Characterization of Telomeric Repeat-Containing RNA (TERRA) localization and protein interactions in Primordial Germ Cells of the mouse

Running tittle: Characterization of TERRA in mouse PGCs

Summary sentence: TERRA transcription and interacting proteins during PGC development are regulated in a dynamic fashion that is dependent on gestational age and sex

Keywords: TERRA, Primordial Germ Cells, Long non-coding RNA, Telomeres, SFPQ, NONO

Miguel A. Brieño-Enríquez[2], Stefannie L. Moak, Anyul Abud-Flores[1], and Paula E. Cohen[2]

Department of Biomedical Sciences and Center for Reproductive Genomics, Cornell University, Ithaca, New York 14853, United States of America

[1] Current Address: Facultad de Medicina de la, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

[2] Corresponding authors: mab587@cornell.edu and paula.cohen@cornell.edu

Research reported in this publication was supported in part by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number K99HD090289 to M.A.B-E. A seed grant from the Cornell Center for Reproductive Genomics to M.A.B-E, using funds obtained as part of the NICHD National Centers for Translational Research in Reproduction and Infertility (NCTRI), award number P50HD076210 to P.E.C. and Empire State Stem Cell Fund through New York State Department of Health Contract # C30293GG. Imaging data was acquired through the Cornell University Biotechnology Resource Center, with NSF funding #1428922 for the shared Zeiss Elyra Microscope. NIH SIG 1S10 OD017992-01 grant support the Orbitrap Fusion mass spectrometer. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Telomeres are dynamic nucleoprotein structures capping the physical ends of linear eukaryotic chromosomes. They consist of telomeric DNA repeats (TTAGGG), the shelterin protein complex, and Telomeric Repeat-Containing RNA (TERRA). Proposed TERRA functions are wide-ranging and include telomere maintenance, telomerase inhibition, genomic stability, and alternative lengthening of telomere. However, the role of TERRA in primordial germ cells (PGCs), the embryonic precursors of germ cells, is unknown. Using RNA-fluorescence in situ hybridization (RNA-FISH) we identify TERRA in PGCs soon after these cells have migrated to, and become established in, the developing gonad. RNA-FISH showed the presence of TERRA transcripts in female PGCs at 11.5, 12.5 and 13.5 days post-coitum. In male PGCs, however, TERRA transcripts are observable from 12.5 dpc. Using qPCR we evaluated chromosome-specific TERRA expression, and demonstrated that TERRA levels vary with sex and gestational age, and that transcription of TERRA from specific chromosomes is sexually dimorphic. TERRA interacting proteins were evaluated using Identification of Direct RNA Interacting Proteins (iDRiP) which identified 48 in female and 26 in male protein interactors specifically within nuclear extracts from PGCs at 13.5 dpc. We validated two different proteins the splicing factor, proline- and glutamine-rich (SFPQ) in PGCs and Non-POU domain-containing octamer-binding protein (NONO) in somatic cells. Our results show that, TERRA interacting proteins are determined by sex in both PGCs and somatic cells. Taken together, our data indicate that TERRA expression and interactome during PGC development are regulated in a dynamic fashion that is dependent on gestational age and sex.

Introduction

Telomeres are dynamic nucleoprotein structures capping the physical ends of linear eukaryotic chromosomes. Telomeres protect the chromosome ends from degradation and erroneous recombination events and, as such, are essential for ensuring genome stability [1]. They consist of DNA, proteins and RNA. Telomeric DNA consists of double strand DNA repeats (TTAGGG) which extend 9-15 kb in length in humans but can be as long as 100 kb in rodents [2-4]. The protein component of the telomere is comprised of the shelterin complex, consisting of TRF1, TRF2, RAP1, TIN2, TPP1 and POT1 [5]. The shelterin complex plays numerous roles in telomere stabilization, including suppression of the DNA damage response (DDR) and the regulation of telomerase activity [6]. The Telomeric RNA component is TERRA (Telomeric repeat-containing RNA) a long non-coding RNA (lncRNA) [3, 7-9]. Long non-coding RNAs are a class of RNAs defined by their size and by their lack of a translation product [10].

TERRA transcripts are comprised of UUAGGG repeats that are transcribed by RNA polymerase II, initiating from the subtelomeric regions of the telomeres and proceeding toward the chromosome ends [5]. TERRA transcription in human cells is regulated by promoters localized at all subtelomeric regions, and by methylation on the CpG islands of these promoters [11, 12]. By contrast, in mice only one promoter has been described on chromosome 18 and its regulation does not appear to dependent on methylation [13]. TERRA expression levels and localization are cell cycle dependent, being high during the G1-S transition, peaking at early S phase and declining as cells transition to G2 and M phase [11]. TERRA was initially described to be associated exclusively with the ends of linear eukaryotic DNA but recently studies have indicated that is also associates with other regions of the genome [14].

Several TERRA functions have been described, including telomere maintenance, telomerase inhibition, telomeric heterochromatin formation, genomic stability, an alternative lengthening mechanism for telomeres [14-16], and regulation of telomerase [14, 17]. More recently, TERRA has been implicated in the protection of telomere stability, in which TERRA competes with ATRX to bind to DNA [14], and in the regulation of sex chromosome pairing during stem cell X chromosome inactivation, where TERRA creates a hub to guide X inactivation center homology searching [18].

TERRA localizes to the telomeres of mammalian germ cells (oocytes and spermatocytes) during all stages of meiotic prophase I, where it appears to be regulated in a sex-specific manner [3, 19]. However, the function and timing of TERRA transcription remains unclear. Preliminary data suggest that TERRA transcription initiates in the primordial germ cell stage [20]. Primordial Germ Cells (PGCs) are the embryonic precursors of the germ cell lineage that will form the oocytes and sperm required for sexual reproduction [21, 22]. In mice, PGCs first become identifiable as a cluster of approximately 40 cells at the base of the allantois at around embryonic day 6.25 days post coitum (dpc). Starting around 7.5 dpc, PGCs begin their migration towards the presumptive gonad, colonizing the genital ridge by ~10.5 dpc [21, 23], and reaching their maximal numbers by ~13.5 dpc through continued migration and proliferation [24]. In the female embryo, at 13.5 dpc PGCs then enter prophase I of meiosis and arrest at diplotene by around the time of birth. Meanwhile, male PGCs at this gestational age undergo mitotic arrest in the G_0/G_1 phase and stay quiescent for the remainder of the embryonic period, only initiating meiosis after birth [21].

In the current study, we have investigated the localization, transcription, and protein interactions of TERRA in male and female PGCs just after sex determination, and prior to the time that female germ cells enter meiosis and male germ cells become quiescent. Our results demonstrate key differences in the expression of TERRA and the localization of TERRA between male and female PGCs and across gestational age. Importantly, we also observe sex and age-dependent variations in the proteins with which TERRA interacts. Our results demonstrate that TERRA expression, localization and interactome in PGCs are sexually dimorphic and dependent on developmental age, suggesting that TERRA regulation may be an important component of the sex-specific development of the mammalian germ line.

Materials and Methods

Mouse Handling and Care

All mouse studies were conducted with the prior approval of the Cornell Institutional Animal Care and Use Committee. 6 to 8 week old wild-type C57Bl/6J mice were mated and checked for vaginal plugs the next morning. Female mice with plugs were moved into a separate cage, and were considered to be at 0.5 days post-coitum (dpc). At 11.5, 12.5 and 13.5 dpc, the mothers were sacrificed to recover gonads from fetuses for the isolation of PGCs.

Extraction of Primordial Germ Cells

PGCs were purified by magnetic cell sorting of gonads from 11.5, 12.5 and 13.5 dpc male and female embryos following a published protocol with some modifications [25, 26]. Briefly, the gonads from 11.5, 12.5 and 13.5 dpc embryos were dissected under a stereomicroscopy in EmbryoMax M2 Medium (Merck-Millipore). Gonads at 12.5 dpc and 13.5 dpc were classified according to the anatomical characteristics. For 11.5 dpc gonads dissected and sexed by PCR following the method described by McClive, P.J. & Sinclair [27]. Testis and ovaries were disaggregated in 0.25% trypsin-EDTA containing 20 µg/ml DNase (Sigma-Aldrich, St. Louis, MO. USA) at RT. The enzymatic reaction was stopped by adding M2 medium with 10% FCS (Gibco, Thermo Fisher, Waltham, MA. USA). The cells were centrifuged at 3220 rcf for 2 min and washed twice with 500 µl of M2 medium containing 10% FCS and 20 µg/ml DNase. The cell pellet was resuspended in 400 µl of M2 medium, mixed with 30 µl anti-SSEA-1 (CD15) MicroBeads (Miltenyi Biotech, Bergisch Gladbach, Germany) and incubated for 45 min at 4°C. PGCs were isolated from somatic cells with a miniMACs column following the manufacturer's instructions and counted (Miltenyi Biotech, Bergisch Gladbach. Germany). The purity of PGCs was verified by, counting cells that stained positive for alkaline phosphatase with the naphtol AS-MX/ FAST-RED procedure (Sigma-Aldrich, St. Louis, MO. USA). In all cases the purity of the PGCs was evaluated from cell counts in 4 different fields of the microscope until a total number of 100 cells were recorded. The purity of PGCs varies according to gestational age: at 11.5 dpc, cell suspensions showed a purity

of 80-85%, at 12.5 dpc the purity of PGCs ranged from 85 to 89%, and at 13.5 dpc PGC there was a recorded 90–95% purity. Typical yield of PGCs per embryo varies by gestational age. At 11.5 dpc the total isolated PGCs was approximately 800, at 12.5 dpc there were approximately 4000 PGCs isolated per embryo and approximately 10000 PGCs per embryo at 13.5 dpc. The somatic cells that were separated in the MS column were also collected for analysis.

Primordial germ cell (PGCs) and somatic cell spreading (Soma)

After extraction, 20 µl of the cell suspension was placed on poly-L-lysine slides that were previously cleaned with RNaseZap (Thermo Fisher, Waltham, MA USA). Cells were permeabilized by incubation for 10 minutes in CSK buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl2, 10 mM PIPES (Sigma-Aldrich, St. Louis, MO. USA), 0.5% Triton X-100 (Fisher Scientific, Pittsburgh, PA. USA) and 10 mM ribonucleoside-vanadyl complex (New England Biolabs, Ipswich, Massachusetts. USA)) [28]. Cells were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA. USA) for 10 min and then washed in 70% ethanol. Slides were stored at -80°C until use.

Immunofluorescence

Immunofluorescence (IF) was performed as previously described with modifications [28]. Slides were blocked 10 min with PTBG (1x PBS, 0.1% Tween-20, 0.2% BSA, 0.2% gelatin) and then incubated overnight at 4°C with mouse anti-*Tert* (dilution 1:100 from Rockland antibodies #600-401-252S), anti-NONO (dilution 1:100 from Proteintech #11058), anti-SFPQ (dilution 1:100 from ProteinTech #15585) and anti-VASA (DDX4) (dilution 1:200 from Abcam #ab13840) in PTBG. Slides were washed three times for 5 minutes in PBST (0.1% Tween-20 in 1x PBS) before incubation at 37°C for 40 minutes with the following secondary antibody: Alexa Fluor 488-AffiniPure F(ab')2 Fragment goat anti-mouse IgG (H+L), Alexa Fluor 488-AffiniPure F(ab')2 Fragment goat anti-rabibit IgG (H+L) and Alexa Fluor 594-AffiniPure F(ab')2 Fragment goat anti-rabbit IgG (H+L), all of them from Jackson Immunoresearch (West Grove, PA. USA). Secondary antibodies were incubated at concentrations of 1:1000 in PTBG at

37°C for 60 minutes. Slides were washed three times in PBST, fixed 10 min in 4% paraformaldehyde in 1x PBS (pH 7), and rinsed with 1x PBS. At the end of the IF procedure, RNA-FISH was performed.

