© 2020. Published by The Company of Biologists Ltd | Development (2020) 147, dev192856. doi:10.1242/dev.192856

RESEARCH REPORT

An autoregulatory switch in sex-specific *phf7* transcription causes loss of sexual identity and tumors in the *Drosophila* female germline

Anne E. Smolko*, Laura Shapiro-Kulnane and Helen K. Salz‡

ABSTRACT

Maintenance of germ cell sexual identity is essential for reproduction. Entry into the spermatogenesis or oogenesis pathway requires that the appropriate gene network is activated and the antagonist network is silenced. For example, in *Drosophila* female germ cells, forced expression of the testis-specific PHD finger protein 7 (PHF7) disrupts oogenesis, leading to either an agametic or germ cell tumor phenotype. Here, we show that PHF7-expressing ovarian germ cells inappropriately express hundreds of genes, many of which are male germline genes. We find that the majority of genes under PHF7 control in female germ cells are not under PHF7 control in male germ cells, suggesting that PHF7 is acting in a tissue-specific manner. Remarkably, transcriptional reprogramming includes a positive autoregulatory feedback mechanism in which ectopic PHF7 overcomes its own transcriptional repression through promoter switching. Furthermore, we find that tumorigenic capacity is dependent on the dosage of *phf7*. This study reveals that ectopic PHF7 in female germ cells leads to a loss of sexual identity and the promotion of a regulatory circuit that is beneficial for tumor initiation and progression.

KEY WORDS: Germ cell fate, Oogenesis, Germline tumors, Sex determination

INTRODUCTION

Germ cell development culminates in the production of sexually dimorphic haploid gametes: sperm and eggs. In most animals, cells destined to become germ cells are set aside during embryogenesis and migrate to the developing gonad. There, they exhibit sex-specific division rates and gene expression programs, ultimately leading to meiosis and differentiation into morphologically and functionally distinct gametes (Lesch and Page, 2012). Germ cell development is not possible when the sexual identity of the germ cells and the surrounding somatic gonadal cells do not match (Salz et al., 2017). Therefore, successful reproduction requires that the appropriate sex-specific expression network be activated and the antagonist network be silenced.

In *Drosophila melanogaster* germ cells, the female/male decision is initially guided by the sex of the developing somatic gonad (Casper and van Doren, 2009; Hashiyama et al., 2011; Horabin et al., 1995; Staab et al., 1996; Wawersik et al., 2005). Extrinsic control is eventually lost, and sexual identity is maintained by cell-intrinsic mechanisms (Casper and van Doren, 2009). In female germ cells, maintenance of the embryonic sex fate decision requires the female-specific RNA-binding protein Sex lethal (SXL) (Chau et al., 2009; Schüpbach, 1985; Shapiro-Kulnane et al., 2015; Smolko et al., 2018). When germ cells lack SXL protein, differentiation is blocked and germ cell tumors are formed. Although loss of SXL leads to the ectopic sex-inappropriate transcription of hundreds of genes, dysregulation of one spermatogenesis gene, *PHD finger protein 7* (*phf7*), was found to be a major driver of the germ cell tumor phenotype (Shapiro-Kulnane et al., 2015).

PHF7 is a predicted chromatin reader that was first identified in a screen for genes expressed in male but not female embryonic germ cells (Yang et al., 2012, 2017). In the adult testes, protein expression is restricted to the nuclei of the undifferentiated germline stem cells and early spermatogonia. However, loss of *phf7* has only minimal impact on spermatogenesis. Mutant males harbor fewer differentiating spermatogonial cysts, resulting in fewer progeny than wild type. Reduced fecundity appears to be caused by a failure to control expression of a small set of male germ cell genes.

Although not essential for germ cell development in males, it is crucial to prevent PHF7 expression in female germ cells. PHF7-expressing ovarian germ cells fail to differentiate, resulting in an agametic or germ cell tumor phenotype (Shapiro-Kulnane et al., 2015; Yang et al., 2012). Interestingly, when PHF7-expressing XX germ cells develop in a sexually transformed somatic environment, they can produce sperm, albeit at a low frequency (Yang et al., 2012). These studies suggest that ectopic PHF7 expression is able to drive the germ cell towards a male developmental program. However, the impact of ectopic PHF7 on the transcriptional landscape is not known.

In this study, we combined genetic and genomic approaches to understand the consequences of ectopic PHF7 expression in ovarian germ cells. As expected, we find that a female to male identity switch underlies tumor formation. However, the majority of genes under PHF7 control in ovarian germ cells are not under PHF7 control in male germ cells. This suggests that PHF7 affects gene expression in a tissue-specific manner. Ectopic transcriptional reprogramming activity includes a positive autoregulatory feedback mechanism in which PHF7 can overcome transcriptional repression of other *phf7* copies in the genome. Together, our work supports a model in which PHF7 reprograms transcription in the female germline by redirecting chromatin remodeling complexes to inappropriately activate male germ cell genes, thus underscoring the importance of preventing the expression of lineage-inappropriate genes for maintaining tissue homeostasis.

Department of Genetics and Genome Sciences, Case Western Reserve University, School of Medicine, Cleveland, OH 44106-4955, USA.
*Present address: Division of Genetics, Brigham and Women's Hospital, Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.

‡Author for correspondence (hks@case.edu)

A.E.S., 0000-0002-1247-4364; H.K.S., 0000-0003-1120-7464

Handling Editor: Cassandra Extavour
Received 13 May 2020; Accepted 7 August 2020

Development (2020) 147, dev192856. doi:10.1242/dev.192856

RESULTS AND DISCUSSION

Deletion of the PHD fingers creates an inactive *phf7* allele

phf7 encodes a 520 amino acid protein with two adjacent N-terminal PHD fingers (Mitchell et al., 2019): a canonical PHD zinc finger (ZNF_PHD) domain and an extended PHD (ePHD) domain (amino acids 2-117). The protein also contains a RING-finger (ZNF_RING) domain (amino acids 132-184). PHD finger proteins are often involved in chromatin and transcriptional regulation. To test whether deleting this region impacts function *in vivo*, we used the tissue-specific GAL4/UAS induction system to force expression of *phf7* cDNA carrying an in-frame deletion of amino acids 68-111 (UASz::*phf7*$^{\Delta PHD}$). This transgene and a control wild-type transgene (UASz::*phf7*) were inserted into the genome via site-specific recombination into the same location to allow direct comparisons (Fig. 1A).

