SUPPLEMENTAL FIGURES

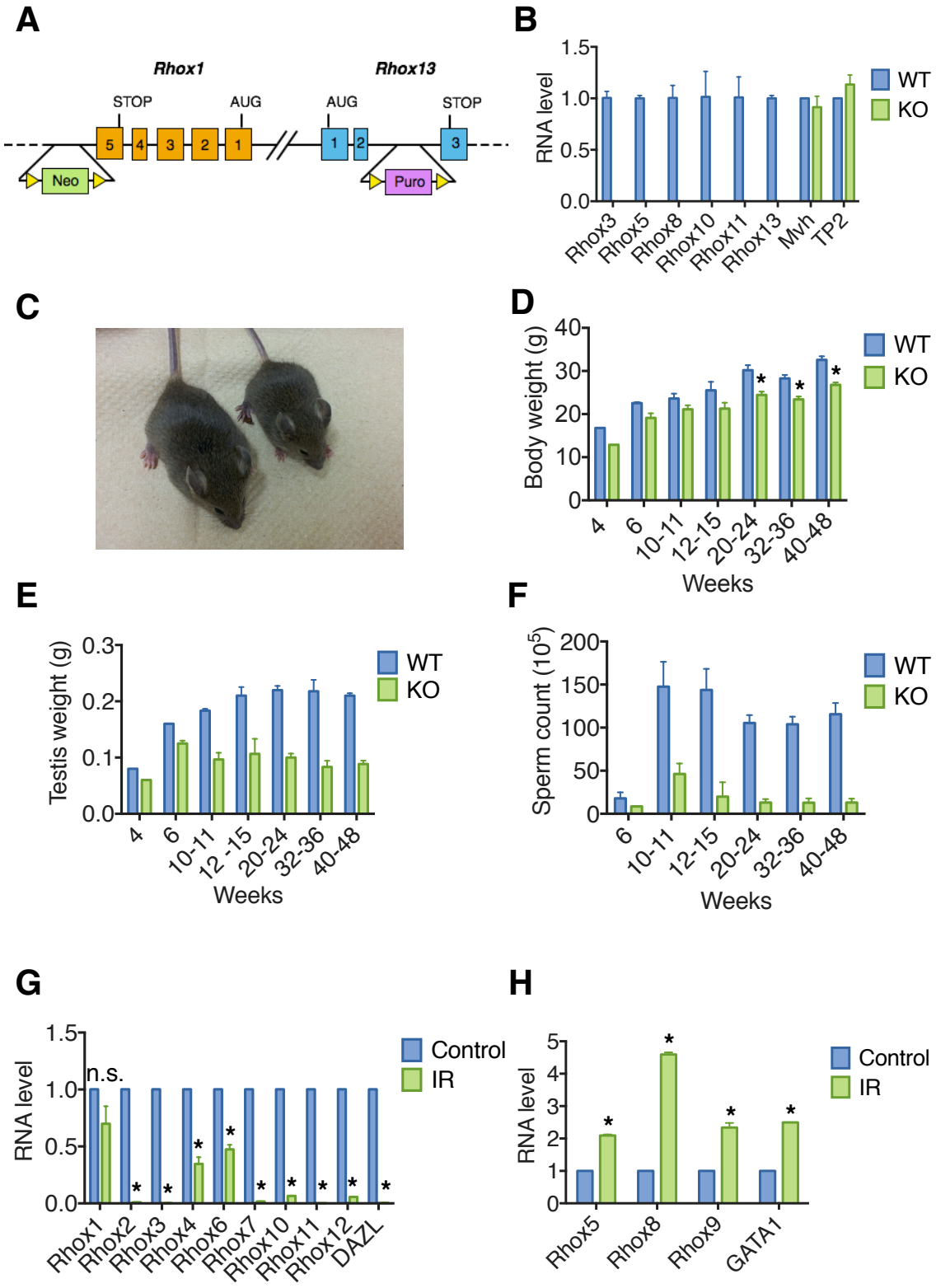

Figure S1. Generation and Phenotypic Analysis of *Rhox-c*-KO Mice, Related to Figure 1

(A) Two independent targeting vectors were used to introduce the indicated loxP sites (yellow arrows) at the beginning and end of the ~920-kb *Rhox* cluster.
(B) qRT-PCR analysis demonstrating loss of Rhox gene expression in testes from 6 week-old Rhox-c-KO mice. WT mice are littermate control mice. Mvh and Tp2 serve as germ cell markers. Values were normalized to Rpl19 mRNA level and denote the mean fold change ± standard error of the mean (SEM).
(C) Body size comparison of a *Rhox-c*-KO (right) and littermate control (WT) mouse (left) at weaning age.
(D) Body weight of *Rhox-c*-KO (KO) and littermate control (WT) mice of the indicated ages.
(E) Testis weight of *Rhox*-c-KO and littermate control (WT) mice of the indicated ages.
(F) Epididymal sperm count of *Rhox*-c-KO and littermate control (WT) mice of the indicated ages
(G, H) qRT-PCR analysis of testes from irradiated adult mice (IR) and non-irradiated mice (control). Testes were collected 5 weeks after abdominal irradiation as previously described (Zhang et al., 2006). Germ cell depletion was confirmed by histological analysis. *Dazl* and *Gata1* serve as a germ cell marker and a somatic cell marker, respectively. Values were normalized to *Rpl19* mRNA level and denote the mean fold change ± standard error of the mean (SEM). Asterisks indicate the difference is statistically significant ($P<0.05$).

