XY DSD and serves as an important resource for understanding the pathophysiology of DSD.

INTRODUCTION

Disorders of sex development (DSD) are characterized by defective gonadal development and/or disorganized genital.[1,2] The prevalence of genital ambiguity at birth varies from 1:1000 to 1:4500 in different studies.[3,4] On the basis of karyotype, DSD are classified into three types which include sex chromosome DSD, 46,XY DSD, and 46,XX DSD.[5] The incidence of patients with 46,XY DSD is estimated to be 1:20 000. However, conditions related to hypospadias and cryptorchidism of DSD are frequently observed, which affects 1:200 to 1:300 juveniles.[6,7] These patients present with testicular-development disorders and androgen biosynthesis or functional deficiency. Heterozygous mutations of the nuclear receptor subfamily 5 group A member 1 (*NR5A1*) lead to a prevalent 46,XY DSD with a mutation frequency of approximately 15%–20%.[8]

NR5A1 (also named steroidogenic factor 1 [SF-1]) is an orphan nuclear receptor composed of 461 amino acids.[9] During the development of male sex organs in 46,XY individuals, NR5A1 firstly appears in the urogenital ridge. The synergistic actions of NR5A1 and the sex-determining region Y (SRY) initiate the male developmental program via the upregulated SRY-box transcription factor 9 (SOX9).[10,11] In the Sertoli cells of the developing prenatal testis, NR5A1 is indispensable for the activation of the anti-Müllerian hormone (AMH) cascade that results in the Müllerian duct degeneration.[12,13] On the other hand, synthesis of steroidogenic enzymes was regulated by NR5A1 for androgen production in prenatal Leydig cells, which is necessary to masculinize the external genitalia and initiate the testicular descent.[14,15]

NR5A1 mutations are common causes of 46,XY DSD associated with various phenotypes, such as infertility, feminized or ambiguous genitalia, and gonadal dysgenesis. A number of studies have demonstrated that 46,XY DSD patients with *NR5A1* mutations commonly present with a sex reversal phenotype with female external genitalia at birth.[16–19] However, the argument on the gender of rearing for these DSD patients is still a trending topic. *NR5A1* mutation carriers could preserve fertility, suggesting that the gender rearing of these patients as a boy may represent an advantage if fertility is of the patient's concern.[20–22]

At present, the reported *NR5A1* gene mutations associated with sex reversal phenotypes include nonsense and frameshift mutations that result in truncated proteins. The missense mutations in the core DNA-binding domain (DBD; between two zinc fingers) affect DNA-binding and/or transcriptional regulator activities.[23] These mutations are implicated in sex determination and development (**Table 1**).

[1]Institute of Reproductive Medicine, Medical School of Nantong University, Nantong 226000, China; [2]Department of Plastic and Reconstructive Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200011, China; [3]Department of Urology, Peking University Third Hospital, Beijing 100191, China; [4]School of Biomedical Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China.
[*]These authors contributed equally to this work.
Correspondence: Dr. H Chen (chenhao@ntu.edu.cn) or Dr. Y Liu (drliuyang9@163.com)

Table 1: Reported clinical patients with disorders of sex determination carrying heterozygous nuclear receptor subfamily 5 group A member 1 mutations in the core DNA-binding domain (between two zinc fingers)

Study	Location of cDNA	Mutation	Karyotype	Diagnostic features	Family data
Kohler *et al.*[69] 2008	c.98G>C	p.Cys33Ser	46,XY	Clitoris hypertrophy, bilateral testes in inguen, urogenital sinus, Sertoli cells rich, spermatogonia and Leydig cell loss	Parents: normal
Wang *et al.*[24] 2018	c.104G>T	p.Gly35Val	46,XY	Small penis (1.5 cm), penoscrotal hypospadias	*De novo*
Eggers *et al.*[70] 2016	c.104G>A	p.Gly35Asp	46,XY	Immature Leydig cells	NA
Achermann *et al.*[71] 1999	c.104_105delinsAA	p.(Gly35Glu)	46,XY	Müllerian structures intact, poorly differentiated tubules, normal breast development induced by 17β-estradiol. Uterus grew and regular menstruation was induced via progestogen treatment	NA
Fabbri *et al.*[72] 2016	c.112A>T	p.Lis38*	46,XY	Short phallus (length in 2.5 cm), single penoscrotal urethral opening, right gonad replaced by fibrous tissue and left gonad was a dysgenetic testis	NA
Fabbri-Scallet *et al.*[73] 2018	c.115C>T	p.Arg39Cys	46,XY	Microphallus (1.5 cm), bilaterally palpable gonads, no Müllerian structures	Mother: carrier
Allali *et al.*[74] 2011 and Philibert *et al.*[75] 2010	c.116G>C	p.Arg39Pro	46,XY	Ambiguous genitalia, microphallus (20 mm), small testis (20 mm) in labia and vaginal rest (15 mm)	NA
Rocca *et al.*[76] 2018	c.119C>G	p.Thr40Arg	46,XY	Bilaterally palpable gonads, no Müllerian structures	NA
Eggers *et al.*[70] 2016	c. 119C>G	p.Thr40Arg	46,XY	Female characteristics, partial gonadal dysgenesis	NA
Werner *et al.*[49] 2017	c.118A>C	p.Thr40Pro	46,XY	Ambiguous genitalia, little vagina, low testosterone, no Müllerian structures	NA
Tajima *et al.*[17] 2009	c.122T>G	p.Val41Gly	46,XY	Clitoris hypertrophy, primary amenorrhea, low testosterone	Parents: normal
The present study	c.124C>G	p.Gln42Glu	46,XY	Clitoromegaly, micropenis (4 cm), no testes, minora and labia majora hypoplasia	NA
Wang *et al.*[24] 2018	c.132-134delCAA	p.N44del (het)	46,XY	Clitoromegaly, bilateral cryptorchidism, asymmetric labia majora	Father: carrier
Rocca *et al.*[76] 2018	c.140A>G	p.Tyr47Cys	46,XY	One urogenital sinus, small vagina without uterus, hypoplasia testis with Leydig-like differentiation	NA
Eggers *et al.*[70] 2016	c.140_163del	p.(Ala48_Pro55del)	46,XY	Female characteristics, complete gonadal dysgenesis	NA
Allali *et al.*[74] 2011	c.141_144dup	p.Cys49Hisfs*39	46,XY	Ambiguous genitalia, Prader III Microphallus, small gonads (18 mm) in scrotum	NA

NA: not available

However, the pathophysiology of DSD in the sex organs at a single-cell resolution has not been described.

Here, we reported the identification of an *NR5A1* heterozygous mutation p.Q42E (c.124C>G) in a patient with 46,XY DSD who exhibited ambiguous genitalia at birth and who is part of the population of East Asia. We have further provided a comprehensive cell atlas of the epididymis and demonstrated the presence of inflammation and fibrosis in the epididymis. Our results provide an important reference for the rearing selection of patients with c.124C>G (p.Q42E) mutation in the *NR5A1* gene.

PATIENT AND METHODS

Patient
An 18-year-old 46,XY DSD patient diagnosed by karyotype testing, computed tomography, and ultrasound was recruited for the current study. This study was approved by the Ethics Committee of the Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine (Shanghai, China; No. SH9H-2020-T138-2). Informed consent was signed by the patient.

Genetic study
Next-generation sequencing (NGS) panels and mutation analysis were performed to explore the abnormal genes and specific mutations, as previously described.[24] Briefly, the patient's genomic DNA was extracted from peripheral blood lymphocytes with the QuickGene DNA whole blood kit L (Kurabo, Osaka, Japan). A total of five genes including steroid 5 alpha-reductase 2 (*SRD5A2*), cytochrome P450 family 17 subfamily A member 1 (*CYP17A1*), androgen receptor (*AR*), *SRY*, and *NR5A1* of a targeted NGS panel were sequenced by the Access Array system (Fluidigm, South San Francisco, CA, USA). Gene mutations

were confirmed by Sanger sequencing of the purified polymerase chain reaction (PCR) products.

