HHS Public Access

Author manuscript

Dev Cell. Author manuscript; available in PMC 2021 January 06.

Published in final edited form as:

Dev Cell. 2020 January 06; 52(1): 38–52.e10. doi:10.1016/j.devcel.2019.11.007.

GCNA preserves genome integrity and fertility across species

Varsha Bhargava[1,*], **Courtney D. Goldstein**[1,*], **Logan Russell**[2,*], **Lin Xu**[3,4,*], **Murtaza Ahmed**[4], **Wei Li**[2], **Amanda Casey**[1], **Kelly Servage**[1,5], **Rahul Kollipara**[6], **Zachary Picciarelli**[2], **Ralf Kittler**[6,7], **Alexander Yatsenko**[2], **Michelle Carmell**[8], **Kim Orth**[1,5], **James F. Amatruda**[1,3,9,§,+], **Judith L. Yanowitz**[2,+], **Michael Buszczak**[1,10,11,+]

[1]Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

[2]Magee-Womens Research Institute, Department of Obstetrics, Gynecology, and Reproductive Sciences, University of Pittsburgh School of Medicine, Pennsylvania 15213, USA.

[3]Quantitative Biomedical Research Center, Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas, USA.

[4]Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, Texas, USA.

[5]Howard Hughes Medical Institute, 6000 Harry Hines Boulevard NA5.120F, Dallas, Texas, USA.

[6]McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

[7]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

[8]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA. RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, MA 01605, USA. Department of Biological Sciences, Wellesley College, Wellesley, MA, USA.

[9]Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas, USA.

[10]Center for Regenerative Science and Medicine, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

[+]Senior author. Correspondance: michael.buszczak@utsouthwestern.edu, yanowitzjl@mwri.magee.edu and jamatruda@chla.usc.edu.
[§]Current address: Departments of Pediatrics and Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, 90027 USA; and Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, CA 90033, USA.
[*]These authors contributed equally.

AUTHOR CONTRIBUTIONS.

Conceptualization, V.B., C.D.G., J.F.A., J.L.Y., M.B.; Methodology, V.B., C.D.G., J.F.A., J.L.Y., M.B.; Investigation, V.B., C.D.G., L.R., L.X., M.A., W.L., A.C., K.S., Z.P., J.L.Y., M.B.; Formal Analysis, L.X., R.K., R.K., A.Y. Resources, M.C., K.O., A.Y.; Writing-Original Draft, J.L.Y., M.B.; Writing-Review and Editing, V.B., C.D.G., K.O., J.F.A., J.L.Y., M.B.; Supervision, K.O., J.F.A., J.L.Y., M.B.; Funding Acquisition, J.F.A., J.L.Y., M.B.;

DECLARATION OF INTERESTS

The authors declare no competing interests.

[11]Lead Contact

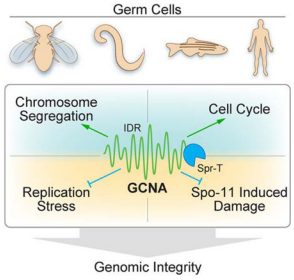
SUMMARY

The propagation of species depends on the ability of germ cells to protect their genome from numerous exogenous and endogenous threats. While these cells employ ubiquitous repair pathways, specialized mechanisms that ensure high-fidelity replication, chromosome segregation, and repair of germ cell genomes remain incompletely understood. We identified Germ Cell Nuclear Acidic Peptidase (GCNA) as a conserved regulator of genome stability in flies, worms, zebrafish and human germ cell tumors. GCNA contains an acidic intrinsically disordered region (IDR) and a protease-like SprT domain. In addition to chromosomal instability and replication stress, *Gcna* mutants accumulate DNA-protein crosslinks (DPCs). GCNA acts in parallel with the SprT domain protein Spartan. Structural analysis reveals that while the SprT domain is needed to limit DNA damage, the IDR imparts significant function. This work shows that GCNA protects germ cells from various sources of damage, providing insights into conserved mechanisms that promote genome integrity across generations.

Precis

Bhargava *et al.* identify GCNA as a critical, conserved germ cell factor that helps to prevent replication stress and DNA-protein crosslink accumulation. Loss of GCNA correlates with increased copy number variation in human germ cell tumors.

Graphical Abstract



INTRODUCTION

Early in the development of metazoans, primordial germ cells are set apart from somatic cells and undergo special programs to preserve genome integrity across generations. These programs include producing haploid gametes through meiosis, inducing and repairing programmed DNA double-strand breaks (DSBs), inhibiting transposable elements, and reprogramming chromatin to an epigenetic state that supports totipotency in the fertilized zygote (Kurimoto and Saitou, 2018; Tang et al., 2016). Facilitating these processes are a subset of germ cell-specific proteins that have been conserved across millions of years, including the DEAD-box helicase VASA, the RNA-binding protein NANOS, and the piRNA processing enzyme PIWI/Argonaute.

The recently identified Germ Cell Nuclear Acidic Peptidase (GCNA), also known as Germ Cell Nuclear Antigen or Acidic Repeat Containing (ACRC), has been conserved across 1.5 billion years of evolution. GCNA has remained tightly associated with sexual reproduction showing enriched expression within germ cells in both invertebrate and vertebrate species (Carmell et al., 2016). GCNA proteins contain an N-terminal acidic intrinsically disordered region (IDR) which is conserved structurally despite amino acid divergence. In most species, but not rodents, GCNA proteins also contain a C-terminal SPARTAN (SprT)-domain which resembles a bacterial metalloprotease, a Zinc finger (ZnF), and an HMG box. Despite this modular domain structure, insights into GCNA function are lacking.

IDR-containing proteins have emerged as important players in cell biology, regulating phase transitions in a number of membrane-less organelles (Banani et al., 2017). In the nucleus, IDR proteins help form nucleoli, speckles, and Cajal bodies. All of these condensates are thought to be macromolecular assembly sites for protein-nucleic acid complexes that control chromatin structure, transcription, and various aspects of RNA processing. IDR proteins are also found in numerous germ cell-specific structures (Seydoux, 2018). For example, the IDR-containing MUT-16 protein phase separates to form mutator bodies in worms (Uebel et al., 2018). VASA also contains an extensive disordered region which contributes to its molecular behavior and function (Nott et al., 2015). Similarly, *C. elegans* MEG-3 and MEG-4 proteins bind to and phase separate RNA to form granules both *in vitro* and *in vivo* (Smith et al., 2016). MEG-3 and MEG-4 are GCNA family members, raising the possibility that GCNA itself may mediate essential germline functions through its IDR.

Potential insight into GCNA function comes from recent investigation into the functions of Spartan proteins for which the SprT-domain gets its moniker. Several independent groups have provided evidence that Spartan proteins specifically cleave DNA-protein crosslinks (DPCs) through their SprT protease domain (Lopez-Mosqueda et al., 2016; Stingele et al., 2016; Vaz et al., 2016). DPCs represent particularly insidious lesions that interfere with almost every chromatin-based process including replication, transcription, and chromatin remodeling (Stingele et al., 2017). The protease activity of Spartan appears highly regulated, and one major target of Spartan proteolysis is Spartan itself. Loss of Spartan in humans and mice results in sensitivity to UV damage, progeroid-like phenotypes, and a predisposition to hepatocellular carcinoma, suggesting the protein plays an essential role in maintaining genome integrity (Davis et al., 2012; Mosbech et al., 2012). Within the germ line, the topoisomerase-like enzyme Spo11 forms DPCs to drive the formation of meiotic DSBs (Keeney et al., 1997). DPCs also occur during mitotic and meiotic DNA replication through the activity of topoisomerases. In addition, epigenetic reprogramming, including histone demethylation, creates cross-linking by-products like formaldehyde (Walport et al., 2012) that can result in DPC formation (Stingele et al., 2017). Inability to remove these DPCs would interfere with the faithful transmission of the genome over generations.

Here, we provide evidence that loss of *Gcna* results in genomic instability in *Drosophila*, *C. elegans*, zebrafish, and human germ cell tumors. GCNA acts to limit Spo11 activity in flies and prevents replication stress in flies and worms. Further analysis shows that GCNA functions in parallel to Spartan proteins within germ cells. Loss of *Gcna* results in the accumulation of DPCs in germ cells and early embryos. Genetic and transgenic analysis

points to distinct roles for the IDR and SprT domains of GCNA. Together, these results reveal a new mechanism by which germ cells ensure the integrity of their genomes from one generation to the next.

RESULTS

Drosophila Gcna mutants exhibit genome instability and chromosome segregation defects

Gcna exhibits enriched expression in germ cells across species (Carmell et al., 2016). We turned to *Drosophila*, *C. elegans*, and zebrafish as model systems in which to characterize the molecular function of GCNA family members. The *Drosophila* genome encodes for three potential GCNA orthologs (Figures 1A,B; S1). Only null mutations in *CG14814* (hereafter called *Gcna*, Fig 1B), resulted in overt phenotypes, while the others appeared viable and fertile (Figure S1A–E). Homozygous $Gcna^{KO}$ mutant females initially laid eggs but stopped after approximately one week. Many of the embryos derived from these eggs exhibited maternal-effect lethality (Figure S1F). Fixed and live-cell imaging experiments revealed that loss of maternal Gcna resulted in numerous mitotic defects during early embryogenesis including chromosome tangling, micronucleus formation, nuclear fusion, chromosome segregation defects, and disruption of cell-cycle synchrony (Figure 1C–E; Movie S1).

The few adult, F1 progeny from *Drosophila* $Gcna^{KO}$ mutant females appeared sickly and sub-fertile. Of these, four percent displayed bilateral gynandromorphism (Figures 1F; S1G) a rare phenotype caused by X chromosome loss during the early embryogenesis (Janning, 1978). Chromosome loss was not X chromosome-specific as shown by fluorescent *in situ* hybridization (FISH). In control female embryos, two discrete X-chromosome foci (359-bp repeat probes) and two chromosome 2 foci ($AACAC_{(n)}$ probes) were observed in dividing nuclei, as expected for a diploid cell (Figure 1G). By contrast, embryos from $Gcna^{KO}$ mutant females displayed X-chromosome bridges and second chromosome missegregation and polyploidy (Figure 1G). Similar chromosomal phenotypes were also observed in *Gcna* mutant ovaries (Figure S1H).

Drosophila Gcna mutant ovaries exhibit increased DNA damage

Loss of fly *Gcna* also resulted in specific ovarian phenotypes (Figure 1H,I; S2A–D). For example, many *Gcna* mutant egg chambers deviated from the normally-invariant number of 16 germ cells per cyst. In aged flies, this phenotype became more penetrant. Labeling of ring canals, which form from arrested cleavage furrows, and the cell death marker Cleaved Caspase 3 suggested that the counting defects arose from abnormal cell divisions and loss of cell-cycle synchrony (Figure S2A, data not shown). We also observed defects in differentiation and delayed oocyte specification (Figure S2B,C).

To test whether *Gcna* regulates meiosis in *Drosophila*, we examined γH2Av, a marker of DNA damage (analogous to γH2AX), and C(3)G, a synaptonemal complex (SC) protein (Jang et al., 2003; Page and Hawley, 2001). In control ovaries, the SC and γH2Av foci were first observed in region 2A of the germarium, as previously described (Jang et al., 2003). In

Gcna mutant germaria as revealed by structured illumination microscopy (SIM), the SC formed normally; however, γH2Av foci appeared larger and more abundant (Figure 1J–L).

In control germaria, DSBs are rapidly repaired. However, in *Gcna* mutant germ cells, γH2Av staining extended into early egg chambers (Figure 1K). Expression of a *p53* reporter, which correlates with DNA damage (Lu et al., 2010; Wylie et al., 2014), was also expanded (Figure S2D). To determine if the persistent DNA damage reflects a defect in DSB repair, we asked whether *Gcna* mutations disrupt homologous recombination (HR). Unlike HR-defective *Rad51* and *Rad54* mutants, which exhibit dorsal appendage defects in 50–60% of their eggs (Abdu et al., 2003; Ghabrial et al., 1998; Staeva-Vieira et al., 2003), only 4–8% of *Gcna* mutant eggs have this phenotype (Figure S2E; n>200 eggs from multiple lays). *Gcna* mutant meiotic nuclei also have more γH2Av foci than HR pathway mutants (Figure S2F; (Jang et al., 2003)). *Gcna* mutant flies also do not display whole body sensitivity to irradiation (IR, Figure S2G), unlike *Rad51* and *Rad54* mutants (Ghabrial et al., 1998; Staeva-Vieira et al., 2003). Together with our observation that meiotic nondisjunction rates did not vary between controls and $Gcna^{KO}$ homozygotes (Figure S2H), these results suggest that Gcna likely does not play an essential role in HR-mediated repair in *Drosophila*.

We next considered that the excess γH2Av foci in *Gcna* mutant ovaries resulted from disruption in spatial or temporal regulation of meiotic DSB induction by Spo11. If correct, this model would predict that loss of either *spo11*, named *mei-W68* in *Drosophila* (McKim and Hayashi-Hagihara, 1998), or *mei-P22*, a gene needed for targeting mei-W68 (Liu et al., 2002), would suppress *Gcna* mutant phenotypes. Loss of either *mei-W68* or *mei-P22* functions in the *Gcna* mutant background resulted in a dramatic suppression of the extra meiotic DSBs in region 2A (Figures 1K,L; S2I). In these double mutants, however, we still observed γH2Av staining in pre-meiotic and later germ cells suggesting a subset of breaks arise independently of the meiotic program. In addition, neither *mei-P22* nor *mei-W68* mutations suppressed other phenotypes associated with loss of *Gcna*, including the germ-cell counting defects and maternal-effect semi-lethality. Thus, it appears that GCNA functions in flies to maintain various aspects of genome integrity during both meiotic and mitotic divisions.

C. elegans gcna-1 mutants exhibit genomic instability in later generations

In parallel with the *Drosophila* experiments described above, we made a null mutation in the *C. elegans Gcna* homolog, CELE_ZK328.4 (now called *gcna-1*) (Figure 2A,B). A previous large-scale RNAi screen noted that *gcna-1* knockdown resulted in a very mild High Incidence of Males (HIM) phenotype (Colaiacovo et al., 2002), which is indicative of X chromosome nondisjunction (Hodgkin et al., 1979). Our CRISPR/Cas9-induced null allele confirmed this mild HIM phenotype, which is exacerbated by growth at higher temperature (25°C) and at later generations (Figure 2C). Moreover, we found that loss of *gcna-1* gives rise to a mortal germ line (MRT) phenotype characterized by transgenerational loss of fecundity and vitality, marked by reduced lifespan, decreased mobility, and loss of fertility in later generations (Figure 2D; Figure S3A,B).

Whereas wild-type worms contain two U-shaped germlines filled with developing oocytes that ultimately arrest at diakinesis of prophase I with 6 bivalent chromosomes, late

generation *gcna-1* mutant germ lines showed a range of phenotypes, from near wild-type to severely runty germ lines (Figure S3C). Diakinesis nuclei of late generation *gcna-1* mutant animals showed chromosomal abnormalities with 4 – 9, often irregularly-shaped, DAPI-stained bodies (Figure 2E), indicative of chromosome fusions and defects in crossover formation (Dernburg et al., 1998; Hillers et al., 2017). Multiple independent lines began to produce excessive male offspring in the several generations before the onset of sterility. In these populations, all worms assayed presented with a consistent karyotype of 5 DAPI-positive bodies suggesting they may have contained X:autosome fusion chromosomes. Similar to the fly embryo, we observed chromosome bridges and chromosomal fragments in *gcna-1* mutant germ cells (Figure 2F). Together, the fly and worm mutant phenotypes indicate that loss of *Gcna* function disrupts reproductive success and chromosome stability across species.

Excessive DNA damage during meiosis was not obvious in *C. elegans gcna-1* mutants (Figure S3D). By crossing *gcna-1* into a *spo-11* mutant background and exposing worms to the DSB-inducing agent IR, we were able to see similar accumulation of RAD-51 in *spo-11* and *gcna-1;spo-11* (Figure S3E), suggesting that early steps in HR are normal in *gcna-1* mutant animals, as they are in *Drosophila Gcna* mutants. Consistent with this observation, *gcna-1* mutant worms did not show increased IR sensitivity (Figure S3F). Nevertheless, RAD-51 foci were present in the pachytene nuclei of the unirradiated *gcna-1;spo-11* mutant controls (Figure 2G) suggesting DNA damage arose during either mitotic divisions of germ cells or meiotic S phase. This “carry through” damage can induce meiotic crossover (CO) formation as seen by a decrease in univalent chromosomes in *gcna-1;spo-11* compared to *spo-11* mutant worms (Figure 2H). Together with the *Drosophila* experiments, these findings indicate that while GCNA plays species-specific roles in the regulation of Spo11 activity during meiosis, loss of *Gcna* function leads to the accumulation of Spo11-independent DSBs within both *Drosophila* and *C. elegans* germ cells.

