A developmental mechanism to regulate alternative polyadenylation in an adult stem cell lineage

Lorenzo Gallicchio,[1] Neuza R. Matias,[1] Fabián Morales-Polanco,[2,3] Iliana Nava,[1] Sarah Stern,[1] Yi Zeng,[3] and Margaret T. Fuller[1,3]

[1]Department of Developmental Biology, Stanford University School of Medicine, Stanford, California 94035, USA; [2]Department of Biology, Stanford University School of Humanities and Sciences, Stanford, California 94035, USA; [3]Department of Genetics, Stanford University School of Medicine, Stanford, California 94035, USA

Alternative cleavage and polyadenylation (APA) often results in production of mRNA isoforms with either longer or shorter 3′ UTRs from the same genetic locus, potentially impacting mRNA translation, localization, and stability. Developmentally regulated APA can thus make major contributions to cell type-specific gene expression programs as cells differentiate. During *Drosophila* spermatogenesis, ∼500 genes undergo APA when proliferating spermatogonia differentiate into spermatocytes, producing transcripts with shortened 3′ UTRs, leading to profound stage-specific changes in the proteins expressed. The molecular mechanisms that specify usage of upstream polyadenylation sites in spermatocytes are thus key to understanding the changes in cell state. Here, we show that upregulation of PCF11 and Cbc, the two components of cleavage factor II (CFII), orchestrates APA during *Drosophila* spermatogenesis. Knockdown of *PCF11* or *cbc* in spermatocytes caused dysregulation of APA, with many transcripts normally cleaved at a proximal site in spermatocytes now cleaved at their distal site, as in spermatogonia. Forced overexpression of CFII components in spermatogonia switched cleavage of some transcripts to the proximal site normally used in spermatocytes. Our findings reveal a developmental mechanism where changes in expression of specific cleavage factors can direct cell type-specific APA at selected genes.

[*Keywords*: alternative polyadenylation; development; cellular differentiation; spermatogenesis; *Drosophila*; mRNA processing; cleavage factor complex II]

Supplemental material is available for this article.

Received March 18, 2024; revised version accepted July 22, 2024.

Alternative cleavage and polyadenylation (APA) of mRNAs results from changes in the site at which the 3′ end cut that terminates nascent transcripts is made (Tian and Manley 2017; Gallicchio et al. 2023). For alternative polyadenylation sites located in the region of a gene encoding the 3′ UTR, APA results in production of alternate mRNA isoforms that differ in 3′ UTR length and content, with potentially profound effects on transcript fate, including translation, stability, and localization (An et al. 2008; Mayr and Bartel 2009; Mansfield and Keene 2012; Berry et al. 2022). More than 70% of human genes are subject to alternative polyadenylation, making APA a widespread biological phenomenon (Gruber and Zavolan 2019). The most intriguing APA instances are those where the choice of 3′ end cut site is developmentally regulated, leading, for example, to production of transcripts with longer 3′ UTRs from specific genes in neurons or shorter 3′ UTRs in differentiating male germ cells compared with precursor cells (Hilgers et al. 2012; Miura et al. 2013; Shan et al. 2017; Vallejos Baier et al. 2017; Grassi et al. 2019; Berry et al. 2022). Changes in 3′ UTR length due to APA also occur in certain pathological states, such as the 3′ UTR shortening of transcripts observed in some cancers (Mayr and Bartel 2009; Morris et al. 2012; Venkat et al. 2020). Because these changes in 3′ UTR content can have dramatic effects on protein expression and thus cell state, cell type-specific regulation of APA is emerging as an important developmental mechanism to promote robust cell fate decisions.

During cleavage and polyadenylation, the last step of mRNA synthesis, the nascent mRNA molecule is cleaved and a string of adenosines is added by nuclear poly(A) polymerase to protect the newly made transcript from degradation and facilitate its export from the nucleus (Jing et al.

Corresponding author: mtfuller@stanford.edu
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.351649.124.

1999). Most mRNAs undergo cleavage and polyadenylation, with a notable exception being histone transcripts, which are not polyadenylated (Marzluff et al. 2008). Four well-conserved multiprotein complexes have been identified as components of the cleavage and polyadenylation machinery (Fig. 1A). The core complex CPSF (cleavage and polyadenylation specificity factor) is responsible for recognizing (via the CPSF30 and WDR33 subunits) the polyadenylation signal (PAS) and inducing cleavage of the nascent RNA molecule (via its CPSF73 subunit). The PAS is a hexamer (most commonly AAUAAA or a close variant) in the nascent mRNA, usually ~15–30 nt upstream of the cleavage site (Clerici et al. 2018; Gallicchio et al. 2023). Another complex, CstF (cleavage and stimulation factor), recognizes U/GU-rich regions downstream from the PAS (MacDonald et al. 1994). CFI (cleavage factor I) recognizes a UGUA motif upstream of the PAS (Yang et al. 2011), and CFII (cleavage factor II) preferentially binds to G-rich sequences via its PCF11 subunit (Schäfer et al. 2018). The cleavage and polyadenylation machinery is also associated with several additional proteins: Symplekin (a scaffolding protein), PABP2, and the nuclear poly(A) polymerase (Fig. 1A; Gallicchio et al. 2023). While the central role of the core complex CPSF in recognizing the PAS and cleaving the nascent mRNA molecule is relatively clear, the specific contributions of the other complexes and associated factors in specifying whether, where, on what specific gene products, and in which cell types a cut is made are still under investigation (Vallejos Baier et al. 2017; Zhu et al. 2018; Turner et al. 2020).

Developmentally regulated changes in the PAS site used can have profound effects on the suite of proteins expressed as cells progress from one state to the next in a developmental sequence. For example, during spermatogenesis in flies and mice, many genes express mRNA isoforms with long 3′ UTRs in proliferating spermatogonia but isoforms with shorter 3′ UTRs at later stages of germ cell development (Li et al. 2016; Berry et al. 2022; Lee et al. 2022). In *Drosophila melanogaster* testes (Fig. 1B), >500 genes express transcripts cleaved at a distal PAS in proliferating spermatogonia but at a more proximal PAS in differentiating spermatocytes, resulting in production of mRNA isoforms with much shorter 3′ UTRs (Fig. 1C).

The switch from distal to proximal cleavage site usage occurs at a time of profound change in cell state in the male *Drosophila* germline adult stem cell lineage. Spermatogenesis in *Drosophila* starts with an oriented stem cell division that causes one of the daughter cells to be displaced away from the somatic hub cells, stop self-renewing, and start a program of transient amplification by mitotic proliferation, founding a clone of interconnected spermatogonia enclosed in two somatic cyst cells (Fig. 1B). After four rounds of mitosis, the spermatogonia in the cyst start differentiation, beginning with premeiotic S phase, producing 16 interconnected spermatocytes. After ~3.5 days of meiotic prophase, during which the germ cells grow 25 times in volume, the spermatocytes undergo meiosis I and II, producing a cyst of 64 haploid spermatids. These then undertake a series of morphological changes that lead to production of mature and functional sperm (Fig. 1B). The switch from spermatogonia to spermatocytes is accompanied by changes in both cell state and gene expression. Almost all the transcripts needed for the changes in cellular morphology that occur after meiosis are expressed in spermatocytes (Raz et al. 2023). Changes in protein expression due to 3′ UTR shortening by APA may thus play key roles not only in setting up differences between proliferating spermatogonia versus differentiating spermatocytes but also in stage-specific translational repression or activation required for proper timing of the subsequent cellular morphogenesis.

As the 3′ UTR of mRNA molecules is a hub for *cis*-regulatory sequences that can attract *trans*-acting factors such as miRNAs and RNA binding proteins to regulate mRNA stability and translation efficiency, it is not surprising that the stage-specific 3′ UTR shortening due to APA at >500 genes observed in spermatocytes compared with spermatogonia (Fig. 1C) resulted in dramatic shifts in the proteins expressed between these two cell types. Polysome fractionation revealed that transcripts from at least 200 genes switched from the long 3′ UTR isoform comigrating with polysomes in spermatogonia to the short 3′ UTR isoform not associated with ribosomes in young spermatocytes. For another ~50 genes, the long 3′ UTR transcript isoforms were not translated in spermatogonia, while the short 3′ UTR isoforms comigrated with polysomes in spermatocytes (Berry et al. 2022). Thus, the developmental program(s) that specify whether nascent transcripts are cleaved at proximal 3′ UTR sites in spermatocytes rather than the distal sites used in spermatogonia can switch expression of many proteins from on in spermatogonia to off in spermatocytes or vice versa, from off in spermatogonia to on in spermatocytes.

Here we show that cell type-specific production of transcripts with short 3′ UTRs from specific genes due to APA in spermatocytes is developmentally regulated by changing the levels of one of the four complexes involved in 3′ end cleavage and polyadenylation. We found that the two protein components of CFII, encoded by the *PCF11* and *crowded by cid (cbc)* genes, are upregulated in spermatocytes compared with spermatogonia. When levels of *PCF11* or *cbc* were knocked down in spermatocytes by cell type-specific RNAi, >240 (~48%) of the ~500 genes for which nascent transcripts would normally be cleaved at a proximal PAS in spermatocytes instead resumed producing transcripts cleaved at the distal PAS used in spermatogonia. Reciprocally, forced overexpression of CFII components in spermatogonia shifted 3′ end cleavage to the more proximal site for some of the transcripts from genes that undergo stage-specific APA identified by Berry et al. (2022). Lowering protein levels in spermatocytes of components of the cleavage machinery-associated complexes CFI or CstF had much less effect on the choice of 3′ end cleavage site, suggesting that CFII plays a special role in regulating APA in spermatocytes compared with spermatogonia. The PCF11 and Cbc proteins interact structurally: Immunoprecipitation of Cbc from testis extracts brought with it a >250 kDa form of PCF11. Reduced

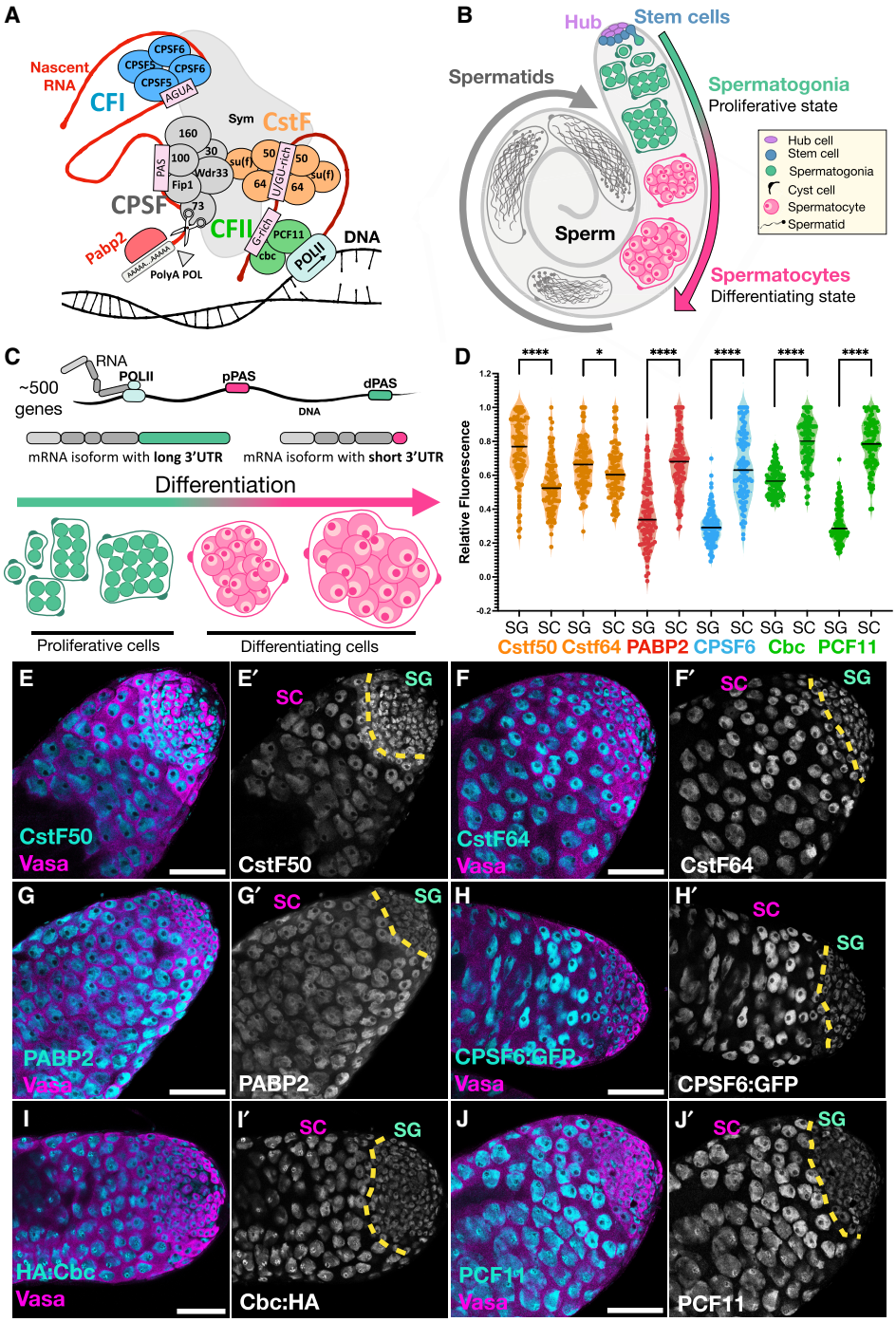

Figure 1. Expression of cleavage factors during *Drosophila* spermatogenesis. (*A*) Schematic representation of the cleavage machinery. Adapted from Gallicchio et al. (2023). The cleavage and polyadenylation factor complex (CPSF) is shown in gray. The cleavage and stimulation factor complex (CstF) is shown in orange. Cleavage factor I (CFI) is shown in blue. Cleavage factor II (CFII) is shown in green. (PAS) Polyadenylation signal. (*B*) Diagram of a wild-type *Drosophila* testis showing key stages of germ cell differentiation. (*C*) Diagram of developmentally regulated APA in the *Drosophila* testes resulting in expression of transcripts with long 3′ UTRs in spermatogonia but short 3′ UTRs in spermatocytes from ~500 genes. (*D*) Quantification of relative immunofluorescence signal per nuclear area from *E*–*J*. Each dot corresponds to a single nucleus. For each testis, 20 spermatogonial nuclei and 20 spermatocyte nuclei were quantified, and $N = 5$ testes were quantified per staining, for a total of 100 spermatogonia and 100 spermatocytes. Statistical significance was calculated using the Kolmogorov–Smirnov *t*-test. (****) *P*-value < 0.0001, (*) $P = 0.0366$. (*E*–*J*) Example immunofluorescence images of testis apical regions from samples quantified in *D*. (Magenta) Anti-Vasa marks germ cells, (cyan) cleavage factor proteins. (*E*–*G*,*J*) Apical region of wild-type (*w1118*) testes stained with anti-CstF50 (*E*), anti-CstF64 (*F*), anti-PABP2 (*G*), and anti-PCF11 (*J*). (*H*) Apical regions of testes from *CPSF6:GFP* transgenic flies stained with anti-GFP. (*I*) Apical region of testes from *HA:cbc/HA:cbc* flies stained with anti-HA. The dotted yellow line indicates the border between proliferating spermatogonia (SG) and differentiating spermatocytes (SC). Scale bars, 50 µm.

expression of PCF11 protein in spermatocytes due to RNAi resulted in concomitant reduction in Cbc, suggesting that Cbc protein was not stable without its binding partner in CFII. However, the reciprocal was not true. Plentiful PCF11 protein persisted in the nuclei of spermatocytes in which expression of *cbc* had been knocked down by RNAi. Strikingly, *PCF11* mRNA is broadly expressed in many cell types in adult flies, while *cbc* expression is low in most adult tissues but upregulated in spermatocytes. These observations suggest a model in which a developmentally regulated increase in expression of *cbc* and *PCF11* in spermatocytes leads to greater activity of CFII. This in turn facilitates recognition of an upstream PAS and consequent 3′ UTR shortening by APA of nascent transcripts from specific genes, indicating CFII as a key player in the developmental regulation of alternative cleavage and polyadenylation.