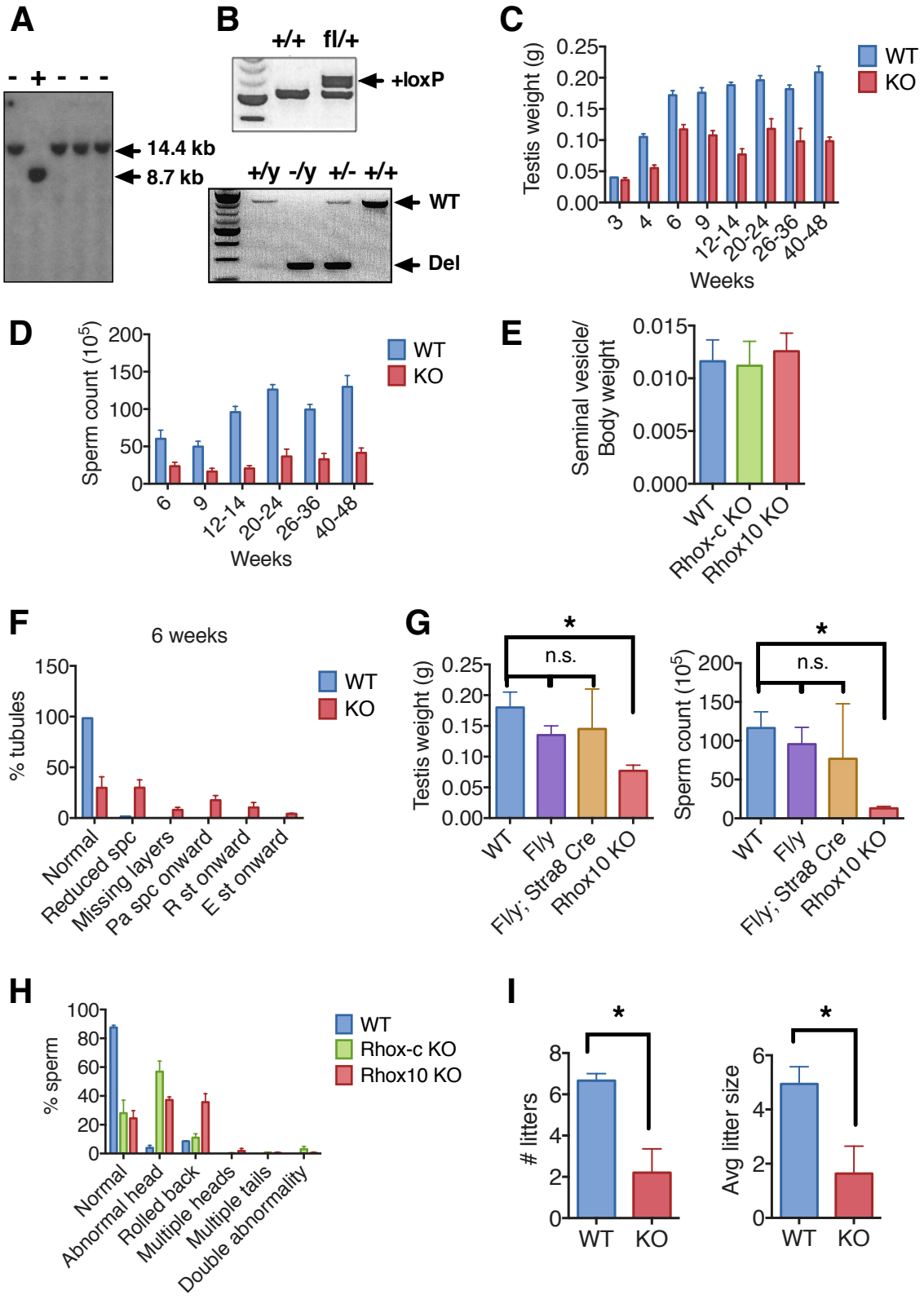

Figure S2. Generation and Phenotypic Analysis of *Rhox10*-KO Mice, Related to Figure 2

(A) Southern blot analysis of genomic DNA (cut with Bst*EII*) from ES cell clones that shows a positive cell clone that has the floxed *Rhox10* allele. The 14.4-kb band is the normal *Rhox10* allele, while the 8.7-kb band is indicative of recombination with the targeting vector and incorporation of the floxed allele.
(B) PCR analysis of genomic DNA indicating germline transmission of the floxed *Rhox10* allele (upper) and deletion of *Rhox10* exon 2 in $Rhox10^{-/y}$ male mice and $Rhox10^{+/-}$ female mice (lower).
(C) Testis weight of *Rhox10*-KO (KO) and littermate control (WT) mice of indicated ages (n=2-16 per genotype for each time points). All values are mean ± SEM.
(D) Epididymal sperm count of *Rhox*-c-KO and littermate control (WT) mice of the indicated ages (n=2-16 per genotype for each time points). All values are mean ± SEM.
(E) Seminal vesicle weight of adult *Rhox-c*-KO, *Rhox10*-KO, and WT mice (n=13-18).
(F) Seminiferous tubule abnormalities in 6 week-old *Rhox10*-KO (KO) and littermate control (WT) mice. spc, spermatocytes; pa spc, pachytene spermatocytes; R st, round spermatids; E st, elongating spermatids.
(G) Testis weight and epididymal sperm count of $Rhox10^{fl/y}$ (Fl/y), $Rhox10^{fl/y}$; Stra8-Cre (Fl/y; Stra8 Cre), *Rhox10*-KO (KO), and control (WT) mice at 12 weeks of age (n=2-10). All values are mean ± SEM. Asterisks denote statistically significant differences ($P<0.05$).
(H) Sperm abnormalities in the indicated mouse strains (3- to 8-months old). Values denote the mean percentage of sperm ± standard error of the mean (SEM). n=2-3, 80-180 sperm per mouse.
(I) Fertility analysis of adult *Rhox10*-KO and littermate control (WT) male mice, each housed with a single BL6 female mouse (initially 8-weeks old) for 7 months. Values denote the mean ± standard error of the mean (SEM). Asterisks indicate that the difference is statistically significant ($P<0.05$).

