LENG8 regulation of mRNA processing is responsible for the control of mitochondrial activity

Yongxu Zhao
SIBS CAS: Chinese Academy of Sciences Shanghai Institutes of Nutrition and Health
Xiaoting Wang
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Yuenan Liu
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Niannian Li
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Shengming Wang
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Zhigang Sun
Jinan Central Hospital Affiliated to Shandong University
Jingyu Zhu
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Zhenfei Gao
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Xiaoxu Zhang
Fudan University
Linfei Mao
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Jian Guan
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Hongliang Yi
Shanghai 6th Peoples Hospital Affiliated to Shanghai Jiaotong University
Qiurong Ding
Shanghai Institutes of Nutrition and Health CAS: Chinese Academy of Sciences Shanghai Institutes of Nutrition and Health
Feng Liu (✉ liufeng@sibs.ac.cn)
Shanghai Sixth People's Hospital https://orcid.org/0000-0002-4442-7028
Nan Zhang
Jinan Central Hospital Affiliated to Shandong University

Research

Keywords: mRNA processing, LENG8, TREX, RNA-IP, mitochondria

Posted Date: November 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1063286/v1

LENG8 regulation of mRNA processing is responsible for the control of mitochondrial activity

Yongxu Zhao[1,3,*], Xiaoting Wang[1,2,*], Yuenan Liu[1,2], Niannian Li[1,2], Shengming Wang[1,2], Zhigang Sun[4], Jingyu Zhu[1,2], Zhenfei Gao[1,2], Xiaoxu Zhang[1,2], Linfei Mao[1,2], Jian Guan[1,2], Hongliang Yi[1,2,#], Qiurong Ding[3,#], Feng Liu[1,2,5,#] , Nan Zhang[4,#]

1. Department of Otolaryngology-Head and Neck Surgery, Otolaryngology Institute of Shanghai Jiao Tong University, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China.
2. Shanghai Key Laboratory of Sleep Disordered Breathing, Shanghai 200233, China.
3. CAS Key Laboratory of Nutrition, Metabolism and Food Safety, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China.
4. Department of Oncology, Jinan Central Hospital, Cheeloo College of Medicine, Shandong University, Jinan 250013, Shandong Province, China.
5. Lead contact

* These author contributed equally to this work.

Correspondence: Dr. Liu Feng liufeng@sibs.ac.cn or liufeng_ent@yeah.net; Dr. Ding Qiurong qrding@sibs.ac.cn, Dr. Zhang Nan zlkzn2016@126.com and Dr. Yi Hongliang yihongl@126.com.

Abstract

Background: The processing of mRNA is essential for the maintenance of cellular

and tissue homeostasis. However, the precise regulation of this process in mammalian cells remains largely unknown.

Method: Bioinformatic alignment tools and Structure modeling were applied to study the evolutionary conservation between LENG8 and its homologs. Tandem affinity purification and co-immune-precipitation approaches were applied to study the LENG8-associated proteins. RNA-precipitation was performed to analyze the RNA molecules bound by LENG8. Morphology and respiration activity of mitochondria from wildtype or *LENG8*-deficient cells were also measured. *Leng8* were deleted in mouse adipose tissues to study the gene effect on adipogenesis.

Results: Here we have found that LENG8 represents the mammalian orthologue of the yeast mRNA processing factor Thp3 and Sac3. We go on to demonstrate that LENG8 binds to mRNAs, associates with components of mRNA processing machinery (the TREX complex) and contributes to mRNA nuclear export to the cytoplasm. Loss of *LENG8*, leads to aberrant accumulation of poly (A) RNA in the nucleus, in both HeLa cells and murine fibroblasts. Furthermore, the precipitation of LENG8 is associated with an enrichment of both mRNAs and lncRNAs, and approximately half of these are also bound by the TREX component, THOC1. However, LENG8 preferentially binds mRNAs encoding for mitochondrial proteins and depletion of this processing factor, causes a dramatic breakdown in mitochondrial ultrastructure and a reduction in mitochondrial respiratory activity. Conditional deletion of *Leng8* in mouse adipose tissues lead to a decreased body weight, and increased adipose thermogenesis.

Conclusion: Our work has found an evolutionarily conserved mRNA processing factor that can control mitochondrial activity.

Key words: mRNA processing, LENG8, TREX, RNA-IP, mitochondria

Introduction

In eukaryotes, the flow of genetic information from DNA to protein requires the correct coupling of RNA transcription and subsequent mRNA processing including steps such as 5'end capping, splicing, 3'end cleavage and polyadenylation, as well as with RNA export. As the nascent pre-mRNA emerges from the RNA polymerase II (RNAPII), it is packed in a messenger ribonucleoparticle (mRNP) whose optimal configuration is critical for the normal pre-mRNA processing and mRNA export (Lee and Tarn, 2013; Wickramasinghe and Laskey, 2015). The biogenesis of mRNP is tightly regulated by the THO complex, which is highly conserved from yeast to mammals (Katahira and Yoneda, 2009; Yuan et al., 2018). In yeast, THO complex are consisted of four tightly interacting subunits Hpr1, Tho2, Mft1, and Thp2. Multiple evidences support a relationship between THO and mRNA processing, as mutations or deletions of each subunit results in defects of mRNA export (Jimeno et al., 2002; Strasser et al., 2002). The core THO complex moreover physically associates with two additional export factors, Aly/Yra1 and RNA-dependent ATPase Sub2/UAP56 and recruits them to the nascent mRNA to form a larger RNA-protein complex termed TREX (Transcription Export) (Xie and Ren, 2019; Zenklusen et al., 2001).

In addition to TREX, another protein complex stepwise involved in regulation of mRNA processing is TREX-2, which are composed of Thp1, Sac3, Sus1, and Cdc31 (Garcia-Oliver et al., 2012; Stewart, 2019). As seen with the THO complex, loss of TREX-2 components in yeast leads to defects in mRNA export. Unlike THO complex, which is primarily associated with the active chromatin, TREX-2 complex, which is at the relative downstream steps of the mRNA processing, locates primarily at the nuclear periphery in association with nuclear pore complex. By biochemistry and genetic analysis approaches, several more proteins have been identified as the mediator of mRNA processing in addition to THO and TREX-2 complex, including Mex67 and Mtr2, which bridges THO and TREX-2 complex (Rondon et al., 2010).

As expected from their fundamental and essential roles in gene expression, both the TREX and TREX-2 complex are conserved from budding yeast to human (Cheng et al., 2006; Dominguez-Sanchez et al., 2011). The human orthologous counterpart of

the THO subunits Hpr1, Tho2 and Tex1 are named THOC1, THOC2 and THOC3, respectively (Kumar et al., 2015; Li et al., 2005; Masuda et al., 2005). However, human THO has three additional subunits THOC5, THOC6 and THOC7 that have no yeast homologs (Pierce et al., 2008; Saran et al., 2016; Tran et al., 2014). Furthermore, as the same with the Yra1 in yeast, it is noted that the human homologs THOC4/ALYREF, also plays essential function to regulate mRNA maturation (Chi et al., 2013; Fan et al., 2019; Shi et al., 2017). However, how the mRNA processing is precisely regulated in mammalian cells still remains largely nebulous.

The super-helical PCI domain, which is firstly found in and named for multi-subunit complexes **p**roteasome, **C**SN and e**I**F3, serves as the principal scaffold for Thp1-Sac3 duo in the TREX-2 complex (Khoshnevis et al., 2014; Kragelund et al., 2016). The Thp3-Csn12 minicomplex, which is also identified as a PCI complex, has been found to regulate transcriptional elongation and mRNA maturation in yeast (Jimeno et al., 2011; Kragelund et al., 2016), but its precise mechanism and the potential existence of homologous proteins in mammals remains unknown. Using bioinformatic alignment tools, we have identified Leukocyte Receptor Cluster Member 8 (LENG8) as the mammalian homologue of Thp3 and found that it can associate with PCI domain containing 2 (PCID2), the mammalian equivalent of Csn12. LENG8 can bind mRNAs associate with the mRNA processing machinery, causing the attenuation of mRNA export. It preferentially binds to mRNAs encoding for proteins that localize to the mitochondria and its activity is required for the maintenance of mitochondrial morphology and respiratory function. This work has revealed evolutionarily conserved mRNA processing machinery, which can control mitochondrial activity.

Materials and Methods

Constructs and Antibodies.

The nucleotide sequence used for *LENG8*, *PCID2*, *THOC1*, *THOC2*, *THOC3*, *THOC5*, *THOC6*, *THOC7* overexpression were *NM_052925.3*, *BC016614*, *BC010381*, *NM_001081550.1*, *NM_032361.3*, *NM_003678.4*, *NM_001142350.1* and

NM_025075.3. All of these constructs were purchased from Sinobiological (Beijing).

The mouse anti-LENG8 were produced by Daian Biotechnology (Wuhan, Hubei) using a recombinant LENG8 fragment (1-300 aa). The rabbit anti-PCID2 were from Abcam; the rabbit anti-THOC1, THOC2, THOC5, ALYREF, Myc-tag were from Abclonal (Wuhan, Hubei); the mouse anti-GFP were from Roche; the mouse anti-FLAG was from Genscript (Nanjing, Jiangsu). FITC-conjugated goat anti-mouse, Cy3 and Cy5 conjugated goat anti-rabbit were from Beyotime biotechnology (Shanghai); HRP conjugated goat anti-rabbit and goat anti-mouse were from Abclonal (Wuhan, Hubei).

Anti-FLAG agarose beads and streptavidin-affinity magnetic beads were from Genscript (Nanjing, Jiangsu); Protein A/G magnetic beads were from Thermo Fisher Scientific.

Cell culture and transfection.

HEK293T and HeLa cells were obtained from cell bank of Shanghai Institute for Biological Sciences (CAS) and maintained in standard conditions (37℃, 5% CO_2) in DMEM/High Glucose medium with 10% FBS, 100 U/ml penicillin and100 mg/ml streptomycin (Life Science). All cell lines have been tested negative for mycoplasma contamination.

Cells were seeded the day before transfection. The next day, when cells were 70%–80% confluent, the medium was changed to penicillin- and streptomycin-free medium. DNA in Opti-MEM (Life Technologies) was mixed with Lipofectamine 2000 or Lipofectamine 3000 (Invitrogen) and then incubated for 20 min at room temperature, then added dropwise to the cells. The medium was changed to complete medium after 6 h and the cells were used in experiments at 48h-72h post-transfection.

shRNA mediated gene knockdown and gRNA-Cas9 mediated gene editing.

For shRNA mediated gene knockdown, more than two effective shRNA clones to each target coding sequence were prepared. shRNA sequences were cloned into the

lentiviral expression plasmid pLKO.1 and transfected into HEK293T cells to generate recombinant lentiviruses. HeLa cells were transduced with the lentiviral supernatants and selected with 1 mM puromycin (MCE). RT-qPCR and western blot analysis was performed to verify significant depletion of each target sequence. The shRNA sequences used are listed in **Supplementary Table 6.**

The *Leng8*$^{-/-}$ mice were generated through the CRISPR/Cas9 method as described previously. Briefly, in vitro-translated Cas9 mRNA and gRNA were co-microinjected into the C57BL/6 zygotes. The pair of gRNA sequences used to generate the knockout mice is GCTATGTGCCACCTTCAGCT and ACTAGGACATGCTAATGTCC. Founders with frame shift mutations were screened by DNA sequencing. One F0 founder, of which 937bp fragment containing exon3 and exon4 of *Leng8* loci was deleted, was crossed with C57BL/6 wildtype and 11 F1 mice were got for the *Leng8*$^{-/-}$ mice. One of the F1 mice was chosen to backcross to the WT mice more than 10 generations to maintain the strain. All protocols were approved by the local ethics committee of Shanghai JiaoTong University Affiliated Sixth People's Hospital.