Using the germ cell-specific driver *nos-Gal4::VP16*, we found that expression delivered from the wild-type UASz-*phf7* transgene caused females to be sterile. Progression of egg chamber formation was analyzed by staining for DNA and Vasa, a germ cell-specific marker (Fig. 1B). The fly ovary is composed of 16 to 20 ovarioles, each of which is organized as an assembly line that generates egg chambers (Hinnant et al., 2020). At the anterior end of the ovariole, a structure called the germarium houses two to three germline stem cells. Each stem cell divides to create one daughter cell that remains a stem cell, and a second daughter cell that enters the differentiation pathway. Differentiation begins with four incomplete mitotic divisions to generate a 16-cell cyst that develops into an egg chamber containing one oocyte and 15 supporting nurse cells. As expected, the ovaries of *nos*>UASz-*phf7* mutant females exhibited morphological defects. The majority (74%, *n*=167) of *nos*>UASz-*phf7* mutant ovarioles were agametic (Fig. 1C). In the remaining 26% of the mutant ovarioles, we observed a tumor phenotype, defined by the accumulation of excess germ cells in the germarium and the failure to form egg chambers with nurse cells.

In sharp contrast, the *nos*>UASz-*phf7*$^{\Delta PHD}$ females produced progeny and their ovaries were similar to wild type. The failure to generate a mutant phenotype in this ectopic expression assay indicates that deleting the PHD finger domain inactivates *phf7*. Overall, these data show that ectopic expression in female germ cells is sufficient to disrupt oogenesis and suggests that the PHD finger domains are required. However, we cannot rule out other explanations for the loss of ectopic *phf7*$^{\Delta PHD}$ activity, such as protein instability, because antibodies against PHF7 are unavailable.

Ectopic *Phf7* functions in a feedback loop

Autoregulatory feedback mechanisms are often used to maintain expression of fate determining genes (Crews and Pearson, 2009). Therefore, we investigated whether forced *Phf7* expression from the transgene could induce testis-like expression at the endogenous locus.

To monitor expression at the endogenous locus we used CRISPR to replace the open reading frame with GFP (*phf7*$^{\Delta ORF}$*::GFP*; Fig. 2A). We determined that the reporter does not interfere with female fertility and that it recapitulates the sex-specific RNA and protein expression patterns of PHF7 (Fig. S1). When ectopic *phf7* expression (nos>UASz-*phf7*) was induced in this background, GFP protein was detected in tumors (Fig. 2B; Fig. S2). We conclude that ectopic PHF7 stimulates testis-like protein expression from an edited allele at the endogenous locus. In contrast, we did not observe GFP staining when ectopic expression of the inactive *phf7*$^{\Delta PHD}$ allele was induced in this background (*phf7*$^{\Delta ORF}$*::GFP; nos>UASz-phf7*$^{\Delta PHD}$), demonstrating that functional PHF7 protein is required for transactivation.

Next, we sought to demonstrate positive autoregulation using a different genetic paradigm. We forced PHF7 expression from the endogenous locus via a UASp-containing EP transposable element insertion, *P{EPgy2}phf7*^{EY03023} (*phf7*^{EY}), located within the first intron (Fig. 2A). Driving expression of *phf7*^{EY} with the germline *nos-GAL4* driver has been shown to drive tumor formation, albeit at a low frequency (Shapiro-Kulnane et al., 2015; Yang et al., 2012). We assessed transactivation with a human influenza hemagglutinin (HA)-tagged *phf7* locus embedded in a 20 kb bacterial artificial chromosome (BAC) rescue construct located on the third chromosome (Fig. 2A). This transgene has been shown to serve

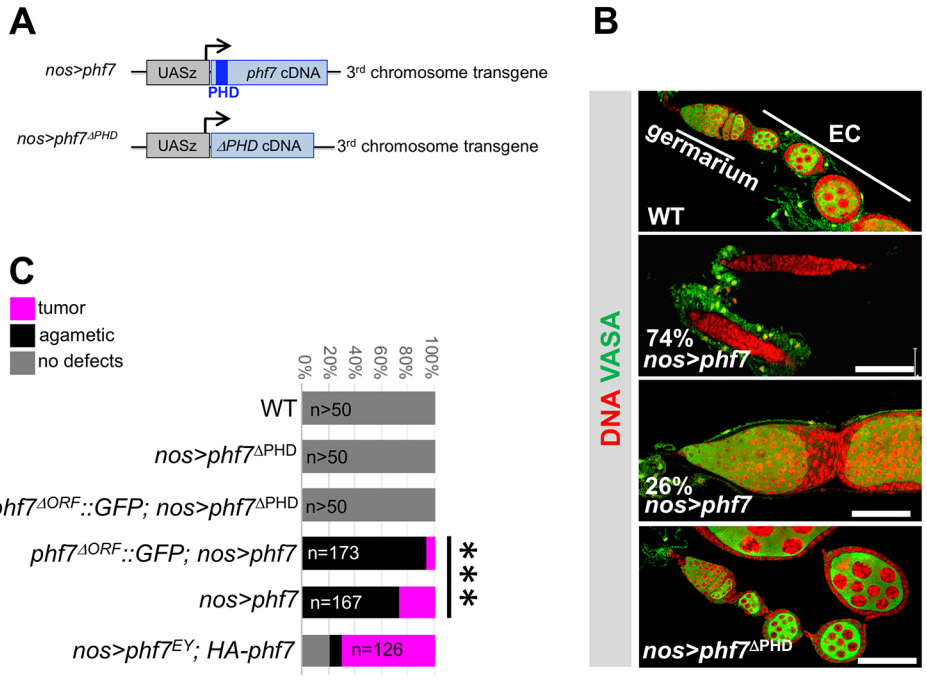

Fig. 1. Deletion of the PHD domain disables ectopic PHF7 activity in female germ cells. (A) Transgenes used to express wild-type and mutant *phf7* cDNAs via the GAL4/UAS system. (B) Expression of wild-type (*nos>phf7*) but not mutant (*nos>phf7*$^{\Delta PHD}$) PHF7 in germ cells disrupts oogenesis. Confocal images of mutant and control ovarioles, including the germarium and egg chambers (EC), stained for Vasa (green) and DNA (red). Scale bars: 50 µm. (C) Quantification of the different mutant ovariole phenotypes. An unpaired two-tailed Fisher's exact test was used to compare the phenotypes of *nos>phf7* and *phf7*$^{\Delta ORF}$*::GFP; nos>phf7*. ***P<0.0001.

