SI Materials and Methods

Mouse models and maintenance. All experiments involving mice conformed to ethical principles and guidelines approved by the Committee on Animal Care at the Massachusetts Institute of Technology or the Cincinnati Children's Hospital Medical Center. Mice carrying the *129P2-Dazl*^{tm1Hjc} allele [*Dazl* mutant, RRID:IMSR_JAX:023802, (2)] were backcrossed to various *M. m. domesticus* strains, including C57BL/6NTac (B6, > 30 generations), FVB/NTac (FVB, > 30 generations), 129S4/SvJae (129S4, > 30 generations), 129S2/SvPasCrl (129S2, backcrossed from 129S4/SvJae for 10 generations), and the *M. m. molossinus* strain derived from wild-trapped Japanese mice, MSM/MsJ (MSM, backcrossed from B6 for 10 generations). The fluorescent Cre-reporter *Cg-Gt(ROSA)26Sor*$^{tm9(CAG-tdTomato)Hze}$ (*LSL-tdTomato*, RRID:IMSR_JAX:007909) allele (7) was maintained on a B6 background (Jackson Laboratory, Bar Harbor ME). The *CBA-Tg(Pou5f1-EGFP)*^{2Mnn} fluorescent reporter allele [*Oct4:EGFP*, RRID:IMSR_JAX:004654, (8)], present as a multi-copy transgene array near the telomere of chromosome 9 (9) was maintained on a B6 background. The *B6-Nanog*^{tm1Hoch} (*Nanog:GFP*, RRID:IMSR_JAX:016233) reporter allele (10) was backcrossed to B6 for seven generations, and to 129S4 for 10 generations. The Cre-recombinase *Mvh*$^{Cre\text{-}mOrange/+}$ [*Ddx4-Cre*, a null-allele (11)] was backcrossed to B6 and 129S4 for 15 generations. Mice carrying both *Sry*^{tm1} deletion (12) and *Tg(Sry)2Ei* transgene [referred to as *TgSry*, RRID:IMSR_JAX:010905, (13)], 129X1-*Bax*^{tm1Sjk} allele [RRID:IMSR_JAX:002994, (14)], and C57BL/6N-*Gcna*$^{tm1.1Dcp}$ allele [RRID:IMSR_JAX:031055, (15)] were each backcrossed to 129S4 for at least 10 generations. CD-1 outbred female mice were obtained from Charles River Laboratories (Wilmington MA).

Generation of Dazl reporter mice. We utilized a CRISPR/Cas9-mediated strategy to introduce a P2A-tdTomato sequence immediately 5' of the endogenous stop codon of *Dazl*. We designed a gRNA to a PAM site near the endogenous stop codon (in the 3'UTR of *Dazl*, target sequence: CTCTGCTAACTCATCTCAGG), and cloned this gRNA into the *BbsI* site of pX458 [also known as pSpCas9(BB)-2A-GFP, RRID:Addgene_48138]. We then performed pronuclear injection of this vector, together with an HDR template encoding *P2A-tdTomato* flanked by 2.89 kb upstream (including exon 10) and 1.98 kb downstream (including exon 11) of the endogenous *Dazl* locus, into one-cell embryos of an F1 B6D2 genetic background. Injected embryos were transferred to pseudopregnant CD-1 female mice. Correctly targeted mice carrying the *Dazl* reporter allele (B6D2-*Dazl*$^{em1(tdTomato)Huyc}$; referred to as *Dazl reporter*) were back-crossed for three generations to a B6 background.

