LENG8 granule regulates alternative polyadenylation in mammals

Zhi-Cheng Wu,[1,6] Chen Du,[2,6] Xue-Ying Huang,[1] You-Wan Qin,[1] Ai Zhong,[3] Ming Rao,[3] Feng-Ming Liu,[4] Hong-Wen Zhu,[5] Yu Zhou,[2*] Peng Dai[1*]

[1]Shanghai Key Laboratory of Maternal and Fetal Medicine, Clinical and Translational Research Center of Shanghai First Maternity and Infant Hospital, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China

[2]College of Life Sciences, TaiKang Center for Life and Medical Sciences, RNA Institute, Hubei Key Laboratory of Cell Homeostasis, Wuhan University, Wuhan, Hubei 430072, China

[3]Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences; University of Chinese Academy of Sciences, Shanghai 200031, China

[4]National Center for Protein Science Shanghai, Shanghai 201210, China

[5]Precise Genome Engineering Center, School of Life Sciences, Guangzhou University, Guangzhou 510006, China

[6]These authors contributed equally: Zhi-Cheng Wu, Chen Du.

* e-mail: yu.zhou@whu.edu.cn; daipeng@tongji.edu.cn

Membraneless organelles participate in multiple cellular processes and layers of gene regulation. Alternative polyadenylation (APA) is emerging as a widespread mechanism for controlling gene expression by generating distinct 3′ untranslated regions (UTRs) in transcripts. However, whether specific membraneless organelles exist in mammals to regulate APA remains largely unknown. Here, we unveil a previously unidentified nuclear body, termed LENG8 granule, which functions as an active hub for APA regulation. *LENG8* deletion leads to extensive 3′UTR lengthening in both cultured cells and male germ cells. LENG8 granules recruit cleavage and polyadenylation (CPA) factors and promote their local assembly, thereby favoring the utilization of proximal poly(A) sites (pPASs). Additionally, LENG8 granules function both independently and cooperatively with nuclear speckles to regulate mRNA 3′ end processing. We also elucidate the pivotal roles of YTHDC1-m^{6}A landscapes in LENG8 granule formation/maintenance and target selection. The 3′UTR length of targets correlates with the formation and properties of LENG8 granules in a cell type-specific and developmental stage-dependent manner. Our findings highlight LENG8 as a key factor in 3′ end processing and uncover a hitherto unrecognized nuclear body involved in

mammalian APA regulation, which broadens our understanding of membraneless organelles and their role in poly(A) site selection.

Alternative polyadenylation (APA), a widespread and dynamic co-transcriptional process that increases the diversity of mRNA transcripts with alternative 3′ untranslated regions (UTRs), has been deemed as a critical layer of gene expression regulation.[1] APA can affect mRNA transcriptional termination, splicing pattern, nuclear export, stability, localization, translation efficiency, protein-protein interaction and functions.[2-5] Due to the cell-type and developmental stage-specific manner of alternative 3′UTRs, dysregulation of APA has been linked to many distinct pathological disorders, including male infertility, cancer, immunological, hematological, and neurological disorders.[6-11]

The process of cleavage and polyadenylation (CPA) is governed by several RNA-binding proteins (RBPs) and four major multiprotein complexes: cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CSTF), cleavage factor I and II (CFI and CFII), through recognition of the polyadenylation signal (PAS, AAUAAA or its variants), downstream U/GU-rich element, and upstream UGUA motif.[5,9,12-16] Gene deletion and the analysis of massive quantitative data of poly (A) site usage have revealed these factors can regulate the selection between proximal and distal poly (A) sites.[9] However, the effects of CPA factors on poly (A) site choice do not always align with changes in their expression levels.[17-19] Recent studies suggest that condensates formed by CPA factors through liquid-liquid phase separation (LLPS) are involved in APA regulation.[4,19-21] Additionally, the RNA-binding protein FCA, involved in flowering time controlling, has the capacity to undergo phase separation with the help of FLL2 to form nuclear bodies to enhance the proximal poly(A) sites (pPASs) usage in *Arabidopsis*.[22] Neither FCA nor FLL2 exhibits direct homology to any component of the mammalian polyadenylation machinery. Nevertheless, the mechanisms through which nuclear bodies regulate APA, the mechanism for target selection, and the presence of nuclear bodies dedicated to APA regulation in mammals remain unclear.

Over the past 20 years, LENG8 was assumed to be potentially involved in glycosylated protein recognition, RNA transport, and its variant/ differential expression may serve as a biomarker for the pathogenesis of Pemphigus foliaceus, kidney renal clear cell carcinoma, breast cancer patients in human, and susceptibility or resistance to *Streptococcus suis* type 2 infection in pigs.[23-29] However, the physiological significance and underlying mechanisms of LENG8 remain a mystery.

In this study, we found that depletion of *LENG8* in both male germ cells and cultured cells causes lengthening of the 3′UTR. Both LENG8 and CPA factors localize in a previous unknown nuclear body (termed LENG8 granule or G granule) in mammalian cells. LENG8 effectively competes with PABPN1 for binding or directly binds to CPSF4, thereby facilitating the assembly of the CPA complex at pPASs in a YTHDC1-m^{6}A-dependent manner. The 3′ UTR length of target genes depends directly on LENG8 nuclear localization and granule assembly, supporting the notion that LENG8 granules

are active sites for facilitating the usage of pPASs. Our studies thus identify a new 3′ end processing factor and unravel the function of a never before described LENG8 nuclear body in APA regulation.

Results

LENG8 can be efficiently precipitated from testis extracts by b-isox

Eukaryotic cells employ liquid-liquid phase separation (LLPS) to form a large number of distinct membraneless organelles (biomolecular condensates), including nuclear speckles, Cajal bodies, nucleoli, stress granules, and germ granules, participating in various cellular processes and layers of gene regulation.[30-33]

Spermatogenesis, a highly intricate and tightly regulated cellular developmental process occurring in the male testes, initiates with the self-renewal and differentiation of spermatogonia stem cells into spermatogonia. These spermatogonia subsequently differentiate into spermatocytes, which undergo two successive meiotic divisions to produce haploid spermatids. Spermatids then undergo morphological changes, passing through stages of round spermatids, elongating spermatids and elongated spermatids, ultimately developing into mature spermatozoa.[34] Accumulating studies demonstrate that membraneless organelles, including germ granules, are crucial for spermatogenesis.[35] To screen potential phase-separation proteins for novel membraneless organelles formation in spermatogenesis, we employed biotinylated isoxazole (b-isox), a small molecule chemical capable of forming microcrystals to precipitate proteins within low complexity (LC) sequences,[36] to obtain pellets from mouse testis extracts at four different developmental stages. This approach yielded a total of 379 shared proteins which were defined as high-potential phase-separation proteins (Fig.1a and Supplementary Table 1). Gene ontology (GO) analysis of these proteins showed that many of them are RNA-binding proteins (RBPs) (Fig.1b). We additionally used PrionW to predict proteins containing Q/N rich prion-like domains (PrLDs), which have been identified as a driver for phase separation of RBPs.[37-39] Our query revealed 4 proteins that have PrLDs in 379 shared proteins (Extended Data Fig.1a).

Among these four candidate proteins, LENG8 caught our attention as it was able to pull down several proteins from CPSF and CSTF complexes (Supplementary Table 2). LENG8 is ubiquitously expressed across mouse tissues, albeit in different amounts and isoforms (Fig.1c and Extended Data Fig.1b). However, the expression of protein rather than RNA is almost constant in testes at different developmental stages (Fig.1d and Extended Data Fig.1c, d). The sequence of LENG8 protein is highly conserved across various species. The variations in the C-terminus imply a potential species functional specificity (Extended Data Fig.1e). LENG8 contains one predicted CSN8_PSD8_EIF3K domain, and harbors an N-terminal intrinsically disordered region (IDR) (Fig.1e, f). Together, our data suggest that LENG8, akin to proteins with LC/IDR or PrLD, interacts with b-isox and exhibits potential for LLPS.

LENG8 forms granule-like condensates in the nucleus

Given that LENG8 was efficiently precipitated by b-isox and that LENG8 consists of PrLD and IDR sequences, we asked whether LENG8 can form liquid-like granules *in vivo*. Using co-immunofluorescence (IF) of LENG8 and DDX25 (a marker of germ granules), we observed that LENG8 localizes to intranuclear foci in spermatocytes and haploid round spermatids, which are distinct from germ granules in cytoplasm (Extended Data Fig.2a). In line with this, LENG8 proteins are detected within the cytoplasm of spermatogonia and subsequently translocate into the nucleus, where they form foci in spermatocytes and round spermatids. In elongating/elongated spermatids, LENG8 proteins become undetectable (Fig.2a, b). To further verify the subcellular localization of LENG8 in germ cells, we generated HA-*LENG8* knock-in mice (HA tag in N-terminus) and conducted IF staining using anti-HA antibodies, revealing consistent results with those obtained using anti-LENG8 antibodies (Extended Data Fig.2b). We then investigated the relationship between LENG8 foci and various nuclear bodies in cultured 293T and HeLa cells. Our findings revealed that LENG8 localizes in a previously undefined granule, termed the LENG8 granule or G granule, which is distinguishable from those well-known nuclear bodies (Fig.2c). Notably, LENG8 forms large, prominent, and spherical foci in 293T, C2C12, U2OS and HeLa cell lines, while generating numerous smaller condensates in the nucleus of H9, N2A, MCF7 and HCT116 cell lines, which closely resemble condensates undergoing LLPS (Fig.2d). Large granules detected in 293T and C2C12 cells likely result from the coalescence of smaller granules (Fig.2d). Additionally, LENG8 also produces a significant number of small condensates in the cytoplasm (Fig.2d). Consistent with these findings in cultured cells, the formation of LENG8 granules and the localization patterns of LENG8 proteins also vary across tissues (Fig.2a, b and Extended Data Fig.2a-c). These results demonstrate that LENG8 has the capacity to form condensates in both the nucleus and cytoplasm, exhibiting distinct cell type-specific and developmental stage-dependent patterns.

To assess the dynamic nature of LENG8 granules, we conducted the fluorescence recovery after photobleaching (FRAP) experiments in 293T cells. The fluorescence of EGFP-LENG8 granules showed rapid recovery after photobleaching, suggesting that LENG8 granules possess dynamic and liquid-like properties. (Fig.2e, f). The foci generated by EGFP-LENG8 displayed mobility and underwent spontaneous fusion upon encountering each other (Extended Data Fig.2d). To further test the physical properties of endogenous LENG8 granules, we conducted experiments wherein cells were treated with varying concentrations of 1,6-Hexanediol (1,6-HD), employed for disrupting LLPS- or low-valency interactions with spatially clustered binding sites (ICBS)-mediated structures both *in vitro* and *in vivo*.[40-45] The results revealed a pronounced reduction in LENG8 granule signal intensity within cells, denoting the presence of hydrophobic interactions within the LENG8 granules (Extended Data Fig.2e). Nuclease digestion employing RNase A resulted in the eradication of the LENG8 granule signal, providing evidence for the involvement of RNA in endogenous LENG8 granules assembly (Extended Data Fig.2f). As with most nuclear bodies,

LENG8 granules also experience disassembly and reassembly during mitosis, underscoring their dynamic behavior (Extended Data Fig.2g). We then asked if LENG8 possessed phase separation potential *in vitro*. We found that purified recombinant LENG8 proteins exhibited strong capability for droplet formation (Fig.2g). This capability is concentration-dependent, and further enhanced by the crowding agent PEG6000 (Fig.2g and Extended Data Fig.2h). Increasing the NaCl concentration and treatment with 5 % 1,6-HD can both impede the phase separation of LENG8 (Fig.2g, h). Moreover, the LENG8 droplets exhibited fusion and growth phenotypes (Fig.2i and Supplementary Movie 1). Together, our investigation unveils a hitherto uncharacterized LENG8 nuclear body generated by LLPS.

LENG8 is essential for mouse spermatogenesis and male fertility

To investigate the function of *LENG8* in male germ cells, we generated *LENG8* conditional knockout mice through the breeding of mice carrying the floxed *LENG8* allele with *Stra8-Cre* knock-in transgenic mice,[46] resulting in the deletion of exon 3 and 4 (Fig.3a). Western blot assay showed that the LENG8 protein level was reduced significantly in *LENG8*$^{flox/\Delta}$ *Stra8-Cre* (*LENG8* cKO) mouse testes compared with wildtype mouse (Extended Data Fig.3a). Although *LENG8* cKO male and female mice were viable and copulated normally, both genders were found to be completely infertile. The testis size and weight were severely reduced in *LENG8* cKO male mice compared with wildtype littermates (Fig.3b, c).

Histological analyses revealed that *LENG8* deletion testes displayed spermatogenesis arrest at early meiotic spermatocytes, with a lack of sperm in the epididymis (Fig.3d and Extended Data Fig.3b, c). We conducted single-cell RNA sequencing (scRNA-seq) to further characterize cell types produced in spermatogenesis without *LENG8* using previously determined cell type markers. As a result, we found that *LENG8* deletion mainly causes cells arrest at leptotene/zygotene stage of meiosis (Fig.3e and Extended Data Fig.3d). Corresponding to this, we conducted the TUNEL (Terminal Deoxynucleotidyl Transferase-Mediated dUTP-Biotin Nick End Labeling) assay, a method designed to detect apoptotic cells by labeling the fragmented DNA. The results of the TUNEL assays showed a significant increase in apoptosis within *LENG8* knockout testes (Fig.3f, g). Thus, these studies reveal that ablation of *LENG8* causes impaired meiosis, indicating an important role of LENG8 in spermatogenesis.

Loss of *LENG8* impacts nuclear RNA metabolism in spermatogenesis

In order to probe the biological functions of LENG8 granules in spermatogenesis, we first initiated the identification of its interacting proteins in mouse testes by co-immunoprecipitation (co-IP) in combination with mass spectrometry (Supplementary Table 2). GO analysis revealed that the interacting proteins of LENG8 are involved in processes such as mRNA processing, splicing, suggesting a potential significant role for LENG8 in nuclear RNA metabolism (Extended Data Fig.3e). To investigate the precise molecular consequences of *LENG8* deletion in spermatogenesis, we isolated mRNAs from wildtype and *LENG8* cKO testes and conducted RNA sequencing (RNA-seq). RNA-seq analyses identified 3,452 upregulated and 3,839 downregulated genes

in *LENG8* cKO testes compared to wildtype controls, indicating that the transcription is dramatically altered in *LENG8* cKO testes (Fig.3h and Supplementary Table 3). The results of the functional analysis showed a number of altered processes, with the regulation of transcription being the most significantly affected process (Extended Data Fig.3f).

Taking into account the indirect regulatory effects on transcription and the tight coupling between transcription, splicing and mRNA 3′-end processing,[47] we also examined the influence of *LENG8* depletion on alternative splicing (AS) and alternative polyadenylation (APA), respectively. A total of 1,297 AS events across five distinct AS types showed significant changes in *LENG8* cKO testes (Fig.3i and Supplementary Table 3). We then systematically analyzed the 3′UTR length between wildtype and *LENG8* cKO testes. The alteration in 3′UTR length was quantitatively evaluated through the calculation of the percentage of distal poly (A) site usage index (PDUI), which can identify lengthening or shortening with 3′UTRs. A total of 915 genes (defined as LENG8 targets) exhibited 3′UTR lengthening, whereas 78 genes displayed 3′UTR shortening after the deletion of *LENG8* (Fig.3j and Supplementary Table 3). We selected 3 transcripts with calculated longer 3′UTRs in *LENG8*-deficient testes for RT-PCR validation. These 3 transcripts preferentially generated the longer isoform in *LENG8* cKO testes due to APA (Fig.3k). The small number of overlapping genes between any two of these four datasets implies that they may not have a direct relationship with each other (Extended Data Fig.3g). Unlike what was observed in cultured cells,[29] the extension of 3′UTR did not notably influence mRNA nuclear export in spermatocytes (Extended Data Fig.3h).

Taking collectively, our results indicate that LENG8 plays a crucial role in nuclear RNA metabolism in spermatogenesis, and its depletion leads to spermatogenesis arrest and male infertility.

LENG8 granule functions as an active hub for APA regulation

Given that LENG8 granules we identified primarily exist in two types within the nucleus: large, but limited in number (ranging from 2 to 10), and smaller, but numerous (Fig.2d). Thus, we assessed the effects of *LENG8* ablation in cultured 293T, HeLa, MCF7 cells and found that *LENG8* knockdown could also influence the transcription, and splicing processes (Extended Data Fig.4a-d and Supplementary Table 4). Considering the distinct localization of LENG8 granules in comparison to other nuclear bodies involved in transcription and splicing regulation (Fig.2c), we speculate that LENG8 protein, along with its associated granules may play a regulatory role in APA (Fig.3j, k and Supplementary Table 2). Indeed, *LENG8* depletion in 293T, HeLa, MCF7 cells also led to extensive 3′UTR lengthening (Fig.4a and Supplementary Table 4).

In order to prove that LENG8 granule is an active site for APA regulation, we first investigated the relationship between RNA and LENG8 granules. We employed the EU (5-ethynyl uridine)-labeling method for detecting nascent RNAs, and our findings confirmed the presence of RNAs within LENG8 granules (Fig.4b). We then inhibited transcription using actinomycin D, which led to the disruption of LENG8 granules into

smaller puncta (Fig.4c). RNase A treatment also results in the disassembly of LENG8 granules (Extended Data Fig.2f). To ascertain the specificity of LENG8 granules in regulating RNA metabolism, we utilized specific probes that bind to LENG8 target *Tpi1* mRNAs to detect their localization within the granules. Our results revealed that target *Tpi1* mRNAs are localized within LENG8 granules (Fig.4d). These data imply that mRNAs are localized within LENG8 granules and play a critical role in the formation or/and maintenance of these granules.

Second, we examined whether LENG8 granules function as active sites for APA regulation. To probe the exact role of LENG8 granules on APA regulation, we constructed a tethered reporter by inserting 6×MS2 sequences (a short RNA stem-loop) near the proximal poly(A) site of LENG8 target *Chchd7*. mCherry was fused to an MCP (bacteriophage MS2 coat protein) tag, which specifically binds to MS2 sequences, thereby allowing for the visualization and tracking of mRNA localization (Fig.4e). We first detected the cellular localization of the *Chchd7* reporter mRNAs by co-transfecting with MCP-mCherry, the *Chchd7* reporter, and EGFP-LENG8. We found that *Chchd7* reporter mRNAs also localized to LENG8 granules in 293T cells (Extended Data Fig.4e). Thus, we can more intuitively detect the effect of *LENG8* knockdown on the length of its 3′UTR. The results revealed that *LENG8* depletion induced target 3′UTR lengthening (Fig.4f). To further substantiate that LENG8 granule is important for APA regulation, we screened and identified that amino acids 24-66 of LENG8 protein are critical for granule formation (Extended Data Fig.4f). We compared the length of the reporter 3′UTR in the presence of MCP-LENG8 or MCP-LENG8$^{\Delta 24\text{-}66}$ (a condensation-deficient mutant). The results showed that, compared to MCP-LENG8$^{\Delta 24\text{-}66}$, the tethering of MCP-LENG8 can significantly enhance the selection of proximal poly(A) site in the *Chchd7* reporter (Fig.4g). Furthermore, we employed *LENG8* siRNAs to deplete endogenous LENG8 in 293T cells and then complemented the cells with either wildtype LENG8 or the condensation-deficient mutant LENG8$^{\Delta 24\text{-}66}$. As expected, wildtype LENG8 was able to restore the usage of proximal poly(A) sites in targets, but LENG8$^{\Delta 24\text{-}66}$ mutant failed to rescue, confirming the indispensable role of LENG8 granule in proximal poly(A) sites utilization (Fig.4h).

Third, we investigated the relationship between granule size and number, and APA regulation. Although these differences do not impact LENG8 granules as APA regulatory hubs, they lead to distinct profiles of regulated genes in 293T, HeLa, and MCF7 cells (Fig.4i). Then, we compared the degree of PDUI changes in LENG8 targets across these cell lines after *LENG8* knockdown relative to control, which can serve as an indicator of the efficiency of LENG8 granule-mediated regulation in the usage of proximal poly(A) sites. Our results suggested that larger LENG8 granules regulate the usage of proximal poly(A) sites with higher efficiency (Extended Data Fig.4g). An increasing number of studies have suggested that RNA contributes to the formation of biomolecular condensates.[48] Thus, we examined the abundance of LENG8 targets mRNA in the nucleus of these cells. We found that the abundance of LENG8 target mRNAs in the nucleus of 293T and HeLa cells was significantly higher than that in MCF7 (Extended Data Fig.4h). This suggests that the higher abundance of mRNAs that

need to be processed requires the recruitment of more LENG8, which in turn contributes to the formation of larger granules. This is in line with the results that actinomycin D and RNase A treatment can both lead to the disruption of LENG8 granules into smaller puncta (Fig.4c and Extended Data Fig.2f). Collectively, these data demonstrate that LENG8 granule functions as an active hub for promoting the usage of pPASs in target transcripts.