Development (2020) 147, dev192856. doi:10.1242/dev.192856

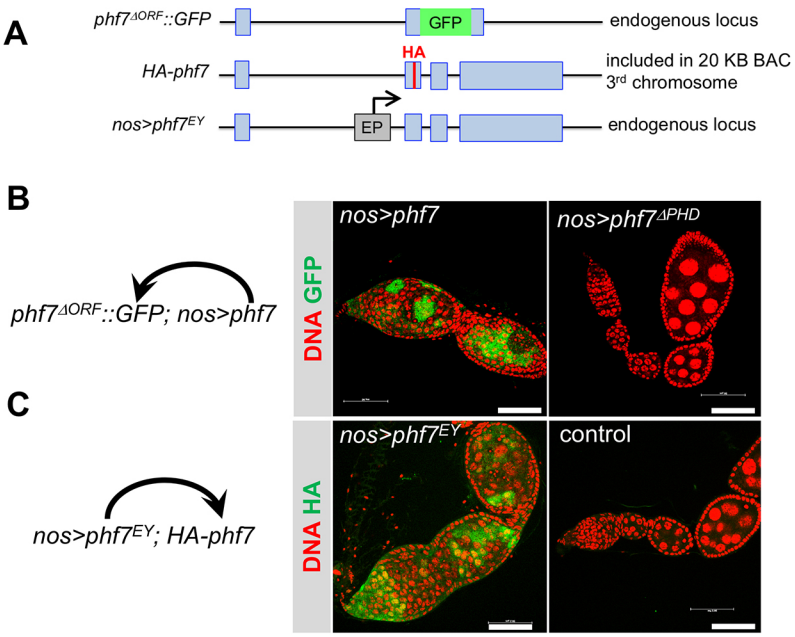

Fig. 2. Ectopic PHF7 stimulates testis-like expression from reporter alleles. (A) Schematic of the knock-in *phf7*$^{\Delta ORF}$*::GFP* allele, the third chromosome HA-tagged *phf7* reporter allele and the *phf7*^{EY} allele. Exons, represented by blue boxes, are connected by black lines representing introns. (B) Ectopic PHF7 can induce testis-like expression from the knock-in *phf7*$^{\Delta ORF}$*::GFP* allele. Confocal images of ovarioles from *phf7*$^{\Delta ORF}$*::GFP*; *nos>phf7* and control *phf7*$^{\Delta ORF}$*::GFP*;*nos>phf7*$^{\Delta PHD}$ females stained for GFP protein (green) and DNA (red). (C) Ectopic PHF7 can induce testis-like expression from a third chromosome HA-tagged reporter construct. Confocal images of ovarioles from mutant *nos>phf7*^{EY}*; HA-phf7* and control *nos; HA-phf7* females stained for HA (green) and DNA (red). Scale bars: 50 µm.

as a faithful reporter of the sex-specific protein expression pattern of PHF7 (Shapiro-Kulnane et al., 2015; Yang et al., 2012). When these two genetic elements were combined (*nos>phf7*^{EY}*; HA-phf7*), ectopic HA-PHF7 protein was detected in tumors (Fig. 2C). We conclude that ectopic PHF7 stimulates testis-like protein expression from the transgenic tagged copy of *phf7*. This shows that ectopic PHF7 can stimulate testis-like expression from any *phf7* allele.

Together, these data suggest that once expressed in ovarian germ cells, PHF7 can increase its own expression via a positive autoregulatory feedback loop. Thus, there may be a correlation between copy number and phenotype. In this context it is interesting to note that in the absence of functional *phf7* copies at the endogenous locus (*phf7*$^{\Delta ORF}$*::GFP; nos>UASz-phf7*), the frequency of tumor formation upon induction with a third chromosome transgene was only 6%, and the majority of the mutant ovarioles were agametic (Fig. 1C). Two functional copies at the endogenous locus, using the same third chromosome transgene to ectopically express *phf7*, significantly shifted the distribution of the mutant phenotypes towards a germ cell tumor phenotype (26% tumors in *nos>UASz-phf7*; $P<0.0001$). It remains unclear why forcing *phf7* expression from a transgene in a *phf7* mutant background is toxic to germ cells, whereas in a wild-type background this same transgene allows germ cell survival and tumor formation. Irrespective of mechanism, we found that the correlation between increased *phf7* copy number and the tumor phenotype extended to a different genetic paradigm in which three full length copies of *phf7* (*nos>phf7*^{EY}*; HA-phf7*) led to a dramatic increase in the penetrance of the tumor phenotype to 70% (Fig. 1C).

Ectopic PHF7 transactivates via promoter switching

Although PHF7 protein is normally limited to male germ cells, *phf7* mRNAs are expressed in both male and female germ cells (Fig. 3A). Sex-specific regulation is achieved by a mechanism that relies primarily on alternative promoter choice and transcription start site

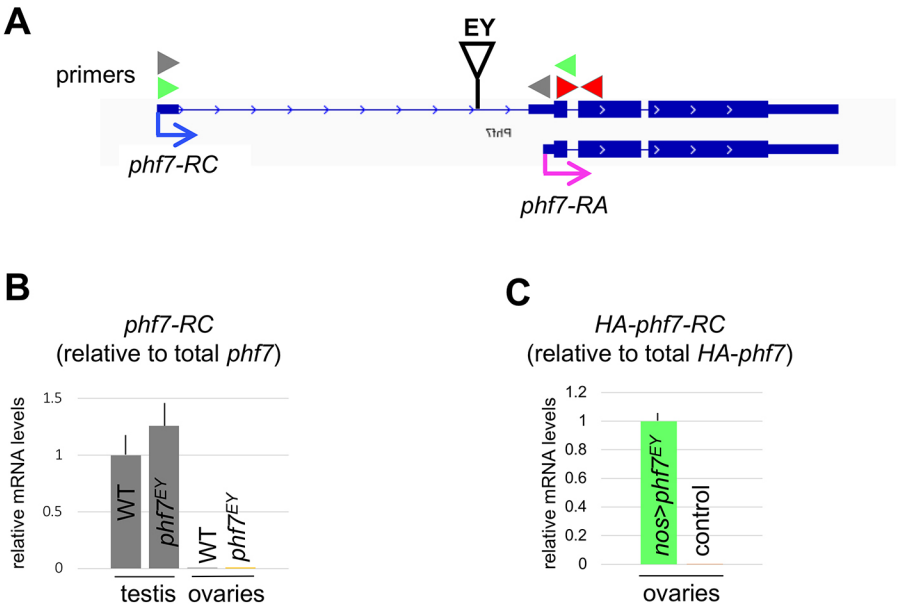

Fig. 3. Ectopic PHF7 transactivates via promoter switching. (A) Genome browser view of the two major *phf7* transcripts, the testis-specific *phf7-RC* (blue arrow) and *phf7-RA* (pink arrow). Exons are represented by blue blocks connected by horizontal lines representing introns. The insertion site of the UAS-containing insertion in *phf7*^{EY} is represented by a triangle. The location of primers for RT-qPCR are indicated by colored arrowheads. (B) RT-qPCR measurements of *phf7-RC* transcript in ovaries and testis from wild-type (WT) and *phf7*^{EY} mutant animals (gray primer pairs). Expression normalized to the total level of *phf7* (red primer pairs). (C) RT-qPCR measurements of transgenic *HA-phf7-RC* transcript in ovaries from mutant *nos>phf7*^{EY}*; HA-phf7* and control *nos; HA-phf7* females. Primers to the HA tag, located at the beginning of the open reading frame, were used to distinguish transgenic *phf7* RNA from the endogenous products: green primer pairs for *HA-phf7-RC* and modified red primer pairs for total *HA-phf7*. Data are mean±s.d. (three biological replicates).

