

Nucleolar URB1 ensures 3' ETS rRNA removal to prevent exosome surveillance

<https://doi.org/10.1038/s41586-023-05767-5>

Received: 17 January 2022

Accepted: 27 January 2023

Published online: 8 March 2023

 Check for updates

Lin Shan^{1,13}, Guang Xu^{1,13}, Run-Wen Yao^{1,12}, Peng-Fei Luan¹, Youkui Huang¹, Pei-Hong Zhang², Yu-Hang Pan¹, Lin Zhang³, Xiang Gao^{1,4}, Ying Li⁵, Shi-Meng Cao¹, Shuai-Xin Gao⁶, Zheng-Hu Yang^{1,4}, Siqi Li¹, Liang-Zhong Yang¹, Ying Wang¹, Catharine C. L. Wong⁷, Li Yu⁸, Jinsong Li^{3,4,9}, Li Yang¹⁰ & Ling-Ling Chen^{1,4,9,11}✉

The nucleolus is the most prominent membraneless condensate in the nucleus. It comprises hundreds of proteins with distinct roles in the rapid transcription of ribosomal RNA (rRNA) and efficient processing within units comprising a fibrillar centre and a dense fibrillar component and ribosome assembly in a granular component¹. The precise localization of most nucleolar proteins and whether their specific localization contributes to the radial flux of pre-rRNA processing have remained unknown owing to insufficient resolution in imaging studies^{2–5}. Therefore, how these nucleolar proteins are functionally coordinated with stepwise pre-rRNA processing requires further investigation. Here we screened 200 candidate nucleolar proteins using high-resolution live-cell microscopy and identified 12 proteins that are enriched towards the periphery of the dense fibrillar component (DFC). Among these proteins, unhealthy ribosome biogenesis 1 (URB1) is a static, nucleolar protein that ensures 3' end pre-rRNA anchoring and folding for U8 small nucleolar RNA recognition and the subsequent removal of the 3' external transcribed spacer (ETS) at the dense fibrillar component–DFC boundary. URB1 depletion leads to a disrupted DFC, uncontrolled pre-rRNA movement, altered pre-rRNA conformation and retention of the 3' ETS. These aberrant 3' ETS-attached pre-rRNA intermediates activate exosome-dependent nucleolar surveillance, resulting in decreased 28S rRNA production, head malformations in zebrafish and delayed embryonic development in mice. This study provides insight into functional sub-nucleolar organization and identifies a physiologically essential step in rRNA maturation that requires the static protein URB1 in the phase-separated nucleolus.

Hundreds of proteins and non-coding RNAs (ncRNAs) constitute the nucleolus, where rRNA biogenesis takes place. These nucleolar proteins and ncRNAs are organized spatially into three sub-nucleolar domains—the fibrillar centre (FC), the dense fibrillar component (DFC) and the granular component (GC)—with coordinated functions¹. Consistent with nucleolar organization, transcription of ribosomal DNA (rDNA) by RNA polymerase I (Pol I) occurs at the FC–DFC border, followed by co-transcriptional movement of the 5' ETS of the nascent 47S pre-rRNA into DFC to initiate processing⁴, leading to the

sequential production of precursors of 18S, 5.8S and 28S rRNAs in the DFC and further assembly of ribosomes in the GC^{1,6} (Extended Data Fig. 1a). Similar to processing of 5' ETS, the proper removal of 3' ETS is an essential step in rRNA processing that ensures 28S rRNA maturation, which is expected to occur before 28S rRNA translocation to the GC^{1,6} (Extended Data Fig. 1b). Notably, little is known about which proteins are required for 3' ETS processing, where this event occurs in a FC–DFC unit, how it is controlled, and its biological consequence.

¹State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China. ²CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China. ³State Key Laboratory of Cell Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China. ⁴School of Life Science and Technology, ShanghaiTech University, Shanghai, China. ⁵Cryo EM facility, Technology Center for Protein Sciences, School of Life Science, Tsinghua University, Beijing, China. ⁶Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing, China. ⁷Clinical Research Institute, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China.

⁸State Key Laboratory of Membrane Biology, Tsinghua University–Peking University Joint Centre for Life Sciences, Beijing Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing, China. ⁹Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China. ¹⁰Center for Molecular Medicine, Children's Hospital, Fudan University and Shanghai Key Laboratory of Medical Epigenetics, International Laboratory of Medical Epigenetics and Metabolism, Ministry of Science and Technology, Institutes of Biomedical Sciences, Fudan University, Shanghai, China. ¹¹New Cornerstone Science Laboratory, Shenzhen, China. ¹²Present address: Department of Biophysics and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA. ¹³These authors contributed equally: Lin Shan, Guang Xu. ✉e-mail: linglingchen@sibcb.ac.cn