Generation of Dazl conditional mice. We utilized a two-step CRISPR/Cas9-mediated strategy to insert *loxP* sites to introns flanking part of the RNA-recognition motif (RRM) of *Dazl*. We first designed gRNAs to target the intron between exons 3 and 4 (intron 3), and between exons 7 and 8 (intron 7; SI Appendix, Table S6). Targeting gRNA oligos were ligated into an expression plasmid pX330 (also known as pX330-U6-Chimeric_BB-CBh-hSpCas9, RRID:Addgene_42230). Cas9 mRNA and gRNAs were generated, and injected together with HDR templates containing *loxP* sequences into one-cell C57BL/6N embryos as previously described (16). After the first round of injections, a single mouse with a correctly targeted *loxP* in intron 7 was used to establish a homozygous colony. A second round of injections into one-cell embryos (homozygous for the *loxP* at intron 7) resulted in incorporation of an additional *loxP* site in intron 3 (the *Dazl* allele with two *loxP* insertions is named B6N-*Dazl*^{em1Dcp}, referred to as *Dazl-2L* in the text), confirmed by PCR and Sanger sequencing (SI Appendix, Fig. S7A). Mice were subsequently backcrossed five additional generations to a C57BL/6N background, and maintained as homozygotes (*Dazl-2L/2L*). We bred mice carrying this conditional allele with a germ cell-specific Cre [*Mvh*$^{Cre\text{-}mOrange}$, referred to as *Ddx4-Cre* in the text (11)] and observed germline recombination of the floxed allele in all Cre-carrying pups (B6N-*Dazl*$^{em1.1Dcp}$, referred to as *Dazl-1L* in the text). This recombined allele creates an out-of-frame mutation (*p.Tyr82Cys*), removing the 35 N-terminal amino acids of the RRM. To ensure that deletion of exons 4 through 7 phenocopied the established null allele [*129P2-Dazl*^{tm1Hjc} (2)], we intercrossed animals hemizygous for the recombined allele (*Dazl-1L/2L*) and analyzed the gonadal phenotype of homozygotes for germline survival on a B6 background (SI Appendix, Fig. S7B). We also back-crossed the *Dazl-1L* allele for five generations to the 129S4 background and assayed teratoma production (SI Appendix, Fig. S7D).

To generate conditional *Dazl* mice, we bred male mice carrying *Ddx4-Cre* and hemizygous for the recombined *Dazl* allele (B6.*Dazl-1L/+;Mvh*$^{+/Cre\text{-}mOrange}$; referred to as B6.*Dazl-1L/+;Cre* in the text) with female mice homozygous for the conditional allele (B6.*Dazl-2L/2L*). Resulting progeny heterozygous for the null allele and the conditional allele of *Dazl*, together with the Cre-recombinase (genotyped as B6.*Dazl-1L/2L*;*Cre*), were considered to be conditional for *Dazl* (referred to as B6.*Dazl* cKO in the text).

Generation of DAZL-deficient pigs. All experiments involving pigs conformed to the standard procedures and protocols approved by International Center for Biotechnology (Mt Horeb WI, formerly MOFA Global). *DAZL*-deficient males were generated as previously described using TALEN-mediated gene editing in porcine fibroblasts (17). In this study, *DAZL*-deficient sows

were generated by transfection of XX-bearing Landrace fetal fibroblasts with transcription activator-like effector nucleases (TALENs) and an oligonucleotide homology directed repair (HDR) template specific to the *DAZL* locus (listed in SI Appendix, Table S6). Briefly, 500,000 cells were transfected with 2 μg of TALENs mRNA and 2 μM oligonucleotide HDR template, and then electroporated using the Neon Transfection system (Thermo Fisher Scientific): input voltage 1800 V; pulse width 20 ms; and pulse number 1. Electroporated cells were cultured for three days at 30° C, followed by dilution cloning and screening of fibroblast clones by PCR amplification of DNA flanking the HDR target site, using primers listed in SI Appendix, Table S6, followed by BamHI digest of the DNA. Candidate cells with at least one HDR allele (BamHI positive) were further characterized by direct sequencing and/or TOPO cloned (Thermo Fisher Scientific) and verified by Sanger sequencing. Cell colonies with bi-allelic frame-shift mutations were pooled and cloned by chromatin transfer under contract with MOFA Global under Animal Welfare Assurance no. A4520/01. Sows were genotyped by the same methods used for colony screening (SI Appendix, Table S3).

Mouse RNA-seq analysis. To define the transcriptional program in the mouse germline as PGCs colonize the nascent gonads, we mapped raw RNA-seq reads from populations of sorted germline cells from GSE41908 (18) to the mouse transcriptome (Ensembl 84) using kallisto (19), with the following options: --bias --single, -l 200 -s 20. Transcript-level estimated counts and transcripts per million (TPM) values were summed to the gene level. We identified genes differentially expressed between E9.5, 11.5 and E13.5 using DESeq2 (20), requiring expression of $\geq$1 TPM at any age for inclusion in our analysis.

