SYCP2 Translocation-Mediated Dysregulation and Frameshift Variants Cause Human Male Infertility

Samantha L.P. Schilit,[1,8] Shreya Menon,[2] Corinna Friedrich,[3] Tammy Kammin,[4] Ellen Wilch,[4] Carrie Hanscom,[5] Sizun Jiang,[6] Sabine Kliesch,[7] Michael E. Talkowski,[5,8,9,10] Frank Tüttelmann,[3,14,]* Amy J. MacQueen,[11,14] and Cynthia C. Morton[4,8,10,12,13,14,]*

Unexplained infertility affects 2%–3% of reproductive-aged couples. One approach to identifying genes involved in infertility is to study subjects with this clinical phenotype and a *de novo* balanced chromosomal aberration (BCA). While BCAs may reduce fertility by production of unbalanced gametes, a chromosomal rearrangement may also disrupt or dysregulate genes important in fertility. One such subject, DGAP230, has severe oligozoospermia and 46,XY,t(20;22)(q13.3;q11.2). We identified exclusive overexpression of *SYCP2* from the der(20) allele that is hypothesized to result from enhancer adoption. Modeling the dysregulation in budding yeast resulted in disrupted structural integrity of the synaptonemal complex, a common cause of defective spermatogenesis in mammals. Exome sequencing of infertile males revealed three heterozygous *SYCP2* frameshift variants in additional subjects with cryptozoospermia and azoospermia. In sum, this investigation illustrates the power of precision cytogenetics for annotation of the infertile genome, suggests that these mechanisms should be considered as an alternative etiology to that of segregation of unbalanced gametes in infertile men harboring a BCA, and provides evidence of *SYCP2*-mediated male infertility in humans.

Introduction

Infertility affects 10%–15% of couples, making it one of the most common disorders for individuals between the ages of 20 and 45 years.[1] While many factors contribute to infertility including anatomical defects, gamete integrity, hormonal dysregulation, environmental exposures, age, and certain genetic syndromes, at least 1 in 5 cases of infertility are "unexplained."[2] Genetic defects may be responsible for many of these idiopathic cases. Indeed, mutations in more than 600 genes have been shown to decrease fertility in animal models, and yet few genetic causes of infertility have been validated in humans.[3,4] This results in part from the decreased reproductive fitness of infertile individuals which reduces the number of large families available for genetic analysis in humans as well as the genetic heterogeneity of the disorder.[5,6]

Male factors contribute to about half of all infertility cases, and 40%–72% of men lack a specific causal diagnosis.[5,7,8] Given that 30%–50% of these cases are estimated to have genetic etiologies, searching for genes involved in unexplained infertility is a rich endeavor.[5,7] Uncovering additional causes not only informs an understanding of mechanisms regulating fertility, but also provides clinical information to support diagnosis, genetic counseling, and therapeutic intervention.

Currently, the most common genetic testing for male infertility in the clinic involves sequencing of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene (pathogenic variants found in 80%–90% of cases of congenital bilateral absence of the vas deferens), assessing Y chromosome microdeletions (accounting for 5%–15% of men with severe oligozoospermia [<5 M/mL] or nonobstructive azoospermia), and karyotype analysis (revealing Klinefelter syndrome in 14% of azoospermic men).[3,7,9,10] In 0.5%–1% of severe oligozoospermic or azoospermic men, a karyotype may reveal an apparently balanced reciprocal translocation.[11] While balanced reciprocal translocations may cause subfertility by doubling the risk of miscarriage, the mechanism for how such chromosomal abnormalities may lead to low sperm counts has not been rigorously investigated.[12,13]

We hypothesize that a chromosomal rearrangement may disrupt or dysregulate genes important for fertility

[1]Program in Genetics and Genomics and Certificate Program in Leder Human Biology and Translational Medicine, Biological and Biomedical Sciences Program, Graduate School of Arts and Sciences, Harvard University, Cambridge, MA 02138, USA; [2]Harvard College, Harvard University, Cambridge, MA 02138, USA; [3]Institute of Human Genetics, University of Münster, 48149 Münster, Germany; [4]Department of Obstetrics and Gynecology, Brigham and Women's Hospital, Boston, MA 02115, USA; [5]Molecular Neurogenetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA; [6]Baxter Laboratory for Stem Cell Biology, Stanford University School of Medicine, Stanford, CA 94305, USA; [7]Centre of Reproductive Medicine and Andrology, Department of Clinical and Surgical Andrology, University Hospital Münster, 48149 Münster, Germany; [8]Harvard Medical School, Boston, MA 02115, USA; [9]Departments of Neurology, Psychology, and Pathology, Massachusetts General Hospital, Boston, MA 02114, USA; [10]Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA; [11]Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06459, USA; [12]Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA; [13]Manchester Academic Health Science Centre, University of Manchester, Manchester M13 9NT, UK
[14]These authors contributed equally to this work
*Correspondence: frank.tuettelmann@ukmuenster.de (F.T.), cmorton@bwh.harvard.edu (C.C.M.)
https://doi.org/10.1016/j.ajhg.2019.11.013.

in the immediate vicinity of rearrangement breakpoints. By using the well-established Developmental Genome Anatomy Project (DGAP) infrastructure, we initiated a study aimed at identifying genes important for fertility and exploring an additional explanation for how balanced reciprocal translocations reduce fertility.

Material and Methods

DGAP Participant Recruitment

Subjects DGAP230, with 46,XY,t(20;22)(q13.3;q11.2), and DGAP278-02, a karyotypically normal age- and sex-matched control subject, were enrolled through the Developmental Genome Anatomy Project (DGAP). DGAP obtained informed consent, medical records, and blood samples under a protocol approved by the Partners HealthCare Systems Institutional Review Board.

Acquisition of Lymphoblastoid Cell Lines

Epstein-Barr virus-transformed lymphoblastoid cell lines (LCLs) were generated at the Genomics and Technology Core in the Center for Human Genetic Research at Massachusetts General Hospital (Boston, MA, USA). Confirmatory GTG-banded karyotyping was performed on the DGAP230 lymphoblastoid cell line at the Brigham and Women's Hospital CytoGenomics Core Laboratory (Boston, MA, USA). Two additional karyotypically normal age- and sex-matched control LCLs, GM20184 and GM20188, were obtained from the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Institute for Medical Research (Camden, NJ, USA).

Genome Jumping Library Sequencing and Analysis

Large-insert ("jumping library") genome sequencing and subsequent Sanger sequencing identified the precise breakpoints of the DGAP230 chromosomal rearrangement (using primers SS_58F-SS_61R from Table S2).[14–16] Chromatograms were analyzed in Geneious (v.7.0, Biomatters) and described using next-generation cytogenetic nomenclature.[17]

Delineation of Topologically Associating Domains Disrupted by DGAP230's Rearrangement Breakpoints

Topologically associating domains (TADs) disrupted by the breakpoints in DGAP230 were identified according to human embryonic stem cell Hi-C domains from the Hi-C project.[18] The University of California Santa Cruz Genome Browser was used to delineate genes residing in these regions.[19]

Culturing of Lymphoblastoid Cell Lines

Lymphoblastoid cells were grown at 37°C, 5% CO$_2$, in R10 media (Roswell Park Memorial Institute [RPMI] 1640 Medium without L-Glutamine supplemented with 10% heat-inactivated fetal bovine serum [FBS], 100 U/mL penicillin-streptomycin-glutamine, and 2 mM L-glutamine [all GIBCO]). Cells were fed fresh R10 every 2 or 3 days. For clonal expansion of single lymphoblastoid cells in the CRISPR/Cas9-mediated genome editing experiment, 30% 0.45 μm-filtered and irradiated conditioned R10 with 45% fresh R10 and 25% FBS was used and feeding only occurred to offset media evaporation.

RNA Extraction, cDNA Synthesis, and Quantitative Real-Time Polymerase Chain Reaction (RT-PCR) of Lymphoblastoid Cells

40–50 million cells were collected for each RNA extraction. Cells were pelleted for 5 min at 200 x g, washed in phosphate-buffered saline (PBS; GIBCO), and resuspended in 1 mL TRIzol (Invitrogen). Cell lysis and phase separation with chloroform (Sigma) were performed as described from the TRIzol manufacturer protocol. Total RNA was isolated using the RNeasy Midi Kit (QIAGEN) with in-column RNase-free DNase I (QIAGEN) digestion. RNA was converted to cDNA using 4 μg of RNA per reaction, random hexamer primers, and the SuperScript III First-Strand Synthesis System (Invitrogen) according to manufacturer instructions. For quantitative RT-PCR, cDNA was amplified with primers SS_36F-SS_55R and SS_62F-SS_147R (exon-spanning when possible; sequences reported in Table S2) and SsoFast EvaGreen Supermix (Bio-Rad) using the CFX Connect Real-Time PCR Detection System (Bio-Rad). Standard curves were employed routinely to quantify amplicons from each primer pair and assess expression of each respective gene in lymphoblastoid cells.

Western Blot Analysis

25 million lymphoblastoid cells were collected for each protein extraction. Cells were pelleted for 5 min at 200 x g, washed in phosphate-buffered saline (PBS; GIBCO), and resuspended in 200 μL Radioimmunoprecipitation assay buffer (RIPA; Sigma-Aldrich) with cOmplete protease inhibitors (Roche). After trituration and agitation for 30 min at 4°C, cellular debris was pelleted for 20 min at 17,000 x g and 6X Laemmli buffer (375 mM Tris-HCl, 9% SDS, 50% glycerol, 9% β-mercaptoethanol, and 0.03% bromophenol blue [all Sigma]) was added to the supernatant to a final concentration of 1×. Samples were passed through a 27-gauge syringe 10 times and boiled at 95°C for 10 min. Samples and a positive control of human testis tissue lysate (Abcam, catalog number ab30257) were loaded into NuPAGE 4%–12% Bis-Tris Protein gels (Novex) and run with NuPAGE MOPS SDS Running Buffer (Novex) in an XCell SureLock Mini-Cell Electrophoresis System (Novex) according to manufacturer instructions. Gels were transferred to 0.2 μm nitrocellulose membranes (Bio-Rad) using NuPAGE Transfer Buffer and an XCell II Blot Module (Invitrogen) overnight at 100 mA. Membranes were probed with primary antibodies anti-SYCP2 (EMD-Millipore, catalog number ABE2622) and anti-GAPDH (Cell Signaling Technology, catalog number 5174) and secondary antibody IRDye 800CW donkey anti-rabbit IgG (LI-COR). Membranes were visualized using the Odyssey Fc Imaging System (LI-COR) and signal intensities were quantified in ImageJ (v.1.46).

Identification and Evaluation of Variable Regions in *SYCP2* and *GAPDH*

SYCP2 and *GAPDH* exons were compared against the database of Single Nucleotide Polymorphisms (dbSNP) to identify variable nucleotides.[20] To assess heterozygosity in DGAP230, genomic DNA of DGAP230 lymphoblastoid cells that had been pelleted and washed with phosphate-buffered saline (PBS, GIBCO) was extracted using the DNeasy Blood and Tissue Kit (QIAGEN). PCR was performed with LongAmp Taq 2X Master Mix (New England Biolabs, NEB) using customized primers (Integrated DNA Technologies). After amplification confirmation with agarose gel electrophoresis and purification with the QIAquick PCR Purification Kit (QIAGEN), Sanger sequencing reactions of PCR products were carried out

with an ABI3730xl DNA analyzer. Chromatograms were aligned and multiple single-nucleotide variants were called using Geneious (v.7.0, Biomatters). The target exonic regions were selected based upon presence of multiple single-nucleotide variants (amplified and Sanger sequenced with primers SS_174F-SS_175R and SS_192F-SS_193R from Table S2). DGAP230 lymphoblastoid cell cDNA was acquired as described above under RNA Extraction, cDNA Synthesis, and Quantitative Real-Time Polymerase Chain Reaction (RT-PCR) of Lymphoblastoid Cells and assessed by the same method as the genomic DNA (amplified and Sanger sequenced with primers SS_175R-SS_193R as reported in Table S2).