Replication stress as a source of DNA damage in *Drosophila* and *C. elegans*

One source of Spo11-independent DSBs are transposable elements (TE) whose mobilization creates DSBs that can induce crossovers (Soper et al., 2008). We carried out experiments to determine if GCNA controls TE surveillance (Figure S4). We saw a modest increase in TE expression in worms (Figure S4A) which suggested a potential role for GCNA in the TE surveillance pathway governed by piRNAs/PIWI. However, *gcna-1* mutations were synthetic sterile with *prg-1*/PIWI mutations (Figure S4B), arguing that GCNA acts in parallel to the canonical TE surveillance pathway. Consistent with these observations, *Gcna* mutant flies did not show altered expression or localization of Aubergine, a piRNA pathway component (Figure S4C). Transcriptomic analyses identified an approximate 2-fold increase in expression of telomere-associated TEs and a concomitant decrease in metabolic genes in *Drosophila Gcna* mutants (Figure S4D–H), raising the possibility that cellular stress pathways may be activated in *Gcna* mutants. These cross-species studies hint at mild effects on TE regulatory pathways, but are unlikely to explain the substantial increase in DNA damage and phenotypic consequences observed in *Gcna* mutants.

We next considered the possibility that *Gcna* mutants experience replicative stress, which is a major cause of endogenous DSBs, chromosome segregation defects, and cell cycle disruption (Magdalou et al., 2014; Zhang et al., 2018). To determine if GCNA participates in replicative repair, we took advantage of an established assay in worms to interrogate microsatellite repeat stability. The DEAD-box helicase *dog-1/FANCJ* is the only known gene required for replication of G-quadraplex-like structures in worms (Kruisselbrink et al., 2008) and is required for accurate replication through GC-rich DNA (Youds et al., 2008). Accordingly, *dog-1* mutations exhibit microsatellite repeat instability (MSI), as readily seen in a PCR-based assay of repeats at the *vab-1* locus ((Youds et al., 2008); Figure S5A). Whereas *gcna-1* mutation did not exhibit repeat instability on its own, it strongly enhanced *dog-1* (Figure 3A). This result is consistent with a role for GCNA in promoting replicative repair and/or replication restart.

When replication forks stall at DNA lesions or aberrant DNA structures, increased stretches of single-stranded DNA (ssDNA) form, which can be visualized as nuclear punctae of RPA (replication protein A), the major single-strand DNA binding protein in cells. In *Drosophila* germ cells, we observed a disruption of RPA expression and localization within *Gcna* mutant ovaries. $Gcna^{KO}$ mutant germ cells displayed discrete RPA nuclear foci that were rarely observed in controls (Figure 3B,C; S5B). Within *Gcna* mutant ovaries, these punctae were present in both mitotic germ cells within germaria and in endocycling nurse cells.

We also tested whether *Drosophila Gcna* mutant germ cells were sensitive to hydroxyurea (HU), an agent that limits the dNTP pools, causing replication fork stalling and DNA damage, particularly in mutants that already suffer from replication stress. Upon HU treatment, *Gcna* mutant germaria accumulated many more γH2Av foci within mitotically active cells relative to controls, indicating that loss of GCNA makes germ cells more sensitive to replicative stress (Figure 3D). In the associated manuscript, the authors found the *gcna-1* mutant worms were also mildly sensitive to HU (Davis et al., 2019). Taken together, these results support a potential role for GCNA in promoting replicative repair across species.

GCNA and Spartan act independently

DNA-protein crosslinks are a significant source of replicative stress (Vaz et al., 2017). GCNA is a Spartan family member, whose namesake removes DPCs (Stingele et al., 2016; Vaz et al., 2016). To determine whether GCNA has a similar function and/or acts together with Spartan, we crossed *Gcna* mutations into *Spartan* mutant backgrounds. In worms, Spartan is encoded by *dvc-1* (DNA damage-associated VCP/p97 Cofactor homolog). The previously described allele is a partial truncation that functions as a mutator, is slow growing and difficult to maintain (Stingele et al., 2016). We generated a full deletion allele, *dvc-1(ea65)* and maintained the stock as a balanced heterozygote. (Figure 3E). *dvc-1(ea65)* homozygotes conferred the reduced brood sizes (Figure 3F) described for *dvc-1(ok260)*. While *dvc-1* and *gcna-1* have near similar affect on brood size, *gcna-1;dvc-1* showed an additive effect, almost doubling embryonic lethality (Figure 3F; n>6 broods/ genotype, 3-way Kruskal-Wallis Test, $p<0.05$). Similarly, diakinesis nuclei of *dvc-1(ea65)* F3 homozygous animals contain aberrant chromosome numbers and morphologies that were

further exacerbated by loss of *gcna-1* (Figure 4G). In *dvc-1* single mutants, chromosome fusions and fragments were seen in 33% and 14% of nuclei, respectively (n= 21; note that some nuclei have both fusions and fragments). In the *gcna-1;dvc-1* double mutants, the frequency of fusions was not statistically different (n= 26; Chi-square, $p>0.1$) yet chromosomal fragments were much more common (62% of nuclei; Chi-Square, $p < 0.001$). These results suggest that GCNA-1 acts independently of DVC-1 to ensure genomic stability.

For *Drosophila*, the Spartan homolog, *maternal haploid* (*mh*), helps maintain paternal chromosome integrity during early embryogenesis (Delabaere et al., 2014). The number of eggs laid per female appeared comparable for *mh* and *Gcna* single mutants and *Gcna mh* double mutants. Of these embryos, 11% from *Gcna* mutant females (n=299) and 1.4% from mh^1 mutant females (n=444) hatched, while none of the embryos from *Gcna mh* double mutants (n=296) completed embryogenesis. Ovaries of *mh* mutants looked comparable to controls, whereas ovaries from *Gcna mh* double mutants exhibited increased γH2Av and RPA foci relative to control and single mutants (Figure S5C,D). These results suggest that Gcna and Mh likely act in parallel to limit DNA damage in the *Drosophila* germline, consistent with our worm data that Spartan and GCNA-1 act independently. These data raise the possibility that GCNA may have an independent role in clearing a subset of DPCs.

Loss of *Gcna* results in an accumulation of DPCs

We used the rapid approach to DNA adduct recovery (RADAR) coupled with SDS-PAGE and silver staining to evaluate whether GCNA influences DPC levels within *Drosophila* ovaries and early embryos (Figure 4A) (Kiianitsa and Maizels, 2013). We observed a modest, but consistently elevated, level of DPCs in the *Gcna* mutant ovaries and early embryos compared to controls (Figure 4B).

To identify proteins that formed DPCs, we analyzed the RADAR samples using mass spectrometry (Figure S6A; Table S1). A number of proteins formed DPCs specifically in *Gcna* mutant embryos or ovaries, or both; these included Histone H2B, Topoisomerase 2 (Top2), SSRP1, MCM3 and Fibrillarin. These proteins were not detected in the wild-type samples although other proteins were, pointing to the specificity of the described effects. We repeated the RADAR assay on isolated nuclei from *Drosophila* embryos and found an enrichment of specific DPCs in *Gcna* mutant samples (Figure 4C; Table S2). The increased levels of Top2 and MCM protein DPCs in the *Gcna* mutant drew our attention. Top2 decatenates supercoiled DNA, while the MCM complex unwinds DNA in front of the replication fork. To begin to test the functional interactions between Gcna and these proteins, we immunoprecipitated a tagged form of Gcna from S2 cells and performed mass spectrometry (Figure 4D; Table S3). Several, but not all, proteins found in DPCs in *Gcna* mutants also physically interact with Gcna protein. These included Top2, as well as multiple components of the MCM complex. These results further link GCNA with regulation of replication in germ cells and early embryos.

To characterize how *Gcna* loss affects the interacting proteins, we examined Top2 expression and localization (Figure 4E). In control embryos, Top2 localized to metaphase chromosomes, as previously reported (Tang et al., 2017). Low levels of cytoplasmic Top2

were also evident. By contrast, *Gcna* mutant embryos displayed a redistribution of Top2, with higher levels on metaphase chromosomes and decreased cytoplasmic localization. *Gcna* mutant egg chambers also exhibited more nuclear Top2 punctae (Figure S6B–D). This redistribution occurred independent of Top2 expression levels which were unchanged in *Gcna* mutants (Figure S6E). Finally, similar increases in chromosome-associated Top2 were observed in worm *gcna-1* mutant germ cells (Figure 4F), indicating that loss of *Gcna* results in increased accumulation of nuclear Top2 across species. Thus, we have demonstrated that GCNA both interacts with and impacts the localization of proteins that become associated with DPCs when *Gcna* function is impaired.

Characterization of *Drosophila* and *C. elegans* GCNA domain function

The accumulation of DPCs in *Gcna* mutants led us to explore the functional contribution of the conserved metallopeptidase zinc-binding SprT domain and IDR to GCNA function. To this end, we made lines that carried a series of *Drosophila* transgenes or tagged endogenous loci (Figure 5A, S6A, B). The Gcna-GFP transgene localized predominantly in cytoplasmic punctae, although discrete nuclear punctae could also be detected. Some foci appeared enriched around the nuclear envelope, but were distinct from the nuage. During mitosis, Gcna-GFP is associated with dividing chromosomes, in addition to its cytoplasmic localization (Figure 5B). The endogenously HA-tagged Gcna protein also exhibited predominantly cytoplasmic localization and association with mitotic chromosomes in ovaries and early embryos (Figure S7).

To test the functionality of the SprT domain in *Drosophila* Gcna, we inserted two different UAS-driven HA-tagged transgenes into the same genomic locus: wild-type Gcna and a HE>AA mutant that altered key residues needed for enzymatic activity, based on the characterization of Spartan proteins (Figure 5C) (Morocz et al., 2017). These transgenes showed similar cytoplasmic and weak nuclear punctae as described above. When expressed in a *Gcna* mutant background using a *nanos-gal4* driver, the HE>AA protein localized to slightly larger cytoplasmic punctae and was more broadly expressed in late stage nurse cells than wild type protein (Figure 5D). Western blot analysis showed that the HE>AA protein was expressed a slightly higher levels than the wild-type transgene (Figure S6C). Together, these experiments suggest that the SprT domain may regulate the size and number of Gcna-labeled condensates and/or overall Gcna protein levels.

Next, we tested the functionality of the *Drosophila* HA-tagged transgenes. The full-length *Gcna* transgene suppressed the excessive formation of both Spo11-dependent and independent breaks that we observed in $Gcna^{KO}$ mutant germ cells (Figure 5E, F). As expected, the HE>AA transgene did not rescue these phenotypes to any appreciable degree. However, the HE>AA transgene partially rescued the maternal-effect semi-lethality of embryos derived from *Gcna* mutant females (Figure 5G). For both transgenes, rescue of the maternal-effect semi-lethality was accompanied by a decrease in chromosome bridges and other mitotic defects based on DAPI labeling. Thus, the HE>AA transgene demonstrates a partial separation-of-function, revealing a requirement for the SprT domain for DNA damage prevention in the fly germ line, while the IDR domain promotes chromosomal stability during early embryogenesis.

We tagged the endogenous *C. elegans gcna-1* locus at the 5' end with either OLLAS (OmpF Linker and mouse Langerin fusion Sequence) or HA tags. The majority of OLLAS::GCNA-1 and HA::GCNA-1 proteins localized to the cytoplasm under steady-state conditions (Figure 5H; S7D). GCNA-1 proteins are maternally-inherited and are ultimately enriched in the primordial germ cell precursors, Z2 and Z3 (Figure S7E) which are readily identified by their size and position. In adult germ cells, chromosome squashes allowed us to see small puncta associated with chromosomes in the nucleus (Figure 5H; S7F), similar to what we observe in *Drosophila*.

We also made a mutant allele, *gcna-1(ea76)* which truncates the C-terminus (of the wild-type protein to retain just the IDR region (Figure 5I). While the most robust phenotype observed with the null allele, *gcna-1(ea43)*, was reduced fecundity that began at the F3 generation and became more severe in later generations (Figure 2D; 5J), *gcna-1(ea76)* had brood sizes greater than or equal to wild type in the F3 and F8 generations. This increase in brood size continued for >15 generations although a subset of the population started to exhibit phenotypes associated with *gcna-1* loss, including a HIM phenotype and sterility. Thus, while loss of catalytic function ameliorates the brood size decline seen in *gcna-1* null mutant animals, the catalytic domain and C-terminus of the protein are required to prevent genome instability across generations. Together with the fly experiments, these data indicate that GCNA is a multi-functional protein with the IDR and SprT domain governing distinct aspects of genome integrity.

Conservation of GCNA function in zebrafish

To characterize *Gcna* function in a vertebrate, we generated *gcna* mutant alleles in zebrafish including a 7 bp deletion (*mut1*) and a complex insertion of 9 bp and 11 bp (*mut2*) both of which lead to early frameshifts, predicted to result in severely truncated proteins (Figure 6A,B). The progeny of *gcna* mutant females displayed widespread morphological defects and cell death (100% penetrant; n>100 embryos) (Figure 6C,D). Close examination of these early embryos revealed asynchronous mitotic divisions and tangled chromosomes, in contrast to wild-type controls (Figure 6E). These results indicate that disruption of zebrafish *gcna* results in maternal-effect lethality marked by chromosome instability remarkably similar to the phenotypes observed in flies and worms. To test whether *gcna* loss also resulted in elevated DPC levels across species, we performed RADAR on zebrafish embryos. This analysis showed that disruption of *gcna* resulted in modestly increased levels of DPCs (Figure 6F).

Loss of *GCNA* correlates with genomic instability in human germ cell tumors

Our work in flies, worms, and zebrafish indicates that GCNA regulates genome stability across species. Whereas in humans, germ cells are largely inaccessible, germ cell tumors provide a window into the genes that control genome integrity. We therefore performed whole-exome and targeted deep sequencing, SNP array, DNA methylation array, and RNA sequencing on a cohort of 233 patients with pediatric germ cell tumors (GCTs, Table S4), and conducted an integrated analysis of the data analyzed (Xu and Amatruda, in preparation). To investigate tumor suppressor genes that are frequently silenced by copy number (CN) loss (based on SNP array data) and promoter hypermethylation (base on DNA

methylation array data) in GCTs, we examined 94 of 233 GCTs in our cohort that were processed by both array technologies. In these 94 GCTs, we calculated the percentage with either CN loss, promoter hypermethylation, or both, for *GCNA* and 441 known cancer genes (from Catalogue of Somatic Mutations in Cancer (COSMIC) database). Interestingly, we found *GCNA* has the highest alteration frequency (66%, 62 out of 94 cases) among all 442 genes studied here (Figure 7A,B), suggesting *GCNA* to be a top candidate for a tumor suppressor involved in GCT development. No somatic protein-altering mutation was found in *GCNA* gene through whole-exome and targeted deep sequencing analysis in 137 GCT cases.

As expected, CN loss and/or promoter hypermethylation in *GCNA* was correlated with significantly lower *GCNA* expression in tumors, compared to GCTs without alterations (Figure 7C). Thus, it appears that genetic and epigenetic alterations are responsible for down-regulation of *GCNA* expression in human GCTs. Notably, we observed a significant association between low *GCNA* expression and poor GCT patient survival (Figure 7D, hazard ratio = 0.66, 95% CI 0.45–0.96, log-rank test, $P = 0.032$), which suggests that *GCNA* might serve as a potential prognostic marker.

To further investigate whether genetic and epigenetic alterations in *GCNA* gene are associated with genome instability, we studied the frequency of copy number amplification and loss on a genome-wide scale. We found that tumors with alterations in *GCNA* gene display significantly elevated frequency of both copy number amplification and loss events (Figure 7E), supporting the idea that loss of *GCNA* expression may contribute to human germ cell tumorigenesis by promoting genome instability.

DISCUSSION

Here, we report the initial functional characterization of GCNA as a key regulator of genome stability within germ cells and early embryos and as a putative tumor suppressor in GCTs. *Gcna* mutants exhibit remarkably similar phenotypes in flies, worms, and zebrafish indicating that GCNA carries out conserved functions throughout the animal kingdom. Germ cells must contend with many distinct challenges imposed by their unique biology and the dangers of various endogenous and exogenous genotoxic threats. Loss of GCNA compromises the ability of germ cells to handle these stresses and disrupts a number of seemingly disparate processes including germ cell development and maintenance, chromosome segregation, the cell cycle, DNA replication, and the formation of programmed DSBs during meiosis. Our findings suggest that these processes may be linked together through GCNA in unexpected ways. Given its enriched germline expression and critical function, GCNA should be considered, alongside with Nanos, Vasa, and Piwi as an essential player in germ cell biology.

