

RESEARCH

Open Access



# Reprogramming skin fibroblasts into Sertoli cells: a patient-specific tool to understand effects of genetic variants on gonadal development

Abhinav Parivesh<sup>1†</sup>, Emmanuèle Délot<sup>1†</sup>, Alejandra Reyes<sup>2</sup>, Janelle Ryan<sup>2</sup>, Surajit Bhattacharya<sup>1</sup>, Vincent Harley<sup>2</sup> and Eric Vilain<sup>3\*</sup> 

## Abstract

**Background** Disorders/differences of sex development (DSD) are congenital conditions in which the development of chromosomal, gonadal, or anatomical sex is atypical. With overlapping phenotypes and multiple genes involved, poor diagnostic yields are achieved for many of these conditions. The current DSD diagnostic regimen can be augmented by investigating transcriptome/proteome in vivo, but it is hampered by the unavailability of affected gonadal tissue at the relevant developmental stage. We try to mitigate this limitation by reprogramming readily available skin tissue-derived dermal fibroblasts into Sertoli cells (SC), which could then be deployed for different diagnostic strategies. SCs form the target cell type of choice because they act like an organizing center of embryonic gonadal development and many DSD arise when these developmental processes go awry.

**Methods** We employed a computational predictive algorithm for cell conversions called Mogrify to predict the transcription factors (TFs) required for direct reprogramming of human dermal fibroblasts into SCs. We established trans-differentiation culture conditions where stable transgenic expression of these TFs was achieved in 46, XY adult dermal fibroblasts using lentiviral vectors. The resulting Sertoli like cells (SLCs) were validated for SC phenotype using several approaches.

**Results** SLCs exhibited Sertoli-like morphological and cellular properties as revealed by morphometry and xCelligence cell behavior assays. They also showed Sertoli-specific expression of molecular markers such as SOX9, PTGDS, BMP4, or DMRT1 as revealed by IF imaging, RNAseq and qPCR. The SLC transcriptome shared about two thirds of its differentially expressed genes with a human adult SC transcriptome and expressed markers typical of embryonic SCs. Notably, SLCs lacked expression of most markers of other gonadal cell types such as Leydig, germ, peritubular myoid or granulosa cells.

**Conclusions** The trans-differentiation method was applied to a variety of commercially available 46, XY fibroblasts derived from patients with DSD and to a 46, XX cell line. The DSD SLCs displayed altered levels of trans-differentiation

<sup>†</sup>Abhinav Parivesh and Emmanuèle Délot contributed equally to the manuscript and Vincent Harley and Eric Vilain contributed equally to the manuscript.

\*Correspondence:

Eric Vilain

evilain@hs.uci.edu

Full list of author information is available at the end of the article



in comparison to normal 46, XY-derived SLCs, thus showcasing the robustness of this new trans-differentiation model. Future applications could include using the SLCs to improve definitive diagnosis of DSD in patients with variants of unknown significance.

### Highlights

- Disorders/differences of sex development (DSD) have low diagnostic rates.
- Validation of variants of unknown significance and discovery of developmental pathways can be improved by utilizing gonadal tissue/cell for genomic and transcriptomic analysis.
- As these cells are not readily accessible, we have developed a novel method for trans-differentiating 46, XY dermal fibroblasts into Sertoli-like cells (SLC).
- Our trans-differentiation method can produce SLCs from dermal fibroblasts belonging to a variety of 46,XY DSDs, thus showing its robustness.
- This method can aid as a tool for precision medicine towards diagnosing DSD.

### Plain language summary

Individuals with disorders/differences of sex development (DSD) frequently do not get a specific genetic diagnostic. A limitation in the field is that the relevant cell types that would be needed to study the molecular events occurring at the time of onset of many DSD are found in the embryonic gonad. This, of course, is not accessible for research or diagnostic purposes. We set out to develop a method for directly transforming more accessible cells, from adult skin, into the cells known to organize the male gonad in the embryo, Sertoli cells. A combination of unique transcription factors was stably integrated into skin fibroblasts, and culture under appropriate conditions allowed differentiation into Sertoli-like cells (SLC), but not other gonadal cell types. The SLCs recapitulated known patterns of gene expression, shape, and behavior of Sertoli cells. The method was also tested on commercially available fibroblasts from a variety of DSD genetic backgrounds. The resulting cells exhibited condition-specific behavior (gene expression, adhesion to substrate, division rate...). This method provides a new tool to study molecular events occurring at the time of onset of DSD in the embryonic gonad, and the impact of patient-specific mutations on those. It could allow identification of new developmental pathways (and, thus, new candidate genes for DSD), as well as provide models to validate the impact of variants of unknown significance, or to test approaches to correct the genetic anomaly in patient cells.

### Introduction

DSD are defined as “congenital conditions in which development of chromosomal, gonadal, or anatomical sex is atypical” [1, 2]. They may present in isolation or as part of complex multiorgan developmental syndromes [3–5]. DSD diagnosis combines several approaches including anatomic description and endocrine and genetic testing. Genetic testing is required to disambiguate many conditions who present with a similar phenotype but can be associated with the dozens of genes identified so far [6, 7].

Genetic testing for DSD spans from traditional karyotyping to chromosomal microarrays, to next-generation sequencing approaches (exome or whole-genome sequencing), as well as some more specific approaches for genes refractory to short-read sequencing such as *CYP21A2* [6, 8]. Pre-clinical technology such as optical genome mapping to detect structural variants [9] or long-read sequencing offer the hope of future increased

diagnosis yield. However, at this time only 30–50% of patients receive a specific molecular diagnosis in the research setting, and much lower in many clinical settings [3, 10, 11]. The low diagnostic rate may be because new DSD-associated genes are waiting to be discovered or because currently available technology misses several types of variants. However, a major limitation is also the difficulty in interpreting the significance of variants.

Transcriptome sequencing from blood, muscle, fibroblasts etc., coupled with genetic testing has been shown to significantly increase diagnostic rates in other rare Mendelian conditions [12–15]. However, procuring a relevant tissue to examine the transcriptome is not possible for conditions where pathogenic mechanisms occur early in development as is the case for most DSD conditions. While a biopsy of embryonic gonad is impossible to obtain, skin biopsies can be readily collected for isolating dermal fibroblasts. If these can be reprogrammed into gonadal cell types in vitro, they

could be used for transcriptomic analysis to identify and validate variants in a patient-specific model. While the fetal gonadal transcriptome is still poorly defined, highly complex and dynamic [16], deriving patient-specific cell types that approximate those developing cells should provide useful tools for DSD diagnosis.

As the main organizing center of the typical male gonad, Sertoli cells (SC) are central to the etiology of many DSD [17–19]. They are therefore the prime target choice into which to reprogram somatic cell types for patient-specific transcriptomic diagnostic of DSD. Sertoli cell growth and development in humans happens in three phases: immature SC proliferate in the prenatal and neonatal phase of development; they then remain quiescent for 8–10 years; in the third phase, proliferation restarts in the peri-pubertal age, making the final number of mature, non-proliferative SCs of the adult testis which are characterized by different transcription patterns than immature SCs [20–22]. Expression of several key genes at the genital ridge of chromosomal males triggers a cascade leading to *SRY* expression and to differentiation of Sertoli cells. For example, *GATA4*-*FOG2* interaction induces *SRY* expression [23]. *WT1* activates *NR5A1*/*SF1* expression early in the bipotential stage of the gonad, which supports induction of *SRY* [24, 25]. In the early, bipotential gonad expression of *SRY* is followed by *SOX9*, which triggers the differentiation of primordial SCs. SCs form the physical framework of seminiferous cords, which provide nutritional and structural support to developing primordial germ cells, peritubular myoid cells, endothelial cells. SCs also drive the differentiation of androgen-secreting Leydig cells, which in turn shapes the male pattern of sexual development [19, 26, 27]. Variants in all these genes can lead to DSD in humans [6, 7]. For example, mutations, deletions, inversions in *SOX9*, as well as in its upstream promoter and enhancer regions cause campomelic dysplasia, a condition associating skeletal malformations and to sex reversal in XY individuals [28–32].

Transient trans-differentiation of pluripotent stem cells (iPSCs) into Sertoli-like fate was achieved with specialized small molecules or biological factor supplementation [33, 34]. A transgenic approach was able to produce potentially stable Sertoli-like cells using murine fibroblasts [35] but this hadn't been reproducible with human cells. A recent report attempted human fibroblasts to Sertoli trans-differentiation using a stable transgenic approach with just two transcription factors: *NR5A1* and *GATA4* [36]. It is however not clearly specified whether the resulting cells express markers of other gonadal cell types, and specificity of the method is difficult to ascertain.

Here we present a method of transgenic expression of a novel combination of transcription factors (TFs) to produce Sertoli-like cells (SLC), as predicted by the computational predictive network Mogrify [37], using lentiviral vectors and suitable culture conditions. We transduced adult 46, XY human dermal fibroblasts, and validated the approach by looking for the appearance of Sertoli-specific molecular signatures and cellular phenotypes. Further, we applied this method to dermal fibroblasts derived from a variety of genetic backgrounds, including 46, XX as well as 46,XY DSD patients to test the robustness of the method in achieving trans-differentiation to a Sertoli-like phenotype.

## Materials and methods

### Plasmids, lentiviral particle generation and titration

Second-generation lentiviral packaging plasmid pMD2.G encoding HIV-1-derived gag (structural protein), pol (retroviral enzyme), tat (transcriptional regulator) and rev (post-transcriptional regulator) and the envelope plasmid psPAX2 encoding VSV-G coat protein were purchased from Addgene repository. Eight different plasmids with human *FGF2*, *GATA6*, *GATA4*, *MXII*, *JUNB*, *NFYB*, *NR5A1* and *EBF1* respectively, cloned into the pEZ-Lv165 lentiviral transfer vector system were purchased from GeneCopoeia, Inc. (Rockville, MD 20850, USA). pEZ-Lv165 vector system allows for robust parallel expression of the cloned gene and GFP marker protein separated by IRES2 sequence under the control of a strong EF1- $\alpha$  bi-cistronic promoter and it lacks mammalian antibiotic selection.