Epididymis collection and single-cell preparation
The epididymis was collected during cryptorchidectomy, and the epididymal tunica adventitia, fat tissue, efferent ductules, and deferent duct were removed under a dissecting microscope (**Supplementary Figure 1a**). The isolated epididymis was placed in GEXSCOPE™ tissue preservation solution (Singleron Biotechnologies, Nanjing, China) and transported to the laboratory. Tissue dissociation was performed as previously described.[25] Briefly, the isolated specimens were minced into 1–2-mm pieces after Hanks Balanced Salt Solution (HBSS) three times washing. A total of 2 ml of GEXSCOPE™ tissue dissociation solution (Singleron Biotechnologies) was added into the minced specimens and digested for 15 min at 37°C. The digested sample was filtered with 40-μm sterile strainers (Corning, New York, NY, USA) and collected by centrifugation (Centrifuge 5804R, Eppendorf, Hamburg, Germany) at 130*g* for 5 min. The cell pellet was treated with 2 ml GEXSCOPE™ red blood cell lysis buffer (Singleron Biotechnologies) for 10 min in order to remove red blood cells. After that, the cells were resuspended in PBS and counted with trypan blue staining (T6146, Sigma, Burlington, VT, USA) and in a TC20 automated cell counter (Bio-Rad, Hercules, CA, USA).

Library preparation and single-cell RNA sequencing (scRNA-seq) data preprocessing
A total of 1×10^5 cells per ml of single-cell suspension was loaded onto the microfluidic plate. scRNA-seq libraries were established according to the instructions of the GEXSCOPE™ single-cell RNA library kit (Singleron Biotechnologies). The qualified individual libraries were diluted into 4 nmol l⁻¹ and pooled for sequencing with Illumina HiSeq ×10 sequencing machine.

Raw reads were processed by the SCOPE-tools pipeline (https://github.com/SingleronBio/SCOPE-tools; last accessed on March 23, 2021) to generate the gene expression matrix. Briefly, poor-quality reads and adapter sequences were removed from the raw fastq files. Clean reads were mapped to GRCh38 of the human genome (version 92). The mapping results were aggregated by unique molecular identifiers (UMIs) extracted from read 1 and converted to the expression matrix.

Dimension reduction and clustering analysis

The analysis for the dimension reduction and cell clustering was carried out in R with the Seurat version 4 package (https://satijalab.org/seurat/).[26] The UMI matrix was filtered and normalized before dimension reduction. The percentages of mitochondrial and hemoglobin reads were calculated in the quality control (QC) process, the upper cutoffs for which were 50% and 10%, respectively. Cells with too many (unique feature count ≥3000 or UMI count ≥20 000) or few (unique feature count ≤200) transcripts were also excluded from downstream analysis. The ratio of mitochondrial reads, along with the unique feature count and UMI count, was regressed out in the process of linear regression to minimize the effect of mitochondrial genes. Principal component (PC) scores were then calculated from the top 2000 variable features. Unsupervised clustering was calculated at a resolution of 0.09 by the default SNN algorithm. Subsequent subcluster analysis adopted a higher clustering resolution of 0.2 to achieve a finer granularity. In addition, the Uniform Manifold Approximation and Projection (UMAP) was applied to the top 50 PCs to visualize clusters and subclusters. Cell clusters were manually annotated using known cell markers from published literature. The expression levels for each cell marker in each cell cluster were plotted using log-normalized counts.

Differential gene analysis and gene enrichment analyses

The "FindAllMarkers" method from Seurat version 4 was used to detect cell-type-specific genes. *P* values were calculated against each gene using a two-sided Wilcoxon rank-sum test, which were adjusted by Bonferroni correction. The criteria for selecting cell-type-specific genes were as follows: the gene detected in at least 25% of cells in the cluster and the average gene expression was higher in the target cluster than that in other clusters ($\log_2$ [fold change] >0.25, adjusted *P* < 0.05). All average expression values were added by 1 before calculation.

Subsequent gene enrichment was performed by the R package clusterProfiler v3.6.1 (https://bioconductor.org/packages/3.11/bioc/html/clusterProfiler.html; last accessed on November 10, 2020).[27] The Gene Ontology (GO) overrepresentation test was implemented in clusterProfiler under the name "enrichGO". GO classifications for human genes were extracted from the Genome-wide annotation for Human database (https://bioconductor.org/packages/3.11/data/annotation/html/org.Hs.eg.db.html, version 3.11.4; last accessed on November 24, 2020). *P* values were corrected by the Benjamini–Hochberg procedure. The *P* value threshold for the selection of significant GO terms was set as 0.05.

Cell–cell communication analysis

Ligand–receptor enrichment analysis was calculated using CellChat (http://www.cellchat.org/; last accessed on May 10, 2021),[28] which uses manifold learning and quantitative contrasts to infer cell–cell communication networks from scRNA-seq data. A total of 1199 unique human ligand–receptor pairs under the "Secreted Signaling" category were extracted from the built-in database for the analysis. Overexpressed ligand–receptor pairs were enriched from normalized expression values, which was used to calculate the strength and probability of interactions. Incoming/outgoing communication patterns were calculated using the "identifyCommunicationPatterns" method.[28]

Transcription factor regulatory network analysis

The transcription factor regulation network was analyzed by Python version of Single-Cell rEgulatory Network Inference and Clustering (pySCENIC),[29] an efficient Python implementation of the original R version SCENIC,[30] which was used in the reconstruction of the transcription factor regulatory network. The pySCENIC workflow consists of four major steps. First, coexpression modules are inferred from the filtered expression matrix using the GRNboost2 algorithm. Second, direct targets are inferred from these modules using cisTarget. Two precomputed genome ranking files (hg38__refseq-r80__500 bp_up_and_100 bp_down_tss.mc9nr.feather and hg38__refseq-r80__10 kb_up_and_down_tss.mc9nr.feather) were used in the motif enrichment procedure. Third, an area under the curve (AUC) score was calculated for all candidate target genes to quantify the activity of each transcription factor (or regulon). Finally, the AUC matrix was binarized to generate a binary (on–off) matrix that represents the activity pattern of regulons. The regulatory network associated with the homeobox (*HOX*) gene family was analyzed and visualized by Cytoscape.[31]

Cell-type-specific regulons were determined using the regulon specificity score (RSS), an entropy-based metric proposed in a previous study,[32] to identify essential regulators for each major cell type across multiple tissues. A different cell pooling strategy was used in this study compared to the original method.[33] In the present study, we take 50 cells from each cluster to generate a new dataset, and it adapts a stratified bootstrapping procedure to correct for cluster size. At each iteration of the bootstrap, the same number of cells was sampled with replacement from each cluster to make an evenly distributed cell population, which was used to calculate the RSS. The calculation of RSS is implemented as described in the original paper.

All results were averaged to calculate the specificity score at the end. Standard errors and deviations from original (unsampled) RSSs were calculated to evaluate the effectiveness of bootstrapping. A total of 1000 replications were performed, and 50 cells were taken from each cell cluster in this study.

Data availability

The raw data of the scRNA-seq reported in the present paper have been deposited in the Genome Sequence Archive[34] at the National Genomics Data Center,[35] China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA001437). The custom codes are available from the corresponding authors upon reasonable requests.

RESULTS

A novel mutation of NR5A1 caused DSD deficiency

An 18-year-old 46,XY DSD patient was recruited. The patient was adopted and raised as a female. After reaching adulthood, the patient presented with amenorrhea. Physical examination revealed a feminization of external genitalia performed as clitoromegaly, perineal hypospadias with a small penis (stretched length of 4 cm), impalpable testes, poor development of labia minora and labia majora, and Tanner stage 1 breast development (data not shown). Pelvic computed tomography revealed testicular and epididymal tissues in the inguinal region where the left gonad resided. Müllerian structures were not detected in the right gonad (**Supplementary Figure 1b**). After discussion with the parents, the patient decided to continue in a female gender and underwent left cryptorchidectomy and feminizing genitoplasty. The removed gonads are shown in

Supplementary Figure 1c and **1d**. Specifically, the testicular tissue was dysplastic (approximately 0.5 ml), and the epididymis and vas deferens were morphologically normal.

NGS panel was performed to examine the pathological gene mutation for the patient. The results revealed a novel c.124C>G heterozygous mutation in the third exon of *NR5A1* of the population of East Asia (**Figure 1a**). The mutation was located in the coding region that led to a missense p.Q42E mutation in the NR5A1 protein (**Figure 1b**). The mutations were located between the two highly conserved zinc finger domains and the core DNA-binding domain of NR5A1 (**Figure 1b**). Furthermore, the protein structure analysis for the p.Q42E mutant revealed the lack of a side chain near the DNA-binding domain (**Figure 1c**).