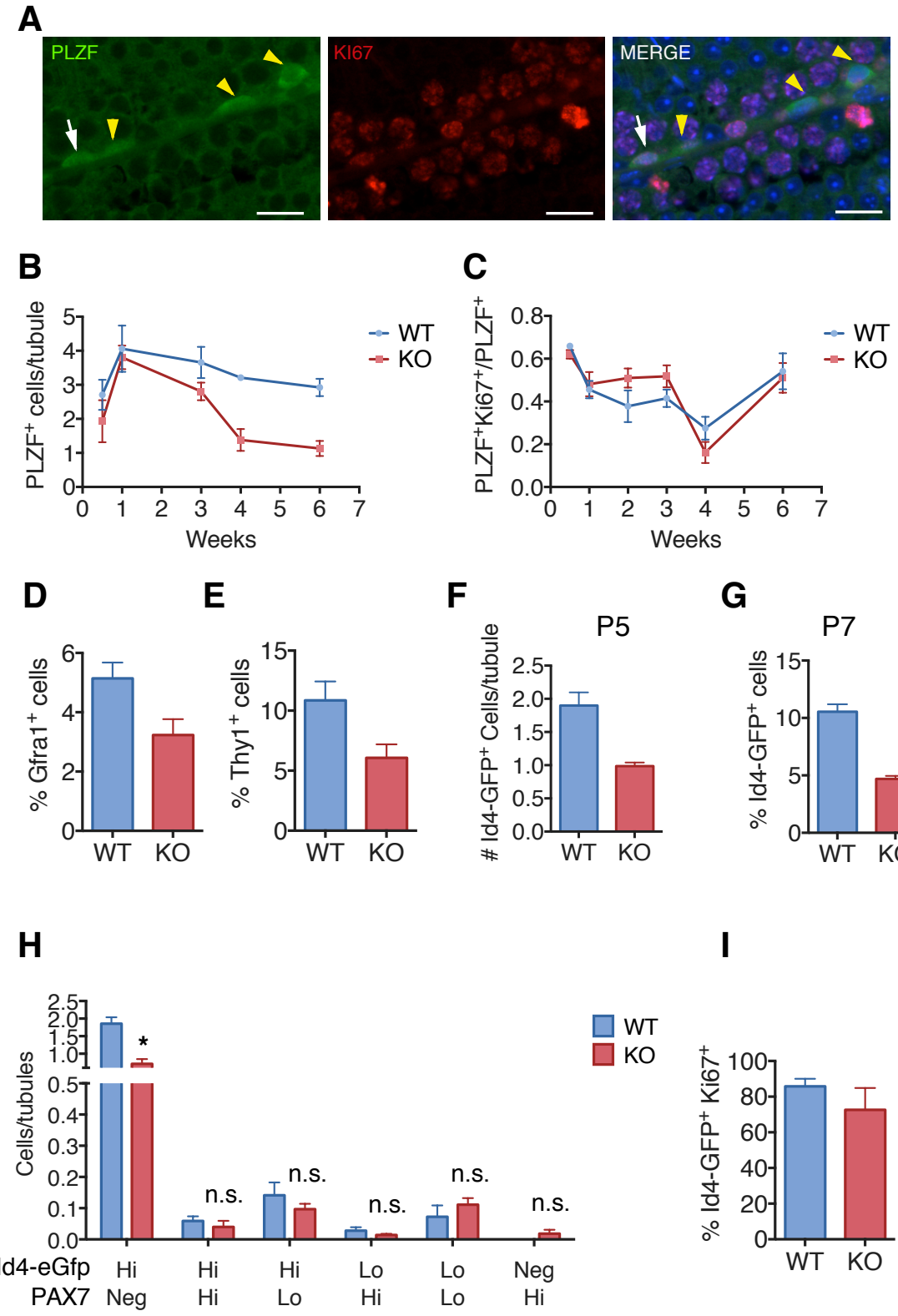

Figure S3. *Rhox10*-KO Mice have a Selective Defect in Cells Labeled with Most SSC Markers, Related to Figure 3

(A) Representative image of 6 week-old adult testes sections stained for the undifferentiated spermatogonial marker, PLZF, and the proliferation marker, KI67. The white arrow indicates a PLZF/KI67

double-positive cell and the yellow arrowheads indicate PLZF single-positive cells. The merged image also shows DAPI staining to indicate location of cells. Scale bars = 20 μ m.

(B, C) Quantification of cells analyzed as in (A) from mice of the indicated ages.

(D, E) Flow cytometric analysis of the percentage GFRα1-positive cells (D) or Thy1-positive cells (E) among all testicular cells from 8 day-old *Rhox10*-KO and littermate control (WT) mice (n=3-5).

(F) Quantification of Id4-eGfp-positive cells per tubule in testes sections from P5 *Rhox10*-KO mice (*Rhox10*$^{-/y}$*;Id4-eGfp*) mice (n=3) and littermate control (*Rhox10*$^{+/y}$*;Id4-eGfp*) mice (n=2). Mice with the same genotype were used for panels G-H.

(G) Flow cytometric analysis of the percentage of Id4-eGfp-positive cells among all testicular cells from 7 day-old *Rhox10*-KO (*Rhox10*$^{-/y}$*;Id4-eGfp*) mice (n=3-5) and littermate control (*Rhox10*$^{+/y}$*;Id4-eGfp*) mice (n=5).

(H) Quantification of Id4-eGfp- and/or PAX7-positive cells in testes sections from P5 *Rhox10*-KO mice (n=3) and littermate control (WT) mice (n=2). Cells were judged as exhibiting high (Hi), low (Lo), or no detectable (Neg) expression.

(I) The percentage of KI67-positive cells among all Id4-eGfp-positive cells in testes sections from P5 *Rhox10*-KO mice (n=3) and littermate control (WT) mice (n=2).