HEK-293 T cells were seeded at  $2.5 \times 10^6$  in 25 mL of DMEM (Genesee Cat. 25-500) completed with 10% heat-inactivated FBS (Thermo Cat. 16000044) in five 150 mm tissue culture dishes (Millipore Sigma Cat. CLS430599) and cultured for three days at 37 °C in humidified tissue culture incubator with 5% CO<sub>2</sub> supply to reach 80% confluency. 2 h before transfections, the media was replaced with 22.5 mL of fresh complete media. Plasmid transfection was carried out by the calcium chloride method. 112.5  $\mu$ g of transfer vector plasmid with gene of interest, 39.5  $\mu$ g of packaging plasmid and 73  $\mu$ g of envelope plasmid were mixed in a conical tube. 3.3 mL of 0.1X Tris-EDTA pH 8.0, 1.75 mL of 2.5 mM HEPES buffered distilled water pH 7.3, and 565  $\mu$ L of 2.5 M CaCl<sub>2</sub> was added and mixed well by pipetting. 5.7 mL of 2 $\times$ HEPES buffered saline (0.28 M NaCl, 50 mM HEPES, 1.5 mM Na<sub>2</sub>HPO<sub>4</sub> dissolved in dH<sub>2</sub>O and pH adjusted to 7.0 with 10N NaOH) was added dropwise to the mixture under vigorous agitation and incubated at room temperature for 15 min. 2.25 mL of this transfection mix was added into each of the five dishes, they were swirled gently to mix and put back into the incubator overnight/12 h. Early the

next morning, the transfection media in each dish was replaced with fresh 14 mL complete media for the transfected cells to generate lentiviral particles and release into the media. Three subsequent harvests of lentiviral particle-containing media were collected 8 h, 12 h and 8 h apart from the five dishes and entire ~210 mL of the harvest was pooled together, stored at 4 °C. The pooled harvest was centrifuged at 500 G for 5 min at 4 °C to pellet the cell debris and filter sterilized using 0.22 microns filter cup (Sigma Millipore Cat. CLS430769). Equal volumes were distributed into 6 ultra-clear thin wall tubes (Beckman Coulter Cat. 344058) and ultra-centrifuged at 50,000 g for 2 h at 16 °C in a Beckman Coulter Optima XE ultracentrifuge to precipitate lentiviral particles. The media supernatant was aspirated out and the lentiviral pellet was left to dry at RT for 10 min. The pellet in each of the six tubes was dissolved in 17.5 µL of DPBS (Genesee Cat. 25-508) by gentle pipetting, avoiding air bubble formation and pooled. The pooled concentrated lentivirus was divided into 20 µL aliquots and stored at -80 °C. All the surfaces, used dishes and spent media were thoroughly decontaminated with bleach before discarding.

The lentivirus was titrated by estimating the GFP positive cell population using flow cytometry.  $1 \times 10^5$  HEK 293 T cells were re-suspended in 0.5 mL of complete media and seeded into each well of a 12-well dish (Millipore Sigma Cat. CLS3513). On the next day, cells were transduced with serial dilutions of 1 µL,  $10^{-1}$  µL,  $10^{-2}$  µL,  $10^{-3}$  µL and  $10^{-4}$  µL lentivirus mixed in 0.5 mL media and 0.5 µL of 100 mg/mL Polybrene (Millipore Sigma Cat. TR-1003-G) in duplicate wells. One of the wells was used to harvest and count cell numbers at the time of transduction (day 1 cell count). The dish was left to incubate for 3 days, the transduced cells were harvested by trypsinization and fixed in 1% formaldehyde-DPBS for 5 min at RT. They were washed once with DPBS and finally re-suspended in DPBS. The samples were analyzed on a Beckman Coulter Cytotflex S flow cytometer instrument for green fluorescence to determine GFP-positive cell percentage. Lentiviral dilution samples yielding 1–20% GFP positives were chosen for titer calculations as any percentage higher than that insinuates multiple expressible viral integration events per cell in the population and less than 1% can be erroneous/false positive. Viral titer concentration was calculated in transducing units / µL as (day 1 cell count x % GFP-positive cells/100)/volume of lentivirus in µL.

#### Cell lines and trans-differentiation culture

Control human adult 46, XY primary dermal fibroblast (HDFa) (Cat. ATCC PCS-201-012) was purchased from American Type Culture Collection repository (ATCC). The following cell lines were purchased from Coriell cell

repository: GM03651: Skin fibroblasts derived from a 25-year-old 46, XX healthy female; GM00083: Gonadal fibroblasts derived from a 1-year-old 46, XY female; SRXY1; GM06938: fibroblasts derived from a 3-week-old male with a *WT1* deletion; GM21894: skin fibroblasts derived from a 46,XY patient with campomelic dysplasia; and GM03368: gonadal fibroblasts derived from a 17-year-old 46, XY woman. HDFa was cultured in Medium 106 (Thermo Cat.M106-500), completed with low serum growth supplement (LSGS) (Thermo Cat. S00310), and 1X penicillin–streptomycin antibiotic (P/S) (Thermo Cat. 15140122). GM03651, GM00083, GM06938 and GM12896 were cultured in Eagle's minimum essential medium (Thermo Cat. M0325) completed with 15% non-inactivated FBS (Thermo Cat. 16000044) and 1X P/S (Thermo Cat. 15140122). GM21894 was cultured in DMEM (Genesee Cat. 25–500) completed with 10% non-inactivated FBS and 1X P/S. All these cells were grown as adherent monolayer cultures in tissue culture-treated T25 flasks (Genesee Cat. 25-207), harvested by trypsinization upon 80–90% confluence, quenched by trypsin neutralization solution (Thermo Cat. R002100), spun down at 150 g for 4 min at RT and sub-cultured in 5 ml of suitable complete media. The media was replaced every alternate day in all these cultures until they reached confluency.

For the trans-differentiation culture, fibroblasts were seeded at  $2.25 \times 10^4$  into each well of 6-well dishes (Millipore Sigma Cat. CLS3516) in 2 mL complete M106 on Day 0 and grown overnight. On Day 1, cells were counted from an extra well to determine the number of lentiviral particles required to transduce the cells at a multiplicity of infection (MOI) of five. Three transduction sets were designed: A. GFP; B. 8 TFs: *FGF2*, *GATA6*, *GATA4*, *MXII*, *JUNB*, *NFYB*, *NR5A1* and *EBF1*; C. 6TFs: *FGF2*, *GATA6*, *GATA4*, *MXII*, *JUNB*, and *NFYB*. The old media in wells was replaced with transduction media consisting of 2 mL complete M106 mixed with an appropriate amount of respective lentivirus cocktail and 2 µL of 100 mg/mL Polybrene (Millipore Sigma Cat. TR-1003-G). The cells were left to transduce for 12 h/overnight. On Day 2, transduction media was replaced with fresh complete M106 and cultured in it for 4 days with alternate days media change. On day 5, M106 was replaced with 2 mL of trans-differentiation media consisting of basal Sertoli Cell Media (SerCM) completed with 5% FBS, 1X Sertoli cell growth supplement (SerCGS) and 1X P/S (ScienCell Cat. 4521). The culture was maintained in this media for another 24 days until day 30 with alternate days media change.

16 transduction conditions were designed for the pilot study. Conditions 1,2,3,4,9,10 exhibited low transduction efficiency while conditions 15, 16 showed high levels of

post-transduction mortality (data not shown, Additional file 1: Fig. S1C). Conditions 5, 6, 7, 8, 11, 12, 13, and 14 showed fairly high levels of transduction efficiency, minimal cell mortality, and variable degrees of morphological change in comparison to their respective GFP controls (Additional file 1: Fig. S1C). Appearance of a few Sertoli markers were explored using qPCR. Mature stage Sertoli marker *AR* (Additional file 1: Table S1) didn't appear in any of the conditions while all-stage marker *BMP4* appeared in conditions 5–8 and 13–16 (Additional file 1: Fig. S1-1B). Immature stage Sertoli marker *PTGDS* appeared in conditions 5, 6, 7, 8 while critical all-stage Sertoli marker *SOX9* appeared to be significantly expressed in conditions 5, 6 and to a lesser extent in 7, 8 (Additional file 1: Fig. S1B).

Stable viral transductions were confirmed by upregulation of each of the 8 TFs or 6 TFs after 1-month transdifferentiation culture. The 8TF and 6TF SLCs exhibited well over 1.5 fold increase for each of the 8 or 6 TFs. Furthermore, *NR5A1* and *EBF1* didn't show an increase in 6TF SLCs, showing specificity of transduction (Additional file 1: Fig. S2E, Fig. S5-3A–D).

#### RT-PCR, qPCR, and target gene expression fold change analysis

Cells were lysed and total RNA was isolated using RNeasy micro kit (Qiagen Cat. 74004) following the manufacturer's instructions. All the qPCR and RT-PCR reagents were purchased from Thermo Scientific. Total RNA was reverse transcribed into cDNA by RT-PCR using High-Capacity cDNA Reverse Transcription Kit (Cat. 4368814) following the manufacturer's instructions on a BioRad C1000 thermal cycler. The following thermocycling conditions were used: 25 °C-10 min, 37 °C-120 min, 85 °C-5 min, 4 °C-hold. The cDNA was used to perform quantitative PCR using Taqman chemistry reagents. qPCR reactions were set up in triplicates using 5 ng of cDNA per reaction using TaqMan™ Fast Advanced Master Mix (Cat. 4444556) with TaqMan™ Gene Expression Assay-VIC probe (Cat. 4448490 Assay ID: Hs02758991\_g1) for *GAPDH* housekeeping gene and respective target genes. Following TaqMan™ Gene Expression Assay (FAM) target gene probes were used: *SOX9* (Cat. 4331182 Assay ID: Hs00165814\_m1), *BMP4* (Cat. 4331182 Assay ID: Hs00370078\_m1), *PTGDS* (Cat. 4331182 Assay ID: Hs00168748\_m1), *AR* (Cat. 4331182 Assay ID: Hs00171172\_m1), *HSD3B2* (Cat. 4331182 Assay ID: Hs00605123\_m1), *DDX4* (Cat. 4331182 Assay ID: Hs00987125\_m1), *FOXL2* (Cat. 4331182 Assay ID: Hs00846401\_s1), *ACTA2* (Cat. 4331182 Assay ID: Hs00426835\_g1). 12 µL reactions were set up following manufacturer's instructions in MicroAmp™ EnduraPlate™ Optical 384-Well Clear

GPLE Reaction Plates with Barcode (Cat. 4483319) on a cooling block, sealed, spun down briefly and subject to PCR on a QuantStudio™ 7 Flex Real-Time PCR platform (Cat. 4485701) using manufacturer suggested default thermocycling conditions for 40 amplification cycles. The target gene expression fold change in experimental over control was calculated using inverse comparative threshold cycle ( $\Delta\Delta C_t$ ) method after normalizing for *GAPDH* as the internal loading control.

#### Live cell imaging and morphometry

GFP, 8TF and 6TF transduced 1-month cultures growing in 6-well dishes and expressing GFP were imaged under the green fluorescence channel of a Leica DMIL LED Inverted Fluorescence Microscope using Leica's LasX imaging software. The images were exported in TIFF format and used for morphometric analysis. Morphometry was performed using MetaMorph software. The freehand tool in the software was used to draw the outlines of GFP expressing cells, which then computed the perimeter as well as area enclosed within that cell perimeter in arbitrary units. An average of 60–80 cells were counted per group and the group statistics were compared using unpaired t test. Cell surface area measured in arbitrary units was used as a readout for cell size as these are two-dimensional adherent cell cultures, while an arbitrary index called shape factor, defined as  $2 \cdot \pi \cdot \text{area} / \text{perimeter}^2$  was utilized as a proxy measurement of cell shape.

