Dual DNA/RNA-binding factor regulates dynamics of hnRNP splicing condensates

Mukulika Ray[*], Julia Zaborowsky[*], Pranav Mahableshwarkar, Smriti Vaidyanathan, Jasmine Shum, Renjith Viswanathan, Annie Huang, Szu-Huan Wang, Victoria Johnson, Noah Wake, Ashley M. Conard, Alexander E. Conicella, Ryan Puterbaugh, Nicolas L. Fawzi[†] and Erica Larschan[†]

[*] Co-First Authors

[†]Co-Corresponding authors

Abstract (70 words)

Despite decades of research, mechanisms by which co-transcriptional alternative splicing events are targeted to the correct genomic locations to drive cell fate decisions remain unknown. By combining structural and molecular approaches, we define a new mechanism by which an essential transcription factor (TF) targets co-transcriptional splicing through physical and functional interaction with RNA and RNA binding proteins (RBPs). We show that an essential TF co-transcriptionally regulates sex-specific alternative splicing by directly interacting with a subset of target RNAs on chromatin and modulating the dynamics of hnRNPA2 homolog nuclear splicing condensates.

Main text (1500 words)

Mechanisms to drive precise alternative splicing at thousands of genomic loci, which establish sex and cell type-specific transcriptome diversity, remain important targets for biological understanding because they establish sex and cell type-specific transcriptome diversity. Chromatin-bound transcription factors (TFs) and RNA binding proteins (RBPs) co-transcriptionally regulate alternative splicing [1-4]. However, the mechanisms by which TFs and RBPs coordinate transcription and alternative splicing at specific genomic locations in a

particular cell type to drive specific alternative splicing events remain poorly understood. Most TFs bind DNA and RNA[5-7] and interact with diverse RBPs[8,9]. Nevertheless, the mechanisms by which interactions between TFs, RNA, and RBPs generate the specific alternatively spliced transcripts essential for developmental decisions remain unknown. We hypothesize that TFs are critical to targeting the correct co-transcriptional splicing events to specific genomic locations due to their unique ability to bind chromatin, RNA, and spliceosomal RBPs.

We test this hypothesis by defining the role of GA-binding TF CLAMP (chromatin-linked adapter for MSL Proteins) in targeting sex-specific splicing to the correct genomic locations on chromatin in *Drosophila.* CLAMP has several properties that suggest that it could target sex-specific alternative splicing events to chromatin: 1) CLAMP binds to specific GA-enriched DNA motifs via its mapped DNA binding domain, and its binding sites differ in males and females[10-13]; 2) CLAMP is a pioneer TF that regulates sex-specific splicing in embryos at genes where it does not regulate transcription[14]; 3) CLAMP is enriched at the intronic region of CLAMP-dependent sex-specifically spliced genes[11,14]; 4) CLAMP is associated with RBPs that are spliceosome components[9].

Here, we define direct CLAMP-RNA interactions and compare them to CLAMP-DNA interactions to identify associations between TF-DNA and TF-RNA binding at genes where CLAMP regulates sex-specific alternative splicing. Furthermore, we determine how CLAMP affects the dynamics of RBPs, which are known to regulate alternative splicing as part of nuclear splicing condensates.

To define RNA species that interact with CLAMP in different cellular compartments at high resolution, we performed fractionation iCLIP (**i**ndividual-nucleotide resolution **C**ross**L**inking and **I**mmuno**P**recipitation) to identify CLAMP RNA targets in chromatin (ChF) and nucleoplasmic (NF) fractions of male (S2) and female (Kc) *Drosophila* cell lines, S2 well established cell lines for studying sex-specific processes such as dosage compensation [15,16].

Most CLAMP interaction with RNA occurs on chromatin with unique sex-specific targets (Fig 1A, Table 1a). Interestingly, CLAMP also directly interacts with spliceosomal RNAs sex-

specifically (Fig 1B). In the male chromatin fraction, CLAMP interacts with the catalytic step 2 spliceosome consisting of U2, U5, and U6 snRNAs (FDR:1.7E-3). In contrast, the female chromatin fraction is enriched for transcripts that encode proteins that bind to the U1-U2 snRNAs (FDR:1.1E-2), suggesting that CLAMP regulates splicing differently in males and females.

Next, we asked how the CLAMP interaction with DNA correlates with its interaction with RNA on chromatin. Therefore, we plotted the frequency of identifying a CLAMP RNA binding peak on chromatin over a region ±1 kb from the closest CLAMP DNA binding peak to define the following categories: a) complete overlap of RNA peaks with DNA peaks; b) partial overlap of RNA peak with the ends of a DNA peak; and c) RNA peaks nearby (±1 kb) DNA peaks (Fig 1C, Fig S1a-c). We found that most overlapping CLAMP RNA peaks are within 250 bp of the middle of the nearest CLAMP DNA peak in both male and female cells (Fig 1C), suggesting that CLAMP links RNA to DNA during co-transcriptional splicing at a subset of its target genes (Fig 1D- E; Table 1b).

Next, we combined *in vivo* and *in vitro* approaches to define how CLAMP regulates alternative splicing through interaction with RNA and RBPs. CLAMP does not have a canonical **R**NA **r**ecognition **m**otif (**RRM**) but contains a prion-like domain (PrLD), a subclass of intrinsically disordered domains (IDR) that are enriched in polar residues and found in yeast prion proteins (Fig 1F-H). Many RBPs and TFs contain PrLD domains[17,18] that can promote RNA-binding and drive the formation of phase-separated biomolecular condensates[19]. PrLD domains in TFs promote co-aggregation with RBPs to regulate transcription[20-22]. However, mechanisms by which the PrLD domains in TFs regulate alternative splicing remained unknown.

We define the role of CLAMP PrLD by generating complete (CLAMP Δ154-290a.a) and partial (CLAMP Δ160-290a.a) PrLD deleted mutant fly lines using CRISPR/Cas9 genomic mutagenesis and homologous repair[39]. Complete deletion of the PrLD ($CLAMP^{delPrLD}$) results in embryonic lethality. In contrast, partial deletion ($clamp^{delPrLD+6Q}$) mutants that retain a stretch of 6 glutamines survive, suggesting that these glutamines (aa154-160) are essential for viability (Fig 1H). Therefore, we assayed sex-specific alternative splicing in trans-heterozygous mutants

(*clamp*^{delPrLD}/*clamp*$^{delPrLD+6Q}$) to assure viability until it is possible to define the sex of animals at the larval stage. Analysis of RNA sequencing data using the time2splice pipeline (Ray, Conard, et al., 2023) from trans-heterozygous mutants (*clamp*^{delPrLD}/*clamp*$^{delPrLD+6Q}$) identifies CLAMP PrLD-dependent female (Table 2a) and male-specific splicing (Table 2b). Therefore, the CLAMP PrLD is essential for viability and impacts sex-specific alternative splicing.

Next, we asked whether the PrLD domain regulates the CLAMP-RNA interaction using NMR and RNA-Protein EMSAs (Electrophoretic Mobility Shift Assays). Both assays demonstrated that the PrLD domain in CLAMP promotes the interaction of CLAMP with RNA (Fig 1I-J). Using NMR spectroscopy to probe a mixture of total yeast RNA extract and the isolated CLAMP PrLD, we find that CLAMP PrLD directly interacts with RNA (Fig 1I). CLAMP is associated with the Male Sex-lethal Complex (MSL), which regulates dosage compensation in males and contains the *roX* long non-coding RNAs[12,23]. However, it was not known whether CLAMP directly interacts with the *roX* RNAs. Because the *roX* RNAs have a defined structure and function[24], we used *roX2,* identified by iCLIP as an *in vivo* male-specific CLAMP interactor (Table 1a), to determine whether CLAMP directly interacts with RNA *in vitro*. Therefore, we used a *roX2* probe (411 nt) for RNA-Protein EMSA gel shift assays with full-length CLAMP^{WT} and CLAMP^{delPrLD} proteins that were expressed and purified from *E. coli*. At the same protein and RNA concentrations, the CLAMP^{delPrLD} protein binds to RNA less efficiently than CLAMP^{WT} (Fig 1J). Furthermore, iCLIP data indicate that CLAMP binds to the *roX2* RNA adjacent (12nt downstream) to a stem-loop region (Fig S2), which is essential for *roX2* interaction with RNA helicase A, a component of both the MSL complex and the spliceosome[24]. Therefore, we hypothesized that the *roX2* stem loop promotes interaction with CLAMP. To test this hypothesis, we also designed and tested the following *roX2* mutant probes: a) a full-length probe lacking the stem-loop and CLAMP binding region (69nt), which reduced CLAMP binding to the *roX2* RNA (Fig S2, Fig 1K), and b) a full-length probe lacking only the *in vivo* CLAMP binding region (25nt) but not the stem loop, which did not qualitatively affect CLAMP binding to the *roX2* RNA *in vitro* (Fig S2, Fig 1L). Therefore, the presence of the roX2 RNA stem-loop and the CLAMP PrLD domain increased the ability of CLAMP to interact with RNA.

To determine whether the CLAMP PrLD domain regulates sex-specific alternative splicing of RNAs that are directly bound by CLAMP (Table 2a, b), we compared the RNAs bound to CLAMP *in vivo* identified by iCLIP (Table 1a) in cell lines with CLAMP PrLD-dependent sex-specifically spliced genes in L3 larvae (Table 2a, b). We determined that 14 CLAMP PrLD-dependent sex-specifically spliced genes are direct CLAMP RNA interactors, identifying a smaller subset of key genes for further analysis (Table 3). Moreover, most of these 14 direct targets are themselves regulators of alternative splicing (Table 3). Therefore, it is possible that CLAMP, which is heavily maternally deposited[25] and regulates 60% of all sex-specific spliced isoforms, functions as an essential upstream splicing regulator by directly regulating the splicing of splicing factors, which then regulates alternative splicing of additional target genes.