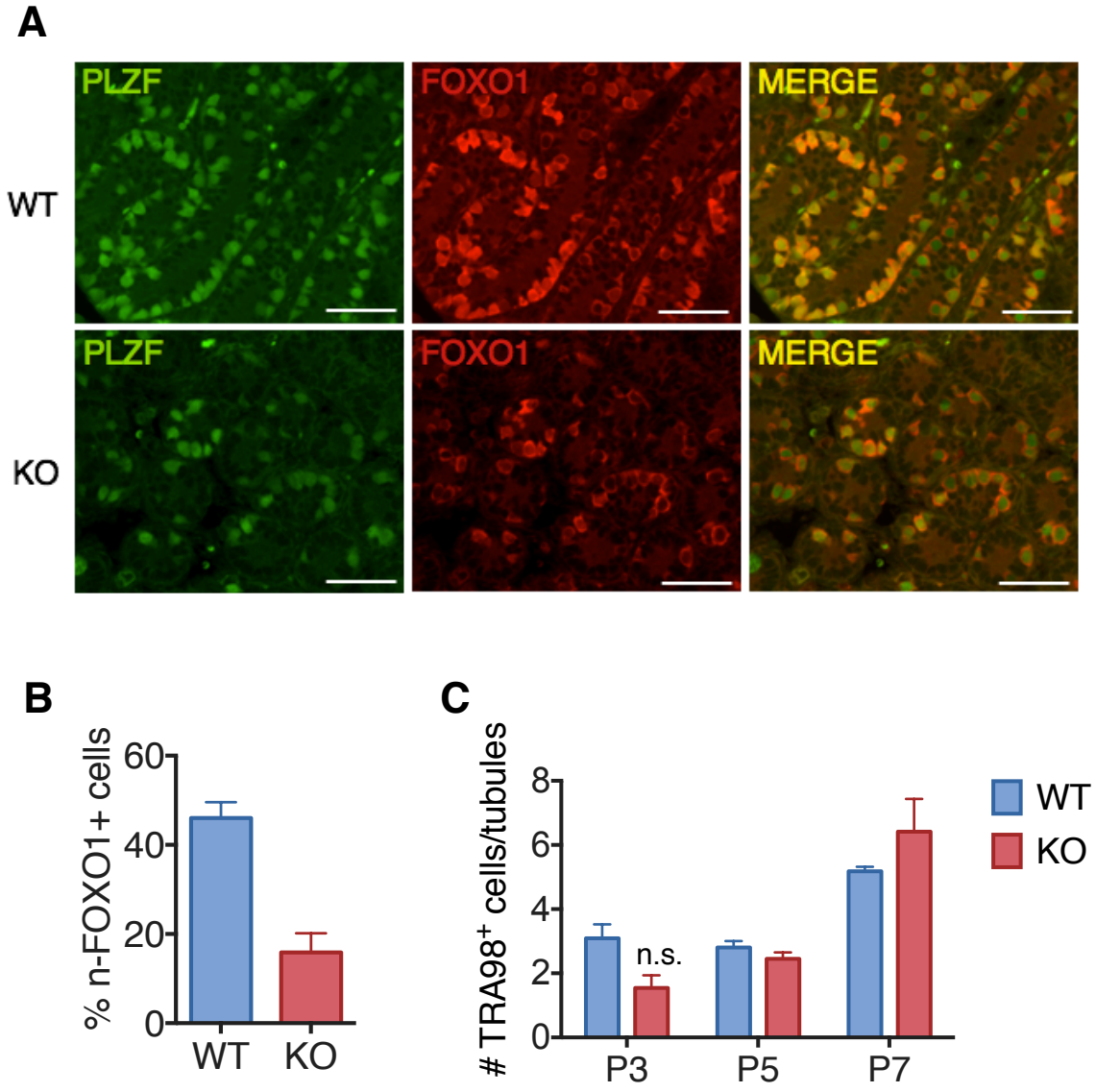

Figure S4. Additional Evidence that *Rhox10* Promotes ProSG-to-SSC Progression, Related to Figure 4

(A) Representative images of PLZF and FOXO1 double-stained testes sections from P5 *Rhox10*-KO mice and littermate control (WT) mice. Scale bars = 60 μ m

(B) Percentage of nuclear n-FOXO1/PLZF double-positive cells among total FOXO1/PLZF double-positive cells in P5 *Rhox10*-KO mice and littermate control mice (n=3) testes sections. All values are mean ± SEM.

(C) Quantification of TRA98-positive cells per tubule in testes sections from *Rhox10*-KO and littermate control (WT) mice of the indicated ages (n=2-5, each time point).

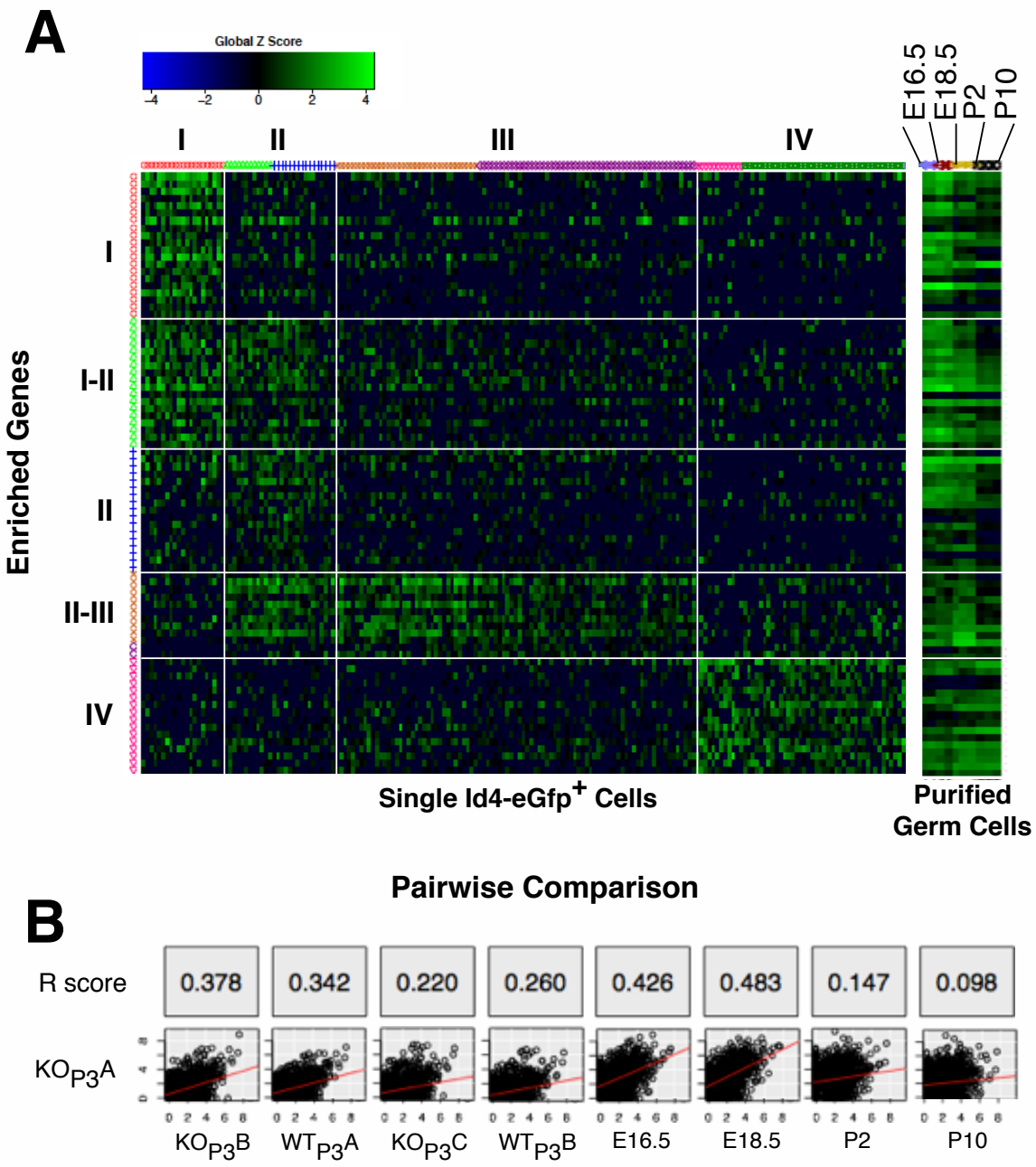

Figure S5. Identification of Id4-eGfp-Positive Germ Cell Clusters, Related to Figure 5

(A) (Left) Heat map depicting gene expression in individual Id4-eGfp-positive cells from the cell clusters described in Figure 5B. x axis=red, $KO_{P3}A$; light green, $KO_{P3}B$; blue, $WT_{P3}A$; orange, $KO_{P3}C$; purple, $WT_{P3}B$; green, $WT_{P7}A$; pink, $WT_{P7}B$. The y axis shows a maximum of 20 genes exhibiting enriched expression in the indicated cell cluster (See also Tables S2 and S3). (Right) Heat map depicting gene expression in purified germ cells from the indicated time points (determined by RNAseq [Pastor et al., 2014]).