#### Immunofluorescence staining and imaging

For immunofluorescence staining, the cells were fixed with 4% paraformaldehyde for seven minutes and washed three times with phosphate-buffered saline (PBS). In brief, slides were blocked for 60 min with 1% donkey serum, followed by incubation with primary antibodies overnight at 4 °C, including anti-SOX9 rabbit polyclonal (Sigma-Aldrich AB5535, 1:400), anti-GATA4 goat polyclonal (Santa Cruz SC-1237, 1:500), anti-DMRT1 rabbit polyclonal (Santa Cruz SC-98341, 1:200), anti-AMH mouse monoclonal (Santa Cruz SC-166752, 1:400), or anti-DHH goat polyclonal (Santa Cruz SC-1193, 1:300). After washing three times in PBS for 5 min each, the slides were incubated with the corresponding secondary antibody at a 1:1000 dilution (ThermoFisher A-21207, ThermoFisher A-11058, ThermoFisher A10037 or ThermoFisher A-31573.). DAPI (4,6-diamidino-2-phenylindole, Invitrogen) was used to label the nuclei. Cells were observed for epifluorescence under a Nikon C1 confocal microscope or Olympus FV1200 confocal microscope under standard conditions.

### RNA sequencing: library preparation with polyA selection, HiSeq sequencing and data analysis

RNA library preparations and sequencing reactions were conducted at GENEWIZ, LLC. (South Plainfield, NJ, USA). RNA samples received were quantified using Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) and RNA integrity was checked using Agilent TapeStation 4200 (Agilent Technologies, Palo Alto, CA, USA).

RNA sequencing libraries were prepared using the NEBNext Ultra RNA Library Prep Kit from Illumina using the manufacturer's instructions (NEB, Ipswich, MA, USA). Briefly, mRNAs were initially enriched with Oligod(T) beads. Enriched mRNAs were fragmented for 15 min at 94 °C. First-strand and second-strand cDNA were subsequently synthesized. cDNA fragments were end-repaired and adenylated at 3'ends, and universal adapters were ligated to cDNA fragments, followed by index addition and library enrichment by PCR with limited cycles. The sequencing library was validated on the Agilent TapeStation (Agilent Technologies, Palo Alto, CA, USA), and quantified by using Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) as well as by quantitative PCR (KAPA Biosystems, Wilmington, MA, USA).

The sequencing libraries were clustered on a single lane of a flowcell. After clustering, the flowcell was loaded on the Illumina HiSeq instrument (4000 or equivalent) according to manufacturer's instructions. The samples were sequenced using a 2×150 bp Paired-End (PE) configuration. Image analysis and base calling were conducted by the HiSeq Control Software (HCS). Raw sequence data (.bcl files) generated from Illumina HiSeq was converted into fastq files and de-multiplexed using Illumina's bcl2fastq 2.17 software. One mismatch was allowed for index sequence identification.

### RNAseq data analysis

BCL files generated from sequencer were converted using bcl2fastq (Illumina) to fastq files, quality-checked using *fastqc* (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Adapter and quality trimming was performed using *trimgalore* ([https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)). Summary QC files for before and after quality trimming was performed using MultiQC [38]. Alignment to hg38 reference was performed using *STAR* [39], followed by count estimation using *RSEM* [40]. For differential expression analysis, read counts were estimated using *tximport* [41], followed by normalization and differential expression analysis performed using *DESeq2* [42]. All significantly expressed genes were above a threshold of absolute log<sub>2</sub> fold change of 1.5 and p-value < 0.05. Gene ontology was

performed using *gProfiler* [43], where we extracted only biological processes. Visualization was performed using heatmap and ggplot2 R packages.

### Adhesion and proliferation xCELLigence assays on SLCs

For the xCELLigence assays Sertoli-like cells were lifted using 0.05% trypsin and counted on a haemocytometer. An xCELLigence E-Plate (ACEA BioSciences, VIEW 96 PET plate) was prepared with 50 µL of complete growth media in each well. All conditions were run in technical triplicate. GFP controls, 8TF SLC and 6TF SLC derived from three cell lines—HDFa (46, XY; healthy male), GM21894 (46, XY; CD) and GM03651 (46, XX; healthy female)—were analyzed using this technique. The E-plate was engaged into the xCELLigence Real Time Cell Analysis (RTCA) Instrument (Agilent) and background measurement of the wells plus growth media was recorded before adding cells in 150 µL of media for each SLC condition (5×10<sup>3</sup> cells/well). The E-plates were re-engaged onto the xCELLigence analyzer and incubated for 48 h at 37 °C in a 5% CO<sub>2</sub> atmosphere. Impedance values were recorded every 2 min for the first 6 h and adhesion data derived. Between 6 and 48 h impedance measurements were recorded every 15 min to derive the proliferation date. The value of the slope, i.e. the change in adhesion/proliferation over time, was used for data analysis. Statistical analysis was performed in Graphpad Prism 5 using a Student's t-test.

## Results

### Adult human 46, XY dermal fibroblast to Sertoli cell trans-differentiation strategy produces Sertoli like cells (SLC) that exhibit a significant change in morphology

We utilized the predictive software Mogrify, to predict the TFs necessary to direct differentiation of fibroblasts into Sertoli cells. Mogrify is a comprehensive atlas of human cell conversions that utilizes gene expression data with regulatory network information from ~300 different cell and tissue FANTOM5 data sets to compute a ranked list of transcription factors that cumulatively exert up to 98% differential regulatory influence on the target cell type in comparison to starting cell type [37]. Mogrify predicted that transgenic expression of eight TFs, *FGF2*, *GATA6*, *GATA4*, *MXII*, *JUNB*, *NFYB*, *NR5A1* and *EBF1*, was required for human dermal fibroblast to Sertoli cell trans-differentiation. It also predicted that just six out of those eight, excluding *NR5A1* and *EBF1*, were enough to achieve a cumulative 95% Sertoli influence on dermal fibroblasts (Fig. 1A). *Fgf2* and *Gata4* are differentially expressed between male and female gonads of mice at developmentally important stages, as revealed by RNA-seq analyses [44, 45]. *NR5A1* is a well-known human DSD gene [46].

In a pilot study, we tested a strategy for trans-differentiation of 46, XY normal adult human dermal fibroblasts using 16 different conditions (Additional file 1: Fig. S1A) stably expressing 8 or 6 TFs (and GFP) at different multiplicity of infections with different transduction and trans-differentiation media. The resulting cells were imaged and harvested to measure expression of Sertoli markers by qPCR. Considering optimal transduction efficiency, cell survival/proliferation, morphological appearance, and expression of Sertoli markers *SOX9*, *PGDS* and *BMP4*, conditions 6 and 5 were chosen to be pursued (and will be called “6TF” and “8TF” respectively below) (Fig. 1B).

One of the defining features of cells is their morphology. Of the two cell types relevant to this study, fibroblasts exhibit an elongated spindle-shaped morphology, while Sertoli cells are known to assume squamous epithelial morphology. Sertoli cells grown in adherent cultures have been reported to appear large, have a polygonal shape and exhibit long cytoplasmic extensions or filamentous processes of irregular lengths radiating out of the cell body, along with irregularly shaped nucleus [47–50]. Cells were imaged weekly under green fluorescence for the duration of the one-month culture (Fig. 1C). The size and shape of cells were investigated by morphometric analysis of cell images with recording of the surface area and perimeter of the individual cells.

Shape factor measurement revealed that while there was no significant difference in the shape between GFP transduced control and 6TF, 8TF SLCs in the first week, the shape of 6TF and 8TF SLCs had diverged significantly from GFP control starting at the 2nd week (Fig. 1D). While the GFP control cells remained elongated in shape, characteristic of fibroblasts, 8TF cells started to assume more polygonal appearance with elongated and irregular cytoplasmic processes, typical of Sertoli cells. These irregular processes were even more accentuated in 6TF cells, as observed visually and quantified by a lower shape value in 6TF SLCs than the more circular or polygonal 8TF SLCs (Fig. 1D). Additionally, the GFP control and 6TF, 8TF cells appeared to be similar in size in the first week but the experimental cells started flattening, spreading and growing in size in comparison to GFP control beginning in the second week and this difference became

even more enhanced with each passing week (Fig. 1C, E). All the cells, including GFP controls exhibited significant growth in size at the end of 1 month in comparison to when they started (Fig. 1E). Together, these morphometric evaluations showed that the 1 month of trans-differentiation culture gradually pushed fibroblasts to change in size and shape and to start exhibiting some morphological features associated with Sertoli cells.

### SLCs exhibit Sertoli molecular signatures

To investigate the Sertoli features of transdifferentiated cells further we compared their transcriptome with that of a published adult Sertoli cell (aSC) transcriptome [48], and curated lists of Sertoli marker genes (Additional file 1: Table S1).

Principal component analysis (PCA) of RNA seq transcriptomes revealed that replicates of 6TF SLC cluster away from GFPC, while 8TF did not (Fig. 2A). Replicates of each category of samples grouped together and away from samples of other category (Fig. 2A).

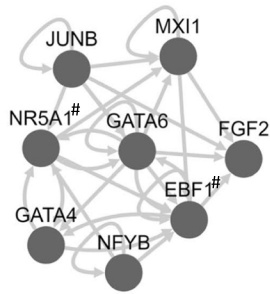
To normalize potential differences owing to their different origins, all three groups (6TF, 8TF and aSC) were first compared with 46,XY-derived GFP control (GFPC) to determine the set of differentially expressed genes (DEGs). Of the 1870 genes differentially expressed between 6TF and control GFPC, 2/3 were in common with the aSC DEGs (Fig. 2C). In 8TF cells, this percentage was almost 3/4 (502/681 genes) (Fig. 2B). Such high percentages of co-differentially expressed genes between the transcriptomes of SLCs and adult Sertoli cells indicated that our experimental SLCs had significant Sertoli-like gene expression patterns.

Gene ontology (GO) analysis of the co-DEGs determined above identified several terms expected to be associated with Sertoli cell as organizing center of male reproductive system structure and seminiferous tubule development (Additional file 1: Fig. S2C, D). The top twenty GO biological processes list included developmental process, organ development, urinogenital system development, tube morphogenesis, external encapsulating structure organization, ECM structure and matrix organization, tube development, with reproductive system development, and reproductive structure development further down the list. It also showed terms for

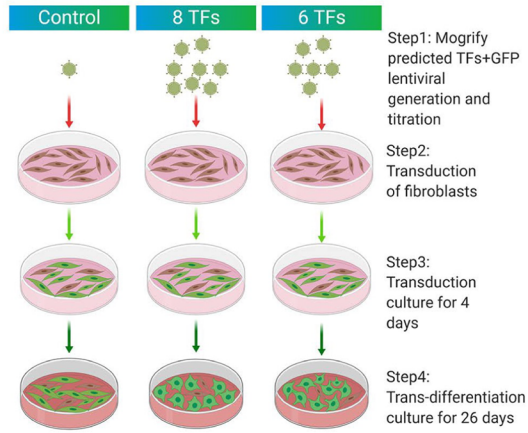
(See figure on next page.)