To assess germ cell-specificity in embryonic mouse gonads, raw RNA-seq reads from Kit^{+}/Kit^{+} (control) and Kit^{W}/Kit^{Wv} (germ cell-depleted) gonads (21) at E14.5, and from *Dazl*-deficient and *Stra8*-deficient ovaries at E14.5 (GSE70361) were mapped as above, and transcript-level estimated counts were summed to the gene level. For each comparison, we used the edgeR package (22) to obtain gene-level counts per million (CPM) values, normalized using the trimmed mean of M-values (TMM) method. Germ cell specificity was calculated by subtracting the CPM of germ cell-depleted (Kit^{W}/Kit^{Wv}) gonads from control gonads, divided by the CPM in control gonads. Where this ratio is < 0, as is the case for genes that increase in CPM in germ cell-depleted gonads compared with controls, the germ cell-specificity was set to 0 as none of the expression in control gonads could be assigned to germ cells. A ratio of 1 indicates germ cell-specific expression.

To assess whether germ cell factors were expressed in *Dazl*-deficient germline cells, we collected Oct4:EGFP-positive cells by FACS from control and *Dazl*-deficient embryonic mice at E10.3 (1-3 tail somites) and E11.5 (17-18 tail somites); *n* = three embryos at each time point. RNA was extracted using a Quick-RNA MiniPrep kit (Zymo Research, Irvine CA). Reverse transcription was performed using the SMARTer PCR cDNA Synthesis Kit (Clontech, Mountain View CA), and sequencing performed using a HiSeq2000 (Illumina). Transcripts were mapped to the mouse transcriptome using kallisto, and TPM values summed to the gene level (as above). Differential gene expression was assessed using DESeq2 (20).

Human single cell RNA-seq analysis. To define the transcriptional program in the human germline as PGCs colonize the nascent gonads, we analyzed raw RNA-seq reads from single cells from GSE86146 (23). Cell barcodes and unique molecular identifiers (UMIs) from read 2 were extracted using the umis python package (24) (https://github.com/vals/umis). We required that the cell barcode be an exact match to one of the 96 possibilities outlined (23). Both read 1 and read 2 were then mapped to the human transcriptome using kallisto (19) to obtain a pseudobam file; this was then passed to the tagcounts function to obtain gene-level UMI counts. The human transcriptome was defined as the subset of transcripts annotated in the GENCODE v24 annotation, comprising the union of "GENCODE Basic" transcripts and transcripts with a Consensus Coding Sequence (CCDS) ID.

We used the Seurat R package (http://satijalab.org/seurat/) to normalize, filter, and cluster single cells on the basis of their expression profiles (25). We first filtered out genes expressed in fewer than 20 single cells, and cells that expressed fewer than 2,000 genes or $1x10^5$ UMIs. After these filters, 1,882 single cells were considered for downstream analysis. We used the NormalizeData function to log-normalize the gene-level counts. We used the FindVariableGenes function with a minimum log-expression value and dispersion of 2 to identify 406 highly variable genes. Data were then scaled using the ScaleData function, and the number of UMIs was regressed out of the data. Principal components analysis (PCA) was performed on the 406 highly variable genes, and a jack-straw analysis with 200 replicates revealed the top 19 principal components (PCs) to be significant. The FindClusters function was used on these 19 PCs to assign single cells to clusters. For the visualization of clusters, tSNE was run on the top 19 PCs, with do.fast = TRUE. TPM values for each gene, in each cell, were calculated by normalizing gene UMI counts to cell library counts, and multiplying by $1x10^7$ (as most libraries had fewer than $1x10^6$ UMIs).

Human embryonic cells were sexed by summing the TPM of all Y-linked genes and defining cells with a total Y-linked gene TPM of less than 10 as XX, and greater than 10 as XY. Where appropriate, we excluded somatic cells on the basis of their expression of *WT1*, and meiotic cells on the basis of their expression of *STRA8*. With these filters applied, the following populations of germline cells were considered for further analysis: (group a) 37 'migratory' XY cells in cluster 5 from week 4, (b) 47 XX cells in cluster 5 from week 5; (c) 175 'early gonadal' XY cells in cluster 1, 11 and 6, (d) 281 XX cells in clusters 2 and 5; (e) 313 'late gonadal' XY cells in cluster 0 derived from weeks 19-25, and (f) 220 XX cells in clusters 7, 10 and 13, derived from weeks 11-26 (SI Appendix, Fig. S1A).