Epididymal cell atlas of DSD patients with NR5A1 mutation

To elucidate the cell types of the patient's epididymis, the scRNA-seq was conducted, from which the efferent ducts were removed. Following cell quality control (**Supplementary Figure 2a**), a total of 13 316 cells from the epididymis were clustered into twelve cell populations (**Figure 2a**). The cell types were annotated using the reported cell markers (**Figure 2b** and **2c**, and **Supplementary Figure 2b**). Cell populations were annotated as fibroblasts (C0), monocytes (C1), epididymis epithelial cells (C2), smooth muscle cells (C3), type 2 classical dendritic cells (DCs; C4), T-cells (C5), clear cells (C6), mast cells (C7), B-cells (C8), gonadal endothelial cells (C9), endothelial cells (C10), and type 1 classical DCs (C11). The top five differentially expressed genes (DEGs) were charted for each cell cluster (**Figure 2c**). Furthermore, the GO terms related to the DEGs in the cell cluster indicated the potential role of cells (**Supplementary Figure 2c**). Unexpectedly, the two major cell types, principal cells and basal cells, which are known to play critical roles in the epididymis, were not identified in the first clustering analysis. Of note, the epididymal epithelial cells comprised only 9.2% of the whole population, while the fibroblasts contributed to 46.5% of the cells (**Figure 2d**), suggesting divergent changes in the cell type in the epididymis of the DSD patient with an *NR5A1* mutant.
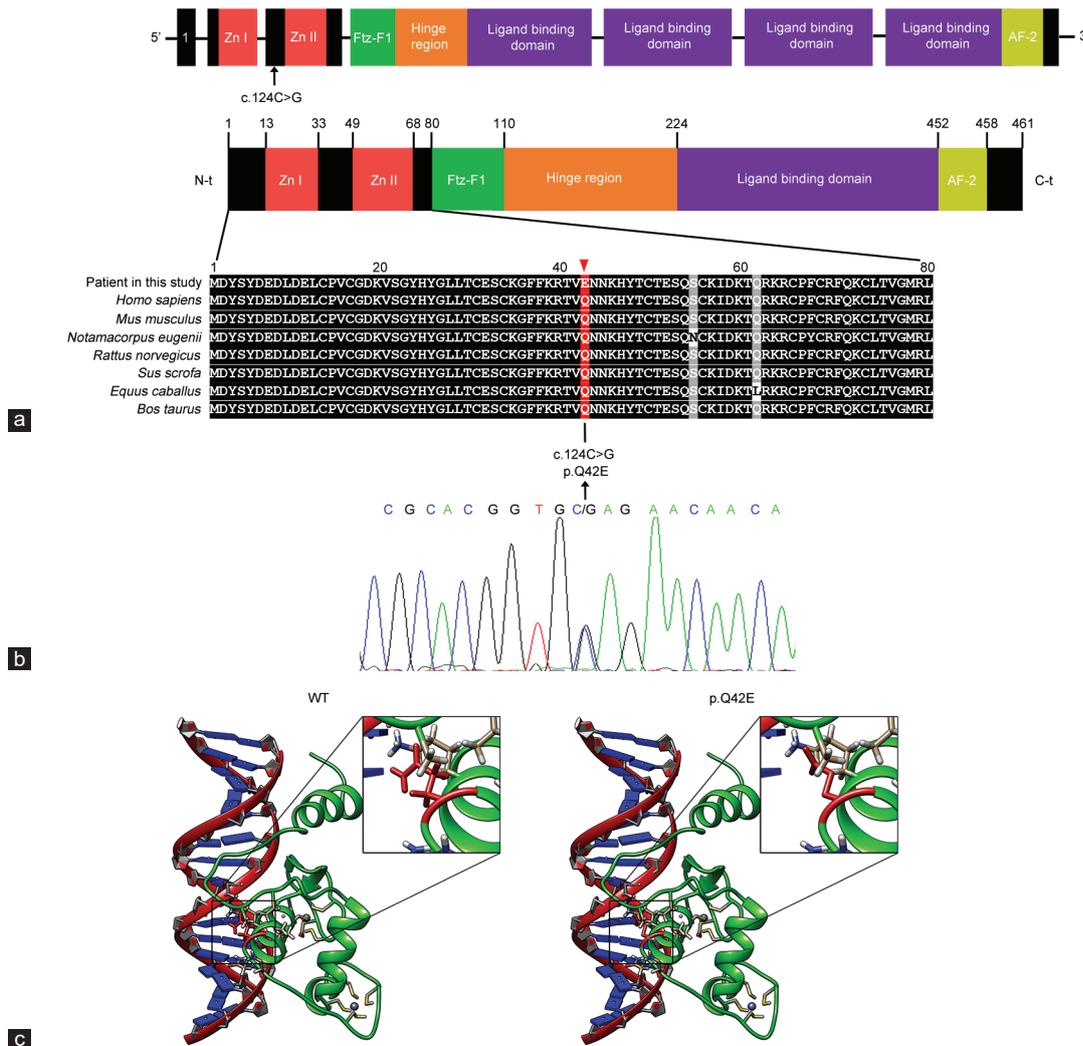
Figure 1: Schematic diagram of the novel mutation site of *NR5A1* gene. (**a**) Schematic graph for the *NR5A1* showing the location of the mutation. (**b**) The mutated p.Q42E was located between two zinc domains, which is important for DNA-binding activity. The alignment of NR5A1 of p.Q42E and its neighboring AAs are displayed in the middle panel. The sequence of the affected AA is conserved. Sanger sequencing showing the c.124C>G substitution is depicted below. (**c**) Predicted structures of the protein encoded by NR5A1 of normal and mutation p.Q42E. *NR5A1*: nuclear receptor subfamily 5 group A member 1; AA: amino acid; N-t: NH$_2$-terminal; C-t: carboxyl-terminal; AF-2: activation function-2 domain; Ftz-F1: Fushi Tarazu factor 1.
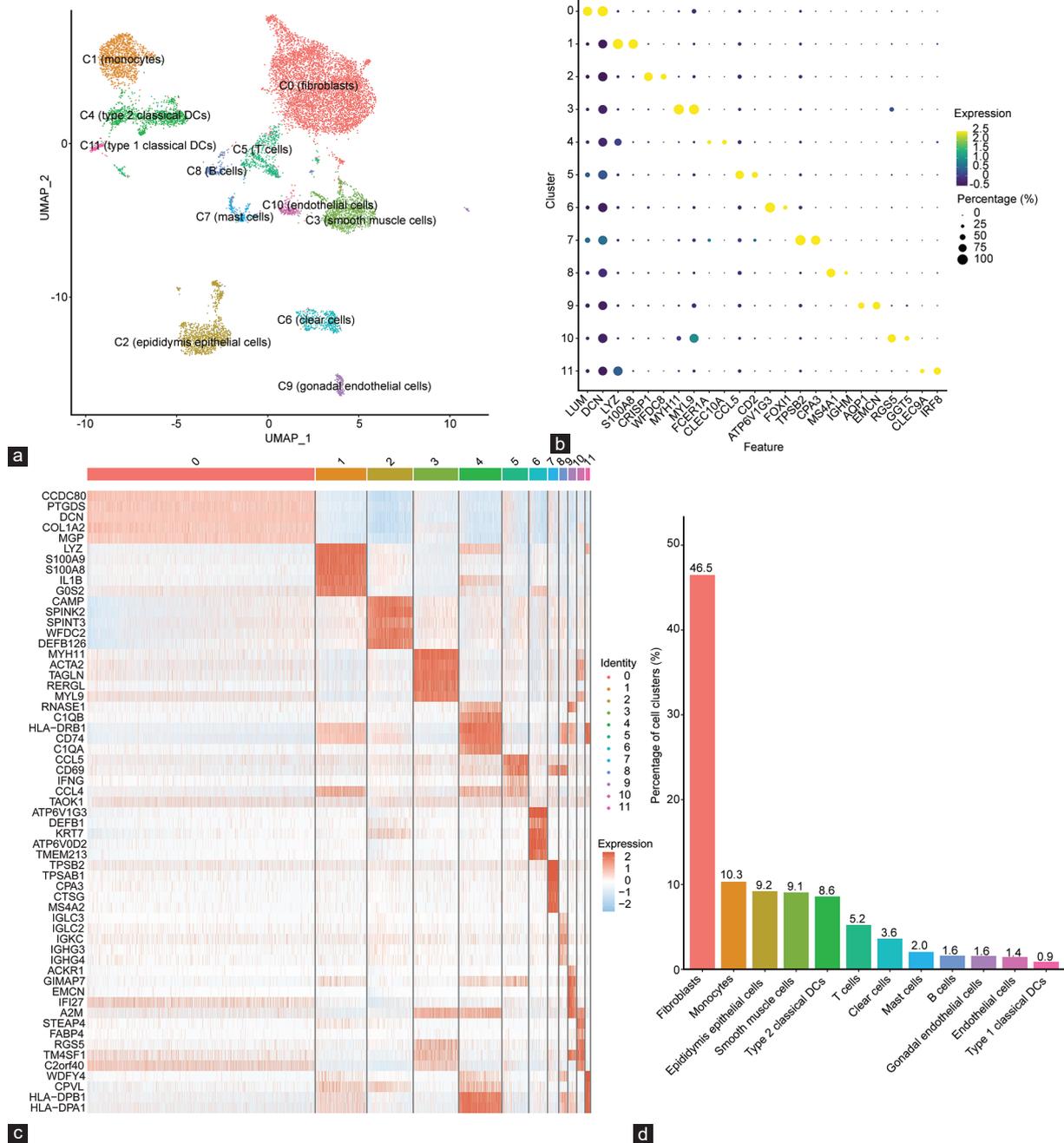
Figure 2: Overview of scRNA sequencing for the human epididymis with the *NR5A1* gene p.Q42E mutation. (**a**) UMAP visualization of the twelve cell clusters annotated in the patients' epididymis after erythrocyte filtration. (**b**) Cell annotation of each epididymal cell population based on cell-type-specific marker genes from the literature (expression values are centered and scaled by row). (**c**) The top 5 marker genes of the heatmap showed for each population (expression scaled by row). (**d**) The cell proportion of each cluster was displayed. *NR5A1*: nuclear receptor subfamily 5 group A member 1; UMAP: uniform manifold approximation and projection; scRNA: single cell RNA.