Immunofluorescence staining and *in situ* hybridization of cultured cells.

Cells were cultured on coverslips or glass bottom dishes, fixed with 4% PFA, permeabilized with 0.1% Triton X-100 in PBS and blocked with 1% BSA. For immunofluorescence staining, cells were incubated with primary antibodies at 4℃ overnight and then secondary antibody or DAPI at room temperature for 30 min. Antibodies were used at the following dilutions: rabbit polyclonal anti-PCID2, THOC1, THOC5, ALYREF, 1:200; mouse monoclonal anti-LENG8, 1:100; Cy3-conjugated goat anti-rabbit IgG, 1:500; and FITC-conjugated goat anti-mouse IgG, 1:500. For *in situ* hybridization, cells were incubated with 5 μM Cy3 labeled oligo-dT (70) in 2× SSC buffer at 42 °C overnight. Samples were examined and the figures were acquired with an LSM 710 confocal laser-scanning microscope (Carl Zeiss, Oberkochen, Germany) at 63 × or 100 × magnification.

RNA-Immunoprecipitation and Sequencing.

HeLa cells expressing GFP-tagged human *THOC1* or *LENG8* were pelleted by centrifugation at 500g for 10 min at 4 °C and washed twice with ice-cold PBS. Cells were lysed in an equal volume of RIP lysis buffer (10 mM HEPES pH 7.0, 100 mM KCl, 5 mM $MgCl_2$, 25 mM EDTA, 0.5% (v/v) Nonidet-P40, 1 mM dithiothreitol, protease inhibitor cocktail (EDTA-free, Beyotime) for 30 min on ice in the presence of 100 U ml^{-1} RNase inhibitor (Sangon)) and lysates clarified by centrifugation at 9,000g and 4 °C for 10 min. Clarified lysates were incubated with anti-GFP with a final concentration at 0.2ug ml^{-1} and the binding reactions were conducted for 2 hours at 4 °C with continuous gentle rotation. The reaction mixture was centrifuged at 2000 × g at 4 °C for 5 minutes to remove debris. The ChIP grade protein A/G magnetic beads were washed with RIP binding/wash buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM MgCl2, 0.05% (v/v) Nonidet-P40) containing 25 mM EDTA, protease inhibitors and 100 U ml^{-1} RNase inhibitor for 3 times. The mixture of RNA-protein complexes was added to 40 µl of 100 % beads, and then the binding was conducted overnight at 4 °C on a rotary wheel, followed by five washes with RIP binding/wash buffer. Beads were then resuspended in Trizol (Invitrogen) and RNA was isolated according to the manufacturer's instructions. Two RNA immunoprecipitations per bait were carried out in parallel. RNA quality was assessed on a Genetic Analyzer (Agilent) and TruSeq RNA library construction and next-generation sequencing were performed by the Lianchuan Biotechnology (Hangzhou, Zhejiang). All samples were sequenced on an Illumina HiSeq2500 platform at 15 million 100-bp single reads per sample. After quality control of the sequencing libraries, reads were trimmed and mapped against the Ensembl genome annotation and the human genome assembly (hg19/GRCh38) using Tophat2. Reads mapping to ribosomal RNAs or the mitochondrial genome were removed. RNAs binding to THOC1 or LENG8 were identified by differential quantification (bait over control) against the Ensembl genome annotation using cuffdiff from the cufflinks package. RNAs with fold changes >2, FDR corrected P values <0.01 and minimal read counts of 10 were considered as enriched. To discover preferences of THOC1 and LENG8 for different RNA species, we extracted RNA types and gene-model-related features from the

Ensembl annotations and plotted them using custom scripts.

Affinity purifications.

For affinity purifications with biotin-labelled mRNA, HeLa cells were pelleted and lysed in an equal volume of AP lysis buffer (10 mM HEPES pH 7.0, 150 mM KCl, 5 mM $MgCl_2$, 25 mM EDTA, 0.5% (v/v) Nonidet-P40, 1 mM dithiothreitol, protease inhibitor cocktail (EDTA-free, Beyotime) for 30 min on ice in the presence of 100 U ml^{-1} RNase inhibitor (Sangon)) and lysates clarified by centrifugation at 9,000g and 4 °C for 10 min. Clarified lysates were incubated with biotin conjugated oligo-dT(25) or non-biotin conjugated control oligo-dT with a final concentration at 5 μM and the binding reactions were conducted for 2 hours at 4 °C with continuous gentle rotation. The reaction mixture was centrifuged at 2000 x g at 4 °C for 5 minutes to remove debris. Streptavidin-affinity magnetic beads were washed with AP binding/wash buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM MgCl2, 0.02% (v/v) Nonidet-P40) containing 25 mM EDTA, protease inhibitors and 100 U ml^{-1} RNase inhibitor for 3 times. The mixture of RNA-protein complexes was added to 20 μl of 100 % beads, and then the binding was conducted overnight at 4 °C on a rotary wheel, followed by five washes with AP binding/wash buffer. The RNA-protein coated beads were boiled in 50 μl 1 x SDS loading buffer, and subjected to SDS–polyacrylamide gel electrophoresis and western blot analysis.

Tandem affinity purification of LENG8 complex

The tandem affinity purification strategy to fractionate the LENG8 complexes from HeLa cells was performed as previously (Tsai and Carstens, 2006). Briefly, LENG8 sequence followed by 3×FLAG, TEV site and dual protein-A (ProtA) immunoglobulin G binding domains (ZZ) was inserted into pCDH-EF1-MCS-T2A-PuroR (#CD520A-1. System Biosciences). The stable cell line capable of expressing LENG8–3×FLAG-ZZ was obtained. Thus, the cells were grown in DMEM with 10% FBS plus 1% penicillin/streptomycin and harvested near confluence. The cell pellet was washed with chilled PBS for three times and then

lysed in an equal volume of TAP lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 25 mM EDTA, 0.5% (v/v) Nonidet-P40, 1 mM dithiothreitol, 1× protease inhibitor cocktail) for 30 min on ice. The homogenate was centrifuged for 20 min at 10,000 g. The supernatant was transferred to a fresh tube. Then 50 μL of packed IgG beads was added to the 4 mg protein extract, followed by gentle rotation overnight at 4 °C and then washed by TAP lysis buffer of 50 mM KCl for three times. The bound protein was eluted by TEV protease cleavage and further purified by anti-FLAG -conjugated beads. The final eluates from the FLAG beads with FLAG peptide were resolved by SDS/PAGE on a 4–12% gradient gel and visualized by silver staining. Specific bands were cut off and subjected to mass spectrometry analysis. The protein interaction network analysis was performed using QIAGEN'S Ingenuity Pathways Analysis (QIAGEN'S Ingenuity pathway analysis, Ingenuity Systems, http://www.qiagen.com/ingenuity, version 52912811).

Nucleocytoplasmic separation, RNA isolation and sequencing.

Cells were washed with cold PBS, and then incubated at -20 °C for 5 minutes. Subsequently, buffer A containing 10 mM HEPES (pH 7.9), 1.5 mM MgCl2, 10 mM KCl, 0.5mM dithiothretol (DTT), and 1 mM PMSF was added. The cytoplasmic and nuclear fractions were separated by centrifugation at 17,000 × g for 30 minutes at 4 °C. Both of the cytoplasmic and nuclear fractions were then resuspended in Trizol (Invitrogen) and RNA was isolated according to the manufacturer's instructions. All samples were sequenced on an Illumina HiSeq2500 platform at 15 million 100-bp single reads per sample. After quality control of the sequencing libraries, reads were trimmed and mapped against the Ensembl genome annotation and the human genome assembly (hg19/GRCh38) using Tophat2. Reads mapping to ribosomal RNAs or the mitochondrial genome were removed.

Western blot

Proteins were lysed from cells using RIPA buffer containing 10 mM Tris-Cl, pH 8.0, 150 mM NaCl, 1% Triton X-100, 1% Na-deoxycholate, 1 mM EDTA, 0.05% SDS

and fresh 1× proteinase inhibitor. The protein concentration was determined via the Bradford method using the Bio-Rad protein assay before proteins were equally loaded and separated in polyacrylamide gels. The proteins were then transferred to a nitrocellulose filter membrane (Millipore) and were incubated overnight with indicated primary antibodies. HRP-conjugated secondary antibodies were then applied to the membrane, and the Western blot signal was detected using auto-radiographic film after incubation with ECL (GE Healthcare) or SuperSignal West Dura reagents (Thermo Scientific).

RT-qPCR

For detection of *LENG8* and *PCID2*, total RNA from HeLa cells was extracted with TRIzol reagent according to the manufacturer's instructions (Invitrogen). One microgram of total RNA was reverse transcribed using the ReverTra Ace® qPCR RT Kit (Toyobo, FSQ-101) according to the manufacturer's instructions. A SYBR RT-PCR kit (Toyobo, QPK-212) was used for quantitative real-time PCR analysis. The relative mRNA expression of different genes was calculated by comparison with the control gene Gapdh (encoding GAPDH) using the 2-△△Ct method. The sequences of primers for the qPCR analysis are shown in **Supplementary Table 6**.

Mitochondrial activity studies.

For OCAR measurement, an XFe24 Extracellular Flux analyzer (Agilent) was used to determine the bioenergetic profile of wildtype and *LENG8* deficient HeLa cells. HeLa cells were plated at 1,000,000 cells per well in XFe24 plates 24 before the Mito Stress Tests. All assays were performed following manufacturer's protocols. Results were normalized to cell number.

For mitochondrial mass measurement, wildtype and *LENG8* deficient HeLa cells were incubated with Mitotracker green and Mitotracker deep red at 50nM for 30 min at 37 ℃. Mitochondria-associated ROS levels were measured by staining cells with MitoSOX at 2.5 μM for 30 min at 37 ℃. Mitochondria membrane potential was measured using the kit from Invitrogen and performed according to the

manufacturer's instructions. Cells were then washed with PBS solution and resuspended in cold PBS solution containing 1% FBS for FACS analysis.

Statistical analysis

The results are represented as the mean ± s.e.m., and statistical significance between groups was determined using an unpaired t-test or the Mann-Whitney U test. GraphPad Prism software 8.0 was used for all analyses, and a *$p<0.05$ was considered statistically significant.

Results

Identification of LENG8 and PCID2 as the mammalian orthologues of the yeast Thp3-Csn12 complex.

To further understand the process of mRNP biogenesis, bioinformatics analyses were performed on the yeast mRNA maturation factors and their mammalian orthologues (**Fig. 1A** and **B**, **Supplementary Table 1**). The yeast Thp3-Csn12 complex is a newly identified protein complex whose activity was required for transcriptional elongation and mRNA processing, however the exact role of Thp3-Csn12 complex or their mammalian counterparts in mRNP biogenesis remains unknown. PCID2, which was previously identified as the mammalian homologous protein of THP1 or SUS1 subunit of yeast TREX2 complex, exhibited much more similarity to CSN12, which also contains PCI domain (Bhatia et al., 2014) (**Fig. 1A, B** and **D**, **Supplementary Table 1** and **Supplementary Fig. 1**). Furthermore, using bioinformatic alignment, we identified LENG8, CG6700 and the Hypersensitive to Pore-forming toxin (HPO-10) as the orthologues of THP3 in mammals, drosophila and nematodes, respectively (**Fig. 1C** and **E**, **Supplementary Table 1** and **Supplementary Fig. 2**). Human LENG8 protein contains three distinct domains: the N-terminal topoisomerase II-associated protein (PAT1) region, mid-HJURP (Holliday junction recognition protein)-associated repeat domain and the C-terminal atypical SAC3-subytpe of PCI domain. All orthologues of THP3 contain this subtype of PCI domain and using algorithmic

analysis, it has been found that the amino acid sequence at the PCI domain, displays 71% similarity between THP3 and LENG8, across all eukaryotes (**Fig. 1C, E & F**). Similar to the PCI domain in SAC3 and THP3, the LENG8 equivalent consists of curved helical repeats, terminating in a globular winged helix(WH) subdomain (**Fig. 1F**). Taken together, these results suggested that eukaryotic LENG8 is highly conserved.