Development (2020) 147, dev192856. doi:10.1242/dev.192856

(TSS) selection. In ovaries, transcription from the downstream TSS produces an mRNA, *phf7-RA*, but no protein is detectable. In testes, transcription from the upstream TSS produces a longer translatable mRNA, called *phf7-RC*. Our discovery that ectopic PHF7 can stimulate protein expression from any *phf7* allele suggests a mechanism that includes transcriptional switching to the male-specific TSS.

With the identification of genetic conditions that increased the penetrance of the tumor phenotype to 70% (*nos>phf7^{EY}; HA-phf7*), we were able to test this hypothesis by assaying for the presence of the testis-specific *phf7-RC* RNA isoform. Using RT-PCR, we found that in control *phf7^{EY}* ovaries, the EP insertion by itself did not interfere with the sexually dimorphic transcription pattern of *phf7* (Fig. 3B). Furthermore, we found that transcription from the transgenic HA-tagged *phf7* locus was also regulated appropriately as no HA-tagged *phf7-RC* (*HA-phf7-RC*) transcript was detected in ovaries (Fig. 3C). However, in *nos>phf7^{EY}; HA-phf7* mutant ovaries, we found that *HA-phf7-RC* was ectopically expressed. This work demonstrates that ectopic PHF7 stimulates testis-like transcription from the transgenic tagged copy of *phf7*.

In agreement with our RT-PCR analysis, alignment of RNA-sequencing (RNA-seq) data showed that the testis-specific *phf7-RC* transcript is ectopically expressed in *nos>phf7^{EY}; HA-phf7* mutant ovaries (Fig. 4A). These data also illustrate that novel RNAs are produced from the region near the EP transposon insertion site within the first intron. We conclude that forced PHF7 expression initiates an autoregulatory feedback loop that overcomes transcriptional repression of the testis-specific promoter.

Ectopic PHF7 provokes testis-specific gene expression programming

To gain a genome-wide view of the expression changes downstream of ectopic PHF7 expression, we compared the transcriptomes of *nos>phf7^{EY}; HA-phf7* mutant ovaries with wild-type ovaries from newborn (0-24 h) females. Because ovaries from newborn females lack late-stage egg chambers, we reasoned that this comparison would minimize the identification of gene expression changes unrelated to the mutant phenotype. In *nos>phf7^{EY}; HA-phf7* mutant ovaries, we identified 835 genes downregulated at least twofold [false discovery rate (FDR)<0.05; Table S1], and 799 genes upregulated at least twofold (FDR<0.05; Table S2).

Of the upregulated genes, 286 were not detectable in wild-type ovaries [fragments per kilobase million (FPKM)<1; Fig. 4B, Table S3]. Hierarchical clustering of the ectopically expressed genes revealed a tissue-specific signature most similar to samples from adult testis (Fig. 4C). Indeed, 44% (126/286) of the ectopically expressed tumor genes were highly expressed in normal testis relative to other tissues (in blue, Table S3). Yet, the *nos>phf7^{EY}; HA-phf7* mutant germ cells resembled undifferentiated germ cells. We therefore investigated whether the ectopic genes expressed in tumors are normally expressed in healthy undifferentiated germ cells. To this end, we compared our gene set with genes expressed in wild-type undifferentiated germ cells obtained by single cell RNA sequencing of adult and larval ovaries (Jevitt et al., 2020; Slaidina et al., 2020). Although we did not observe any overlap between our set of ectopically expressed genes and expressed genes in undifferentiated germ cells isolated from adult ovaries, we did identify an overlap of nine genes (3% of the ectopic gene set) expressed in larval primordial germ cells (Table S3). Thus, the vast majority of the genes identified as ectopically expressed in tumors are induced (either directly or indirectly) by inappropriate *phf7* expression. Strikingly, several of the aberrantly expressed genes encode components of the specialized transcription and translation machinery that operates in the male germline (White-Cooper and Caporilli, 2013). These include Taf12L, Dany, RpL37b, eIF4E-3 and eIF4E-6 (Fig. 4D). The acquisition of these male germ cell features reveals that forcing PHF7 expression leads to a loss of sexual identity and adoption of a male fate.

Given that ectopic expression of PHF7 in ovarian germ cells can drive transcription of normally silenced male germ cell genes, we hypothesized that PHF7 controls expression of the same set of genes in male germ cells. In an analysis focused on *phf7* function in the male germline, only 45 genes were found to be differentially expressed in *phf7* loss-of-function testis (Yang et al., 2017). Only one of these genes (CG15599) was ectopically expressed in PHF7-expressing ovarian germ cells. An additional 16 genes were