How can GCNA influence so many different processes within germ cells? Our data indicate that the phenotypes exhibited by *Gcna* mutants do not extend from one specific malfunction, but rather from disruption of a number of distinct functions. Our genetic experiments in flies and worms indicates specific functions of GCNA can be attributed to distinct domains (Figure 5). For example, SprT enzymatic activity appears necessary for the prevention of

excessive Spo11-dependent and -independent DNA damage during *Drosophila* oogenesis. By contrast, the SprT mutant transgene partially rescues the maternal-effect semi-lethality of *Gcna* mutants, indicating this enzymatic activity is not essential for the proper regulation of chromosome segregation and cell cycle during early embryogenesis. Similarly, in *C. elegans*, *gcna-1* alleles which encode just the amino-terminal IDR domain do not exhibit a reduction in average brood size over the first eight generations. This contrasts with the loss of fecundity conferred by the null allele, again indicating that many germ cell activities depend on the IDR. This separation of structure and function provides an important framework for further understanding how GCNA ensures genomic stability across species. Of note, the mouse protein is comprised of just an IDR domain (Carmell et al., 2016), raising the possibility that the chromosome segregation and cell cycle functions may represent the core activities of these proteins.

Given the modest increase of DPCs within *Gcna* mutants (Figure 4) and the failure of the HE>AA mutant transgene to rescue Spo11-dependent and -independent damage in *Drosophila* ovaries (Figure 5), it is tempting to speculate that the SprT domain of GCNA serves the same function as it does in Spartan, namely to regulate DPCs that form on chromosomes (Stingele et al., 2017). Indeed, a recent study found that human GCNA/ACRC is recruited to DPCs in a SUMO-dependent manner (Borgermann et al., 2019). Consistent with these results, we find GCNA associates with replication machinery, based on its ability to immunoprecipitate components of the MCM complex (Figure 4) and *Gcna* mutants likely suffer from replicative stress based on increased RPA foci in *Drosophila* cells (Figure 3), microsatellite instability in worms (Figure 3), and copy number amplification and loss in human germ cell tumors (Figure 7). This raises the question regarding the separate requirements for GCNA and Spartan. Germ cells may have an increased DPC load compared to somatic cells, and therefore require alternative mechanisms for dealing with these lesions. For example, germ cells undergo extensive epigenetic reprogramming. These reactions, such as histone demethylation, produce cross-linking agents as by-products (Stingele et al., 2017). In addition, germ cells encounter enzymatic DPCs during meiotic DSB induction when Spo11 acts through a topoisomerase-like mechanism and becomes covalently attached to DNA at the break site (Keeney et al., 1997) While the MRN complex clears these adducts through endonuclease cleavage (Neale et al., 2005) perhaps GCNA represents an alternative mechanism for clearing Spo11 off of meiotic chromosomes in order to ensure that all lesions are repaired prior to embryogenesis. Given the increase of Spo11-induced breaks in flies, where very few meiotic breaks are made, GCNA may have evolved to limit either Spo11 activity or the amount of Spo11 that reaches the DNA, perhaps through sequestration of Spo11 or its accessory factors in the cytoplasm. The dramatic accumulation of Top2 in a subset of nuclei in *gcna-1* mutant worms and on mitotic DNA in mutant flies is also consistent with a sequestration model. Lastly, germ cells, particularly oocytes, experience extensive periods of cell cycle arrest during which they may accumulate DPCs that would otherwise interfere with chromatin-based processes upon fertilization. Perhaps GCNA also provides a replication-independent means for clearing or preventing such lesions.

IDR proteins have emerged as important regulators of germ cell biology. Many proteins that play essential roles in germ cells, such as Vasa, Oskar, Bucky ball and MEG-3, contain IDRs. These IDRs often control the ability of these proteins to undergo phase transitions,

allowing for the compartmentalization of various RNAs and proteins (Pritesh and Roland, 2018). We are just beginning to understand the functional significance of this molecular behavior. The IDR of GCNA is essential for its function within germ cells. While the primary amino acid sequence of GCNA's IDR has diverged, this region has continued to retain a high percentage of aspartic acid and glutamic acid residues. The importance of this distinct composition remains unknown, but perhaps it regulates the stability of GCNA or its ability to form condensates. Our expression studies indicate that GCNA localizes in a discrete cytoplasmic punctae in interphase and on chromosomes during mitosis. In flies and worms, the GCNA IDR promotes proper chromosome segregation and cell cycle regulation. Perhaps GCNA undergoes phase transitions and thereby controls the availability, assembly, or function of factors needed for these processes.

In flies and zebrafish, loss of *Gcna* leads to profound maternal-effect phenotypes, marked by chromosome segregation defects, chromosome bridges, and cell cycle asynchrony. Worm *gcna-1* mutants also exhibit chromosome bridges and X-chromosome loss, albeit at a lower frequency or only in later generations. The chromosome bridges that form in $Gcna^{KO}$ fly embryos often contain the heterochromatic 359-Bp repeats found on the X chromosome. This satellite sequence is responsible for the chromatid separation defects that occur in hybrid progeny of *D. melanogaster* and *D. simulans* (Ferree and Barbash, 2009). In addition, recent work has shown that loss of *mh*, which encodes the *Drosophila* homolog of Spartan, also results in chromosome bridges that contain 359-bp sequences (Tang et al., 2017). While, our genetic experiments indicate that GCNA and Spartan proteins act in parallel pathways, they may both converge on a mechanism that helps to resolve segregation defects involving heterochromatic sequences. Interestingly, we observe the formation of micronuclei and chromosome fragmentation in both *Drosophila* and *C. elegans Gcna* mutants. Similar events are thought to presage chromothripsis in cancer cells (Stephens et al., 2011) Sequencing data in the accompanying paper indicate that *C. elegans Gcna* mutants display molecular signatures consistent with chromothripsis (Davis et al. (2019)). Thus, the further study of GCNA may provide a model for understanding the origins of this newly recognized process and for determining how germ cells protect themselves from widespread chromosome re-arrangements in the face of DNA damage.

Underscoring the importance of GCNA in cancer, we have found that *GCNA* is among the most highly mutated genes in pediatric germ cell tumors and its loss is associated with pathogenicity. The platinum-resistance of certain GCTs (Batool et al., 2019) necessitates the discovery of novel druggable targets, and our studies suggest GCNA may be such a therapeutic target. The association of GCNA with both gene amplification and loss in GCT samples suggests that tight regulation of GCNA may be critical. Future studies on gain and loss of GCNA function will help elucidate its role as a driver of tumorigenesis.

Germ cells have many unique features in regards to reprogramming, regulation of the cell cycle and DNA repair. The study of *GCNA* will provide further insights into how these processes are coordinated with each other to ensure the faithful transmission of genetic material from one generation to the next. In addition to the observed correlation between loss of *GCNA* and human germ cell tumors, we anticipate *GCNA* function likely influences

other aspects of human fertility and transgenerational inheritance across sexually reproducing species.

STAR METHODS

I. LEAD CONTACT AND MATERIALS AVAILABILITY.

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Michael Buszczak (Michael.buszczak@utsouthwestern.edu)

II. EXPERIMENTAL MODEL AND SUBJECT DETAILS.

A) D. melanogaster—All *Drosophila* genetics followed standard procedures. Fly stocks were maintained at room temperature on standard cornmeal molasses-yeast food unless otherwise mentioned. All flies expressing a transgene and the corresponding controls were maintained at 25°C. w^{Berlin} strain was used as a wild type control.

All embryos used in experiments were 0–2 hours old collected on grape juice agar plates with wet yeast. For the fertility assays, depending on the availability of desired genotype, 5–10 flies were set up in each cage with grape juice agar plates with wet yeast, and the same number of flies were set up in the control cages to allow for comparable counting of eggs laid and fertility. For comparison of rescue genotypes, minimum of 615 eggs were counted and a maximum of 1059. For chemical treatment of flies, 15–25 flies of each genotype were placed in vials and starved for 16–18 hours before being fed on whatman paper soaked in 1% sucrose mixed with the drug (hydroxyurea). HU is water soluble, so the dilutions were made in 1% sucrose directly. For irradiation, adult flies were placed in a bottle for 24 hours and flipped out before exposing to 0 or 10 Gy of irradiation. For RNA sequencing, 30 females of w^{Berlin} and $GCNA^{KO}$ phenotype were fed on standard cornmeal molasses supplemented with wet yeast to allow for fatting. All flies were 1–3 days old and allowed to fat for 24 hours before dissection. To generate the $GCNA^{KO}$ allele Rainbow Transgenics injected 250 embryos of the Nanos-Cas9 expressing line and sent us the surviving larvae.

B) C. elegans—All strains were established and maintained on NGM plates seeded with OP50. Strains were kept at 20°C for long-term passaging under standard conditions (Brenner, 1974). For all experiments, unless otherwise stated, F1 homozygous *gcna-1(ea43)* worms were shifted as L4/young adults to 25°C and grown for two generations before analysis, since brood analysis showed that F1 and F2 populations sizes were near wild-type, suggesting maternal rescue. Wild type refers to the *C. elegans* variety Bristol, strain N2. The stocks utilized in this study are provided in the Reagents and Resources Tables. Primers and PCR conditions for genotyping are provided in Table S5. Several of the stocks were provided by the *Caenorhabditis* Genome Center (University of Minnesota) which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

C) D. rerio—Zebrafish were maintained in a recirculating aquatics facility at 28°C on a standard diet as described previously (Westerfield, 2007). Vertebrate animal work was accredited by AALAC and overseen by the UT Southwestern IACUC committee. The WIK wild-type strain was used and was obtained from the Zebrafish International Resource

Center (https://zebrafish.org). Embryos were collected at the 64–128 cell stage for DAPI staining, and at 25 hours post-fertilization for morphologic characterization.

D) H. sapiens—Tumor samples and clinical information used in this study were obtained under informed consent and approval by the Institutional Review Board of the participating facility. Samples were assembled from collections at the University of Texas Southwestern Medical Center, Dallas, TX USA; Children's Oncology Group; Boston Children's Hospital, Boston, MA USA; the Erasmus University Medical Center, Rotterdam, Netherlands; and the Hospital Sant Joan de Déu, Barcelona, Spain. All samples were de-identified at the source. Genomic DNA and RNA were extracted using the QIAamp DNA Mini kit (Qiagen) or Gentra PureGene kit (Qiagen) and the RNeasy Mini kit (Qiagen), respectively.

III. METHOD DETAILS.

A) D. melanogaster:

Immunofluorescence: Adult ovaries were stained according to (Tastan et al., 2010). Ovaries were dissected in 1x PBS. Tissue was fixed for 10 minutes with gentle rocking in 4% formaldehyde (EM grade) in PBS. After fixation, ovaries were washed four times in PBT (PBS + 0.5% BSA + 0.3% Triton-X 100) at RT for 10 minutes. Primary antibodies were incubated overnight at 4°C. Ovaries were then washed four times with PBT for 10 minutes, incubated for five hours with secondary antibodies. Ovaries were then washed and mounted in VectaShield Mounting medium with DAPI (Vector Laboratories).

Drosophila embryos were stained according to (Mani et al., 2014). Embryos were collected, dechorionated in 50% bleach, and fixed in 50% heptane, 50% fixative (3 parts fixing buffer, 1.33X PBS and 67mM EGTA :1 part 37% formaldehyde) for 10 mins. Embryos were then washed and devitellinized in methanol (MeOH) and stored at −20°C. Before staining, embryos were washed in a rehydration series consisting of 70%MeOH: 30%PBST, 50%MeOH: 50%PBST, 30%MeOH:70% PBST and finally 100% PBST for 5 minutes each, where PBST is PBS with 0.2% Triton X. Embryos were blocked in 10% normal goat serum for 1 hour.

The following primary antibodies were used: mouse anti gamma-H2AV (DSHB unc93.5.3.1) (30:200), rabbit anti-gamma-H2Av (Kim McKim) (1:500), rabbit anti-C(3)G (Mary Lilly) (1:3000) (Hong et al., 2003), rabbit anti-RPA (1:500) (Terry Orr-Weaver from Fisher and Cotterill labs), rabbit anti-Top2 (T. Hsieh and D. Ardnt-Jovin) (1:400), mouse anti-Hts (1B1) (DSHB) (1:20), rat anti-Vasa (DSHB) (1:20), rabbit anti-GFP (1:1,000) (Life Technologies), rat anti-HA 3F10 (Roche), and fluorescence-conjugated secondary antibodies (Jackson Laboratories)(1:300).

Western blot analysis: Proteins extracts were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. The following primary antibodies were used: rabbit anti-Top2 (T. Hsieh and D. Ardnt-Jovin) (1:2000), rat anti-Vasa (DSHB) (1:200), mouse anti-actin (DSHB) (1:100), rat anti-HA 3F10 (Roche) (1:2000). The secondary antibodies were anti-mouse IgG HRP (Jackson Laboratories) (1:2000), anti-rabbit IgG HRP (Jackson Laboratories) (1:2000), and anti-rat IgG HRP (Jackson Laboratories) (1:2000).

DNA-protein crosslink isolation: DPCs were isolated and detected using a modified rapid approach to DNA adduct recovery (RADAR) assay wherein the tissue was lysed in 10 mM Tris-HCl pH 6.8, 6 M GTC, 1% dithiothreitol, 20 mM EDTA, 4% Triton X, 1% sarkosyl (Kiianitsa and Maizels, 2013; Vaz et al., 2016). The DNA was ethanol precipitated by adding an equal volume of 100% ethanol to the lysis buffer and incubated at −20°C for five minutes. The pellet was washed three times in wash buffer (20 mM Tris-HCL pH 6.8, 150 mM NaCl and 50% ethanol). The nucleic acid pellet was solubilized in 8 mM NaOH. For *Drosophila*, the ovaries were dissected in cold Graces media and lysed for RADAR. Embryos were 0–2hrs old. Embryos were dechorionated in 50% bleach for 2 minutes and then rinsed thoroughly with water before lysis. The zebrafish embryos were lysed at the 1000 cell stage. DNA concentration was measured using PicoGreen (Invitrogen) according to the manufacturer's instructions.

DNA and DNA-protein crosslink detection: DNA was detected by blotting the DNA with a slot blot vacuum manifold (Biorad) onto a positively charged nylon membrane. The membrane was probed with mouse anti-dsDNA (Abcam) (1:2000). Specific proteins were detected by normalizing to DNA concentration and digesting with benzonase. The proteins were separated on a polyacrylamide gel and silver stained (Sigma).

Immunoprecipitation: For overexpression in S2 cells, the Gcna construct was cloned into pAFHW (Drosophila Gateway Vector Collection). S2 cells were transfected with the expression constructs using Effectene (Qiagen). Cells were lysed (50mM Tris pH8.0, 137 mM NaCl, 1mM EDTA, 10mM NaF, 1% Triton X100, 10% glycerol, Roche protease inhibitor cocktail) and were applied to FLAG sepharose beads (Sigma) for 6.5 hours. The beads were washed three times in ice cold lysis buffer and the protein was eluted with 0.5 mg/mL 3x FLAG peptide (Sigma) overnight. The protein was applied to HA sepharose beads (Roche) for 8 hours. The beads were washed three times in ice cold lysis buffer, and bound proteins were retrieved by boiling the beads. Samples were run on SDS-PAGE gel and stained with Coomassie blue dye prior to analysis by mass spectrometry.

Mass spectrometry analysis: Protein gel bands were excised before being reduced with DTT and alkylated with iodoacetamide. Samples were digested overnight at 37°C using trypsin. Tryptic peptides were de-salted via solid phase extraction (SPE). LC-MS/MS experiments were performed on a Thermo Scientific EASY-nLC liquid chromatography system coupled to a Thermo Scientific Orbitrap Fusion Lumos mass spectrometer. To generate MS/MS spectra, MS1 spectra were first acquired in the Orbitrap mass analyzer (resolution 120,000). Peptide precursor ions were then isolated and fragmented using high-energy collision-induced dissociation (HCD). The resulting MS/MS fragmentation spectra were acquired in the ion trap. MS/MS spectral data was searched using Proteome Discoverer 2.1 software (Thermo Scientific) against entries included in the *Drosophila melanogaster* Uniprot protein database. Search parameters included setting Carbamidomethylation of cysteine residues (+57.021 Da) as a static modification and oxidation of methionine (+15.995 Da) and acetylation of peptide N-termini (+42.011 Da) as dynamic modifications. Precursor and product ion mass tolerances of 15 ppm and 0.6 Da were used, respectively. Peptide spectral matches were adjusted to a 1% false discovery rate (FDR) and additionally

proteins were filtered to a 5% FDR. Proteins were quantified by area values determined via label-free quantitation using Proteome Discoverer 2.1 software. Complete protein lists from raw data are included as Tables S1–3.

Generating the *Gcna^{KO}* alleles: To generate the *Gcna^{KO}* allele, guide RNAs were designed using http://tools.flycrispr.molbio.wisc.edu/targetFinder and synthesized as 5'-unphosphorylated oligonucleotides (see key reagents), annealed, phosphorylated, and ligated into the BbsI sites of the pU6-BbsI-chiRNA plasmid (Gratz et al., 2013). Homology arms were PCR amplified and cloned into pHD-dsRed-attP (Gratz et al.,2014). Guide RNAs and the donor vector were co-injected into nosP>Cas9 attP embryos at the following concentrations: 250 ng/ml pHD-DsRed-attP donor vector and 20 ng/ml of each of the pU6-BbsI-chiRNA plasmids containing the guide RNAs (Rainbow Transgenics). Injected embryos were allowed to develop into larvae and crossed to wild type strains. The progeny of this cross was screened by a fluorescent fly microscope to allow for detection of positive KO/KI events by presence of DsRed.