LENG8 granules orchestrate CPA complex assembly and crosstalk with nuclear speckles

To investigate how LENG8 regulates pPAS usage, we first examined the interaction of LENG8 with CPSF4, FIP1L1, and CPSF2 in cells forming prominent, large granules (293T, HeLa, U2OS, C2C12) and in cells containing numerous smaller condensates (MCF7, HCT116). Results revealed that LENG8 can interact with these proteins, in both types of cells, implying that these interactions are independent of the size and number of LENG8 granules (Fig.5a). We also validated the interaction between LENG8 and these CPA factors by co-transfected Myc-LENG8 and EGFP-CPA factors into 293T cells. Our results demonstrated that Myc-LENG8 efficiently pulled down these EGFP-CPA factors (Fig.5b). CPSF4, a core CPSF complex component, directly interacts with PAS in conjunction with WDR33 to determine RNA cleavage sites.[14,49-51] *In vitro* pull-down assay demonstrated that there is a direct interaction between LENG8 and CPSF4 (Fig.5c). FIP1L1, another essential constituent of the CPSF complex, binds to U/AU-rich regions between the PAS and the cleavage site, stimulating poly (A) polymerase activity and facilitating of pPAS usage.[52-56] Considering the pivotal role of the CPSF complex assembly in 3′ end CPA processes, we studied the involvement of LENG8 in the interaction between CPSF4 and FIP1L1. We found that knocking down endogenous *LENG8* significantly reduced the interaction between CPSF4 and FIP1L1(Fig.5d). IF data revealed that CPSF4 assembles into small foci which are embedded within LENG8 granules (Fig.5e). Consistent with these findings, the interaction between LENG8$^{\Delta 24\text{-}66}$ and these CPA factors is notably weaker than that of wildtype LENG8 (Fig.5f). Furthermore, CPSF6, a core component of the CFI complex within the CPA machinery, [5,9] also exhibits localization to LENG8 granules (Fig.5g). These data establish LENG8 granules as critical storage reservoirs for CPA factors and demonstrate their essential role in orchestrating CPA complex assembly.

PABPN1 is involved in numerous RNA metabolism processes, including poly (A) tail length regulation, mRNA transportation, facilitation of splicing for weak 3′ splice sites in last introns, maternal mRNA decay, and translation.[57-63] PABPN1 can prevent the binding of the CPSF complex to pPASs and enhance the formation of long 3′UTR transcripts.[20,64,65] Considering the opposing roles of PABPN1 and LENG8 in poly (A) site selection, we used the STRING protein-protein interaction database to speculate on the interaction between PABPN1 and CPSF complex factors, suggesting that PABPN1 might associate with CPSF complex via CPSF4 (Extended Data Fig.5a). This is consistent with a recent report that identified an interaction between PABPN1 and CPSF4.[20] To validate the interaction between PABPN1 and CPSF complex factors, we co-expressed Flag-PABPN1 along with EGFP-CPA factors in 293T cells for co-IP.

Western blot analysis showed that PABPN1 exclusively interacted with CPSF4, unlike LENG8 which can bind to several CPA factors (Extended Data Fig.5b). Furthermore, we observed that both LENG8 and PABPN1 interacted with CPSF4 in both 293T cells and testes (Extended Data Fig.5c, d). IF data revealed distinct patterns of condensation for LENG8 and PABPN1(a component of nuclear speckles). However, there exists partial or docking-like co-localization patterns (51% in 293T, 43% in HeLa) between LENG8 granules and nuclear speckles, implying a functional connection between them (Extended Data Fig.5e). Reducing *PABPN1* expression significantly enhanced the interaction between LENG8 and CPSF4, without overtly affecting the formation/ maintenance of LENG8 granules (Fig.5h and Extended Data Fig.5f). The restoration of PABPN1 expression was able to suppress the interaction between them (Extended Data Fig.5g). Of note, high-resolution imaging data demonstrate that CPSF4 localizes to both nuclear speckles and LENG8 granules, with striking accumulation at the docking sites between these two biomolecular condensates (Fig.5i). To investigate the competitive binding mechanism, we screened the CPSF4 regions responsible for binding to both LENG8 and PABPN1, and identified the fourth zinc finger (ZF4) domain as crucial for their interaction (Extended Data Fig.5h). This suggests that ZF4 domain is a key region for the competitive relationship between LENG8 and PABPN1. These findings unveil that LENG8 competes with PABPN1 for binding or directly binds to CPSF4, thereby facilitating the assembly of the CPA complex at the pPASs.

To further investigate the functional interplay between LENG8 granules and nuclear speckles, we performed double knockdown of SRRM2 and SON to disrupt nuclear speckles, triggering LENG8 granules disassembly into smaller granules (Fig.5j). Therefore, we study the relationship between PABPN1-regulated targets and LENG8 granule-regulated targets. Our data also demonstrated that *PABPN1* downregulation caused 3′UTR shortening (Extended Data Fig.5i). Indeed, LENG8 and PABPN1 can modulate poly (A) sites selection of a common set of genes both in cultured cells and testes (Extended Data Fig.5j, k). Meanwhile, LENG8 can also independently regulate pPAS selection of a group of genes (Extended Data Fig.5j, k).

In conclusion, these findings indicate that LENG8 granules exert their functions both autonomously and cooperatively with nuclear speckles to regulate mRNA 3' end processing.

LENG8 promotes proximal poly (A) site selection in a YTHDC1-m^{6}A dependent manner

To characterize the targets regulated by LENG8, we analyzed the features of flanking regions of proximal poly (A) signals. Notably, the majority of LENG8 targets contain a typical m^{6}A consensus sequence in both testicular and 293T cells (Fig.6a). Indeed, LENG8 targets exhibit a higher m^{6}A signals near the pPASs compared to random transcripts (Extended Data Fig.6a, b). These results are in agreement with the observation that transcripts with m^{6}A modifications display a tendency for utilizing pPASs.[66] A growing body of evidence suggests that YTHDC1, a nuclear m^{6}A binding protein, is involved in regulating mRNA alternative splicing, mediating mRNA

transport from the nucleus to the cytoplasm, influencing mRNA stability, facilitating carRNA decay, silencing retrotransposons, regulating the scaffold function of LINE1 RNA, and modulating APA.[67-75] Both *YTHDC1* knockout mouse oocytes and *YTHDC1* knockdown cultured cells exhibited alterations in the 3′UTR length of hundreds of mRNAs, suggesting that YTHDC1 may exert bidirectional control over APA regulation.[69,75] However, the molecular mechanism by which YTHDC1 regulates APA remains ambiguous. According to our mass spectrometry data, we detected a potential interaction between YTHDC1 and LENG8, prompting us to investigate their association in APA regulation. Co-IP assays in several cell types validated their interaction (Fig.5a). *In vitro* pull-down assay demonstrated that there is a direct interaction between LENG8 and YTHDC1 (Fig.6b). Using co-immunostaining of LENG8 and YTHDC1, we observed that some YTHDC1 small puncta were found to be dotted within the LENG8 granules in 293T and HeLa cells, indicating that YTHDC1 is also a component protein of LENG8 granules (Fig.6c). A recent study has found that YTHDC1 can undergo LLPS to form numerous small nuclear puncta, with m^6A binding being required for this process.[72] Based on this, we sought to determine whether the formation and function of LENG8 granules depend on YTHDC1. To address this question, we silenced *YTHDC1* in 293T cells and observed the dissolution of LENG8 granules (Fig.6d). The absence of *YTHDC1* resulted in either a shortened or extended 3′UTR length, consistent with previous reports (Fig.6e).[69,75] However, most of the genes tended to have longer 3′UTRs under our experimental conditions. Both the overall transcriptome and LENG8 targets displayed a trend of selecting dPASs upon *YTHDC1* knockdown (Fig.6f, g). The majority of LENG8 targets also overlapped with those targets bound by YTHDC1 (Fig.6h). Moreover, a similar scenario was observed in testes (Extended Data Fig.6c).

Deletion of *METTL3*, an m^6A writer, leads to significant m^6A loss in germ cells and influences spermatogenesis.[46,76] Analysis of transcriptome data from wildtype and *METTL3* cKO testes revealed that *METTL3* deletion also increased 3′UTR length (Extended Data Fig.6d). Moreover, both the overall transcriptome and LENG8 targets exhibit a tendency to preferentially select dPASs upon *METTL3* knockout (Extended Data Fig.6e, f). Meanwhile, the loss of function of *ALKBH5*, an m^6A eraser, impacts proper splicing and causes a significant decrease in transcript length, ultimately leading to male infertility in mice.[77] An opposite pattern was identified, with transcripts displaying shortened 3′UTR upon *ALKBH5* knockout and extended 3′UTR upon *LENG8* absence (Extended Data Fig.6g, h). Furthermore, we carried out double knockdown experiments on 293T cells to verify whether this is only a correlation or a direct effect. Compared with the knockdown of *LENG8* alone, a total of 1,671 genes exhibited 3′UTR lengthening, whereas 591 genes displayed 3′UTR shortening after the double knockdown of *LENG8* and *METTL3* (Fig.6i). Notably, 592 out of 1,451 LENG8 targets exhibited further 3′UTR lengthening (Fig.6j). These results demonstrate that *METTL3* knockdown further intensifies the decrease in the utilization of proximal poly(A) sites caused by *LENG8* knockdown. Meanwhile, a total of 1,284 genes exhibited 3′UTR shortening, whereas 187 genes displayed 3′UTR lengthening after the double knockdown of *LENG8* and *ALKBH5* (Fig.6k). 574 out of 1,451 LENG8 targets

exhibited 3′UTR shortening (Fig.6l). These data indicate that *ALKBH5* knockdown can attenuate the decrease in the utilization of proximal poly(A) sites caused by *LENG8* knockdown. These data demonstrate that LENG8/METTL3/ALKBH5 can regulate the APA process within the same m^{6}A-dependent pathway. In contrast to the multifaceted functions of METTL3 and ALKBH5 in RNA metabolism, LENG8 might have a more specific regulatory function for APA.

Taken together, these results demonstrate that YTHDC1-m^{6}A landscapes plays crucial regulatory roles in LENG8 granule formation and the selection of targets for proximal poly (A) site usage (Fig.6m).

LENG8 granules regulate APA in a cell-type specific and developmental stage-dependent manner

APA is highly regulated in a tissue/cell-type- and/or developmental stage-specific manner.[5] A systematic comparison of mRNAs expressed across different tissues revealed distinct utilization of pPASs and dPASs. The brain exhibits a propensity for the elongated 3′UTR isoforms, whereas the testis shows extensive 3′ end shortening (Extended Data Fig.7a), consistent with previous studies.[78-85] Interestingly, the targets with extended 3′UTR lengths resulting from *LENG8* deletion in testes display a more prominent preference for pPAS selection compared to other tissues, in concordance with the intracellular localization pattern of LENG8 (Fig.2a and Extended Data Fig.2a-c, 7b). These findings promoted us to hypothesize a potential correlation between LENG8 localization, LENG8 granule assembly, and pPAS selection. By comparing poly (A) site selection in germ cells at different stages, we found a progressive shortening of 3′UTR length in transcripts from spermatogonia to spermatocytes then to round spermatids (Fig.7a). As expected, the targets of LENG8 exhibit more pronounced alteration in 3′UTR length, with the shortest isoforms in round spermatids (Fig.7b).

While previous studies have noted this trend of 3′UTR shortening in testes,[86,87] the underlying regulatory mechanism remains largely unknown. Therefore, we focused on spermatocytes and round spermatids to investigate the mechanisms of LENG8 granules in pPAS selection. We observed that LENG8, CPSF4, FIP1L1, YTHDC1 and PABPN1 transition from multiple small puncta in spermatocytes to a limited number of larger granules in round spermatids (Fig.7c-f). LENG8 exhibits partial or docking-like co-localization patterns with CPSF4 and FIP1L1, especially in round spermatids, indicating an increased propensity for CPA complex assembly and pPAS selection (Fig.7c, d). Compared to LENG8, the co-localization of PABPN1 and CPSF4 in both spermatocytes and round spermatids is more pronounced (Extended Data Fig.7c), indicating that CPSF4 is also a component of nuclear speckles in germ cells. Remarkably, the strong co-localization of YTHDC1 in LENG8 granules in both spermatocytes and round spermatids implies that YTHDC1 is also a component of LENG8 granules in germ cells (Fig.7e). These results reveal that LENG8 granules in germ cells are also capable of regulating the APA process and engaging in cross-talk with nuclear speckles, akin to that in cultured cells (Fig.7g and Fig.5).

Transcripts with diverse 3′UTR lengths through APA harbor specific cis-regulatory

elements, such as microRNA-binding sites, transposable elements, and RBP binding sites, which affect mRNA metabolism.[3] Thus, we assessed the impact of LENG8 granules on the fate of target mRNAs. We initially analyzed the effect of *LENG8* deletion on protein levels by mass spectrometry, identifying 146 proteins with elevated levels and 161 proteins with reduced levels among those detected (Extended Data Fig.7d and Supplementary Table 5). However, the protein levels of LENG8 targets remained largely unaltered (Extended Data Fig.7d). Integrated analysis of RNA and protein expression data suggested that LENG8 has a minor influence on the RNA levels and translation of its targets (Extended Data Fig.7e). Considering the indirect effects of developmental arrest caused by *LENG8* knockout, we evaluated the expression of its targets during normal spermatogenesis. RNA-seq and mass spectrometry data from purified germ cells were used to analyze the expression alteration of LENG8 targets. Based on the change characteristics in RNA and protein levels from spermatogonia to spermatocytes then to round spermatids, LENG8 targets can be categorized into four main clusters: cluster 1, exhibiting a negative correlation between protein and RNA levels; cluster 2, showing a positive correlation between protein and RNA levels with a gradual increase during development; cluster 3, displaying a positive correlation between protein and RNA levels with opposite trends before and after spermatocytes; cluster 4, demonstrating a positive correlation between protein and RNA levels with a gradual decrease during development (Fig.7h). Bioinformatic analysis of hexamers as potential cis-regulatory elements in the sequences between the pPAS and dPAS of LENG8 targets revealed different cis-regulatory elements in these four clusters (Extended Data Fig.7f, g). Cluster 2 and 4 were selected as examples due to their opposite expression trends (Fig.7h). It is intriguing to note that cluster 4 exhibits a significant enrichment of AU-rich elements (AREs), such as AUUUAA, AUUUUU, UUUUUA (Extended Data Fig.7g, h). HuR, a canonical ARE-binding protein, is recognized for a positive role in mRNA stability and translation in spermatogenesis.[88-90] Concomitant with 3′UTR shortening, the loss of HuR-binding sites results in destabilization of the target mRNAs, leading to decreased protein output. Conversely, GCUG motif in cluster 2 targets is likely involved in diminishing target stability (Extended Data Fig.7g, h). As the 3′UTR undergoes shortening, the elimination of these elements leads to increased targets stability and elevated protein yield. Ribosome profiling analysis also revealed an elevation in mRNA levels and translation efficiency of cluster 2 targets compared to cluster 4 targets during spermatogenesis (Extended Data Fig.7i). However, we cannot exclude the impact of transcriptional regulation or other factors on the expression of these targets. The synergistic or antagonistic relationship between these cis-regulatory elements and 3′UTR shortening in regulating the fate of target RNAs required further investment. Taken together, these data support the notion that the YTHDC1-m^{6}A-dependent regulation of target 3′UTR shortening by LENG8 granules collaborates with other regulatory mechanisms to finely tune the fate of these target RNAs, which are essential for spermatogenesis.

Discussion

Membraneless organelles formed through LLPS are implicated in gene regulation. Recent studies highlight the involvement of condensates formed by CPA factors in APA regulation. However, it's unclear if there are dedicated nuclear bodies for APA regulation in mammalian cells, and the mechanisms of their regulatory effects remain unknown. Our multiple lines of evidence demonstrate that a poorly understood protein LENG8 forms dynamic liquid-like nuclear granules by LLPS. Here, we designate it as the LENG8 granule (G granule), setting it apart from other nuclear bodies. We found that LENG8 granule functions as an active hub for promoting the usage of pPASs in target transcripts.

In this study, we elucidate the function of LENG8 granules in APA regulation. Deletion of *LENG8* results in a notable shift to dPAS usage in numerous genes in both cultured cells and male mouse germ cells, suggesting its conserved function as a 3′ end processing factor. LENG8 granules recruit CPA factors and promote their local assembly, thereby favoring the utilization of pPASs for a specific group of genes (Fig.7i). LENG8 can also compete with PABPN1 for binding to CPSF4, thereby promoting the assembly of the CPSF complex at the pPASs. Decreasing PABPN1 expression significantly strengthens the interaction between LENG8 and CPSF4, with no discernible impact on the formation or stability of LENG8 granules. Indeed, LENG8 and PABPN1 can modulate poly (A) sites selection of a common set of genes both in 293T and testes. Nuclear speckles are phase-separated biomolecular condensates that orchestrate RNA transcription, splicing, APA and other RNA metabolic pathways.[91,92,93] *LENG8* ablation in cultured cells and germ cells could also influence the transcription and splicing processes. We found that there are partial or docking-like co-localization patterns between LENG8 granules and nuclear speckles, which may provide an interface for processing these common targets. This docking-type colocalization, similar to the observations of MIWI2 granules and MILI granules in male mouse germ cells,[94] P-bodies and transport ribonucleoprotein particles in dendrites of mature hippocampal neurons,[95] TDRD7 granules and P-bodies in the lens,[96] Chromatoid bodies and eIF3f granules in male mouse germ cells,[90] and stress granules and P-bodies in cells under stress conditions,[97] may facilitate the exchange of protein and RNA components between these two biomolecular condensates. Moreover, the non-overlapping targets between LENG8 and PABPN1 indicate that LENG8 granules can also independently regulate the pPAS usage of this subset of targets, leading to the generation of short isoforms.

How are the targets regulated by LENG8 granules selected? Our findings indicate that YTHDC1-m^{6}A landscapes is of utmost importance in shaping LENG8 granules and determining targets. YTHDC1 is a component protein of LENG8 granules and its deletion induces the dissolution of these granules, along with 3′UTR lengthening. Meanwhile, interfering with m^{6}A writing or erasing can also have an effect on LENG8's regulation of target 3′UTR length. The heterogeneity of LENG8 targets among different cell types could be linked to cell type-specific m^{6}A landscape. Additionally, it is

plausible to infer that YTHDC1-m^{6}A facilitates the recruitment of LENG8 to targets, as well as the ensuing granule formation mediated by LENG8 LLPS. However, the mechanisms underlying LENG8 granule formation require further in-depth investigation.

Despite the consistent findings of a progressive decrease in 3′UTR length during spermatogenesis in previous studies,[86,87,98] the precise significance and mechanisms of the shift towards shorter 3′UTR expression have remained largely elusive. Our observations reveal a sequential and dynamic shift in the cellular positioning of LENG8, transitioning from cytoplasmic localization in spermatogonia to the formation of LENG8 granules in spermatocytes and round spermatids, concomitant with a stepwise reduction in the target 3′UTR length. Conditional deletion of *LENG8* using *Stra8-Cre* results in the blockade of spermatogenesis at the spermatocyte stage, impeding the assessment of 3′UTR length changes in round spermatids. Therefore, the targets identified in this study may only represent a subset of genes modulated by LENG8 granules.

The relationship between LENG8 granule properties (e.g., size and number) and the pPAS selection has also been studied. Our results indicated that LENG8 granules of varying sizes and numbers in cells can all fulfill their roles in APA regulation, yet with distinct profiles of regulated genes across cells. Similarly, biomolecular condensates such as stress granules, nuclear speckles, and P-bodies, can also execute their respective functions irrespective of their sizes and numbers within cells.[99] At present, it is challenging to clearly understand the relationship between the structure and function of biomolecular condensates due to their number, size, dynamics and heterogeneity.[99-101] However, we found that larger LENG8 granules regulate the usage of proximal poly(A) sites with higher efficiency, which is consistent with the results in spermatogenesis. In addition, the abundance of LENG8 target mRNAs in the nucleus of cells with larger granules was higher than that in cells with smaller granules, which suggests that the higher abundance of mRNAs contributes to the formation of larger granules. This aligns with the findings that both actinomycin D and RNase A treatment can lead to the disruption of LENG8 granules into smaller puncta. However, whether other properties of RNAs, including the composition, length, structure, modifications, affect the biophysical features of LENG8 granules needs to be tested in the further. Furthermore, we observed that only the LENG8 protein was significantly enriched in LENG8 granules, while other proteins such CPSF4 and YTHDC1 showed no remarkable enrichment. Based on these findings, we speculate that the LENG8 protein may function as a scaffold protein critical for the formation and maintenance of LENG8 granules. This could also explain why LENG8 granules have remained undiscovered for such a long time.

Overall, we unveil a newly discovered 3′ end processing factor and shed light on the function of a previously unrecognized nuclear body in APA regulation. It is reasonable to speculate that LENG8 plays essential roles in the nucleus, with its functions being executed and finely tuned through LENG8 granules assembly, properties and functional interplay with nuclear speckles in a cell-type specific and developmental stage-

dependent manner.