Loss of epithelial cells in the NR5A1 p.Q42E mutated epididymis

We reasoned that the principal cells and basal cells in DSD patients might have a distorted transcriptome and thus were missed in the routine clustering analysis. To further validate the presence of the principal cells and basal cells in the epididymis of this patient, a higher shared nearest neighbor (SNN) resolution was applied to clarify the subpopulations in epididymis epithelial cells. Three subpopulations of epididymis epithelial cells were identified based on cell annotation using cell markers reported in the literature (**Figure 3a** and **3b**, and **Supplementary Figure 3a**).

These included apical and narrow cells (82.4%), basal cells (15.4%), and principal cells (2.2%), as shown in **Figure 3c**. These results suggested that the major cell type of the epididymis, the principal cell, was lost in the NR5A1 p.Q42E mutated epididymis.

Epithelial cell–fibroblast cell transition determined by cell–cell communication analysis

Since the epididymal epithelial cells were markedly lost and replaced by fibroblasts and other cells, we sought to decipher the signals involved
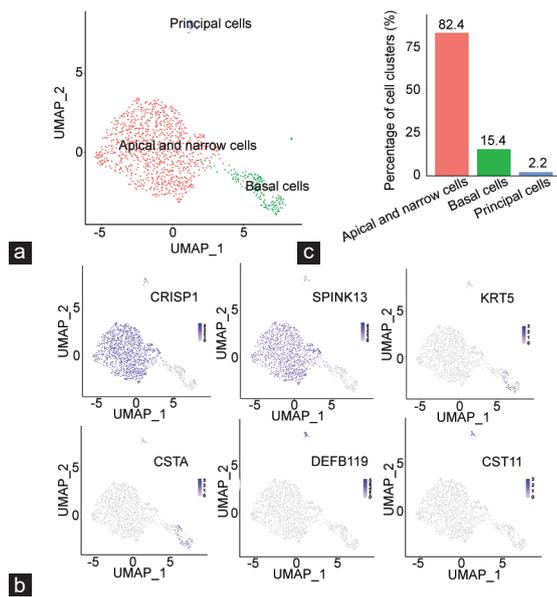
Figure 3: Features of epididymis epithelial cell subpopulations. (**a**) UMAP graph of the subclusters in the epididymal epithelial cells. Three subclusters of epididymis epithelial cells were identified. (**b**) The proportion of each epithelial cell subpopulation is illustrated. (**c**) Representative cell-specific marker illustrated by UMAP of epididymy epithelial cell clusters (color values are expression). UMAP: uniform manifold approximation and projection; CRISP1: cysteine-rich secretory protein 1; SPINK13: serine peptidase inhibitor Kazal type 13; KRT5: keratin 5; CSTA: cystatin A; DEFB119: defensin beta 119; CST11: cystatin 11.

in this transition. The transition of epithelial cells to fibroblasts or fibroblast-like cells is well established in lung fibrosis and several cancer metastases that are accompanied by the epithelial-to-mesenchymal transition (EMT) program.[36–38] Therefore, we postulated the involvement of EMT in the pathophysiology of this DSD patient. To elucidate the possible cell–cell interactions and the underlying signaling pathways, we probed for known cell–cell communications using CellChat. As expected, a complex cell–cell interaction network was found between the twelve cell types (**Figure 4a**). In particular, the communication between epididymal epithelial cells and fibroblast cells was among the top interactions identified (**Supplementary Figure 3b**). Notably, the CellChat-inferred IGF signaling pathway network was enriched between epithelial cells and fibroblast cells (**Figure 4b**). IGF signaling is well reported to induce EMT, which contributes to growth, cancer metastasis, and fertility.[39] We further probed the expression patterns of the molecules involved in the insulin-like growth factor (IGF) signaling cascade and found that IGF1 receptor 2 (IGF2R) was only expressed in epididymis epithelial cells, while its ligands IGF1 and IGF2 were mainly expressed in fibroblasts (**Figure 4c**). Notably, consistent with the involvement of IGF1-IGF1R signaling in the EMT process,[40] IGF1-IGF1R was the primary contributor to fibroblast–epithelial cell communication (**Figure 4d** and **Supplementary Figure 4**). We next demonstrated the crucial signaling cascades with the function of the CellChat analysis module. During the outgoing communication of signaling, the fibroblast growth factor (FGF), gastrin (GAS), and platelet-derived growth factor (PDGF) signals were predicted to be secreted by fibroblasts, whereas IGF and Wingless (WNT) signals were secreted by epididymal epithelial cells. Meanwhile, the incoming communication of signaling demonstrated that epididymis epithelial cells and fibroblasts were regulated by IGF and WNT signaling

(**Figure 4e**). Together, these analyses uncovered EMT-related signaling between epididymal epithelial cells and fibroblasts in DSD.

Transcription factor analysis uncovers the cell state of EMT
Gene expression is primarily regulated by transcription factors (TFs). With the robust gene regulatory network (GRN) analysis tool SCENIC, we evaluated the activity of GRNs and the stable cell state in the *NR5A1* mutant epididymis. The binary regulon activity matrix is charted in **Figure 5a**, and the fibroblast cell state was driven by EMT TFs, including members of the *HOX* gene family. Furthermore, *HOX* genes, including *HOXA6*, *HOXC6*, *HOXC8*, and *HOXC9* (**Figure 5a** and **5b**), were mainly activated in fibroblast and epididymis epithelial cells with EMT genes.[41] To further explore the regulatory functions of *HOX* genes,[42] GO terms and motif analysis of their target genes were calculated accordingly. The gene interaction networks and biological processes in which *HOX* genes participate are charted in **Figure 5c**. In addition, the enriched sequences on the promoter region of TFs targeting genes such as *HOXB7* and *HOXB8* indicated that the epididymis epithelial cells were in the EMT cell state.

DISCUSSION
NR5A1 is well characterized to contribute to sex development both in humans and mice.[12,13,43] It was reported that NR5A1 was upregulated in the undifferentiated stage of gonads, where it stimulates SOX9 and AMH expression during male development and sex differentiation.[11,44] In humans, mutations of *NR5A1* led to various phenotypes in gonadal and adrenal development, including 46,XY DSD with phenotypes of female/androgen insensitivity syndrome (AIS)-like males and 46,XX DSD with ovotesticular or premature ovarian insufficiency.

