

RESEARCH ARTICLE

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. Ruthig¹, Matthew B. Friedersdorf², Jason A. Garness¹, Steve C. Munger³, Corey Bunce¹, Jack D. Keene² and Blanche Capel^{1,*}

ABSTRACT

The adult spermatogonial stem cell population arises from pluripotent primordial germ cells (PGCs) that enter the fetal testis around embryonic day (E)10.5. PGCs undergo rapid mitotic proliferation, then enter prolonged 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 male germ cells (MGCs) fail to enter G1/G0 and instead form teratomas: tumors containing many embryonic cell types. To investigate the origin of these tumors, we sequenced the MGC transcriptome in *Dnd1^{Ter/Ter}* mutants at E12.5, E13.5 and E14.5, immediately prior to teratoma formation, and correlated this information with DO-RIP-Seq-identified DND1 direct targets. Consistent with previous results, we found DND1 controls downregulation of many genes associated with pluripotency and active cell cycle, including mTor, Hippo and Bmp/Nodal signaling pathway elements. However, DND1 targets also include genes associated with male differentiation, including a large group of chromatin regulators activated in wild-type but not mutant MGCs during the E13.5 and E14.5 transition. Results suggest multiple DND1 functions and link DND1 to initiation of epigenetic modifications in MGCs.

KEY WORDS: DO-RIP-Seq, *Dnd1*, Germ cells, RNAseq, Teratoma

INTRODUCTION

Primordial germ cells (PGCs) are specified at the base of the allantois in mouse embryos at embryonic day (E) 6.75. They proliferate, migrate through the gut mesentery and arrive in the gonad between E10.0 and E11.0, expressing stem cell markers that include 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 Schöler, 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, tumors 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-Marín et al., 2013; Dawson

et al., 2018). However, between E12.5 and E15.5, the efficiency of both EGC derivation and teratoma induction declines, reflecting epigenetic and gene expression changes that lead to suppression of the underlying pluripotent state of germ cells (Gu et al., 2018; Ng et al., 2013). 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, 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 crucial process is controlled 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 (Li et al., 2018; Zechel et al., 2013), and by a conditional deletion at E13.5 in mouse (Suzuki et al., 2016). Loss-of-function alleles of *Dnd1* led to lethality or complete germ cell loss. However, the increase in teratoma incidence is specific to the presence of the *Dnd1^{Ter}* mutation on the 129 genetic background. With the goal of understanding the role of DND1 in the transition from PGC to teratoma, we focused our efforts on characterizing the transcriptome in germ cells from *Dnd1^{Ter/Ter}* 129SvT2/SvEMSJ (129T2) mutant embryos immediately prior to the time when teratomas form (Cook et al., 2009, 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* (cyclin-dependent kinase inhibitor 1B, a negative regulator of the cell cycle) and *Lats2* (large tumor suppressor 2, 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 *Cdkn1b* and *Ezh2* (enhancer of zeste homolog 2), a mediator of H3K27me3 repression (Cook et al., 2011; Gu et al., 2018). Recent work in *Xenopus* showed that DND1 promotes translation of *nanos1* (C2HC type zinc finger 1) by relieving the

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

*Author for correspondence (blanche.capel@duke.edu)

© V.A.R., 0000-0003-1126-6106; M.B.F., 0000-0001-9999-8517; B.C., 0000-0002-6587-0969

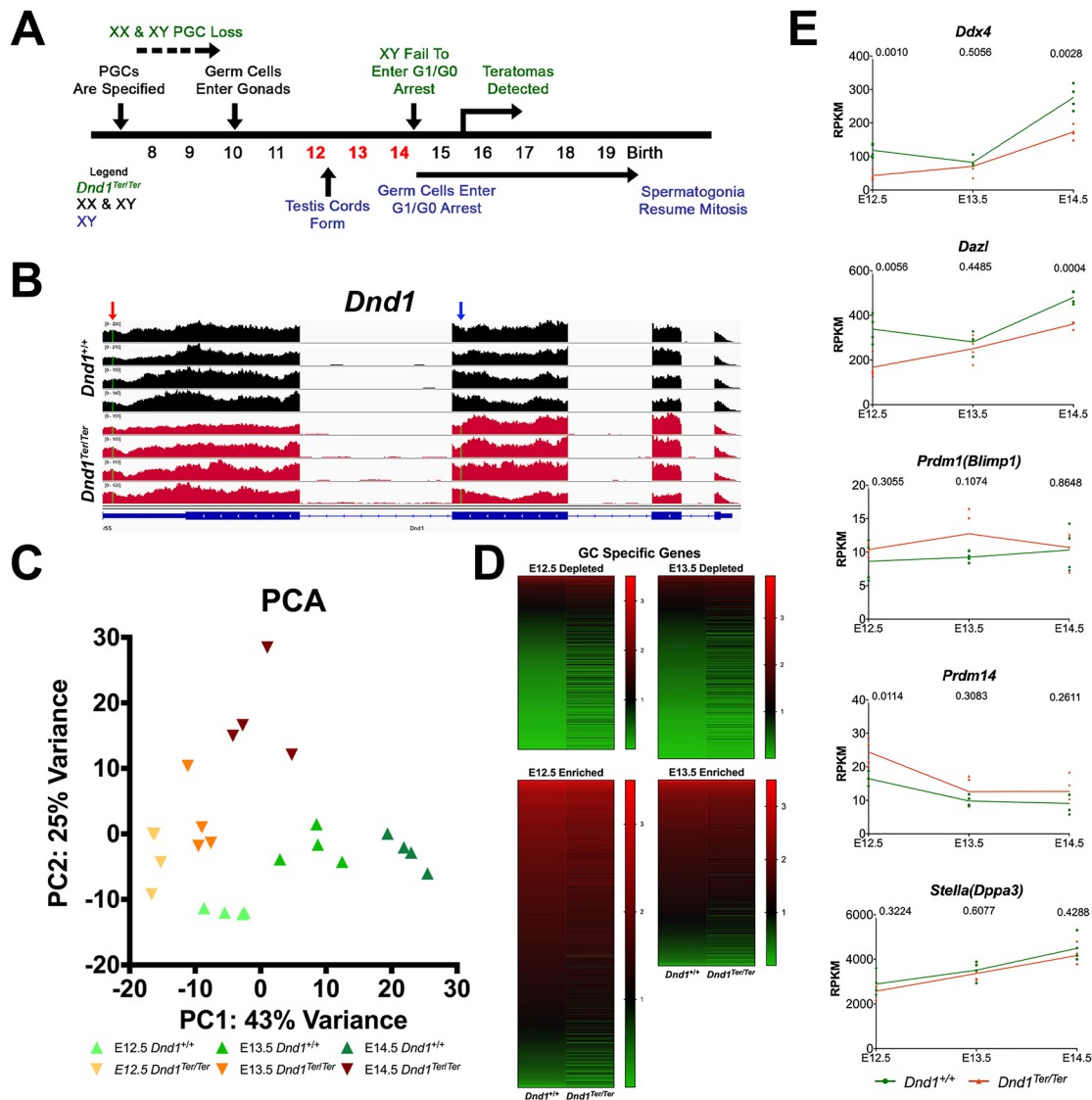


Fig. 1. Male germ cells in *Dnd1^{Ter/Ter}* mutants retain a germ cell identity. (A) Time line of male germ cell development and defects in *Dnd1^{Ter/Ter}* mutants. Primordial germ cells (PGCs) are specified at E6.75, and migrate and arrive in the gonad at E10.5. In the male gonad, they are enclosed 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 ~E15.5 (shown 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 upregulated (bottom) at E12.5 and E13.5. (E) Specific examples of gene expression across the time course (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). Although both *Ddx4* (*Vasa*) and *Dazl* show a delay in activation, similar levels at E13.5 and a slight downregulation at E14.5, *Prdm1* (*Blimp1*), *Prdm14* and *Stella* (*Dppa3*) show no significant differences from wild type at any stage. Mean data are plotted with individual data points shown.