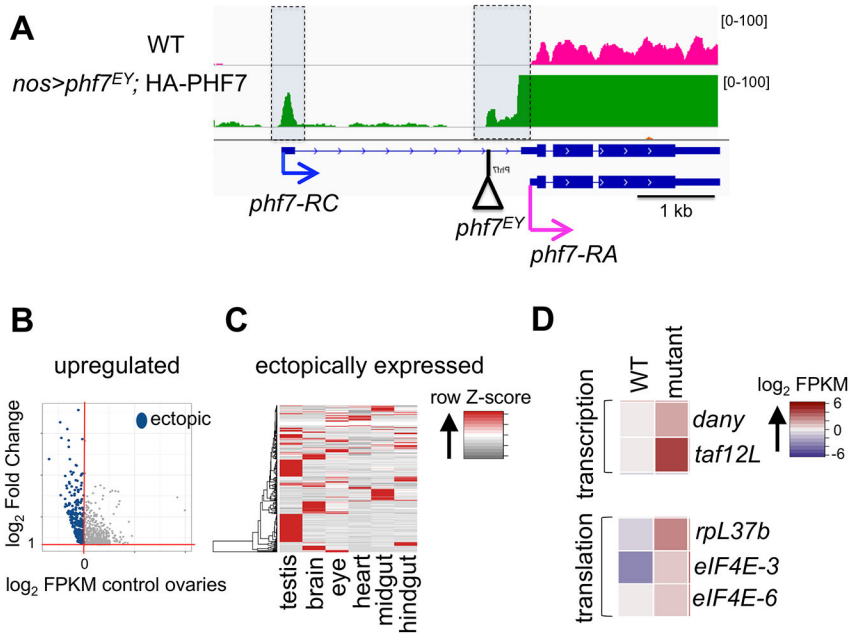

Fig. 4. Ectopic *phf7* leads to female-to-male reprogramming of many genes including itself. (A) Genome browser view of the *phf7* locus. *phf7-RC* is represented by a blue arrow and *phf7-RA* is represented by a pink arrow. Exons are represented by blue blocks connected by horizontal lines representing introns. The insertion site of the UAS-containing insertion in *phf7^{EY}* is represented by a triangle. Wild-type (WT) RNA-seq reads are in pink and *nos>phf7^{EY}; HA-phf7* RNA-seq reads are in green. All tracks are viewed at the same scale. (B) Scatter plots of significantly upregulated genes in *nos>phf7^{EY}; HA-phf7* mutant ovaries. The $\log_2$ fold change in gene expression is plotted against the $\log_2$ of the FPKM values in wild-type ovaries. Blue points indicate ectopically expressed genes (genes that are not expressed in wild-type ovaries). (C) Tissue expression clustering of the ectopically expressed genes in *nos>phf7^{EY}; HA-phf7* mutant ovaries, displayed as a Z-score heatmap. Each column is an adult tissue. Each row is an ectopically expressed gene. (D) Expression values of selected testis-specific genes in wild-type ovaries and *nos>phf7^{EY}; HA-phf7* mutant ovaries displayed as a heatmap of log2 FPKM values from RNA-seq analysis.

Development (2020) 147, dev192856. doi:10.1242/dev.192856

significantly affected in both PHF7-expressing ovarian germ cells and *phf7* loss-of-function male germ cells (Table S4). Thus, contrary to our expectations, we found that the majority of genes under PHF7 control were different in male and female germ cells. These observations imply that tissue-specific factors define the genes and pathways under PHF7 control and raises the possibility that the role of PHF7 in female germ cells might be different from its role in male germ cells.

In summary, our studies highlight the importance of preventing expression of lineage-inappropriate genes for maintaining tissue homeostasis. We demonstrate that once expressed in female germ cells, PHF7 can increase its own expression via a positive autoregulatory feedback loop. Our data further suggest that ectopic PHF7 is prerequisite to activate a male germ cell transcriptional program that drives tumor initiation and progression in female germ cells. PHF7 is a presumed chromatin reader as it preferentially binds to H3K4me2 *in vitro*, a mark generally enriched at transcription start sites (Yang et al., 2012). However, the mechanism by which PHF7 controls gene expression *in vivo* has remained elusive owing to the lack of experimentally confirmed target genes. We speculate that, in female germ cells, PHF7 redirects chromatin remodeling complexes to inappropriately activate male germ cell genes. Future studies will focus on identifying the genes directly controlled by PHF7 to reveal the mechanism by which PHF7 fuels the oncogenic gene expression program.

MATERIALS AND METHODS

Drosophila stocks and culture conditions

The wild-type reference strain was y^1, w^1 [Bloomington *Drosophila* Stock Center (BDSC), 1495]. The following stocks were used to ectopically express *phf7* in female germ cells: *P*{GAL4::VP16-nos.UTR} (BDSC, 4937); *P*{EPgy2}Phf7^{EY03023} (BDSC, 15894); *P*{UASz-Phf7} (this study); and *P*{UASz-Phf7$^{\Delta PHD}$} (this study). The following stocks were used to report on *phf7* gene activity: The HA-tagged *phf7* transgene PBac{3XHA-PHF7}, generated in the Van Doren lab by tagging the *phf7* locus included in the 20 kb BAC construct CH322-177L19 and inserting it into the 65B2 PBac{y[+]-attP-3B}VK00033 site via *phi-C31* catalyzed integration (Yang et al., 2012); and the *phf7$^{\Delta ORF}$::GFP* allele (this study).

Drosophila stocks were fed and maintained with molasses and kept at 25°C. Crosses to drive ectopic expression with the UASz transgenes were set up at 29°C and the adults were aged 3 to 5 days before gonad dissection. Crosses to generate ectopic expression with *P*{EPgy2}Phf7^{EY03023} were set up at 18°C and the adults were transferred to 29°C for 10 days before gonad dissection.

Generation of transgenic lines

Constructs were generated in the UASz expression vector to maximize expression in the female germline (DeLuca and Spradling, 2018). The *P*{UASz-Phf7} transgene was constructed by cloning the *phf7*cDNA (LD43541) into the mini-white containing *pUASz1.1* transformation vector [*Drosophila* Genomics Resource Center (DGRC), 1433]. To construct the *P*{UASz-Phf7$^{\Delta PHD}$} transgene, a New England Biolabs Q5 mutagenesis kit was used to generate a deletion in the coding region using the primers: F-5′-TGCCATCAGCATGTGCTG-3′ and R-5′-AAGCAAACGGCAGCGGTT-3′. The transgenic constructs were sent for *phi-C31* catalyzed integration into the 65B2 PBac{y[+]-attP-3B}VK00033 site (Rainbow Transgenic Flies).

Generation of the *phf7$^{\Delta ORF}$-GFP* allele

The *phf7$^{\Delta ORF}$::GFP* allele was generated using CRISPR to replace the *phf7* open reading frame with GFP. To generate the *phf7$^{\Delta ORF}$* deletion allele, the following guide RNAs were synthesized and ligated into the pU6-BbsI-chiRNA vector (Addgene #45946): gRNA1, F-5′-CTTCGGTCACCGGAAACGCATCCA-3′ and R-5′-AAACTGGATGCGTTTCCGGTGACC-3′; and gRNA2: F-5′-CTTCGAATCCTTGCGGCTGGCCATG-3′ and R-5′-AAACCATGGCCAGCCGCAAGGATTC-3′.