To define how CLAMP regulates sex-specific splicing beyond directly interacting with the RNA of target genes, we determined how CLAMP regulates the splicing and dynamics of one of its essential target genes and interactors, the hnRNPA2B1 orthologue, *hrp38*. One of the 14 CLAMP PrLD-dependent spliced genes that is a direct CLAMP RNA target is the *hrp38* transcript (Table 2a, 3). Furthermore, we previously demonstrated using proteomics that CLAMP sex-specifically interacts with the Hrp38 protein that regulates alternative splicing[9,26].

Due to the functional interaction between CLAMP and Hrp38 at both the RNA and protein levels, we further defined how CLAMP regulates both the sex-specific splicing of the *hrp38* transcript and the sex-specific dynamics of Hrp38 which is a component of highly mobile nuclear splicing speckles which regulate alternative splicing[27-30]. First, we defined how CLAMP interacts with *hrp38* RNA using RNA EMSA. Integrating iCLIP with alternative splicing analysis suggested that CLAMP binds to exon1 of the *hrp38* transcript in males, which undergoes splicing to remove exon2. In contrast, in $CLAMP^{delPrLD}$/$CLAMP^{delPrLD+6Q}$ mutants, the *hrp38* transcript retains exon2 (Fig 2A, B). Therefore, we compared $CLAMP^{WT}$ binding to *hrp38* exon1 and exon2 probes using RNA EMSAs and found that at the same protein and RNA concentrations, CLAMP binds strongly to *hrp38* exon1 RNA with no unbound RNA remaining (Fig 2C). Furthermore, the PrLD domain of CLAMP is important for CLAMP-*hrp38* RNA binding (Fig 2D). Interestingly, $CLAMP^{WT}$ still binds to *hrp38* exon2 *in vitro* even though it does

not bind *in vivo,* suggesting that other RBPs modulate the specificity of CLAMP-RNA interactions *in vivo*.

Next, we validated that the alternative splicing of the *hrp38* gene is sex-specifically regulated by the CLAMP PrLD domain *in vivo* during development at the third instar larval stage in males and females. We found that in *clamp*^{delPrLD}/*clamp*$^{delPrLD+6Q}$ mutants, splicing of intron AF:r6:3R:28600568:28600718:+ between exon2 and exon4 (Fig 2B) is regulated by CLAMP as predicted by genome-wide splicing analysis (Table 2b). In both female and male mutants, splicing of intron AF:r6:3R:28600568:28600718:+ occurs However, in males, splicing was more efficient than in females (Fig 2E, F). Our *in vivo* splicing analysis demonstrates that the CLAMP PrLD domain is important for *hrp38* RNA splicing during male and female development. PrLD domains can drive the phase separation of proteins[19] and promote the formation of subnuclear bio-condensates involved in transcription[20,31] and splicing[19,32], therefore, we performed biochemical phase separation assays to determine how the CLAMP PrLD domain regulates the phase separation behavior of CLAMP (Fig 2G). We found that the CLAMP PrLD domain contributes to phase separation of the N-terminal half of CLAMP (residues 1-300), which is important because many RBPs involved in splicing form splicing condensates, and TFs form transcription condensates, which interact during co-transcriptional splicing[33]. Although the PrLD domains of TFs can promote co-aggregation with RBPs, current work does not yet explain how TF-RBP interaction regulates the dynamics of splicing condensates, which is known to regulate their function.

Interestingly, one splicing component that is part of dynamic nuclear splicing condensates or speckles is Hrp38[27-30], which is associated at three levels with CLAMP: 1) CLAMP interacts with *hrp38* RNA transcript (Table 1a); 2) CLAMP regulates alternative splicing of the *hrp38* transcript, and 3) CLAMP interacts with Hrp38 protein[9] (Fig. 1; Table 1a). Therefore, we hypothesized that CLAMP PrLD-mediated phase separation affects the phase separation properties of Hrp38 condensates. We tested our hypothesis by combining *in vitro* and *in vivo* approaches to measure condensate formation, stability, and dynamics. In *vitro*, CLAMP 1-300^{WT} and Hrp38 co-phase-separate into liquid-like (Fig 2H). To complement *in vitro* studies, we analyzed the dynamics of Hrp38 nuclear speckles/condensates *in vivo* during development in

live tissues isolated from males and females. Therefore, we compared the dynamics of Hrp38-GFP nuclear speckles/condensates in the presence of CLAMP^{WT} and CLAMP^{delPrLD} proteins. FRAP analysis shows that in both *clamp*^{delPrLD}/*clamp*$^{delPrLD+6Q}$ mutant male and female individuals, Hrp38 remains in a more immobile phase compared with matched wild-type controls (Fig 2I). Next, we analyzed the trajectories of moving Hrp38 speckles over time to determine whether CLAMP regulates their mobility. When we analyze trajectories of moving Hrp38 speckles over time, more restricted tracks denote more immobile fractions, while freely moving particles represent mobile fractions[34]. Interestingly, we found that even in wild-type controls, female Hrp38 speckles are more immobile than male Hrp38 speckles. Furthermore, in the absence of the CLAMP PrLD domain, Hrp38 speckles in males become more immobile (Fig 2J), becoming more female-like. Therefore, the CLAMP PrLD domain regulates the dynamics of Hrp38 speckles sex-specifically in live tissues, promoting their enhanced mobility in males.

Prior studies[35-38] suggest that larger aggregates of hnRNPs are more immobile and have decreased function. In CLAMP^{delPrLD} mutants, we observe bigger Hrp38 speckle size (Fig 2K) and loss of mobility (Fig 2I, J). We thus conclude that the CLAMP PrLD domain regulates Hrp38 splicing condensate size and, therefore, is likely to modulate the ability of Hrp38 condensates to regulate alternative splicing. Furthermore, the PrLD domain in CLAMP promotes CLAMP liquid droplet formation (Fig 2G), which co-localizes with Hrp38 splicing condensates (Fig 2H). Hence CLAMP PrLD may help maintain Hrp38's dynamic properties, preventing them from forming large non-functional aggregates. Therefore, our data support a model in which sex-specific differences in alternative splicing arise from differential TF-RBP and TF-RNA interactions (Fig. 1). How the sex-specific differential interactions are established requires future investigation.

Thus, we show for the first time that an intrinsically disordered domain (IDR) in a TF regulates alternative splicing by regulating the dynamics of splicing condensates. Also, our data suggest that TFs that bind to specific DNA motifs and have RNA binding properties are the best candidates for understanding how specific alternative splicing events are regulated in specific cellular contexts.

Methods:

Cell culture: Kc and S2 cells were maintained at 25°C in Schneider's media supplemented with 10% Fetal Bovine Serum and 1.4X Antibiotic-Antimycotic (Thermofisher Scientific, USA). Cells were passaged every three days to maintain an appropriate cell density.

Fly strains and husbandry: *Drosophila melanogaster* fly stocks were maintained at 24°C on standard corn flour sucrose media. Fly strains with complete (CLAMP Δ154-290a.a) and partial (CLAMP Δ160-290a.a) PrLD deleted mutant fly lines using CRISPR/Cas9 genomic mutagenesis and homologous repair[39]. We used the flyCRISPR Optimal Target Finder tool from the University of Wisconsin to design a CRISPR target sequence for $clamp^{delPrLD}$ and $clamp^{delPrLD+6Q39}$. We cloned target sequence oligonucleotides (one gRNA) for $clamp^{delPrLD}$ (sense:5'CTTCGGGTACAACGCCAAAGCGAG3';antisense:3'CCCATGTTGCGGTTTCGCTC-CAAA5') and two gRNAs for $clamp^{delPrLD+6Q}$ (sense: 5'CTTCG-AATAGAATCGCCGCCCGCT3';antisense:3'CTTATCTTAGCGGCGGGCGACAAA5')and(se nse:5'CTTC-GTTGTGGCTGCACAGACTGG3';antisense:3'CAACACCGACGTGTCTGACC-CAAA5') into the pU6-BbsI-chiRNA plasmid (Addgene no. 45946), following the protocol outlined on the flyCRISPR website. We validated the correct ligation of the *clamp* CRISPR target sequence into the pU6-BbsI-chiRNA plasmid by Sanger sequencing using universal M13 primers. For homologous repair, we used ssODN15'ACATAAGCTTTAAGTGTGACGTATGTTCAGATATGTTCCCTCATTTGGCAC TTCTTAATGCTCATAGAGAGGCGAGCAGTGGGTCTGGCCATCATCCTGTGAAAAAAC GAAATTCCCAGCAGATGACCAAAT3' for $clamp^{delPrLD}$andssODN2 5'CTTCTTAATGCTCATAAGCGGATGCATACAGACGGGGAACAGCAGCAACAACAGC AACATAACGCCCAAGCTGGCGGTACAACGCCAAAGCGAGAGGCGAGCAGTGGGTCT GGCCATCATCCTGTGAAAAAA3' for $clamp^{delPrLD+6Q}$.The commercial service, BestGene

Inc., microinjected the validated pU6-BbsI-chiRNA plasmid containing the *clamp* target sequence into germline-expressing Cas9 flies Bloomington stock #51324. Flies containing a single mutation were returned balanced over the *Curly of Oster* (*CyO)* second chromosome balancer. From these progenies, we identified the CRISPR/Cas9-generated mutations by PCR across the target region (forward:5′-GATATGTTCCCTCATTTGGCAC-3′, reverse:5′-CACTCCCATGCTTCACACAG-3′). We isolated two independent clamp alleles from this validation: (1) y^1, w^{1118}; $clamp^{delPrLD}/CyO$; and (2) y^1, w^{1118}; $clamp^{delPrLD+6Q}/CyO$.These were crossed to obtain male and female $clamp^{delPrLD+6Q}/clamp^{delPrLD}$ genotypes. Crisper/Cas9 generated fly mutants and fly strain y^1 w^{1118}; P{w[+mC] =PTT-GC} Hrb98DE[ZCL0588] expressing Hrp38GFP (Bloomington #6822) were crossed obtained $clamp^{delPrLD+6Q}/clamp^{delPrLD}$male and female expressing GFP tagged Hrp38 protein.