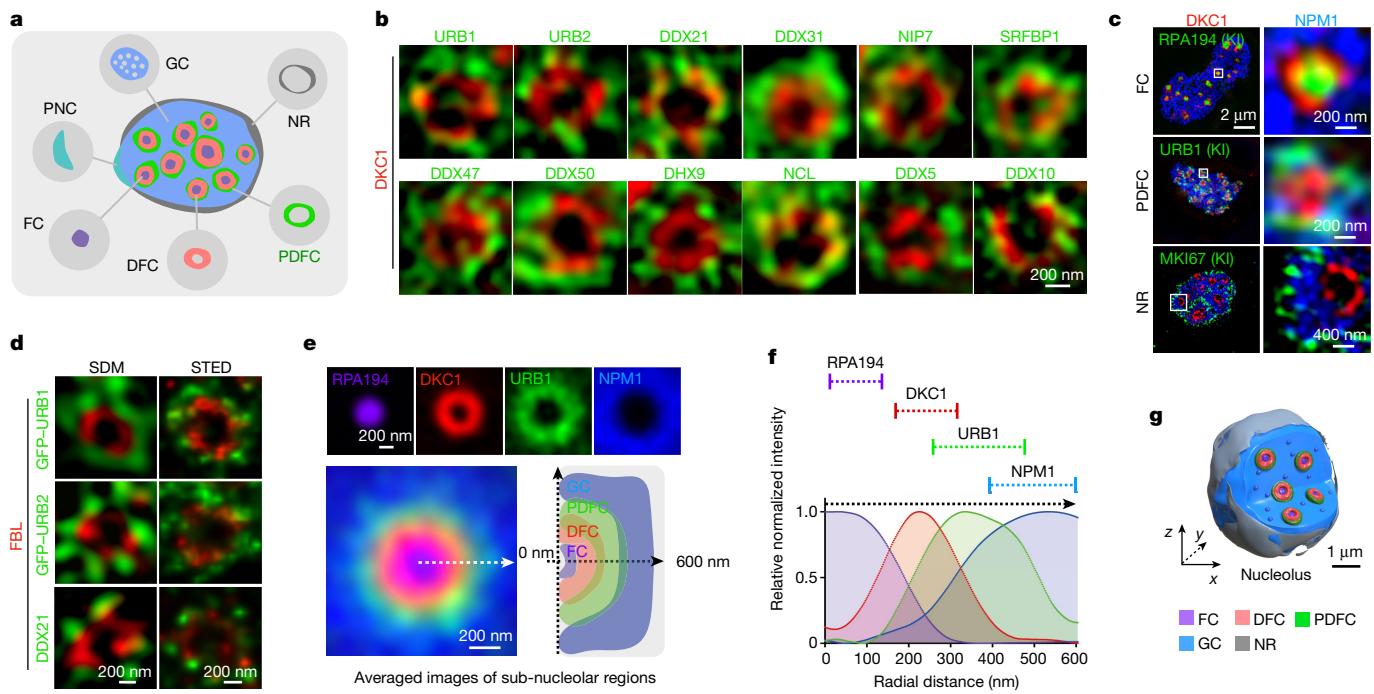


Fig. 1 | Image-based screening identifies 12 nucleolar proteins that are enriched in the PDFC. **a**, Schematic showing different sub-regions of the nucleolus. **b**, Representative structured illumination microscopy (SIM) images of indicated PDFC proteins and DKC1 (DFC marker) in fixed cells. **c**, Representative images of the nucleolus of HeLa cells with eGFP knock-in to create the indicated fusion proteins and also expressing mRuby3–DKC1 and mTagBFP2–NPM1. Right, magnified view of the outlined region of the main image (left). RPA194 marks FC, URB1 marks PDFC and MKI67 marks nucleolar rim. **d**, Representative superresolution immunofluorescence images of eGFP-knock-in fusions of URB1 and URB2 (detected with anti-GFP antibody),

and DDX21 (detected with anti-DDX21 antibody). Images were acquired by spinning disk (left) or STED (right) microscopy. **e**, The PDFC is a nucleolar sub-region located between the DFC and the GC. Top, aligned and averaged images of different nucleolar proteins show the radial flux structure of the nucleolus in HeLa cells. RPA194, DKC1, URB1 and NPM1 were used as markers for FC, DFC, PDFC and GC, respectively. $n = 60$ images. Bottom, schematic of nucleolar sub-regions. **f**, Variation of relative intensities of proteins with radial position in the nucleolus, quantified from images in **e**. **g**, Three-dimensional rendering of a representative human nucleolus, indicating the different nucleolar sub-regions in 3D. NR, nucleolar rim.

We first sought to examine systematically the precise 3D localization of a subset of potential nucleolar proteins identified by mass spectrometry and imaging approaches^{2,3,7,8}. This effort identified 12 proteins that were enriched towards the PDFC, where the static protein URB1 has an essential role in anchoring, folding and processing of the 3' ETS at PDFC to ensure 28S rRNA production during brain development in zebrafish and during embryonic development in mouse.

Localization of nucleolar proteins

To assess the 3D localization of nucleolar proteins, we first established a reference HeLa cell line, in which the DFC and GC were visualized on the basis of the expression of the marker proteins DKC1 and NPM1, respectively^{4,6,9}. mRuby3 and mTagBFP2 were inserted into the N termini of DKC1 and NPM1, respectively, using CRISPR–Cas9 (Extended Data Fig. 1c). Next, we collected a pool of 200 proteins of interest (POIs) (Extended Data Fig. 1d) that have been reported to localize^{2,7,10} or to have roles in the nucleolus^{11,12}, and constructed plasmids encoding these POIs with an eGFP tag at their 5' ends, driven by the CMV promoter. After transfecting each plasmid into the dual-labelled reference cell line, we performed high-resolution live-cell imaging with z-stacks. (Extended Data Fig. 1c).

We analysed the distribution data by comparing POI signals (eGFP) to reference signals in the DFC and GC (marked by mRuby3 and mTagBFP2, respectively), using Mander's overlap coefficient (MOC) as the established parameter (Extended Data Fig. 1e). Using this pipeline, we acquired a pilot dataset of all 200 POIs. The multi-colour distribution analysis examined more than six cells per POI, providing a large dataset for the spatial distribution-based clustering analysis (Extended Data Fig. 1e,f).

