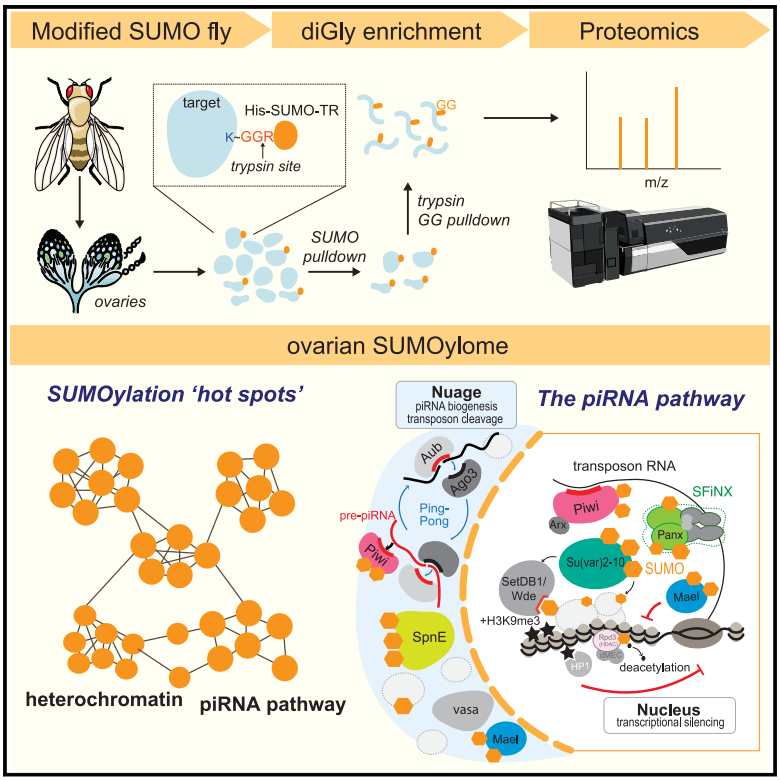
Cell Genomics

Resource

Pervasive SUMOylation of heterochromatin and piRNA pathway proteins

Graphical abstract



Highlights

- diGly proteomics identifies ~1,500 SUMO sites in ~840 proteins in the fly ovary
- SUMO targets are highly enriched among heterochromatin and piRNA pathway proteins
- Loss of SUMO disrupts the nuage
- Piwi is required for the SUMOylation of Spn-E and Mael, but not Panx, in the germline

Authors

Maria Ninova, Hannah Holmes, Brett Lomenick, Katalin Fejes Tóth, Alexei A. Aravin

Correspondence

mninova@ucr.edu (M.N.),
aaa@caltech.edu (A.A.A.)

In brief

Adapting diGly proteomics for SUMO target discovery, Ninova et al. identify numerous SUMOylated proteins among heterochromatin and piRNA pathway factors, indicating unexpectedly broad implications of protein modification by SUMO in cellular pathways governing heterochromatin regulation and transposon control.

Ninova et al., 2023, Cell Genomics *3*, 100329
July 12, 2023
https://doi.org/10.1016/j.xgen.2023.100329

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Resource

Pervasive SUMOylation of heterochromatin and piRNA pathway proteins

Maria Ninova,[1,4,*] Hannah Holmes,[1] Brett Lomenick,[3] Katalin Fejes Tóth,[2] and Alexei A. Aravin[2,*]
[1]Department of Biochemistry, University of California Riverside, 3401 Watkins Drive, Boyce Hall, Riverside, CA 92521, USA
[2]Division of Biology and Biological Engineering, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[3]Proteome Exploration Laboratory of the Beckman Institute, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[4]Lead contact
*Correspondence: mninova@ucr.edu (M.N.), aaa@caltech.edu (A.A.A.)
https://doi.org/10.1016/j.xgen.2023.100329

SUMMARY

Genome regulation involves complex protein interactions that are often mediated through post-translational modifications (PTMs). SUMOylation—modification by the small ubiquitin-like modifier (SUMO)—has been implicated in numerous essential processes in eukaryotes. In *Drosophila*, SUMO is required for viability and fertility, with its depletion from ovaries leading to heterochromatin loss and ectopic transposon and gene activation. Here, we developed a proteomics-based strategy to uncover the *Drosophila* ovarian "SUMOylome," which revealed that SUMOylation is widespread among proteins involved in heterochromatin regulation and different aspects of the Piwi-interacting small RNA (piRNA) pathway that represses transposons. Furthermore, we show that SUMOylation of several piRNA pathway proteins occurs in a Piwi-dependent manner. Together, these data highlight broad implications of protein SUMOylation in epigenetic regulation and indicate novel roles of this modification in the cellular defense against genomic parasites. Finally, this work provides a resource for the study of SUMOylation in other biological contexts in the *Drosophila* model.

INTRODUCTION

Post-translational modifications (PTMs) affect a broad range of cellular functions such as protein turnover, localization to different subcellular compartments, or specific interactions, and are utilized in the regulation of various molecular pathways. PTMs involve the covalent attachment of diverse moieties from small chemical groups to entire modifier proteins to the main polypeptide chain. The best-known protein modifier is ubiquitin, a ~9-kDa unit that gets conjugated to target lysine side chains via an isopeptide bond between its C-terminal carboxyl group and the lysine epsilon amino group. Following the discovery of ubiquitin, other small proteins that can act as modifiers have emerged, including the small ubiquitin-like modifier (SUMO). SUMO is a ~11-kDa protein that shares structural and sequence homology with ubiquitin and similarly gets covalently attached to target lysines. However, SUMO conjugation is mediated by different enzymes and serves distinct and non-redundant functions (reviewed in Flotho and Melchior[1]). In brief, the SUMOylation cascade involves activation by a dedicated E1 heterodimer followed by transfer to the SUMO E2 conjugating enzyme Ubc9. Ubc9 catalyzes SUMO transfer to the final acceptor lysine, which often (but not always) resides within a consensus motif ψKxE (ψ is a hydrophobic amino acid) and is sufficient for SUMOylation *in vitro*.[2–4] Nevertheless, non-catalytic SUMO E3 ligases can facilitate Ubc9 or enable substrate specificity, and they seem to be required for SUMOylation in some contexts and perhaps for non-consensus sites *in vivo*.[5–8] SUMO is primarily nuclear, and since its discovery has emerged as an important regulator of different nuclear processes (reviewed in Xhao[9]) such as transcription factor activity, DNA repair, rRNA biogenesis, chromosome organization, and segregation. Mechanistically, SUMOylation may lead to diverse consequences, including changes in protein conformation or localization, masking or competing with other PTMs, and, most famously, regulating protein-protein interactions. SUMO-mediated interactions typically involve an aliphatic stretch flanked by acidic amino acids (SUMO interacting motif [SIM]) in the partner protein and, although individually weak in their nature, multivalent SUMO-SIM interactions were proposed to stabilize large molecular complexes[10,11] and promote the formation of phase-separated compartments such as promyelocytic leukemia (PML) protein bodies.[12] However, despite being implicated in a myriad of biological processes, our understanding of SUMO's role within different molecular and cellular contexts is far from complete.

Previous work in *Drosophila* implicated the SUMO pathway in the regulation of heterochromatin establishment and Piwi-interacting small RNA (piRNA)-mediated transposon (transposable element, TE) silencing.[13–16] In germ cells, Piwi clade proteins (Piwi, Aub, and Ago3 in flies) and piRNAs cooperate in intimately linked processes that ensure transcriptional and

post-transcriptional transposon silencing and continuous production of mature piRNAs (reviewed in Czech and Hannon[17]). Mature piRNA production and post-transcriptional cleavage of transposon RNAs by Aub and Ago3 occur in a dedicated perinuclear structure, the nuage, similar to "germ granules" in other systems. Antisense piRNAs produced in the nuage also become loaded in Piwi, which then enters the nucleus to find transposon nascent RNA and enforce co-transcriptional silencing at target loci.[18–20] To date, SUMOylation is known to participate in the nuclear piRNA pathway in several ways. First, the SUMO E3 ligase Su(var)2-10 was found to interact with piRNA pathway and heterochromatin proteins and play an essential role in the recruitment of the enzymatic complex SetDB1/Wde, which deposits the silencing epigenetic mark H3K9me3,[14] as well as with the MEP-1/Mi-2 chromatin remodeler complex.[16] Second, SUMOylation of Panoramix (Panx)—a co-repressor required for H3K9me3 deposition downstream of Piwi—was found to mediate its interaction with the general heterochromatin effector Sov.[15] In addition to silencing of transposons, Su(var)2-10 and SUMO were found to control H3K9me3 deposition at piRNA-independent genomic loci such as developmentally silenced tissue-specific genes.[21] The pervasive effect of SUMO depletion on global H3K9me3 levels at diverse classes of genomic targets suggests that SUMOylation is a critical process in the regulation of repressive chromatin integrated within different regulatory pathways. However, further understanding of the role of SUMOylation on chromatin requires knowledge of the full spectrum of SUMO targets.

We developed a proteomics approach that enables the identification of SUMOylated proteins with amino acid-level site predictions from different tissues in the classic model for piRNA and heterochromatin studies, *Drosophila melanogaster*. Here, we report a comprehensive dataset of SUMO targets in the fly ovary. Notably, we identified strong enrichment of heterochromatin factors among SUMOylated proteins, supporting the notion that SUMO plays a complex role in heterochromatin regulation that extends to multiple targets and protein complexes. Moreover, we find a striking enrichment of proteins specific to the piRNA pathway among SUMO targets, including the central effector of epigenetic silencing by piRNAs, Piwi, and several proteins that localize to the nucleus and nuage. We further validated the SUMOylation of selected piRNA pathway factors—Piwi, Panx, Spindle-E (Spn-E), and Maelstrom (Mael)—in germ cells and showed that the modification of Mael and Spn-E, but not Panx, is Piwi dependent, indicative of multiple SUMO roles in distinct steps of transposon silencing. Altogether, our findings point to a previously unappreciated multilayered role of SUMOylation in the piRNA pathway and heterochromatin regulation that provide important clues toward our understanding of the molecular mechanisms of genome regulation and transposon control.

RESULTS

Establishing a system for the detection of SUMOylated proteins in *Drosophila*

Proteome-wide studies of PTMs have benefited from the development of methods and reagents that enable specific enrichment of modified proteins or peptides from total protein lysates. Basic methods for the enrichment of SUMO-modified proteins involve pull-down with anti-SUMO antibodies. However, this approach prohibits the use of stringent washing conditions and is therefore prone to high background. Furthermore, endogenous SUMO does not have trypsin cleavage sites close to its C terminus, and trypsin digestion of SUMOylated proteins generates large, branched peptides from modified regions that are incompatible with conventional bottom-up proteomics (Figure 1A). Accordingly, specific modified residues remain unknown and SUMOylation events can only be inferred indirectly from the abundance of other peptides from purified SUMOylated proteins.

To overcome these obstacles and obtain a high-confidence dataset of the SUMO-modified proteins (herein referred to as the SUMOylome) in *Drosophila*, we adapted an approach that allows stringent purification, enrichment, and proteomic detection of peptides containing a modified SUMO remnant.[22] This approach, originally developed for the study of SUMOylation in human cells,[22] employs ectopically expressed SUMO protein with 6xHis-tag and a point mutation (T > R substitution, herein referred to as SUMO-TR) that enables trypsin cleavage before the C-terminal GG motif at the conjugation site (Figures 1A and 1B). In a two-step purification process, SUMO-modified proteins are first selected based on the His tag by nickel affinity under denaturing conditions, which eliminates the activity of SUMO proteases and most background from noncovalently bound proteins. The resulting SUMOylated protein-enriched fraction is then trypsinized to generate a mixture of peptides including branched peptides with short di-glycine (diGly) remnant from cleaved SUMO-TR moiety. These diGly remnant-containing peptides are further enriched by pull-down with a specific antibody, purified, and analyzed by mass spectrometry (MS). As the diGly remnant alters the molecular mass and charge of a given peptide, diGly-modified species can be distinguished from other background peptides with high confidence. Of note, ubiquitin and other ubiquitin-like protein modifiers naturally have an arginine residue before the terminal diGly motif; therefore, any ubiquitinated protein that unspecifically co-purifies with SUMOylated proteins during the His-based enrichment step can also generate diGly peptides. However, this background can be accounted for through negative control samples from tissues that do not express SUMO-TR (Figure 1B). To enable this method for SUMOylation detection in *Drosophila* tissues, we created a transgenic line that carries a modified copy of the *smt3* locus (which encodes the single SUMO homolog in this species) where the SUMO coding sequence has a 6xHis N-terminal tag, T86 > R substitution before the C-terminal GG motif, and a ~2.5-kb upstream region containing the putative endogenous promoter. The His-tagged SUMO-TR protein was detectable by western blotting (WB) (Figure 1C) and, importantly, this transgene completely rescues the lethality of the null allele $smt3^{04493}$, confirming that it encodes a fully functional protein.