3C-PCR

3C-PCR was performed as previously described.[21] In brief, standard 3C libraries were generated[22–26] and subjected to rearrangement-spanning nested PCR and subsequent Sanger sequencing using primers SS_216R-SS_227R (Table S2).

Circular Chromatin Conformation Capture Sequencing (4C-Seq) and Analysis

4C-seq was adapted from previously described protocols with the following specifications.[23–25] Ten million cell aliquots of lymphoblastoid cells were crosslinked by adding formaldehyde (MilliporeSigma) to a final concentration of 2% for 10 min before the reaction was quenched with glycine (Promega) at a final concentration of 125 mM. After lysis, chromatin was further released by douncing with a 21.5-gauge needle and digested using 1,500 U HindIII (NEB) per reaction at 37°C overnight. After enzyme inactivation, ligation, reverse crosslinking, DNA isolation, and purification, DNA was subjected to a second digestion by 50 U Csp6I (Thermo Fisher Scientific) per reaction at 37°C overnight. Enzyme inactivation, ligation, DNA isolation, and purification were next performed to create the 4C libraries. Each library was amplified by inverse PCR with primers containing multiplexed overhangs (SS_198F-SS_212R as reported in Table S2) using 50 ng of 4C template per reaction and purified by Qiaquick PCR purification kit (QIAGEN). Samples were submitted to The Biopolymers Facility in the Department of Genetics at Harvard Medical School (Boston, MA, USA) for quality control on an Agilent 2200 TapeStation D1000 HS ScreenTape and by SYBR qPCR assay. Samples were spiked with 30% PhiX prior to sequencing on an Illumina HiSeq 2500 sequencer using the HiSeq Single-Read Rapid Cluster Kit v2 (Illumina) for 100 cycles. At least 14 million pass-filter reads were acquired per sample.

4C-seq datasets were imported into the Galaxy-BioTeam Appliance and demultiplexed using Barcode Splitter (Galaxy v.1.0.0).[27] Reads were aligned to GRCh37/hg19 and a custom-made derivative chromosome 20 (der(20)) hg19 reference genome using Bowtie2 (Galaxy v.0.2). The der(20) was made using a Perl script that parsed through chromosomes 20 and 22 and concatenated parts of the chromosomes at the resolved breakpoints (Web Resources). Bam files were then analyzed using FourCSeq (v.1.16.0) with normal parameters established by the FourCSeq protocol with the exception of using the der(20) for some analyses.[28] The interactions between fragments were then mapped to a circular plot using Circos (v.0.69-6), by applying the relative coordinates of each fragment within the respective chromosome.[29] p-adjusted significance values were used to superimpose a colored heatmap onto the plot. These were then transformed into linear plots using Photoshop (version CC 2015) by converting the image from Rectangular mode to Polar mode.

CRISPR/Cas9-Mediated Genome Editing in Lymphoblastoid Cells

CRISPR/Cas9-mediated deletion of chromosome 22 putative enhancers in the DGAP230 LCL was performed as previously described with the following parameters.[30] A Cas9-expressing stable DGAP230 cell line was developed by transduction with lenti-Cas9-Blast lentivirus, selection in blasticidin S HCl (GIBCO), and validation by western blot with primary antibodies anti-Cas9 (Abcam, catalog number 7A9-3A3) and anti-GAPDH (Cell Signaling Technology, catalog number 5174) and secondary antibodies IR-Dye 680RD Donkey anti-Mouse IgG and IRDye 800CW Donkey anti-Rabbit IgG (both LI-COR). This cell line was transduced with lentivirus derived from lentiviral vectors containing custom-made dual targeting sgRNAs that flank the putative enhancer region as well as a lentiGuide-Puro empty vector control. The dual targeting sgRNA vector was constructed and validated with primers SS_278F-SS_289F (Table S2) and the previously described H1 bridge dsDNA block.[30] Selection of infected cells in puromycin (GIBCO) yielded heterogeneous cultures, which were then either grown for 1 to 4 months or sorted into single cells in 96-well cell culture plates (Corning) using a MoFlo Astrios EQ cell sorter with a 100 μm nozzle at the Harvard Medical School Division of Immunology Flow Cytometry Core (Boston, MA, USA). Single-cell clones were PCR validated for homozygous deletions after genomic DNA extraction with the DNeasy Blood and Tissue Kit (QIAGEN) by using LongAmp Taq 2X Master Mix (NEB) and primers SS_304F-SS_309R (Table S2). To determine the relative amount of putative enhancer deletion in heterogeneous cultures derived from the dual targeting sgRNA lentivirus transduction and grown for different lengths of time, quantitative RT-PCR of 12 ng genomic DNA was performed with primers SS_308F-SS_309R and SS_52F-SS_53R (Table S2) and SsoFast EvaGreen Supermix (Bio-Rad) using the CFX Connect Real-Time PCR Detection System (Bio-Rad). Standard curves were employed routinely to quantify amplicons from each primer pair.

Construction of Yeast Strains

All strains used in this study (Table S1) are isogenic to BR1919-8B.[31] Strains were created by standard genetic crosses and transformation procedures. For the development of strains AM3762 and AM4282, a TRP1:P_{GAL1} promoter cassette was amplified from pFA6a-TRP1-P_{GAL1}[32] using Velocity polymerase (Bioline) and primers AJM1741-AJM1742 (Table S2).

Cytological Analysis and Imaging

Induction, sporulation, chromosome spreading, immunostaining, imaging, and analysis were performed on diploid strains as previously described[33,34] using the following parameters. Overnight cultures of YAM2592 (wild-type) and AM3762 (P_{GAL1}-RED1) were resuspended in 2% potassium acetate and split into two cultures per strain. Immediately after resuspension, one culture per strain was induced with 2 μm β-estradiol (Sigma E2257, prepared in DMSO) and the second received the corresponding volume of DMSO as an uninduced control. Cells were collected at 26 h post sporulation and induction for chromosome spreading. The following primary antibodies were used for immunostaining at a 1:100 dilution: rabbit anti-Red1 (kind gift of G.S. Roeder)[35] and affinity-purified rabbit anti-Zip1 (raised at YenZym Antibodies against a C-terminal fragment of Zip1).[36] The secondary antibody, donkey anti-rabbit conjugated to Alexa Fluor 488 (Abcam), was used at a 1:200 dilution. Imaging was carried out for eight different chromosome spreads derived from two distinct cultures per

condition using a Deltavision RT imaging system (Applied Precision) adapted to an Olympus (IX71) microscope. Zip1 lengths were measured using the Softworx Measure Distance Tool.

RNA Extraction, cDNA Synthesis, and Quantitative RT-PCR of *S. cerevisiae*

Induction and sporulation were performed on diploid strains as previously described.[33,34] Overnight cultures of YAM2592 (wild-type), AM4063 (wild-type), AM3762 (P_{GAL1}-*RED1*), AM4282 (P_{GAL1}-*RED1*), AM4283 (Δ*zip1*), AM4284 (Δ*red1*), and AM4286 (Δ*red1*) were resuspended in 2% potassium acetate and split into two cultures per strain. Immediately after resuspension, one culture per strain was induced with 2 μm β-estradiol (Sigma E2257, prepared in DMSO) and the second received the corresponding volume of DMSO as an uninduced control. Cells from 5 mL cultures were collected at 0, 6, and 26 h post sporulation and induction and resuspended in 500 μL TRIzol (Invitrogen). RNA extraction, cDNA synthesis, and quantitative RT-PCR in *S. cerevisiae* were performed as described in the corresponding materials and methods section for LCLs with the exception of the following differences. Cells were disrupted by vortex with 0.5 mm glass beads (BioSpec Products) for 30 min at 4°C before phase separation by MaXtract High Density Tubes (QIAGEN). Total RNA was isolated using the RNeasy Mini Kit (QIAGEN) and RNA was converted to cDNA using 60 ng of RNA per reaction, oligo(dT) primers, and the SuperScript III First-Strand Synthesis System (Invitrogen). For quantitative RT-PCR, cDNA was amplified with primers SS_326F-SS_341R (sequences reported in Table S2). Standard curves were employed routinely to quantify amplicons from each primer pair and assess expression of each respective gene in *S. cerevisiae* cells.

Münster Male Infertility Cohort Recruitment

Males attended the Centre of Reproductive Medicine and Andrology (CeRA) of the University Hospital Münster (UKM) to assess their fertility and were diagnosed by means of semen analysis performed according to guidelines of the World Health Organization (WHO).[37] After routine clinical diagnostics, males with known causes for their infertility including malignant disease, exposure to chemotherapy or radiation, structural and numerical chromosomal aberrations, and Y chromosome microdeletions were excluded from the study. All participants who gave written informed consent for evaluation of their clinical data and genetic analysis of their DNA samples were included in the MERGE (Male Reproductive Genomics) study of the Institute of Human Genetics, University of Münster. The study was approved by the Ethics Committee of the State Medical Board and the Medical Faculty in Münster (Kennzeichen 2010-578-f-S).

Exome Sequencing and Analysis

Exome sequencing (ES) was performed in 627 patients with diverse infertility phenotypes from the MERGE study to identify possible deleterious sequence variants which might be causal for male infertility. Genomic DNA was isolated using standard procedures as previously described.[38] Samples were prepared, enriched, and indexed for ES according to the manufacturer's protocol for SureSelect[QXT] Target Enrichment for Illumina Multiplexed Sequencing Featuring Transposase-Based Library Prep Technology (Agilent). For multiplexed sequencing, libraries were index-tagged using appropriate pairs of index primers. To capture libraries, SureSelect[XT] Human All Exon Kits (v.4, v.5, and v.6) were used. Quan-

tity and quality of libraries were determined with an Agilent TapeStation 2200 and the final concentration was adjusted to 1.6 pM. Sequencing was performed on the Illumina HiScanSQ System, the Illumina NextSeq500 System, or the Illumina HiSeqX System using the TruSeq SBS Kit v3 - HS (200 cycles), the NextSeq 500 v2 High-Output Kit (300 cycles), or the HiSeq Rapid SBS Kit v2 (300 cycles), respectively.

Reads were trimmed with Cutadapt v.1.15[39] and aligned to GRCh37.p13 using BWA-MEM v.0.7.17.[40] Base quality recalibration and variant calling were performed with the GATK toolkit v.3.8[41] with haplotype caller according to their best practice recommendations. Resulting variants were annotated with Ensembl Variant Effect Predictor.[42]

Before evaluation of variants in *SYCP2* (GenBank: NM_014258.3), participants with likely pathogenic or pathogenic variants in *TEX11*, *NR5A1*, and *DMRT1* were excluded from the study.[38,43,44] Rare variants in *SYCP2* were selected by a minor allele frequency (MAF) ≤ 1% in the Genome Aggregation Database (gnomAD) browser[45] and assessed for functional consequences at the protein level.

Confirmatory Sanger Sequencing

Identified variants in *SYCP2* were verified by Sanger sequencing according to standard procedures using primers SS_369F-SS_372R and SS_375F-SS_376R (Table S2) and a 3730 DNA Analyzer (Applied Biosystems).[38] When parental DNA was available, segregation was analyzed within the family using Sanger sequencing. Sequence analysis and visualization of the chromatograms was performed with CodonCode Aligner software (v.8.0.1).

Testicular Biopsy Histopathology

Testicular biopsies were fixed overnight in Bouin's solution, washed with 70% ethanol, and embedded in paraffin. Subsequently, 5 μm sections were stained with Periodic acid-Schiff (PAS) and hematoxylin according to previously published protocols.[46] Slides were evaluated and documented using an Axioskop microscope (Zeiss, Oberkochen, Germany).