This screening identified around 70% of POIs (140 proteins) with a nucleolar localization, which were enriched in 6 nucleolar sub-regions, including FC, DFC, GC, nucleolar rim and peri-nucleolar compartment (PNC), and the sub-region that we termed PDFC. Around 30% of the POIs (60 proteins) exhibited no nucleolar signals (Extended Data Fig. 1g). The remaining 140 proteins included 97 GC-localized proteins, consistent with the preferential enrichment of ectopically expressed proteins in the largest nucleolar sub-region, the GC^{6,9}; 10 FC-localized proteins, including Pol I subunits and co-factors that facilitate transcription by Pol I; 13 DFC-localized proteins, including small nucleolar ribonucleoproteins (snoRNPs) and co-factors involved in the processing and modification of pre-rRNAs; 2 PNC-localized proteins involved in RNA splicing; and 5 nucleolar rim proteins involved in heterochromatin modification and the nucleolar rim protein MKI67^{13,14}. Of note, 12 proteins displayed a specific PDFC localization, including 7 DExD/H-box proteins and 5 pre-ribosomal particle-associated proteins (Fig. 1a and Extended Data Fig. 1h–l). Finally, some POIs displayed varied localization patterns; representative images for all POIs are collected in Extended Data Fig. 2.

Validation of PDFC proteins

All 12 POIs associated with the PDFC, which forms a discontinuous shell surrounding the DFC, partially overlap with DFC signals (Fig. 1b and Extended Data Fig. 3a). We next verified the sub-nucleolar localization of PDFC proteins, using mainly URB1 as an example. First, knock-in (KI) of eGFP to the N terminus URB1 in HeLa cells confirmed that URB1 surrounds the DFC (Fig. 1c). As controls, similar experiments confirmed the localization of RPA194 (the largest component of Pol I) in the FC

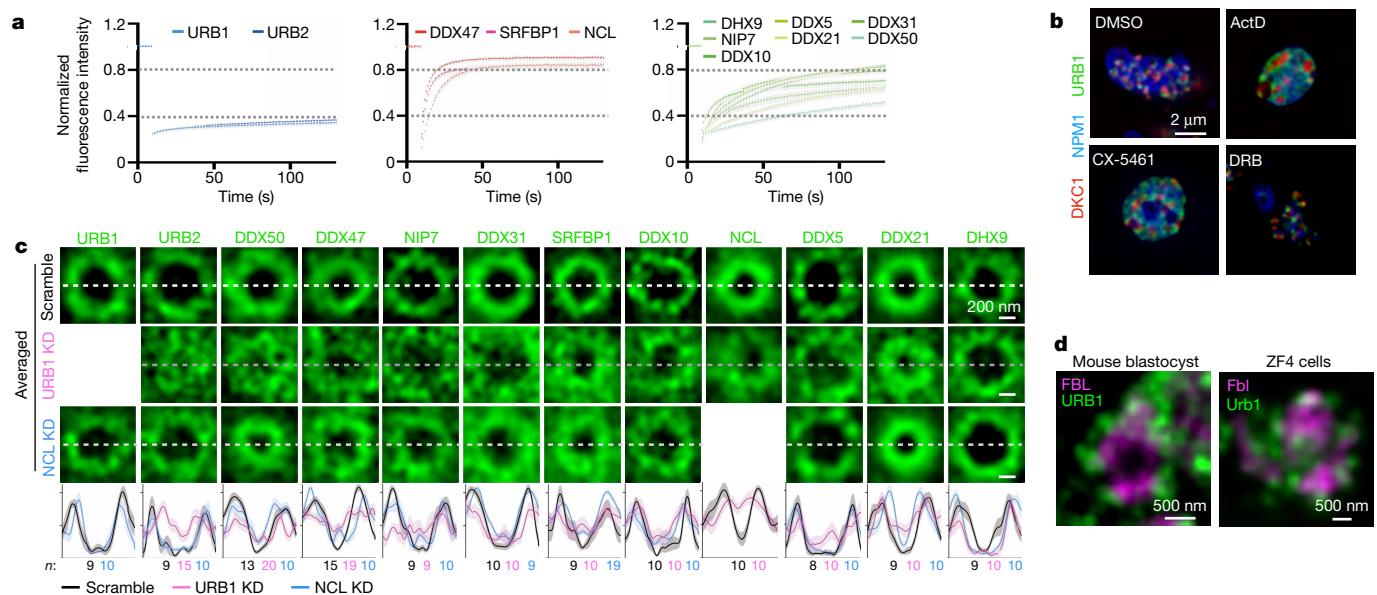


Fig. 2 | URB1 is a conserved PDFC protein and is required for PDFC integrity. **a**, FRAP assays of individual PDFC proteins in live HeLa cells, indicating their differential mobilities. $n = 10$. Proteins are classified into three groups—left, middle and right, in order of increasing relative mobility. Data are mean \pm s.e.m. **b**, URB1 surrounds the DFC under different conditions. Representative images of DFC (DKC1), PDFC (URB1) and GC (NPM1) proteins following treatment of live HeLa cells with DMSO (control), ActD, CX-5461 or DRB. mRuby3–DKC1 and mTagBFP2–NPM1 were ectopically expressed in mNeoGreen–URB1 knock-in HeLa cells, followed by the treatment (ActD: 10 ng ml $^{-1}$, 3 h; DRB: 50 μ M, 1.5 h;