Identification and characterization of the *Drosophila* ovarian SUMOylome

We used flies expressing SUMO-TR to optimize the two-step purification procedure described above (Figure 1B) and obtain

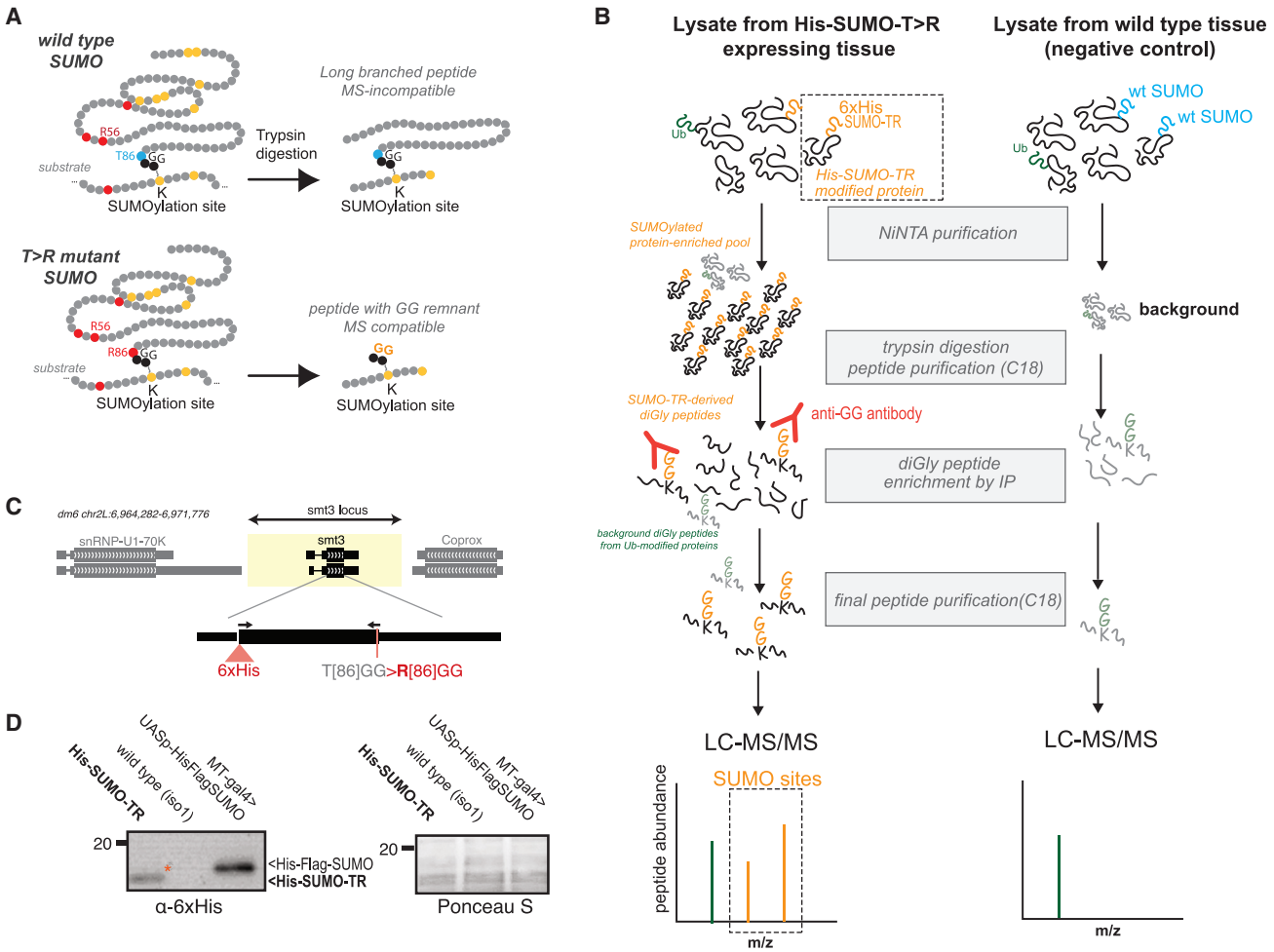


Figure 1. Proteomics approach for SUMO site detection

(A) Diagram of the peptides generated after trypsin digestion of proteins modified by wild-type and T > R mutant SUMO.

(B) Diagram of the experimental design and work flow for SUMOylation-derived diGly-modified peptide MS analysis.

(C) Diagram of the recombinant SUMO construct used to generate SUMO-TR expressing transgenic flies.

(D) Western blot of total ovary lysate confirming the expression of recombinant SUMO. Ovaries from flies overexpressing 6xHis-FLAG-tagged SUMO are used as a positive control.

SUMOylation site-derived peptides from ovarian tissue (see STAR Methods for details). Following this procedure, we performed three independent experiments, where each replica involved parallel sample preparation from SUMO-TR ovaries and wild-type controls followed by label-free tandem MS. Each sample yielded ~1,000–2,500 diGly sites (Figure 2A) and, altogether, we detected 3,159 exact diGly sites mapping to the protein products of 1,295 genes. In comparison, MS analyses of input samples prior to diGly enrichment yielded less than 100 diGly sites (data not shown), highlighting the importance of the diGly enrichment step for capturing SUMOylation sites. For each diGly site, we calculated the normalized intensity ratios between SUMO-TR and control samples. The ratios showed a prominent bimodal distribution: approximately half of the sites in each experiment were detected exclusively in SUMO-TR samples, indicative of genuine SUMOylation targets (Figure 2B). The diGly sites detected with similar intensities in SUMO-TR and control samples or biased to control samples likely originate from unspecifically bound ubiquitinated proteins. Consistent with this, motif analysis showed that diGly sites detected in SUMO-TR samples are enriched in the canonical SUMOylation motif, while diGly sites from the negative control do not have any motif enrichment (Figures 2B and S1).

The sets of predicted *bona fide* SUMO targets (SUMO-TR/control intensity ratio >10) were highly reproducible between the three experiments, with ~50% overlap on the specific site level, and a 75% overlap on the target protein-coding gene level (Figure 2C). Previously known SUMO targets, including RanGAP1, PCNA, and Su(var)2-10/dPIAS, are present in the high-confidence set. Specific sites that do not appear in all experiments tend to have the lowest intensities (Figure S2A). Considering that SUMOylation is often transient and affects a minuscule fraction of a given protein, it is likely that sensitivity is the major limitation to the reproducibility of detected sites

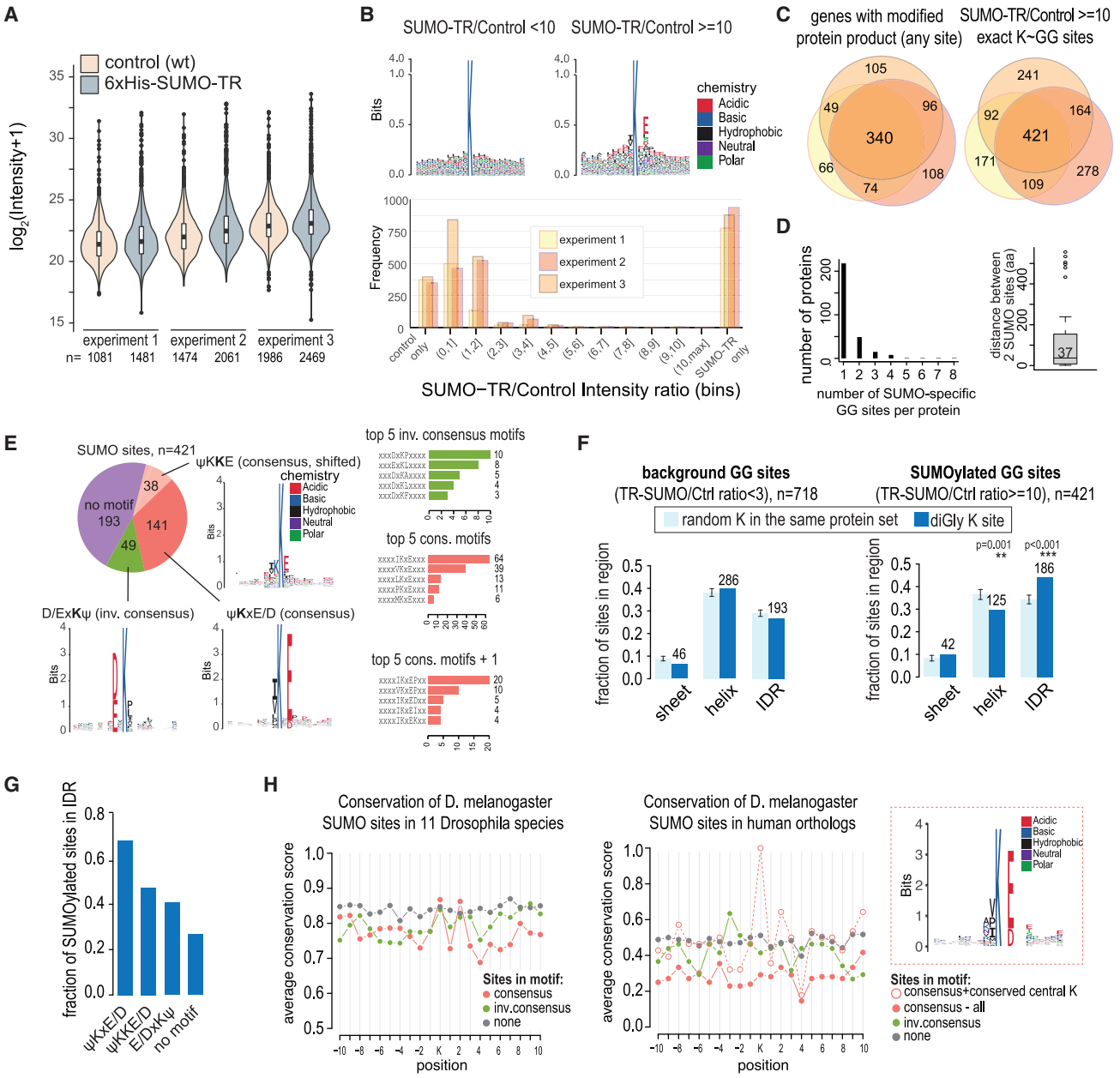


Figure 2. Characteristics of the ovarian SUMOylome

(A) Distribution of diGly site intensities in different samples.

(B) (Top) Amino acid frequencies flanking predicted K-ε-GG, for sites with high and low SUMO-TR:control intensity ratios in a representative replicate. Also see Figure S1. (Bottom) Distribution of SUMO-TR:Control intensity ratios of diGly sites in each experiment.

(C) Overlap of exact SUMO sites or SUMO-modified genes (irrespective of exact sites) between replicates.

(D) Numbers of predicted SUMO sites per protein and distance between sites for proteins with two sites.

(E) Sequence motifs at 421 high-confidence SUMO sites. The predicted diGly-modified lysine is in bold. Bar graphs show the counts of topmost frequent amino acids at positions −1 and +2, and −2, −1, and +2 from the modified lysine.

(F) Localization of diGly sites with high or low intensity ratios in predicted structured regions or IDRs, compared with randomly selected lysines from the same protein over 1,000 iterations. Error bars represent SD.

(G) Localization of diGly SUMO sites embedded in different motifs to IDRs.

(H) Conservation of predicted SUMO sites and flanking regions embedded in indicated motifs in 11 other *Drosophila* species and human. Sequence logo shows the amino acid frequencies of homologous regions in human proteins, where the *Drosophila* site is embedded in a consensus SUMOylation motif and the central lysine is conserved between fly and human.

between replicates. About ~25% of the proteins with diGly remnant have more than one modification site, with widely variable distances between two sites (Figure 2D). This pattern could indicate that proteins can be alternatively SUMOylated at different residues, multi-SUMOylated, or even SUMO- and ubiquitin-modified at the same time. As an aside, we detected several diGly sites on SUMO itself (Table S1); as some of these sites can be detected in the negative control, this pattern is indicative of hybrid SUMO-ubiquitin chains.