Homology arms (1 kb) were generated by PCR and cloned into the pHD-dsRed-attP (Addgene #51019). Guide RNAs and the donor vector were co-injected into *vas-Cas9* embryos (BDSC, 51324; Rainbow Transgenic Flies). To insert GFP, the *gfp* coding sequence was cloned into the attB-containing RIV-white$^+$ transformation vector (DGRC, 1330) and sent for *phi-C31* catalyzed integration into the attP site present in *phf7$^{\Delta ORF}$* (Rainbow Transgenic Flies).

Immunofluorescence and image analysis

Ovaries and testes were fixed and stained according to standard procedures with the following primary antibodies: rat anti-HA (1:500; Roche, 11867423001; RRID: AB_390919); rat anti-Vasa (1:100; Developmental Studies Hybridoma Bank; RRID, AB_760351); and rabbit anti-GFP (1:2500; Thermo Fisher Scientific, A-11122; RRID, AB_221569). Staining was detected with the following conjugated antibodies: fluorescein (FITC) anti-rat (1:200; Jackson ImmunoResearch, A-21434; RRID, AB_2535855); FITC anti-rabbit (1:200; Jackson ImmunoResearch, 111-095-003; RRID, AB_2337972), Alexa Fluor 555 anti-rat (1:200; Thermo Fisher Scientific, A-21434; RRID, AB_2535855) or Alexa Fluor 555 anti-rabbit (1:200; Thermo Fisher Scientific, A-21428; RRID, AB_2535849). TO-PRO-3 iodide (Thermo Fisher Scientific, T3605) was used to stain DNA.

Images were taken using a Leica SP8 confocal microscope with 1024×1024 pixel dimensions, a scan speed of 600 Hz, and a frame average of three. Sequential scanning was performed for each channel and three *z* stacks were combined for each image. Processed images were compiled using the Gnu Image Manipulation Program and Microsoft PowerPoint.

qRT-PCR and data analysis

RNA was extracted from dissected gonads using TRIzol (Thermo Fisher Scientific, 15596026) and DNase RQ1 (Promega, M6101). Quantity and quality were measured using a NanoDrop spectrophotometer. cDNA was generated by reverse transcription using a SuperScript First-Strand Synthesis System for RT-PCR kit (Thermo Fisher Scientific, 11904018) using random hexamers. qPCR was performed using Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, 4367659) with an Applied Biosystems 7300 Real-Time PCR system. PCR steps were as follows: 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Melt curves were generated with the following parameters: 95°C for 15 s, 60°C for 1 min, 95°C for 15 s, and 60°C for 15 s. Measurements were taken in biological triplicate with two technical replicates each. Relative transcript levels were calculated using the 2-ΔΔCt method (Livak and Schmittgen, 2001). To measure RNA levels in *phf7$^{\Delta ORF}$::GFP* gonads, the following primers were used: for *phf7-RC*, F-5′-AGTTCGGGAATTCAACGCTT-3′ and R-5′-GAGATAGCCCTGCAGCCA-3′; and for *gfp*, F-5′-ACGTAAACGGCCACAAGTTC-3′ and R-5′-AAGTCGTGCTGCTTCATGTG-3′. To measure RNA levels in wild-type and *phf^{EY}* gonads, the following primers were used: for *phf7-RC*, F-5′-AGTTCGGGAATTCAACGCTT-3′ and R-5′-GAGATAGCCCTGCAGCCA-3′; and for total *phf7*, F-5′-GAGCTGATCTTCGGCACTGT-3′ and R-5′-GCTTCGATGTCCTCCTTGAG-3′. To measure RNA levels from the PBac{3XHA-PHF7} transgene, the following primers were used: for *HA-phf7-RC*, F-5′-CTGCAGGGCTATCTCCGATA-3′ and R-5′-TAGCCCGCATAGTCAGGAAC-3′; and for total *HA-phf7*, F-5′-CGATGTTCCTGACTATGCGG-3′ and R-5′-ACAGTGCCG-AAGATCAGCT-3′.

RNA-seq and data analysis

Total RNA was extracted from dissected ovaries using TRIzol. RNA quality was assessed using a Qubit fluorometer and an Agilent Bioanalyzer. Libraries were generated using a TruSeq Stranded Total RNA kit (Illumina, 20020599). Sequencing was completed on two biological replicates of each genotype with the Illumina HiSeq 2500 v2 with 100 bp paired end reads. Sequencing reads were aligned to the *Drosophila* genome (UCSC dm6) using TopHat (2.1.0) and duplicate reads were removed (Trapnell et al., 2009). The number of reads generated were as follows: wild type-1, 12,598,832 (93.7% unique); wild type-2, 13,855,129 (93.5% unique); *phf7*-1, 12,804,022 (93.6% unique); and *phf7*-2, 12,219,346 (92.9% unique). The correlation between biological replicates was assessed using the

plotCorrelation tools in deepTools, with a 50 bp bin size. The correlation coefficient of the biological replicates was R=0.99. Differential analysis was completed using Cuffdiff (2.2.1) (Trapnell et al., 2012). Genes were considered differentially expressed if they exhibited a twofold or higher change relative to wild type with FDR<0.05.

The screenshot of the expression data is from Integrated Genome Viewer (IGV). To account for the differences in sequencing depth when creating IGV screenshots, the processed RNA-seq alignment files were scaled to the number of reads in the wild-type file. This was carried out with deepTools bigwigCompare using the scale factors parameter with a bin size of five. Scatter plots were generated using the ggplot function in R. Genes that were expressed in mutant ovaries (FPKM≥1) but not wild-type ovaries (FPKM<1) were called ectopic.

Tissue expression clustering of the ectopically expressed genes was performed to identify tissue-specific signatures. Expression values normalized to the whole fly were extracted from FlyAtlas (versions 1 and 2) (Leader et al., 2018; Robinson et al., 2013). Genes were considered to be testis enriched if their enrichment score was at least 1.5. The heatmap to compare the tissue expression profile of these genes per tissue was generated in R with heatmap.2 (gplots). Genes were clustered and normalized per row.

Acknowledgements
We thank Dr Mark Van Doren, the Bloomington *Drosophila* Stock Center and the Iowa Developmental Studies Hybridoma Bank for fly stocks and antibodies. We also thank the Genomics Core at the Case Western Reserve University for performing the RNA-seq, and Jane Heatwole for fly food. Imaging was performed using equipment purchased through the National Institutes of Health (S10OD016164).

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: A.E.S., H.K.S.; Methodology: A.E.S., L.S.-K., H.K.S.; Validation: A.E.S., L.S.-K.; Formal analysis: A.E.S., L.S.-K.; Investigation: A.E.S., L.S.-K.; Resources: A.E.S.; Data curation: A.E.S., H.K.S.; Writing - original draft: H.K.S.; Writing - review & editing: A.E.S., L.S.-K., H.K.S.; Visualization: A.E.S., L.S.-K.; Supervision: H.K.S.; Project administration: H.K.S.; Funding acquisition: H.K.S.