Statistical Analyses

For quantitative RT-PCR experiments, cDNA for three DGAP230 LCL replicates and three distinct age- and sex-matched control LCLs were each assessed in technical duplicates, normalized to *GAPDH* expression, and evaluated using an unpaired two-tailed t test (Excel). For western blot analysis, signal intensities of SYCP2 relative to GAPDH for three DGAP230 LCL replicates and three distinct age- and sex-matched control LCLs were assessed and evaluated using an unpaired one-tailed t test (Excel). Statistical significance for RNA and protein analyses were determined by a p value of $p < 0.05$ and figure error bars show standard error of the mean. For identifying and assessing variable regions in DGAP230 LCLs, Sanger sequencing was performed on at least three different preparations of genomic DNA and/or cDNA. 3C-PCR and 4C-seq were each performed in triplicate, creating both 3C and 4C libraries from three independent cultures for the DGAP230 LCL and three different control LCLs. For differential analysis of 4C-seq datasets using FourCSeq (v.1.16.0),[28] read-counts were variance stabilized and fit with a monotonic decay that was used to calculate z-scores. Z-scores were converted to p values using a normal cumulative distribution curve and adjusted for multiple testing by the Benjamini-Hochberg method. A differential analysis was then conducted and significant interactions were detected by a p-adjusted value of $p_{adj} < 0.01$. For quantitative RT-PCR of the
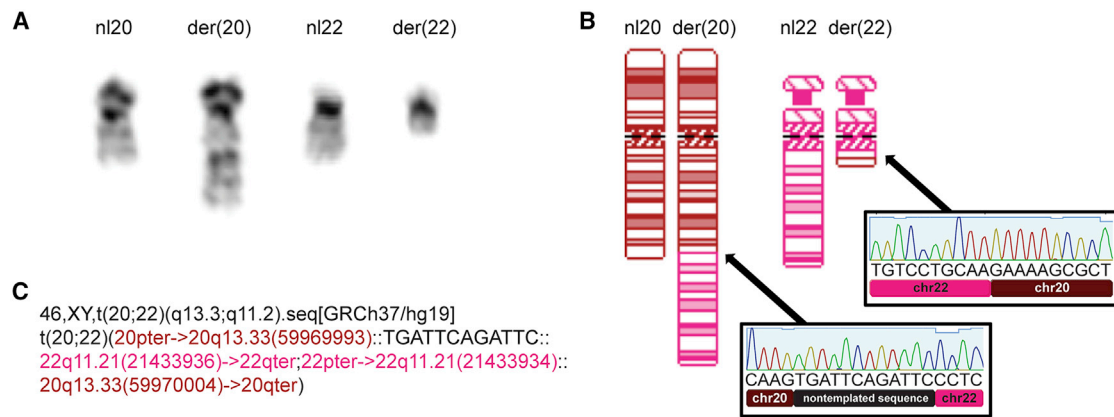
Figure 1. DGAP230 Breakpoint Characterization
(A) DGAP230 composite partial karyotype of normal (left) and derivative (right) GTG-banded chromosomes 20 and 22.
(B) The DGAP230 ideogram indicates a simple long arm (q-q) translocation between chromosomes 20 (nl20) and 22 (nl22). Chromatograms from Sanger sequencing show rearrangement breakpoints at nucleotide-level resolution.
(C) Next-generation cytogenetic nomenclature is provided for DGAP230.[17]

putative enhancer deletion, genomic DNA for two biological replicates of DGAP230 LCLs transduced with the dual targeting sgRNA lentivirus and grown for 1 month, 15 biological replicates of DGAP230 LCLs transduced with the dual sgRNA lentivirus and each grown for 4 months, and two biological replicates of DGAP230 LCLs transduced with the lentiGuide-Puro empty vector control were each assessed in technical triplicates, normalized to an intronic region of genomic DNA, and evaluated using a heteroscedastic one-tailed t test for 1 versus 4 months and homoscedastic one-tailed t test for dual targeting sgRNA lentivirus versus empty vector control (Excel). Statistical significance for quantitative RT-PCR analyses on genomic DNA was determined by a p value of $p < 0.05$ and figure error bars show standard error of the mean. For yeast cytology, at least 50 surface-spread meiotic nuclei were assessed per condition using eight different chromosome spreads derived from two distinct cultures. Graphpad Prism7 software was used for scatterplot generation of Zip1 filament length (with error bars indicating standard error of the mean) and statistical significance was determined by a Mann-Whitney U test, using VassarStats Concepts and Applications of Inferential Statistics. Yeast qPCR experiments were performed in triplicate in two different strain backgrounds (n = 6), normalized to *ACT1* expression, and evaluated using an unpaired one-tailed t test for *RED1* and an unpaired two-tailed t test for *ZIP1* (Excel). Statistical significance for RNA analyses was determined by a p value of $p < 0.05$ and figure error bars show standard error of the mean. To determine the prevalence of *SYCP2* loss-of-function (LoF) variants in the Münster infertility cohort compared to the general population, 74 *SYCP2* LoF variants were found to be listed in gnomAD, resulting in an overall prevalence of 0.05% (v.2.1.1, queried 9 September 2019, containing information on 141,456 individuals). Prevalence was also calculated with low confidence variants excluded, leaving 14 variants and resulting in a carrier prevalence of 0.01%. Chi-square analysis was used to compare these frequencies to that of the Münster infertility cohort (3 out of 627 men, or 0.48%).

Results

Clinical Report for DGAP230
We sought to identify the genetic etiology of infertility for a male research participant, designated DGAP230, who pre-

sented with a 2-year history of infertility at age 28. His evaluation showed severe oligozoospermia (<2 M/mL) with normal semen volume. GTG-banded metaphase chromosomes revealed the apparently balanced reciprocal translocation 46,XY,t(20;22)(q13.3;q11.2) (Figure 1A). Microarray analysis using a 135K-feature whole-genome microarray (SignatureChip Oligo Solution v.2.0, Signature Genomics Laboratories) indicated no clinically significant abnormalities and no detectable gains or losses of genomic material at either breakpoint. Genetic testing for the Y chromosome microdeletions (YCMD) AZFa, AZFb, and AZFc as well as *CFTR* pathogenic variants were negative. DGAP230 has no dysmorphic features and has normal serum levels of follicle-stimulating hormone (FSH), luteinizing hormone (LH), and testosterone. Given DGAP230's normal hormonal levels, normal semen volume, and absence of *CFTR* pathogenic variants, we predicted that the etiology of his severe oligozoospermia likely originates in the testes.[47] Based on these initial observations, we hypothesized that the t(20;22) breakpoints disrupt or dysregulate gene(s), which may be important in fertility, perhaps by impacting gametogenesis.

Identification of Candidate Genes for DGAP230's Severe Oligozoospermia
To identify gene(s) disrupted or potentially dysregulated by the t(20;22) breakpoints, the translocation was refined to nucleotide resolution using large-insert ("jumping library") sequencing and subsequent Sanger sequencing (Figures 1B and 1C).

Breakpoint locations were then used to identify candidate genes potentially etiologic in DGAP230's severe oligozoospermia. Genetic regions specifically disrupted by the t(20;22) may be impacted in two ways. If the breakpoint occurs within the open reading frame of any gene, the gene is considered disrupted. Alternatively, positional effects on genes near breakpoints could occur due to loss or gain of regulatory elements in the genetic neighborhood. We defined the potential extent of these effects by
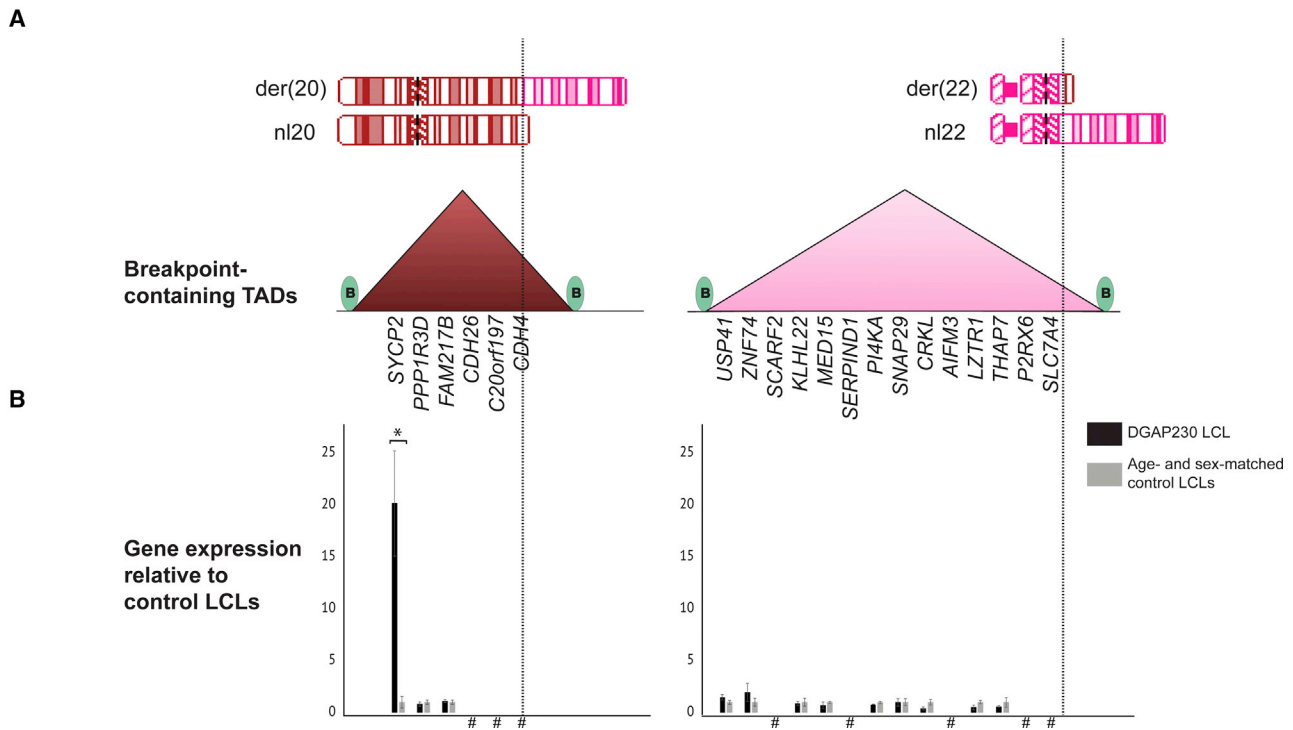
Figure 2. Topologically Associating Domains Disrupted by DGAP230's t(20;22) Define Candidate Genes
(A) Vertical lines indicate positions of rearrangement breakpoints across ideograms of normal and derivative chromosomes 20 (mahogany) and 22 (pink). Genes residing in breakpoint-containing topologically associating domains (TADs) are listed below each TAD (triangle) defined by boundary regions (green ovals).
(B) qPCR of candidate genes in the DGAP230 LCL relative to age- and sex-matched control LCLs. Statistical significance was determined by an unpaired two-tailed t test (n = 3; p < 0.02) and all results display mean ± standard error. #genes deemed not expressed in LCLs by qPCR standard curve; *p < 0.05.

assessing the breakpoint-residing topologically associating domains (TADs), compartments of chromatin with frequent interactions in three-dimensional space, because TAD disruption by structural rearrangements can rewire gene expression and induce pathogenicity.[18,48,49] As a result, our list of candidate genes includes all genes disrupted or potentially dysregulated that reside within the breakpoint-containing TADs.

We computationally assessed predicted TADs by using stem cell Hi-C domains from the Hi-C project and converting them to hg19 for comparison to the breakpoints.[18] We identified six genes in the chromosome 20 (chr20) breakpoint-containing TAD and 14 genes in the chromosome 22 (chr22) breakpoint-containing TAD (Figure 2A). While one gene, *CDH4* (MIM: 603006), is directly disrupted by the breakpoint on chromosome 20, it is not a strong candidate for severe oligozoospermia. *CDH4* encodes cadherin-4 or R-cadherin, which has predominant expression in the brain and plays a role in retinal axon outgrowth and visual system development.[50–52] The disruption is likely not pathogenic because it has a high haploinsufficiency score (%HI = 37), suggesting that disruption of only one copy of *CDH4* is not sufficient to induce pathogenicity.[53] In addition, paternally inherited *CDH4* deletions are reported in DECIPHER, which would be incompatible with a phenotype of male infertility.[54]

To determine whether DGAP230's translocation dysregulates candidate genes, we assessed expression of the 20 genes. Using an Epstein-Barr virus-transformed lymphoblastoid cell line (LCL), which has shown prior utility in deducing the effects of structural rearrangements in DGAP subjects,[55,56] we hypothesized that we would be able to detect genes that exhibit constitutive misexpression or overexpression compared to age- and sex-matched control subjects.[52,53] By assessing expression of every candidate gene by qPCR, we identified one gene on chr20, *SYCP2* (MIM: 604105), whose expression is increased more than 20-fold in DGAP230 LCLs relative to age- and sex-matched control subjects, a result that was confirmed to be statistically significant by an unpaired two-tailed t test (p < 0.02). No other significant differences were noted in expression between DGAP230 and control LCLs (Figure 2B).