To gain further insight into the sequence features of the fly SUMOylome, we assessed the presence of amino acid motifs in the conservative set of 421 distinct high-confidence SUMOylation sites detected in all three experiments. Sequence pattern searches identified the canonical consensus and inverse consensus SUMOylation motifs (ψKxE/D and E/DxKψ, respectively) at 45% of the sites (141 and 49) in this set (Figure 2E). Approximately 10% of the sites have the motif ψKKE/D with the modification assigned to the second lysine, a pattern that might arise from uncertainty in diGly position prediction in the case of neighboring lysines. The remaining sites were not enriched for any known motif, and *de novo* motif search via MoMo did not identify any significant sequence bias around the central lysine (not shown), suggesting that, as in other systems, the SUMOylation consensus motif is not obligatory. Like observations in human cells,[17] consensus sites displayed a marked preference for isoleucine and valine as hydrophobic residues. Additionally, a sizable fraction of the sites with strong SUMOylation consensus had a downstream proline, an extended motif associated with SUMOylation/acetylation switch[23] (Figure 2E). Finally, we assessed the position of SUMOylation sites with respect to predicted structural features based on deep learning language models,[24] including α sheets, β helices, and intrinsically disordered regions (IDRs). This analysis demonstrated that SUMOylated sites are significantly overrepresented in IDRs and under-represented in structured regions compared with randomly selected lysines from the same protein or the background diGly sites (TR-SUMO/control ratio <3 in all replicates). Similar results were obtained using IDR predictions generated with IUPred2A[25] (Figure S2B). Notably, SUMOylation sites within canonical motifs showed the strongest bias toward IDRs (Figure 2G).

Finally, we analyzed the conservation of predicted *bona fide* SUMO sites (n = 421 in 292 genes) and flanking amino acids in closely related species (11 other *Drosophila* species with sequenced and well-annotated genomes, separated between 1 million and ~65 million years of evolution; see STAR methods) and humans. To this end, we performed protein sequence alignments and calculated the conservation score (i.e., the percentage of homologous positions in the alignments that carry the same amino acid as *D. melanogaster*). Out of 292 genes with *bona fide* SUMO sites, 246 have an annotated homolog in all other *Drosophila* species, and the target lysine carrying diGly remnant is highly conserved (Figure 2H). Interestingly, flanking amino acids for regions with non-consensus SUMO sites are more highly conserved than those flanking consensus sites. We speculate that this pattern may be related to the fact that consensus motifs are more frequently found in IDRs (Figure 2G), which in turn tend to be less conserved. Conversely, for regions containing the consensus site, only the central K and E/D residue two positions downstream has higher conservation score, indicating selective pressure to maintain the SUMO consensus motif. Two-hundred and twenty-three genes with 311 SUMO sites have homologs in human per DIOPT annotations retrieved from FlyBase.[26] As within the *Drosophila* genus, non-consensus sites generally have higher conservation scores. When all sites are considered, average levels of conservation at central lysines are not higher than other flanking amino acids. However, when sites in consensus motifs with conserved lysines between *Drosophila* and human are considered, the surrounding consensus is also conserved in the human homolog (Figure 2H). Finally, we queried UniProt for experimental evidence for SUMO modification at the human orthologs of *D. melanogaster* SUMO targets, with the caveat that human studies are performed on different tissue types. Remarkably, for 120 out of 223 human-fly ortholog pairs, the human counterpart is SUMO target, with the majority of positionally conserved consensus SUMO sites also modified in human cells (Table S2).

Altogether, these results show that SUMOylation affects a large number of proteins in the *Drosophila* ovary, with the SUMOylome displaying conserved sequence and structural features with other organisms. Importantly, the breadth of the SUMOylome, together with the observations that a substantial fraction of SUMOylation sites do not have a specific association with a motif or structural features, highlights the need for unbiased experimental approaches to identifying SUMOylation targets in different cellular and organismal contexts.

Functional groups of SUMOylated proteins

To gain further insights into the ovarian SUMOylome, we first analyzed the expression patterns of the 292 genes that that carry reproducible SUMO sites in all experiments using modEncode RNA sequencing (RNA-seq) data retrieved from FlyBase.[26] A large proportion of the detected SUMO targets are ubiquitously expressed, with a subset biased toward the ovary and early embryo (Figure S3). Additionally, the majority of targets are expressed in both somatic cell lines derived from the ovarian follicle cells (OSS, OSC) and pre-zygotic embryos (0–2 h) reflective of the germline/oocyte transcriptome. Thus, the SUMO pathway does not display a particular tissue or cell type bias.

To investigate the possible roles of SUMOylation in the ovary, we carried out functional enrichment and interactome analyses for proteins with *bona fide* SUMO sites using Gene Ontology (GO) and STRING (Figures 3A and 3B; Table S3). Results showed that SUMOylation is common among specific physically and functionally linked groups of ribosomal and nuclear proteins (Figures 3A and 3B; Table S3). Ribosomal proteins are highly abundant and hence a frequent source of background in proteomics. However, the number of reproducible diGly sites mapping to ribosomal proteins in the SUMO-TR samples is significantly above the background (ribosomal proteins carry 57 out of the 421 *bona fide* SUMO sites [13%], versus 37 out of the 714 background diGly sites [5%]) and could therefore represent genuine targets. Indeed, the SUMOylation machinery has been implicated in ribosomal protein regulation.[27] Beyond the ribosome, SUMO targets are common among proteins participating in rRNA processing, mRNA splicing, and DNA

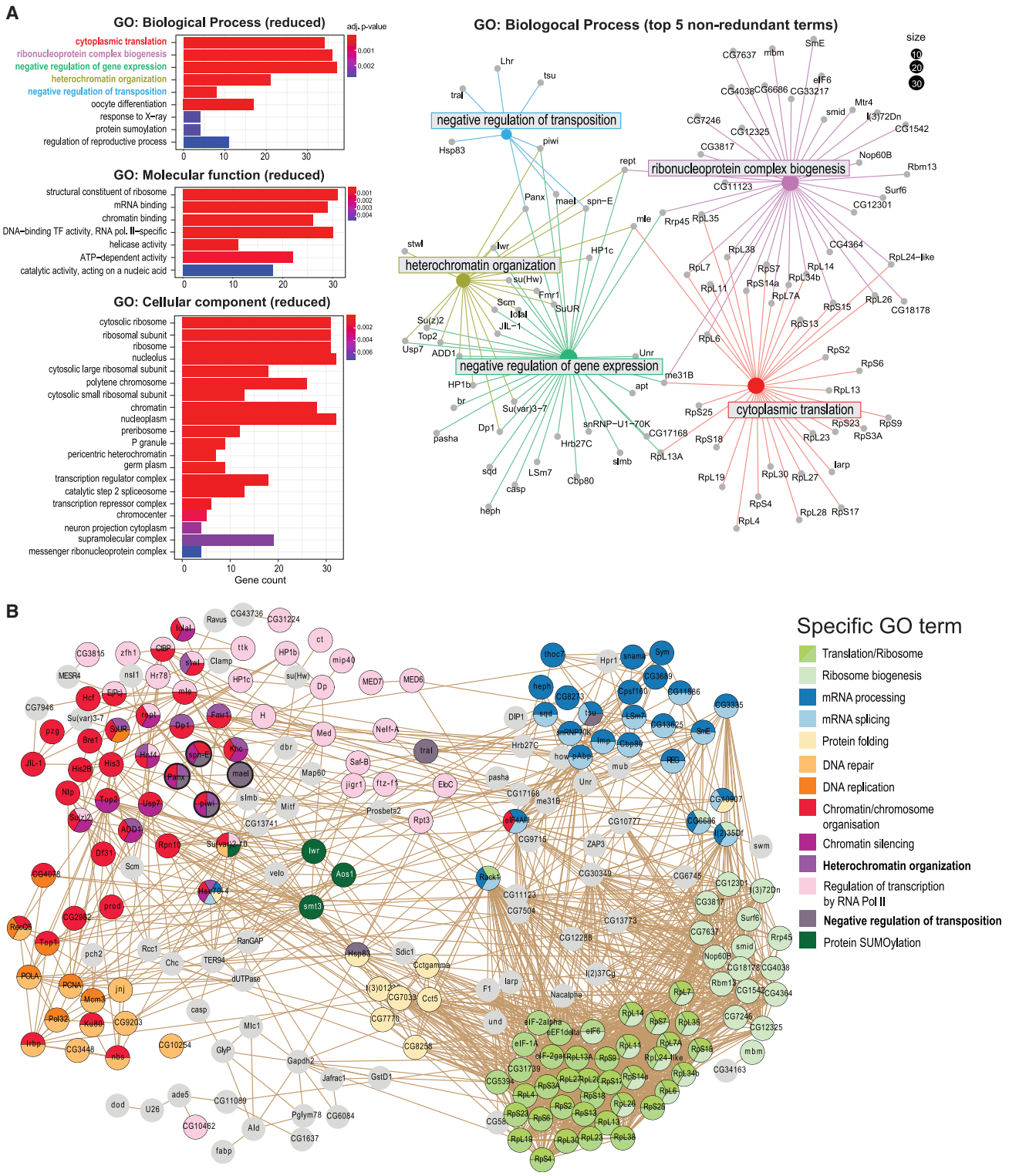


Figure 3. Functional groups of SUMOylated proteins

(A) Summary of GO analysis. (Left) Bar charts of the most enriched GO terms and corresponding gene numbers among the high-confidence SUMOylated targets after semantic simplification, ordered by adjusted p value. (Right) Network diagram of proteins annotated in the top five enriched Biological Process GO categories. (B) STRING diagram of the interactions between identified high-confidence SUMOylation targets. Single nodes are not shown. Nodes are colored according to selected GO terms and manually grouped according to functional annotations.

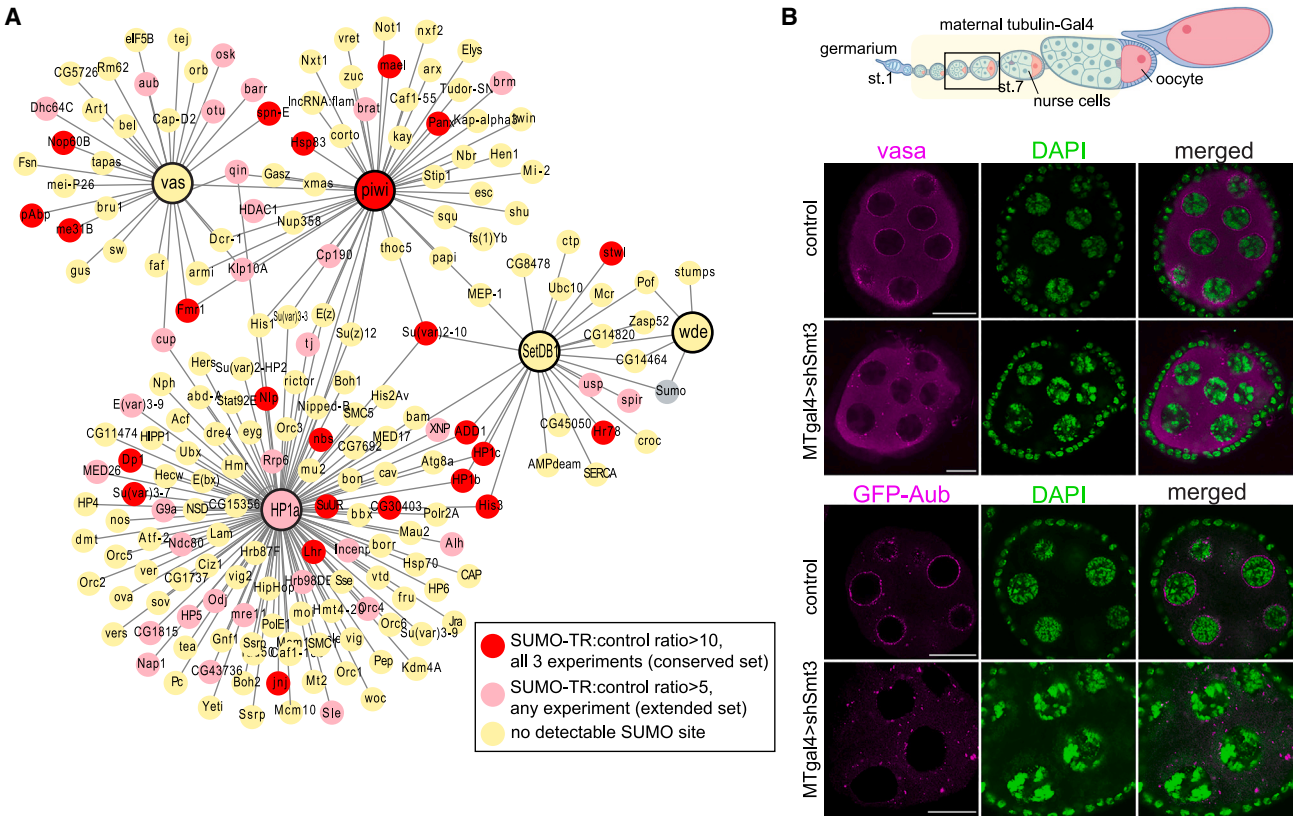


Figure 4. SUMO targets nuclear and cytoplasmic piRNA factors

(A) Network diagram showing physical interacting partners of Piwi, HP1, SetDB1, Wde, and Vasa proteins retrieved from FlyBase. Nodes are colored according to detection frequency of SUMOylation sites in the MS data. Additional interactions between partner nodes are not shown.