Cloning of *Drosophila Gcna* transgene: PCR products were cloned into pENTR (Life Technologies) and swapped into pAHW, pAWG (attB added by Tony Harris) or pAFHW (Drosophila Gateway Vector Collection) using an LR reaction. Using this approach, we isolated clones corresponding to the *CG14814* cDNA (accession number BDGP:RE06257) transcripts.

Live imaging: 3–5 day old males and virgin females were mated in mating cages containing grape juice (3%) agar plates with a little bit of wet yeast. The flies were allowed to lay eggs for 1–2 hrs at 25°C. Eggs were carefully collected and dechorionated by rolling them on double-sided tape pasted on a slide. Dechorionated eggs were then mounted using oil and non-auto-fluorescent glue. Live imaging was conducted every 15 seconds using a Zeiss LSM800 microscope.

Fertility assays: 0–2 day-old wild type males and virgin females of the genotype being tested were mated in mating cages with grape juice (3%) agar plates with a little bit of wet yeast. The flies were allowed to lay eggs for 48–72 hrs at 25°C before switching out the plates. Flies were allowed to lay eggs for a total of 7 days before eggs were counted.

Super-resolution imaging structured illumination microscopy (SIM) imaging and image processing: Ovaries were dissected according to standard protocol and mounted in Prolong Gold antifade reagent (Life technologies). Nikon N-SIM system (Nikon, Tokyo, Japan) equipped with a 100×/1.49 TIRF oil immersion objective lens (Nikon), the iXon + electron multiplying charged-coupled device camera (Andor) and an excitation laser unit of 405 nm, 488 nm, 561 nm and 640 nm (Coherent) was used for a superresolution optical imaging. Z-stacks of SIM optical sections were acquired with a 120 nm Z-step size. Image processing, including 3-dimensional reconstruction and colocalization analysis, were carried out using the NIS-Element Advanced Research software (Nikon).

Irradiation of flies: Adult flies were placed in a bottle for 24 hours and flipped out before exposing to 0 or 10 Gy from a cesium source. The progeny that survived to eclosion were

counted and noted genotype. The ratio of homozygotes to heterozygotes was used to determine radiation sensitivity.

Hydroxyurea exposure: 0–2 day-old wild-type and *Gcna*^{KO} female flies were collected and fed on wet yeast for 24 hrs. They were then starved for 16–18 hrs. Whatman paper was soaked in either 1% sucrose alone or 1% sucrose with 50mM Hydroxyurea. Flies were allowed to feed on sucrose with solvent or drug for 24 hours before being dissected and immunostained.

Scanning electron microscopy: Eggshells were mounted on SEM stubs and sputter coated with gold/palladium in a Cressington 108 auto sputter coater. Images were acquired on a Field-Emission Scanning Electron Microscope (ZeissSigma, Carl Zeiss, Thornwood, NY) at 10.0 kV accelerating voltage for WT embryos and 8.0 kV accelerating voltage for *Gcna*^{KO} embryos.

FISH: 0–2 hour embryos were fixed according to “Immunofluorescence”. FISH was performed on embryos according to Beliveau *et al* (2015). Embryos were washed once in the following: 1x PBS, 1x PBS + 0.1% Tween for 1 minute, 1x PBS + .5% Triton for 10 minutes, 1x PBS + 0.1% Tween for 1 minute, and .1N HCl for five minutes. Embryos were washed three times in 2X SSC + 0.1% Tween (SSCT). Embryos were washed in 2X SSCT/50% formamide (Sigma) for 5 minutes. Embryos were incubated in 2X SSCT/50% formamide at 60°C for 20 minutes. Probes were prepared in so lution of 10% dextrose, 2X SSC, 50% formamide and 30 pmol of oligopaint probe. Probes were added to the samples and denatured at 78°C for 2.5 minutes. Samples were incubated overnight at 42°C. Embryos were washed i n 2X SSCT/50% formamide at 60°C and then room temperature. Embry os were washed in 0.2X SSC and mounted in Vectashield mounting medium + DAPI.

RNA sequencing: 30 ovaries were dissected from *Gcna*^{KO} and control (*w*^{Berlin}) flies. There were three biological replicates produced from independent backcrosses of the KO into *w*^{Berlin} background for each sample. Qiagen miRNeasy Mini kit was used for RNA extraction and the RNA was sent for library preparation that selected for total RNA at McDermott Sequencing core at UTSW. The samples were sequenced in Illumina HiSeq 2500 cycle single-read platform. The sequencing reads were checked for quality using FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). For analyzing differential gene expression, STAR aligner (version 2.5) (Dobin et al., 2013) was used to map the RNA-Seq reads to the Drosophila reference genome (genome assembly BDGP6.88) with an additional flag -- outFilterMultimapNmax 100. TEtranscripts (Jin et al., 2015) was used to generate the read counts that mapped to TEs and genes. DESeq2 (Love et al., 2014) was used for differential expression analyses of TEs with a FDR of 5%.

DNA extraction and genotyping: A single fly from each genotype were ground and DNAzol and pelleted to clear debris. The DNA was then ethanol precipitated with 100% ethanol and incubated at −20°C for 20 minutes. Genotyping PCR was performed using the SapphireAmp fast PCR from Takara.

B) C. elegans

Generation of *C. elegans* CRISPR alleles: Null, tagged, and mutated versions of *gcna-1* were created by CRISPR/Cas9 genome engineering. For null alleles *ea43* and linked double mutant lines, 5' *gcna-1* crRNA and 3' *gcna-1* crRNA were utilized with *gcna-1* null repair ultramer. For *ea76*, internal crRNA was used with the 5' *gcna-1* crRNA. For tagging with OLLAS and HA, 5' *gcna-1 crRNA* was injected with either OLLAS::GCNA or HA::GCNA ultramer. The *dvc-1(ea65)* null was created with 5' *dvc-1* crRNA and 3' *dvc-1* crRNA and *dvc-1* null repair ultramer.

Injection mixes containing crRNA, tracrRNA, ssDNA ultramers, and house-purified Cas9 protein were prepared as described (Paix et al., 2017). crRNA and ssDNA repair template for *dpy-10(cn64)* co-injection marker were also included (Arribere et al., 2014). Both Roller and Dumpy (*dpy-10*) transformants were individually plated and F2 Dumpy progeny were screened by PCR for relevant insertions and deletions (see Resources tables for all crRNA, ultramer, and primer sequences). Specifically, three to four F2 Dumpy progeny of Rolling or Dpy mothers were pooled in 14µl worm lysis buffer (10mm Tris-HCl, 50mm KCl, 2.5mm MgCl2, 0.45% NP40, 0.45% Tween20, 0.01% gelatin with 5ng/µl proteinase K). Worms were lysed at 65°C for 1 hours followed by 95°C for 15 minutes and then cooled to 10°C. 1µl of lysate was mixed in a 10µl PCR reaction with 0.5µM of each primer, 5µl 2x MasterAmp PCR mastermix, and 0.25µl Taq DNA polymerase. Reaction conditions were 94°C 2 minutes, 40 cycles of 94°C 1 5 sec, 53°C for 30 sec, and 72 °C for 30 sec, and followed by an extension at 72°C for 5 minutes. Products were run on a 2–3% agarose gel next to wild type control. PCR primers and conditions at listed in Table S5. Putative mutant lines were outcrossed to N2 at least twice prior to sequencing at the University of Pittsburgh Genomic Core, Pittsburgh, PA.

Purification of Cas9 protein: DE3 GOLD cells were transformed with nm2973 plasmid (Xu Hua Fu et al., 2014) and plated on LB + 50ug/mL Carbenicilin. Fresh transformants were inoculated into 5ml LB/Carbenicillin, grown at at 37°C overnight, and transferred to 1L LB + 0.1% glucose + 50ug/mL Carbenicilin and grown at 25°C to OD_{600}=~0.5. Cultures were shifted to 18°C for 15–25 minutes, and IPTG was added to 0.2 mM. Cultures were incubated overnight, pelleted and wet weight was measured. Cells were resuspended at ~6mL/g cells with Buffer A (20mM Tris ph 8.0, 250 mM KCl, 20 mM imidazole, 10% glycerol, 1 mM TCEP) +protease inhibitor (Roche, #11836170001) + 1mM PMSF. Cells were sonicated 6 × 45 sec (setting 3 at 30%, 1 second pulse-2 second pause) with 1 minute cooling in between and then spun 30 minutes at 16000×g. Supernatants were transferred to a fresh tube. A 5mL Ni-agarose column (Qiagen, #30410) was equilibrated with Buffer A (at least 25mL). The clarified lysate was batch bound onto the Ni-agarose for 45 minutes at 4°C and the column was then washed with 100mL of Buffer B (20mM Tris ph 8.0, 800 mM KCl, 20 mM imidazole, 10% glycerol, 1mM TCEP). Protein was eluted with Buffer C (20mM HEPES ph 8.0, 500 mM KCl, 250 mM imidazole, 10% glycerol). Fractions were checked by SDS-PAGE gel and fractions with Cas9 proteins were pooled and mixed 1:1 with Buffer D (20mM HEPES pH 8.0). To remove contaminating DNA in the prep. The eluate was run of a 5mL Q Sepharose (Sigma, #Q1126) column equilibrated with 25ml 1M KCl and then 25ml Buffer D. Flow-through was collected and dialyzed into 1L Buffer E (20mM HEPES, 500

mM KCl, 20% glycerol) for 5 hours at 4°C, transferred into 1L Buffer E, and dialyzed overnight. Protein was concentrated to ~10 mg/mL using a 100K centrifugal filter (Milipore, UFC910024). Aliquots and flash-freeze in liquid nitrogen and stored at −80°C.

Fertility Assays: Brood sizes were performed by individually plating L4 animals on center-seeded 3cm plates and transferring every 24 hours until the cessation of egg-laying. Total numbers of L4 and adult hermaphrodite and male offspring were counted 48–72 hours later.

Transgenerational assays were performed by starting 3cm or 12-well plates with single F1 progeny of balanced heterozygous moms. Each generation, the first F1 progeny to reach L4 were transferred. Plates were examined after 24 hours to determine if the worm was sterile and replaced with an adult from the parental plate if no eggs were observed in the uterus or on the plate. Brood sizes were binned into ranges shown in Figure 2. Presence of males and mutant animals (Dumpy, Uncoordinated, etc) was noted. Sterile populations were reassessed by plating additional worms from the parental plate and classified as sterile when no further offspring could be attained from the previous generation. This method ensures that sterility is inherent to the population and not simply a subset of offspring. All experiments utilized non-starved populations of worms. Data is shown for 25°C as sterility did not arise in populations of *gcna-1(ea43)* grown at 20°C.

Fecundity of *gcna-1;prg-1* hermaphrodites was tested by shifting *gcna-1(ea43)*, *prg-1*, or *prg-1;gcna-1(ea43)* animals from 20°C to 25°C as early L4 larvae and counting offspring 3 – 4 days later.

Detecting microsatellite deletions: The *dog-1(gk10)* mutations was outcrossed to N2 twice before mating with *gcna-1(ea43). dog-1, dog-1;gcna-1*, and *gcna-1* lines were generated from the cross and grown three generations at 25°C. For each genotype 100–200 F4 worms were collected in 14μl of lysis buffer (50 mm KCL, 10 mm, Tris pH 8.3, 2.5 mm $MgCl_2$, 0.45% NP-40, 0.45% Tween-20, 0.01% gelatin) with 5 mg/ml freshly added proteinase K and individually lysed at 60° for 60 min and then at 95° for 15 min. Deletions upstream of the *vab-1* locus were detected using a nested PCR reaction as described (Youds et al., 2006): 94°C for 4 min followed by 34 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 1 min 30 sec, and a final elongation step of 72°C for 10 min. 1μl of external PCR product was used in the internal PCR reactions with similar conditions except 62°C annealing temperature and only 1 min ex tension time. PCR products were run on a 2–3% agarose gel and scored for unique DNA fragments less than 500bp.

Irradiation sensitivity assay: Day 1 adults were exposed to increasing doises of ionizing radiation using a ^{137}Cs source (Gammacell 1000 Elite; Nordion International). Embryos and L1 larvae from individually-plated animals at t= 24–36 hours post-irradiation were collected and counted. Viable offspring were counted 2.5 – 3 days later. Data is normalized to the hatching rates of the unexposed animals of each genotype.

Quantitative PCR: Approximately 100 young adults of each genotype were washed three times in 1x M9 buffer (3 g/L KH_2PO_4, 6 g/L Na_2HPO_4, 5 g/L NaCl, 1 mM $MgSO_4$), resuspended in Trizol (Invitrogen), and vortexed for ~60 seconds before being flash frozen

and stored at −80°C. Once all the samples were collected, samples were thawed on ice, sonicated, and RNA was isolated by chloroform extraction and isopropanol precipitation. Samples were treated with DNase (Sigma #AMPD1) and reverse transcribed into cDNA (Protoscript m-MuLV First Strand cDNA Synthesis kit, NEB #E6300S) according to manufacturer's instructions. Quantitative real-time PCRs were performed on the Applied Bio Systems 7300 Real Time PCR System using Sybr Green chemistry (SensiMix SYBR Hi-ROX kit, Bioline #QT-605) with transcript-specific primers for Tc1, Tc2, Tc3, and Tc4v as described in Table S2. The reference genes *rpl-32* (Hoogewijs et al., 2008) was used for normalization across samples and gene expression was analyzed using the ΔC_T method (Livak and Schmittgen, 2001). Results are presented as the average of combined data from three independent biological replicates that in turn is comprised of three technical replicates each.

Immunostaining: One day-old adult worms were dissected in 3.5μL 1x sperm salts (50mM PIPES, pH 7.0, 25 mM KCL, 1 mM $MgSO_4$, 45 mM NaCl, 2 mM $CaCl_2$.) +0.2μL 10mg/mL levamisole. Fixation and pre-hybridization varied for different primary antibodies:

α-RAD-51: 7μL of 2% paraformaldehyde was added to slides and incubated 5 minutes in a humidity chamber at room temperature before freezing on a metal block on dry ice for 10 minutes. Coverslips were flicked off and slides were then submerged in −20°C methanol for 5 minutes and dipped in −20°C acetone for 5 seconds. Slides were air dried, a wax box was drawn around the samples, and dissected worms were prehybridized for 3× 10 minutes in Phosphate buffered saline (PBS) + 0.1% Tween 20 +0.1% BSA (PBSTB). Primary antibody (courtesy of Sarit Smolikove) was diluted 1:20,000 and incubated overnight at 4°C.

α-FLAG (for visualization for TOP-2::FLAG): 3.5μL of 2% Triton and 7μL 2% PFA were added to dissected samples (fixed and frozen as above). Slides were submerged in −20°C methanol for 1 minute and washed as above in PBSTB. Mouse α-FlagM2 (Sigma F1804) was diluted 1:500 and incubated overnight at 4°C.

α-HA (for visualization of HA::GCNA-1): 3.5μL of 2% Triton and 7μL 2% PFA were added to dissected samples (fixed and frozen as above). After fixation, slides were submerged in −20°C Methanol for 5 minutes. Mouse α-HA antibody (Santa Cruz F7) was diluted 1:500 and incubated overnight at 4°C.

The next day, slides were washed 3 ×10 min each with PBSTB and then incubated with secondary antibodies diluted in PBSTB (goat α-Rabbit Alexa 568, goat anti-mouse Alexa 488, diluted 1:2000) for 2 – 4 hours at room temperature. Slides were then washed once with PBSTB for 10 min, once with PBS+DAPI (4',6-diamidino-2-phenylindole) for 15 min, and again with PBSTB for at least 10 min prior to mounting in Prolong Gold Antifade Mountant with DAPI (ThermoFisher Scientific, P36931). Slides were cured overnight prior to confocal imaging.

Whole-mount staining for diakinesis analysis: One-day old adult worms were fixed in Carnoy's solution (three parts absolute ethanol; two parts chloroform; one part glacial acetic acid), stained with DAPI (4',6-diamidino-2-phenylindole) for at least fifteen minutes, then

mounted in Prolong Gold Antifade Mountant with DAPI (ThermoFisher Scientific, P36931). Slides were cured overnight prior to confocal imaging. Statistical analyses were performed as described in Macaisne et al., 2018 using GraphPad Prism software.

Imaging and Quantification of Staining: RAD-51 foci and diakinesis nuclei were quantified by collecting Z-stack images on a Nikon A1r confocal microscope with 0.2μm sections. Three dimensional stacks were visualized using Volocity imaging software (Quorum Technologies). For RAD-51 foci, nuclei from the transition zone nuclei through the pachytene/ diplotene border were quantified. The pachytene region was divided into 6 equal parts according to number of rows of nuclei in this region. Diakinesis nuclei were rotated in three dimensions to attain the number of DAPI-staining bodies in the −1 and −2 nuclei (i.e. the two oocytes preceding the spermatheca).