Determination of the Cytogenetic Etiology of *SYCP2* Dysregulation

To determine whether *SYCP2* misexpression is caused by the t(20;22), we next characterized the mechanism by which *SYCP2* is dysregulated. Given that *SYCP2* is not expressed in karyotypically normal LCLs, we hypothesized that only the der(20)-residing allele would be expressed if the t(20;22) were responsible. To test the hypothesis that
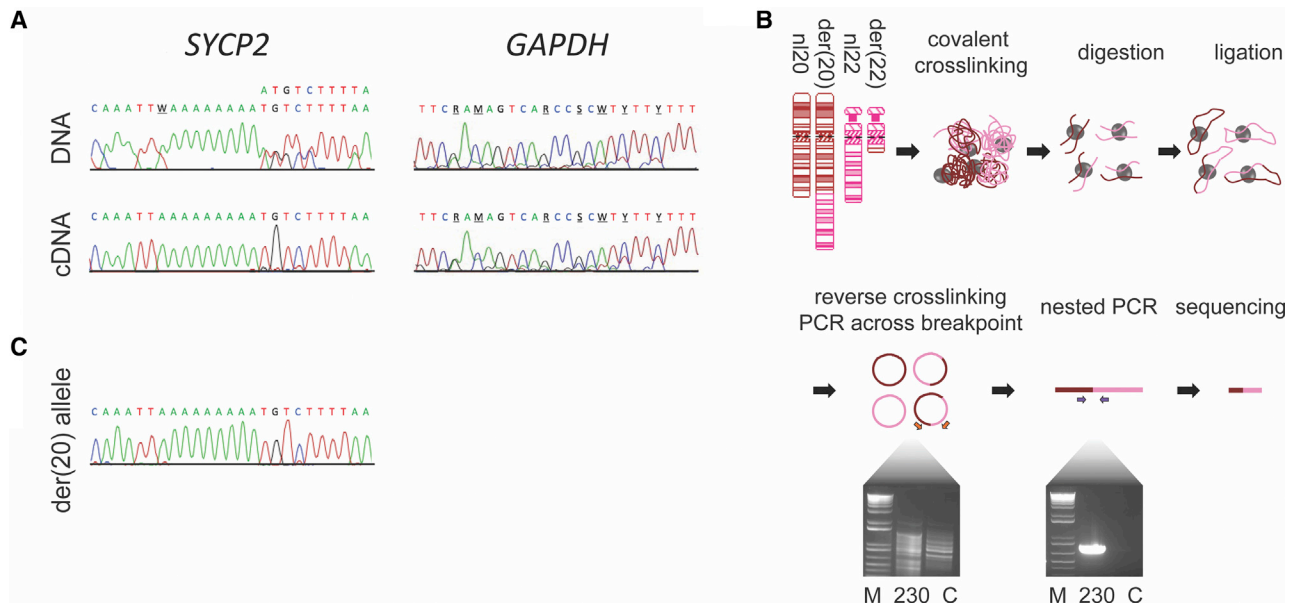
Figure 3. *SYCP2* Is Overexpressed Exclusively from the der(20) Allele
(A) Comparison of DNA and RNA from polymorphic regions in *SYCP2* and *GAPDH* exons. Sequence traces are shown for genomic DNA (top) and cDNA (bottom) from the DGAP230 LCL. The *SYCP2* DNA trace shows heterozygosity at SNPs rs568645874 and rs199662252 while the *SYCP2* cDNA trace only detects one haplotype. In contrast, the *GAPDH* control shows heterogeneity in both DNA and cDNA (SNPs rs45568532, rs551180067, and rs11549332). W, A/T; R, A/G; M, A/C; S, C/G; Y, C/T.
(B) Overview of 3C-PCR, a protocol that couples proximity ligation with breakpoint-spanning PCR to capture *cis* sequences distant from the chromosomal rearrangement.[21] Gel electrophoresis of PCR products from the first PCR across the breakpoint (bottom left) and the second nested PCR using amplification products from the first PCR as a substrate (bottom right) is shown. M, marker; C, control LCLs.
(C) Sanger sequencing traces for the der(20) allele, compared to the DNA and cDNA Sanger sequencing traces in Figure 3A (n = 3).

only one allele of *SYCP2* is expressed, we identified an exonic polymorphic region in *SYCP2* from DGAP230 genomic DNA and assessed expression of both allelic variants in DGAP230 LCL cDNA.[57] Sequence traces detected only one *SYCP2* haplotype in DGAP230 LCL cDNA, suggesting expression from a single allele in the DGAP230 LCL (Figure 3A).

To examine whether the expressed *SYCP2* allele resides in *cis* with the der(20) translocation breakpoint, we sought to amplify selectively the der(20) allele using primers that span the translocation junction for subsequent Sanger sequencing. To overcome the technical challenge presented by *SYCP2*'s position 1.5 Mb proximal to the translocation breakpoint, we applied 3C-PCR.[21] 3C-PCR capitalizes on principles underlying chromosome conformation capture (3C) to bring fragments containing the translocation junction and the der(20)-residing *SYCP2* allele closer together, thus enabling PCR across the junction of a ligation product including the *cis* allele (Figure 3B).[26,58] This resulted in production of amplicons in DGAP230 but not in karyotypically normal LCLs, demonstrating specificity for the t(20;22) (Figure 3B). Sanger sequencing of this amplicon revealed that it contains the expressed allele (Figure 3C), suggesting that *SYCP2* expression is detected exclusively from the der(20) allele in the DGAP230 LCL.

Given that misexpression of *SYCP2* derives from the der(20), we hypothesized that *SYCP2* dysregulation is mediated by an "enhancer adoption" mechanism, a long-

range *cis*-regulatory mutation that results in a gain of regulatory elements and subsequent promiscuous gene expression.[59,60] To explore this hypothesis, we performed Circular Chromatin Conformation Capture sequencing (4C-seq) using the *SYCP2* promoter as bait in DGAP230 and age- and sex-matched control LCLs and FourCSeq analysis for both normal chromosome 20 (nl20) and der(20) chromosomes.[28] A single statistically significant fragment was identified with 210-fold increased interaction in the DGAP230 LCL (n = 3; p_{adj} control = 0.0053), which maps to a genomic region on chr22, 8 Mb downstream of the *SYCP2* promoter (Figure 4A). A closer look at this region revealed many signatures of enhancer activity including a high ratio of H3K4me1 to H3K4me3, H3K27ac, and two regions that demonstrate DNaseI hypersensitivity in fetal testis tissue (Figure 4B).[61,62] Taken together, our findings support an enhancer adoption model, where an active enhancer residing in the segment of chr22 translocated to the der(20) may enter a newly formed chromatin contact encompassing the *cis*-residing *SYCP2* allele, resulting in illegitimate overexpression.

In order to implicate the genomic region identified by 4C-seq as the adopted enhancer for *SYCP2* dysregulation, we attempted to delete the putative enhancer region using CRISPR/Cas9 with the goal of assessing the subsequent impact on *SYCP2* expression.[30] Our experiments suggest that the putative enhancer region targeted by CRISPR/Cas9 is important for cell growth or viability (Figure S1), as we were unable to generate DGAP230 LCLs with a
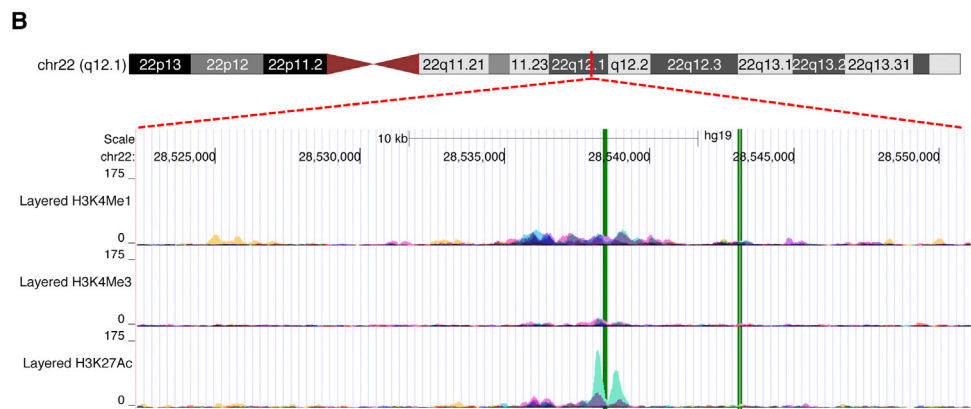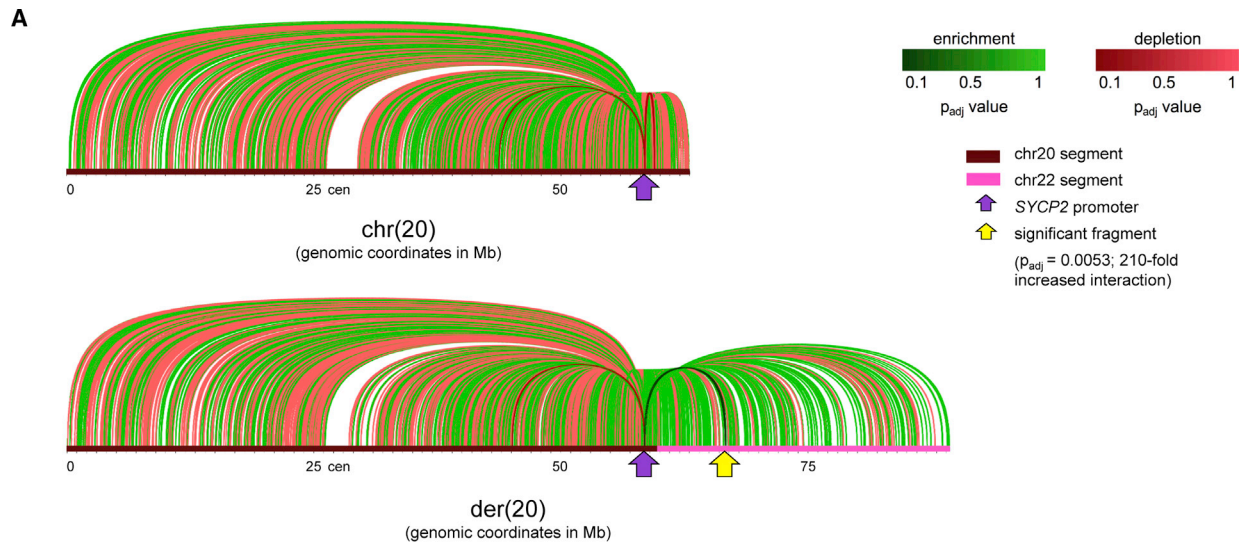
Figure 4. **Investigation of Chromatin Contacts with the *SYCP2* Promoter**
Circular Chromatin Conformation Capture sequencing (4C-seq) was performed on three biological replicates of the DGAP230 LCL and control LCLs from three different karyotypically normal age- and sex-matched control subjects.
(A) FourCSeq analysis for chr20 (top) and der(20) (bottom) reveals differential interactions of DNA fragments with the *SYCP2* promoter (purple arrow).[28] Plots were made in Circos and linearized using Adobe Photoshop. Lines represent each DNA fragment found to interact with the *SYCP2* promoter. A single statistically significant fragment (yellow arrow) with DGAP230 enrichment appears in the der(20) analysis (210-fold enrichment, p_{adj} control = 0.0053). cen = centromere.
(B) The significant DNA fragment and flanking enriched regions were consolidated to define a 29 kb putative interacting region. Tracks indicating enhancer activity from seven different cell lines are displayed in the UCSC genome browser and green bars highlight two regions of DNaseI hypersensitivity in fetal testis tissue according to ENCODE.[61,62,63]

deleted putative enhancer region in order to study further its role in *SYCP2* overexpression.