**Fig. 1** 46,XY fibroblast derived SLCs exhibit significant change in morphology over the 1-month trans-differentiation and show subcellular expression of Sertoli markers. **A** Mogrify-predicted transcription factors required for dermal fibroblasts to Sertoli cell trans-differentiation, # = not required to reach 95% coverage of the required genes. **B** Outline of the trans-differentiation strategy employed throughout this report. **C** Representative images of live cells belonging to the indicated groups expressing GFP across the 1-month trans-differentiation culture of 46, XY dermal fibroblasts. Morphometric analysis of cells in **(C)** showing shape factor **(D)** and area quantification **(E)** (N = 3, n = 50–60 for each group). \*Represent p values calculated from Mann Whitney statistical tests conducted between the indicated groups

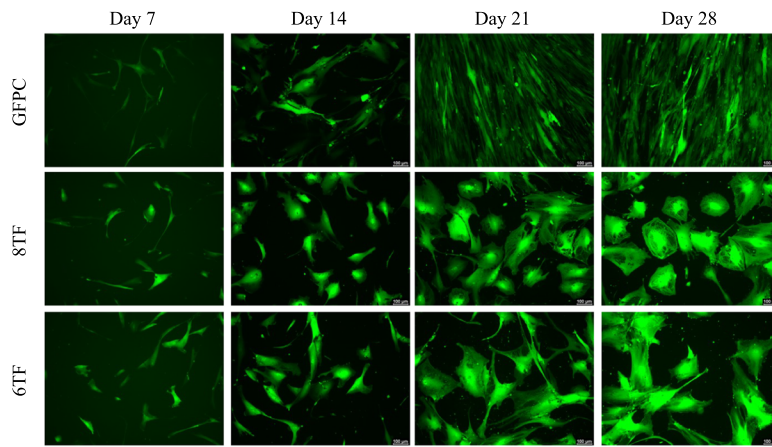
A



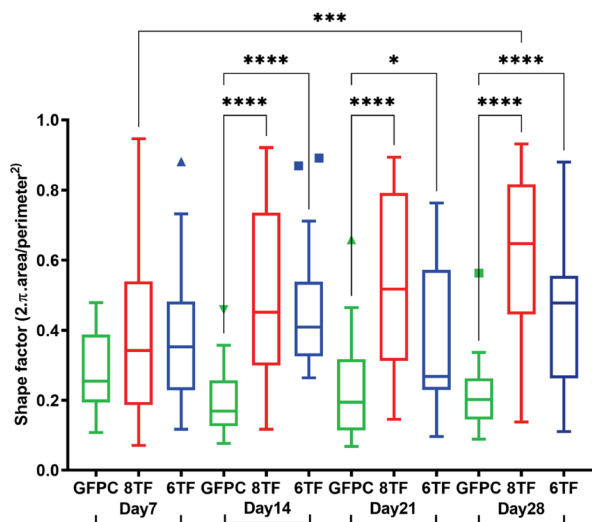
B



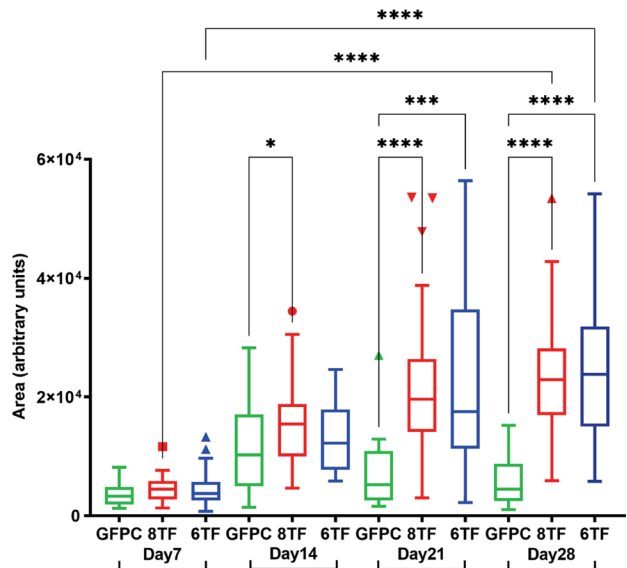
C



D



E



**Fig. 1** (See legend on previous page.)



positive and negative regulation of lipid biosynthetic and metabolic processes, congruent with the known lipid metabolism and lipid droplet formation in Sertoli cells [50, 51].

Next, we compared the gene expression profile of the 1253 co-DEG in 6TF SLC, GFPC and aSC in a heatmap (Additional file 1: Fig. S2A). [A similar heatmap looking at expression of the 502 8TF-GFPC co-DEG genes is shown in Additional file 1: Fig. S2B]. Each 6TF or 8TF SLC replicate appears quite different from GFPC and notably, resemble aSC expression to a large extent.

In order to investigate the molecular phenotype, a gene list of immature, mature, and all-stage Sertoli markers was drawn up from literature (Additional file 1: Table S1). The list also includes markers for non-Sertoli male gonadal cell types such as Leydig, myeloid, and germ cells, as well as for granulosa cells, the female homologue of Sertoli cells.

6TF SLC exhibited an increase in expression of bipotential gonad markers *GADD45G*, *ZFPM*; immature Sertoli markers *KRT18*, *INHBA*, *NACAM1/2*, *TGFA*; mature Sertoli markers *IL1A*, *CX43/GJA1*; and all-stage-Sertoli markers *SOX9*, *KITLG*, *GDNF*, *BMP4* over GFPC (Fig. 2D, Additional file 1: Fig. S2E).

In contrast, myeloid (*ACTA2*) and granulosa (*FOXL2*) cell markers did not show an increase; Leydig cell markers *CYP11A1* and *STAR* decreased (Fig. 2D). (Expression of germ cell marker *DDX4* was undetected; data not shown). This indicates that the six transcription factors *GATA4*, *GATA6*, *MXII*, *JUNB*, *NFYB* and *FGF2* were sufficient to push dermal fibroblasts towards a Sertoli fate with predominant expression of immature and all-stage markers.

8TF SLC exhibited a significant increase in the expression of bipotential gonad marker *GADD45G*, immature Sertoli markers *CYP26B1*, *PTGDS*, *KRT18*, and *NCAM1/2*, and all-stage Sertoli markers *SOX9* and *BMP4* over GFPC (Additional file 1: Fig. S1E). 8TF SLC also exhibited a similar pattern for mature markers (*CTSL*, *IL1A* and *CX43/GJA1*) with lower fold increase than for immature markers. While 8TF SLCs also didn't show an increase in expression of myeloid marker *ACTA2* or granulosa marker *FOXL2*, they did exhibit a

strong increase in expression of Leydig markers *HSD3B2*, *CYP11A* and *STAR* (Fig. 2D, Additional file 1: Fig. S1E). This suggests that 6TF is a better condition for producing SLCs as 8TF additionally produce Leydig (or adrenal) cell characteristics.

Expression of several markers was further validated by qPCR, which confirmed the RNASeq findings. All-stage Sertoli marker *SOX9* and *BMP4* showed 46-fold and 3.5-fold increases in 6TF SLCs over GFPC, respectively (Additional file 1: Fig. S1F). Immature marker *PTGDS* showed a twofold increase on linear scale. 6TF SLC showed no change in germ cell marker *DDX4* or in granulosa marker *FOXL2* and myeloid marker *ACTA2*. (Similar results were found for 8TF; Additional file 1: Fig. S1F).

Finally, the molecular phenotype was ascertained at the protein level. Sertoli cells express the transcription factors *SOX9* and *DMRT1* in the nucleus across all stages [52]. Sertoli cells also exhibit cytoplasmic expression of signaling molecules AMH during the immature stage and DHH in all stages (Additional file 1: Table S1). Immunofluorescent imaging revealed that all these four markers exhibited Sertoli-specific sub-cellular expression in 6TF SLCs as well as in 8TF SLC (Fig. 2E–H). The percentage of the cell population showing these appropriate expression patterns was higher in 6TF than in 8TF confirming that 6TF is likely the better transduction condition for Sertoli-like cells.

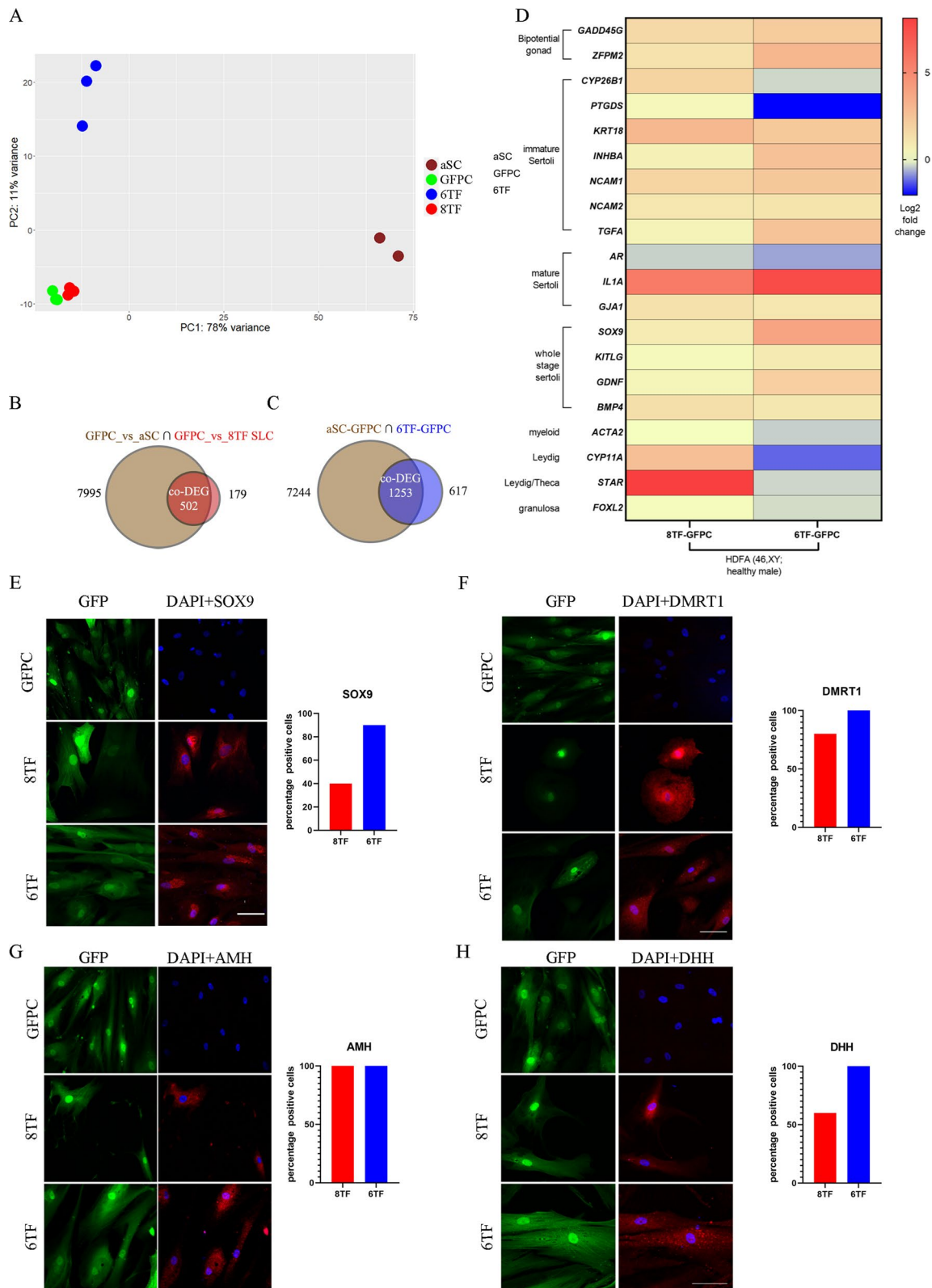
#### Trans-differentiation strategy applied to dermal fibroblasts of different 46, XY DSD genetic backgrounds and 46,XX produced SLCs that exhibit morphological changes

The method described here was able to trans-differentiate dermal fibroblasts derived from a 46, XY adult male into Sertoli-like cells. To test the robustness of the method, it was applied to commercially available dermal fibroblasts from individuals with various DS including:

- 46, XY SRXY1 (46,XY female “sex reversal” of unknown genetic etiology. No variant in known DSD genes was found by exome sequencing (not shown).
- 46,XY CD: Campomelic Dysplasia. A 3.2 Mb heterozygous deletion including *SOX9* was identified

(See figure on next page.)