RNA-Fluorescence In Situ Hybridization

TERRA focus detection was performed by RNA-FISH [28, 29]. RNA-fluorescence in situ hybridization (RNA-FISH) was performed immediately after IF. Briefly, after dehydration through a ice cold graded ethanol series (70%, 80%, 90% and 100%, during 5min each), cells were hybridized overnight at 37°C with a 25 nM $(CCCTAA)^3$ oligonucleotide probe conjugated with Cy3 (Integrated DNA Technologies IDT, Coralville, IO. USA) in hybridization buffer (10% of 20x SSC (Sigma-Aldrich, St. Louis, MO. USA), 20% 10 mg/ml BSA (Sigma-Aldrich, St. Louis, MO. USA), 20% of 50% dextran sulfate (Fisher Scientific, Pittsburgh, PA. USA), and 5% deionized formamide (Fisher Scientific, Pittsburgh, PA. USA). Next, slides were washed with 50% formamide/1xSSC during 5 min followed by two washes of 2xSSC at 39°C. Finally, cell nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI) diluted in Vectashield (Vector Laboratories, Burlingame, CA). For each sample, a negative control was included, consisting of slides treated 10 min with RNAse A (Sigma-Aldrich, St. Louis, MO. USA) at a concentration of 100 µg/ml prior to RNA-FISH performance. Foci were counted as signals that appeared as discrete dots or signals within the cell. At least 75 cells were counted per each PGC pool and 50 for each somatic cells pool.

Immunohistochemistry

Immunohistochemistry was performed on paraffin-embedded sections from 13.5 dpc ovaries and testis. Slides were deparaffinized and rehydrated with 3 washes of Safeclear (Fisher Scientific, Pittsburgh, PA. USA) for 5 minutes each, followed by 3 washes of each concentration in a graded series of ethanol (100%, 95%, 80%, 70%). After rinsing the slides 2 times for 5 minutes in distilled water, the slides were incubated in sodium citrate pH 6.0 during 40 min at 95 °C. Permeabilization was performed in 0.2% of Triton-X 100 in PBS for 30 min. Section were blocked for 4 h in blocking solution (2.52 mg/ml glycine,

10% goat serum, 3% BSA in PBS-T) and then incubated with primary antibody overnight at RT. After 2 washes with PBST, the slides were incubated with secondary antibodies for 2 h at RT. The slides where rinsed in PBST, cell nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI) and mounted in Vectashield (Vector Laboratories, Burlingame, CA).

Imaging

Imaging of PGCs and somatic cells was performed using ELYRA 3D-Structured Illumination Super resolution Microscopy (3D-SIM) from Carl Zeiss with ZEN Black software (Carl Zeiss AG, Oberkochen. Germany). Images are show as maximum intensity projections of z-stack images- 3D-SIM. To reconstruct high-resolution images, raw images were computationally processed by ZEN Black. Channel alignment was used to correct for chromatic shift. The brightness and contrast of images were adjusted using ImageJ (National Institutes of Health, USA). Image acquisition of the tissue sections was performed using a Zeiss Imager Z1 microscope under 20X, 40X or 63X magnifying objectives, at RT. Images were processed using ZEN 2 (Carl Zeiss AG, Oberkochen. Germany).

RNA isolation and reverse transcription

Total RNA from each gestational age (11.5, 12.5 and 13.5 dpc for both male and female gonads) was extracted using TRIzol (Invitrogen Co, Carlsbad, CA, USA). Total RNA was re-suspended in 40 µl of RNase free water. RNA concentration was then determined spectrophotometrically using a NanoDrop 2000 (Thermo Fisher, Waltham, MA. USA). 1.5µg of total RNA was reverse transcribed using and Superscript III First-Strand Synthesis System (Invitrogen Co, Carlsbad, CA, USA) using TERRA reverse primer or 1.65 µM random hexamers/1.25 µM Oligo(dt) for the housekeeping gene (Supplemental Table 1). cDNA was kept at -20°C until used in qPCR.

qPCR analysis

Real-time PCR amplification and analysis was performed following the protocol previously described by Feretzaki and Lingner [30]. Primers were designed to amplify the subtelomeric region of different chromosomes using the sequences previously published by Lopez de Silanes et al., 2014 [13] (Supplemental Table 1) (Subtelomeric regions with active transcription of functional genes were not included). Gene expression was normalized to succinate dehydrogenase complex flavoprotein subunit A (SDHA) expression [31]. Each reaction mixture consisted of, 1 µl of cDNA, 0.5 µl of forward primer (0.2 µM), 0.5 µl of reverse primer (0.2 µM), 10 µl of Roche FastStart SYBR Green Master (Sigma-Aldrich, St. Louis, MO. USA) and 7 µl of nuclease-free water. qRT-PCR amplification reaction was performed with specific primers (Supplemental Table 1). PCR conditions were the same as those used by Feretzaki and Lingner [30]: 30 s at 95°C, followed by 40 cycles at 95°C for 1 s and 60°C for 60 s. After PCR, melting curve analyses were performed to verify specificity and identity of the PCR products. All data were analyzed with the CFX-manager Bio-Rad (Bio-Rad Laboratories). All analyzed genes were performed in triplicate for each one of the 3 biological samples of the three different developmental ages (11.5, 12.5 and 13.5 dpc) and both sexes (female and male). qPCR data for TERRA quantification are analyzed using the relative quantification method [32]. This method feeds the Ct values obtained from the qPCR experiment into a series of subtractions to calculate the relative gene expression of the gene of interest (TERRA) normalized against a reference gene (SDHA) in different conditions as was described previously by [14, 18, 30] .

Identification of Direct RNA Interacting Proteins (iDRiP)

Proteins that directly interact with TERRA were identified through method called iDRiP, following a published protocol with some adaptations [33]. The original iDRiP method utilized large numbers of cultured somatic cells (around 30 million). In the current study, we adapted the conditions to reduce the input of cells around 10-fold and to use primary PGCs from 13.5 dpc male and female gonads, along with gonadal somatic cells as controls. After cell isolation, cells were rinsed with cold PBS 3 times and

the plated in a petri dish for 30 min at 37°C 5% CO_2. Excess PBS was removed and the cells were irradiated with UV light at 200 mJ/cm^2 energy (Spectrolinker XL, UV crosslinker. Spectronics Corporation. Westbury, NY. USA). The cells were transferred to an eppendorf tube and spun down at 1258 rcf for 5 min at 4°C. The pellet was re-suspended in 300 µl cold cytoskeleton buffer (CSKT) (0.5% Triton X-100, 0.5% PIPES, 100 mM NaCl, 3 mM MgCl2, 0.3 M sucrose and 1 mM PMSF) (all from Fisher Scientific, Pittsburgh, PA. USA) with protease inhibitor (1x Roche Protease Inhibitor Cocktail Tablets) and incubated for 10 min at 4°C on a rocker. Cells were spun down again at 453 rcf at 4°C for 5 minutes, supernatant was removed and re-suspended in nuclear isolation buffer (10 mM Tris pH 7.5, 10 mM KCl, 0.5% Nonidet-P 40, 1x protease inhibitors, 1 mM PMSF), before spinning again at 453 rcf at 4°C for 10 min. Supernatant was removed and the cell pellets were flash frozen in liquid nitrogen and stored at -80°C until use. Cells were pooled from number of ≅300 female and ≅300 male embryos and thawed at 37°C. 500 µl Turbo DNaseI buffer, 50 µl Turbo DNaseI enzyme (2U/µl), 10 µl superaseIN (Thermo Fisher, Waltham, MA USA) and 5 ul of 50x protease inhibitor was added to the cell suspension. Samples were incubated at 37°C for 45 min on a rocker. The nuclear lysates were further solubilized by adding 1% sodium lauryl sarcosine, 1x protease inhibitor, 0.3 M lithium chloride, 25 mM EDTA and 25 mM EGTA to final concentrations (all of them from Sigma-Aldrich, St. Louis, MO. USA). Samples were mixed well and incubated again at 37°C for 15 min. As a positive control, the highly expressed RNA U6 was used, and RNAse A treated samples as a negative control. U6 and TERRA-specific biotinylated probes (Integrated DNA Technologies IDT, Coralville, IO. USA), were conjugated to streptavidin beads (MyOne streptavidin C1 Dyna beads, Invitrogen) for a 30 min incubation period at RT. The conjugated beads were mixed with the lysates and incubated at 55°C for one hour before overnight incubation at 37°C in a hybridization chamber. After incubation, the beads were washed three times in wash buffer (10 mM Tris, pH 7.5, 0.3 M LiCl, 1% LDS, 0.5% Nonidet-P 40, 1x protease inhibitor) at RT then treated with DNase I digestion buffer, Turbo DNase I, 0.3 M LiCl, protease inhibitors, and SuperaseIn (Thermo Fisher, Waltham, MA USA) at 37°C for 20 min. Beads were re-

suspended and washed two more times in the wash buffer. For mass spectrometry analysis, proteins were eluted in Elution Buffer (10 mM Tris, pH 7.5, 1 mM EDTA) at 70°C for 4 min.

Nano-scale reverse phase chromatography and tandem MS (nanoLC-MS/MS)

Mass spectrometry of iDRiP-derived proteins was performed in the Cornell University Proteomics and Mass Spectrometry facility. The nanoLC-MS/MS analysis was carried out using an Orbitrap Fusion (Thermo-Fisher Scientific, San Jose, CA) mass spectrometer equipped with a nanospray Flex Ion Source using high energy collision dissociation (HCD) and coupled with the UltiMate3000 RSLCnano (Dionex, Sunnyvale, CA). Each reconstituted samples for both PGC and SOMA (18 ul) was injected onto a PepMap C-18 RP nano trap column (3 µm, 100 µm × 20 mm, Dionex) with nanoViper Fittings at 20 µL/min flow rate for on-line desalting and then separated on a PepMap C-18 RP nano column (3 µm, 75 µm x 25 cm), and eluted in a 120 min gradient of 5% to 35% acetonitrile (ACN) in 0.1% formic acid at 300 nL/min. The instrument was operated in data-dependent acquisition (DDA) mode using FT mass analyzer for one survey MS scan for selecting precursor ions followed by 3 second "Top Speed" data-dependent HCD-MS/MS scans in Orbitrap analyzer for precursor peptides with 2-7 charged ions above a threshold ion count of 10,000 with normalized collision energy of 38.5%. For label-free protein analysis, one MS survey scan was followed by 3 second "Top Speed" data-dependent CID ion trap MS/MS scans with normalized collision energy of 30%. Dynamic exclusion parameters were set at 1 within 45s exclusion duration with ±10 *ppm* exclusion mass width. Two samples from each group PGC and SOMA were analyzed in Orbitrap in the order of female followed by male samples for data acquisition. All data are acquired under Xcalibur 3.0 operation software and Orbitrap Fusion Tune 2.0 (Thermo-Fisher Scientific).

NanoLC-MS/MS data processing, protein identification and data analysis

All MS and MS/MS raw spectra from each experiment were processed and searched using the Sequest HT search engine within the Proteome Discoverer 2.2 (PD2.2, Thermo). The default search settings

used for relative protein quantitation and protein identification in PD2.2 searching software were: two mis-cleavage for full trypsin with fixed carbamidomethyl modification of cysteine and oxidation of methionine and demaidation of asparagine and glutamine and acetylation on N-terminal of protein were used as variable modifications. Identified peptides were filtered for maximum 1% false discovery rate (FDR) using the Percolator algorithm in PD 2.2. The relative label free quantification method within Proteome Discoverer 2.2 software was used to calculate the protein abundances. The intensity values of peptides, which were summed from the intensities values of the number of peptide spectrum matches (PSMs), were summed to represent the abundance of the proteins. For relative ratio between the two groups, here PGC female/male and Soma female/male, no normalization on total peptide amount for each sample was applied. Protein ratios are calculated based on pairwise ratio, where the median of all possible pairwise ratios calculated between replicates of all connected peptides.

Statistical Analysis

Statistical analyses were performed using GraphPad Prism version 6.00 for Macintosh (GraphPad Software, San Diego California USA, www.graphpad.com). Specific analyses are described within the text and the corresponding figures. Alpha value was established at 0.05.

Results

Developmental changes in TERRA localization are different in male and female PGCs

Using RNA-FISH and 3D-SIM microscopy, we evaluated the presence of TERRA at three developmental stages (11.5 12.5 and 13.5 dpc) in both sexes. The selection of these stages was based on the timing of sex determination in the mouse, the methylation status of PGCs, and the entrance into meiosis of female PGCs after 13.5 dpc [21]. Quantitation of TERRA focus numbers was performed in both male and female SSEA-1 positive PGCs as well as in SSEA-1 negative cells (somatic cells). As a negative control, cells treated with RNAse A were used in which TERRA RNA should be completely degraded. A total number of 1352 PGCs and 973 somatic cells were analyzed. Female and male

PGCs, and somatic cells were obtained from at least 3 different pools at the different gestational ages. Pools consisted of PGCs from between 20 and 40 gonad pairs (1 pair per embryo).