Analysis of the Impact of *SYCP2* Misexpression on Severe Oligozoospermia

We next wanted to understand how *SYCP2* dysregulation can explain DGAP230's phenotype of severe oligozoospermia. *SYCP2* encodes synaptonemal complex protein 2, a component of the lateral element substructure of the synaptonemal complex (SC).[64] SYCP2 nonspecifically interacts with the minor groove of DNA and serves as a scaffold for recruiting SYCP3 through its coiled-coil domain, thus facilitating formation of meiotic chromosome axes that are competent to assemble the SC.[65–67] SC assembly (synapsis) is a meiosis-specific process that plays a role in pairing, recombination, and segregation of homologous chromosomes during meiosis I.[67,68] SYCP2 is important for spermatogenesis, as male mice homozygous for coiled-coil domain-deficient *Sycp2* exhibit diminished homologous chromosome synapsis, apoptosis within the developing germline, and infertility.[65]

In response to the intriguing role that SYCP2 plays in meiosis and spermatogenesis, we pursued identification of SYCP2 overexpression at the protein level. We found by western blot that DGAP230 lymphoblastoid cells have greater than five times more SYCP2 than age- and sex-matched controls, which achieved statistical significance by an unpaired one-tailed t test (p < 0.0032) (Figure S2).

Based upon DGAP230's phenotype, his overexpression of *SYCP2* at the RNA and protein levels in LCLs, and current literature on *SYCP2*, we hypothesized that *SYCP2*
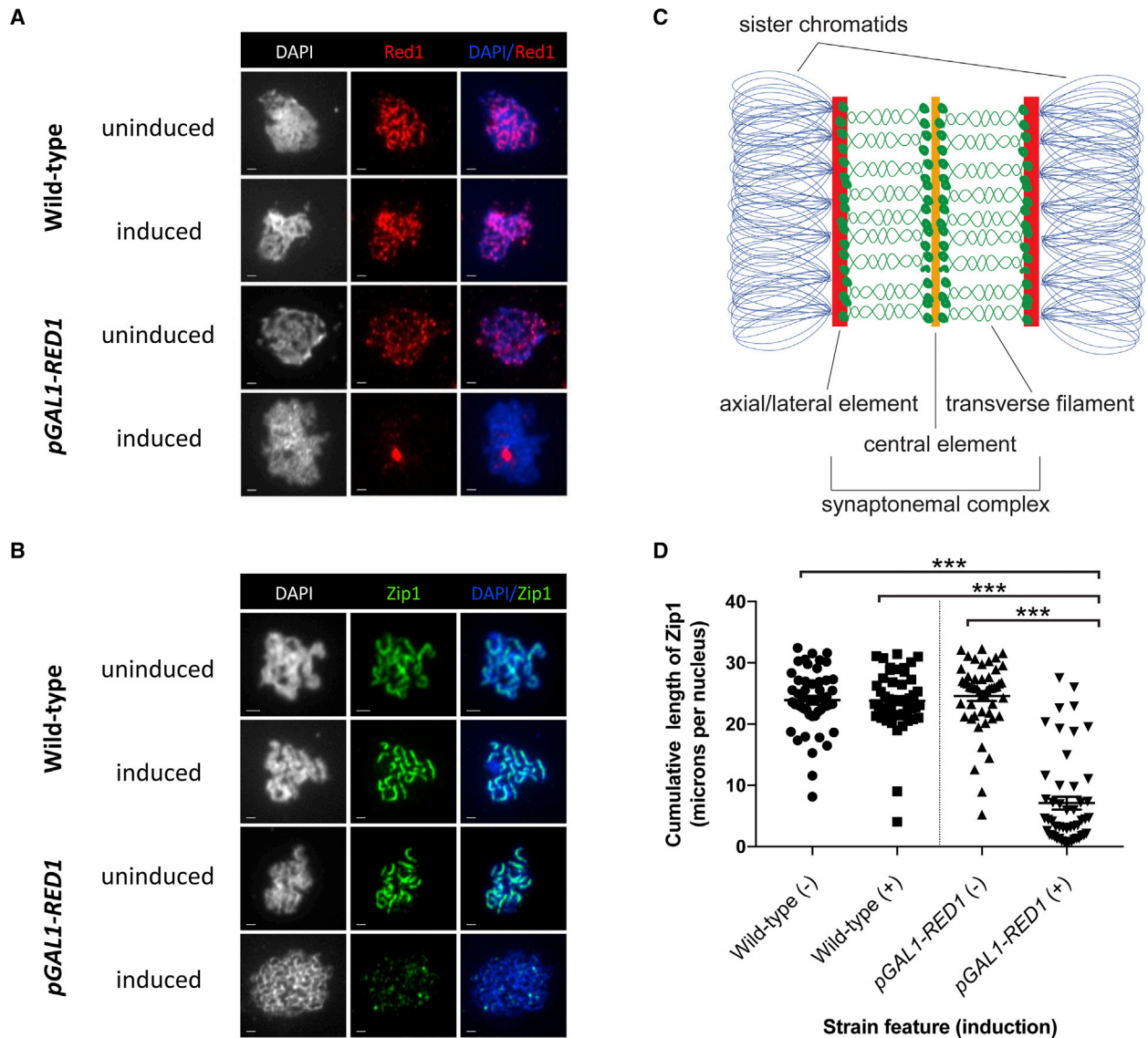
Figure 5. Consequences of Red1 Misexpression on Synaptonemal Complex Assembly in *S. cerevisiae*

(A) Anti-Red1 immuno- and DAPI-staining of surface-spread pachytene chromosomes reveal that constitutive *RED1* induction during sporulation leads to Red1 polycomplex formation (scale bar represents one micron in length).

(B) Anti-Zip1 immuno- and DAPI-staining of surface-spread pachytene chromosomes show that constitutive *RED1* induction during sporulation obliterates Zip1 formation at the interface of homologous chromosomes (scale bar represents one micron in length).

(C) The SC is evolutionary conserved between mammals and budding yeast. SYCP2 and Red1 are axial element functional homologs (red) and SYCP1 and Zip1 are transverse filament homologs (green) in the mammalian and *S. cerevisiae* SC, respectively. The central element substructure is colored yellow.

(D) Quantification of the cumulative length of synaptonemal complex Zip1 in surface-spread meiotic nuclei demonstrates a statistically significant decrease upon induction of *RED1* (Figures S3A and S3B), as analyzed by Mann-Whitney U test (n = 50). Error bars represent mean ± standard error. ***p < 0.0001.

misexpression may lead to defects in meiosis, resulting in problems with spermatogenesis, thus leading to the phenotype of severe oligozoospermia.

SYCP2 misexpression may inhibit proper meiosis in DGAP230 because of its role in facilitating homologous chromosome synapsis in meiosis I. It could influence meiosis in two opposing ways. First, it is possible that accumulation of SYCP2 leads to excess synaptonemal complex formation, perhaps preventing proper separation of homologs in anaphase I. Alternatively, excess SYCP2 may lead to an altered stoichiometry that might prevent proper binding of other axial element proteins such as synaptonemal complex protein 3 (SYCP3), causing poor integrity of the synaptonemal complex with resulting asynapsis.

Because meiosis is evolutionarily conserved, the unicellular eukaryote *Saccharomyces cerevisiae* serves as an excellent model organism to study the molecular mechanisms of meiosis.[69] We positioned the yeast functional homolog of *SYCP2*, *RED1*, under control of the inducible promoter

P_{GAL1} in a strain background that contains *GAL4-ER*, which induces constitutive strong expression from P_{GAL1} promoters in the presence of β-estradiol (Figures S3A and S3B). To learn how induction of excess Red1 impacts its function as an axial element protein, we used immunolocalization to label Red1 on surface-spread meiotic nuclei. While normal Red1 localizes along the lengths of pachytene chromosomes with stretches of staining and non-staining regions in meiotic prophase nuclei, induced P_{GAL1}-*RED1* meiocytes exhibit an aggregation of Red1 protein associated with meiotic chromosomes, which we interpret to be a polycomplex structure (Figure 5A). We hypothesized that mislocalization of Red1 may influence its ability to serve as a scaffold for the SC. To assess structural integrity of the SC, we immunostained surface-spread meiotic nuclei for Zip1, a transverse filament in the yeast SC that forms linear structures at the interface of each synapsed chromosome pair (Figures 5B and 5C). We discovered diminished Zip1 assembly in P_{GAL1}-*RED1* meiotic nuclei compared to control strains, a result that was found to be statistically significant after quantifying the cumulative length of Zip1 per nucleus in a population of 50 cells (Figure 5D). We confirmed that loss of Zip1 on meiotic chromosomes is not a result of decreased Zip1 expression (Figure S3C). We conclude that overexpression of meiotic chromosome axis/lateral element component Red1 can result in aberrant assembly of synaptonemal complex in budding yeast meiotic nuclei.

These results support the possibility that an abnormal abundance of SYCP2 axial element protein disrupts homologous chromosome synapsis in DGAP230, which would be expected to promote germline apoptosis.[70]

Identification of Additional Male Infertility Case Subjects with *SYCP2* Loss-of-Function Variants

Current evidence from our work and the literature suggests that proper SYCP2 dosage is important for its function.[65] In this project, we have found that misexpression of *SYCP2* leads to a loss of function by stoichiometric imbalance. On the other hand, mice lacking the SYCP2 coiled-coil domain also have an infertility phenotype.[65] Ultimately, it appears that either too much or too little SYCP2 results in a loss of function leading to infertility.

Under the assumption that genetic variants causing infertility would be eliminated from the general population by natural selection, we analyzed allele frequencies of *SYCP2* variants from more than 140,000 genomes in the gnomAD browser.[45] We found that *SYCP2* is severely depleted for LoF variants including stop-gained and essential splice site variants, as demonstrated by a *pLI* (probability of loss-of-function intolerance) of 1.000 and *oe* (observed/expected metric) of 0.105.[45] This extreme intolerance to loss of function could be explained by the inability to segregate such variants due to a phenotype of infertility, which would further support *SYCP2* pathogenicity in humans. While we have found

that *SYCP2* is significantly depleted for LoF variants, these aggregation datasets lack phenotypic information that would be necessary to assess whether LoF variants lead to infertility. To test the model that LoF *SYCP2* variants are more prevalent in infertile men than in fertile men, we searched for *SYCP2* (GenBank: NM_014258.3) variants in the exome sequencing (ES) data from the MERGE (Male Reproductive Genomics) study of the Institute of Human Genetics, University of Münster, comprising a male infertility cohort from the Centre of Reproductive Medicine and Andrology (CeRA). This search revealed three males with heterozygous LoF variants. The frequency of 0.48% found in the Münster male infertility cohort is significantly higher than the overall prevalence of *SYCP2* LoF variants in the population according to gnomAD (0.05% for all *SYCP2* LoF variants [chi-square p = 0.0002] and 0.01% for high confidence *SYCP2* LoF variants [chi-square p < 0.0001]).

Participant M1686 presented with a 1.5-year history of infertility at age 29 and was diagnosed with cryptozoospermia (sperm concentration < 0.1 M/mL) according to the current World Health Organization (WHO) reference ranges for semen parameters.[37] He had normal testicular volume (right: 17 mL, left: 16 mL [reference range > 12 mL]), borderline elevated serum FSH levels (7.1 IU/L [reference range 1–7 IU/L]), and normal LH (9.5 IU/L [reference range 2–10 IU/L]) and testosterone (39.7 nmol/L [reference range > 12 nmol/L]). ES and subsequent validation by Sanger sequencing revealed a heterozygous deletion in exon 24 of *SYCP2*: c.2022_2025del causing p.Lys674AsnfsTer8 (Figure 6A, ClinVar: SCV000920592). This variant is absent in both gnomAD and Trans-Omics for Precision Medicine (TOPMed) databases.[45]

Participant M1581 presented with a 2.5-year history of infertility at age 27. He was diagnosed with cryptozoospermia (sperm concentration < 0.1 M/mL), had normal testicular volume (right: 14 ml, left: 16 ml), and normal serum levels of FSH (3.3 IU/l), LH (3.0 IU/l) and testosterone (14.5 nmol/l). ES and subsequent validation by Sanger sequencing revealed a heterozygous deletion in exon 31 of *SYCP2*: c.2793_2797del causing p.Lys932SerfsTer3 (Figure 6B, ClinVar: SCV000920594). This variant is absent in both gnomAD and TOPMed databases.[45] Segregation was assessed in all available family members, which revealed that the deletion was inherited from the mother while the father is wild type at this position (Figure 6B).

