bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The RNA-Binding Protein DND1 Acts Sequentially as a Negative Regulator of Pluripotency and a Positive Regulator of Epigenetic Modifiers Required for Germ Cell Reprogramming

Victor A. Ruthig[1], Matthew B. Friedersdorf[2], Jason A. Garness[1], Steve C. Munger[3], Corey Bunce[1], Jack D. Keene[2], and Blanche Capel[1]*

1 Department of Cell Biology, Duke University Medical Center, Durham, NC 27710
2 Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710
3 The Jackson Laboratory, Bar Harbor, ME, 04609

* Corresponding author: Blanche Capel, E-mail: blanche.capel@duke.edu

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

The adult spermatogonial stem cell population arises from pluripotent primordial germ cells (PGCs) that enter the fetal testis around embryonic day 10.5 (E10.5). These cells undergo rapid mitotic proliferation, then enter a prolonged period of cell cycle arrest (G1/G0) during which they transition to pro-spermatogonia. In mice homozygous for the *Ter* mutation in the RNA-binding protein *Dnd1* (*Dnd1*$^{Ter/Ter}$), many germ cells fail to enter G1/G0, and give rise to teratomas, tumors in which many embryonic cell types are represented. To investigate the origin of these tumors, we sequenced the transcriptome of male germ cells in *Dnd1*$^{Ter/Ter}$ mutants at E12.5, E13.5, and E14.5, just prior to the formation of teratomas, and correlated this information with direct targets of DND1 identified by DO-RIP-Seq. Consistent with previous results, we found that DND1 controls the down regulation of many genes associated with pluripotency and active cell cycle, including elements of the mTor, Hippo and Bmp/Nodal signaling pathways. However, DND1 targets also include genes associated with male differentiation including a large group of chromatin regulators activated in wild type but not mutant germ cells during the transition between E13.5 and E14.5. These results suggest multiple functions of DND1, and link DND1 to the initiation of epigenetic modifications in male germ cells.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Primordial germ cells (PGCs) are specified at the base of the allantois in mouse embryos at embryonic day (E)6.75, proliferate, migrate through the gut mesentery, and arrive in the gonad between E10.0-E11.0, expressing stem cell markers including SOX2 (SRY-box 2), NANOG (Nanog homeobox), and OCT4 (also known as POU5F1, POU domain, class 5, transcription factor 1) (Matsui et al. (1992); (McLaren, 1995); (Yeom et al., 1996); (Pesce and Scholer, 2000; Yamaguchi et al., 2005). Prior to E12.5, germ line stem cells (EGCs) can be derived readily from the germ cell population (Matsui et al., 1992), and teratomas, a tumor in which all embryonic cell types are represented, spontaneously arise from germ cells in males with a frequency of 1-10% in some strains (Stevens and Little, 1954); (Bustamante-Marin et al., 2013); (Dawson et al., 2018). However, between E12.5 and E15.5, the efficiency of both EGC derivation and teratoma induction declines, presumably reflecting changes that lead to suppression of the underlying pluripotent state of germ cells. These changes are coincident with the entry of male germ cells into G1/G0 cell cycle arrest and their fate transition to pro-spermatogonia (Fig. 1A) (Stevens, 1966); (Western et al., 2011; Western et al., 2010).

A mutation, called *Ter*, that arose spontaneously during routine breeding of 129/SvJ mice, caused a severe reduction in germ cell numbers soon after their specification, and acted as a strong modifier that increased the incidence of spontaneous testicular teratomas to >90% in 129/SvJ mutants (Fig. 1A)(Stevens, 1973). In 2005, *Ter* was mapped to a point mutation that introduced a premature stop codon in the RNA-binding protein (RBP) Dead End 1 *Dnd1* (*Dnd1*^{Ter}) (Fig. 1B) (Youngren et al., 2005). Understanding how DND1 is involved in stabilization of germ cell fate will provide important insights into how this critical process is effected in germ cells and, by analogy, in other stem cell populations from which tumors arise.

The function of DND1 has been studied by knockouts in zebrafish (Gross-Thebing et al., 2017; Weidinger et al., 2003) and mouse (Zechel et al., 2013), and by a conditional deletion at E13.5 in mouse (Suzuki et al., 2016), but none of these mutants formed teratomas, indicating that there is something unique about the *Dnd1*^{Ter} mutation and its

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

129/SvJ genetic background. We have focused our efforts on characterizing the transcriptome in germ cells from *Dnd1*$^{Tet/Ter}$ 129SvT2/SvEMSJ (129T2) mutant embryos just prior to the time when teratomas form (Cook et al., 2009; Cook et al., 2011; Heaney et al., 2012), and cross-referencing this information with direct binding targets of DND1.

Previous studies have implicated DND1 in both positive and negative regulatory roles. The molecular role of DND1 was first investigated in a human tumor cell line where the protein was shown to bind to the mRNAs of *Cdkn1b* (a negative regulator of the cell cycle) and *Lats2* (a tumor suppressor and negative regulator of p53) to protect these transcripts from miRNA-mediated translational repression (Kedde et al., 2007) perhaps through interaction with APOBEC3 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) (Bhattacharya et al., 2008). Other studies identified targets protected by DND1 in NIH3T3 or HEK293 cells, including the negative cell cycle regulator *Cdkn1b* (cyclin dependent kinase inhibitor 1B)(Cook et al., 2011) and *Ezh2* (enhancer of zeste homolog 2), a mediator of H3K27me3 repression (Gu et al., 2018). Subsequently, two labs showed that DND1 acts as a negative regulator of mRNAs by recruiting targets to the CCR4-NOT deadenylase complex for degradation (Suzuki et al., 2016); (Yamaji et al., 2017). Suzuki et al. showed that DND1 acts as an essential partner of NANOS2 (C2HC-type zinc finger 2, another male-specific RBP) to form P-bodies, and to select and recruit mRNA for degradation. In these studies, conditional deletion of *Dnd1* at E13.5 led to up-regulation of proliferation, meiosis markers, and Caspase3, and the down regulation of several NANOS2 targets associated with male differentiation, similar to *Nanos2* mutants (Suzuki et al., 2016).

Early studies relied on a candidate approach to identify targets of DND1, which included pluripotency factors and cell cycle genes (Cook et al., 2011; Kedde et al., 2007; Zhu et al., 2011). More recently, Yamaji and co-workers performed photoactivatable-Ribonuleoside-Enhanced Cross-Linking and Immunoprecipitation (PAR-CLIP) assays in HEK293 cells, and investigated the expression of identified targets in induced primordial germ like cells (PGLCS) and germ line stem cells derived from adult spermatogonia. Consistent with Suzuki's results, Yamaji et al. showed that DND1 acts with CCR4-NOT to destabilize and repress RNAs associated with apoptosis, inflammation and signaling pathways including

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TgfB, Wnt, and PI3K-AKT. Extrapolating results from these in vitro cell types, Yamaji et al. suggested that activation of pluripotency genes, as well as TgfB and inflammatory signaling genes, contribute to the formation of teratomas in gonadal germ cells (Yamaji et al., 2017).

We sequenced the transcriptome of male germ cells in *Dnd1*$^{Ter/Ter}$ mutants at E12.5, E13.5, and E14.5, just prior to the formation of teratomas, and correlated this information with direct targets of DND1. We identified direct transcript targets of DND1 using an independent method, Digestion Optimized-RIP-seq (DO-RIP-seq), after expressing a tagged DND1 protein in HEK293 cells. We cross-referenced our DO-RIP-seq targets with the PAR-CLIP results from Yamaji et al., and compared the transcript levels of nontargets and targets common to both RNA-immunoprecipitation assays in isolated germ cells from 129T2-*Dnd1*$^{Ter/Ter}$ and *Dnd1*$^{+/+}$ littermates at E12.5, E13.5 and E14.5. Consistent with our previous results and the molecular models proposed by Suzuki and Yamaji, between E12.5-E13.5 DND1 controls the down regulation of many genes associated with pluripotency and cell cycle, including elements of the both the Hippo and Bmp/Nodal signaling pathways. However, we found that DND1 targets include many genes associated with male differentiation including a large group of chromatin regulators normally activated during the second transition between E13.5 and E14.5. These results suggest multiple functions of DND1, and link DND1 to the initiation of epigenetic modifications in male germ cells.

Materials and Methods

Mice, timed matings, and genotyping

Dnd1$^{Ter/+}$ mice were kindly provided by Dr. Joseph Nadeau and maintained on a 129T2/SvEmsJ background. 129T2/SvEmsJ (hereafter referred to as "129T2"; https://www.jax.org/strain/002065) was predicted to be the closest living strain at The Jackson Laboratory (JAX) to the original 129/SvJ strain on which *Ter* arose. Because the formation of teratomas is a 129-strain-specific phenotype, to generate mutants in which germ cells were fluorescently tagged, Oct4-EGFP (originally on a mixed genetic background) (Yoshimizu et al., 1999) was back-crossed to 129T2-*Dnd1*$^{Ter/+}$ for 9

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

generations, then crossed to 129T2-*Dnd1*$^{Ter/Ter}$ mutant females to generate 129T2-*Dnd1*$^{Ter/+}$*;Oct4-EGFP* male and female heterozygotes. For timed matings, heterozygotes were intercrossed, females were checked for plugs, and staged as day E0.5 if positive. For genotyping, tail DNA was extracted using standard methods and genotyped using the following primer sets: GFP-F 5- AAG TTC ATC TGC ACC ACC G and GFP-R 5- TCC TTG AAG AAG ATG GTG CG and Ter-F 5-GTA GTT CAG GAA CTC CAC TTG TG-3, Ter-R 5-GCT CAA GTT CAG TAC GCA C-3. *Dnd1*^{Ter} mice were genotyped by PCR using an annealing temperature of 62°C. The PCR product (145 bp) was digested overnight at 37°C with the restriction enzyme *DdeI* and run on a 4% agarose gel or 6% acrylamide gel. *DdeI* digestion of DNA from mice with the *Dnd1*^{Ter} mutation produces 123 bp and 22 bp products.

FACS, RNA isolation, Library Preparation

Global gene expression was quantified by RNA sequencing (RNA-seq) in germ cells at multiple embryonic stages. To isolate germ cells, embryos were collected at E12.5, E13.5, and E14.5. The gonad was isolated from XY embryos, and incubated in 250 ml 0.25% Trypsin EDTA (Gibco #25200) at 37ºC for 5–10 minutes. The trypsin was removed and replaced with 400 ml PBS with or without 4 ml RNase-free DNase (Promega #M6101). The tissue was dissociated, and the cells were passed through a 20µm cell strainer (BD Falcon #352235). Fluorescence activated cell sorting (FACS) was performed by the Duke Comprehensive Cancer Center Flow Cytometry Shared Resource. Between 200-900 EGFP+ cells were sorted into 100 µl Qiagen Buffer RLT (from RNeasy Micro Kit, Qiagen #74004) +DTT, without carrier RNA and frozen at -80ºC.

Total RNA was extracted from purified germ cell aliquots and eluted in 13 µl of RNase/DNase free water using the Qiagen RNeasy Micro Kit (Qiagen #74004) with on column DNase digestion following manufacturer's protocols. Next, 10 µl of RNA eluate went directly into first strand synthesis, using oligo-dT primers according to the Smart-seq2 protocol (Picelli et al., 2014) scaled for input. Pre-amplification mix was prepared according to the Smart-seq2 protocol, scaled to 27 µl input from the first strand synthesis reaction. DNA was purified using the recommended Agencourt AMPure XP bead kit (Beckman Coulter #A63880), DNA was quantified by fluorometric quantitation (QUBIT).

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

0.5 ng from each sample was brought forward for tagmentation using the Nextera XT DNA Sample Prep Kit (illumina #15032352). Samples were size-selected using a 0.6 ratio of SPRIselect beads (Beckman Coulter #B23317) to DNA to obtain fragments >300 bp. Samples were eluted from beads according to protocol, and submitted to the Duke Sequencing Facility for QC by QUBIT and by traces run on an Agilent 2100 bioanalyzer to determine library sizes. The concentration of each library was normalized and 8 pmoles of the pool were sequenced using 50bp SR on the Illumina Hi-seq 2500. Eight samples (4 mutant and 4 wild type) were pooled for each stage and run in technical replicate in 2 lanes of a flow cell. The Duke Sequencing Facility used FastQC for preliminary quality control.

Transcriptome processing and mapping reads

There were ~22M reads/sample at E12.5, ~15M reads/sample at 13.5, and ~38M reads/sample at 14.5. Bowtie 2 was used to align the RNA-seq reads to mm9 using default settings. >81% of E12.5 transcripts aligned to the genome, whereas that number for >87% at E13.5, and >92% at E14.5. The Cufflinks pipeline (version 2.2.0) (Top-Hat, Cufflinks, Cuffmerge) was used to assemble transcripts using default settings. We considered only those transcripts previously annotated using Ensembl75. Gene levels were quantified using Cuffquant, and differential expression analysis was performed using Cuffdiff. This data was used for further analyses.

Filtering the datasets We found that some of the most significant outlier genes are strongly expressed in Leydig and Sertoli cells based on the Jameson datasets (Jameson et al., 2012) (Suppl. Fig. 1). Most of these genes were upregulated between E13.5 and E14.5 in mutant samples. We considered the possibility that this might reflect partial transdifferentiation of germ cells to express genes associated with somatic lineages, but using fluorescent immunocytochemistry we did not detect any co-localization of markers (data not shown). Because the intensity of *Oct4-EGFP* fluorescence declines, and the number of germ cells in mutants is reduced at E14.5, we suspected that contamination by somatic cell types in the gonad during fluorescence activated cell-sorting (FACS) accounted for this problem. As a means to eliminate any transcripts that might be due to contamination by another cell type, cell-type-specific microarray data from the Jameson

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

et al. study, for stages E11.5, E12.5 and E13.5, was used to generate lists of all genes enriched in the male germ cell population. Lists from E11.5 and E12.5 in the Jameson array were applied to the E12.5 and E13.5 transcriptome data sets. The list from E13.5 in the Jameson array was applied to the E13.5 and E14.5 transcriptome data sets. Only genes appearing on these lists were retained in the filtered RNA-seq data (1213 genes for E12.5-E13.5, and 1207 genes for E13.5-E14.5). The filtered datasets were replotted and used for this analysis. A similar method was used to generate lists of genes specifically enriched and depleted in all germ cells, female germ cells and male germ cells at E12.5 and E13.5 (E12.5 or E13.5 lists from Jameson Array applied to E12.5 or E13.5 transcriptome data sets respectively) . These lists were used for heat map analysis.