Analysis of female PGCs and somatic cells showed the presence of discrete foci of TERRA (Fig. 1A and 1B). The mean focus number observed in 11.5 dpc female PGCs was 0.76 ± 0.37 per cell, rising to 3.88 ± 0.73 at 12.5 dpc, and 8.76 ± 4.03 at 13.5 dpc. Statistical analysis revealed significant differences among the three gestational ages (p=0.001 ANOVA; Fig. 1B). Meanwhile, TERRA focus counts from female gonadal somatic cells at the same gestational ages revealed no statistical differences from 11.5 dpc (19.76 ± 1.143), 12.5 dpc (19.86 ± 1.52), and 13.5 dpc (19.78 ± 1.44; Fig. 1C and 1D). The number of TERRA foci in somatic cells was considerably higher than that observed in the neighboring germ cells at each gestational age. These results indicate that TERRA focus numbers alter with gestational in female PGCs but not in neighboring somatic cells.

In male PGCs, the dynamics of TERRA focus accumulation were very different to that seen in females. We were not able to detect TERRA signal in male 11.5 dpc PGCs. Instead, the earliest detection TERRA in male PGCs was at 12.5 dpc where we observed cells with either zero or one TERRA focus (0.73 ± 0.45 foci/cell). A statistically significant increase in the TERRA foci was observed in male 13.5 dpc PGCs, where the mean focus number rose to 4.03 ± .82 (p=0.001; Fig. 1A). Adjacent somatic cells of the male gonad showed TERRA focus numbers that were indistinguishable at all ages from that of female gonads, with no statistically significant differences found between sex or gestational age (11.5 dpc: 19.66 ± 1.02 foci/cell, 12.5 dpc: 20.05 ± 1.71 foci/cell, and 13.5 dpc: 19.35 ± 1.76 foci/per cell; Fig. 1D). These results indicate that in male PGCs, like in female PGCs, the gestational age is a key factor in TERRA localization in PGCs but not in somatic cells.

Comparison of TERRA foci numbers in both female PGCs and male PGCs reveled a sex bias. Compared to male PGCs, female PGCs showed significantly more TERRA foci at all the stages of development (p=0.001; Fig. 1B), while no differences were observed at any gestational age between male and female somatic cells (Fig. 1D).

Differential transcription of TERRA in male and female PGCs

In human cell lines, transcription of TERRA is regulated by chromosome specific promoters that are repressed by CpG methylation [34, 35]. However, previous reports in mouse cell lines showed that almost all the TERRA transcripts are transcribed from the subtelomeric region of chromosome 18, with some minor contribution from chromosome 9 [13]. However, nothing is known about TERRA transcription during this unique period of PGC development in which distinct changes in DNA methylation are occurring. Thus, we hypothesized that the loss of methylation that is a unique feature of PGCs may result in increased TERRA transcription, resulting in the increased localization of TERRA that we observed through this gestational time span. We performed qPCR in isolated female and male PGCs at 11.5, 12.5 and 13.5 dpc using the subtelomeric sequences of each chromosome previously published by Lopez de Silanes et al., 2014 [13] (Supplemental Table 1). As expected from the localization of TERRA, our results showed higher levels of TERRA expression in female PGCs compared to male PGCs at each gestational age (Fig. 2A-2O). Furthermore, and in contrast to previous results, we observed that TERRA is transcribed from multiple telomeres in a gestational and sex-dependent manner, though not all telomeres were found to be transcriptionally active (only those with any expression of TERRA are shown in Figure 2).

Female PGCs showed TERRA expression at 11.5 dpc from 8 different chromosome subtelomeric regions (Chromosomes 1, 2, 7,9, 11,13, 17 and 18; Fig. 2A, 2B, 2E, 2G, 2I, 2J, 2L and 2M, respectively), while in male PGCs at the same age, TERRA expression was confined to the chromosome 17 subtelomeric region (Figure 2L). At 12.5 dpc, we detected increased expression from chromosomes 5, 6, 8, 10, 15, 19 in both female and male PGCs (Fig. 2C, 2D, 2F, 2H, 2K and 2N, respectively). The exception was the subtelomeric region of chromosome X, from which transcription of TERRA was only detected for female PGCs (Fig. 2O). Transcription of TERRA from the single X chromosome of male PGCs only became evident at 13.5 dpc, but at a very low level compared to that of female PGCs at this gestational age (Fig, 2O). Most TERRA transcription in 13.5 dpc male PGCs arose from chromosomes 2 and 6, and only in the case of the latter was there higher transcription in

male PGCs than female PGCs (Fig. 2B, 2D). Indeed, at all gestational ages, we observed higher transcription of TERRA at each telomere in female PGCs, except for the subtelomeric region of chromosomes 2, 6 and 10. These results indicate that the transcription of TERRA is differentially regulated in the male and female germ line, and that transcription of TERRA in male PGCs is developmentally delayed compared to that in female PGCs. Moreover, our results demonstrate that TERRA is transcribed from multiple subtelomeric regions in mouse PGCs.

TERRA and the catalytic subunit of the enzyme telomerase (TERT) colocalization and expression are regulated by gestational age and sex

There is a lot of controversy regarding the role of TERRA in the regulation of Telomerase (TERT). The most common idea suggests that TERRA expression down-regulates or inhibits the catalytic subunit of the TERT [17, 36]. Previous reports have indicated that there is decay in the *Tert* expression with age in male germ cells, including PGCs [37, 38]. Confounding this, however, are the suggestions that high levels of telomerase are required to maintain spermatogonia in their undifferentiated state [39]. Using SIM microscopy we evaluated the colocalization of TERT and TERRA at 13.5 dpc, and plotted the percent colocalization in PGCs from both sexes (percentages were obtained from the number of TERRA-TERT foci divided by the total TERRA foci, and multiplied by 100) (Fig. 3A and 3B). In female PGCs, 63.5% of the TERRA foci colocalized with TERT but only the 36.1% in 13.5 male PGCs (Fig. 3C; p=0.0001).

We evaluated the expression of *Tert* using qPCR, and showed a statistically significant decrease of *Tert* expression 11.5 dpc to 13.5 dpc in both male and female PGCs (Fig. 3D; p=0.001). Consistently, however, *Tert* expression was significantly higher in male PGCs than in age-matched female PGCs (Fig. 3D). Thus, the increasing TERRA focus count with gestational age, and the relatively increased number of TERRA foci/cell in female germ cells compared to male germ cells are both inversely correlated with *Tert* expression, which declines with gestational age and which is higher in male PGCs.

The TERRA interactome is sexually dimorphic

To understand TERRA function, it is important to identify key TERRA interacting proteins in the cell type of interest. Given that we obtain so few PGCs at specific developmental ages, we focused on only one gestational age in which to investigate the TERRA interactome in PGCs and somatics cells: 13.5 dpc. We performed iDRiP (Identification of Direct RNA Interacting Proteins), using three million PGCs from both male and female gonads, along with comparable neighboring somatic cells. A total of 48 proteins were identified in female PGCs and 26 in male PGCs, of which 32 (55.2%) were unique to female PGCs and 10 (17.2%) were specific for male PGCs (Fig. 4A and 4B). The remaining 16 (27.6%) TERRA-associated proteins were shared between female and male PGCs (Fig. 4A). Using the relative label free quantification method within Proteome Discoverer 2.2 software, we calculated the protein abundances. The intensity values of peptides, which were summed from the intensities values of the number of peptide spectrum matches (PSMs), were summed to represent the abundance of the proteins. The results of these relative label free quantitations showed different relative levels of protein between female and male PGCs (Fig. 4B) (Supplemental Table 2).

In somatic cells, we observe far higher numbers of TERRA interacting proteins, but with greater number of proteins in male somatic cells than in female somatic cells (Fig. 4C). 118 proteins were obtained from female somatic cells and 158 were obtained from male somatic cells (Fig. 4C), with 31 (16.4%) and 71 (37.6%) proteins, respectively, being unique to one sex. Overall, female and male somatic cells shared 87 (46%) TERRA interacting proteins (Fig. 4D). Similar to PGCs, we used the relative label free quantification method to compare the relative abundance of proteins in both female and male somatic cells. Our results showed a different distribution of the relative protein abundance in female gonads compared to male gonads (Fig. 4D, and Supplemental Table 3).

SFPQ interacts with TERRA in PGCs

To validate the interactions of TERRA we used two different approaches. First, we immunolocalized the protein of interest on 13.5 dpc ovaries and testis, and secondly, we co-localized TERRA and the protein

on isolated cells. Based on our iDRIP-identified interacting proteins, we decided to analyze proteins with the most extreme ratios between male and female. After antibody testing and standardization, we selected two proteins for validation one for PGCs and another for somatic cells. The splicing factor, proline- and glutamine-rich (SFPQ) showed one of the lowest male/female ratios (0.01) (Supplemental Table 3). Using antibodies against the germ cell lineage marker VASA (DDX4/MVH) and SFPQ, we first determined the specific localization to the nucleus in both female and male PGCs by immunofluorescence on tissue sections (Fig. 5A and B). We observed the presence of SFPQ signal on the nuclei of the VASA positive cells (PGCs), but not in the VASA negative cells (somatic cells). Thereafter, we performed IF followed by TERRA RNA-FISH to evaluate the cololocalization of SPFQ and TERRA on isolated female and male 13.5 dpc PGCs. Since the number of TERRA foci varies within female/male, we compared the percentages of colocalizing foci between male and female PGCs. Percentages were obtained from the number of TERRA-SFPQ foci divided by the total TERRA foci, and multiplied by 100). Interestingly, we observed colocalization in both female and male PGCs (Fig. 5C). However, female PGCs showed increased percentage of colocalization of TERRA-SFPQ (36%) compared to male PGCs (30%) (Fig. 5D; p=0.01), indicating although TERRA interacts with SFPQ in both sexes, there is an increased interaction in female PGCs

NONO interacts with TERRA in somatic cells of the fetal gonad

Similar to our approach for SFPQ, above, we selected one TERRA-ineracting protein to validate on somatic cells. We analyzed the localization and interaction of TERRA with Non-POU domain-containing octamer-binding protein (NONO). First, we evaluated the localization of NONO in female and male gonads, the use of VASA antibody allowed us to identify PGCs from somatic cells. In both sexes, we observed a positive NONO staining in somatic cells with a clear absence of NONO staining on the nuclei of the VASA positive cells (Fig. 6A and 6B). Using IF followed by TERRA RNA-FISH we evaluated the colocalization of NONO with TERRA on isolated somatic cells (Fig. 5C). Colocalization analysis showed that the 42% of the TERRA foci colocalized with NONO on 13.5 female somatic cells,

however on the male somatic only the 25% of the TERRA foci colocalized with NONO (percentages were obtained from the number of TERRA-NONO foci divided by the total TERRA foci, and multiplied by 100) (Fig. 5 C and D; p=0.01). These results indicate that the colocalization TERRA and NONO is restricted to the somatic cell lineages of the fetal gonad.

Discussion

To date, there is a very limited understanding of the function of TERRA, especially as it pertains to germ cells. Herein, we describe for first time the localization, expression and the protein interactions of TERRA in PGCs during the critical period of epigenetic reprogramming and the onset of sex determination. TERRA transcription, at least in human cells, was described to be regulated by methylation status [15, 40]. During PGC development, genome-wide DNA demethylation in mid-gestation results in the lowest levels of methylated DNA at around 13 dpc [21, 41]. Therefore, we hypothesized that the transcription and localization of TERRA within primordial germ cells will increase inversely to the drop in DNA methylation. At the same time, this period is associated with gonadal sex determination and the concomitant entry of female PGCs into meiosis, with male PGCs entering a period of mitotic quiescence. Thus, we suggest that this period of development would provide for key differences in TERRA regulation between the male and female germ lineages. Our results showed that both TERRA focus distribution and TERRA transcription increases with gestational age in both female and male PGCs. However, the increases are significantly higher in PGCs isolated from ovaries compared to those obtained from testis, and there is a delay by one day in the onset of TERRA transcription in males.