Funding
This work was supported by the National Institutes of Health (R01GM129478 to H.K.S. and T32GM008056 to A.E.S.). Deposited in PMC for release after 12 months.

Data availability
The wild-type ovary mRNA-seq data sets have been deposited in GEO under accession number GSE109850, and the mutant ovary mRNA-seq data sets are available under accession number GSE150213.

Supplementary information
Supplementary information available online at
https://dev.biologists.org/lookup/doi/10.1242/dev.192856.supplemental

Peer review history
The peer review history is available online at
https://dev.biologists.org/lookup/doi/10.1242/dev.192856.reviewer-comments.pdf

References
Casper, A. L. and van Doren, M. (2009). The establishment of sexual identity in the Drosophila germline. *Development* **136**, 3821-3830. doi:10.1242/dev.042374

Chau, J., Kulnane, L. S. and Salz, H. K. (2009). Sex-lethal facilitates the transition from germline stem cell to committed daughter cell in the Drosophila ovary. *Genetics* **182**, 121-132. doi:10.1534/genetics.109.100693

Crews, S. T. and Pearson, J. C. (2009). Transcriptional autoregulation in development. *Curr. Biol.* **19**, R241-R246. doi:10.1016/j.cub.2009.01.015

DeLuca, S. Z. and Spradling, A. C. (2018). Efficient expression of genes in the Drosophila germline using a UAS promoter free of interference by Hsp70 piRNAs. *Genetics* **209**, 381-387. doi:10.1534/genetics.118.300874

Hashiyama, K., Hayashi, Y. and Kobayashi, S. (2011). Drosophila sex lethal gene initiates female development in germline progenitors. *Science* **333**, 885-888. doi:10.1126/science.1208146

Hinnant, T. D., Merkle, J. A. and Ables, E. T. (2020). Coordinating proliferation, polarity, and cell fate in the drosophila female germline. *Front. Cell Dev. Biol.* **8**, 19. doi:10.3389/fcell.2020.00019

Horabin, J. I., Bopp, D., Waterbury, J. and Schedl, P. (1995). Selection and maintenance of sexual identity in the Drosophila germline. *Genetics* **141**, 1521-1535.

Jevitt, A., Chatterjee, D., Xie, G., Wang, X.-F., Otwell, T., Huang, Y.-C. and Deng, W.-M. (2020). A single-cell atlas of adult Drosophila ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLoS Biol.* **18**, e3000538. doi:10.1371/journal.pbio.3000538

Leader, D. P., Krause, S. A., Pandit, A., Davies, S. A. and Dow, J. A. T. (2018). FlyAtlas 2: a new version of the Drosophila melanogaster expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic Acids Res.* **46**, D809-D815. doi:10.1093/nar/gkx976

Lesch, B. J. and Page, D. C. (2012). Genetics of germ cell development. *Nat. Rev. Genet.* **13**, 781-794. doi:10.1038/nrg3294

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) Method. *Methods* **25**, 402-408. doi:10.1006/meth.2001.1262

Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., Brown, S. D., Chang, H.-Y., El-Gebali, S., Fraser, M. I. et al. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351-D360. doi:10.1093/nar/gky1100

Robinson, S. W., Herzyk, P., Dow, J. A. T. and Leader, D. P. (2013). FlyAtlas: database of gene expression in the tissues of Drosophila melanogaster. *Nucleic Acids Res.* **41**, D744-D750. doi:10.1093/nar/gks1141

Salz, H. K., Dawson, E. P. and Heaney, J. D. (2017). Germ cell tumors: insights from the Drosophila ovary and the mouse testis. *Mol. Reprod. Dev.* **84**, 200-211. doi:10.1002/mrd.22779

Schüpbach, T. (1985). Normal female germ cell differentiation requires the female X chromosome to autosome ratio and expression of sex-lethal in Drosophila melanogaster. *Genetics* **109**, 529-548.

Shapiro-Kulnane, L., Smolko, A. E. and Salz, H. K. (2015). Maintenance of Drosophila germline stem cell sexual identity in oogenesis and tumorigenesis. *Development* **142**, 1073-1082. doi:10.1242/dev.116590

Slaidina, M., Banisch, T. U., Gupta, S. and Lehmann, R. (2020). A single-cell atlas of the developing Drosophila ovary identifies follicle stem cell progenitors. *Genes Dev.* **34**, 239-249. doi:10.1101/gad.330464.119

Smolko, A. E., Shapiro-Kulnane, L. and Salz, H. K. (2018). The H3K9 methyltransferase SETDB1 maintains female identity in Drosophila germ cells. *Nat. Commun.* **9**, 4155. doi:10.1038/s41467-018-06697-x

Staab, S., Heller, A. and Steinmann-Zwicky, M. (1996). Somatic sex-determining signals act on XX germ cells in Drosophila embryos. *Development* **122**, 4065-4071.

Trapnell, C., Pachter, L. and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *J. Gerontol.* **25**, 1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L. and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562-578. doi:10.1038/nprot.2012.016

Wawersik, M., Milutinovich, A., Casper, A. L., Matunis, E., Williams, B. and van Doren, M. (2005). Somatic control of germline sexual development is mediated by the JAK/STAT pathway. *Nature* **436**, 563-567. doi:10.1038/nature03849

White-Cooper, H. and Caporilli, S. (2013). Transcriptional and post-transcriptional regulation of Drosophila germline stem cells and their differentiating progeny. *Adv. Exp. Med. Biol.* **786**, 47-61. doi:10.1007/978-94-007-6621-1_4

Yang, S. Y., Baxter, E. M. and van Doren, M. (2012). Phf7 controls male sex determination in the Drosophila germline. *Dev. Cell* **22**, 1041-1051. doi:10.1016/j.devcel.2012.04.013

Yang, S. Y., Chang, Y.-C., Wan, Y. H., Whitworth, C., Baxter, E. M., Primus, S., Pi, H. and van Doren, M. (2017). Control of a novel spermatocyte-promoting factor by the male germline sex determination factor PHF7 of Drosophila melanogaster. *Genetics* **206**, 1939-1949. doi:10.1534/genetics.117.199935

Fig. S1 The *phf7$^{\Delta ORF}$::GFP* allele is sex-specifically regulated. (A) Schematic of the *phf7$^{\Delta ORF}$::GFP* allele in which the open reading frame was replaced by GFP (green box). Primers for RT-qPCR are indicated by arrowheads: Gray for *phf7-RC,* and Green for GFP. (B) RT-qPCR measurements of *phf7-RC* transcript in *phf7$^{\Delta ORF}$::GFP* ovaries and testis. Error Bars indicate standard deviation of three biological replicates. (C) Confocal images of testis from hemizygous *phf7$^{\Delta ORF}$::GFP* males and ovaries from homozygous *phf7$^{\Delta ORF}$::GFP* females stained for GFP protein (green) and DNA (red). Scale bar, 50 μm.