RepEnrich Analysis To evaluate the status of repetitive elements (specifically transposable elements), the original unfiltered raw data for wild type and *Dnd1*$^{Ter/Ter}$ samples were compared using RepEnrich (Criscione et al., 2014). The protocol for library preparation, mapping and differential enrichment analysis (Robinson et al., 2010) was followed as described by (Greco et al., 2016) and detailed by SW Criscione (https://github.com/nskvir/RepEnrich) with minor adjustments. Briefly, Bowtie 1.2.1 was used (Langmead et al., 2009) and the custom repetitive annotation was prepared using parameters: -p 6 -q -m 1 -S --max. For statistical analysis, a FDR (false discovery rate) of less than 0.05 was used as the cutoff for significance (Greco et al., 2016); (Criscione et al., 2014).

Gene Set Enrichment Analysis and Ingenuity Pathway Analysis

Gene set enrichment analysis was performed using GSEA v3.0, build 0160 (http://software.broadinstitute.org/gsea/) as described in (Subramanian et al., 2005). "Log2 ratio of classes" was used to rank changes in gene expression between developmental stages as well as between wild type and *Dnd1*$^{Ter/Ter}$ samples. Normalized Enrichment Scores (NES) and FDR values were used to describe the magnitude of change and the significant enrichment of the gene sets respectively.

Ingenuity Pathway Analysis. Differentially expressed genes (DEGs) were assessed for statistical enrichment/depletion of specific pathways using the Qiagen Ingenuity Pathway

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Analysis (IPA) database (Application build: 470319M; Content version: 43605602). DEGs included those that were differentially expressed 1) between developmental stages within a *Ter* genotype class, and 2) between *Ter* genotype classes at a specific stage. The cutoff for inclusion in the IPA analysis was set at a differential expression FDR < 0.01.

Dnd1 DO-RIP-seq. DO-RIP-seq was performed as described in (Nicholson et al., 2017). Briefly, 5 - 15 cm^2 plates of HEK293 cells were transfected with ~35 µg/plate FLAG-DND1 or untagged DND1 (as a negative control), grown for 24hrs, then harvested in PLB. Lysates were treated with 50 U/µg of MNase for 5 min at 30^{o}C, and FLAG immunoprecipitations were performed. RNA from IPs was radiolabeled and run on a 10% TBE-Urea gel. RNA fragments from ~25-75 nucleotides were excised and extracted from the gel. cDNA libraries were prepared using the NEBNext Multiplexing Small RNA Library Prep Set for Illumina with 14-17 rounds of amplification, following the manufacturer's protocol with the exception that the 5' RNA adapter was replaced with a custom 5' adapter compatible with UMI (Unique Molecular Identifier) tagging (Kivioja et al., 2011) (RNA adapter sequence GUUCAGAGUUCUACAGUCCGACGAUCNNNNN). In addition to DND1, negative control IPs input samples were also prepared for sequencing in a similar fashion using 17 rounds of amplification. An aliquot of the MNase-digested lysate prior to IP was set aside for generating input libraries. rRNA depletion was performed on input samples using the Epicenter Ribo-Zero Gold rRNA removal kit. Input RNA was size selected and libraries were prepared as described above for the IP samples. All experiments were performed as biological duplicates.

Libraries were sequenced using Illumina HiSeq 2500. Processing of data was performed as described in Nicholson et al. After removal of adapter and UMI sequences, reads were mapped to hg19 using TopHat2. Binding sites were identified using the binning and normalization procedure of Nicholson et al. (Nicholson et al., 2017), IPs were compared to input samples to identify sites enriched in DND1 IPs. Reads with a low pass filter of >3 reads and >1.5 fold enrichment over input were considered DND1 binding sites.

Results

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$Dnd1^{Ter/Ter}$ mutant germ cells are delayed in the germ cell transcriptional profile relative to wild type at E12.5. To investigate the changes in $Dnd1^{Ter/Ter}$ mutant germ cells relative to wild type, 129T2-$Dnd1^{Ter/+}$; *Oct4-EGFP* males and females were intercrossed, and ~600 germ cells were isolated by FACS from 4 independent homozygous and wild type fetal gonads at each of 3 stages, E12.5, E13.5 and E14.5, after germ cells have populated the gonads, and prior to teratoma detection (Fig. 1A; (Cook et al., 2011)). Sequencing was performed, reads were mapped to the genome, and a reads per kilobase transcript per million (RPKM) value was calculated for each transcript. RPKM values used in our analysis are provided with the capability to generate an expression graph for any gene, as a user-friendly resource for the community (Suppl. Table 1).

The *Dnd1* transcript was detected at similar levels in wild type and homozygous mutant germ cells at E12.5 (Fig. 1B), but showed a rapid decline at later stages (data not shown). The *Ter* mutation, predicted to lead to a premature stop codon (Youngren et al., 2005) is evident in the third exon of *Dnd1* and efficiently distinguishes wild type and mutant samples (Fig. 1B, blue arrow). The 129T2 variant of a SNP that differs between B6 and 129 was detected in the 3'UTR of all samples confirming the 129T2 strain background (red arrow).

Dnd1 is expressed in both XX and XY germ cells from early migratory stages (Youngren et al., 2005). In $Dnd^{Ter/Ter}$ mutants, many germ cells are lost prior to arrival in the gonad (Cook et al., 2009; Noguchi and Noguchi, 1985). Given that loss of RBPs in germ cells of other species results in transdifferentiation to somatic fate (Ciosk et al., 2006; Gross-Thebing et al., 2017; Updike et al., 2014), we investigated the possibility that $Dnd1^{Ter/Ter}$ mutant germ cells were significantly altered by the time they populated the gonad, and did not retain their germ cell identity. The top two principle components account for 43% and 25% of the observed variation in the top 500 most variable protein-coding genes. Based on the variability captured by Principle Component (PC) 1, it appears that E13.5 mutant germ cells cluster more closely with E12.5 wild type germ cells while E12.5 mutant germ cells are delayed. Further, PC2 appears to stem mainly from expression variation in E14.5 $Dnd1^{Ter/Ter}$ germ cells, suggesting that the transcription profile in these mutant

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

germ cells continues to diverge from a normal trajectory (Fig. 1C). To investigate these differences further, we compared expression in *Dnd1*$^{Ter/Ter}$ mutant germ cells with a list of genes specific to germ cells (Fig. 1D) and a list of genes that distinguish *male* germ cells (Suppl Fig. 2A) at E12.5 and E13.5. These lists include genes that characterize germ cells by their specific expression in this lineage, and those that characterize germ cells by their specific repression (eg. somatic genes)(Jameson et al., 2012). Based on heat map analysis, *Dnd1*$^{Ter/Ter}$ mutant germ cells show an overall pattern similar to wild type male germ cells for both up- and down-regulated genes at E12.5 and E13.5. For example, both *Ddx4* (dead box helicase 4)(*Vasa*) and *Dazl* (deleted in azoospermia like) show a delay in activation, similar levels at E13.5, and a down regulation at E14.5, while *Stella (Dppa3* (developmental pluripotency associated protein 3)*)*, a germ cell specific gene, and *Prdm1* (PR/SET domain 1) (*Blimp1*) show no significant differences from wild type at any stage (Fig. 1E). These data indicate that although *Dnd1*$^{Ter/Ter}$ germ cells have a delayed gonadal differentiation program, they have a transcriptome similar to wild type germ cells soon after they arrive in the gonad.

Neither the female germ cell program nor other embryonic pathways are activated in XY *Dnd1*$^{Ter/Ter}$ germ cells. Because *Dnd1* expression becomes specific to the male pathway after germ cells enter the gonad (Youngren et al., 2005), we next investigated whether there was evidence for a shift to the female germ cell-specific program in *Dnd1*$^{Ter/Ter}$ male germ cells by comparing expression in mutants to a list of genes specific to female germ cells (Jameson et al., 2012). A heat map comparison showed little evidence for ectopic activation of female pathway genes (Suppl. Fig. 2B). Based on a cumulative distribution plot, few female-specific genes were significantly altered in mutants at E13.5 (Suppl. Fig. 2C). In contrast to their pattern in wild type female germ cells, genes that are associated with entry into meiosis (eg. *Stra8* (stimulated by retinoic acid 8)*, Sycp3* (synaptonemal complex protein 3)) were not upregulated in *Dnd1*$^{Ter/Ter}$ mutant XY germ cells (Suppl. Fig. 2D). However, some genes that become female-specific by downregulation in male germ cells (eg. *Rhox9* (reproductive homeobox 9) and *Rhox6*) showed a less robust downregulation at the stage when male and female germ cell pathways diverge (E12.5) (Suppl. Fig. 2E). In general, these were genes associated with the pluripotent state of early germ cells.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Many genes associated with developmental pathways carry bivalent marks in germ cells (Lesch and Page, 2014). Because teratoma development is associated with the upregulation of developmental pathways, we investigated whether specific groups of bivalent genes showed upregulation in *Dnd1*$^{Ter/Ter}$ germ cells. We found no evidence of lineage infidelity based on the activation of any subgroup of bivalent genes associated with embryonic pathways (Suppl. Fig. 3A), or significant differences in specific factors associated with differentiation (Suppl. Fig. 3B).

Homozygosity for the *Dnd1*^{Ter} mutation is associated with subtle changes in expression of many genes in male germ cells between E12.5-E14.5.

Next we compared the effect of the *Ter* mutation on the temporal gene expression of germ cells during the transition between E12.5 and E13.5 and between E13.5 and E14.5. While most expressed genes retain a similar temporal pattern in the mutant compared to wild type over this period, hundreds are up- and down-regulated between E12.5-13.5 and/or E13.5-14.5 (Fig. 2A). More genes are up-regulated in mutant PGCs relative to wildtype during the first transition between E12.5-13.5 (151 up v. 58 down), whereas more genes are down-regulated in mutant PGCs during the second transition between E13.5-14.5 (43 up v. 266 down). Transcripts for many pluripotency genes are enriched in mutant PGCs throughout the time course (Fig. 2B). A large group of genes associated with the initiation of male germ cell development, including *Nanos2*, *Nanos3* (C2HC-type zinc finger 3), *Ret* (rearranged during transfection – a receptor tyrosine kinase), *Ddx25* (dead box helicase 25), and *Dppa4* (developmental pluripotency-associated protein 4), a marker of pluripotent cells required for differentiation (Madan et al., 2009) fail to be activated in mutants (Fig. 2C). Based on Gene Set Enrichment Analysis (GSEA), targets of DPPA4 (including *Gtsf1* (gametocyte specific factor 1), *Stk31* (serine threonine kinase 31), *Ddx4* (dead box helicase 4), *Iqcg* (IQ motif containing G), *Mael* (maelstrom spermatogenic transposon silencer), *Slc25A31* (solute carrier famly 25), *Rnf17* (ring finger protein 17), and *Mep1b* (meprin A subunit beta), Suppl. Table 2) are enriched in wildtype PGCs (Fig. 2D, left panel) but not in mutant PGCs (Fig. 2D, right panel) during the E13.5-E14.5 transition, further suggesting that the *Ter* mutation disrupts *Dppa4* expression and its

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

339 downstream function in PGCs.

340

341 Given the critical roles of *Nanos2* in establishing the male pathway in mouse germ cells
342 (Suzuki and Saga, 2008), we asked whether most transcriptional changes in *Dnd1*$^{Ter/Ter}$
343 mutants could be accounted for by failure to activate *Nanos2* (Fig. 2C). To investigate this
344 possibility, we compared the transcriptome changes between the two mutants. Although
345 some genes changed in a similar manner (eg. genes associated with pluripotency), many
346 genes exhibit opposing patterns (Fig. 2E). For example, *Ret* was not affected in *Nanos2*
347 mutants whereas *Sycp3* and *Stra8* were elevated (Saba et al., 2014). In *Nanos2* mutants,
348 *Nan*os3 was upregulated, whereas in *Dnd1* mutants, it was not (Fig. 2C; see Suppl. Table
349 3 and for full data set). This indicates that DND1 is upstream of both Nanos genes, and is
350 consistent with the finding that *Nanos3* can partially rescue the *Nanos2* mutants (Saba et
351 al, 2014).
352

353 **Transposable Elements and piRNA pathways were not strongly affected by**
354 **E14.5, despite the failure to activate their repressors.**

355 Several genes involved in repression of transposable elements including *Piwil4* (piwi like
356 RNA-mediated gene silencing 4), *Tdrd5* (tudor domain containing 5), *Morc1* (MORC
357 family CW-type zinc finger 1), and *Mael* were either not activated or expressed at much
358 lower levels in mutants (Fig. 2F). Proteins encoded by these genes normally silence LINE
359 and other transposable elements. Our transcriptome data was based on oligo-dT priming,
360 however some transposable elements contain polyadenylation sites (Lee et al., 2008). We
361 interrogated the raw transcriptome data at E14.5 for activation of transposable/repetitive
362 elements using RepEnrich, a program optimized to detect differential enrichment
363 between two RNA-seq datasets (Criscione et al., 2014). Very few transposable element
364 reads showed a significant difference between wild type and mutant germ cells (FDR <
365 0.05) (Fig. 2G). This trend of similarity between wild type and mutant germ cells was also
366 observed when we interrogated repetitive elements and transposable elements more
367 generally (Suppl. Fig. 4A,B).
368

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DND1 was previously reported to protect transcripts from miR470 (in zebrafish) and mir-1, miR-206, and miR-221 (in human cell lines) -mediated degradation (Kedde et al., 2007). Therefore, we predicted that targets of a mouse ortholog of these miRNAs such as miR302 (the ortholog of zebrafish miR470) would be down-regulated in the absence of DND1 protection. However, only *Lats2* (large tumor suppressor kinase 2) showed this behavior; other targets of miR302 were elevated in *Dnd1*$^{Ter/Ter}$ germ cells relative to wildtype (Suppl. Fig. 5A). We used Ingenuity Pathway Analysis (IPA-Qiagen) to infer activation of individual miRNAs on a global basis. IPA compares the behavior of sets of miRNA targets among samples, calculates a p-value based on this overlap, and returns a Z-score reflecting the probability that a particular miRNA pathway is activated or repressed based on the expression of target genes. Based on IPA analysis, only one miRNA (mir223) cleared the activation cut-off of 2 between E13.4 and E14.5 in mutant cells (Suppl. Fig. 5B). However, investigation of the expression of predicted targets of mir223 revealed no consistent pattern (data not shown). Overall, we discovered no strong miRNA regulatory signature in this mRNA-seq analysis.