inhibitory function of the eukaryotic initiation factor 3f (eIF3f), a repressive component of the preinitiation complex (Aguero et al., 2017), and that the second RNA recognition motif (RRM2) exhibits ATPase activity required for this function (Aguero et al., 2018). In contrast, 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., (2016) 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

upregulation of proliferation, meiosis markers and caspase 3, and the downregulation of several NANOS2 targets associated with male differentiation, similar to *Nanos2* mutants (Suzuki et al., 2016). CRISPR-Cas9 analysis identified four missense mutations in the first RNA recognition motif (RRM1) that decrease the interactions of DND1 with NANOS, but not with the CNOT complex. All of these mutations led to a failure of germ cell development by E12.5 (Li et al., 2018). The NMR spectroscopy-based solution structure of DND1 bound to a target RNA revealed that RRM1 acts as the main binding platform with the cooperation of RRM2 to stabilize binding to targets, and a role for the double-

stranded RNA-binding domain in the recruitment of factors that mediate target repression (Duszczyc et al., 2019 preprint).

Most studies have relied on a candidate approach to identify targets of DND1, which included pluripotency factors and cell cycle genes (Cook et al., 2011; Gu et al., 2018; Kedde et al., 2007; Zhu et al., 2011). More recently, Yamaji and co-workers performed photoactivatable-ribonucleoside-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 germline 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 TGF β (transforming growth factor β), Wnt (wingless-type MMTV integration site family) and PI3K-AKT (serine-threonine protein kinase). Extrapolating results from these *in vitro* cell types, Yamaji et al. suggested that activation of pluripotency genes, as well as TGF β 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, immediately 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 tagged DND1 proteins 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 with the molecular models proposed by Suzuki and Yamaji, between E12.5 and E13.5 DND1 controls the downregulation of many genes associated with pluripotency and the 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 biphasic functions of DND1, and link DND1 to the initiation of epigenetic modifications in male germ cells.

RESULTS

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 four 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 before teratoma detection (Fig. 1A; Cook et al., 2011; Dawson et al., 2018; Gu et al., 2018). 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 (Table S1).

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 (Fig. S7A). The *Ter* mutation, which is 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 *Dnd1^{Ter/Ter}* mutants, many germ cells are lost before arrival in the gonad (Cook et al., 2009; Noguchi and Noguchi, 1985). Given that loss of RBPs (including DND) 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 mouse gonad, and did not retain their germ cell identity. The top two principal 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 principal component (PC) 1, it appears that E13.5 mutant germ cells cluster more closely with E12.5 wild-type germ cells whereas E12.5 mutant germ cells are delayed. Furthermore, PC2 appears to stem mainly from expression variation in E14.5 *Dnd1^{Ter/Ter}* germ cells, suggesting that the transcription profile in these mutant 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 (Fig. S2A) 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 (e.g. 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 downregulated 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 downregulation at E14.5, whereas *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). We performed GO analysis and ingenuity pathway analysis (IPA) on this data, but neither approach provided further insight into the mutant phenotype beyond that presented in other contexts within this manuscript. Overall (Table S2), 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 (Fig. S2B). Based on a cumulative distribution plot, few female-specific genes were significantly altered in mutants at E13.5 (Fig. S2C). In contrast to their pattern in wild-type female germ cells, genes that are associated with entry into meiosis [e.g. *Stra8* (stimulated by retinoic acid 8) and *Sycp3* (synaptonemal complex protein 3)] were not upregulated in *Dnd1^{Ter/Ter}* mutant XY germ

cells (Fig. S2D). However, some genes that become female specific by downregulation in male germ cells [e.g. *Rhox9* and *Rhox6* (reproductive homeobox 9 and 6)] showed a less robust downregulation at the stage when male and female germ cell pathways diverge (E12.5) (Fig. S2E). In general, these were genes associated with the pluripotent state of early germ cells.

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 somatic 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 (Fig. S3A) and no significant differences in specific factors associated with differentiation (Fig. S3B).

Homozygosity for the *Dnd1^{Ter}* mutation is associated with subtle changes in expression of many genes in male germ cells between E12.5 and 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. Although most expressed genes retain a similar temporal pattern in the mutant compared with wild type over this period, hundreds are up- and downregulated at E12.5-E13.5 and/or E13.5-E14.5 (Fig. 2A). More genes are upregulated in mutant PGCs relative to wild type during the first transition between E12.5 and E13.5 (151 upregulated versus 58 downregulated), whereas more genes are downregulated in mutant PGCs during the second transition between E13.5 and E14.5 (43 upregulated versus 266 downregulated). 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) 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 family 25), *Rnf17* (ring finger protein 17) and *Mep1b* (meprin A subunit beta)] (Table S3) are enriched in wild-type 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 downstream function in PGCs.

Given the crucial roles of *Nanos2* in establishing the male pathway in mouse germ cells (Suzuki and Saga, 2008), we asked whether most transcriptional changes in *Dnd1^{Ter/Ter}* mutants could be accounted for by failure to activate *Nanos2* (Fig. 2C). To investigate this possibility, we compared the transcriptome changes between the two mutants. Although some genes changed in a similar manner (e.g. genes associated with pluripotency), many genes exhibited opposite patterns (Fig. 2E). For example, *Ret* was not affected in *Nanos2* mutants, whereas *Sycp3* and *Stra8* were elevated (Saba et al., 2014). In *Nanos2* mutants, *Nanos3* was upregulated, whereas in *Dnd1* mutants, it was not (Fig. 2C; see Table S4A,B for full dataset). This indicates that DND1 is upstream of both *Nanos* genes, and is consistent with the finding that *Nanos3* can partially rescue *Nanos2* mutants (Saba et al., 2014).

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

Several genes involved in repression of transposable elements, including *Piwil4* (piwi like RNA-mediated gene silencing 4), *Tdrd5* (tudor domain containing 5), *Morc1* (MORC family CW type zinc finger 1) and *Mael* were either not activated or were expressed at much lower levels in mutants (Fig. 2F). Proteins encoded by these genes normally silence LINE and other transposable elements. Our transcriptome data was based on oligo-dT priming; however, some transposable elements contain polyadenylation sites (Lee et al., 2008). We interrogated the raw transcriptome data at E14.5 for activation of transposable/repetitive elements using RepEnrich, a program optimized to detect differential enrichment between two RNA-seq datasets (Criscione et al., 2014). Very few transposable element reads showed a significant difference between wild-type and mutant germ cells (FDR<0.05) (Fig. 2G). This trend of similarity between wild-type and mutant germ cells was also observed when we interrogated repetitive elements and transposable elements more generally (Fig. S4A,B).