CX-5461: 2 μ M, 1 h). **c**, URB1 knockdown (KD) disrupts PDFC integrity. Top, averaged images of PDFC proteins in HeLa cells treated with scramble shRNA (control) or shRNA targeting URB1 or NCL. Bottom, profile of the proteins along the dashed line in the images. The number of cells for each treatment is shown in the corresponding colour below the graph. Data are mean \pm s.e.m. **d**, URB1 surrounds DFC in the nucleolus of cells of different species. Left, immunofluorescence of FBL and URB1 in mouse blastocysts (E4.5). Right, exogenous expression of fluorescent proteins tagged Fbl and Urb1 in zebrafish ZF4 cells.

and MKI67 in the nucleolar rim (Fig. 1c). Second, antibody staining of eGFP–URB1, eGFP–URB2 and eGFP–DDX21 5 followed by visualization by spinning disk microscopy or stimulated emission depletion microscopy (STED) confirmed their PDFC localization (Fig. 1d). Third, correlative light and electron microscopy (CLEM) further showed that eGFP–URB1 is preferentially localized to the outside of the DFC (Extended Data Fig. 3b). Fourth, PDFC proteins were detected in other human cell lines (Extended Data Fig. 3c,d).

Next, we generated a HeLa cell line with four labels, each representing a nucleolar sub-region (RPA194 for FC, DKC1 for DFC, URB1 for PDFC and NPM1 for GC), to calculate the size of each region using averaged images (Fig. 1e). The diameters of the FC, DFC and GC were consistent with a previous report 4 , whereas the diameter of PDFC was in the range of 500 to 900 nm. The FC–DFC–PDFC unit was always embedded within the GC and has a radial profile in 3D, with FC in the centre, DFC around FC with a thickness of around 150 nm and PDFC outside DFC with a thickness of approximately 200 nm. All four radial compartments were partially overlapping (Fig. 1e,f and Extended Data Fig. 3e). The 3D model of the nucleolus model was generated on the basis of representative images and measured diameters of different sub-regions (Fig. 1g and Extended Data Fig. 3f,g). Together these analyses reveal the existence of PDFC as a layer in each FC–DFC unit.

Characterization of PDFC proteins

Ten out of the twelve PDFC proteins contained intrinsically disordered regions 15 (IDRs), although the two largest PDFC proteins—URB1 (around 2,250 amino acids) and URB2 (around 1,500 amino acids)—had no obvious IDRs (Extended Data Fig. 4a,b). IDRs have a role in the formation of layered droplets 16 . As the length and strength of IDRs in proteins are associated with protein dynamics 4 , we ectopically expressed eGFP-tagged PDFC proteins in HeLa cells and used fluorescence recovery after photobleaching (FRAP) to characterize their

mobilities (Fig. 2a). All IDR-containing proteins displayed mobility, with NCL, SRFBP1 and DDX47 showing the fastest movement, followed by DDX10, DDX21, DDX50, DDX31, DDX5, NIP7 and DHX9, whereas URB1 and URB2 remained static within PDFC. The lack of IDRs and mobility are distinct features of URB1 and URB2—which have different sequences but similar predicted structures (Extended Data Fig. 4c)—among the PDFC proteins.

URB1 persistently coated the DFC in cells grown under different conditions (Extended Data Fig. 4d) and during inhibition of Pol I transcription (Fig. 2b). Inhibition of Pol I transcription with a low dose (10 ng ml $^{-1}$) of actinomycin D 17 (ActD) or with CX-5461 18 resulted in a typical DFC segregation from GC, while the shell-like PDFC structure, indicated by URB1 (Fig. 2b), URB2, and other PDFC proteins (Extended Data Fig. 4e,f), still surrounded DFC. The Pol II inhibitor 5,6-dichloro- β -D-ribofuranosylbenzimidazole (DRB) induced nucleolar disorganization, indicated by the altered localization of rRNA processing factors (such as DKC1; Fig. 2b) into the nucleoplasm 19 , but URB1 still surrounded the DFC (Fig. 2b). Together, these results suggest that the PDFC is functionally connected to FC–DFC units.

Given its large size and immobility, we investigated the possible contribution of URB1 to PDFC integrity. URB1 depletion resulted in a dispersed distribution of most PDFC proteins in the reference HeLa cell line (Fig. 2c and Extended Data Fig. 4g). For example, the shell-like structure of URB2 and DDX50 was lost, and extended into the DFC and GC. DDX47, NIP7, DDX31, SRFBP1, DDX10, NCL and DDX5 also diffused into DFC and GC, but partially retained the shell-like organization. The PDFC localization of DDX21 and DHX9 was maintained, DDX21 formed a smaller shell-like structure and whereas the shell-like structure of DHX9 was increased in size (Fig. 2c). Each examined PDFC protein displayed a shell-like organization surrounding the DFC in scramble short hairpin RNA (shRNA)-treated control cells (Fig. 2c). Further, knockdown of the PDFC protein NCL had a much milder effect on the localization of PDFC proteins compared with URB1 knockdown (Fig. 2c), although

