Chromatin and gene expression changes during female Drosophila germline stem cell development illuminate the biology of highly potent stem cells

Liang-Yu Pang[2], Steven DeLuca[3], Haolong Zhu, John M. Urban, and Allan C. Spradling[1]

Howard Hughes Medical Institute

Carnegie Institution for Science

Baltimore, MD 21218

[1]corresponding author: spradling@carnegiescience.edu

[2]current address: <gugulypang@gmail.com>.

[3]current address: Montana State University, Bozeman, MT

Abstract

Highly potent animal stem cells either self renew or launch complex differentiation programs, using mechanisms that are only partly understood. Drosophila female germline stem cells (GSC) perpetuate without change over evolutionary time and generate cystoblast daughters that develop into nurse cells and oocytes. Cystoblasts initiate differentiation by generating a transient syncytial state, the germline cyst, and by increasing pericentromeric H3K9me3 modification, actions likely to suppress transposable element activity. Relatively open GSC chromatin is further restricted by Polycomb repression of testis or somatic cell-expressed genes briefly active in early female germ cells. Subsequently, Neijre/CBP and Myc help upregulate growth and reprogram GSC metabolism by altering mitochondrial transmembrane transport, gluconeogenesis and other processes. In all these respects GSC differentiation resembles development of the totipotent zygote. We propose that the totipotent stem cell state was shaped by the need to resist transposon activity over evolutionary time scales.

Introduction

The animal zygote launches embryonic development from a totipotent chromatin state in a process that has been intensively studied using model invertebrate and vertebrate organisms (reviewed in Macrae et al. 2023; Zhou and Cho, 2022; Ahmad and Henikoff, 2022; Gleason and Chen, 2022). Pluripotent cell lines derived from embryos (Thomson et al. 1998) or reprogramed somatic cells (Takahashi and Yamanaka, 2006) provide technically favorable material for analyzing the gene expression, chromatin and metabolic states that drive stem cells to diverse differentiated derivatives (Rafalski et al. 2012; Kinoshita and Smith, 2018; Omole et al. 2018; Soldner and Jaenisch, 2018). Highly potent stem cells in vivo and in cell lines cycle in an unusual manner with very short G1 phases, express genes in a variable manner, and show "bivalent" chromatin marks

(Hsu et al. 2008; Pauklin and Vallier, 2013; ter Huurne and Stunnenberg, 2021). During subsequent development daughter cells drastically alter their chromatin, gene expression, organelles, and metabolism. These changes parallel reduced developmental potency but the evolutionary forces that gave rise to animal stem cells remain little understood.

Drosophila female germline stem cells provide an attractive system for studying totipotency maintenance and differentiation (reviews: Fuller and Spradling, 2007; Drummond-Barbosa, 2019; Hinnant et al 2020). New female germ cells continuously arise downstream from GSCs located within the adult germarium, and develop into ovarian follicles ordered by age within ovarioles (Figure 1A). GSC daughters initially divide synchronously to build 16-cell cysts that specify an oocyte and 15 nurse cells, traverse meiotic prophase, become associated with somatic follicle cells and leave the germarium as stage 2 (S2) follicles. Over the next 3 days, follicle development increases the oocyte volume to more than 10,000 times that of a GSC (Figure S1). Following fertilization of the mature S14 oocyte, the resulting totipotent zygote develops as a syncytial embryo before germ cells bud from the posterior pole, migrate to the gonad and complete the generational cycle by forming new GSCs (Fig. S2). Previous studies (reviewed in Matova and Cooley, 2001; Lu et al. 2017; Hinnant et al 2020; Spradling et al. 2022) show that aspects of oocyte development have been conserved between Drosophila, mouse and other animals, including germline cyst and nurse cell formation (Lei and Spradling, 2016; Niu and Spradling, 2022), nurse cell turnover (Lebo and McCall, 2019; Niu and Spradling, 2022) and Balbiani body production (Mahowald and Strassheim, 1970; Pepling et al. 2007).

Gene expression during oogenesis has been previously characterized in purified GSCs (Kai et al. 2005; DeLuca et al. 2020), downstream follicles (Sieber et al. 2016; Greenblatt et al 2019; DeLuca et al. 2020); and also in single ovarian cells (Rust et al. 2020; Jevitt et al. 2020; Slaidina et al. 2020; Slaidina et al. 2021) or nuclei (Li et al. 2022) . Early germ cells express conserved germline

genes (CGGs) (Fierro-Constain et al. 2017) encoding essential Piwi-piRNA pathway components that repress transposon transcription and promote transposon mRNA slicing and piRNA amplification in nuage (reviews: Huang et al. 2017; Czech et al. 2018; Wang and Lin, 2021). Many important steps in germ cell differentiation are controlled at the translational levels (Slaidina and Lehmann, 2014; Ma et al. 2017; Kong et al. 2019; Blatt et al. 2020; Mercer et al. 2021). For example, the GSC differentiation factor Bam stimulates exit from the stem cell state by repressing high level ribosome production in the nucleolus (Zhang et al. 2014; Sanchez et al. 2016; review: Breznak et al. 2023) and by lifting the Nanos-controlled translation block of differentiation genes (review: Mercer et al. 2021).

GSCs and developing follicles sense environmental nutrients using insulin/IGS and steroid signaling to maximize oocyte output (LaFever and Drummond-Barbosa, 2005; review: Drummond-Barbosa, 2019). An important regulator of resource adaptation is Myc, whose translation becomes derepressed as early as the 4-cell cyst (Harris et al. 2011) and increases further under favorable conditions starting in region 2b (Wang et al. 2019). As in other rapidly growing systems, Myc plays an essential role (Maines et al. 2004) by controlling hundreds of target genes (Wang et al. 2019) that regulate both ribosome production and metabolism.

Here we extend our knowledge of GSC chromatin modifications and gene expression. Upon leaving the GSC state, daughters form germline cysts and upregulate H3K9me3-rich heterochromatin on transposon sequences in centric chromosome regions. A cyst or syncytium likely protects against transposable element (TE) activity by sharing piRNAs, whereas heterochromatin formation blocks TE transcription. Polycomb complexes dependent on the H3K27 methylase Enhancer of zeste (E(z)) repress male germ cell and somatic cell genes that are transiently active in GSCs and early daughters. Metabolism is reprogrammed downstream from GSCs to enhance the production in mitochondria of biosynthetic precursors for follicle growth. Our results argue that

GSCs and many other highly potent stem cells modulate the mechanisms they use to control transposable elements as the initial step in their differentiation programs.

Results

Mapping changes in GSC chromatin and gene expression

Chromatin was characterized in purified germ cells or in follicles (DeLuca et al. 2020). GSCs, and nurse cells labeled using Germline-GAL4::UASz-tomato were purified from specific ploidy classes using fluorescence-activated cell sorting (FACS) (Figure 1B; Figure S3). GSC, nurse cell (NC) and follicle cell (FC) ploidy classes (4C, 8C, etc), and whole follicles were collected and analyzed by Chip-seq for H3K9me3, H3K27ac, H3K27me3, and H3K4me3. ATACseq (Buenrostro et al. 2015) was applied to interrogate changes in chromatin accessibility.

Gene expression (Table S1) was analyzed by RNAseq using three independent preparations of purified GSCs. To identify gene expression changes in older nurse cells, we prepared three samples of RNA from wild type or E(z) germline knockdown (GLKD) ovaries (Deluca et al. 2020) isolated less than 8 hours after eclosion. At this time ovaries lack follicles older than S6, and are weighted by mass toward stage 4-6 follicles (OVS6). We also analyzed three samples of adult ovaries from well fed females whose RNA derives mostly from large stage 9-14 follicles undergoing vitellogenesis and eggshell production (WTOV). Finally, we utilized gene expression data from three preparations of mature follicles ("S14") or from S14 oocytes after removing follicle cells by collagenase digestion and washing ("S14nfc") (Greenblatt et al. 2019). Processed data from these experiments is summarized in Tables S1-S5 as described (Methods; Data and software availability).

Female germ cells do not induce many new genes downstream from the GSC. In fact, the diversity of expressed genes (defined as genes with RNA> 1 TPM and SD/mean <0.67) is highest in GSCs and decreases thereafter, while inactive genes increase (Figure 1C). A parallel decline is seen in total ATAC peaks, most of which are found around gene promoters and potentially at enhancers (Figure 1D). Among transcripts expressed in mature S14 oocytes, 75/79 genetically defined maternal effect genes (95%) and 6,520/6,597 (99%) total transcripts are already expressed in GSCs, consistent

with the importance of translational regulation (Mercer et al. 2021). The RNA levels of about 50% of maternal effect genes are already within twofold of their final values at the GSC stage, and 80% increase less than four fold (Figure 1E).

Heterochromatin forms in centric chromosome regions and on transposons downstream from the GSC

Studies of H3K9me3 modification were carried out to look for changes in heterochromatin within early germ cells. GSCs appear to contain less H3K9me3-rich heterochromatin than germ cells within 16-cell cysts based on their lower immunofluorescence staining for the H3K9me3-binding protein HP1a (Rangan et al. 2011; Figure 2A). Nuclear H3K9me3 staining was visible in GSCs, but levels increase in region 2a and 2b germ cells (Figure 2B). Chip-seq experiments for H3K9me3 supported these and earlier observations (Clough et al. 2007; Yoon et al. 2008; Rangan et al 2011; Clough et al. 2014). H3K9me3 Chip-seq was quantitated in 250 kb bins after screening out ambiguously mapped reads. Figure 2C shows that minimal H3K9me3 enrichment relative to chromosome arms was detected in GSCs on chromosome 2 and 3 centric sequences. However, centric H3K9me3 modification increased in 4C NC (nurse cells and oocytes) and considerably more in 8C NC, which derive from nurse cells in stage 2-3 follicles. Rising H3K9me3 levels were also found on transposable element (TE) sequences (Figure 2D) which are enriched in centric regions, but studies of satellite DNAs were inconclusive. Thus, germ cells begin to increase their centric heterochromatin (HET) as 16-cell cysts form, and H3K9me3 levels rise further during meiotic prophase and early follicular development.

To investigate whether increased H3K9me3 correlates with gene repression, we used the system of individual hsGFP expression reporters that were previously employed to document the onset of Polycomb domain repression in follicular nurse cells (DeLuca et al. 2020). These lines typically do not express GFP in the absence of heat shock (Figure 2E, " -heat shock") but after a

brief standard heat shock, reporters located in euchromatin are readily induced at most germ cell stages (Figure 2E, "+ heat shock"). In contrast, most insertions located in centric regions were only inducible in GSCs and early cysts, but not in later germ cells (Figure 2E, F). This provides evidence that increasing H3K9me3-rich heterochromatin represses gene expression beginning around the time of meiotic onset. When the H3K9 methylase encoded by *eggless (egg)/Setdb1* (Clough et al. 2007; Yoon et al. 2008; Rangan et al 2011; Clough et al. 2014), was knocked down, centromeric reporters could still be expressed in meiotic germ and nurse cells (Figure 2F). Reporter gene repression also depended on Piwi. The relative expression of hsGFP in *piwi*[GLKD] compared to control ovaries depended on reporter location and developmental time (Figure 2G). Reporters located on chromosome arms (Figure 2G, pink) were unaffected by *piwi*[GLKD), whereas only after meiotic onset, those positioned in centric regions (Figure 2G, blue) were upregulated in *piwi*[GLKD] compared to control (Figure 2G, green), consistent with Chip-seq data (Figure 2C-D).

The fact that heterochromatin begins to form downstream from the GSC in germline cysts and repress gene transcription extends a parallel previously noted between stem cell differentiation downstream from the GSC during oogenesis and differentiation downstream from the zygote during early embryogenesis (DeLuca et al. 2020). Both take place in a syncytial environment generated either by germline cyst formation or by synchronous embryonic cleavage division, and both include upregulated heterochromatin formation (Lu et al. 1998; Gu and Elgin, 2013; Czech et al. 2018; Fabry et al. 2021). Figure 2H summarizes the state of germ cell chromatin during the maternal and zygotic portions of the generational cycle. Highly open chromatin characteristic of the GSC state (pink) lasts only briefly before heterochromatin increases (blue). Repression in the oocyte likely becomes widespread as its nucleus condenses to form a karyosome beginning in S3, before meiotic progression resumes about S8. Zygote formation re-establishes stem cell chromatin, which persists during syncytial embryonic cleavage divisions 1-9 and in newly budded pole cells. Only then do

cleavage divisions begin to slow, and transposon-directed, Piwi/piRNA-dependent heterochromatin (blue) forms during divisions 10-14 (Fabry et al. 2021; Wei et al. 2021; Zenk et al. 2021). The cellularized embryo goes on to produce diverse somatic cell types using Polycomb repression (green). In a further parallel, follicular ovarian nurse cells of the GSC lineage repress Polycomb domains and differentiate as somatic cells (DeLuca et al. 2020). During embryogenesis the PGCs migrate to the gonads, proliferate during larval stages and complete the cycle by forming new GSCs at the onset of pupal development (Bhaskar et al. 2022; Grmai et al. 2022; Figure S2).

E(z)-dependent repression contributes to germline- and sex-differential gene expression

We investigated whether some of the more than 900 genes that turn off between the GSC and early follicle stages are repressed by Polycomb. Previous studies identified differences in PRC1 and/or PRC2 (Figure 3A) downstream from the GSC (DeLuca et al. 2020). GSCs contain high amounts of the alternative PRC2 component Pcl, but relatively low levels of Scm as well as canonical PRC1 and PRC2 subunits. As a result, while H3K27me3 modifications are present on GSC chromatin, site selection by Pcl-dominated PRC2 is stochastic and common foci are largely lacking across all GSCs. After follicles form, Pcl levels fall, Scm levels increase, and more efficient targeting allows canonical Polycomb domains to become repressed in follicular nurse cells (DeLuca et al. 2020). We investigated whether individual genes also undergo PRC2-mediated repression as GSC daughters differentiate.

Genes repressed by E(z) were identified by analyzing GSC-expressed genes that are downregulated in stage 4-6 follicles, but whose RNA persists at higher levels in *E(z)*[GLKD] animals. 423 candidate genes showed either 1) >3-fold increased expression in *E(z)*[GLKD] compared to wild type OVS6, or 2) gene expression after the GSC was downregulated at least 20% less in *E(z)*[GLKD] at S6 compared to OVS6 (Table S2). Normally, 78% of these genes downregulate more than 10-fold in OVS6 compared to GSC; most no longer express at OVS6 (88%)

or in the mature oocyte (98%). Thus, genes regulated in whole or in part through the direct or indirect action of E(z), make up a substantial fraction (40-50%) of the genes that shut off in germ cells between the GSC and early follicular stages.

Some E(z)-downregulated genes function within the GSC itself, and/or in early germ cell development. For example *bgcn* (Figure 3B), facilitates the switch between GSC and cystoblast (Mercer et al. 2021). *Bgcn* chromatin shows strong promoter region peaks of ATAC in GSCs and early nurse cells, and high promoter H3K27ac in GSCs. In 32C and later nurse cells H3K27ac has been lost, and H3K27me3 has risen from low GSC levels (GSC: 0.057; 32C NC: 0.584; 512C NC: 0.479; Table S1), consistent with a direct E(z)-dependent shutoff. Despite a tenfold increase in *bgcn* RNA levels following *E(z)*[GLKD], S6 expression remains at only 15% of GSC levels, suggesting that *bgcn* is controlled by other pathways as well. Other E(z)-downregulated genes that may function in GSCs encode proteins supporting gene transcription (HP1Dcsh, HP1Lcsd, TfIIA-S-2), or translation (RpL22-like, RpL37b, RpS10a, RpS19b, eIF3f2, eIF4E3, eIF4E7, EF1a2, Rcd-1r) (see Kong et al. 2019; Hopes et al. 2022).

Analyzing the tissue expression of E(z)-downregulated genes using Fly Atlas (Leader et al. 2018) showed that 49% are predominantly expressed in the testis or male accessory gland (Figure 3C), including most of the general gene expression isoforms discussed above. Many of the Polycomb-repressed testis-expressed genes such as *CG43088*, *CG15599*, and *CG43843* present little evidence of ovarian function and some may be recently evolved (Su et al. 2021). *CG43088* is related to human harbinger-transpose-derived-1, and shows promoter-associated peaks of ATAC and H3K27ac only in GSCs (Figure 3D). The *CG43088* gene body contains significant levels of H3K27me3 in GSCs (0.976) that change little. In *E(z)*[GLKD] early follicles *CG43088* becomes 3-fold overexpressed relative to GSC levels (127+/-10.2 vs 40+/-9.46 tpm) and 400-fold overexpressed relative to early wild type follicles (0.28 +/- 0.080 tpm).

A much larger group of the testis-expressed genes that are only transiently expressed in early female germ cells are known or likely to have a testis function. *escargot*, encoding a snail class transcription factor needed to maintain male (but not female) GSCs, retains an open promoter in early follicles, but is downregulated (GSC: 91+/- 0.11 tpm; OVS6: 0.070+/- 0.020 tpm), loses H3K27ac and acquires H3K27me3 at later stages (GSC: 1.54; 32C NC: 16.7; 512C NC: 13.7) (Figure 3E). *Chinmo* (Figure 3F) encodes a major transcription factor maintaining sexual identity in testis (Flaherty et al. 2010; Grmai et al. 2018). It undergoes E(z)-dependent shutdown (DeLuca et al. 2020) and accumulates H3K27me3 (GSC: 0.753; 32C NC: 11.2; 512C NC: 11.3). *Always early (aly)* is essential for the spermatocyte gene program, has no known female function, and is likewise repressed in early female germ cells supported by E(z) (H3K27me3 GSC: 0.387; 32C NC: 1.25; 512C NC: 0.901) (Figure 3G). *Tdrd5l* (Figure 3H, a gene implicated in germline male sex regulation (Primus et al. 2019), is also strongly expressed in female GSCs (82.4+/-7.02 tpm), and is downregulated before OVS6 (1.72+/0.528 tpm) dependent on E(z) and associated with increasing H3K27me3 (GSC: 0.631; 32C NC: 2.87; 512C NC: 2.30). In addition to controlling many highly testis-enriched genes, E(z) also downregulates other genes with sexually dimorphic expression, including genes preferentially expressed in male accessory gland. However, only about 25% of the male-biased genes affected by E(z) are primarily repressed by Polycomb (i.e. >50%), or show increases in gene body H3K27me3 levels as high ass those of Polycomb domain targets, suggesting that the majority of the E(z)-dependent gene repression is indirect.