C) D. rerio

Generation of *gcna* mutant zebrafish: Target sites for sgRNAs were selected using the online software CRISPR DESIGN (http://crispr.mit.edu). One sgRNA was designed to target a site on exon 3 (5'-GAAGACCAGACGTCCAGCTT-3') of the zebrafish *gcna* gene. The sgRNA was synthesized as previously described (Gagnon et al., 2014). Briefly, the gene-specific oligonucleotide, consisting of an upstream SP6 promoter (5'-GCGATTTAGGTGACACTATA-3') followed by the 20-base target sequence (5'-GAAGACCAGACGTCCAGCTT-3') and a sequence complementary to the reverse tracrRNA tail oligonucleotide (5'-GTTTTAGAGCTAGAA-3'), was annealed to the reverse tracrRNA tail oligonucleotide, followed by incubation with T4 DNA polymerase to fill the ssDNA overhangs. The resulting DNA template was then purified using QIAquick PCR Purification Kit (Qiagen) and used for sgRNA transcription using the Megascript T7 Kit (Ambion). The sgRNA was then treated with DNase and precipitated with LiCl/ethanol.

Microinjections: Zebrafish embryos were injected at the one-cell stage with 2nl of a mix consisting of sgRNA (83 ng/ul), 1.2ul Cas9 protein (500 ng/ul) (PNA Bio Inc) and 0.08% phenol red dye. The embryos were injected with the PLI-90A picoinjector (Warner Instruments).

DNA extraction and PCR genotyping: Genomic DNA was extracted from either a pool of three 24hpf larvae or a single caudal fin of adult zebrafish. Tissue was incubated in 50 ul 50mM NaOH at 95°C for 30min. 10ul Tris-HCl (pH 8.0) was added to the lysate and vortexed to neutralize it. 1ul of the lysate was used for each PCR reaction with the forward (5'- GCTTAGGATCGGTAGTTTTCCG -3') and reverse (5'-GCAGGAGTCCATGTATGGAC -3') primers. For the PCR reactions the samples were denatured at 95°C for 3 min followed by 40 cycles consisting of 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 se c and a final step at 72°C for 5 min. To identify founders and determine germline transmission of indels, PCR products from embryo lysates were either run directly on a 3% agarose gel or the T7E1 Assay was performed as described previously (Kim et al., 2013b) and the digest products ran on a 1.5% agarose gel. To identify and sequence specific indels in the F1 generation, adult zebrafish PCR products were either

sequenced directly with the primers listed above or were cloned into the pCR4-TOPO vector (Thermo-Fisher) and sequenced with M13 forward and reverse primers.

Establishment and propagation of *gcna*^{KO} zebrafish lines: Adult mosaic fish were out-crossed to wild-type AB fish to genotype embryos and identify fish with germline transmission. Confirmed founders were subsequently out-crossed to wild-type AB fish and the progeny were genotyped at adulthood.

Immunohistochemistry: Embryos at stages prior to 24 hours post-fertilization were fixed overnight at 4°C. in 4% paraformaldehyde/1xPBS (PFA) and manually dechorionated. Embryos at stages >24 hpf were enzymatically dechorionated with Pronase (Sigma-Aldrich) prior to overnight fixation in PFA. For DAPI staining, the embryos were deyolked post-fixation, incubated for 5 min in 300 nM DAPI in 1xPBS with 0.1%Tween (PBST) and subsequently washed six times in PBST. Embryos were mounted in 3% methyl cellulose in E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$) and images acquired on a Zeiss Confocal Microscope.

For whole-mount phospho-Histone H2Av staining, fixed, dechorionated embryos were permeabilized with 100% acetone (7 minutes, −20°C.), washed 1 × 5 mins. in H_2O, 4 × 5 mins. in PBST, and blocked in 1% BSA in PBST for 30minutes at room temperature. Embryos were incubated overnight at 4 °C. in a 1:1000 dilution of rabbit anti-zebrafish histone H2Av (Sidi et al., 2008) and washed 4 × 5 mins. in PBST. Washed embryos were incubated 2hrs. room temperature in a 1:300 dilution of HRP-conjugated goat anti-rabbit IgG (Bio-Rad), washed 4 × 5 mins.in PBST and developed in 3,3′-Diaminobenzidine tetrahydrochloride (Sigma-Aldrich). Stained embryos were mounted in 90% glycerol/10% PBS and images were acquired a Leica MZ12.5 stereomicroscope attached to a Nikon E4500 camera.

Embryonic phenotype imaging and classification: To analyze the embryonic phenotype, 27 hpf embryos were dechorionated with Pronase, anesthetized with 0.015% MS-222 and mounted in 3% methyl cellulose in E3 medium. Images were taken at 3.2X magnification on a Leica MZ12.5 stereomicroscope attached to a Nikon E4500 camera. Embryos were classified as abnormal if they displayed significant morphologic defects in comparison to wild type embryos, as indicated by ventral body axis curvature posteriorly, shortened body axis and the presence of dark necrotic tissue.

D) H. sapiens

Characterization of human tumors: To detect somatic mutations, exome capture was carried out using SureSelect Human All Exon v4+UTRs (Agilent Technologies), and sequencing was performed with a HiSeq 2000 instrument (Illumina) with 100-bp paired-end reads to a mean coverage of 130× for exomes. Raw reads were mapped to human reference genome (hg19) using BWA (Li and Durbin, 2009). Matched tumor-normal BAM files were used as input for VarScan software (Koboldt et al., 2012) to identify somatic single-nucleotide variants (SNVs) and small-scale insertion/deletions (INDELs).

Detection of copy-number variation: Genomic DNA from GCT samples was analyzed by SNP array technologies using the Illumina Omni 2.5M SNP array and Affymetrix OncoScan array, according to the manufacturers' recommendations. We used Nexus Copy Number Discovery 7.0 software (BioDiscovery, Inc.), which can process raw data from both platforms with the same algorithm and procedure. In this software, the data were corrected for GC content and segmented by using SNP-FASST2 algorithm with default parameters.

Detection of promoter methylation: Genome-wide methylation analysis was performed using the Infinium HumanMethylation450 BeadChip array (Illumina, San Diego, CA) following Illumina's standard protocol. Raw intensity (idat) files were converted by using the methylumi package (Triche et al., 2013). Combined with IMA package (Wang et al., 2012), DNA methylation sites with missing values, cross hybridizing probes, located within repeat regions or on sex chromosomes were excluded, resulting in a total of 392,714 probes retained. Methylation data were subsequently converted into β values, ranging from 0 (unmethylated) to 1 (fully methylated), and these values were normalized using a betamixture quantile normalization method (BMIQ) (Teschendorff et al., 2013). To detect gene expression and further conduct analysis on association between gene expression and genetic/epigenetic alterations, RNA of GCT samples was sequenced on Illumina HiSeq2000 according to the manufacturer's protocol (Illumina). 100-bp paired-end reads were assessed for quality and reads were mapped using CASAVA (Illumina). The generated FASTQ files were aligned by Bowtie2 (Langmead and Salzberg, 2012) and TopHat2 (Kim et al., 2013a). Cufflinks (Roberts et al., 2011; Trapnell et al., 2010) was used to assemble and estimate the relative abundances of transcripts at the gene and transcript level.

Survival association analysis.: Survival association analysis between gene expression and GCT patients' survival was calculated based on 108 GCT cases measured by Affymetrix U133A microarray platform (Korkola et al., 2015; Korkola et al., 2009). Signal intensity CEL files were downloaded from Gene Expression Omnibus (GEO) repository at http://www.ncbi.nlm.nih.gov/geo/, data set GSE3218 and GSE10783. CEL files were then processed by Affymetrix Power Tools (APT) with Robust Multiarray Average (RMA) method. Cox proportional hazards model was used to calculate the statistical significance, as well as hazard ratios and 95% confidence intervals of the associations between the gene expression and survival. Kaplan-Meier curves were generated based on gene expression values dichotomized into over- and under-expressed groups using the within cohort median expression value as a cutoff.

IV. QUANTIFICATION AND STATISTICAL ANALYSIS.

A) D. melanogaster

Quantifications by Figure: All statistical analysis was performed using GraphPad/Prism.

- Fig 1 H: 100 different ovarioles counted across 4 pairs of ovaries for each genotype.
- Fig 1 L: 10–12 nuclei counted for each across 4 different pairs of ovaries for each sub-type (Meiotic and Non-Meiotic).

- Fig 3 C: 30–40 ovarioles were counted for each genotype, across three independent experiments. P value (two-tailed t-test) = 0.0012 by. P value (one-tailed) = 0.0055. Pairing was significant by both tests (P<0.05).
- Fig 3D: 37–52 ovarioles/germaria for each genotype were counted each experiment, and the experiment performed in three biological replicates.
- Figure 5 E: Number of nuclei counted for WT = 16, $Gcna^{KO}$ = 18, WT transgene rescue = 18 and HE>AA transgene rescue = 15.

  Unpaired t test, wild-type versus $Gcna^{KO}$ p<0.0001, $Gcna^{KO}$ versus $Gcna^{KO}$;*nos*>$Gcna^{WT}$ p=0.03, $Gcna^{KO}$;*nos*>$Gcna^{WT}$ versus $Gcna^{KO}$; *nos*>$Gcna^{HE>AA}$ p=0.205 and $Gcna^{KO}$ versus $Gcna^{KO}$; *nos*>$Gcna^{HE>AA}$ p=0.7.
- Figure 5 F: Number of germaria counted for each genotype = 8, across 4 pairs of ovaries.

  Unpaired t test, wild-type versus $Gcna^{KO}$ p= 0.0045, $Gcna^{KO}$ versus $Gcna^{KO}$;*nos*>$Gcna^{WT}$ p=0.0424, WT versus $Gcna^{KO}$;*nos*>$Gcna^{WT}$ p=0.3749, $Gcna^{KO}$;*nos*>$Gcna^{WT}$ versus $Gcna^{KO}$;*nos*>$Gcna^{HE>AA}$ p=0.0145 and $Gcna^{KO}$ versus $Gcna^{KO}$;*nos*>$Gcna^{HE>AA}$ p=0.5798.
- Figure 5 G: Unpaired t test, $Gcna^{KO}$; *nos*>$Gcna^{WT}$ versus $Gcna^{KO}$; *nos*>$GCNA^{HE>AA}$ ovaries, p=0.1322.
- Figure 5 G: Unpaired t test: wild-type vs $Gcna^{KO}$ p<0.0001, and $Gcna^{KO}$ versus $Gcna^{KO}$; *nos*>$Gcna^{WT}$ p=0.0004.
- Figure 5 G : Number of eggs counted for WT = 1035, $Gcna^{KO}$;*nosgal4*= 749, $Gcna^{KO}$; *nos*>$Gcna^{WT}$ = 1059, $Gcna^{KO}$;*nos*>$Gcna^{HE>AA}$ = 615.
- Fig S1 F: 8 individual single pair matings were set up.
- Fig S5 B and Fig S3 C: 3–40 ovarioles counted at each stage, $Gcna^{KO}$ samples had fewer later stages. For counting puncta per nucleus- between 14–51 nuclei were counted for $Gcna^{KO}$ samples, and represented with average and standard deviation.
- Fig S3 D: 40–50 germaria were counted across at least 5 pairs of ovaries.
- Fig S5 C: counted 30–50 of each stage for each genotype.
- Fig S6 C: Between 26–54 germaria counted for each genotype, done in three independent experiments. P value = 0.0003. Test used = two tailed t – test.
- Fig S6 D: single count. n=50 ovarioles across 5 pairs of ovaries for each genotype.

B) C. elegans

Quantifications by Figure: All statistical analysis was performed using GraphPad/Prism.

- Figure 2 C: n=12 parental worms for wt and n= 21, 21, 34 for *gcna-1* F1, F3, and F18, respectively

- Figure 2 D: n=12 parental worms for wt, and 34 each for *gcna-1* F1, F8, and F18
- Figure 2 G: Three germ lines of each genotypes were analyzed for RAD-51 foci in the meiotic germ line.
- Figure 2 H: *spo-11*, n= 20; *gcna-1;spo-11*, n=42
- Figure 3 A: n = numbers of individual F3 progeny scored for changes in repeat length at the *vab-1* locus. For *gcna-1, dog-1, dog-1;gcna-1*, n= 76, 162, 189, respectively.
- Figure 3 F: n= number of individual animals whose progeny were assessed. For N2, *gcna-1, dvc-1, gcna-1;dvc-1*, n = 12, 8, 6, 10 respectively.
- Figure 4 F: greater than 15 germ lines of each genotype were analyzed for TOP-2 localization. Representative images are shown.
- Figure 5 H: At least 10 squashes were analyzed for GCNA localization.
- Figure 5 J: n = numbers of adults F3 animals whose brood sizes were assessed.
- Figure S3 A: Fifty animals on each of 4 plates were examined for lifespan at the indicated generation.
- Figure S3 B: n values represent the number of animals of each age and generation tested for location, as indicated in the figure.
- Figure S3 D, E: At least 10 germ lines of each genotype are visualized.
- Figure S3 F: Percent survival indicates the number of L4 and adult offspring from eggs laid after exposure to control and increasing doses of IR > 500 eggs/ genotype from at least 10 offspring were counted.
- Figure S4 A: RT-qPCR was performed on 3 independent lines for each of 2 biological replicates.
- Figure S4 B: n = animals shifted to 25°C whose brood was subsequently analyzed. For N2, *gcna-1*, *prg-1*, *prg-1;gcna-1*, n = 10, 10, 10, and 20, respectively.

V. DATA AND CODE AVAILABILITY

- The RNA-seq datasets generated during this study are available at Gene Expression Omnibus under accession GSE127220.

VI. ADDITIONAL RESOURCES.

N/A

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We would like to thank K. McKim, T. Orr-Weaver, D. Ardnt-Jovin, M. Osterfield, K. Dubrovinski, M. Lilly, J. Abrams, H. Richardson, the Bloomington *Drosophila* Stock Center and the Developmental Studies Hybridoma Bank for *Drosophila* reagents and S. Smolikove and A. Jaramillo-Lambert for worm antibodies. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We also thank D. Jangam and E. Bertran for their help with RNA Sequencing data analysis. We thank members of the Buszczak and Yanowitz labs for comments and advice, and Jose Cabrera for help in preparing the graphical abstract. V.B. was supported by HHMI (Grant#56006776), NIGMS (T32GM10977601) and a fellowship from the Center for Regenerative Science and Medicine at UT Southwestern. C.D.G. was supported by NIGMS (5T32GM008203-30). This work was also supported in various phases by awards from NIGMS (R01GM086647 and R01GM125812), NIA (R01AG047318) and the Simmons Comprehensive Cancer Center to M.B. and by NIGMS (R01GM104007) to J.L.Y, as well as NIGMS (R01GM127569) to J.L.Y. and M.B. J.F.A. was supported by grant RP110394 from the Cancer Prevention and Research Institute of Texas and a Consortium Research Grant from the St. Baldrick's Foundation. The EM core at UT Southwestern is supported by NIH S10 OD020103-01.

REFERENCES

Abdu U, Gonzalez-Reyes A, Ghabrial A, and Schupbach T (2003). The Drosophila spn-D gene encodes a RAD51C-like protein that is required exclusively during meiosis. Genetics 165, 197–204. [PubMed: 14504227]

Arribere JA, Bell RT, Fu BX, Artiles KL, Hartman PS, and Fire AZ (2014). Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics 198, 837–846. [PubMed: 25161212]

Banani SF, Lee HO, Hyman AA, and Rosen MK (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18, 285–298. [PubMed: 28225081]

Batool A, Karimi N, Wu XN, Chen SR, and Liu YX (2019). Testicular germ cell tumor: a comprehensive review. Cell Mol Life Sci 76, 1713–1727. [PubMed: 30671589]

Borgermann N, Ackermann L, Schwertman P, Hendriks IA, Thijssen K, Liu JC, Lans H, Nielsen ML, and Mailand N (2019). SUMOylation promotes protective responses to DNA-protein crosslinks. EMBO J 38.

Brenner S (1974). The genetics of Caenorhabditis elegans. Genetics 77, 71–94. [PubMed: 4366476]

Carmell MA, Dokshin GA, Skaletsky H, Hu YC, van Wolfswinkel JC, Igarashi KJ, Bellott DW, Nefedov M, Reddien PW, Enders GC, et al. (2016). A widely employed germ cell marker is an ancient disordered protein with reproductive functions in diverse eukaryotes. Elife 5.

Colaiacovo MP, Stanfield GM, Reddy KC, Reinke V, Kim SK, and Villeneuve AM (2002). A targeted RNAi screen for genes involved in chromosome morphogenesis and nuclear organization in the Caenorhabditis elegans germline. Genetics 162, 113–128. [PubMed: 12242227]

Davis EJ, Lachaud C, Appleton P, Macartney TJ, Nathke I, and Rouse J (2012). DVC1 (C1orf124) recruits the p97 protein segregase to sites of DNA damage. Nat Struct Mol Biol 19, 1093–1100. [PubMed: 23042607]

Davis GM, Dokshin GA, Sawle AD, Eldridge MD, Romer KA, Gourley TE, Molesworth LW, Tatnell HR, Ozturk AR, de Rooij DG, et al. (2019). GCNA interacts with Spartan and Topoisomerase II to regulate genome stability. Dev Cell XXX, XXXX.