**Fig. 2** 46,XY SLCs show Sertoli marker gene expression: Whole genome transcriptional profiling of HDFa SLCs. **A** Principal Component Analysis of all the transcriptomes under consideration in this analysis. **B** Venn diagrams displaying differentially expressed genes (DEGs, fold change > 1.5, p adjusted < 0.05) between GFPC (N = 3) vs adult Sertoli cells (aSC, in brown) (N = 2) and GFPC (N = 3) vs 8TF (red) and **C** 6TF (blue) SLCs (N = 3). Intersection areas show co-differentially expressed genes (co-DEGs). **D** Heat map showing 6TF and 8TF vs. GFPC log2 fold change for indicated gonadal marker genes measured by RNAseq (n = 3). **E–H** Representative immuno-fluorescence images showing subcellular expression and quantification of percentage of cells showing such expression for 6TF SLCs, 8TF SLCs, and GFP controls (N = 3, n = 50–60 for each group) for **(E)** *SOX9* **(F)** *DMRT1* **(G)** AMH and **(H)** DHH



**Fig. 2** (See legend on previous page.)

by whole genome sequencing (Additional file 1: Fig. S5-S4),

- 46, XY WT1, where a typical large deletion including the *WT1* gene was demonstrated by the company.
- a 46,XX healthy female.

These four cell lines were separately seeded, transduced and cultured in the same way as described previously to produce 6TF and 8TF SLC and GFPC for each line. The live cultures were imaged on Day 7 and Day 28 of the culture under the GFP filter and the images were analyzed for changes in shape (shape factor) and size (area).

The 6TF SLCs for three 46, XY DSDs showed significant changes in shape (Fig. 3A–C) and size (Additional file 1: Fig. S3A–C) over the 1-month differentiation period. In contrast, the 46, XX 6TF SLC did not exhibit a significant shape change but did show a size increase like the others (Fig. 3D, Additional file 1: Fig. S3D). [Data for 8TF are shown in Additional file 1: Fig. S3A–D].

We also compared these parameters in 1-month-old experiment groups of all these DSD and female lines with that of control 46,XY. While the GFPC of the four lines had significantly different shapes at the start of the culture, these differences ebbed after the 6TF trans-differentiation process (Fig. 3E). Interestingly, this was coupled with persistent differences in sizes of XY DSD and XX-derived GFPC and 6TF cells with those of control 46,XY (Additional file 1: Fig. S3E). This showed that cells derived from DSD individuals can also undergo transdifferentiation but with different characteristics than the control 46,XY cells and each other, possibly reflecting the different etiology of the DSD condition.

#### Adhesion and proliferation of SLCs from XY, XX, and XY DSD fibroblast lines

A key feature of Sertoli cells is their ability to adhere to form tubules then proliferate. SLCs derived from 46,XY, 46,XY with CD and 46,XX were analysed using xCELLigence Real Time Cell Analysis (RTCA) assays to measure their adhesion and proliferation properties. 0–6 h covers the time for adhesion of SLC suspensions to culture surface, while 6–48 h covers the period for cell proliferation. The Cell Index (CI) was plotted over time (Additional file 1: Fig. S4B–F) and the slope of CI curve was used as a measure to discern the differences or similarities between different experimental groups in their adhesion and proliferation phenotypes. The 46, XY 6TF SLCs showed lower adhesion (Fig. 4A) and slower proliferation (Fig. 4B) than the GFP control. This might be indicative of cell differentiation, which usually correlates negatively with proliferation. By contrast, 46 XY 8TF SLCs exhibited higher adhesion than GFPC while showing slower proliferation (Additional file 1: Fig. S4A).

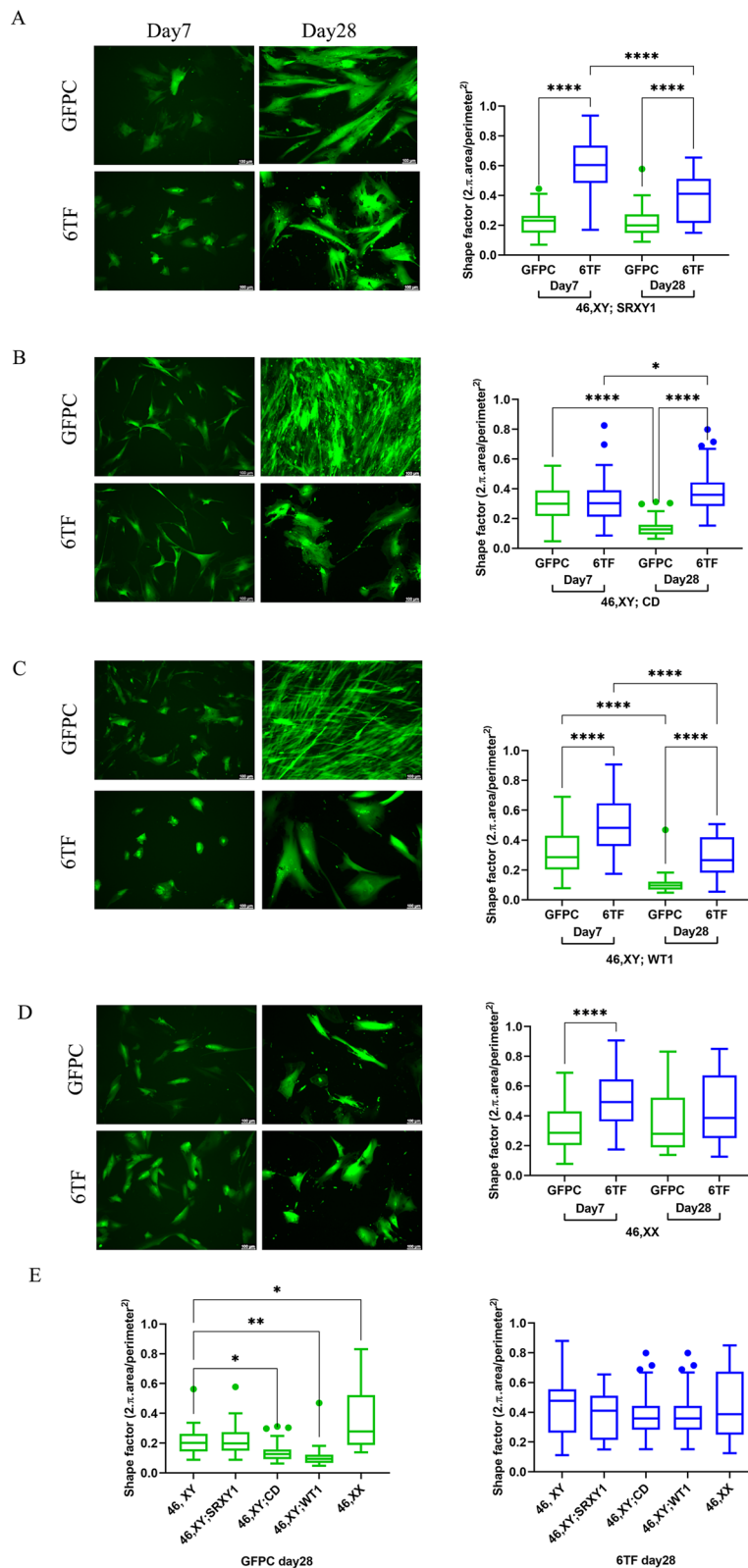
We also analyzed the adhesive and proliferative abilities of SLCs derived from the 46, XY with CD and 46,XX cells. The GFP control fibroblasts of all three cell lines showed similar adhesion and proliferation profiles (Fig. 4A, B) despite their genetic differences. Sex dimorphism was observed in the adhesive abilities of the 46,XY SLCs and 46,XX SLCs, with the 46,XX SLCs showing a significantly higher adhesive ability (Fig. 4A). The 46,XY;CD SLCs showed an intermediate adhesive ability, significantly higher than the 46,XY SLCs but significantly lower than the 46,XX SLCs. A trend towards sex dimorphism was also observed when comparing the proliferation of 46,XY SLCs and 46, XX SLCs although this did not reach statistical significance (Fig. 4B). The 46,XY;CD SLCs once again showed significantly higher proliferation than the 46,XY SLCs, to a comparable level with the 46,XX SLCs (Fig. 4B).

#### XX 46,XX and 46,XY DSD fibroblasts derived SLC exhibited altered levels of Sertoli marker gene expression

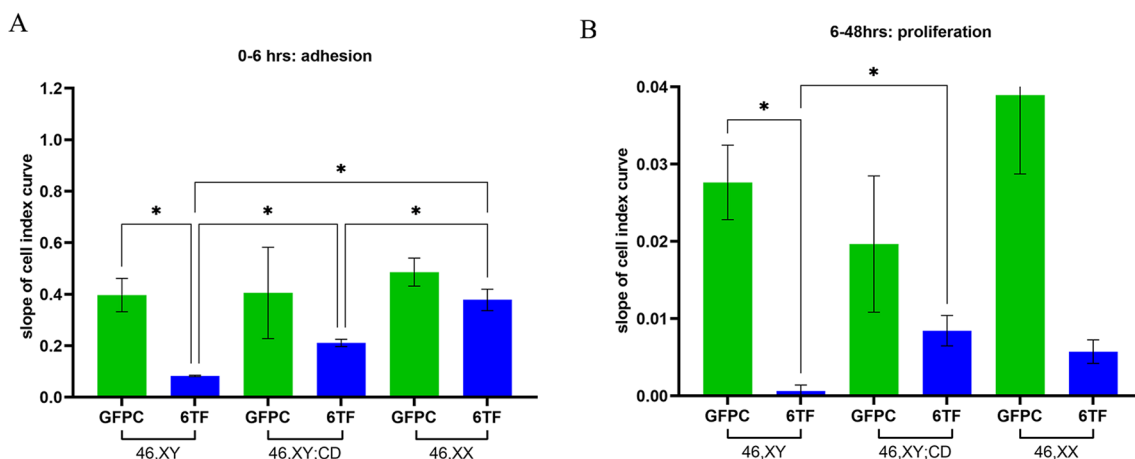
The 46,XX- and 46,XY DSD-derived SLCs were evaluated for expression of Sertoli cell markers. Principal component analysis (PCA) of transcriptomes showed that all the samples belonging to each of the four cell lines clustered away from aSC as observed earlier for 46,XY SLCs (Fig. 2A). The 6TF and 8TF samples clustered away from GFPC in 46, XY; SRY1, 46,XY; WT1 and 46, XX but they didn't separate as much in 46,XY;CD (Additional file 1: Fig. S5-S1). The correlation plot for GFPC, 6TF SLC and 8TF SLC derived from each of the four cell lines showed that the replicate samples correlated with each other and each of the three experimental groups within each line were distinct from one another (Additional file 1: Fig. S5-2B,C,D). Together these data suggest that the experimental procedure had a high level of reproducibility.