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 downregulated 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 wild type (Fig. S5A). We used IPA 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 (Fig. S5B). However, investigation of the expression of predicted targets of mir223 revealed no consistent pattern (data not shown).

Identification of DND1 direct targets

We identified hundreds of genes that were up- and downregulated between E12.5 and E14.5 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 downregulated 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 ribonucleoprotein R). All members of this subfamily contain three RNA recognition motifs (RRMs), except DND1, which contains two RRM. Additionally, three members of the family, including DND1, contain a C-terminal double-stranded RNA-binding domain (DSR). (Fig. 3A).

The mutation in *Dnd1^{Ter}* is predicted to lead to a premature STOP codon that disrupts the second RNA-recognition domain (arrow) and eliminates the DSR-binding domain. Whether or not the mutant allele functions as a dominant negative remains unclear. To identify binding sites of DND1, we performed DO-RIP-seq (Nicholson

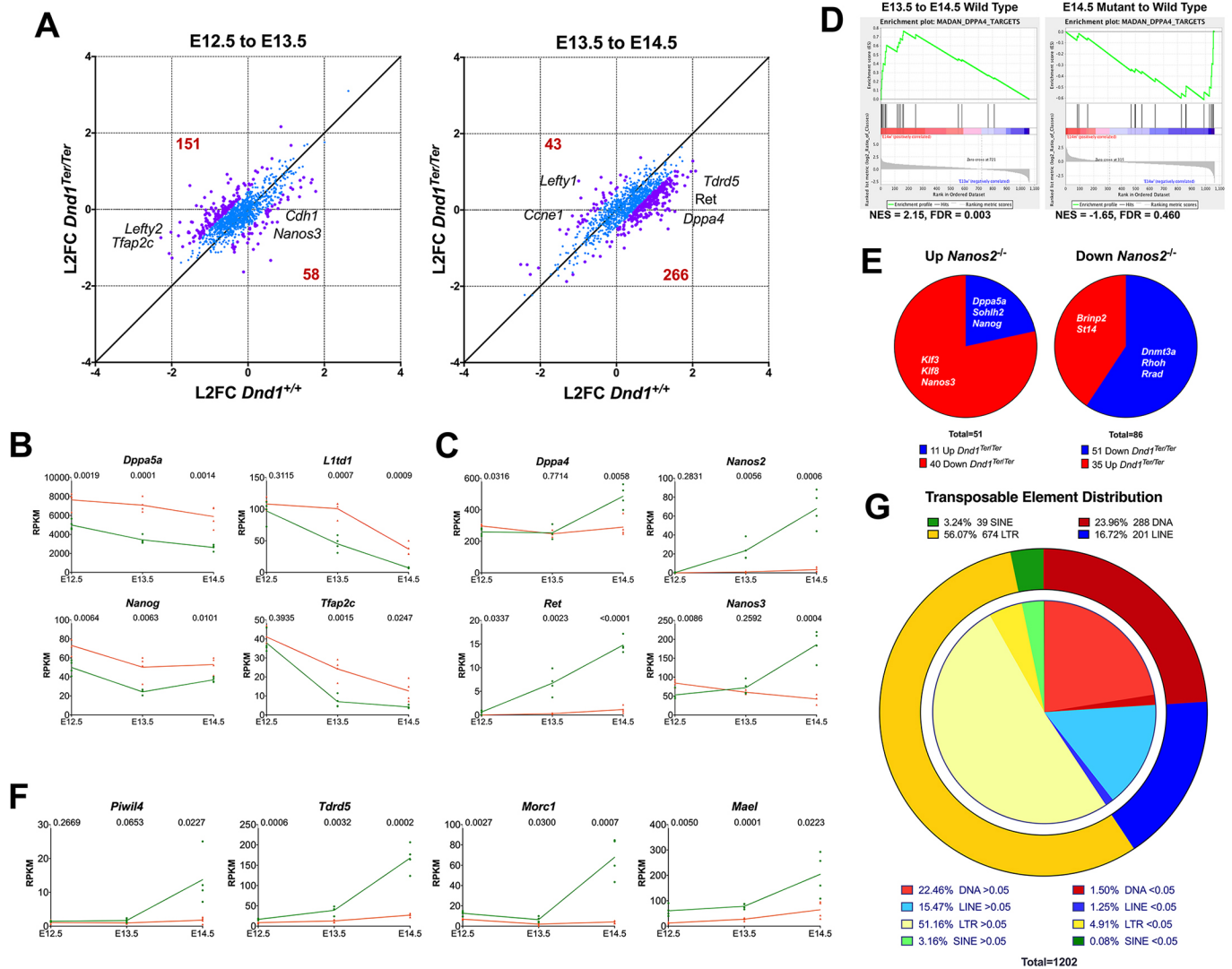


Fig. 2. *Dnd1^{TerTer}* mutant germ cells show delayed suppression of pluripotency, and a failure to activate the male differentiation program. (A) Scatter plots illustrating dynamic changes in expression of germ cell-specific genes (data from Jameson et al., 2012) between mutant (y-axis) and wild type (x-axis) during the transitions between E12.5 and E13.5, and E13.5 and E14.5. Between E12.5 and E13.5, more genes (151) show a higher log₂ fold change (L2FC) in mutants than in wild type, whereas only 58 genes lag behind. In the transition between E13.5 and 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 other signaling pathways. Points with L2FC between mutant and wild type: >0.5 in purple; <0.5 in blue. (B) Examples of four genes associated with pluripotency (*Dppa5a*, *L1td1*, *Nanog* and *Tfap2c*) that show a delay in downregulation in *Dnd1^{TerTer}* mutant germ cells. (C) Examples of four genes associated with initiation of the male pathway (*Dppa4*, *Nanos2*, *Ret* and *Nanos3*) that fail to be activated. Mean data are plotted with individual data points shown. (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^{TerTer}* mutants. Of genes upregulated in *Nanos2* mutants, 11 behave the same (blue) and 40 (red) are downregulated in *Dnd1^{TerTer}*. Of genes downregulated in *Nanos2* mutants, 51 are also downregulated in *Dnd1^{TerTer}* (blue) and 35 are upregulated (red) (Saba et al., 2014). Only genes with significant L2FC are shown ($P < 0.05$). See also Table S5. (F) Examples of genes associated with suppression of transposable elements and the piRNA pathway are activated in wild type but not mutants. (G) Although many silencers of repetitive elements were downregulated, 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 wild type (bold colors: $FDR < 0.05$; pale colors: $FDR > 0.05$).

et al., 2017) in HEK293 cells, using a tagged mouse DND1 and a tagged DND1-Ter allele. We introduced expression plasmids for both proteins and performed western blot analysis to show that both FLAG-DND1 and FLAG-DND1-Ter proteins were produced in the cells, although FLAG-DND1-Ter was detected at significantly lower levels (Fig. S7B). In accordance with this finding, after pull-down, the autoradiograph of the RNA bound to DND1-Ter was indistinguishable from the negative control, suggesting that little, if any, RNA immunoprecipitated with DND1-Ter (Fig. S7C,D). A

single DO-RIP-seq against DND1-Ter was marginally successful with RNA enrichment that was distinct from the negative control. Enriched RNAs in the DND1-Ter pull-down overlapped with 75% of enriched RNAs from DO-RIP-seq against DND1. As efforts to repeat this assay were not successful, there was not sufficient evidence to decisively support or refute the inability of DND1-Ter to efficiently bind transcripts affecting the regulation of these targets. Conclusive functional evidence on DND1-Ter will require further studies employing alternative methods.