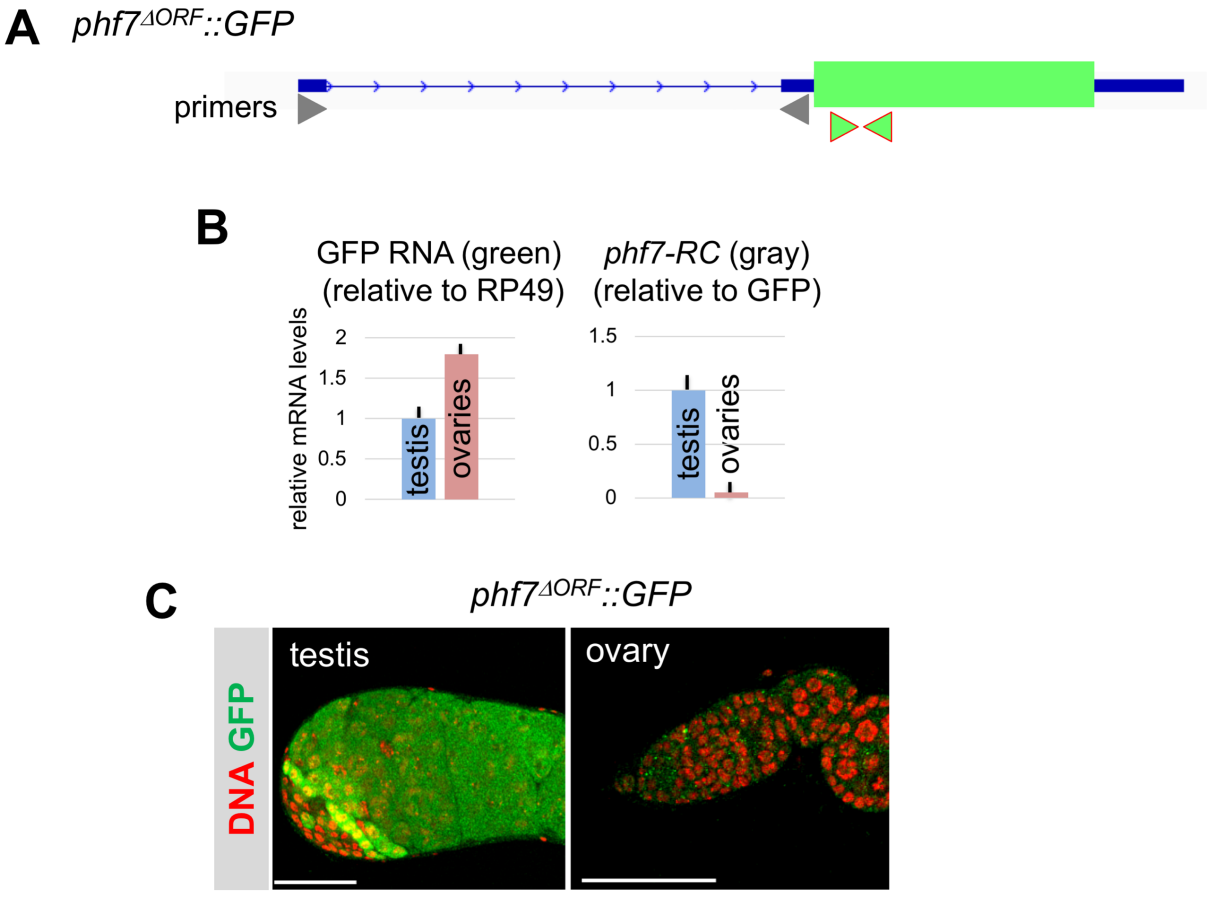

Fig. S2. Ectopic PHF7 stimulates testis-like expression from the knock-in *phf7$^{\Delta ORF}$::GFP* allele Ectopic PHF7 can induce testis-like expression from the knock-in *phf7$^{\Delta ORF}$::GFP* allele. Confocal images of ovarioles from *phf7$^{\Delta ORF}$::GFP*; *nos>phf7* and control *phf7$^{\Delta ORF}$::GFP*;*nos>phf7$^{\Delta PHD}$* females stained for GFP protein (green) and DNA (red). Scale bar, 50 μm.

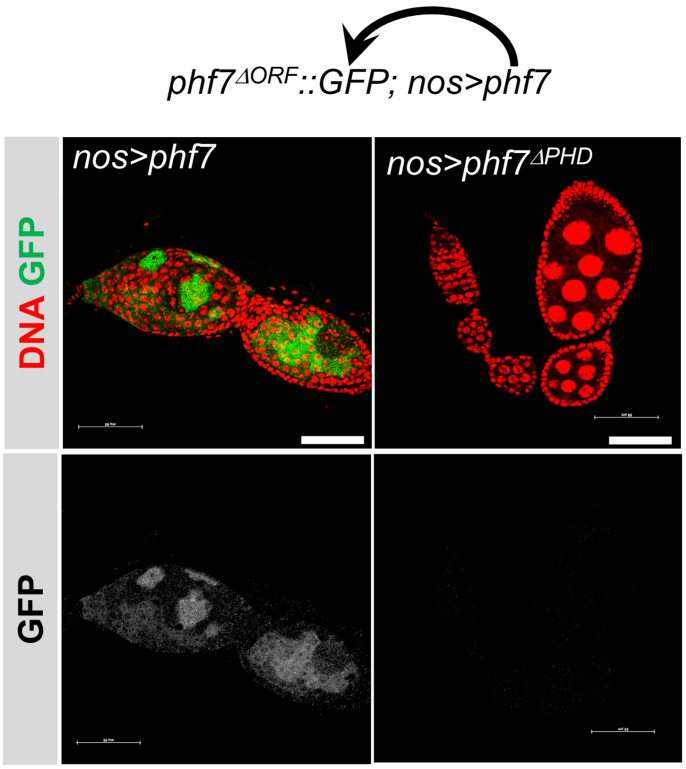

Table S1

Click here to Download Table S1

Table S2

Click here to Download Table S2

Table S3

Click here to Download Table S3

Table S4

Click here to Download Table S4