Results

Specific cleavage factor proteins are upregulated in spermatocytes

Immunostaining of whole-mount *Drosophila* testes revealed that some but not all cleavage factor proteins were upregulated in spermatocytes (Fig. 1D–J). Relative expression levels of six protein components of various subcomplexes of the cleavage factor machinery were measured by immunostaining using available antibodies against endogenous proteins (for CstF50, CstF64, PABP2, and PCF11), anti-HA (for HA:Cbc, an available *Drosophila* line tagged with HA at the endogenous *cbc* locus by CRISPR) (Wu et al. 2021), or anti-GFP (for an available *Drosophila* line carrying a Fosmid-based transgenic copy of CPSF6 tagged with sGFP [superfolder GFP]) (Sarov et al. 2016). In comparing nuclear signal in spermatogonia and spermatocytes, each measurement was divided by the area of the corresponding nucleus to account for the great difference in nuclear size between the two cell types. Immunofluorescence signal from antibodies recognizing two members of the CstF complex, CstF50 (Fig. 1E) and CstF64 (Fig. 1F), decreased slightly or remained almost the same in nuclei in spermatocytes compared with nuclei in spermatogonia (quantified in Fig. 1D). In contrast, immunofluorescence signal from antibodies recognizing PABP2 (Fig. 1G), the CFI component CPSF6 (Fig. 1H), and the two members of CFII, Cbc (Fig. 1I) and PCF11 (Fig. 1J), increased substantially in nuclei from spermatocytes compared with nuclei from spermatogonia (quantified in Fig. 1D).

Reduction of CFII component levels in spermatocytes shifted APA to more distal sites

To lower expression of cleavage factor components in spermatocytes, females carrying a *bamGal4* expression driver and *UAS > dicer2* were crossed with males carrying RNAi constructs under control of a *UAS* array with an *hsp70* promoter. The *bamGal4* flies express Gal4 under control of the *bag-of-marble (bam)* promoter, which fires in mid to late spermatogonial and early spermatocyte stages. The crosses were incubated for 3 days at 25°C, and then the progeny were shifted to 29°C to facilitate effective transcriptional activation by Gal4.

Immunofluorescence staining revealed that *bamGal4*-driven knockdown using the *VDRC-103710* RNAi line lowered levels of *PCF11* protein detected in spermatocyte nuclei to a level similar to that detected in spermatogonia (Fig. 2B). Testes from *UAS > PCF11-RNAi-VDRC-103710; bamGal4* males viewed in live squash preparations by phase contrast microscopy had plentiful spermatocytes but few elongating spermatid bundles (Fig. 2E). An RNAi line directed against a different region of the PCF11 transcript, *UAS* > PCF11-RNAi-*NIG-10228R-4*, also lowered the level of anti-PCF11 immunofluorescence signal in spermatocytes when driven using *bamGal4* (Fig. 2C). Testes from *UAS* > PCF11-RNAi-*NIG-10228R-4;bamGal4* males viewed by phase contrast microscopy had some elongating spermatid bundles (Fig. 2F). In both cases, PCF11 protein was still detected in early spermatogonia and in nuclei of the somatic cyst cells that accompany spermatocyte cysts, as expected because the *bamGal4* driver is not expressed in these cells (Fig. 2B,C, arrowheads).

Knockdown of *cbc*, the partner of PCF11 in CFII, also caused defects in spermatogenesis. Immunofluorescence staining of testes from males heterozygous for a *cbc* allele tagged at the endogenous locus with the HA epitope (*HA:cbc*) showed strong anti-HA signal in spermatocyte nuclei, compared with little or no signal detected in spermatogonia. However, for two RNAi lines directed against different regions of *cbc*, knockdown under control of *bamGal4* reduced the level of anti-HA signal in spermatocytes to below the level of detection, as in spermatogonia (Fig. 2G–I). In both cases, analysis of unfixed squash preparations by phase contrast microscopy showed that the knockdown testes had abundant spermatocytes but lacked elongating spermatid bundles (Fig. 2K,L).

In situ hybridization revealed that knockdown of *PCF11* or *cbc* in spermatocytes by RNAi affected APA at both the *Kpc2* (also named *isopeptidase-T-3*) and the *Red* genes (Fig. 3). Both *Kpc2* and *Red* had been found to express transcripts with a long 3′ UTR in testes filled with spermatogonia but a much shorter 3′ UTR in testes filled with mid-stage differentiating spermatocytes (Fig. 3A,E,J). The relative levels of transcripts with long versus short 3′ UTRs expressed in spermatogonia were compared by hybridization chain reaction fluorescence in situ hybridization (HCR-FISH), a technique that can be used to quantify relative RNA levels in a sample, as the amplified fluorescent signal has been shown to be proportional to the number of mRNA molecules in the quantified area (Choi et al. 2018; Trivedi et al. 2018; Schulte et al. 2024). Comparing signal in the same sample from two different sets of HCR-FISH probes (Supplemental Table S1), one recognizing the common coding sequence and one recognizing the 3′ UTR extension (Fig. 3A), revealed the proportion of the total transcripts from a given gene (recognized by the common probe set) that contained the 3′ UTR extension (recognized by the 3′ UTR extension probe). For both *Kpc2* and *Red*, knockdown of either

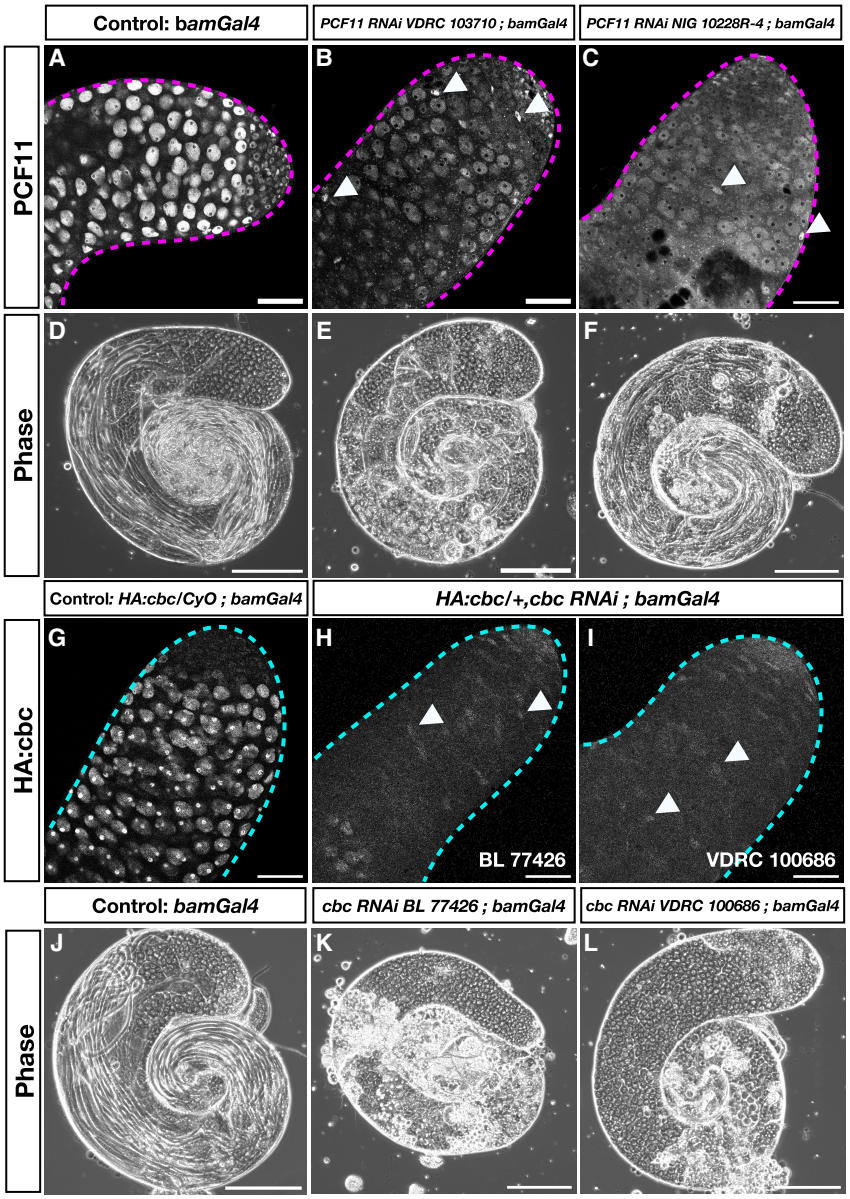

Figure 2. Knockdown of *PCF11* and *cbc* in spermatocytes. (*A–C*) Apical tips of testes stained with anti-PCF11. (*A*) Control *bamGal4*. (*B*) *PCF11-RNAi-VDRC-103710;bamGal4*. (*C*) *PCF11-RNAi NIG-10228R-4;bamGal4*. (*D–F*) Phase contrast microscope images of whole-mount live testes from *Drosophila* lines in *A–C*, respectively. (*G–I*) Apical tips of testes stained with anti-HA. (*G*) Control *HA:cbc/CyO;bamGal4*. (*H*) *HA:cbc/+,cbc-RNAi-BL-77426;bamGal4*. (*I*) *HA:cbc/+,cbc-RNAi-VDRC-100686;bamGal4*. (*J–L*) Phase contrast microscope images of whole-mount live testes. (*J*) Control *bamGal4*. (*K*) *cbc-RNAi-BL-77426;bamGal4*. (*L*) *cbc-RNAi-VDRC-100686;bamGal4*. Arrowheads indicate somatic cyst cell nuclei. Scale bars: *A–C*, *G–I*, 50 µm; *D–F*,*J–L*, 200 µm. Note that in all the immunofluorescence experiments on *HA:cbc/+* flies, the level of immunofluorescence signal with anti-HA in spermatogonia was below the detection limit for the experimental conditions. In contrast, for the testes quantified in Figure 1I, which were from homozygous *HA:cbc/HA:cbc* flies, the level of HA:Cbc in nuclei of spermatogonia was high enough to be detected (for comparison between homozygous and heterozygous testes, see Supplemental Fig. S1).

PCF11 or *cbc* in spermatocytes under control of *bamGal4* substantially increased the level of signal from the 3′ UTR extension detected in spermatocytes (Fig. 3C,D,H,I) compared with testes from control flies carrying the *bamGal4* driver but not the RNAi transgene raised under the same temperature regimen (Fig. 3B,G). Quantification of fluorescent signal confirmed that for both *Kpc2* and *Red*, the ratio of fluorescence from probes recognizing the long 3′ UTR versus probes against the coding sequence (CDS) was significantly higher in the *PCF11* or *cbc* knockdown spermatocytes than in spermatocytes from control testes (Fig. 3F,K).

Analysis of cleavage site usage by 3′ seq from testes in which either *PCF11* or *cbc* had been knocked down in spermatocytes by RNAi under control of *bamGal4* showed that the increase in signal from the 3′ UTR extension observed in knockdown spermatocytes was due to a shift from cleavage at the proximal site normally used in spermatocytes to the more distal cut site normally used in spermatogonia for both *Kpc2* and *Red* (Fig. 4A,B). Knockdown of *cbc* had the strongest effect. Knockdown of *PCF11* shifted 3′ end cleavage to the same distal site as *cbc* knockdown but to a lesser degree, especially with the *UAS>PCF11-RNAi-NIG-10228R-4* construct (Fig. 4A,B), consistent with the less efficient reduction in PCF11 protein levels observed with that particular line (Fig. 2C).