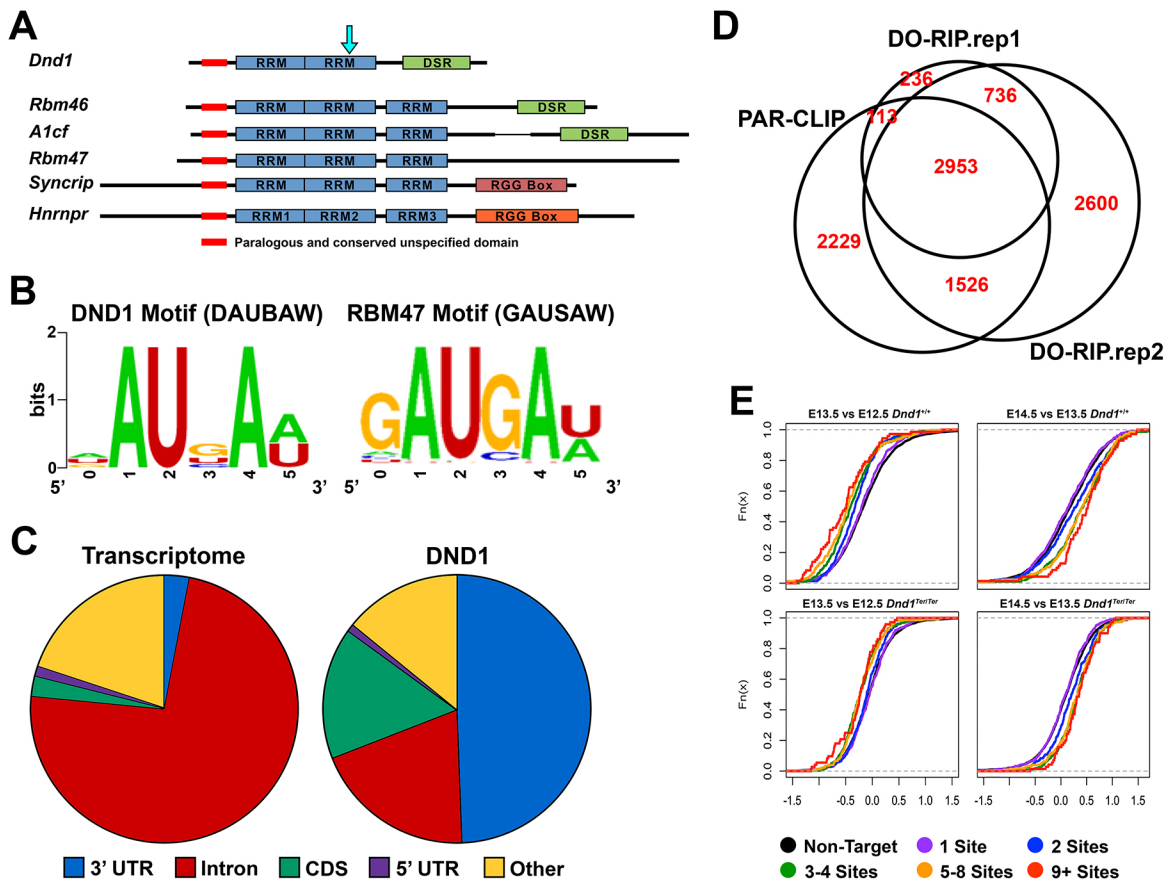


Fig. 3. DND1 binds to the DAUBAW motif preferentially localized to 3' UTRs, and regulates the up- and downregulation of targets. (A) DND1 and five other RBPs (RBM46, A1CF, RBM47, SYNCRIP and HNRNPR) are part of a subfamily containing RNA recognition motifs (RRMs) sharing sequence similarity. Three of these subfamily members (DND1 included) also have similar C-terminal double-stranded RNA-binding domains (DSRs). The blue arrow indicates the position of the *Ter* mutation, partially truncating the second RRM and completely removing the DSR. (B) k-mer analysis was performed to identify the most enriched motif in Dnd1 DO-RIP-Seq-binding sites (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 with 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 sequence (CDS). (D) Comparison of DND1-binding sites identified in the two replicates of the present study and binding sites identified by PAR-CLIP by Yamaji et al. (2017) (see Table S5). (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, one site, two 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 E13.5 and E14.5, 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.

In contrast, DO-RIP-seq for the wild-type DND1 allele was very successful. 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 *Hnmpnr* are all expressed in germ cells in the gonad, and could partially compensate for loss of DND1 (Fig. S6).

Compared with 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 Yamaji

et al., with 2953 mRNAs shared in common (Yamaji et al., 2017) (Fig. 3D, Table S5). 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. We suspect that the relatively minor differences between the studies are the result of differences in experimental approaches, and targets that are unique to one study are likely just subthreshold in the other.

Direct targets of DND1 are both up- and downregulated 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 and E13.5, and E13.5 and 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 downregulation relative to transcripts with few or no target sites during the transition between E12.5 and E13.5, whereas the opposite pattern was observed between E13.5 and E14.5 – target transcripts with more DND1-binding sites were more likely to be upregulated 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 S5). 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 *Trp53* (transformation related protein 53), 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 (Fig. S8).

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 Ci) 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-