Some of the "testis" genes expressed in GSCs function in early female germ cells. *ord, corolla,* and *vilya,* which function in meiosis, are all expressed in GSCs and downregulated in young follicles, although the contribution of E(z) to gene downregulation varies: *vilya* (38%), *ord* (20%), and *corolla* (6.6%). *Phf7* (Figure 3I) is expressed in GSCs (75.1+/-9.00), and at higher levels in OVS6 in *E(z)*[GLKD] (62.8+/7.15) compared to control (32.4+/3.51). However, *Phf7* continues to

be expressed throughout oogenesis (S14nfc: 99.3+/-5.36). Thus, E(z) represses a variety of genes downstream from GSCs in female germ cells (Figure 3J-K), where it contributes to normal sex- or tissue-specific gene regulation (Figure 3J-K).

Maintenance of the GSC state

High levels of gene expression are often associated with promoter H3K27 acetylation (H3K27ac) catalyzed by CREB-binding protein (CBP), a conserved histone acetyl transferase (Holmqvist and Mannervick, 2013; reviewed in Lerhmann et al., 2002). We selectively knocked down the Drosophila gene encoding CBP, *nejire (nej)*, in germ cells using specific GAL4 drivers (DeLuca et al. 2020). While germline knockdown of the *white* gene in control ovaries had no effect, knockdown of *nej* in germ cells greatly reduced H3K27ac staining, and arrested follicle growth and development beyond stage 1 (Figure 4A).

The chromatin state of several genes highly expressed in the GSC including the GSC differentiation gene *bgcn* (Figure 3B) and the core germline gene *piwi* (Figure 4B) show strong ATAC peaks in the vicinity of their promoters, high levels of transcription start site (TSS)-associated H3K27ac, and low GSC levels of repressive H3K27me3. To see if this is generally true of GSC genes, we plotted the entire GSC transcriptome as a function of mRNA abundance and H3K27ac (Figure 4C). Highly expressed genes usually had high levels of H3K27ac (and low levels of H3K27me3), while genes expressed at low levels did not. H3K27ac peaks in GSCs were almost always (91%) associated with transcription start sites (TSS; Figure 4C, 4D). Most of these genes also had ATAC peaks near the TSS that accounted for 68% of total ATAC peaks; the remaining ATAC peaks were located in other gene regions, and indicate the location of candidate enhancers. ATAC peaks likely reflect the effects of transcription factor binding at enhancers and promoters, as well as general transcription and splicing factors.

To investigate the GSC transcription program, we identified 101 genes with the highest H3K27ac levels in GSCs (Table S3). These genes cluster in autosomal pericentric regions, as well as in two euchromatic regions, 31B-F on chromosome 2L, and in region 13F-14C on the X chromosome (Figure S4). They encode multiple regulators of germ cell development, the cell cycle (including cdk1 and CycD), translation, protein stability, and signaling. Thus, a significant number of genes whose expression is stimulated in GSCs by H3K27 acetylation reside in relatively few clusters, primarily in centric chromosome regions.

Young nurse cells metabolically transition

Cells that develop from highly potent stem cells often alter their metabolic state (Rafalski et al. 2012; Hayashi et al. 2017; Hayashi and Matsui, 2022). Downstream from GSCs in young follicles, we identified more than 600 genes that are increased at least three fold (Table S4). H3K27ac is elevated in nurse cell nuclei from young follicles (Figure 4F) and at the promoters of many upregulated genes (Figure 4G, pink). Conversely, the expression of genes with reduced acetylation in nurse cells (green) decreased or changed little. Genes acetylated at similar levels in both cell types (black) rarely changed expression.

Consistent with previous work (Zhang et al. 2019), Myc is among the transcription factors found to increase sharply in growing follicles (Figure 5A, OVS6/GSC: 5.1-fold). FoxO, the mediator of insulin signaling, is also much higher than in GSCs (7.6-fold). Mitochondrial oxidative phosphorylation is substantially increased beginning in region 2b (Zhang et al. 2019). Previous genetic studies also found that ATP synthase function during oogenesis is largely independent of ATP production (Teixeira et al. 2015).

We looked for additional metabolic changes by analyzing the genes in Table S4 using the gene ontology program Metascape (Zhou et al. 2019; Figure 5B). All the top GO categories suggest that nurse cell mitochondrial metabolism undergoes reprograming in young nurse cells to upregulate

production of precursor molecules including amino acids, nucleotides, carbohydrates and lipids to support high levels of growth, changes characteristic of aerobic glycolysis. Consistent with upregulation of fatty acid production we observed increases in acetyl CoA carboxylase mRNA (*ACC*, 3.1-fold) which catalyzes production of malonyl-CoA, the rate-limiting substrate for fatty acid biosynthesis. Nurse cell gene expression shows other indications of increased lipid metabolism (see Figure 5B, "metabolism of lipids"). *CG3902* encoding an Acyl-CoA dehydrogenase, the first step in mitochondrial fatty acid beta-oxidation is upregulated more than 400-fold. RNA levels increase for both *Lpr1* (12-fold) and *Lpr2* (87-fold). *Hnf4*, encoding a nuclear hormone receptor that senses fatty acid levels, increases 4.9-fold. In mammals, Hnf4 is considered a master regulator of the liver, where it induces genes (including *FoxO*) involved in lipid metabolism, gluconeogenesis, and amino acid metabolism (Sieber and Spradling, 2017), processes that also appear to be upregulated in nurse cells.

Carbohydrate metabolism is also likely to be strongly increased in young nurse cells. Sugar transporters are sometime direct Myc targets, and in young NCs *Glut1* sugar transporter expression increases 26-fold, as does *sut1* (5.0-fold). Glycolysis genes often upregulated in Myc-expressing cells are also increased (Figure 5C). These include genes encoding Hexokinase (*Hex-A*, 5.1-fold), Triose phosphate isomerase (*Tpi*, 10.5 fold), and Phosphoglyceromutase (*Pglym78*, 14-fold). Myc may also increase glutamine uptake and metabolism, which synergizes with increased glycolysis to maintain TCA cycle flux and stimulate lipid metabolism. *Ldh* mRNA production in young nurse cells is virtually shut down (*Ldh*, 0.011-fold), suggesting that glycolysis is not needed to generate ATP and that lactate is not being generated from pyruvate in large amounts. This would be consistent with upregulation of gluconeogenesis for carbohydrate production (Figure 5B). In addition, increased expression of Glyceraldehyde-3-phosphate dehydrogenases (*Gapdh1*, 2.5-fold) and (*Gapdh2*, 1.9-fold) may provide an alternative way to prevent excess NADH generated by

glycolysis from accumulating (Li et al. 2019). During larval development the metabolic switch to aerobic glycolysis is mediated by the "estrogen related receptor" encoded by *ERR* (Tennessen et al. 2011), but in young nurse cells *ERR* levels were unchanged (0.96) relative to GSCs.

A closer examination of 20 of the more than 40 transmembrane transporters that were upregulated at least tenfold (Table S5) is shown in Figure 5D. Each of these genes has a strong mammalian ortholog. For example, *Sfxn1-3* (upregulated 15-fold) whose mouse ortholog is *Sfxn1*, encodes the mitochondrial serine transporter located in the mitochondrial inner membrane. Serine import by Sfxn1 supplies the sole substrate for one-carbon metabolism, which generates S-adenosyl methionine (using *Sam-S*, upregulated 28-fold), glutathione, and NADPH that are needed for methylation, nucleotide biosynthesis, gluconeogenesis and many other processes (Kory et al. 2018). The *Dic1/Slc25a10* transporter (23-fold increased) encodes another inner membrane protein. It catalyzes exchanges of dicarboxylates for inorganic phosphate and sulfur-containing compounds in support of gluconeogenesis. The changes in substrate import alone (Figure 5D) reveal how extensively young nurse cells remodel metabolism downstream from the GSC in support of rapid germ cell growth.

Discussion

<u>Drosophila female GSCs: a model in vivo totipotent stem cell</u>

Several new insights into highly potent stem cells resulted from these studies using Drosophila ovarian GSCs as a model (Table S1). 1) Polycomb contributes to repressing a class of genes expressed in GSC daughters for no apparent reason; 2) genes encoding many proteins needed and expressed by stem cells are clustered in centric heterochromatic domains and highly modified with H3K27ac in GSCs; 3) a novel mechanism of transposon control used by stem cells is implicated by the initial focus of their daughter cells on syncytial interconnection and H3K9me3-mediated repression. 4) animal life cycles require two connected totipotent differentiation processes, rather

than just one. The ability to study totipotent GSCs and genetically manipulate their operating instructions will further advance our understanding of these topics. The resources presented here (Tables S1-S5) expand genomic data on GSCs and extend developmental studies of their daughters that are highly annotated (Flybase: Gramates et al. 2022).

E(z)-dependent repression downstream from the female GSC modulates sex- and tissue-specific gene expression

GSCs transiently express more than 400 diverse genes whose downregulation depends on E(z), including ~50% that are primarily found in male germ cells (Table S2). The functional significance of their Polycomb-mediated repression, and the pathways by which many targets are indirectly controlled, remain to be determined. Polycomb repression is essential for follicular nurse cell differentiation, for follicle development beyond about S5 and for female fertility (DeLuca et al. 2020). Early repression of the genes in Table S2 may also be important without arresting oocyte development before S5.

The Polycomb-associated downregulation of genes in early Drosophila female germ cells may be analogous to Polycomb regulation during mammalian PGC development (review: Loh and Veenstra, 2023). Following induction in the proximal epiblast at about E7.0, mouse PGCs migrate to the gonad which they reach between E10.5 and E11.5. During migration, the PGCs significantly reduce their DNA methylation and repressive H3K9me2, but upregulate H3K27me3 (Seki et al. 2005). The timing of sex-specific PGC differentiation depends on the PRC2 component EED, which may interact with Dnmt1 to regulate this process (Lowe et al. 2022). During part of this period they arrest in G2 and modify RNA polymerase tail phosphorylation to repress transcription (Seki et al. 2007), changes that also take place in migrating Drosophila PGCs (Su et al.1998; Deshpande et al. 2004; Martinho et al. 2004). These events allow mouse PGCs to reacquire totipotency, reactivate conserved germ cell genes such as *Ddx4*, and to silence gene expression resulting from somatic

signals received prior to their reprogramming as germ cells (Hill et al. 2018). Polycomb dependent silencing in Drosophila may likewise "clean up" extraneous transcription to compensate for stochastic or incomplete programming at earlier stages.

Our finding that GSCs and early female germ cells normally express many male germ cell genes calls into question elevated male gene expression as a criterion for identifying germline sex determination genes (Bhaskar et al. 2022; Grimai et al. 2022). Mutants that increase testis-related gene expression and also arrest female germ cell development or generate GSC-like tumors may not determine germline sex, because elevated male gene expression is normal in female germ cells at early stages. For example, over-expression of the proposed germline sex determination regulator Phf7 (Yang et al. 2012) in female germ cells disrupts their differentiation and induces expression of several hundred testis genes (Shapiro-Kulnane et al. 2015; Smolko et al. 2022). We find that these genes extensively overlap the testis genes normally downregulated by E(z). Mutations of *stonewall (stwl)* encoding a heterochromatin-interacting factor required to maintain GSCs (Clark and McKearin, 1996), also disrupt differentiation, alter germline gene expression and overproduce Phf7 (Zinshteyn and Barbash, 2022). The testis genes upregulated by *stwl* mutation strongly overlap those found in normal female GSCs, which itself "overproduces" *Phf7* transcripts.

The SKI complex is required for follicle development beyond early follicle stages and degrades some early female germ cell RNAs (Blatt et al. 2021). We find that about 10% of SKI-complex target gene overlap with the E(z)-repressed genes, including *RpL22-like*, and *dany*.. However, most SKI-complex target genes are not even expressed in GSCs (tpm <1) and the reason for this small overlap is unknown.

<u>A connection between totipotent stem cells and genes in heterochromatic domains.</u>

GSCs express many important genes residing in centric domains that display high levels of promoter-associated H3K27 acetylation, suggesting they may share some regulation in common

(Table S3). *ovaries absent (ova)* is needed for heterochromatin formation and transposon repression in GSC daughters (Yang et al. 2019). *vasa* controls mRNA translation in female germ cells and represses TEs in nuage. *me31B* is required to form germ cell P granules (Nakamura et al. 2001), while insulin pathway genes *chico*, *Pten* and *Reptor-BP* couple GSC activity to nutrients. *daughterless* is maternally required for sex determination (Cronmiller and Cline, 1987). One of the two "euchromatic" zones housing genes in this group, located in region 31, is connected to heterochromatin. This region forms the unusual "gooseneck" region of polytene chromosomes and binds high level of HP1a (James et al. 1989; de Wit et al. 2007). Many conserved genes in this domain are located in heterochromatin in *D. virilis* and *D. pseudobscura* (Caizzi et al. 2016). This chromatin context in cells with relatively low H3K9me3 levels may be advantageous for genes with critical functions in stem cells, and may facilitate access by CBP and transcription factors. Some germline-specific genes located in H3K9me-rich chromatin in *C. elegans* are also distinctively regulated (Methot et al. 2021; Padeken et al. 2022).

Formation of germline cysts and heterochromatin downstream from totipotent stem cells suppresses transposon activity

The formation of syncytial germline cysts downstream from GSCs or PGCs in the male and female gonads of diverse species helps protect gametes from meiotic drive and allows cyst cytoskeletal polarity to specify which female cyst cells will be oocytes and which nurse cells (review: Spradling et al. 2022). However, the syncytial environment in the germline cyst, whose cells express high levels of piwi/piRNA pathway components (Figure 6B), can also help control TE activity. The presence of a cyst or syncytium would allow piRNAs induced by TE activity in one cyst cell to be shared with the other cyst cells (Figure 6A), enhancing resistance to that TE in the other cells before transposition there can occur (reviews: Huang et al. 2017; Czech et al. 2018; Wang

and Lin, 2021). The high similarity of sister cells recently derived from a common precursor likely enhances the efficacy of such sharing.

Activation of the nuclear arm of the piRNA pathway contributes to the H3K9me3-modified chromatin that increases in GSC daughters as they prepare and enter meiosis. These increases utilize *piwi* and *eggless* and target transposon-related as well as uniquely mapping sequences (Figure 2F, 2G; Clough et al. 2007; Clough et al. 2014; Ninova et al. 2020; Chen and Aravin, 2023). Very similar heterochromatin formation takes place during the last few syncytial cleavages of the Drosophila embryo downstream from the zygote (Lu et al. 1998; Gu and Elgin, 2013). Heterochromatin formation during early embryogenesis in *Drosophila miranda* also involves piRNA-directed H3K9me3 deposition that may start even earlier in the cleaving pre-blastoderm embryo (Wei et al., 2021). These embryos remain syncytial and likely share piRNAs between preblastoderm nuclei, prior to cellularization. Similar pathways of heterochromatin formation act in diverse organisms (Liu et al. 2014; Onishi et al. 2021).

The formation of interconnected germline cysts or syncytia just downstream from highly potent stem cells has been extensively conserved in male and female germ cells throughout animal phylogeny (review: Spradling, 2023). These events can be traced back to the highly potent stem cells (i.e. interstitial stem cells (ISCs), neoblasts) found in basal animals including hydra, planaria, parasitic flatworms and other organisms (Zheng et al. 2018; Issigonis and Newmark, 2019; Reddien, 2022; Littlefield, 1994). Stem cells in basal eukaryotes express largely the same set of CGGs present in bilaterian animal germ cells (Fierro-Constain et al 2017), and frequently also form cysts downstream from highly potent germline and somatic stem cells. Notably, cysts similar to those produced by Drosophila GSC daughters not only arise downstream from Hydra GSCs (Littlefield, 1994), but immediately downstream from somatic interstitial stem cell (ISC) daughters which generate the cnidocytes required for feeding (David, 2012). The intercellular bridges (IBs) in

cnidoblast cysts resemble germline IBs and were observed to share cytoplasm between cyst members (Fawcett, et al. 1959). Hydra in ISCs, GSCs and some downstream cells express *piwi* genes carrying piRNAs that mostly target transposons (Kim et al. 2014), suggesting that cnidoblast cysts also help control TE activity within and downstream from highly potent somatic and germline stem cells.