CUT&RUN in cell lines: Cells were allowed to grow to confluency and harvested. An equal number of cells for each category were suspended in wash buffer and subjected to Cut&Run assay according to Skene et al. 2018[40] using rabbit anti-CLAMP (5µg) to immunoprecipitate CLAMP bound DNA fragments from male (S2) and female (Kc) cell lines. Three replicates for males and females were run, but one female sample was dropped during later stages due to insufficient starting material. Rabbit IgG was used as a control for each male and female cell line sample. 1ng CUT&RUN DNA was used to generate libraries using the Kapa Hyper prep kit (Roche, USA) and SeqCapAdapter Kit A (Roche, USA). 14 PCR cycles were used to amplify the libraries. AMPure XP beads (Beckman Coulter, USA) were used for library purification, and fragment analysis was performed to check the library quality. Paired-end 2x25 bp Illumina Hi-seq sequencing was performed.

RNA-seq in third instar larvae (L3): Total RNA was extracted from control (y^1 w^{1118}; *hrp38GFP*) and *clamp* mutant (*yw,* $clamp^{delPrLD+6Q}/clamp^{delPrLD}$; *hrp38GFP*) male and female third instar larvae (3 each) using Trizol (Invitrogen, USA). Four replicates were in each category prepared. Messenger RNA was purified from total RNA using poly-Toligo-attached magnetic beads. After fragmentation, the first strand of cDNA was synthesized using random hexamer primers, followed by the second cDNA synthesis. The library was ready after end repair, A-tailing, adapter ligation, size selection, amplification, and purification, followed by paired-end

RNA-sequencing in Illumina Novaseq 6000. The sequencing data was run through a SUPPA-based time2splice pipeline[14] to identify CLAMP-dependent sex-specific splicing events. Data is to be submitted to the GEO repository.

iCLIP: Cells were allowed to grow to confluency, and UV crosslinked using 254 nm UV light in Stratalinker 2400 on ice (Stratagene, USA). UV-treated cells were lysed to get different cellular fractions (Cytoplasmic, Nucleoplasmic, and Chromatin) according to the Fr-iCLIP (fractionation-iCLIP) protocol from Brugiolo et al 2017[42]. Chromatin and nucleoplasmic fractions were sonicated with a Branson digital sonicator at 30% amplitude for 30 seconds (10 sec on and 20 sec off) to disrupt DNA before IP. All three fractions were separately centrifuged at 20,000 xg for 5 min at 4℃. Fractions were tested by Western blotting using RNApolI for Chromatin Fraction and Actin for Cytoplasmic Fraction. Protein quantification for each fraction was done using the manufacturer's protocol for Pierce 660 nm protein assay reagent (Thermo Scientific, USA). Each Fraction was subjected to iCLIP protocol as described in Huppertz et al. 2014[41] using rabbit-CLAMP antibody to immunoprecipitate bound RNAs extracted using proteinase K and phenol: chloroform. Custom cDNA libraries prepared according to Huppertz et al. 2014[41] using distinct primers Rt1clip-Rt16clip for separate samples containing individual 4nt-barcode sequences that allow multiplexing of samples. cDNA libraries for each sample were amplified separately using 31 cycles of PCR, mixed later, and sequenced using standard Illumina protocols. Heyl et al. 2020[42] methods using the Galaxy CLIP-Explorer were followed to preprocess, perform quality control, post-process, and peak calling.

Fluorescence recovery after photobleaching (FRAP):

In vivo (Malphighian tubule principle cell nucleus expressing Hrp38GFP in control and $clamp^{delPrLD+6Q}/clamp^{delPrLD}$male and female third instar larvae) FRAP was performed on a Nikhon Spin-disc Confocal Microscope with a 488 nm laser on a 60× objective taking frames without delay (short time course and fast-recovering control) with 3-sec acquisition pre-bleaching, bleaching using 473 nm laser at 100% and2 minutes acquisition post-bleach. Malphighian tubules (MTs) were dissected in Grace's Insect media with 1:50 dilution of ProlongLive (antifading agent). A 6% slurry of low melting agarose (A9414-5G) in Grace's

Insect media was used to stabilize the MTs during imaging on a bridge slide with a cavity to pour the agarose, which was allowed to solidify, forming a soft base to mount the tissue. The intensities recorded on selected regions of interest were obtained using NIS element software. Data fitting and immobile fraction analysis were obtained with NIS element software FRAP analysis module.

***In-vivo* live imaging:** Salivary gland expressing Hrp38GFP (nuclear) in control and $clamp^{delPrLD+6Q}/clamp^{delPrLD}$male and female third instar larvae were dissected in Grace's Insect media with 1:50 dilution of ProlongLive (antifading agent) and mounted in the dissecting media on a bridge slide with low melting agarose (A9414-5G) in Grace's Insect media as a base. Moving Hrp38GFP condensates in the nucleus were imaged using a 488 nm laser on Nikon Spin-disc Confocal Microscope at 60X magnification for 2 minutes each without any delay in acquisition.

Validation of hrp38 splicing using RT-PCR assay: Total RNA was extracted from 5 third instar larvae (L3) female and male embryos expressing $clamp^{delPrLD+6Q}/clamp^{delPrLD}$and y^1, w^{1118}. Following the manufacturer's protocol, we reverse-transcribed one microgram of total RNA using the SuperScript VILO cDNA Synthesis Kit (Life Technologies, USA). We amplified target sequences by PCR using primers designed to span alternatively spliced junctions (FP-5'AGAACGGCAACTCCAATGGC3' and RP-5'GCCAGTCTCCTTGTCAATGA3') and Quick load Taq 2X Master mix (#M0271L, NEB, USA) according to the manufacturer's protocol (28 cycles). 10ul of PCR product of each replicate for each gene was loaded in separate wells in 2% agarose gels and imaged using a ChemiDoc™ MP Imaging system (BioRad, USA). All replicates for each gene were loaded on the same gel. The gel images were quantified using the densitometry steps with the Fiji image analysis tool. Student's t-tests were performed to determine significant differences between groups (two samples at a time). Three replicates for RT-PCR samples were performed.

CLAMP Protein Expression and Purification: Vectors encoding Maltose binding protein (MBP)-tagged CLAMP 1-300 and MBP-tagged CLAMP 1-300 delPrLD were produced by cloning into the pTHMT vector. Plasmids were transformed into BL21 cells and bacteria cultures

were grown in M9 minimal medium supplemented with ^{15}N ammonium chloride. Cultures were grown at 37°C and 200 RPM to an optical density of 0.6-0.8 and subsequently induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 4 hr at 37°C. Cell pellets were harvested by centrifugation at 6,000 RPM, resuspended in lysis buffer (20 mM Tris pH 8, 1 M NaCl, 10 mM Imidazole, 1 mM DTT, and one EDTA-free protease inhibitor tablet (Roche) per liter of culture), and lysed by sonication. The lysed cell suspension was centrifuged at 20,000 RPM for 50 min. The supernatant was filtered using a 0.2 μm syringe filter, and loaded onto a HisTrap HP 5 ml column. The HisTrap column was first washed with 20 mM Tris pH 8, 1 M NaCl, 10 mM Imidazole, 1 mM DTT Buffer, and then the bound protein was eluted with 20 mM Tris pH 8, 1 M NaCl, 500 mM Imidazole, 1 mM DTT buffer. Fractions containing MBP-CLAMP 1-300 were collected and purified on a Superdex 200 (26/600) column equilibrated in 20 mM Tris pH 8, 1.0 M NaCl buffer. CLAMP PrLD (154-290) was expressed and purified using Histrap HP as above. Following that (instead of size exclusion chromatography), CLAMP PrLD was collected from the Histrap HP, concentrated to 1 mL, and cleaved overnight with TEV protease to separate CLAMP PrLD from the MBP in 20 mM Tris pH 8, 1 M NaCl, 10 mM Imidazole, and 1 mM DTT Buffer. The cleaved protein was separated from the HisTag by a HisTrap subtraction with the same Tris buffer previously described. MBP-tagged Full-Length CLAMP was grown following the same methods above. Once the Full-Length CLAMP cell lysate supernatant was loaded onto HisTrap HP 5 ml column, the protein was eluted with a gradient from 10 to 500 mM imidazole in pH 5.5 20 mM MES and 361 mM $CaCl_2$. Fractions containing MBP-CLAMP FL were collected and purified on a Superdex 200 (26/600) column equilibrated in pH 5.5, 20mM MES, and 361mM CaCl2 buffer. For each CLAMP construct, fractions containing the desired protein were verified using SDS-PAGE, concentrated using a 10 kDa centrifugation filter (Millipore), aliquoted, and frozen.