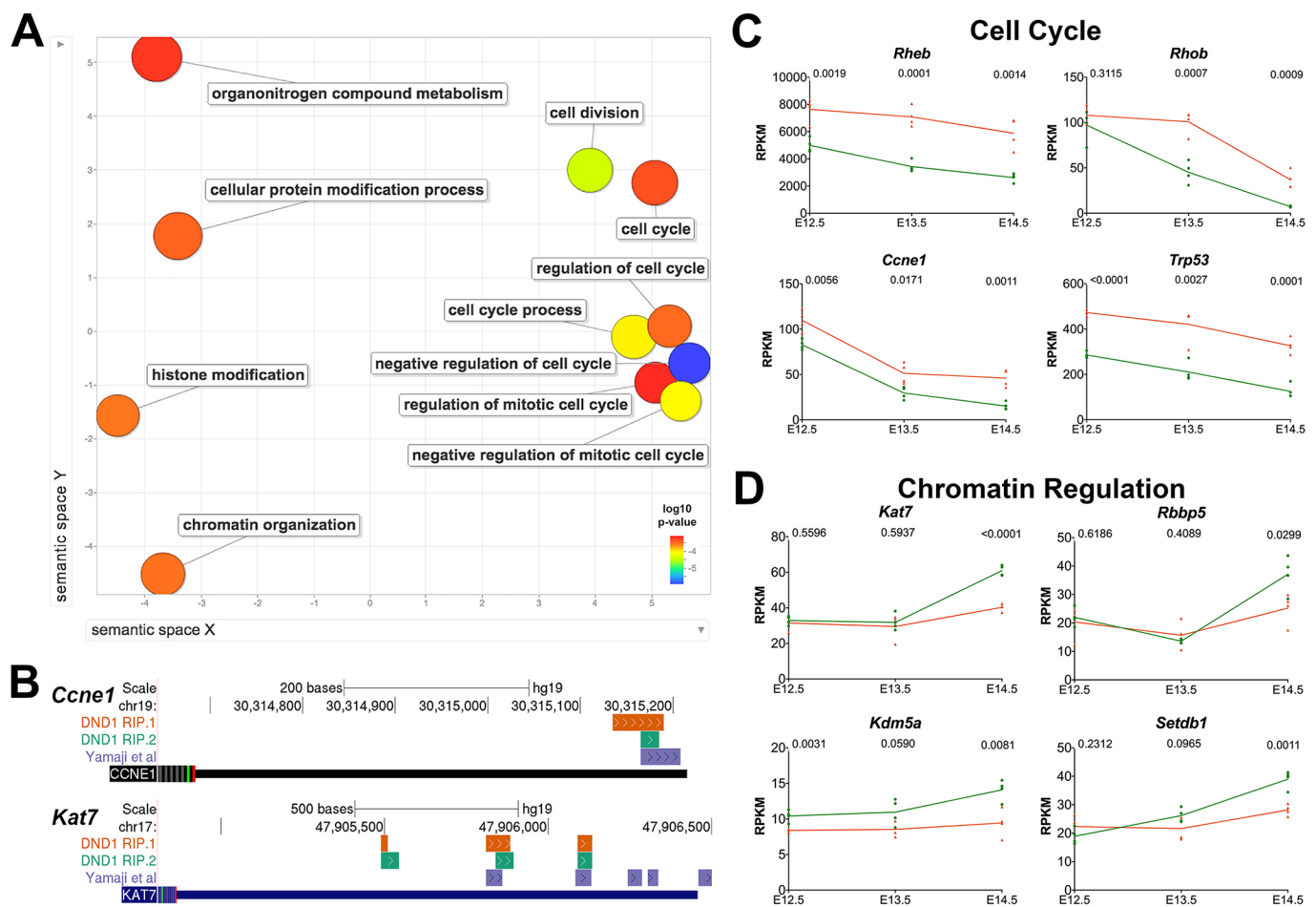


Fig. 4. DND1 targets include cell cycle genes and chromatin regulators, which are both up- and downregulated in mutant germ cells. (A) Semantic similarity clustering of GO terms enriched in DND1 DO-RIP-seq targets indicated that the most common categories of targets of DND1 are genes associated with 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 upregulated in mutants at E14.5 (*Rheb*, *Rhob*, *Ccne1* and *Trp53*). (D) Examples of chromatin regulators that are downregulated in mutants relative to wild type by E14.5 (*Kat7*, *Rbbp5*, *Kdm5a* and *Setdb1*).

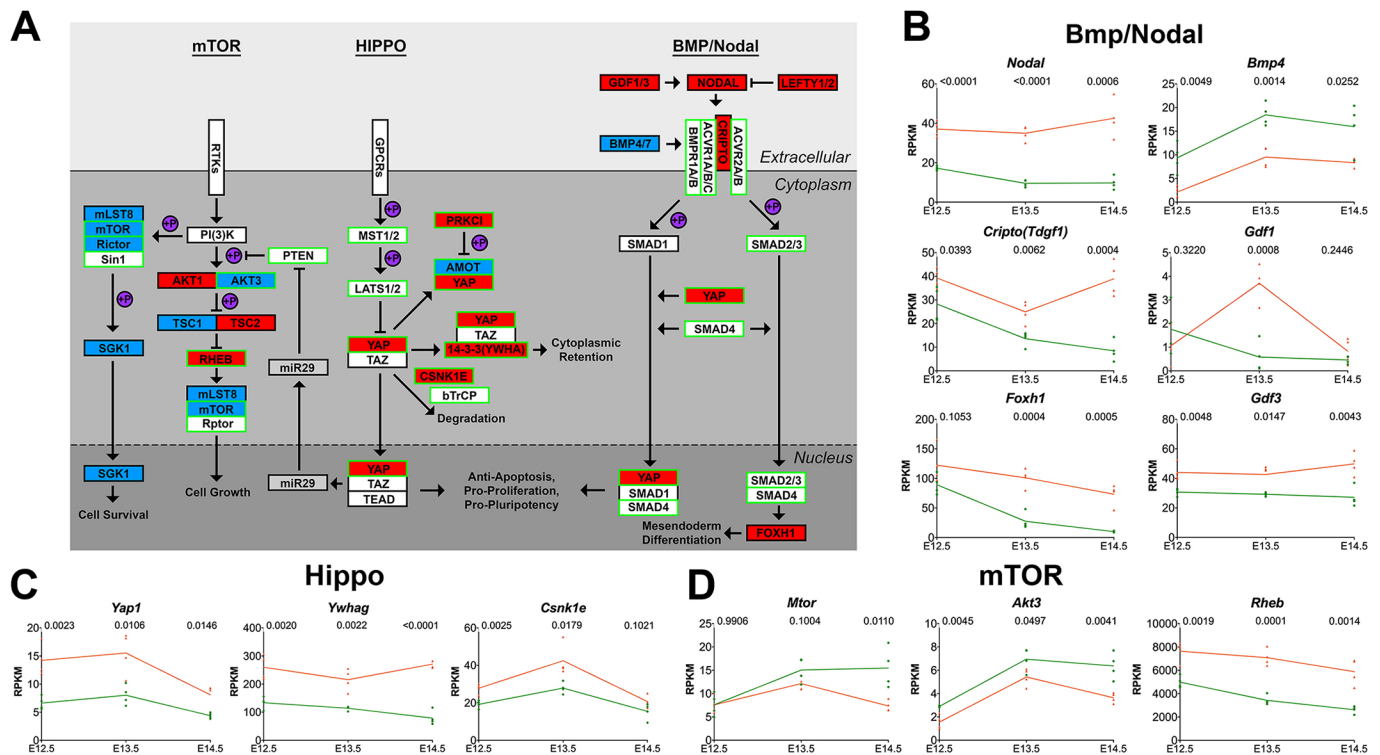


Fig. 5. In the presence of the *Ter* mutation in *Dnd1*, members of mTOR, Hippo and BMP/Nodal pathways have altered expression, with some transcripts indicated as direct targets of DND1. (A) Interactions for mTOR, Hippo and BMP/Nodal signaling pathways based primarily on KEGG analysis (Kanehisa et al., 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 (RNA-seq). Genes that map as transcript targets of DND1 based on RIP/RIPseq are boxed in green. (B-D) Graphs for representative genes in pathways. (B) *Nodal*, *Cripto*, *Foxh1* and *Gdf3* show strong upregulation, whereas *Bmp4* is downregulated in the mutant transcriptome data. *Gdf1* has a spike in upregulation after E12.5. (C) Genes in the Hippo pathway, including *Yap1*, *Ywhag* and *Csnk1e*, are targets of DND1 that are upregulated in mutants. (D) Genes mapping as DND1 targets in the mTOR pathway, including *Mtor* and *Akt3*, are downregulated, while the target *Rheb* is strongly upregulated in mutants.

insensitive companion of mTOR) and *Mtor* (mammalian target of rapamycin), were downregulated in mutants (Fig. 5D).

