Online Supplemental

Methods

CNV screening

A qPCR approach was used to identifying further patients with CNVs at Xp21.2. 168 Patients with 46,XY DSD of unknown molecular genetic origin were studied. qPCRs were performed on a LightCycler 96 (Roche, Basel, Switzerland) with a CYBR-Green system using the Takyon No Rox SYBR MasterMix dTTP Blue (Eurogentec, Seraing, Belgium). Six primer pairs were designed to cover both the up- and downstream regions of *NR0B1* and located in the exons of *IL1RAPL1, MAGEB1, NR0B1, TASL* (also known as *CXorf21*), *GK* and *TAB3*. Primer validation consisted of standard curves for efficiency calculation (2.00 ± 0.10), melting curve analysis and gel electrophoresis of PCR products. Fertile males and females were included in each run as positive and negative controls for the x-chromosomal locus. *ZNF80* and *GPR15* were used as reference genes with primers published in RT-PrimerDB (ID 1021 and 1022, respectively) (Pattyn et al., 2003).

Breakpoint Sequencing

Primers at either end of the constructed breakpoint sequences were designed using Lasergene PrimerSelect (DNASTAR, Wisconsin, USA). Optimal annealing temperature was determined through gradient PCR and electrophoresis. Consequently, PCR setup was 35 cycles with 30 sec. at 95°C for denaturation, 30 sec. at primer specific temperature (Table S2, Supporting Information) for annealing and 1 min. extension at 72°C. Sanger Sequencing of amplicons was performed on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, USA).

Meinel JA, *et al. J Med Genet* 2023; 60:469–476. doi: 10.1136/jmg-2022-108635

Supporting Figures

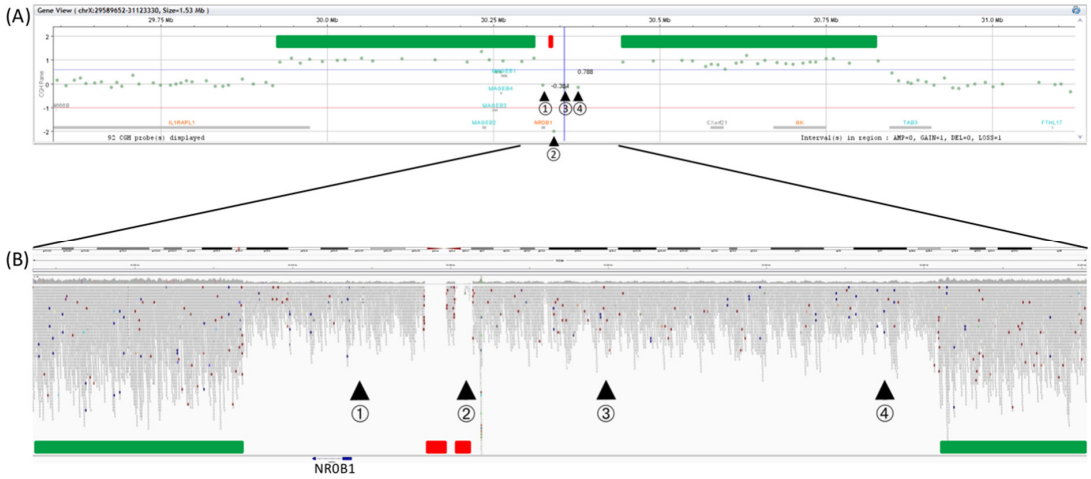

Figure S1. aCGH and GS comparison in P1 at *NR0B1* locus (A) aCGH results for P1. Green dots indicate probes used for hybridization. A pooled sample of 10 normal 46,XY males was used as control. Results show two separate duplications as illustrated by increased hybridization of probes and emphasized by green bars. To the left a 389kb duplication (chrX:29,924,420-30,313,761; GRCh37/hg19) was detected by increased hybridization of 17 probes. The second duplication was detected as 447kb of size (chrX: 30,401,819-30,848,988; GRCh37/hg19) indicated by 23 probes. The duplications are separated by four probes labelled 1 – 4. Probes 1, 3 and 4 hybridized only to a single copy. Probe 2 showed no hybridization. (B) Probes 1 to 4 have been compared to genome sequencing data. For visualisation of WGS data of P2 (region ChrX:30,292,562-30,423,695; GRCh37/hg19) the Integrative Genome Viewer (IGV) was used (Robinson et al., 2011). Increased number of reads (grey bars) at both sides of the viewing window signify duplication detected by WGS (green bars). Deletions are evident through the lack of reads in the two regions marked with red bars. Single hybridized CGH probes 1, 3 and 4 map to areas not duplicated according to WGS data. Probe 2 maps to a deletion and correspondingly did not hybridize. Note that the aCGH could be easily misinterpreted as one large duplication, if the different hybridisation of probes 1-4 is not recognized.