We also found that the ectopically expressed LENG8, as well as the endogenous LENG8, are predominantly located in the cell nucleus, especially nucleosomes and nuclear speckles (**Supplementary Fig. 3A** and **B**). However, the truncated mutant LENG8 (1-560) which lack the PCI domain was localized in the cytosol, indicating that the CPE domain were pivotal for LENG8's nuclear localization (**Supplementary Fig. 3C** and **D**). However, the true functional relevance of this domain in LENG8 remains unknown.

Here, we hypothesized that LENG8 also associated with PCID2 and contributed to mRNA export. To test this, FLAG-tagged LENG8 were co-expressed with MYC-tagged PCID2 in HEK293T cells, and co-immunoprecipitation (coIP) experiment was performed to detect the interaction between LENG8 and PCID2, and the results showed that PCID2 were detected in the pellet fraction of FLAG-LENG8 (**Fig. 2A**). Furthermore, poly (A) RNP complex were purified using biotin labeled oligo-dT and then streptavidin conjugated beads and sent to western blotting analysis, the results showed that LENG8, PCID2, as well as THOC1 and ALYREF, two of the dominant THO components, were detected in the fraction of mRNP (**Fig. 2B**). We therefore performed fluorescence *in situ* hybridization (FISH) analysis of poly (A) mRNA using Cy3-labeled oligo-dT and the results showed that LENG8 was localized with PCID2, and that both of them co-localized with poly (A) RNAs (**Fig. 2C-E**). Taken together these results indicated that LENG8 and PCID2 associated with RNAs.

To verify the functional role of *LENG8* in mRNA export, we used shRNAs to inhibit *THOC1*, or *LENG8* expression in HeLa cells, and the FISH analysis results showed that in *THOC1* or *LENG8* deficient cells, the cytosolic mRNAs were markedly

decreased compared to the control cells, in the meantime, the nuclear mRNAs were aberrantly accumulated (**Fig. 2F**). We then generated *Leng8* knockout mice using the CRISPR/Cas9 genome editing method (Wang et al., 2013) and isolated the tail fibroblast from wild type and *Leng8*$^{+/-}$ mice. Likewise, *Leng8*$^{+/-}$ fibroblast exhibited accumulation of poly (A) mRNAs in nucleus compared to wild type cells (**Fig. 2G**). Furthermore, HeLa cells that transfected with a pool of small interference RNAs targeting *PCID2*, exhibited aberrant accumulation of poly (A) mRNA in nucleus (**Fig. 2H**). Taken together, these results suggested that LENG8 and PCID2 were evolutionarily conserved mRNA processing factors and regulated mRNAs exporting in mammalian cells.

LENG8 associates with mRNA processing factors.

To identify LENG8 associated proteins in mRNA factors, we generated a cell line from human HeLa cells that is stably transfected with ZZ-LENG8-FLAG (Sun et al., 2008). The tagged LENG8 was purified from cell extracts by sequential affinity chromatography steps, and the final FLAG peptide elutes was subjected to 4–12% gradient SDS-PAGE and visualized by silver staining. The indicated bands were excised and analyzed by mass spectrometry (**Fig. 3A** and **Supplementary Table 2**).

The results showed that LENG8 significantly enriched for 243 proteins that are mostly nuclear protein (**Fig. 3B**) and related to mRNA processing, particularly mRNA splicing and trafficking (**Fig. 3C** and **D, Supplementary Table 3**). Beside of PCID2, LENG8 prominently interacted with components of the TREX/THO complex, for example, THOC1, THOC2, ALYREF, DDX39B, which is consistent with the findings in yeast that Thp3 associated with TREX/THO complex. Furthermore, LENG8 associated numerous components of spliceosome (for example, SNRNP200, SNRPA, U2AF1, SRSF1, HNRNPA1) and related to spliceosome cycle according to IPA pathway analysis. In addition, LENG8 precipitates also contained the paraspeckle components SFPQ and NONO (Benegiamo et al., 2018; Bond and Fox, 2009), or several previously identified pre-mRNA processing factors (for example, pre-mRNA capping factor NCBP2, THRAP3, RNA helicase DHX9), or DNA repair factors (**Fig.**

3E and **Supplementary Fig. 4**). Considering the functions of PCI domain as scaffold for assembly for protein complex, LENG8 most likely functions as a platform that facilitates multi-steps of the pre-mRNA processing.

We next determined a potential association between LENG8 and the THO complex in mammalian cells. FLAG-tagged LENG8 were co-expressed with HA-tagged THOs protein THOC1/3/5/6/7 in HEK293T cells, and coIP experiment was performed to detect the interaction between LENG8 and THOC components. 48 hours after transfection, the cell lysates were prepared and subjected to IP with anti-FLAG M2 beads. The western blotting results showed that THOC1, THOC5 and THOC6, but not THOC3 and THOC7, were detected in the pellet fraction of FLAG-LENG8 (**Fig. 4A**). We then re-performed the coIP assay between HA-LENG8 and FLAG-THOC1 or THOC5, and the pellet were incubated with RNase A, the western blotting results showed that RNase A treatment dramatically attenuated the interaction between LENG8 and THOC1 or THOC5 (**Fig. 4B** and **C**), suggesting that the association of LENG8 and THOCs was dependent on the existence of RNA. Subsequently, immunofluorescence microscopy analysis also revealed that either ectopic expressed or endogenous LENG8 co-localized with THOC1, ALYREF and THOC5 (**Fig. 4D-F**). Interestingly, depletion of *LENG8* in HeLa cells dramatically decreased the co-localization of PCID2 with poly (A) RNAs, but had no effect on THOC1 and poly (A) RNAs co-localization **(Supplementary Fig. 5)**, suggesting that LENG8-PICD2 complex and THOC1 may exert their functions at different steps of mRNA processing. Taken together, all of these results suggested that LENG8 was a novel THO complex-associated protein.

RNA bound by LENG8 and THOC1.

Given that LENG8 was associated with THO complex and required for mRNA processing, to identify specific types of LENG8 binding RNA and test the possibility that distinct types of RNA associate with either LENG8 or THOC1, we precipitated GFP-tagged THOC1, LENG8 or control GFP and sequenced bound RNA by deep sequencing (RNA immune-precipitation followed by deep sequencing, RIP-Seq). The

results showed that 13,328 transcripts of 5,796 genes were enriched by LENG8, while THOC1 precipitates contained 32,542 transcripts of 6,150 genes. There were 5,179 transcripts of 2,979 genes enriched by both LENG8 and THOC1 (**Fig. 5A**). The individual RNAs enriched by THOC1 and LENG8 were plotted, and color coding represents the RNA species (**Fig. 5B**). Most of the LENG8 and THOC1-binding sites were located in protein-coding transcripts (73.5 % and 75.6%, respectively)， long non-coding RNAs (15.9% and 14.2%, respectively) and unprocessed RNAs(10.2% and 9.4%, respectively). In contrast, LENG8 and THOC1 barely bound pseudogene (0.6% and 0.2%, respectively), ncRNA (0.1% and 0.1%, respectively) and TEC (To be Experimentally Confirmed) (0.1% and 0.1%, respectively) (**Fig. 5C**). Metagene profiling indicated that, compared to THOC1, LENG8 binding sties were located more in 5' UTR (27.1% *vs* 13.4%) and 3' UTR (36.9% *vs* 28.2%) region of the mRNA, but less in exon (30.9% *vs* 39.6%) and intron region (4.7% *vs* 17.7%) (**Fig. 5D** and **Supplementary Fig. 6**). Gene ontology analysis indicated that LENG8 preferentially enriched mRNA that encoding proteins primarily located in cytosolic organelles, especially mitochondria, and involved in the biological processes of mitochondria membrane assembly, mitochondria translation, and metabolism (**Fig. 5E**). In contrast, proteins encoded by THOC1 enriched mRNAs were preferentially located in chromatin and nucleosome, and involved in regulation gene silencing and chromatin function (**Fig. 5F**). HOMER motif analysis showed that LENG8 preferentially bind to the 'C(U/A)GG(A/U)G' consensus sequence contained in both mRNAs and lncRNAs (**Fig. 5G**), while THOC1 prefers "CAGCAG" consensus sequence in both mRNAs and lncRNAs (**Fig. 5H**). Taken together, these data indicated that LENG8 and THOC1 might have distinct functions and associate with different steps in mRNA processing.

We found that LENG8 and THOC1 bound with high confidence to *VDAC1* and *VDAC2* mRNA, and this finding was validated using precipitated and quantitative RT-PCR protocols (**Fig. 6A**). To determine which exported mRNAs were influenced by LENG8, we extracted both cytosolic and nuclear mRNAs from Hela cells stably

transfected with *LENG8* shRNA and identified them using deep sequencing. Plots of individually sequenced RNAs from both the nuclear and cytosolic fractions after *LENG8* knockdown, revealed that 2,608 transcripts (18.1% of total transcripts sequenced), the nuclear-cytosol export of which were reduced (using a cutoff fold change > 2) after *LENG8* depletion (**Fig. 6B**). However, when the cutoff of fold change was set at 1.5, *LENG8* knockdown inhibited the nuclear export of 8,664 transcripts (59.54% of total transcripts sequenced) (**Supplementary Fig. 7A**). Through IPA signaling pathway analysis, signaling pathways related to autophagy and mitochondria damage response were enriched after inhibition of *LENG8* (**Supplementary Fig. 7B**). Furthermore, *LENG8* deficiency resulted in 68.5% of those mRNA bound by LENG8 or 50.8% of those bound by THOC1 detained in nucleus of HeLa cells (**Fig. 6C**). In addition, the aberrant nuclear detain of *VDAC1* and *VDAC2*, two high-confidence targets of LENG8 and THOC1, were validated by quantitative RT-PCR and FISH assay (**Fig. 6D** and **E**). Thus, this data is in strong agreement with the FISH results that LENG8 is required for mRNA export (**Fig. 6F**).

LENG8 controls mitochondrial activity

Given that LENG8 preferentially bound mRNAs encoding mitochondrial proteins, we therefore hypothesized that LENG8 was required for maintenance of mitochondrial activity. To test this, we generated *LENG8* knockdown and knockout HeLa cells by shRNAs and CRISPR/Cas9 (**Supplementary Fig. 8**) and then examined the morphology and respiration activity of mitochondria. The mitochondrial activity was determined using three types of mitochondria-specific labels that distinguish respiration (Mitotracker Deep Red), total (Mitotracker Green) and ROS-generating mitochondria (MitoSOX) (Zhou et al., 2011). Either knockdown or knockout of *LENG8* in HeLa resulted in reduction of ROS production and loss of mitochondria membrane potential (**Fig. 7A** and **B**). Likewise, silencing of THOC1, THOC2 & ALYREF/THOC4 in HeLa cells lead to a reduction of Mitotracker Deep Red staining, indicating an essential role of THO complex in regulation of mitochondria respiratory activity.