Ingenuity pathway analysis identified apoptotic and cell cycle pathways as strongly affected in mutants

We used IPA 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 (Fig. S9A-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 for investigating how stem cells embark on a highly specific differentiation pathway while restricting their potential for spontaneous differentiation or tumorigenesis. Although most germ cells appear to strictly control this potential and to 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 for investigating the genetic regulation of germ cell differentiation. What are the crucial 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 before overt transformation. We found that DND1 downregulates 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 crucial for initiating 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; as a result of its early role in germ cell development, many germ cells undergo apoptosis, or are otherwise lost, before arrival in the gonad (Cook et al., 2009; Noguchi and Noguchi, 1985). However, this transcriptome analysis reveals that, overall, the 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 in their differentiation trajectory.

DND1 likely performs several roles in post-transcriptional regulation. As previously shown (Ng et al., 2013), 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-spermatogonial 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* (*Bcl2l1*; B-cell lymphoma-extra 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 in this study, and all of these genes were strongly downregulated in *Dnd1^{Ter/Ter}* mutants, suggesting that DND1 acts as a positive regulator of the pro-spermatogonial program. How DND1 switches between negative and positive roles, and whether it performs both functions in individual cells, or changes roles at different stages of development, remain to be determined.

One possibility is that partners of DND1 regulate its negative or positive function. 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 downregulation 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 upregulation 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 and 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 pathways, as has often been proposed (Dawson et al., 2018; Gu et al., 2018; Heaney et al., 2012; Ng et al., 2013), 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 knockout 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). Although links between epigenetic regulation and RBPs have been well described in *C. elegans* (Strome and Updike, 2015), they are poorly understood in mammalian systems.

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. 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. 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.

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 genes in this category. We were also surprised by the absence of significant changes in miRNA regulatory signatures, given the evidence that DND1 protects transcripts from miRNA-mediated degradation (Kedde et al., 2007). One possibility is that

the protective activity of DND1 occurs at later stages of germ cell development. However, we are cautious about drawing strong conclusions about these data, as Z-scores used in the IPA may not be based on important targets specific to germ cells. 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 wild-type 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.

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

Finally, the role played by the *Dnd1^{Ter}* allele in the teratoma phenotype remains unclear. Recent work has implicated the RRM2 and DSR domains, both disrupted by the *Ter* mutation, as important mediators of DND1 activity (Aguero et al., 2018; Duszczuk et al., 2019preprint). We show that RNA from the *Dnd1^{Ter}* allele was produced at near normal levels when germ cells first arrive in the gonad (Fig. 1), but declined to very low levels at E13.5 and E14.5 (Fig. S10); thus, if the DND1-Ter allele plays a role in driving the teratoma phenotype that appears after E15.5, it is likely acting in the regulation of germ cells prior to or just after their arrival in the gonad. Although we found that the protein was translated in HEK293 cells, it was present at much lower levels and bound RNAs much less efficiently than the wild-type allele. However, based on a single successful DO-RIP-seq assay, ~75% of the RNAs bound by DND1 also came down with DND1-Ter. Although this limited finding is in no way conclusive, it seems likely that DND1-Ter binds targets less efficiently, but could interfere with the ability of DND1 to regulate the stabilization or degradation of targets. Further experiments would be required to validate this finding.

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’; 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 germ background) (Yoshimizu et al., 1999) was back-crossed to 129T2-*Dnd1*^{Ter/+} for nine 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 G-3'; GFP-R, 5'-TCC TTG AAG ATG GTG CG-3'; 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-mediated digestion of DNA from mice with the *Dnd1*^{Ter} mutation produces 123 bp and 22 bp products.

FACS, RNA isolation and 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 µl 0.25% Trypsin EDTA (Gibco, 25200) at 37°C for 5-10 min. 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 20 µm cell strainer (BD Falcon, 352235). Fluorescence-activated cell sorting (FACS) was performed by the Duke Comprehensive Cancer Center Flow Cytometry Shared Resource. Between 200 and 900 EGFP+ cells were sorted into 100 µl Qiagen Buffer RLT (from RNeasy Micro Kit, Qiagen, 74004)+DTT, without carrier RNA and frozen at -80°C.

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). 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 50 bp SR on the Illumina Hi-seq 2500. Eight samples (four mutant and four wild type) were pooled for each stage and run in technical replicate in two lanes of a flow cell. The Duke Sequencing Facility used FastQC for preliminary quality control.

Transcriptome processing and mapping reads

There were ~22 M reads/sample at E12.5, ~15 M reads/sample at 13.5, and ~38 M reads/sample at 14.5. Bowtie 2 was used to align the RNA-seq reads to mm9 using default settings. Over 81% of E12.5 transcripts aligned to the genome, whereas over 87% aligned at E13.5 and over 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. These 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) (Fig. S1). 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 colocalization 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 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 datasets. The list from E13.5 in the Jameson array was applied to the E13.5 and E14.5 transcriptome datasets. 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 datasets, 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 previously (Greco et al., 2016) and detailed by S. W. Criscione (github.com/nskvr/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 (software.broadinstitute.org/gsea/) as described previously (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 score (NES) and FDR values were used to describe the magnitude of change and the significant enrichment of the gene sets, respectively.

Ingenity pathway analysis

Differentially expressed genes (DEGs) were assessed for statistical enrichment/depletion of specific pathways using the Qiagen 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 previously (Nicholson et al., 2017). Briefly, 5-15 cm² plates of HEK293 cells were transfected with ~35 µg/plate FLAG-DND1 or untagged DND1 (as a negative control), grown for 24 h, then harvested in PLB. Lysates were treated with 50 U/µg of MNase for 5 min at 30°C and FLAG immunoprecipitation was 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., 2012) (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 immunoprecipitation 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 immunoprecipitation samples. All experiments were performed as biological duplicates.

Libraries were sequenced using Illumina HiSeq 2500. Processing of data was performed as described by 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. (2017). IPs were compared with input samples to identify sites enriched in DND1 immunoprecipitates. Reads with a low pass filter of >3 reads and >1.5-fold enrichment over input were considered DND1-binding sites.

Acknowledgements

We thank the teams in the FACS and Genomics Cores at Duke, especially Mike Cook, Nicolas Devos and Olivier Fedrigo, who helped with these experiments and their analysis. We are also grateful to Alex Bortvin, who provided information for the RepEnrich analysis; to Cindo Nicholson who provided technical advice for DO-RIP-seq experiments; to Joe Nadeau, who supplied the mice that founded our *Dnd1^{Ter}* colony; to past and present members of the Capel lab, especially Jordan Batchvarov, who contributed to the management of the strain over the last 10 years; and to Mike Czerwinski, who transferred bioinformatics expertise to J.A.G.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: B.C.; Methodology: J.A.G., J.D.K.; Software: S.C.M., C.B., J.D.K.; Validation: J.A.G.; Formal analysis: V.A.R., M.B.F., J.A.G., S.C.M., C.B.; Investigation: M.B.F., J.A.G.; Data curation: V.A.R., M.B.F., S.C.M., C.B.; Writing - original draft: V.A.R., M.B.F., B.C.; Writing - review & editing: S.C.M., C.B., J.D.K.; Visualization: V.A.R.; Supervision: J.D.K., B.C.; Project administration: J.D.K., B.C.; Funding acquisition: J.D.K., B.C.