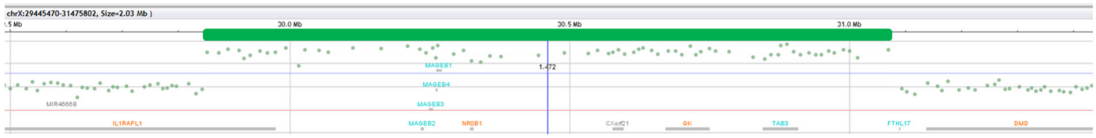

Figure S2. aCGH in P2 at *NR0B1* locus. Green dots indicate probes used for hybridization. A pooled sample of 10 normal 46,XY males was used as control. Results show a 1,2 Mb triplication (ChrX:29,851,537-31,069,736; GRCh37/hg19) emphasized by green bar. The genes *MAGEB1-4*, *NR0B1, TASL* (also known as *CXorf21*) and *TAB3* are fully triplicated whereas *IL1RAPL1* is only partially triplicated.

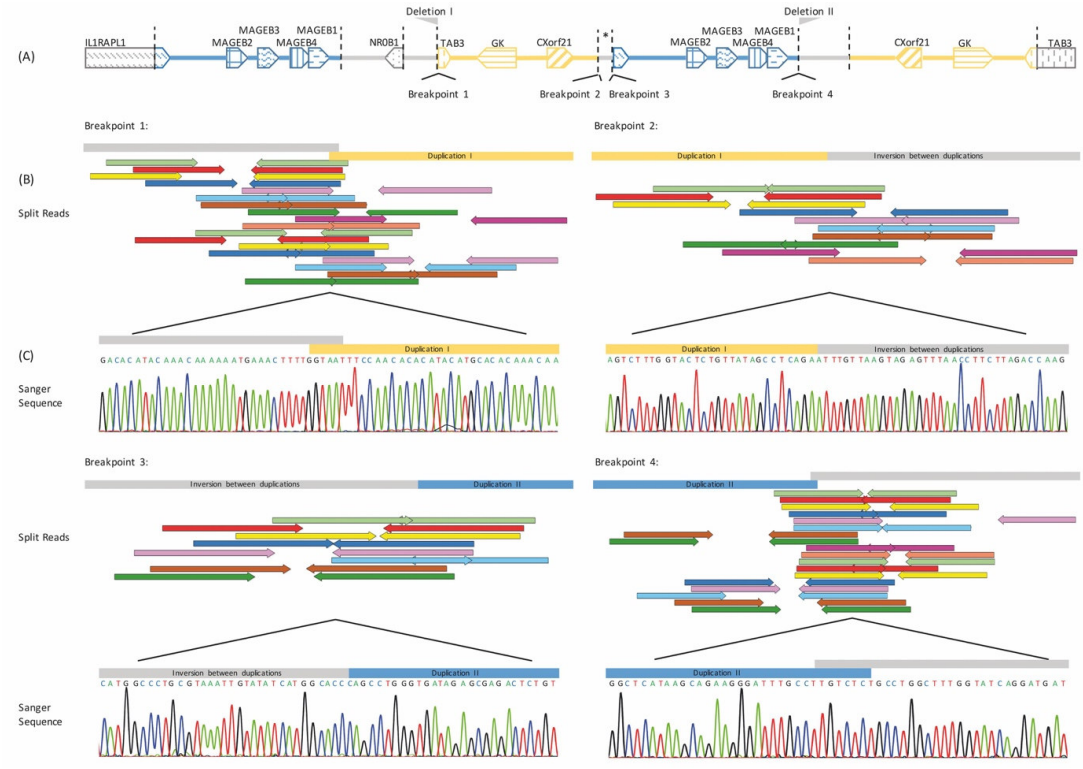

Figure S3. Structural Variation of Patient 1 (A) Overview of the structural variation upstream of the *NR0B1* gene. The genomic distances are not to scale. A 447kb duplication of *TASL* (*CXorf21)*, *GK* and part of *TAB3* (yellow) originally mapping to 30,401,819-30,848,988 is inserted in an inverted manner between breakpoint 1 (bp1; chrX:30,336,745; GRCh37/hg19) and bp2 (chrX: 30,340,605; GRCh37/hg19) 9.25kb upstream of the *NR0B1* reading frame. * marks a 1.2kb piece of reference sequence inverted between bp2 and bp3 separating the two large duplications. Both duplications are flanked by deletions indicated by the grey flags denoted *Deletion I & II*. A 389kb duplication of the *MAGEB* genes 1-4 and part of *IL2RAPL2* (blue) originally mapping to chrX: 29,924,420-30,313,761 (GRCh37/hg19) is inserted between breakpoint 