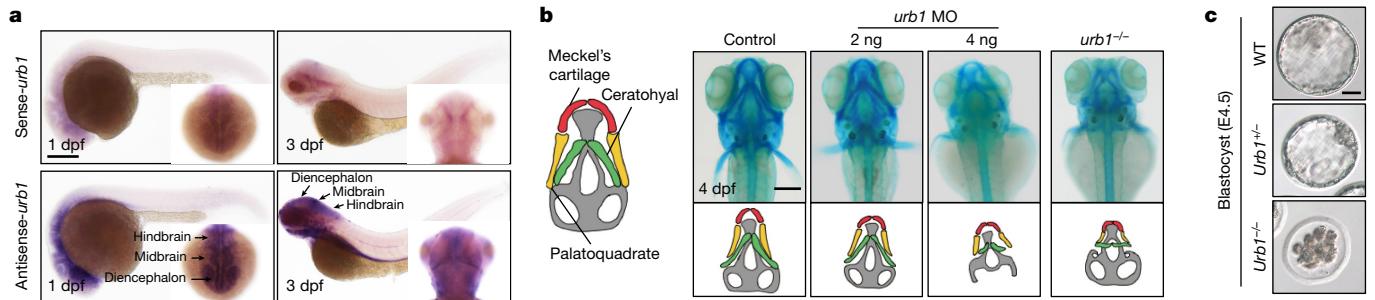


Fig. 3 | URB1 is required during early development. **a**, *urb1* is expressed in zebrafish brain from 1 dpf to 3 dpf, indicated by WISH with sense (top) or antisense (bottom) *urb1* probes. Arrows mark diencephalon, midbrain and hindbrain. $n = 4$ embryos. **b**, URB1 is required for zebrafish embryonic development. Left, schematic of craniofacial bone structures in zebrafish. Right (top row), representative dorsal view images of larvae cranial cartilages

at 4 dpf with different morpholino (MO) dosages, or with *urb1* knockout (*urb1*^{-/-}). Right (bottom row), traces display the Meckel's and ceratohyal defects in *urb1*-knockdown (*urb1* MO) and *urb1*^{-/-} larvae. $n = 6$ fish. **c**, URB1 is required for mouse embryonic development. Representative E4.5 blastocysts isolated from *Urb1*^{-/-} intercrosses. Embryos were collected from three biologically independent experiments. Scale bars, 100 μ m (a,b), 20 μ m (c).

the DDX50 and DDX21 shells appeared smaller (Fig. 2c). These results suggest that URB1 acts as a static nucleolar protein with a role in PDFC organization. Furthermore, URB1 was localized to the PDFC in mouse blastocysts and Urb1 was localized to the PDFC in zebrafish ZF4 cells, which contain a merged FC–DFC bipartite nucleolar structure (Fig. 2d and Extended Data Fig. 5a). This conserved nature of URB1 indicates its physiological importance in nucleolar function.

URB1 is required for development

Depletion of the URB1 and URB2 homologues Npa1p and Npa2p, respectively, in yeast is lethal, and Npa1p appears to be essential for pre-rRNA processing and ribosome biogenesis, but the precise roles of Npa1p and Npa2p are unclear^{20,21}. Whole-mount *in situ* hybridization (WISH) in zebrafish larvae revealed that *urb1* was expressed in a specific time window from 1 to 3 days post-fertilization (dpf) (Fig. 3a and Extended Data Fig. 5b), with a specific enrichment in the zebrafish brain.

Consistently, knockdown of *urb1* expression by morpholinos (Extended Data Fig. 5c,d) resulted in severe head hypoplasia and deformity of mandibular and hyoid stream cartilage structures at 1 dpf, which further developed into small head deformity at 4 dpf in *urb1*-knockdown (KD) larvae (Fig. 3b). Notably, we observed more severe defects in head cartilage development and higher malformation rates at higher morpholino doses (Fig. 3b and Extended Data Fig. 5e), and these defects were largely rescued by injecting *urb1* mRNA into *urb1*-knockdown embryos (Extended Data Fig. 5f,g). *urb1*^{-/-} zebrafish embryos generated by CRISPR–Cas9 (Extended Data Fig. 5h,i) also displayed head hypoplasia and severe defects in the development of Meckel's, ceratohyal and palatoquadrate cartilage structures (Fig. 3b and Extended Data Fig. 5j,k). Compared to non-injected morpholino controls (Fig. 3b) and *urb1*^{-/-} siblings (Extended Data Fig. 5l,m), the developmental delay in Meckel's and ceratohyal cartilage of *urb1*^{-/-} zebrafish was significant at 3–4 dpf. The developmental defects observed with *urb1* depletion suggest a crucial role of *urb1* in ribosome biogenesis linking to ribosomopathies²².

We next generated *Urb1*-knockout mice (Extended Data Fig. 5n). *Urb1*^{-/-} heterozygous mice were viable, with no apparent phenotypic abnormalities. However, heterozygote intercrosses did not produce viable homozygous *Urb1*^{-/-} offspring. At embryonic day 4.5 (E4.5), *Urb1*^{-/-} embryos were arrested at the morula stage and did not form distinct blastocoels (Fig. 3c and Extended Data Fig. 5o), indicating a pre-implantation arrest and lethality caused by URB1 depletion. URB1 depletion in human pluripotent stem cells also resulted in a substantial delay in colony formation (Extended Data Fig. 5p). Together, these results suggest that URB1 has an essential and conserved role in early development.