Hrp38 protein purification and phase separation assay: Maltose binding protein (MBP)-tagged Hrp38 was expressed in BL21 *E. coli* cells. The cells were resuspended in 20 mM NaPi at pH 7.4 with 300 mM NaCl and 10 mM imidazole, and the lysate was cleared by centrifuging at 20,000 rpm for 60 mins at 4° C. The supernatant was filtered using 0.2 μM filters and loaded on a 5 ml Histrap HP column. The protein was eluted using an imidazole gradient of 10-300 mM over 5-column volumes. The protein fractions were pooled and loaded on a Superdex 200

(26/600) column for size exclusion. 20 mM NaPi with 300 mMNaCl at pH 7.4 was used for size exclusion chromotography and storage. Protein was flash frozen as aliquots of 1 mM. Fluorescent labeling of the CLAMP 1-300 and CLAMP 1-300 dPrLD was done using AlexaFluor 488 maleimide dye, while Hrp38 was labeled using 555 maleimide. The protein was diluted to 100 µM in 20 mM Tris buffer at pH 7.4 with 50 mMNaCl, and a 5-fold concentration (500 µM) of the dye dissolved in DMSO was added to 5% of the total volume. The reaction mixture was incubated for 1 hour, and then, unbound AlexaFluor was removed using 1 ml Zeba spin desalting columns. The proteins were concentrated to 1 mM, flash frozen, and stored. For phase separation assay and microscopy Hrp38 was buffer exchanged to 20 mM Tris buffer at pH 7.4 with 50 mM NaCl to a final concentration of 20 µM. We calculated all concentrations on NanoDrop by using the extinction coefficient 130,000 M^{-1} cm^{-1} for Hrp38 and 78270 for MBP CLAMP 1-300 and MBP CLAMP 1-300 dPrLD. 1 uM of AlexaFluor labeled CLAMP was added to Hrp38 and TEV protease was added to cleave the MBP tag. Less than 1% fluorescent protein was used in all samples. A Nikon spin-disc confocal microscope was used for imaging at 20X magnification with 1.5X zoom. The images were processed using ImageJ.

NMR Sample Preparation and Spectroscopy: ^{15}N Isotopically labeled samples of CLAMP 1-300 were prepared at 50 µM in a buffer containing 20 mM MES pH 5.5, 100 mM NaCl, 1 mM DTT, and 5% D_2O and then moved into a 5 mm NMR tube using a glass pipette. CLAMP concentration was determined by measuring absorbance at 280 nm (and then dividing absorbance by the extinction coefficient estimated by the Expasy ProtParam). NMR spectra were recorded on Bruker Avance 850 MHz ^{1}H Larmor frequency spectrometer with HCN TCI z-gradient cryoprobes. A two-dimensional $^{1}H^{15}N$ HSQC was acquired using spectral widths of 10.5 ppm and 30.0 ppm in the direct and indirect dimensions, with 3072 and 512 total points and acquisition time of 172 ms and 99.0 ms, respectively. Samples of isotopically labeled CLAMP PrLD were prepared at 20 µM in a buffer containing 50 mM MES pH 5, and 5% D_2O. Because the PrLD domain of CLAMP contains no tyrosines or tryptophans, the A_{280} absorbance could not be measured to determine the protein concentration. Instead, sample concentration was estimated by measuring the absorbance at 230 nm, and calculating the extinction coefficient to be 300 $M^{-1}cm^{-1}$ per peptide bond (40.5 $mM^{-1}cm^{-1}$ for CLAMP PrLD). Torula yeast was added to one PrLD sample at a 1:1 Protein to RNA ratio by weight and NMR spectra were acquired in 5 mm NMR

tubes on a Bruker Avance 850 MHz ^{1}H Larmor frequency spectrometer. Both two-dimensional $^{1}H^{15}N$ HSQC were acquired using spectral widths of 14.0 ppm and 25.0 ppm in the direct and indirect dimensions, with 2048 and 256 total points and acquisition time of 86.0 ms and 59.4 ms, respectively. All data was processed and analyzed using NMRPipe and CCPNMR Analysis 2.5 software[50,51].

RNA Electrophoretic Mobility Shift Assays: *rox2*RNA probes at 100 nM and *hrp38* RNA probes at 50nM were incubated with MBP-tagged FL CLAMP^{WT} protein or MBP-tagged FL CLAMP^{delPrLD} protein in REMSA binding buffer provided with the LightShift Chemiluminescent RNA EMSA kit (Thermo Scientific, USA) at room temperature for 30 min according to manufacturer's protocol. Reactions were loaded onto 6% TBE retardation gels (Thermo Fisher Scientific) and run in 0.5× Tris–borate–EDTA buffer for one hour. RNA-Protein complex was transferred to the Nylon membrane using the iBlot transfer system (ThermoFisher Scientific), and the probe signal was detected using a Chemiluminescent Nucleic acid detection module (#80880, ThermoFisher Scientific).

Computational Methods

CUT&RUN Data analysis: Sequenced reads were run through FASTQC[43](fastqc replicate_R1_001.fastq.gz replicate_R2_001.fastq.gz) with default parameters to check the quality of raw sequence data and filter out any sequences flagged for poor quality. Sequences were trimmed and reassessed for quality using TrimGalore (https://github.com/FelixKrueger/TrimGalore/issues/25) and FastQC[43], respectively. All Illumina lanes of the same flow cell. fastq files were merged, and sequenced reads were mapped to release 6 *Drosophila melanogaster* genome (dm6). We compared Bowtie2[44], HISAT2[45], and BWA[46]. We found the best alignment quality with BWA and thus used this method's results downstream. Next, we performed conversion to bam and sorting (e.g., using: bowtie2 -x dm6_genome -1 replicate_R1_001.fastq.gz -2 replicate_R2_001.fastq.gz -S out.sam> stout.txt 2> alignment_info.txt; samtools view -bSout.sam>out.bam; rm -rfout.sam; samtools sort out.bam -o out.sorted.bam). We removed reads (using samtools) with a MAPQ less than 30 and any reads with PCR duplicate reads (identified using MarkDuplicates Picard -2.20.2). Peaks identified

using MACS2[47](macs2 callpeak -t out.sorted.bam -B -f BAM --nomodel --SPMR --keep-dup all -g dm --trackline -n outname --cutoff-analysis --call-summits -p 0.01 --outdiroutdir) and keep duplicates separate**.** To calculate fold-enrichment macs2 is rerun (macs2 bdgcmp -t $treat -c $control -o $out.sorted.bam_FE.bdg -m FE 2> $ out.sorted.bam_FE.log; macs2 bdgcmp -t $treat -c $control -o $out.sorted.bam_logLR.bdg -m logLR -p 0.00001 2). For motif analysis, the MEME[48] suite was used. Data was submitted to the GEO repository (#GSE174781, #GSE220981 and #GSE220053).

iCLIP Data analysis: The method from Heyl et al. 2020[42] using the Galaxy CLIP-Explorer was followed to preprocess, perform quality control, post-process and perform peak calling. UMI tools were used for preprocessing, and then UMI tools and Cutadapt were used for Adapter, Barcode, and UMI removal. Cutadapt (Galaxy version 3.5) was used for filtering with a custom adaptersequenceAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTCGTAT GCCGTCTTCTGCTTG. All other settings followed the Heyl et al. 2020 Galaxy iCLIP-explorer workflow. UMI-Tools Extract (Galaxy Version 1.1.2+galaxy2) was then used with a barcode pattern of NNNXXXXNN. No unpaired reads were allowed. The barcode was on the 3' end. Je-Demultiplex (Galaxy Version 1.2.1) was then used for demultiplexing. FastQC was used for quality control. Mapping was done by RNA STAR (Galaxy version 2.5.2b-2) using dm6. All settings were chosen based on the existing parameters from the iCLIP-explorer settings. We selected FALSE for the option to use end-to-end read alignments with no soft-clipping. bedtools used for Read-Filtering, and UMI-Tools (Galaxy version 0.5.3.0) for de-duplication. PEAKachu was used for Peak Calling to generate bed files. The PEAKachu settings were followed using the Galaxy CLIP-explorer workflow. The maximum insert size was set to 150, the minimum cluster expression fraction was set to 0.01, the minimum block overlap was set to 0.5, and the minimum block expression was set to 0.1. The Mad Multiplier was set to 0.0, the Fold Change Threshold was set to 2.0, and the adjusted p-value threshold was set to 0.05. Peaks were annotated using RCAS[49] (RNA Centric Annotation System), an R package using Rstudio. MEME Suite is used for motif detection. RCAS was used for functional analysis of the transcriptomes isolated by iCLIP, such as transcript features. Data was submitted to the GEO repository (#GSE205987).

Integrating CUT&RUN and iCLIP data: A Python script was created that iterates through all of the DNA peak bed files for CLAMP DNA binding sites in Kc and S2 cell lines (CUT&RUN data, #GSE220053) as a reference and tests for overlap with CLAMP-bound RNA peaks (each sequence is between 25-50bp in size) in the Kc and S2 (iCLIP data, (#GSE205987). The overlaps are categorized into four main categories based upon the location of the overlap: 1) completely overlapping (purple lines in frequency plot), 2) partially overlapping at the DNA peak start site (red lines in frequency plot); 3) partially overlapping at the DNA peak end site (blue lines in frequency plot) and 4) non-overlapping, i.e., when there is an overlap in a region outside the DNA binding site (yellow lines in frequency plot). This extended region is defined by the *scope* variable in the script, allowing the overlap to look for binding sites near the DNA binding site (this scope is 2 kb, including the DNA binding site). We note that multiple RNA peaks can be found on one DNA peak. These overlaps are placed onto a [-scope, scope] region. Then, each type of overlap shown with a different color is overlaid and plotted onto a frequency plot. So, if the frequency at a given base pair is 5, five overlaps contain that base pair within the region defined by the scope.