3 (chrX: 30,339,452; GRCh37/hg19) and 4 (chrX: 30,342,785; GRCh37/hg19) in the same orientation as its reference sequence. (B) Shows position of extracted and aligned WGS split reads crossing the four breakpoints. Split reads mapped at varying distances apart for each breakpoint. Each pair was separated by at least >29kb. Sequences at either side of bp1 and bp4 showed homology, thus no exact definition of the breakpoint position was

possible, as depicted by the overlap of the grey and yellow or blue bars. (C) Shows verification of sequences deduced from WGS reads through Sanger sequencing.

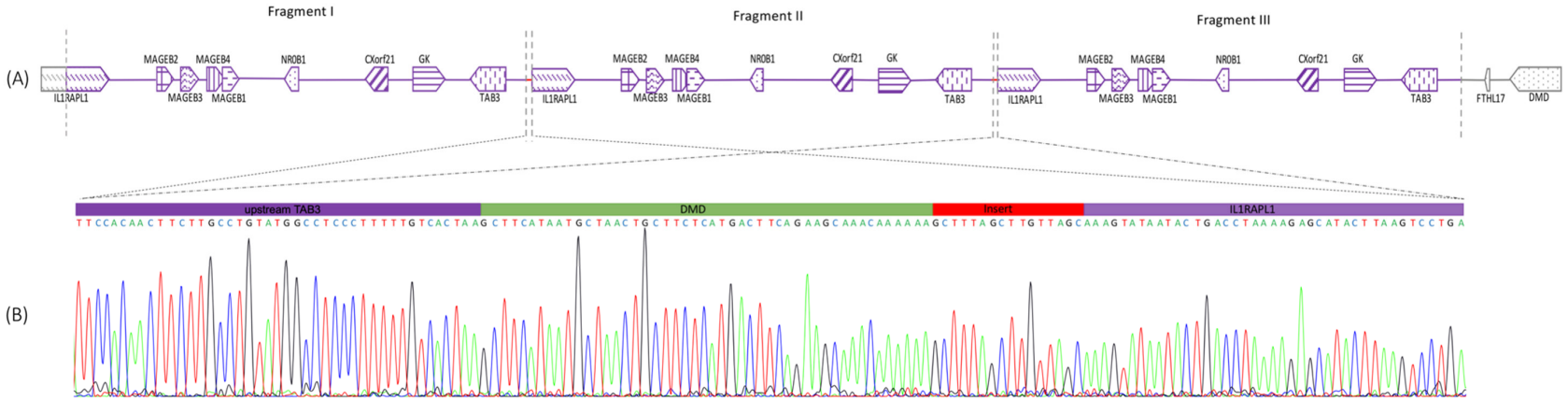

Figure S4. Structural Variation of Patient 2. (A) Overview of the structural variation at Xp21.2 locus. Genomic distances are not to scale. The chromosome segment is drawn with the distal chromosome arm to the left and centromere to the right. WGS showed the triplication originally mapped to chrX: 29,849,782 – 31,088,713 (GRCh37/hg19) and is arranged in a tandem manner. The triplicated fragments are connected by two identical 49 bp inserts at the breakpoints. This sequence originates from an intron of the *DMD* gene, located to the centromere of the SV event. *CXorf21* corresponds to *TASL* (B)Verification of sequences deduced from WGS reads through Sanger sequencing.