Imaging of Mitotracker green staining using confocal microscopy found that after application of RNAi directed towards *LENG8*, the cells exhibited enhanced disruption of the mitochondrial network and increased mitochondrial depolarization. (**Supplementary Fig. 9**). Furthermore, electron microscopy of *LENG8* deficient HeLa cells revealed complete absence of "normal" mitochondria in *LENG8* deficient cells, including swollen morphology and dramatically decreased numbers of cristae (**Fig. 7C**). Consistent to the morphological change, *LENG8* deficient HeLa cells exhibited a dramatically reduced oxygen consumption rate (OCR), both in the presence and absence of oligomycin (**Fig. 7D** and **E**).

Mitochondrial activity is tightly associated with thermogenesis in adipose tissue. To explore the physiological function of LENG8, we generated the adipose tissue specific *Leng8* knockout mice by crossing mice carrying gene-targeted floxed *Leng8* alleles ($Leng8^{fl/fl}$) with Cre recombinase transgenic mice driven by Adipoq gene promoter ($Adipoq^{Cre}$). When fed with high fat diet (HFD), $Leng8^{fl/fl}Adipoq^{Cre}$ mice exhibited a decreased body weight compared to $Leng8^{fl/fl}$ mice (**Fig. 7F**). Furthermore, $Leng8^{fl/fl}Adipoq^{Cre}$ mice displayed marked reduction in adipocyte area and perimeter in sections of brown adipose tissue (BAT) and inguinal white adipose tissue (iWAT) (**Fig. 7G-J**). Consistently, $Leng8^{fl/fl}Adipoq^{Cre}$ mice displayed marked increase in UCP1 immuno-labeling in BAT (**Fig. 7K-N**). Taken together, our finding suggests an essential role of LENG8 in maintenance of mitochondrial integrity and activity.

Discussion

Using genetic or biochemistry approach, a few regulatory factors in mRNA processing have been found in yeast, and some of them have their mammalian counterpart identified and characterized (Cheng et al., 2006; Dominguez-Sanchez et al., 2011; Katahira and Yoneda, 2009; Yuan et al., 2018). However, it is still little known how the mRNA processing is regulated in mammalian cells. Thp3-Csn12 protein complex has been found to regulate transcriptional elongation and mRNA maturation in yeast, but its specific mechanism and whether there are homologous proteins in mammalian remains unknown (Jimeno et al., 2011; Kragelund et al., 2016).

Our finding identified LENG8-PCID2 complex, as the mammalian homologous of Thp3-Csn12. LENG8 is a novel regulator of mRNA processing in mammals. It binds mRNA, associates with mRNA processing machinery, and regulates mRNA export. We have also found that LENG8 preferentially bound mRNA encoding mitochondria-localized proteins, and its activity is required for the maintenance of mitochondria morphology and function (**Supplementary Fig. 11**). Importantly, this work has revealed novel aspects of an evolutionarily conserved mRNA processing complex with the ability to control mitochondrial activity.

Both in yeast and mammals, TREX/THO complex provides a connection between transcription, RNA processing and genome integrity (Dominguez-Sanchez et al., 2011; Gaillard et al., 2007). Although the functional role of LENG8 remains uncharacterized, several interactome studies conducted previously have already revealed the existence of LENG8 in RNA polymerase II complex (Baillat et al., 2005), RNA processing machinery (Cano et al., 2015; Gebhardt et al., 2015; Viita et al., 2019) and DNA repair complex (Alsulami et al., 2019; Hu et al., 2019), suggesting a possible function of LENG8 as a master coordinator of these processes. Clearly, to determine the potential ability of LENG8 to synchronize mRNA processing with genome integrity warrants further investigation.

Surprisingly, human mitochondrial genome includes only 13 coding genes while nuclear-encoded genes account for 99% of mitochondrial proteins (Hendrickson et al., 2010). It is thus widely recognized that expression of nuclear genes controls all of the aspects of mitochondria activity, including morphology, redox regulation, and energetics (Karakaidos and Rampias, 2020). Our finding suggested an essential role of LENG8 regulated mRNA processing in maintenance of mitochondria activity. To our knowledge, this is the very first time that the maturation process of nuclear gene mRNA links to the regulation of mitochondrial activity. Our results showed that LENG8 associates a large number of mitochondria proteins, including PHB1, a key regulator in mitochondrial homeostasis, which responds to mitochondria stress, translocate from mitochondria to nuclear and regulates the expression of nuclear

genes essential for mitochondrial biogenesis, regeneration and degradation (Hernando-Rodriguez and Artal-Sanz, 2018). This result is consistent to the previous finding that LENG8 was also highly enriched in the PHB1 interactome (Xu et al., 2016). Taken together, there might be a close association between LENG8 and PHB1. Whether PHB1 may exert its function in the mito-stress response through its association with LENG8 remains to be explored.

N6-methyl-adenosine (m^6A) is a newly characterized RNA fate determiner and affects multiple aspects of mRNA processing including splicing, translation and decay through various m^6A recognition proteins, the so-called m^6A readers. Several studies have shown that the m^6A reader YTHDC1 recruits TREX complex to mRNA and its deficiency lead to aberrant nuclear accumulation of mRNA, suggesting an essential role of m^6A modification for mRNA nuclear export. Here we showed that THOC1 itself, which is the most well characterized TREX component, is a potential m^6A reader. Furthermore, YTHDC1, as well as another m^6A reader YTHDF2, were found to co-precipitate with LENG8 in this study. These finding strengthens the link between m^6A modification and mRNA export. However, our RIP-seq results do not support LENG8 as a significant m^6A reader. Thus, whether LENG8 is involved in m6A mediated mRNA export, as well as other processes of RNA biology remains to be explored.

We conclude that LENG8 is a fundamental factor involved in the support of mitochondrial activity. Furthermore, this has been seen previously in *C. elegans*, where Hpo-10 was shown to be the nematode homologue of LENG8/Thp3 and may be a candidate for the regulation of the mitochondrial unfolded protein response (Liu et al., 2014). Furthermore, *Thoc-1* was also found to regulate mitoUPR in their reverse-genetic screening. Collectively, these findings suggested an essential role of TREX/THO-mediated mRNA maturation in regulation of mitochondria homeostasis in different species.

Conclusion

LENG8 is evolutionarily conserved mRNA processing factor that regulate mRNP biogenesis, and its regulation of mRNA export is required for the control of mitochondrial activity.

Ethics Approval And Consent To Participate

Not applicable

Consent For Publication

Not applicable

Availability of data and material

Data have been deposited in the Gene Expression Omnibus under accession code GSE171126. Other data that support the findings of this study are available from the corresponding author upon request.

Competing interests

Not applicable

Funding

This study was supported by National Science Foundation (81971240 to Liu F and 82070824 to Zhao YX); China Postdoctoral Science Foundation (No. 2020M671248 and No.2020T130672 to Zhao YX), Shanghai Yangfan Talents Program (20YF1456200 to Zhao YX), Shanghai Municipal Commission of Science and Technology (No.18DZ2260200), Jinan Science and Technology Development Program (No. 201907018 to Zhang N) and Shandong Provincial Key Research and Development Program (No. 2017G006037 and ZR2020MH201 to Zhang N)

Authors' contributions

Y.X.Z. X.T.W. H.L.Y. and F.L. designed experiments. Y.X.Z. X.T.W., Y.N.L., N.N.L.,

S.M.W., N.Z., Z.F.G, W.Y.X., L.F.M, C.Y.L., J.G. and Z.L.C. performed experiments and analyzed data. Y.X.Z., Q.R.D. and F.L. prepared the figures and wrote the manuscript.

Acknowledgements

Not applicable

REFERENCES

Alsulami, M., Munawar, N., Dillon, E., Oliviero, G., Wynne, K., Alsolami, M., Moss, C., P, O.G., O'Meara, F., Cotter, D.*, et al.* (2019). SETD1A Methyltransferase Is Physically and Functionally Linked to the DNA Damage Repair Protein RAD18. Mol Cell Proteomics *18*, 1428-1436.

Baillat, D., Hakimi, M.A., Naar, A.M., Shilatifard, A., Cooch, N., and Shiekhattar, R. (2005). Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell *123*, 265-276.

Benegiamo, G., Mure, L.S., Erikson, G., Le, H.D., Moriggi, E., Brown, S.A., and Panda, S. (2018). The RNA-Binding Protein NONO Coordinates Hepatic Adaptation to Feeding. Cell Metab *27*, 404-418 e407.

Bhatia, V., Barroso, S.I., Garcia-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature *511*, 362-365.

588 Bond, C.S., and Fox, A.H. (2009). Paraspeckles: nuclear bodies built on long
589 noncoding RNA. J Cell Biol *186*, 637-644.

590 Cano, F., Rapiteanu, R., Sebastiaan Winkler, G., and Lehner, P.J. (2015). A
591 non-proteolytic role for ubiquitin in deadenylation of MHC-I mRNA by the
592 RNA-binding E3-ligase MEX-3C. Nat Commun *6*, 8670.

593 Cheng, H., Dufu, K., Lee, C.S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human
594 mRNA export machinery recruited to the 5' end of mRNA. Cell *127*, 1389-1400.

595 Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., Chang, X., and Cheng, H.
596 (2013). Aly and THO are required for assembly of the human TREX complex and
597 association of TREX components with the spliced mRNA. Nucleic Acids Res *41*,
598 1294-1306.

599 Dominguez-Sanchez, M.S., Barroso, S., Gomez-Gonzalez, B., Luna, R., and Aguilera,
600 A. (2011). Genome instability and transcription elongation impairment in human cells
601 depleted of THO/TREX. PLoS Genet *7*, e1002386.

602 Fan, J., Wang, K., Du, X., Wang, J., Chen, S., Wang, Y., Shi, M., Zhang, L., Wu, X.,
603 Zheng, D.*, et al.* (2019). ALYREF links 3'-end processing to nuclear export of
604 non-polyadenylated mRNAs. EMBO J *38*.

605 Gaillard, H., Wellinger, R.E., and Aguilera, A. (2007). A new connection of mRNP
606 biogenesis and export with transcription-coupled repair. Nucleic Acids Res *35*,
607 3893-3906.

608 Garcia-Oliver, E., Garcia-Molinero, V., and Rodriguez-Navarro, S. (2012). mRNA
609 export and gene expression: the SAGA-TREX-2 connection. Biochim Biophys Acta
610 *1819*, 555-565.

611 Gebhardt, A., Habjan, M., Benda, C., Meiler, A., Haas, D.A., Hein, M.Y., Mann, A.,
612 Mann, M., Habermann, B., and Pichlmair, A. (2015). mRNA export through an
613 additional cap-binding complex consisting of NCBP1 and NCBP3. Nat Commun *6*,
614 8192.

615 Hendrickson, S.L., Lautenberger, J.A., Chinn, L.W., Malasky, M., Sezgin, E.,
616 Kingsley, L.A., Goedert, J.J., Kirk, G.D., Gomperts, E.D., Buchbinder, S.P., *et al.*
617 (2010). Genetic variants in nuclear-encoded mitochondrial genes influence AIDS
618 progression. PLoS One *5*, e12862.

619 Hernando-Rodriguez, B., and Artal-Sanz, M. (2018). Mitochondrial Quality Control
620 Mechanisms and the PHB (Prohibitin) Complex. Cells *7*.

621 Hu, K., Wu, W., Li, Y., Lin, L., Chen, D., Yan, H., Xiao, X., Chen, H., Chen, Z.,
622 Zhang, Y., *et al.* (2019). Poly(ADP-ribosyl)ation of BRD7 by PARP1 confers
623 resistance to DNA-damaging chemotherapeutic agents. EMBO Rep *20*.