Loss of URB1 impairs 3' ETS removal

Next, we investigated whether and how URB1 regulates the pre-rRNA processing at the PDFC that ultimately contributes to its crucial role in early development (Fig. 3). We used northern blot probes²³ that recognize 5' ETS, internal transcribed spacer 1 (ITS1) or ITS2 to detect and quantify pre-rRNA intermediates. Loss of URB1 resulted in a marked reduction of 32S and 12S pre-rRNAs (precursors of 28S and 5.8S rRNAs, respectively) (Fig. 4a and Extended Data Fig. 6a) whereas 30S and 21S pre-rRNAs (precursors of 18S rRNAs) remained largely unaltered (Extended Data Fig. 6b,c). As a result, we observed a notable reduction of 28S rRNA, but not 18S rRNA, upon URB1 depletion (Fig. 4b,c). Further, northern blot analysis with the 5.8S probe showed that the 32S pre-rRNA-associated aberrant rRNA intermediates were derived from the erroneous processing of 32S pre-rRNA from its 3' end upon URB1 depletion (Extended Data Fig. 6d).

The eukaryotic 47S pre-rRNA transcript contains a 3' ETS next to the 3' end of 28S rRNA (Fig. 4d and Extended Data Fig. 1b). Unlike the 5' ETS processing at the FC–DFC border¹, how the 3' ETS processing is controlled has remained largely unclear. Using a northern blot probe that recognizes 3' ETS, we found that pre-rRNA intermediates ranging from 47S to 28S generated upon URB1 knockdown all contained the 3' ETS, whereas the hybridization signals for 3' ETS were barely detected in cells treated with scramble shRNA (Fig. 4d). Quantification revealed an approximately sixfold increase in the level of 32S pre-rRNA with attached 3' ETS in URB1-KD cells compared with control cells (Fig. 4d). Notably, these 32S pre-rRNAs with attached 3' ETS (named as 32S*) accounted for the majority of the detected 32S pre-rRNAs and migrated more slowly than the wild-type isoform in URB1-KD cells (Extended Data Fig. 6e). In sum, these analyses confirmed that URB1 is involved in 3' ETS removal, which is essential for 28S rRNA production (Fig. 4a–d).

The essential role of URB1 in 3' ETS processing was confirmed in 293FT cells (Extended Data Fig. 6f). Of note, although URB1 and URB2 had similar physical properties (Fig. 2a and Extended Data Fig. 4), the loss of URB2 or several other PDFC proteins barely impaired 32S pre-rRNA processing (Extended Data Fig. 6g,h), confirming the unique role of URB1 in controlling 3' end processing of 47S pre-rRNA.

URB1 regulates 3' ETS removal

We explored how URB1 controls 3' ETS processing. Individual-nucleotide resolution cross-linking and immunoprecipitation (iCLIP) revealed that URB1 robustly and specifically interacts with a region (nucleotides 12878–12966 of the 47S pre-rRNA) at the 3' end of 28S rRNA, which is adjacent to the upstream 3' ETS (Fig. 4e). Following loss of URB1, we