Participant M1401 presented with a 17-year history of infertility at age 39. He was diagnosed with azoospermia, and histopathological analysis from testicular sperm extraction (TESE) revealed a phenotype of meiotic arrest at the pachytene spermatocyte stage (Figure 6C). Participant M1401 had borderline testicular atrophy (right: 12 mL, left: 11 mL), elevated serum levels of FSH (44.9 IU/L) and LH (12.4 IU/L), and low serum levels of testosterone (8.2 nmol/L). ES and subsequent validation by Sanger sequencing revealed a heterozygous deletion in exon 33 of *SYCP2*: c.3067_3071del causing
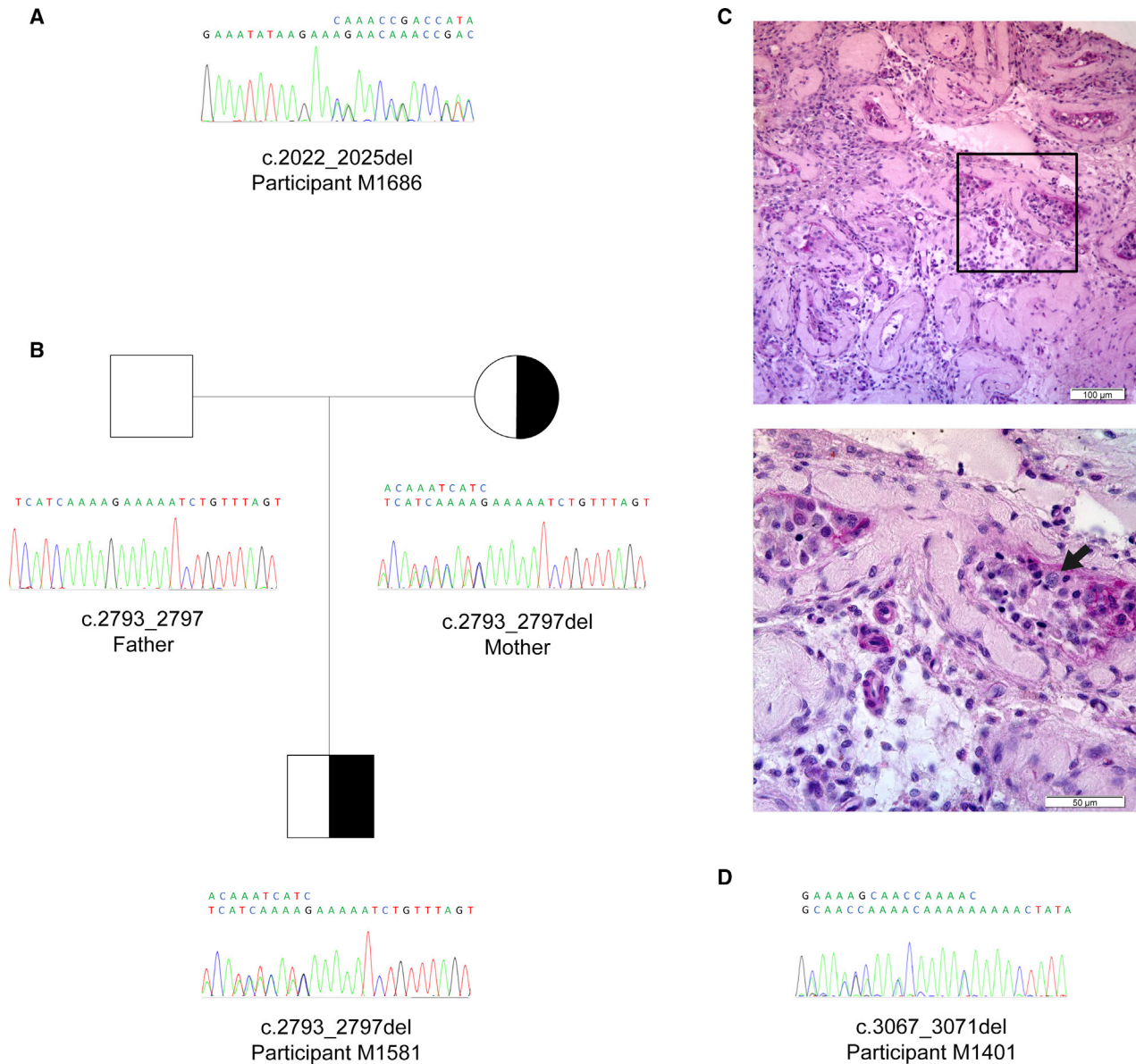
Figure 6. Identification of Heterozygous Frameshift Mutations in *SYCP2* from Münster Male Infertility Cohort Participants

(A) Participant M1686 with cryptozoospermia has a heterozygous deletion in *SYCP2* (c.2022_2025del [p.Lys674AsnfsTer8]).

(B) Participant M1581 with cryptozoospermia has a heterozygous deletion in *SYCP2* (c.2793_2797del [p.Lys932SerfsTer3]) inherited from his mother.

(C) Histological PAS staining of a testis biopsy from participant M1401 shows a phenotype of meiotic arrest at the pachytene spermatocyte stage leading to nonobstructive azoospermia. The majority of the tubules in an overview section (top) are degenerated to the phenotype of tubular ghosts with a few tubular cross sections that present with spermatocytes. A closer look at the boxed section (magnified in the bottom image) shows disorganized seminiferous epithelium with a single pachytene spermatocyte (black arrow). Scale bar is as indicated.

(D) Participant M1401 has a heterozygous deletion in *SYCP2* (c.3067_3071del [p.Lys1023LeufsTer2]).

p.Lys1023LeufsTer2 (Figure 6D, ClinVar: SCV000920593). This variant is absent in both gnomAD and TOPMed databases.[45]

To evaluate the significance of our findings in the context of gene-disease association, we performed two independent clinical validity assessments by two researchers using different methods, blinded to each other's results.[71,72] The approaches classify *SYCP2* as a gene with limited to moderate evidence for a gene-dis-

ease relationship with human male infertility (Tables S3 and S4).

Discussion

Male infertility is a common disorder among reproductive-aged couples and the majority of subjects lack a specific etiologic diagnosis.[5] Understanding the precise causes of

male infertility may directly inform therapies for infertile couples.[73] For example, an azoospermic male with a complete deletion of AZFc has a 50% success rate for obtaining sperm by TESE, while TESE would not be recommended or successful for a man with a complete deletion of AZFa.[7]

Identifying genetic etiologies for human male infertility has been hindered by smaller pedigrees inherent to decreased reproductive fitness and genetic heterogeneity of the disorder. In addition, some genetic evidence of a disorder may not be investigated rigorously. For example, balanced reciprocal translocations identified by karyotype in infertile men are rarely followed up beyond reporting a risk for segregation of unbalanced gametes. As a result, a deep investigation into single-case studies is critical for uncovering additional genetic etiologies for male infertility.

In this study, we identified a balanced reciprocal translocation in a severe oligozoospermic male designated DGAP230. While it is generally thought that balanced reciprocal translocations may reduce fertility due to production of unbalanced gametes[13] or meiotic silencing of unsynapsed chromatin,[74] this does not account for the specific phenotype of severe oligozoospermia or azoospermia because the majority of men with balanced reciprocal translocations have normal sperm counts.[75] In addition, men with low sperm counts and a balanced reciprocal translocation have rearrangement breakpoints that sometimes cluster in distinct genomic regions, suggesting that as opposed to a nonspecific mechanism of meiotic segregation, there may be something intrinsic to these genomic regions important for fertility.[76,77]

In the case of DGAP230, the structural rearrangement leads to dysregulation of SYCP2, which resides 1.5 Mb proximal to one of the rearrangement breakpoints. This cytogenomic influence on gene expression supports the finding that translocation breakpoints can influence gene expression by dysregulating genes residing within the same TAD.[60] As a result, in men with balanced chromosomal rearrangements and a phenotype of infertility, pathogenesis by specific breakpoints should be considered as an alternative etiology to that of segregation of unbalanced gametes.

While it is known that many different SC proteins have intrinsic ability to self-assemble into polycomplexes when overexpressed, mutated, or expressed in mitotic cells,[78–82] this study demonstrates the first observation of Red1 overexpression leading to its polycomplex formation. We predict that aggregation may be mediated by misexpression before formation of meiotic chromosomes and expression of other axial element proteins as well as the presence of a coiled-coil domain, which facilitates protein complex interactions. Red1 overexpression in budding yeast also decreases double-strand break formation and meiotic recombination initiation, a requirement for SC assembly in budding yeast and most mammals.[83] The resulting asynapsis phenocopies red1 mutants in S. cerevisiae as well as coiled-coil domain-deficient Sycp2 mice, suggesting that synaptonemal complex formation is sensitive to dosage of axial elements.[65,84]

We believe that our finding of asynapsis resulting from axial element misexpression is directly related to severe oligozoospermia, because asynapsis triggers checkpoint-mediated apoptosis of spermatocytes during spermatogenesis,[70,85,86] which reduces sperm count and has been shown to lead to male-specific infertility.[65] Therefore, DGAP230's phenotype of severe oligozoospermia and infertility is likely due to asynapsis-triggered cell death in spermatocytes.

The identification of three frameshift variants in SYCP2 from men with cryptozoospermia and meiotic arrest further supports the role of SYCP2 in human male fertility. All variants are extremely rare, consistent with the inability to segregate these mutations in the general population due to a phenotype of infertility. This is also supported by the maternal inheritance of the variant in participant M1581, as SYCP2-mediated pathogenicity has been shown to cause male infertility but not female infertility in a mouse model.[65]

Despite the predicted pathogenicity of all three variants, it is important to note that all cases represent heterozygous variants which would support an autosomal-dominant disease model. This is discordant with the coiled-coil domain-deficient Sycp2 mouse model that only demonstrates a male infertility phenotype in homozygotes.[65] However, there are many explanations for this discrepancy. First, haploinsufficiency discordance has been observed between human and mouse for the male infertility gene SYCP3, the other mammalian axial element in the synaptonemal complex as well as numerous other mouse models.[87,88] It is also possible that the discordance merely reflects different effects of the truncated SYCP2 protein modeled in the mouse compared to the predicted nonsense-mediated decay of SYCP2 transcripts from Münster infertility cohort subjects, as all of the premature termination codons reside more than 50 bp upstream of the last exon-exon junction.[89] There is also support for haploinsufficiency of SYCP2 in humans. SYCP2's $pLI > 0.9$ and $oe < 0.35$ support a haploinsufficient model, as extreme intolerance to LoF according to constraint analysis is strongly correlated with haploinsufficiency.[45] In addition, SYCP2 variants have an autosomal-dominant estimated mode of inheritance (p_AD 0.77, p_AR 0.23).[90] Moreover, while less common than autosomal-recessive forms of male infertility, autosomal-dominant forms have been identified for pathogenic variants in PIWIL1,[91] KLHL10,[92] PLK4,[93] SYCP3,[87] SPINK2,[94] NR5A1,[38] and DMRT1.[44]

Another potential concern is that the phenotypes of severe oligozoospermia, cryptozoospermia, and meiotic arrest are distinctive from each other. It is possible that the differences reflect variable expressivity, as has been observed in the male infertility genes DDX3Y (Sertoli cell-only syndrome and severe hypospermatogenesis), KLHL10 (severe oligozoospermia and oligozoospermia),

TAF4B (azoospermia and oligozoospermia), *TDRD9* (azoospermia and cryptozoospermia), and *TEX11* (complete meiotic arrest and mixed testicular atrophy).[43,92,95–98]

It is well known that homologous chromosome synapsis is critical for spermatogenesis. Indeed, several genes implicated in human male infertility encode proteins that are members of the synaptonemal complex (*SYCP3* and *SYCE1*) or are otherwise required for synapsis (*SPO11*, *MEIOB*, *TEX11*, and *TEX15*).[6,43,87,98–101] Before this study, *SYCP2* was considered a strong candidate gene for human male infertility because it encodes a protein that interacts directly with protein products of the human male infertility genes *SYCP3* and *TEX11*, serves as an axial element in the synaptonemal complex, and is required for male fertility in the mouse.[65,66,102] DGAP230 and the three participants from the Münster cohort represent the first case subjects of putative *SYCP2*-mediated male infertility in humans.