Identification of DND1 direct targets. We identified hundreds of genes that were up- and down-regulated between E12.5-14.5 in DND1 mutant PGCs (Fig. 2A). DND1 has been shown to both promote degradation of target transcripts (Suzuki et al., 2014; Yamaji et al., 2017) and to stabilize other target transcripts by competing for miRNA binding sites (Kedde et al., 2007; Zhu et al., 2011). We were interested in whether genes in both up- and down-regulated categories were direct targets of DND1.

Dnd1 shares sequence similarity with genes encoding five other RNA binding proteins (*Rbm46* (RNA binding motif protein 46), *A1cf* (APOBEC complementation factor), *Rbm47*, *Syncrip* (synaptotagmin binding cytoplasmic RNA interacting protein) and *Hnrnpr* (heterogeneous nuclear ribonucleprotein R). All members of this subfamily contain three RNA recognition motifs (RRMs) except DND1 which contains two RRMs. Additionally, three members of the family, including DND1, contain a C-terminal double stranded RNA binding domain (DSR). (Fig. 3A). The *Ter* mutation generates a premature stop codon in the middle of RRM2 (arrow). To identify binding sites of DND1 we performed DO-RIP-seq (Nicholson et al., 2017) in HEK293 cells, using a tagged mouse

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DND1. The most enriched motif in DND1 binding sites was DAUBAW (D=A, G or U, B=C, G or U and W=A or U), a motif that is very similar to the known RNA motif of RBM47, its closest relative (Fig. 3B). The RBM47 motif defined by RNA-Compete (Ray et al., 2013) is GAUSAW (S=G or C and W=A or U), both the DND1 and RBM47 motifs contain AU in the second and third position and AW in the fifth and sixth, but differ in the first and fourth position where the recognition sequence is more degenerate. Although *Rbm47* expression declines sharply between E12.5 and E14.5, *Rbm46*, *Rbm47*, *Syncrip*, and *Hnrnpr* are all expressed in germ cells in the gonad, and could partially compensate for loss of DND1 (Suppl. Fig. 6).

Compared to their proportional representation across the whole transcriptome, DND1 binding sites were strongly over-represented in 3' untranslated regions (UTRs) and to a lesser extent in the coding sequence (CDS) (Fig. 3C). Targets identified by DO-RIP-seq largely agree with PAR-CLIP identified DND1 targets from Yamji et al, with 2953 mRNAs shared in common (Yamaji et al., 2017)(Fig. 3D). This represents a significant overlap ($p<0.001$) based on 1000 iterations of randomly shuffled binding sites matched for size and restricted to expressed 3'UTRs.

Direct targets of DND1 are both up- and down-regulated in *Dnd1*$^{Ter/Ter}$ mutants. Because RIP assays were not conducted in germ cells due to the absence of appropriate antibodies or transgenic tags, we first identified the set of conserved targets expressed in germ cells (1347 out of 2953 targets). Next we investigated the relationship between number of DND1 binding sites in the target transcript (1 to 9+ binding sites) and its expression in wild type and 129T2 *Dnd1*$^{Ter/Ter}$ mutant germ cells between E12.5-E13.5 and E13.5-E14.5. Cumulative distribution function (CDF) plots were generated comparing RNA abundance changes for transcripts classified by number of DND1 binding sites (non-targets, 1 site, 2 sites, etc). Changes in RNA abundance were evaluated in this manner for four conditions: E12.5-E13.5, and E13.5-E14.5 in both mutant and wild type germ cells. In wild type germ cells, the presence of more DND1 binding sites in a target transcript was correlated with a higher likelihood of down-regulation relative to transcripts with few or no target sites during the transition between E12.5-E13.5, while the opposite pattern was observed between E13.5-E14.5 – target

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

transcripts with more DND1 binding sites were more likely to be up-regulated relative to non-targets. During this later transition, DND1 targets tend to change faster than the rest of the transcriptome (Fig. 3E). In 129T2-*Dnd1*$^{Ter/Ter}$ mutants, DND1 targets behave more like non-targets, reflecting the effect of the *Ter* mutation on DND1 function.

Overall, among targets of DND1, cell cycle genes and chromatin regulators were over-represented (Fig. 4A; Table 4). Overlapping binding sites from all three RIP assays were present in the 3'UTR of the cell cycle gene, *Ccne1*, and throughout the CDS and 3'UTR of the chromatin regulator, *Kat7* (lysine acetyltransferase 7) (Fig. 4B). In addition to *Ccne1*, *Rheb* (Ras homolog enriched in brain) and *Rhob* (Ras homolog family member B), two GTP-binding proteins associated with active cell cycle were targets of DND1, and were elevated in the mutant transcriptome along with *Trp57* (*Cdkn1c*, cyclin dependent kinase inhibitor P57), which was not expressed in HEK293 cells (Fig. 4C). *Kat7* and other chromatin regulators, including *Setdb1* (set domain bifurcated 1), *Kdm5a* (lysine demethylase 5a), and *Rbbp5* (RB Binding Protein 5, Histone Lysine Methyltransferase Complex Subunit) mapped as targets, and all lagged behind wild type levels (Fig. 4D), as did many other chromatin regulators not expressed in HEK293 cells (Suppl. Fig. 7).

Many genes associated with signaling pathways were identified as targets of DND1, and were significantly changed in mutants (Fig. 5A). A large group of genes associated with the BMP signaling pathway, were elevated in mutant germ cells (Fig. 5B), as also reported by Yamaji (Yamaji et al., 2017). We identified elements of other signaling pathways elevated in mutants that map as targets of DND1, including *Ywhah* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein eta), *Ywhag* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma), *Csnk1e* (casein kinase 1 epsilon), *Yap1* (yes-associated protein 1), and *Prkci* (protein kinase C lota) in the Hippo pathway (Fig. 5C). In contrast, transcripts for two central genes in the mTOR pathway that mapped as direct targets of DND1, *Rictor* (rapamycin-insensitive companion of mTOR) and *Mtor* (mammalian target of rapamycin), were down-regulated in mutants (Fig. 5D).

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ingenuity Pathway Analysis identified apoptotic and cell cycle pathways as strongly affected in mutants. We used Ingenuity Pathway Analysis (Qiagen) to identify curated pathways most strongly affected in the mutant transcriptome. This analysis resulted in the identification of Cellular Growth and Proliferation (482 genes) and Cell Death and Survival (433 genes) as the most affected pathways (Suppl. Fig. 8A-C). These results are consistent with previous findings that many *Dnd1*$^{Ter/Ter}$ germ cells do not enter cell cycle arrest, but are eliminated through apoptotic pathways (Cook et al., 2009; Dawson et al., 2018; Yamaji et al., 2017).

Discussion

Germ cells are an important model to investigate how stem cells embark on a highly specific differentiation pathway while restraining their potential for spontaneous differentiation or tumorigenesis. While most germ cells appear to strictly control this potential and successfully navigate this transition, a small subset of 129T2 *Dnd1*$^{Ter/Ter}$ mutant germ cells fail this test and spontaneously differentiate into teratomas - and in so doing provide a tool to investigate the genetic regulation of germ cell differentiation. What are the critical roles of DND1, and why does mutation in this RBP lead specifically to teratoma development in a subset of mutant germ cells? To address these questions, we investigated the binding targets of the DND1 protein and analyzed the transcriptome of *Dnd1*$^{Ter/Ter}$ germ cells relative to wild type over the critical period following residency in the gonad but prior to overt transformation. We found that DND1 down-regulates genes associated with pluripotency, including multiple components of the Hippo and Bmp pathways, mediates levels of many genes involved in the control of cell cycle, and is critical to initiate male germ cell differentiation, which involves the activation of many chromatin regulators.

DND1 is expressed from the time germ cells are allocated at the base of the allantois, and as a result of its early role in germ cell development, many germ cells undergo apoptosis, or are otherwise lost, prior to arrival in the gonad (Cook et al., 2009; Noguchi and Noguchi, 1985). However, this transcriptome analysis reveals that overall, the

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

transcriptome of the group of mutant germ cells that survive to populate the gonad at E12.5 is similar to wild type male germ cells, albeit somewhat delayed.

DND1 likely performs several roles in posttranscriptional regulation. E13.5 is a clear transition point between the embryonic and fetal programs: functions of DND1 may segregate into early exit from the pluripotency program, involving the degradation of transcripts for many pluripotency genes, and subsequent activation of transcripts associated with the male pro-spermatogonia pathway. Although we could not detect pluripotency targets in HEK293 cells, Matin and co-workers previously reported *Oct4, Nanog, Sox2, LIN28* (LIN28 family RNA-binding protein)*, Bax* (BCL2 Associated X apoptosis regulator), and *Bclx* (B-cell lymphoma-estra large, apoptosis regulator) as targets of DND1 in ESCs (Zhu et al., 2011). A large group of chromatin regulators mapped as targets of DND1 in HEK293 cells, and all of these genes were strongly down-regulated in *Dnd1*$^{Ter/Ter}$ mutants. However, it is unclear whether this down-regulation is the result of post-transcriptional control or a failure to launch the male-specific differentiation program.

There is strong evidence that DND1 interacts with the CNOT degradation complex, and provides target specificity for NANOS2 (Suzuki et al., 2014; Yamaji et al., 2017). Because *Nanos2* is not expressed in mutants, we asked whether the *Dnd1*$^{Ter/Ter}$ mutant transcriptome phenocopied the *Nanos2* mutant transcriptome (Saba et al., 2014). We found that many pluripotency genes showed a similar delay in down regulation during the early transition (E12.5-E13.5). However, *Nanos2* mutant germ cells enter meiosis as though they have lost their male identity (Suzuki and Saga, 2008), whereas *Dnd1*$^{Ter/Ter}$ germ cells do not. Key differences may be the up-regulation and partial compensation by *Nanos3* in *Nanos2* mutants, and/or the retention of *Ret*. Neither *Nanos3* nor *Ret* is expressed in *Dnd1*$^{Ter/Ter}$ germ cells.

A key finding of this study is the global identification of many chromatin regulators that are activated in wild type male germ cells between E13.5-E14.5, and the observation that nearly all of these fail to be activated in *Dnd1*$^{Ter/Ter}$ mutants. If the role of these chromatin regulators is to stabilize the spermatogonial genome and silence somatic differentiation

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pathways, as has often been proposed (Dawson et al., 2018; Gu et al., 2018; Heaney et al., 2012), the failure to activate these genes could explain the tendency of mutant germ cells to initiate differentiation to multiple cell types characteristic of teratomas. The failure to activate chromatin regulators was not reported in the original *Nanos2* mutants (Saba et al., 2014). However, heterozygosity for a CRISPR/*Cas9*-mediated knock out of *Nanos2* on the 129/SvJ background led to a decrease in DNMT3L expression, a defect in re-methylation of the male germ cell genome, elevated levels of *Line-1*, and an increase in teratoma incidence (Dawson et al., 2018).

Many genes associated with the cell cycle are targets of DND1 and are dysregulated in mutants (this study;(Cook et al., 2011; Yamaji et al., 2017). Several signaling pathways, including Kit, Fgf, Nodal, and Hippo are also elevated and could be responsible for driving ongoing proliferation. The association of FOXH1 with the YAP/SMAD complex is believed to drive the switch from pluripotency to differentiation of mesendoderm (Slagle et al., 2011), but further studies would be required to determine whether this finding has relevance for the differentiation of cell types in teratomas. A failure to exit active cell cycle and enter G0 arrest is a characteristic of mutant germ cells and teratoma formation, but whether this failure is causal remains unclear.

Based on the failure to activate many repressors of transposable elements such as *Piwil4, Morc1, Mael*, and *Tdrd5*, we anticipated a strong activation of LINES and other transposons. However, at the latest stage of our analysis (E14.5), we did not detect significant changes in this category. Similarly, despite the absence of many epigenetic regulators that normally characterize the male germ cell transcriptome at these stages, we found no evidence of lineage infidelity. One possibility is that E14.5 (the latest stage of this analysis) is too early to detect these changes at the population level. There is significant variability in the response of individual germ cells to the $Dnd1^{Ter}$ mutation. $Dnd1^{Ter}$ mutant germ cells transform and form teratomas at a much higher rate than wildtype germ cells. However, even though the "per mouse" teratoma rate is high, the "per germ cell" transformation rate is still quite low. Most $Dnd1^{Ter/Ter}$ germ cells undergo apoptosis, as evidenced by the strong footprint of cell death and cell cycle disruption at the population level. A molecular explanation for this cell heterogeneity remains elusive.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

559

560 For stem cell biology, it would be very instructive to know the steps involved in the
561 transition from a germ cell to a teratoma. Dawson and colleagues have recently suggested
562 that this involves a transition through a primed pluripotent EC cell state (Dawson, 2018).
563 The fact that only a few germ cells form teratomas in a *Dnd1*$^{Ter/Ter}$ testis, while other germ
564 cells nearby either die or continue to express normal germ cell markers (Cook et al., 2011;
565 Dawson et al., 2018), suggests that the trigger must be a threshold event that is influenced
566 by cell autonomous fluctuations of factors in germ cells and/or by the microenvironment
567 in the testis of some strains. Teratoma formation likely competes with apoptotic pathways
568 that are strongly activated at all stages of our analysis. In this study, we have identified
569 many targets of DND1 and captured their transcriptional changes in *Dnd1*$^{Ter/Ter}$ germ
570 cells. However, the specific transcriptional events in the rare cells undergoing the
571 transition to teratoma may be obscured by the transcriptomes of the large number of cells
572 that are not. We plan to use a single cell analysis in future experiments to resolve this
573 problem.