Delabaere L, Orsi GA, Sapey-Triomphe L, Horard B, Couble P, and Loppin B (2014). The Spartan ortholog maternal haploid is required for paternal chromosome integrity in the Drosophila zygote. Curr Biol 24, 2281–2287. [PubMed: 25242033]

Dernburg AF, McDonald K, Moulder G, Barstead R, Dresser M, and Villeneuve AM (1998). Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell 94, 387–398. [PubMed: 9708740]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Ferree PM, and Barbash DA (2009). Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLoS Biol 7, e1000234. [PubMed: 19859525]

Fu BX, Hansen LL, Artiles KL, Nonet ML, and Fire AZ (2014). Landscape of target:guide homology effects on Cas9-mediated cleavage. Nucleic Acids Res 42, 13778–13787. [PubMed: 25399416]

Gagnon JA, Valen E, Thyme SB, Huang P, Akhmetova L, Pauli A, Montague TG, Zimmerman S, Richter C, and Schier AF (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS One 9, e98186. [PubMed: 24873830]

Gemkow MJ, Dichter J, and Arndt-Jovin DJ (2001). Developmental regulation of DNA-topoisomerases during Drosophila embryogenesis. Exp Cell Res 262, 114–121. [PubMed: 11139335]

Ghabrial A, Ray RP, and Schupbach T (1998). okra and spindle-B encode components of the RAD52 DNA repair pathway and affect meiosis and patterning in Drosophila oogenesis. Genes Dev 12, 2711–2723. [PubMed: 9732269]

Hillers KJ, Jantsch V, Martinez-Perez E, and Yanowitz JL (2017). Meiosis. WormBook 2017, 1–43.

Hodgkin J, Horvitz HR, and Brenner S (1979). Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS. Genetics 91, 67–94. [PubMed: 17248881]

Hong A, Lee-Kong S, Iida T, Sugimura I, and Lilly MA (2003). The p27cip/kip ortholog dacapo maintains the Drosophila oocyte in prophase of meiosis I. Development 130, 1235–1242. [PubMed: 12588841]

Hoogewijs D, Houthoofd K, Matthijssens F, Vandesompele J, and Vanfleteren JR (2008). Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in C. elegans. BMC Mol Biol 9, 9. [PubMed: 18211699]

Inaba M, Buszczak M, and Yamashita YM (2015). Nanotubes mediate niche-stem-cell signalling in the Drosophila testis. Nature 523, 329–332. [PubMed: 26131929]

Jang JK, Sherizen DE, Bhagat R, Manheim EA, and McKim KS (2003). Relationship of DNA double-strand breaks to synapsis in Drosophila. J Cell Sci 116, 3069–3077. [PubMed: 12799415]

Janning W (1978). Gynandromorph fate maps in Drosophila. Results Probl Cell Differ 9, 1–28. [PubMed: 373036]

Jaramillo-Lambert A, Fabritius AS, Hansen TJ, Smith HE, and Golden A (2016). The Identification of a Novel Mutant Allele of topoisomerase II in Caenorhabditis elegans Reveals a Unique Role in Chromosome Segregation During Spermatogenesis. Genetics 204, 1407–1422. [PubMed: 27707787]

Jin Y, Tam OH, Paniagua E, and Hammell M (2015). TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics 31, 3593–3599. [PubMed: 26206304]

Keeney S, Giroux CN, and Kleckner N (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88, 375–384. [PubMed: 9039264]

Kiianitsa K, and Maizels N (2013). A rapid and sensitive assay for DNA-protein covalent complexes in living cells. Nucleic Acids Res 41, e104. [PubMed: 23519618]

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, and Salzberg SL (2013a). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14, R36. [PubMed: 23618408]

Kim Y, Kweon J, and Kim JS (2013b). TALENs and ZFNs are associated with different mutation signatures. Nat Methods 10, 185. [PubMed: 23396284]

Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, and Wilson RK (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22, 568–576. [PubMed: 22300766]

Korkola JE, Heck S, Olshen AB, Feldman DR, Reuter VE, Houldsworth J, Bosl GJ, and Chaganti RS (2015). Development and Validation of a Gene-Based Model for Outcome Prediction in Germ Cell Tumors Using a Combined Genomic and Expression Profiling Approach. PLoS One 10, e0142846. [PubMed: 26624623]

Korkola JE, Houldsworth J, Feldman DR, Olshen AB, Qin LX, Patil S, Reuter VE, Bosl GJ, and Chaganti RS (2009). Identification and validation of a gene expression signature that predicts outcome in adult men with germ cell tumors. J Clin Oncol 27, 5240–5247. [PubMed: 19770384]

Kruisselbrink E, Guryev V, Brouwer K, Pontier DB, Cuppen E, and Tijsterman M (2008). Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCJ-defective C. elegans. Curr Biol 18, 900–905. [PubMed: 18538569]

Kurimoto K, and Saitou M (2018). Epigenome regulation during germ cell specification and development from pluripotent stem cells. Curr Opin Genet Dev 52, 57–64. [PubMed: 29908427]

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. [PubMed: 22388286]

Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. [PubMed: 19451168]

Liu H, Jang JK, Kato N, and McKim KS (2002). mei-P22 encodes a chromosome-associated protein required for the initiation of meiotic recombination in Drosophila melanogaster. Genetics 162, 245–258. [PubMed: 12242237]

Livak KJ, and Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408. [PubMed: 11846609]

Lopez-Mosqueda J, Maddi K, Prgomet S, Kalayil S, Marinovic-Terzic I, Terzic J, and Dikic I (2016). SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. Elife 5.

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. [PubMed: 25516281]

Lu WJ, Chapo J, Roig I, and Abrams JM (2010). Meiotic recombination provokes functional activation of the p53 regulatory network. Science 328, 1278–1281. [PubMed: 20522776]

Magdalou I, Lopez BS, Pasero P, and Lambert SA (2014). The causes of replication stress and their consequences on genome stability and cell fate. Semin Cell Dev Biol 30, 154–164. [PubMed: 24818779]

Mani SR, Megosh H, and Lin H (2014). PIWI proteins are essential for early Drosophila embryogenesis. Dev Biol 385, 340–349. [PubMed: 24184635]

Marton RF, Thommes P, and Cotterill S (1994). Purification and characterisation of dRP-A: a single-stranded DNA binding protein from Drosophila melanogaster. FEBS letters 342, 139–144. [PubMed: 8143866]

McKim KS, and Hayashi-Hagihara A (1998). mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes Dev 12, 2932–2942. [PubMed: 9744869]

Mehrotra S, and McKim KS (2006). Temporal analysis of meiotic DNA double-strand break formation and repair in Drosophila females. PLoS Genet 2, e200. [PubMed: 17166055]

Morocz M, Zsigmond E, Toth R, Enyedi MZ, Pinter L, and Haracska L (2017). DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA. Nucleic Acids Res 45, 3172–3188. [PubMed: 28053116]

Mosbech A, Gibbs-Seymour I, Kagias K, Thorslund T, Beli P, Povlsen L, Nielsen SV, Smedegaard S, Sedgwick G, Lukas C, et al. (2012). DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. Nat Struct Mol Biol 19, 1084–1092. [PubMed: 23042605]

Neale MJ, Pan J, and Keeney S (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. Nature 436, 1053–1057. [PubMed: 16107854]

Nishida KM, Saito K, Mori T, Kawamura Y, Nagami-Okada T, Inagaki S, Siomi H, and Siomi MC (2007). Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. RNA (New York, NY) 13, 1911–1922.

Nott TJ, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, Craggs TD, Bazett-Jones DP, Pawson T, Forman-Kay JD, et al. (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell 57, 936–947. [PubMed: 25747659]

Page SL, and Hawley RS (2001). c(3)G encodes a Drosophila synaptonemal complex protein. Genes Dev 15, 3130–3143. [PubMed: 11731477]

Paix A, Folkmann A, and Seydoux G (2017). Precision genome editing using CRISPR-Cas9 and linear repair templates in C. elegans. Methods 121–122, 86–93.

Pritesh K, and Roland D (2018). Germ Cell Specification: The Evolution of a Recipe to Make Germ Cells.

Roberts A, Trapnell C, Donaghey J, Rinn JL, and Pachter L (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol 12, R22. [PubMed: 21410973]

Sander M, and Hsieh T (1983). Double strand DNA cleavage by type II DNA topoisomerase from Drosophila melanogaster. The Journal of biological chemistry 258, 8421–8428. [PubMed: 6305984]

Seydoux G (2018). The P Granules of C. elegans: A Genetic Model for the Study of RNA-Protein Condensates. J Mol Biol.

Sidi S, Sanda T, Kennedy RD, Hagen AT, Jette CA, Hoffmans R, Pascual J, Imamura S, Kishi S, Amatruda JF, et al. (2008). Chk1 suppresses a caspase-2 apoptotic response to DNA damage that bypasses p53, Bcl-2, and caspase-3. Cell 133, 864–877. [PubMed: 18510930]

Smith J, Calidas D, Schmidt H, Lu T, Rasoloson D, and Seydoux G (2016). Spatial patterning of P granules by RNA-induced phase separation of the intrinsically-disordered protein MEG-3. Elife 5.

Soper SF, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, and Bortvin A (2008). Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. Dev Cell 15, 285–297. [PubMed: 18694567]

Staeva-Vieira E, Yoo S, and Lehmann R (2003). An essential role of DmRad51/SpnA in DNA repair and meiotic checkpoint control. EMBO J 22, 5863–5874. [PubMed: 14592983]

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, et al. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144, 27–40. [PubMed: 21215367]

Stingele J, Bellelli R, Alte F, Hewitt G, Sarek G, Maslen SL, Tsutakawa SE, Borg A, Kjaer S, Tainer JA, et al. (2016). Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. Mol Cell 64, 688–703. [PubMed: 27871365]

Stingele J, Bellelli R, and Boulton SJ (2017). Mechanisms of DNA-protein crosslink repair. Nat Rev Mol Cell Biol 18, 563–573. [PubMed: 28655905]

Tang WW, Kobayashi T, Irie N, Dietmann S, and Surani MA (2016). Specification and epigenetic programming of the human germ line. Nat Rev Genet 17, 585–600. [PubMed: 27573372]

Tang X, Cao J, Zhang L, Huang Y, Zhang Q, and Rong YS (2017). Maternal Haploid, a Metalloprotease Enriched at the Largest Satellite Repeat and Essential for Genome Integrity in Drosophila Embryos. Genetics 206, 1829–1839. [PubMed: 28615282]

Tastan OY, Maines JZ, Li Y, McKearin DM, and Buszczak M (2010). Drosophila ataxin 2-binding protein 1 marks an intermediate step in the molecular differentiation of female germline cysts. Development 137, 3167–3176. [PubMed: 20724451]

Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, and Beck S (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 29, 189–196. [PubMed: 23175756]

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, and Pachter L (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28, 511–515. [PubMed: 20436464]

Triche TJ Jr., Weisenberger DJ, Van Den Berg D, Laird PW, and Siegmund KD (2013). Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res 41, e90. [PubMed: 23476028]

Uebel CJ, Anderson DC, Mandarino LM, Manage KI, Aynaszyan S, and Phillips CM (2018). Distinct regions of the intrinsically disordered protein MUT-16 mediate assembly of a small RNA amplification complex and promote phase separation of Mutator foci. PLoS Genet 14, e1007542. [PubMed: 30036386]

Vaz B, Popovic M, Newman JA, Fielden J, Aitkenhead H, Halder S, Singh AN, Vendrell I, Fischer R, Torrecilla I, et al. (2016). Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. Mol Cell 64, 704–719. [PubMed: 27871366]

Vaz B, Popovic M, and Ramadan K (2017). DNA-Protein Crosslink Proteolysis Repair. Trends Biochem Sci 42, 483–495. [PubMed: 28416269]

Wagner CR, Kuervers L, Baillie DL, and Yanowitz JL (2010). xnd-1 regulates the global recombination landscape in Caenorhabditis elegans. Nature 467, 839–843. [PubMed: 20944745]

Walport LJ, Hopkinson RJ, and Schofield CJ (2012). Mechanisms of human histone and nucleic acid demethylases. Curr Opin Chem Biol 16, 525–534. [PubMed: 23063108]

Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, Johnson CS, Smiraglia DJ, and Liu S (2012). IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. Bioinformatics 28, 729–730. [PubMed: 22253290]

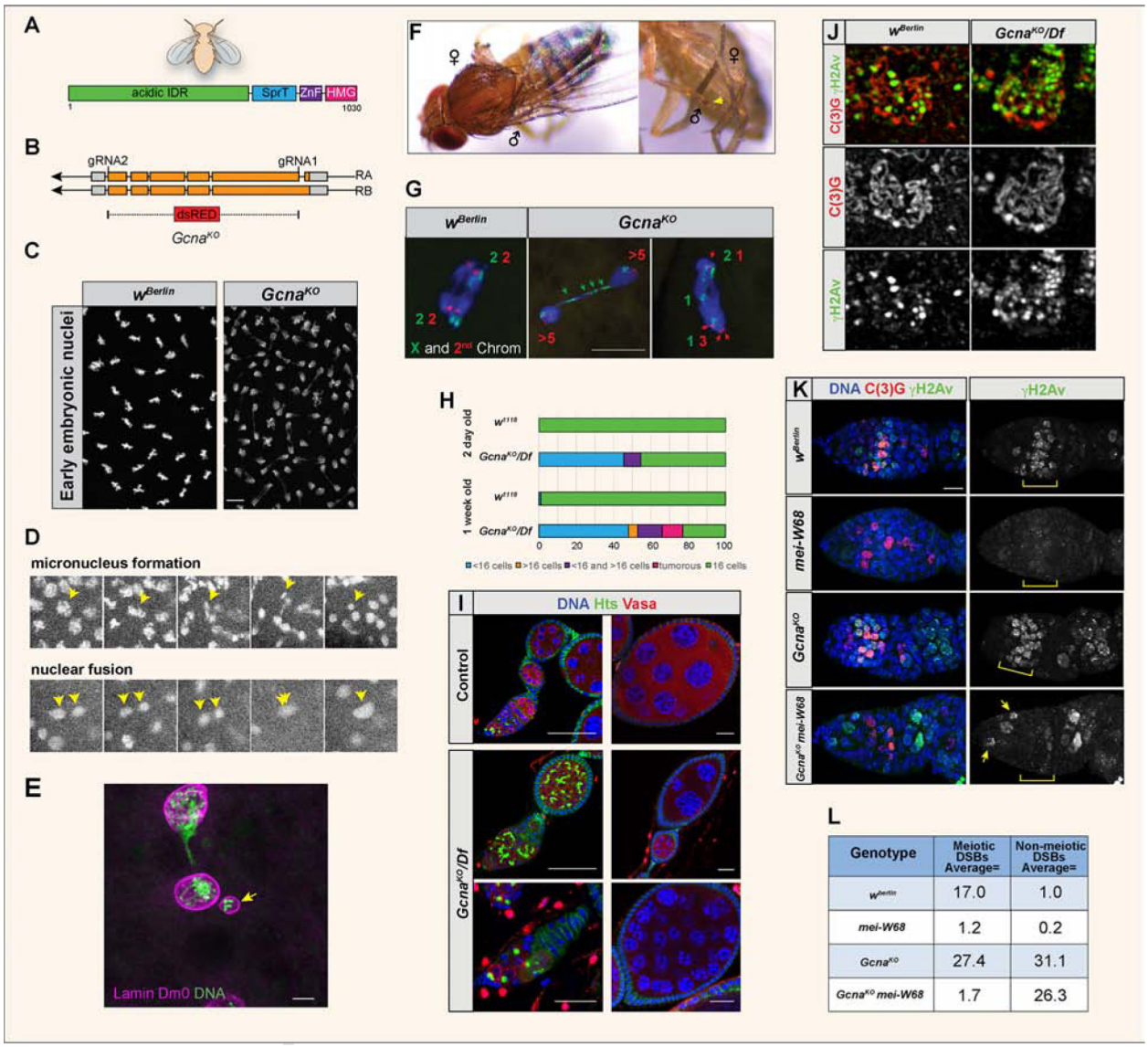
Westerfield M (2007). The Zebrafish Book A Guide for the Laboratory Use of Zebrafish (Danio rerio), Vol 5th Edition (Eugene: University of Oregon Press).