We used the SCDE R package (http://hms-dbmi.github.io/scde/) to analyze differential expression (26) between XY cells in group (a) and (c), and separately, between XX cells in (b) and (d). *P* values for 44 orthologous genes significantly upregulated in mouse were extracted and corrected for multiple hypothesis testing using the Benjamini-Hochberg procedure.

To assess germ cell specificity in humans, we first defined four populations of cells: (group g) XY germ cells (clusters 1, 11 and 6, *WT1*-negative), (h) XY somatic cells (cluster 4, *WT1*-positive), (i) XX germ cells (clusters 2, 5, 7, 10, and 13, *WT1*-negative), and (j) XX somatic cells (cluster 3, *WT1*-positive). We used the scde.posteriors function from SCDE to obtain the maximum likelihood estimate (MLE) of gene expression in each cluster. For a given gene, XY germ cell specificity was defined as: (MLE in XY germ cells; group g) / (MLE in XY germ cells; group g, + MLE in XY somatic cells; group h), and XX germ cell specificity was defined as: (MLE in XX germ cells; group i) / (MLE in XX germ cells; group i, + MLE in XX somatic cells; group j). A ratio of 1 indicates germ cell-specific expression, a ratio of 0 indicates somatic cell-specific expression and/or not detected in germ cells. Custom scripts for scRNA-seq are available at https://github.com/snaqvi1990/pgc_scrnaseq/.

Cross-species single gene analysis of testis specificity. To determine gonad-specific gene expression in tetrapods, we analyzed published RNA-seq datasets from nine adult tissues (heart, muscle, kidney, liver, lung, colon, brain, spleen, and testis) isolated from seven species. For rhesus macaque, mouse, rat, bull, and chicken, raw reads were analyzed from GSE41637 (27), and for frog (*Xenopus laevis*), raw reads were analyzed from GSE73419 (28). For human, raw reads were obtained from the Genotype-Tissue Expression Consortium [GTEx, midpoint v6 release, dbGap accession phs000424.v6.p1, (29)]. For each of the corresponding nine adult tissues in human, we selected the 10 male samples with the highest RNA Integrity Number (RIN; sample IDs are provided in SI Appendix, Table S5). We used Ensembl 84 transcriptome

annotations for mouse, rat, and bull; Ensembl 88 annotation for rhesus (as they were derived from an improved genome assembly); GENCODE v24 annotation for human (as described above); and Xenbase v9.1 primary transcript assemblies for frog. Transcript-level TPM values were obtained using kallisto (19) in paired-end mode with the --bias option enabled, summed to the gene level, and summarized by the median TPM across all individuals in each tissue. Amniote orthology relationships were obtained from Ensembl Biomart, and human-frog orthologies were obtained from Xenbase. In all analyses, we required one-to-one orthology between human and mouse, but for comparisons of testis specificity, we allowed one-to-many orthology relationships, averaging testis specificity values in species where multiple orthologs were present.

Gonad specificity was calculated as previously described by dividing the testis TPM value by the sum of TPM values across all nine adult tissues, and repeated for each species (30). A ratio of 1 indicates gonad-specific expression; a ratio of 0 indicates no gonadal expression. Where no orthologous gene exists, or there is no annotation in the reference genome, the label 'no data' was applied (see Fig. 1D and SI Appendix, Fig. S2A). Violin plots were generated using the 'vioplot' package (https://CRAN.R-project.org/package=vioplot) and significance determined by Wilcoxon rank-sum test in R (https://www.r-project.org/).

Cross-species gene set analysis of testis specificity. The curated set of PGC factors (set iii, SI Appendix, Fig. S2A-D) was identified by survey of the literature. For each factor, we required that the gene be necessary for PGC or PGC-like cell development from pluripotent stem cell lines in either mouse or human, or has been reported as a marker of PGCs in any mammal *in vivo*.