3′ seq data revealed that loss of function of *cbc* or *PCF11* affected APA at many loci. For roughly half of the ~500 genes previously identified by Berry et al. (2022) as undergoing 3′ UTR shortening by APA in spermatocytes compared with spermatogonia, knockdown of either *PCF11* or *cbc* in spermatocytes resulted in increased usage of the distal cleavage site normally used in spermatogonia

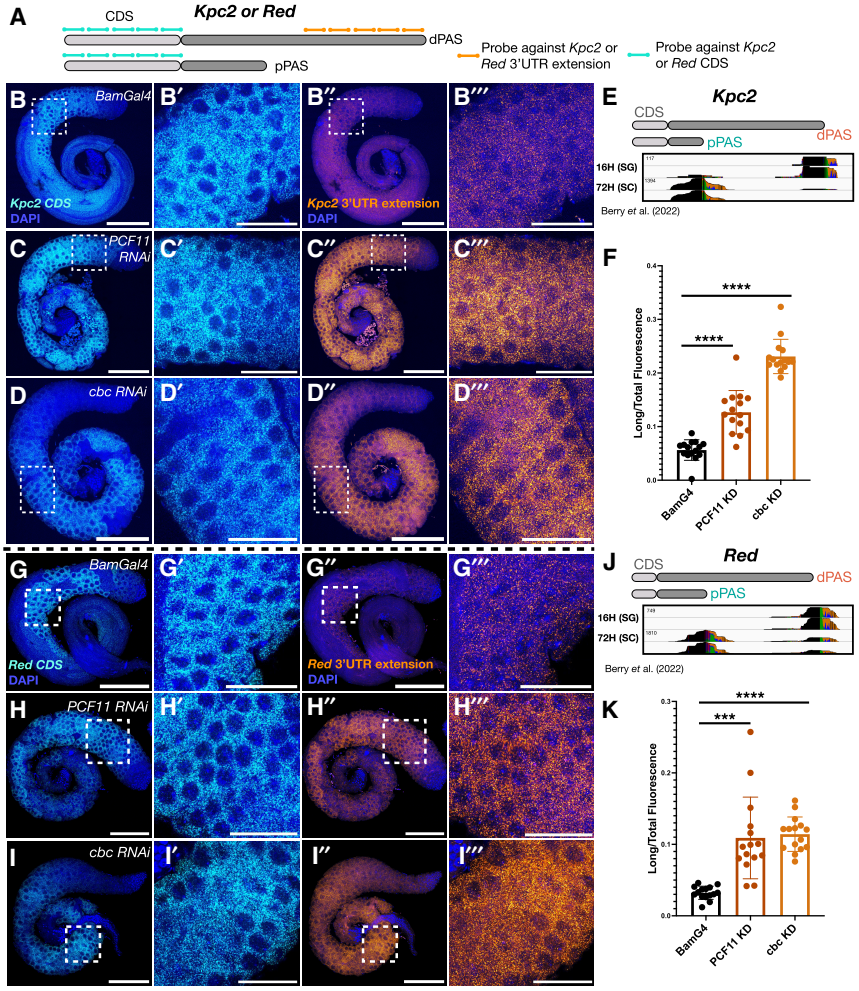

Figure 3. Knockdown of either *PCF11* or *cbc* reduces cleavage at the proximal PAS in spermatocytes. (*A*) Diagram illustrating the design of the two probe sets. (Light blue) Probe against the protein-coding sequence (CDS), (orange) probes against the 3′ UTR extension. (*B–D*′) Confocal microscope images of *bamGal4* (control) (*B*,*B*″), *PCF11-RNAi-VDRC-103710; bamGal4* (*C*,*C*″), and *cbc-RNAi-BL-77426; bamGal4* (*D*,*D*″) testes hybridized with probes recognizing the CDS of the *Kpc2* transcripts (cyan) and DAPI (blue) (*B–D*), and with probes recognizing the 3′ UTR extension of the *Kpc2* transcripts (orange) and DAPI (blue) (*B*″,*C*″, *D*″). (*B*′–*D*′,*B*‴–*D*‴) Magnified images of spermatocytes of corresponding regions from *B–D* and *B*″–*D*″, respectively. (*E*) Integrative Genomics Viewer (IGV) plots of 3′ seq pileup of reads for the gene *Kpc2* from a sample enriched for spermatogonia (SG; 16H PHS [post heat shock]) (Kim et al. 2017; Berry et al. 2022) and a sample enriched for spermatocytes (SC; 72H PHS) to show the APA switch from the distal PAS (dPAS) used in spermatogonia and the proximal PAS (pPAS) used in spermatocytes. (*F*) Quantification of fluorescence ratio from probes recognizing the *Kpc2* 3′ UTR extension (orange) compared with probes recognizing the *Kpc2* CDS (cyan) in spermatocytes. For each genotype, $N = 3$ testes were quantified. For each image, five measurements were taken in different spermatocyte areas, which are plotted on the graph. (*G–K*) Analogous to *A–F* for the *Red* gene. Scale bars: *B*, *B*″,*C*,*C*″,*D*,*D*″,*G*,*G*″,*H*,*H*″,*I*,*I*′, 150 µm; *B*′, *B*‴,*C*′,*C*‴,*D*′,*D*‴,*G*′,*G*‴,*H*′,*H*‴,*I*′,*I*‴, 50 µm. (****) $P < 0.0001$, (***) $P = 0.0001$. For statistical analysis, we used the Welch's *t*-test.

(Fig. 4C). When 3′ seq data were analyzed via PolyA-miner (Yalamanchili et al. 2020), providing as input the list of pairs of proximal PAS (pPAS) and distal PAS (dPAS) from the >500 genes undergoing APA identified by Berry et al. (2022), many genes showed increased usage of the distal PAS in knockdown testes compared with control testes. Orange in Figure 4C indicates that the frequency of nascent transcripts from a given gene being cleaved at the distal PAS in knockdown testis samples was higher than in control testes [positive poly(A) index], while blue indicates that the frequency of the nascent transcripts being cleaved at the proximal PAS in knockdown testes was higher than in control testes [negative poly(A) index]. Noticeably, a negative poly(A) index does not indicate a shortening event but indicates a gene for which the proximal PAS was overrepresented in knockdown compared with control samples. This could be caused either by different expression of the gene compared with in control testes or by the different cell composition in wild-type versus knockdown testes, as generally knockdown testes do not have as many elongating spermatids as control testes. Again, knockdown of *cbc* in spermatocytes had the strongest effect on the APA switch, affecting APA at the highest number of genes. Importantly, >90% of the genes where APA was affected by knockdown of *PCF11* [poly(A) index >0.5, indicating more cleavage at the distal PAS] also showed a switch in APA to more cleavage at the distal site with knockdown of *cbc*, consistent with these two CFII components acting together in the choice of where to make the 3′ end cut for transcripts from specific genes (Fig. 4F). Analysis of RNA-seq libraries from testis samples in which *cbc* or *PCF11* was knocked down under *bamGal4* revealed that in general the genes that had a poly(A) index either >0.5 or <−0.5 were not expressed at dramatically different levels in knockdown versus control testes (Supplemental Fig. S3).

3′ seq revealed that lowering the levels of the nuclear PABP2 protein in spermatocytes by RNAi also frequently affected 3′ UTR shortening by APA (Fig. 4A–C). Immunofluorescence staining showed effective knockdown of PABP2 protein to below detection in spermatocytes from flies carrying *UAS>PABP2-RNAi-VDRC-106466;bamGal4*, although PABP2 protein was still abundant in nuclei of somatic cyst cells, as expected (Supplemental Fig. S4B). For two different RNAi lines, testes from *UAS>PABP2-RNAi;bamGal4* flies had plentiful spermatocytes but

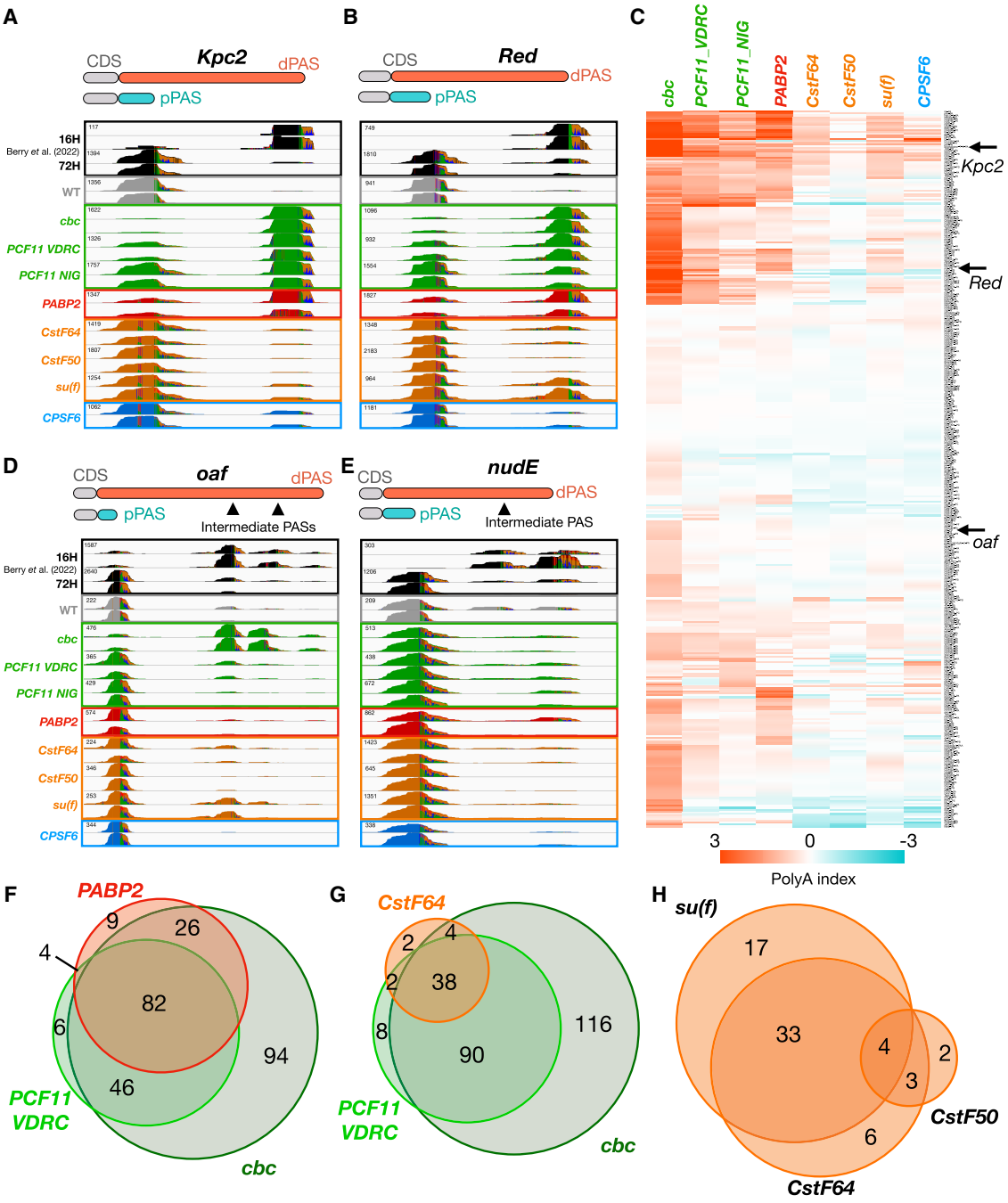

Figure 4. 3′ end RNA-seq data reveal the effects on APA of knockdowns in spermatocytes of different components of the cleavage machinery. (*A*,*B*) Integrative Genomics Viewer (IGV) plots of 3′ end RNA sequencing reads for regions encoding the 3′ UTR for the *Kpc2* (*A*) and *Red* (*B*) genes. Regions of peaks that change in color show mismatches between the poly(A) tail and the genome to highlight the 3′ end of the mRNA. (*C*) Heat map of poly(A) index quantified with PolyA-miner across 3′ seq libraries. Orange indicates increased cleavage at the distal PAS in knockdowns compared with control, and blue indicates increased cleavage at the proximal PAS in knockdowns compared with control. Each row represents one of the >500 genes that switch to a proximal PAS in wild-type spermatocytes compared with spermatogonia from Berry et al. (2022). Each column represents a component of the cleavage machinery that was knocked down in spermatocytes. The interval chosen to plot the poly(A) index (−3,3) does not represent the maximum and minimum poly(A) index observed. In this heat map, genes that showed a statistically nonsignificant poly(A) index in all the samples or a negative poly(A) index in all the samples were removed, and in all other cases, poly(A) indexes that were nonsignificant were put equal to 0 (another heat map with these genes shown is reported in Supplemental Fig. S2). (*D*,*E*) IGV plots of 3′ end RNA sequencing libraries for the genes *oaf* (*D*) and *nudE* (*E*). (*F*–*H*) Venn diagrams representing the number of genes with poly(A) index >0.5 in *cbc* knockdown libraries (dark green), *PCF11* VDRC knockdown libraries (light green), *PABP2* knockdown libraries (red), and components of CstF complex knockdowns (orange).

lacked elongating spermatid bundles (Supplemental Fig. S4D,E). 3′ seq revealed that in testes from *UAS > PABP2-RNAi-VDRC-106466;bamGal4*, a substantial percentage of transcripts from the *Kpc2* and *Red* loci switched 3′ end cleavage from the proximal site normally used in spermatocytes to the more distal site normally used in spermatogonia (Fig. 4A,B). Analysis of the 3′ seq data showed that for the >500 genes identified previously as undergoing APA in spermatocytes compared with spermatogonia, ~22% showed increased usage of the distal PAS when PABP2 was knocked down in spermatocytes (Fig. 4C). Furthermore, most (~92%) of the genes that switched toward the more distal PAS [poly(A) index >0.5] after knocking down *PABP2* also showed a shift [poly(A) index >0.5] toward the distal PAS when either *PCF11* or *cbc* was knocked down in spermatocytes (Fig. 4F).

In contrast, knockdown of any of the three components of the CstF complex [*CstF64*, *CstF50*, and *su(f)*] or a component of CFI (*CPSF6*) in spermatocytes had much less effect on the APA profile of mRNAs from the ~500 genes previously shown to undergo 3′ end shortening due to APA in spermatocytes, as assessed by 3′ seq (Fig. 4C). Immunofluorescence staining with antibodies against CstF64 or CstF50 showed that levels of the proteins were substantially reduced in spermatocytes from testes where expression of the respective genes had been knocked down by RNAi under control of *bamGal4* (Supplemental Fig. S5B, E). In both cases, signal remained in the nuclei of early spermatogonia and cyst cells, as expected from the cell type specificity of the *bamGal4* expression driver. No antibody reagents against *su(f)* were available to assess effectiveness of the RNAi construct directed against *su(f)*. However, the defects in spermatogenesis in *UAS > su(f)-RNAi;bamGal4* testes observed by phase contrast microscopy (Supplemental Fig. S5G) and the effects of knocking down *su(f)* on protein levels of CstF50 in spermatocyte nuclei (Supplemental Fig. S6) suggested that the knockdown of *su(f)* was effective. Levels of nuclear CstF50 protein drastically decreased in spermatocytes from males carrying an RNAi construct directed against *su(f)* expressed under control of *bamGal4* (Supplemental Fig. S6D). Knockdown of *CstF64* under control of *bamGal4* also lowered the levels of CstF50 in spermatocyte nuclei (Supplemental Fig. S6F). Notably, when either *su(f)* or *CstF64* was knocked down in spermatocytes, nuclear CstF50 staining decreased, although some signal remained in the cytoplasm. Therefore, the protein expression data suggest that components of the CstF complex may depend on each other for protein stability and/or nuclear localization, as previously reported in HeLa cells (Grozdanov et al. 2018). Consistent with the knockdowns affecting function, squash preparations of testes from *UAS > su(f)-RNAI-BL-65693;bamGal4* showed abundant spermatocytes but no elongating spermatid stages (Supplemental Fig. S5G). Testes from *UAS > CstF64-RNAi-BL65987/bamGal4* males also showed defective spermatogenesis, with plentiful spermatocytes but extremely aberrant elongation stages (Supplemental Fig. S5F). Surprisingly, even though CstF50 protein was clearly strongly reduced in spermatocytes from *UAS > CstF50-RNAi-BL77377;bamGal4* males, spermatogenesis appeared largely normal by phase contrast microscopy (Supplemental Fig. S5C).