(B) (Top) Diagram of an ovariole with indicated stages where the maternal tubulin-Gal4 (MT-gal4) driver is active; stages of imaging are highlighted. (Bottom) Confocal images of Vasa (immunostaining) and Aub (GFP-tagged Aub) localization in stage 5–7 egg chambers from flies expressing SUMO shRNA under the control of the MT-gal4 driver, or no-driver control siblings. Scale bar, 20 μm.

damage response and repair, consistent with a conserved role of the SUMO pathway in these processes from yeast to mammals (reviewed in Zhao[9]). Notably, SUMOylation is widespread among chromatin-associated proteins. High-confidence SUMOylation sites can be found in several core histones and the HP1 paralogs HP1b and HP1c; a weaker site was also detected in the central heterochromatin effector HP1a/Su(var)205 at a position homologous to the SUMO site in its mammalian homolog HP1α/CBX5 (K32). Strikingly, among the most enriched GO terms in the SUMOylated set are terms related to heterochromatin and transposon repression; for example, 21 out of the 96 genes associated with the "heterochromatin organization" GO term, and eight out of 25 genes associated with the "negative regulation of transposition" GO term are in the high-confidence SUMO target set.

To further examine the breadth of SUMOylation within heterochromatin- and piRNA pathway-associated factors that control transposons in the ovary, we mapped the presence of SUMOylation sites among proteins that physically interact with some of the best-established effectors of heterochromatin and the nuclear and cytoplasmic components of the piRNA pathway: the H3K9me3 writer enzyme SetDB1/Eggless and its cofactor Windei (Wde); heterochromatin protein 1 (HP1a), which binds H3K9me3 and promotes chromatin compaction; the effector of piRNA-mediated transcriptional silencing Piwi; and Vasa, central to the assembly of the perinuclear nuage granules where piRNA biogenesis and transposon RNAi-mediated cleavage take place[28–30] (Figure 4A). Several HP1a interactors are modified by SUMO, including the split orthologs of the mammalian chromatin remodeler ATRX dADD1 and XNP, pointing to a multifaceted role of SUMOylation in transcriptional silencing and heterochromatin regulation. Furthermore, several high-confidence SUMOylation sites are present in Piwi, as well as several other nuclear piRNA pathway proteins, including Panoramix (Panx), a component of the so-called SFiNX/Pandas/PICTS[31–33] complex required for H3K9me3 deposition downstream of Piwi (also recently reported as SUMO target in another study[15]), and Maelstrom (Mael), essential for transcriptional repression and piRNA biogenesis from dual-stranded clusters.[18,34] Additionally, we detected SUMO sites at several Vasa-interacting proteins, including Spindle-E (Spn-E), a germline-specific DEAD-box helicase that localizes to nuage granules and is essential for piRNA production.[35–38] Aub, the cytoplasmic ping-pong mechanism effector, also appears as a SUMO target in some datasets;

however, in contrast to Spn-E and Mael (see below), we were unable to detect modified Aub by WB (data not shown), suggesting Aub modification may occur very transiently. Several additional piRNA biogenesis proteins, including thoc7, CG13741/Bootlegger, and Hel25E, also have detectable SUMO sites (Table S1). The presence of SUMO at a wide range of proteins involved in piRNA biogenesis and function suggests that this modification may be involved in the repression of TE activity on multiple levels.

Conditional knockdown of SUMO in the female germline results in a collapse of oogenesis associated with loss of heterochromatin and strong TE upregulation. Previous studies identified direct mechanistic roles of SUMOylation in the piRNA/Piwi-mediated formation of repressive chromatin at TE loci[15,21] that may explain this phenotype. Since we found several SUMO targets among nuage and piRNA biogenesis proteins, we wondered if SUMO loss also affects the structure of this essential compartment, thereby further contributing to defects in TE silencing. To this end, we examined the subcellular localization of well-established nuage components in SUMO depleted ovaries. We employed a previously characterized small hairpin RNA (shRNA) against SUMO[14] under the control of maternal tubulin Gal4 driver, which is active from stages 2–3 of oogenesis onward. In this condition, oogenesis fails to complete, with nurse cell nuclei collapsing around stage 7; however, prior to that, egg chambers appear morphologically normal and can be analyzed (Figures S4 and S4A). The nuage is a highly dynamic structure where multiple piRNA effector proteins, including Spn-E, Aub, and Ago3, are recruited in a hierarchical manner dependent on the conserved germinal granule protein Vasa.[29,39] We found that SUMO knockdown leads to markedly delocalized Vasa in dispersed granules throughout the cytoplasm compared with its normal localization as nearly continuous perinuclear layer (Figure 4B). Notably, Vasa is not a direct SUMO target per our data; however, its delocalization demonstrates that the SUMO pathway is genetically indispensable for the nuage integrity. As expected, other piRNA pathway proteins that depend on Vasa for their nuage recruitment, including the central effector of TE post-transcriptional cleavage and secondary piRNA production, Aub, Mael, and Spn-E, also display aberrant localization upon SUMO loss (Figures 4B and S4B). It is important to note that, although Mael and Spn-E, and possibly Aub, can be SUMOylated, their dependence on Vasa precludes understanding of whether SUMO may also be directly required for their normal localization. Collectively, these results indicate that the SUMO pathway may support transposon silencing not only through its involvement in heterochromatin formation but also through disruption of the nuage by various indirect and perhaps direct mechanisms.

SUMOylation of piRNA pathway factors in the female germline

To further understand SUMO's roles in the piRNA pathway, we sought to validate the SUMOylation of four essential piRNA pathway proteins: Piwi, Panx, Spn-E, and Mael. As SUMOylation typically affects only a small fraction of the total protein pool, and the detection of such species is limited by antibody availability and affinities, we devised a sensitive system to analyze SUMOylation of proteins of interest in the germ cells of the ovary, which express the full cytoplasmic and nuclear piRNA pathway. Specifically, we utilized UASp/Gal4 to express His-FLAG-tagged SUMO (herein referred to as FLAG-SUMO) and GFP-tagged target protein in nurse cells from stages 2–3 of oogenesis and later using the maternal tubulin-Gal4 driver (except GFP-Piwi, which was under the control of its native regulatory region). In this system, the GFP tag and high-affinity anti-GFP nanobody allow protein purification under stringent washing conditions, while the sensitive and specific monoclonal anti-FLAG antibody maximizes the detection of small amounts of SUMO-modified proteins by WB.

Analyses of immunopurified GFP-tagged Spn-E, Piwi, Panx, and Mael proteins from ovaries showed FLAG-SUMO-conjugated higher-molecular-weight forms consistent with SUMO modification in all cases (Figure 5A). Note that free FLAG-tagged SUMO migrates at about ~17 kDa on WB, slightly higher than its predicted molecular weight of ~12 kDa. Each target displayed multiple higher-molecular-weight bands, supporting the presence of multiple SUMO moieties or mixtures of SUMO and other protein PTMs on the same protein. The migration pattern of Panx indicates the existence of mono-SUMOylated and an array of poly-modified forms. A similar pattern of multiple modified Panx species was reported in another recent study where authors used a custom-made antibody against the native protein.[15] The observed shift in molecular weight for GFP-Piwi from ~120 to ~160 kDa and above indicates that it carries two or more protein modifiers, at least one of which is SUMO. Mael's and Spn-E's SUMO-modified forms have molecular weights of ~40–50 kDa greater than their unmodified forms, also consistent with at least two or three modified sites within the same protein. Of note, the number of higher-molecular-weight bands in Mael and Panx exceeds the number of predicted diGly sites. While our proteomics detection is limited to peptides within a particular size range and thus is unlikely to cover all possible modified residues, we also cannot rule out the existence of poly-SUMO or hybrid SUMO-ubiquitin chains (see Discussion). Notably, only a small fraction of each protein is SUMOylated, as the SUMOylated forms were only detectable with the sensitive anti-FLAG antibody, but not the anti-GFP antibody (Figure 5A), suggesting that the modification may occur transiently.

We also tested whether ectopically overexpressed piRNA pathway proteins can become SUMOylated in S2 cells, a somatic cell line from embryonic hematocyte origin that does not have an active piRNA pathway. While SUMOylation was still detectable, its pattern was drastically different from the characteristic "ladder" observed in the female germline (Figure S5A). The tissue-specific patterns of Piwi, Panx, Mael, and Spn-E SUMOylation, therefore, suggest that this modification is involved in specific regulatory steps in the germline in the context of active piRNA pathway response to genomic parasites.

Previous work identified the SUMO E3 ligase Su(var)2-10 as an effector of transposon silencing and heterochromatin establishment that acts downstream of Piwi.[14] To test if Su(var)2-10 regulates the SUMOylation of Mael, Panx, Spn-E, or Piwi, we induced its germline-specific knockdown via previously characterized UASp-controlled shRNA.[14] Su(var)2-10 depletion did not abolish the SUMOylation of any of the investigated proteins, despite

CellPress
OPEN ACCESS

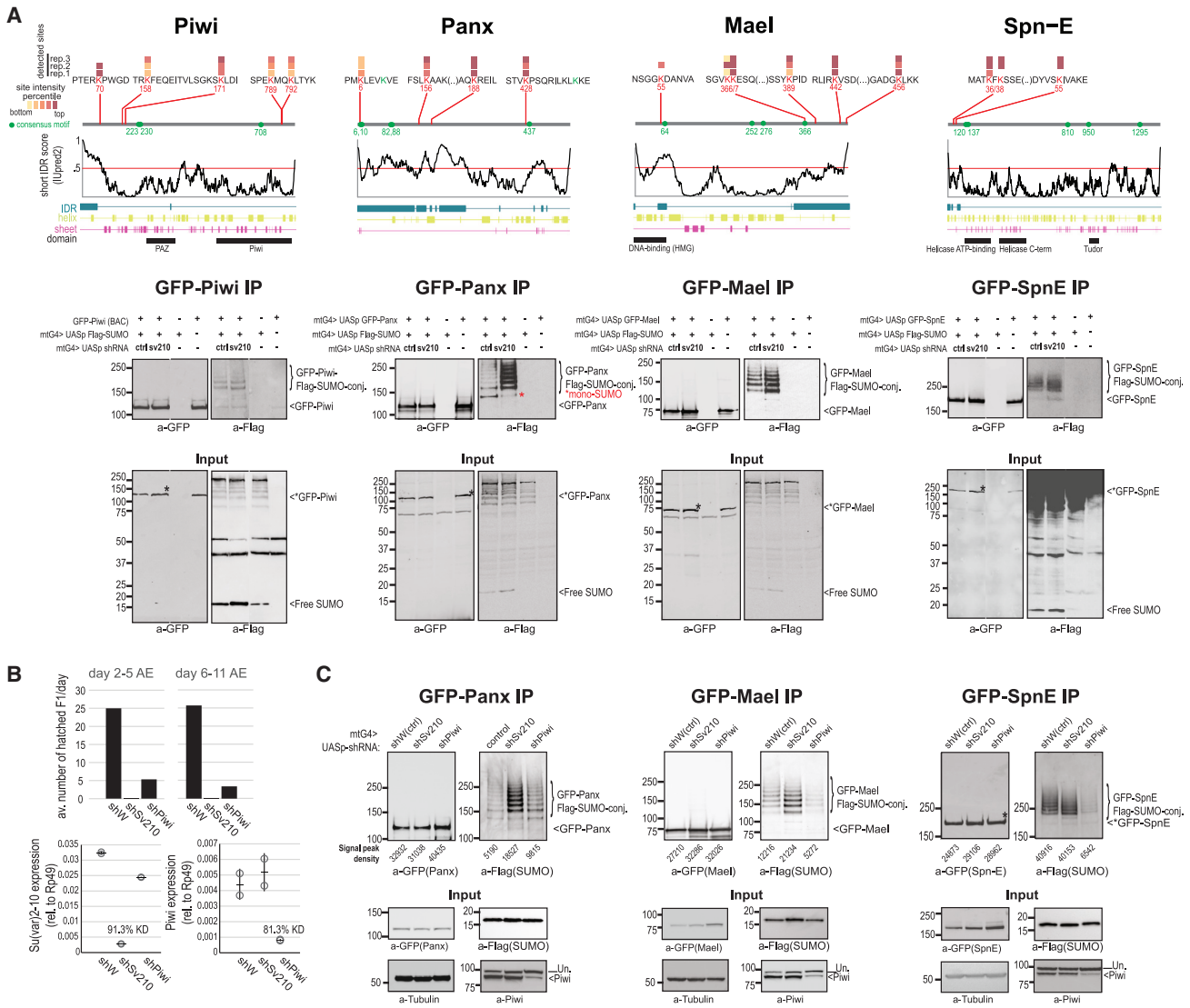


Figure 5. SUMOylation of Piwi, Mael, Spn-E, and Panx

(A) (Top) Diagrams of structural features, consensus SUMOylation motifs, and experimentally detected diGly sites with TR-SUMO:control intensity ratios >10. Color indicates diGly site intensity percentiles. (Bottom) WB analysis of target SUMOylation in the female germline. Ovary lysates from flies carrying indicated transgenes (mtG4 = maternal tubulin Gal4 driver) were subjected to GFP immunoprecipitation (IP) followed by WB detection first with anti-FLAG antibody and, after stripping, anti-GFP antibody. Note that SUMOylated species are only detectable by the highly sensitive anti-FLAG antibody, but not anti-GFP antibody. Lanes in Piwi and Spn-E gels were cropped and flipped for sample order consistency. Images are representative of three or four independent experiments per target.