Sexual dimorphism in the epigenetic marker and DNA methylation patterns in PGCs has been extensively studied at different stages of gestation in the mouse [21, 22]. In both sexes, the lowest levels of genome-wide methylation are observed at 13.5 dpc, but the recovery of the epigenetic marks takes longer in the female germ line compared to male[42]. Concommitant with the return of epigenetic marks in female PGCs is their synchronous entry into the meiotic program. Thus, the increased TERRA

localization may be highly influenced by the onset of prophase I in the female germ line but we cannot discard the role of the changes in methylation status. Indeed, since one proposed function of TERRA is to regulate/suppress DNA damage repair processes, and since prophase I is characterized by the induction and orchestrated repair of hundreds of DNA double strand breaks (DSBs), it is tempting to speculate that one role of TERRA is to prevent DSB induction/repair at telomeres and instead to direct DSB events to more proximal chromosomal locations. The purpose of prophase I DSB induction/repair is to generate a highly regulated number of crossovers that serve as tethering points to maintain homologous chromosome interactions until the first meiotic division. Thus the placement of DSB events at telomeres would not be ideal for this purpose.

Interestingly, TERRA focus numbers in both male and female PGCs are significantly lower than in somatic cells, and represent a frequency that is less than a quarter of the number of telomeres present in the nucleus. In the soma, by contrast, only half the diploid number of chromosomes (40 in mouse) appear to be associated with a TERRA focus. It is important to note that somatic cells of the gonad do not undergo to epigenetic mark erasure, indicating that TERRA transcription in PGCs and somatic cells at these particular stages of development is likely to be regulated by different mechanisms.

Controversial results have been published about the expression of TERRA and *Tert*, where it has been described that TERRA either recruits *Tert* to the telomere to promote its enzymatic function, or conversely TERRA is acting as a competitive inhibitor that is competing for access to telomeric DNA[17, 36]. Our results showed that in PGCs the expression of TERRA and *Tert* decreases in relation to gestational age in agreement with previous reports [37, 38, 43]. Our results demonstrate intermittant colocalization of TERT and TERRA, and also showed that when transcription of TERRA increases the *Tert* transcription decreases. However, more studies analyzing the interaction of TERRA with TERT are required to evaluate whether TERRA affects the enzymatic activity or the transcription of *Tert* in PGCs.

A previous study had investigated TERRA-protein interactions in MEFs using a pre-existing RNA-IP technique[44] only 41 proteins were identified as part of the TERRA interactome. Analysis of TERRA interacting proteins also has been performed using SILAC-labeled nuclear cell lysates, pooled

results from 2 independent pull downs showed approximately 924 interacting proteins [45]. More recently iDRiP was used to examine TERRA interactions in $15x10^7$ mouse embryonic stem cells and the authors report 134 interacting proteins ranging from components of the shelterin complex, to chromatin associated proteins, DNA repair proteins, and cell cycle regulators, to name just a few [14, 18]. In all these reports, the authors use cultured somatic cells or ESCs, allowing for high cell input (≈30 million cells). By contrast, the current study analyzed TERRA interacting proteins starting with 3 million PGCs from each sex and an equivalent number of somatic cells, representing cell purifications from 300 pups for each sex. Despite the low input, we were able to identify 48 proteins were identified in female PGCs and 26 in male PGCs. 37.5 % of these proteins were shared with the iDRiP study of Chu *et al.* (2017) using ES cells [14]. These common protein interactors included several heterogeneous ribonucleoproteins, histones, chromatin associated proteins, RNAP II and DNA repair proteins, all found in PGCs and in somatic cells. Interestingly, the SFPQ-TERRA interaction that we identify in female and male PGCs was not observed in ES cells suggesting potential specificity of this interaction for germ cells. By contrast, in somatic cells of the gonad, we identified NONO as an interactor of TERRA, and this interaction was also identified in ES Cells by Chu *et al*. (2017). The comparison between our data and previous reports suggest that there are interactions conserved in different cells types, however TERRA also have specific protein interactions depending of the cell type and at least in PGCs the sex is another determinant factor.

To validate our iDRIP data, we chose to examine further one PGC-specific (SFPQ) and somatic cell-specific (NONO) interacting protein. We did not observe clear differences in protein localization on histology sections in 13.5 dpc ovary and 13.5 dpc testis. We analyzed the colocalization of TERRA-SFPQ and TERRA-NONO, in both cases the percentage of colocalization TERRA-Protein was higher in females compared to males, indicating that at least at this specific point of development, the interaction/function of the TERRA-Protein complex is different. SFPQ is DNA- and RNA binding protein, while NONO is an RNA splicing factor [46]. Both proteins have been implicated in a range of DNA/RNA metabolic processes, including ssDNA invasion to generate a D-loop, non-homologous end joining

(NHEJ), and mRNA processing and stabilization[47]. However, their function in the developing mouse gonad including meiosis, and particularly in connection with TERRA, remains unknown.

Taken together, herein we described for first time the presence of TERRA, in mouse primordial germ cells. Our results also showed that in mouse PGCs TERRA foci number, expression and protein interactions depend on the gestational age and sex suggesting that the role of TERRA telomere dynamics is related to sexually dimorphic changes in germ cell development with age.

Competing interests

The authors declare no competing interest.

Author contribution

M.A.B-E and P.E.C designed experiments. M.A.B-E, S.L.M and A.B-F carried out experiments. M.A.B-E, S.L.M, and P.E.C analyzed and interpreted data. M.A.B-E and P.E.C wrote the manuscript.

Acknowledgements

We are grateful to Dr. Jen Grenier, Dr. Rita Reig-Viader, Dr. Anand Minajigi and Dr. Jeannie Lee for their advice and protocol sharing to perform the qPCRs, RNA-FISH and iDRiP. We thank Dr. Sheng Zhang and Dr. Ruchika Bhawal from Cornell Proteomics and Mass Spectrometry Facility for his advice in development of mass spectrometry experiments. We are also thankful to Dr. Stephen Gray, Dr. Kathryn Grive, Carolyn Milano, Amanda Touey and Jeffrey Pea for providing critical feedback for the manuscript.

References

1. Maicher, A., Lockhart, A. & Luke, B. Breaking new ground: digging into TERRA function. *Biochimica et biophysica acta* **1839**, 387-394 (2014).
2. Blasco, M.A. Telomeres and human disease: ageing, cancer and beyond. *Nature reviews. Genetics* **6**, 611-622 (2005).
3. Reig-Viader, R., Garcia-Caldes, M. & Ruiz-Herrera, A. Telomere homeostasis in mammalian germ cells: a review. *Chromosoma* **125**, 337-351 (2016).
4. Mason, J.M.O. & McEachern, M.J. Mild Telomere Dysfunction as a Force for Altering the Adaptive Potential of Subtelomeric Genes. *Genetics* **208**, 537-548 (2018).
5. Schoeftner, S. & Blasco, M.A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nature cell biology* **10**, 228-236 (2008).
6. Hu, C. *et al.* Structural and functional analyses of the mammalian TIN2-TPP1-TRF2 telomeric complex. *Cell research* **27**, 1485-1502 (2017).
7. Jain, D. & Cooper, J.P. Telomeric strategies: means to an end. *Annual review of genetics* **44**, 243-269 (2010).
8. Sfeir, A. & de Lange, T. Removal of shelterin reveals the telomere end-protection problem. *Science* **336**, 593-597 (2012).
9. Diman, A. & Decottignies, A. Genomic origin and nuclear localization of TERRA telomeric repeat-containing RNA: from Darkness to Dawn. *The FEBS journal* **285**, 1389-1398 (2018).
10. Kung, J.T., Colognori, D. & Lee, J.T. Long noncoding RNAs: past, present, and future. *Genetics* **193**, 651-669 (2013).
11. Rippe, K. & Luke, B. TERRA and the state of the telomere. *Nature structural & molecular biology* **22**, 853-858 (2015).

12. Zhang, B. *et al.* Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies. *PloS one* **4**, e5232 (2009).

13. Lopez de Silanes, I. *et al.* Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nature communications* **5**, 4723 (2014).

14. Chu, H.P. *et al.* TERRA RNA Antagonizes ATRX and Protects Telomeres. *Cell* **170**, 86-101 e116 (2017).

15. Deng, Z., Campbell, A.E. & Lieberman, P.M. TERRA, CpG methylation and telomere heterochromatin: lessons from ICF syndrome cells. *Cell cycle* **9**, 69-74 (2010).

16. Morris, K.V. & Mattick, J.S. The rise of regulatory RNA. *Nature reviews. Genetics* **15**, 423-437 (2014).

17. Redon, S., Reichenbach, P. & Lingner, J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic acids research* **38**, 5797-5806 (2010).

18. Chu, H.P. *et al.* PAR-TERRA directs homologous sex chromosome pairing. *Nature structural & molecular biology* **24**, 620-631 (2017).

19. Reig-Viader, R. *et al.* Telomeric repeat-containing RNA (TERRA) and telomerase are components of telomeres during mammalian gametogenesis. *Biology of reproduction* **90**, 103 (2014).

20. Reig-Viader, R. *et al.* Telomere homeostasis is compromised in spermatocytes from patients with idiopathic infertility. *Fertility and sterility* **102**, 728-738 e721 (2014).

21. Saitou, M. & Yamaji, M. Primordial germ cells in mice. *Cold Spring Harbor perspectives in biology* **4** (2012).

22. Saitou, M. & Yamaji, M. Germ cell specification in mice: signaling, transcription regulation, and epigenetic consequences. *Reproduction* **139**, 931-942 (2010).

23. Saitou, M. *et al.* Specification of germ cell fate in mice. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **358**, 1363-1370 (2003).

593 24. McLaren, A. & Lawson, K.A. How is the mouse germ-cell lineage established? *Differentiation; research in biological diversity* **73**, 435-437 (2005).

595 25. Pesce, M. & De Felici, M. Purification of mouse primordial germ cells by MiniMACS magnetic separation system. *Developmental biology* **170**, 722-725 (1995).

597 26. Brieño-Enriquez, M.A. *et al.* Exposure to endocrine disruptor induces transgenerational epigenetic deregulation of microRNAs in primordial germ cells. *PloS one* **10**, e0124296 (2015).

599 27. McClive, P.J. & Sinclair, A.H. Rapid DNA extraction and PCR-sexing of mouse embryos. *Molecular reproduction and development* **60**, 225-226 (2001).

601 28. Reig-Viader, R. *et al.* Telomeric repeat-containing RNA and telomerase in human fetal oocytes. *Human reproduction* **28**, 414-422 (2013).

603 29. Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798-801 (2007).

606 30. Feretzaki, M. & Lingner, J. A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA. *Methods* (2016).

608 31. van den Bergen, J.A., Miles, D.C., Sinclair, A.H. & Western, P.S. Normalizing gene expression levels in mouse fetal germ cells. *Biology of reproduction* **81**, 362-370 (2009).

610 32. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**, 402-408 (2001).

612 33. Minajigi, A. *et al.* Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349** (2015).

614 34. Negishi, Y., Kawaji, H., Minoda, A. & Usui, K. Identification of chromatin marks at TERRA promoter and encoding region. *Biochemical and biophysical research communications* **467**, 1052-1057 (2015).

617 35. Thijssen, P.E. *et al.* Chromatin remodeling of human subtelomeres and TERRA promoters upon
618 cellular senescence Commonalities and differences between chromosomes. *Epigenetics* **8**, 512-
619 521 (2013).
620 36. Cusanelli, E., Romero, C.A.P. & Chartrand, P. Telomeric Noncoding RNA TERRA Is Induced by
621 Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres. *Molecular cell* **51**,
622 780-791 (2013).
623 37. Dolci, S. *et al.* Stem cell factor activates telomerase in mouse mitotic spermatogonia and in
624 primordial germ cells. *Journal of cell science* **115**, 1643-1649 (2002).
625 38. Coussens, M. *et al.* Regulation and effects of modulation of telomerase reverse transcriptase
626 expression in primordial germ cells during development. *Biology of reproduction* **75**, 785-791
627 (2006).
628 39. Pech, M.F. *et al.* High telomerase is a hallmark of undifferentiated spermatogonia and is
629 required for maintenance of male germline stem cells. *Genes & development* **29**, 2420-2434
630 (2015).
631 40. Nergadze, S.G. *et al.* CpG-island promoters drive transcription of human telomeres. *Rna-a*
632 *Publication of the Rna Society* **15**, 2186-2194 (2009).
633 41. Guibert, S., Forne, T. & Weber, M. Global profiling of DNA methylation erasure in mouse
634 primordial germ cells. *Genome research* **22**, 633-641 (2012).
635 42. Smallwood, S.A. & Kelsey, G. De novo DNA methylation: a germ cell perspective. *Trends in*
636 *Genetics* **28**, 33-42 (2012).
637 43. Coussens, M. *et al.* Regulation and effects of modulation of telomerase reverse transcriptase
638 expression in primordial germ cells during development. *Biology of reproduction* **75**, 785-791
639 (2006).
640 44. de Silanes, I.L., d'Alcontres, M.S. & Blasco, M.A. TERRA transcripts are bound by a complex
641 array of RNA-binding proteins. *Nature communications* **1** (2010).