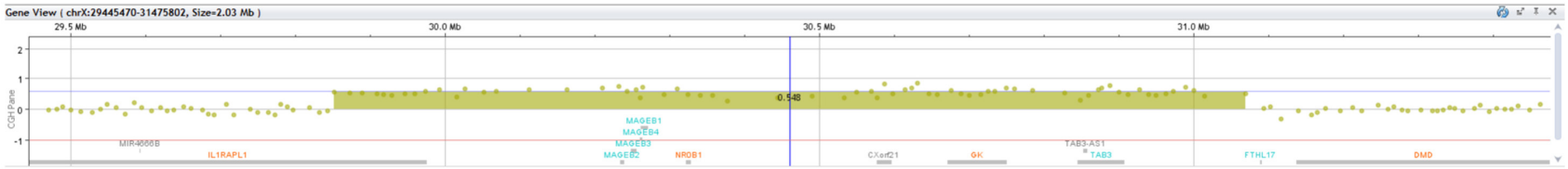

Figure S5. aCGH of mother of P2 at *NR0B1* locus. Green dots indicate probes used for hybridization. A pooled sample of 10 normal 46,XX females was used as control. Results show a 1,2 Mb duplication (ChrX:29,851,537-31,069,736; GRCh37/hg19) in a 46,XX karyotype. The size of the copy number gain corresponds to that of P2 i.e. the genes *MAGEB1-4*, *NR0B1, TASL (CXorf21)* and *TAB3* are fully duplicated whereas *IL1RAPL1* is only partially duplicated.

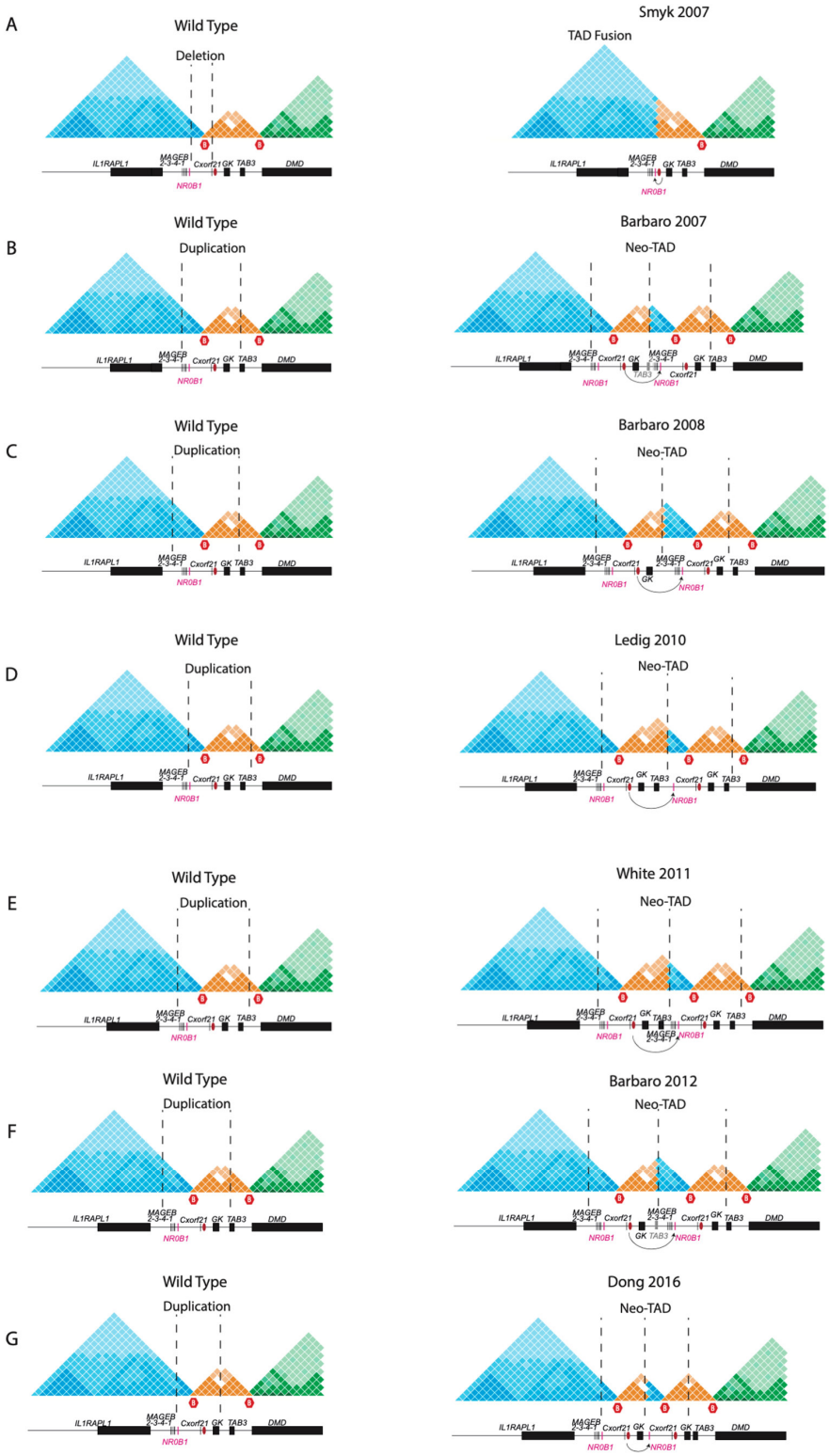
A
Wild Type
Deletion
Smyk 2007
TAD Fusion
B
Wild Type
Duplication
Barbaro 2007
Neo-TAD
C
Wild Type
Duplication
Barbaro 2008
Neo-TAD
D
Wild Type
Duplication
Ledig 2010
Neo-TAD
E
Wild Type
Duplication
White 2011
Neo-TAD
F
Wild Type
Duplication
Barbaro 2012
Neo-TAD
G
Wild Type
Duplication
Dong 2016
Neo-TAD
IL1RAPL1
MAGEB 2-3-4-1
Cxorf21
GK
TAB3
DMD
NR0B1