(B) Efficiency of *su(var)2-10* and *piwi* germline knockdown estimated by fertility test (top) and RT-qPCR (bottom). Results are from samples from the same genetic crosses used for WB in (C), GFP-Panx. Error bars represent SD.

(C) WB of target SUMOylation in the female germline upon *white* (control), *su(var)2-10*, and *piwi* germline knockdown. Images are representative of two independent experiments. Un., unspecific band.

efficient knockdown shown by female sterility and RT-qPCR (Figure 5B). In fact, we systematically observed (n = 2–3 independent experiments for each protein; Figures 5A, 5C, and S5B) a slight increase of SUMOylated species upon Su(var)2-10 germline knockdown. Of these, the most prominent change was seen for Panx, which displayed a marked increase in its poly-modified forms and a relative decrease in the mono-SUMOylated form (Figure 5A, asterisk; Figure S5B). Thus, it appears that Su(var)2-10 is not a SUMO E3 ligase for Piwi, Mael, Spn-E, or Panx, but the modification of these proteins in germ cells is regulated in a manner influenced by Su(var)2-10 loss.

As Su(var)2-10 loss leads to strong transposon upregulation,[14] the increased SUMOylation of piRNA pathway proteins upon Su(var)2-10 knockdown may indicate SUMO implication in protein complexes actively engaged in transposon response. To explore this possibility, we decided to analyze the

SUMOylation of Panx, Spn-E, and Mael upon Piwi depletion, which disrupts the nuclear piRNA-dependent silencing machinery. To this end, we utilized an UASp-controlled shRNA against Piwi, which induces efficient Piwi germline knockdown (Figures 5B and 5C). Piwi depletion resulted in a significant loss of Mael and Spn-E SUMOylated forms (Figures 5C and S5B), suggesting that the modification of both proteins in germ cells occurs in a regulated manner genetically dependent of Piwi. Notably, these changes in SUMOylated form levels are not associated with obvious changes in the protein levels or subcellular localization (Figures S4B and S4C), supporting a transient nature of this modification. Interestingly, we noticed that Panx co-localization with chromatin (per DAPI stain, Figure S4C) is higher in Su(var)2-10 depleted nurse cells, in line with a model where transposon upregulation leads to increased Piwi-dependent Panx association with targets.[15] However, in contrast to Spn-E and Mael, Piwi germline knockdown did not affect the number and levels of SUMO-modified Panx forms, indicating that the bulk of Panx SUMOylation may be governed by an alternative mechanism independent of the piRNA-Piwi complex in this cellular context. Altogether, these results point to multiple roles of protein SUMOylation in the piRNA pathway extending to both piRNA biogenesis and silencing factors.

DISCUSSION

The SUMO pathway is essential for normal cell function, as shown by the severe phenotypes of loss-of-function mutants of SUMO and SUMO ligases in various systems.[40] Our mechanistic understanding of SUMOylation, however, has remained limited, partly owing to the technical difficulties of detecting this modification. Here, we adapted the diGly remnant enrichment method for the unbiased discovery of SUMO-modified proteins with specific site predictions in *Drosophila*. Altogether, our proteomics survey of the *Drosophila* ovary uncovered a large SUMOylated complement that displays conserved features of this modification, including sequence and structural biases of preferred SUMOylation sites, collective targeting of proteins from the same process/complex, and enrichment of SUMO targets among proteins involved in various aspects of nucleic acid metabolism. We detected both SUMO sites within well-established SUMO consensus motifs biased toward protein IDRs, and non-consensus sites. Previous studies suggested that consensus sites can be directly targeted by the SUMO E2 conjugation enzyme Ubc9, while non-consensus sites require an E3 ligase for their modification.[5,7,8] In the future, it will be interesting to unravel the regulation, kinetics, and functional implications of different SUMO sites. Notably, we found SUMO targets with complex modification patterns among a variety of piRNA pathway proteins functioning in both transcriptional and post-transcriptional transposon silencing, pointing to a multifaceted role of this modification in the cellular response to genomic parasites.

Limitations of the study

Even though diGly peptide enrichment improves the sensitivity and specificity of SUMOylated protein detection, as with other bottom-up proteomics methods, detection is only possible for sites residing in protease-generated peptides within a specific size range. Also, exact site assignment might be imprecise when two lysine residues are in proximity and in the event a protein simultaneously carries SUMO and another modification that leaves a diGly remnant after trypsin cleavage, such as ubiquitin (in such cases, the target protein would be enriched in the SUMOylated pool but detected diGly remnants might come from any concomitant modification). Despite these limitations, the absolute identification of SUMO protein targets remains of high confidence, and predicted sites are informative to narrow down exact modified residues for future mutational studies.

Complexity of the SUMO-modified proteome

Approximately a third of all SUMO targets in our high-confidence set contained two or more diGly remnants, and analysis of selected proteins in germ cells by WB (Figure 5) demonstrated a characteristic ladder of multiple modified forms. These patterns are indicative of multi-SUMOylation (when multiple residues in the same protein are SUMOylated), and, possibly, concomitant modification by SUMO and other protein modifiers. In addition to multiple modifications at different residues, studies in yeast and mammalian systems have identified a variety of homotypic and heterotypic SUMO and ubiquitin chains, as well as hybrid SUMO-ubiquitin chains that could create a complex regulatory "code."[41,42] We detected several diGly sites on the endogenous SUMO proteins in control samples (Table S1), confirming the presence of hybrid SUMO-ubiquitin chains in various configurations in flies. *In vitro* studies and comparative genomics suggested that SUMO-only chains—common in yeast, plants, and mammals—have been evolutionarily lost in flies.[43] Distinguishing SUMO chains from SUMO-ubiquitin hybrid chains is not possible with our approach. However, considering the presence of diGly remnants at syntenic positions to typical sites of SUMO chain formation in other species, the existence of SUMO chains *in vivo* merits future investigation.

Heterochromatin as a SUMOylation hotspot

A prominent feature of the identified SUMO targets is that they are often found among physically and functionally related proteins. This is consistent with previous notions of an unusual property of SUMOylation compared with other PTMs; namely, that SUMO ligases can modify entire groups of physically interacting proteins at multiple and perhaps redundant sites (reviewed in Jentsch and Psakhye[11]). Such group SUMOylation is thought to create multiple SUMO-SIM interactions within large molecular complexes that act synergistically to facilitate their assembly and function. Previous examples of SUMO hotspots include DNA repair pathway effectors in yeast, ribosome biogenesis, or PML bodies.[11] The SUMO pathway has long been linked to heterochromatin: SUMOylated histones, SUMO, and SUMO ligases were shown to localize to heterochromatic regions in various systems, and several essential heterochromatin proteins, including HP1, H3K9 methyltransferase, and histone deacetylase complex components, were identified as SUMO targets or interactors (reviewed in Cubeñas-Potts and Matunis[44]). Our data identifying dozens of heterochromatin proteins as SUMOylation targets suggest that heterochromatin can be considered another hotspot of group SUMOylation. Moreover,

SUMO was recently shown to affect interconnected repressive chromatin factors in mouse embryonic stem cells,[45] pointing to an evolutionarily conserved role of SUMOylation in the regulation of chromatin organization. Notably, SUMO polymers were shown to drive phase separation in different subcellular contexts.[12,46] In the future, it will be interesting to establish the functional significance of the heterochromatic SUMO target spectra, particularly from the perspective of the biophysical properties of heterochromatin and heterochromatin-related multi-subunit regulatory complexes.

A multifaceted role of SUMO in the piRNA pathway

The discovery of a large set of SUMO-modified proteins among piRNA pathway factors paints a complex picture of the roles of SUMOylation in the cellular response to transposons. First hints on the implication of SUMO in the piRNA pathway emerged from genetic screens to identify factors required for transposon repression.[13] More recently, functional studies demonstrated that SUMO is essential for the nuclear compartment of the piRNA pathway that enforces co-transcriptional silencing of transposons by the installation of the silencing mark H3K9me3 at target loci. First, transcriptional silencing was shown to involve Su(var)2-10/SUMO-dependent recruitment of chromatin-modifying complexes downstream of Piwi and the co-repressor Panx.[14,16] Additionally, transposon silencing requires a SUMO-mediated interaction between Panx and the general heterochromatin factor Sov.[15] The proteomic identification of numerous general heterochromatin proteins as well as Panx itself as SUMO targets emphasizes the possibility that SUMOylation plays a multifaceted role in the transcriptional silencing of transposons by piRNAs. However, in addition to that, we also uncovered SUMO targets among piRNA proteins that belong to spatially and functionally distinct piRNA pathway compartments such as the nuage and found that SUMO depletion disrupts the nuage structure upstream of Vasa, suggesting that SUMOylation has even broader implications in transposon silencing.

While dissecting the multiple roles of SUMOylation in the piRNA pathway will require extensive future work, observations of the SUMOylation patterns of selected piRNA pathway proteins provide several interesting clues. First, the SUMO-modified forms of all four proteins are of relatively low abundance and only detectable by highly sensitive antibodies, indicative of transient nature. Dynamic and reversible combinatorial modifications via SUMO and SUMO-ubiquitin chains can create unique surfaces for multivalent protein interactions. Complex chains were previously shown to act as signals in the regulated recruitment and assembly of protein complexes involved in centromere organization and DNA damage repair (reviewed in Keiten-Schmitz et al.[41]). It is possible that, analogously, complex SUMOylation patterns may be involved in dynamically regulated steps of piRNA biogenesis and function. A second line of clues comes from the different effects of Piwi and Su(var)2-10 germline knockdown on the SUMOylation of the examined piRNA pathway proteins. The SUMO E3 ligase Su(var)2-10 is essential for H3K9me3 deposition at transposon targets.[14] The finding that Piwi, Panx, Spn-E, and Mael SUMOylation do not depend on Su(var)2-10 is consistent with a model where Su(var)2-10 acts at downstream steps of heterochromatin establishment, but Su(var)2-10-independent SUMOylation (solely by Ubc9 or involving other E3 ligases) is also involved in upstream or parallel processes related to piRNA biogenesis and function. Moreover, as the SUMO-modified forms of the four proteins increase upon Su(var)2-10 knockdown, it is tempting to speculate that SUMOylation might be associated with increased piRNA pathway activity in response to the strong transposon upregulation that occurs upon this genetic perturbation. Nevertheless, it is important to consider that Su(var)2-10 depletion causes substantial transcriptomic changes beyond transposon activation, including the up- and downregulation of hundreds of protein-coding genes, including *smt3* upregulation,[14,21] and sometimes increase in free SUMO protein (see Figures 5 and S5), with possible indirect consequences on the SUMOylome.