624 Jimeno, S., Rondon, A.G., Luna, R., and Aguilera, A. (2002). The yeast THO complex
625 and mRNA export factors link RNA metabolism with transcription and genome
626 instability. EMBO J *21*, 3526-3535.

627 Jimeno, S., Tous, C., Garcia-Rubio, M.L., Ranes, M., Gonzalez-Aguilera, C., Marin,
628 A., and Aguilera, A. (2011). New suppressors of THO mutations identify Thp3
629 (Ypr045c)-Csn12 as a protein complex involved in transcription elongation. Mol Cell
630 Biol *31*, 674-685.

631 Karakaidos, P., and Rampias, T. (2020). Mitonuclear Interactions in the Maintenance
632 of Mitochondrial Integrity. Life (Basel) *10*.

633 Katahira, J., and Yoneda, Y. (2009). Roles of the TREX complex in nuclear export of
634 mRNA. RNA Biol *6*, 149-152.

635 Khoshnevis, S., Gunisova, S., Vlckova, V., Kouba, T., Neumann, P., Beznoskova, P.,
636 Ficner, R., and Valasek, L.S. (2014). Structural integrity of the PCI domain of
637 eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes.
638 Nucleic Acids Res *42*, 4123-4139.

639 Kragelund, B.B., Schenstrom, S.M., Rebula, C.A., Panse, V.G., and
640 Hartmann-Petersen, R. (2016). DSS1/Sem1, a Multifunctional and Intrinsically
641 Disordered Protein. Trends Biochem Sci *41*, 446-459.

642 Kumar, R., Corbett, M.A., van Bon, B.W., Woenig, J.A., Weir, L., Douglas, E., Friend,
643 K.L., Gardner, A., Shaw, M., Jolly, L.A.*, et al.* (2015). THOC2 Mutations Implicate
644 mRNA-Export Pathway in X-Linked Intellectual Disability. Am J Hum Genet *97*,
645 302-310.

646 Lee, K.M., and Tarn, W.Y. (2013). Coupling pre-mRNA processing to transcription on
647 the RNA factory assembly line. RNA Biol *10*, 380-390.

648 Li, Y., Wang, X., Zhang, X., and Goodrich, D.W. (2005). Human hHpr1/p84/Thoc1
649 regulates transcriptional elongation and physically links RNA polymerase II and RNA
650 processing factors. Mol Cell Biol *25*, 4023-4033.

651 Liu, Y., Samuel, B.S., Breen, P.C., and Ruvkun, G. (2014). Caenorhabditis elegans
652 pathways that surveil and defend mitochondria. Nature *508*, 406-410.

653 Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005).
654 Recruitment of the human TREX complex to mRNA during splicing. Genes Dev *19*,
655 1512-1517.

656 Pierce, A., Carney, L., Hamza, H.G., Griffiths, J.R., Zhang, L., Whetton, B.A.,
657 Gonzalez Sanchez, M.B., Tamura, T., Sternberg, D., and Whetton, A.D. (2008).
658 THOC5 spliceosome protein: a target for leukaemogenic tyrosine kinases that affects
659 inositol lipid turnover. Br J Haematol *141*, 641-650.

660 Rondon, A.G., Jimeno, S., and Aguilera, A. (2010). The interface between
661 transcription and mRNP export: from THO to THSC/TREX-2. Biochim Biophys Acta
662 *1799*, 533-538.

663 Saran, S., Tran, D.D., Ewald, F., Koch, A., Hoffmann, A., Koch, M., Nashan, B., and
664 Tamura, T. (2016). Depletion of three combined THOC5 mRNA export protein target
665 genes synergistically induces human hepatocellular carcinoma cell death. Oncogene
666 *35*, 3872-3879.

667 Shi, M., Zhang, H., Wu, X., He, Z., Wang, L., Yin, S., Tian, B., Li, G., and Cheng, H.
668 (2017). ALYREF mainly binds to the 5' and the 3' regions of the mRNA in vivo.

669 Nucleic Acids Res *45*, 9640-9653.

670 Stewart, M. (2019). Structure and Function of the TREX-2 Complex. Subcell
671 Biochem *93*, 461-470.

672 Strasser, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S.,
673 Rondon, A.G., Aguilera, A., Struhl, K., Reed, R.*, et al.* (2002). TREX is a conserved
674 complex coupling transcription with messenger RNA export. Nature *417*, 304-308.

675 Sun, Q., Fan, W., Chen, K., Ding, X., Chen, S., and Zhong, Q. (2008). Identification
676 of Barkor as a mammalian autophagy-specific factor for Beclin 1 and class III
677 phosphatidylinositol 3-kinase. Proc Natl Acad Sci U S A *105*, 19211-19216.

678 Tran, D.D., Saran, S., Williamson, A.J., Pierce, A., Dittrich-Breiholz, O., Wiehlmann,
679 L., Koch, A., Whetton, A.D., and Tamura, T. (2014). THOC5 controls
680 3'end-processing of immediate early genes via interaction with polyadenylation
681 specific factor 100 (CPSF100). Nucleic Acids Res *42*, 12249-12260.

682 Viita, T., Kyheroinen, S., Prajapati, B., Virtanen, J., Frilander, M.J., Varjosalo, M., and
683 Vartiainen, M.K. (2019). Nuclear actin interactome analysis links actin to KAT14
684 histone acetyl transferase and mRNA splicing. J Cell Sci *132*.

685 Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and
686 Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes
687 by CRISPR/Cas-mediated genome engineering. Cell *153*, 910-918.

688 Wickramasinghe, V.O., and Laskey, R.A. (2015). Control of mammalian gene
689 expression by selective mRNA export. Nat Rev Mol Cell Biol *16*, 431-442.

690 Xie, Y., and Ren, Y. (2019). Mechanisms of nuclear mRNA export: A structural
691 perspective. Traffic *20*, 829-840.

692 Xu, Y., Yang, W., Shi, J., and Zetter, B.R. (2016). Prohibitin 1 regulates tumor cell
693 apoptosis via the interaction with X-linked inhibitor of apoptosis protein. J Mol Cell
694 Biol *8*, 282-285.

Yuan, X., Zhang, T., Yao, F., Liao, Y., Liu, F., Ren, Z., Han, L., Diao, L., Li, Y., Zhou, B., *et al.* (2018). THO Complex-Dependent Posttranscriptional Control Contributes to Vascular Smooth Muscle Cell Fate Decision. Circ Res *123*, 538-549.

Zenklusen, D., Vinciguerra, P., Strahm, Y., and Stutz, F. (2001). The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. Mol Cell Biol *21*, 4219-4232.

Zhou, R., Yazdi, A.S., Menu, P., and Tschopp, J. (2011). A role for mitochondria in NLRP3 inflammasome activation. Nature *469*, 221-225.

ACKNOWLEDGEMENT

This study was supported by National Science Foundation (No. 81971240 to Liu F); China Postdoctoral Science Foundation (No. 2020M671248 and No.2020T130672 to Zhao YX), Shanghai Yangfan Talents Program (20YF1456200 to Zhao YX), Shanghai Municipal Commission of Science and Technology(No.18DZ2260200), Jinan Science and Technology Development Program (No. 201907018 to Zhang N) and Shandong Provincial Key Research and Development Program (No. 2017G006037 to Zhang N)

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

FIGURE LEGEND

Figure 1 Conservation of mRNP biogenesis machinery from yeast to human. (**A** and **B**) mRNP biogenesis machinery in yeast and human. (**C**) Alignment of human LENG8 with its orthologues in other species. (**D**) Alignment of human PCID2 with its orthologues in other species. (**E**) Alignment of the amino acid sequences of the PCI domain. Color coding represents the identical or similar amino acids. (**F**) Ribbon representation of the structure of human LENG8, yeast THP3 and SAC3.

Figure 2 Requirement of LENG8-PCID2 complex in mRNA export. (**A**) HEK293T cells were transfected with FLAG-LENG8 and MYC-PCID2 and then the lysates were sent to immunoprecipitation using anti-FLAG, and western blot using anti-FLAG and anti-MYC. (**B**) HeLa cell lysates were sent to streptavidin-affinity purification using biotin-labeled or non-biotin oligo-dT, and western blot using anti-PCID2, anti-THOC1, anti-ALYREF or anti-LENG8. (**C** and **D**) HeLa cells were

sent to in situ hybridization using Cy3-labelled oligo-dT and immuno-fluorescence staining using anti-LENG8 (**C**) and anti-PCID2 (**D**). (**E**) HeLa cells were transfected with FLAG-LENG8, and sent to in situ hybridization using Cy3-labelled oligo-dT and immuno-fluorescence staining using mouse anti-FLAG and rabbit anti-PCID. (**F**) HeLa cells stably expressing shRNA targeting *LENG8* or *THOC1* were sent to in situ hybridization using Cy3-labelled oligo-dT. (**G**) Mouse tail fibroblasts were isolated from wildtype or *Leng8*$^{+/-}$ mice, and then sent to in situ hybridization using Cy3-labelled oligo-dT. (**H**) HeLa cells were transfected with a pool of small interferencing RNA (siRNAs) targeting *PCID2* were sent to in situ hybridization using Cy3-labelled oligo-dT. Unprocessed scans of western blot analysis are available in **Supplementary Figure 10**. Bar = 2 μm in **C - E** and 5 μm in **F - H.** Data are representative of at least three independent experiments.

Figure 3 LENG8 associates with mRNA processing factors. (**A**) Silver staining of LENG8 associated proteins by tandem affinity purification. (**B**) Subcellular localization of LENG8 associated proteins. (**C - E**) Molecular function category (**C**), enriched signal pathways (**D**) and protein interaction network (**E**) of LENG8 associated proteins analyzed by Ingenuity Pathway Analysis. Source data of **A** and **C** are in **Supplementary Table 2** and **3**.

Figure 4 LENG8 associates with TREX components. (**A**) HEK293T cells were transfected with FLAG-LENG8 and indicated plasmids of HA tagged THO components and then the lysates were sent to immunoprecipitation using anti-FLAG, and western blot analysis using anti-FLAG and anti-HA. (**B** and **C**) HEK293T cells were transfected with FLAG-LENG8 and HA tagged THOC1 (**B**) or THOC5 (**C**) and then the lysates were sent to immunoprecipitation using anti-FLAG. The precipitates were treated with RNase A and then sent to western blot analysis using anti-FLAG and anti-HA. (**D - F**) HeLa cells transfected with FLAG-LENG8 were sent to

immuno-fluorescence staining using anti-THOC1 (**D**), ALYREF (**E**) or THOC5 (**F**) and in situ hybridization using Cy3-labelled oligo-dT. Unprocessed scans of western blot analysis are available in **Supplementary Figure 10**. Bar = 2 μm in **D - F**. Data are representative of at least three independent experiments.

Figure 5 Global analysis of RNAs bound by LENG8 or THOC1. (**A**) Venn diagram showing the numbers of shared high-confidence genes or transcripts enriched by LENG8 and THOC1. (**B**) Scatter plot showing enrichment of transcripts binding to LENG8 (x axis) and THOC1 (y axis) quantified on the gene level. RNAs are color-coded according to their annotated RNA types in Ensembl. (**C**) Percentage of enriched RNA types binding to LENG8 or THOC1. (**D**)The distribution of LENG8 or THOC1 binding peaks within different gene regions. (**E** and **F**) Gene ontology analysis of LENG8 (**E**) or THOC1 (**F**) enriched RNAs. (**G** and **H**)Top consensus sequences of LENG8 (**G**) or THOC1 (**H**) binding sites detected by HOMER Motif analysis. Source data of **B** are in **Supplementary Table 4**.