Figure S6. Proposed effects of inter-TAD deletions and duplications on Xp21.2 chromatin organisation.

Deletions (A) and duplications (B-F) at Xp21.2 are associated with gonadal dysgenesis and their chromatin landscape was modelled based on the TAD structure observed in our study. Since not for all duplications the exact breakpoints have been determined, tandem duplications have been assumed for all. The deletion (A) (Smyk et al., 2007) and all Inter-TAD duplications (B-F) (Barbaro et al., 2008; Barbaro et al., 2012; Barbaro et al., 2007; Dong et al., 2016; Ledig et al., 2010; White et al., 2011) containing *NR0B1* upstream boundary result in the incorporation of *NR0B1* into new chromatin domains (neo-TADs) with a novel regulatory landscape that can result in *NR0B1* upregulation. Red oval represents enhancers between *TASL* (*CXorf21)* and *GK*. Partially duplicated genes are in in light grey.

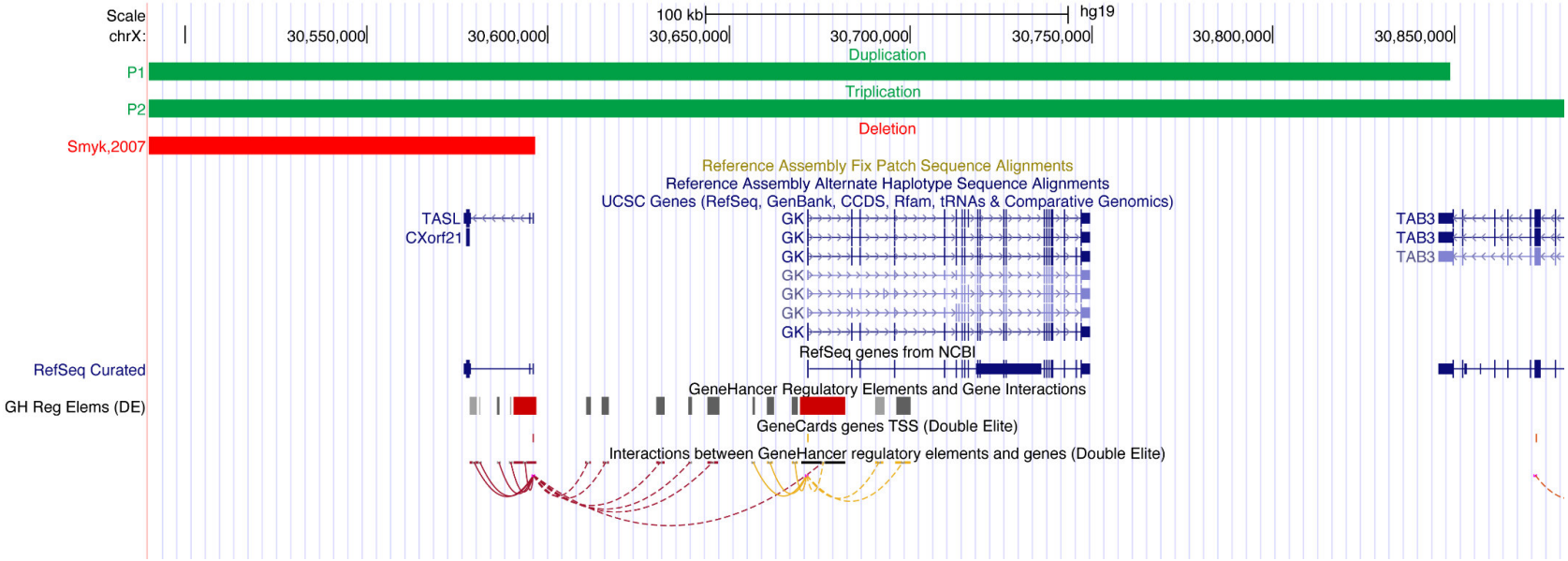

Figure S7. GeneHancer Promoters and Enhancers upstream of *NR0B1*. Data from the GeneHancer database (Fishilevich et al., 2017) was visualized in the UCSC genome browser (Kent et al., 2002). Promoters and enhancers are distinguished by red and grey colours respectively in the *GH Reg Elements (DE)* track. Colour intensity shows confidence score of the respective enhancers. According to Fishilevich et al. this is based on: "the number of supporting data sources, the number of unique TFBSs contained, and the overlap with ultra-conserved non-coding genomic elements".(Fishilevich et al., 2017) Arcs in this tack show proposed gene-enhancer interaction in the reference genome.