GSCs and highly potent stem cells contain an enhanced system of transposon resistance

The timing of cyst and heterochromatin formation just after the stem cell state is surprising. Potent stem cells and GSCs are collectively immortal and responsible for protecting a species' genome from TEs over developmental time, spanning potentially millions of generations. Animal genomes contain diverse TEs, that in many groups now constitute a substantial fraction of the total (review: Wells and Feschotte, 2020). Stem cells and germ cells would seem to require the highest levels of TE defense consistent with their high expression of CGG genes. Yet female Drosophila GSCs display relatively open chromatin (DeLuca et al. 2020) and like the Drosophila zygote only form a syncytial environment and upregulate heterochromatin in downstream cells. The timing of these events argues that stem cells use additional mechanisms in addition to the cytoplasmic *piwi* pathway, for transposon control, that render H3K9me3-mediated heterochromatin formation unnecessary. However, this stem cell repression mechanism must fade quickly in daughter cells, necessitating them to increase alternative transposon defense mechanisms.

Does evidence exist that transposition is relatively low in totipotent stem cells, but increases in early derivative cells ? Cluster analysis, in theory, can measure when a TE moves during germ cell development. In the simplified lineage (with 4-cell cysts and 1 meiotic division) shown in Figure 6C, the TE1 (blue) insertion at the 2-cell cyst stage will generate 4 sperm with the same genetic change, whereas the TE2 (orange) insertion at the 4-cell stage will induce two identical mutant sperm, and the TE3 (insertion) at the sperm stage will result in one unique insertion. In practice, the

low rate of transposon movement in animals during normal development makes such studies difficult.

A Drosophila genetic screen (Karpen and Spradling, 1992) generated single P element transposition events in the progeny of individual males that were subsequently analyzed to determine cluster size. 100-200 progeny were analyzed from each of ~15,000 individual males. About 6,800 had 0 transposon-containing progeny, 6,173 had 1 offspring with a jump, 1,024 had two such progeny, 17 had 3 and 2 had 4. Progeny with transpositions were then analyzed by enhancer trapping and limited insertion sequencing, indicating that overall about 7,909 transpositions occurred singly and that 367 occurred as doublets. From the 19 males with 3 or 4 transposed offspring, only 2 contained a single sibling with one identical transposition; thus no clusters larger than 2 were observed. If P elements had jumped during preadult germ cell development or at the stem cell stage, then each mutant stem cell (from 10-20 total) would have generated clusters amounting to 5-10% of progeny, i.e. 5-20 offspring with identical insertions. The fact that only an estimated 25-50% of sperm (randomly) survive in recipient females to fertilize an egg, further advantages the detection of GSC insertions. These results argue that P elements transpose much less frequently at or before the stem cell stage, but show that transposition has begun before the end of cyst formation.

Consistent with a loss of effective stem cell repression, transposon activity increases downstream from the zygote. TE transcription is high in early Drosophila embryos (Scherer et al. 1982) and elevated TE expression and activity are widely observed in early mouse and human embryos (Xiang and Liang, 2021; reviewed in Friedli and Trono, 2020). Particular classes of retroelements drive expression in temporal and tissue-specific patterns (Göke et al. 2015). Programmed changes in TE expression patterns can accurately indicate the developmental state of stem cell lines that capture stages downstream from full pluripotency (Theunissen et al. 2016). TE-dependent transcription in early embryonic cells may even be functionally important (Percharde et

al. 2018). An association of intrinsic TE resistance with high potency states might also explain why DNA methylation levels can be reduced, as during the reprogramming of PGCs back to a pluripotent state to generate gonocytes (Hill et al. 2018).

Altered cell cycle and DNA replication behavior may enhance stem cell transposon resistance

The most likely sources of intrinsic TE resistance in potent stem cells are their unusual cell cycles. Such stem cells and their immediate progeny in gonads, embryos and cell lines frequently have very short G1 and long G2 phases (Hsu et al. 2008; Sheng and Matunis, 2011; Pauklin and Vallier, 2013; ter Huurne and Stunnenberg, 2021; Ables and Drummond-Barbosa 2013; Sellers et al. 2019; Vogg et al. 2021). DNA replication during S phase is also altered, due to changes in the number and order of origin activation (Hyrien and Mechali, 1995; Ferreira and Carmo-Fonseca, 1997; Seller and O'Farrell, 2018; Macheret and Halazonetis, 2018; Rausch et al. 2020; Klein et al. 2021), replication elongation rate (Nakatani et al. 2021) or because of other changes in DNA replication and repair (Wooten et al. 2017; Sjakste and Riekstiņa, 2021). These differences may affect TE movement because many transposons are sensitive to cellular replication. P elements, for example, insert preferentially near replication origins (Spradling et al. 2011; Cao et al. 2023) and this may allow them to transpose after duplication to a new unfired origin, often in a late replicating region (Figure 6D). Altering replication timing to eliminate late S phase might disrupt this pathway by eliminating available targets before they can be used (Figure 6E). Thus, highly potent stem cells may alter cell cycle and replication properties to enhance their ability to suppress TE movement.

Differences between highly potent stem cells and later more differentiated derivatives frequently involve the loss of late replication. Early in Drosophila cleavage there is no late S phase, but after cycle 9 derepression of Rap1-binding factor (Rip1) allows Rip1 to bind to and repress satellite DNA duplication, prolonging S phase (Seller and O'Farrell, 2018). Pericentromeric heterochromatin replicates in early to mid S phase in mouse ES cells, but replicates late upon

differentiation, a change that is associated with reduced acetylation of H3K9 and H3K8 in these regions (Rausch et al. 2020). The reason DNA replication is temporally regulated is not understood although it is proposed to facilitate chromatin epigenetic regulation (see Klein et al. 2021). Why differentiated cells acquire a distinct late S phase also remains a mystery, despite its apparent universality. Finally, it is not at all clear that avoiding late replication would be sufficient to reduce TE movement enough to maintain genomes over evolutionary timescales. However, the striking organization of eukaryotic genomes in two compartments, one of which replications late in S phase and houses most genomic transposons, reminds us that there is much about replication we do not understand and that these questions deserve greatly increased investigation.

Several other properties of highly potent stem cells may contribute to TE resistance. Despite a relatively open chromatin state, highly potent cells like Drosophila female GSCs might exclude expression of transcription factors that recognize and activate TE promoters. The special isoforms of general transcription, splicing and translation machinery expressed by GSCs (Kai et al. 2005; Shigenobu et al. 2006) might interfere with TE gene expression. Except for RpS5B (Kong et al. 2019), production of most of these isoforms shuts off downstream from female GSCs (Figure 3K, Table S2).

Animal generational cycles involve sequential germline and zygotic programs of stem cell differentiation

The process of transitioning from a stem cell to a more differentiated state represents a fundamental unit of development. Our studies emphasize that two such stem cell differentiation units are required each generational cycle, rather than one as often visualized (Figure 2H). Maternal and zygotic stem cell development act over parts of two animal generations (zygote to zygote) to complete a single generational cycle (GSC to GSC). While each unit involves common regulatory events, including exit from a totipotent stem cell state, chromatin remodeling, growth and

reprogramming of metabolism, there are many important differences between the maternal and zygotic cycles that make both essential. The embryo cannot overcome or replace all the events that took place in its mothers germline.

Restoring potential immortality likely represents the major focus of the maternal stem cell program. Rejuvenation occupies most of oogenesis, but continues during zygotic development in the embryo. Perturbations in metabolism during oogenesis can have lasting effects on the zygote (Hocaoglu et al. 2021). Maternal Piwi is essential for normal embryonic germ cell development and female fertility (Gonzalez et al. (2021). Polycomb function during oogenesis is needed for normal mouse embryonic development (Posfai et al. 2012). A better understanding of the processes that link maternal and zygotic pluripotent stem cell differentiation promises to illuminate the full range of mechanisms that sustain animal species and to further improve the cell therapies that depend on such an understanding.

Materials and Methods

Drosophila strains

Mutant Alleles	Source	Stock Number		Citation	Other Info
bamΔ86	Bloomington Drosophila Stock Center: BDSC NIH P40OD018537	5427		Flybase (Gramates 2022)	
yl^{13} (K621)	BDSC	4320		Flybase (Gramates 2022)	yl^{13} v^{24}/FM3
yl^{15} (11-380)	BDSC	4621		Flybase (Gramates 2022)	y^1 cv^1 v^1 yl^{15} f^1/FM0
Cap-$H2^{TH1}$	BDSC	2608		Flybase (Gramates 2022)	Df(3L)W10, ru^1 h^1 Cap-$H2^{TH1}$ Sb^{sbd-2}/TM6B, Tb^1
Cap-$H2^{TH2}$	BDSC	12126		Flybase (Gramates 2022)	y^1 w^*; P{w[+mC]=lacW}l(3)j9A5^{j9A5} Cap-$H2^{TH2}$/TM6B, Tb^1
RNAi Lines	Source	Stock Number	Insertion Site	Vector	Other info
UASp-E(z)^{RNAi}	BDSC	36068	attp40	Valium22	TRiP.GL00486
UASp-nej^{RNAi}	???				
Gal4 Drivers	Source	Stock Number	Insertion site	Citation	Genotype
Nos-Gal4	BDSC	4937	3R:10407270	Flybase (Gramates 2022)	P{GAL4::VP16- nos.UTR}
MTD-Gal4	BDSC	31777	X, 2, 3R:10,407,270	Flybase (Gramates 2022)	P{otuGAL4::VP16}; P{GAL4-nos.NGT}; P{GAL4::VP16nos.UTR}
Fluorescent Lines	Source	Stock Number	Insertion Site	Citation	Genotype
UASz-tdTomato	S. DeLuca		VK33	DeLuca, et al. 2020	

Primary antibody Table

Epitope	Species	Source/Reference	Concentration for IF
Hts (1B1)	mouse	Developmental Studies (Zaccai and Lipshitz, 1996)	1:50

H3K27ac	rabbit	Active Motif (#39135) Lot 17513002	1:1000
vasa	rabbit	Lab generated (Final bleed FB87)	1:1000

Immunofluorescent staining and microscopy

Flies are fed on wet yeast for 3 days before dissection. Dissected ovaries are fixed in 4% PFA and 0.01% PBST for 30 min, blocked with 5% Normal Goat Serum (NGS) before overnight incubation in primary antibodies. Please see Table Primary antibodies used for antibody sources and their appropriate dilution. Secondary antibodies are incubated overnight at 1:1000 dilution. DNA is stained with DAPI (0.5 μg/ml) and the ovaries are mounted in 50% glycerol for imaging. Images are acquired using Leica TCS SP5 with 63X oil objective.

Volume measurement

MTD-Gal4 > UASz-tdTomato flies with labeled germline cells are fed on wet yeast for 3 days. Ovaries are dissected carefully with forceps to break apart the surrounding muscle sheets. Dissected ovaries are fixed in 4% PFA for 10 min, stained with DAPI, and mounted in VECTASHIELD®. For yl mutant and control lines, dissected ovaries are stained with anti-vasa to mark germline cell volumes (see immunofluorescent staining method for details). Full z-stacked images of the ovary are taken under 40X oil objective with Leica TCS SP5 Confocal Microscope. Reconstruction of 3D images is performed with Oxford Instruments Imaris software. Volumes and total DAPI signal pixels of individual follicles are recorded and developmental stages of the follicles are manually noted. Growth rate changes are plotted and modeled with R.

Growth modeling

Local polynomial regression was used to accommodate for the complexity of the data. Smoothing factor =0.75 and degree of polynomial =2 were used. The data set is fairly dense with n>20 for each stage and is therefore suitable for local polynomial regression. However, the result of LOESS can be difficult to interpret because it cannot be written down as simple equations (Guthrie, 2002). To further test the developmental growth rate changes, multiple linear regression was applied with two conditional terms. Condition1 distinguishes follicles before (stage 1 to 5) and after (stage 6 to 10) nurse cell chromatin dispersion. Condition2 distinguishes follicles before (stage 1 to 7) and after (stage 8 to 10) onset of vitellogenesis. The model to be tested is as following:

log10 Volume = a0*log2DNA +a1*log2DNA*condition1 +a2*log2DNA*condition2
+b0+b1*condition 1 +b2*condition2