Figure 6 LENG8 is required for *VDAC1* and *VDAC2* mRNA export. (**A**) Validation of RIP-sequencing data by RT-qPCR. RNA in LENG8, THOC1 or Control precipitates was amplified by RT-qPCR using specific primers for two mRNAs (*VDAC1* and *VDAC2*). (**B**) HeLa cells stably expressing shRNA targeting *LENG8* were sent to cytosolic-nuclear fractioning, and total RNA from each fraction was extracted and sent to RNA sequencing. Scatter plot indicates individual RNAs sequenced. X axis shows log2FC (fold change) of the ratio of cytosolic/nuclear RNAs after *LENG8* knockdown. (**C**) Analysis of the cytosolic/nuclear ratio of mRNA enriched by LENG8, THOC1 or both. (**D** and **E**) RT-qPCR analysis of cytosolic or nuclear *VDAC1* (**D**) and *VDAC2* (**E**) mRNA after *LENG8* knockdown. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by the unpaired t-test (**A, D** and **E**). (**F**) FISH analysis of *VDAC1* mRNA after *LENG8* or *PCID2* knockdown. Data are from three independent

787 experiments (means ± s.e.m.).

788

789 **Figure 7 LENG8 controls mitochondrial activity.** (**A** and **B**) HeLa cells of *LENG8*
790 knockdown or knockout were stained with MitoSOX (**A**) or Mitotracker Green and
791 Mitotracker Deep Red (**B**) for 30 min and analyzed by flow cytometry. (**C** - **E**) HeLa
792 cells of *LENG8* knockdown or knockout were sent to electron microscopy imaging (**C**)
793 or OCR assay (**D** and **E**). O for oligomycin, F for FCCP, A for antimycin and R for
794 retenone. (**F**) Relative body weight of $Leng8^{fl/fl}$ $Adipoq^{Cre}$ and $Leng8^{fl/fl}$ mice after fed
795 with high fat diet. (**G** - **J**) HE staining of brown adipose tissues (**G** and **H**) and
796 inguinal white adipose tissue of $Leng8^{fl/fl}$ $Adipoq^{Cre}$ and $Leng8^{fl/fl}$ mice after fed with
797 high fat diet (**I** and **J**). (**K** - **N**) BAT from $Leng8^{fl/fl}$ $Adipoq^{Cre}$ and $Leng8^{fl/fl}$ mice after
798 fed with HFD were sent to immune-staining against UCP1. Data are representative of
799 at least three independent experiments.

800

Figures

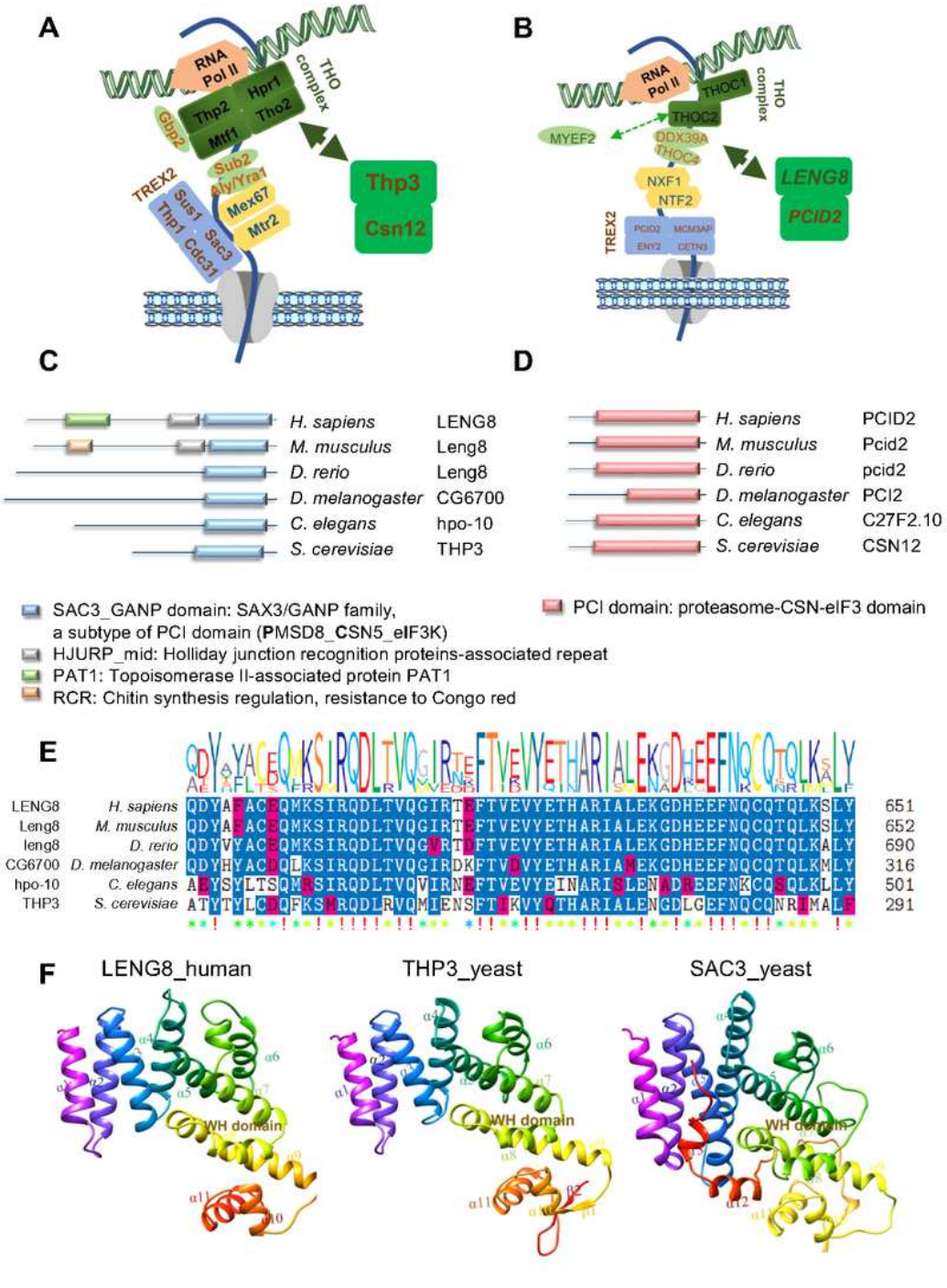

Figure 1

Conservation of mRNP biogenesis machinery from yeast to human. (A and B) mRNP biogenesis machinery in yeast and human. (C) Alignment of human LENG8 with its orthologues in other species. (D) Alignment of human PCID2 with its orthologues in other species. (E) Alignment of the amino acid

sequences of the PCI domain. Color coding represents the identical or similar amino acids. (F) Ribbon representation of the structure of human LENG8, yeast THP3 and SAC3.

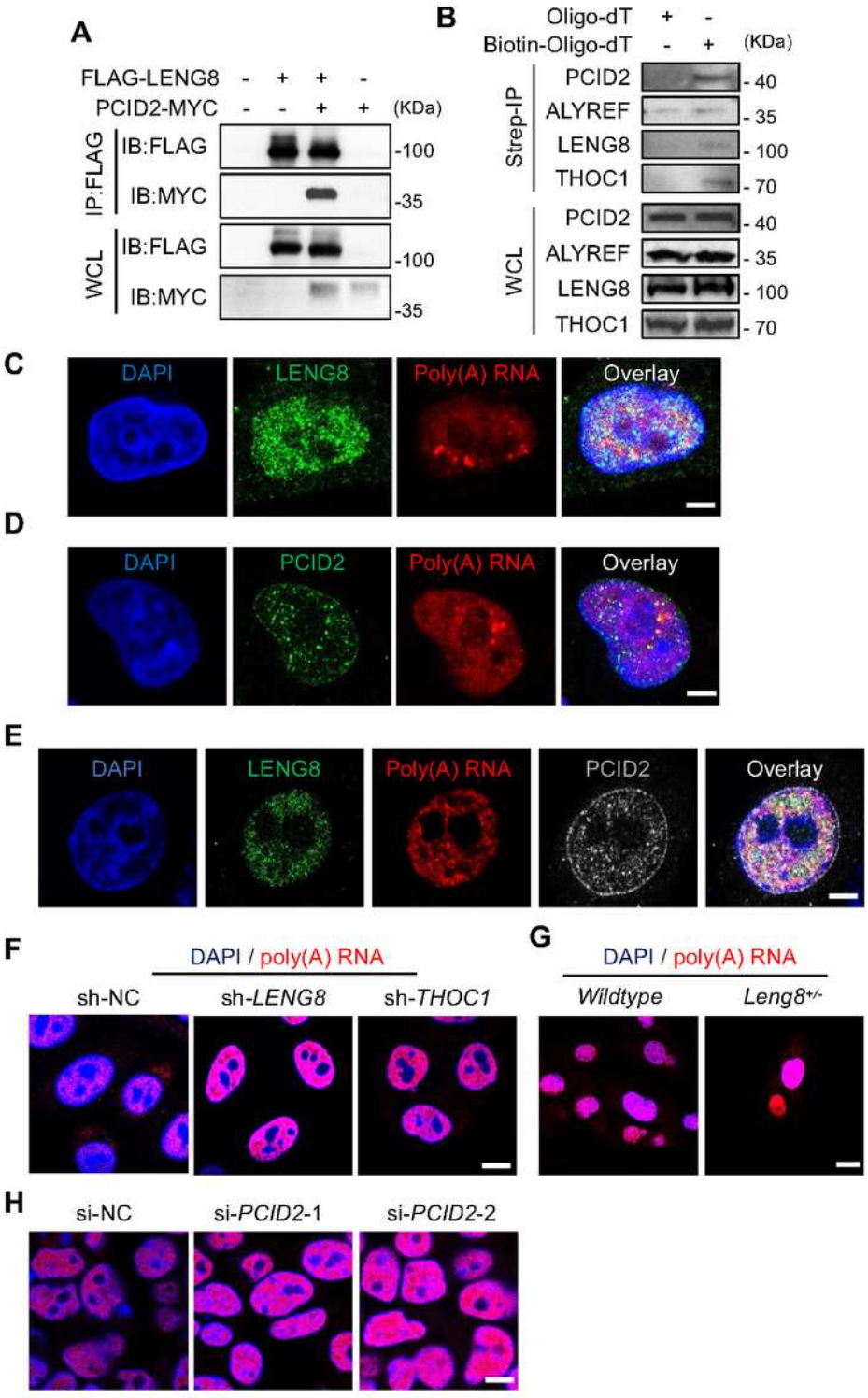

Figure 2

Requirement of LENG8-PCID2 complex in mRNA export. (A) HEK293T cells were transfected with FLAG-LENG8 and MYC-PCID2 and then the lysates were sent to immunoprecipitation using anti-FLAG, and western blot using anti-FLAG and anti-MYC. (B) HeLa cell lysates were sent to streptavidin-affinity

purification using biotin-labeled or non-biotin oligo-dT, and western blot using anti-PCID2, anti-THOC1, anti-ALYREF or anti-LENG8. (C and D) HeLa cells were sent to in situ hybridization using Cy3-labelled oligo-dT and 731 immuno-fluorescence staining using anti-LENG8 (C) and anti-PCID2 (D). (E) HeLa cells were transfected with FLAG-LENG8, and sent to in situ hybridization using Cy3-labelled oligo-dT and immuno-fluorescence staining using mouse anti-FLAG and rabbit anti-PCID. (F) HeLa cells stably expressing shRNA targeting LENG8 or THOC1 were sent to in situ hybridization using Cy3-labelled oligo-dT. (G) Mouse tail fibroblasts were isolated from wildtype or Leng8+/- mice, and then sent to in situ hybridization using Cy3-labelled oligo-dT. (H) HeLa cells were transfected with a pool of small interferencing RNA (siRNAs) targeting PCID2 were sent to in situ hybridization using Cy3-labelled oligo-dT. Unprocessed scans of western blot analysis are available in Supplementary Figure 10. Bar = 2 μm in C - E and 5 μm in F - H. Data are representative of at least three independent experiments.