642 45. Scheibe, M. *et al.* Quantitative interaction screen of telomeric repeat-containing RNA reveals
643 novel TERRA regulators. *Genome research* **23**, 2149-2157 (2013).
644 46. Lee, M.W. *et al.* The structure of human SFPQ reveals a coiled-coil mediated polymer essential
645 for functional aggregation in gene regulation. *Nucleic acids research* **43**, 3826-3840 (2015).
646 47. Jaafar, L., Li, Z.T., Li, S.Y. & Dynan, W.S. SFPQ center dot NONO and XLF function separately
647 and together to promote DNA double-strand break repair via canonical nonhomologous end
648 joining. *Nucleic acids research* **45**, 1848-1859 (2017).

649

650 **Figure legends**

651 **Figure 1. TERRA presence in mouse PGCs differs by sex and developmental age.** A) Super
652 resolution imaging microscopy of TERRA RNA-FISH (red) on female and male primordial germ cells at
653 different stages of development (11.5, 12.5 and 13.5 dpc). Control negative correspond to 13.5 dpc
654 female and male PGCs treated with RNAse before the RNA-FISH procedure. Scale bar 10 μm; B)
655 Quantitation of TERRA foci numbers on female and male PGCs at different stages of development
656 (11.5, 12.5 and 13.5 dpc). Statistical analysis was performed using ANOVA followed by Kruskal-Wallis
657 multiple comparison analysis; C) Super resolution imaging microscopy of TERRA RNA-FISH (red) on
658 female and male somatic cells at different stages of development (11.5, 12.5 and 13.5 dpc). Negative
659 controls used were 13.5 dpc female and male somatic cell treated with RNAse before the RNA-FISH
660 procedure. Scale bar 10 μm; D) Quantitation of TERRA foci numbers on female and male somatic cells
661 at different stages of development (11.5, 12.5 and 13.5 dpc). Statistical analysis was performed using
662 ANOVA followed by Kruskal-Wallis multiple comparison analysis. p value was set at 0.05.

663

664 **Figure 2. TERRA expression is different depending on gestational age and sex.** A to O, qPCR
665 analysis of the expression of TERRA from subtelomeric regions of chromosome 1, 2, 5, 6, 7, 8, 9, 10,
666 11, 13, 15, 17, 18, 19 and X. The results are expressed as relative TERRA expression normalized to
667 *Sdha* at the different developmental ages (11.5, 12.5 and 13.5 dpc). Columns in red represent female

668 PGCs and blue represent male PGCs Statistical analysis was performed using one-way ANOVA with
669 multiple comparison. p value was set at 0.05
670
671 **Figure 3. TERRA and the catalytic subunit of the enzyme telomerase (TERT) colocalization and**
672 **expression are regulated by gestational age and sex.** A) Super resolution imaging microscopy of
673 TERT (green) immunofluorescence followed by TERRA RNA-FISH (red) on female 13.5 dpc PGCs.
674 Scale bar 10 μm; in the inset the white arrowhead shows TERRA-TERT interaction, yellow arrow head
675 shows only TERRA foci; B) Super resolution imaging microscopy of TERT (green) immunofluorescence
676 followed by TERRA RNA-FISH (red) on male 13.5 dpc primordial germ cells. Scale bar 10 μm; in the
677 inset the white arrowhead shows TERRA-TERT interaction, yellow arrow head shows only TERRA foci;
678 C) Analysis of the percentage of TERRA-TERT colocalization on female and male PGCs. Percentages
679 were obtained from the number of TERRA-TERT foci divided by the total TERRA foci, and multiplied by
680 100. Statistical analysis was performed with unpaired t-test. p value was set at 0.05; D) qPCR analysis
681 of the expression of *Tert* at different stages of development (11.5, 12.5 and 13.5 dpc). The results are
682 expressed as the *Tert* relative expression normalized by *Sdha*. Red columns indicated the analysis
683 performed in female PGCs and blue columns indicate the analysis performed in male PGCs. Statistical
684 analysis was performed using one-way ANOVA with multiple comparisons. p value was set at 0.05
685
686 **Figure 4. The TERRA interactome of the gonad is sexually dimorphic.** A) Analysis of the male and
687 female PGCs protein distribution obtained from iDRiP; B) Heat map analysis of the relative protein
688 concentration in PGCs obtained by iDRiP-mass spec; The relative label free quantification method
689 within Proteome Discoverer 2.2 software was used to calculate the protein abundances. C) Analysis of
690 the male and female somatic protein distribution obtained from iDRiP; D) Heat map analysis of the
691 relative protein concentration in somatic obtained by iDRiP-mass spec. The relative label free
692 quantification method within Proteome Discoverer 2.2 software was used to calculate the protein
693 abundances.

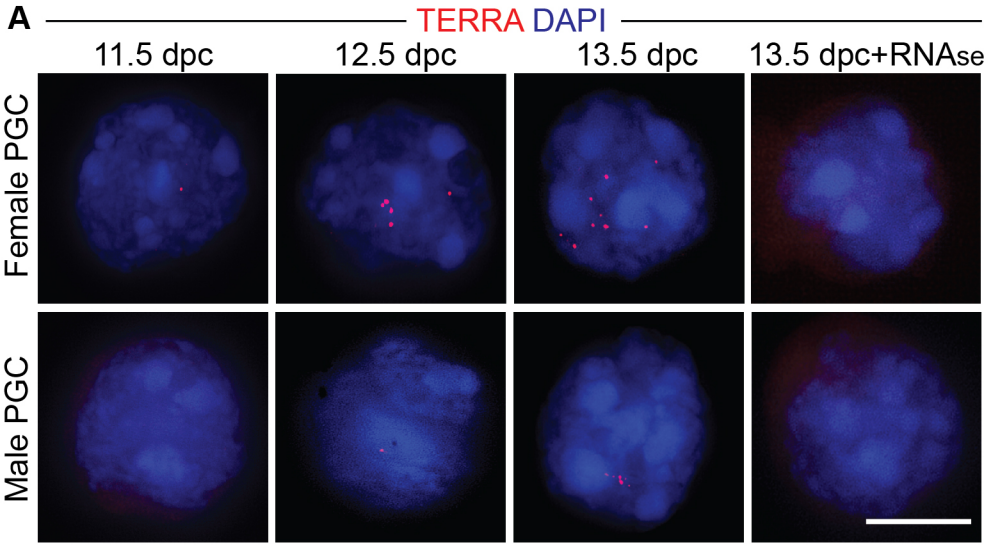
Figure 5. SFPQ interacts TERRA on PGCs. A) Immunofluorescence on 13.5 dpc ovary with antibodies against SFPQ (green), VASA (red) and DAPI. Scale bar 20μ; B) Immunofluorescence on 13.5 dpc testis with antibodies against SFPQ (green), VASA (red) and DAPI. Scale bar 20 μm; C) Super resolution imaging microscopy of SFPQ (green) immunofluorescence followed by TERRA RNA-FISH (red) on female and male 13.5 dpc primordial germ cells. Scale bar 10 μm; in the inset the white arrowhead shows TERRA-SFPQ interaction, yellow arrow head shows only TERRA foci D) Analysis of the percentage of colocalization of TERT-TERRA on male and female PGCs. Percentages were obtained from the number of TERRA-SFPQ foci divided by the total TERRA foci, and multiplied by 100. Statistical analysis was performed using t-test analysis. p value was set at 0.05

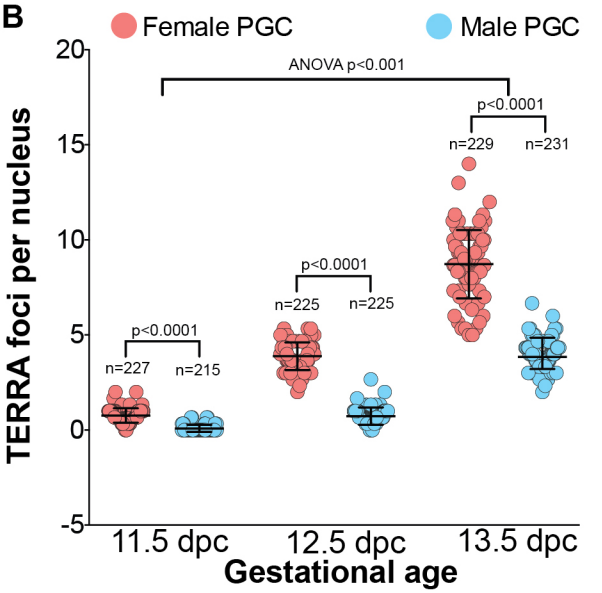
Figure 6. NONO interacts with TERRA on gonadal somatic cells. A) Immunofluorescence on 13.5 dpc ovary with antibodies against NONO (green), VASA (red) and DAPI. Scale bar 20 μm; B) Immunofluorescence on 13.5 dpc testis with antibodies against NONO (green), VASA (red) and DAPI. Scale bar 20 μm; C) Super resolution imaging microscopy of NONO (green) immunofluorescence followed by TERRA RNA-FISH (red) on female and male 13.5 dpc somatic cells. Scale bar 10μm; in the inset the white arrowhead shows TERRA-NONO interaction, yellow arrow head shows only TERRA foci D) Analysis of the percentage of colocalization of TERRA-NONO on male and female somatic cells. Percentages were obtained from the number of TERRA-NONO foci divided by the total TERRA foci, and multiplied by 100. Statistical analysis was performed using t-test analysis. p value was set at 0.05

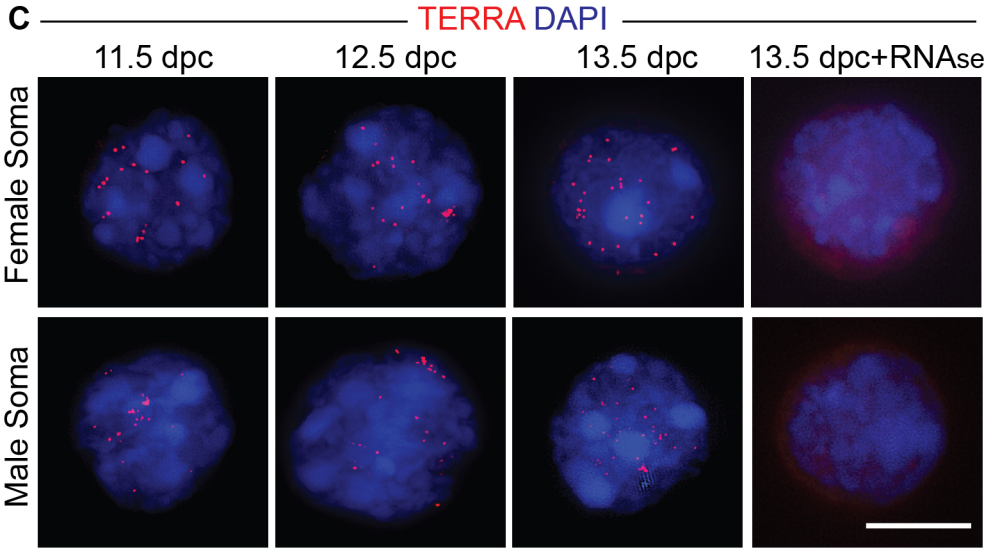
Supplementary Table 1. TERRA and specific subtelomeric chromosome primers