The fitting result is as following:

```
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept) (b0)     3.47678  0.03855 90.186 < 2e-16 ***
## log2norm.DAPI (a0)   0.23343  0.01152 20.263 < 2e-16 ***
## condition1 (b1)     -0.13196  0.08155 -1.618 0.106569
## condition2 (b2)      1.45582  0.33120  4.396 1.49e-05 ***
## log2norm.DAPI:condition1 (a1) 0.05044  0.01544  3.266 0.001204 **
## log2norm.DAPI:condition2 (a2) -0.13115  0.03544 -3.701 0.000252 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

FACS sorting

The protocol is adapted from DeLuca et al (2020). Briefly, females fed on wet yeast are dissected and collagenase treated and fixed in 2% PFA. Fixed nuclei are isolated, stained with DAPI, and filtered before sorted with an a Becton-Dickenson FACSAria™ III cell sorter.

ChIP-seq

ChIp-seq was carried out as described in Deluca et al (2020). Briefly, we standardized input to the number of genomes rather than the number of cells, as we worked with cells with various ploidy levels spanning from 2C to 512C. For each IP, the equivalent of 500,000 2C sorted nuclei were used. Samples were mixed with a specific amount of fixed mouse 3T3 cells as a spike-in if applicable. Mixed samples were then sonicated to acquire mono-nucleosome fragments. Fragment sizes were confirmed with the Bioanalyzer (Agilent). 1% of each sample was set aside and later was reverse cross-linked to be used as input control. The rest of the nucleosome fragments are incubated with antibody-conjugated beads against the corresponding target protein for chromatin immunoprecipitation. Afterward, all samples were treated with RNAse A, and subsequently proteinase K. DNA was then extracted with phenol/chloroform, precipitated with NaAc/ethanol, washed with 70% ethanol wash, and resuspended samples in 10 µl water and of which all 10 µl was used for library preps. The library is prepared with Takara Bio ThurPLEX® DNA seq kit and sequenced for 75 b.p. single-end reads on an Illumina NextSeq 500

ATAC-seq

ATAC-seq libraries was originally prepared with the illumina Nextera™ DNA Sample Preparation Kits. Later experiments used the Illumina DNA Prep kit with a modified protocol. Sorted nuclei are resuspended in the Tagmentation Buffer 1 (TB1) and incubated with DNA Tagment DNA Enzyme (TDE1) (available as illumina Tagment DNA TDE1 Enzyme and Buffer

Kits) for 30 min at 37□. The reaction is then stopped with Tagment Stop Buffer (TSB) at 65□ overnight and digested with RNAase and Proteinase K for 1 and 2 hours respectively. Tagged DNA is extracted with phenol/chloroform and amplified with Enhanced PCR Mix (EPM), i7 and i5 index adapters for 10 PCR cycles. Amplified DNA is cleaned up with the Sample Purification Beads (SPB) or alternatively with the Beckman Coulter AMPure XP beads. The finished library is checked for size and concentration with Agilent Bioanalyzer and sequenced for 75 b.p. single-end reads on an Illumina NextSeq 500.

RNA-seq analysis

75 b.p. single-ended RNA-seq reads are aligned to the Drosophila melanogaster release 6 reference genome (dm6) and Ensembl annotation release 98 (dm6.22.Ensembl.28) with RSEM quantification package (1.3.0) (Li, and Dewey, 2011) paired with STAR RNA-seq aligner (2.7.3a) (Dobin et al. 2013). Differential expression analysis is performed with DESeq2 (1.26.0) (Love et al. 2014). Aligned reads are imported into DESeq2 with the tximport package (1.14.2) (Soneson et al. 2015). Normalized bigwig data for visualization is generated with deepTools (3.1.3) (Ramirez, 2016). GO analysis was performed using DAVID (da Huang et al. 2009a,b).

ChIP-seq analysis

ChIP-seq analysis was performed as described in Deluca et al (2020). 75 b.p. single-ended ChIP-seq reads were aligned to either the Drosophila melanogaster release six genome (w/o spike in) or a hybrid genome of Drosophila melanogaster and Mus musculus (w/ spike in) using Bowtie2 (Langmead and Salzberg, 2012) and processed using SAMtools (Li et al. 2009). To remove ambiguously mapped reads, only mapped reads that pass the mapping quality filtering were used for subsequent analysis. The corresponding BedGraph and bigWig files were generated using the bedtools genomecov function (Quinlan & Hall, 2010) and kentutils wigToBigWig function (citation)

respectively. Differential ChIP-seq analysis was performed with MACS2 bdgdiff function (Zhang et al. 2008). Signal distribution was modeled with mixtools using CPM normalized read counts (Benaglia et al. 2009). The read depth across a genomic region of interest was visualized and presented in the Integrative Genomics Viewer (Robinson et al. 2011).

Chromatin state calling

Bowtie2 was used to map 75 b.p. ChIP and Input reads (--very-sensitive -N 1) from GSCs, NCs, and FCs to the Drosophila genome (dm6), and subsequently filtered and sorted using SAMtools "view" (-q 10 -hF 4) and "sort" subcommands. BEDTools (Quinlan, 2010). "makewindows" was used to bin the genome into 500 bp non-overlapping bins (-w 500 -s 500) and the "coverage" subcommand was used to count the number of reads in each bin (-counts). For each ChIP/Input pair of bedGraphs, bins with 0 reads in either sample were discarded from both. The fold-change for each bin between the ChIP and Input was then taken. Raw fold-change values were between-sample normalized to the median fold-change value, similar to DEseq2 (ref), and then log10 transformed. A 3-state Hidden Markov Model (HMM) was trained on the log10 FC ratios using PufferFish v0.1.20200925 (https://github.com/JohnUrban/pufferfish; commit. f0c862a122357c062316d833f58ed6f6955c141c). It was initialized by learning the Viterbi Encoded state path starting with uniform initial and transition probabilities and semi-informed normally distributed emission probabilities that modeled a depleted state, a no-difference state, and an enriched state. The parameters and statepath were iteratively updated with Viterbi Training until the convergence of the log-likelihood of the Viterbi Encoded state path given the parameters. The final state path that reached convergence was used as the final segmentation of the ChIP data into 3 domain types: depleted, no-difference, or enriched.

For mapping between euchromatin and heterochromatin (Fig. 3B) the following limits of euchromatin were used:

chr = "2L" & start < 21400000; chr = "2R" & start > 5970000; chr = "3L" & start < 22270000

chr = "3R" & start > 4191000; chr == "X" & start < 21400000

ATAC-seq analysis

75 b.p. single-ended ATAC-seq reads are aligned and analyzed similarly as ChIP-seq data. Additionally, ATAC-seq peak calling across multiple samples is performed with the Genrich package ATAC-seq mode (0.6) (available at https://github.com/jsh58/Genrich).

Heterochromatin analysis based on H3K9me3 ChIp-seq

To examine the heterochromatin distribution on the chromosome scale, processed H3K9me3 ChIp-seq data were further segmented into 250kb bins using the bedtools multicov function (Quinlan & Hall, 2010). Reads were scaled by the sequencing depth. Corresponding input controls were used to eliminate backgrounds. Signals were then normalized with the euchromatin regions in between stages. The following limits of euchromatin at the 250kb bin resolution (deduced from the nearby reported euchromatin limits) were used: chr2L, 0 to 21,500,000; chr2R, 5,500,000 to 25,286,936; chr3L, 0 to 23,000,000; chr3R, 4,250,000 to 32,079,331; chrX, 0 to 21,500,000. Processed data were visualized on the line charts.

To examine the heterochromatin distribution on repetitive sequences, reads were reprocessed with the similar pipeline as detailed above with a few alterations. Specifically, reads were remapped to a customized masked genome (repetitive sequences masked) with the addition of individual

repetitive sequences as separated entries. Reads were scaled and normalized with the sequencing depth of euchromatin regions and corresponding input controls. Signals were quantified for individual repetitive sequences including transposons (TE) and satellites and visualized on the scatter plots.

Data and software availability

High throughput sequencing data and processed data is available under accession GSE145282 and GSE229943. Software code is available at https://github.com/ciwemb/polycomb-development (DeLuca et al. , 2020; copy archived at https://github. com/elifesciences-publications/polycomb-development) add new information.

ACKNOWLEDGEMENTS

We thank Dr. Fredrick Tan and Allison Pinder for assistance with sequencing and genomic analysis. Mamhud Siddiqi provided valuable assistance with microscopy. We thank Drs. Rebecca Frederick and Lucy Morris for assistance in analyzing Janelia lines. We thank Gerry Rubin, Todd Laverty and the Janelia Farm external investigator program for supporting the Spradling lab analysis of ovary expression using a sample of the Rubin enhancer collection. We thank current Spradling lab members for comments on the manuscript.

Additional information:

Funding: Steve Z DeLuca was a postdoctoral fellow of the Helen Hay Whitney Foundation. Haolong Zhu and Liang-Yu Pang are/were graduate students in the Cell, Molecular and Developmental Biology Program at Johns Hopkins University, Department of Biology. Allan Spradling is an Investigator and John Urban is an Associate of the Howard Hughes Medical Institute.

Author contributions: Steven Z. DeLuca, Conceptualization, Resources, Data curation, Software, formal analysis, Writing- review and editing; Liang-Yu Pang, Resources, Data curation, Software, Investigation, Visualization, Methodology; Writing, review and editing; Haolong Zhu Conceptualization, Resources, Data curation, formal analysis, Writing- review and editing; John Urban, Data curation, software, visualization, methodology; Writing- review and editing; Allan C Spradling, Conceptualization, Resources, Formal analysis, Supervision, Writing- original draft; review and editing.

REFERENCES

Ables,ET., Drummond-Barbosa D. (2013). Cyclin E controls Drosophila female germline stem cell maintenance independently of its role in proliferation by modulating responsiveness to niche signals. Development 140, 530–540.

Ahmad K, Henikoff S. (2022). The Drosophila embryo as a tabula rasa for the epigenome. Fac Rev. 11, 40. doi: 10.12703/r/11-40.

Benaglia T, Chauveau D, Hunter DR, and Young DS (2009). mixtools: an R Package for Analyzing Mixture Models. *J. Stat. Softw* 32, 29.

Bhaskar PK, Southard S, Baxter K, Van Doren M. (2022). Germline sex determination regulates sex-specific signaling between germline stem cells and their niche. Cell Rep. 39, 110620. doi: 10.1016/j.celrep.2022.110620.

Blatt P, Martin ET, Breznak SM, Rangan P. (2020). Post-transcriptional gene regulation regulates germline stem cell to oocyte transition during Drosophila oogenesis. Curr Top Dev Biol. 140, 3-34. doi: 10.1016/bs.ctdb.2019.10.003.

Blatt P, Wong-Deyrup SW, McCarthy A, Breznak S, Hurton MD, Upadhyay M, Bennink B, Camacho J, Lee MT, Rangan P. (2021). RNA degradation is required for the germ-cell to maternal transition in Drosophila. Curr Biol. 31, 2984-2994e7. d oi: 10.1016/j.cub.2021.04.052.

Breznak SM, Kotb NM, Rangan P. (2023). Dynamic regulation of ribosome levels and translation during development. Semin Cell Dev Biol. 136, 27-37. doi: 10.1016/j.semcdb.2022.06.004.

Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 109, 21.29.1-21.29.9. doi: 10.1002/0471142727.mb2129s109.

Caizzi R, Moschetti R, Piacentini L, Fanti L, Marsano RM, Dimitri P.(2016). Comparative Genomic Analyses Provide New Insights into the Evolutionary Dynamics of Heterochromatin in Drosophila. PLoS Genet. 12, e1006212. doi: 10.1371/journal.pgen.1006212.