Funding

This work was supported by the National Institutes of Health (GM087500 to B.C.); by Bridge Funding from Duke University Medical Center (to B.C.); and by the National Cancer Institute (CA157268 to J.D.K.). Deposited in PMC for release after 12 months.

Data availability

The raw RNA-Seq and DO-RIP-seq datasets have been deposited in GEO as a single unified SuperSeries with accession number GSE132719.

Supplementary information

Supplementary information available online at <http://dev.biologists.org/lookup/doi/10.1242/dev.175950.supplemental>

References

- Aguero, T., Jin, Z., Chorghade, S., Kalsotra, A., King, M. L. and Yang, J. (2017). Maternal Dead-end 1 promotes translation of nanos1 by binding the eIF3 complex. *Development* **144**, 3755-3765. doi:10.1242/dev.152611
- Aguero, T., Jin, Z., Owens, D., Malhotra, A., Newman, K., Yang, J. and King, M. L. (2018). Combined functions of two RRMs in Dead-end1 mimic helicase activity to promote nanos1 translation in the germline. *Mol. Reprod. Dev.* **85**, 896-908. doi:10.1002/mrd.23062
- Beyer, T. A., Weiss, A., Khomchuk, Y., Huang, K., Ogunjimi, A. A., Varelas, X. and 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. doi:10.1016/j.celrep.2013.11.021
- Bhattacharya, C., Aggarwal, S., Kumar, M., Ali, A. and 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. doi:10.1371/journal.pone.0002315
- Bustamante-Marín, X., Garness, J. A. and Capel, B. (2013). Testicular teratomas: an intersection of pluripotency, differentiation and cancer biology. *Int. J. Dev. Biol.* **57**, 201-210. doi:10.1387/ijdb.130136bc
- Ciosk, R., DePalma, M. and Priess, J. R. (2006). Translational regulators maintain totipotency in the *Caenorhabditis elegans* germline. *Science* **311**, 851-853. doi:10.1126/science.1122491
- Cook, M. S., Coveney, D., Batchvarov, I., Nadeau, J. H. and Capel, B. (2009). BAX-mediated cell death affects early germ cell loss and incidence of testicular teratomas in *Dnd1(Ter/Ter)* mice. *Dev. Biol.* **328**, 377-383. doi:10.1016/j.ydbio.2009.01.041
- Cook, M. S., Munger, S. C., Nadeau, J. H. and Capel, B. (2011). Regulation of male germ cell cycle arrest and differentiation by DND1 is modulated by genetic background. *Development* **138**, 23-32. doi:10.1242/dev.057000
- Criscione, S. W., Zhang, Y., Thompson, W., Sedivy, J. M. and Neretti, N. (2014). Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics* **15**, 583. doi:10.1186/1471-2164-15-583
- Dawson, E. P., Lanza, D. G., Webster, N. J., Benton, S. M., Suetake, I. and Heaney, J. D. (2018). Delayed male germ cell sex-specification permits transition into embryonal carcinoma cells with features of primed pluripotency. *Development* **145**, dev156612. doi:10.1242/dev.156612
- Duszczak, M. M., Wischniewski, H., Kazeeva, T., Loughlin, F. E., von Schroetter, C., Pradère, U., Hall, J., Ciaudo, C. and Allain, F. H.-T. (2019). The solution structure of Dead End bound to AU-rich RNA reveals an unprecedented mode of tandem RRM-RNA recognition required for mRNA repression. *bioRxiv*, 572156. doi:10.1101/572156
- Fuerer, C., Nostro, M. C. and Constam, D. B. (2014). Nodal Gdf1 heterodimers with bound prodomains enable serum-independent nodal signaling and endoderm differentiation. *J. Biol. Chem.* **289**, 17854-17871. doi:10.1074/jbc.M114.550301
- Greco, C. M., Kunderfranco, P., Rubino, M., Larcher, V., Carullo, P., Anselmo, A., Kurz, K., Carell, T., Angius, A., Latronico, M. V. G. et al. (2016). DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nat. Commun.* **7**, 12418. doi:10.1038/ncomms12418
- Gross-Thebing, T., Yigit, S., Pfeiffer, J., Reichman-Fried, M., Bandemer, J., Ruckert, C., Rathmer, C., Goudarzi, M., Stehling, M., Tarbashevich, K. et al. (2017). The vertebrate protein dead end maintains primordial germ cell fate by inhibiting somatic differentiation. *Dev. Cell* **43**, 704-715.e705. doi:10.1016/j.devcel.2017.11.019
- Gu, W., Mochizuki, K., Otsuka, K., Hamada, R., Takehara, A. and Matsui, Y. (2018). Dnd1-mediated epigenetic control of teratoma formation in mouse. *Biology open* **7**, bio032318. doi:10.1242/bio.032318
- Heaney, J. D., Anderson, E. L., Michelson, M. V., Zechel, J. L., Conrad, P. A., Page, D. C. and Nadeau, J. H. (2012). Germ cell pluripotency, premature differentiation and susceptibility to testicular teratomas in mice. *Development* **139**, 1577-1586. doi:10.1242/dev.076851
- Jameson, S. A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D. M., Mork, L., Munger, S. C. and Capel, B. (2012). Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet.* **8**, e1002575. doi:10.1371/journal.pgen.1002575
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353-D361. doi:10.1093/nar/gkw1092
- Kedde, M., Strasser, M. J., Boldajipour, B., Vrielink, J. A. F. O., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P. M., van Duijse, J., Ørom, U. A. et al. (2007). RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell* **131**, 1273-1286. doi:10.1016/j.cell.2007.11.034
- Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S. and Taipale, J. (2012). Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **9**, 72-74. doi:10.1038/nmeth.1778
- Langmead, B., Trapnell, C., Pop, M. and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25. doi:10.1186/gb-2009-10-3-r25
- Lee, J. Y., Ji, Z. and 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 Res.* **36**, 5581-5590. doi:10.1093/nar/gkn540
- Lesch, B. J. and Page, D. C. (2014). Poised chromatin in the mammalian germ line. *Development* **141**, 3619-3626. doi:10.1242/dev.113027
- Li, Q., Li, Y., Yang, S., Huang, S., Yan, M., Ding, Y., Tang, W., Lou, X., Yin, Q., Sun, Z. et al. (2018). CRISPR-Cas9-mediated base-editing screening in mice identifies DND1 amino acids that are critical for primordial germ cell development. *Nat. Cell Biol.* **20**, 1315-1325. doi:10.1038/s41556-018-0202-4
- Madan, B., Madan, V., Weber, O., Tropel, P., Blum, C., Kieffer, E., Viville, S. and 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. doi:10.1128/MCB.01970-08
- Matsui, Y., Zsebo, K. and Hogan, B. L. M. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841-847. doi:10.1016/0092-8674(92)90317-6
- McLaren, A. (1995). Germ cells and germ cell sex. *Philos. Trans. R. Soc. Lond.* **350**, 229-233. doi:10.1098/rstb.1995.0156
- Ng, J.-H., Kumar, V., Muratani, M., Kraus, P., Yeo, J.-C., Yaw, L.-P., Xue, K., Lufkin, T., Prabhakar, S. and Ng, H.-H. (2013). In vivo epigenomic profiling of germ cells reveals germ cell molecular signatures. *Dev. Cell* **24**, 324-333. doi:10.1016/j.devcel.2012.12.011
- Nicholson, C. O., Friedersdorf, M. B., Bisogno, L. S. and Keene, J. D. (2017). DO-RIP-seq to quantify RNA binding sites transcriptome-wide. *Methods* **118-119**, 16-23. doi:10.1016/j.ymeth.2016.11.004