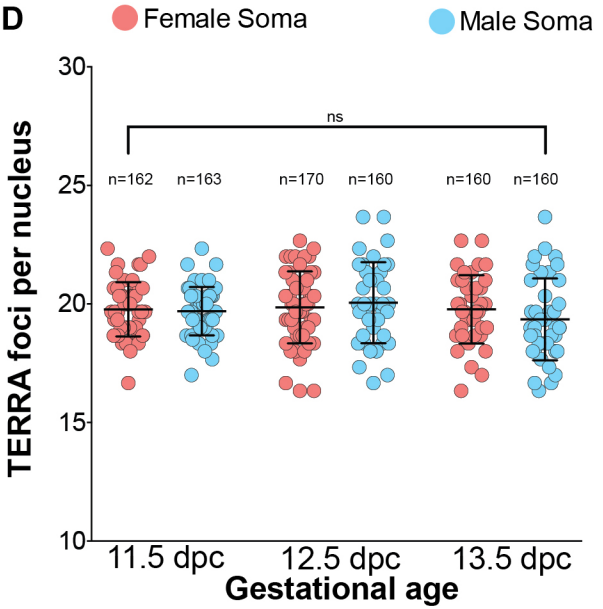
Supplementary Table 2. PGCs protein ratios obtained from iDRiP (m/f)

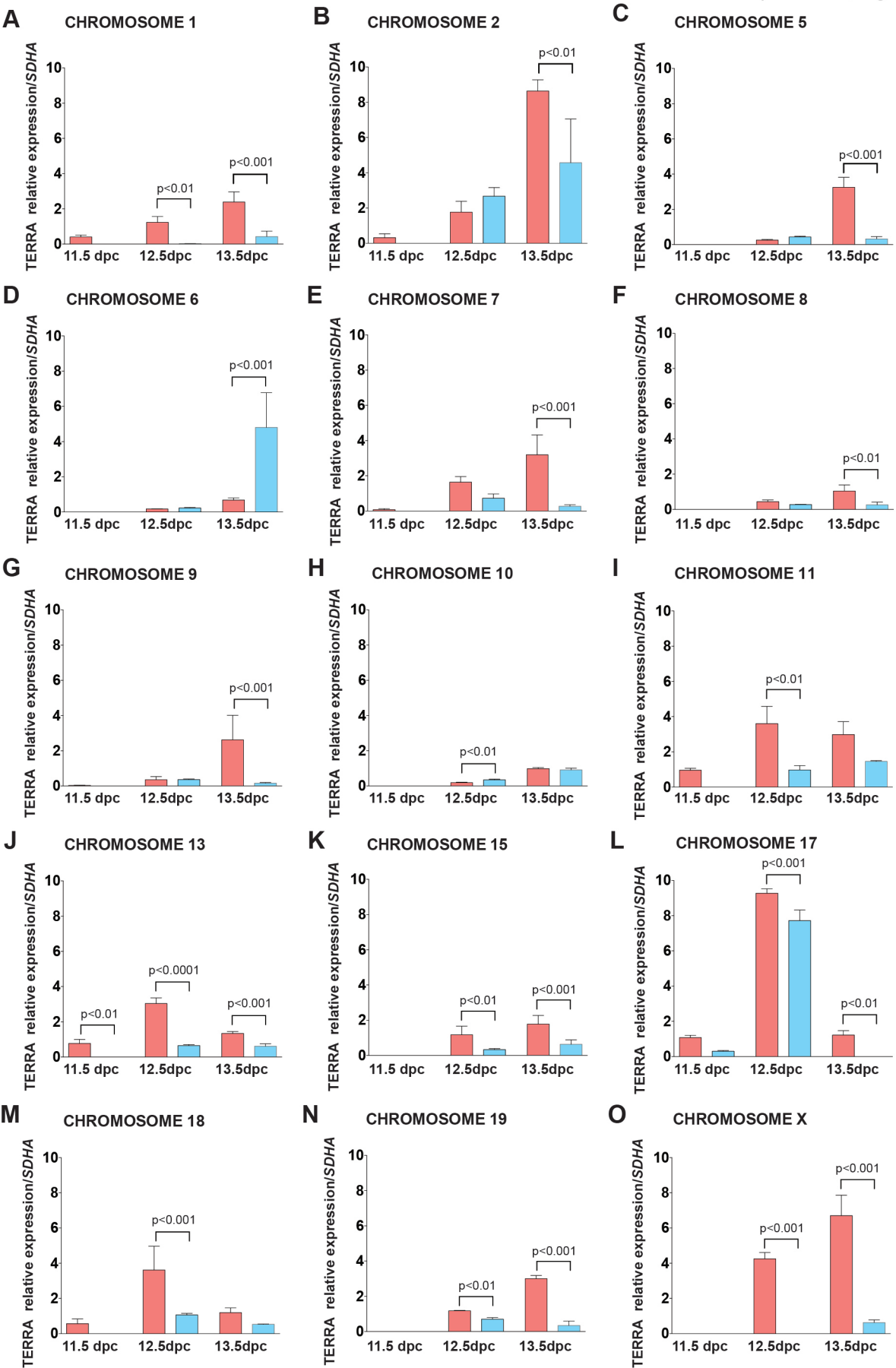
Supplementary Table 3. Somatic cells protein ratios obtained from iDRiP (m/f)

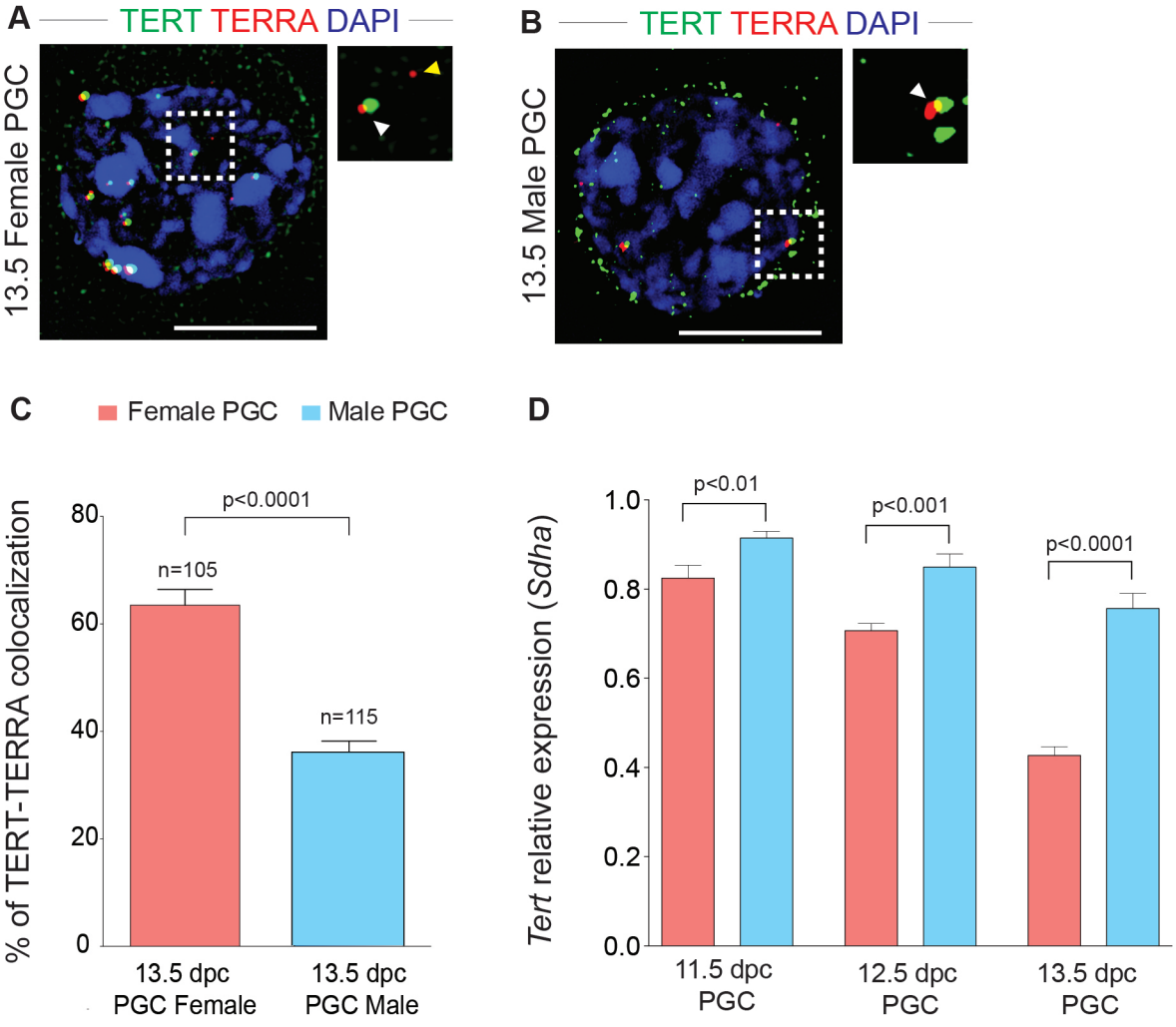

A
TERRA DAPI
11.5 dpc
12.5 dpc
13.5 dpc
13.5 dpc+RNAse
Female PGC
Male PGC

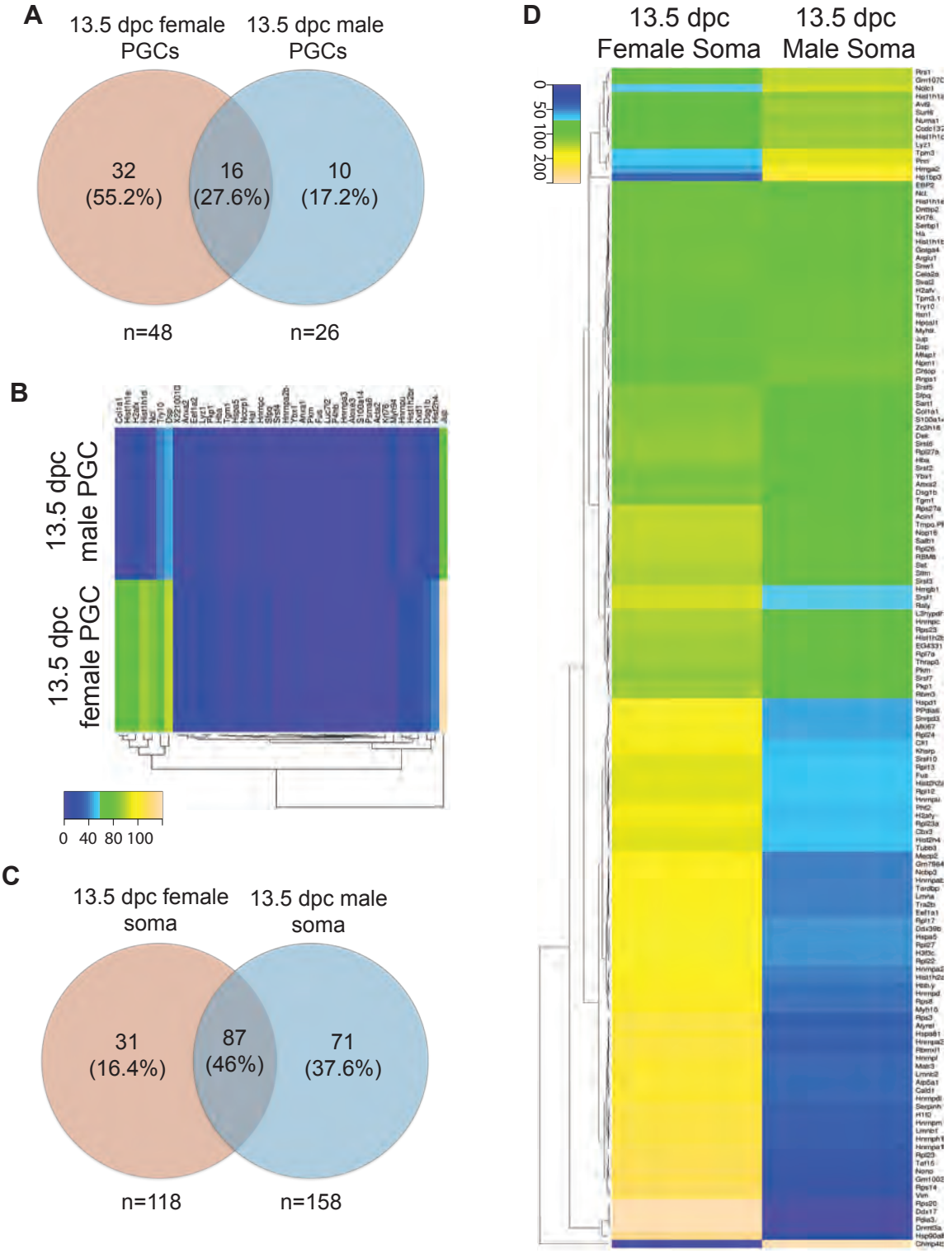

B
Female PGC
Male PGC
ANOVA p<0.001
p<0.0001
p<0.0001
p<0.0001
n=227
n=215
n=225
n=225
n=229
n=231
TERRA foci per nucleus
20
15
10
5
0
-5
11.5 dpc
12.5 dpc
13.5 dpc
Gestational age