References

1. Tian, B., and Manley, J. L. Alternative polyadenylation of mRNA precursors. Nature reviews. *Nat. Rev. Mol. Cell Biol*. **18**, 18–30 (2017).
2. Mayr C. Regulation by 3'-Untranslated Regions. *Annu. Rev. Genet*. **51**, 171–194 (2017).
3. Mayr C. What Are 3' UTRs Doing? *Cold Spring Harb. Perspect. Biol*. **11**, a034728 (2019).
4. Chen, X., and Mayr, C. A working model for condensate RNA-binding proteins as matchmakers for protein complex assembly. *RNA* **28**, 76–87 (2022).
5. Gallicchio, L., Olivares, G. H., Berry, C. W., and Fuller, M. T. Regulation and function of alternative polyadenylation in development and differentiation. *RNA Biol*. **20**, 908–925 (2023).
6. Derti, A., Garrett-Engele, P., Macisaac, K. D., Stevens, R. C., Sriram, S., Chen, R., Rohl, C. A., Johnson, J. M., and Babak, T. A quantitative atlas of polyadenylation in five mammals. *Genome Res*. **22**, 1173–1183 (2012).
7. Shi Y. Alternative polyadenylation: new insights from global analyses. *RNA* **18**, 2105–2117 (2012).
8. Lianoglou, S., Garg, V., Yang, J. L., Leslie, C. S., and Mayr, C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev*. **27**, 2380–2396 (2013).
9. Gruber, A. J., and Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet*. **20**, 599–614 (2019).
10. MacDonald C. C. Tissue-specific mechanisms of alternative polyadenylation: Testis, brain, and beyond (2018 update). *Wiley Interdiscip. Rev. RNA* **10**, e1526 (2019).
11. Fansler, M. M., Zhen, G. and Mayr, C. Quantification of alternative 3′UTR isoforms from single cell RNA-seq data with scUTRquant. Preprint at bioRxiv, 10.1101/2021.11.22.469635 (2021).
12. Shi, Y., Di Giammartino, D. C., Taylor, D., Sarkeshik, A., Rice, W. J., Yates, J. R., 3rd, Frank, J., and Manley, J. L. Molecular architecture of the human pre-mRNA 3' processing complex. *Mol. Cell* **33**, 365–376 (2009).
13. Schönemann, L., Kühn, U., Martin, G., Schäfer, P., Gruber, A. R., Keller, W., Zavolan, M., and Wahle, E. Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev*. **28**, 2381–2393 (2014).
14. Chan, S. L., Huppertz, I., Yao, C., Weng, L., Moresco, J. J., Yates, J. R., 3rd, Ule, J., Manley, J. L., and Shi, Y. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev*. **28**, 2370-2380 (2014).
15. Bogard, N., Linder, J., Rosenberg, A. B., and Seelig, G. A Deep Neural Network for Predicting and Engineering Alternative Polyadenylation. *Cell* **178**, 91–106 (2019).
16. Liu, H., and Moore, C. L. On the Cutting Edge: Regulation and Therapeutic Potential of the mRNA 3' End Nuclease. *Trends Biochem. Sci*. **46**, 772–784 (2021).
17. Mitra, M., Johnson, E. L., Swamy, V. S., Nersesian, L. E., Corney, D. C., Robinson,

D. G., Taylor, D. G., Ambrus, A. M., Jelinek, D., Wang, W., Batista, S. L., and Coller, H. A. Alternative polyadenylation factors link cell cycle to migration. *Genome Biol*. **19**, 176 (2018).

18. Tan, S., Zhang, M., Shi, X., Ding, K., Zhao, Q., Guo, Q., Wang, H., Wu, Z., Kang, Y., Zhu, T., Sun, J., and Zhao, X. CPSF6 links alternative polyadenylation to metabolism adaption in hepatocellular carcinoma progression. *J. Exp. Clin. Cancer Res*. **40**, 85 (2021).
19. Liu, S., Wu, R., Chen, L., Deng, K., Ou, X., Lu, X., Li, M., Liu, C., Chen, S., Fu, Y., and Xu, A. CPSF6 regulates alternative polyadenylation and proliferation of cancer cells through phase separation. *Cell Rep*. **42**, 113197 (2023).
20. Dai, X. X., Pi, S. B., Zhao, L. W., Wu, Y. W., Shen, J. L., Zhang, S. Y., Sha, Q. Q., and Fan, H. Y. PABPN1 functions as a hub in the assembly of nuclear poly(A) domains that are essential for mouse oocyte development. *Sci Adv*. **8**, eabn9016 (2022).
21. Hu, Z., Li, M., Huo, Z., Chen, L., Liu, S., Deng, K., Lu, X., Chen, S., Fu, Y., and Xu, A. U1 snRNP proteins promote proximal alternative polyadenylation sites by directly interacting with 3' end processing core factors. *J. Mol. Cell Biol*. **14**, mjac054 (2022).
22. Fang, X., Wang, L., Ishikawa, R., Li, Y., Fiedler, M., Liu, F., Calder, G., Rowan, B., Weigel, D., Li, P., and Dean, C. Arabidopsis FLL2 promotes liquid-liquid phase separation of polyadenylation complexes. *Nature* **569**, 265–269 (2019).
23. Wende, H., Volz, A., and Ziegler, A. Extensive gene duplications and a large inversion characterize the human leukocyte receptor cluster. *Immunogenetics*, **51**, 703–713 (2000).
24. Baillat, D., Hakimi, M. A., Näär, A. M., Shilatifard, A., Cooch, N., and Shiekhattar, R. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265–276 (2005).
25. Gaur, U., Xiong, Y. Y., Luo, Q. P., Yuan, F. Y., Wu, H. Y., Qiao, M., Wimmers, K., Li, K., Mei, S. Q., and Liu, G. S. Breed-specific transcriptome response of spleen from six-to eight-week-old piglet after infection with Streptococcus suis type 2. *Mol. Biol. Rep*. **41**, 7865–7873 (2014).
26. Ye, X., Zhang, Y., He, B., Meng, Y., Li, Y., and Gao, Y. Quantitative proteomic analysis identifies new effectors of FOXM1 involved in breast cancer cell migration. *Int. J. Clin. Exp. Pathol*. **8**, 15836–15844 (2015).
27. Farias, T., Augusto, D. G., de Almeida, R. C., Malheiros, D., and Petzl-Erler, M. L. Screening the full leucocyte receptor complex genomic region revealed associations with pemphigus that might be explained by gene regulation. *Immunology*, **156**, 86–93 (2019).
28. Song, J., Liu, Y. D., Su, J., Yuan, D., Sun, F., and Zhu, J. Systematic analysis of alternative splicing signature unveils prognostic predictor for kidney renal clear cell carcinoma. *J. Cell Physiol*. **234**, 22753–22764 (2019).
29. Zhao, Y.X., Wang, X.T., Liu, Y.N., Li, N.N., Wang, S.M., Sun, Z.G., Gao, Z.F., Zhang, X.X., Mao, L.F., Tang, R., Xue, W.Y., Li, C.Y., Guan, J., Yi, H.L., Zhang, N., Ding, Q.R., Liu, F. LENG8 regulation of mRNA processing, is responsible for

the control of mitochondrial activity. Preprint at bioRxiv, 10.1101/2021.07.17.452750 (2021).

30. Gomes, E., and Shorter, J. The molecular language of membraneless organelles. *J. Biol. Chem*. **294**, 7115–7127 (2019).
31. Dodson, A. E., and Kennedy, S. Phase Separation in Germ Cells and Development. *Dev. Cell* **55**, 4–17 (2020).
32. Ripin, N., and Parker, R. Formation, function, and pathology of RNP granules. *Cell* **186**, 4737–4756 (2023).
33. Hirose, T., Ninomiya, K., Nakagawa, S., and Yamazaki, T. A guide to membraneless organelles and their various roles in gene regulation. *Nat. Rev. Mol. Cell Biol*. **24**, 288–304 (2023).
34. Griswold M. D. Spermatogenesis: The Commitment to Meiosis. *Physiological reviews*, **96**, 1–17 (2016).
35. Lehtiniemi, T., & Kotaja, N. Germ granule-mediated RNA regulation in male germ cells. *Reproduction* **155**, R77–R91 (2018).
36. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R., and McKnight, S. L. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
37. Zambrano, R., Conchillo-Sole, O., Iglesias, V., Illa, R., Rousseau, F., Schymkowitz, J., Sabate, R., Daura, X., and Ventura, S. PrionW: a server to identify proteins containing glutamine/asparagine rich prion-like domains and their amyloid cores. *Nucleic Acids Res*. **43**, W331–W337 (2015).
38. Sabate, R., Rousseau, F., Schymkowitz, J., and Ventura, S. What makes a protein sequence a prion?. *PLoS Comput. Biol*. **11**, e1004013 (2015).
39. Mittag, T., and Parker, R. Multiple Modes of Protein-Protein Interactions Promote RNP Granule Assembly. *J. Mol. Biol*. **430**, 4636–4649 (2018).
40. Ribbeck, K., and Görlich, D. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J*. **21**, 2664–2671(2002).
41. Patel, S. S., Belmont, B. J., Sante, J. M., and Rexach, M. F. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* **129**, 83–96 (2007).
42. Kroschwald, S., Maharana, S., Mateju, D., Malinovska, L., Nüske, E., Poser, I., Richter, D., and Alberti, S. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *eLife*, **4**, e06807 (2015).
43. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T., and Taylor, J. P. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
44. Wheeler, J. R., Matheny, T., Jain, S., Abrisch, R., and Parker, R. Distinct stages in stress granule assembly and disassembly. *eLife* **5**, e18413 (2016).
45. Muzzopappa, F., Hummert, J., Anfossi, M., Tashev, S. A., Herten, D. P., & Erdel, F. Detecting and quantifying liquid-liquid phase separation in living cells by model-

free calibrated half-bleaching. *Nature communications* **13**, 7787 (2022).

46. Lin, Z., Hsu, P. J., Xing, X., Fang, J., Lu, Z., Zou, Q., Zhang, K. J., Zhang, X., Zhou, Y., Zhang, T., Zhang, Y., Song, W., Jia, G., Yang, X., He, C., and Tong, M. H. Mettl3-/Mettl14-mediated mRNA N6-methyladenosine modulates murine spermatogenesis. *Cell Res*. **27**, 1216–1230 (2017).
47. Mitschka, S., and Mayr, C. Context-specific regulation and function of mRNA alternative polyadenylation. *Nat. Rev. Mol. Cell Biol*. **23**, 779–796 (2022).
48. Roden, C., & Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol*. **22**, 183–195 (2021).
49. Shimberg, G. D., Michalek, J. L., Oluyadi, A. A., Rodrigues, A. V., Zucconi, B. E., Neu, H. M., Ghosh, S., Sureschandra, K., Wilson, G. M., Stemmler, T. L., and Michel, S. L. Cleavage and polyadenylation specificity factor 30: An RNA-binding zinc-finger protein with an unexpected 2Fe-2S cluster. *Proc. Natl. Acad. Sci. USA* **113**, 4700–4705 (2016).
50. Clerici, M., Faini, M., Muckenfuss, L. M., Aebersold, R., and Jinek, M. Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat. Struct. Mol. Biol*. **25**, 135–138 (2018).
51. Sun, Y., Zhang, Y., Hamilton, K., Manley, J. L., Shi, Y., Walz, T., and Tong, L. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proc. Natl. Acad. Sci. USA* **115**, E1419–E1428 (2018).
52. Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and Keller, W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J*. **23**, 616–626 (2004).
53. Martin, G., Gruber, A. R., Keller, W., and Zavolan, M. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep*. **1**, 753–763 (2012).
54. Lackford, B., Yao, C., Charles, G. M., Weng, L., Zheng, X., Choi, E. A., Xie, X., Wan, J., Xing, Y., Freudenberg, J. M., Yang, P., Jothi, R., Hu, G., and Shi, Y. Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J*. **33**, 878–889 (2014).
55. Li, W., You, B., Hoque, M., Zheng, D., Luo, W., Ji, Z., Park, J. Y., Gunderson, S. I., Kalsotra, A., Manley, J. L., and Tian, B. Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet*. **11**, e1005166 (2015).
56. Davis, A. G., Johnson, D. T., Zheng, D., Wang, R., Jayne, N. D., Liu, M., Shin, J., Wang, L., Stoner, S. A., Zhou, J. H., Ball, E. D., Tian, B., and Zhang, D. E. Alternative polyadenylation dysregulation contributes to the differentiation block of acute myeloid leukemia. *Blood* **139**, 424–438 (2022).
57. Wahle E. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* **66**, 759–768 (1991).
58. Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J*. **22**, 3705–3714 (2003).

59. Kühn, U., Gündel, M., Knoth, A., Kerwitz, Y., Rüdel, S., and Wahle, E. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *The Journal of biological chemistry*, **284**, 22803–22814 (2009).
60. Shi, M., Zhang, H., Wu, X., He, Z., Wang, L., Yin, S., Tian, B., Li, G., and Cheng, H. ALYREF mainly binds to the 5' and the 3' regions of the mRNA in vivo. *Nucleic Acids Res*. **45**, 9640–9653 (2017).
61. Mohibi, S., Zhang, J., and Chen, X. PABPN1, a Target of p63, Modulates Keratinocyte Differentiation through Regulation of p63α mRNA Translation. *The Journal of investigative dermatology*, **140**, 2166–2177.e6 (2020).
62. Zhao, L. W., Zhu, Y. Z., Wu, Y. W., Pi, S. B., Shen, L., and Fan, H. Y. Nuclear poly(A) binding protein 1 (PABPN1) mediates zygotic genome activation-dependent maternal mRNA clearance during mouse early embryonic development. *Nucleic Acids Res*. **50**, 458–472 (2022).
63. Huang, L., Li, G., Du, C., Jia, Y., Yang, J., Fan, W., Xu, Y. Z., Cheng, H., and Zhou, Y. The polyA tail facilitates splicing of last introns with weak 3' splice sites via PABPN1. *EMBO Rep*. **24**, e57128 (2023).
64. Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kühn, U., Menzies, F. M., Oude Vrielink, J. A., Bos, A. J., Drost, J., Rooijers, K., Rubinsztein, D. C., and Agami, R. The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**, 538–553 (2012).
65. Chen, L., Dong, W., Zhou, M., Yang, C., Xiong, M., Kazobinka, G., Chen, Z., Xing, Y., and Hou, T. PABPN1 regulates mRNA alternative polyadenylation to inhibit bladder cancer progression. *Cell Biosci*. **13**, 45 (2023).
66. Molinie, B., Wang, J., Lim, K. S., Hillebrand, R., Lu, Z. X., Van Wittenberghe, N., Howard, B. D., Daneshvar, K., Mullen, A. C., Dedon, P., Xing, Y., and Giallourakis, C. C. m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. *Nat. Methods* **13**, 692–698 (2016).
67. Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., Sun, H. Y., Li, A., Ping, X. L., Lai, W. Y., Wang, X., Ma, H. L., Huang, C. M., Yang, Y., Huang, N., Jiang, G. B., et al. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell* **61**, 507–519 (2016).
68. Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P., He, E., Shen, B., and He, C. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *eLife* **6**, e31311 (2017).
69. Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., Schultz, R. M., and Wang, P. J. Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet*. **14**, e1007412 (2018).
70. Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M. M., Zhao, S., Shen, B., Gao, Y., Han, D., and He, C. N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580–586 (2020).
71. Chen, C., Liu, W., Guo, J., Liu, Y., Liu, X., Liu, J., Dou, X., Le, R., Huang, Y., Li, C., Yang, L., Kou, X., Zhao, Y., Wu, Y., Chen, J., Wang, H., Shen, B., Gao, Y., and

Gao, S. Nuclear m6A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. *Protein Cell* **12**, 455–474 (2021).

72. Cheng, Y., Xie, W., Pickering, B. F., Chu, K. L., Savino, A. M., Yang, X., Luo, H., Nguyen, D. T., Mo, S., Barin, E., Velleca, A., Rohwetter, T. M., Patel, D. J., Jaffrey, S. R., and Kharas, M. G. N6-Methyladenosine on mRNA facilitates a phase-separated nuclear body that suppresses myeloid leukemic differentiation. *Cancer cell* **39**, 958–972.e8 (2021).
73. Liu, J., Gao, M., He, J., Wu, K., Lin, S., Jin, L., Chen, Y., Liu, H., Shi, J., Wang, X., Chang, L., Lin, Y., Zhao, Y. L., Zhang, X., Zhang, M., Luo, G. Z., Wu, G., et al. The RNA m6A reader YTHDC1 silences retrotransposons and guards ES cell identity. *Nature* **591**, 322–326 (2021).
74. Zhu, W., Ding, Y., Meng, J., Gu, L., Liu, W., Li, L., Chen, H., Wang, Y., Li, Z., Li, C., Sun, Y., and Liu, Z. Reading and writing of mRNA m6A modification orchestrate maternal-to-zygotic transition in mice. *Genome Biol*. **24**, 67 (2023).
75. Chen, L., Fu, Y., Hu, Z., Deng, K., Song, Z., Liu, S., Li, M., Ou, X., Wu, R., Liu, M., Li, R., Gao, S., Cheng, L., Chen, S., and Xu, A. Nuclear m6 A reader YTHDC1 suppresses proximal alternative polyadenylation sites by interfering with the 3' processing machinery. *EMBO Rep*. **23**, e54686 (2022).
76. Xu, K., Yang, Y., Feng, G. H., Sun, B. F., Chen, J. Q., Li, Y. F., Chen, Y. S., Zhang, X. X., Wang, C. X., Jiang, L. Y., Liu, C., Zhang, Z. Y., Wang, X. J., Zhou, Q., Yang, Y. G., and Li, W. Mettl3-mediated m6A regulates spermatogonial differentiation and meiosis initiation. *Cell Res*. **27**, 1100–1114 (2017).
77. Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., Zheng, H., Klungland, A., and Yan, W. ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc. Natl. Acad. Sci. USA*. **115**, E325–E333 (2018).
78. Zhang, H., Lee, J. Y., and Tian, B. Biased alternative polyadenylation in human tissues. *Genome Biol*. **6**, R100 (2005).
79. Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
80. Ji, Z., Lee, J. Y., Pan, Z., Jiang, B., and Tian, B. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc. Natl. Acad. Sci. USA*. **106**, 7028–7033 (2009).
81. Smibert, P., Miura, P., Westholm, J. O., Shenker, S., May, G., Duff, M. O., Zhang, D., Eads, B. D., Carlson, J., Brown, J. B., Eisman, R. C., Andrews, J., Kaufman, T., Cherbas, P., Celniker, S. E., Graveley, B. R., and Lai, E. C. Global patterns of tissue-specific alternative polyadenylation in Drosophila. *Cell Rep*. **1**, 277–289 (2012).
82. Ulitsky, I., Shkumatava, A., Jan, C. H., Subtelny, A. O., Koppstein, D., Bell, G. W., Sive, H., and Bartel, D. P. Extensive alternative polyadenylation during zebrafish development. *Genome Biol*. **22**, 2054–2066 (2012).
83. Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O., and Lai, E. C. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res*. **23**, 812–825 (2013).

84. Sanfilippo, P., Wen, J., and Lai, E. C. Landscape and evolution of tissue-specific alternative polyadenylation across Drosophila species. *Genome Biol*. **18**, 229 (2017).
85. Lee, S., Chen, Y. C., FCA Consortium, Gillen, A. E., Taliaferro, J. M., Deplancke, B., Li, H., and Lai, E. C. Diverse cell-specific patterns of alternative polyadenylation in Drosophila. *Nat. Commun*. **13**, 5372 (2022).
86. Li, W., Park, J. Y., Zheng, D., Hoque, M., Yehia, G., and Tian, B. Alternative cleavage and polyadenylation in spermatogenesis connects chromatin regulation with post-transcriptional control. *BMC Biol*. **14**, 6 (2016).
87. Zhang, Y., Tang, C., Yu, T., Zhang, R., Zheng, H., and Yan, W. MicroRNAs control mRNA fate by compartmentalization based on 3' UTR length in male germ cells. *Genome Biol*. **18**, 105 (2017).
88. Nguyen-Chi, M., Chalmel, F., Agius, E., Vanzo, N., Khabar, K. S., Jégou, B., and Morello, D. Temporally regulated traffic of HuR and its associated ARE-containing mRNAs from the chromatoid body to polysomes during mouse spermatogenesis. *PLoS One* **4**, e4900 (2009).
89. Nguyen-Chi, M., Auriol, J., Jégou, B., Kontoyiannis, D. L., Turner, J. M., de Rooij, D. G., and Morello, D. The RNA-binding protein ELAVL1/HuR is essential for mouse spermatogenesis, acting both at meiotic and postmeiotic stages. *Mol. Biol. Cell* **22**, 2875–2885 (2011).
90. Dai, P., Wang, X., Gou, L. T., Li, Z. T., Wen, Z., Chen, Z. G., Hua, M. M., Zhong, A., Wang, L., Su, H., Wan, H., Qian, K., Liao, L., Li, J., Tian, B., Li, D., Fu, X. D., Shi, H. J., Zhou, Y., and Liu, M. F. A Translation-Activating Function of MIWI/piRNA during Mouse Spermiogenesis. *Cell* **179**, 1566–1581.e16 (2019).
91. Galganski, L., Urbanek, M. O., & Krzyzosiak, W. J. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res*. **45**, 10350–10368 (2017).
92. Bhat, P., Chow, A., Emert, B., Ettlin, O., Quinodoz, S. A., Strehle, M., Takei, Y., Burr, A., Goronzy, I. N., Chen, A. W., Huang, W., Ferrer, J. L. M., Soehalim, E., Goh, S. T., Chari, T., Sullivan, D. K., Blanco, M. R., & Guttman, M. Genome organization around nuclear speckles drives mRNA splicing efficiency. *Nature* **629**, 1165–1173 (2024).
93. Yoon, Y., Bournique, E., Soles, L. V., Yin, H., Chu, H. F., Yin, C., Zhuang, Y., Liu, X., Liu, L., Jeong, J., Yu, C., Valdez, M., Tian, L., Huang, L., Shi, X., Seelig, G., Ding, F., Tong, L., Buisson, R., & Shi, Y. RBBP6 anchors pre-mRNA 3' end processing to nuclear speckles for efficient gene expression. *Molecular cell* **85**, 555–570 (2025).
94. Aravin, A. A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K. F., Bestor, T., & Hannon, G. J. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* **31**, 785–799 (2008).
95. Zeitelhofer, M., Karra, D., Macchi, P., Tolino, M., Thomas, S., Schwarz, M., Kiebler, M., & Dahm, R. Dynamic interaction between P-bodies and transport ribonucleoprotein particles in dendrites of mature hippocampal neurons. *J. Neurosci*. **28**, 7555–7562 (2008).
96. Lachke, S. A., Alkuraya, F. S., Kneeland, S. C., Ohn, T., Aboukhalil, A., Howell, G.