Interestingly, Piwi depletion has differential effects on Mael, Spn-E, and Panx SUMOylation: Mael and Spn-E markedly lose their SUMOylation in Piwi-knockdown germ cells, while Panx SUMOylation remains unaffected (Figure 5C, also see Figure S5B for extended data). Previous studies recognized Mael as an essential factor for transcriptional repression at piRNA targets as well as piRNA-independent genomic loci, although its precise mechanistic role has remained unclear.[18,34] Considering the nuclear functions of Mael, the loss of its SUMOylation upon Piwi knockdown might reflect a regulatory step following piRNA/Piwi recognizing their targets. Nevertheless, Mael also localizes to the nuage,[18] and further work will be required to establish whether Mael SUMOylation is related to its roles in the nucleus or an as-yet unknown function in the nuage. A possible role of SUMOylation in the cytoplasmic piRNA pathway branch is further highlighted by the modification of Spn-E. Spn-E is a putative RNA helicase that, despite having an unclear mechanistic role, is a well-established nuage component essential for piRNA biogenesis.[35–38] Since Piwi operates in the nucleus, yet its loading with piRNA occurs in the nuage, the dependence of Spn-E SUMOylation on Piwi raises the intriguing possibility that SUMO-mediated interactions may be involved in certain aspects of piRNA biogenesis. Future work combining our highly sensitive and robust germline SUMOylation assay with different genetic perturbations and SUMOylation-deficient mutants promises to improve our understanding of the precise mechanisms that orchestrate the piRNA pathway in its different subcellular contexts.

Finally, the lack of effect of Piwi loss on the SUMOylation of Panx—one of the core components of the RNA-binding SFiNX/Pandas/PITSC complex that is essential for H3K9me3 deposition downstream of Piwi—suggests Panx is modified in a process that occurs independently of Piwi-mediated target recognition. Surprisingly, this result contrasts with recent data from OSC cells—an immortalized cell line derived from the ovarian somatic follicle cells that expresses only the nuclear portion of the piRNA pathway—where Panx SUMOylation was found to completely depend on Piwi.[15] While we cannot rule out that residual levels of Piwi may be sufficient to maintain Panx SUMOylation in our system, it is also possible that the regulation of Panx differs between the germline and soma, or even at different stages of oogenesis. A potential mechanistic difference in the nuclear piRNA pathway between germline and soma has critical implications for our understanding of this process and will be important to address in the future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Drosophila stocks and husbandry
- METHOD DETAILS
  - Proteomics sample preparation
  - LC-MS/MS and raw data processing
  - Bioinformatics analysis of SUMOylation sites
  - Detection of protein SUMOylation by immunoprecipitation (IP) and western blotting
  - RT-qPCR
  - Protein localization imaging
  - Fertility test
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2023.100329.

ACKNOWLEDGMENTS

We thank former Caltech Protein Exploration Laboratory (PEL) members Dr. Michael Sweredoski and Dr. Annie Moradian for their advice on diGly proteomics, and Corinne Karalun (former laboratory assistant in M.N. laboratory, UCR), Hannah Ryon (former rotation student in K.F.T. laboratory, Caltech), and Matea Ibrahim (undergraduate student at UC Riverside) for assistance with WB, fly dissections, fixation, and genotyping. This work was supported by grants from the NIH (K99/R00 HD099316) to M.N.; the NIH (R01 GM097363) and the Howard Hughes Medical Institute Faculty Scholar Award to A.A.A.; and the NIH (R01 GM110217) and Ellison Medical Foundation Awards to K.F.T.

AUTHOR CONTRIBUTIONS

M.N., K.F.T., and A.A.A. conceptualized the proteomics study. M.N. designed and performed experiments, data curation, and formal analysis, except LC-MS/MS runs and raw data processing, which were performed at the Caltech PEL facility by B.L. H.H. performed GFP-Aub localization experiments. M.N. prepared figures and drafted the manuscript. M.N. and A.A.A. edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 29, 2022
Revised: March 24, 2023
Accepted: April 26, 2023
Published: May 18, 2023

REFERENCES

1. Flotho, A., and Melchior, F. (2013). Sumoylation: a regulatory protein modification in health and disease. Annu. Rev. Biochem. *82*, 357–385. https://doi.org/10.1146/annurev-biochem-061909-093311.
2. Johnson, E.S., and Blobel, G. (1997). Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. J. Biol. Chem. *272*, 26799–26802. https://doi.org/10.1074/jbc.272.43.26799.
3. Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. J. Biol. Chem. *276*, 12654–12659. https://doi.org/10.1074/jbc.M009476200.
4. Sampson, D.A., Wang, M., and Matunis, M.J. (2001). The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. J. Biol. Chem. *276*, 21664–21669. https://doi.org/10.1074/jbc.M100006200.
5. Gareau, J.R., and Lima, C.D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat. Rev. Mol. Cell Biol. *11*, 861–871. https://doi.org/10.1038/nrm3011.
6. Bernier-Villamor, V., Sampson, D.A., Matunis, M.J., and Lima, C.D. (2002). Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. Cell *108*, 345–356. https://doi.org/10.1016/s0092-8674(02)00630-x.
7. Li, C., McManus, F.P., Plutoni, C., Pascariu, C.M., Nelson, T., Alberici Delsin, L.E., Emery, G., and Thibault, P. (2020). Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. Nat. Commun. *11*, 834. https://doi.org/10.1038/s41467-020-14581-w.
8. Yunus, A.A., and Lima, C.D. (2009). Structure of the siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA. Mol. Cell *35*, 669–682. https://doi.org/10.1016/j.molcel.2009.07.013.
9. Zhao, X. (2018). SUMO-mediated regulation of nuclear functions and signaling processes. Mol. Cell *71*, 409–418. https://doi.org/10.1016/j.molcel.2018.07.027.
10. Psakhye, I., and Jentsch, S. (2012). Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. Cell *151*, 807–820. https://doi.org/10.1016/j.cell.2012.10.021.
11. Jentsch, S., and Psakhye, I. (2013). Control of nuclear activities by substrate-selective and protein-group SUMOylation. Annu. Rev. Genet. *47*, 167–186. https://doi.org/10.1146/annurev-genet-111212-133453.
12. Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional control of phase-separated cellular bodies. Cell *166*, 651–663. https://doi.org/10.1016/j.cell.2016.06.010.
13. Muerdter, F., Guzzardo, P.M., Gillis, J., Luo, Y., Yu, Y., Chen, C., Fekete, R., and Hannon, G.J. (2013). A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila. Mol. Cell *50*, 736–748. https://doi.org/10.1016/j.molcel.2013.04.006.
14. Ninova, M., Chen, Y.-C.A., Godneeva, B., Rogers, A.K., Luo, Y., Fejes Tóth, K., and Aravin, A.A. (2020). Su(var)2-10 and the SUMO pathway link piRNA-guided target recognition to chromatin silencing. Mol. Cell *77*, 556–570.e6. https://doi.org/10.1016/j.molcel.2019.11.012.
15. Andreev, V.I., Yu, C., Wang, J., Schnabl, J., Tirian, L., Gehre, M., Handler, D., Duchek, P., Novatchkova, M., Baumgartner, L., et al. (2022). Panoramix SUMOylation on chromatin connects the piRNA pathway to the cellular heterochromatin machinery. Nat. Struct. Mol. Biol. *29*, 130–142. https://doi.org/10.1038/s41594-022-00721-x.
16. Mugat, B., Nicot, S., Varela-Chavez, C., Jourdan, C., Sato, K., Basyuk, E., Juge, F., Siomi, M.C., Pélisson, A., and Chambeyron, S. (2020). The Mi-2 nucleosome remodeler and the Rpd3 histone deacetylase are involved in piRNA-guided heterochromatin formation. Nat. Commun. *11*, 2818. https://doi.org/10.1038/s41467-020-16635-5.

17. Czech, B., and Hannon, G.J. (2016). One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. Trends Biochem. Sci. *41*, 324–337. https://doi.org/10.1016/j.tibs.2015.12.008.

18. Sienski, G., Dönertas, D., and Brennecke, J. (2012). Transcriptional silencing of transposons by piwi and Maelstrom and its impact on chromatin state and gene expression. Cell *151*, 964–980. https://doi.org/10.1016/j.cell.2012.10.040.

19. Rozhkov, N.V., Hammell, M., and Hannon, G.J. (2013). Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev. *27*, 400–412. https://doi.org/10.1101/gad.209767.112.

20. Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Tóth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. *27*, 390–399. https://doi.org/10.1101/gad.209841.112.

21. Ninova, M., Godneeva, B., Chen, Y.-C.A., Luo, Y., Prakash, S.J., Jankovics, F., Erdélyi, M., Aravin, A.A., and Fejes Tóth, K. (2020). The SUMO ligase su(var)2-10 controls hetero- and euchromatic gene expression via establishing H3K9 trimethylation and negative feedback regulation. Mol. Cell *77*, 571–585.e4. https://doi.org/10.1016/j.molcel.2019.09.033.

22. Impens, F., Radoshevich, L., Cossart, P., and Ribet, D. (2014). Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli. Proc. Natl. Acad. Sci. USA *111*, 12432–12437. https://doi.org/10.1073/pnas.1413825111.

23. Stankovic-Valentin, N., Deltour, S., Seeler, J., Pinte, S., Vergoten, G., Guérardel, C., Dejean, A., and Leprince, D. (2007). An acetylation/deacetylation-SUMOylation switch through a phylogenetically conserved ψKXEP motif in the tumor suppressor HIC1 regulates transcriptional repression activity. Mol. Cell Biol. *27*, 2661–2675. https://doi.org/10.1128/MCB.01098-06.

24. Dallago, C., Schütze, K., Heinzinger, M., Olenyi, T., Littmann, M., Lu, A.X., Yang, K.K., Min, S., Yoon, S., Morton, J.T., and Rost, B. (2021). Learned embeddings from deep learning to visualize and predict protein sets. Curr. Protoc. *1*, e113. https://doi.org/10.1002/cpz1.113.

25. Mészáros, B., Erdős, G., and Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. *46*, W329–W337. https://doi.org/10.1093/nar/gky384.

26. Attrill, H., Falls, K., Goodman, J.L., Millburn, G.H., Antonazzo, G., Rey, A.J., and Marygold, S.J.; FlyBase Consortium (2016). FlyBase: establishing a gene group resource for Drosophila melanogaster. Nucleic Acids Res. *44*, D786–D792. https://doi.org/10.1093/nar/gkv1046.

27. Panse, V.G., Kressler, D., Pauli, A., Petfalski, E., Gnädig, M., Tollervey, D., and Hurt, E. (2006). Formation and nuclear export of preribosomes are functionally linked to the small-ubiquitin-related modifier pathway. Traffic *7*, 1311–1321. https://doi.org/10.1111/j.1600-0854.2006.00471.x.

28. Xiol, J., Spinelli, P., Laussmann, M.A., Homolka, D., Yang, Z., Cora, E., Couté, Y., Conn, S., Kadlec, J., Sachidanandam, R., et al. (2014). RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. Cell *157*, 1698–1711. https://doi.org/10.1016/j.cell.2014.05.018.

29. Lim, A.K., and Kai, T. (2007). Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA *104*, 6714–6719. https://doi.org/10.1073/pnas.0701920104.

30. Nishida, K.M., Iwasaki, Y.W., Murota, Y., Nagao, A., Mannen, T., Kato, Y., Siomi, H., and Siomi, M.C. (2015). Respective functions of two distinct siwi complexes assembled during PIWI-interacting RNA biogenesis in Bombyx germ cells. Cell Rep. *10*, 193–203. https://doi.org/10.1016/j.celrep.2014.12.013.

31. Batki, J., Schnabl, J., Wang, J., Handler, D., Andreev, V.I., Stieger, C.E., Novatchkova, M., Lampersberger, L., Kauneckaite, K., Xie, W., et al. (2019). The nascent RNA binding complex SFiNX licenses piRNA-guided heterochromatin formation. Nat. Struct. Mol. Biol. *26*, 720–731.

32. Zhao, K., Cheng, S., Miao, N., Xu, P., Lu, X., Zhang, Y., Wang, M., Ouyang, X., Yuan, X., Liu, W., et al. (2019). A Pandas complex adapted for piRNA-guided transcriptional silencing and heterochromatin formation. Nat. Cell Biol. *21*, 1261–1272. https://doi.org/10.1038/s41556-019-0396-0.

33. Fabry, M.H., Ciabrelli, F., Munafò, M., Eastwood, E.L., Kneuss, E., Falciatori, I., Falconio, F.A., Hannon, G.J., and Czech, B. (2019). piRNA-guided co-transcriptional silencing coopts nuclear export factors. Elife *8*, e47999. https://doi.org/10.7554/eLife.47999.

34. Chang, T.H., Mattei, E., Gainetdinov, I., Colpan, C., Weng, Z., and Zamore, P.D. (2019). Maelstrom represses canonical polymerase II transcription within Bi-directional piRNA clusters in Drosophila melanogaster. Mol. Cell *73*, 291–303.e6. https://doi.org/10.1016/j.molcel.2018.10.038.