574

575 **Acknowledgments**

576 We would like to thank the teams in the FACS and Genomics Cores at Duke, especially
577 Mike Cook, Nicolas Devos, and Olivier Fedrigo, who helped with these experiments and
578 their analysis. We are also grateful to Alex Bortvin, who provided information for the
579 RepEnrich analysis, to Joe Nadeau, who supplied the mice that founded our *Dnd1*^{Ter}
580 colony, and to past and present members of the Capel lab, especially Jordan Batchvarov,
581 who contributed to the management of the strain over the last 10 years and Mike
582 Czerwinski, who transferred bioinformatics expertise to JG. This work was funded by a
583 grant to BC from NIH (GM087500), Bridge Funding from DUMC, and a grant to JK from
584 NCI (CA157268) and Bridge Funding from DUMC.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

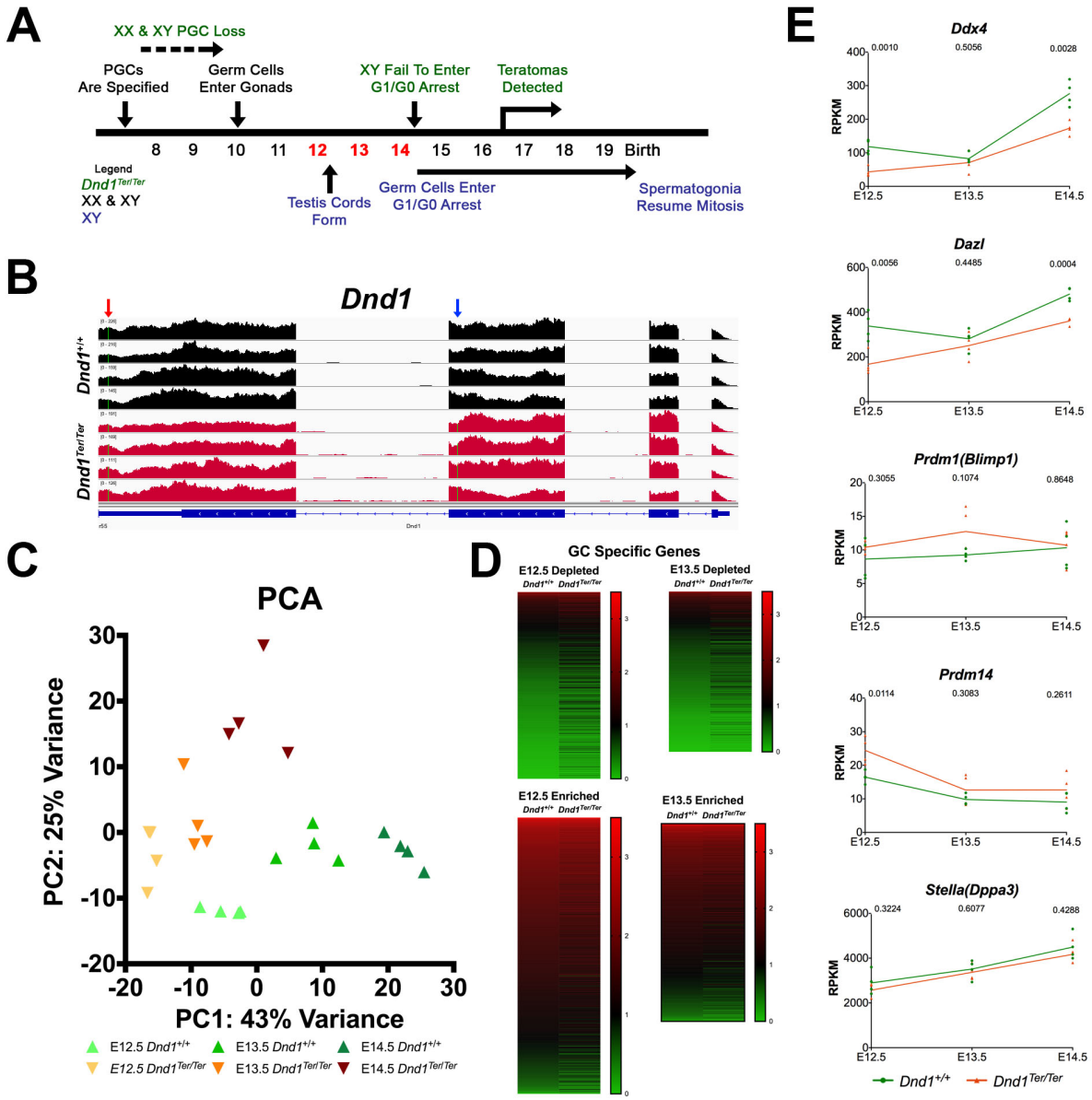


Figure 1. Male germ cells in *Dnd1*$^{Ter/Ter}$ mutants retain a germ cell identity. **(A)** Time line of male germ cell development and defects in *Dnd1*$^{Ter/Ter}$ mutants. Primordial germ cells (PGCs) are specified at ~E7.5, migrate and arrive in the gonad at E10.5. In the male gonad, they are enclosed inside testis cords by E12.5, and undergo G0 arrest between E14.5-E15.5, remaining in arrest until after birth. In *Dnd1*$^{Ter/Ter}$ mutants, many PGCs are lost soon after specification. Remaining germ cells enter the gonad, but fail to arrest in G0. Teratomas are detected in the fetal testis beginning at ~E16.5 (shown

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in green above time line). **(B)** *Dnd1* is transcribed in mutants similar to wild type levels at E12.5. The *Ter* mutation is visible in the 2nd exon of all mutant samples (blue arrow, red samples), whereas the 129T2-specific SNP (red arrow) is present in all samples in the 5' non-coding region. **(C)** A PCA analysis revealed that mutant germ cells are more similar to wild type at E12.5 and E13.5 and diverge further at E14.5. **(D)** Nonetheless, male germ cells show an overall expression pattern characteristic of germ cells with respect to genes that are normally depleted (top) and genes that are normally up-regulated (bottom) at E12.5 and E13.5. **(E)**. Specific examples of gene expression across the timecourse (green line = wild type samples; orange line = $Dnd1^{Ter/Ter}$ in all graphs in the manuscript; significance values for each time point are shown at top of graph). While both *Ddx4* (*Vasa*) and *Dazl* show a delay in activation, similar levels at E13.5, and a slight down regulation at E14.5, *Prdm1* (*Blimp1*), *Prdm14*, and *Stella (Dppa3)* show no significant differences from wild type at any stage.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

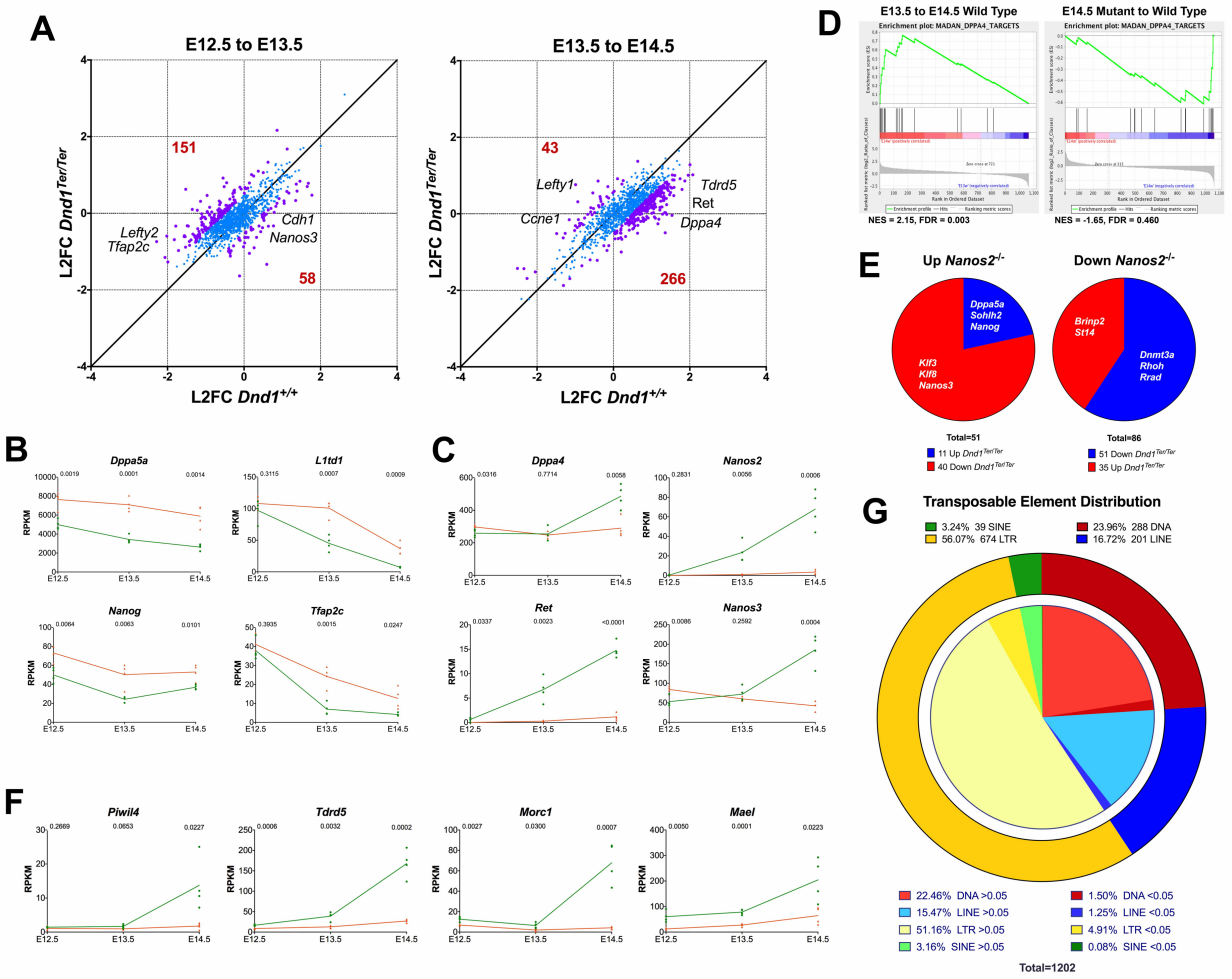


Figure 2. *Dnd1*$^{Ter/Ter}$ mutant germ cells show delayed suppression of pluripotency, and a failure to activate the male differentiation program. (A) Scatter plots illustrating dynamic changes in expression of germ cell specific genes (data from Jameson et al., 2012) between mutant (Y axis) and wild type (X axis) during the transitions between E12.5-E13.5 and E13.5-E14.5. Between E12.5-E13.5 more genes (151) show a higher $\log_2$ fold change (L2FC) in mutants than in wild type, while only 58 genes lag behind. In the transition between E13.5-E14.5, the reverse is true: more genes lag behind wild type expression (266). Of the 43 genes that show a higher L2FC in this interval, many belong to cell cycle or elevated signaling pathways. Points with L2FC between mutant and wild type >0.5 in purple, <0.5 in blue. **(B)** Examples of 4 genes associated with pluripotency (*Dppa5a*, *L1td1*, *Nanog*, and *Tfap2c*) that show a delay in down-regulation in *Dnd1*$^{Ter/Ter}$ mutant germ cells. **(C)** Examples of 4 genes associated

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with initiation of the male pathway (*Dppa4, Nanos2*, *Ret*, and *Nanos3*) that fail to be activated. **(D)** Consistent with this, based on GSEA analysis, targets of *Dppa4* are not enriched in mutant germ cells during the E13.5-E14.5 transition. **(E)** Some transcriptional changes in *Nanos2* mutants are discordant with changes in *Dnd1*$^{Ter/Ter}$ mutants. Of genes up in *Nanos2* mutants, 11 behave the same (blue) and 40 (red) are down in *Dnd1*$^{Ter/Ter}$. Of genes down in Nanos2 mutants, 51 are also down in *Dnd1*$^{Ter/Ter}$ (blue) and 35 are up (red)(Saba et al., 2014). Only genes with significant L2FC are shown ($p<0.05$). See also Suppl. Table 3. **(F)** Examples of genes associated with suppression of transposable elements and the piRNA pathway activated in wild type but not mutants. **(G)** Although many silencers of repetitive elements were down-regulated, RepEnrich analysis showed that transposable elements did not exhibit significant aberrant activity at E14.5. Transposable element distribution organized by class: Outer ring: class distribution of elements represented in reads at E14.5 (yellow=LTRs; green=SINEs; Red=DNA transposons; Blue=LINEs). Inner circle divides specific classes into reads with significant differences between mutant and wildtype (bold colors: FDR<0.05; pale colors: FDR>0.05).