Accession Numbers

DGAP230's reported breakpoints and clinical interpretations have been submitted to dbVar: nsv2781820 and ClinVar: SCV000320840.1. Frameshift variants in *SYCP2* identified in participants M1401, M1581, and M1686 and the clinical interpretation have been submitted to ClinVar: SCV000920593, SCV000920594, SCV000920592.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2019.11.013.

Acknowledgments

Declaration of Interests

The authors declare no competing interests.

Web Resources

BLA(S)T Output Sequence Tool of Nomenclature (BOSToN), http://BOSToN.bwh.harvard.edu/

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/

Custom script used to create the derivative chromosomes 20 and 22, https://github.com/mnnshreya/Morton-DGAP230

Cytogenetic Data Analysis System (CyDAS), http://www.cydas.org/OnlineAnalysis/

dbSNP, https://www.ncbi.nlm.nih.gov/projects/SNP/

dbVar, https://www.ncbi.nlm.nih.gov/dbvar/

DECIPHER, https://decipher.sanger.ac.uk/

Galaxy, https://usegalaxy.org/

GenBank, https://www.ncbi.nlm.nih.gov/genbank/

Geneious (v.0, Biomatters), https://www.geneious.com/

gnomAD Browser, https://gnomad.broadinstitute.org/

Haploinsufficiency Index (%HI), https://decipher.sanger.ac.uk/files/downloads/HI_Predictions_Version3.bed.gz

Hi-C project, http://chromosome.sdsc.edu/mouse/hi-c/download.html

OMIM, https://www.omim.org/

Trans-Omics for Precision Medicine (TOPMed), https://bravo.sph.umich.edu

UCSC Genome Browser, https://genome.ucsc.edu

VassarStats Concepts and Applications of Inferential Statistics, http://vassarstats.net/utest.html

References

1. Kliesch, S. (2014). Diagnosis of male infertility: diagnostic work-up of the infertile man. Eur. Urol. Suppl. *13*, 73–82.

2. Shah, K., Sivapalan, G., Gibbons, N., Tempest, H., and Griffin, D.K. (2003). The genetic basis of infertility. Reproduction *126*, 13–25.

3. Matzuk, M.M., and Lamb, D.J. (2002). Genetic dissection of mammalian fertility pathways. Nat. Cell Biol. *4* (Suppl), s41–s49.

4. Oud, M.S., Volozonoka, L., Smits, R.M., Vissers, L.E.L.M., Ramos, L., and Veltman, J.A. (2019). A systematic review and standardized clinical validity assessment of male infertility genes. Hum. Reprod. *34*, 932–941.

5. Tüttelmann, F., Ruckert, C., and Röpke, A. (2018). Disorders of spermatogenesis: Perspectives for novel genetic diagnostics after 20 years of unchanged routine. Med. Genetik *30*, 12–20.

6. Fakhro, K.A., Elbardisi, H., Arafa, M., Robay, A., Rodriguez-Flores, J.L., Al-Shakaki, A., Syed, N., Mezey, J.G., Abi Khalil, C., Malek, J.A., et al. (2018). Point-of-care whole-exome sequencing of idiopathic male infertility. Genet. Med. *20*, 1365–1373.

7. Okutman, O., Rhouma, M.B., Benkhalifa, M., Muller, J., and Viville, S. (2018). Genetic evaluation of patients with non-syndromic male infertility. J. Assist. Reprod. Genet. *35*, 1939–1951.

8. Kumar, N., and Singh, A.K. (2015). Trends of male factor infertility, an important cause of infertility: A review of literature. J. Hum. Reprod. Sci. *8*, 191–196.

9. Massart, A., Lissens, W., Tournaye, H., and Stouffs, K. (2012). Genetic causes of spermatogenic failure. Asian J. Androl. *14*, 40–48.

10. Rives, N., Joly, G., Machy, A., Siméon, N., Leclerc, P., and Macé, B. (2000). Assessment of sex chromosome aneuploidy in sperm nuclei from 47,XXY and 46,XY/47,XXY males: comparison with fertile and infertile males with normal karyotype. Mol. Hum. Reprod. *6*, 107–112.

11. Ferlin, A., Raicu, F., Gatta, V., Zuccarello, D., Palka, G., and Foresta, C. (2007). Male infertility: role of genetic background. Reprod. Biomed. Online *14*, 734–745.

12. Donker, R.B., Vloeberghs, V., Groen, H., Tournaye, H., van Ravenswaaij-Arts, C.M.A., and Land, J.A. (2017). Chromosomal abnormalities in 1663 infertile men with azoospermia: the clinical consequences. Hum. Reprod. *32*, 2574–2580.

13. Kavalier, F. (2005). Investigation of recurrent miscarriages. BMJ *331*, 121–122.

14. Redin, C., Brand, H., Collins, R.L., Kammin, T., Mitchell, E., Hodge, J.C., Hanscom, C., Pillalamarri, V., Seabra, C.M., Abbott, M.A., et al. (2017). The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. Nat. Genet. *49*, 36–45.

15. Talkowski, M.E., Ernst, C., Heilbut, A., Chiang, C., Hanscom, C., Lindgren, A., Kirby, A., Liu, S., Muddukrishna, B., Ohsumi, T.K., et al. (2011). Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. Am. J. Hum. Genet. *88*, 469–481.

16. Hanscom, C., and Talkowski, M. (2014). Design of large-insert jumping libraries for structural variant detection using Illumina sequencing. Curr. Protoc. Hum. Genet. *80*, 1–9.

17. Ordulu, Z., Wong, K.E., Currall, B.B., Ivanov, A.R., Pereira, S., Althari, S., Gusella, J.F., Talkowski, M.E., and Morton, C.C. (2014). Describing sequencing results of structural chromosome rearrangements with a suggested next-generation cytogenetic nomenclature. Am. J. Hum. Genet. *94*, 695–709.

18. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

19. Rosenbloom, K.R., Armstrong, J., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2015). The UCSC Genome Browser database: 2015 update. Nucleic Acids Res. *43*, D670–D681.

20. Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. *29*, 308–311.

21. Schilit, S.L.P., and Morton, C.C. (2018). 3C-PCR: a novel proximity ligation-based approach to phase chromosomal rearrangement breakpoints with distal allelic variants. Hum. Genet. *137*, 55–62.

22. Miele, A., Gheldof, N., Tabuchi, T.M., Dostie, J., and Dekker, J. (2006). Mapping chromatin interactions by chromosome conformation capture. Curr Protoc Mol Biol *Chapter 21*, 21.11.01.

23. van de Werken, H.J., de Vree, P.J., Splinter, E., Holwerda, S.J., Klous, P., de Wit, E., and de Laat, W. (2012). 4C technology: protocols and data analysis. Methods Enzymol. *513*, 89–112.

24. Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods *58*, 221–230.

25. Gheldof, N., Leleu, M., Noordermeer, D., Rougemont, J., and Reymond, A. (2012). Detecting long-range chromatin interactions using the chromosome conformation capture sequencing (4C-seq) method. Methods Mol. Biol. *786*, 211–225.

26. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306–1311.

27. Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. *46* (W1), W537–W544.

28. Klein, F.A., Pakozdi, T., Anders, S., Ghavi-Helm, Y., Furlong, E.E., and Huber, W. (2015). FourCSeq: analysis of 4C sequencing data. Bioinformatics *31*, 3085–3091.

29. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645.

30. Jiang, S., Wang, L.W., Walsh, M.J., Trudeau, S.J., Gerdt, C., Zhao, B., and Gewurz, B.E. (2018). CRISPR/Cas9-mediated genome editing in Epstein-Barr virus-transformed lymphoblastoid B-cell lines. Curr. Protoc. Mol. Biol. *121*, 1–23.

31. Rockmill, B., and Roeder, G.S. (1998). Telomere-mediated chromosome pairing during meiosis in budding yeast. Genes Dev. *12*, 2574–2586.

32. Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953–961.

33. MacQueen, A.J., and Rockmill, B. (2017). Analysis of meiotic chromosome-associated protein dynamics using conditional expression in budding yeast. Methods Mol. Biol. *1471*, 157–174.

34. Voelkel-Meiman, K., Moustafa, S.S., Lefrançois, P., Villeneuve, A.M., and MacQueen, A.J. (2012). Full-length synaptonemal complex grows continuously during meiotic prophase in budding yeast. PLoS Genet. *8*, e1002993.

35. Smith, A.V., and Roeder, G.S. (1997). The yeast Red1 protein localizes to the cores of meiotic chromosomes. J. Cell Biol. *136*, 957–967.

36. Sym, M., Engebrecht, J.A., and Roeder, G.S. (1993). ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. Cell *72*, 365–378.

37. World Health Organization (2010). WHO laboratory manual for the examination and processing of human semen (Geneva: World Health Organization).

38. Röpke, A., Tewes, A.C., Gromoll, J., Kliesch, S., Wieacker, P., and Tüttelmann, F. (2013). Comprehensive sequence analysis of the *NR5A1* gene encoding steroidogenic factor 1 in a large group of infertile males. Eur. J. Hum. Genet. *21*, 1012–1015.

39. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17*, 10–12.

40. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

41. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

42. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl variant effect predictor. Genome Biol. *17*, 122.

43. Yatsenko, A.N., Georgiadis, A.P., Röpke, A., Berman, A.J., Jaffe, T., Olszewska, M., Westernströer, B., Sanfilippo, J., Kurpisz, M., Rajkovic, A., et al. (2015). X-linked *TEX11* mutations, meiotic arrest, and azoospermia in infertile men. N. Engl. J. Med. *372*, 2097–2107.

44. Tewes, A.C., Ledig, S., Tüttelmann, F., Kliesch, S., and Wieacker, P. (2014). *DMRT1* mutations are rarely associated with male infertility. Fertil. Steril. *102*, 816–820.e3.

45. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.

46. Brinkworth, M.H., Weinbauer, G.F., Schlatt, S., and Nieschlag, E. (1995). Identification of male germ cells undergoing apoptosis in adult rats. J. Reprod. Fertil. *105*, 25–33.

47. Gudeloglu, A., and Parekattil, S.J. (2013). Update in the evaluation of the azoospermic male. Clinics (São Paulo) *68* (*Suppl 1*), 27–34.

48. Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012–1025.

49. Ordulu, Z., Kammin, T., Brand, H., Pillalamarri, V., Redin, C.E., Collins, R.L., Blumenthal, I., Hanscom, C., Pereira, S., Bradley, I., et al. (2016). Structural chromosomal rearrangements require nucleotide-level resolution: lessons from next-generation sequencing in prenatal diagnosis. Am. J. Hum. Genet. *99*, 1015–1033.

50. Andrews, G.L., and Mastick, G.S. (2003). R-cadherin is a Pax6-regulated, growth-promoting cue for pioneer axons. J. Neurosci. *23*, 9873–9880.

51. Babb, S.G., Kotradi, S.M., Shah, B., Chiappini-Williamson, C., Bell, L.N., Schmeiser, G., Chen, E., Liu, Q., and Marrs, J.A. (2005). Zebrafish R-cadherin (Cdh4) controls visual system development and differentiation. Dev. Dyn. *233*, 930–945.

52. GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585.

53. Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. *6*, e1001154.

54. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. Am. J. Hum. Genet. *84*, 524–533.

55. Kim, H.G., Herrick, S.R., Lemyre, E., Kishikawa, S., Salisz, J.A., Seminara, S., MacDonald, M.E., Bruns, G.A., Morton, C.C., Quade, B.J., and Gusella, J.F. (2005). Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for *FGFR1*. J. Med. Genet. *42*, 666–672.