R., Saadi, I., Cavallesco, R., Yue, Y., Tsai, A. C., Nair, K. S., Cosma, M. I., Smith, R. S., Hodges, E., Alfadhli, S. M., Al-Hajeri, A., Shamseldin, H. E., Behbehani, A., Hannon, G. J., Bulyk, M. L., … Maas, R. L. Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. *Science* **331**, 1571–1576 (2011).

97. Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M. J., Scheuner, D., Kaufman, R. J., Golan, D. E., & Anderson, P. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol*. **169**, 871–884 (2005).
98. Bao, J., Vitting-Seerup, K., Waage, J., Tang, C., Ge, Y., Porse, B. T., and Yan, W. UPF2-Dependent Nonsense-Mediated mRNA Decay Pathway Is Essential for Spermatogenesis by Selectively Eliminating Longer 3'UTR Transcripts. *PLoS Genet*. **12**, e1005863 (2016).
99. Ruan, K., Bai, G., Fang, Y., Li, D., Li, T., Liu, X., Lu, B., Lu, Q., Songyang, Z., Sun, S., Wang, Z., Zhang, X., Zhou, W., & Zhang, H. Biomolecular condensates and disease pathogenesis. *Sci. China Life Sci*. **67**, 1792–1832 (2024).
100. Forman-Kay, J. D., Ditlev, J. A., Nosella, M. L., & Lee, H. O. What are the distinguishing features and size requirements of biomolecular condensates and their implications for RNA-containing condensates?. *RNA* **28**, 36–47 (2022).
101. Sahin, C., Leppert, A., & Landreh, M. Advances in mass spectrometry to unravel the structure and function of protein condensates. *Nature protocols* **18**, 3653–3661 (2023).

Methods

Mammalian cell culture

All human and mouse cell lines used in this study have been authenticated and were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco, cat# C11995500BT) or RPMI1640 (Gibco, cat# C22400500CP) supplemented with 10% fetal bovine serum (FBS) (TransGen Biotech, cat# FS201-01) and 1% penicillin/streptomycin (PS) (Gibco, cat# 15140122), and maintained at 37℃ in an incubator with 5% CO2. Also, we used trypsin-EDTA (Gibco, cat# 25200072) for cell passaging and digestion.

Mice

The wildtype male C57BL/6J mice at different ages used in this study were purchased from Shanghai laboratory animal center (SLAC) and GemPharmatech company. The C57BL/6J *LENG8*-flox mouse strain was constructed by Cyagen Biosciences using CRISPR/Cas9 technology, with two loxp sites inserted into the intron before the third exon and after the fourth exon of *LENG8*. *Stra8-Cre* knock-in transgenic mice were provided as a gift by Professor Minghan Tong at the Center for Excellence in Molecular Cell Science, CAS. The *Cre* recombinase excised the third and fourth exons of *LENG8*, resulting in frameshift mutations and leading to the knockout of *LENG8*. $LENG8^{flox/\Delta}$ *Stra8-Cre* (*LENG8* cKO) conditional knockout male mice were obtained by mating $LENG8^{flox/+}$ *Stra8-Cre* with $LENG8^{flox/flox}$ mice. HA-*LENG8* knock-in mice were constructed by GemPharmatech company, and adult male mice were used for the experiments described in this study. All the animal experiments were approved by the Committee on Science and Technology Ethics of Tongji University.

Plasmids and primers

DNA fragments were amplified and inserted into pCMV-Myc (Clontech, 631604), p3×Flag-CMV-10 (Sigma, E7658), pET-28a (Novagen, 69864-3), pEGFP-C1 (Clontech, 6084-1), pEGFP-N2 (Clontech, 6081-1) and pCOLD-GST (Takara, 3372) vectors, respectively, using ClonExpress II One Step Cloning Kit (Vazyme, C112-01). The primer sequences used to detect the length of targets 3′UTR are seen in Supplementary Table 6.

Biotinylated isoxazole (B-isox)-mediated precipitation from testis extracts

B-isox precipitation in mouse testes at different developmental stages was performed with a procedure modified from the method described previously.[36] In brief, testes were homogenized in lysis buffer [20 mM Tris-HCl, pH=7.4, 150 mM NaCl, 5 mM $MgCl_2$, 20 mM BME;1% TritonX-100, RNase inhibitor (Takara) and protease inhibitor cocktail (MCE)], followed by the addition of B-isox at a final concentration of 200 μM and then incubated at 4 °C for 1 hour. The B-isox-bound protein pellets were obtained by centrifugation at 14,000 rpm for 15 min and subsequently washed twice with B-isox lysis buffer. The pellets were used for western blotting and mass spectrometry analysis.

Histological analysis and apoptosis assay

LENG8 cKO and wildtype mouse testicular tissues were fixed with Bouin's solution at room temperature, followed by sequential washing with 70%, 80%, 90% and 100%

ethanol. Subsequently, tissues were rendered transparent using xylene and then embedded in paraffin. After deparaffinization and rehydration, the sections (4 μm) were subjected to with hematoxylin and eosin (H&E) staining.

Germ cell apoptosis was evaluated in testis cryosections using TdT-mediated dUTP nick end Labeling (TUNEL) assays with a commercial kit (Beyotime Biotechnology) following the manufacturer's instruction.

Immunoprecipitation (IP), RNA immunoprecipitation (RIP) and immunoblotting

IP assays were conducted following previously established protocols.[90] First, mouse testes or cultured cells were lysed with pre-cooled lysis buffer [50 mM Tris-HCl (pH 7.4), 1% Triton X-100, 150 mM NaCl, 5 mM EDTA] contained protease inhibitor cocktail (MedChemExpress) and RNase inhibitor (MedChemExpress) on ice for 20 min. After centrifugation at 12,000 rpm for 10 min at 4 °C, the supernatant was collected. Cell lysates were then incubated with primary antibodies (IgG, anti-LENG8, anti-PABPN1, anti-YTHDC1, anti-Myc, anti-Flag, anti-EGFP) pre-coupled Protein A/G beads at 4 °C. The beads were washed four times for 5 min each with low salt washing buffer [50 mM Tris-HCl (pH 7.4), 0.1% Triton X-100, 150 mM NaCl, 5 mM EDTA, and protease inhibitor cocktail], followed by one wash with lysis buffer. RNAs bound to target proteins in IP pellets were extracted using RNAiso Plus (Takara) according to the manufacturer's protocol. The extracted RNAs were used for subsequent RNA sequencing analysis. For immunoblotting analysis, cell/tissue extracts or IP pellets were mixed with SDS loading buffer, followed by standard SDS-PAGE and western blotting procedures. The images were obtained utilizing the Tanon-5200 Chemiluminescent Imaging System (Tanon).

Fluorescence in situ hybridization (FISH) and immunofluorescent microscopy (IF)

To detect global mRNAs or LENG8 target *Tpi1* mRNAs, cultured cells, testicular germ cells or testis cryosections seeded onto glass slides were washed once with pre-cooled PBS and fixed with 4% paraformaldehyde (PFA). Cells were washed three times with PBS and permeabilized with 0.5% Triton X-100 in PBS for 5 min, followed by a single 15 min wash with 2×saline-sodium citrate buffer (SSC). Then, cells were incubated overnight at 42 ℃ with the Cy3-conjugated-oligo dT probes or *Tpi1* antisense probes in the hybridization buffer. At last, cells were washed twice with 2×SSC and twice with 0.5×SSC, respectively, followed by DAPI staining. For IF, cultured cells, testicular germ cells or testis cryosections seeded onto glass slides were fixed with 4% PFA 15 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS. The fixed cells were then blocked with 1% BSA and 0.5% NDS (normal donkey serum) in PBS for 1 h, followed by incubation with primary antibodies and then with either Alexa Fluor 488/594 (JACKSON ImmunoResearch) or IFKine Green/Red secondary antibodies (Abbkine). *In vivo* 5-ethynyluridine (EU) labeling assay was performed according to the manufacturer's protocol. Both 293T and HeLa cells were treated with 500 μM 5-EU for 24 h prior to detection. The cell nuclei were stained with DAPI or Hoechst 33342. All images were acquired using Leica SP8, SpinSR10 or OLYMPUS CKX53SF microscope. The fluorescence densities were quantified using ImageJ, and statistical analyses were conducted using GraphPad Prism 7.0.

Antibodies

Antibodies were used for western blot, IF, and IP: anti-LENG8 (our lab), anti-PABPN1 (Beyotime Biotechnology, AG3616), anti-YTHDC1(Santa Cruz, sc-136428), anti-CPSF4 (Santa Cruz, sc-390643), anti-SYCP3(Santa Cruz, sc-74569), anti-NONO (Santa Cruz, sc-166702), anti-DDX25 (Santa Cruz, sc-166289), anti-CPSF4 (Proteintech, 15023-1-AP), anti-COIL (Proteintech, 67218-1-Ig), anti-YTHDC1 (ABclonal, A22992), anti-FIP1L1(ABclonal, A7138), anti-GFP (Abways, AB0005), anti-β-actin (Abways, Ab2001), anti-Myc (Abways, AB0001), anti-Flag (Abways, AB0008), anti-HA (Cell Signaling Technology, 3724), anti-SC35 (Abcam, Ab11826), anti-ALKBH5 (ABclonal, A22137), anti-METTL3 (ABclonal, A8370), anti-CPSF2 (ABclonal, A9297), anti-SC35 (Santa Cruz, sc-53518), Mouse polyclonal anti-IgG (Merck, 12-371), Rabbit polyclonal anti-IgG (Merck, 12-370), HRP, Goat Anti-Mouse IgG (Abbkine, A21010), HRP, Goat Anti-Mouse IgG LCS (Abbkine, A25012), HRP, Goat Anti-Rabbit IgG (Abbkine, A21020), Alexa Fluor 488 AffiniPure Donkey Anti-MOUSE (JACKSON ImmunoResearch, 715-545-150), Alexa Fluor 594 AffiniPure Donkey Anti-Rabbit (JACKSON ImmunoResearch, 711-585-152), IFKine Red Donkey Anti-Mouse (Abbkine, A24411) and IFKine Green Donkey Anti-Rabbit (Abbkine, A24221).

Transfection and RNAi

Cells were cultured in DMEM or RPMI1640 supplemented with 10% FBS. Lipofectamine 2000 (Thermo Fisher Scientific) transfection reagent was used to transfect plasmids or siRNAs according to the manufacturer's protocol. For RNAi, 40-80 nM siRNAs against *LENG8*, *PABPN1*, *YTHDC1*, *METTL3*, *ALKBH5*, *SRRM2*, *SON* and control siRNAs (GenePharma or Biotend) were used for transfection. Cells were harvested for IP or IF 48 h after transfection. For PABPN1 rescue experiments, siRNAs against *PABPN1* were transfected into 293T cells, followed by transfection of Flag-PABPN1 plasmids 12 h post siRNA transfection. For wildtype LENG8 or condensation-deficient mutant rescue experiments, siRNAs against *LENG8* were transfected into 293T cells, followed by transfection of wildtype LENG8 or condensation-deficient mutant plasmids 24 h post siRNA transfection. Cells were harvested for WB or RNA-seq 72 h later.

siRNA sequence information is as follows.

si*LENG8*-1: GCCGUGGACUUGGGUUGUAAA;
si*LENG8*-2: GAAGCGACCCUUUGCUGUUAC;
si*LENG8*-3: GGACAAGGACCGCACGGAA;
si*PABPN1*-1: UUAGAUGAGUCCCUAUUUA;
si*PABPN1*-2: AGUCAACCGUGUUACCAUA;
si*YTHDC1*-1: UGGAUUUGCAGGCGUGAAUU;
si*YTHDC1*-2: UGCCUCCAGAGAACCUUAUA;
si*METTL3*: CUGCAAGUAUGUUCACUAUGA;
si*ALKBH5*-1: ACAAGUACUUCUUCGGCGA;
si*ALKBH5*-2: GCGCCGUCAUCAACGACUA;
si*SRRM2*: CGACGAAGAUCCCGGUCAA;
si*SON*: GGAUAAGGCUCAAUUACUU.

Fluorescence recovery after photobleaching (FRAP)

293T cells were seeded onto Nunc cell dishes and allowed to adhere for 24 h, followed by transfection with pEGFP-LENG8 plasmids. After 24 h of transfection, EGFP-LENG8 puncta were photobleached using a 488 nm laser beam with the FRAP module of the Leica SP8 confocal microscopy system. GFP fluorescence signal was collected before and after photobleaching over a period of time. Fluorescence intensity was measured by ImageJ. The pre-bleached fluorescence intensity was set to 1, and the post-bleach signal was normalized to the pre-bleach level.

Mouse germ cells isolation

Isolation of mouse germ cells from testes was performed as described previously.[90] The testes were dissected form male mice and the tunica albuginea was carefully removed. The testicular tissue was placed in 1×PBS, and the seminiferous tubules were torn apart using fine forceps. The seminiferous tubules were subsequently digested into single-cell suspension using 1 mg/ml of type I collagenase and 2.5 mg/ml of trypsin. Finally, applied the cell suspension onto slides evenly and fixed them with 4% paraformaldehyde.

Ribosome profiling

We modified previous described methods to perform ribosome profiling.[90] Briefly, homogenization of mouse testes from different ages was performed in lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, 100 μg/ml cycloheximide, 25U/ml Turbo DNase I, 1% Triton X-100). The DNase-treated lysates were carefully layered onto 10%–60% sucrose gradients to isolate polysome fractions. Further ultracentrifugation of the polysome fractions was performed to obtain ribosome pellets. Ribosome pellets were then treated with RNase I to obtain ribosome-protected RNA fragments of ~30 nt in length. Subsequently, the underwent RNA sequencing analysis utilizing Illumina sequencing technology.

LENG8 granule dissolution by RNase A, 1,6-hexanediol (1,6-HD), actinomycin D treatment

For the treatment with RNase A, cells were grown on coverslips were rinsed twice with PBS and then permeabilized with a 0.1% Triton X-100 solution diluted in reaction buffer (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$) for 10 min at room temperature. Cells were incubated with 500 μg/ml RNase A for 20 min. The 1,6-HD treatment involved the addition of 1,6-HD dissolved in DMEM medium to the cells at 2%, 4%, 6%, 8%, 10% concentrations for an incubation of 10 min. After RNase A or 1,6-HD treatment, the cells were fixed with 4% PFA and subjected to subsequent IF analysis. For actinomycin D treatment (MCE), 2 μM actinomycin D was used to inhibit transcription in both 293T and HeLa cells for 4 h.

Protein expression, purification, *in vitro* LLPS assay, and *in vitro* pull-down assay

The full-length *LENG8* was PCR amplified from a human cDNA library and subsequently cloned into the pCOLD-GST vectors. For *in vitro* pull-down assay, the full-length *LENG8*, *YTHDC1*, and *CPSF4* were cloned into pET-28a vectors, respectively. The constructed plasmids were then transformed into *Escherichia coli*

BL21 (DE3) cells (Tsingke) and cultured overnight at 37 ℃. A single colony was selected and added to the medium for expanded incubation until OD600 reached 0.6-0.8. 0.4 mM IPTG was added to the medium to induce recombinant proteins expression at 18 ℃ for 24 h and collected by centrifugation. To purify GST-LENG8 proteins, the induced BL21 (DE3) cells were resuspended in binding buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, and 2 mM DTT). To purify 6×His-LENG8/YTHDC1/CPSF4 proteins, the induced BL21 (DE3) cells were resuspended in binding buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM DTT and 10 mM imidazole). These cells were lysed with ultrasound and then sedimented at 12,000 rpm for 30 min to pellet debris. The supernatant lysates were loaded onto Glutathione agarose beads (Yueke) to purify the recombinant GST-LENG8 proteins, and onto Ni-NTA agarose beads (Vazyme) to collect 6×His-LENG8/YTHDC1/CPSF4 proteins, respectively. After five washes with their respective binding buffers, GST-LENG8 proteins were eluted with reduced glutathione (GSH) elution buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM DTT, and 30 mM GSH), and 6×His-LENG8/YTHDC1/CPSF4 proteins were eluted with 6×His elution buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM DTT and 500 mM imidazole). Subsequently, we concentrated the eluted GST-LENG8 proteins using Ultra Centrifugal Filters (MERCK) and exchanged the elution buffer to the storage buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl). The collected proteins were assessed using SDS-PAGE gel by Coomassie staining, followed by snap-freezing in liquid nitrogen and storage at −80 °C for later use.

The purified GST-LENG8 proteins were diluted into varying concentrations (0 μM, 1.2 μM, 2.4 μM, 6 μM, 12 μM) in a buffer containing 50 mM Tris-HCl pH 8.0, 150 mM NaCl and 6% (w/v) polyethylene glycol 6000 (PEG6000). The purified GST-LENG8 proteins were also diluted to a concentration of 6 μM in buffers containing 0 %, 1.5 %, 3 %, 6 %, 12 % of PEG6000 or in buffers with 150 mM, 250 mM, 375 mM, 500 mM NaCl concentrations. For the 1,6-HD treatment assay, 12 μM purified proteins were mixed with a final concentration of 5 % (w/v) 1,6-HD. *In vitro* fusion assay was conducted in a buffer containing 6 μM purified proteins, 6 % PEG6000, 50 mM Tris-HCl pH 8.0, and 150 mM NaCl. All these samples were then dripped onto a glass slide to conduct *in vitro* phase separation assay. Images were captured with OLYMPUS CKX53SF microscope.

For *in vitro* pull-down assay, the eluted 6×His-LENG8 proteins were first bound to beads using LENG8 antibodies, with IgG serving as a control. Subsequently, the eluted 6×His-YTHDC1/CPSF4 proteins were incubated with these beads. After five washes with pull-down binding buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, and 0.1 % Triton X-100), the beads were used for WB analysis.

Tethering reporter assay

The LENG8 target gene *Chchd7* containing its CDS and 3′UTR, along with 6×MS2 (a short RNA stem-loop) sequences, was constructed on the pEGFP-N2 vector. MCP (bacteriophage MS2 coat protein)-tagged wildtype LENG8 and a condensation-deficient mutant (LENG8$^{\Delta 24\text{-}66}$) were inserted into the pCMV-Myc vector. The reporter plasmid (*Chchd7*-6×MS2) and the effecter plasmid (MCP-LENG8 or MCP- LENG8$^{\Delta 24\text{-}66}$) were used to co-transfected 293T cells. For *LENG8* knockdown assay, siRNAs

against *LENG8* were first transfected into 293T cells, followed by transfection of the reporter plasmid 24 h post siRNA transfection. After 48 h, the cells were harvested for RNA extractions. Then, semi-quantitative PCR assays were performed targeting the control β-actin, as well as the long and short 3′UTRs of reporter mRNAs.

RNA-seq data analysis

The RNA-seq libraries were sequenced in the PE150 mode. The raw pair-end reads were preprocessed to remove the adapters with cutadapt (v.3.5). The quality of the resulting reads was evaluated using FastQC (v.0.11.9). Reads were firstly mapped to rRNA sequences from NCBI, and those reads not mapped to rRNA sequences were aligned to the mouse genome (mm10) with GENCODE vM10 gene annotation or human genome (hg38) with GENCODE v32 gene annotation using STAR aligner (v.2.7.9a). Then, PCR duplicated reads were removed using Picard (v. 2.26.3) program. Gene expression matrixes were counted by featureCounts (v.2.0.1) command, the differential expression analysis was performed by DeSeq2 R package. Threshold of differential expression genes is p-value < 0.05 and fold change > 2 or < 0.5.