(B) Pairwise comparison of gene expression profiles of the $KO_{P3}A$ cell subset with that of (i) the other cell subsets defined herein, or (ii) purified germ cells from the time points indicated (Pastor et al., 2014). The genes examined were the 1009 differentially expressed genes described in Figure 7A.

A

RHOX10-positively regulated genes

	GO	Fold enrichment
ProSG subset	**Trabecula formation;** *Ovol2, Fhl2, Mmp2*	14.308
	Phospholipid transport; *Atp10a, Atp11a, Atp10d, Abca1*	9.124
	Developmental maturation; *Kcnma1, Gata2, Sept4, Epas1, Xbp1, Mmp2, Trip13*	3.636
	Regulation of cell motion; *Selp, Cdkn1b, Lama5, Pecam1, Abi3, Hbegf, Itgb3*	3.432
	Kidney development; *Lama5, Myo1e, Ilk, Pdgfrb, Anxa4, Tenc1, Gzf1*	3.432
	Vasculature development; *Mef2c, Aldh1a2, Epas1, Ovol2, Myo1e, Hbegf, Zc3h12a, Anpep, Sox18, Nos3, Mmp2, Epha2, Vash1, Hdac7*	2.938
	Tissue morphogenesis *Aldh1a2, Ddr1, Bcl10, Wnt3, Ovol2, Lama5, Ilk, Fzd3, Epha2, Etv4, Gzf1, Nkx2-3*	2.645
SSC subset	**Melanocyte differentiation;** *Myo5A, Mitf, Hps1, Mreg, Cited1*	6.347
	Polyol metabolic process; *Galk1, Impa1, Got1, Dgat2, Dak, Ppip5k2, Inpp5a, Gdpd1*	3.779
	Positive regulation of locomotion; *Col18a1, Coro1a, Il16, Ptp4a1, Hbegf, Cbll1, Pik3r1*	3.231
	Regulation of myeloid cell differentiation; *Zfp36, Ankrd54, Tesc, Ikzf1, Id2, Mitf, Jag1, Pik3r1*	3.186
	Steroid biosynthetic process; *Stard3, Cyb5r1, Stard5, Cyp39a1, Mvd, Stard6, Hsd17b1, Hsd17b3, Cftr, Sigmarl1*	2.861
	Response to hypoxia; *Kcnma1, Nol3, Epas1, Il18, Edn1, Pml, Ryr2, Trf, Cited2*	2.856
	Vitamin metabolic process; *Aldh1a1, Aldh1a2, Nampt, Rbp4, Rdh10, Ncf2, Folr1, Rfk, Cyr2r1*	2.649
	Regulation of cell adhesion; *Icam1, Ddr1, Tesc, Egflam, Pik3cb, Kifap3, Il18, Cbll1, Ecm2, Pik3r1, Cyr61, Cited2*	2.593

B

RHOX10-negatively regulated genes

	GO	Fold Enrichment
ProSG subset	**Base-excision repair;** *Mutyh, Neil3, Ung, Rad51l3*	21.509
	Negative regulation of kinase activity; *Prkca, Dusp19, Mllt1, Cdk5rap1*	8.013
	Placenta development; *Fgfr2, Plk4, Ccnf, Gab1, Plcd1*	5.872
	Response to DNA damage stimulus; *Mutyh, Aen, Neil3, Mre11a, Ung, Rad18, Rad51l3, Chaf1b, Atrip*	3.204
	DNA metabolic process; *Peo1, Ccne1, Mutyh, Neil3, Mre11a, Ung, Rad18, Cidea, Rad51l3, Parp4, Chaf1b, Atrip*	2.912
	Cell division; *Cdc7, Ccne1, Ccnb3, Usp9x, Ccnf, 1700017B05rik, Zbtb16, 2610039C10rik*	2.909
	Cellular response to stress; *Dusp19, Mutyh, Aen, Neil3, Mre11a, Ung, Gab1, Rad18, Rad51l3, Chaf1b, Atrip*	2.782

Figure S6. GO analysis of RHOX10-Regulated Genes in Id4-eGfp-Positive Germ Cell Subsets Defined by SC-RNAseq Analysis, Related to Figure 6

(A, B) Significantly enriched GO categories corresponding to genes regulated by RHOX10 in the cell subsets shown (defined in Figure 5) (>2.5 fold enrichment, P<0.02). Note that there is no significantly enriched GO category corresponding to *Rhox10*-negatively regulated genes in the "SSC subset"