We investigated the fold change in expression of known DSD genes and Sertoli markers.

**46,XX cell line:** Notably, while 46, XX SLCs expressed the expected Mogrify TFs in 6TF and 8TF SLCs (Additional file 1: Fig. S5-3D) they did not show increased *SOX9* expression over GFPC, either in the RNA seq data (Fig. 5A) or in qPCR data (Additional file 1: Fig. S5-3E). Even though they didn't express *SOX9*, they did express bipotential gonad markers *GADD45G*, *ZFPM2*, *NR0B1* and immature SC markers *KRT18*, *INHBA*, *NCAM1/2*, *HSD17B3* and the all-stage marker *GDNF*. They barely expressed any mature markers, unlike 46,XY SLCs. Even though their SC marker expression profile was subdued in comparison to 46,XY SLCs, the presence of *NR5A1* and *EBF1* in the transduction mix had a similar effect on 8TF XX SLC in inducing Leydig markers *HSD3B2* and *CYP11A1* (Additional file 1: Fig. S5-3D). Lack of induction of *SOX9* in 46,XX-derived SLCs interestingly



**Fig. 3** 46, XY DSD and 46,XX fibroblast derived SLCs exhibit varying degrees of change in morphology after 1 month trans-differentiation. Representative images of GFPC and 6TF SLCs on Day7 and Day28; Shape factor quantifications (N = 3, n = 50–60 for each group). \*Represent p values calculated from Mann Whitney statistical tests conducted between the indicated groups for **(A)** 46, XY; SRXY1 **(B)** 46,XY; CD; **(C)** 46,XY; WT1 and **(D)** 46,XX **(E)** Shape factor quantifications comparing GFPC and 6TF SLC among all the four genetic backgrounds



**Fig. 4** xCELLigence assay shows that 6TF SLCs exhibit distinct adhesion and proliferation phenotype than GFP controls across different genetic backgrounds. Slope of 0–6 h cell index (CI) adhesion curves (**A**) and slope of 6–48 h cell index (CI) proliferation curves (**B**) for 1 month old GFPC (green) and 6TF (blue) SLC derived from fibroblasts of indicated genetic backgrounds. All the experimental readings represent an average of three biological replicates (N = 3), \* represent p values calculated from one way ANOVA test conducted amongst the three groups within each graph

coincides with lack of significant morphometric change (Fig. 3D).

**46,XY, SRXY1 DSD cell line:** 6TF and 8TF SLCs derived from 46, XY, SRXY1 DSD cells showed induction of *SOX9* by both RNAseq (Fig. 5A) and qPCR (Additional file 1: Fig. S5-3E) as observed in the wild-type 46,XY-derived SLCs. These SLCs exhibited similar profiles of expression of bipotential gonad markers *LHX9*, *GADD45G*, *ZFPM2*; immature SC markers *KRT18*, *INHBA/B*, *NCAM1* and all-stage marker *GDNF*. This showed that the trans-differentiation method was capable of pushing cells from DSD background genotype towards a Sertoli fate.

**46,XY, campomelic dysplasia cell line:** 46, XY;CD-derived 6TF SLCs showed increased *SOX9* expression (Fig. 5A; 8TF shown in Additional file 1: Fig. S5-3B,E); over the respective GFP control which was notably much smaller than that observed in control 46, XY-derived SLCs. This is evident in both the qPCR data (Additional file 1: Fig. S5-3E) and in the RNAseq data (Additional file 1: Fig. S2E vs Additional file 1: Fig. S5-3B). 46, XY;CD-derived 6TF SLCs also showed a modest, not statistically significant rise in the expression of *PTGDS*, *NCAM1/2*, *INHBA* and *BMP4*.

**46,XY;WT1 DSD cell line:** *SOX9* was also expressed in 46, XY;WT1-derived 6TF and 8TF SLCs at levels comparable to control 46, XY-derived SLCs (Fig. 5A, Additional file 1: Fig. S5-3C,E). These SLCs expectedly exhibited an increase in expression of other immature SC markers like *CYP26B1*, *PTGDS*, *KRT18*, *INHBA* and *NCAM2*. They additionally expressed mature markers *CLDN11*, *CTSL* and another all-stage marker *BMP4*.

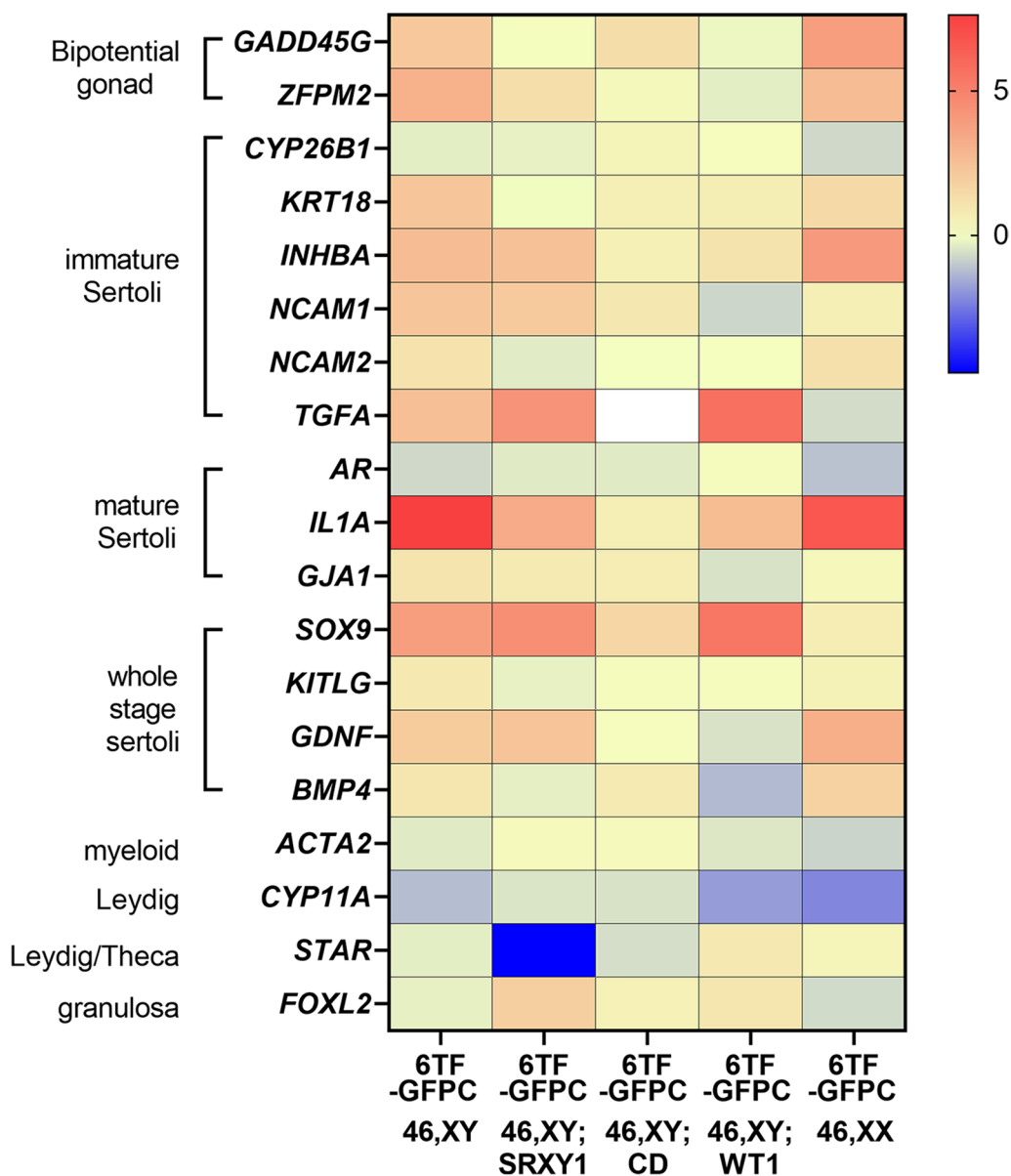
8TF SLCs derived from all the three DSD fibroblasts also exhibited higher levels of expression of Leydig

markers *HSD3B2* and *CYP11A1* in comparison to 6TF SLCs, just like observed for control 46,XY SLCs.

The transcriptomic profiling of all these cell lines shows that the trans-differentiation method is robust enough to bring about fibroblast to Sertoli cell fate differentiation in cells belonging to different genetic backgrounds with some properties sexually dimorphic and DSD SLCs showing intermediate, distinct phenotypes.

## Discussion

Cell differentiation can be achieved by utilizing a transgene-free approach of supplementing the culture media with suitable growth and differentiation supplements. Sertoli and gonadal cell types have been produced using such approaches from different cell types like human embryonic stem cells (hESC), human induced pluripotent stem cells (iPSC) and human umbilical cord-derived perivascular cells (HUCPVC) [33, 34, 49, 51]. A major drawback of such an approach is that they are transient, and target differentiation fate may not pass down to the daughter cells optimally, thus limiting the scope of utilization of resultant cells. Creating a cell model to aid DSD diagnosis in a patient-specific manner warrants a transgene-based approach to produce genetically stable Sertoli-like cells. One of the most critical considerations in such operations is getting the right combination of transcription factors (TFs) and culture conditions that could drive the targeted cell fate change. Murine and human fibroblasts have been trans-differentiated into Sertoli-like cells using such a transgene approach. Murine fibroblasts were reprogrammed to a Sertoli fate using transgenic expression of *Nr5a1*, *Wt1*, *Dmrt1*, *Gata4*, and *Sox9* [35]. Transgenic expression of *NR5A1*



**Fig. 5** 46, XY DSD and 46, XX fibroblasts-derived 6TF SLCs show varying levels of Sertoli specific marker gene expression. Heat map showing 6TF vs. GFPC log2 fold change for indicated gonadal marker genes as measured by RNAseq for each of the indicated genetic backgrounds: 46,XY (N=3), 46,XY; SRXY1 (N=2); 46,XY; CD (N=3), 46,XY; WT1 (N=3), 46,XX (N=3)

and *GATA4* in human fibroblasts achieved expression of Sertoli markers, but specificity wasn't clearly reported [36], and collateral induction of other gonadal cell types could not be ruled out.