Wylie A, Lu WJ, D'Brot A, Buszczak M, and Abrams JM (2014). p53 activity is selectively licensed in the Drosophila stem cell compartment. Elife 3, e01530. [PubMed: 24618896]

Youds JL, Barber LJ, Ward JD, Collis SJ, O'Neil NJ, Boulton SJ, and Rose AM (2008). DOG-1 is the Caenorhabditis elegans BRIP1/FANCJ homologue and functions in interstrand cross-link repair. Mol Cell Biol 28, 1470–1479. [PubMed: 18086896]

Youds JL, O'Neil NJ, and Rose AM (2006). Homologous recombination is required for genome stability in the absence of DOG-1 in Caenorhabditis elegans. Genetics 173, 697–708. [PubMed: 16547095]

Zhang BN, Bueno Venegas A, Hickson ID, and Chu WK (2018). DNA replication stress and its impact on chromosome segregation and tumorigenesis. Semin Cancer Biol.



Genotype	Meiotic DSBs Average=	Non-meiotic DSBs Average=
w^{Berlin}	17.0	1.0
mei-W68	1.2	0.2
$Gcna^{KO}$	27.4	31.1
$Gcna^{KO}$ *mei-W68*	1.7	26.3

Figure 1. Loss of *Gcna* results in chromosome instability in *Drosophila*.

(**A**) Domain organization of the *Drosophila* Gcna protein.

(**B**) *Drosophila Gcna* gene locus showing the two isoforms and design of the $Gcna^{KO}$ allele with insertion of a dsRED cassette. UTRs in grey, exons in orange.

(**C**) DAPI staining of control (w^{Berlin}) and $Gcna^{KO}$ (maternal null) embryos reveals mitotic defects caused by loss of maternal *Gcna*. Scale bar represents 10μM.

(D) Still images from movies of $Gcna^{KO}$ maternal-null embryos carrying a *HistoneH2Av-mRFP* transgene to visualize chromosomes. Yellow arrowheads indicate micronucleus formation (top) and nuclear fusion event (bottom).

(E) Early *Gcna* maternal mutant *Drosophila* embryo stained for Lamin Dm0 (magenta) and DNA (green) to show that a nuclear envelope forms around a micronucleus (yellow arrow). Scale bar represents 5 μm.

(**F**) Gynandromorph phenotypes in progeny of $Gcna^{KO}$ females. Left, note the line of dark and light pigmentation bisecting the thorax, reflecting the presence or absence of the *yellow*

marker carried on one of the X chromosomes. Right, sex combs are seen on one forelimb (yellow arrow) but are missing from the other, reflecting adoption of male and female fates respectively.

(**G**) Embryonic nuclei from control (w^{Berlin}) and $Gcna^{KO}$ mutant females were labeled with FISH probes to the X (green) and second (red) chromosomes. The green and red numbers refer to number of X and second chromosome-specific foci observed in each half of the dividing nuclei. Controls show chromosomes dividing equally into daughter cells. $Gcna^{KO}$ cells have aberrant chromosome numbers and lagging chromosomes (green arrows). Scale bar represents 5 μm

(**H**) Quantification and **(I)** corresponding images of *Drosophila* ovarian phenotypes. Hts marks the fusome; Vasa marks germ cells. Control (w^{1118}) ovarioles contain the expected 16-cell cysts. $Gcna^{KO}/Df$ ovarioles have many cysts with abnormal numbers of cells, as well as tumors. The frequency of these phenotypes worsens with age. Scale bars represent 20 μm.

(**J**) SIM images of w^{Berlin} and $Gcna^{KO}$ mutant meiotic nuclei stained for synaptonemal complex marker C(3)G (red) and DNA damage marker γH2Av (green).

(**K**) w^{Berlin}, $Gcna^{KO}$, $mei\text{-}W86^{0949}$, and $Gcna^{KO}$; $mei\text{-}W68^{0949}$ ovarioles stained for C(3)G (red), γH2Av (green), and DNA (DAPI, blue). Scale bar represents 10 μm.

(**L**) Average number of γH2Av labeled foci per nucleus in the indicated genotypes. 10 meiotic and non-meiotic nuclei were examined per genotype.

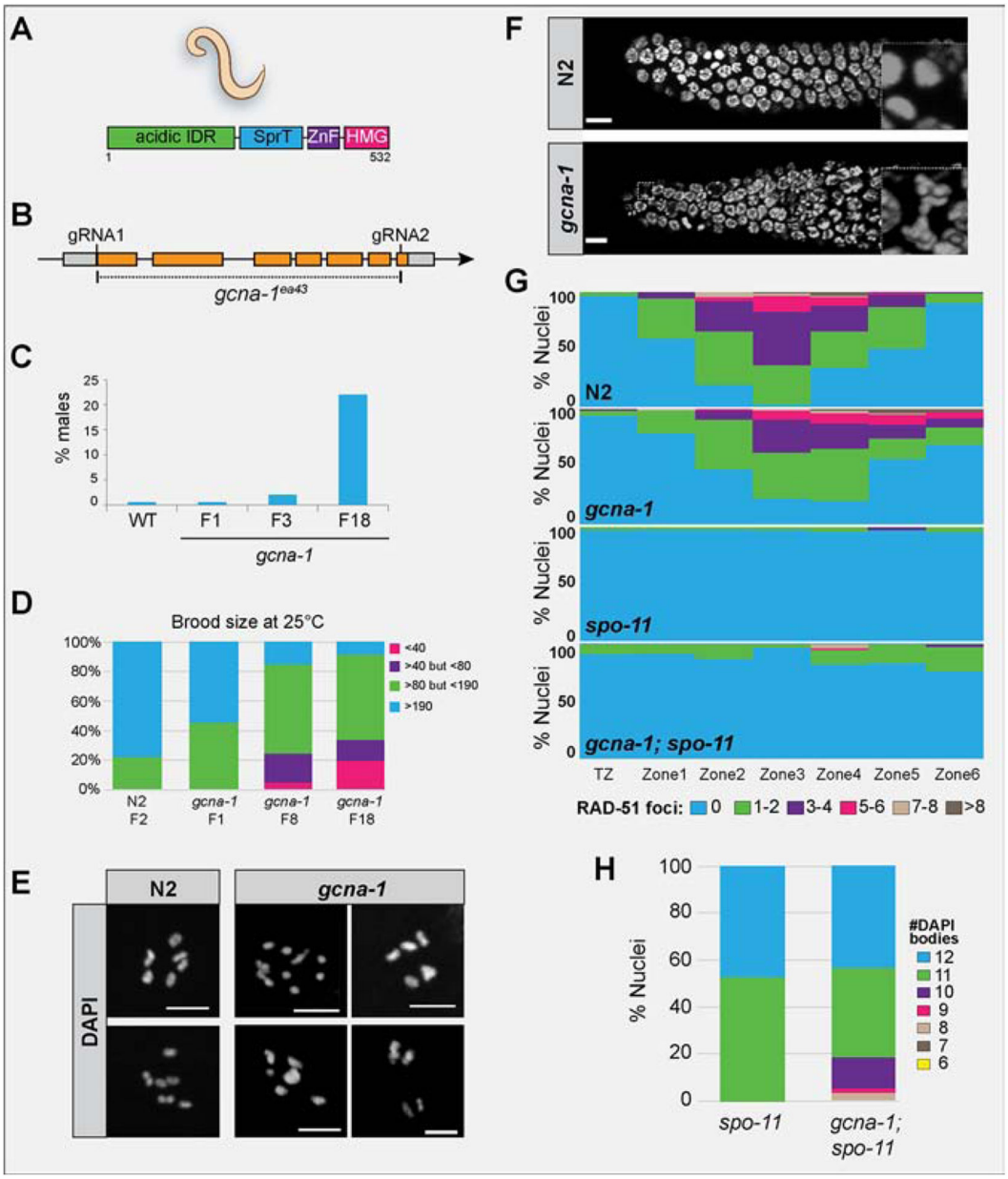


Figure 2. Loss of *gcna-1* results in genomic instability in *C. elegans*.

(**A**) Domain structure of *C. elegans* GCNA-1 protein.

(**B**) The *C. elegans gcna-1* locus detailing the design of the *gcna-1(ea43)* allele.

(**C**) The HIM phenotype increases over generations (n=12 parental worms for N2 (wild type control) and n= 21, 21, 34 for *gcna-1* F1, F3, and F18, respectively)

(**D**) *gcna-1* mutant worms exhibit decreases in brood size with successive passaging (n=12 parental worms for wt, and 34 each for *gcna-1* F1, F8, and F18).

(**E**) DAPI staining of diakinesis nuclei from control (N2, wild type) and *gcna-1* mutants. Controls showed the expected 6 DAPI^{+} bivalents; *gcna-1* mutant nuclei contained mixtures of bivalents, univalents, and fused chromosomes. Scale bars represent 5 μm.

(**F**) Mitotic germ cells from *gcna-1* mutants exhibit chromosome bridges as seen by DAPI-stained DNA. Insets are examples of nuclei in anaphase. Scale bars represent 20 μm.
(**G**) Quantification of meiotic (leptotene through onset of diplotene) RAD-51 foci in *N2, gcna-1, spo-11*, and *gcna-1;spo-11* mutant *C. elegans* germ lines (n = 3 germ lines/genotype).
(**H**) Quantification of DAPI positive bodies in diakinesis-stage, −1 oocytes of *C. elegans* (*spo-11*, n= 20; *gcna-1;spo-11*, n=42).

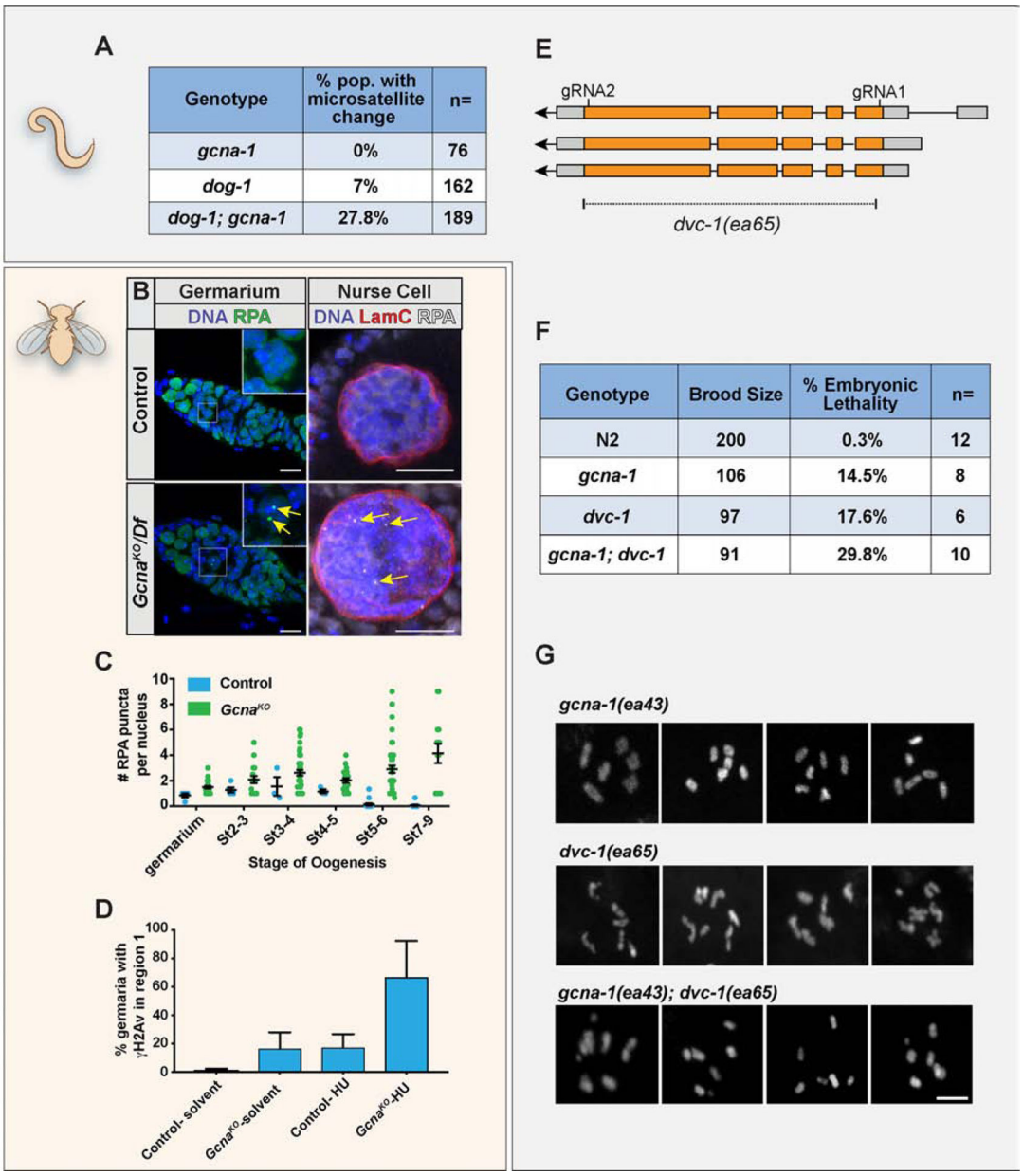


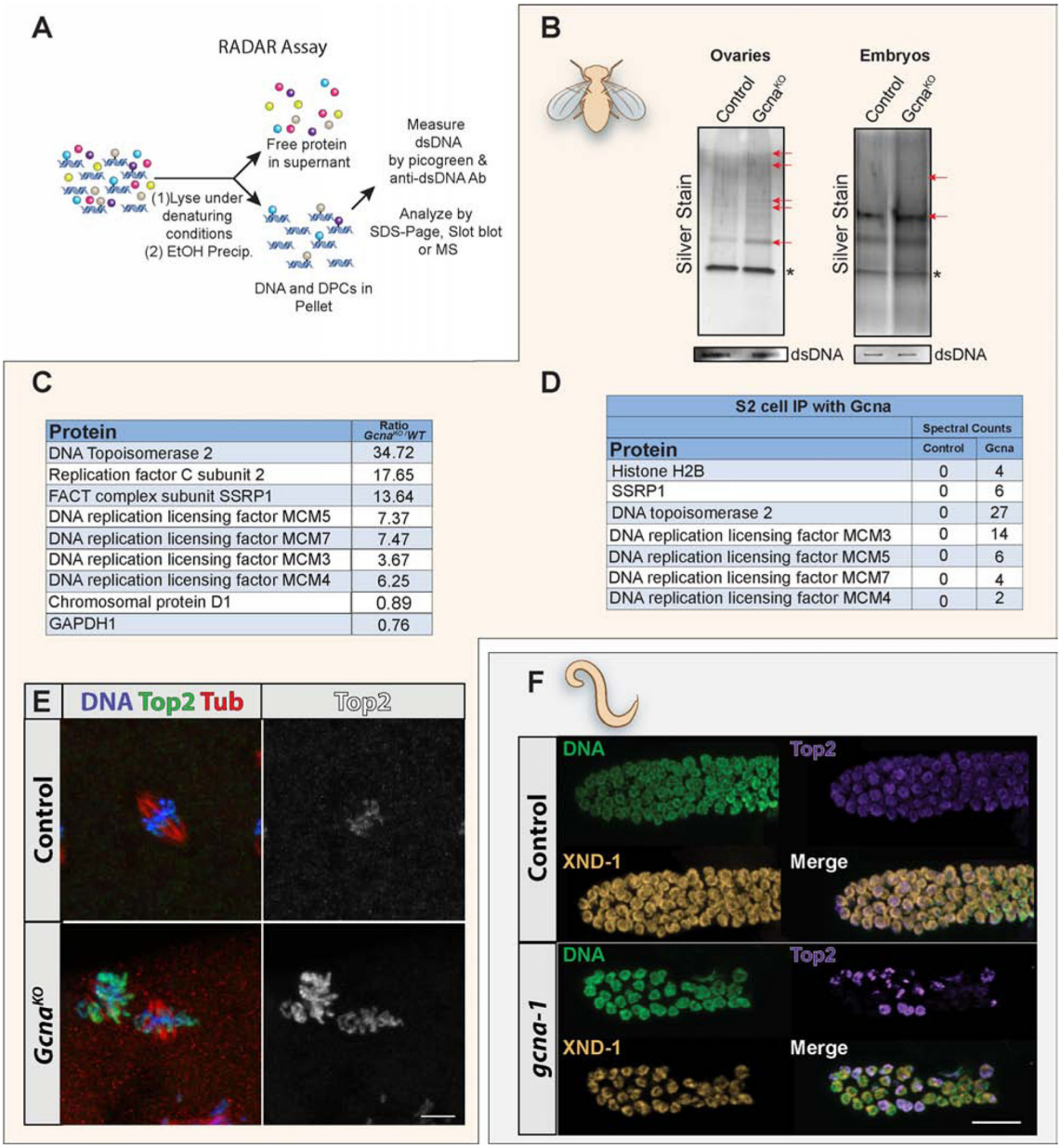
Figure 3. *GCNA* mutant germ cells experience replication stress

(**A**) Frequency of microsatellite repeat length changes within the *C. elegans vab-1* locus. (**B-C**) RPA foci (yellow arrows) accumulate in *Drosophila* $Gcna^{KO}$ germ cells. (**B**) Control and $Gcna^{KO}/Df$ germaria (left) and nurse cell nuclei (right) stained for DNA (blue), RPA (green), and Lamin C (red) as indicated. Scale bars represent 10 μm. (**C**) Quantification of RPA punctae per nucleus.