Analysis of Imaging data: Live-image 2 minutes movies acquired through confocal microscopy were pre-processed using Fiji (ImageJ) Jython macro script, involving Fiji Plugins and built-in software on the Hrp38 condensate (green channel) to distinguish Hrp38 phase condensates from the cell background. Simple Ratio bleach correction was applied, and the minimum and maximum intensity adjustment was executed based on the mean intensity and standard deviation of intensity. The specific adjustment method varied depending on the image quality, ensuring optimal visibility of Hrp38 condensate speckles. Then, background subtraction was performed with a rolling ball radius of 120 pixels. The conversion to a binary image involved max entropy thresholding, although the minimum thresholding method was employed for some images to accommodate diverse lighting conditions. The thresholded images were saved as TIFF files for use in the TrackIt tracking software. In total, 15 male control movies, 28 male CLAMP del PrLD mutant movies, 15 female control movies, and 38 female CLAMP del PrLD movies were utilized in subsequent analyses.

For tracking analysis, TIFF movies from each experimental condition were loaded into TrackIt[34], a program for tracking and analyzing single-molecule experiments developed by the Gebhardt lab. The analysis was conducted using MATLAB on a Windows 10 PC, utilizing default settings with specific parameter adjustments: threshold (4), tracking radius (8), minimum track length (2), one gap frame allowed, and minimum track length before the gap (3). Each experimental condition was analyzed using TrackIt's MATLAB-based data visualization tool, with settings further edited for formatting in a MATLAB script.

Bound fraction analysis was done according to TrackIt package[34]. TrackIt provided information on tracked events, including long, short, linked, and nonlinked events. The calculation of bound fractions by TrackIt involved determining the ratio of the sum of short and long events to non-linked events. Movie-wise means were plotted, and the mean across all movies and the pooled fraction (sum over all movies for each type of event) were plotted to ensure that one movie did not unduly influence the mean.

Competing Interest Statement

The authors declare no conflicting interests.

Acknowledgments

This work was supported by R35GM126994 to E.N.L. from NIH and by National Science Foundation (Directorate for Biological Sciences) 1845734 to N.L.F. A.M.C is funded by the NSF Graduate Research Fellowship and CCMB, Brown University. We thank Bloomington stock center for fly lines. We thank Steve Henikoff for pAMNase protein and spike-in DNA for Cut and Run.

Author Contributions

M.R., **N. L. F** and **E.N.L**. planned experiments, analyzed results and wrote the manuscript. **J.Z** did all protein purifications, the NMR and phase separation assays for CLAMP. **M.R.** carried out the experimental work and collected data for Cut and Run, iCLIP, FRAP, Live imaging, RNA-EMSA, RNA isolation for RNA-sequencing, and generated fly mutants. **P.M** analyzed third instar larval RNA-seq data using time2splice pipeline to identify CLAMP dependent splicing

events, integrated CLAMP iCLIP data with CLAMP CUT&RUN in cell lines and performed the splicing assay for *hrp38*. **S.V** analyzed the live imaging data sets. **R.V** performed the co-LLPS assay with Hrp38 and CLAMP. **J.S** cloned CLAMP1-300 and CLAMP1-300del PrLD clones and helped with RNA-EMSA assays. **A.H** analyzed the iCLIP-seq data. **A.M.C** analyzed CUT&RUN data. **S.H.W** helped with NMR analysis. **V.J.** and **N.W.** helped with CLAMP Protein Expression and Purification. **A.E.C** and **R.P** helped with the protein and NMR assays.

References

1. Boumpas, P., Merabet, S. & Carnesecchi, J. Integrating transcription and splicing into cell fate: transcription factors on the block. *Wiley Interdisciplinary Reviews: RNA***14**, e1752 (2023).
2. Han, H. et al. Multilayered control of alternative splicing regulatory networks by transcription factors. *Molecular cell***65**, 539-553. e7 (2017).
3. Saulnier, O. et al. ERG transcription factors have a splicing regulatory function involving RBFOX2 that is altered in the EWS-FLI1 oncogenic fusion. *Nucleic Acids Research***49**, 5038-5056 (2021).
4. Jabre, I. et al. Does co-transcriptional regulation of alternative splicing mediate plant stress responses? *Nucleic Acids Research***47**, 2716-2726 (2019).
5. Oksuz, O. et al. Transcription factors interact with RNA to regulate genes. *Molecular Cell***83**, 2449-2463. e13 (2023).
6. He, C. et al. High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. *Molecular cell***64**, 416-430 (2016).
7. Cassiday, L.A. & Maher III, L.J. Having it both ways: transcription factors that bind DNA and RNA. *Nucleic acids research***30**, 4118-4126 (2002).
8. Göös, H. et al. Human transcription factor protein interaction networks. *Nature communications***13**, 766 (2022).
9. Urban, J.A., Urban, J.M., Kuzu, G. & Larschan, E.N. The Drosophila CLAMP protein associates with diverse proteins on chromatin. *PloS one***12**, e0189772 (2017).
10. Kuzu, G. et al. Expansion of GA dinucleotide repeats increases the density of CLAMP binding sites on the X-chromosome to promote Drosophila dosage compensation. *PLoS genetics***12**, e1006120 (2016).
11. Kaye, E.G. et al. Differential occupancy of two GA-binding proteins promotes targeting of the drosophila dosage compensation complex to the male X chromosome. *Cell reports***22**, 3227-3239 (2018).
12. Soruco, M.M. et al. The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. *Genes & development***27**, 1551-1556 (2013).
13. Albig, C. et al. Factor cooperation for chromosome discrimination in Drosophila. *Nucleic acids research***47**, 1706-1724 (2019).
14. Ray, M. et al. Sex-specific splicing occurs genome-wide during early Drosophila embryogenesis. *eLife***12**, e87865 (2023).
15. Straub, T., Zabel, A., Gilfillan, G.D., Feller, C. & Becker, P.B. Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. *Genome research***23**, 473-485 (2013).
16. Alekseyenko, A.A. et al. A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. *Cell***134**, 599-609 (2008).

17. March, Z.M., King, O.D. & Shorter, J. Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain research***1647**, 9-18 (2016).
18. Harrison, A.F. & Shorter, J. RNA-binding proteins with prion-like domains in health and disease. *Biochemical Journal***474**, 1417-1438 (2017).
19. Gotor, N.L. et al. RNA-binding and prion domains: the Yin and Yang of phase separation. *Nucleic acids research***48**, 9491-9504 (2020).
20. Wang, Y. et al. A prion-like domain in transcription factor EBF1 promotes phase separation and enables B cell programming of progenitor chromatin. *Immunity***53**, 1151-1167. e6 (2020).
21. Yang, J. et al. A prion-like domain of TFEB mediates the co-aggregation of TFEB and mHTT. *Autophagy***19**, 544-550 (2023).
22. Brodsky, S., Jana, T. & Barkai, N. Order through disorder: The role of intrinsically disordered regions in transcription factor binding specificity. *Current opinion in structural biology***71**, 110-115 (2021).
23. Eggers, N., Gkountromichos, F., Krause, S., Campos-Sparr, A. & Becker, P.B. Physical interaction between MSL2 and CLAMP assures direct cooperativity and prevents competition at composite binding sites. *Nucleic Acids Research***51**, 9039-9054 (2023).
24. Ilik, I.A. et al. A mutually exclusive stem–loop arrangement in roX2 RNA is essential for X-chromosome regulation in Drosophila. *Genes & development***31**, 1973-1987 (2017).
25. Duan, J. et al. CLAMP and Zelda function together to promote Drosophila zygotic genome activation. *Elife***10**, e69937 (2021).
26. Li, S. et al. Genetic interaction of hnRNPA2B1 and DNAJB6 in a Drosophila model of multisystem proteinopathy. *Human molecular genetics***25**, 936-950 (2016).
27. Blanchette, M. et al. Genome-wide analysis of alternative pre-mRNA splicing and RNA-binding specificities of the Drosophila hnRNP A/B family members. *Molecular cell***33**, 438-449 (2009).
28. Borah, S., Wong, A.C. & Steitz, J.A. Drosophila hnRNP A1 homologs Hrp36/Hrp38 enhance U2-type versus U12-type splicing to regulate alternative splicing of the prospero twintron. *Proceedings of the National Academy of Sciences***106**, 2577-2582 (2009).
29. Herold, N. et al. Conservation of the protein composition and electron microscopy structure of Drosophila melanogaster and human spliceosomal complexes. *Molecular and cellular biology* (2009).
30. Singh, A.K. & Lakhotia, S.C. Dynamics of hnRNPs and omega speckles in normal and heat shocked live cell nuclei of Drosophila melanogaster. *Chromosoma***124**, 367-383 (2015).
31. Mann, R. & Notani, D. Transcription factor condensates and signaling driven transcription. *Nucleus***14**, 2205758 (2023).
32. Hennig, S. et al. Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *Journal of Cell Biology***210**, 529-539 (2015).
33. Guo, Y.E. et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature***572**, 543-548 (2019).