To identify factors up-regulated on derivation of mouse PGC-like cells, we analyzed microarray data from GSE61924 (31). Gene expression was analyzed with NCBI's GEO2R tool, comparing day four cytokine-induced PGC-like cells to 2i + LIF-cultured mESCs. For human, raw reads from GSE60138 (32) were obtained and mapped to the above-described version of the human transcriptome using kallisto, with the following options: --bias -s 20 -l 200 --single. Transcript-level estimated counts and TPM values were summed to the gene level, and DESeq2 was used to perform differentially expression analysis between day four PGCLCs and day two pre-induced hESCs. Genes were required to be expressed >1 TPM in either condition to be considered for differential expression.

To identify genes commonly upregulated upon PGC-like cell induction in both human and mouse (set iv, SI Appendix, Fig. S2D), we used a procedure analogous to that used to identify genes induced on PGC colonization of the gonad. First, 101 genes upregulated in mouse (Benjamini-Hochberg-adjusted P value < 0.05, $\log_2$ fold-change > 1) were identified, of which 62

had one-to-one orthologs in human. Next, these 62 genes were tested for differential expression upon human PGC-like cell derivation as described above; of these 62 genes, 23 were also upregulated in human with adjusted *P* value < 0.05 (we did not require a $\log_2$ fold-change cutoff in humans as no such criteria were applied when delineating the gonadal PGC program).

Random sampling gene set analysis of testis specificity. To compare the testis specificity of the set of 13 genes commonly activated on PGC colonization to a cohort of unbiased gene sets (expressed at similar levels in both human and mouse PGCs), we sampled 500,000 random sets of 13 genes, obtained from the list of genes expressed at >1 TPM in both human and mouse migratory PGCs. To ensure equivalent expression levels in the germline, we recorded median testis specificity only if the set's expression levels in migratory PGCs did not differ significantly (two-sided Wilcoxon rank-sum test, *P* value > 0.1) from the set of 13 genes activated on PGC colonization (after arrival at the gonad). Using these criteria, 86,550 gene sets were identified for comparison of median testis specificity.

Analysis of pluripotency network expression in the germline. Human pluripotency factors were collated from the union of two analyses of naïve-like human ES cells (33, 34). Of this gene list, *CDHR1, DNMT3L*, *GCM1*, *LIPH, PRAP1*, *SOX2*, *ZFN57*, and *ZIC2* did not meet the criterion of TPM >10 in the human germline. Mouse 'naïve' and 'general' pluripotency factors were collated as previously described (35); *Klf4* and *Tbx3* did not meet the criterion of TPM >1 in sorted mouse germline cells; *Nr0b1* and *Klf2* were excluded due to expression in somatic lineages of the gonad.

Mouse genotyping. A small ear biopsy was taken prior to weaning. Genomic DNA was extracted in lysis buffer (100 mM Tris pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, and 100 μg/ml Proteinase K) at 65°C overnight. DNA was precipitated with an equal volume of isopropanol and centrifuged. The pellet was then washed in 70% v/v ethanol, centrifuged, and resuspended in TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). Genotyping was performed using the PCR primers and conditions outlined in SI Appendix, Table S6, using Phusion DNA polymerase (New England Biolabs Inc, Ipswich MA).

Flow cytometry. Embryonic urogenital ridges or gonads (where appropriate) were dissected into phosphate buffered saline (PBS). A single-cell suspension was generated using trypsin, with 20 μg/ml DNase (Sigma, St Louis MO), in PBS. Cells were washed once with 20% v/v serum in

PBS, centrifuged at 500 *g*, resuspended in 1% v/v serum in PBS with DNase, and the suspension passed through a 40 μm filter. Oct4:EGFP-positive cells were collected using a FACSAria cell sorter (BD, Franklin Lakes NJ) into cell culture media. Expression from *Nanog:GFP*, *DAZL:tdTomato*, *LSL-tdTomato* or *Oct4:EGFP* reporters in germline cells was detected with a LSRFortessa cell analyzer (BD), and cytometry data analyzed with FlowJo software (v10.2, FlowJo, LLC., Ashland OR).