35. Andress, A., Bei, Y., Fonslow, B.R., Giri, R., Wu, Y., Yates, J.R., III, and Carthew, R.W. (2016). Spindle-E cycling between nuage and cytoplasm is controlled by Qin and PIWI proteins. J. Cell Biol. *213*, 201–211. https://doi.org/10.1083/jcb.201411076.

36. Aravin, A.A., Klenov, M.S., Vagin, V.V., Bantignies, F., Cavalli, G., and Gvozdev, V.A. (2004). Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line. Mol. Cell Biol. *24*, 6742–6750. https://doi.org/10.1128/MCB.24.15.6742-6750.2004.

37. Gillespie, D.E., and Berg, C.A. (1995). Homeless is required for RNA localization in Drosophila oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases. Genes Dev. *9*, 2495–2508. https://doi.org/10.1101/gad.9.20.2495.

38. Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522–535. https://doi.org/10.1016/j.cell.2009.03.040.

39. Webster, A., Li, S., Hur, J.K., Wachsmuth, M., Bois, J.S., Perkins, E.M., Patel, D.J., and Aravin, A.A. (2015). Aub and Ago3 are recruited to nuage through two mechanisms to form a ping-pong complex assembled by krimper. Mol. Cell *59*, 564–575. https://doi.org/10.1016/j.molcel.2015.07.017.

40. Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nat. Rev. Mol. Cell Biol. *8*, 947–956. https://doi.org/10.1038/nrm2293.

41. Keiten-Schmitz, J., Schunck, K., and Müller, S. (2020). SUMO chains rule on chromatin occupancy. Front. Cell Dev. Biol. *7*.

42. Komander, D., and Rape, M. (2012). The ubiquitin code. Annu. Rev. Biochem. *81*, 203–229. https://doi.org/10.1146/annurev-biochem-060310-170328.

43. Ureña, E., Pirone, L., Chafino, S., Pérez, C., Sutherland, J.D., Lang, V., Rodriguez, M.S., Lopitz-Otsoa, F., Blanco, F.J., Barrio, R., and Martín, D. (2016). Evolution of SUMO function and chain formation in insects. Mol. Biol. Evol. *33*, 568–584. https://doi.org/10.1093/molbev/msv242.

44. Cubeñas-Potts, C., and Matunis, M.J. (2013). SUMO: a multifaceted modifier of chromatin structure and function. Dev. Cell *24*, 1–12. https://doi.org/10.1016/j.devcel.2012.11.020.

45. Theurillat, I., Hendriks, I.A., Cossec, J.-C., Andrieux, A., Nielsen, M.L., and Dejean, A. (2020). Extensive SUMO modification of repressive chromatin factors distinguishes pluripotent from somatic cells. Cell Rep. *32*, 108146. https://doi.org/10.1016/j.celrep.2020.108146.

46. Keiten-Schmitz, J., Röder, L., Hornstein, E., Müller-McNicoll, M., and Müller, S. (2021). SUMO: glue or solvent for phase-separated ribonucleoprotein complexes and molecular condensates? Front. Mol. Biosci. *8*, 673038.

47. Chen, Y.-C.A., Stuwe, E., Luo, Y., Ninova, M., Le Thomas, A., Rozhavskaya, E., Li, S., Vempati, S., Laver, J.D., Patel, D.J., et al. (2016). Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. Mol. Cell *63*, 97–109. https://doi.org/10.1016/j.molcel.2016.05.010.

48. Rogers, A.K., Situ, K., Perkins, E.M., and Toth, K.F. (2017). Zucchini-dependent piRNA processing is triggered by recruitment to the

cytoplasmic processing machinery. Genes Dev *31*, 1858–1869. https://doi.org/10.1101/gad.303214.117.

49. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for comparing biological themes among gene clusters. OMICS A J. Integr. Biol. *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.
50. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504. https://doi.org/10.1101/gr.1239303.
51. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340.
52. Zhao, Q., Xie, Y., Zheng, Y., Jiang, S., Liu, W., Mu, W., Liu, Z., Zhao, Y., Xue, Y., and Ren, J. (2014). GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res *42*, W325–W330. https://doi.org/10.1093/nar/gku383.
53. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372. https://doi.org/10.1038/nbt.1511.
54. Nie, M., Xie, Y., Loo, J.A., and Courey, A.J. (2009). Genetic and proteomic evidence for roles of Drosophila SUMO in cell cycle control, ras signaling, and early pattern formation. PLoS One *4*, e5905. https://doi.org/10.1371/journal.pone.0005905.
55. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805. https://doi.org/10.1021/pr101065j.
56. Cheng, A., Grant, C.E., Noble, W.S., and Bailey, T.L. (2019). MoMo: discovery of statistically significant post-translational modification motifs. Bioinformatics *35*, 2774–2782. https://doi.org/10.1093/bioinformatics/bty1058.
57. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-Flag (HRP conjugated)	Sigma	Cat# A8592; RRID:AB_439702
Rabbit polyclonal anti-GFP	Abcam	Cat# ab290; RRID:AB_303395
Rabbit polyclonal anti-GFP	Chen et al., 2016[47]	N/A
Anti-Tubulin	Sigma-Aldrich	Cat# T5168; RRID:AB_477579
HRP-conjugated anti-rabbit	Cell Signaling	Cat# 7074; RRID:AB_2099233
HRP-conjugated anti-mouse	Cell Signaling	Cat# 7076; RRID:AB_330924
IRDye®-conjugated anti-mouse	LiCor	Cat# 925-68070; RRID:AB_2651128
IRDye®-conjugated anti-rabbit	LiCor	Cat# 925-68071; RRID:AB_2721181
mouse anti-Piwi	Santa Cruz	Cat# sc-390946
Rat anti-vasa	DSHB	AB_760351; RRID:AB_760351
Anti-rat Alexa Fluor 594	Invitrogen	Cat# A-11007; RRID:AB_141374
Mouse 6x-His Tag Monoclonal Antibody	ThermoFisher	Cat# His.H8; RRID:AB_2536988
GFP-Trap®	ChromoTek	Cat# gtma-20; RRID:AB_2631358
Critical commercial assays		
PTMScan® HS Ubiquitin/SUMO Remnant Motif (K-ε-GG) Kit	Cell Signaling	59322
Deposited data		
Mass-spec raw and MaxQuant processed data	PRIDE	PXD037421
Experimental models: Cell lines		
Drosophila S2 cells	DGRC	S2-DRSC #181
Experimental models: Organisms/strains		
6xHis-smt3[T86R]}attP2	This study	N/A
UASp-shWhite	BDSC	33623
UASp-shSUMO	Ninova et al. 2020[14,21]	N/A
UASp-shPiwi	BDSC	33724
UASp-shSv210-2	BDSC	32956
UASp-GFP-Panx	Rogers et al. 2017[48]	N/A
UASp-GFP-Mael	Webster et al. 2015[39]	N/A
UASp-Spindle-E	Webster et al. 2015[39]	N/A
GFP-Piwi	Le Thomas et al. 2013[20]	N/A
GFP-Aub	Webster et al. 2015[39]	N/A
P{ry[+t7.2] = PZ}Sumo[04493] cn[1]/CyO; ry[506]	BDSC	11378
Iso-1	BDSC	2057
w[*]; P{w[+mC] = matalpha4-GAL-VP16}V37	BDSC	7063
Oligonucleotides		
Su(var)2-10-F	IDT	CCAGCACAGGACGAACAGCCC
Su(var)2-10-R	IDT	CGTGGAACTGGCGACGGCTT
Piwi-F	IDT	CTGCTGATCTCCAAAAATAGGG
Piwi-R	IDT	TCGCGTATAACTGCTCATGG
rp49-F	IDT	CCGCTTCAAGGGACAGTATCTG
rp49-R	IDT	ATCTCGCCGCAGTAAACGC

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
Drosophila Gateway Vector collection	DGRC	N/A
pActin-3xFlag3xHA-SUMO	Ninova et al. 2020[14,21]	N/A
pUbi-GFP-Piwi	This study	N/A
pUbi-GFP-Panx	This study	N/A
pUbi-GFP-Mael	This study	N/A
pUbi-GFP-SpnE	This study	N/A
Software and algorithms		
Rmd notebook	This study	https://doi.org/10.5281/zenodo.7834381
clusterProfiler	Yu et al. 2012[49]	N/A
cytoScape	Shannon et al. 2003[50]	v. 3.9.1
BioEmbeddings	Dallago et al. 2021[24]	N/A
MUSCLE	Edgar, 2004[51]	N/A
GPS-SUMO	Zhao et al., 2014[52]	http://sumosp.biocuckoo.org/
MaxQuant	Cox et al. 2008[53]	v. 1.6.10.43
IUPred2A	Mészáros et al. 2018[25]	N/A
FlyBase		http://flybase.org
Other		
Schneider's *Drosophila* Medium	Gibco (Life Technologies)	21720–024
Fetal Bovine Serum	GEMINI bio-products	100–106
Penicillin/Streptomycin	Gibco (Life Technologies)	15140–122

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Maria Ninova (mninova@ucr.edu).

Materials availability

Fly lines and plasmids generated in this study are available upon request.

Data and code availability

The mass spectrometry raw and processed data are deposited to the ProteomeXchange Consortium (https://www.ebi.ac.uk/pride/) via the PRIDE repository with the dataset identifier PRIDE: PXD037421 and are publicly available as of the date of publication. All original code is available at Zenodo: https://doi.org/10.5281/zenodo.7834381.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Drosophila stocks and husbandry

Drosophila lines encoding GFP-tagged Piwi under the native promoter (third chromosome), UASp- GFP -Panoramix, -Maelstrom, and -Spn-E, on the second chromosome, and small hairpin RNA against the *smt3* gene, and GFP-tagged Aub under the control of the endogenous Aub promoter, were previously generated in Aravin/Fejes Toth laboratories.[14,20,39] Flies encoding UASp-driven His-Flag-tagged SUMO on the third chromosome were a gift from Courey Laboratory.[54] Flies encoding shRNA against *su(var)2-10*, *piwi*, and *white*, the maternal tubulin(Mt)-Gal4 driver, the iso-1 Celera sequencing strain, and smt[04493] were obtained from the *Drosophila* Bloomington stock center (#32956, #33724, #33623, #7063, #2057, #11378). To study SUMOylation in germ cells, four transgenes including UASp-HisFlag-SUMO, GFP-tagged target, UASp-shRNA and Mt-Gal4 driver were combined by crossing balanced parental lines encoding GFP-tagged target and Mt-Gal4 driver, and lines carrying UASp-shRNA and UASp-HisFlag-SUMO.

To generate SUMO-TR flies, the *smt3* locus + ~2kb upstream region fragments (chr2R:21704005–21717574, dm6 reference genome) were amplified with the introduction of a 6xHis tag and a T86 > R mutation using Gibson assembly, subcloned into a phiC31 backbone based on the pCasper5 vector, and integrated into the P{CaryP}attP2 locus at BestGene Inc.

All stocks were maintained at 25°C on standard molasses-based media. Experiments were performed on samples from 3–5-day old females maintained on standard media supplemented with yeast for 2 days prior to dissections. Ovaries were manually dissected in cold PBS, flash frozen in liquid nitrogen and stored at −80°C.

METHOD DETAILS

Proteomics sample preparation

The preparation of SUMO-derived GG-modified peptide samples was adapted from Impens et al. with some modifications.[22] All procedures were performed using low binding plasticware, HPLC grade water, and freshly prepared solutions. Hand dissected ovaries from 600 2-5-day old *D.melanogaster* individuals of smt3[04493]/CyO; {6xHis-smt3[T86R]}attP2 genotype (SUMO-TR) or iso-1 (Bloomington #2057) (control) were used in each experiment. Frozen samples were immediately lysed in 2.5 mL denaturing buffer (6M guanidium HCl, 10 mM Tris, 100 mM sodium phosphate buffer pH 8.0) using glass Potter-Elvehjem tissue grinder. Lysates were cleared by centrifugation at 20,000rpm for 10 min at 4°C. Cleared lysates were treated by 5mM tris(2-carboxyethyl)phosphine for 30 min at 37°C with gentle rotation, followed by 10 mM N-ethylmaleimide (Sigma) for 30 min at room temperature, and finally, 10 mM Dithiothreitol (Sigma). Lysate volume was brought to 8 mL with lysis buffer and imidazole was added to 5 mM final concentration. Lysates were incubated with 1 mL HisPur Ni-NTA Resin (ThermoFisher) overnight at 4°C with gentle rotation. After incubation, the Ni-NTA slurry was washed once with lysis buffer supplemented with 10 mM imidazole, once with wash buffer 1 (8M urea, 10 mM Tris, 100 mM sodium phosphate (pH 8.0), 0.1% Triton X-100, 10 mM imidazole), and three times with wash buffer 2 (8 M urea, 10mM Tris, 1 mM sodium phosphate buffer pH 6.3, 0.1% Triton X-100, 10mM imidazole), followed by two rounds of elution with elution buffer (300mM imidazole, sodium phosphate buffer pH 6.8) for 2 h at 4°C to a final eluate volume of 1.5 mL. The eluate volume was increased to 10 mL with 50 mM ammonium bicarbonate and treated with sequencing grade trypsin (Promega, V5111) at 1:50 trypsin:protein ratio overnight at 37°C with gentle agitation. The following steps were performed according to the manual of the PTMScan HS Ubiquitin/SUMO Remnant Motif (K-ε-GG) Kit (Cell Signaling #59322). In brief, trypsinized samples were acidified to pH 2–3 by the addition of 0.5 mL 20% trifluoroacetic acid (TFA), kept on ice for 15 min, and centrifuged to remove potential precipitates. Peptides were then purified on a C18 column according to the kit manual, and lyophilized for >24hrs. Dry peptides were then resuspended in an IAP buffer (Cell Signaling #59322) and incubated with anti-diGly antibody-conjugated slurry (Cell Signaling #59322) at 4°C overnight, followed by two washes with IAP buffer, and 3 washes with HPLC-grade water. Finally, peptides were eluted with 0.15% TFA, purified using C18 StageTips, vacuum dried, and submitted to the Caltech Proteome Exploration Laboratory for Mass spectrometry analysis.