Computational algorithms using published literature and datasets to predict TFs required for directed cellular reprogramming have been successfully used for other cell types [37, 53, 54]. We utilized one such predictive tool, Mogrify, which suggested 8 or 6 TFs as being sufficient to transdifferentiate dermal fibroblasts to

Sertoli cells (Fig. 1): *FGF2*, *GATA6*, *GATA4*, *MXII*, *JUNB*, *NFYB*+*NR5A1* and *EBF1*. This method, when applied on 46,XY fibroblasts, produced Sertoli-like cells that exhibited significant change in morphology (Fig. 1), as well as expression of Sertoli molecular signatures as measured by RNAseq, qPCR as well as IF (Fig. 2, Additional file 1: Fig. S2), and cell behavior (Fig. 4). They also exhibited upregulation of several genes associated with known Sertoli phenotype including cell adhesion, proliferation, and lipid metabolism as revealed by GO analysis of

co-DEG between them and an adult Sertoli transcriptome (Fig. 2—Additional file 1: Fig. S2). Both 6TF and 8TF SLCs exhibited predominantly immature Sertoli molecular signatures rather than mature SC signatures (Fig. 2). We found that the 6TF trans-differentiation condition was likely best suited to create SLCs, as the additional *NR5A1* and *EBF1* in 8TF conditions induced expression of unwanted Leydig cell markers (Fig. 2—Additional file 1: Fig. S2). The method, when applied to fibroblasts of different 46, XY DSD and 46, XX genotypes, produced varying levels of success in achieving the Sertoli phenotype. While 46, XY; SRXY1 and WT1 SLCs achieved comparable levels of change in morphometry and Sertoli marker expression, 46, XY; CD and 46,XX SLCs exhibited significantly less change than was observed in 46, XY SLCs (Figs. 3, 5, Additional file 1: Fig. 5–S3). When comparing adhesion and proliferation of SLCs derived from different backgrounds also, the 46, XX SLCs showed quite a difference from 46, XY SLCs whereas 46,XY CD SLCs were somewhat in between (Fig. 4). The varying levels of success in achieving that phenotype in different DSD cells might be reflective of their different genetic backgrounds and of which level of the genetic networks are affected by the mutations carried by those DSD cells. This trans-differentiation method can serve as a robust resource for precision medicine as it can be employed to transduce DSD patient-derived fibroblasts into Sertoli cells, which in turn can be used for relevant gonadal developmental stage-specific transcriptomic and phenotypic analysis, thus aiding DSD molecular diagnostics. The differentiated cells could be used for high-throughput screening to observe the effects on cell phenotypes and molecular pathways of genetic variants of unknown significance, either in patient cells or by creating variants in genetically modified cells (e.g. using CRISPR). Finally, it could be used to test how correction of suspected causative variants in patient-derived SLCs rescues the transcriptomic or cell phenotype differences.

### Perspectives and significance

The ability to reproduce in vitro a specialized cellular state (such as Sertoli cells) from a skin fibroblast opens the possibility to identify pathophysiological mechanisms of diseases without invasive procedures to access a specific tissue. The molecular pathways leading to formation of the testes or ovaries happen in utero and models to study these early developmental stages are limited to immortalized cells typically derived from reproductive cancers and animal models. It is being increasingly recognized that early human and mouse gonadal development are not highly conserved. Thus, this human-specific model would allow understanding of how variants affect molecular pathways in the developing gonad, and may

allow discovery of new pathways. This would both validate variants of unknown significance and possibly reveal new candidate genes increasing the diagnostic yield in DSD. It would also provide a model to test strategies for potential targeted therapeutic interventions. The strategy used here to generate SLCs could be applied to other testicular cell types (e.g. Leydig cells or germ cells) and ovarian cells to extend the applicability of the model to 46,XX and other types of 46,XY DSD. These cells can also be utilized in three dimensional cultures using appropriate ECM components to look for formation of functional seminiferous tubules. They can further be co-cultured with similarly in situ produced gonadal cell types such as Leydig, germ, myeloid etc. in attempts to produce functional gonadal organoids or to preserve future fertility. Such a development could expand the scope of these cells from diagnostics to functional gonad replacement therapeutics.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13293-024-00599-y>.

**Additional file 1: Fig. S1.** Standardization of trans-differentiation protocol. (A) Outline of the initial 16 culture conditions designed to attempt fibroblast-Sertoli cell like trans-differentiations. (B) The outcome of these conditions was evaluated by looking for Sertoli marker expression in each of the 16 conditions by qPCR for: *SOX9*, *BMP4*, *PTGDS* and *AR*.  $N = 2$ ,  $n = 6$  for each of the qPCRs, error bars represent SEM. (C) Representative images of 1-month live cultures for all these conditions. Missing images either had too low green fluorescence or extensive cell death. **Figure S2.** 46,XY SLC characterization. Heatmaps showing expression of co-DEGs for 8TF, GFPC, aSC (A) or 6TF, GFPC, aSC (B). Biological process gene ontology (GO) analyses for the co-DEGs of 8TF (C, red) and 6TF (D, blue). (E) RNAseq analysis showing linear scale fold change in expression of indicated markers in 46,XY derived 8TF (red) and 6TF (blue) SLCs over GFPC  $N = 3$ , \* represent adjusted p value of fold change calculations, ns not shown. (F) Linear scale fold change for indicated markers as determined by qPCR for 8TF (red) and 6TF (blue). The values represent mean  $\pm$  SEM of the following number of samples ( $N =$  biological replicate,  $n =$  technical replicate): For 8 TF, *PTGDS*:  $N = 9$ ,  $n = 18$ ; *AR*:  $N = 9$ ,  $n = 12$ ; *FSHR*:  $N = 4$ ,  $n = 16$ ; *SOX9*:  $N = 12$ ,  $n = 25$ ; *BMP4*:  $N = 9$ ,  $n = 18$ ; *ACTA2*:  $N = 3$ ,  $n = 3$ ; *DDX4*:  $N = 3$ ,  $n = 3$ ; *FOXL2*:  $N = 3$ ,  $n = 6$ . For 6 TF: *PTGDS*:  $N = 6$ ,  $n = 12$ ; *AR*:  $N = 6$ ,  $n = 6$ ; *FSHR*:  $N = 9$ ,  $n = 15$ ; *SOX9*:  $N = 6$ ,  $n = 12$ ; *BMP4*:  $N = 6$ ,  $n = 12$ ; *ACTA2*:  $N = 3$ ,  $n = 3$ ; *DDX4*:  $N = 3$ ,  $n = 3$ ; *FOXL2*:  $N = 3$ ,  $n = 3$ . \* represent p values calculated from unpaired t test conducted between each 8TF or 6TF and GFPC. **Figure S3.** Morphometry of 46,XY DSD and 46,XX SLC. Representative images of GFPC and 8TF SLCs on Day7 and Day28; Shape factor and area quantifications for GFPC, 8TF and 6TF SLCs ( $N = 3$ ,  $n = 50$ – $60$  for each group). \* represent p values calculated from Mann-Whitney statistical tests conducted between the indicated groups for (A). 46, XY; SRXY1 (B) 46,XY; CD; (C) 46,XY; WT1 and (D) 46,XX; non significant p value comparisons are omitted. (E) GFPC area factor, 8TF shape factor, 8TF area factor, 6TF area factor quantifications comparing 46, XY DSD and 46,XX genetic backgrounds with 46, XY. **Figure S4.** Cell index (CI) curves of xCELLigence assays. (A) Slope of 0–6 h cell index (CI) adhesion and 6–48 h proliferation curves for 1-month-old GFPC, 8TF and 6TF SLC derived from fibroblasts of indicated genetic backgrounds. All the experimental readings represent an average of three biological replicates ( $N = 3$ ), \* represent p values calculated from one way ANOVA test conducted amongst the three groups within each graph, ns is not shown. 0–6 h CI adhesion curves and 6–48 h CI proliferation curves for 1-month-old GFPC, 8TF SLC and 6TF SLC derived from 46,XY (B), 46,XY;CD (C) and 46, XX (D). (E): Alternative representation of

0–6 h cell index (CI) adhesion curves and (F) 6–48 h CI proliferation curves for 46,XY, 46,XY with campomelic dysplasia (CD), or 46,XX-derived GFPC, 8TF SLCs and 6TF SLCs. All the data points represent an average plus SEM of three biological replicates (N = 3). **Figure S5.** S1: Principal Component Analysis of RNAseq transcriptomes comparing all the replicate samples for aSC (in brown), GFPC (green), 8TF (red), 6TF (blue) derived from (A) GM00083: 46,XY; SRXY1, (B) GM21894: 46,XY; CD, (C) GM06938: 46,XY; WT1 and (D) GM03651: 46,XX. S2: RNAseq sample replicates cluster together. Correlation plots for each of the 8TF (red bars along Y axes), 6TF (blue bars) and GFPC (green bars) RNA seq samples of (A) 46,XY, (B) 46,XY; SRXY1, (C) 46,XY; CD, (D) 46,XY; WT1 and (E) 46,XX. S3: A wide range of gonadal markers and transduced Mogrify TFs show expression in SLC samples. RNAseq analysis shows linear scale fold change in expression of indicated markers in 8TF- (red) or 6TF- (blue) derived SLCs in comparison to GFPC for (A) 46,XY; SRXY1 (B) 46,XY; CD (C) 46,XY; WT1 and (D) 46,XX. N = 3, \* represent adjusted p value of fold change calculations between experimental and respective control, ns not shown. (E): Fold change for *SOX9* as determined by qPCR, in both 8TF (red) or 6TF (blue) SLC over respective GFPC for the indicated cell lines. N = 5, n = 25 for 46,XY, N = 3, n = 3 each for 46,XY; SRXY1, 46,XY; CD, 46,XY; WT1 and 46,XX, error bars represent SEM. \* represent p values calculated from unpaired t test conducted between each experimental and its corresponding GFPC, ns not shown. S4: Identification of the *SOX9* variant in the 46,XY CD cell line. DNA extracted from the Coriell cell lines GM03368 (46,XY; SRXY1 of unknown cause) and GM21894 (46,XY; CD) was subjected to exome sequencing, using the Agilent SureSelect v.6 capture kit. Sequencing was performed at Novogene. Average depth of coverage was 174x and 189x, respectively. Exome did not identify a causative variant in either sample. GM21894 (46,XY; CD) was then subjected to PCR-free whole-genome sequencing (performed at the Broad Institute, Cambridge, MA), with a resulting average read depth of 45x. Sequence analysis was performed on the Franklin (Genoox, Inc.) platform, using both the single nucleotide variants and structural variants pipelines (A). This identified a 3.2 Mb heterozygous deletion on chromosome 17q24.3–25.1 (B), including the entire *SOX9* gene (and 44 other genes), at chromosomal coordinates chr17:68,100,001–71,280,001 (on the hg19 human genome reference) (C). **Table S1.** List of marker genes.

### Acknowledgements

The authors thank Mr. Miguel Almalvez for help with submission of the manuscript and laboratory management.

### Author contributions

AP conceptualized and designed the trans-differentiation protocol, conducted most of the wet lab and molecular biology benchwork, analyzed data, and drafted the manuscript. ED provided mentorship on experimental design and results interpretation, helped analyze the data, troubleshoot experiments, contributed to and edited the manuscript, and supervised revision submissions. AR performed the immunofluorescence imaging. JR conducted the xCELLigence assay, wrote its method section, and contributed to all manuscript revisions. SB conducted the qPCR and RNA-seq bio-informatic data analysis, composed the RNA-seq analysis method section, and deposited raw data in GEO. VH supervised JR and AR, funded the immunofluorescence and xCELLigence experiments, provided critical revision and inputs for important content. EV conceived the study, provided funding, mentorship, and edits throughout the manuscript preparation process.

### Funding

Funding for the US-based team included support by funding at Children's National Hospital and the Disorders/Differences of Sex Development (DSD)—Translational Research Network NIH R01HD093450 grant. Funding for the Australian team included: National Health and Medical Research Council Ideas Grant 2002426 and NHMRC Fellowship APP1154870 (to V.H.) and support from the Victorian Government's Operational Infrastructure Support Program.