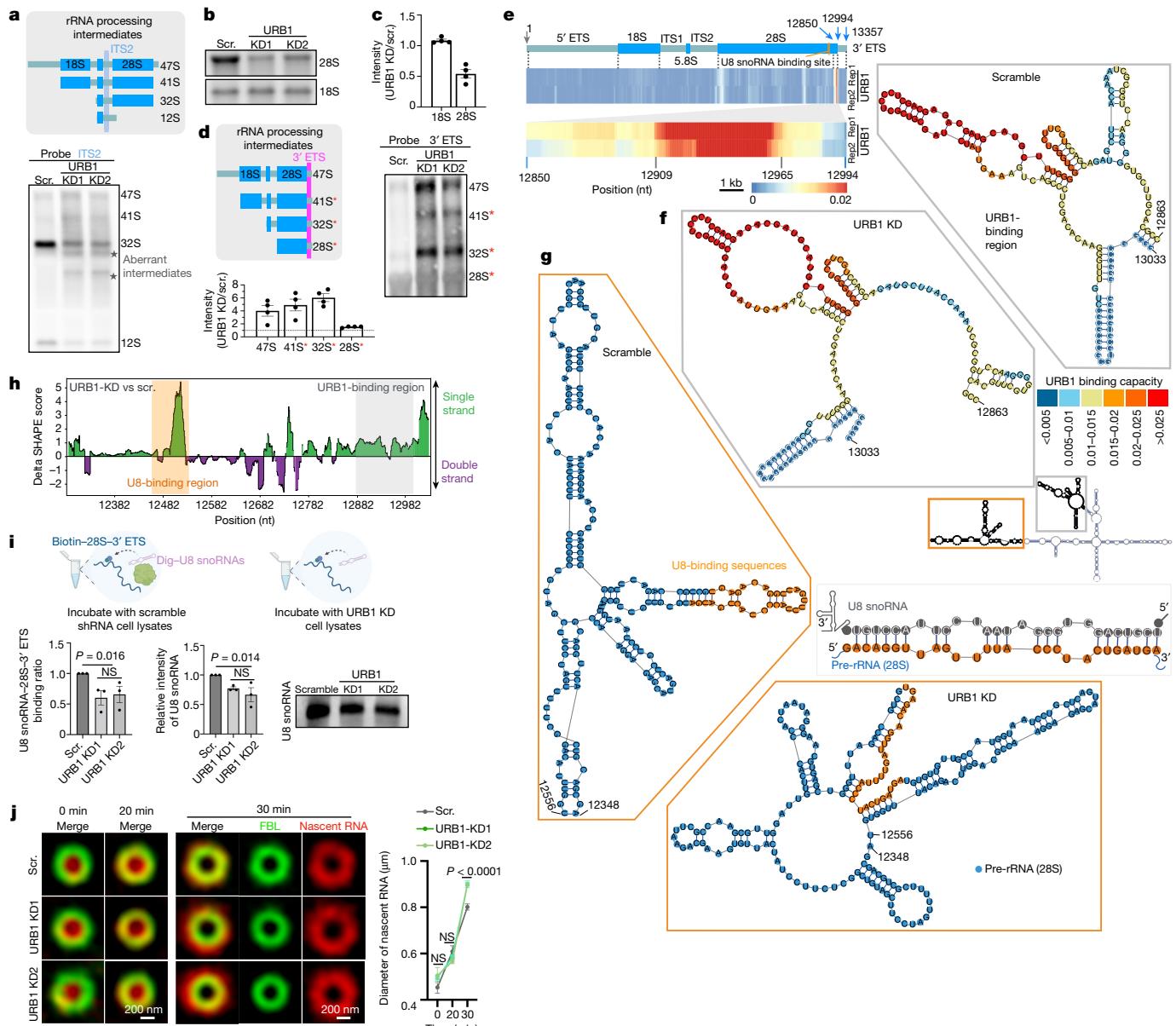


Fig. 4 | URB1 depletion disrupts 3' ETS removal by altering U8 snoRNA binding to pre-32S rRNA. **a**, Top, schematic of human pre-rRNA processing and probes used in northern blot experiments. Bottom, northern blot of pre-rRNA intermediates in HeLa cells treated with scramble shRNA (scr.) and URB1-KD HeLa cells, using probes targeting ITS2. Aberrant intermediates are indicated with grey stars. **b**, Ethidium bromide staining of 28S and 18S rRNAs is shown as loading control for equal amounts of RNA as in **a**. **c**, Relative intensities of 28S and 18S rRNAs in **b**. $n = 4$ experiments. **d**, Left, schematic of human 3' ETS-related pre-rRNA intermediates. Right, northern blot analyses with a 3' ETS probe detected pre-rRNA intermediates in scramble shRNA-treated and URB1-KD HeLa cells. Bottom, relative intensities of detected pre-rRNA intermediates. $n = 4$ experiments. 41S*, aberrant 3' ETS-attached 41S pre-rRNA; 32S*, aberrant 3' ETS-attached 32S pre-rRNA; 28S*, aberrant 3' ETS-attached 28S rRNA.

e, Anti-Flag iCLIP using mNeonGreen-Flag-URB1 knock-in HeLa cells. Bottom, magnified view of the URB1-binding region in the 3' terminus of 28S rRNA.

f, Top, the secondary conformation of the URB1-binding region analysed in **e**.

scramble shRNA-treated and URB1-KD HeLa cells by in-cell SHAPE-MaP. Bottom right, schematic of the URB1-binding region (grey rectangle) and the upstream sequence (orange rectangle). **g**, Right, predicted U8 snoRNA-binding sequences in 28S rRNA³⁷. Left, the secondary conformation of pre-rRNA targeted by U8 analysed in scramble shRNA-treated (top) and URB1-KD (bottom) HeLa cells.

h, DeltaSHAPE comparison of the SHAPE value between scramble shRNA-treated and URB1-KD HeLa cells.

i, Top, schematic of the in vitro binding assay for biotin-28S-3' ETS and digoxigenin (Dig)-U8 snoRNA. Bottom left, quantitative PCR with reverse transcription (RT-qPCR) of pre-rRNA-bound U8 fragments incubated with scramble shRNA-treated or URB1-KD cell lysates. Bottom right, the corresponding northern blot. U8 was detected with an anti-Dig antibody. Three independent experiments.

j, Left, representative averaged images of the distribution of nascent pre-rRNA at indicated time points ($n = 10$). Right, corresponding analyses of the distribution of nascent pre-rRNA. **c,d,i,j**, Data are mean \pm s.e.m. **i,j**, Two-tailed Student's *t*-test.

detected an increased proportion of single-stranded RNA at the 28S–3' ETS boundary region (nucleotides 12683–13036 of the 47S pre-rRNA) where URB1 binds (Fig. 4e) using in-cell selective 2'-hydroxyl acylation analysed by primer extension and mutational profiling (SHAPE-MaP)

with 2-methylnicotinic acid imidazolide (NAI) (Fig. 4f and Extended Data Fig. 7). Application of dimethyl sulfate (DMS), which is less prone than NAI reagents to association with RNA-binding proteins²⁴, also revealed an increased proportion of single-stranded RNA upon URB1