Alternative polyadenylation events were analyzed using Dapars program through RNA-seq aligned files. Threshold for cell lines of change events is p-value < 0.05 and absolute value of ΔPDUI > 0.1, and we used a stricter threshold in tissues which is p-value < 0.01 and absolute value of ΔPDUI > 0.2. For a specific gene, the pair of PASs with the greatest difference in PDUI between the treatment and control groups were selected.

Alternative splicing events were analyzed using rMATS (v4.1.2) program, including RI, A3SS, A5SS, SE, and MXE. We used a Δpsi cutoff of 0.15, and an FDR cutoff of 0.05 to estimate the significance of change.

Ribo-seq analysis

The Ribo-seq libraries were sequenced in the PE150 mode. First, we extracted the insert fragments from sequenced reads using custom python scripts, which limited the minimum length of insert fragment to 18 nt. Then, reads were mapped to mm10 genome using STAR aligner, PCR duplicated reads were removed using Picard, gene expression matrix was counted by featureCounts in CDS regions. Differential translation efficiency analysis was performed by Xtail R package. Quality control was performed by ribosomeProfilingQC R package on fragments length of 28-29 nt.

Single-cell RNA-seq analysis

Single-cell RNA sequencing was performed on SeekGene platform. Data was first mapped to mm10 genome and counted by cellranger (v. 7.0.0) program. Then the count matrix was read into R by Seurat R package. Data of WT sample was downloaded from GEO dataset GSE174731. WT and KO samples matrixes were normalized, then removed batch effects and integrated by IntegrateData function. Finally, the resolution was set to 1.5 by using FindClusters function for clustering. *Ddx4*, *Pecam1*, *Vwf*, *Igf2*, *Sox9*, *Cd74*, *Ccnd2*, *Cenpa*, *Stra8*, *Gm960*, *Meiob*, *Psma8*, *Piwil1*, *Pou5f2*, *Ccna1*, *Tex36*, *Sun5* and *Prm1* were used to identify spermatogenic and somatic cells. A1: A1 type spermatogonia, ePL: early preleptotene, mPL: middle preleptotene, L: leptotene, eP: early pachytene, mP: middle pachytene, lP: late pachytene, D: diplotene, M:

metaphase, RS2: steps 1-2 spermatids; RS4: steps 3-4 spermatids; RS6: steps 5-6 spermatids; RS8: steps 7-8 spermatids.

RIP-seq analysis

The RIP-seq libraries were sequenced in the PE150 mode. Reads were mapped to mm10 genome using STAR aligner, PCR duplicated reads were removed using Picard, transcript abundance was quantified by stringtie (v.2.1.5). Transcripts with the TPM fold change over 2 between IP and input were identified as targets of bait protein, and corresponding genes were identified as target genes. Fisher's exact test was used to evaluate the statistical significance for the overlapping of a Venn diagram on the 2×2 contingency table, using the number of protein coding genes not in the diagram as the double negative set.

YTHDC1 CLIP-seq analysis

We downloaded HEK293T YTHDC1 iCLIP data from GEO dataset, then reads were mapped to hg38 genome using STAR aligner. We quantified reads mapped to 3′UTR by featureCounts for each gene, and defined 5763 YTHDC1 target genes whose 3′UTR TPM over 35.

K-mer differential analysis

We counted the frequency of k-mer (6-mer) occurrence in each sequence set and used fisher's exact test for significance test. The odds ratio value was used as a statistical indicator of difference. K-mers whose p-value < 0.05 and odds ratio > 1.5 or $< 1/1.5$ between groups were defined as differential ones.

Targeted genes clustered analysis

LENG8 targeted genes were clustered into 4 groups based on RNA and protein expression matrix using Mfuzz R package to perform c-means clustering. Mass spectrum data at different developmental stages was download from a previous study.[97] We filled missing values with the lower quartile, then each sample was normalized according to the total amount. RNA FPKM and normalized protein quantified data at different developmental stages were scaled respectively and then combined into a matrix to perform clustering.

Calculating nuclear abundance of targets in different cell lines

We collected publicly available nuclear RNA-seq data for the respective cell lines to calculate the nuclear abundance of targets in different cell lines. The gene expression levels were normalized using TPM (Transcripts Per Million). Specifically, the dataset of MCF7 was sourced from GSE139151, the HeLa dataset was sourced from GSE74397, and the 293T dataset was sourced from GSE195493.

Statistical analyses

Statistical analysis was performed in Python (v.3.9.7) and graphs were plotted in RStudio (R version 4.3.0) using ggplot2 (v.3.4.3) R package. The k-mer differential significance test was performed by fisher's exact test through scipy (v. 1.10.1) package. GraphPad Prism software was applied for the weight and $TUNEL^+$ cell/tubule of WT and *LENG8* cKO testis analysis. More detailed information could be found in related

figure legends and Methods.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The RNA-seq data generated in this study have been deposited in China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (CRA015512 and HRA006998). The shared URLs for reviewers are https://ngdc.cncb.ac.cn/gsa/s/JfiCY7VS and https://ngdc.cncb.ac.cn/gsa-human/s/1xRXKGEE. The public datasets used in this study include: GSE195493, GSE74397, GSE139151, GSE78030, GSE112223, SRP125927, GSE161059, GSE125304, GSE132054, GSE99773, GSE174731, Fang et al[102]. All other data are available from the corresponding authors upon reasonable request.

References

102. Fang, K., Li, Q., Wei, Y., Zhou, C., Guo, W., Shen, J., Wu, R., Ying, W., Yu, L., Zi, J., Zhang, Y., Yang, H., Liu, S., and Chen, C. D. Prediction and Validation of Mouse Meiosis-Essential Genes Based on Spermatogenesis Proteome Dynamics. *Mol. Cell Proteomics* **20**, 100014 (2021).

Acknowledgements

We thank members of the P.D. and Y.Z. laboratories for helpful discussions and assistance. We express our gratitude to Mo-Fang Liu, Lan-Tao Gou, Dang-Sheng Li, Ming-Han Tong (Center for Excellence in Molecular Cell Science, CAS), Ling-Bo Wang (Fudan University, China), Juan Tang, Shan Bian, Jia-Min Zhang, Zhi-Ying He and Jian-Feng Lu (Tongji University, China), Lai-Xing Zhang (University of California Los Angeles, USA), Fu-Dong Li (University of Science and Technology of China), Jing Fan (Southeast University, China) for their experimental help and invaluable advice. We thank the staff members of the HPC & Storage facility at Center for Excellence in Molecular Cell Science (CEMCS) and the Integrated Laser Microscopy System and Mass Spectrometry System at the National Facility for Protein Science in Shanghai, Shanghai Advanced Research Institute, CAS for their instrumental support and technical assistance. This work was supported by grants from the National Key R&D Program of China (2022YFA1305300), National Natural Science Foundation of China (32171287 and 82341023), the Fundamental Research Funds for the Central Universities (22120240435), Peak Disciplines (Type IV) of Institutions of Higher Learning in Shanghai, Shanghai Rising-Star Program (21QA1409400), Science and Technology Commission of Shanghai Municipality (228001437).

Author contributions

Z.C.W., and P.D. conceived the study. Z.C.W., C.D., Y.Z., P.D. designed the

experiments. Z.C.W., C.D., X.Y.H., Y.W.Q., A.Z., M.R., H.W.Z. conducted the experiments. Z.C.W., C.D. performed bioinformatics analyses. Z.C.W., P.D. wrote the manuscript. All authors discussed the results. Y.Z. was responsible for bioinformatics studies. Y.Z. and P.D. supervised the project.

Competing interests

The authors declare no competing interests.

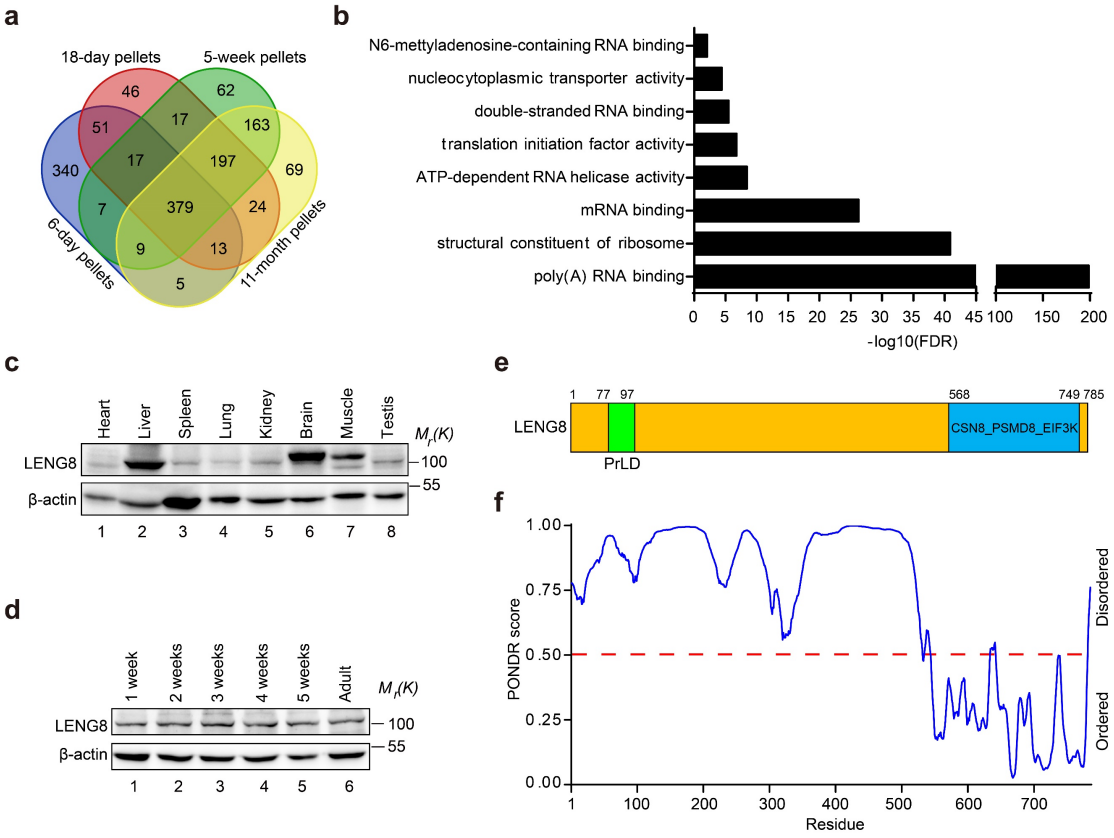


Fig. 1| LENG8 is precipitated by b-isox in testis lysates. a, Venn diagram illustrating the overlap of proteins in the b-isox-mediated pellets obtained from male mouse testes at different developmental stages. **b,** GO term analysis of the 379 overlapping proteins. **c,** Western blot analysis of LENG8 expression in several adult mouse tissues. β-actin served as a control. **d,** Developmental expression of LENG8 protein in postnatal testes. Testis extracts were prepared from different ages. β-actin served as a control. **e,** Predicted domain structure of mouse LENG8 based on the analysis with conserved domain database. PrLD is located between residues 77 and 97 and is flanked by a high likelihood of disordered regions. **f,** PONDR analysis of LENG8 sequence predicted a high likelihood of a disordered region at its N-terminus.

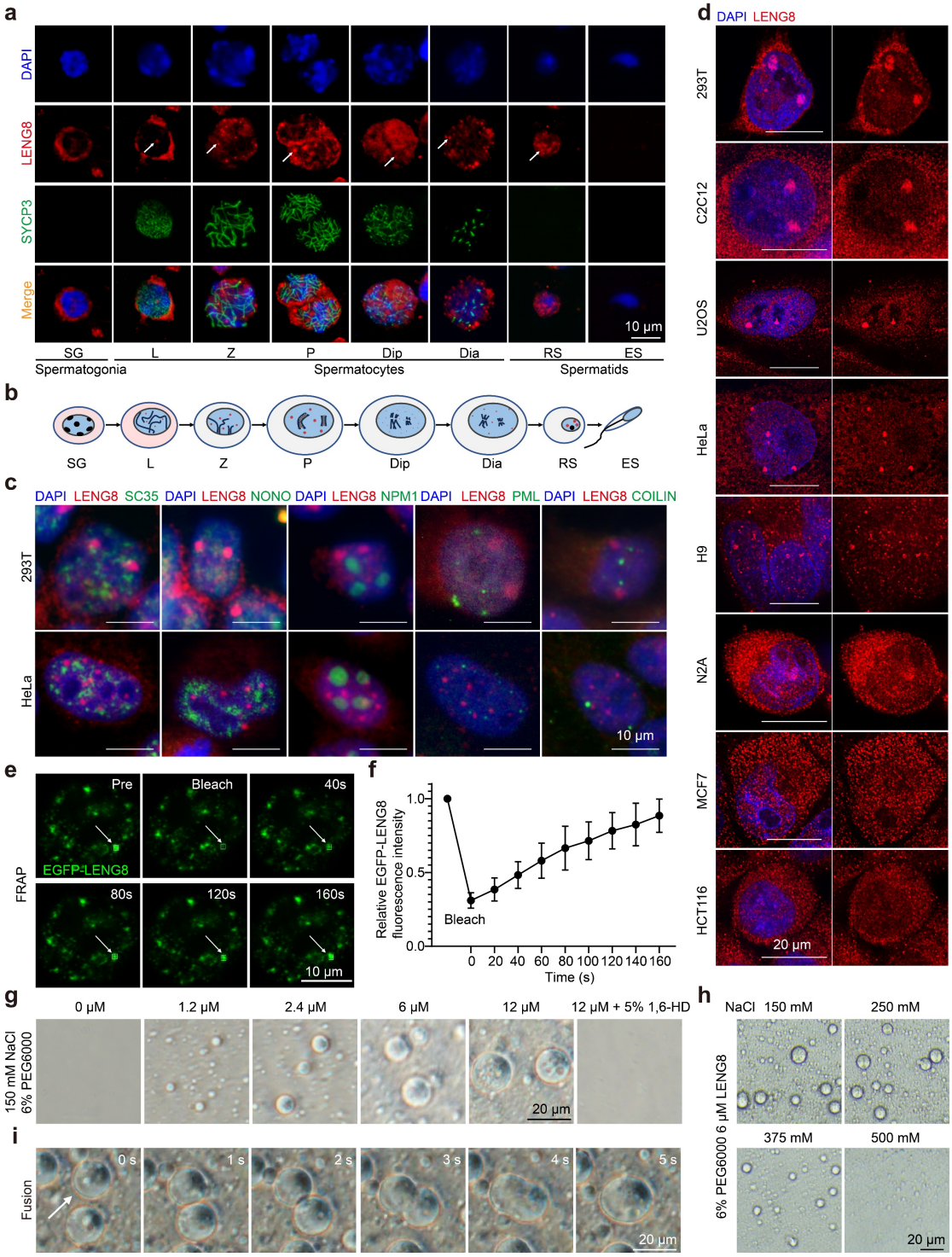


Fig. 2| LENG8 forms liquid-like granules in the nucleus. a, Immunofluorescent assay of LENG8 (red) and SYCP3 (green) on germ cells at different stages of development. Nuclei were counterstained with DAPI (blue). The arrow indicates one of the LENG8 granules in germ cells. SG: spermatogonia, L: leptotene, Z: zygotene, P: pachytene, Dip: diplotene, Dia: diakinesis, RS: round spermatids, ES: elongating/elongated spermatids. SYCP3, synaptonemal complex protein 3, a marker of meiosis. **b,** A schematic summary depicting the dynamic localizations of LENG8 in germ cells during spermatogenesis, based on the staining patterns observed in Fig 2a.

LENG8 granules are in red dots. **c,** Immunofluorescence images of the relationships between LENG8 (red) and markers (green) of different nuclear bodies in cultured 293T (a human embryonic kidney cell line) and HeLa (a human cervical cancer cell line) cells. SC35, nuclear speckle; NONO, paraspeckle; NPM1, nucleolus; PML, PML nuclear body. COILIN, Cajal body. **d,** Immunofluorescence images of LENG8 (red) in different cultured human and mouse cell lines. C2C12: a mouse myoblast cell line; U2OS: a human osteosarcoma cell line; H9: a human embryonic stem cell line; N2A: a mouse neuroblastoma cell line; MCF7: a breast cancer cell line; HCT116: a human colorectal carcinoma cell line. **e,** FRAP assays showed that fluorescence recovery of EGFP-LENG8 granules after photobleaching in 293T cells. The arrow indicates the area being tested. **f,** Quantification of fluorescence recovery of the EGFP-LENG8 granules. The error bars denote the standard deviation calculated from five independent measurements at each time point (mean ± SEM, n=5). **g,** Images of varying concentrations of purified recombinant LENG8 protein droplet formation in the presence of 6% (w/v) PEG6000 and 150 mM NaCl. 5% 1,6-HD treatment impedes the phase separation of LENG8. **h,** Representative images of 6 μM recombinant LENG8 droplet formation at different NaCl concentrations. **i,** Images depicting the fusion event of LENG8 droplets at a concentration of 6 μM.

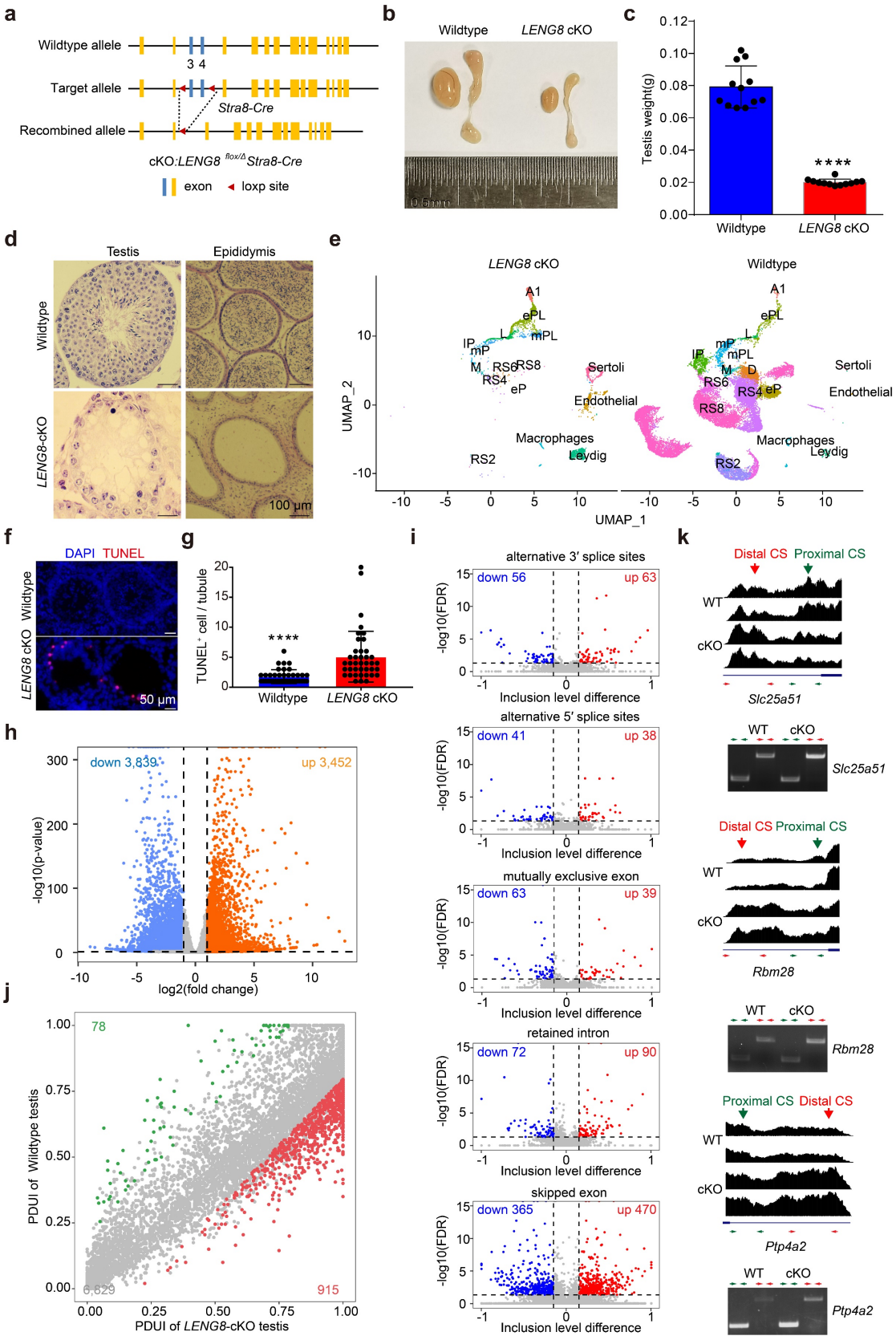


Fig. 3| Deletion of mouse *LENG8* results in male sterility and RNA metabolic disorder. a, Schematic diagram illustrating the construction of *LENG8* conditional knockout mice. Loxp sites were inserted between exons 3 and 4. Mice harboring the

floxed allele were crossed with *Stra8-Cre* transgenic mice to generate germ cell-specific deletion of *LENG8*. **b,** Morphology of the testis and the epididymis in wildtype and *LENG8* cKO adult mice. **c,** Histogram showing the testis weights in wildtype and *LENG8* cKO adult mice. Data are presented as mean ± SEM, n=12. ****P<0.0001 by Student's *t*-test. **d,** H&E staining of testis and epididymis sections from wildtype and *LENG8* cKO adult mice. **e,** Single-cell uniform manifold approximation and projection (UMAP) plot of testis cells from wildtype and *LENG8* cKO adult mice. **f,** TUNEL analysis of apoptotic germ cells in wildtype and *LENG8* cKO adult mouse testis sections. **g,** Number of TUNEL-positive cells per seminiferous tubules. Data are presented as mean ± SEM, n=40. ****P<0.0001 by Student's *t*-test. **h,** Volcano plot describing differentially expressed transcripts in *LENG8* cKO testes compared with 14-day-old wildtype testes. Logarithmic fold changes in expression are plotted on the x-axis, with normalized p-values on the y-axis. Horizontal line shows the negative base 10 logarithmic p-values (0.05) cutoff and vertical lines denote the fold change cutoff (log2 fold change≧1). Blue scatter dots represent downregulated genes and red scatter dots represent upregulated genes. **i,** Volcano plot depicting significant changes (-log10(FDR)) verse inclusion level difference in five alternative splicing types, including A3SS (alternative 3′ splice sites), A5SS (alternative 5′ splice sites), RI (retained intron), MXE (mutually exclusive exon), and SE (skipped exon) between *LENG8* cKO and wildtype testes. Red dots indicate genes with significantly increased inclusion levels and blue dots indicate those with significantly decreased inclusion levels. **j,** Scatter plots showing the percentage of distal poly (A) site usage index (PDUI) in *LENG8* cKO and wildtype testes. A total of 78 shortened and 915 lengthened gene transcripts are highlighted with green and red dots, respectively. Threshold of change events is p-value < 0.05 and absolute value of ΔPDUI > 0.2. **k,** Gene track view and RT-PCR validation of alternative polyadenylation in three examples. Dark blue bars represent the last exons, and dark blue lines indicate the introns. The regions within the red or green arrows delineate the sizes of the PCR amplification fragments.