NR5A1 protein is composed of five elements which include zinc finger I and zinc finger II consisting of DNA-binding domains, a hinge region containing an activation function domain (AF-1), a ligand-binding domain (LBD), and the other activation function domain (AF-2). To date, approximately 190 mutations in *NR5A1* have been described.[45,46] Among the reports, 97% of cases presented with heterozygous mutations of *NR5A1*, and few cases were homozygous mutations,[46] suggesting that heterozygous mutations may be more frequently seen in people. A few studies have reported that NR5A1 presents a dose-dependent pattern.[47,48]

Of note, the present study reported a novel heterogynous mutation in *NR5A1* (c.124C>G) in a patient with 46,XY DSD. This novel c.124C>G (p.Q42E) mutation in the *NR5A1* gene is located between the two zinc finger domains (**Figure 1b**). Similar missense mutations in amino acid residues of close proximity demonstrated the loss of function to activate cytochrome P450 family 19 subfamily A member 1 (*CYP19*) or anti-Müllerian hormone (*AMH*) promotor,[17,49] strongly suggesting that the p.Q42E mutation might impair DNA-binding and transcriptional activity.

This 46,XY DSD patient showed dysplastic testis tissue. Interestingly, an epididymis with normal size and morphology was found when the patient underwent cryptorchidectomy surgery (**Supplementary Figure 1c** and **1d**). We further investigated the cell atlas of the remaining epididymis by single-cell sequencing. Although the patient showed a normal epididymis with distinct segments, the cell composition was distorted. In healthy individuals, principal cells are the major cell type in the epididymis, which accounts for approximately 60% of the human epididymal epithelium.[50,51] However, the percentages of principal cells and basal cells in the patient's epididymis were only 0.2% and 1.4%, respectively. In contrast, fibroblasts represented the
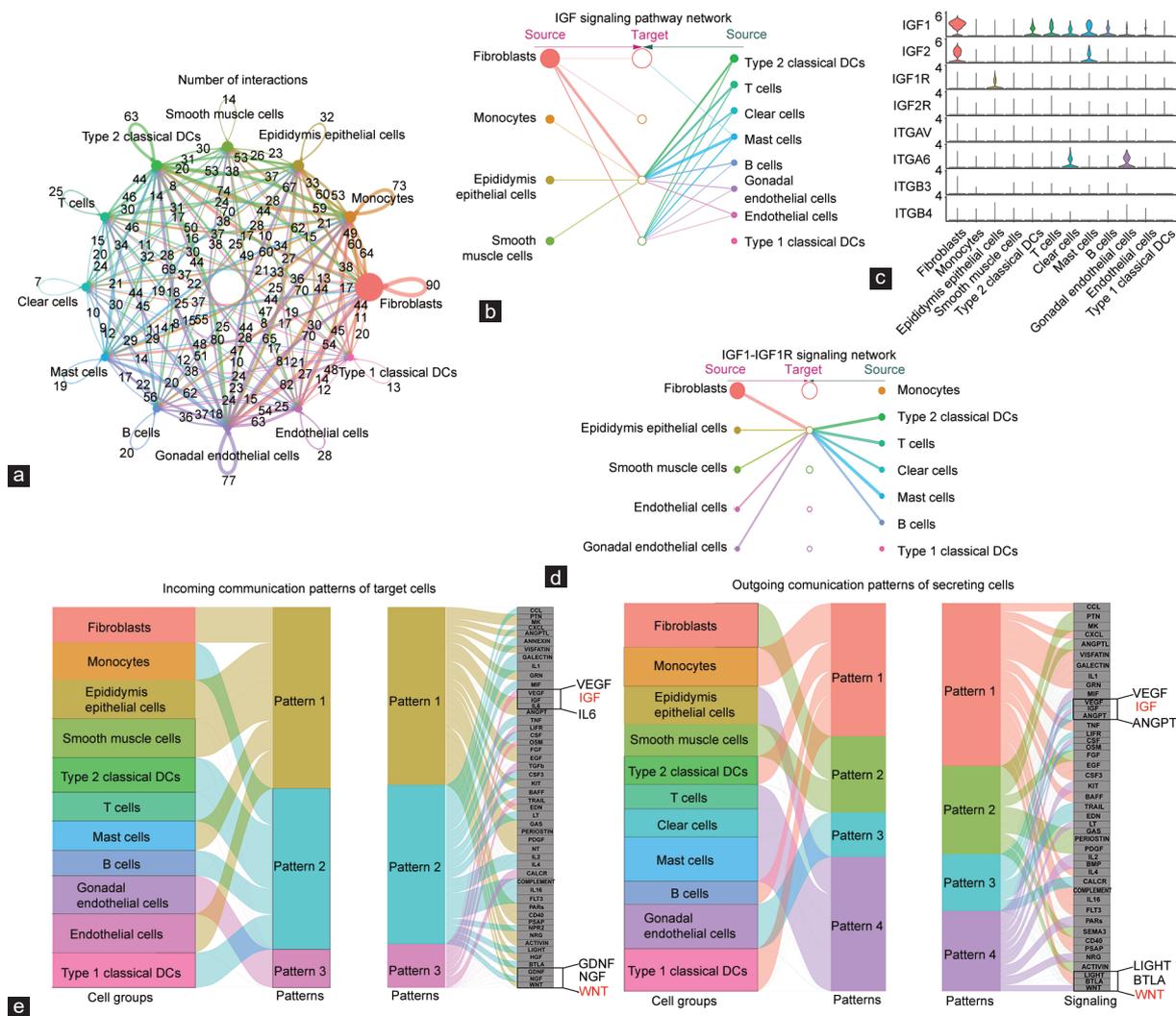
Figure 4: CellChat analysis between fibroblasts and epididymis epithelial cells. (**a**) Overview of pronounced ligand–receptor pairs in cell clusters. The thickness of the line represents ligand–receptor pair number. (**b**) Hierarchical plot representing the deductive cell communications for the IGF signaling pathway. (**c**) Violin plot displaying the expression of IGF1 signaling genes in each cell cluster. (**d**) Hierarchical plot showing the interactions for IGF1-IGF1R signaling. (**e**) Alluvial plot showing the incoming and outgoing signaling patterns of cell groups. The inferred outgoing communication patterns of secreting cells and incoming patterns of target cells suggest the correlation between the inferred potential cell groups, patterns, and signaling pathways. IGF1R: insulin-like growth factor 1 receptor; DC: dendritic cell; IGF: insulin-like growth factor; ITGB: integrin subunit beta; VEGF: vascular endothelial growth factor; ANGPT: angiopoietin; GDNF: glial cell-derived neurotrophic factor; NGF: nerve growth factor; WNT: Wingless; BTLA: B and T lymphocyte associated; LIGHT: TNF superfamily member 14.

most abundant cell cluster, accounting for 46% of the total cells. This phenomenon suggested that the epididymal epithelial cells might have degenerated and/or transited into fibroblasts via the EMT program. With the novel and powerful tool for cell–cell communication named CellChat, cell–cell communication networks were constructed. Notably, several EMT signaling pathways were identified between epithelial cells and fibroblast cells, including WNT and IGF signaling. EMT is a critical step during biological and pathological processes, including embryonic development, cancer metastasis,[52] and endometriosis.[53] It has been reported that EMT is involved in many tissue fibroses in which fibroblasts or myofibroblasts are derived from epithelial cells and cause permanent damage or organ malfunction.[54,55] Many studies have demonstrated the involvement of the IGF1-IGF1R signaling pathway in EMT in different organs.[40,56–58] For example, Li *et al.*[40] demonstrated that IGF signaling induced EMT in alveolar epithelial type II (ATII) cells

and formed fibroblast-like, ECM-producing cells. De Vincenzo *et al.*[59] reported that the IGF1-IGF1R axis was involved in breast epithelial cell and stromal fibroblast transition. Intriguingly, we found that the IGF1-IGF1R signaling pathway mediated cell communication between epididymis epithelial cells and fibroblasts. These results suggest that the onset of EMT may contribute, at least partially, to the loss of epididymal epithelial cells in this 46,XY DSD patient.