Cao J, Yu T, Xu B, Hu Z, Zhang XO, Theurkauf WE, Weng Z. (2023). Epigenetic and chromosomal features drive transposon insertion in Drosophila melanogaster. Nucleic Acids Res. 51, 2066-2086. doi: 10.1093/nar/gkad054.

Chen P, Aravin AA. (2023). Genetic control of a sex-specific piRNA program. Curr Biol. Apr 6, S0960-9822(23)00377-9. doi: 10.1016/j.cub.2023.03.057.

Clark KA, McKearin DM. (1996). The Drosophila stonewall gene encodes a putative transcription factor essential for germ cell development. Development 122, 937-50. doi: 10.1242/dev.122.3.937.

Clough E, Moon W, Wang S, Smith K, Hazelrigg T (2007) Histone methylation is required for oogenesis in Drosophila. Development 134: 157–165.

Clough E, Tedeschi T, Hazelrigg T. (2014). Epigenetic regulation of oogenesis and germ stem cell maintenance by the Drosophila histone methyltransferase Eggless/dSetDB1. Dev Biol. 388, 181-91. doi: 10.1016/j.ydbio.2014.01.014.

Cronmiller C, Cline TW. (1987). The Drosophila sex determination gene daughterless has different functions in the germ line versus the soma. Cell 48, 479-87. doi: 10.1016/0092-8674(87)90198.

Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, Hannon GJ. (2018). piRNA-Guided Genome Defense: From Biogenesis to Silencing. Annu Rev Genet. 52, 131-157. doi: 10.1146/annurev-genet-120417-031441.

da Huang W, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13.

da Huang W, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4, 44–57.

David CN (2012). Interstitial stem cells in Hydra: multipotency and decision-making. Int. J. Dev. Biol. 56, 489-497. doi: 10.1387/ijdb.113476cd.

DeLuca SZ, Ghildiyal M, Pang LY, Spradling AC. (2020). Differentiating *Drosophila* female germ cells initiate Polycomb silencing by regulating PRC2-interacting proteins. Elife. 9, e56922. doi: 10.7554/eLife.56922.

Deshpande, G., Calhoun, G. and Schedl, P. (2004). Overlapping mechanisms function to establish transcriptional quiescence in the embryonic Drosophila germline. Development 131, 1247-1257.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. (2012). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21. doi: 10.1093/bioinformatics/bts635.

de Wit E, Greil F, van Steensel B. (2007). High-resolution mapping reveals links of HP1 with active and inactive chromatin components. PLoS Genet. 3, e38. doi: 10.1371/journal.pgen.0030038.

Drummond-Barbosa D. (2019). Local and physiological control of germline stem cell lineages in *Drosophila melanogaster*. Genetics 213, 9-26. doi:10.1534/ genetics.119.300234

Fabry MH, Falconio FA, Joud F, Lythgoe EK, Czech B, Hannon GJ. (2021). Maternally inherited piRNAs direct transient heterochromatin formation at active transposons during early *Drosophila* embryogenesis. Elife 10, e68573. doi: 10.7554/eLife.68573.

Fawcett DW, Ito S, Slauterback D. (1959). The occurrence of intercellular bridges in groups of cells exhibiting synchronous differentiation. J Biophys Biochem Cytol. 5, 453-60. doi: 10.1083/jcb.5.3.453.

Ferreira. J. Carmo-Fonseca, M. (1997). Genome replication in early mouse embryos follows a defined temporal and spatial order. J. Cell Sci. 110, 889-897.

Fierro-Constain, L., Schenkelaars, Q., Gazave, E., Haguenauer, A., Rocher, C., Ereskovsky A., Borchiellini, C., Renard, E. (2017). The Conservation of the Germline Multipotency Program, from Sponges to Vertebrates: A Stepping Stone to Understanding the Somatic and Germline Origins. Genome Biol. Evo. 9, 474-88.

Flaherty MS, Salis P, Evans CJ, Ekas LA, Marouf A, Zavadil J, Banerjee U, Bach EA. (2010). Chinmo is a functional effector of the JAK/STAT pathway that regulates eye development, tumor formation, and stem cell self-renewal in Drosophila. Developmental Cell 18, 556–568. DOI: https://doi.org/10.1016/j.devcel.2010.02.006

Friedli M, Trono D. (2020). Developmental control of transposable elements and the evolution of higher species. Annu. Rev. Cell. Dev. Biol. 31, 429-51.

Fuller MT, Spradling AC. (2007). Male and female Drosophila germline stem cells: two versions of immortality. Science 316, 402-4. doi: 10.1126/science.1140861.

Gabut M, Bourdelais F, Durand S.(2020). Ribosome and translational control in stem cells. Cells. 9, 497. doi: 10.3390/cells9020497.

Gleason RJ, Chen X. (2022). Epigenetic dynamics during germline development: insights from Drosophila and *C. elegans*. Curr Opin Genet Dev. 78, 102017. doi: 10.1016/j.gde.2022.102017.

Göke J, Lu X, Chan YS, Ng HH, Ly LH, Sachs F, Szczerbinska I. (2015). Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. Cell Stem Cell 16, 135-41. doi: 10.1016/j.stem.2015.01.005.

Gonzalez LE, Tang X, Lin H. (2021). Maternal Piwi regulates primordial germ cell development to ensure the fertility of female progeny in Drosophila. Genetics 219, iyab091. doi: 10.1093/genetics/iyab091.

Gramates LS, Agapite J, Attrill H, Calvi BR, Crosby MA, dos Santos G, Goodman JL, Goutte-Gattat D, Jenkins VK, Kaufman T, Larkin A, Matthews BB, Millburn G, Strelets VB and the FlyBase Consortium (2022). FlyBase: A guided tour of highlighted features. Genetics 220, iyac035 DOI:10.1093/genetics/iyac035.

Greenblatt EJ, Obniski R, Mical C, Spradling AC. (2019). Prolonged ovarian storage of mature *Drosophila*oocytes dramatically increases meiotic spindle instability. Elife. 8, e49455. doi: 10.7554/eLife.49455.

Grmai L, Hudry B, Miguel-Aliaga I, Bach EA. (2018). Chinmo prevents transformer alternative splicing to maintain male sex identity. PLoS Genet. 14e1007203

Grmai L, Pozmanter C, Van Doren M.(2022). The Regulation of Germline Sex Determination in Drosophila by Sex lethal. Sex Dev doi: 10.1159/0005212351-6.

Gu T, Elgin SC. (2013). Maternal depletion of Piwi, a component of the RNAi system, impacts heterochromatin formation in Drosophila. PLoS Genet. 9, e1003780. doi: 10.1371/journal.pgen.1003780.

Guthrie W, Filliben JJ, Hecker A.. (2002). "NIST/SEMATECH e-Handbook of Statistical Methods Chapter 4." from http://www.itl.nist.gov/div898/handbook/.

Harris RE, Pargett M, Sutcliffe C, Umulis D, Ashe HL. (2011). Brat promotes stem cell differentiation via control of a bistable switch that restricts BMP signaling. Dev Cell. 20, 72-83. doi: 10.1016/j.devcel.2010.11.019.

Hayashi Y, Otsuka K, Ebina M, Igarashi K, Takehara A, Matsumoto M, Kanai A, Igarashi K, Soga T, Matsui Y. (2017). Distinct requirements for energy metabolism in mouse primordial germ cells and their reprogramming to embryonic germ cells. Proc Natl Acad Sci U S A. 114, 8289-8294. doi: 10.1073/pnas.1620915114.

Hayashi Y, Matsui Y. (2022). Metabolic Control of Germline Formation and Differentiation in Mammals. Sex Dev. 27:1-16. doi: 10.1159/000520662.

Hill PWS, Leitch HG, Requena CE, Sun Z, Amouroux R, Roman-Trufero M, Borkowska M, Terragni J, Vaisvila R, Linnett S, Bagci H, Dharmalingham G, Haberle V, Lenhard B, Zheng Y, Pradhan S, Hajkova P. (2018). Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. Nature 555, 392-396. doi: 10.1038/nature25964.

Hinnant TD, Merkle JA, Ables ET. (2020). Coordinating Proliferation, Polarity, and Cell Fate in the Drosophila Female Germline. Front Cell Dev Biol. 8, 19.

Hocaoglu H, Wang L, Yang M, Yue S, Sieber M. (2021). Heritable shifts in redox metabolites during mitochondrial quiescence reprogram progeny metabolism. Nat Metab. 3, 1259-1274. doi: 10.1038/s42255-021-00450-3.

Holmqvist PH, Mannervik M. (2013). Genomic occupancy of the transcriptional co-activators p300 and CBP. Transcription 4, 18-23.

Holstein TW. (2023). The Hydra stem cell system - Revisited. Cells Dev. 174, 203846. doi: 10.1016/j.cdev.2023.203846.

Hopes T, Norris K, Agapiou M, McCarthy CGP, Lewis PA, O'Connell MJ, Fontana J, Aspden JL. (2022). Ribosome heterogeneity in Drosophila melanogaster gonads through paralog-switching. Nucleic Acids Res. 50, 2240-2257. doi: 10.1093/nar/gkab606.

Hsu HJ, LaFever L, Drummond-Barbosa D. (2008) Diet controls normal and tumorous germline stem cells via insulin-dependent and -independent mechanisms in Drosophila. Dev. Biol. 313: 700–712. https://doi.org/10.1016/j.ydbio. 2007.11.006

Huang X, Fejes Tóth K, Aravin AA. (2017). piRNA Biogenesis in Drosophila melanogaster. Trends Genet. 33, 882-894. doi: 10.1016/j.tig.2017.09.002.

Hyrien, Ol, Mechali, M. (1995). Transition in specification of embryonic metazoan DNA replication origins. Science 270, 994-7.

Issigonis M, Newmark PA. (2019). From worm to germ: Germ cell development and regeneration in planarians. Curr Top Dev Biol. 135:127-153. doi: 10.1016/bs.ctdb.2019.04.001.

James TC, Eissenberg JC, Craig C, Dietrich V, Hobson A, Elgin SC. (1989). Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of Drosophila. Eur J Cell Biol. 50, 170-80.

Jenett A, Rubin GM, Ngo TT, Shepherd D, Murphy C, et al. (2012). A GAL4-driver line resource for Drosophila neurobiology. Cell Rep. 2, 991-1001.

Jevitt A, Chatterjee D, Xie G, Wang XF, Otwell T, Huang YC, Deng WM. (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.

Kai T., Williams D, Spradling AC. (2005). The expression profile of purified Drosophila germline stem cells. Dev. Biol 283, 486-502.

Karpen GH, Spradling AC. (1992). Analysis of subtelomeric heterochromatin in the Drosophila minichromosome Dp1187 by single P element insertional mutagenesis. Genetics 132, 737-53. doi: 10.1093/genetics/132.3.737.

Kinoshita M, Smith A. Pluripotency Deconstructed. (2018). Dev Growth Differ. 60, 44-52. doi: 10.1111/dgd.12419.

Klein KN, Zhao PA, Lyu X, Sasaki T, Bartlett DA, Singh AM, Tasan I, Zhang M, Watts LP, Hiraga SI, Natsume T, Zhou X, Baslan T, Leung D, Kanemaki MT, Donaldson AD, Zhao H, Dalton S, Corces VG, Gilbert DM. (2021). Replication timing maintains the global epigenetic state in human cells. Science 372, 371-378. doi: 10.1126/science.aba5545.

Kong J, Han H, Bergalet J, Bouvrette LPB, Hernández G, Moon NS, Vali H, Lécuyer É, Lasko P. (2019). A ribosomal protein S5 isoform is essential for oogenesis and interacts with distinct RNAs in Drosophila melanogaster. Sci Rep. 24, 13779. doi: 10.1038/s41598-019-50357.

Kory N, Wyant GA, Prakash G, Uit de Bos J, Bottanelli F, Pacold ME, Chan SH, Lewis CA, Wang T, Keys HR, Guo YE, Sabatini DM. (2018). SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism. Science 362, eaat9528. doi: 10.1126/science.aat9528.

LaFever L, Drummond-Barbosa D. (2005). Direct control of germline stem cell division and cyst growth by neural insulin in Drosophila. Science 309, 1071-1073.

Leader DP, Krause SA, Pandit A, Davies SA, Dow JAT. (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.

Lebo DPV, McCall K. (2021). Murder on the Ovarian Express: A Tale of Non-Autonomous Cell Death in the Drosophila Ovary. Cells 10, 1454. doi: 10.3390/cells10061454.

Lehrmann H, Pritchard LL, Harel-Bellan A. (2002). Histone acetyltransferases and deacetylases in the control of cell proliferation and differentiation. Adv Cancer Res. 2002;86:41-65.

Lei L, Spradling AC. (2016). Mouse oocytes differentiate through organelle enrichment from sister cyst germ cells. Science 352, 95-99.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009). 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-9. doi: 10.1093/bioinformatics/btp352.

Li B, Dewey CN. (2011). Rsem: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics. 12, 323.

Li H, Rai M, Buddika K, Sterrett MC, Luhur A, et al. (2019). Lactate dehydrogenase and glycerol-3-phosphate dehydrogenase cooperatively regulate growth and carbohydrate metabolism during Drosophila melanogaster larval development. Development 146: dev175315.

Li H, Janssens J, De Waegeneer M, Kolluru SS, Davie K, Gardeux V, Saelens W, David FPA, Brbić M, Spanier K, Leskovec J et. al. (2022). Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. Science 375, eabk2432. doi: 10.1126/science.abk2432. .