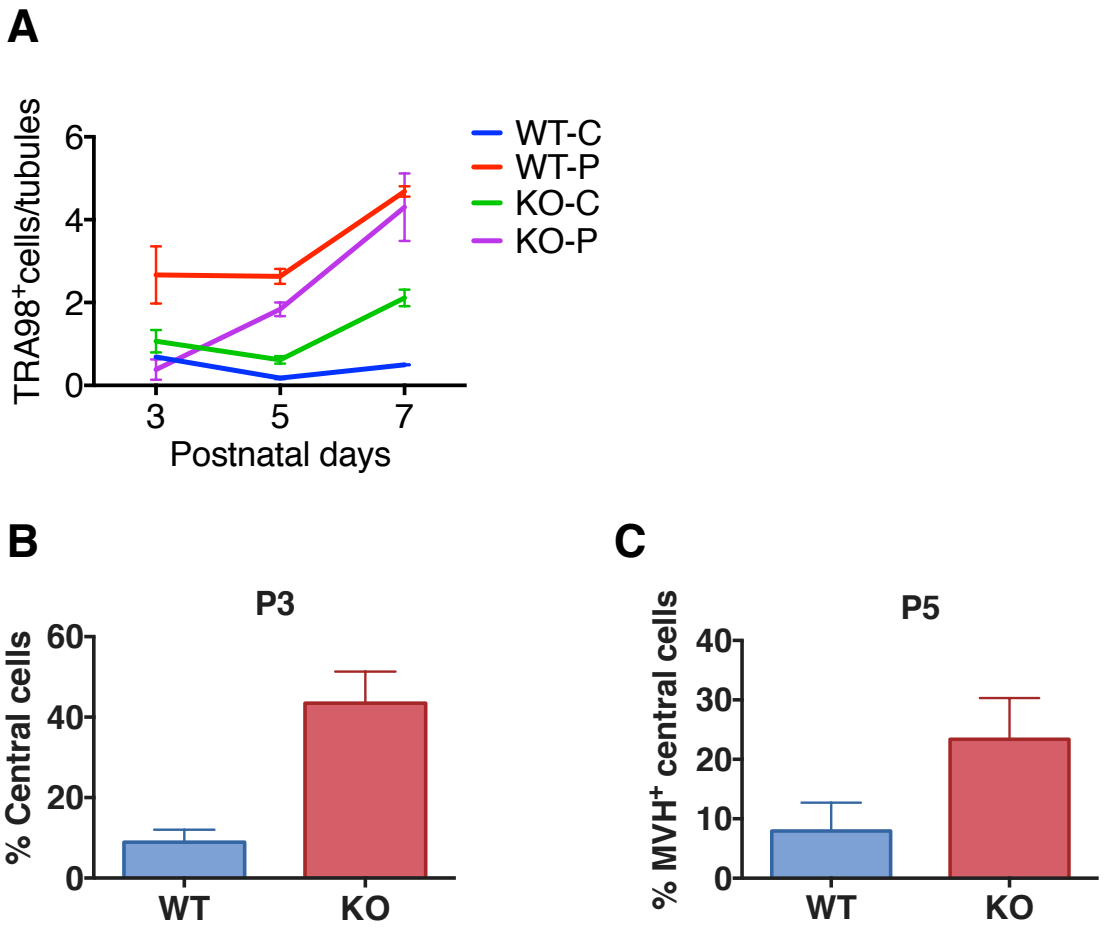

Figure S7. Additional Evidence that *Rhox10* Promotes ProSG Migration into the SSC Niche, Related to Figure 7

(A) Quantification of TRA98 positive germ cells located in the center (C) or periphery (P) of seminiferous tubules in *Rhox10*-KO mice and littermate control (WT) mice of the indicated age (n=2-5 each time point).
(B) Percentage of germ cells (identified by morphological criteria) located in the center of seminiferous tubules among total germ cells in P3 *Rhox10*-KO mice and littermate control (WT) mice (n=2-3).
(C) Percentage of MVH positive germ cells located in the center of seminiferous tubules among total germ cells in P5 *Rhox10*-KO mice and littermate control (WT) mice (n=3-5). All values are mean ± SEM.

SUPPLEMENTAL TABLE TITLES

Table S1. SC-RNAseq Library Mapping Statistics, Related to Experimental Procedures and Figure 5

Table S2. Fraction of Positive Values for Each Subgroup Using the Top (1009) Differentially Expressed Genes Identified by SCDE Analysis, Related to Experimental Procedures and Figures 5 and S5

Table S3. Fraction of Positive Values for Top Type-Specific Genes (from Supplemental Table 2), Related to Experimental Procedures and Figures 5 and S5

Table S4. SCDE Analysis Results Comparing *Rhox10*-KO and Control within Common Type Groupings, Related to Experimental Procedures and Figure 6 and S6

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Animals

Mouse colonies were maintained in agreement with protocols approved by the Institutional Animal Care and Use Committee at the University of California, San Diego. All animals were housed under a 12h light/12h dark cycle and provided with food and water *ad libitum*.

Generation of Rhox-c-KO mice

To generate *Rhox-c*-floxed mice, two targeting vectors were designed that insert *loxP* sites at the 5' and 3' ends of the *Rhox* cluster, respectively. The first vector contains two *loxP* sites surrounding a neomycin resistant cassette and two homologous arms for homologous recombination with the upstream intergenic region of *Rhox1* (Vega Biolab, Philadelphia, PA). The second vector contains two *loxP* sites surrounding a puromycin resistant cassette and two homologous arms for recombination with the second intronic region of *Rhox13* (Vega Biolab, Philadelphia, PA) (Figures 1A and S1A). The University of Michigan Transgenic Animal Model Core electroporated the first vector into embryonic stem (ES) cells and selected with G418. Southern blotting was performed to identify correctly targeted ES cell clones. Several positive cell clones were selected to use for the second targeting. After electroporation of the second vector and puromycin selection, the ES cell clones with both targeting vector recombination were screened by Southern blotting analysis. Two ES cell clones were selected to be injected into mouse blastocyst by the UCSD Transgenic Core. Male chimeras with the highest fraction of agouti fur from both parent ES cell clones were mated to *C57BL/6* females (Jackson Laboratory) and germline transmission was determined in the F1 generation by the presence of agouti fur color and a positive band corresponding to the mutant allele by PCR. To generate complete knockout mice, the F1 heterozygote mice were mated to animals from the mouse line *B6.FVB-Tg(EIIa-Cre)C5379mgd/J* (Jackson Laboratory), which express CRE at an early stage of embryogenesis. Because these *EIIa-Cre*-driver mice express CRE at only modest levels, this allowed us to select for F2 mice with both neomycin cassette and puromycin cassette deletion, leaving only a single *loxP* site at each end of the *Rhox* cluster. The F2 mice were backcrossed with BL6 mice to eliminate the *Cre* allele, and then the F3 mice were selected for the genotype of *Rhox-c*-floxed mice ($Rhox\text{-}c^{fl/y}$ males and $Rhox\text{-}c^{fl/+}$ females). To generate global- and germ cell specific- KO mice, *Rhox-c*-floxed mice were bred with *EIIa-Cre* and *Vasa-Cre* driver mice, respectively. The global *Rhox-c*-KO mice used for our studies were in a mixed genetic background (50-93.75% BL6) because mice with higher % *BL6* background were not observed amongst the progeny of our backcross matings (see Results). The germ cell-specific *Rhox-c*-KO mice used for our studies were >99% BL6 (as they were progeny of *Rhox-c*-floxed mice backcrossed until they were 99.1% BL6 and *Vasa-Cre* driver mice that were backcrossed until they were 99.2% *BL6*).