Apart from the cell–cell communication analysis, transcription factor analysis for single-cell sequencing was performed to elucidate each cell state and TF gene regulation. A novel SCENIC analysis revealed that many *HOX* genes, including *HOXC6*, *HOXA5*, *HOXA10*, *HOXC8*, and *HOXA11*, were turned on in the fibroblast cluster. *HOX* genes are the most conserved homeodomains and are involved in DNA-specific binding.[60] *HOX* genes were also reported to be expressed in adult epididymis epithelial cells[61] and

Figure 5: Transcription factor regulatory networks derived from SCENIC analysis. (**a**) Regulon activity for the epididymal cell clusters. The color bar indicates cell clusters. The on/off state of the cluster-specific regulons are displayed by heatmap. Black represents the state in which the regulon was activated, and white represents the state in which the regulon was inactivated. The *HOX* gene family is magnified. (**b**) Heatmap showing the expression level of the *HOX* gene family in each population. (**c**) Cytoscape visualization revealed the networks of transcriptional factors and target genes of the *HOX* gene family. Biological processes and motif sequences of target genes for *HOX* transcription factors were listed. g: targeting gene; HOX: homeobox; MF: molecular function; SCENIC: Single-Cell rEgulatory Network Inference and Clustering.

play a role in the coiling of the epididymis ductus and segmental function regulation in the epididymis.[62,63] Of note, HOX proteins, as transcription factors, contribute to EMT in the cancers.[41,60] In the reproductive system, *HOX* genes were found to be involved in EMT in ovarian tumors[64] and endometrial carcinomas.[65] Consistent with previous studies, the *HOX* gene regulons were "on" in fibroblasts, which suggested that HOX regulated EMT during the epithelial cell to fibroblast cell transition.

Although the novel mutation of *NR5A1* in a patient and its cell composition in the epididymis were examined, there are still several limitations in the present study. First, this novel mutation was found in one patient, and there was a lack of tissues for validation after single-cell sequencing. Second, the epididymis of the patient was found in the inguinal region, which is very similar to what is observed in cryptorchidism. Interestingly, some *NR5A1* mutations caused the cryptorchidism phenotype,[66] which is often associated with peritubular and interstitial fibrosis,[67,68] suggesting that the signaling cascade for EMT identified in the present study may be associated with DSD with a cryptorchidism-like phenotype only. Further studies are warranted to better clarify the underlying mechanisms in *NR5A1* mutant mice as well as the role of the p.Q42E polymorphism in clinical manifestations.

CONCLUSIONS

Our current data provide an indispensable resource for the cell atlas of a patient epididymis with a novel c.124C>G (p.Q42E) mutation in the *NR5A1* gene at a single-cell resolution. Advanced and powerful bioinformatics analysis of cell–cell communications and GRNs at the single-cell level revealed that epididymal epithelial cell loss and fibroblast occupation are associated with the EMT process. The present study not only elucidates the abnormal epididymis functionalities of *NR5A1* mutation but also provides a reference for the diagnosis and therapy of patients with an *NR5A1* gene mutation.

AUTHOR CONTRIBUTIONS

HC and YL conceived and supervised the project. HC, YL, and HJ designed the experiments. YWZ and YL diagnosed the patient, applied for the ethical approval, performed the genetic study and the surgery, and collected the human sample. JWS, YWZ, YFC, JWT, FQ, MYZ, and MY performed the experiments. JWS and YFC performed the data analysis. KLF, HC, JWS, and YWZ wrote the manuscript. KLF, HC, HCL, YL, YWZ, HJ, and GCX revised the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

All authors declare no competing interests.

ACKNOWLEDGMENTS

Supplementary Information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

REFERENCES

1 Hughes IA, Houk C, Ahmed SF, Lee PA; Lawson Wilkins Pediatric Endocrine Society/European Society for Paediatric Endocrinology Consensus Group. Consensus statement on management of intersex disorders. *J Pediatr Urol* 2006; 2: 148–62.

2 Nabhan ZM, Lee PA. Disorders of sex development. *Curr Opin Obstet Gynecol* 2007; 19: 440–5.

3 Aydin BK, Saka N, Bas F, Bas EK, Coban A, *et al.* Frequency of ambiguous genitalia in 14,177 newborns in Turkey. *J Endocr Soc* 2019; 3: 1185–95.

4 Ohnesorg T, Vilain E, Sinclair AH. The genetics of disorders of sex development in humans. *Sex Dev* 2014; 8: 262–72.

5 Pasterski V, Prentice P, Hughes IA. Impact of the consensus statement and the new DSD classification system. *Best Pract Res Clin Endocrinol Metab* 2010; 24: 187–95.

6 Bashamboo A, Donohoue PA, Vilain E, Rojo S, Calvel P, *et al.* A recurrent p.Arg92Trp variant in steroidogenic factor-1 (NR5A1) can act as a molecular switch in human sex development. *Hum Mol Genet* 2016; 25: 3446–53.

7 Lee PA, Nordenstrom A, Houk CP, Ahmed SF, Auchus R, *et al.* Global disorders of sex development update since 2006: perceptions, approach and care. *Horm Res Paediatr* 2016; 85: 158–80.

8 Nagy O, Karteszi J, Hartwig M, Bertalan R, Javorszky E, *et al.* The importance of the multiplex ligation-dependent probe amplification in the identification of a novel two-exon deletion of the *NR5A1* gene in a patient with 46,XY differences of sex development. *Mol Biol Rep* 2019; 46: 5595–601.

9 Sablin EP, Blind RD, Krylova IN, Ingraham JG, Cai F, *et al.* Structure of SF-1 bound by different phospholipids: evidence for regulatory ligands. *Mol Endocrinol* 2009; 23: 25–34.

10 Hatano O, Takakusu A, Nomura M, Morohashi K. Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. *Genes Cells* 1996; 1: 663–71.

11 Sekido R, Lovell-Badge R. Sex determination involves synergistic action of SRY and SF1 on a specific *Sox9* enhancer. *Nature* 2008; 453: 930–4.

12 Anamthathmakula P, Miryala CS, Moreci RS, Kyathanahalli C, Hassan SS, *et al.* Steroidogenic Factor 1 (*Nr5a1*) is required for Sertoli cell survival post sex determination. *Sci Rep* 2019; 9: 4452.

13 Karpova T, Ravichandiran K, Insisienmay L, Rice D, Agbor V, *et al.* Steroidogenic factor 1 differentially regulates fetal and adult Leydig cell development in male mice. *Biol Reprod* 2015; 93: 83.

14 Parker KL, Rice DA, Lala DS, Ikeda Y, Luo X, *et al.* Steroidogenic factor 1: an essential mediator of endocrine development. *Recent Prog Horm Res* 2002; 57: 19–36.

15 Zimmermann S, Schwarzler A, Buth S, Engel W, Adham IM. Transcription of the Leydig insulin-like gene is mediated by steroidogenic factor-1. *Mol Endocrinol* 1998; 12: 706–13.

16 Siklar Z, Berberoglu M, Ceylaner S, Camtosun E, Kocaay P, *et al.* A novel heterozygous mutation in steroidogenic factor-1 in pubertal virilization of a 46,XY female adolescent. *J Pediatr Adolesc Gynecol* 2014; 27: 98–101.

17 Tajima T, Fujiwara F, Fujieda K. A novel heterozygous mutation of steroidogenic factor-1 (SF-1/Ad4BP) gene (*NR5A1*) in a 46,XY disorders of sex development (DSD) patient without adrenal failure. *Endocr J* 2009; 56: 619–24.

18 Tantawy S, Lin L, Akkurt I, Borck G, Klingmuller D, *et al.* Testosterone production during puberty in two 46,XY patients with disorders of sex development and novel *NR5A1* (SF-1) mutations. *Eur J Endocrinol* 2012; 167: 125–30.

19 Woo KH, Cheon B, Kim JH, Cho J, Kim GH, *et al.* Novel heterozygous mutations of *NR5A1* and their functional characteristics in patients with 46,XY disorders of sex development without adrenal insufficiency. *Horm Res Paediatr* 2015; 84: 116–23.

20 Bashamboo A, Ferraz-de-Souza B, Lourenco D, Lin L, Sebire NJ, *et al.* Human male infertility associated with mutations in *NR5A1* encoding steroidogenic factor 1. *Am J Hum Genet* 2010; 87: 505–12.