Lim RS, Anand A, Nishimiya-Fujisawa C, Kobayashi S, Kai T. (2014). Analysis of Hydra PIWI proteins and piRNAs uncover early evolutionary origins of the piRNA pathway. Dev Biol. 386, 237-51. doi: 10.1016/j.ydbio.2013.12.007.

Lin H, Spradling AC. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development 124, 2463-76. doi: 10.1242/dev.124.12.2463.

Liu S, Brind'Amour J, Karimi MM, Shirane K, Bogutz A, Lefebvre L, Sasaki H, Shinkai Y, Lorincz MC. (2014). Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. Genes Dev. 28, 2041–2055. https://doi.org/10.1101/gad.244848.114.

Littlefield LC. (1994). Cell-cell interactions and the control of sex determination in hydra. Dev. Biol. 5, 13-20.

Loh CH, Veenstra GJC. (2022). The Role of Polycomb Proteins in Cell Lineage Commitment and Embryonic Development. Epigenomes. 6, 23. doi: 10.3390/epigenomes6030023.

Love MI, Huber W, Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550.

Lu BY, Ma J, Eissenberg JC. (1998). Developmental regulation of heterochromatin-mediated gene silencing in Drosophila. Development 125, 2223–2234.

Lu K, Jensen L, Lei L, Yamashita YM. (2017). Stay Connected: A Germ Cell Strategy. Trends Genet. 33, 971-978. doi: 10.1016/j.tig.2017.09.001.

Ma X, Zhu X, Han Y, Story B, Do T, Song X, Wang S, Zhang Y, Blanchette M, Gogol M, Hall K, Peak A, Anoja P, Xie T. (2017). Aubergine controls germline stem cell self-renewal and progeny differentiation via distinct mechanisms. Dev Cell. 41, 157-169.e5. doi: 10.1016/j.devcel.2017.03.023.

Macheret,M. and Halazonetis,T.D. (2018) Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. Nature 555, 112-116.

Macrae TA, Fothergill-Robinson J, Ramalho-Santos M. (2023. Regulation, functions and transmission of bivalent chromatin during mammalian development. Nat Rev Mol Cell Biol. 24, 6-26. doi: 10.1038/s41580-022-00518-2.

Mahowald AP, Strassheim JM. (1970). Intercellular migration of centrioles in the germarium of Drosophila melanogaster. An electron microscopic study. J Cell Biol. 45, 306-20. doi: 10.1083/jcb.45.2.306.

Maines JZ, Stevens LM, Tong X, Stein,D. (2004). Drosophila dMyc is required for ovary cell growth and endoreplication. Development 131, 775--786.

Martinho RG, Kunwar PS, Casanova J, Lehmann R. (2004). A noncoding RNA is required for the repression of RNApolII-dependent transcription in primordial germ cells. Curr Biol. 14, 159-65. doi: 10.1016/j.cub.2003.12.036.

Matova N, Cooley L. (2001). Comparative aspects of animal oogenesis. Dev Biol. 231(2):291–320.

Mercer M, Jang S, Ni C and Buszczak M (2021) The Dynamic Regulation of mRNA Translation and Ribosome Biogenesis During Germ Cell Development and Reproductive Aging. Front. Cell Dev. Biol. 9:710186. doi: 10.3389/fcell.2021.710186.

Methot SP, Padeken J, Brancati G, Zeller P, Delaney CE, Gaidatzis D, Kohler H, van Oudenaarden A, Großhans H, Gasser SM. (2021). H3K9me selectively blocks transcription factor activity and ensures differentiated tissue integrity. Nat Cell Biol. 23, 1163-1175. doi: 10.1038/s41556-021-00776-w.

Nakamura A, Amikura R, Hanyu K, Kobayashi S. (2001). Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. Development 128, 3233-42. doi: 10.1242/dev.128.17.3233.

Nakatani T, Lin J, Ji F, Ettinger A, Pontabry J, Tokoro M, Altamirano-Pacheco L, Fiorentino J, Mahammadov E, Hatano Y, Van Rechem C, Chakraborty D, Ruiz-Morales ER, Arguello Pascualli PY, Scialdone A, Yamagata K, Whetstine JR, Sadreyev RI, Torres-Padilla ME. (2022). DNA replication fork speed underlies cell fate changes and promotes reprogramming. Nat Genet. 54, 318-327. doi: 10.1038/s41588-022-01023-0.

Ninova M, Chen YA, Godneeva B, Rogers AK, Luo Y, Fejes Tóth K, Aravin AA. (2020). Su(var)2-10 and the SUMO Pathway Link piRNA-Guided Target Recognition to Chromatin Silencing. Mol Cell 77, 556-570.e6. doi: 10.1016/j.molcel.2019.11.012.

Niu W, Spradling AC. (2022). Mouse oocytes develop in cysts with the help of nurse cells. Cell 185, 2576-2590.

O'Farrell, P.H., Stumpff, J., Su, T.T. (2004). Embryonic cleavage cycles: how is a mouse like a fly? Curr. Biol. 14, R35-45.

Omole AE, Fakoya AOJ. (2018). Ten years of progress and promise of induced pluripotent stem cells: historical origins, characteristics, mechanisms, limitations, and potential applications. Peer J .6, e4370. doi: 10.7717/peerj.4370.

Onishi R, Yamanaka S, Siomi MC. (2021). piRNA- and siRNA-mediated transcriptional repression in Drosophila, mice, and yeast: new insights and biodiversity. EMBO Rep. 22, e53062. doi: 10.15252/embr.202153062.

Padeken J, Methot SP, Gasser SM. (2022). Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. Nat Rev Mol Cell Biol. 23, 623-640. doi: 10.1038/s41580-022-00483-w.

Pauklin, S. and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. Cell 155, 135-147. https://doi.org/10.1016/j.cell.2013.08.031.

Pepling ME, Spradling AC. (1998). Female mouse germ cells form synchronously dividing cysts. Development 125, 3323-8. doi: 10.1242/dev.125.17.3323.

Pepling ME, Wilhelm JE, O'Hara AL, Gephardt GW, Spradling AC. (2007). Mouse oocytes within germ cell cysts and primordial follicles contain a Balbiani body. Proc Natl Acad Sci U S A. 104, 187-92. doi: 10.1073/pnas.0609923104.

Percharde M, Lin CJ, Yin Y, Guan J, Peixoto GA, Bulut-Karslioglu A, Biechele S, Huang B, Shen X, Ramalho-Santos M. (2018). A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. Cell. 2018 Jul 12;174(2):391-405.e19. doi: 10.1016/j.cell.2018.05.043.

Posfai E, Kunzmann R, Brochard V, Salvaing J, Cabuy E, Roloff TC, Liu Z, Tardat M, van Lohuizen M, Vidal M, Beaujean N, Peters AH. (2012). Polycomb function during oogenesis is required for mouse embryonic development. Genes Dev. 26, 920-32. doi: 10.1101/gad.188094.112.

Primus et al. (2019) Tudor-domain containing protein 5-like promotes male sexual identity in the Drosophila germline and is repressed in females by Sxl. PLoS Genet. 15, e1007617.

Quinlan, A. R. and I. M. Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features." Bioinformatics 26(6): 841-842.

Rafalski VA, Mancin E, Brunet A (2012). Energy metabolism and energy sensing pathways in mammalian embryonic and adult stem cell fate. J. Cell Science 125, 5597-5608.

Ramirez, F., D. P. Ryan, B. Gruning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dundar and T. Manke (2016). "deepTools2: a next generation web server for deepsequencing data analysis." Nucleic Acids Res 44(W1), W160-165.

Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ, Lehmann R. (2011). piRNA production requires heterochromatin formation in Drosophila. Curr Biol. 21, 1373-9. doi: 10.1016/j.cub.2011.06.057.

Rausch C, Weber P, Prorok P, Hörl D, Maiser A, Lehmkuhl A, Chagin VO, Casas-Delucchi CS, Leonhardt H, CardosoMC (2020). Nuc. Acids Res.48, 12751-12777.

Reddien PW. (2022). Positional Information and Stem Cells Combine to Result in Planarian Regeneration. Cold Spring Harb Perspect Biol. 14, a040717. doi: 10.1101/cshperspect.a040717.

Rust K, Byrnes LE, Yu KS, Park JS, Sneddon JB, Tward AD, Nystul TG. (2020). A single-cell atlas and lineage analysis of the adult Drosophila ovary. Nat Commun. 11, 5628.

Sanchez CG, Teixeira FK, Czech B, Preall JB, Zamparini AL, Seifert JR, Malone CD, Hannon GJ, Lehmann R. (2016). Regulation of ribosome biogenesis and protein synthesis controls germline stem cell Differentiation. Cell Stem Cell 18, 276-90. doi: 10.1016/j.stem.2015.11.004.

Scherer G, Tschudi C, Perera J, Delius H, Pirrotta V. (1982). B104, a new dispersed repeated gene family in Drosophila melanogaster and its analogies with retroviruses. J Mol Biol. 157, 435-51. doi: 10.1016/0022-2836(82)90470-3.

Seller CA, O'Farrell PH. (2018). Rif1 prolongs the embryonic S phase at the Drosophila mid-blastula transition. PLoS Biol. 16, e2005687. doi: 10.1371/journal.pbio.2005687.

Seller CA, Cho CY, O'Farrell PH. (2019). Rapid embryonic cell cycles defer the establishment of heterochromatin by eggless/SetDB1 in Drosophila. Genes & Development 33:403–417.

Shapiro-Kulnane, L., Smolko, A.E., Salz, H.K. (2015). Maintenance of Drosophila germline sex cell identity in oogenesis and tumorigenesis. Development 147, dev192856.

Sheng XR, Matunis E. (2011). Live imaging of the Drosophila spermatogonial stem cell niche reveals novel mechanisms regulating germline stem cell output. Development 138, 3367-76. doi: 10.1242/dev.065797.

Shigenobu, S., Kitadate, Y., Noda, C. & Kobayashi, S. (2006). Molecular characterization of embryonic gonads by gene expression profiling in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA 103, 13728–13733 (2006).

Sieber, MH., Thomsen, MB. and Spradling, A.C. (2016). Electron transport chain remodeling by GSK3 during oogenesis connects nutrient state to reproduction. Cell 164: 420-32.

Sieber, M., Spradling, A.C. (2017). The role of metabolic states in development and disease. Curr. Opinion Genet. Dev. 45, 58-68.

Sjakste N, Riekstiņa U. (2021). DNA damage and repair in differentiation of stem cells and cells of connective cell lineages: A trigger or a complication? Eur J Histochem. 65, 3236. doi: 10.4081/ejh.2021.3236.

Slaidina M, Lehmann R. (2014). Translational control in germline stem cell development. J Cell Biol. 207, 13-21. doi: 10.1083/jcb.201407102.

Slaidina M, Banisch TU, Gupta S, Lehmann R. (2020). A single-cell atlas of the developing Drosophila\ ovary identifies follicle stem cell progenitors. Genes Dev. 34, 239-249.

Slaidina M, Gupta S, Banisch TU, Lehmann R. (2021). A single-cell atlas reveals unanticipated cell type complexity in Drosophila ovaries. Genome Res. 31, 1938-1951. doi: 10.1101/gr.274340.120.

Slautterbck DB, Fawcett DW. (1959). The development of the cnidoblasts of Hydra; an electron microscope study of cell differentiation. J Biophys Biochem Cytol. 5, 441-52. doi: 10.1083/jcb.5.3.441.

Smolko, A.E., Shapiro-Kulnane, L., Salz, H.K. (2020). An autoregulatory switch in sex-specific phf7 transcription causes loss of sexual identity and tumors in the Drosophila female germline. Development 147, dev192856.

Soldner F, Jaenisch R. (2018). Stem Cells, Genome Editing, and the Path to Translational Medicine. Cell 175, 15-632. doi: 10.1016/j.cell.2018.09.010.

Soneson, C., M. I. Love and M. D. Robinson (2015). "Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences." F1000Res 4: 1521.

Spradling AC, Bellen HJ, and Hoskins RA. (2011). Drosophila *P* elements preferentially transpose to replication origins. Proc. Natl. Acad. Sci., USA 108, 15948-15953**.** doi: 10.1073/pnas.1112960108.

Spradling AC, Niu W, Yin Q, Pathak M, Maurya B. (2022). Conservation of oocyte development in germline cysts from Drosophila to mouse. Elife. 11: e83230. doi: 10.7554/eLife.83230.

Spradling A. (2023). The origin and function of germline cysts. Results and Problems in Cell Differentiation 72, in press.

Su TT, Campbell SD, and O'Farrell PH. (1998). The cell cycle program in germ cells of the Drosophila embryo. Dev. Biol. 196, 160-170.

Su Q, He H, Zhou Q. (2021). On the Origin and Evolution of Drosophila New Genes during Spermatogenesis. Genes (Basel) 12, 1796.

Takahashi K, Yamanaka S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663-676.

Teixeira FK, Sanchez CG, Hurd TR, Seifert JRK, Czech B, Preall JB, et al. (2015). ATP synthase promotes germ cell differentiation independent of oxidative phosphorylation. Nat. Cell Biol. 17, 689–696. doi: 10.1038/ ncb3165

Tennessen, J.M., Baker, K.D., Lam, G., Evans, J., and Thummel, C.S. (2011). The Drosophila estrogen-related receptor directs a metabolic switch that supports developmental growth. Cell Metab 13, 139-148.