34. Kuhn, T., Hettich, J., Davtyan, R. & Gebhardt, J.C.M. Single molecule tracking and analysis framework including theory-predicted parameter settings. *Scientific Reports***11**, 9465 (2021).
35. Forman-Kay, J.D., Ditlev, J.A., Nosella, M.L. & Lee, H.O. What are the distinguishing features and size requirements of biomolecular condensates and their implications for RNA-containing condensates? *RNA***28**, 36-47 (2022).
36. Deshaies, J.-E. et al. TDP-43 regulates the alternative splicing of hnRNP A1 to yield an aggregation-prone variant in amyotrophic lateral sclerosis. *Brain***141**, 1320-1333 (2018).
37. Shin, Y. et al. Spatiotemporal control of intracellular phase transitions using light-activated optoDroplets. *Cell***168**, 159-171. e14 (2017).
38. Shin, Y. & Brangwynne, C.P. Liquid phase condensation in cell physiology and disease. *Science***357**, eaaf4382 (2017).
39. Gratz, S.J., Rubinstein, C.D., Harrison, M.M., Wildonger, J. & O'Connor-Giles, K.M. CRISPR-Cas9 genome editing in Drosophila. *Current protocols in molecular biology***111**, 31.2. 1-31.2. 20 (2015).
40. Skene, P.J., Henikoff, J.G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nature protocols***13**, 1006-1019 (2018).
41. Huppertz, I. et al. iCLIP: protein–RNA interactions at nucleotide resolution. *Methods***65**, 274-287 (2014).
42. Heyl, F., Maticzka, D., Uhl, M. & Backofen, R. Galaxy CLIP-Explorer: a web server for CLIP-Seq data analysis. *GigaScience***9**, giaa108 (2020).
43. Andrews, S. FastQC: a quality control tool for high throughput sequence data. 2010. (2017).
44. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature methods***9**, 357-359 (2012).
45. Kim, D., Paggi, J.M., Park, C., Bennett, C. & Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology***37**, 907-915 (2019).
46. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics***25**, 1754-1760 (2009).
47. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology***9**, 1-9 (2008).
48. Bailey, T.L., Johnson, J., Grant, C.E. & Noble, W.S. The MEME suite. *Nucleic acids research***43**, W39-W49 (2015).
49. Uyar, B. et al. RCAS: an RNA centric annotation system for transcriptome-wide regions of interest. *Nucleic acids research***45**, e91-e91 (2017).
50. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR***6**, 277-293 (1995).
51. Skinner, S.P. et al. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *Journal of biomolecular NMR***66**, 111-124 (2016).

Figure legends

Figure. 1. **The PrLD domain is important for CLAMP-RNA interaction and phase separation. A**. Venn diagrams showing the distribution of CLAMP RNA targets between male and female cell types and between chromatin and nucleoplasm fractions (**four replicates** for each category performed except **three replicates** for the Kc nucleoplasm fraction). **B.** Venn diagrams showing the distribution of CLAMP snRNA targets between male and female cell types in chromatin and nucleoplasm fractions. The corresponding bar plots denote the total number of snRNAs CLAMP binds to in respective fractions and cell types. **C.** Frequency distribution of CLAMP RNA binding peaks (iCLIP data, **four replicates**) plotted over a region (±1kb across the middle of the DNA peak) spanning CLAMP DNA binding peaks (black line, CUT&RUN data, **three replicates** for males, **two replicates** for females). Complete overlaps are denoted by magenta, non-overlaps in yellow, partial overlaps in red (near the starting boundary of DNA peaks), and blue (near the ending of DNA peaks). **D-E.** Bar plots (**D**) and pie-chart (**E**) show the distribution of the number of CLAMP-RNA peaks that overlap with CLAMP-DNA peaks or are within ±1kb region of CLAMP-DNA peaks (proximal peaks, PXP). Overlapping peaks are sub-categorized into complete RNA peaks overlapping with DNA peaks (CRO), overlapping RNA peaks with DNA peak front (5' end, ORF), and overlapping RNA peaks with DNA peak 3' end (ORE). **F.** The $^{1}H^{15}N$ HSQC spectrum of 1-300 CLAMP at 50 μM visualizes the largely disordered nature of the protein. Dispersed resonances in the region from 7.9 – 7.0 ppm in the ^{1}H and 115-122 ppm in the ^{15}N likely arise from the folded zinc finger **G.** Diagram of full-length CLAMP depicting the location of the PrLD (residues 154-290) and other important features (**top**). Regions of full-length CLAMP (**below**) are predicted to resemble yeast prion protein (enriched in polar residues with little to no aliphatic or charged residues). **I.** $^{1}H^{15}N$ HSQC spectra of CLAMP PrLD at 20uM with (red) and without (blue) torula yeast RNA. **J.** RNA Electrophoretic mobility shift assay (RNA-EMSA) showing the binding of increasing amounts of MBP fusion CLAMP-Full length and CLAMP PrLD domain deleted protein to *rox2* RNA biotinylated probes (100 nM). Concentrations (μM) of CLAMP protein increase from left to right. **K.** RNA-EMSA showing difference in binding of full-length (FL) CLAMP (fused to MBP) to *rox2-RNA* probe with (411nt) and without (342nt) stem-loop and CLAMP-rox2 RNA (iCLIP

data) binding region. Concentrations (μM) of CLAMP FL increase from left to right. **L.** RNA-EMSA showing difference in binding of MBP-fusions of full-length (FL) CLAMP and CLAMP PrLD domain deleted protein to *rox2-RNA* probe without (399nt) CLAMP-rox2 RNA (iCLIP) binding region. Concentrations (μM) of CLAMP FL and CLAMPdel PrLD increase from left to right.

Figure. 2. **The CLAMP PrLD domain regulates hnRNPA2 homolog Hrp38 biomolecular condensate dynamics. A.** CLAMP-DNA binding (Cut&Run data) and CLAMP-RNA binding in chromatin fraction (iCLIP data) peaks visualized in the IGV genome browser at the *hrp38* gene location. **B.** Schematicof *hrp38* transcript showing CLAMP binding at exon1 and alternative splicing of intron AF:r6:3R:28600568:28600718:+ in $clamp^{delPrLD}/clamp^{delPrLD+6Q}$ mutant male. **C-D.** RNA EMSA mobility shift showing a difference in the efficiency of CLAMP full-length protein-*hrp38* exon1 and exon2 binding (**C**) and that of CLAMP-PrLD domain deleted and CLAMP full-length protein with *hrp38* exon1(**D**). $CLAMP^{delPrLD}$ and $CLAMP^{WT}$-Full length MBP fusions were used at increasing concentrations of 0, 3.1, 6.2, and 12.5 μM with 50 nM biotinylated RNA probes. **E-F.** Bar plot showing the change in levels of intron AF:r6:3R:28600568:28600718:+ spliced isoforms resulting from alternative splicing events in male and female L3 larvae under control $clamp^{WT}$(green) and $clamp^{delPrLD}/clamp^{delPrLD+6Q}$ (orange) conditions. The isoform transcript levels are normalized by the levels of *gapdh* housekeeping gene transcript. p-values (paired student's t-test) for groups showing significant differences are noted at the top of the line connecting the compared groups (**three replicates** for each category). The corresponding agarose electrophoretic gel image is shown in **F**. **G**. DIC micrograph of CLAMP 1-300 with or without PrLD domain (100 μM) at 150 mM salt concentration with 5% PEG. **H.** Single plane confocal images showing the co-LLPS of Alexa Fluor 488 maleimide dye-labeled Hrp38 with a single cystine added for labeling (green) and Alexa Fluor 555 maleimide dye-labeled 1-300 CLAMP (red) under phase separating conditions for Hrp38. **I**. Bar plots showing the percentage of Hrp38-GFP in the immobile phase (FRAP analysis) in male and female L3 larvae under control $clamp^{WT}$(green) and $clamp^{delPrLD}/clamp^{delPrLD+6Q}$ (orange) conditions. p-values (paired student's t-test) for groups showing significant differences marked by asterisk are noted at the top of the line connecting the compared groups—**n**=Number of Malphighian tubule principal cell nuclei, **N**= Number of

individuals. **J**. Mean bound fraction of Hrp38GFP speckles (Number of Hrp38GFP speckles bound, i.e., immobile track lengths/the total number of Hrp38GFP speckles with all types of track lengths over a fixed time-period) in salivary gland nuclei expressing Hrp38-GFP in male and female *clamp*^{WT}and *clamp*^{delPrLD}/*clamp*$^{delPrLD+6Q}$ individuals. Each dot represents each individual (movie), and '**n**' denotes the number of Hrp38 speckles identified in each category. **K.** Confocal image of a single plane of salivary gland nuclei expressing Hrp83-GFP in *clamp*^{WT}(green) and *clamp*^{delPrLD}/*clamp*$^{delPrLD+6Q}$ (orange) conditions.

Fig S1 **CLAMP binds to both DNA and RNA at specific gene locii A-C.** CLAMP-DNA binding (Cut & Run data) and CLAMP-RNA binding in chromatin fraction (iCLIP data) peaks visualized in the IGV genome browser at the *mrj, sf3b3* and *sqd*gene location. Arrow in **A** denotes CLAMP RNA peak completely overlapping with CLAMP DNA peak (CRO) and in **B** denotes CLAMP RNA peaks overlapping with 5' and 3' ends of DNA peaks (ORF and ORE). RNA peaks in **C** show CLAMP-RNA peaks proximal to DNA peaks (PXP)

Fig S2 ***rox2* RNA probes for CLAMP-*rox2* binding assays** *rox2* gene region showing putative CLAMP binding site (iCLIP data), stem loop region and *rox2* RNA probe used for CLAMP-*rox2* RNA-EMSA assay.

Table legends

Table 1a: CLAMP RNA targets identified in the nuclear fractions of male (S2) and female (Kc) cells.

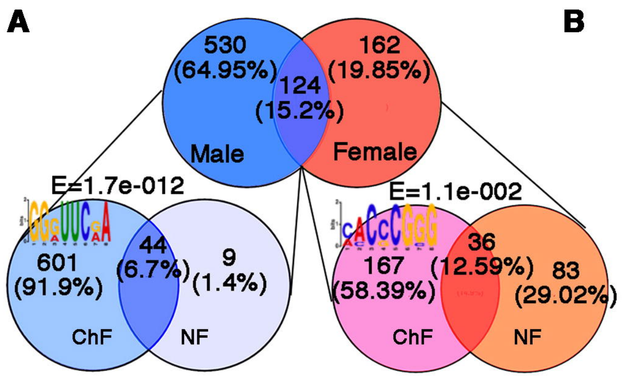
Table 1b: List ofgenes and genomic locations where CLAMP RNA peak completely overlapping with CLAMP DNA peak (CRO), CLAMP RNA peaks overlapping with 5' and 3' ends of DNA peaks (ORF and ORE) and CLAMP RNA peaks are proximal to DNA peaks (PXP) in male (S2) and female (Kc) cells.