21 Philibert P, Polak M, Colmenares A, Lortat-Jacob S, Audran F, *et al.* Predominant Sertoli cell deficiency in a 46,XY disorders of sex development patient with a new *NR5A1/SF-1* mutation transmitted by his unaffected father. *Fertil Steril* 2011; 95: 1788.e5–9.

22 Yagi H, Takagi M, Kon M, Igarashi M, Fukami M, *et al.* Fertility preservation in a family with a novel *NR5A1* mutation. *Endocr J* 2015; 62: 289–95.

23 Little TH, Zhang Y, Matulis CK, Weck J, Zhang Z, *et al.* Sequence-specific

24 Wang H, Zhang L, Wang N, Zhu H, Han B, *et al.* Next-generation sequencing reveals genetic landscape in 46,XY disorders of sexual development patients with variable phenotypes. *Hum Genet* 2018; 137: 265–77.

25 Shi JW, Fok KL, Dai PY, Qiao F, Zhang MY, *et al.* Spatio-temporal landscape of mouse epididymal cells and specific mitochondria-rich segments defined by large-scale single-cell RNA-seq. *Cell Discov* 2021; 7: 34.

26 Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, *et al.* Comprehensive integration of single-cell data. *Cell* 2019; 177: 1888–902.e21.

27 Yu GC, Wang LG, Han YY, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012; 16: 284–7.

28 Jin SQ, Guerrero-Juarez CF, Zhang LH, Chang I, Ramos R, *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* 2021; 12: 1088.

29 Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, *et al.* A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc* 2020; 15: 2247–76.

30 Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017; 14: 1083–6.

31 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498–504.

32 Suo SB, Zhu Q, Saadatpour A, Fei LJ, Guo GJ, *et al.* Revealing the critical regulators of cell identity in the mouse cell atlas. *Cell Rep* 2018; 25: 1436–45.e3.

33 Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, *et al.* Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* 2018; 174: 1293–308.e36.

34 Chen T, Chen X, Zhang S, Zhu J, Tang B, *et al.* The Genome Sequence Archive family: toward explosive data growth and diverse data types. *Genomics Proteomics Bioinformatics* 2021; 19: 578–83.

35 CNCB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2021. *Nucleic Acids Res* 2021; 49: D18–28.

36 Aiello NM, Kang Y. Context-dependent EMT programs in cancer metastasis. *J Exp Med* 2019; 216: 1016–26.

37 Mittal V. Epithelial mesenchymal transition in tumor metastasis. *Annu Rev Pathol* 2018; 13: 395–412.

38 Salton F, Volpe MC, Confalonieri M. Epithelial-mesenchymal transition in the pathogenesis of idiopathic pulmonary fibrosis. *Medicina (Kaunas)* 2019; 55: 83.

39 Hakuno F, Takahashi SI. IGF1 receptor signaling pathways. *J Mol Endocrinol* 2018; 61: T69–86.

40 Li HM, Batth IS, Qu XJ, Xu L, Song N, *et al.* IGF-IR signaling in epithelial to mesenchymal transition and targeting IGF-IR therapy: overview and new insights. *Mol Cancer* 2017; 16: 6.

41 Jonkers J, Pai P, Sukumar S. Multiple roles of HOX proteins in metastasis: let me count the ways. *Cancer Metastasis Rev* 2020; 39: 661–79.

42 Xie X, Liu M, Zhang Y, Wang B, Zhu C, *et al.* Single-cell transcriptomic landscape of human blood cells. *Natl Sci Rev* 2021; 8: nwaa180.

43 El-Khairi R, Achermann JC. Steroidogenic factor-1 and human disease. *Semin Reprod Med* 2012; 30: 374–81.

44 Shen WH, Moore CC, Ikeda Y, Parker KL, Ingraham HA. Nuclear receptor steroidogenic factor 1 regulates the Mullerian inhibiting substance gene: a link to the sex determination cascade. *Cell* 1994; 77: 651–61.

45 Askari M, Rastari M, Seresht-Ahmadi M, McElreavey K, Bashamboo A, *et al.* A missense mutation in *NR5A1* causing female to male sex reversal: a case report. *Andrologia* 2020; 52: e13585.

46 Fabbri-Scallet H, de Sousa LM, Maciel-Guerra AT, Guerra-Junior G, de Mello MP. Mutation update for the *NR5A1* gene involved in DSD and infertility. *Hum Mutat* 2020; 41: 58–68.

47 Lourenco D, Brauner R, Lin L, De Perdigo A, Weryha G, *et al.* Mutations in *NR5A1* associated with ovarian insufficiency. *N Engl J Med* 2009; 360: 1200–10.

48 Soardi FC, Coeli FB, Maciel-Guerra AT, Guerra-Junior G, Mello MP. Complete XY gonadal dysgenesis due to p.D293N homozygous mutation in the *NR5A1* gene: a case study. *J Appl Genet* 2010; 51: 223–4.

49 Werner R, Monig I, Lunstedt R, Wunsch L, Thorns C, *et al.* New *NR5A1* mutations and phenotypic variations of gonadal dysgenesis. *PLoS One* 2017; 12: e0176720.

50 Cornwall GA. New insights into epididymal biology and function. *Hum Reprod Update* 2009; 15: 213–27.

51 Leir SH, Yin S, Kerschner JL, Cosme W, Harris A. An atlas of human proximal epididymis reveals cell-specific functions and distinct roles for CFTR. *Life Sci Alliance* 2020; 3: e20200744.

52 Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. *Nat Rev Cancer* 2018; 18: 128–34.

53 Wang CQ, Zhang JT, Fok KL, Tsang LL, Ye M, *et al.* CD147 induces epithelial-

deoxyribonucleic acid (DNA) recognition by steroidogenic factor 1: a helix at the carboxy terminus of the DNA binding domain is necessary for complex stability. *Mol Endocrinol* 2006; 20: 831–43.

to-mesenchymal transition by disassembling cellular apoptosis susceptibility protein/E-cadherin/beta-catenin complex in human endometriosis. *Am J Pathol* 2018; 188: 1597–607.

54 Li MR, Luan FX, Zhao YL, Hao HJ, Zhou Y, *et al*. Epithelial-mesenchymal transition: an emerging target in tissue fibrosis. *Exp Biol Med* 2016; 241: 1–13.

55 Nishioka M, Venkatesan N, Dessalle K, Mogas A, Kyoh S, *et al*. Fibroblast-epithelial cell interactions drive epithelial-mesenchymal transition differently in cells from normal and COPD patients. *Respir Res* 2015; 16: 72.

56 Wang XH, Wu HY, Gao J, Wang XH, Gao TH, *et al*. IGF1R facilitates epithelial-mesenchymal transition and cancer stem cell properties in neuroblastoma via the STAT3/AKT axis. *Cancer Manag Res* 2019; 11: 5459–72.

57 Zhao CZ, Wang Q, Wang B, Sun Q, He ZB, *et al*. IGF-1 induces the epithelial-mesenchymal transition via Stat5 in hepatocellular carcinoma. *Oncotarget* 2017; 8: 111922–30.

58 Zhou J, Wang JJ, Zeng YY, Zhang X, Hu QT, *et al*. Implication of epithelial-mesenchymal transition in IGF1R-induced resistance to EGFR-TKIs in advanced non-small cell lung cancer. *Oncotarget* 2015; 6: 44332–45.

59 De Vincenzo A, Belli S, Franco P, Telesca M, Iaccarino I, *et al*. Paracrine recruitment and activation of fibroblasts by c-Myc expressing breast epithelial cells through the IGFs/IGF-1R axis. *Int J Cancer* 2019; 145: 2827–39.

60 Taniguchi Y. Hox transcription factors: modulators of cell-cell and cell-extracellular matrix adhesion. *Biomed Res Int* 2014; 2014: 591374.

61 Bomgardner D, Hinton BT, Turner TT. 5' *hox* genes and *Meis 1*, a Hox-DNA binding cofactor, are expressed in the adult mouse epididymis. *Biol Reprod* 2003; 68: 644–50.

62 Bomgardner D, Hinton BT, Turner TT. Hox transcription factors may play a role in regulating segmental function of the adult epididymis. *J Androl* 2001; 22: 527–31.

63 Snyder EM, Small CL, Bomgardner D, Xu BF, Evanoff R, *et al*. Gene expression in the efferent ducts, epididymis, and vas deferens during embryonic development of the mouse. *Dev Dyn* 2010; 239: 2479–91.