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

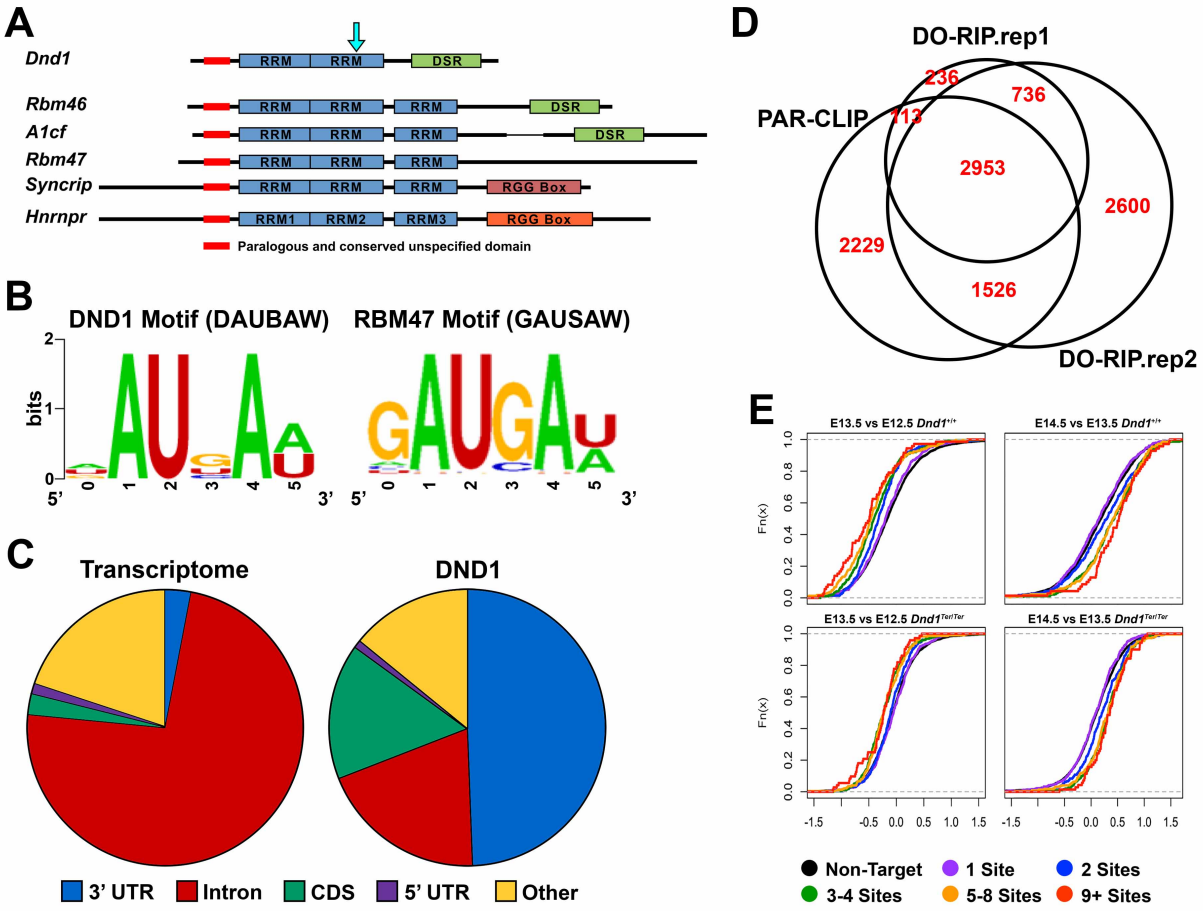


Figure 3. DND1 binds to the DATBAW motif preferentially localized to 3' UTRs, and regulates the up- and down-regulation of targets. (A) DND1 and five other RBPs (RBM46, A1CF, RBM47, SYNCRIP and HNRNPR) are part of a subfamily containing RNA recognition motifs (RRMs) sharing sequence similarity. Three of these subfamily member (DND1 included) also have similar C-terminal double stranded RNA binding domains (DSR) The black arrow indicates the position of the Ter mutation, partially truncating the second RRM and completely removing the DSR. **(B)** k-mer analysis was preformed to identify the most enriched motif in Dnd1 DO-RIP-Seq binding sites (http://weblogo.berkeley.edu), this motif is highly similar to a previously identified motif of a related RBP, Rbm47 (Ray et al., 2013). **(C)** The locations of DND1 binding sites were broken down into regions (3'UTR, intron, CDS, 5'UTR, and other) and compared to the proportion of the regions present in the expressed transcriptome. DND1 has a strong preference for binding 3'UTR sequences and a slightly weaker preference for coding

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sequence (CDS). **(D)** Comparison of DND1 binding sites identified in the two replicates of the present study and binding sites identified by PAR-CLIP in Yamaji et al., 2017 (Yamaji et al., 2017)**(E)** Cumulative distribution function (CDF) plots evaluating RNA abundance under four conditions (E12.5-E13.5, and E13.5-E14.5 in both mutant and wild type germ cells) with changes for transcripts classified by the number of DND1 binding sites (non-targets, 1 site, 2 sites, etc). In wild type germ cells RNA abundance of DND1 targets is more likely to decrease between E12.5 and E13.5 than for non-targets. In contrast, between 13.5 and 14.5 RNA abundance of DND1 targets generally increases. The effects are amplified by increasing the number of DND1 binding sites. In *Dnd1* mutant germ cells, RNA abundance changes for both E12.5-E13.5 and E13.5-E14.5 are similar for DND1 targets and non-targets.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

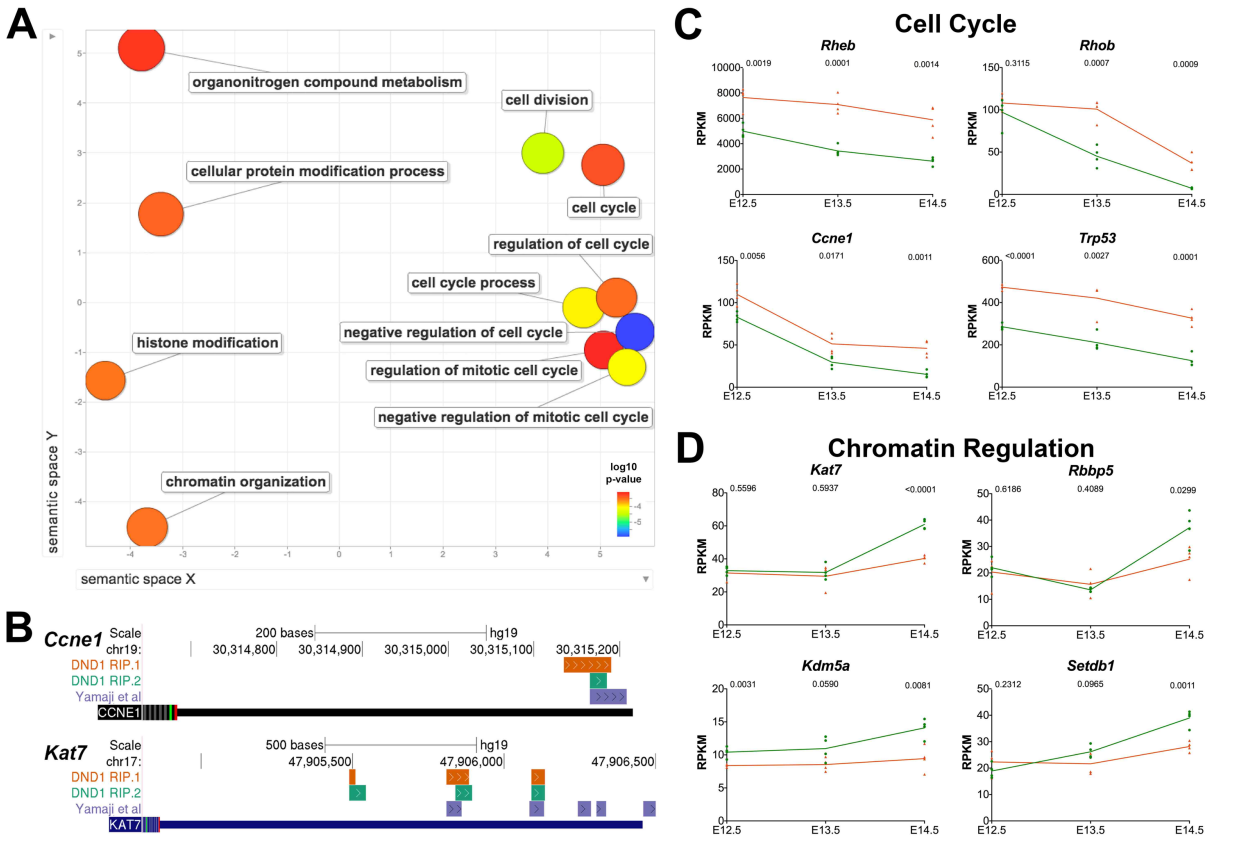


Figure 4. DND1 targets include cell cycle genes and chromatin regulators, which are both up and down-regulated in mutant germ cells. (A) Semantic similarity clustering of GO terms enriched in DND1 DO-RIP-seq targets indicated that the most common categories of targets of DND1 are genes associated with cell cycle and chromatin regulation. **(B)** Two examples of DND1 targets, *Ccne1* and *Kat7*. Binding sites for DND1 are present in the 3'UTR of *Ccne1* and *Kat7* in both DO-RIP-seq experiments and in the Yamaji dataset (Yamaji et al., 2017). **(C)** Examples of cell cycle gene targets that are up-regulated in mutants at E14.5 (*Rheb*, *Rhob*, *Ccne1*, and *Trp53*). **(D)** Examples of chromatin regulators that are down-regulated in mutants relative to wild type by E14.5 (*Kat7*, *Rbbp5*, *Kdm5a*, and *Setdb1*).

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

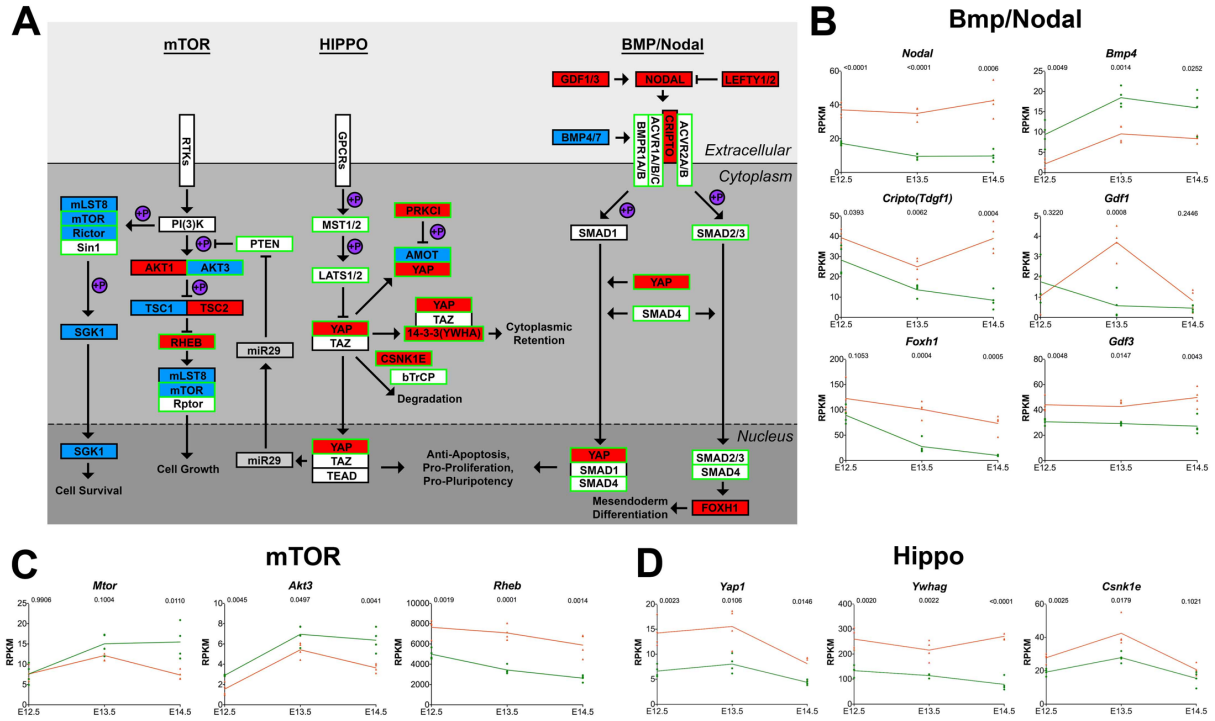


Figure 5. In the presence of the *Ter* mutation in *Dnd1*, members of mTOR, Hippo and Bmp/Nodal pathways have altered expression with some transcripts indicated as direct targets of DND1. (A) Interactions for mTOR, Hippo, and BMP/Nodal signaling pathways based primarily on KEGG analysis (Kanehisa et al., 2017) with input from primary literature (Beyer et al., 2013; Fuerer et al., 2014; Shimobayashi and Hall, 2014; Slagle et al., 2011; Spiller et al., 2017; Tanaka et al., 2007; Yu and Guan, 2013). Genes in pathways showing altered expression are colored blue for depression and red for elevation (RNAseq). Genes that map as transcript targets of DND1 based on RIP/ RIPseq are boxed in green. **(B,C,D)** Graphs for representative genes in pathways. **(B)** *Nodal*, *Cripto*, *Foxh1*, and *Gdf3* show strong up-regulation while *Bmp4* is down-regulated in the mutant transcriptome data. *Gdf1* has a spike in up-regulation after E12.5. **(C)** Genes mapping as DND1 targets in the mTor pathway including *Mtor* and *Akt3* are down-regulated while the target *Rheb* is strongly up-regulated in mutants. **(D)** Genes in the Hippo pathway, including *Yap1*, *Ywhag*, and *Csnk1e* are targets of DND1 that are up-regulated in mutants.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Figure Legends

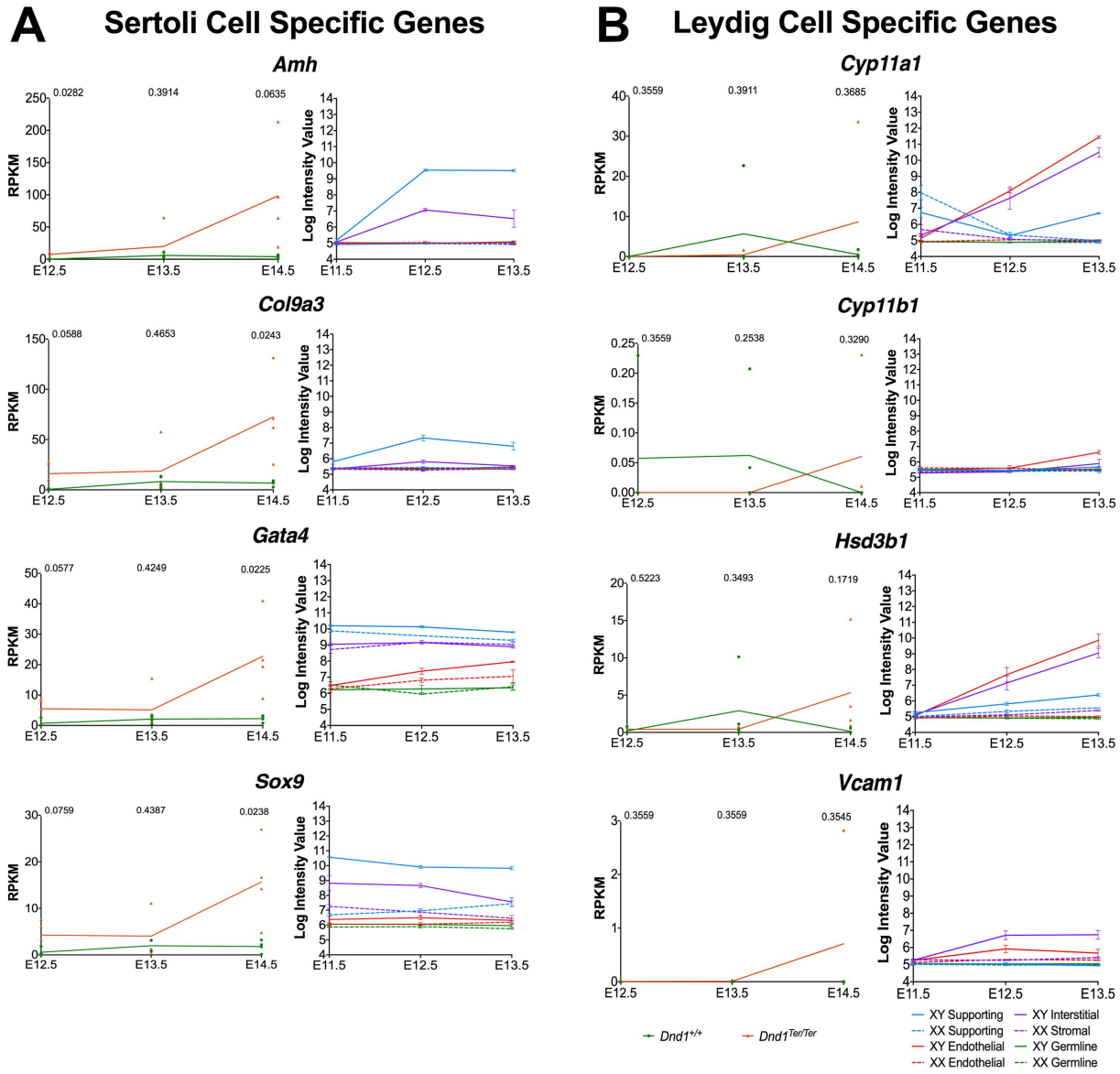
Suppl. Table 1. RNAseq gene expression graphing. We provide an excel file with **(A)** the RPKM values used in our analysis and **(B)** the graphic format used to display gene expression in wild type and *Dnd1*$^{Ter/Ter}$ mutants over time. The user can copy any from **(A)** with values into row 2 of **(B)** to generate the graph for that gene.