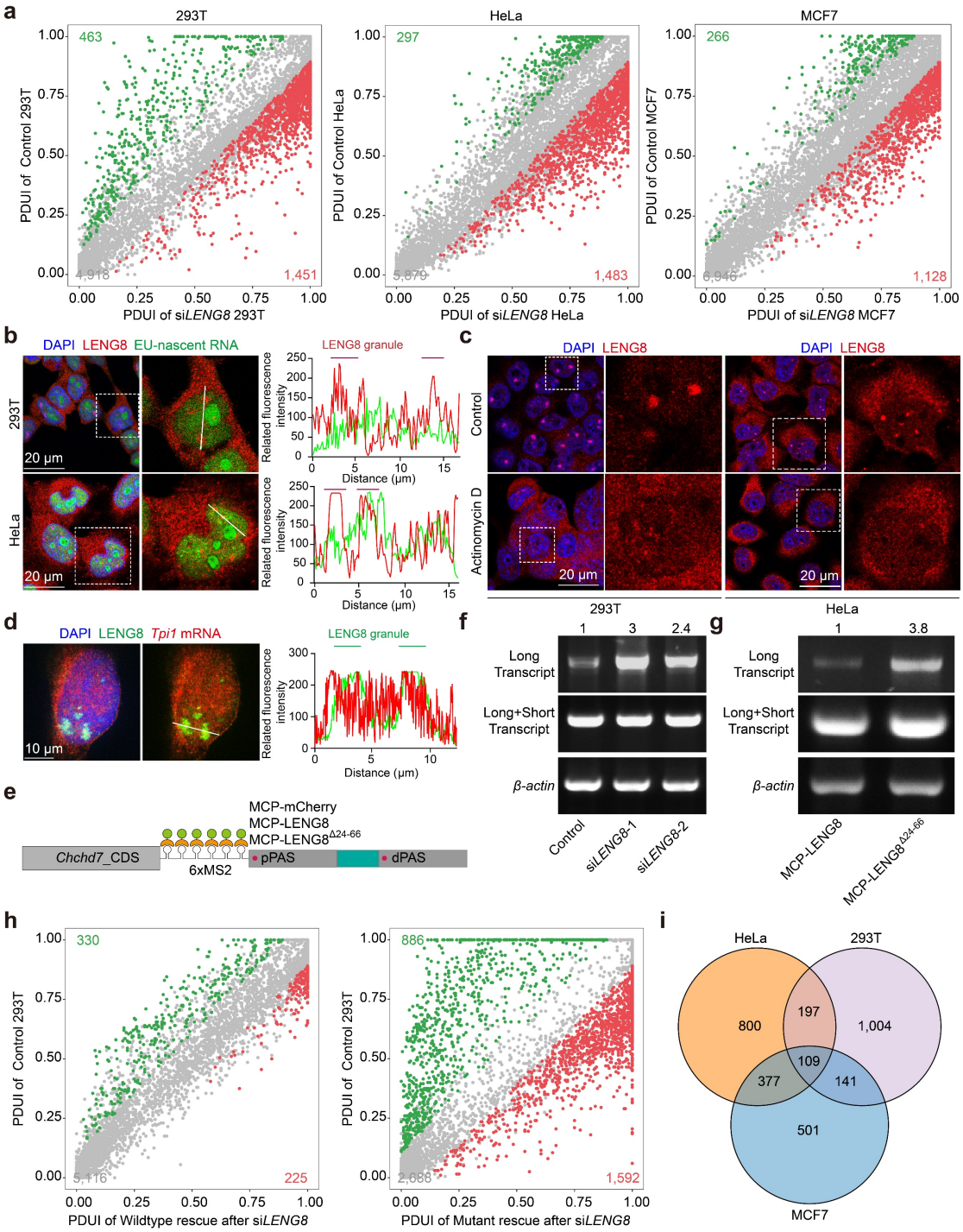


Fig. 4| LENG8 granule is an active site for APA regulation. a, Scatter plots showing the percentage of distal poly (A) site usage index (PDUI) in *LENG8* knockdown and control 293T, HeLa, and MCF7 cells. A total of 463 shortened and 1,451 lengthened genes in 293T cells, 297 shortened and 1,483 lengthened genes in HeLa cells, and 266 shortened and 1,128 lengthened genes in MCF7 cells are highlighted with green and red dots, respectively. **b,** Imaging the localization relationship between 5-EU labeled nascent RNAs and LENG8 granules in 293T and HeLa cells. Middle: magnified views of the areas enclosed by dashed boxes. Right: line scans display the related fluorescence intensity profiles of nascent RNAs (green) and LENG8 (red). The x-axis indicated

distance, and the y-axis indicates fluorescence intensity. **c,** Representative images showing the effects of transcription inhibition on LENG8 granules. 2 μM actinomycin D was used to inhibit transcription in both 293T and HeLa cells for 4 h. Right: magnified views of the areas enclosed by dashed boxes. **d,** Representative images of the localization relationship between LENG8 granules and target *Tpi1* mRNAs. Right: line scans display the related fluorescence intensity profiles of LENG8 (green) and target *Tpi1* mRNAs (red). **e,** Construct of the tethering reporter assay. The mRNA reporter contains the CDS and 3′UTR of LENG8 target gene *Chchd7*, along with 6×MS2 sequences. MCP peptides were fused with wildtype LENG8, a condensation-deficient mutant (LENG8$^{\Delta 24\text{-}66}$) or mCherry. **f,** RT-PCR analysis to assess the effect of *LENG8* knockdown on the 3′UTR length of the reporter mRNAs. **g,** RT-PCR analysis to assess the effect of wildtype LENG8, LENG8$^{\Delta 24\text{-}66}$ mutant on the 3′UTR length of the reporter mRNAs. **h,** Scatter plots showing the percentage of distal poly (A) site usage index (PDUI) in wildtype LENG8 (left) or LENG8$^{\Delta 24\text{-}66}$ mutant (right) rescue after *LENG8* knockdown and control 293T cells. **i,** Venn diagram illustrates the intersection of LENG8 targets in 293T, HeLa, and MCF7 cells.

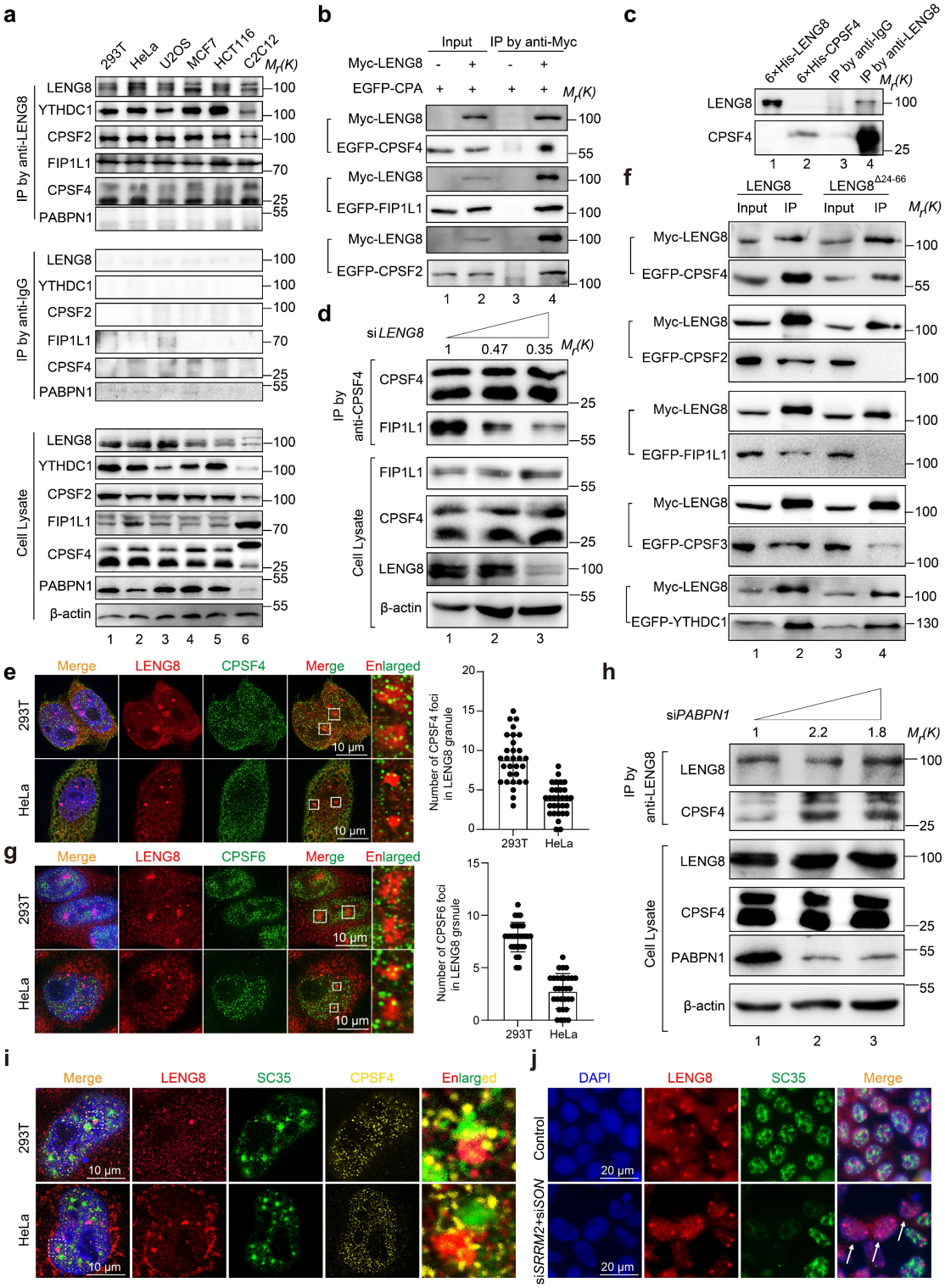


Fig. 5| LENG8 granules mediate CPA local assembly and functional interaction with nuclear speckles. a, Co-IP assay of the interaction between endogenous LENG8 with YTHDC1, CPSF2, FIP1L1, CPSF4, and PABPN1 in several cell types. Anti-LENG8 was used to pellet proteins from cell extracts. IgG was used as the negative control. **b,** Co-IP assay of the interaction between Myc-LENG8 and three EGFP-CPA factors co-transfected in 293T cells. **c,** *In vitro* pull-down assay showing the direct interaction between LENG8 and CPSF4. Anti-LENG8 was used to pellet recombinant

LENG8 proteins. IgG was used as the negative control. **d,** Co-IP assay of the interaction between endogenous CPSF4 and FIP1L1 in the context of *LENG8* knockdown. Anti-CPSF4 was used to pellet proteins from 293T cell extracts with varying levels of LENG8. **e,** High-resolution imaging data of the relationship between CPSF4 (green) and LENG8 (red) in cultured 293T and HeLa cells. Nuclei were counterstained with DAPI (blue). Enlarged: magnified views of the areas enclosed by dashed boxes. Right: quantitative summary of the number of CPSF4 foci in LENG8 granules (mean ± SEM, n=30). **f,** Co-IP assay of the interaction between wildtype LENG8 or LENG8$^{\Delta 24\text{-}66}$ mutant and four CPA factors and YTHDC1. Anti-Myc was used to pellet proteins from cell extracts. **g,** High-resolution imaging data of the relationship between CPSF6 (green) and LENG8 (red) in cultured 293T and HeLa cells. Nuclei were counterstained with DAPI (blue). Enlarged: magnified views of the areas enclosed by dashed boxes. Right: quantitative summary of the number of CPSF6 foci in LENG8 granules (mean ± SEM, n=30). **h,** Co-IP assay of the interaction between endogenous CPSF4 and LENG8 in the context of *PABPN1* knockdown. Anti-LENG8 was used to pellet proteins from 293T cell extracts with varying levels of PABPN1. **i,** High-resolution imaging data of the relationship among SC35 (green, a nuclear speckle marker protein), LENG8 (red) and CPSF4 (yellow) in cultured 293T and HeLa cells. **j,** Immunofluorescence images of the effect of nuclear speckle disruption by double knockdown SRRM2 and SON on LENG8 granules in 293T cells. Arrows point to the cells exhibiting disrupted nuclear speckles.

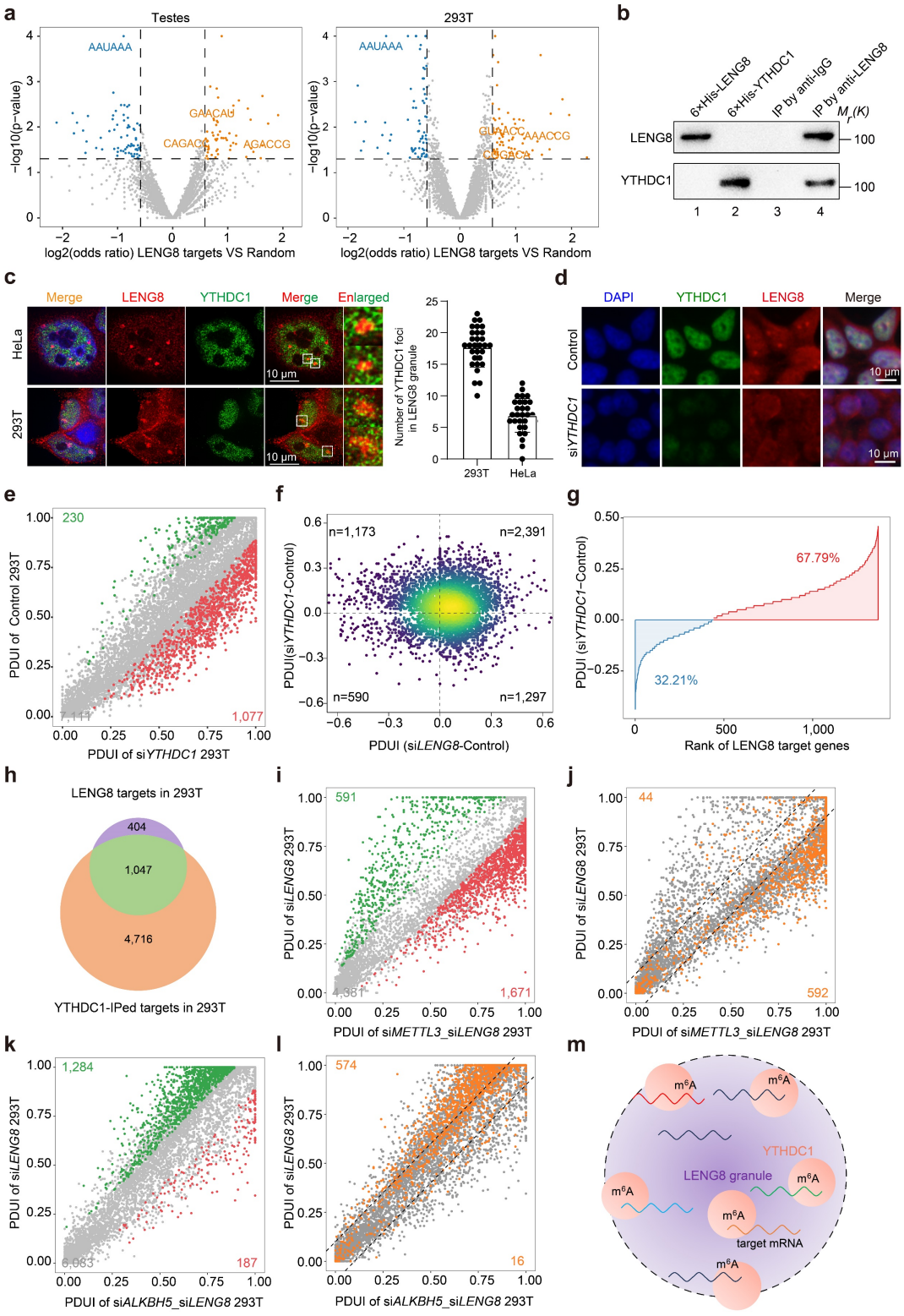


Fig. 6| YTHDC1-m^{6}A facilitates LENG8 granule formation and proximal poly (A) site usage. a, Scatter plot illustrating the odds ratio of hexamers near the pPASs in LENG8 targets from testes (left) and 293T (right) compared with non-LENG8 target random transcripts. P-value was calculated by Fisher's exact test. The orange letters represent the DRACH motif (D=A/G/T, R=A/G, H=A/C/T), where m^{6}A is installed. **b,** *In vitro* pull-down assay showing the direct interaction between LENG8 and YTHDC1.

Anti-LENG8 was used to pellet recombinant LENG8 proteins. IgG was used as the negative control. **c,** Immunofluorescence images of the relationship between YTHDC1 (green) and LENG8 (red) in 293T and HeLa cells. Enlarged: magnified views of the areas enclosed by dashed boxes. Right: quantitative summary of the number of YTHDC1 foci in LENG8 granules (mean ± SEM, n=30). **d,** The effect of *YTHDC1* knockdown on LENG8 granules in 293T cells. **e,** Scatter plots showing PDUI in *YTHDC1* knockdown compared with control 293T cells. 230 shortened and 1,077 lengthened transcripts are highlighted with green and red dots, respectively. **f,** Scatter plots showing PDUI in *YTHDC1* knockdown compared with *LENG8* knockdown 293T cells. **g,** Ranking the changes in PDUI of LENG8 target genes between *YTHDC1* knockdown and control 293T cells. A total of 32.21% shortened and 67.79% lengthened transcripts are highlighted with blue and red, respectively. **h,** Venn diagram illustrating the overlapping genes between LENG8 targets and genes identified in YTHDC1-IPed data in 293T cells. **i,** Scatter plots showing PDUI in *METTL3-LENG8* double knockdown compared with *LENG8* knockdown 293T cells. 591 shortened and 1,671 lengthened transcripts are highlighted with green and red dots, respectively. **j,** Scatter plots showing 44 LENG8 targets shortened and 592 lengthened transcripts in *METTL3-LENG8* double knockdown 293T cells, compared with *LENG8* knockdown. **k,** Scatter plots showing PDUI in *ALKBH5-LENG8* double knockdown compared with *LENG8* knockdown 293T cells. 1,284 shortened and 187 lengthened transcripts are highlighted with green and red dots, respectively. **l,** Scatter plots showing 574 LENG8 targets shortened and 16 lengthened transcripts in *ALKBH5-LENG8* double knockdown 293T cells, compared with *LENG8* knockdown. **m,** Schematic showing YTHDC1-m^{6}A target mRNAs in LENG8 granule.