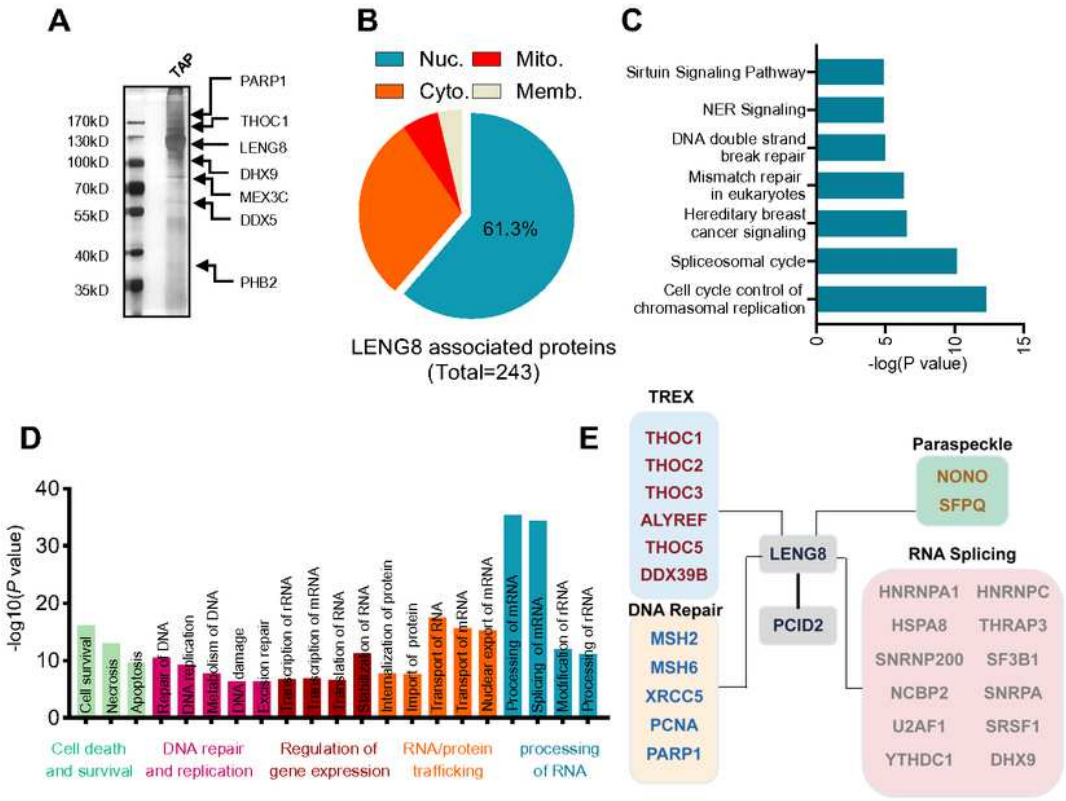

Figure 3

LENG8 associates with mRNA processing factors. (A) Silver staining of LENG8 associated proteins by tandem affinity purification. (B) Subcellular localization of LENG8 associated proteins. (C - E) Molecular function category (C), enriched signal pathways (D) and protein interaction network (E) of LENG8 associated proteins analyzed by Ingenuity Pathway Analysis. Source data of A and C are in Supplementary Table 2 and 3.

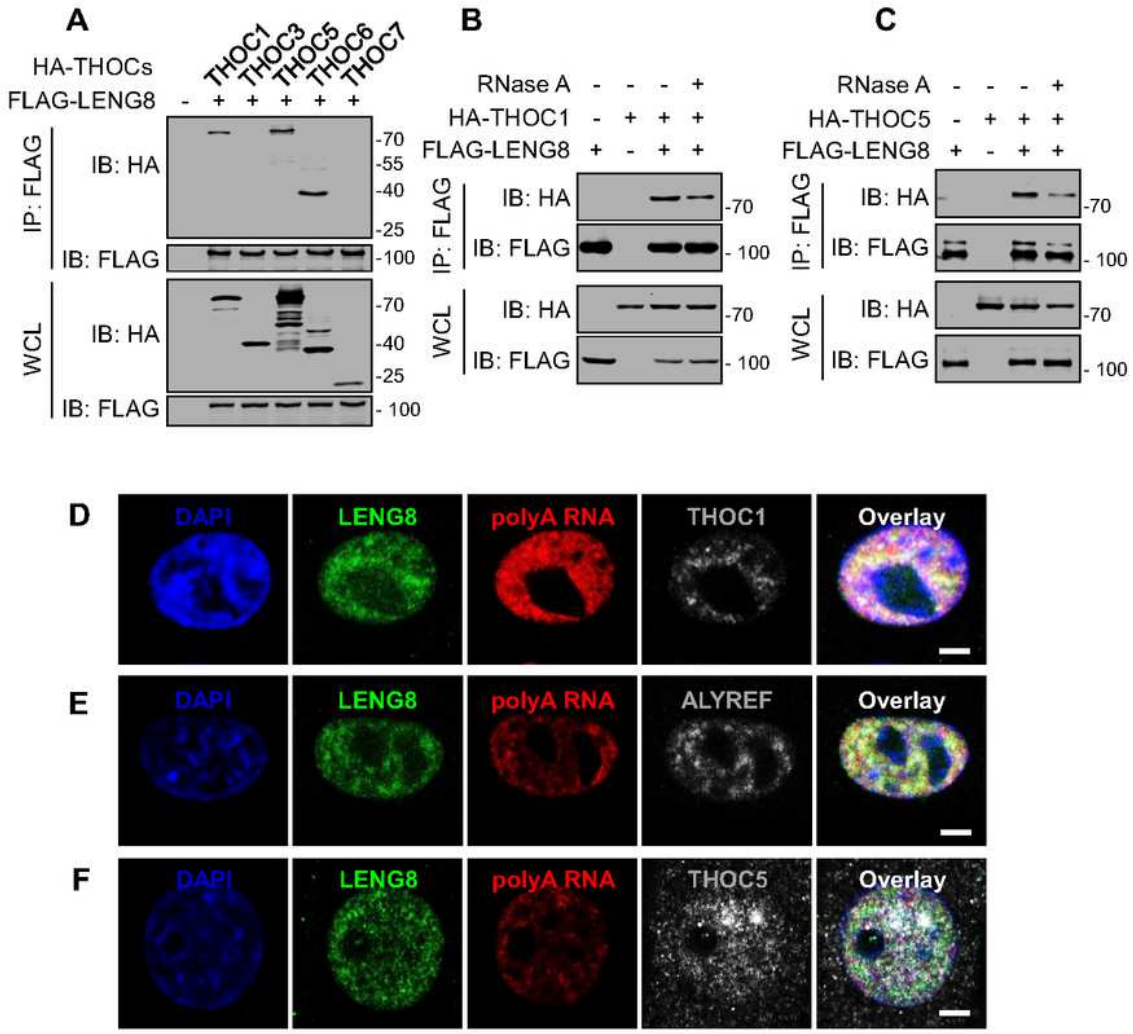

Figure 4

LENG8 associates with TREX components. (A) HEK293T cells were transfected with FLAG-LENG8 and indicated plasmids of HA tagged THO components and then the lysates were sent to immunoprecipitation using anti-FLAG, and western blot analysis using anti-FLAG and anti-HA. (B and C) HEK293T cells were transfected with FLAG-LENG8 and HA tagged THOC1 (B) or THOC5 (C) and then the lysates were sent to immunoprecipitation using anti-FLAG. The precipitates were treated with RNase A

and then sent to western blot analysis using anti-FLAG and anti-HA. (D - F) HeLa cells transfected with FLAG-LENG8 were sent to immuno-fluorescence staining using anti-THOC1 (D), ALYREF (E) or THOC5 (F) and in situ hybridization using Cy3-labelled oligo-dT. Unprocessed scans of western blot analysis are available in Supplementary Figure 10. Bar = 2 µm in D - F. Data are representative of at least three independent experiments.

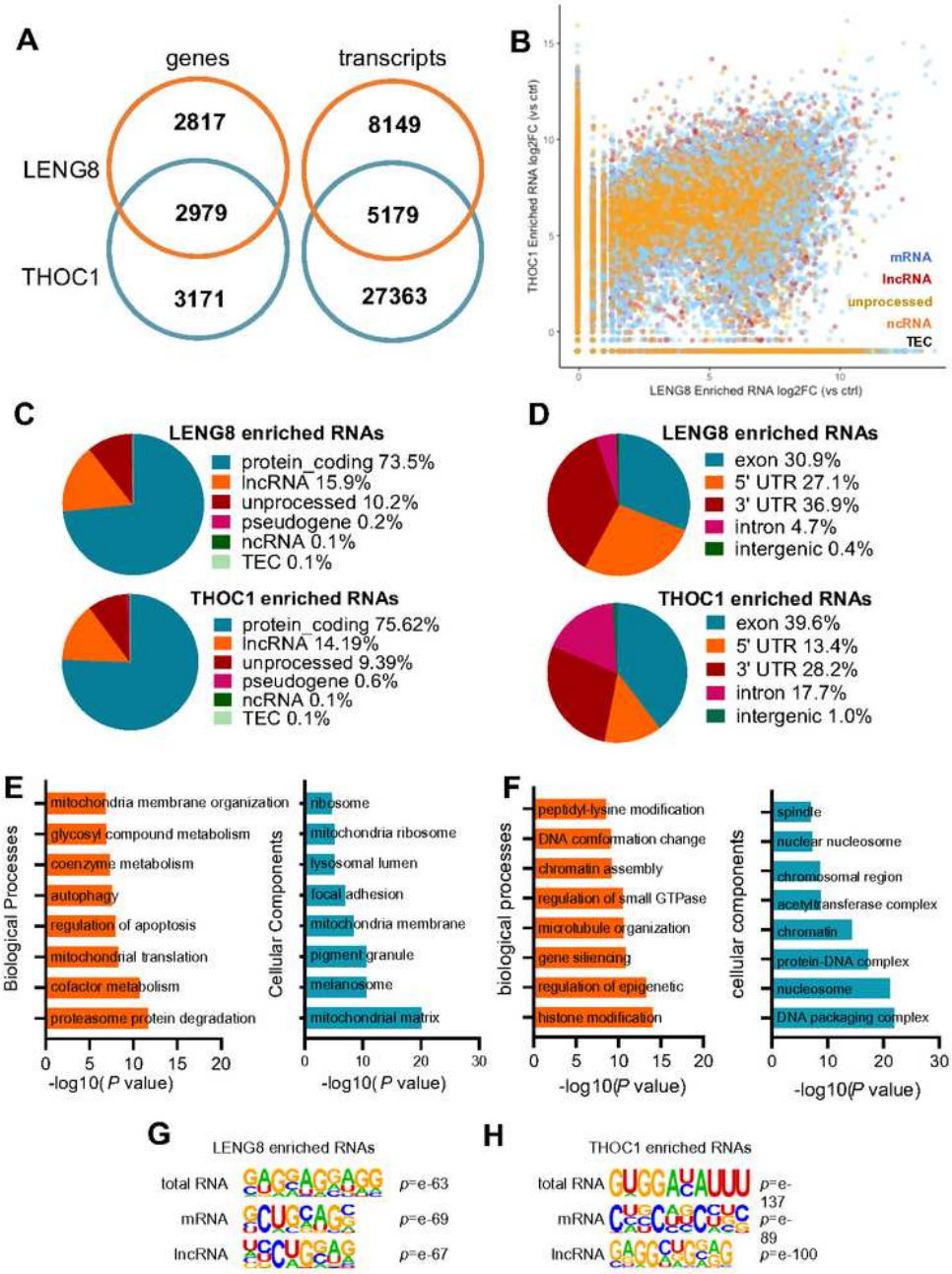

Figure 5

Global analysis of RNAs bound by LENG8 or THOC1. (A) Venn diagram showing the numbers of shared high-confidence genes or transcripts enriched by LENG8 and THOC1. (B) Scatter plot showing enrichment of transcripts binding to LENG8 (x axis) and THOC1 (y axis) quantified on the gene level. RNAs are color-coded according to their annotated RNA types in Ensembl. (C) Percentage of enriched RNA types binding to LENG8 or THOC1. (D)The distribution of LENG8 or THOC1 binding peaks within different gene regions. (E and F) Gene ontology analysis of LENG8 (E) or THOC1 (F) enriched RNAs. (G and H)Top consensus sequences of LENG8 (G) or THOC1 (H) binding sites detected by HOMER Motif analysis. Source data of B are in Supplementary Table 4.