Knockdown of either *CstF64* or *su(f)* in spermatocytes shifted the site of 3′ end formation toward the more distal site used in spermatogonia at some of the >500 genes previously identified as subject to 3′ UTR shortening in spermatocytes due to developmentally regulated APA (Fig. 4C). These genes tended to be among the same set of genes where APA was affected by knockdown of components of CFII: Approximately 95% of the genes where APA was affected by knockdown of *CstF64* were also affected by knockdown of either *PCF11*, *cbc*, or both (Fig. 4G). However, the effects of loss of function of CstF components on APA in spermatocytes was small compared with the effect of loss of CFII components: Only 44 of the 258 genes where APA was strongly affected by knockdown of either *cbc* or *PCF11* in spermatocytes also showed a switch toward the more distal cut site in testes where function of *CstF64* had been knocked down in spermatocytes (Fig. 4G). More than 85% of the genes affected [poly(A) index >0.5] by *CstF64* knockdown were also affected either by *CstF50* or *su(f)* knockdown, and 68% of the genes affected by *su(f)* knockdown were also affected by either *CstF64* or *CstF50* knockdown (Fig. 4H), consistent with these three components acting together. For *CPSF6*, a component of the CFI complex, which binds nascent transcripts upstream of the PAS, knockdown under the control of *bamGal4* caused extensive defects in spermatid differentiation for one RNAi line (VDRC 107147) and variable or weaker defects for the other RNAi line tested (BL 34804, used for 3′ seq), suggesting partial loss of function (Supplemental Fig. S7). Analysis of cut sites used by 3′ end seq indicated that for the >500 genes previously identified as undergoing APA in spermatocytes, the effect of *CPSF6* knockdown on APA using the BL 34804 RNAi line was weak compared with knockdowns of *cbc* or *PCF11* (Fig. 4C). However, the weak effect could be due in part to a weaker knockdown.

De novo genome-wide calling of PAS sites located on annotated 3′ UTRs revealed that knockdown of components of the cleavage machinery affected APA at additional loci (Supplemental Figs. S8, S9) in addition to the >500 genes from Berry et al. (2022). For this analysis, we only considered genes in which the most upregulated PAS and the most downregulated PAS in a knockdown compared with a wild-type sample were located in the annotated gene 3′ UTR. Again, the strongest effects were detected in *cbc* knockdown testes, followed by *PCF11* knockdown and *PABP2* knockdown testes (Supplemental Fig. S8), while *CstF64*, *CstF50*, *su(f)*, and *CPSF6* knockdowns had much milder effects on the global APA profile (Supplemental Fig. S9).

Of the components of the cleavage machinery tested, the 3′ seq data indicated that the 3′ UTR shortening of transcripts from selected genes due to APA in spermatocytes was most strongly dependent on levels of CFII (composed of Cbc and PCF11) or *PABP2.* For almost half of the >500 genes previously identified as undergoing the developmentally regulated APA, lowering function of one or the other of these proteins in spermatocytes resulted in

a shift toward usage of the more distal 3′ end cut site typical of spermatogonia. Several genes seemed to be more affected by knockdown of *cbc* than by knockdown of *PCF11*; for example, *oaf* (Fig. 4D). Notably, over half of the genes previously identified as subject to 3′ UTR shortening in spermatocytes compared with spermatogonia due to APA were not substantially affected by knockdown of any of the cleavage factor machinery that we tested; for example, see *nudE* (Fig. 4E; Supplemental Fig. S2).

Transcription of the CFII components PCF11 *and* cbc *is strongly upregulated in spermatocytes*

According to FlyBase, the *PCF11* locus can give rise to seven different RNA isoforms (Fig. 5A), grouped in three classes for simplicity: a short isoform (*PCF11-RC*) of 819 nt predicted to encode a small protein (~10 kDa), a medium isoform (*PCF11-RB*) of 2647 nt encoding an ~63 kDa predicted protein, and a group of ~6000 nt transcripts (*PCF11-RD-RE-RF-RG-RH*) predicted to encode a family of >200 kDa proteins that have subtle differences in a central region. All the transcript isoforms share the same 5′ UTR but differ in their 3′ UTRs (as well as in exon combinations). PCF11 has a CID (Pol II C-terminal binding domain) present in both the PCF11-long and the 63 kDa PCF11-RB isoforms but only partially present in the short protein encoded by *PCF11-RC*. PCF11 homologs in mammals and yeast have been shown to bind Cbc homologs through a region close to the C terminus of PCF11 (Sadowski et al. 2003; Schäfer et al. 2018). Notably, while the *PCF11-RNAi-NIG10228-R4* line targets a region common to all *PCF11* mRNA isoforms, *PCF11-RNAi-VDRC107310* targets a region present only in the *PCF11-long* isoforms. The *cbc* locus is simpler, with only one predicted transcript 1654 nt long, encoding a 46.8 kDa predicted protein (Fig. 5A).

The two *Drosophila* components of CFII show substantially different patterns of transcript expression in adult flies. In the Fly Cell Atlas single-nucleus RNA-seq analysis of tissues from adult flies (Li et al. 2022), transcription of *PCF11* was detected at generally high levels in many cell types (Fig. 5B). In contrast, *cbc* expression was detected at substantial levels in only a small subset of nuclei from adult tissues, with the clusters showing the highest

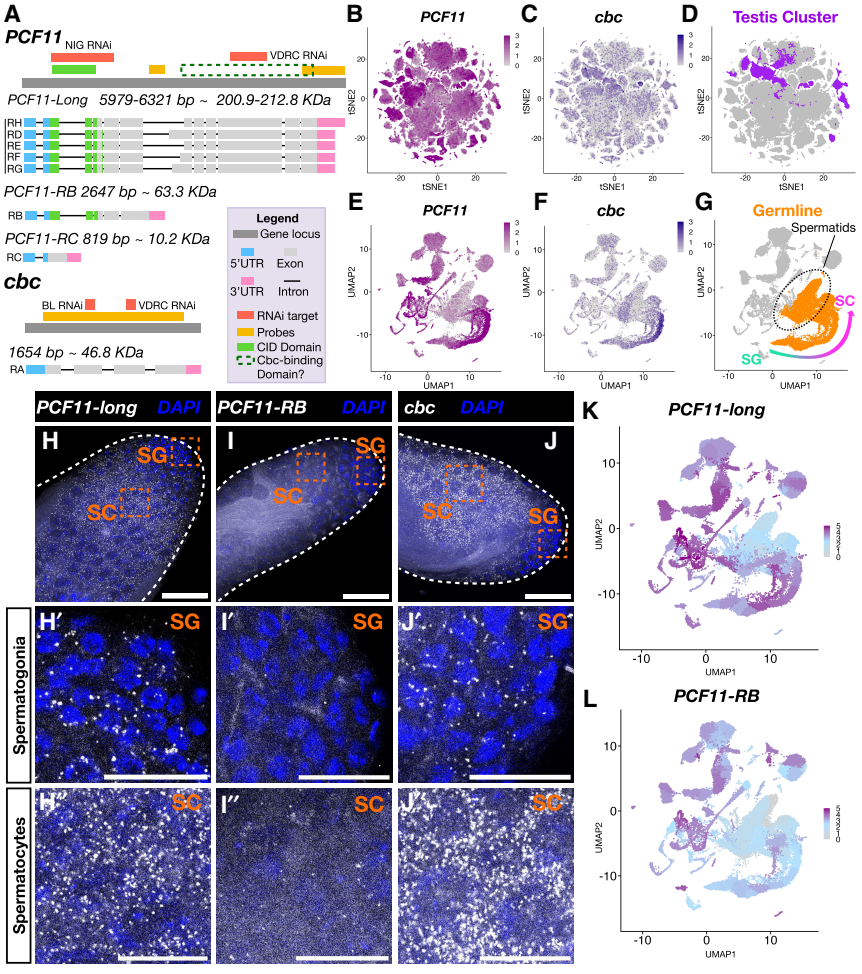

Figure 5. Both *PCF11* and *cbc* transcripts are upregulated in spermatocytes, with *cbc* particularly highly expressed in testes compared with other tissues. (*A*) Diagram adapted from FlyBase of *PCF11* and *cbc* loci. Light-blue boxes represent 5′ UTRs, green and gray boxes represent exons, and pink boxes represent 3′ UTRs. Introns are represented with a straight black line. The transcript regions targeted by HCR-FISH probes for RNA in situs are shown in yellow. The areas of transcripts targeted by RNAi are shown in red. Green and dotted dark-green boxes *above* the *PCF11* locus represent the CID and a region in which the Cbc binding domain could be located, respectively. (*B–G*) Scope diagrams of Fly Cell Atlas single-nucleus RNA-seq data (Li et al. 2022) showing *PCF11* (*B*) and *cbc* (*C*) expression across all of the analyzed adult tissues. (*D*) Purple indicates nuclei from the testis plus seminal vesicle sample across all analyzed adult tissues. (*E–G*) UMAP plots showing *PCF11* (*E*) and *cbc* (*F*) expression in the testis plus seminal vesicle sample; orange in *G* indicates nuclei from the male germline clusters. (*H–J*) Confocal microscope images of apical tips of wild-type (*w1118*) testes hybridized with probes against *PCF11* long isoforms (white) and DAPI (blue) (*H*), probes against *PCF11*-RB isoform (white) and DAPI (blue) (*I*), and probes against *cbc* (white) and DAPI (blue) (*J*). In *H–J*, the boxes highlight a region of spermatogonia (SG) magnified in *H′*, *I′*, and *J′*, and a region of spermatocytes (SC) magnified in *H″*, *I″*, and *J″*. (*K*,*L*) UMAP from Fly Cell Atlas single-nucleus data from testis plus seminal vesicle sample showing *PCF11-long* (*K*) and *PCF11-RB* (*L*) expression across the different clusters. Scale bars: *H–J*, 50 µm; *H′*,*H″*,*I′*,*I″*,*J′*,*J″*, 20 µm.

cbc transcription belonging to the testis plus seminal vesicle sample (Fig. 5C,D). Plotting expression on the UMAP of snRNA-seq data from testis plus seminal vesicle revealed that *PCF11* is transcribed in many somatic cell types in the testes, including terminal epithelia, seminal vesicle, nuclei from testis muscle cells, and prominently in the somatic cyst cells that enclose germline cysts. *PCF11* transcripts were also expressed in germ cells, with the level increasing substantially in spermatocytes compared with spermatogonia (Fig. 5E). Transcription of *cbc* also increased substantially in spermatocytes compared with spermatogonia (Fig. 5F).

Analysis of RNA expression in wild-type testes by HCR-FISH revealed that the long transcripts from *PCF11* were upregulated in spermatocytes compared with spermatogonia (Fig. 5H,H′,H″). In contrast, the intermediate-sized transcript isoform *PCF11-RB*, detected with a set of probes recognizing the 3′ UTR of *PCF11-RB*, which derives from an intronic region of the *PCF11-long* isoforms and so should not recognize fully processed *PCF11-long* isoforms in the cytoplasm, was extremely low in both spermatogonia and spermatocytes (Fig. 5I–I″). Single-nucleus RNA-seq from the Fly Cell Atlas confirmed that while the collective *PCF11-long* isoforms are highly expressed in spermatocytes, cyst cells, and somatic cell nuclei in the testis (Fig. 5K), the *PCF11-RB* transcript was only lowly expressed in the germline, with higher but still modest expression in cyst cell and other somatic cell nuclei in the testes (Fig. 5L). Similar HCR-FISH using a set of probes against the *cbc* protein-coding region confirmed that expression of *cbc* mRNA was strongly upregulated in spermatocytes compared with spermatogonia (Fig. 5J–J″).

PCF11 and Cbc coimmunoprecipitate, and knockdown of PCF11 *causes a decrease of Cbc protein in spermatocyte nuclei*

Analysis of proteins expressed in wild-type testes by polyacrylamide gel electrophoresis (PAGE) followed by Western blot probed with antibodies against PCF11 showed a predominant high-molecular-weight (>250 kDa) band (Fig. 6A), consistent with the expression of *PCF11-long* mRNA isoforms detected by HCR-FISH and snRNA-seq (Fig. 5H,K). Anti-PCF11 also detected an ~70 kDa protein isoform in testis extracts, as predicted from the *PCF11-RB* mRNA isoform. Although the antibody used was raised against amino acids 1–283 of *Drosophila* PCF11 (Zhang and Gilmour 2006), no protein product corresponding to PCF11-RC was observed despite a seemingly high level of the *PCF11-RC* transcript isoform detected by snRNA-seq of *Drosophila* testis tissue (Supplemental Fig. S10). Western blots of testes from flies homozygous for the *HA:cbc* allele probed with anti-HA showed expression of the expected ~50 kDa isoform (Fig. 6B).

Immunoprecipitation with anti-HA antibodies from extracts of testes from flies homozygous for *HA:cbc* coimmunoprecipitated the >250 kDa band from the PCF11 isoforms, consistent with the Cbc and PCF11 proteins interacting in a complex (Fig. 6C). The 70 kDa isoform of PCF11 was not detected in the immunoprecipitate (entire gels of Western blots and co-IP experiments are reported in Supplemental Fig. S11). Consistent with physical interaction between the Cbc and PCF11 proteins, accumulation of Cbc protein in spermatocytes was at least partially sensitive to the levels of PCF11 protein. Analysis by RNA-seq showed that when *PCF11* expression was knocked down in late spermatogonia and spermatocytes by RNAi under control of *bamGal4*, *cbc* transcript was still expressed in *PCF11* knockdown testes at a similar level compared with control testes (Fig. 6D). Likewise, in testes in which *cbc* had been knocked down by RNAi under control of *bamGal4*, transcripts from the *PCF11* locus were still expressed (Fig. 6E). However, the situation was different at the level of protein accumulation in spermatocytes. Simultaneous immunostaining of testis from *HA:cbc/+;bamGal4* males with anti-PCF11 and anti-HA antibodies showed both proteins present in spermatocyte nuclei from control males (Fig. 6F′,F″). However, for two different RNAi lines directed against *PCF11*, knockdown of *PCF11* under control of *bamGal4* resulted in a drastic reduction of signal from HA:Cbc (Fig. 6G′,H′, quantified in K). Similar reduction of Cbc signal in *UAS>PCF11_VDRC-RNAi;bamGal4* and *UAS>PCF11_NIG-RNAi;bamGal4* spermatocytes was observed in flies homozygous for the *HA:cbc*-tagged allele (Supplemental Fig. S12). In contrast, knockdown of *cbc* by RNAi under control of *bamGal4* did not cause loss of signal from anti-PCF11 (Fig. 6I″,J″, quantified in L).

Forcing high expression of cbc *and/or* PCF11 *in spermatogonia induced cleavage at proximal PASs normally used in spermatocytes*

Our data from the knockdown experiments in spermatocytes suggest that the high levels of CFII components expressed in spermatocytes are required for cleavage of nascent transcripts of certain genes at an upstream PAS rather than at the distal PAS used in spermatogonia. Forced overexpression of CFII components in spermatogonia triggered a fraction of the nascent transcripts from certain genes to be cleaved at the upstream site normally used in spermatocytes.