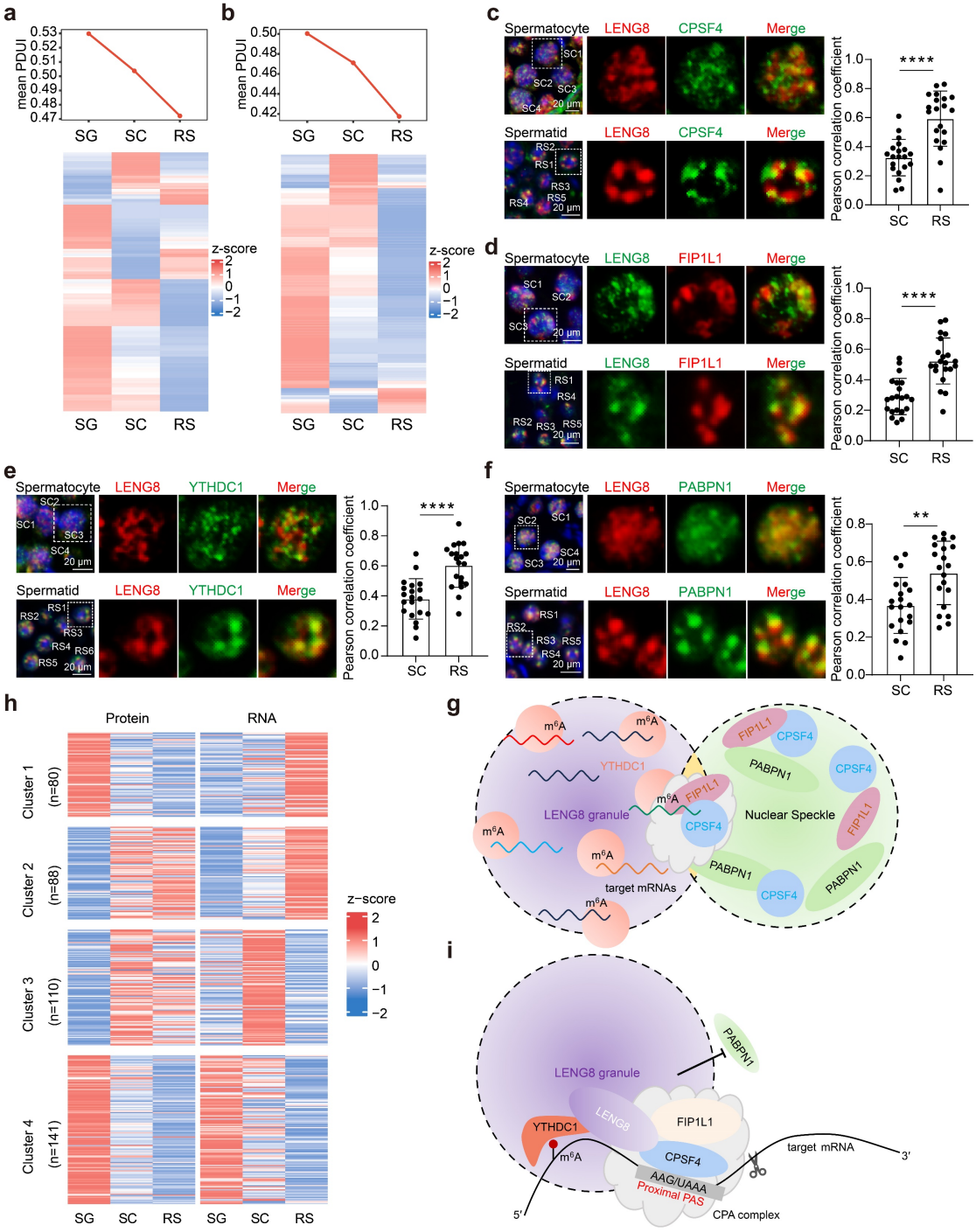
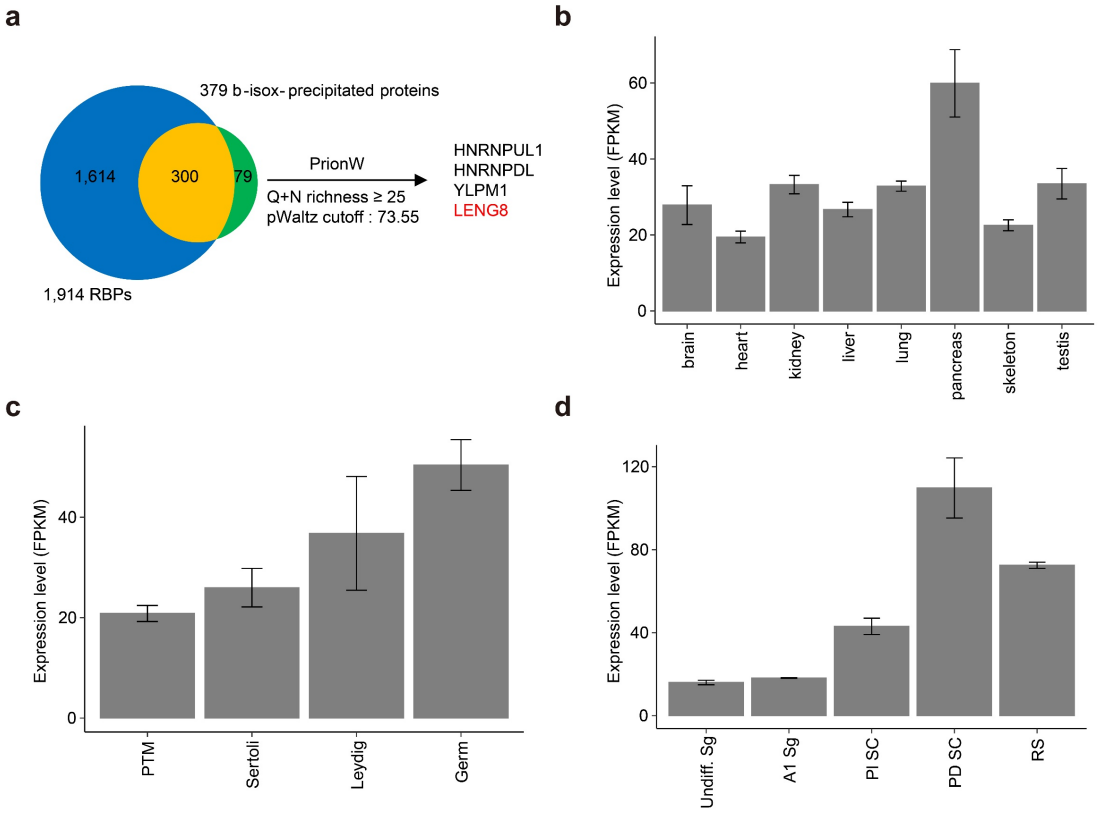
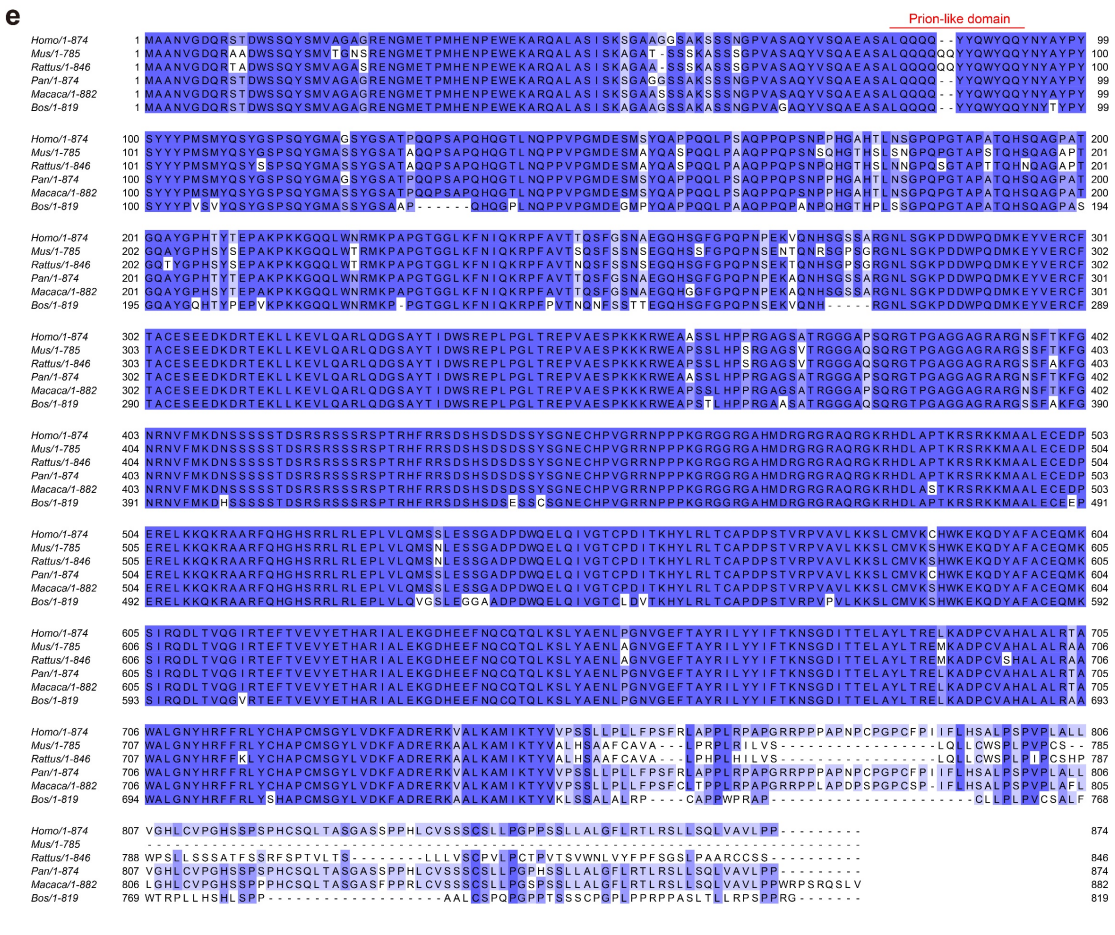


Fig. 7| LENG8 granules are involved in target 3′UTR shortening during spermatogenesis. a, b, Heatmap showing total (**a**) and LENG8 target genes (**b**) with shortened or lengthened 3′UTR in three types of germ cells. Line plot on the top showing the mean PDUI in spermatogonia (SG), spermatocytes (SC), and round spermatids (RS). **c-f,** Immunofluorescence images of the relationship between CPSF4 (green) and LENG8 (red) (**c**), LENG8 (green) and FIP1L1 (red) (**d**), LENG8 (red) and YTHDC1 (green) (**e**), LENG8 (red) and PABPN1(green) (**f**) in spermatocytes (top) and round spermatids (bottom). Nuclei were counterstained with DAPI (blue). Middle: magnified views of the areas enclosed by dashed boxes. Right: Pearson correlation coefficient showing co-localization patterns between LENG8 and CPSF4, FIP1L1,

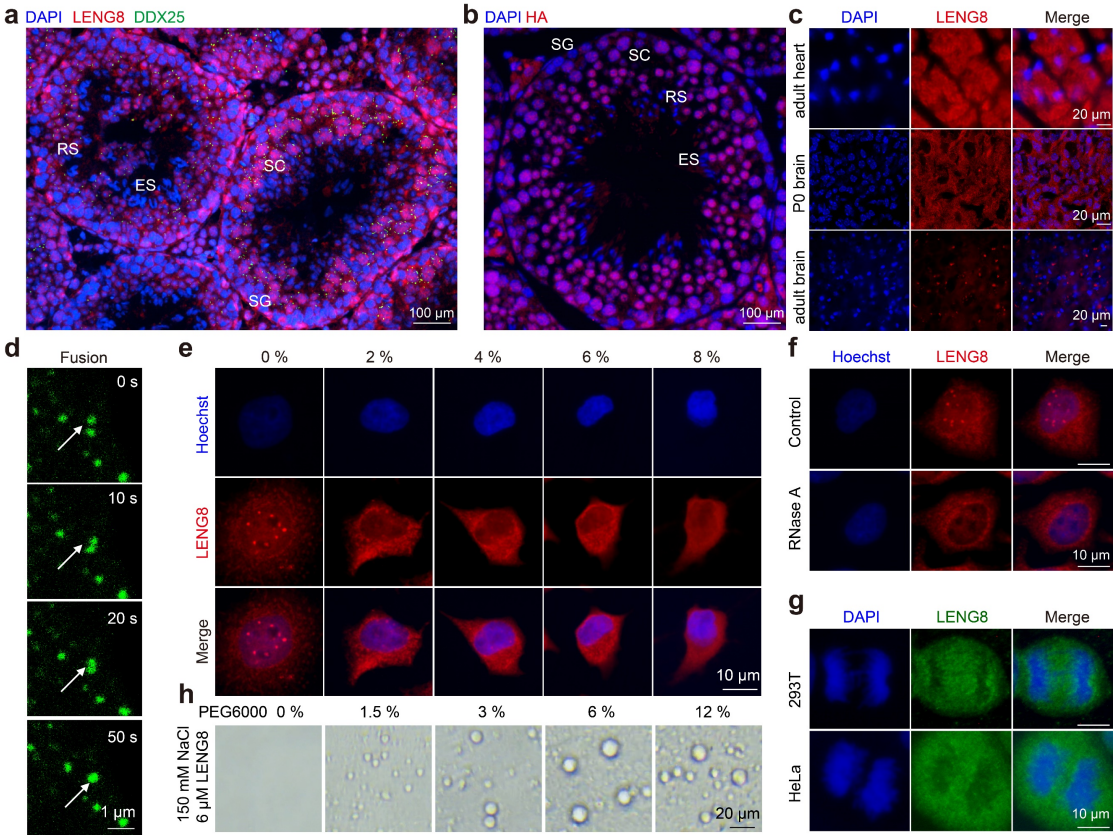
YTHDC1, and PABPN1. n=20, **P<0.01, ****P<0.0001 by Student's *t*-test. **g,** Schematic showing the relationship between LENG8 granule and nuclear speckle in male germ cells. **h,** Heatmap illustrating the protein abundance and RNA level of targets in four clusters across spermatogonia, spermatocytes and round spermatids. **i,** Model for LENG8 granule-mediated APA regulation. LENG8 granules compete with PABPN1 or directly promote CPA complex assembly at pPASs in a YTHDC1-m^{6}A dependent manner, ultimately enhancing the usage of these sites.





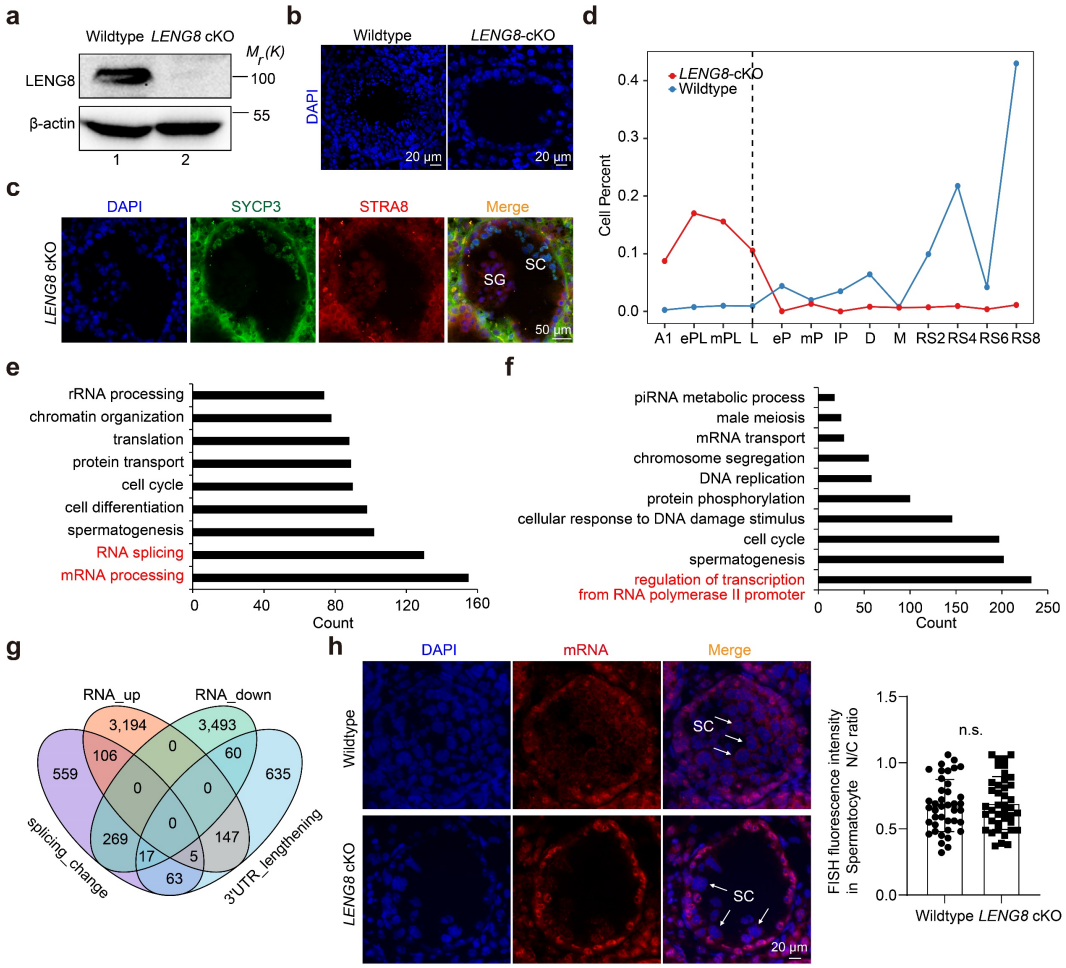
Extended Data Fig. 1| The expression pattern of prion-like domain-containing protein LENG8 in different cell types. a, PrionW predictions of prion-like domain-

containing proteins from 379 overlapped proteins. 1,914 mouse RBPs are defined through the use of computational analyses. **b,** FPKM-normalized expression levels for *LENG8* from different adult mouse tissues. FPKM, fragments per kilobase of transcript per million fragments mapped. **c,** FPKM-normalized expression levels for *LENG8* from germ cells and three testis somatic cell types. PTM: peritubular myoid cell. **d,** FPKM-normalized expression levels for *LENG8* from five spermatogenic stages cell types. Undiff. Sg: undifferentiated spermatogonia; A1 Sg: A1 spermatogonia; PI SC: preleptotene spermatocyte; P/D SC: pachytene/diplotene spermatocyte; RS: round spermatid. **e,** Amino acid sequence alignment of LENG8 from six different mammalian species. The alignment was generated with the Clustal series of programs and visualized with Jalview.

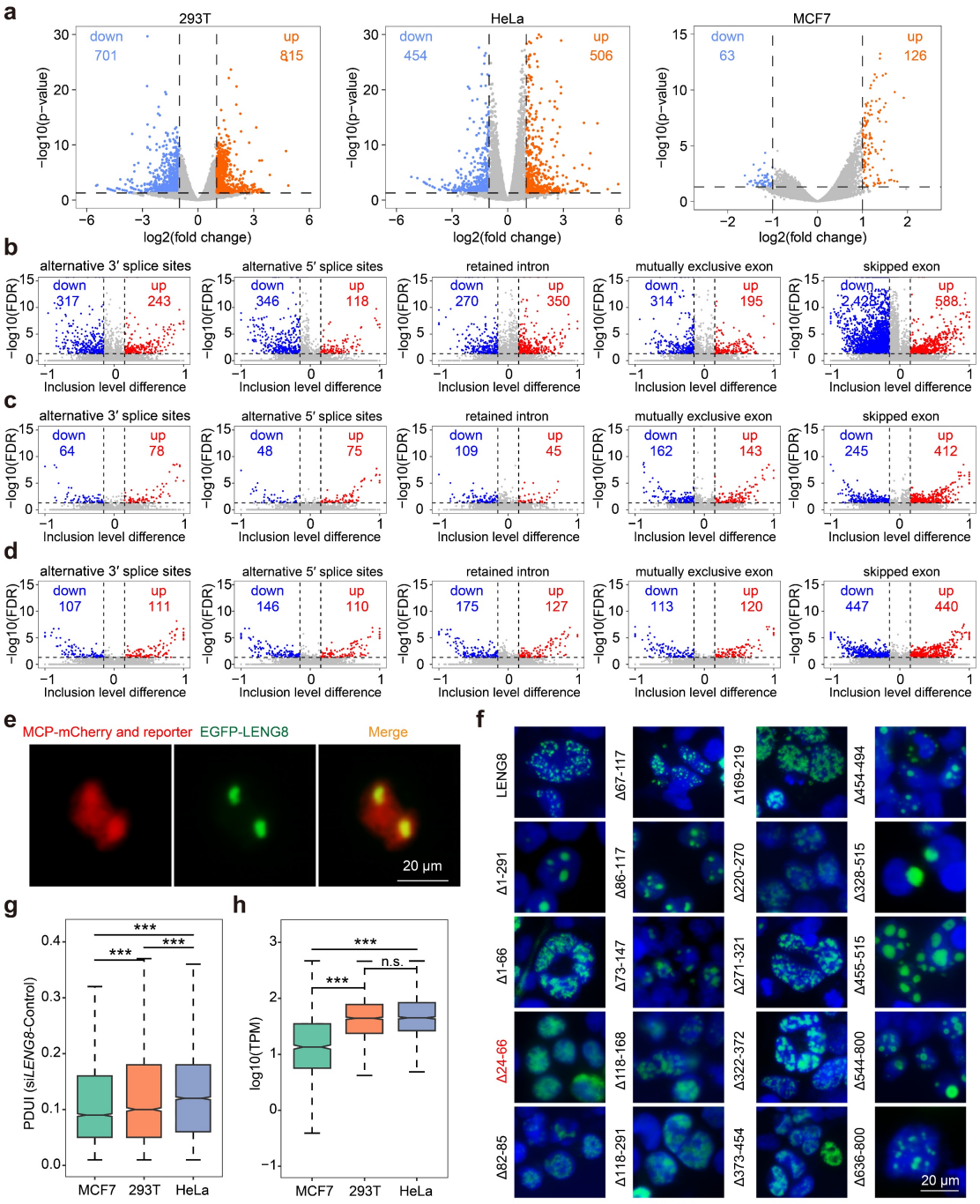


Extended Data Fig. 2| LENG8 forms condensates through LLPS in cells. a, Double immunostaining with LENG8 (red) and DDX25 (green) in adult testis sections. Nuclei were counterstained with DAPI (blue). DDX25 labels the germ granules in spermatocytes and round spermatids. SG: spermatogonia; SC: spermatocytes; RS: round spermatids; ES: elongating/elongated spermatids. Scale bar: 100 μm. **b,** Immunofluorescence image depicting HA-LENG8 localization in adult HA-*LENG8* knock-in mouse testis sections. Scale bar: 100 μm. **c,** Immunofluorescence images of LENG8 in heart and brain tissue cells. Scale bar: 20 μm. **d,** Images of two EGFP-LENG8 foci undergoing fusion upon in contact. Scale bar: 1 μm. **e,** The localization of HeLa endogenous LENG8 proteins is shown either without treatment or in the presence of 2%, 4%, 6%, 8% 1,6-Hexanediol for 10 min. Scale bar: 10 μm. **f,** Fluorescence images of endogenous LENG8 granules in HeLa cells treated with RNase A. **g,**

Endogenous LENG8 granules disperse at the metaphase of mitosis in both 293T and HeLa cells. **h,** Representative images of 6 μM recombinant LENG8 droplet formation at different PEG6000 concentrations.