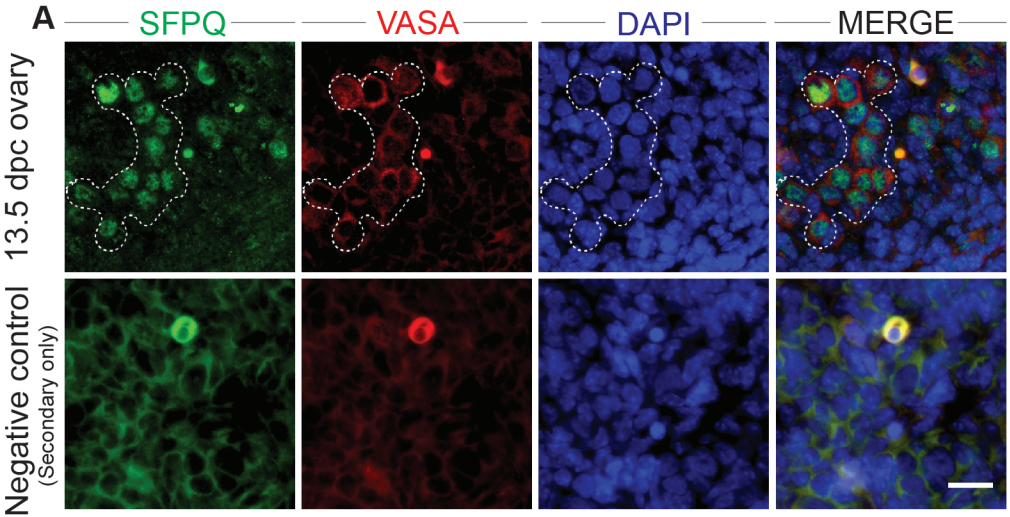

C
TERRA DAPI
11.5 dpc
12.5 dpc
13.5 dpc
13.5 dpc+RNAse
Female Soma
Male Soma

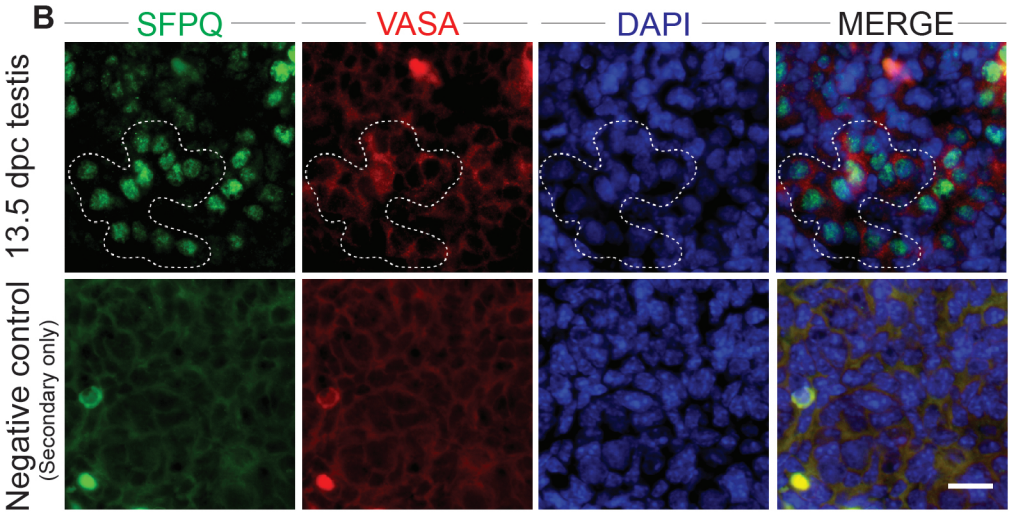

D
Female Soma
Male Soma
ns
n=162
n=163
n=170
n=160
n=160
n=160
TERRA foci per nucleus
30
25
20
15
10
11.5 dpc
12.5 dpc
13.5 dpc
Gestational age

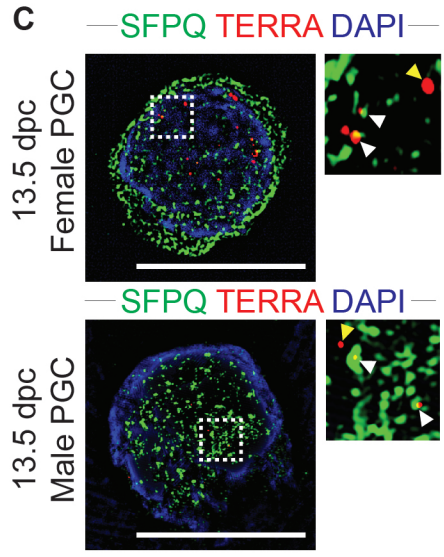

A
CHROMOSOME 1
B
CHROMOSOME 2
C
CHROMOSOME 5
D
CHROMOSOME 6
E
CHROMOSOME 7
F
CHROMOSOME 8
G
CHROMOSOME 9
H
CHROMOSOME 10
I
CHROMOSOME 11
J
CHROMOSOME 13
K
CHROMOSOME 15
L
CHROMOSOME 17
M
CHROMOSOME 18
N
CHROMOSOME 19
O
CHROMOSOME X
TERRA relative expression/SDHA
11.5 dpc
12.5dpc
13.5dpc
p<0.01
p<0.001
p<0.0001

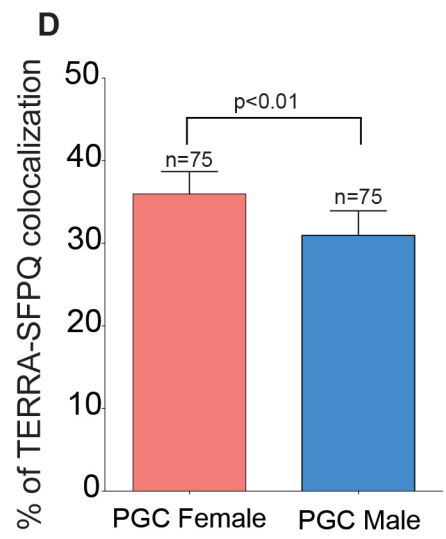

A
TERT TERRA DAPI
13.5 Female PGC
B
TERT TERRA DAPI
13.5 Male PGC
C
Female PGC
Male PGC
% of TERT-TERRA colocalization
p<0.0001
n=105
n=115
13.5 dpc PGC Female
13.5 dpc PGC Male
D
Tert relative expression (Sdha)
p<0.01
p<0.001
p<0.0001
11.5 dpc PGC
12.5 dpc PGC
13.5 dpc PGC

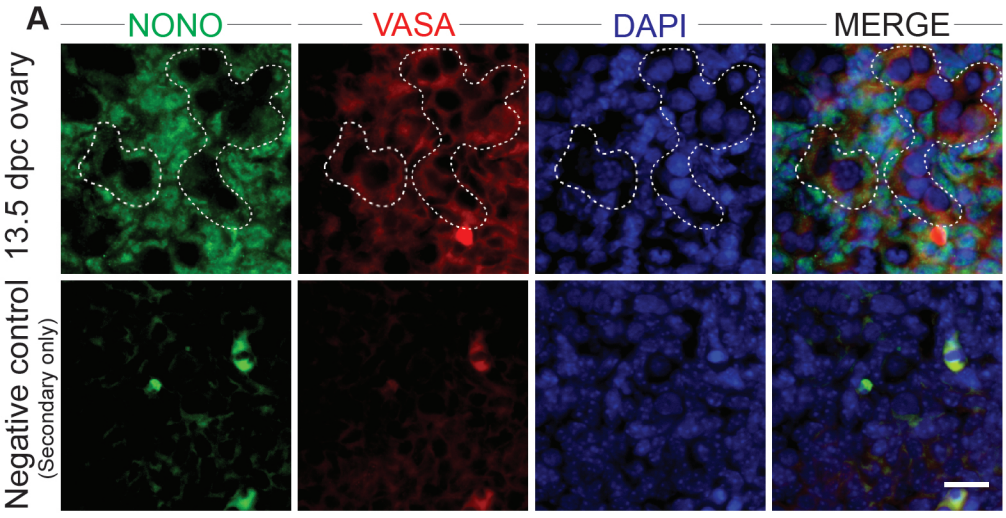

A
13.5 dpc female PGCs
13.5 dpc male PGCs
32 (55.2%)
16 (27.6%)
10 (17.2%)
n=48
n=26
B
13.5 dpc male PGC
13.5 dpc female PGC
0 40 80 100
C
13.5 dpc female soma
13.5 dpc male soma
31 (16.4%)
87 (46%)
71 (37.6%)
n=118
n=158
D
13.5 dpc Female Soma
13.5 dpc Male Soma
0 50 100 200

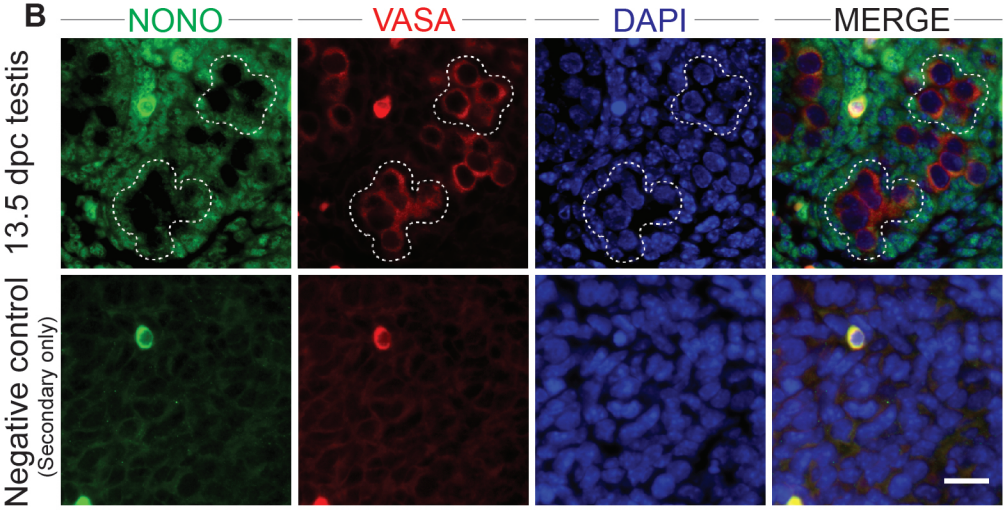

A
SFPQ
VASA
DAPI
MERGE
13.5 dpc ovary
Negative control
(Secondary only)

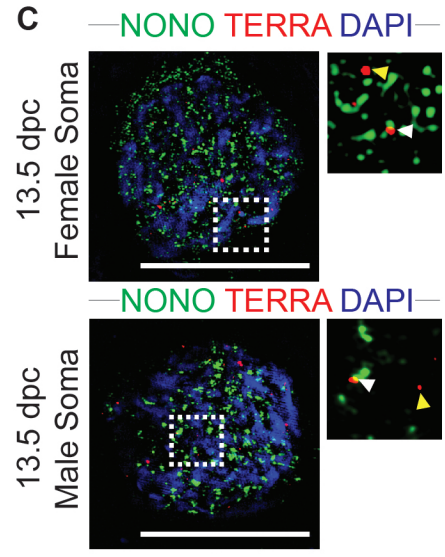

B
SFPQ
VASA
DAPI
MERGE
13.5 dpc testis
Negative control
(Secondary only)

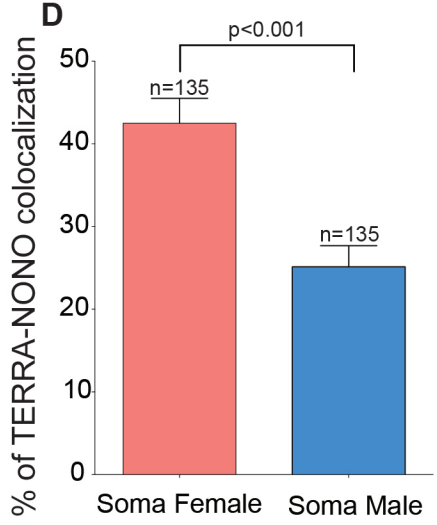

C
SFPQ TERRA DAPI
13.5 dpc
Female PGC
SFPQ TERRA DAPI
13.5 dpc
Male PGC


D
% of TERRA-SFPQ colocalization
50
40
30
20
10
0
p<0.01
n=75
n=75
PGC Female
PGC Male


A
NONO
VASA
DAPI
MERGE
13.5 dpc ovary
Negative control
(Secondary only)


B
NONO
VASA
DAPI
MERGE
13.5 dpc testis
Negative control
(Secondary only)


C
NONO TERRA DAPI
13.5 dpc
Female Soma
NONO TERRA DAPI
13.5 dpc
Male Soma


D
p<0.001
n=135
n=135
% of TERRA-NONO colocalization
50
40
30
20
10
0
Soma Female
Soma Male

SUPLEMENTARY TABLE 1. PRIMERS USED IN THIS STUDY.