Generation of Rhox10-KO mice

To generate *Rhox10*-floxed mice, a targeting vector was designed that inserts a *loxP* site on either side of *Rhox10* exon 2. The targeting vector had 2.8 and 4.9-kb homologous arms upstream and downstream, respectively, of this exon. The targeting vector also included a neomycin resistance cassette flanked by two *FRT* sites (Figure 2A). The UCSD Transgenic Core electroporated the targeting vector into embryonic stem (ES) cells and selected with G418. Southern blotting was performed to identify correctly targeted ES cell clones (Figure S2A). Two ES cell clones were selected to be injected into mouse blastocyst by the UCSD Transgenic Core. Male chimeras with the highest fraction of agouti fur from both parent ES cell clones

were mated to BL6 females (Jackson Laboratory) and germline transmission was determined in the F1 generation by the presence of agouti fur color and a positive band corresponding to the mutant allele by PCR (Figure S2B). To generate complete knockout mice, the F1 heterozygote mice were mated to animals from the mouse line *B6.FVB-Tg(EIIa-Cre)C5379mgd/J* (Jackson Laboratory), which express CRE recombinase at an early stage of embryogenesis. The F2 mice were backcrossed with *C57BL/6* mice >5 times to generate *Rhox10*-KO mice with >90% *BL6* genetic background. Phenotypic data were obtained from 92.2-99.5 % *BL6* genetic background mice.

Histology and immunofluorescence

For histological analysis, mice testes were fixed for 2-6 hr in Bouin's fixative, transferred to 70% Ethanol, paraffin embedded using standard procedures, 5 m sections cut, deparaffinized, rehydrated, and stained for hematoxylin and eosin. For immunofluorescence analysis, mice testes were fixed overnight in 10% neutral-buffered formalin, transferred to 70% Ethanol, paraffin embedded using standard procedures, deparaffinized, rehydrated, and subjected to antigen retrieval by boiling in "Antigen unmasking solution, citric acid based" (Vector Laboratory, Inc., Burlingame, CA) for 14 min. Sections were blocked with 3% normal serum (from the species that the secondary antibody was raised in), 5% BSA, and 0.25% Triton X100 in PBS for 1 hr at room temperature. Primary antibodies were diluted in PBS containing 3% serum and 1% BSA and added to the samples for overnight incubation at 4°C. For double staining, another primary antibody was diluted in PBS containing 3% serum and 1% BSA and added to the samples for 2 hr at room temperature. The following primary antibodies were used: goat polyclonal anti-PLZF (AF2944, R&D Systems, 1:200 dilution); rabbit monoclonal anti-FOXO1 (C29H4; 2880, Cell Signaling, 1:200 dilution); mouse monoclonal anti-PAX7 (Developmental Studies Hybridoma Bank, Iowa City, IA, 1:50 dilution); rabbit monoclonal anti-KI67 (SP6; RM-9106, Thermo Scientific, 1:200 dilution); goat polyclonal anti-GFP (ab6662, Abcam, 1:300 dilution); goat polyclonal anti-GFR α 1 (AF560, R&D systems, 1:200 dilution); rabbit polyclonal anti-STRA8 (ab49405, Abcam, 1:200 dilution); and rat monoclonal anti-TRA98 (#73-003, B-bridge, 1:500 dilution). The samples were then incubated for 1 hr at room temperature in the dark with the appropriate fluorescence-conjugated secondary antibodies diluted in PBS containing 5% BSA. Vectashield anti-fade mounting medium containing DAPI (Vector Laboratory, Inc., Burlingame, CA) was applied to the samples and sections were viewed using a Leica DMI6000B inverted microscope (Leica, Boorburg, The Netherlands).

Phenotypic analysis

Whole body weight, testes weight, and seminal vesicle weight were measured using a chemical balance. For sperm analysis, caudal epididymides were placed in 1 ml PBS, minced with scissors, incubated for 20 min at room temperature, the cells filtered through a 70- μ m cell strainer to remove debris, and the sperm were diluted 2-10 times in water. Sperm counts were done using a hemocytometer. To assess sperm morphology, the sperm were smeared on coated glass slides, dried and stained with hematoxylin.

RNA isolation and q-RT-PCR analysis

Total RNA was extracted from samples using Trizol reagent following the manufacturer's protocol (Invitrogen, Carlsbad, CA). cDNAs were generated from 0.5 to 1 μ g total RNA using the Iscript reverse transcriptase (RT) kit, following the manufacturer's protocol (Bio-Rad). Primer pairs were designed to be complementary with sequences in separate exons in order to amplify only spliced mRNA (and thus largely avoid genomic DNA amplification). The RT product and primer pair were mixed with SYBR Green supermix (Bio-Rad) and PCR was performed using a Step One qPCR machine following the manufacturer's protocol (Applied Biosystems). The production of the amplicon was measured by SYBR green fluorescence and the threshold cycle (Ct) values were calculated. The Ct values obtained were normalized to that for the transcript encoding the ribosomal protein RPL19.

Cell analysis

Testes were decapsulated in HBSS (Mediatech Inc.) and digested with collagenase IV (1 mg/ml in HBSS) at 37 °C for 3 min with gentle agitation. After washing with HBSS, samples were digested with 0.25 ml of DNase I (7 mg/ml, Sigma Inc.) and 1 ml of Trypsin/EDTA (0.25%, Invitrogen Inc.) at 37 °C for 10 min. Digestion was stopped by adding 1 ml of media (MEMα + 10% FBS), the cells were filtered through a 40-mm cell strainer and resuspended in PBS with 3% FBS. For THY1 and GFRα1 staining, single testicular cells were stained with PE anti-CD90.2 (553005, BD Biosciences, 1:50 dilution) or anti-GFRα1 (AF560, R&D systems, 1:50 dilution) for 20 min on ice. After washing with PBS with 3% FBS, donkey anti-goat

IgG alexa Fluor 647 (1:200, A21447, Invitrogen) was added to GFRα1-stained cells and incubated for 20 min on ice. After washing, cells were either directly analyzed by FACS or fixed in 1% paraformaldehyde (PFA) for 15 min on ice and then analyzed by FACS. For Annexin-V staining, single testicular cells from *Id4-eGfp* mice were stained using the PE Annexin-V apoptosis detection kit (559763, BD Biosciences) following the manufacture's protocol. For cell cycle analysis, single testicular cells from *Id4-eGfp* mice were fixed in 4% PFA for 15 min on ice and stained with DAPI (3 μM concentration) for 15 min at room temperature and analyzed. All flow cytometry analysis was performed using LSRFortessa (BD Biosciences) and the Flowjo program.