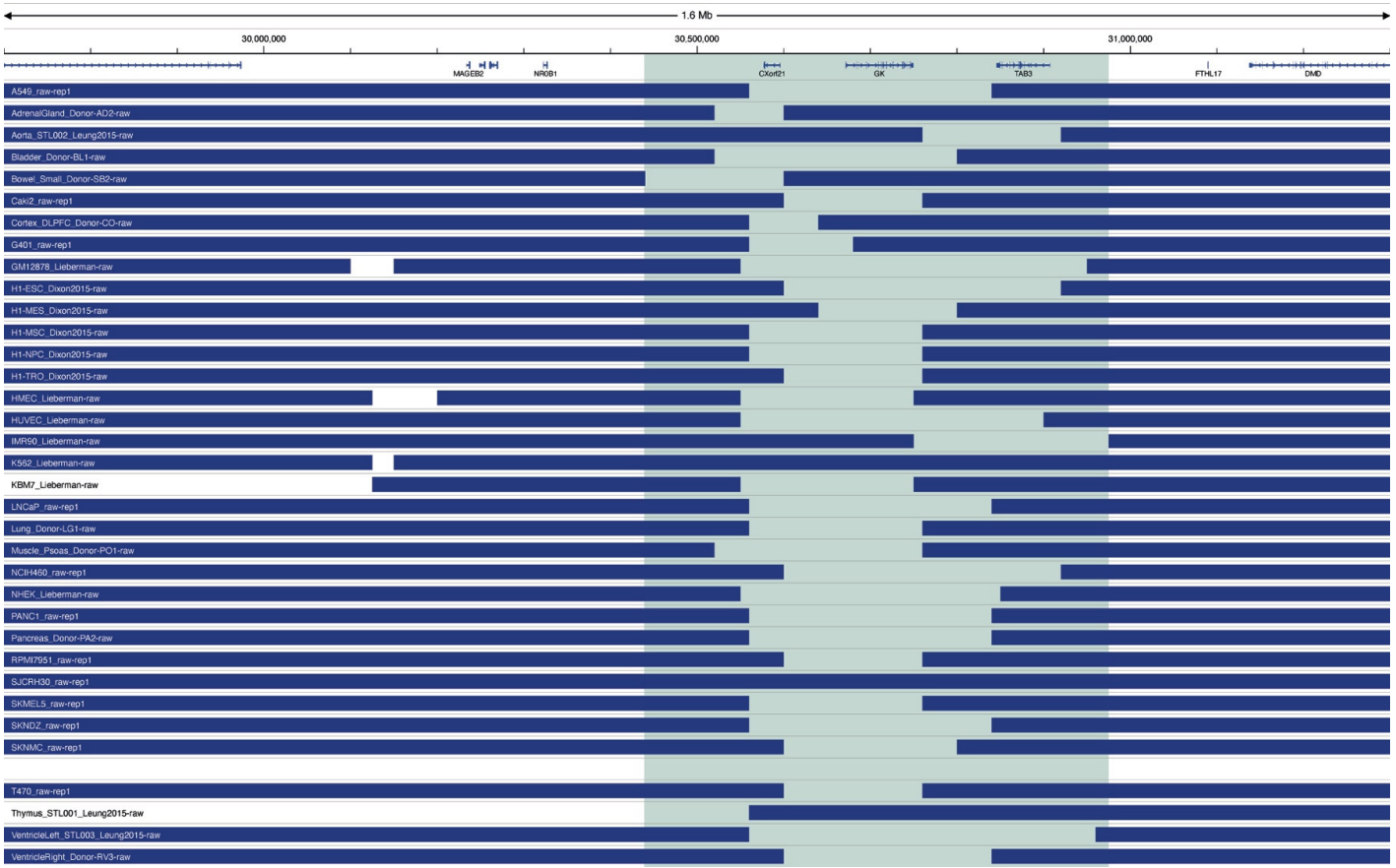

Figure S8: *NR0B1* TAD conformation of different human cell lines: Representation of TADs (rectangles) and their boundaries (empty spaces between rectangles) taken from 3D Genome Browser (Wang et al., 2018). The region encompassing all TAD boundaries consensus is shaded as light green.

Supporting Tables

Table S1. **Mutations in Genes Associated With 46,XY DSD Found in Patients 1 and 2.**

Patient	Gene	dbSNP ID	ExAc All (Lek et al., 2016)	ExAc European Non-Finish (Lek et al., 2016)	SIFT (Vaser et al., 2016)	Polyphen2 (Adzhubei et al., 2010)	MutationTaster (Schwarz et al., 2014)	ClinVar Miner (Henrie et al., 2018)	cDNA change	Protein Change
P1	ZFPM2	rs202217256	0.00366	0.00567	T (0.422)	P (0.914)	D (1.0)	benign	NM_012082.3: c.292G>A	NP_036214.2: p.Asp98Asn
P2	ESR2	rs367855747	0.0002	0.0004	D (0.021)	D (0.975)	D (0.999)	n.a.	NM_001214902.1: c. 1331G>A	NP_001201831.1: p.Ser444Asn

All three mutations are heterozygous missense SNVs.

SIFT: D = Deleterious, T = Tolerated; Polyphen2: B = benign, D = probably damaging, P = possibly damaging; MutationTaster: D = disease causing

References:

Adzhubei, I.A., S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, and S.R. Sunyaev. 2010. A method and server for predicting damaging missense mutations. *Nat Methods*. 7:248-249.

Barbaro, M., A. Cicognani, A. Balsamo, A. Lofgren, L. Baldazzi, A. Wedell, and M. Oscarson. 2008. Gene dosage imbalances in patients with 46,XY gonadal DSD detected by an in-house-designed synthetic probe set for multiplex ligation-dependent probe amplification analysis. *Clinical genetics*. 73:453-464.

Barbaro, M., J. Cook, K. Lagerstedt-Robinson, and A. Wedell. 2012. Multigeneration Inheritance through Fertile XX Carriers of an NR0B1 (DAX1) Locus Duplication in a Kindred of Females with Isolated XY Gonadal Dysgenesis. *International journal of endocrinology*. 2012:504904.

Barbaro, M., M. Oscarson, J. Schoumans, J. Staaf, S.A. Ivarsson, and A. Wedell. 2007. Isolated 46,XY gonadal dysgenesis in two sisters caused by a Xp21.2 interstitial duplication containing the DAX1 gene. *The Journal of clinical endocrinology and metabolism*. 92:3305-3313.

Dong, Y., Y. Yi, H. Yao, Z. Yang, H. Hu, J. Liu, C. Gao, M. Zhang, L. Zhou, Asan, X. Yi, and Z. Liang. 2016. Targeted next-generation sequencing identification of mutations in patients with disorders of sex development. *BMC medical genetics*. 17:23.