EG cell derivation. We used a modified EG derivation procedure (36). Briefly, Oct4:EGFP-positive cells were collected by flow cytometry and cultured in N2B27 medium, consisting of DMEM/F12 and Neurobasal, with N2 and B27 supplements, 0.1 mM nonessential amino acids, 2 mM L-Glutamine, 0.05% bovine albumin fraction V (Cat# 15260037), penicillin/streptomycin (each from Thermo Fisher Scientific, Waltham MA), and 0.1 mM 2-mercaptoethanol (Sigma-Aldrich). Cells were plated on 6-well culture dishes first coated with 10 μg/ml human plasma fibronectin for 1 h (Cat# FC010, EMD Millipore) on a monolayer of Sl220 feeder cells previously inactivated with mitomycin C (Cat# M4287, Sigma-Aldrich). For the first 48 h, cells were cultured in N2B27, with the addition of 20 ng/ml bFGF (Cat# PHG0261, Thermo Fisher Scientific), 50 ng/ml SCF (Cat# 455-MC, R&D systems), 1 μM ATRA (Cat# R2625, Sigma-Aldrich), the GSK inhibitor CHIR99021 at 3 μM (Cat# 4423, R&D systems) and 1,000 U/ml LIF (ESGRO, Cat# ESG1106, EMD Millipore) at 37°C in 5% CO_2 (EG derivation media). Cells were then subjected to half media changes every 48 h in N2B27, supplemented with 3 μM CHIR99021, the MEK inhibitor PD0325901 at 1 μM (Cat# 13034, Cayman Chemical, Ann Arbor MI) and 1,000 U/ml LIF (N2B27/2iLIF). Cultures were maintained in these conditions, and EG colony formation was monitored for 10 days, at which point the number of discrete colonies was counted and the efficiency of EG colony formation determined.

After 10 days, EG colonies were picked and cultured on CF6Neo MEF 7M MitoC feeder cells (Cat# GSC-6005M, MTI-GlobalStem, Gaithersburg MD) in DMEM/F12 supplemented with 10% fetal calf serum, 0.1 mM nonessential amino acids, 2mM L-Glutamine, 0.1 mM 2-mercaptoethanol and Penicillin/Streptomycin, with 3 μM CHIR99021, 1μM PD0325901 and 1,000 U/ml LIF. An early passage of EG cells were cryopreserved in 10% v/v DMSO in serum and stored under liquid nitrogen.

Histological analysis. Dissected tissues were fixed in 4% paraformaldehyde (PFA) or Bouin's solution for 16 h at 4°C. Tissues were washed in 70% ethanol, embedded in paraffin, and 5-μm sections cut for histology. Periodic acid-Schiff (PAS) staining was performed using a kit from

Sigma-Aldrich. Immunohistology was performed by dewaxing and dehydration of sections, followed by antigen retrieval in citrate buffer (10 mM sodium citrate, 0.05% Tween 20, pH6.0) by microwaving. Sections were then blocked in 5% donkey serum (Jackson ImmunoResearch Laboratories, West Grove PA), and incubated with primary antibodies (DAZL: #sc-27333, Santa Cruz Biotechnology, Dallas TX; DDX4: Cat# AF2030, RRID:AB_2277369, R&D Systems, Minneapolis MN; GCNA: a gift from George Enders; GFP: Cat# ab13970, RRID:AB_300798, Abcam, Cambridge, United Kingdom; RFP: Cat #600-401-379, RRID:AB_2209751, Rockland Immunochemicals Inc., Limerick PA; SOX9: Cat# AB5535, RRID:AB_2239761, EMD Millipore, Billerica MA). Sections were then washed in PBS, and later incubated with secondary antibodies (Jackson ImmunoResearch Laboratories). A glass coverslip was applied with ProLong Gold Antifade reagent with DAPI (Thermo Fisher Scientific), and immunohistology detected by confocal microscopy (Zeiss 700, Jena Germany). Meiotic spreads were prepared as previously described (37). 5-ethynyl-2'-deoxyuridine (EdU, Cat# A10044, Thermo Fisher Scientific) was diluted in sterile saline and injected into pregnant mice at 10 μg/g. After 3 h, embryos were dissected and histological sections prepared (as above), and EdU incorporation detected using a Click-iT imaging kit (Cat# C10337, Thermo Fisher Scientific).