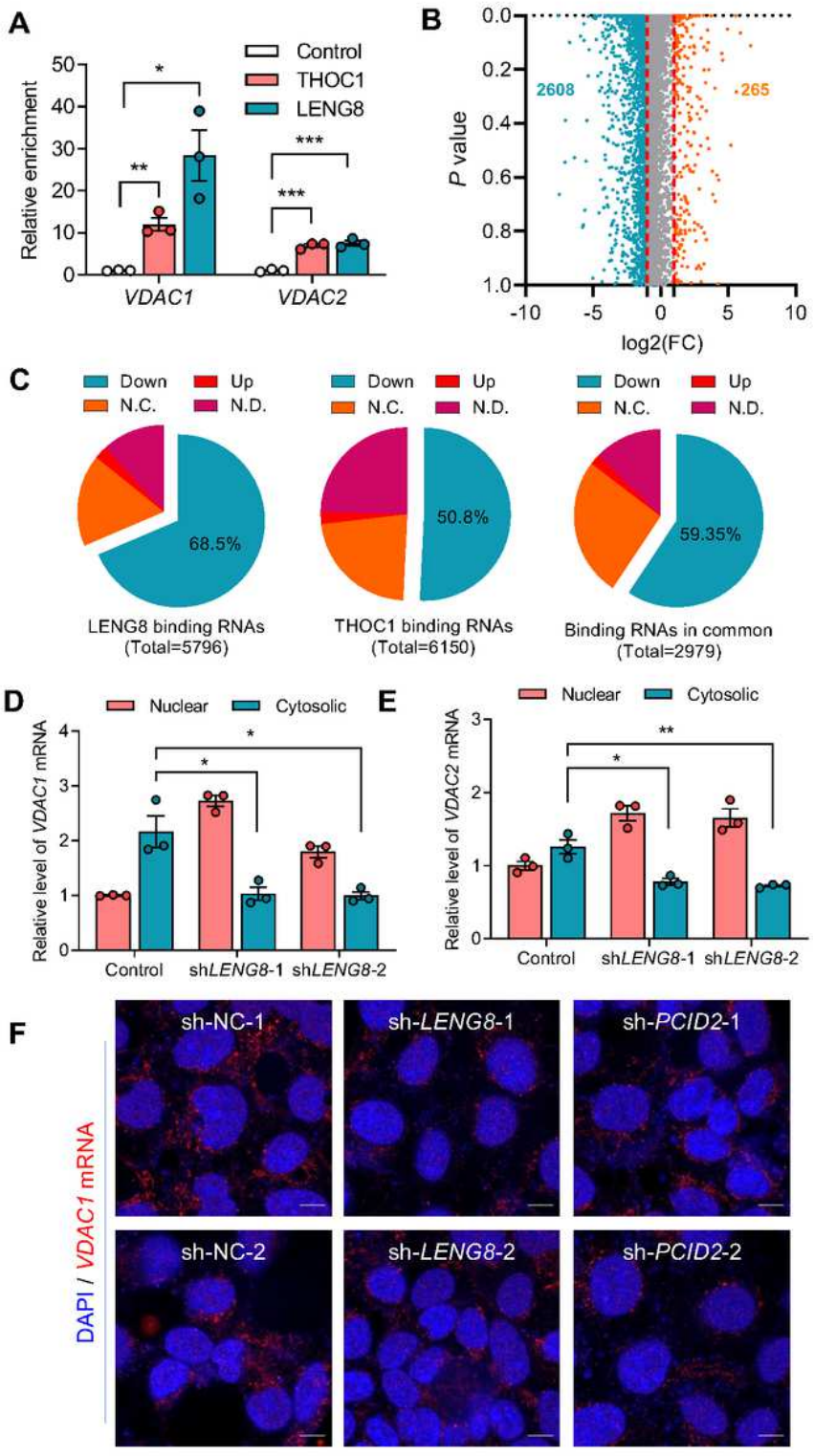

Figure 6

LENG8 is required for VDAC1 and VDAC2 mRNA export. (A) Validation of RIP-sequencing data by RT-qPCR. RNA in LENG8, THOC1 or Control precipitates was amplified by RT-qPCR using specific primers for two mRNAs (VDAC1 and VDAC2). (B) HeLa cells stably expressing shRNA targeting LENG8 were sent to cytosolic-nuclear fractioning, and total RNA from each fraction was extracted and sent to RNA sequencing. Scatter plot indicates individual RNAs sequenced. X axis shows log2FC (fold change) of the

ratio of cytosolic/nuclear RNAs after LENG8 knockdown. (C) Analysis of the cytosolic/nuclear ratio of mRNA enriched by LENG8, THOC1 or both. (D and E) RT-qPCR analysis of cytosolic or nuclear VDAC1 (D) and VDAC2 (E) mRNA after LENG8 knockdown. * p < 0.05, ** p < 0.01, *** p < 0.001 by the unpaired t-test (A, D and E). (F) FISH analysis of VDAC1 mRNA after LENG8 or PCID2 knockdown. Data are from three independent experiments (means ± s.e.m.).

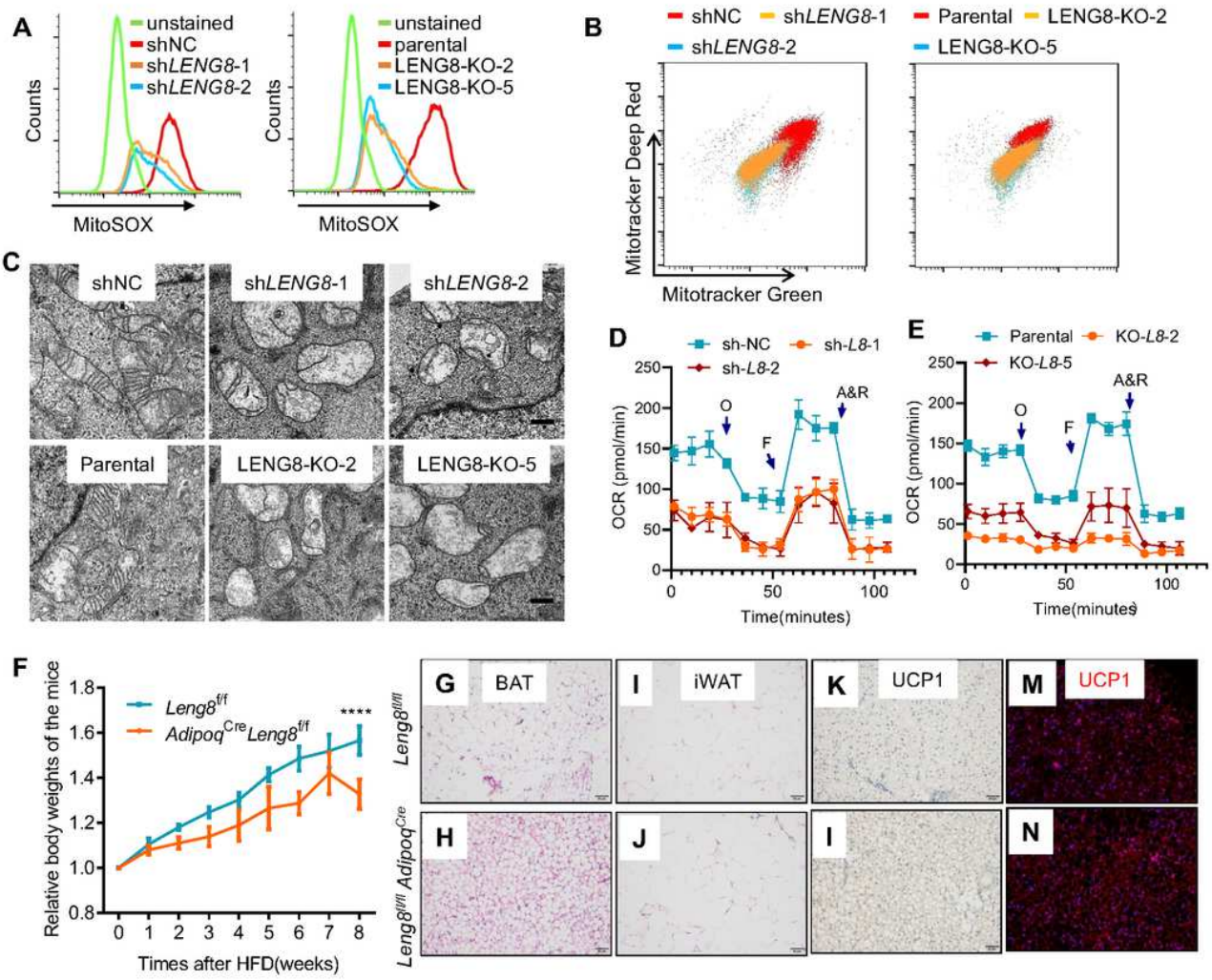

Figure 7

LENG8 controls mitochondrial activity. (A and B) HeLa cells of LENG8 knockdown or knockout were stained with MitoSOX (A) or Mitotracker Green and Mitotracker Deep Red (B) for 30 min and analyzed by flow cytometry. (C - E) HeLa cells of LENG8 knockdown or knockout were sent to electron microscopy imaging (C) or OCR assay (D and E). O for oligomycin, F for FCCP, A for antimycin and R for retenone. (F)

Relative body weight of Leng8fl/fl AdipoqCre and Leng8fl/fl mice after fed with high fat diet. (G - J) HE staining of brown adipose tissues (G and H) and inguinal white adipose tissue of Leng8fl/fl AdipoqCre and Leng8fl/fl mice after fed with high fat diet (I and J). (K - N) BAT from Leng8fl/fl AdipoqCre and Leng8fl/fl mice after fed with HFD were sent to immune-staining against UCP1. Data are representative of at least three independent experiments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFile.pdf
- SupplementaryTable1.xlsx
- SupplementaryTable2.xlsx
- SupplementaryTable3.xlsx
- SupplementaryTable4.xlsx
- SupplementaryTable5.xlsx
- SupplementaryTable6.xlsx