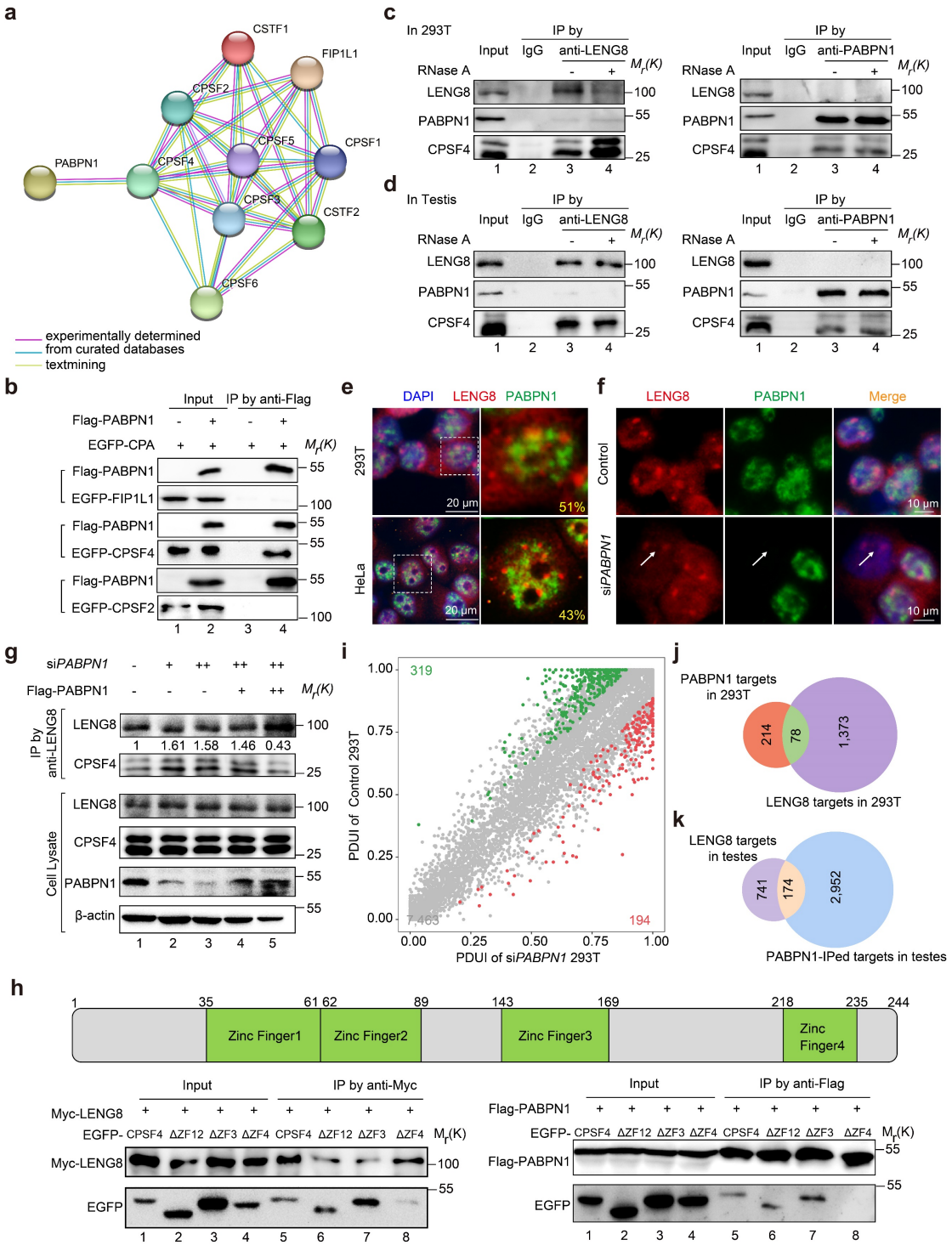


Extended Data Fig. 3| The depletion of *LENG8* leads to spermatogenesis arrest and disruptions in RNA metabolism. a, Western blot analysis of LENG8 expression from wildtype and *LENG8* cKO adult mouse testes. **b,** Immunostaining of testis sections from wildtype and *LENG8* cKO adult mice with nuclei counterstained by DAPI (blue). Scale bar: 20 μm. **c,** Immunostaining of SYCP3 (marker of spermatocytes, green) and STRA8 (marker of spermatogonia, red) showing spermatogenesis arrest at meiotic spermatocytes in *LENG8* cKO adult testes. Scale bar: 50 μm. **d,** Percentage of cells at different stages in wildtype and *LENG8* cKO mouse testes, respectively. **e,** GO term (biological process) enrichment analysis of LENG8-interacting proteins. **f,** GO term (biological process) enrichment analysis of genes with differential expression. **g,** Venn diagram of the intersection of genes from four datasets. **h,** FISH analysis of total mRNA localization in *LENG8* cKO and wildtype testis germ cells. Right: quantification of the ratio of mRNA fluorescence intensity between the nucleus and the cytoplasm. n= 40; n.s., not significant. P value by Student's *t*-test.



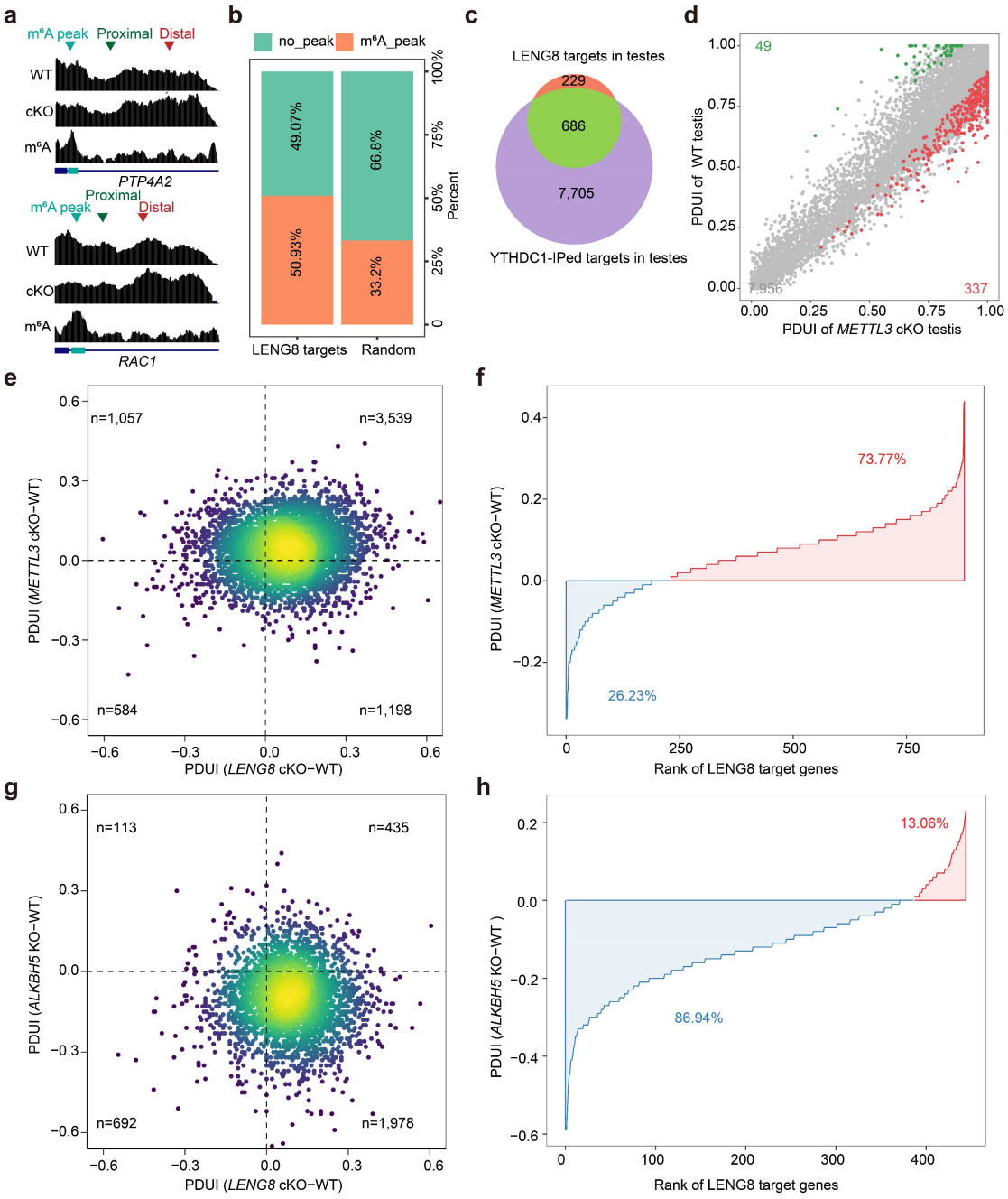
Extended Data Fig. 4| LENG8 granules play a significant role in APA regulation. **a,** Volcano plot describing differentially expressed transcripts in 293T, HeLa, and MCF7 cells with *LENG8* knockdown compared to the control group. Logarithmic fold changes in expression are plotted on the x-axis, with normalized p-values on the y-axis. Horizontal line shows the negative base 10 logarithmic p-values (0.05) cutoff and vertical lines denote the fold change cutoff (log2 foldchange≧1). Blue scatter dots represent downregulated genes and red scatter dots represent upregulated genes. **b-d,** Volcano plot depicting significant changes (-log10(FDR)) verse inclusion level difference in five alternative splicing types in 293T (**b**), HeLa (**c**), and MCF7 cells (**d**), including A3SS (alternative 3′ splice sites), A5SS (alternative 5′ splice sites), RI (retained intron), MXE (mutually exclusive exon), and SE (skipped exon) between

LENG8 knockdown and control cells. Red dots indicate genes with significantly increased inclusion levels and blue dots indicate those with significantly decreased inclusion levels. **e,** Representative images showing that the reporter mRNAs localize in the LENG8 granules. **f,** Representative images illustrating the intracellular localization of various LENG8 mutants under conditions of equivalent expression levels. **g,** Bar plot showing the comparison of PDUI changes between *LENG8* knockdown and control among 293T, HeLa, and MCF7 cells. **h,** Bar plot illustrating the abundance of LENG8 targets mRNA in the nucleus of 293T, HeLa and MCF7 cells.



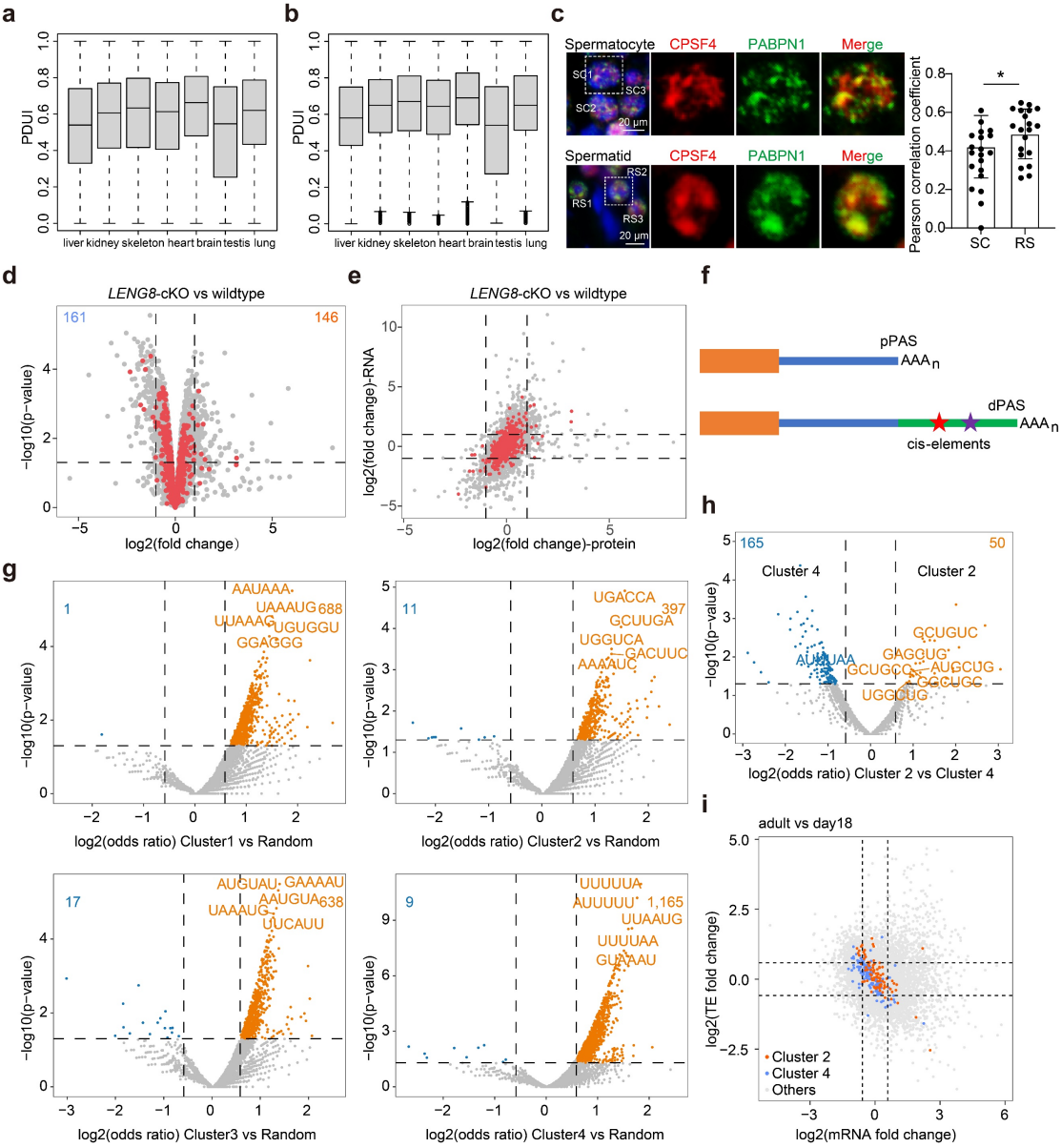
Extended Data Fig. 5| Competition between LENG8 and PABPN1 for binding to CPSF4 regulates the choice of poly (A) sites. a, STRING analysis the relationship of LENG8-interacting proteins identified using Mass spectrum with PABPN1 by Version 11.5.**b,** Co-IP assay of the interaction between Flag-PABPN1 and three EGFP-CPA factors co-transfected in 293T cells. **c, d,** Co-IP assay of the interaction between endogenous LENG8, PABPN1, and CPSF4 in 293T (**c**) and testes (**d**). Anti-LENG8 (left) and anti-PABPN1 (right) were used to pellet proteins from RNase A untreated (lane 3) or treated (lane 4) cell extracts, respectively. IgG was used as the negative

control. **e,** Immunofluorescence images of the relationship between PABPN1(green) and LENG8 (red) in cultured 293T and HeLa cells. Nuclei were counterstained with DAPI (blue). Right: magnified views of the areas enclosed by dashed boxes. The relative association between LENG8 granules and nuclear speckles were quantified (51% in 293T; 43% in HeLa). **f,** Immunofluorescence images of the effect of *PABPN1* knockdown on LENG8 granules in 293T cells. Arrows point to the cell with down-regulated expression of PABPN1. **g,** Co-IP assay of the influence of rescue expression of PABPN1 on the interaction between LENG8 and CPSF4. **h,** Co-IP results showing the fourth zinc finger (ZF4) domain of CPSF4 mediated the interaction between CPSF4 and LENG8 or PABPN1. Schematic diagram of four zinc finger domains in CPSF4 protein (top). Myc-LENG8 (left) or Flag-PABPN1 co-transfected with various CPSF4 mutants in 293T cells, and pelleted by anti-Myc or anti-Flag, respectively. **i,** Scatter plots showing PDUI in *PABPN1* knockdown and control 293T cells. A total of 319 shortened and 194 lengthened transcripts are highlighted with green and red dots, respectively. **j,** Venn diagram illustrating the overlapping genes between targets with shortened 3′UTRs upon *PABPN1* knockdown and LENG8 targets in 293T cells. **k,** Venn diagram illustrating the overlapping genes between LENG8 targets and genes identified in PABPN1-RIP data in the testes.



Extended Data Fig. 6| YTHDC1-m^{6}A is essential for LENG8 granule formation and the selection of proximal poly (A) sites. a, Integrative Genomics Viewer (IGV) tracks of two examples showing the correlation between m^{6}A peaks and poly (A) sites selection, comparing wildtype and *LENG8* cKO testes. PTP4A2 and RAC1 are LENG8 targets. **b,** Bar plot showed the percentage of genes' 3′UTR with m^{6}A peaks in LENG8 targets and random genes, respectively. **c,** Venn diagram illustrating the overlapping genes between LENG8 targets and YTHDC1-IPed targets in the testes. **d,** Scatter plots showing the PDUI in *METTL3* cKO and wildtype testes. A total of 49 shortened and 337 lengthened transcripts are highlighted with green and red dots, respectively. **e,** Scatter plots showing PDUI in *METTL3* cKO compared with *LENG8* cKO testes. **f,** Ranking the changes in PDUI of LENG8 target genes between *METTL3* cKO and wildtype testes. A total of 26.23% shortened and 73.77% lengthened transcripts are

highlighted with blue and red, respectively. **g,** Scatter plots showing PDUI in *ALKBH5* KO compared with *LENG8* cKO testes. **h,** Ranking the changes in PDUI of LENG8 target genes between *ALKBH5* KO and wildtype testes. A total of 86.94% shortened and 13.06% lengthened transcripts are highlighted with blue and red, respectively.



Extended Data Fig. 7| LENG8 granules contribute to the choice of proximal poly (A) sites during spermatogenesis. a, Box plot representing PDUI values across different tissues. Inside each box, the horizontal black bold lines denote the median values. Vertical extending lines represent adjacent values. **b,** Box plot representing PDUI values of LENG8 targets across different tissues. Dots denote observations outside the range of adjacent values. **c,** Immunofluorescence images of the relationship between PABPN1 (green) and CPSF4 (red) in spermatocytes (top) and round spermatids (bottom). Nuclei were counterstained with DAPI (blue). Middle: magnified views of the areas enclosed by dashed boxes. Right: Pearson correlation coefficient showing co-localization pattern between PABPN1 and CPSF4. SC: spermatocyte; RS: round spermatid. n=20, *P<0.05, ****P<0.0001 by Student's *t*-test. **d,** Volcano plot

illustrating the changes in protein levels between *LENG8* cKO and wildtype mouse testes. Horizontal line shows the negative base 10 logarithmic p-values (0.05) cutoff and vertical lines denote the fold change cutoff (log2FC≧1). Red points represent detected LENG8 targets. 161 dots represent downregulated proteins and 146 dots represent upregulated proteins. **e,** Integrated analysis of RNA and protein levels between wildtype and *LENG8* cKO mouse testes. Red points represent detected LENG8 targets. **f,** Schematic diagram showing transcripts with alternative poly (A) sites, as well as cis-regulatory elements in the sequences between the pPAS and dPAS. **g,** Scatter plot illustrating odds ratio of hexamers in the sequences between pPAS and dPAS of four clusters compared with random transcripts (n=1,000). Top5 hexamers with the minimum p-value are labelled in orange letters. **h,** Scatter plot illustrating odds ratio of hexamers in the sequences between pPAS and dPAS of cluster 2 targets compared with cluster 4 targets. **i,** Scatter plot showing mRNA level changes (x-axis) against translational efficiency (TE) changes (y-axis) between 18-day and adult testes. Orange dots represent cluster 2 targets, and blue dots represent cluster 4 targets.