(**D**) Quantification of γH2Av foci in region 1 of the *Drosophila* germarium in control and $Gcna^{KO}$ mutant germaria in response to HU treatment.

(**E**) The *dvc-1/Spartan* locus of *C. elegans* showing the strategy for creating a complete null, *dvc-1(ea65)*.

(**F**) Brood analysis (total viable adult offspring) and embryonic lethality associated with N2, *dvc-1* and *gcna-1* single mutants, and *dvc-1;gcna-1* double mutant worms.
(**G**) Examples of DAPI-stained, diakinesis oocytes from *dvc-1*, *gcna-1*, and *dvc-1; gcna-1* mutant germ cells show chromosome fragmentation in the double mutant. Scale bars represent 5 μm.



Protein	Ratio $Gcna^{KO}/WT$
DNA Topoisomerase 2	34.72
Replication factor C subunit 2	17.65
FACT complex subunit SSRP1	13.64
DNA replication licensing factor MCM5	7.37
DNA replication licensing factor MCM7	7.47
DNA replication licensing factor MCM3	3.67
DNA replication licensing factor MCM4	6.25
Chromosomal protein D1	0.89
GAPDH1	0.76

S2 cell IP with Gcna		
	Spectral Counts	
Protein	Control	Gcna
Histone H2B	0	4
SSRP1	0	6
DNA topoisomerase 2	0	27
DNA replication licensing factor MCM3	0	14
DNA replication licensing factor MCM5	0	6
DNA replication licensing factor MCM7	0	4
DNA replication licensing factor MCM4	0	2

Figure 4. *GCNA* mutants exhibit increased DPC levels

(**A**) Schematic illustrating principle of the RADAR assay.

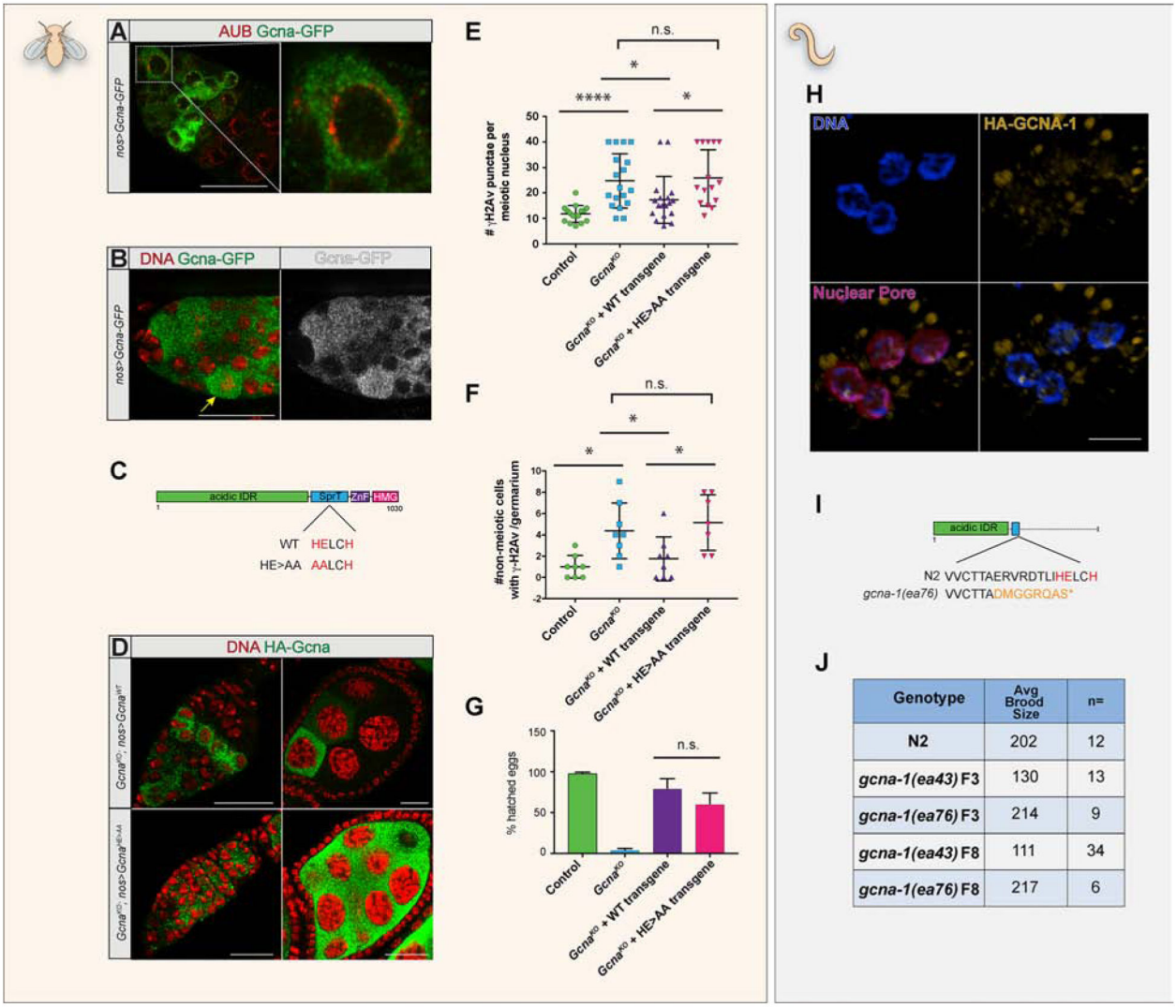
(**B**) RADAR was performed on *Drosophila* ovaries and embryos. Samples were normalized to dsDNA, separated by SDS-PAGE, and silver stained. Red arrows mark examples of differentially enhanced bands in the *Gcna* mutant samples. * marks benzonase added to the prep.

(**C**) Proteins enriched in DPCs from $Gcna^{KO}$ *Drosophila* embryos were analyzed by mass spectrometry and presented as enriched over wild-type controls.

(**D**) List of proteins that co-immunoprecipate with Gcna expressed in *Drosophila* S2 cells.

(**E, F**) Top2 is mislocalized in *Gcna* mutant flies and worms. (**E**) Early *Drosophila* embryos derived from control and $Gcna^{KO}$ females stained for DNA (blue), Top2 (green) and Tubulin

(Tub, red). Grayscale shows Top2 alone. Scale bar represents 5 μm. (**F**) Control (N2) and *gcna-1* mutant *C. elegans* gonads stained for DNA (green), TOP-2 (magenta), and staining control XND-1 (yellow). Scale bar represents 20 μm.



Genotype	Avg Brood Size	n=
N2	202	12
gcna-1(ea43) F3	130	13
gcna-1(ea76) F3	214	9
gcna-1(ea43) F8	111	34
gcna-1(ea76) F8	217	6

Figure 5. Functional analysis of GCNA domain structure

(**A, B**) Expression of a *UAS-Gcna*^{WT} transgene tagged at the C-terminus driven by *nanos* (*nos*)-*gal4* shows prominent cytoplasmic staining. Germaria co-labeled for the Gcna-GFP transgene and (**A**) the nuage component, Aubergine (Red) or (**B**) DAPI (red) to visualize DNA. Yellow arrow points to a mitotic germ cell. Scale bars represent 20 μm.

(**C**) Structure of the *Drosophila Gcna*^{WT} and catalytic dead, *Gcna*$^{HE>AA}$, mutant transgenes.

(**D**) Comparison the Gcna^{WT} and Gcna$^{HE>AA}$ proteins tagged at their N-termini. Note the abundant Gcna$^{HE>AA}$ punctae in the germarium and germ cells. DNA (red), Gcna (anti-HA, green). Scale bars represent 20 μm.

(**E-G**) The catalytic dead transgene confers a separation-of-function phenotype and (**E, F**) cannot suppress the accumulation of γH2Av foci in (**E**) meiotic and (**F**) non-meiotic nuclei from wild-type, *Gcna*^{KO}, *Gcna*^{KO}; *nos>Gcna*^{WT}, and *Gcna*^{KO}; *nos>Gcna*$^{HE>AA}$ ovaries, but (**G**) can partially rescue the maternal-effect lethality conferred by loss of *Gcna*. Plotted are the percentage of eggs derived from females of the indicated genotypes that hatch into larvae.

(**H**) Expression of *C. elegans* OLLAS::GCNA-1. Germ cells are labeled for DNA (blue), GCNA-1 (anti-OLLAS, yellow), and the nuclear pore antibody mAb414 (magenta). Scale bar represents 5 μm.
(**I**) Structure of the *C. elegans* "IDR-only", C-terminal truncation allele, *gcna-1(ea76)*.
(**J**) Comparison of brood sizes from N2 control, and homozygous, F3 and F8 *gcna-1(ea43)* null and *gcna-1(ea76)* truncation mutants. The IDR-only mutation does not exhibit the rapid transgenerational sterility seen in the null.

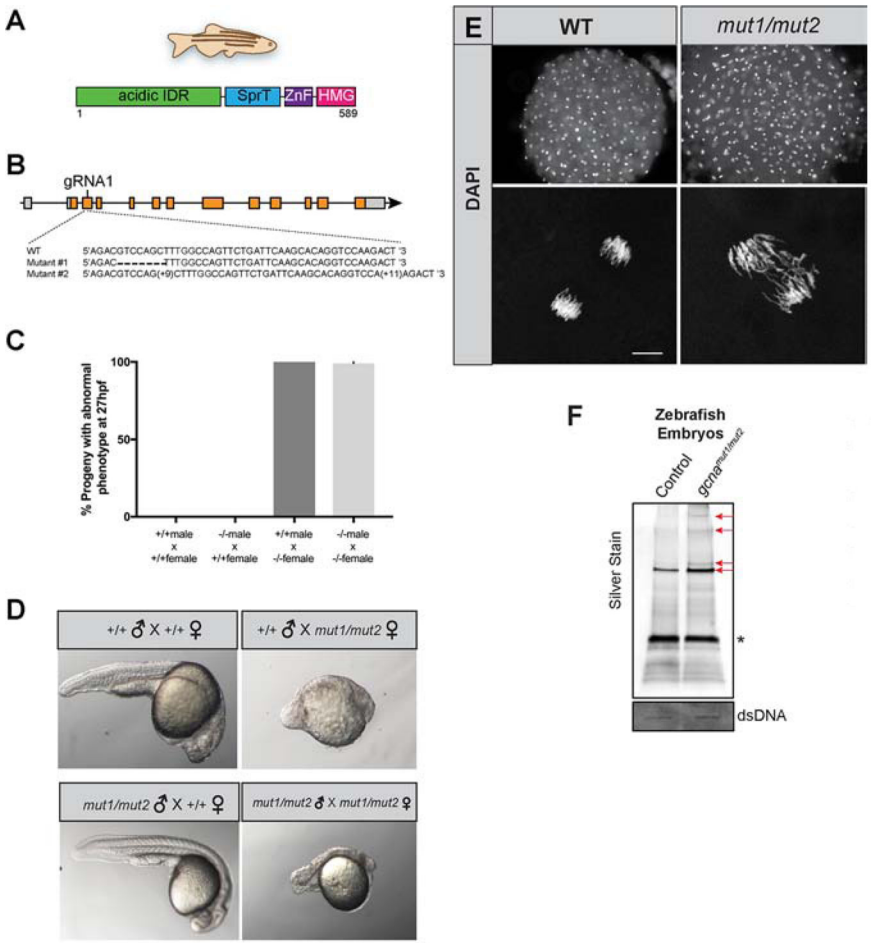


Figure 6. GCNA function is conserved in vertebrates

(**A**) Domain structure of the zebrafish GCNA protein.

(**B**) Organization of the zebrafish *gcna* gene showing the lesions in the *mut1* and *mut2* alleles.

(**C, D**) Zebrafish embryos derived from mutant *gcna* mutant females exhibit profound maternal-effect defects, regardless of the paternal genotype. (**C**) Quantification of defects and (**D**) Representative embryos from indicated crosses.

(**E**) Fixed, DAPI-stained embryos derived from mutant *gcna* mutant females exhibit asynchronous divisions and extensive chromosome bridges during early development. Scale bar represents 10 μm.

(**F**) RADAR was performed on zebrafish embryos, as described in Figure 4A, B. Samples were normalized to dsDNA input, separated by SDS-PAGE, and silver stained. Red arrows mark bands increased in *gcna* mutants. * marks benzonase.

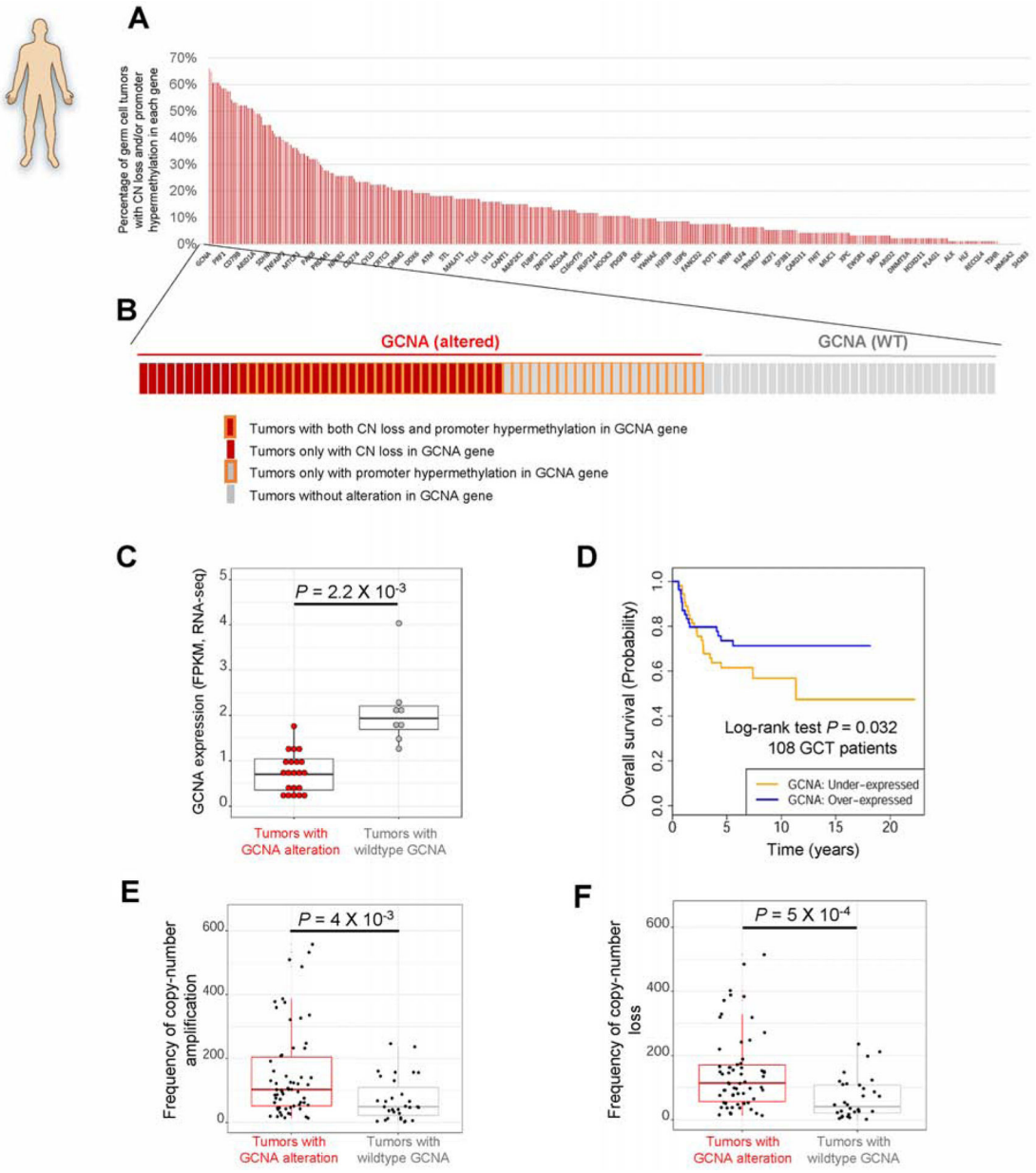


Figure 7. *GCNA* is frequently altered in pediatric germ cell tumors (GCTs), and is associated with poor patient survival and genomic instability.

(**A**) *GCNA* genes display the most frequent copy-number loss and promoter hypermethylation in 94 childhood patients with GCTs.

(**B**) Frequency of copy-number loss and promoter hypermethylation of *GCNA* gene in pediatric GCTs.

(**C**) Low GCNA expression in pediatric GCTs is found in tumors with alterations of the *GCNA* locus.

(**D**) Low GCNA expression is significantly associated with poor survival of GCT patients.

(**E,F**) GCTs with copy number loss of *GCNA* exhibit elevated frequencies of both copy number (**E**) amplification and (**F**) loss relative to tumors with normal *GCNA* copy number.