Minigenes with the coding sequences of *cbc* with a V5 tag or *PCF11-RF* with an HA-FLAG tag, both under the control of two repeats of 5XUAS and the *hsp70* promoter, were introduced into the *Drosophila* genome. These transgenes allow spermatogonial-specific overexpression of *cbc-V5* and/or *PCF11-RF-HA* by crossing males carrying the designed UAS constructs with females carrying a combination of *nanosGal4* (to drive in stem cells and early spermatogonia) and *bamGal4* (to drive in mid to late spermatogonia). Transgenes and Gal4-expressing lines were introduced into a $bam^{-/-}$ genetic background in which early stage spermatogonia overproliferate and fail to differentiate into spermatocytes, and expression from the transgenes was scored by immunofluorescence staining with anti-V5 and anti-HA antibodies (Fig. 7A–D). Anti-HA staining to visualize the overexpressed PCF11-RF-HA showed that the protein was highly nuclear in

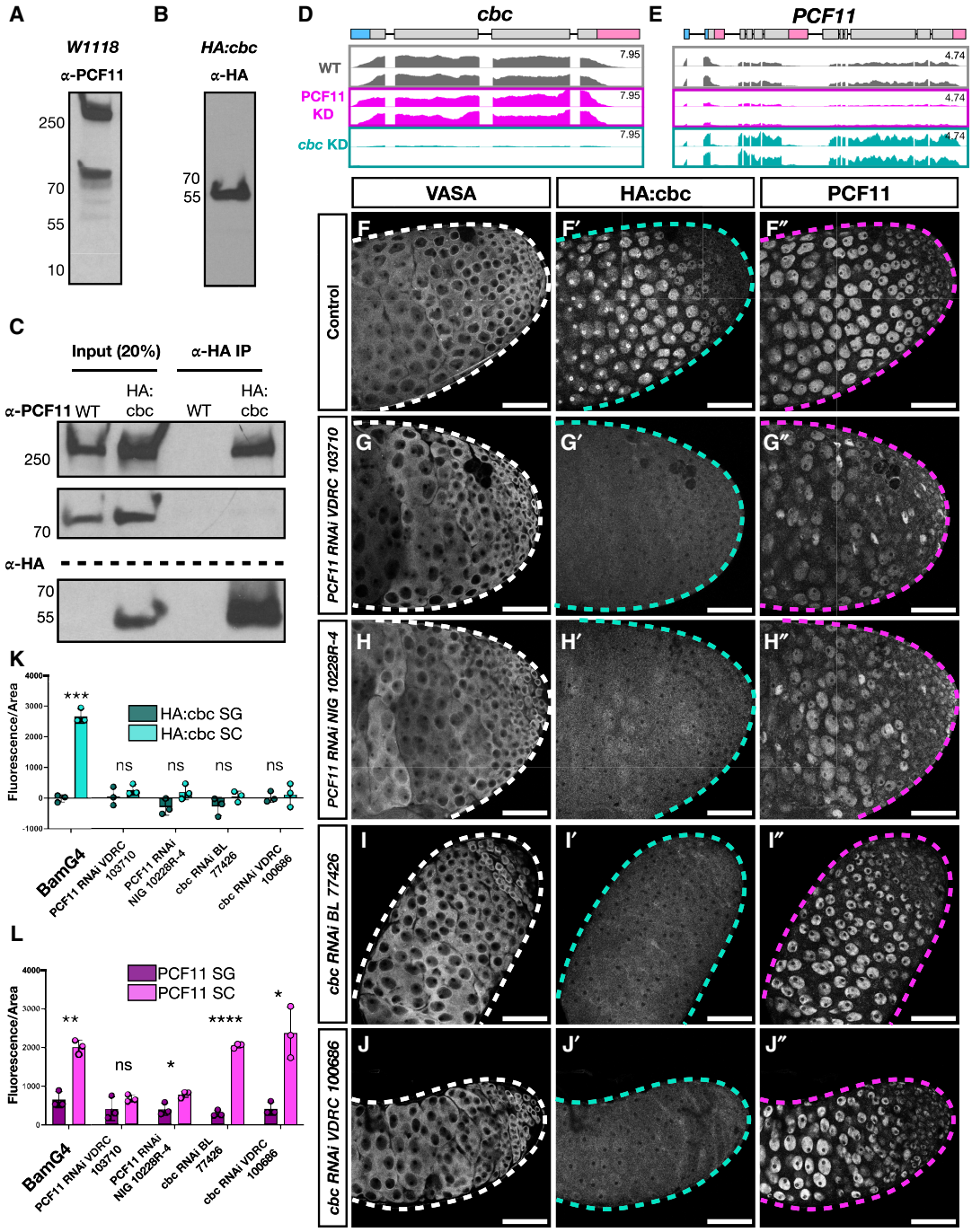

Figure 6. Knockdown of *PCF11* caused a decrease of Cbc protein. (*A*) Western blots of wild-type (*w1118*) testis lysate probed with anti-PCF11. (*B*) Western blots of *HA:cbc* homozygous testis lysate probed with anti-HA. (*C*) Immunoprecipitation (IP) of HA:Cbc in wild-type (*w1118*) and *HA:cbc* homozygous testes followed by PAGE and Western blots probed with anti-PCF11 or anti-HA. Input was 20% of lysate before IP. (*D*,*E*) Integrative Genomics Viewer (IGV) plots of *cbc* (*D*) and *PCF11* (*E*) loci of RNA-seq data from *bamGal4* (control) samples (gray), *PCF11-RNAi-VDRC-103710;bamGal4* samples (magenta), and *cbc-RNAi-BL-77426;bamGal4* samples (teal). (*F–J*) Immunofluorescence images showing apical tips of testes stained with anti-Vasa to mark germ cells, anti-HA, and anti-PCF11. (*F–F″*) *HA:cbc/CyO;bamGal4* (control), (*G–G″*) *HA:cbc/+,PCF11-RNAi-VDRC-103710;bamGal4*. (*H–H″*) *HA:cbc/+,PCF11-RNAi-NIG-10228R-4;bamGal4*. (*I–I″*) *HA:cbc/+,cbc-RNAi-BL-77426;bamGal4*. (*J–J″*) *HA:cbc/+,cbc-RNAi-VDRC-100686;bamGal4*. (*K*,*L*) Quantification of fluorescence/area from HA antibodies (teal) (*K*) and PCF11 antibodies (magenta) (*L*). For each testis, 10 spermatogonial and 10 spermatocyte nuclei were quantified for both anti-PCF11 and anti-HA fluorescence and then averaged. For each genotype, $N = 3$ testes were quantified. Statistical analysis was performed using the Welch's *t*-test. (**) $P = 0.0011$ for *bamGal4_PCF11*, (***) $P = 0.0005$ for b*amGal4_cbc*, (*) $P = 0.0341$ for *PCF11-RNAi-NIG10228R-4_PCF11*, (****) $P < 0.0001$ for *cbc-RNAi-BL77427_PCF11*, (*) $P = 0.0303$ for *cbc-RNAi-VDRC100686_PCF11*. Scale bars: *F–J′*, 50 µm. Entire gels for Western blots and co-IP are reported in Supplemental Figure S11.

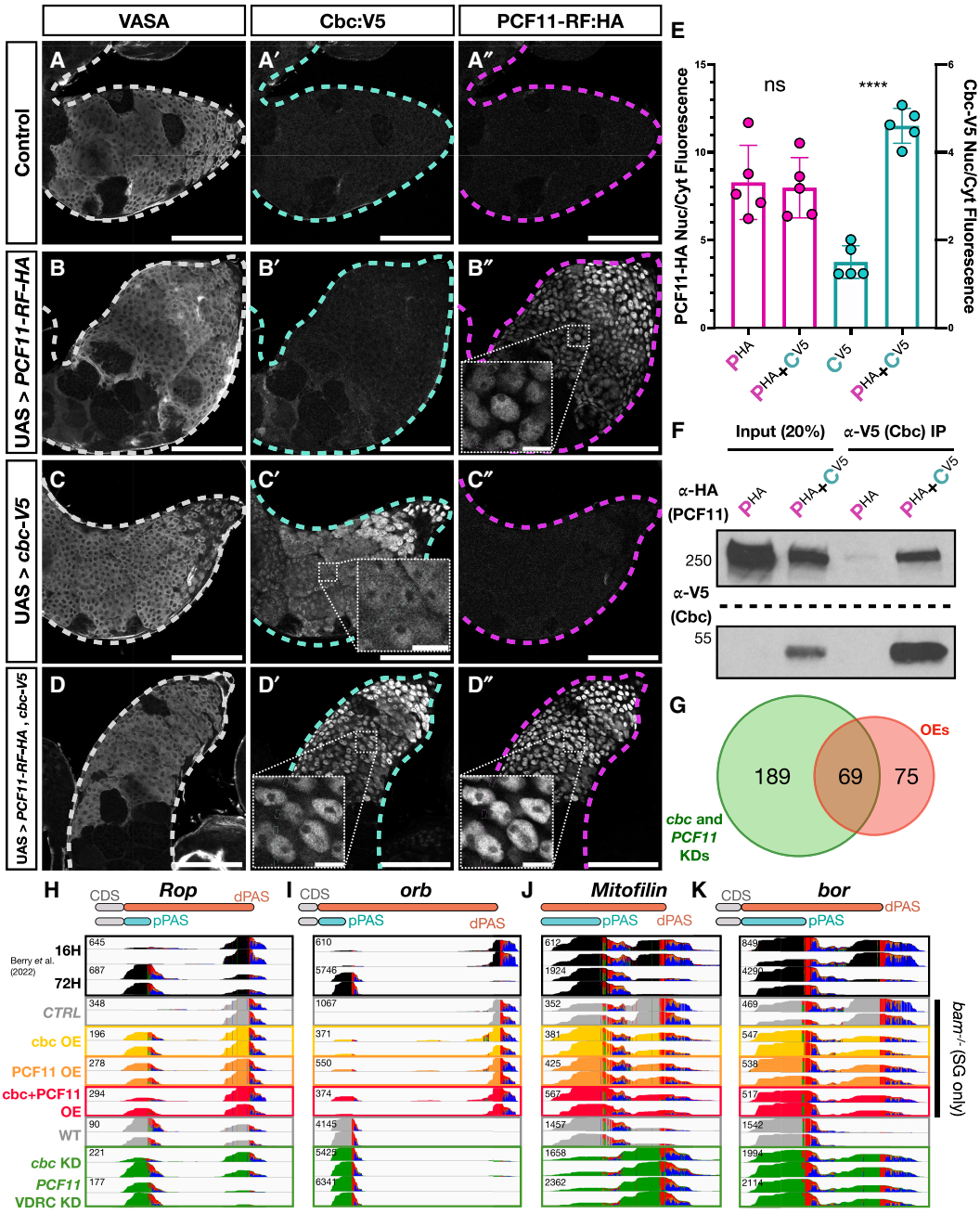

Figure 7. Overexpression of PCF11 and Cbc in spermatogonia shifted cleavage toward the more proximal site for some transcripts. (*A–D*) *bam*$^{-/-}$ testes stained with anti-Vasa, anti-HA, and anti-V5. (*A–A″*) *bam* mutant control lacking overexpression constructs (*nosGal4/Y; +/CyO; bam*^{86}*,bamGal4/bam*^{1}). (*B–B″*) *bam* mutant testes overexpressing *PCF11-RF-HA (nosGal4/Y; UAS > PCF11-RF-HA/+; bam*^{86}*, bamGal4/bam*^{1}*)*. (*C–C″*) *bam* mutant testes overexpressing *cbc-V5 (nosGal4/Y; UAS > cbc-V5/+; bam*^{86}*,bamGal4/bam*^{1}*)*. (*D–D″*) *bam* mutant testes overexpressing both *PCF11-RF-HA* and *cbc-V5 (nosGal4/Y; UAS > PCF11-RF-HA, UAS > cbc-V5/+; bam*^{86}*,bamGal4/bam*^{1}*)*. (*E*) Pink indicates quantification of the nuclear/cytoplasmic ratio of PCF11-RF-HA in *bam*$^{-/-}$ testes overexpressing *PCF11-RF-HA* alone or *PCF11-RF-HA* plus *cbc-V5*. Blue indicates quantification of the nuclear/cytoplasmic ratio of cbc-V5 in *bam*$^{-/-}$ testes overexpressing *cbc-V5* alone or *PCF11-RF-HA* plus *cbc-V5*. For each sample, $N = 5$ testes were quantified. Statistical analysis was performed using the Welch's *t*-test. (ns) $P = 0.8103$, (****) $P < 0.0001$. (*F*) Immunoprecipitation (IP) of *cbc-V5* from lysates of *bam*$^{-/-}$ testes in which either *PCF11-RF-HA* plus *cbc-V5* or *PCF11-RF-HA* alone (control) was overexpressed, followed by PAGE and Western blot probed with anti-HA and anti-V5 antibodies. Input was 20% of lysate before IP. (*G*) Venn diagram showing the number of genes that are affected in opposite APA directions by CFII knockdowns (green) or CFII overexpression (red). The green circle represents all the genes that had a poly(A) index >0.5 (lengthening) in either *PCF11* or *cbc* knockdowns, while the red circle represents all the genes that had a poly(A) index <−0.5 (shortening) in either *PCF11-RF-HA* overexpression, *cbc-V5* overexpression, or *PCF11-RF-HA* plus *cbc-V5* overexpression. (*H–K*) IGV plots showing 3′ end seq reads from libraries for overexpression and knockdown samples for the *Rop* (*H*), *orb* (*I*), *Mitofilin* (*J*), and *bor* (*K*) loci. (OE) Overexpression, (SG) spermatogonia, (CTRL) *nosGal4/Y; +/+; bam*^{86}*,bamGal4/bam*^{1}, (*cbc* OE) *nosGal4/Y; UAS > cbc-V5/+; bam*^{86}*,bamGal4/bam*^{1}, (*PCF11* OE) *NosGal4/Y; UAS > PCF11-RF-HA/+; bam*^{86}*,bamGal4/bam*^{1}, (*cbc + PCF11* OE) *nosGal4/Y; UAS > PCF11-RF-HA, UAS > cbc-V5/+; bam*^{86}*,bamGal4/bam*^{1}.

spermatogonia from flies overexpressing *PCF11-HA-RF* (Fig. 7B″) and from flies simultaneously overexpressing *cbc-V5* and *PCF11-RF-HA* (Fig. 7D″), recapitulating PCF11 nuclear localization in wild-type spermatogonia (Fig. 1J). Anti-V5 staining to visualize the overexpressed Cbc-V5 was instead both nuclear and cytoplasmic in spermatogonia from transgenic flies overexpressing *cbc-V5* (Fig. 7C′). However, in spermatogonia overexpressing both *PCF11-RF-HA* and *cbc-V5*, more of the anti-V5 signal was nuclear and less was cytoplasmic (Fig. 7D′). Quantification of the fluorescent signal from the nucleus and cytoplasm showed that the ratio of anti-HA signal in the nucleus compared with the cytoplasm remained constant when *PCF11-RF-HA* was overexpressed alone or with *cbc-V5*. However, the ratio of nuclear to cytoplasmic anti-V5 signal increased when *cbc-V5* was overexpressed with *PCF11-RF-HA* compared with when *cbc-V5* was overexpressed alone (Fig. 7E). Immunoprecipitation of Cbc-V5 using anti-V5 antibody from extracts of $bam^{-/-}$ testes in which both *PCF11-RF-HA* and *cbc-V5* were overexpressed in spermatogonia under control of *nanosGal4* and *bamGal4* brought down PCF11-RF-HA, indicating that the overexpressed tagged proteins interact, similar to their counterparts expressed in wild-type testes (Fig. 7F; entire gels in Supplemental Fig. S13).