SC-RNAseq analysis

Single testicular cells from *Id4-eGfp* mice were prepared from P7 WT (7 pups pooled), P3 WT (3 pups pooled), and P3 *Rhox10*-KO mice (4 pups pooled) and GFP^+ cells were sorted using a FACSAria II machine (BD Biosciences). Individual GFP^+ cells were captured on a C1 chip using the C1 Single-Cell Microfluidic Auto Prep System (Fluidigm) and reverse transcription and cDNA amplification was performed on the C1 chip using the SMARTer Ultra Low RNA Kit for Illumina Sequencing (Clontech Inc.). Modifications from the standard protocol (Fluidigm) include: 1) addition of ERCC RNA Spike-In Mix (4456740, Ambion) at 1:40,000 dilution in the lysis reaction buffer; 2) inclusion of a supplemental random primer in addition to the provided Poly-dT primer to enhance efficiency of reverse transcription. On chip LIVE/DEAD® Cell Viability staining (Life Technologies) was used to select only libraries from single live cells for further processing. cDNA was quantified using the Quant-iT PicoGreen dsDNA Assay Kit (P7589, Life Technologies). The concentration of single-cell libraries and a limited number of uncaptured chambers were normalized using the Eppendorf epMotion automated pipetting system. Final libraries were prepared for sequencing using the NexteraXT dual barcoding system (Illumina, Inc.). Sequencing was performed using Illumina's HiSeq2500 (La Jolla Institute for Allergy and Immunology and University of California, San Diego) and reads (50 bp single end) were mapped to the mouse genome (UCSC version mm10, GRCm38) using STAR (v2.3.0). Overall, 201 single cell libraries were sequenced to an average depth of 3.1 million reads which allowed for detection of an average of 6737 genes per cell (TPM $\geq$ 4, Table S1). Minimal technical variation across cells was demonstrated by the high correlation of ERCC spike-in RNA transcripts with their input quantities (mean R value of 0.894). Gene read counts were calculated based on exon mapping using HTSeq (v0.6.1). TPM values were calculated as read counts for all isoforms of an annotated gene divided by all reads mapped to all annotated genes.

Transcriptome data analysis

RNA sequencing data was analyzed and outlier samples excluded using the Singular 3.5.2 software. Genes having log2 transformed TPM values less than one were excluded (limit of detection or LOD of 1). Principal component analysis was used for each sample condition to identify the top 200 genes that best account for variance in the first three principal components for that sample set. These variant genes were used in hierarchical clustering analysis using the R package gplots to split each sample into subgroups. Differentially expressed genes were obtained from pairwise comparison of subgroups ($WT_{P3}A$, $WT_{P3}B$, $KO_{P3}A$, $KO_{P3}B$, $WT_{P7}A$ and $WT_{P7}B$) using the Single-Cell Differential Expression (SCDE) software version 1.2.1(Kharchenko et al., 2014) on raw gene counts. Compilation of the top differentially expressed genes or DEGs involved combining the top 50 genes (p value < 0.05) ranked on their maximum-likelihood estimation (MLE) from each comparison as well as all genes with MLE $\geq$ 3.3, representing fold-change value $\geq$ 10, for the following comparisons: $WT_{P7}A$ versus $WT_{P3}B$ or $KO_{P3}A$; $KO_{P3}B$ or $WT_{P3}A$ versus $WT_{P3}B$ or $KO_{P3}A$; $KO_{P3}B$ or $WT_{P3}A$ versus $WT_{P7}A$; $KO_{P3}B$ versus $WT_{P3}A$; $KO_{P3}A$ versus $WT_{P3}B$. This gene list was further trimmed using the pairwise anova function in the Singular software to include only genes with an anova p value < 0.01 to obtain a final list of 1009 DEGs (Table S2). PCA analysis using this gene subset revealed a potential for further splitting $KO_{P3}B$ into two new subgroups ($KO_{P3}B$ and $KO_{P3}C$), which was achieved as described above. PCA plots (Figures 5A and 5B) and scatter plots (Figure 5D) using the 1009 DEG list were generated using Singular 3.5.2. To identify genes associated with possible cell type, fraction of positive (FoP) values for the top 1009 DEGs were calculated from TPM values for each of the final subgroups using an established excel macro(Usoskin et al., 2014) and with a threshold of expression of five percent of the calculated maximum (Table S2). Unique positive expression was predicted as FoP values of 0.5 or greater for "Type" groupings (Type I: $KO_{P3}A$; Type II: $KO_{P3}B$ and $WT_{P3}A$; Type III: $KO_{P3}C$ and $WT_{P3}B$; Type IV: $WT_{P7}A$ and $WT_{P7}B$) and FoP values less than 0.5 for remaining subgroups in the analysis (See Table S2). Type FoP values were ranked and the top 20 were selected for each type grouping for generation of a hierarchical clustering plot (Figure S5, Table S3) using Singular

3.5.2. To identify potential Rhox10 target genes, pairwise SCDE analysis was performed within similar type subgroups (Type II: $WT_{P3}A$ versus $KO_{P3}B$; Type III: $WT_{P3}B$ versus $KO_{P3}C$). Upregulated genes in each sample subgroup (which can also be considered as downregulated in the comparison subgroup) were selected based on p values < 0.05 and MLE values ≥ 1 representing at least a 2-fold change in expression (Table S4). Developmental control RNA sequencing data from E16.5, E18.5, P2.5 and P10.5 mouse male germ cells (Pastor et al., 2014) were obtained from the Gene Expression Omnibus (GEO Series GSE63048) and mapped as indicated above. To identify classes of genes regulated by RHOX10, DAVID gene ontology analysis was performed (Huang et al., 2009).

Germ cell transplantation analysis

Donor testicular cells for germ cell transplantation analysis were isolated from P7-8 male mice progeny from *C57BL/6-Tg(UBC-GFP)30Scha/J* male mice (Jackson Laboratories) bred with *Rhox10*-heterozygote female mice. All testicular cells from these mice express GFP, allowing identification of these donor germ cells in recipient seminiferous tubules. Four *Rhox10*-KO pups and five WT littermate control pups were used for transplantation. The recipient mice (129XC57; 5-6 weeks old; Jackson Laboratories) were treated with busulfan (50 mg/kg; Sigma, St. Louis, MO) to remove endogenous germ cells for donor germ cell engraftment. Approximately 6 weeks after busulfan treatment, the donor testis cells suspension was introduced into recipient seminiferous tubules by efferent duct injection as described previously (Ogawa et al., 1997). A single-cell suspension of testicular cells from each donor mouse was transplanted into 3-4 recipient mice. Testes of recipient mice were collected 3 month after transplantation and analyzed for donor-derived colonies of spermatogenesis, as described (Ogawa et al., 1997).

SUPPLEMENTAL REFERENCES

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kharchenko, P. V, Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. Nat. Methods *11*, 740–742.

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P. V, et al. (2014). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat. Neurosci. *18*, 145–153.