ter Huurne M, Stunnenberg HG (2021). G1 phase progression in pluripotent stem cells. Cellular and Molecular Life Sciences (2021) 78:4507-4519.

Theunissen TW, Friedli M, He Y, Planet E, O'Neil RC, Markoulaki S, Pontis J, Wang H, Iouranova A, Imbeault M, Duc J, Cohen MA, Wert KJ, Castanon R, Zhang Z, Huang Y, Nery JR, Drotar J, Lungjangwa T, Trono D, Ecker JR, Jaenisch R. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell. 19, 502-515. doi: 10.1016/j.stem.2016.06.011.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. (1998). Embryonic stem cell lines derived from human blastocysts. Science 282, 1145-1147.

Vogg MC, Buzgariu W, Suknovic NS, Galliot B. (2021). Cellular, Metabolic, and Developmental Dimensions of Whole-Body Regeneration in *Hydra*. Cold Spring Harb Perspect Biol. 13, a040725. doi: 10.1101/cshperspect.a040725.

Wang ZH, Liu Y, Chaitankar V, Pirooznia M, Xu H. (2019). Electron transport chain biogenesis activated by a JNK-insulin-Myc relay primes mitochondrial inheritance in Drosophila. Elife. 8, e49309.

Wang C, Lin H. (2021). Roles of piRNAs in transposon and pseudogene regulation of germline mRNAs and lncRNAs. Genome Biol. 22, 27. doi: 10.1186/s13059-020-02221-x.

Wei KH-C, Chan C, Bachtrog D (2021). Establishment of H3K9me3-dependent heterochromatin during embryogenesis in Drosophila miranda. Elife. 10, e55612. doi: 10.7554/eLife.55612.

Wells JN, Feschotte C. (2020). A Field Guide to Eukaryotic Transposable Elements. Annu Rev Genet. 54, 539-561. doi: 10.1146/annurev-genet-040620-022145.

Wooten M, Snedeker J, Nizami ZF, Yang X, Ranjan R, Urban E, Kim JM, Gall J, Xiao J, Chen X. Asymmetric histone inheritance via strand-specific incorporation and biased replication fork movement. Nat Struct Mol Biol. 2019 Aug;26(8):732-743. doi: 10.1038/s41594-019-0269-z. Epub 2019 Jul 29. PMID: 31358945; PMCID: PMC6684448..

Xiang Y, Liang H. (2021). The Regulation and Functions of Endogenous Retrovirus in Embryo Development and Stem Cell Differentiation. Stem Cells Int. Feb 27; 6660936. doi: 10.1155/2021/6660936.

Yang SY, Baxter EM, Van Doren M. (2012). Phf7 controls male sex determination in the Drosophila germline. Dev Cell 22, 1041–51.

Yang, F., Quan, Z., Huang, H., He, M., Liu, X., Cai, T., Xi, R. (2019). Ovaries absent links dLsd1 to HP1a for local H3K4 demethylation required for heterochromatic gene silencing. eLife 8, e40806.

Yoon J, Lee K-S, Park JS, Yu K, Paik S-G, Kang Y-K (2008) dSETDB1 and SU(VAR)3–9 Sequentially Function during Germline-Stem Cell Differentiation in Drosophila melanogaster. PLoS ONE 3(5): e2234. https://doi.org/10.1371/journal.pone.0002234.

Zaccai M, Lipshitz HD. (1996). Role of Adducin-like (hu-li tai shao) mRNA and protein localization in regulating cytoskeletal structure and function during Drosophila Oogenesis and early embryogenesis. Dev Genet. 19, 249-57.

Zeng A, Li H, Guo L, Gao X, McKinney S, Wang Y, Yu Z, Park J, Semerad C, Ross E, Cheng LC, Davies E, Lei K, Wang W, Perera A, Hall K, Peak A, Box A, Sánchez Alvarado A. (2018). Prospectively Isolated Tetraspanin^{+} Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration. Cell 173, 1593-1608.e20. doi: 10.1016/j.cell.2018.05.006.

Zenk F, Zhan Y, Kos P, Löser E, Atinbayeva N, Schächtle M, Tiana G, Giorgetti L, Iovino N. (2021). HP1 drives de novo 3D genome reorganization in early Drosophila embryos. Nature 593, 289-293. doi: 10.1038/s41586-021-03460-z.

Zhang Q, Shalaby NA, Buszczak M. (2014). Changes in rRNA transcription influence proliferation and cell fate within a stem cell lineage. Science. 2343, 98-301. doi: 10.1126/science.1246384. PMID: 24436420.

Zeng A, Li H, Guo L, Gao X, McKinney S, Wang Y, Yu Z, Park J, Semerad C, Ross E, Cheng LC, Davies E, Lei K, Wang W, Perera A, Hall K, Peak A, Box A, Sánchez Alvarado A. (2018). Prospectively Isolated Tetraspanin+ Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration. Cell 173, 1593-1608.e20. doi: 10.1016/j.cell.2018.05.006.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK.(2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 10, 1523. doi: 10.1038/s41467-019-09234-6.

Zhou JJ, Cho KWY. (2022). Epigenomic dynamics of early Xenopus Embryos. Dev Growth Differ. 64, 508-516. doi: 10.1111/dgd.12813.

Zinshteyn, D., Barbash, D.A. (2022). Stonewall prevents expression of ectopic genes in the ovary and accumulates at insulator elements in D. melanogaster.

FIGURE LEGENDS

Figure 1. Analyzing Drosophila female germline chromatin

A. Stages of female gamete development in an ovariole. A single germline stem cell (GSC) and germline cyst are shown. To the right, a germarium is illustrated showing regions 1-3. Follicles in an ovariole are pictured at a lower magnification starting with stage 2 (S2). Below the diagrams, major events are summarized. B. FACS purification (upper panel) of 4c-512c germ cells for ATAC after earlier separation of 2c and 4c GSCs, and 2c-16c follicle cells (see Figure S2). DAPI, DNA content; Inset shows further purification based on germ cell marker expression (tomato). Below, follicle stages collected for RNAseq and Chip-seq are show by black bars. C. The number of genes expressed (>1 tpm, blue) and off (<1 tpm, orange) in GSCs decreases as germ cell development proceeds downstream. D. Box plot showing a decline in total ATAC peaks from GSCs to young ovaries (before S6) to whole ovaries (S6-9). E. Stem cells already express a high fraction of 79 oocyte maternal effect genes defined in Drosophila genetic screens.

Figure 2. Heterochromatin forms in cysts downstream from the GSC

A,B. Germaria stained for (A) H3K9me3-binding protein HP1a (pink) and CENP-A (green), or (B) H3K9me3 (green), fusome antibody 1B1 (pink) and DAPI (blue). C. H3K9me3 Chip-seq of unambiguously mapped reads spanning chromosome 2 (lower) and 3 (upper) in 250 kb bins from GSC, 4C NC or 8C NC. D. Plots showing relative enrichment of reads mapping to TEs or to satellite sequences in 4C, and 8C germ cells relative to GSCs. E. A representative euchromatic hsGFP insertion after heat shock expresses GFP (green) in GSCs and nearly all downstream cyst and follicular NC (arrows). In contrast a typical heterochromatic insertion after heat shock only expresses in GSCs and early cysts (upper arrow), but has become repressed in 16-cell cysts and NC

in meiotic and later follicles (lower arrow; other labeled cells are somatic). F. Repression of heterochromatic hsGFP expression after heat shock (+hs) (Control) requires the H3K9 methylase *eggless/SETDB1* (eggRNAi). G. Diagram shows the ratio of hsGFP expression in *piwi*(GLKD) or control genetic backgrounds from individual hsGFP insertions (points) at the indicated developmental stages (Figure 1A) colored based on insert location in euchromatin (pink), pericentric chromatin (green) or centric heterochromatin (blue). note: SC = GSC; CB = cystoblast (1st cell of germline cyst). H. Summary diagram of Drosophila female germ cell generational cycle showing two highly potent stem cell chromatin states (pink) beginning with the GSC or zygote and their downstream developmental trajectories. Both the GSC (maternal) and zygotic trajectories quickly undergo heterochromatin formation (blue) and then somatic cell production dependent on Polycomb repression (green). see text and Figure S2 for further details).

Figure 3. E(z)-dependent repression contributes to sex-differential gene expression

A. Diagram of Drosophila PRC1 and PRC2. B. Chromatin changes around the GSC differentiation gene *bgcn*. The promoter-associated H3K27ac peak is lost by 32cNCs, when H3K27me3 has begun to accumulate. C. The number of genes strongly downregulated by E(z) in young ovaries and their predominant tissue of expression. D. Chromatin changes around the CG43088 testis gene related to a transposon. E. The male GSC gene *escargot (esg)* downregulated by OVS6 shows increased H3K27me3. F. The *chinmo* gene encoding a transcription regulator required in male and female germ cells is repressed as in E. G. The testis regulatory gene *aly*, is repressed as in E. H. The tdrdl5 male germ cell regulator is repressed as in C. I. Phf7 gene regulation. J. Summary of E(z)-dependent repression of male and somatic genes during early female gene expression. K. Three sub classes of genes downregulated by E(z).

Figure 4. Maintenance and exit from the GSC state

A. Ovarioles subjected to white gene (w^{GLKD}) or *nejire* gene ($nejire^{GLKD}$) germline knockdown. Levels of H3K27ac (green) are greatly reduced in nej^{GLKD} and germ cell development is arrested at stage 1-2. Hts staining (1B1, purple) highlights cell membranes and fusomes. B. Chromatin tracks around the CGG gene *piwi* which is expressed throughout oogenesis. H3K27me3 does not increase. C. mRNA abundance vs the transcription start site (TSS) levels of H3K27ac are plotted versus GSC gene expression. Genes are colored according to hidden Markov model (HMM) calls of H3K27me3 levels. D. The number of gene-associated H3K27ac peaks (upper) or ATAC peaks (lower) in GSCs is plotted versus $\log_{10}$ peak score. 81% of H3K27ac peaks and 68% of ATAC peaks are associated with TSSs; the remaining peaks may be associated with enhancer regions. E. Examples GSC expressed genes that are downregulated in young ovaries in a E(z)-dependent manner (Table S2). The H3K27ac patterns of two such genes, blanks and eIF4E3 are shown. F.Germarisum immunostained for H3K29ac showing low levels in early germ cells (dashed pink circle) and increase in S1-S3 nurse cells (purple dashed circles). G. Plot of GSC mRNA abundance vs expression change in young ovary compared to GSCs. A strong correlation can be seen between genes with changes in mRNA levels and changes in H3K27ac levels. Genes (dots) are colored according to changes in ac = H3K27ac levels.

Figure 5 Genes induced in germ cells downstream from GSCs mediate a metabolic and growth transition

A. Germ cell-expressed transcription regulators with the largest fold change increases in young NC (OVS6) compared to GSCs. B. GO terms for genes upregulated in young ovaries vs GSCs. The GO categories reveal that a dramatic metabolic shift occurs between GSCs and OVS6 nurse cells. C.

Expression values (tpm) in GSCs and OVS6 NCs of glycolytic genes showing mostly upregulation but a shutoff of Ldh. D. Diagram showing upregulated transmembrane transport genes (1st GO category in B). Sub-cellular localization of a sample (20) of membrane transporters (Dros name/Mouse name) is indicated (arrows) on the cell drawing showing plasma membrane (red), mitochondrion (blue) or Golgi (orange). along with and their fold change (increase) in NC. Many are key regulators of cellular import or mitochondrial metabolism that help generate the precursor nucleotides, lipids and carbohydrates needed for rapid cellular growth. For complete list (Table S5).

Figure 6. The GSC state provides high immunity to transposable elements (TEs)

A. Proposed amplification of TE resistance by a germline cyst or syncytial stage downstream from a highly potent stem cell by amplifying and sharing piRNA (blue lines) between all connected cells following TE1 movement in one cell. B. Summed expression (tpm) of conserved germline genes (CGG) genes (blue) involved in transposon regulation and nuclear piwi genes (orange). C. Illustrative diagram of cluster analysis in a male GSC lineage with two cyst divisions and a single meiotic division leading to 8 sperm. The number of sperm with identical TE insertions depends of the time of insertion during cyst formation, as shown by 3 different TE examples that insert at different stages. D. Model of P element (triangles) copy number increase on chromosomes (horizontal lines) during a cyst cell or meiotic germ cell cycle by "replication timing". red = ORC proteins, Green = MCM proteins, black line shows postulated repression prior to loss of pre-replication complex in S phase by origin activation or fork passage. blue arrow shown transposition from a replicated to an unreplicated region (recognized by unfired pre-replication complex during later S phase (modified from Spradling et al. 2011). E. Like D, but in a GSC whose short G1 is proposed activate all origins simultaneously. F. An electron micrograph of a Drosophila female GSC