Blastocyst injections. Chimeras were generated by injecting 129SB6F1.*Dazl*$^{tm1Hjc/tm1Hjc}$; *CBA-Tg(Pou5f1-EGFP)*^{2Mnn} EG cells derived from E15.5 testes (resulting from a 129S4 x B6 cross) into blastocysts from wildtype C57BL/6NTac embryos. Blastocysts were implanted into pseudopregnant CD-1 female mice and carried to term. To assess chimerism, various tissues were dissected from chimeric mice and genotyped for the *Dazl*^{tm1Hjc} and *CBA-Tg(Pou5f1-EGFP)*^{2Mnn} alleles (see SI Appendix, Table S6). To determine if EG cells contributed to the germline, gonads were dissected and prepared for flow cytometric detection of the *Oct4:EGFP* transgene (expressed by undifferentiated spermatogonia, Fig. S4H). *Dazl*'s essential role in spermatogenesis precludes an analysis of germline transmission (SI Appendix, SI Discussion).

Teratoma formation. To assess teratoma formation, animals were dissected at the age indicated in the text, and gonads isolated and fixed in either PFA or Bouin's fixative. Each gonad was prepared for histological analysis, and teratoma formation confirmed by the presence of derivatives from each somatic germ layer. A clinically trained pathologist examined 20 mouse gonadal tumors, and independently confirmed that each contained differentiated lineages of all three germ layers. Some teratomas in four-week-old animals were composed primarily of

neuroepithelium, consistent with the histology of immature teratomas. We did not identify tumors at extra-gonadal locations in any of our mice.

Detection of SNPs in 129S2 and 129S4 mice, and genotyping of teratomas. Genomic DNA was extracted from wildtype 129S2 and 129S4 mice and hybridized to the Mouse Diversity Genotyping Array (Cat# 901615, Thermo Fisher Scientific). We visually inspected normalized summary values for all variants to identify homozygous variants that differentiated the 129S2 and 129S4 substrains. We identified 411 SNPs that met our criteria, predominantly located on Chromosomes 1, 3 and 4 (204, 36 and 132 SNPs, respectively, Dataset S4). We then validated 11 of these SNPs that differentiate between 129S2 and 129S4 mice by Sanger sequencing (SI Appendix, Table S2).

To determine the cellular origin of spontaneous gonadal teratomas, genomic DNA was extracted from multiple pieces of each gonadal teratoma from 129SF1.*Dazl*-deficient mice, in parallel with host DNA (ear sample). For each sample, a PCR was performed at 11 discriminative loci (using the primers outlined in SI Appendix, Table S2), and analyzed for the presence of SNPs originating from 129S2 and 129S4 strains by PCR using Phusion DNA polymerase, followed by Sanger sequencing, and visualized using SnapGene Viewer software (v3.0.2, GSL Biotech, LLC., Chicago IL).

Calculation of mammalian evolutionary relationships. The evolutionary relationship between mammalian species was calculated using TimeTree (38), which estimates the timing of speciation based on a synthesis of the published literature.

Statistical analysis. Where applicable, the efficiency of EG cell derivation was compared using a t-test (Prism, v10.2 GraphPad Software Inc., La Jolla, CA). Where no EG cells were derived, a Fisher's exact test was employed (Prism). Statistical tests for the incidence of teratoma formation were performed using a Fisher's exact test using a two-tailed distribution (Prism). For all experiments, a P value of ≤ 0.05 was used as a measure of statistical significance. Statistical tests for RNA-seq data are described above.

Data availability. Data generated using the Mouse Diversity Genotyping Array (Thermo Fisher Scientific) for the SNP genotyping of 129S2 and 129S4 substrains has been deposited at the Gene Expression Omnibus under accession number GSE87771. Data generated from control and *Dazl*-deficient germline cells at E10.3 and E11.5 have been deposited under the SRA BioProject accession number PRJNA434733. Raw reads from RNA-seq experiments were downloaded from

publically available datasets (GSE41637, GSE41908, GSE60138, GSE70361, GSE86146, GSE73419 and phs000424.v6.p1). Array data (GSE61924) was analyzed with NCBI's GEO2R tool. Subsequent re-analyses of RNAseq of the mouse and human germline can be found in Datasets S1-S3.