Analysis of 3′ seq libraries from $bam^{-/-}$ testes in which *cbc-V5* alone, *PCF11-RF-HA* alone, or *cbc-V5* and *PCF11-RF-HA* together were overexpressed in spermatogonia compared with 3′ seq libraries from $bam^{-/-}$ testes in which neither was overexpressed revealed that of the ~500 genes identified by Berry et al. (2022) as producing mRNAs with different 3′ UTR lengths due to APA, ~140 showed upregulation of 3′ end cleavage at the proximal PAS (for example, see the genes *Rop*, *orb*, *Mitofilin*, and *bor*) (Fig. 7H–K). Forty-seven percent of the genes that showed an increase in usage of the proximal PAS normally used in spermatocytes [poly(A) index <−0.5] in samples in which either *cbc-V5*, *PCF11-RF-HA*, or both were overexpressed in $bam^{-/-}$ spermatogonia also showed higher usage of the distal PAS in samples in which *cbc* or *PCF11* had been knocked down in spermatocytes (Fig. 7G). The increased processing at proximal PASs in $bam^{-/-}$ testes in which CFII components were overexpressed only affected some of the nascent transcripts from a given locus and was never an APA switch as striking as the one observed when *cbc* or *PCF11* was knocked down in spermatocytes (Figs. 3, 4; Supplemental Fig. S14).

Discussion

Recent results show that stage-specific alternative cleavage and polyadenylation (APA) can regulate dramatic changes in the suite of proteins expressed as cells advance stepwise through a differentiation sequence (Berry et al. 2022). The developmental mechanisms that regulate cell type- and stage-specific APA can thus be key contributors to cellular differentiation. Our analysis has revealed that cell type-specific differential expression of components of cleavage factor complex II (CFII) of the 3′ end cleavage machinery regulates alternative polyadenylation in the *Drosophila* male germline adult stem cell lineage. Results of cell type-specific knockdown experiments suggest that the developmentally specified increase in expression of CFII components PCF11 and Cbc is important for the shift to 3′ end cleavage at a more proximal site in spermatocytes compared with spermatogonia at >200 loci during *Drosophila* spermatogenesis. Forcing overexpression of CFII components in spermatogonia resulted in increased 3′ end cleavage at the proximal site more typical of spermatocytes for a fraction of transcripts from some genes. Knockdown of several other cleavage factors in spermatocytes had much milder effects on shifting the 3′ end cleavage from the proximal site to the more distal site normally used in spermatogonia, singling out CFII for a key role in the developmentally regulated 3′ UTR shortening due to APA characteristic of wild-type *Drosophila* spermatocytes.

Cell type-specific upregulation of *cbc* may be a key factor in developmental regulation of APA in differentiating male germ cells. While *PCF11* transcripts are broadly expressed in fly adult tissues, *cbc* seems to be particularly highly expressed mostly in spermatocytes. In addition, previous studies involving knockdown of *PCF11* in cancer cells and zebrafish embryos showed that downregulating *PCF11* broadly induced 3′ UTR lengthening. On the other hand, knocking down *CLP1* (the mammalian homolog of *cbc*) in mammalian cells did not show remarkable effects on APA (Li et al. 2015; Ogorodnikov et al. 2018; Kamieniarz-Gdula et al. 2019). However, it is possible that *cbc* was not highly expressed or did not play a key role in regulating APA in the cell types tested. In contrast, knockdown of *cbc* in *Drosophila* spermatocytes strongly affected the APA profile, resulting in cleavage at the distal site typical of spermatogonia for almost half of the genes that would instead normally be cleaved at proximal sites in spermatocytes.

Several lines of evidence suggest that PCF11 may carry out functions independently of CFII. PCF11 has been found to bind the C-terminal domain (CTD) of RNA Pol II in vitro and in yeast (Sadowski et al. 2003; Zhang and Gilmour 2006), where it acts as a termination factor (Guéguéniat et al. 2017; Schäfer et al. 2018; Kamieniarz-Gdula et al. 2019). ChIP-seq of Pol II in control versus *PCF11* knockdown human cells showed that Pol II fails to terminate efficiently when PCF11 is reduced (Kamieniarz-Gdula et al. 2019). Accordingly, downregulating *PCF11* expression levels has been shown to induce 3′ UTR lengthening in cancer cells and zebrafish embryos (Li et al. 2015; Ogorodnikov et al. 2018; Kamieniarz-Gdula et al. 2019). Whether skipping upstream cleavage sites might be linked to the dynamics of Pol II termination and what other functions PCF11 might carry out in association with RNA Pol II still need to be elucidated. PCF11 has been shown to bind RNA in vitro, with a preference for G-rich regions. On the other hand, Cbc does not have a recognized RNA binding domain, and immunoprecipitation of Cbc in vitro failed to detect bound RNA (Schäfer et al. 2018). Our findings raise the possibility that Cbc might act as a developmentally regulated factor that, when upregulated in certain cell types, might bind to

PCF11 to form CFII, use the ability of PCF11 to recognize RNA, and modulate PCF11 activity to enhance 3′ end cleavage at specific sites.

Notably, not all of the ~500 genes found to undergo cell type-specific 3′ UTR shortening due to APA in *Drosophila* spermatocytes were affected by *PCF11* or *cbc* knockdown, suggesting that other developmental mechanisms (such as changes in transcription rates and/or activity of other RNA binding proteins) might play a role in regulating stage-specific APA at some genes. In addition, forced overexpression of CFII components in spermatogonia increased 3′ end cleavage at the proximal PAS normally used in spermatocytes only for some genes, suggesting that additional factors (perhaps PABP2, for example) may help drive cleavage at the proximal PAS. It is also possible that CFII has spermatocyte-specific components or interacting proteins that still need to be identified.

PCF11 has been found to be part of an autoregulatory feedback loop in cancer cells and zebrafish, where it autoregulates its own protein levels by activating a premature PAS (Kamieniarz-Gdula et al. 2019). Our 3′ end sequencing data revealed that a PAS located very far 5′ in the *PCF11-RC* coding sequence becomes highly used in spermatocytes compared with spermatogonia. Interestingly, cleavage at this premature PAS was dependent on *PCF11* and *cbc* levels, as the upstream PAS was no longer highly used if *PCF11* or *cbc* were knocked down in spermatocytes (Supplemental Fig. S15). This could mimic what has been observed in human cells and zebrafish embryos and suggests that the proposed PCF11 autoregulatory loop might be dependent on both components of the CFII complex rather than on PCF11 alone.

Developmentally regulated switches in the site at which nascent mRNAs are terminated, such as those we have observed to be dependent on components of CFII, may serve as a mechanism to quicky change the menu of proteins expressed as cells progress through sequential steps in differentiation. By affecting the length and content of 3′ UTRs, APA can change whether mRNAs are translated, blocked from translation, differentially localized, stabilized, or degraded. Since APA is a cotranscriptional process that does not require opening or closing of chromatin or reinitiation of transcription, it may provide a more rapid way to turn expression of some proteins on and other proteins off, changing cell state and canalizing cell fate decisions. As such, understanding how the alternative choices in 3′ end cut sites are regulated is key for understanding developmental regulation of cell states.

Materials and methods

Fly strains, husbandry, and design of UAS transgenes for overexpression.

Flies were grown on standard molasses media and kept at 25°C, 29°C, or room temperature as indicated. For RNAi-induced knockdowns, fly crosses (*bamGal4* driver females × RNAi line males) were kept for 3 days at 25°C. Adults were then shifted to a new media bottle and kept at 25°C while the eggs and early larvae were shifted to 29°C to increase knockdown efficiency. To further increase knockdown penetrance, we used *Gal4* drivers that also expressed *UAS > dicer2*. For overexpression in spermatogonia of *PCF11* and *cbc*, male flies expressing minigenes of *PCF11-RF*-HA-FLAG and/or *cbc*-V5 under the control of two 5XUAS sequences and a *hsp70* promoter were crossed with females bearing both the *nanos*Gal4 and *bam*Gal4 transgenes to induce expression in both early (*nanosGal4*) and late (*bamGal4*) spermatogonia. Crosses were kept for 3 days at 25°C. The adults were then removed, and the eggs and larvae were shifted to 29°C.

Plasmid manipulation was performed following standard cloning strategies. In brief, for the plasmid to force expression of *PCF11-RF-HA*, the plasmid LD11480 from the *Drosophila* Genomics Resource Center (DGRC) was used as a template, while for *cbc-V5* forced expression, the plasmid UFO01678 from the DGRC was used as a template. Plasmid injection and final line balancing were performed by BestGene. The construct to overexpress *PCF11-RF-HA* was inserted into an attP landing site on the left arm of the second chromosome, while the construct to overexpress *cbc-V5* was inserted into an attP landing site on the right arm of the second chromosome. Plasmid maps are reported in the Supplemental Material. Fly lines used in this study were *W1118* (Bloomington *Drosophila* Stock Center [BL] 3605), *bamGal4* (Chen and McKearin 2003), *nanosGal4-VP16* (Van Doren et al. 1998), bam^{1}/TM6B (McKearin and Spradling 1990), bam^{86}/TM6B (McKearin and Ohlstein 1995), *CPSF6:GFP* (Vienna *Drosophila* Resource Center [VDRC] 318105), *HA:Cbc* (N-terminal CRISPR tag; a kind gift from the Wang laboratory) (Wu et al. 2021), *PCF11*-RNAi_1 (VDRC 103710), *PCF11*_RNAi_2 (NIG 10228R4), *cbc*-RNAi_1 (BL 77426), *cbc*-RNAi_2 (VDRC 100686), *CPSF6*-RNAi_1 (BL 34804), *CPSF6*-RNAi_2 (VDRC 107147), *CstF64*-RNAi_1 (BL 65987), *CstF50*-RNAi_1 (BL 77377), *PABP2*-RNAi_1 (VDRC 106466), *PABP2*-RNAi_2 (VDRC 33499), *su(f)*-RNAi_1 (BL 55693), UAS > *PCF11-RF*-HA-FLAG (this study), and UAS > *cbc*-V5 (this study).

Immunofluorescence staining and signal quantification

Testes from 0–3 day old males were removed by dissection in 1× PBS on a cyclops dissecting dish over 30 min or less at room temperature, and testes were immediately fixed in a 1.7 mL Eppendorf tube with 1 mL of 5% formaldehyde in 1× PBS for 30 min at room temperature with rocking. If needed, fixed samples were stored in 1× PBS for up to 1 week at 4°C. Testes were then permeabilized in 0.3% Triton X-100 and 0.6% sodium deoxycholate in 1× PBS for 30 min at room temperature, rinsed twice in 1× PBS + 0.05% Tween-20 (PBT), blocked for 30 min in 1× Western blocking reagent (WBR; Roche 11921673001) in PBT at room temperature, and incubated with primary antibody mixed in 1× WBR + PBT overnight at 4°C. After primary antibody incubation, testes were rinsed twice in PBT, blocked again for 30 min in 1× WBR + PBT at room temperature, and incubated with secondary antibody mixed in 1× WBR + PBT for 2 h at room temperature in a dark container. Testes were

then rinsed twice in PBT and mounted on a glass slide in 10 µL of ProLong Diamond antifade with DAPI (Thermo Fisher P36962). Primary antibodies used were guinea pig anti-CstF50 (1:300; a kind gift from John T. Lis) (Ni et al. 2008), rabbit anti-PABP2 (1:500; a kind gift from Martine Simonelig) (Benoit et al. 1999), goat anti-Vasa (1:200; Santa Cruz Biotechnology sc-26877, discontinued), rat anti-Vasa (1:10 for supernatant or 1:100 for the concentrated version; both from Developmental Studies Hybridoma Bank antibody ID 760351), rabbit anti-CstF64 (1:400; a kind gift fom Zbigniew Dominski) (Skrajna et al. 2018), rabbit anti-PCF11 (1:200; a kind gift from David Scott Gilmour) (Zhang and Gilmour 2006), rabbit anti-HA (1:500; Cell Signaling Technologies 3724), mouse anti-HA (1:200; Biolegends 901501), chicken anti-GFP (1:10,000; Abcam 13970), and mouse anti-V5 (1:200; Invitrogen 46-0705). Secondary antibodies used were donkey or goat antirabbit, antimouse, anti-guinea pig, antichicken, antirat, and antigoat, all used at 1:500 and conjugated with either Alexa 488, Alexa 568, Alexa 594, or Alexa 647, depending on the combination of antibodies used in each experiment (all from Thermo Fisher). Imaging of IF experiments was performed on a Leica SP8 confocal microscope. Protein levels in Figures 1 and 6 were quantified in Fiji by manually delineating the cell nuclei and applying the formula Integrated_Density – (Area × Background_Mean). The nuclear:cytoplasmic ratio of Cbc-V5 and PCF11-RF-HA proteins in Figure 7 was quantified in Fiji by measuring the mean intensity of protein signal in 10 different uniform pixels areas per image in the central region of the sample and applying the formula $(\text{Mean\_protein} - \text{Mean\_background})_{\text{Nucleus}}/(\text{Mean\_protein} - \text{Mean\_background})_{\text{Cytoplasm}}$.