showing the spectrosome (ss) adjacent to cap cells (cap) and near terminal filament cells (tf). Note, section does not include nucleus.

Supplemental Material

Supplementary Materials and Methods

Volume measurement

MTD-Gal4 > UASz-tdTomato flies with labeled germline cells are fed on wet yeast for 3 days. Ovaries are dissected carefully with forceps to break apart the surrounding muscle sheets. Dissected ovaries are fixed in 4% PFA for 10 min, stained with DAPI, and mounted in VECTASHIELD®. For *yl* mutant and control lines, dissected ovaries are stained with anti-vasa to mark germline cell volumes (see immunofluorescent staining method for details). Full z-stacked images of the ovary are taken under 40X oil objective with Leica TCS SP5 Confocal Microscope. Reconstruction of 3D images is performed with Oxford Instruments Imaris software. Volumes and total DAPI signal pixels of individual follicles are recorded and developmental stages of the follicles are manually noted. Growth rate changes are plotted and modeled with R.

Growth modeling

Local polynomial regression was used to accommodate for the complexity of the data. Smoothing factor =0.75 and degree of polynomial =2 were used. The data set is fairly dense with n>20 for each stage and is therefore suitable for local polynomial regression. However, the result of LOESS can be difficult to interpret because it cannot be written down as simple equations (Guthrie et al. 2002). To further test the developmental growth rate changes, multiple linear regression was applied with two conditional terms. Condition1 distinguishes follicles before (stage 1 to 5) and after (stage 6 to 10) nurse cell chromatin dispersion. Condition2 distinguishes follicles before (stage 1 to 7) and after (stage 8 to 10) onset of vitellogenesis. The model to be tested is as following:

log10 Volume = a0*log2DNA +a1*log2DNA*condition1 +a2*log2DNA*condition2

+b0+b1*condition 1 +b2*condition2

Supplementary Tables

Table S1. All processed data

Table S2. E(z)-dependent gene downregulation.

Table S3. Mostly heterochromatic GSC genes with high H3K27ac

Table S4 Genes upregulated in young follicles relative to GSCs.

Table S5. 41 membrane transporters upregulated more than 10-fold in young follicles.

Supplementary Figure Legends

Figure S1. Growth profiles (total germ cell volume) of Drosophila ovarian follicles

A. Growth of wild type follicles shown by plotting volume (log10) determined as described in Methods vs integrated DAPI (log10). An example of follicles with GFP labeled germ cells is shown at the right. Stages of the follicles are indicated by color key on the right.. B. Similar methods applied to reveal the growth of control vs yolkless follicles. C. Similar methods used to determine the growth of control and Cap/Df follicles.

Figure S2. The Drosophila female germ cell cycle.

A. The Drosophila female germ cell generational cycle. Starting with adult germline stem cells (GSCs, magenta, upper right) germ cells form germline cysts (tan) that move through germarium regions 1-3 and bud off as stage 2 follicles (S2). Mature S14 oocytes are fertilized and undergo 14 preblastoderm syncytial divisions, with pole cells budding off at division 9 to form primordial germ

cells (PGCs). About 10 PGCs complete migration to the embryonic ovary, that develops new germaria prior to adulthood. The full cycle entails 19-20 germ cell divisions entailing an estimated >10,000-fold cytoplasmic volume increase.

Figure S3. Isolation of GSC and NC nuclei by FACs sorting.

A. The Drosophila female germ cell generational cycle. Starting with adult germline stem cells (GSCs, magenta, upper right) germ cells form germline cysts (tan) that move through germarium regions 1-3 and bud off as stage 2 follicles (S2). Mature S14 oocytes are fertilized and undergo 14 preblastoderm syncytial divisions, with pole cells budding off at division 9 to form primordial germ cells (PGCs). About 10 PGCs complete migration to the embryonic ovary, that develops new germaria prior to adulthood. The full cycle entails 19-20 germ cell divisions entailing an estimated >10,000-fold cytoplasmic volume increase.

Figure S4. Genes with high H3K27ac levels in GSCs cluster in heterochromatin and two

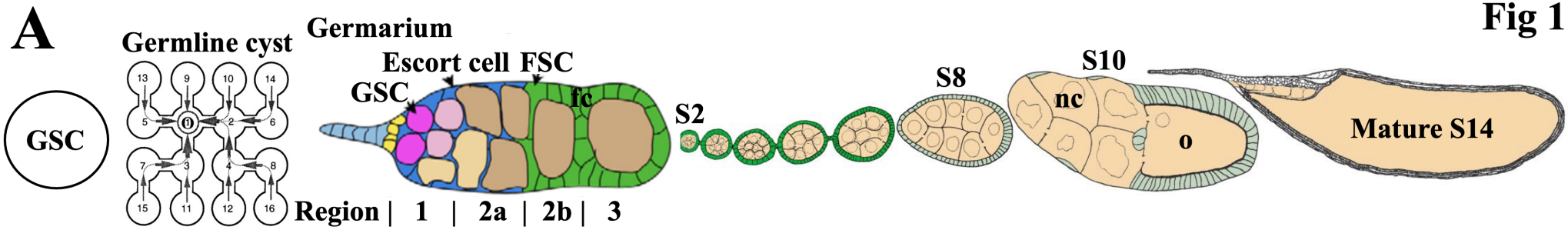
A. . Diagram of the Drosophila chromosomes showing zones of clustered genes with high H3K27ac that are expressed in GSCs and downstream germ cells. heterochromatin = blue boxes. numbers below line are coordinates in Mb. numbers above line are cytogenetic regions.

Figure S5. Analysis of chromatin changes at enhancers of Myc during oogenesis

A. myc region chromatin accessibility assayed by ATACseq in GSCs and 4C-32C nurse cells (NC). B. Genomic DNA fragments from this region (1-11, indicated) cloned and integrated within Janelia GAL4 lines (Jennet et al. 2012) are shown. C. A summary of the GFP expression driven by 8 tested clones fragments. D. Enlarged view of GFP expression from clone 8 (R10G10), which is similar to

clone 7 and is shown below. The pattern of expression is indicated to show low level (clones 7 and 8) and higher level (clones 9 and 10) GFP expression starting about stage 4.

Fig 1



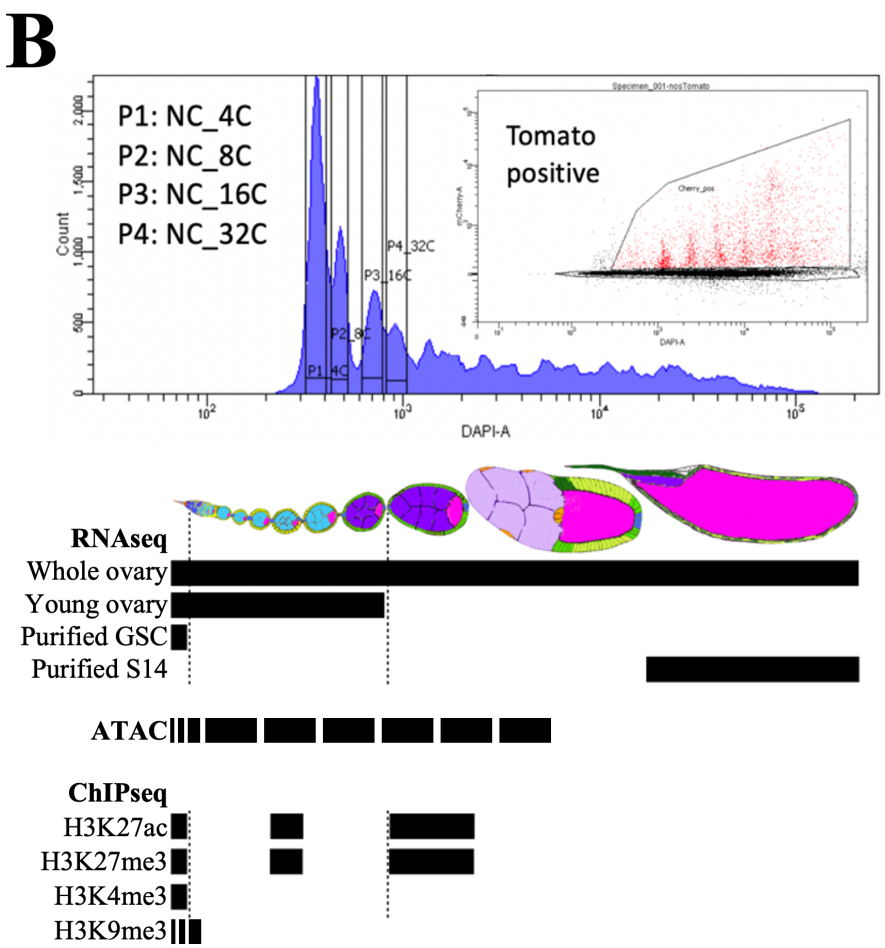
GSC: Open chromatin Bam-off Nanos-on.

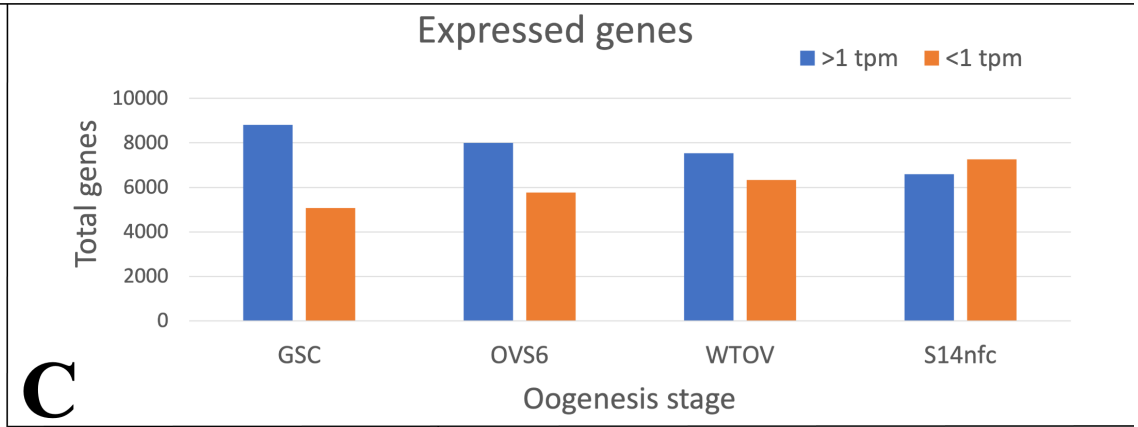
Cyst: 1st cell becomes oocyte (O), rest are nurse cells (nc) due to cyst polarity (arrows); Bam-on Nanos-off.

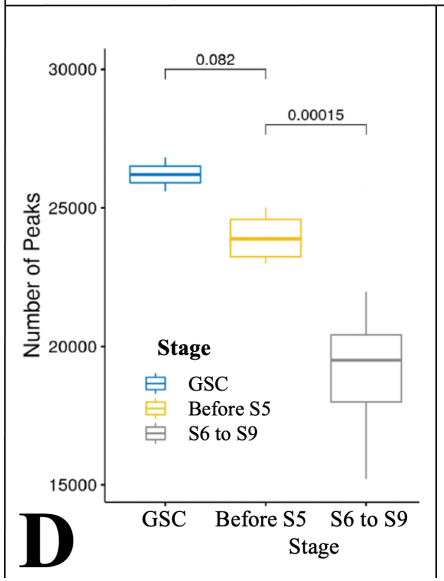
Germarium: region 1: cyst formation; 2a: oocyte, nurse cell specification, meiosis; 2b: organelle migration; 3: posterior MTOC, Balbiani body, follicle formation.

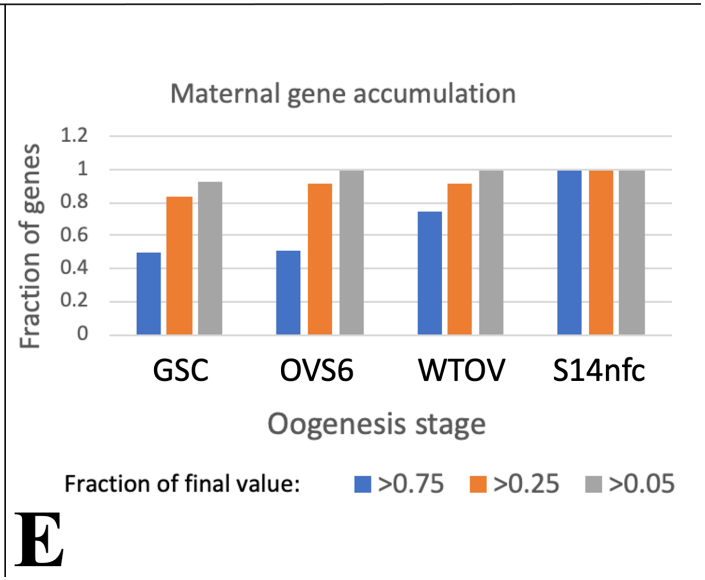
Follicle stages:
S2-S8: Rapid follicle growth; S8: growth checkpoint;
S8-10: vitellogenesis, growth ends, nc dumping;
S11-14: eggshell production, nc turnover;
Mature S14: fertilization, corpus luteum formation.

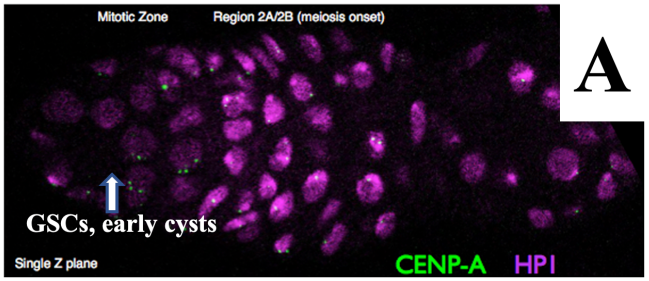
Key: escort cells: early somatic cells (blue); FSC: follicle stem cell generates somatic follicle cells (fc) (green); germline cells (tan).

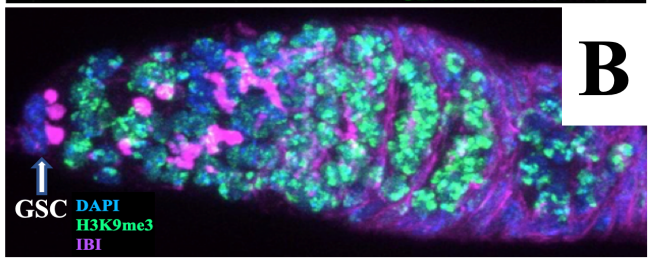


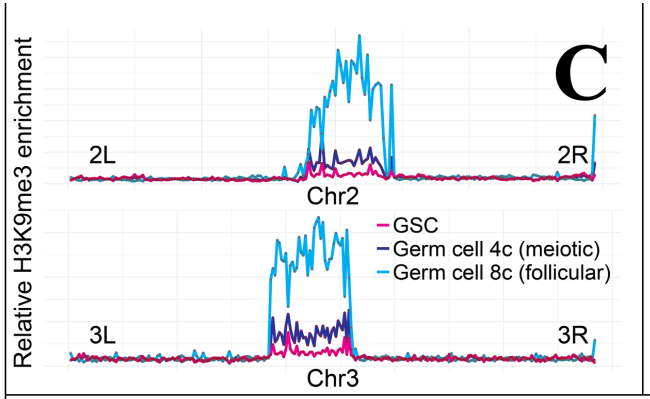


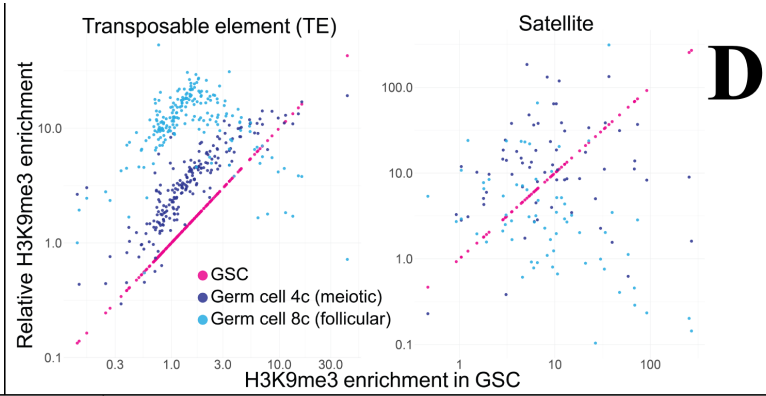


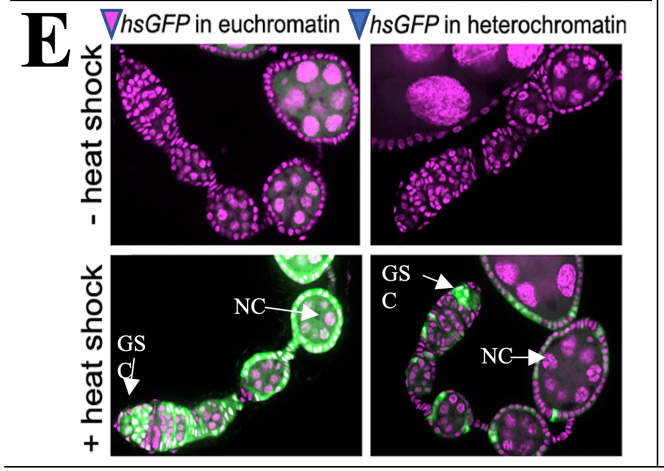


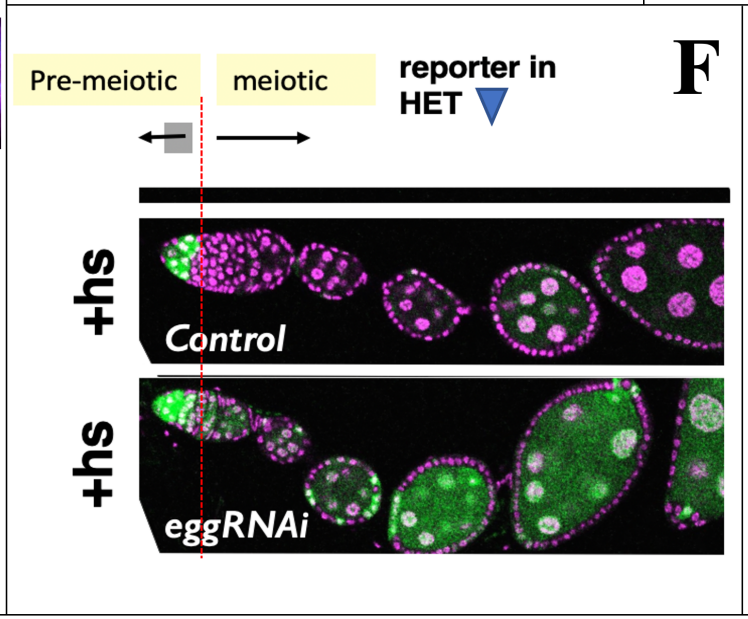

Mitotic Zone
Region 2A/2B (meiosis onset)
A
GSCs, early cysts
Single Z plane
CENP-A
HP1

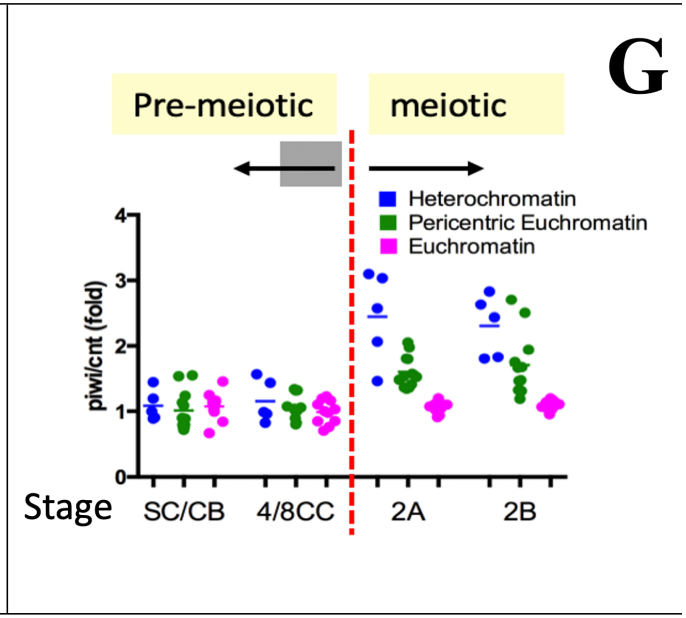

B
GSC
DAPI
H3K9me3
IBI


C
Relative H3K9me3 enrichment
2L
2R
Chr2
3L
3R
Chr3
GSC
Germ cell 4c (meiotic)
Germ cell 8c (follicular)


D
Transposable element (TE)
Satellite
Relative H3K9me3 enrichment
H3K9me3 enrichment in GSC
GSC
Germ cell 4c (meiotic)
Germ cell 8c (follicular)


E
hsGFP in euchromatin
hsGFP in heterochromatin
- heat shock
+ heat shock
NC
GSC
GSC
NC


F
Pre-meiotic
meiotic
reporter in HET
+hs
Control
+hs
eggRNAi


G
Pre-meiotic
meiotic
Heterochromatin
Pericentric Euchromatin
Euchromatin
piwi/cnt (fold)
Stage
SC/CB
4/8CC
2A
2B

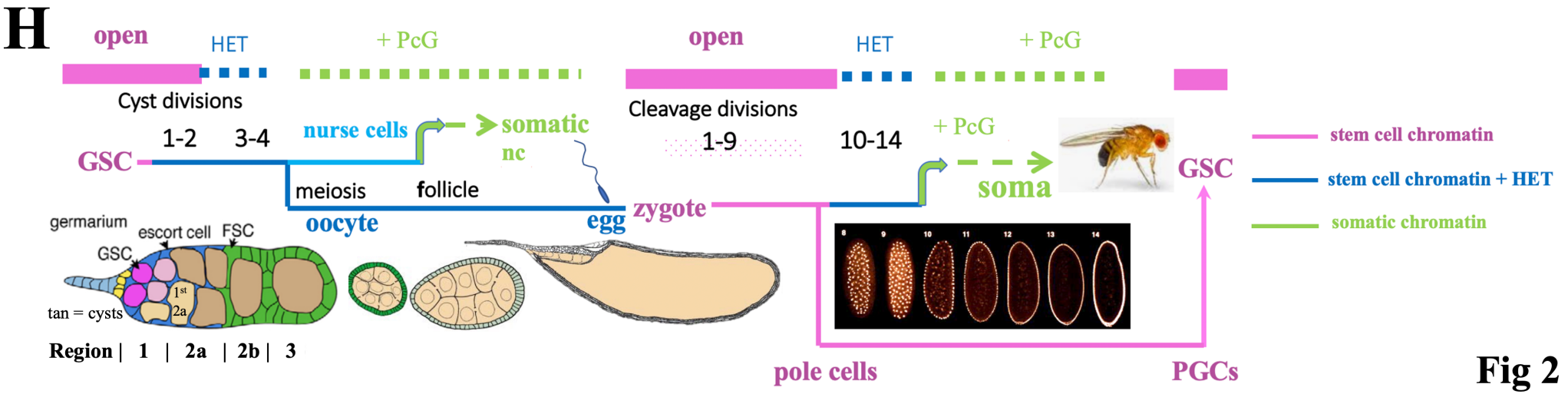

H
open
HET
+ PcG
Cyst divisions
1-2
3-4
nurse cells
somatic
nc
GSC
meiosis
follicle
oocyte
egg
zygote
open
HET
+ PcG
Cleavage divisions
1-9
10-14
+ PcG
soma
GSC
stem cell chromatin
stem cell chromatin + HET
somatic chromatin
germarium
escort cell
FSC
GSC
1st
2a
tan = cysts
Region | 1 | 2a | 2b | 3
pole cells
PGCs

Fig 2

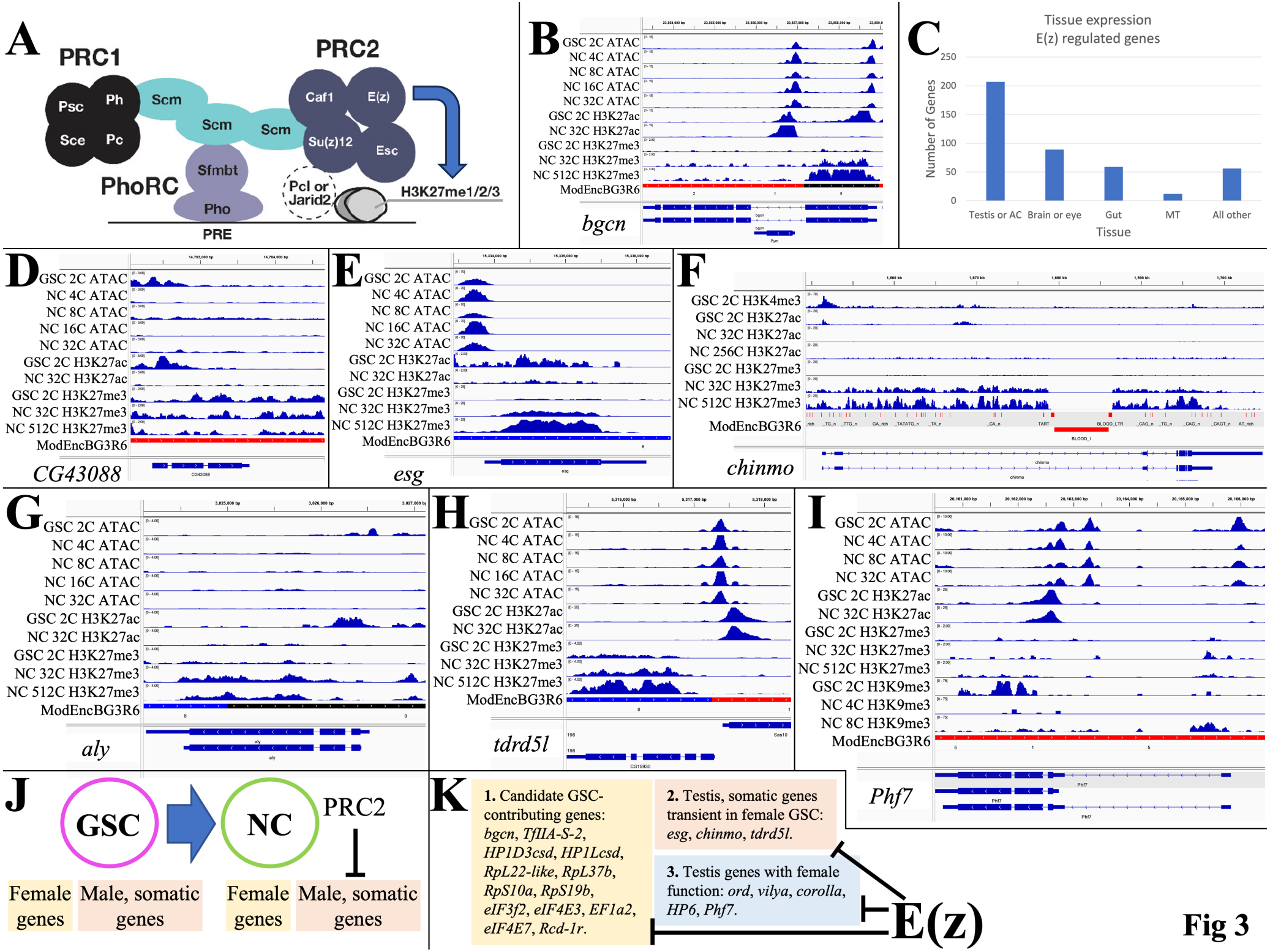

A
PRC1
PRC2
Psc
Ph
Scm
Caf1
E(z)
Sce
Pc
Scm
Scm
Su(z)12
Esc
Sfmbt
PhoRC
Pcl or Jarid2
H3K27me1/2/3
Pho
PRE
B
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 16C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
bgcn
C
Tissue expression
E(z) regulated genes
Number of Genes
250
200
150
100
50
0
Testis or AC
Brain or eye
Gut
MT
All other
Tissue
D
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 16C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
CG43088
E
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 16C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
esg
F
GSC 2C H3K4me3
GSC 2C H3K27ac
NC 32C H3K27ac
NC 256C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
chinmo
G
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 16C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
aly
H
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 16C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
tdrd5l
I
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
GSC 2C H3K9me3
NC 4C H3K9me3
NC 8C H3K9me3
ModEncBG3R6
Phf7
J
GSC
NC
PRC2
Female genes
Male, somatic genes
Female genes
Male, somatic genes
K
1. Candidate GSC-contributing genes: bgcn, TfIIA-S-2, HP1D3csd, HP1Lcsd, RpL22-like, RpL37b, RpS10a, RpS19b, eIF3f2, eIF4E3, EF1a2, eIF4E7, Rcd-1r.
2. Testis, somatic genes transient in female GSC: esg, chinmo, tdrd5l.
3. Testis genes with female function: ord, vilya, corolla, HP6, Phf7.
E(z)

Fig 3

Fig 4

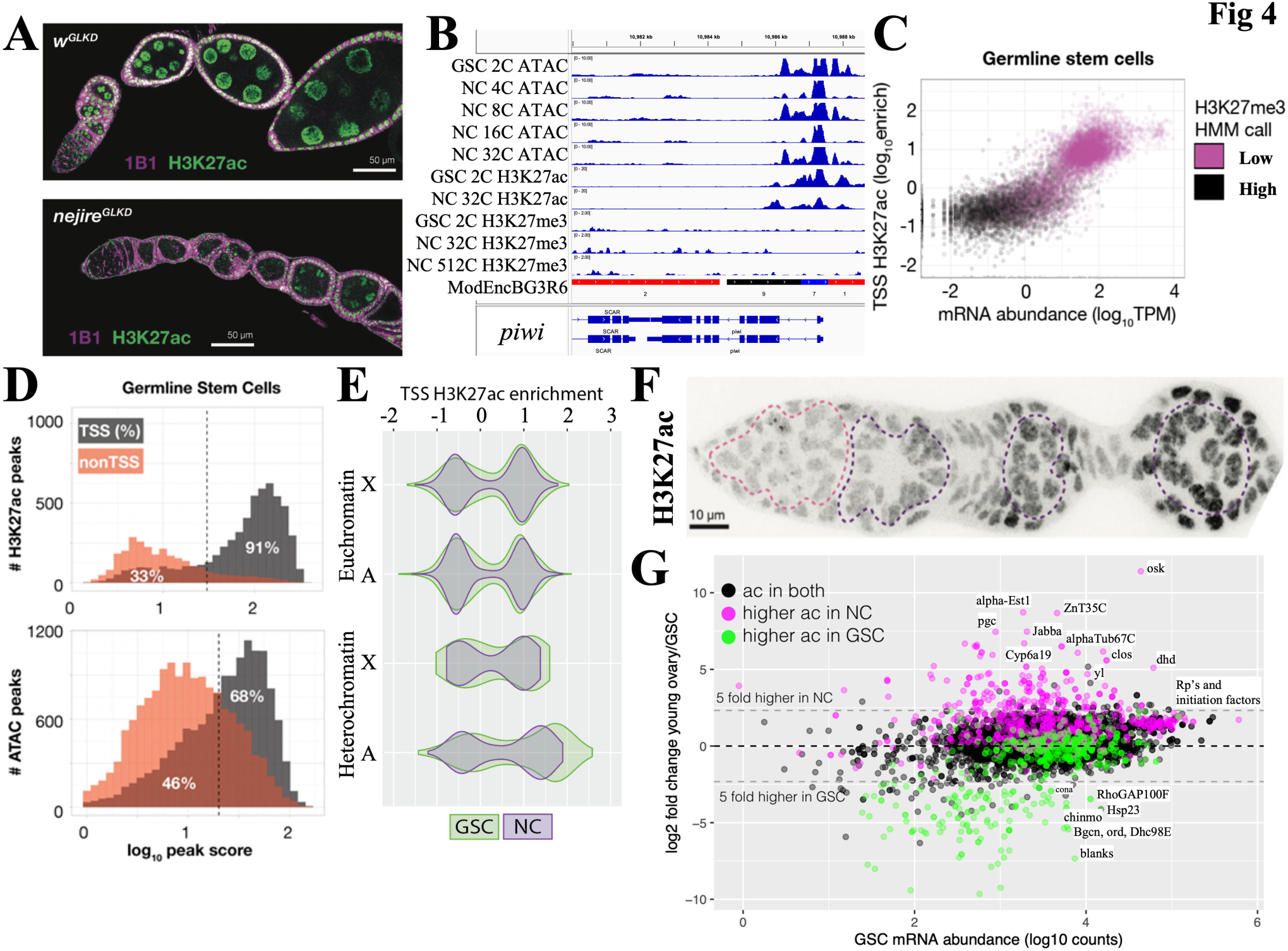

A
wGLKD
1B1 H3K27ac
50 µm
nejireGLKD
1B1 H3K27ac
50 µm
B
GSC 2C ATAC
NC 4C ATAC
NC 8C ATAC
NC 16C ATAC
NC 32C ATAC
GSC 2C H3K27ac
NC 32C H3K27ac
GSC 2C H3K27me3
NC 32C H3K27me3
NC 512C H3K27me3
ModEncBG3R6
piwi
SCAR
piwi
C
Germline stem cells
TSS H3K27ac (log10enrich)
mRNA abundance (log10TPM)
H3K27me3
HMM call
Low
High
D
Germline Stem Cells
TSS (%)
nonTSS
H3K27ac peaks
91%
33%
ATAC peaks
68%
46%
log10 peak score
E
TSS H3K27ac enrichment
Euchromatin
Heterochromatin
X
A
GSC
NC
F
H3K27ac
10 µm
G
ac in both
higher ac in NC
higher ac in GSC
5 fold higher in NC
5 fold higher in GSC
osk
alpha-Est1
ZnT35C
pgc
Jabba
alphaTub67C
Cyp6a19
clos
dhd
yl
Rp's and
initiation factors
cona
RhoGAP100F
Hsp23
chinmo
Bgcn, ord, Dhc98E
blanks
log2 fold change young ovary/GSC
GSC mRNA abundance (log10 counts)

Fig 5

A

Gene	NC/GSC_FC	Gene	NC/GSC_FC
pgc	>100	Hnf4	4.91
hb	>100	Trl	4.56
EcR	18.22	emc	4.01
cad	11.74	cwo	3.91
foxo	7.65	cbt	3.85
Mnt	7.09	retn	3.71
CHES-1-like	5.23	zfh1	3.41
Myc	5.13	scf	3.15

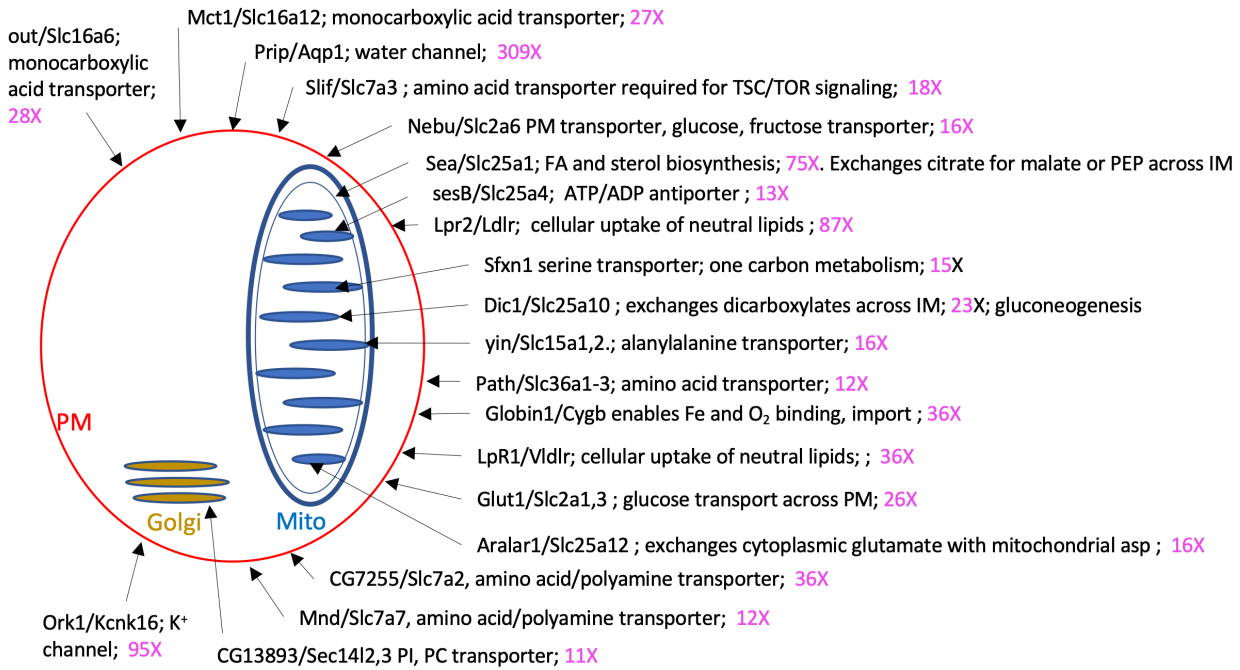
B

Description	LogP	Log(q-value)	In/Total entry
Transmembrane transport	-11.9983	-8.030	32/678
Metabolism of amino acids and derivatives	-7.50335	-4.314	14/198
Gluconeogenesis	-6.65012	-3.682	6/25
Small molecule metabolic process	-5.17649	-2.413	23/757
Fatty acid degradation	-4.57674	-1.888	5/33
Metabolism of lipids	-4.42276	-1.817	16/458
Nucleobase-containing small molecule metabolism	-4.27605	-1.688	12/279
Biosynthesis of amino acids	-4.02578	-1.489	6/68

C

Glycolysis gene		GSC	OVS6
Hex-A	Hexokinase A	31.5	114
Pgi	Phosphoglucose isomerase	176	113
pfk	Phosphofructokinase	54	34
Ald	Aldolase	61	302
Tpi	Triose phosphate isomerase	34	254
Gapdh1	Glyceraldehyde 3P dehydrog	301	545
Gapdh2	Glyceraldehyde 3P dehydrog	567	780
Pgk	Phosphoglycerate kinase	201	268
CG9961	Phosphoglycerate kinase	28	0.49
Pglym78	Phosphoglyceromutase	29	276
Eno	Enolase	361	525
PyK	Pyruvate kinase	344	315
Ldh (Impl3)	Lactate dehydrogenase	121	0.95

D



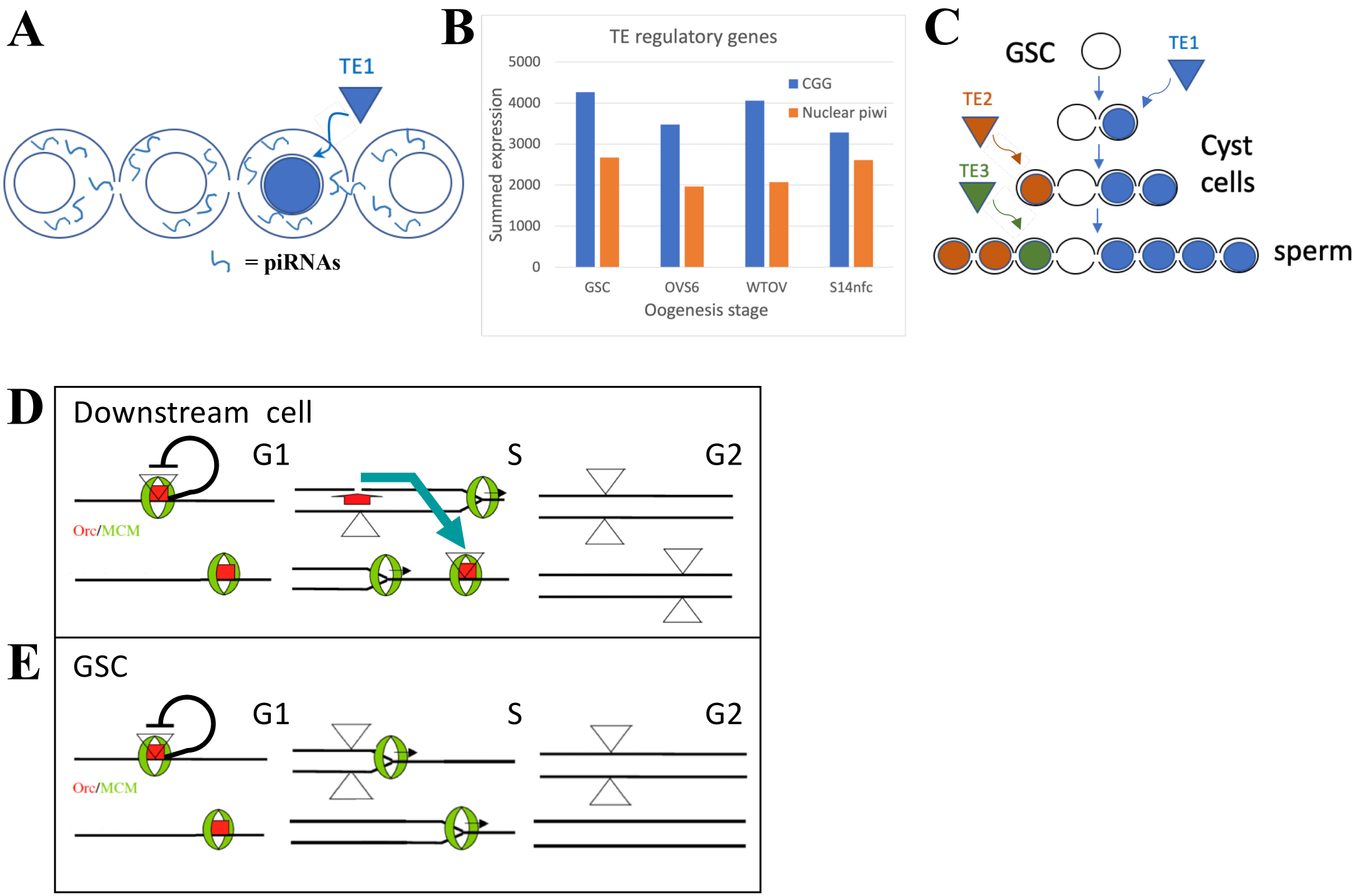

A
TE1
= piRNAs
B
TE regulatory genes
Summed expression
5000
4000
3000
2000
1000
0
CGG
Nuclear piwi
GSC
OVS6
WTOV
S14nfc
Oogenesis stage
C
GSC
TE1
TE2
TE3
Cyst cells
sperm
D
Downstream cell
G1
S
G2
Orc/MCM
E
GSC
G1
S
G2
Orc/MCM

Fig 6