Suppl. Table 2. The list of changed genes defined by GSEA as part of the "DPPA4 Targets".

Suppl. Table 3. List of genes that are **(A)** up-regulated or **(B)** down-regulated in *Nanos2* mutants, annotated for activity in *Dnd1*$^{Ter/Ter}$ mutants.

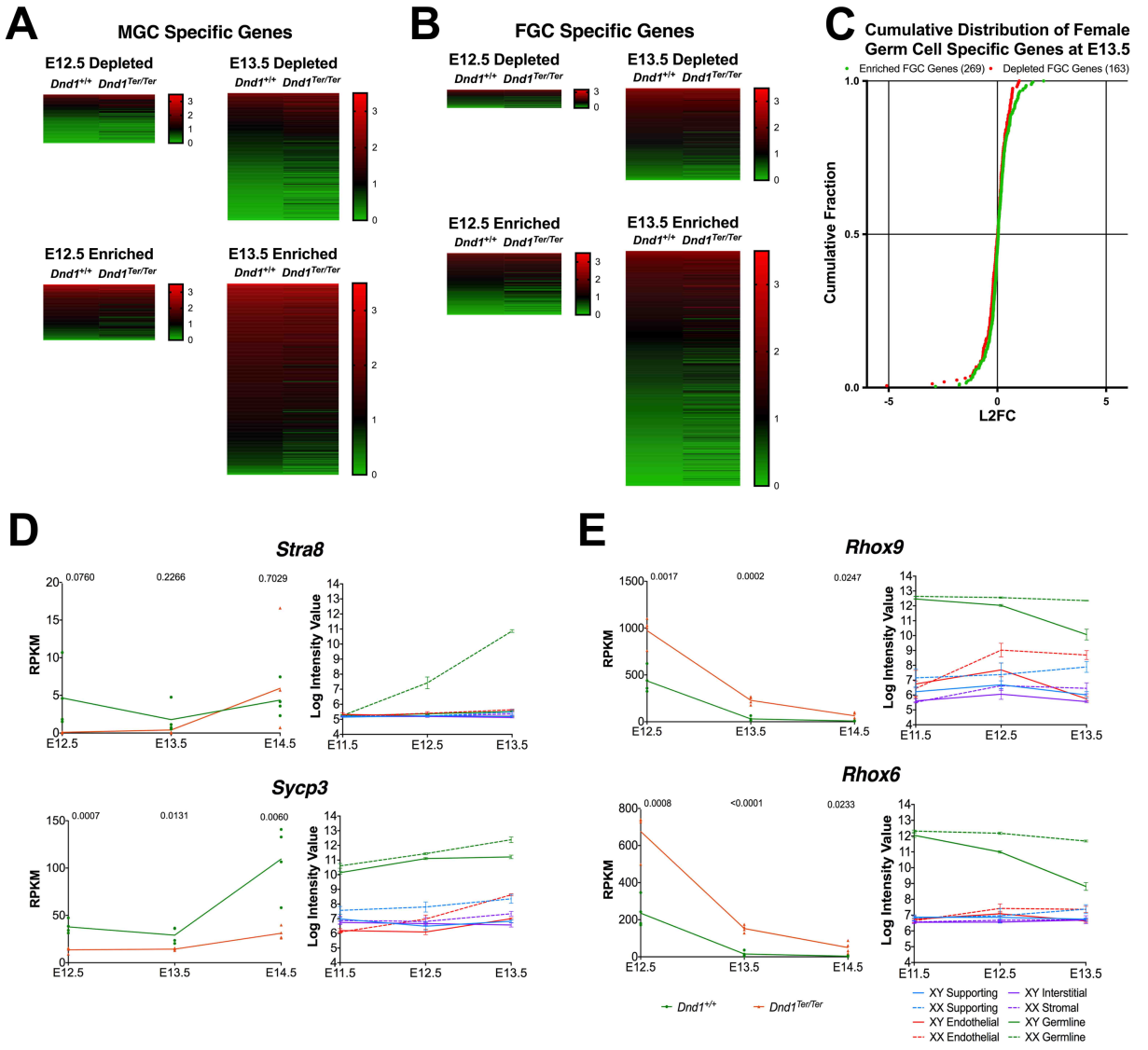
Suppl. Table 4. (A) List of all genes (expressed or not expressed in HEK293 cells), indicating genes identified in DO-RIP and/or PAR-CLIP (Yamaji et al., 2017). **(B)** List of all binding sites enriched in DO-RIP-seq. **(C)** List of genes with reproducible 3'UTR binding sites enriched in both DO-RIP-seq experiments. This table was used to label targets in Figure 5A.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Suppl. Fig. 1. Examples of genes detected in the *Dnd1*$^{Ter/Ter}$ transcriptome data that are specific to Sertoli or Leydig cell lineages in the gonad based on the Jameson data (Jameson et al., 2012). Typically, expression of these genes spikes in the E14.5 *Dnd1*$^{Ter/Ter}$ data, a stage when *Dnd1*$^{Ter/Ter}$ germ cell numbers are reduced and expression of Oct4-GFP declines, which may have led to errors in threshold settings during FACS. Graphs of transcriptome data from each cell type in the gonad performed at E11.5, E12.5 and E13.5 are shown for comparison. The germ cell lineage is green; supporting cells are blue; interstitial/stromal cells are purple; endothelial cells are red.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

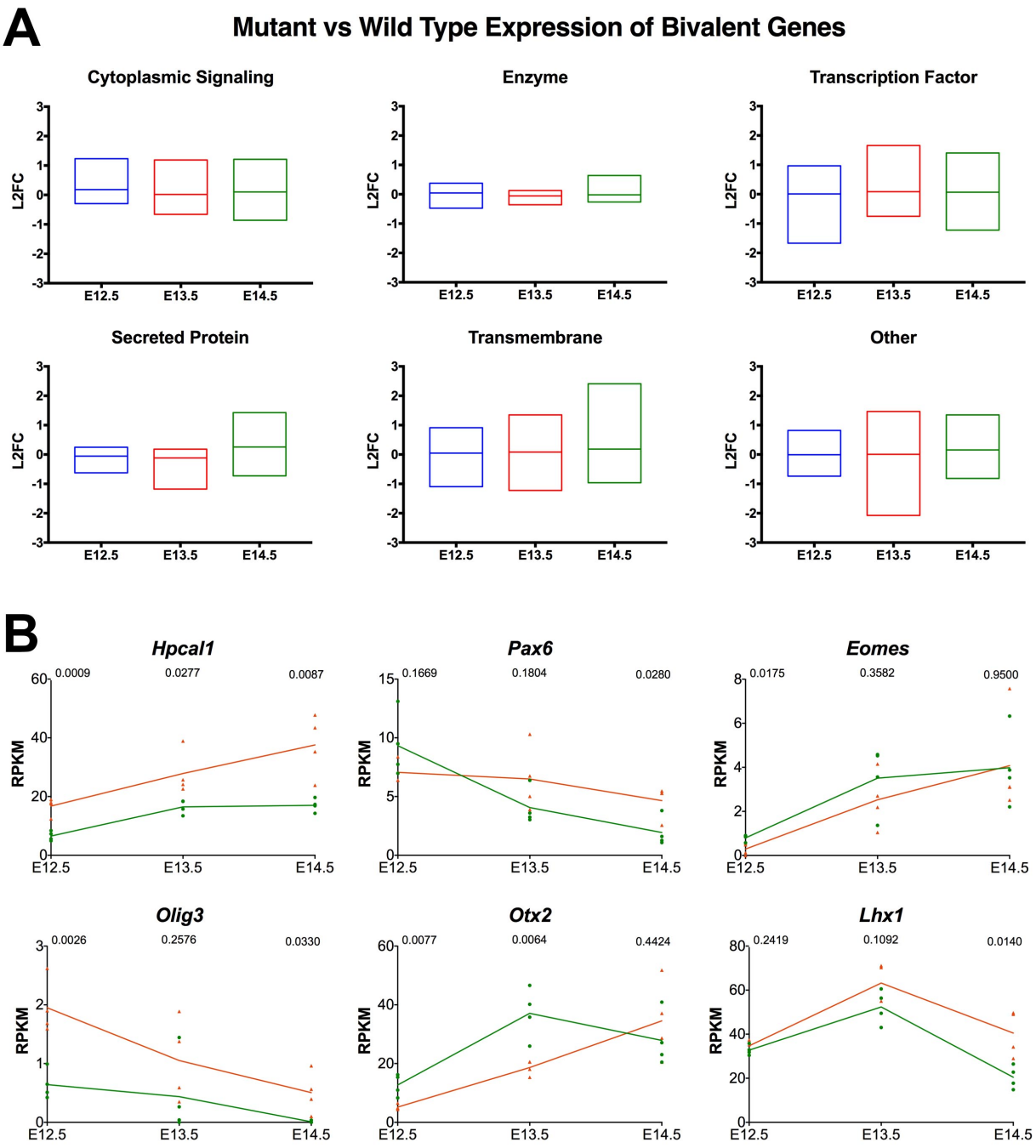


Suppl. Fig. 2. Male germ cells do not acquire a female (or meiotic) identity, but retain expression of pluripotency genes shared with female germ cells. **(A)** Heat map analysis showing that *Dnd1*$^{Ter/Ter}$ male germ cells retain a profile very similar to wild type at E12.5 and E13.5 with respect to genes that are specifically depleted in male germ cells, and genes that are specifically enriched. **(B)** Similarly, heat maps comparing expression of female-specific genes in XY *Dnd1*$^{+/+}$ to XY *Dnd1*$^{Ter/Ter}$ samples show few changes for depleted or enriched genes at E12.5 or E13.5. **(C)** A cumulative distribution plot of the log2 fold change between mutant male germ cells and wild type male germ cells at E13.5 for female germ cell enriched (green) or depleted (red) genes. Genes typically enriched (green) or depleted (red) in female germ cells show no clear bias

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for the direction of change in mutant germ cells. **(D)** Genes specifically upregulated as female germ cells enter meiosis (*Stra8* and *Sycp3*) were not activated by E14.5 in *Dnd1*$^{Ter/Ter}$ mutants. Graphs to the left show expression of the gene in all gonadal lineages (green broken line = XX germ cells; green solid line = XY germ cells; data from Jameson et al., 2012). **(E)** *Rhox9* and *Rhox6*, which normally become specific to XX germ cells by abrupt down-regulation in male germ cells, show a delay in down-regulation similar to other pluripotency genes.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

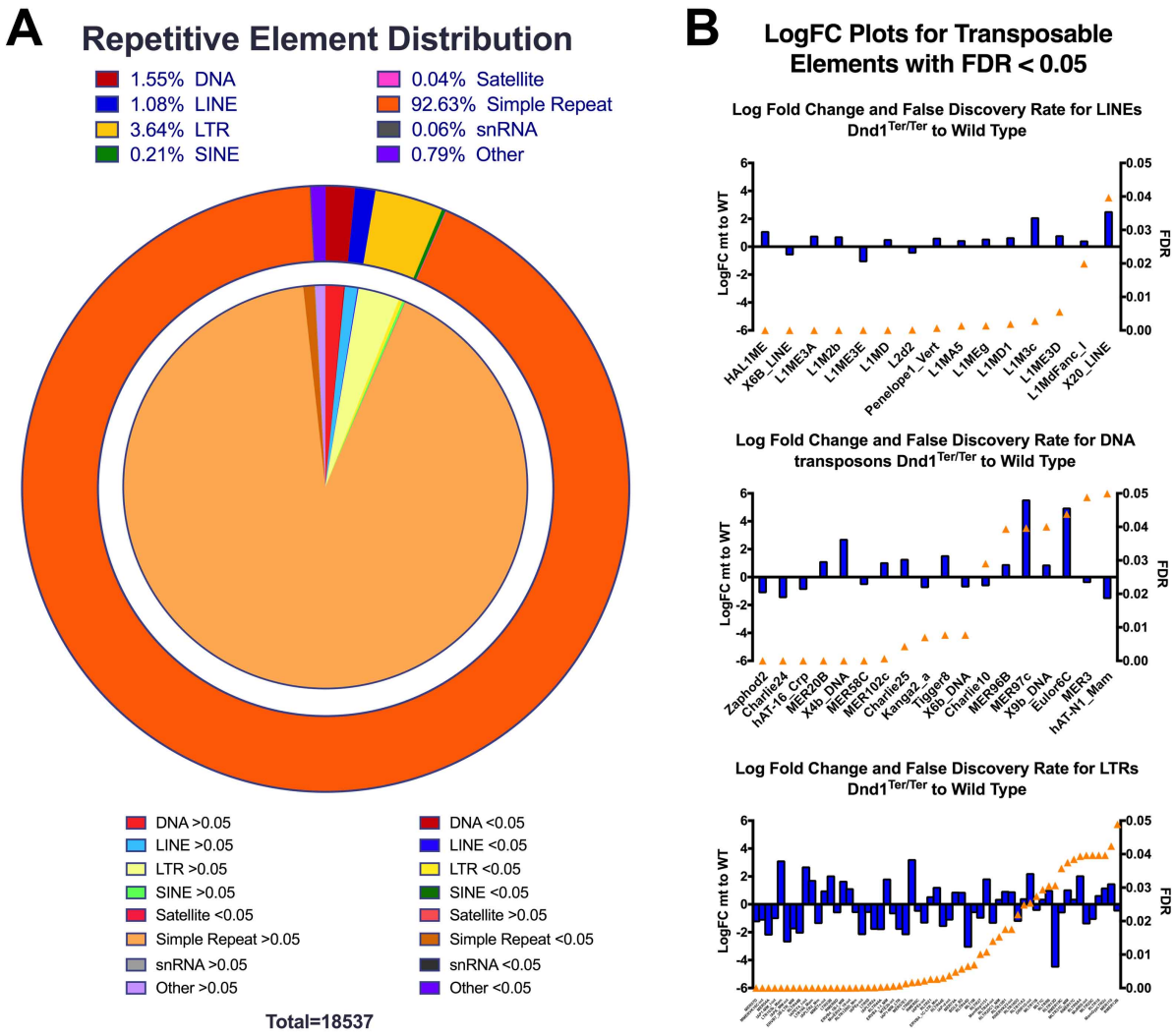


Suppl. Fig. 3. *Dnd1*$^{Ter/Ter}$ germ cells show no evidence of lineage infidelity between E12.5-E14.5. Despite the failure to activate many chromatin regulators in *Dnd1*$^{Ter/Ter}$ mutants, **(A)** expression of bivalent genes grouped by category (cytoplasmic signaling proteins, enzymes, transcription factors, secreted proteins, transmembrane proteins, and other) was not elevated, and **(B)** with the exception of *Hpcal1,* neither neuronal genes, nor other transcription factors associated with somatic cell