TERT
FWD TGAAAGTAGAGGATTGCCACTG
REV CCTCAGACGGTGCTCTGC

18S
FWD GCAATTATTCCCCATGAACG
REV GGGACTTAATCAACGCAAGC

Chromosome 18
L-qPCR-Chr18 GCAAGCTCCGAAGTTGTGAT
R-qPCR-Chr18 CGCTTCTGTGAAGGATCCAG

TERRA
REV CCCTAACCCTAACCCTAACCCTAA

U6:
FWD GCATGACGTCTGCTTTGGA
REV primer 5´-CCACAATCATTCTGCCATCA

L-Chr1-1 ATACACCCGGTTTGGATG
R-Chr1-1 CTTGTGGCTCTTGGTATCTT

L-Chr1-2 ACAAGCTGACCTGTGAGTTC
R-Chr1-2 ATCTGGGACAAAGGCAAG

L-chr2_6 GAGTGCCTTACTATCTCCTAAGTTTT
R-chr2_6 TGGAGTTAATTTTGTGGAGGTTG

L-Chr2-13 CGAGTCTCAGAAAAGGGCAG
R-Chr2-13 TCACCAATTCTCCAACAGGG

L-Chr5-7 GAAGACTGAGGCAGAAGAGTT
R-Chr5-7 TGAGATCCCTTGCTACTTGC

L-Chr5-8 CTGAACCTTGTAGTTCCTCTG
R-Chr5-8 CTGTGAAAGGGTCATTCAAC

L-chr6_3 AGGTGTTCTGAGGCAAGCTC
R-chr6_3 AAGACCCACCACACCAGTTC

L-Chr6-4 CTGCAATTGGGAAGAGCA
R-Chr6-4 CAGAGGGTTTCTCTCCAGTA

L-Chr7-1 GGGCTTTGTTATCTTGCTG
R-Chr7-1 AGAGTCATGTGGCTATCAGTC

L-Chr7-2 TAAGGACACAGGCCATACAC
R-Chr7-2 CCACTCAGCTCAGACATACA

L-Chr9-1 GCATGCTCATTGAAGACCAG
R-Chr9-1 ATTGGTCTCTGAGGTCTTGC

L-Chr9-2 TGCCTCTCAAGTGCTGTT
R-Chr9-2 GTAGGCATTGTGTCAGTCTCA

L-Chr11-7 TGTATTTGAAATTGATTATGAGAAGT
R-Chr11-7 TAAAGCCTTTGTACATCAATATG

L-Chr11-8 GGCTTTACCACATTGACTACA
R-Chr11-8 CTTTGCTTCTCCTGCTCAAC

L-Chr12-1 TTTGGCTACTGAGCAACTG
R-Chr12-1 CCAGAAACGTTCCATACTAAC

L-Chr12-2 CAAGCTTTGTTCCCTCAG
R-Chr12-2 AAAGCAATTTAAGATTTGTAGCA

L-Chr13-4 CAATGAAACTCCACCCACAG
R-Chr13-4 TCTGCCTCACCATCTGTGTA

L-Chr13-5 CCAAGACCCAGCTTCTGATTC
R-Chr13-5 TTTTGATGGTGTCTCTTGGGA
L-Chr14-1 GAGCTCAGATGCATCACTGT
R-Chr14-1 TTGATGAACAAAGGTCACTG

L-Chr14-2 TACCACCTACTAGCCTCCAGT
R-Chr14-2 GAGTATGACCTAACCCTGACTC

L-Chr15-1 CAGAATGGGTCAAGGATAGC
R-Chr15-1 CTCAAGTGCTGAAGGATGAC

L-Chr15-2 CTTCCCTTCCCTCTTCTGTA
R-Chr15-2 TGGGTTAGACCTTATCTGAGTT

L-Chr16-1 TGTAATTGGAGAGTACTATGATATGC
R-Chr16-1 ATTACACATACTGGAGAGAAATCTTAC

L-Chr16-2 CATGTGTAAGGTTTCTCTCCAGT
R-Chr16-2 CAATGTAGTAAATCCTTTGCTTCTC

L-Chr19-1 GTCTCCAACATCACATCTCTG
R-Chr19-1 GATGTGCATGTGAACTTTACTC

L-Chr19-2 CTTACATTCATAGGGTTTCTCTCTAGT
R-Chr19-2 GTGGTAAAGCCTTTGCATATCATAC

L-ChrX-1 GAGGTTCCTGTAAGTCTCCA
R-ChrX-1 CCTATGATGATGTGCATGTG

L-ChrX-2 GTTACAGAGAACTCGCATCTTC
R-ChrX-2 GGGCAATACCCTTTGTCC

L_Chr8-1 TCAATGCATCAGCAGCAG
R_Chr8-1 GGTAATGTTGCCTGATGAGC

L-Chr8-B ATGAGCAGGTAATGAACTCTG
R_Chr8-B GTGTGGGAGGATAGTCATGTA

L_Chr10-seg1 GTCTCAAGTGAAACAGACTGC
R_Chr10-seg1 GCAGTTCCAGAAAGATCACTG
L_Chr10-seg2 TCAGCAAATCATGGTTCAGAT

R_Chr17_1 TCTCAAGTGAAACAGACTGC
L-_Chr17EST TGTCAAAGTTCCAGAAAACATGG
R-Chr17EST CTTTTGGGGATGACTGTGACAT

L-qPCR-Chr18 GCAAGCTCCGAAGTTGTGAT
R-qPCR-Chr18 CGCTTCTGTGAAGGATCCAG

L-qPCR-Chr8 TCCCACTGTCAATAACAGAC
R-qPCR-Chr8 CAAGCACAGGCTAGAAGTG

L-qPCR-Chr10 TCAGCAAATCATGGTTCAGAT
R-qPCR-Chr10 TGCATTGCATTTGACAACAG

L-qPCR_Chr17 TGTCAAAGTTCCAGAAAACATGG
R-qPCR_Chr17 CTTTTGGGGATGACTGTGACAT

L-Jarid1d CTGAAGCTTTTGGCTTTGAG
R-Jarid1d CCACTGCCAAATTCTTTGG

L-SRY GAGAGCATGGAGGGCCATG
R-SRY GAGTACAGGTGTGCAGCTC

L-Sry2 CTCTGAAGAAGAGACAAGTT
R-Sry2 CTGTGTAGGATCTTCAATC

SUPPLEMENTARY TABLE 2. PGC PROTEIN RATIOS FROM IDRIP (m/f)

Checked	Protein FDI	Accession	Description	Sum PEP Sc	Coverage [%	# Peptides	# PSMs	# Unique P	# Protein G	# AAs	MW [kDa]	calc. pI	Score Sequ	# Peptides	Abundance Ratio: (F32, Sa
TRUE	High	P04104	Keratin, type II cytoskeletal 1 OS=Mus musculus GN=Krt1 PE=1 SV=4	37.142	7	8	94	1	1	637	65.6	8.15	212.15	8	
TRUE	High	E9Q557	Desmoplakin OS=Mus musculus GN=Dsp PE=1 SV=1	35.364	5	12	24	12	1	2883	332.7	6.8	61.43	12	0.648
TRUE	High	Q02257	Junction plakoglobin OS=Mus musculus GN=Jup PE=1 SV=3	29.84	21	12	21	12	1	745	81.7	6.14	58.11	12	0.469
TRUE	High	Q3UV17	Keratin, type II cytoskeletal 2 oral OS=Mus musculus GN=Krt76 PE=1 SV=1	23.5	12	10	31	1	1	594	62.8	8.43	74.03	10	4.188
TRUE	High	P09405	Nucleolin OS=Mus musculus GN=Ncl PE=1 SV=2	20.188	17	9	9	9	1	707	76.7	4.75	26.79	9	0.01
TRUE	High	P43274	Histone H1.4 OS=Mus musculus GN=Hist1h1e PE=1 SV=2	20.005	18	8	11	2	1	219	22	11.11	29.87	8	0.047
TRUE	High	Q149Z9	Histone H1.3 OS=Mus musculus GN=Hist1h1d PE=1 SV=1	18.988	19	9	12	3	1	221	22.1	11.03	30.29	9	0.108
TRUE	High	P97350	Plakophilin-1 OS=Mus musculus GN=Pkp1 PE=1 SV=1	10.687	7	3	6	3	1	728	80.8	8.91	21.96	3	0.499
TRUE	High	P11087	Collagen alpha-1(I) chain OS=Mus musculus GN=Col1a1 PE=1 SV=4	9.173	3	3	11	3	1	1453	137.9	5.85	33.95	3	0.027
TRUE	High	P07356	Annexin A2 OS=Mus musculus GN=Anxa2 PE=1 SV=2	7.363	12	3	4	3	1	339	38.7	7.69	11.22	3	0.424
TRUE	High	B2RTM0	Histone H4 OS=Mus musculus GN=Hist2h4 PE=1 SV=1	7.197	21	2	4	2	1	103	11.4	11.36	11.04	2	0.205
TRUE	High	Q792Z1	MCG140784 OS=Mus musculus GN=Try10 PE=1 SV=1	6.316	15	2	4	2	1	246	26.2	5.83	12.05	2	0.569
TRUE	High	P62960	Nuclease-sensitive element-binding protein 1 OS=Mus musculus GN=Ybx1 PE=1 SV=3	6.274	20	2	2	2	1	322	35.7	9.88	7.24	2	0.01
TRUE	High	B2RXW1	Histidine ammonia-lyase OS=Mus musculus GN=Hal PE=1 SV=1	5.71	3	2	2	2	1	657	72.2	6.34	5.98	2	0.598
TRUE	High	P20029	78 kDa glucose-regulated protein OS=Mus musculus GN=Hspa5 PE=1 SV=3	5.104	4	2	3	2	1	655	72.4	5.16	10.52	2	0.265
TRUE	High	A0A0A0MC	F-box only protein 50 OS=Mus musculus GN=Nccrp1 PE=1 SV=1	4.933	5	2	2	2	1	291	33	6.89	7.74	2	0.397
TRUE	High	P62737	Actin, aortic smooth muscle OS=Mus musculus GN=Acta2 PE=1 SV=1	4.648	7	2	2	1	1	377	42	5.39	5.95	2	100
TRUE	High	B2RQH0	Desmoglein 1 beta OS=Mus musculus GN=Dsg1b PE=1 SV=1	4.449	2	2	4	2	1	1060	114.4	4.84	14.92	2	0.411
TRUE	High	Q8VIJ6	Splicing factor, proline- and glutamine-rich OS=Mus musculus GN=Sfpq PE=1 SV=1	4.371	4	2	2	2	1	699	75.4	9.44	5.86	2	0.01
TRUE	High	P27661	Histone H2AX OS=Mus musculus GN=H2afx PE=1 SV=2	3.495	11	2	3	2	1	143	15.1	10.74	7.44	2	0.198