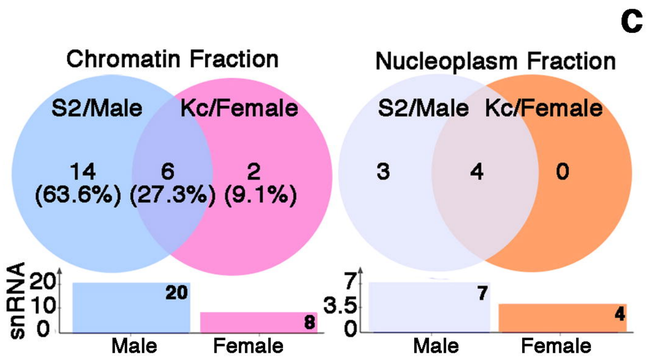
Table 2a: List of all CLAMP PrLD dependent differentially spliced genes in *Drosophila* female third instar larvae (L3).

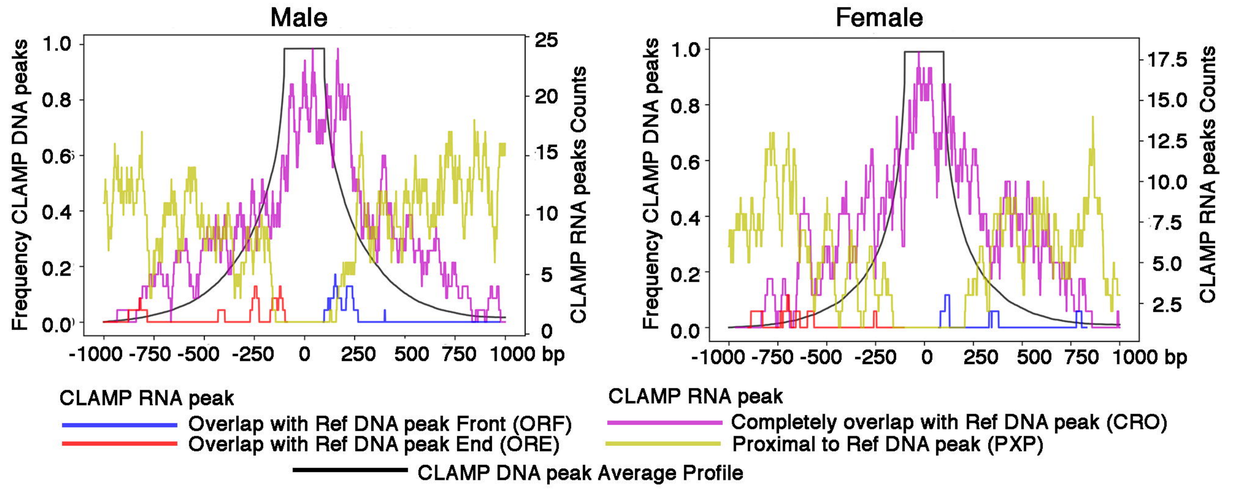
Table 2b: List of all CLAMP PrLD dependent differentially spliced genes in *Drosophila* male third instar larvae (L3).

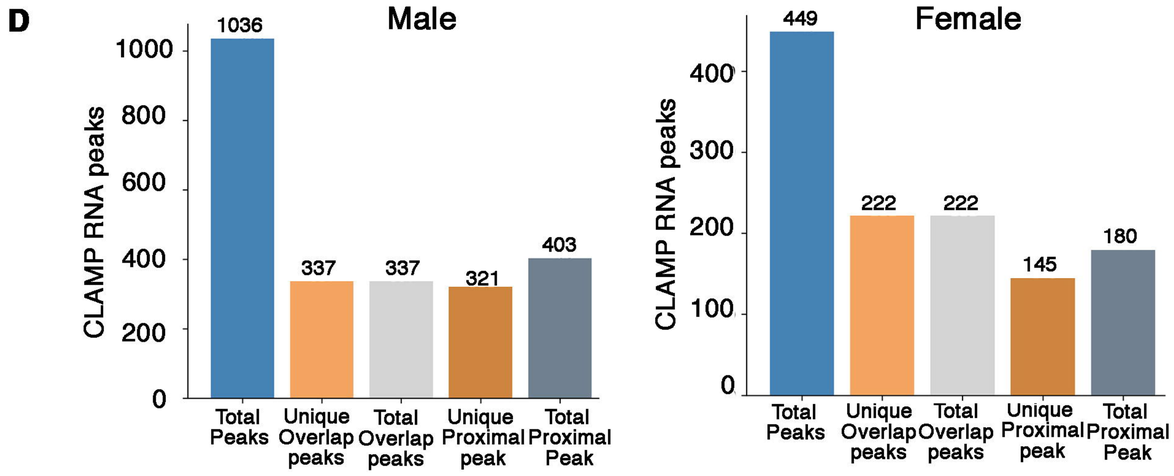
Table 3: List of CLAMP PrLD dependent male and female specifically spliced genes whose RNA isoforms are direct targets of CLAMP protein.