Fishilevich, S., R. Nudel, N. Rappaport, R. Hadar, I. Plaschkes, T. Iny Stein, N. Rosen, A. Kohn, M. Twik, M. Safran, D. Lancet, and D. Cohen. 2017. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*. 2017.

Henrie, A., S.E. Hemphill, N. Ruiz-Schultz, B. Cushman, M.T. DiStefano, D. Azzariti, S.M. Harrison, H.L. Rehm, and K. Eilbeck. 2018. ClinVar Miner: Demonstrating utility of a Web-based tool for viewing and filtering ClinVar data. *Human mutation*. 39:1051-1060.

Kent, W.J., C.W. Sugnet, T.S. Furey, K.M. Roskin, T.H. Pringle, A.M. Zahler, and D. Haussler. 2002. The human genome browser at UCSC. *Genome research*. 12:996-1006.

Ledig, S., O. Hiort, G. Scherer, M. Hoffmann, G. Wolff, S. Morlot, A. Kuechler, and P. Wieacker. 2010. Array-CGH analysis in patients with syndromic and non-syndromic XY gonadal dysgenesis: evaluation of array CGH as diagnostic tool and search for new candidate loci. *Hum Reprod*. 25:2637-2646.

Lek, M., K.J. Karczewski, E.V. Minikel, K.E. Samocha, E. Banks, T. Fennell, A.H. O'Donnell-Luria, J.S. Ware, A.J. Hill, B.B. Cummings, T. Tukiainen, D.P. Birnbaum, J.A. Kosmicki, L.E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D.N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M.I. Kurki, A.L. Moonshine, P. Natarajan, L. Orozco, G.M. Peloso, R. Poplin, M.A. Rivas, V. Ruano-Rubio, S.A. Rose, D.M. Ruderfer, K. Shakir, P.D. Stenson, C. Stevens, B.P. Thomas, G. Tiao, M.T. Tusie-Luna, B. Weisburd, H.H. Won, D. Yu, D.M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J.C. Florez, S.B. Gabriel, G. Getz, S.J. Glatt, C.M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M.I. McCarthy, D. McGovern, R. McPherson, B.M. Neale, A. Palotie, S.M. Purcell, D. Saleheen, J.M. Scharf, P. Sklar, P.F. Sullivan, J. Tuomilehto, M.T. Tsuang, H.C. Watkins, J.G. Wilson, M.J. Daly, D.G. MacArthur, and Exome Aggregation Consortium. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 536:285-291.

Pattyn, F., F. Speleman, A. De Paepe, and J. Vandesompele. 2003. RTPrimerDB: the real-time PCR primer and probe database. *Nucleic Acids Res*. 31:122-123.

Robinson, J.T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, and J.P. Mesirov. 2011. Integrative genomics viewer. *Nature biotechnology*. 29:24-26.

Schwarz, J.M., D.N. Cooper, M. Schuelke, and D. Seelow. 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 11:361-362.

Smyk, M., J.S. Berg, A. Pursley, F.K. Curtis, B.A. Fernandez, G.A. Bien-Willner, J.R. Lupski, S.W. Cheung, and P. Stankiewicz. 2007. Male-to-female sex reversal associated with an approximately 250 kb deletion upstream of NR0B1 (DAX1). *Hum Genet*. 122:63-70.

Vaser, R., S. Adusumalli, S.N. Leng, M. Sikic, and P.C. Ng. 2016. SIFT missense predictions for genomes. *Nat Protoc*. 11:1-9.

Wang, Y., F. Song, B. Zhang, L. Zhang, J. Xu, D. Kuang, D. Li, M.N.K. Choudhary, Y. Li, M. Hu, R. Hardison, T. Wang, and F. Yue. 2018. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome biology*. 19:151.

White, S., T. Ohnesorg, A. Notini, K. Roeszler, J. Hewitt, H. Daggag, C. Smith, E. Turbitt, S. Gustin, J. van den Bergen, D. Miles, P. Western, V. Arboleda, V. Schumacher, L. Gordon, K. Bell, H. Bengtsson, T. Speed, J. Hutson, G. Warne, V. Harley, P. Koopman, E. Vilain, and A. Sinclair. 2011. Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. *PloS one*. 6:e17793.