64 Naora H, Montz FJ, Chai CY, Roden RB. Aberrant expression of homeobox gene *HOXA7* is associated with Mullerian-like differentiation of epithelial ovarian tumors and the generation of a specific autologous antibody response. *Proc Natl Acad Sci U S A* 2001; 98: 15209–14.

65 Yoshida H, Broaddus R, Cheng W, Xie S, Naora H. Deregulation of the *HOXA10* homeobox gene in endometrial carcinoma: role in epithelial-mesenchymal transition. *Cancer Res* 2006; 66: 889–97.

66 Song YN, Fan LJ, Gong CX. Phenotype and molecular characterizations of 30 children from China with *NR5A1* mutations. *Front Pharmacol* 2018; 9: 1224.

67 Rodprasert W, Virtanen HE, Makela JA, Toppari J. Hypogonadism and cryptorchidism.

Front Endocrinol (Lausanne) 2019; 10: 906.

68 Krishna OH, Kotaiah MT, Geetha K, Kiran GS, Reddy PS, *et al*. Cryptorchidism and its effects on histomorphology of testis in paediatric age group. *J Evol Med Dent Sci* 2019; 8: 2480–4.

69 Kohler B, Lin L, Ferraz-de-Souza B, Wieacker P, Heidemann P, *et al*. Five novel mutations in steroidogenic factor 1 (SF1, *NR5A1*) in 46,XY patients with severe underandrogenization but without adrenal insufficiency. *Hum Mutat* 2008; 29: 59–64.

70 Eggers S, Sadedin S, van den Bergen JA, Robevska G, Ohnesorg T, *et al*. Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. *Genome Biol* 2016; 17: 243.

71 Achermann JC, Ito M, Ito M, Hindmarsh PC, Jameson JL. A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. *Nat Genet* 1999; 22: 125–6.

72 Fabbri HC, Ribeiro de Andrade JG, Maciel-Guerra AT, Guerra-Junior G, de Mello MP. *NR5A1* loss-of-function mutations lead to 46,XY partial gonadal dysgenesis phenotype: report of three novel mutations. *Sex Dev* 2016; 10: 191–9.

73 Fabbri-Scallet H, de Mello MP, Guerra G, Maciel-Guerra AT, de Andrade JG, *et al*. Functional characterization of five *NR5A1* gene mutations found in patients with 46,XY disorders of sex development. *Hum Mutat* 2018; 39: 114–23.

74 Allali S, Muller JB, Brauner R, Lourenco D, Boudjenah R, *et al*. Mutation analysis of *NR5A1* encoding steroidogenic factor 1 in 77 patients with 46,XY disorders of sex development (DSD) including hypospadias. *PLoS One* 2011; 6: e24117.

75 Philibert P, Leprieur E, Zenaty D, Thibaud E, Polak M, *et al*. Steroidogenic factor-1 (SF-1) gene mutation as a frequent cause of primary amenorrhea in 46,XY female adolescents with low testosterone concentration. *Reprod Biol Endocrinol* 2010; 8: 28.

76 Rocca MS, Ortolano R, Menabo S, Baronio F, Cassio A, *et al*. Mutational and functional studies on *NR5A1* gene in 46,XY disorders of sex development: identification of six novel loss of function mutations. *Fertil Steril* 2018; 109: 1105–13.
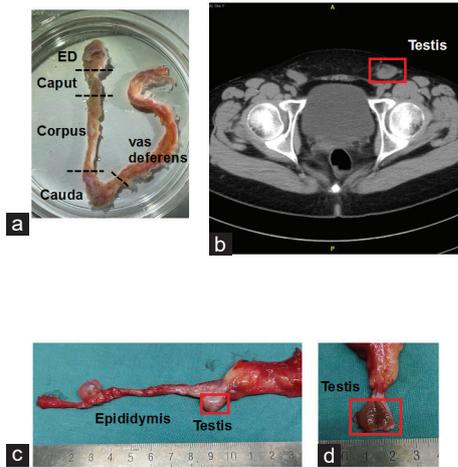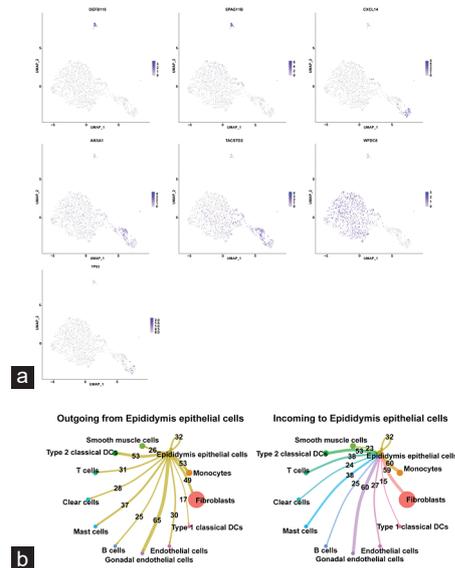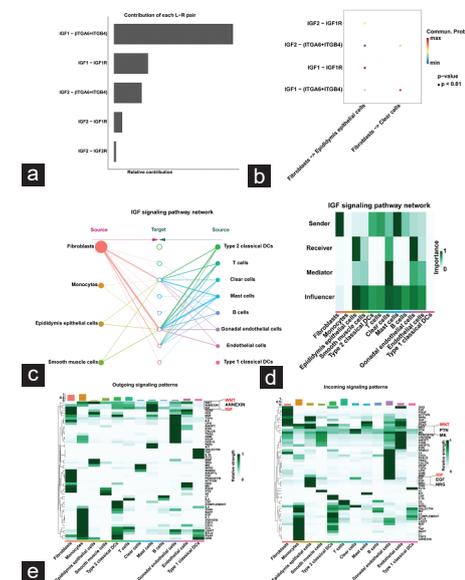
Supplementary Figure 1: Epididymis sample collection, CT graphs, and photographs of the patient. (**a**) Image of dissected human epididymis with the ED, caput, corpus, cauda, and vas deferens regions labeled. (**b**) CT results of the patient with the p.Q42E NR5A1 mutation. (**c** and **d**) A photograph of resected gonads. CT: computed tomography; CD: efferent duct.
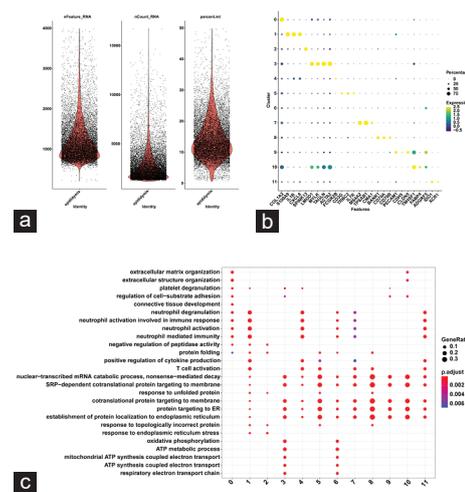
Supplementary Figure 2: QC for the scRNA-seq datasets. (**a**) Distribution of gene numbers, number of UMIs, and percentage of mitochondria (before cell filtering). (**b**) Cell annotation of each epididymal cell population based on cell-type-specific marker genes from the literature (log1pRPM values are centered and scaled by row). (**c**) GO enrichment analysis for the twelve cell clusters. GO: gene ontology; QC: quality control; UMI: unique molecular identifier.

Supplementary Figure 3: (**a**) Cell-type-specific markers in the epididymis epithelial cell cluster (color values are log1pRPM). (**b**) Detailed view of ligand–receptor pairs between epididymis epithelial cells and the other cell types. DC: dendritic cell.

Supplementary Figure 4: CellChat analysis between fibroblasts and epididymis epithelial cells. (**a**) The relative contribution of each ligand–receptor pair of the IGF signaling pathway. (**b**) The communication probability analysis of the IGF signaling pathway between fibroblasts and epididymis epithelial cells, fibroblasts, and clear cells. (**c**) The IGF signaling pathway in the twelve cell clusters. (**d**) The relative importance of cell clusters based on sender, receiver, mediator, and influencer measurement of the IGF signaling network. (**e**) Outgoing and incoming signaling patterns in the twelve cell clusters. DC: dendritic cell; IGF: insulin-like growth factor; ITGA: integrin subunit alpha; ITGB: integrin subunit beta.