56. Lu, W., van Eerde, A.M., Fan, X., Quintero-Rivera, F., Kulkarni, S., Ferguson, H., Kim, H.G., Fan, Y., Xi, Q., Li, Q.G., et al. (2007). Disruption of *ROBO2* is associated with urinary tract anomalies and confers risk of vesicoureteral reflux. Am. J. Hum. Genet. *80*, 616–632.

57. Kim, H.G., Kishikawa, S., Higgins, A.W., Seong, I.S., Donovan, D.J., Shen, Y., Lally, E., Weiss, L.A., Najm, J., Kutsche, K., et al. (2008). Disruption of neurexin 1 associated with autism spectrum disorder. Am. J. Hum. Genet. *82*, 199–207.

58. Selvaraj, S., R Dixon, J., Bansal, V., and Ren, B. (2013). Whole-genome haplotype reconstruction using proximity-ligation and shotgun sequencing. Nat. Biotechnol. *31*, 1111–1118.

59. Lettice, L.A., Daniels, S., Sweeney, E., Venkataraman, S., Devenney, P.S., Gautier, P., Morrison, H., Fantes, J., Hill, R.E., and FitzPatrick, D.R. (2011). Enhancer-adoption as a mechanism of human developmental disease. Hum. Mutat. *32*, 1492–1499.

60. Lupiáñez, D.G., Spielmann, M., and Mundlos, S. (2016). Breaking TADs: how alterations of chromatin domains result in disease. Trends Genet. *32*, 225–237.

61. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

62. Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G., et al. (2013). ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res. *41*, D56–D63.

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

64. Schalk, J.A., Offenberg, H.H., Peters, E., Groot, N.P., Hoovers, J.M., and Heyting, C. (1999). Isolation and characterization of the human SCP2 cDNA and chromosomal localization of the gene. Mamm. Genome *10*, 642–644.

65. Yang, F., De La Fuente, R., Leu, N.A., Baumann, C., McLaughlin, K.J., and Wang, P.J. (2006). Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis. J. Cell Biol. *173*, 497–507.

66. Offenberg, H.H., Schalk, J.A., Meuwissen, R.L., van Aalderen, M., Kester, H.A., Dietrich, A.J., and Heyting, C. (1998). SCP2: a major protein component of the axial elements of synaptonemal complexes of the rat. Nucleic Acids Res. *26*, 2572–2579.

67. Winkel, K., Alsheimer, M., Ollinger, R., and Benavente, R. (2009). Protein SYCP2 provides a link between transverse filaments and lateral elements of mammalian synaptonemal complexes. Chromosoma *118*, 259–267.

68. Kolas, N.K., Yuan, L., Hoog, C., Heng, H.H., Marcon, E., and Moens, P.B. (2004). Male mouse meiotic chromosome cores deficient in structural proteins SYCP3 and SYCP2 align by homology but fail to synapse and have possible impaired specificity of chromatin loop attachment. Cytogenet. Genome Res. *105*, 182–188.

69. Kerr, G.W., Sarkar, S., and Arumugam, P. (2012). How to halve ploidy: lessons from budding yeast meiosis. Cell. Mol. Life Sci. *69*, 3037–3051.

70. Burgoyne, P.S., Mahadevaiah, S.K., and Turner, J.M. (2009). The consequences of asynapsis for mammalian meiosis. Nat. Rev. Genet. *10*, 207–216.

71. Strande, N.T., Riggs, E.R., Buchanan, A.H., Ceyhan-Birsoy, O., DiStefano, M., Dwight, S.S., Goldstein, J., Ghosh, R., Seifert, B.A., Sneddon, T.P., et al. (2017). Evaluating the clinical validity of gene-disease associations: an evidence-based framework developed by the Clinical Genome Resource. Am. J. Hum. Genet. *100*, 895–906.

72. Smith, E.D., Radtke, K., Rossi, M., Shinde, D.N., Darabi, S., El-Khechen, D., Powis, Z., Helbig, K., Waller, K., Grange, D.K., et al. (2017). Classification of genes: standardized clinical validity assessment of gene-disease associations aids diagnostic exome analysis and reclassifications. Hum. Mutat. *38*, 600–608.

73. Schilit, S.L.P. (2019). Recent advances and future opportunities to diagnose male infertility. Curr. Sex. Health Rep. Published online October 26, 2019. https://doi.org/10.1007/s11930-019-00225-8.

74. Mitchell, M.J., Metzler-Guillemain, C., Toure, A., Coutton, C., Arnoult, C., and Ray, P.F. (2017). Single gene defects leading to sperm quantitative anomalies. Clin. Genet. *91*, 208–216.

75. Vozdova, M., Oracova, E., Kasikova, K., Prinosilova, P., Rybar, R., Horinova, V., Gaillyova, R., and Rubes, J. (2013). Balanced chromosomal translocations in men: relationships among semen parameters, chromatin integrity, sperm meiotic segregation and aneuploidy. J. Assist. Reprod. Genet. *30*, 391–405.

76. Zhang, H.G., Wang, R.X., Li, L.L., Sun, W.T., Zhang, H.Y., and Liu, R.Z. (2015). Male carriers of balanced reciprocal translocations in Northeast China: sperm count, reproductive performance, and genetic counseling. Genet. Mol. Res. *14*, 18792–18798.

77. Vogt, P.H., Bender, U., Zimmer, J., and Strowitzki, T. (2017). Human Y chromosome and male infertility: forward and back from azoospermia factor chromatin structure to azoospermia factor gene function. Monogr. Hum. Genet. *21*, 57–73.

78. Zickler, D., and Kleckner, N. (1999). Meiotic chromosomes: integrating structure and function. Annu. Rev. Genet. *33*, 603–754.

79. Sym, M., and Roeder, G.S. (1995). Zip1-induced changes in synaptonemal complex structure and polycomplex assembly. J. Cell Biol. *128*, 455–466.

80. Ollinger, R., Alsheimer, M., and Benavente, R. (2005). Mammalian protein SCP1 forms synaptonemal complex-like structures in the absence of meiotic chromosomes. Mol. Biol. Cell *16*, 212–217.

81. Jeffress, J.K., Page, S.L., Royer, S.K., Belden, E.D., Blumenstiel, J.P., Anderson, L.K., and Hawley, R.S. (2007). The formation of the central element of the synaptonemal complex may occur by multiple mechanisms: the roles of the N- and C-terminal domains of the *Drosophila* C(3)G protein in mediating synapsis and recombination. Genetics *177*, 2445–2456.

82. Yuan, L., Pelttari, J., Brundell, E., Björkroth, B., Zhao, J., Liu, J.G., Brismar, H., Daneholt, B., and Höög, C. (1998). The synaptonemal complex protein SCP3 can form multistranded, cross-striated fibers *in vivo*. J. Cell Biol. *142*, 331–339.

83. Bailis, J.M., Smith, A.V., and Roeder, G.S. (2000). Bypass of a meiotic checkpoint by overproduction of meiotic chromosomal proteins. Mol. Cell. Biol. *20*, 4838–4848.

84. Rockmill, B., and Roeder, G.S. (1990). Meiosis in asynaptic yeast. Genetics *126*, 563–574.

85. Champion, M.D., and Hawley, R.S. (2002). Playing for half the deck: the molecular biology of meiosis. Nat. Cell Biol. *4* (Suppl), s50–s56.

86. Morelli, M.A., and Cohen, P.E. (2005). Not all germ cells are created equal: aspects of sexual dimorphism in mammalian meiosis. Reproduction *130*, 761–781.

87. Miyamoto, T., Hasuike, S., Yogev, L., Maduro, M.R., Ishikawa, M., Westphal, H., and Lamb, D.J. (2003). Azoospermia in patients heterozygous for a mutation in *SYCP3*. Lancet *362*, 1714–1719.

88. Yuan, L., Liu, J.G., Zhao, J., Brundell, E., Daneholt, B., and Höög, C. (2000). The murine *SCP3* gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. Mol. Cell *5*, 73–83.

89. Popp, M.W., and Maquat, L.E. (2013). Organizing principles of mammalian nonsense-mediated mRNA decay. Annu. Rev. Genet. *47*, 139–165.

90. Cassa, C.A., Weghorn, D., Balick, D.J., Jordan, D.M., Nusinow, D., Samocha, K.E., O'Donnell-Luria, A., MacArthur, D.G., Daly, M.J., Beier, D.R., and Sunyaev, S.R. (2017). Estimating the selective effects of heterozygous protein-truncating variants from human exome data. Nat. Genet. *49*, 806–810.

91. Gou, L.T., Kang, J.Y., Dai, P., Wang, X., Li, F., Zhao, S., Zhang, M., Hua, M.M., Lu, Y., Zhu, Y., et al. (2017). Ubiquitination-deficient mutations in human *Piwi* cause male infertility by impairing histone-to-protamine exchange during spermiogenesis. Cell *169*, 1090–1104.e13.

92. Yatsenko, A.N., Roy, A., Chen, R., Ma, L., Murthy, L.J., Yan, W., Lamb, D.J., and Matzuk, M.M. (2006). Non-invasive genetic diagnosis of male infertility using spermatozoal RNA: *KLHL10* mutations in oligozoospermic patients impair homodimerization. Hum. Mol. Genet. *15*, 3411–3419.

93. Miyamoto, T., Bando, Y., Koh, E., Tsujimura, A., Miyagawa, Y., Iijima, M., Namiki, M., Shiina, M., Ogata, K., Matsumoto, N., and Sengoku, K. (2016). A *PLK4* mutation causing azoospermia in a man with Sertoli cell-only syndrome. Andrology *4*, 75–81.

94. Kherraf, Z.E., Christou-Kent, M., Karaouzene, T., Amiri-Yekta, A., Martinez, G., Vargas, A.S., Lambert, E., Borel, C., Dorphin, B., Aknin-Seifer, I., et al. (2017). *SPINK2* deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes. EMBO Mol. Med. *9*, 1132–1149.

95. Foresta, C., Ferlin, A., and Moro, E. (2000). Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for *DBY* in male infertility. Hum. Mol. Genet. *9*, 1161–1169.

96. Ayhan, Ö., Balkan, M., Guven, A., Hazan, R., Atar, M., Tok, A., and Tolun, A. (2014). Truncating mutations in *TAF4B* and *ZMYND15* causing recessive azoospermia. J. Med. Genet. *51*, 239–244.

97. Arafat, M., Har-Vardi, I., Harlev, A., Levitas, E., Zeadna, A., Abofoul-Azab, M., Dyomin, V., Sheffield, V.C., Lunenfeld, E., Huleihel, M., and Parvari, R. (2017). Mutation in *TDRD9* causes non-obstructive azoospermia in infertile men. J. Med. Genet. *54*, 633–639.

98. Yang, F., Silber, S., Leu, N.A., Oates, R.D., Marszalek, J.D., Skaletsky, H., Brown, L.G., Rozen, S., Page, D.C., and Wang, P.J. (2015). *TEX11* is mutated in infertile men with azoospermia

and regulates genome-wide recombination rates in mouse. EMBO Mol. Med. *7*, 1198–1210.

99. Gershoni, M., Hauser, R., Yogev, L., Lehavi, O., Azem, F., Yavetz, H., Pietrokovski, S., and Kleiman, S.E. (2017). A familial study of azoospermic men identifies three novel causative mutations in three new human azoospermia genes. Genet. Med. *19*, 998–1006.

100. Maor-Sagie, E., Cinnamon, Y., Yaacov, B., Shaag, A., Goldsmidt, H., Zenvirt, S., Laufer, N., Richler, C., and Frumkin, A. (2015). Deleterious mutation in *SYCE1* is associated with non-obstructive azoospermia. J. Assist. Reprod. Genet. *32*, 887–891.

101. Colombo, R., Pontoglio, A., and Bini, M. (2017). Two novel *TEX15* mutations in a family with nonobstructive azoospermia. Gynecol. Obstet. Invest. *82*, 283–286.

102. Yang, F., Gell, K., van der Heijden, G.W., Eckardt, S., Leu, N.A., Page, D.C., Benavente, R., Her, C., Höög, C., McLaughlin, K.J., and Wang, P.J. (2008). Meiotic failure in male mice lacking an X-linked factor. Genes Dev. *22*, 682–691.