LC-MS/MS and raw data processing

Label-free peptide samples were subjected to LC-MS/MS analysis on an EASY-nLC 1200 (Thermo Fisher Scientific, San Jose, CA) coupled to a Q Exactive HF Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a Nanospray Flex ion source. Samples were directly loaded onto an Aurora 25 cm × 75 μm ID, 1.6μm C18 column (Ion Opticks) heated to 50°C. The peptides were separated with a 60 min gradient at a flow rate of 220 nL/min for the in-house packed column or 350 nL/min for the Aurora column. The gradient was as follows: 2–6% Solvent B (3.5 min), 6–25% B (42 min), 25–40% B (14.5 min), 40–98% B (1 min), and held at 100% B (12 min). Solvent A consisted of 97.8% H2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8% H2O, 80% ACN, and 0.2% formic acid. The Q Exactive HF Orbitrap was operated in data dependent mode with the Tune (version 2.7 SP1build 2659) instrument control software. Spray voltage was set to 1.5 kV, S-lens RF level at 50, and heated capillary at 275°C. Full scan resolution was set to 60,000 at *m/z* 200. Full scan target was 3 × 106 with a maximum injection time of 15 ms. Mass range was set to 400−1650 *m/z*. For data dependent MS2 scans the loop count was 12, AGC target was set at 1 × 105, and intensity threshold was kept at 1 × 105. Isolation width was set at 1.2 *m/z* and a fixed first mass of 100 was used. Normalized collision energy was set at 28. Peptide match was set to off, and isotope exclusion was on. Data acquisition was controlled by Xcalibur (4.0.27.13), with ms1 data acquisition in profile mode and ms2 data acquisition in centroid mode.

Thermo raw files were processed and searched using MaxQuant (v. 1.6.10.43).[55,53] Spectra were searched against *D. melanogaster* UniProt entries plus the His-SUMO-TR sequence and a common contaminant database. Trypsin was specified as the digestion enzyme and up to two missed cleavages were allowed. False discovery rates were estimated using a target-decoy approach, where the decoy database was generated by reversing the target database sequences. Protein, peptide, and PSM scores were set to achieve a 1% FDR at each level. Carbamidomethylation of cysteine was specified as a fixed modification. Protein N-terminal acetylation, methionine oxidation, and diGly remnant on lysine were specified as variable modifications with a maximum of 5 mods per peptide. Precursor mass tolerance was 4.5 ppm after mass recalibration and fragment ion tolerance was 20 ppm. Search type was specified as Standard with multiplicity of 1. Fast LFQ and normalization were used, and both re-quantify and match-between-runs were enabled.

Bioinformatics analysis of SUMOylation sites

Summary tables of the normalized diGly sites output from MaxQuant (GlyGly (K)Sites.txt, Table S1) were analyzed and figures were generated using custom R scripts (see key resources table for code availability). In brief, SUMO sites were considered sites with MS intensity ratios in the SUMO-TR to corresponding control samples >10, and background diGly sites were considered sites with

SUMO-TR/control ratios <3. Motif searches were performed using regular expressions and the MoMo suite[56] using a 11 amino acid window centered on the predicted GG-modified lysine. Protein structure predictions were performed using the 'bio_embeddings' package with default parameters,[24] or IUPred2 with default parameters.[25] To test whether the numbers of diGly sites falling within a specific structural region are different than expected by random chance, we created a bootstrap distribution (n = 1,000) of the fraction of randomly selected lysines from the same set of proteins falling within each region. Functional enrichment analyses presented on Figure 3 were carried out using the clusterProfiler R package[49] with annotations from the OrgDb package (GO source date 2021-Sep13). Semantic simplification was used to merge highly redundant terms. We note that similar results were obtained using TopGO, STRING and the bINGO Cytoscape plugin (not shown). STRING Interactions were plotted using CytoScape at 0.6 confidence cutoff and nodes were grouped manually according to association with specific GO categories flagged by STRING annotations. Interaction networks of Piwi, HP1a, Egg, Wde and Vasa were retrieved from FlyBase and nodes present in the diGly SUMO dataset were custom colored in CytoScape.[50]

The evolutionary conservation of SUMO sites using annotated homologs in *Drosophila sechellia, Drosophila simulans, Drosophila erecta, Drosophila yakuba, D, ananassae, Drosophila willistoni, Drosophila pseudoobscura, Drosophila persimilis, Drosophila virilis, Drosophila grimshawi, Drosophila mojavensis, and H. sapiens* on FlyBase.[26] For conservation analysis in the *Drosophila* genus, amino acid sequences of protein isoforms for the homologous genes were retrieved from FlyBase, and multiple sequence alignments for each gene were performed using Muscle with default parameters,[51] and alignment blocks of 10 amino acid upstream and downstream the '*bona fide*' SUMO sites of the *D. melanogaster* protein were extracted using custom R script. Protein isoforms from the same gene and species with identical sequence in this region were considered redundant and merged. For the alignment blocks around the SUMO site for genes conserved in all 11 *Drosophila* species, we calculated the ''conservation score'' – the percentage of sites in other species that have the same amino acid as in *D. melanogaster*. Human conservation analysis was performed similarly but based on pairwise alignments between the *D. melanogaster* protein and the Flybase-annotated human homolog with highest DIOPT score. Human protein sequences were retrieved from uniprot.org.

Detection of protein SUMOylation by immunoprecipitation (IP) and western blotting

All samples were prepared in parallel with respective controls. Hand dissected ovaries from 100 to 150 flies (depending on the expressed protein) of appropriate genotypes were lysed in RIPA buffer supplemented with Complete Protease Inhibitor cocktail (Roche, 11836170001) and 20 mM NEM. Debris were removed by centrifugation at 20,000 rpm for 10 min at 4°C. Cleared lysates were incubated with GFP-Trap magnetic agarose beads (ChromoTek, gtma-20) for 1–2 h with end-to-end rotation at 4°C. Beads were washed 5 times with ''harsh'' wash buffer (20 mM Tris pH 7.4, 0.5 M NaCl, 1% NP40 (Igepal), 0.5% Sodium deoxycholate, 1% SDS). Finally, samples were transferred to fresh tubes and boiled in LDS sample buffer (Invitrogen, NP0007) for 5 min at 95°C. For optimal separation of high molecular weight SUMOylated species, IP samples were analyzed on 3–8% tris-glycine gels (Invitrogen). Additionally, input samples were separately analyzed at higher percent gels to capture low molecular weight free SUMO. After electrophoresis, proteins were transferred to 0.45μm nitrocellulose membrane (BioRad) and transfer was verified by Ponceau S stain. Membranes were then blocked with 5% nonfat milk in PBS-T(0.1% Tween in PBS) for >30 min at room temperature (RT), followed by incubation with primary antibodies for 2 h at RT or overnight at 4°C. The following antibodies were used for detection: HRP-conjugated anti-Flag antibody (Sigma, A8592), rabbit anti-GFP (Abcam, ab290), mouse anti-Tubulin (Sigma, T5168), mouse anti-Piwi (Santa Cruz, sc-390946), HRP-conjugated anti-mouse IgG (Cell Signaling, 7076), HRP-conjugated anti-rabbit IgG (Cell Signaling, 7074), IRDye 800CW anti-rabbit IgG (Licor). For SUMOylation analysis in S2 cells, plasmids encoding GFP-tagged Piwi, Mael, Spn-E and Panx and 3xFlag3xHA-tagged SUMO under the control of ubiquitin and actin promoter, respectively, were co-transfected using TransIT-Insect Transfection reagent (Mirus, MIR6105). Cells were harvested 48 h post-transfection and samples were processed and analyzed following the procedure described above.

RT-qPCR

RNA was extracted from 20 to 25 pairs of ovaries from appropriate genotypes using Trizol (Invitrogen, 15596018) according to the manufacturer instructions. Purified RNA was treated with DNAse and subjected to reverse transcription with the Super-Script IV with ezDNAse kit (Invitrogen, 11766050) according to the manufacturer instructions. qPCR was performed using SybrGreen and the following primers: Su(var)2-10: CCAGCACAGGACGAACAGCCC and CGTGGAACTGGCGACGGCTT, Piwi: CTGCTGATCTCCA AAAATAGGG and TCGCGTATAACTGCTCATGG; Rp49 (internal control): CCGCTTCAAGGGACAGTATCTG and ATCTCGCCGCA GTAAACGC. Relative expression was calculated with respect to the Rp49 expression using the delta Ct method.

Protein localization imaging

Ovaries were dissected in cold PBS and fixed in 4% formaldehyde in PBS-T (0.1% Tween 20 in PBS) with end-to-end rotation for 20 min at room temperature, following by 3 washes with PBS-T for >10 min each. For Vasa detection, fixed ovaries were incubated with blocking solution (0.3% Triton X-, 0.1% Tween 20, 3% Bovine serum albumin (BSA)) for 1 h with end-to-end rotation, followed by an overnight incubation with 1:10 dilution of anti-vasa antibody (DSHB anti-vasa supernatant, AB_760351) at 4°C with rotation. Next, ovaries were washed 3 times with washing solution 0.3% Triton X-, 0.1% Tween 20 in PBS) for 10 min and incubated with 1:400 diluted AlexaFluor 594 secondary antibody (Invitrogen A-11007) in blocking solution for 2 h at RT in the dark. Finally, ovaries were washed with washing solution for 10 min 3 times, 1 μg/mL DAPI solution for 5 min, and a final rinse of 5 min. For proteins

GFP-Aub, fixed ovaries were only stained with DAPI. Samples were mounted on glass slides in Vectashield Antifade Mounting Medium (Vectorlabs, H-1200-10) and analyzed on an Upright Zeiss LSM 880 Confocal Microscope at the UCR Microscopy and Imaging Facility. Final images were processed and assembled with Fiji.[57] Shown are representative images of 10 or more analyzed ovaries.

Fertility test

10 2-day old females of appropriate genotype and 5 wild type males were maintained at standard fly media for 3 and 5 days, and the numbers of viable adult offspring from each vial was manually counted. Results are presented as viable offspring per day. Additionally, vials where females expressing shRNAs against Piwi and Su(var)2-10 from different crosses were maintained prior dissection were inspected for viable progeny to verify expected sterility phenotypes.

QUANTIFICATION AND STATISTICAL ANALYSIS

Detailed statistical analysis and software packages used for proteomics data are documented within the Rmarkdown notebooks available at https://doi.org/10.5281/zenodo.7834381. Western blot signal of modification of purified proteins was quantified using Fiji[57] gel analysis function, where multiple bands of Flag-SUMO-modified proteins were treated as a single, broad band. The relative densities of signal peaks were obtained by dividing the values of knockdown samples by the values of control samples separately for GFP and Flag signal (unmodified and modified proteins), and then changes of modified proteins normalized to unmodified proteins were expressed as relative to the control (shW), i.e. $(KD_{FlagSUMO\_IP}/Control_{FlagSUMO\_IP})/(KD_{GFP\_IP}/Control_{GFP\_IP})$. For Western blot signal data, statistical analyses were performed by one-way Anova followed by pairwise comparisons (for experiments with 3 conditions, followed by Dunnett's multiple comparisons test), or Student's *t* t-test (for experiments with 2 conditions).