- Noguchi, T. and 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. doi:10.1093/jnci/75.2.385
- Pesce, M. and Schöler, H. R.** (2000). Oct-4: control of totipotency and germline determination. *Mol. Reprod. Dev.* **55**, 452-457. doi:10.1002/(SICI)1098-2795(200004)55:4<452::AID-MRD14>3.0.CO;2-S
- Picelli, S., Faridani, O. R., Björklund, A. K., Winberg, G., Sagasser, S. and Sandberg, R.** (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171-181. doi:10.1038/nprot.2014.006
- Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A. et al.** (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172-177. doi:10.1038/nature12311
- Robinson, M. D., McCarthy, D. J. and Smyth, G. K.** (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140. doi:10.1093/bioinformatics/btp616
- Saba, R., Kato, Y. and Saga, Y.** (2014). NANOS2 promotes male germ cell development independent of meiosis suppression. *Dev. Biol.* **385**, 32-40. doi:10.1016/j.ydbio.2013.10.018
- Shimobayashi, M. and Hall, M. N.** (2014). Making new contacts: the mTOR network in metabolism and signalling crosstalk. *Nat. Rev. Mol. Cell Biol.* **15**, 155-162. doi:10.1038/nrm3757
- Slagle, C. E., Aoki, T. and Burdine, R. D.** (2011). Nodal-dependent mesendoderm specification requires the combinatorial activities of FoxH1 and Eomesodermin. *PLoS Genet.* **7**, e1002072. doi:10.1371/journal.pgen.1002072
- Spiller, C., Burnet, G. and Bowles, J.** (2017). Regulation of fetal male germ cell development by members of the TGFbeta superfamily. *Stem Cell Res.* **24**, 174-180. doi:10.1016/j.scr.2017.07.016
- 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. doi:10.1093/jnci/50.1.235
- Stevens, L. C. and Little, C. C.** (1954). Spontaneous Testicular Teratomas in an Inbred Strain of Mice. *Proc. Natl. Acad. Sci. USA* **40**, 1080-1087. doi:10.1073/pnas.40.11.1080
- Strome, S. and Updike, D.** (2015). Specifying and protecting germ cell fate. *Nat. Rev. Mol. Cell Biol.* **16**, 406-416. doi:10.1038/nrm4009
- Subramanian, A., Tamayo, R., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. and Mesirov, J. P.** (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545-15550. doi:10.1073/pnas.0506580102
- Suzuki, A. and Saga, Y.** (2008). Nanos2 suppresses meiosis and promotes male germ cell differentiation. *Genes Dev.* **22**, 430-435. doi:10.1101/gad.1612708
- Suzuki, A., Niimi, Y. and 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. *Biol. Open* **3**, 1207-1216. doi:10.1242/bio.20149308
- Suzuki, A., Niimi, Y., Shinmyozu, K., Zhou, Z., Kiso, M. and Saga, Y.** (2016). Dead end1 is an essential partner of NANOS2 for selective binding of target RNAs in male germ cell development. *EMBO Rep.* **17**, 37-46. doi:10.15252/embr.201540828
- Tanaka, C., Sakuma, R., Nakamura, T., Hamada, H. and Saijoh, Y.** (2007). Long-range action of Nodal requires interaction with GDF1. *Genes Dev.* **21**, 3272-3282. doi:10.1101/gad.1623907
- Updike, D. L., Knutson, A. K., Egelhofer, T. A., Campbell, A. C. and Strome, S.** (2014). Germ-granule components prevent somatic development in the *C. elegans* germline. *Curr. Biol.* **24**, 970-975. doi:10.1016/j.cub.2014.03.015
- Weidinger, G., Stebler, J., Slanchev, K., Dumstrei, K., Wise, C., Lovell-Badge, R., Thisse, C., Thisse, B. and Raz, E.** (2003). dead end, a novel vertebrate germ plasm component, is required for zebrafish primordial germ cell migration and survival. *Curr. Biol.* **13**, 1429-1434. doi:10.1016/S0960-9822(03)00537-2
- Western, P. S., van den Bergen, J. A., Miles, D. C. and Sinclair, A. H.** (2010). Male fetal germ cell differentiation involves complex repression of the regulatory network controlling pluripotency. *FASEB J.* **24**, 3026-3035. doi:10.1096/fj.09-151555
- Western, P. S., Ralli, R. A., Wakeling, S. I., Lo, C., van den Bergen, J. A., Miles, D. C. and Sinclair, A. H.** (2011). Mitotic arrest in teratoma susceptible fetal male germ cells. *PLoS ONE* **6**, e20736. doi:10.1371/journal.pone.0020736
- Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N. and Tada, T.** (2005). Nanog expression in mouse germ cell development. *Gene Expr. Patterns* **5**, 639-646. doi:10.1016/j.modgep.2005.03.001
- Yamaji, M., Jishage, M., Meyer, C., Suryawanshi, H., Der, E., Yamaji, M., Garzia, A., Morozov, P., Manickavel, S., McFarland, H. L. et al.** (2017). DND1 maintains germline stem cells via recruitment of the CCR4-NOT complex to target mRNAs. *Nature* **543**, 568-572. doi:10.1038/nature21690
- Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hubner, K. and 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. and Matsui, Y.** (1999). Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41**, 675-684. doi:10.1046/j.1440-169x.1999.00474.x
- 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. et al.** (2005). The *Ter* mutation in the dead end gene causes germ cell loss and testicular germ cell tumours. *Nature* **435**, 360-364. doi:10.1038/nature03595
- Yu, F.-X. and Guan, K.-L.** (2013). The Hippo pathway: regulators and regulations. *Genes Dev.* **27**, 355-371. doi:10.1101/gad.210773.112
- Zechel, J. L., Doerner, S. K., Lager, A., Tesar, P. J., Heaney, J. D. and 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. doi:10.1186/1471-2156-14-54
- Zhu, R., Iacovino, M., Mahen, E., Kyba, M. and Matin, A.** (2011). Transcripts that associate with the RNA binding protein, DEAD-END (*DND1*), in embryonic stem (ES) cells. *BMC Mol. Biol.* **12**, 37. doi:10.1186/1471-2199-12-37