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

differentiation (*Pax6*, *Eomes*, *Olig3*, *Otx2*, and *Lhx1*) were significantly elevated in mutants.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

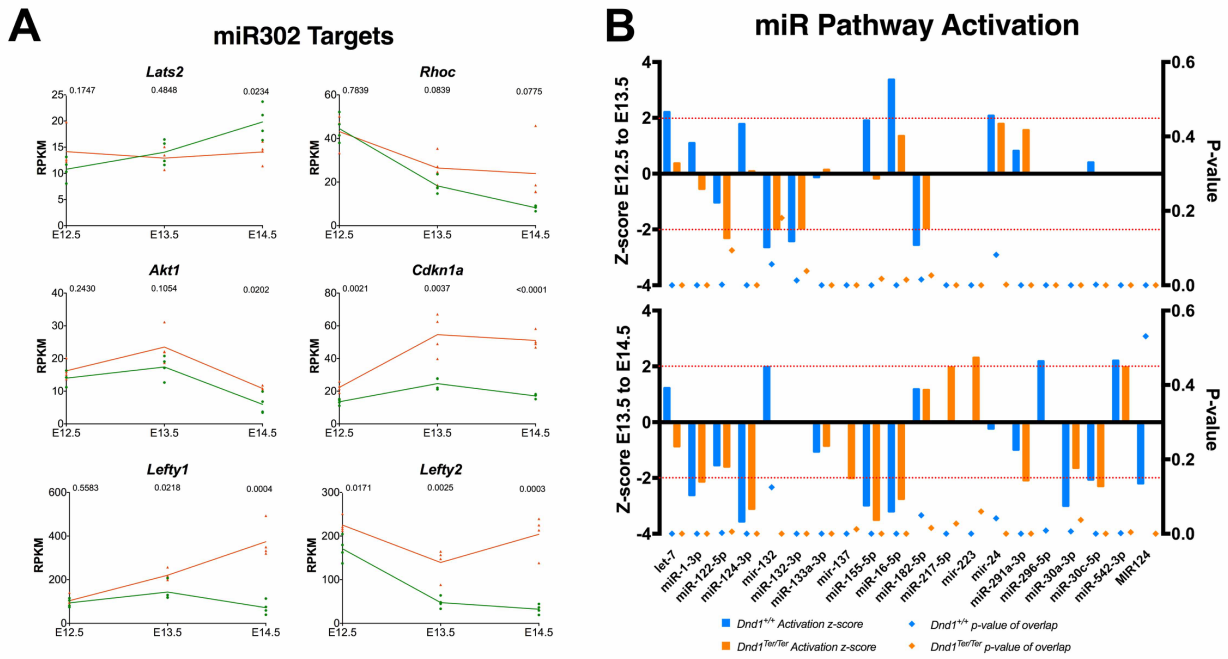


Suppl. Fig. 4. Repetitive elements in general do not exhibit aberrant activity. **(A)** A survey of repetitive element distribution organized by class. Outer ring: class distribution of elements represented in reads at E14.5 (Red=DNA transposons; Blue=LINEs; gold=LTRs; green=SINEs; Magenta=satellite; Orange=simple repeats; Dark Green=snRNAs; Purple=other). Inner circle divides specific classes into reads with significant differences between mutant and wildtype and reads without significant differences between mutant and wildtype (bold colors: FDR<0.05, significant; pale colors: FDR>0.05). **(B)** Waterfall plots of Log Fold Change (LogFC) between mutant and wildtype in each transposable element class (LINEs, LTRs, and DNA transposons) with

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

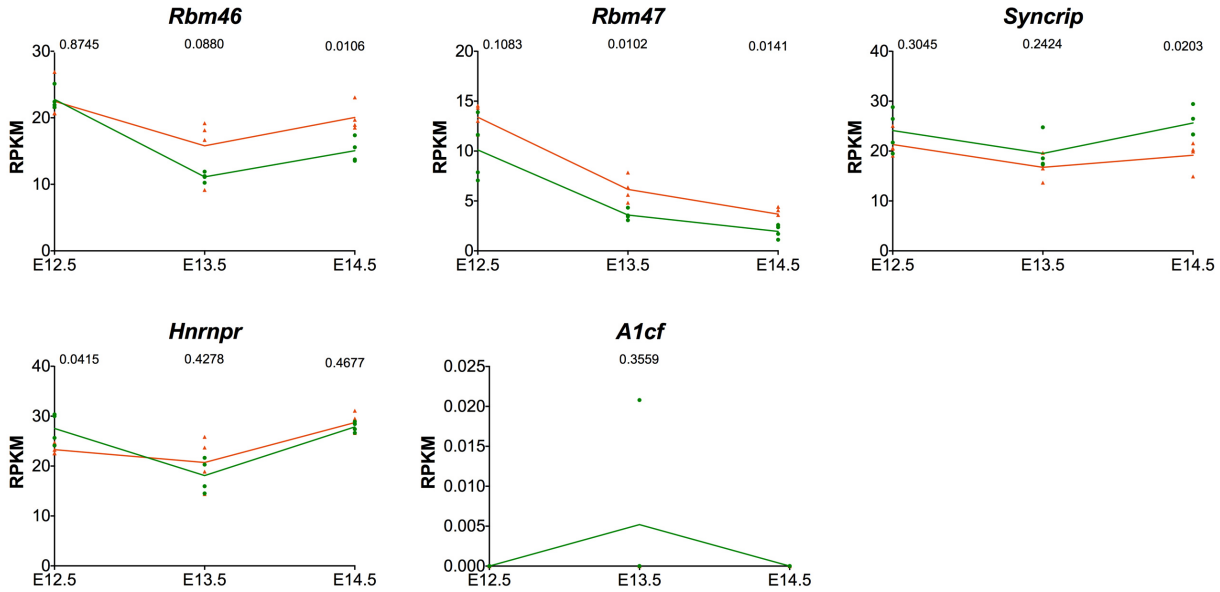
FDR<0.05 (significant). LogFC shown to left of plots; FDR shown by orange triangle placement relative to right scale on each graph.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



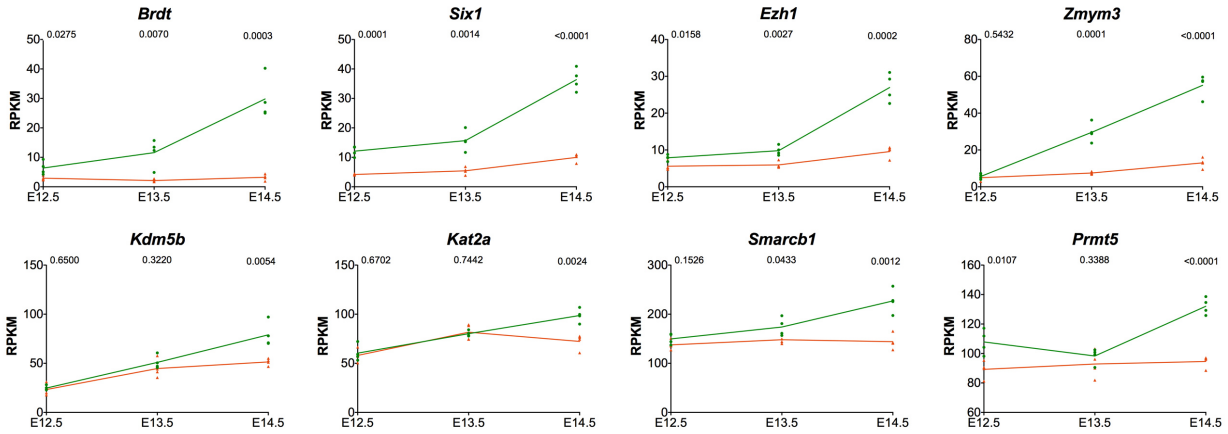
Suppl. Fig. 5. In general, miRNAs did not show activation based on downstream target analysis. (A) Although DND1 is believed to block degradation of zebrafish miR430 targets, with the exception of *Lats2*, targets of the orthologous miRNA in mice (miR320) are not down-regulated in mutants as predicted. **(B)** Waterfall graphs of Ingenuity Pathway Analysis (Qiagen) of miRNA activation based on expression changes in predicted targets at E12.5-E13.5 (top) and E13.5-E14.5 (bottom) for wild type (blue) and mutant (orange) samples. Only miRNA pathways with an activation Z-score of >2 or <-2 (left y-axis) in at least one transition period for mutant or wild type data are shown. P-values are plotted on the right Y-axis and represented as diamonds.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Suppl. Fig 6. Four RBPS closely related to DND1 are expressed in germ cells. Four of the five RNA binding proteins (*Rbm47*, *Rbm46*, *Syncrip*, *Hnrnpr*, but not *A1cf*) related to DND1 (Fig. 3) are expressed in male germ cells during the transcriptome timecourse.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

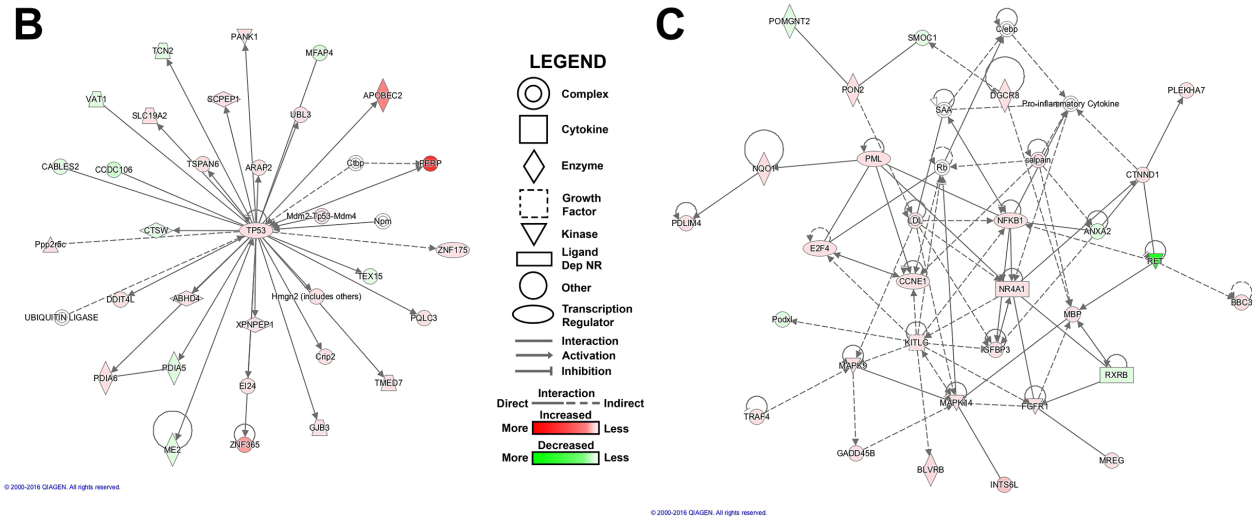


Suppl. Fig. 7. Many chromatin regulators that are not expressed in HEK293 cells, thus did not map as targets of DND1, fail to be activated in mutants. Nine additional chromatin regulators (*Brdt*, *Six1*, *Ezh1*, *Zmym3*, *Kdm5b*, *Kat2a*, *Smarcb1*, and *Prmt5*) are all expressed at lower levels in $Dnd1^{Ter/Ter}$ mutants likely reflecting a failure to activate the male pathway.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