### Availability of data and materials

The raw fastq data and the normalized read counts are available on Gene Expression Omnibus (GEO accession number GSE246236).

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they do not have any competing interests.

### Author details

<sup>1</sup>Center for Genetic Medicine Research, Children's National Hospital, Washington D.C. 20010, USA. <sup>2</sup>Centre for Endocrinology and Metabolism, Hudson Institute of Medical Research, Melbourne, VIC 3168, Australia. <sup>3</sup>Institute for Clinical and Translational Science, University of California, Irvine, CA, USA.

Received: 17 February 2023 Accepted: 22 February 2024

Published online: 22 March 2024

## References

- Hughes IA, Houk C, Ahmed SF, Lee PA. Consensus statement on management of intersex disorders. *Arch Dis Child.* 2006;91:554–63.
- Lee PA, Houk CP, Ahmed SF, Hughes IA, International Consensus Conference on Intersex organized by the Lawson Wilkins Pediatric Endocrine S, the European Society for Paediatric E Consensus statement on management of intersex disorders. International Consensus Conference on Intersex. *Pediatrics.* 2006;118(2):e488–500.
- Délot EC, Papp JC, Délot EC, Fox M, Grody W, Lee H, et al. Genetics of disorders of sex development: the DSD-TRN experience. *Endocrinol Metab Clin N Am.* 2017;46(2):519–37.
- Hutson JM, Grover SR, O'Connell M, Pennell SD. Malformation syndromes associated with disorders of sex development. *Nat Rev Endocrinol.* 2014;10(8):476–87.
- Baetens D, Stoop H, Peelman F, Todeschini AL, Rosseel T, Coppieters F, et al. NRS1A1 is a novel disease gene for 46, XX testicular and ovotesticular disorders of sex development. *Genet Med.* 2016;4(19):367.
- Délot EC, Vilain E. Towards improved genetic diagnosis of human differences of sex development. *Nat Rev Genet.* 2021;22(9):588–602.
- Parivesh A, Barseghyan H, Délot E, Vilain E. Translating genomics to the clinical diagnosis of disorders/differences of sex development. *Curr Top Dev Biol.* 2019;134:317–75. <https://doi.org/10.1016/bs.ctdb.2019.01.005>.
- León NY, Reyes AP, Harley VR. A clinical algorithm to diagnose differences of sex development. *Lancet Diabetes Endocrinol.* 2019;7(7):560–74.
- Barseghyan H, Délot EC, Vilain E. New technologies to uncover the molecular basis of disorders of sex development. *Mol Cell Endocrinol.* 2018;468(January):60–9.
- Croft B, Ayers K, Sinclair A, Ohnesorg T. Review disorders of sex development: the evolving role of genomics in diagnosis and gene discovery. *Birth Defects Res C Embryo Today.* 2016;108(4):337–50.
- Délot E, Vilain E. Chapter 17: Differences of sex development. In: J. F. Strauss, R. L. Barbieri, A. Dokras, C. J. Williams & S. Zev Williams editors. Ninth edition of Yen & Jaffe's Reproductive Endocrinology. Physiology, Pathophysiology, and Clinical management. Elsevier.
- Hamanaka K, Miyatake S, Koshimizu E, Tsurusaki Y, Mitsuhashi S, Iwama K, et al. RNA sequencing solved the most common but unrecognized NEB pathogenic variant in Japanese nemaline myopathy. *Genet Med.* 2019;21(7):1629–38.
- Kernohan KD, Frésard L, Zappala Z, Hartley T, Smith KS, Wagner J, et al. Whole-transcriptome sequencing in blood provides a diagnosis of spinal muscular atrophy with progressive myoclonic epilepsy. *Hum Mutat.* 2017;38(6):611–4.
- Kremer LS, Bader DM, Mertes C, Kopajtic R, Pichler G, Iuso A, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun.* 2017;8(1):1–11.
- Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with



- transcriptome sequencing. *Sci Transl Med*. 2017;9(386):eaal5209. <https://doi.org/10.1126/scitranslmed.aal5209>.
16. Lecluze E, Rolland AD, Filis P, Evrard B, Leverrier-Penna S, Ben MM, et al. Dynamics of the transcriptional landscape during human fetal testis and ovary development. *Hum Reprod*. 2020;35(5):1099–119.
  17. Ross AJ, Capel B. Signaling at the crossroads of gonad development. *Trends Endocrinol Metab*. 2005;16(1):19–25.
  18. Sekido R, Bar J, Narváez V, Penny G, Lovell-Badge R. SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli cell precursors. *Dev Biol*. 2004;274(2):271–9.
  19. Svingen T, Koopman P. Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. *Genes Dev*. 2013;27(22):2409–26.
  20. Mackay S. Gonadal development in mammals at the cellular and molecular levels. *Int Rev Cytol*. 2000;200:47–99.
  21. McLaren A. Germ and somatic cell lineages in the developing gonad. *Mol Cell Endocrinol*. 2000;163(1–2):3–9.
  22. Cool J, Capel B. Mixed signals: development of the testis. *Semin Reprod Med*. 2009;27:5–13.
  23. Tevosian SG, Albrecht KH, Crispino JD, Fujiwara Y, Eicher EM, Orkin SH. Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development*. 2002;129(19):4627–34.
  24. Wilhelm D, Englert C. The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sfl. *Genes Dev*. 2002;16(14):1839–51. <https://doi.org/10.1101/gad.220102>.
  25. Hanley NA, Ball SG, Clement-Jones M, Hagan DM, Strachan T, Lindsay S, et al. Expression of steroidogenic factor 1 and Wilms' tumour 1 during early human gonadal development and sex determination. 1999. [www.elsevier.com/locate/modo](http://www.elsevier.com/locate/modo). Accessed 8 Dec 2020.
  26. Tarulli GA, Stanton PG, Meachem SJ. Is the adult sertoli cell terminally differentiated? *Biol Reprod*. 2012;87(1):1–11.
  27. Sharpe RM, McKinnell C, Kivlin C, Fisher JS. Proliferation and functional maturation of Sertoli cells, and their relevance to disorders of testis function in adulthood. *Reproduction*. 2003;125(6):769–84.
  28. Sekido R, Lovell-Badge R. Genetic control of testis development. *Sex Dev*. 2013;7(1–3):21–32.
  29. Cameron FJ, Sinclair AH. Mutations in SRY and SOX9: testis-determining genes. *Hum Mutat*. 1997;9:388–95.
  30. Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, et al. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell*. 1994;79(6):1111–20.
  31. Croft B, Ohnesorg T, Hewitt J, Bowles J, Quinn A, Tan J, et al. Human sex reversal is caused by duplication or deletion of core enhancers upstream of SOX9. *Nat Commun*. 2018;9(1):1–10. <https://doi.org/10.1038/s41467-018-07784-9>.
  32. Knower KC, Kelly S, Harley VR. Turning on the male—SRY, SOX9 and sex determination in mammals. *Cytogenet Genome Res*. 2003;101(3–4):185–98. <https://doi.org/10.1159/000074336>.
  33. Rodríguez Gutiérrez D, Eid W, Biason-Lauber A. A human gonadal cell model from induced pluripotent stem cells. *Front Genet*. 2018;9(October):1–14. <https://doi.org/10.3389/fgene.2018.00498/full>.
  34. Knarston IM, Pachernegg S, Robevska G, Ghobrial I, Er PX, Georges E, et al. An in vitro differentiation protocol for human embryonic bipotential gonad and testis cell development. *Stem Cell Rep*. 2020;15(6):1377–91. <https://doi.org/10.1016/j.stemcr.2020.10.009>.
  35. Buganim Y, Itskovich E, Chiang HY, Cheng AW, Ganz K, Sarkar S, et al. Direct reprogramming of fibroblasts into embryonic sertoli-like cells by defined factors. *Cell Stem Cell*. 2012;11:373–86.
  36. Liang J, Wang N, He J, Du J, Guo Y, Li L, et al. Induction of sertoli-like cells from human fibroblasts by nr5a1 and gata4. *Elife*. 2019;8:1–27.
  37. Rackham OJL, Firas J, Fang H, Oates ME, Holmes ML, Knaupp AS, et al. A predictive computational framework for direct reprogramming between human cell types. *Nat Genet*. 2016;48(3):331–5. <https://doi.org/10.1038/ng.3487>.
  38. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047–8.
  39. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15.
  40. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform*. 2011;12(1):1–16.
  41. Soneson C, Love MI, Robinson MD, Floor SN. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*. 2016;4:1521.
  42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):1–21.
  43. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019;47(W1):W191–8.
  44. Jameson SA, Natarajan A, Cool J, DeFalco T, Maatouk DM, Mork L, et al. Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet*. 2012;8(3): e1002575. <https://doi.org/10.1371/journal.pgen.1002575>.
  45. Zhao L, Wang C, Lehman ML, He M, An J, Svingen T, et al. Transcriptomic analysis of mRNA expression and alternative splicing during mouse sex determination. *Mol Cell Endocrinol*. 2018;15(478):84–96.
  46. Buonocore F, Achermann JC. Human sex development: targeted technologies to improve diagnosis. *Genome Biol*. 2016;17(1):257.
  47. Lakpour M, Aghajanzpour S, Koruji M, Shahverdi A, Sadighi-Gilani M, Sabaghian M, et al. Isolation, culture and characterization of human sertoli cells by flow cytometry: development of procedure. *J Reprod Infertil*. 2017;18(2):213–7.
  48. Wen L, Yuan Q, Sun M, Niu M, Wang H, Fu H, et al. Generation and characteristics of human Sertoli cell line immortalized by overexpression of human telomerase. *Oncotarget*. 2017;8(10):16553–70.
  49. Shlush E, Maghen L, Swanson S, Kenigsberg S, Moskovtsev S, Barretto T, et al. In vitro generation of Sertoli-like and haploid spermatid-like cells from human umbilical cord perivascular cells. *Stem Cell Res Ther*. 2017;8(1):1–16.
  50. Chui K, Trivedi A, Cheng CY, Cherbavaz DB, Dazin PF, Huynh ALT, et al. Characterization and functionality of proliferative human sertoli cells. *Cell Transplant*. 2011;20(5):619–35.
  51. Bucay N, Yebra M, Cirulli V, Afrikanova I, Kaido T, Hayek A, et al. A novel approach for the derivation of putative primordial germ cells and sertoli cells from human embryonic stem cells. *Stem Cells*. 2008;27:68–77.
  52. Ono M, Harley VR. Disorders of sex development: new genes, new concepts. *Nat Rev Endocrinol*. 2013;9(2):79–91. <https://doi.org/10.1038/nrendo.2012.235>.
  53. Alessio ACD, Fan ZP, Wert KJ, Baranov P, Cohen MA, Saini JS, et al. A systematic approach to identify candidate transcription factors that control cell identity. *Stem Cell Rep*. 2015;5(5):763–75. <https://doi.org/10.1016/j.stemcr.2015.09.016>.
  54. Morris SA, Cahan P, Li H, Zhao AM, San Roman AK, Shivdasani RA, et al. Dissecting engineered cell types and enhancing cell fate conversion via Cellnet. *Cell*. 2014;158(4):889–902.

## Publisher's Note

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