HCR RNA-FISH

Testes were dissected from 0–3 day old males in 1× PBS on a cyclops dissecting dish over a maximum of 30 min at room temperature, fixed, and permeabilized as done for immunofluorescence. The protocol for "generic sample in solution" from the Molecular Instruments website (https://www.molecularinstruments.com/hcr-rnafish-protocols) was followed. Hybridization buffer, wash buffer, amplification buffer, and hairpins were all from Molecular Instruments (HCR RNA-FISH bundle). After permeabilization as described for immunofluorescence above, samples were incubated with 200 µL of probe hybridization buffer for 30 min at 37°C. Probe mixes were prepared by mixing 3.2 µL of 1 µM gene-specific probe mix (one probe mix per gene or transcript) in 200 µL of probe hybridization buffer. Hybridization buffer was removed, probe mix was added, and the samples were incubated overnight in probe mix at 37°C in a water bath. The following day, samples were washed four times for 15 min each in 200 µL of probe wash buffer at 37°C and then washed three times for 5 min each in 5× SSCT (5× sodium chloride/sodium citrate (SSC; Invitrogen AM9770), 0.1% Tween-20 in H_2O) at room temperature. Samples were then incubated in 500 µL of amplification buffer for 30 min at room temperature. In the meantime, hairpin mix was prepared: 30 pmol per 5 µL sample of harpin h1 and hairpin h2 (each fluorophore had its own harpin combination; in our case, because we always imaged two probe sets at the same time, we used two fluorophores and therefore four hairpins) were heated for 90 sec at 95°C in a thermocycler in the dark and cooled in a drawer for 30 min at room temperature (each hairpin was heated and cooled in a separate tube). To prepare hairpin mixes, 5 µL of each harpin was added to 250 µL of amplification buffer. The samples were then incubated in harpin mix in the dark overnight at room temperature in a drawer with no rocking. The following day, samples were washed with 500 µL of 5× SSCT twice for 5 min each, twice for 30 min each, and finally once for 5 min, always in the dark and at room temperature. Samples were then mounted on a glass slide in 10 µL of ProLong Diamond antifade with DAPI (Thermo Fisher P36962). Probes were designed as in Bedbrook et al. (2023) and were obtained from Integrated DNA Technologies as oligo pools (50 pmol/oligo; oPools). Harpins h1 and h2 were ordered from Molecular Instruments conjugated with either 488 or 546 fluorophores. Probe sequences are listed in Supplemental Table S1. Imaging was performed on a Leica SP8 confocal microscope. For each testis, five areas of spermatocyte cytoplasm from different testis regions were quantified on Fiji from both probe channels, the background was subtracted, and then the ratio between the probe recognizing the 3′ UTR extensions and the probe recognizing the protein-coding region was calculated. The following formula was used: (Area_Mean_p2 – Back_Mean_p2)/(Area_Mean_p1 – Back_Mean_p1), where p2 is the probe set recognizing the 3′ UTR extension, and p1 is the probe set recognizing the protein-coding region of the transcript. The "Area" size varied between pictures but was constant between the five measurements from the same image.

3′ end sequencing and analysis

Testes were dissected from 0–3 day old males in S2 media in batches of 20–80 flies in a cyclops dissecting dish over a maximum of 1 h at room temperature per batch. Testes in the batch were then transferred together to a 1.7 mL Eppendorf tube containing 700 µL of 1× PBS. The 1× PBS was then immediately removed, and the testes were snap-frozen in liquid nitrogen and stored at −80°C. RNA was extracted from ∼150 testes pairs using the RNAeasy minikit (Qiagen 74104). Frozen tissue was dissociated using a 1 mL syringe with a 27 gauge needle aspirating up and down ∼10 times in RNAeasy minikit lysis buffer supplemented with 1:100 β-mercaptoethanol as suggested by the kit manual. The quality of extracted RNA was assessed via Bioanalyzer, and the concentration was measured by NanoDrop and Qubit RNA HS assay kit (Life Technologies Q32852). 3′ end RNA-seq libraries were prepared using the Quant Seq (FWD) kit from Lexogen using 500 ng of total extracted RNA per replicate. The quality of DNA libraries was checked via Bioanalyzer, and the concentration was measured with Qubit dsDNA HS assay kit (Life Technologies Q32854). Sequencing was performed by the Stanford University Genomics Facility on an Illumina NextSeq 500 platform as in Berry et al. (2022). Each biological sample was performed in duplicate. Reads

were filtered to retain only reads with ≥10 consecutive As. Reads were then trimmed to remove adaptor sequences and aligned to the *Drosophila* dm6 genome using bowtie 0.12.8 with a value of $e = 5000$ to allow for the expected mismatch in the poly(A) tail sequence, as in Berry et al. (2022). Quantification of PAS usage for the ∼500 genes identified by Berry et al. (2022) was performed using PolyA-miner (Yalamanchili et al. 2020) using as input a bed file containing the proximal and distal PASs of the ∼500 genes identified by Berry et al. (2022) as cleaved at a proximal PAS in spermatocytes and a distal PAS in spermatogonia. To consider various spermatocyte stages, we generated a master bed file by merging the proximal and distal cleavage sites from three different lists: $bam^{-/-}$ (enriched for spermatogonia) versus 32H (very early spermatocytes), $bam^{-/-}$ versus 48H (polar spermatocytes), and $bam^{-/-}$ versus 72H (mid-stage maturing spermatocytes). 32H, 48H, and 72H indicate hours after heat shock, which causes a pulse of bam expression in $bam^{-/-}$ flies, inducing differentiation as explained by Kim et al. (2017). As in these lists, the coordinates of the cleavage sites did not match perfectly; some genes had more than two cleavage sites considered. However, PolyA-miner was developed to consider genes with multiple poly(A) sites.

De novo PAS calling and APA analysis

For analysis of changes in PAS usage genome-wide between wild-type and knockdown samples, we adapted the analysis pipeline from Zeng et al. (2024). The sequenced 3′ end-seq libraries were adapter-trimmed and quality-filtered using bbduk. The filtered reads were mapped to the dm6 *Drosophila* genome using STAR and then extracted for unique mapping reads using Samtools. If a read was mapped to a region that was immediately upstream of six consecutive genomically encoded As or a 10 bp region with at least 60% As, it was considered misprimed and removed. The resulting filtered reads were used for de novo PAS calling using LAPA[61] with the default setting, except requiring a replication rate cutoff of 0.95. For each gene, lengthening or shortening was quantified by considering the ratio between the two highest changing PASs: the PAS that was more upregulated in the test sample versus the control sample and the PAS that was more downregulated in the test sample versus the control sample. For this analysis, we considered only genes for which both PASs were located in the 3′ UTR.

RNA-seq and analysis

The same RNA extractions used for 3′seq were also used for preparation of short-read RNA-seq libraries. One microgram of total extracted RNA was used to prepare libraries following the protocol from the NEBNext Ultra II direction RNA library preparation kit for Illumina (E7760 and E7765) according to the section for PolyA-selected RNA library preparation. Each biological sample was performed in duplicate. The quality and concentration of the libraries were measured via Bioanalyzer and Qubit, as for 3′seq libraries. Sequencing was performed by Novagene on a NOVAseq, PE 150 Illumina platform. Raw reads were trimmed to remove the Illumina adapters using Trim Galore and then aligned to the *Drosophila* dm6 genome using STAR. Differential expression analysis was performed using the package DEseq2.

Fly Cell Atlas single-nucleus RNA-seq analysis

Using the Fly Cell Atlas "testis" polydT-primed 10× single-nucleus RNA sequencing data set (Li et al. 2022), for each of the 111 Leiden 6.0 clusters (as in Berry et al. 2022), we aggregated all reads from nuclei in that cluster. For each cluster, reads that uniquely mapped to one of the isoform-specific *PCF11* 3′ UTRs were counted and normalized by the length of that 3′ UTR as well as the total number of reads derived from that cluster. This produced an isoform-specific RPKM value for each Leiden 6.0 cluster that was then log-normalized and plotted using Seurat. Gene expression and cell type annotation plots across all of the adult *Drosophila* tissues analyzed in the Fly Cell Atlas project were generated in Seurat using published code modified to visualize relevant genes and cell identities of interest (Raz et al. 2023).

Western blots

Testes from 20–50 (depending on the genotype and the protein needed to be detected) 0–3 day old males were quickly dissected into 1× PBS in a cyclops dissecting dish at room temperature and transferred to a 1.7 mL Eppendorf tube containing 700 µL of 1× PBS. The 1× PBS was then removed, and the tube was snap-frozen in liquid nitrogen and stored at −80°C until needed. Frozen tissue was lysed in 40 µL of cold lysis buffer—either RIPA buffer (50 mM TRIS at pH 7.5, 250 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, to which was added one tablet per 50 µL of protease inhibitor cocktail [cOmplete ultra tablets, Mini, EDTA-free, EASY pack; Roche 05892791001]) or JUME buffer (10 mM MOPS, 10 mM EDTA, 8 M urea, 1% SDS). Tissue was dissociated using a 1 mL syringe with a 27 gauge needle aspirating up and down ∼10 times and then incubated for 20 min at 4°C with rocking. Following incubation, lysate was centrifuged at 15,000 rpm for 5 min at 4°C. The supernatant was transferred to a new tube, mixed with 4× Laemli buffer with β-mercaptoethanol, boiled for 7 min at 85°C, and then run on a Mini-Protean TGX precast protein gel (4%–15% Mini-Protean TGX precast protein gels, 10 well plate, 50 µL; Bio-Rad 4561084). The separated proteins were transferred from the gel to a PVDF membrane (Bio-Rad 1620177) overnight at 4°C. The following day, the membrane was blocked overnight at 4°C in 5% nonfat milk powder dissolved in PBST (1× PBS, 0.1% Tween-20). After blocking, the membrane was incubated for 1 h at room temperature with primary antibody mix in 2% nonfat milk powder dissolved in PBST, washed three times for 20 min each with PBST at room temperature, incubated with secondary antibody mix for 1 h at room

temperature in 2% nonfat milk powder dissolved in PBST, and washed again three times for 20 min each with PBST at room temperature. Bound antibodies were then revealed using Revvity Health Sciences Western Lightning Plus-ECL (enhanced chemiluminescence substrate; PerkinElmer NEL104001EA). Primary antibodies used for Western blots were rabbit anti-HA (1:5000 or 1:1000; Cell Signaling Technologies 3724), rabbit anti-PCF11 (1:10.000; a kind gift from David Scott Gilmour) (Zhang and Gilmour 2006), and mouse anti-V5 (1:1000; Invitrogen 46-0705). Secondary antibodies used for Western blots were peroxidase AffiniPure goat antirabbit IgG (H+L; 1:10,000; Jackson ImmunoResearch AB_2313567) and peroxidase AffiniPure goat antimouse IgG (H+L; 1:10,000; Jackson ImmunoResearch AB_10015289).

Immunoprecipitation

For immunoprecipitation (IP) and co-IP experiments, testes were dissected into 1× PBS in a cyclops dissecting dish in batches of 50 flies for a maximum of 30 min at room temperature per batch and transferred into a 1.7 mL Eppendorf tube with 700 µL of 1× PBS. The 1× PBS was removed, and the tube was snap-frozen in liquid nitrogen and stored at −80°C. For one sample, ∼100 pairs of testes were used. The day before the actual IP, 2 µL per sample of rabbit anti-HA antibodies (Cell Signaling Technologies 3724) or 2 µL of mouse anti-V5 antibodies (Invitrogen 46-0705) was cross-linked to magnetic beads (Dynabeads protein A [Invitrogen 10002D] for antibodies raised in rabbits, and Dynabeads pan mouse IgG [Invitrogen 11041] for antibodies raised in mice) as described by Baker et al. (2023). Throughout this procedure, each of the following washes of the magnetic beads was done by gently resuspending the beads in solution, placing the 1.7 µL Eppendorf tubes containing beads on a magnetic rack, and removing the liquid. The day of the experiment, beads cross-linked with antibody were rinsed once in 200 mM glycine (pH 2.5) and washed once for 5 min in 200 mM glycine (pH 2.5) to remove noncross-linked antibodies, blocked for 1 h in 10% BSA in 50 mM Tris (pH 8), resuspended in 20 µL of RIPA buffer, and stored at 4°C while the testis lysate was prepared. Frozen testis tissue was dissociated in RIPA buffer using a 1 mL syringe and a 27 gauge needle as in the Western blot experiments. The lysate was rocked for 20 min at 4°C and then centrifuged at maximum speed for 5 min in a 1.7 mL Eppendorf tube at 4°C to separate the lysate from tissue debris. The supernatant was transferred to a new 1.7 mL Eppendorf tube and incubated for 45 min at 4°C with beads with no antibody. Tubes containing the lysate and beads were placed on a magnetic rack, and 40 µL of precleared lysate was saved as input sample. The volume was brought back to 220 µL by adding 40 µL of cold RIPA buffer and pipetted into the tubes containing beads cross-linked to antibody and left incubating for 4 h at 4°C with rocking. Because we noticed that HA:Cbc in testis lysates tended to form aggregates (see Supplemental Fig. S11), we protein-precipitated (see the protocol below) the input sample and resuspended the formed protein pellet in JUME buffer. Laemli sample buffer (4×) (Bio-Rad 1610747) mixed with β-mercaptoethanol was added to the input sample to a final 1× concentration, incubated for 7 min at 85°C, and frozen. After 4 h of rocking at 4°C, IP samples were placed on a magnet to separate beads from lysate, and 40 µL of the supernatant was saved as the supernatant sample, mixed with 4× Laemli sample buffer with β-mercaptoethanol, incubated for 7 min at 85°C, and frozen. The rest of the supernatant was discarded. The beads with attached protein were immediately resuspended in RIPA buffer and washed with RIPA buffer twice for 5 min at room temperature with rocking. Tubes were placed on a magnet to separate beads from RIPA buffer, RIPA buffer was removed, and the beads were resuspended in 40 µL of elution buffer (1× cOmplete Ultra tablets, 1% SDS, 10 mM EDTA, 50 mM Tris at pH 8), vortexed, and incubated for 30 min at 70°C with frequent (every 5 min) vortexing or shaking. Tubes were then centrifuged at maximum speed for 10 sec at room temperature and then placed on the magnet to separate beads from the eluate. Forty microliters of eluate was moved to another 1.7 mL tube (IP sample) to which 4× Laemli sample buffer with β-mercaptoethanol was added, and the sample was incubated for 7 min at 85°C and frozen. The beads were then mixed with 40 µL of elution buffer, mixed with 4× Laemli sample buffer with β-mercaptoethanol, incubated for 7 min at 85°C, and frozen. Samples were then run on a Mini-Protean TGX precast protein gel, and the above protocol for Western blots was followed. If the same membrane needed to be restained with another antibody, the membrane was stripped by incubation for 30 min at 55°C in 62.5 mM Tris (pH 6.8), 2% SDS, and 100 mM β-mercaptoethanol; washed three times for 10 min each at room temperature in PBST; and either stored in 1× PBS or blocked again before being restained as above. Alternatively, for a milder stripping, the membrane was incubated overnight at 4°C with 5% nonfat milk in PBST + 0.1 mM sodium azide. The membrane was blocked again and then restained as above.