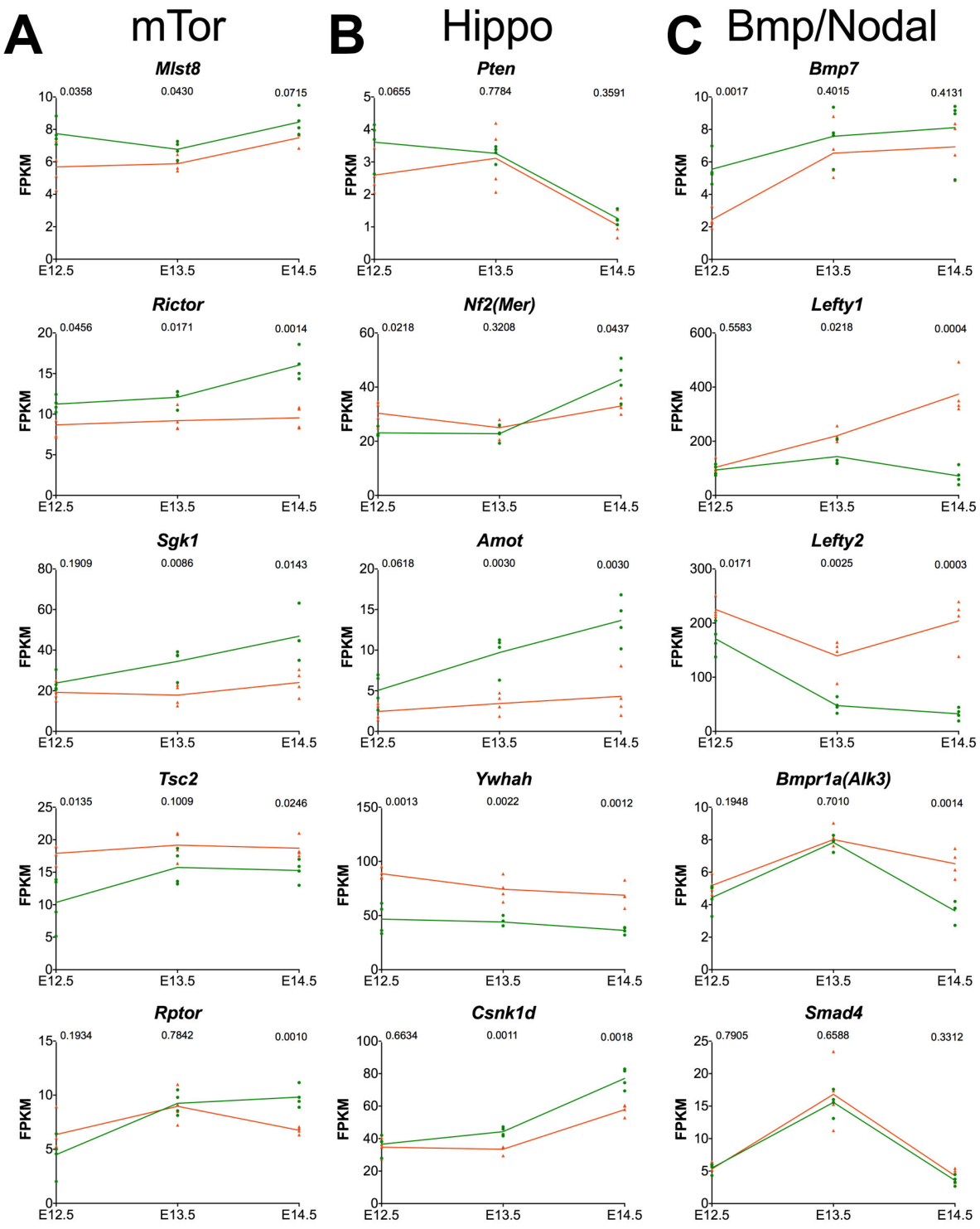
A

Pathway	P-value	# Affected Genes
Cell Death and Survival	$1.05E^{-04} – 3.21E^{-22}$	433
Cellular Growth and Proliferation	$1.06E^{-04} – 5.18E^{-20}$	482



Suppl. Fig. 8. Cell death and cell cycle pathways are strongly affected in *Dnd1*$^{Ter/Ter}$ mutants. (A) Ingenuity pathway analysis identified cell death (433 genes) and cell cycle (482 genes) pathways as the most strongly affected in *Dnd1*$^{Ter/Ter}$ mutants. **(B)** Network diagram centered on *Tp53*, depicting the relationship of many cell death and survival genes. **(C)** Network diagram centered on *Ccne1*, depicting the relationship of many cell cycle genes, most of which are up-regulated.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Suppl. Fig. 9. Expression of genes in the mTor, Hippo and Bmp/Nodal pathways are altered in *Dnd1*$^{Ter/Ter}$ mutants. Expression of other members of the **(A)** mTor, **(B)** Hippo and **(C)** Bmp/Nodal pathways from Fig. 5 is misregulated in mutants with some mapping as transcript targets of DND1 (Fig. 5A).

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Beyer, T.A., Weiss, A., Khomchuk, Y., Huang, K., Ogunjimi, A.A., Varelas, X., Wrana, J.L., 2013. Switch enhancers interpret TGF-beta and Hippo signaling to control cell fate in human embryonic stem cells. Cell reports 5, 1611-1624.

Bhattacharya, C., Aggarwal, S., Kumar, M., Ali, A., Matin, A., 2008. Mouse apolipoprotein B editing complex 3 (APOBEC3) is expressed in germ cells and interacts with dead-end (DND1). PloS one 3, e2315.

Bustamante-Marin, X., Garness, J.A., Capel, B., 2013. Testicular teratomas: an intersection of pluripotency, differentiation and cancer biology. The International journal of developmental biology 57, 201-210.

Ciosk, R., DePalma, M., Priess, J.R., 2006. Translational regulators maintain totipotency in the Caenorhabditis elegans germline. Science 311, 851-853.

Cook, M.S., Coveney, D., Batchvarov, I., Nadeau, J.H., Capel, B., 2009. BAX-mediated cell death affects early germ cell loss and incidence of testicular teratomas in Dnd1(Ter/Ter) mice. Developmental biology 328, 377-383.

Cook, M.S., Munger, S.C., Nadeau, J.H., Capel, B., 2011. Regulation of male germ cell cycle arrest and differentiation by DND1 is modulated by genetic background. Development 138, 23-32.

Criscione, S.W., Zhang, Y., Thompson, W., Sedivy, J.M., Neretti, N., 2014. Transcriptional landscape of repetitive elements in normal and cancer human cells. BMC genomics 15, 583.

Dawson, E.P., Lanza, D.G., Webster, N.J., Benton, S.M., Suetake, I., Heaney, J.D., 2018. Delayed male germ cell sex-specification permits transition into embryonal carcinoma cells with features of primed pluripotency. Development 145.

Fuerer, C., Nostro, M.C., Constam, D.B., 2014. Nodal.Gdf1 heterodimers with bound prodomains enable serum-independent nodal signaling and endoderm differentiation. The Journal of biological chemistry 289, 17854-17871.

Greco, C.M., Kunderfranco, P., Rubino, M., Larcher, V., Carullo, P., Anselmo, A., Kurz, K., Carell, T., Angius, A., Latronico, M.V., Papait, R., Condorelli, G., 2016. DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. Nature communications 7, 12418.

Gross-Thebing, T., Yigit, S., Pfeiffer, J., Reichman-Fried, M., Bandemer, J., Ruckert, C., Rathmer, C., Goudarzi, M., Stehling, M., Tarbashevich, K., Seggewiss, J., Raz, E., 2017. The Vertebrate Protein Dead End Maintains Primordial Germ Cell Fate by Inhibiting Somatic Differentiation. Developmental cell 43, 704-715 e705.

Gu, W., Mochizuki, K., Otsuka, K., Hamada, R., Takehara, A., Matsui, Y., 2018. Dnd1-mediated epigenetic control of teratoma formation in mouse. Biology open 7.

Heaney, J.D., Anderson, E.L., Michelson, M.V., Zechel, J.L., Conrad, P.A., Page, D.C., Nadeau, J.H., 2012. Germ cell pluripotency, premature differentiation and susceptibility to testicular teratomas in mice. Development 139, 1577-1586.

Jameson, S.A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D.M., Mork, L., Munger, S.C., Capel, B., 2012. Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. PLoS genetics 8, e1002575.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., Morishima, K., 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic acids research 45, D353-D361.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kedde, M., Strasser, M.J., Boldajipour, B., Vrielink, J.A., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P.M., van Duijse, J., Orom, U.A., Lund, A.H., Perrakis, A., Raz, E., Agami, R., 2007. RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell 131, 1273-1286.

Kivioja, T., Vaharautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., Taipale, J., 2011. Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods 9, 72-74.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10, R25.

Lee, J.Y., Ji, Z., Tian, B., 2008. Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3'-end of genes. Nucleic acids research 36, 5581-5590.

Lesch, B.J., Page, D.C., 2014. Poised chromatin in the mammalian germ line. Development 141, 3619-3626.

Madan, B., Madan, V., Weber, O., Tropel, P., Blum, C., Kieffer, E., Viville, S., Fehling, H.J., 2009. The pluripotency-associated gene Dppa4 is dispensable for embryonic stem cell identity and germ cell development but essential for embryogenesis. Mol Cell Biol 29, 3186-3203.

Matsui, Y., Zsebo, K., Hogan, B.L., 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell 70, 841-847.

McLaren, A., 1995. Germ cells and germ cell sex. Philosophical Transactions of the Royal Society of London 350, 229-233.

Nicholson, C.O., Friedersdorf, M.B., Bisogno, L.S., Keene, J.D., 2017. DO-RIP-seq to quantify RNA binding sites transcriptome-wide. Methods 118-119, 16-23.

Noguchi, T., Noguchi, M., 1985. A recessive mutation (ter) causing germ cell deficiency and a high incidence of congenital testicular teratomas in 129/Sv-ter mice. J Natl Cancer Inst 75, 385-392.

Pesce, M., Scholer, H.R., 2000. Oct-4: control of totipotency and germline determination. Molecular reproduction and development 55, 452-457.

Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., Sandberg, R., 2014. Full-length RNA-seq from single cells using Smart-seq2. Nature Protocols 9, 171-181.

Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L.H., Dale, R.K., Smith, S.A., Yarosh, C.A., Kelly, S.M., Nabet, B., Mecenas, D., Li, W., Laishram, R.S., Qiao, M., Lipshitz, H.D., Piano, F., Corbett, A.H., Carstens, R.P., Frey, B.J., Anderson, R.A., Lynch, K.W., Penalva, L.O., Lei, E.P., Fraser, A.G., Blencowe, B.J., Morris, Q.D., Hughes, T.R., 2013. A compendium of RNA-binding motifs for decoding gene regulation. Nature 499, 172-177.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Saba, R., Kato, Y., Saga, Y., 2014. NANOS2 promotes male germ cell development independent of meiosis suppression. Developmental biology 385, 32-40.

Shimobayashi, M., Hall, M.N., 2014. Making new contacts: the mTOR network in metabolism and signalling crosstalk. Nature reviews. Molecular cell biology 15, 155-162.

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Slagle, C.E., Aoki, T., Burdine, R.D., 2011. Nodal-dependent mesendoderm specification requires the combinatorial activities of FoxH1 and Eomesodermin. PLoS genetics 7, e1002072.

Spiller, C., Burnet, G., Bowles, J., 2017. Regulation of fetal male germ cell development by members of the TGFbeta superfamily. Stem Cell Res 24, 174-180.

Stevens, L.C., 1966. Development of resistance to teratocarcinogenesis by primordial germ cells in mice. J Natl Cancer Inst 37, 859-867.

Stevens, L.C., 1973. A new inbred subline of mice (129-terSv) with a high incidence of spontaneous congenital testicular teratomas. J Natl Cancer Inst 50, 235-242.

Stevens, L.C., Little, C.C., 1954. Spontaneous Testicular Teratomas in an Inbred Strain of Mice. Proceedings of the National Academy of Sciences of the United States of America 40, 1080-1087.

Suzuki, A., Niimi, Y., Saga, Y., 2014. Interaction of NANOS2 and NANOS3 with different components of the CNOT complex may contribute to the functional differences in mouse male germ cells. Biology open 3, 1207-1216.

Suzuki, A., Niimi, Y., Shinmyozu, K., Zhou, Z., Kiso, M., Saga, Y., 2016. Dead end1 is an essential partner of NANOS2 for selective binding of target RNAs in male germ cell development. EMBO reports 17, 37-46.

Suzuki, A., Saga, Y., 2008. Nanos2 suppresses meiosis and promotes male germ cell differentiation. Genes & development 22, 430-435.

Tanaka, C., Sakuma, R., Nakamura, T., Hamada, H., Saijoh, Y., 2007. Long-range action of Nodal requires interaction with GDF1. Genes & development 21, 3272-3282.

Updike, D.L., Knutson, A.K., Egelhofer, T.A., Campbell, A.C., Strome, S., 2014. Germ-granule components prevent somatic development in the C. elegans germline. Current biology : CB 24, 970-975.

Weidinger, G., Stebler, J., Slanchev, K., Dumstrei, K., Wise, C., Lovell-Badge, R., Thisse, C., Thisse, B., Raz, E., 2003. dead end, a novel vertebrate germ plasm component, is required for zebrafish primordial germ cell migration and survival. Current biology : CB 13, 1429-1434.

Western, P.S., Ralli, R.A., Wakeling, S.I., Lo, C., van den Bergen, J.A., Miles, D.C., Sinclair, A.H., 2011. Mitotic arrest in teratoma susceptible fetal male germ cells. PloS one 6, e20736.

Western, P.S., van den Bergen, J.A., Miles, D.C., Sinclair, A.H., 2010. Male fetal germ cell differentiation involves complex repression of the regulatory network controlling pluripotency. FASEB J 24, 3026-3035.

Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N., Tada, T., 2005. Nanog expression in mouse germ cell development. Gene Expr Patterns 5, 639-646.

Yamaji, M., Jishage, M., Meyer, C., Suryawanshi, H., Der, E., Yamaji, M., Garzia, A., Morozov, P., Manickavel, S., McFarland, H.L., Roeder, R.G., Hafner, M., Tuschl, T., 2017. DND1 maintains germline stem cells via recruitment of the CCR4-NOT complex to target mRNAs. Nature 543, 568-572.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., Scholer, H.R., 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development 122, 881-894.

Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y.I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Scholer, H.R., Matsui, Y., 1999. Germline-specific expression of the Oct-

bioRxiv preprint doi: https://doi.org/10.1101/402008; this version posted September 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4/green fluorescent protein (GFP) transgene in mice. Dev Growth Differ 41, 675-684.

Youngren, K.K., Coveney, D., Peng, X., Bhattacharya, C., Schmidt, L.S., Nickerson, M.L., Lamb, B.T., Deng, J.M., Behringer, R.R., Capel, B., Rubin, E.M., Nadeau, J.H., Matin, A., 2005. The Ter mutation in the dead end gene causes germ cell loss and testicular germ cell tumours. Nature 435, 360-364.

Yu, F.X., Guan, K.L., 2013. The Hippo pathway: regulators and regulations. Genes & development 27, 355-371.

Zechel, J.L., Doerner, S.K., Lager, A., Tesar, P.J., Heaney, J.D., Nadeau, J.H., 2013. Contrasting effects of Deadend1 (Dnd1) gain and loss of function mutations on allelic inheritance, testicular cancer, and intestinal polyposis. BMC Genet 14, 54.

Zhu, R., Iacovino, M., Mahen, E., Kyba, M., Matin, A., 2011. Transcripts that associate with the RNA binding protein, DEAD-END (DND1), in embryonic stem (ES) cells. BMC Mol Biol 12, 37.