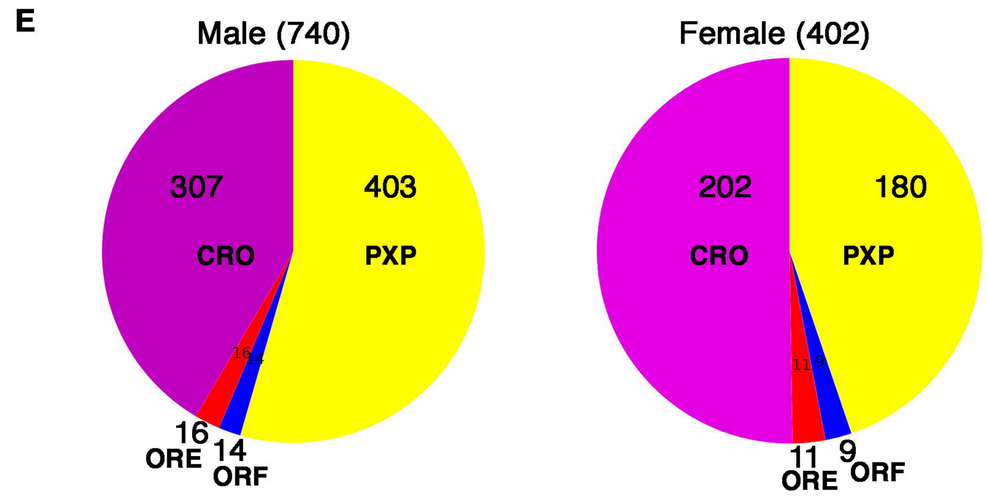
Figure 1.Ray, Zaborowsky et al 2024

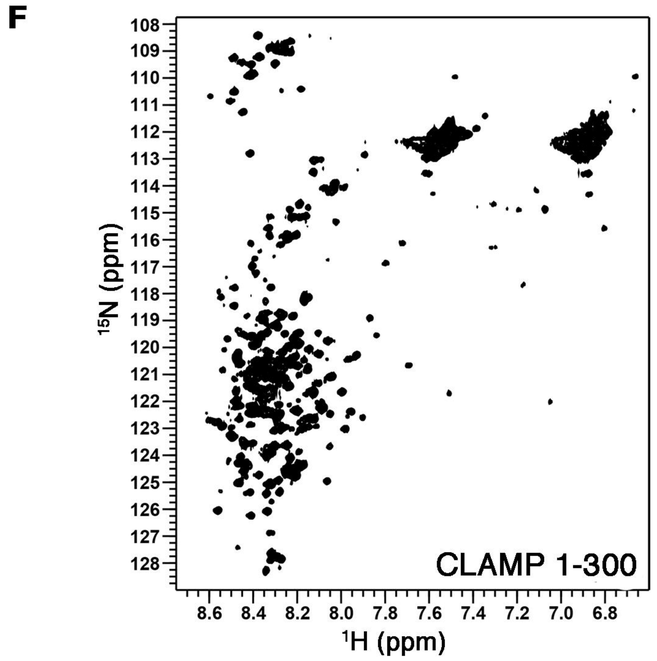


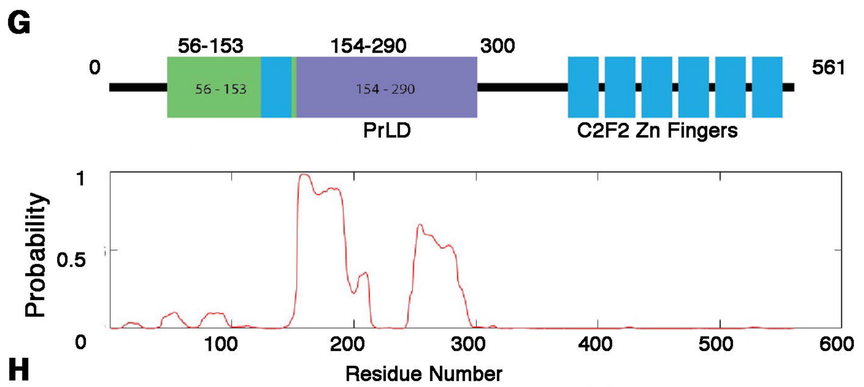


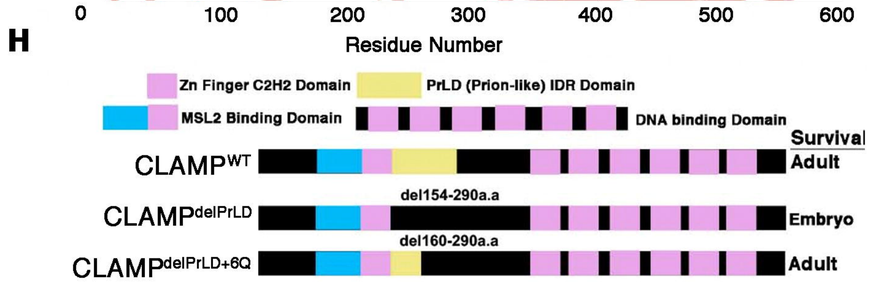


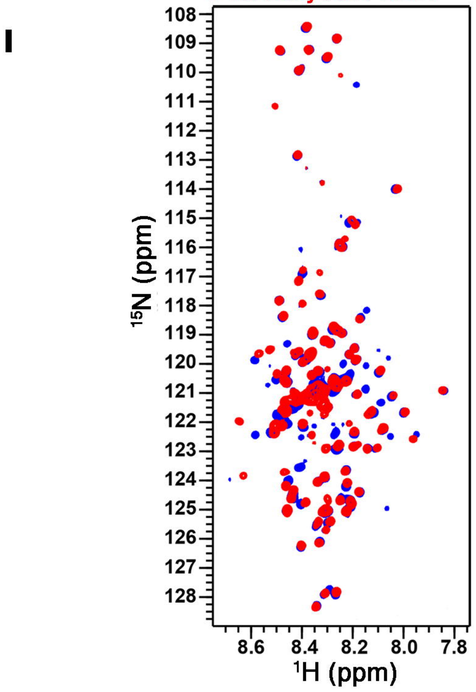


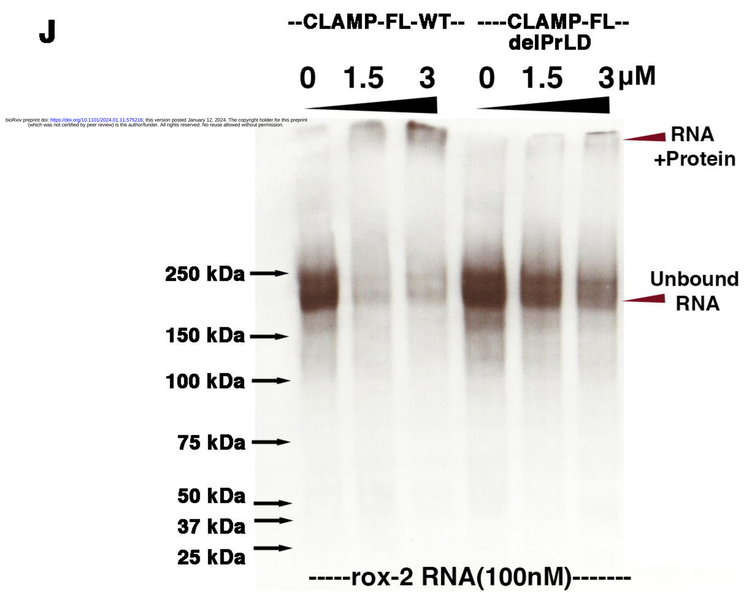


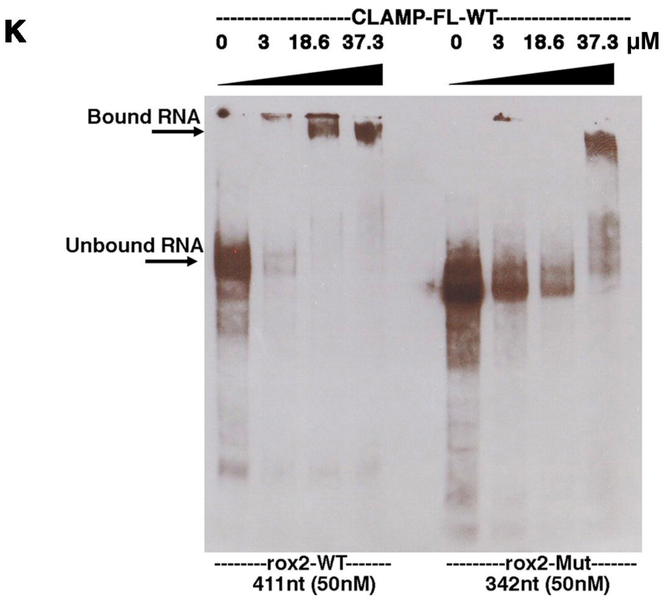


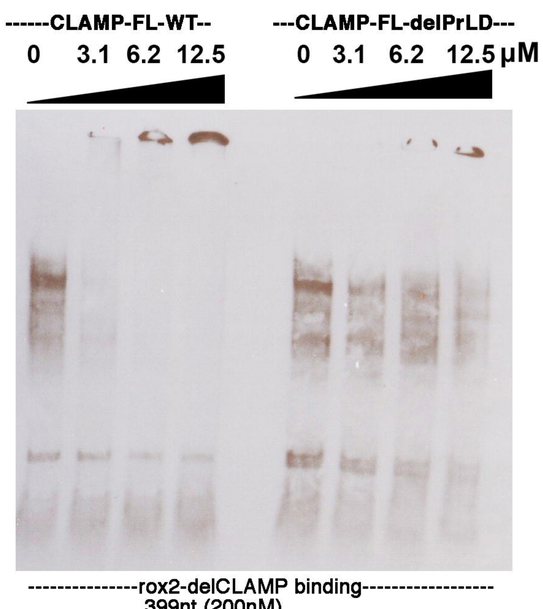












A
chr3R:28,598,407-28,607,777
9,372 bp
Gene
Noa36
Exon1 Exon2
Hrp38
CG9986
CLAMP DNA
Female (Kc)
Male (S2)
MSPC_Peak_1071
MSPC_Peak_1331
MSPC_Peak_1328
CLAMP RNA
peak_region
B
Exon1 Exon4
CLAMPWT
CLAMPdelPrLD
C
CLAMPWT-FL
12.5 6.2 3.1 0 uM
Protein + RNA
Unbound RNA
Hrp38 exon2 50nM
Hrp38 exon1 50nM
D
CLAMPdelPrLD-FL
CLAMPWT-FL
Bound RNA
unbound RNA
Hrp38 Exon1 RNA (50nM)
E
Ratio between spliced vs unspliced hrp38 isoform
p=0.0292
p=1.02012e-06
p=0.02119
p=0.0937
F
CLAMPWT
CLAMPdelPrLD
Female Male
R1 R2 R3
unspliced isoform
spliced intron isoform
GAPDH
G
CLAMP 1-300
Control CLAMP
CLAMP 1-300delPrLD
50 µm
H
Hrp38
CLAMP
MERGED
1-300a.a CLAMPWT +Hrp38+TEV
I
p=9.4585 E-05
p=7.76067 E-06
Average Percentage intensity of Hrp38GFP in immobile phase
CLAMPWT
CLAMPdelPrLD
Female n=77, N=13
Male n=86, N=18
Female n=51, N=9
Male n=90, N=14
J
Mean per individual + standard error
Mean per speckles + standard error
Fraction of Bound Hrp38 speckles
CLAMPWT (n=52900)
CLAMPdelPrLD (n=98048)
CLAMPWT (n=56092)
CLAMPdelPrLD (n=121629)
Male
Female
K
Hrp38GFP
clamp/clamp;hrp38GFP
clampdelPrLD/clampdelPrLD+6Q; hrp38GFP

Figure S1. Ray, Zaborowsky et al 2024

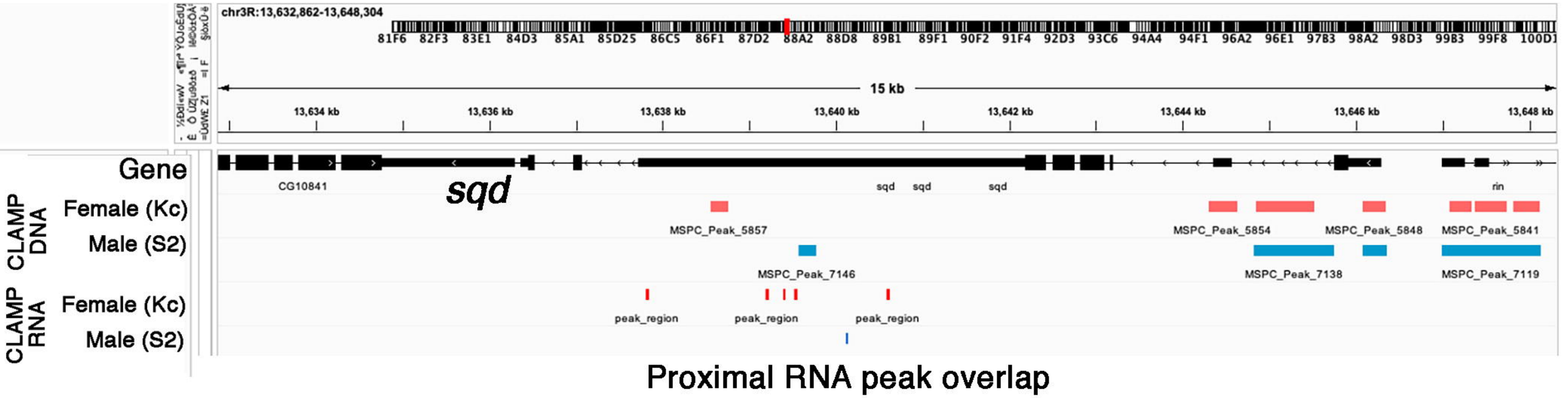
A

Complete RNA peak overlap

B

5’ and 3’ End RNA peak overlap

C



Proximal RNA peak overlap

Figure S2 Ray Zaborowsky et al 2024