Protein precipitation

To precipitate proteins, 4 vol of ice-cold methanol was added to 1 vol of testis lysate and vortexed, and 1 vol of chloroform was added. The sample was vortexed again, mixed with 3 vol of H_2O, vortexed, and centrifuged in a 1.7 mL Eppendorf tube at maximum speed for 5 min at 4°C. After centrifugation, two liquid layers formed, with proteins at the interface. The upper liquid layer was removed, and 3 vol of ice-cold methanol was added to the sample. After a quick vortex, the sample was centrifuged at maximum speed for 10 min at 4°C to pellet the proteins. After centrifugation, all liquid was removed, and the white protein pellet was air-dried and resuspended in 40 µL of JUME buffer.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank all the members of the Fuller laboratory for suggestions and comments on this manuscript, and Dr. Nicole Martinez and Dr. Lauren Goins for their valuable feedback on the manuscript and suggested experiments. We thank Dr. Catherine Baker for suggestions on immunoprecipitations and for proofreading the manuscript, and Benjamin Bolival for helping in maintaining fly crosses. We thank Dr. Jingxun Chen from the Brunet laboratory at Stanford University for her help in designing HCR-FISH probes and for sharing some of the reagents, the Wang laboratory for the *HA:cbc* CRISPR tag flies, the Lis laboratory for anti-CstF50, the Simonelig laboratory for anti-PABP2, the Dominski laboratory for anti-CstF64, and the Gilmour laboratory for anti-PCF11 antibodies. We thank Daniel Mokhtari for his help parsing the Fly Cell Atlas data set, and Dr. Hari Krishna Yalamanchili for his help in running PolyA-miner. We thank the Cell Science Imaging Facility (CSIF) at Stanford University, supported in part by award number 1S10OD010580-01A1 from the National Center for Research Resources (NCRR), particularly the CSIF manager Kitty Lee. (The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the NCRR or the National Institutes of Health.) We thank the Stanford Fly Media Center as well and all the staff working in the Department of Developmental Biology. L.G. was supported by an American Italian Cancer Foundation (AICF) postdoctoral fellowship (2021–2023). This research was supported by National Institutes of Health grant R35 GM136433 and funds from the Katherine D. McCormick and Stanley McCormick Memorial Professorship and the Reed-Hodgson Professorship in Human Biology to M.T.F. This manuscript is dedicated to Cuchuflí.

Author contributions: L.G. and M.T.F. conceived the project and wrote the manuscript. L.G. designed the experimental work and performed or supervised all experimental and computational procedures. N.R.M. contributed to experimental design and bioinformatic analysis. F.M.-P. contributed to experimental design and performed molecular cloning. I.N. performed immunofluorescence experiments. S.S. performed bioinformatic analysis of the Fly Cell Atlas data. Y.Z. performed bioinformatic analysis to de novo map PASs. Funding was acquired by L.G. and M.T.F.

References

An JJ, Gharami K, Liao GY, Woo NH, Lau AG, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, et al. 2008. Distinct role of long 3′ UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* **134:** 175–187. doi:10.1016/j.cell.2008.05.045

Baker CC, Gallicchio L, Matias NR, Porter DF, Parsanian L, Taing E, Tam C, Fuller MT. 2023. Cell-type-specific interacting proteins collaborate to regulate the timing of cyclin B protein expression in male meiotic prophase. *Development* **150:** dev201709. doi:10.1242/dev.201709

Bedbrook CN, Nath RD, Nagvekar R, Deisseroth K, Brunet A. 2023. Rapid and precise genome engineering in a naturally short-lived vertebrate. *Elife* **12:** e80639. doi:10.7554/eLife.80639

Benoit B, Simonelig M, Nemeth A, Aulner N, Kühn U, Wahle E, Bourbon HM. 1999. The *Drosophila* poly(A)-binding protein II is ubiquitous throughout *Drosophila* development and has the same function in mRNA polyadenylation as its bovine homolog in vitro. *Nucleic Acids Res* **27:** 3771–3778. doi:10.1093/nar/27.19.3771

Berry CW, Olivares GH, Gallicchio L, Ramaswami G, Glavic A, Olguín P, Li JB, Fuller MT. 2022. Developmentally regulated alternate 3′ end cleavage of nascent transcripts controls dynamic changes in protein expression in an adult stem cell lineage. *Genes Dev* **36:** 916–935. doi:10.1101/gad.349689.122

Chen D, McKearin DM. 2003. A discrete transcriptional silencer in the *bam* gene determines asymmetric division of the *Drosophila* germline stem cell. *Development* **130:** 1159–1170. doi:10.1242/dev.00325

Choi HMT, Schwarzkopf M, Fornace ME, Acharya A, Artavanis G, Stegmaier J, Cunha A, Pierce NA. 2018. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145:** dev165753. doi:10.1242/dev.165753

Clerici M, Faini M, Muckenfuss LM, Aebersold R, Jinek M. 2018. Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat Struct Mol Biol* **25:** 135–138. doi:10.1038/s41594-017-0020-6

Gallicchio L, Olivares GH, Berry CW, Fuller MT. 2023. Regulation and function of alternative polyadenylation in development and differentiation. *RNA Biol* **20:** 908–925. doi:10.1080/15476286.2023.2275109

Grassi E, Santoro R, Umbach A, Grosso A, Oliviero S, Neri F, Conti L, Ala U, Provero P, DiCunto F, et al. 2019. Choice of alternative polyadenylation sites, mediated by the rna-binding protein Elavl3, plays a role in differentiation of inhibitory neuronal progenitors. *Front Cell Neurosci*. **12:** 518. doi:10.3389/fncel.2018.00518

Grozdanov PN, Masoumzadeh E, Latham MP, MacDonald CC. 2018. The structural basis of CstF-77 modulation of cleavage and polyadenylation through stimulation of CstF-64 activity. *Nucleic Acids Res* **46:** 12022–12039. doi:10.1093/nar/gky862

Gruber AJ, Zavolan M. 2019. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet* **20:** 599–614. doi:10.1038/s41576-019-0145-z

Guéguéniat J, Dupin AF, Stojko J, Beaurepaire L, Cianférani S, Mackereth CD, Minvielle-Sébastia L, Fribourg S. 2017. Distinct roles of PCF11 zinc-binding domains in pre-mRNA 3′-end processing. *Nucleic Acids Res* **45:** 10115–10131. doi:10.1093/nar/gkx674

Hilgers V, Lemke SB, Levine M. 2012. ELAV mediates 3′ UTR extension in the *Drosophila* nervous system. *Genes Dev* **26:** 2259–2264. doi:10.1101/gad.199653.112

Jing Z, Linda H, Claire M. 1999. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **63:** 405–445. doi:10.1128/MMBR.63.2.405-445.1999

Kamieniarz-Gdula K, Gdula MR, Panser K, Nojima T, Monks J, Wiśniewski JR, Riepsaame J, Brockdorff N, Pauli A, Proudfoot NJ. 2019. Selective roles of vertebrate PCF11 in premature and full-length transcript termination. *Mol Cell* **74:** 158–172.e9. doi:10.1016/j.molcel.2019.01.027

Kim J, Lu C, Srinivasan S, Awe S, Brehm A, Fuller MT. 2017. Blocking promiscuous activation at cryptic promoters directs cell type-specific gene expression. *Science* **356:** 717–721. doi:10.1126/science.aal3096

Lee S, Chen Y-C, Gillen AE, Taliaferro JM, Deplancke B, Li H, Lai EC. 2022. Diverse cell-specific patterns of alternative polyadenylation in *Drosophila*. *Nat Commun* **13:** 5372. doi:10.1038/s41467-022-32305-0

Li W, You B, Hoque M, Zheng D, Luo W, Ji Z, Park JY, Gunderson SI, Kalsotra A, Manley JL, et al. 2015. Systematic profiling of poly(A)$^+$ transcripts modulated by core 3′ cnd processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet* **11:** e1005166. doi:10.1371/journal.pgen.1005166

Li W, Park JY, Zheng D, Hoque M, Yehia G, Tian B. 2016. Alternative cleavage and polyadenylation in spermatogenesis connects chromatin regulation with post-transcriptional control. *BMC Biol* **14:** 6. doi:10.1186/s12915-016-0229-6

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

MacDonald CC, Wilusz J, Shenk T. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol Cell Biol* **14:** 6647–6654. doi:10.1128/mcb.14.10.6647-6654.1994

Mansfield KD, Keene JD. 2012. Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation. *Nucleic Acids Res* **40:** 2734–2746. doi:10.1093/nar/gkr1114

Marzluff WF, Wagner EJ, Duronio RJ. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* **9:** 843–854. doi:10.1038/nrg2438

Mayr C, Bartel DP. 2009. Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138:** 673–684. doi:10.1016/j.cell.2009.06.016

McKearin D, Ohlstein B. 1995. A role for the *Drosophila* Bag-of-marbles protein in the differentiation of cystoblasts from germline stem cells. *Development* **121:** 2937–2947. doi:10.1242/dev.121.9.2937

McKearin DM, Spradling AC. 1990. bag-of-marbles: a *Drosophila* gene required to initiate both male and female gametogenesis. *Genes Dev* **4:** 2242–2251. doi:10.1101/gad.4.12b.2242

Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC. 2013. Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. *Genome Res* **23:** 812–825. doi:10.1101/gr.146886.112

Morris AR, Bos A, Diosdado B, Rooijers K, Elkon R, Bolijn AS, Carvalho B, Meijer GA, Agami R. 2012. Alternative cleavage and polyadenylation during colorectal cancer development. *Clin Cancer Res* **18:** 5256–5266. doi:10.1158/1078-0432.CCR-12-0543

Ni Z, Saunders A, Fuda NJ, Yao J, Suarez J-R, Webb WW, Lis JT. 2008. P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. *Mol Cell Biol* **28:** 1161–1170. doi:10.1128/MCB.01859-07

Ogorodnikov A, Levin M, Tattikota S, Tokalov S, Hoque M, Scherzinger D, Marini F, Poetsch A, Binder H, Macher-Göppinger S, et al. 2018. Transcriptome 3′end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat Commun* **9:** 5331. doi:10.1038/s41467-018-07580-5

Raz AA, Vida GS, Stern SR, Mahadevaraju S, Fingerhut JM, Viveiros JM, Pal S, Grey JR, Grace MR, Berry CW, et al. 2023. Emergent dynamics of adult stem cell lineages from single nucleus and single cell RNA-seq of *Drosophila* testes. *eLife* **12:** e82201. doi:10.7554/eLife.82201

Sadowski M, Dichtl B, Hübner W, Keller W. 2003. Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination. *EMBO J* **22:** 2167–2177. doi:10.1093/emboj/cdg200

Sarov M, Barz C, Jambor H, Hein MY, Schmied C, Suchold D, Stender B, Janosch S, KJ VV, Krishnan RT, et al. 2016. A genome-wide resource for the analysis of protein localisation in *Drosophila*. *eLife* **5:** e12068. doi:10.7554/eLife.12068

Schäfer P, Tüting C, Schönemann L, Kühn U, Treiber T, Ihling C, Graber A, Keller W, Meister G, Sinz A, et al. 2018. Reconstitution of mammalian cleavage factor II involved in 3′ processing of mRNA precursors. *RNA* **24:** 1721–1737. doi:10.1261/rna.068056.118

Schulte SJ, Fornace ME, Hall JK, Shin GJ, Pierce NA. 2024. HCR spectral imaging: 10-plex, quantitative, high-resolution RNA and protein imaging in highly autofluorescent samples. *Development* **151:** dev202307. doi:10.1242/dev.202307

Shan L, Wu C, Chen D, Hou L, Li X, Wang L, Chu X, Hou Y, Wang Z. 2017. Regulators of alternative polyadenylation operate at the transition from mitosis to meiosis. *J Genet Genomics* **44:** 95–106. doi:10.1016/j.jgg.2016.12.007

Skrajna A, Yang X, Dadlez M, Marzluff WF, Dominski Z. 2018. Protein composition of catalytically active U7-dependent processing complexes assembled on histone pre-mRNA containing biotin and a photo-cleavable linker. *Nucleic Acids Res* **46:** 4752–4770. doi:10.1093/nar/gky133

Tian B, Manley JL. 2017. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18:** 18–30. doi:10.1038/nrm.2016.116

Trivedi V, Choi HMT, Fraser SE, Pierce NA. 2018. Multidimensional quantitative analysis of mRNA expression within intact vertebrate embryos. *Development* **145:** dev156869. doi:10.1242/dev.156869

Turner RE, Henneken LM, Liem-Weits M, Harrison PF, Swaminathan A, Vary Robert, Nikolic I, Simpson KJ, Powell DR, Beilharz TH, et al. 2020. Requirement for cleavage factor II_m in the control of alternative polyadenylation in breast cancer cells. *RNA* **26:** 969–981. doi:10.1261/rna.075226.120

Vallejos Baier R, Picao-Osorio J, Alonso CR. 2017. Molecular regulation of alternative polyadenylation (APA) within the *Drosophila* nervous system. *J Mol Biol* **429:** 3290–3300. doi:10.1016/j.jmb.2017.03.028

Van Doren M, Williamson AL, Lehmann R. 1998. Regulation of zygotic gene expression in *Drosophila* primordial germ cells. *Curr Biol* **8:** 243–246. doi:10.1016/S0960-9822(98)70091-0

Venkat S, Tisdale AA, Schwarz JR, Alahmari AA, Carlo Maurer H, Olive KP, Eng KH, Feigin ME. 2020. Alternative polyadenylation drives oncogenic gene expression in pancreatic ductal adenocarcinoma. *Genome Res* **30:** 347–360. doi:10.1101/gr.257550.119

Wu J, Li X, Gao Z, Pang L, Liu X, Huang X, Wang Y, Wang Z. 2021. RNA kinase CLP1/Cbc regulates meiosis initiation in spermatogenesis. *Hum Mol Genet* **30:** 1569–1578. doi:10.1093/hmg/ddab107

Yalamanchili HK, Alcott CE, Ji P, Wagner EJ, Zoghbi HY, Liu Z. 2020. PolyA-miner: accurate assessment of differential alternative poly-adenylation from 3′seq data using vector projections and non-negative matrix factorization. *Nucleic Acids Res* **48:** e69. doi:10.1093/nar/gkaa398

Yang Q, Gilmartin GM, Doublié S. 2011. The structure of human cleavage factor I_m hints at functions beyond UGUA-specific RNA binding. *RNA Biol* **8:** 748–753. doi:10.4161/rna.8.5.16040

Zeng Y, Lovchykova A, Akiyama T, Liu C, Guo C, Jawahar VM, Sianto O, Calliari A, Prudencio M, Dickson DW, et al. 2024. TDP-43 nuclear loss in FTD/ALS causes widespread alternative polyadenylation changes. bioRxiv doi:10.1101/2024.01.22.575730

Zhang Z, Gilmour DS. 2006. Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol Cell* **21:** 65–74. doi:10.1016/j.molcel.2005.11.002

Zhu Y, Wang X, Forouzmand E, Jeong J, Qiao F, Sowd GA, Engelman AN, Xie X, Hertel KJ, Shi Y. 2018. Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. *Mol Cell* **69:** 62–74.e4. doi:10.1016/j.molcel.2017.11.031