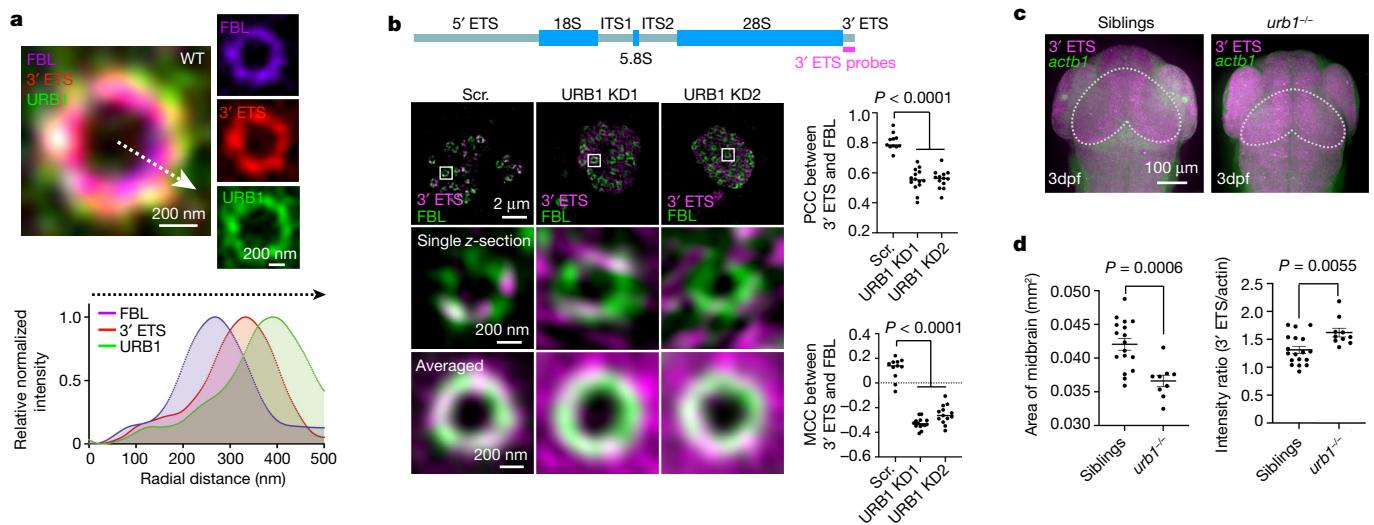


Fig. 5 | URB1 depletion impairs the distribution of 3' ETS-containing pre-rRNA molecules and nucleolar morphology. **a**, 3' ETS-attached pre-rRNA molecules localize at the DFC–PDFC boundary in wild-type (WT) HeLa cells. Top, representative aligned and averaged SIM images of FBL, 3' ETS (the FISH probe is shown in **b**) and URB1. $n = 10$. Bottom, relative intensities of FBL, 3' ETS and URB1. **b**, URB1 knockdown impairs the radial flux processing of pre-rRNA. Top, schematic of 3' ETS smFISH probes. Bottom, representative and averaged SIM images in scramble shRNA-treated and URB1-KD HeLa cells. $n = 20$ cells.

Right, statistics of Pearson correlation coefficient (PCC) and Manders' colocalization coefficient (MCC) of FBL and 3' ETS-attached pre-rRNAs. $n = 10$. Two-tailed Student's *t*-test. **c**, Aberrantly processed 3' ETS-attached pre-rRNAs accumulate in *urb1*^{-/-} zebrafish. 3' ETS-attached pre-rRNAs in siblings and *urb1*^{-/-} mutants at 3 dpf. *actb1* is shown as control. **d**, Midbrain (dashed line in **c**) size and intensity of signal for 3' ETS-attached pre-rRNA from **c**. $n = 26$ fish. Data are mean \pm s.e.m. Two-tailed Student's *t*-test.

loss, indicating that the observed alteration in RNA conformation was independent of binding to RNA-binding proteins (Extended Data Fig. 7c,d).

The altered pre-rRNA conformation upon URB1 loss could be extended to around 500 nucleotides (nt) upstream of the URB1-binding site (around nucleotide 12328 of 47S pre-rRNA) (Fig. 4e–h and Extended Data Fig. 7e), indicating that the binding of the 3' end of 28S rRNA to URB1 may provide a docking site to ensure the proper pre-rRNA folding, and may thereby be critical for the subsequent pre-rRNA processing.

Box C/D U8 small nucleolar RNAs (snoRNAs) are required for 5'-end processing of 5.8S RNA, 3'-end remodelling of 28S RNA and 3' ETS removal^{25,26}. The U8 snoRNA-binding site on rRNA is located 300 nt upstream of the URB1-binding site and falls in the region that exhibited altered conformation upon URB1 loss (Fig. 4e,g,h). In particular, the U8-binding site (nucleotides 12328–12638 of the 47S pre-rRNA) exhibited increased single-stranded structure (Fig. 4g,h and Extended Data Fig. 7e), which could be owing to the loss of U8 binding as a result of altered pre-rRNA folding upon URB1 depletion, suggesting a potential role of URB1 in facilitating U8 binding. To test this hypothesis, we incubated U8 snoRNA and 3' ETS-attached 28S RNA in vitro and added cell lysates obtained from with scramble shRNA-treated control cells or URB1-KD cells. Indeed, U8 binding to 28S RNA was reduced in the absence of URB1 (Fig. 4i).

Next, we measured the pace of the radial flux nascent pre-rRNAs movement towards the GC by metabolic chasing and measuring their most frontier fluorescence signals (Extended Data Fig. 8a). In the absence of URB1, the speed of radial flux of pre-rRNAs across the PDFC was increased within 30 min (Fig. 4j). These observations indicate that URB1 binding to the 3' end of pre-rRNA is essential for pre-rRNA folding and U8 snoRNA recognition, as well as ensuring sufficient time for these events to occur. This view is further supported by the observation that many PDFC POIs, including four DExD/H-box proteins known to be involved in pre-rRNA processing^{27–30}, preferentially bind to the 28S and 3' ETS regions of pre-rRNAs (Extended Data Fig. 8b). Collectively, the key role of URB1 in 3' ETS removal in the PDFC

highlights the importance of the layered nucleolar sub-regions for stepwise pre-rRNA processing.

Removal of 3' ETS at the DFC–PDFC boundary

Given that 5' ETS processing occurs co-transcriptionally, accompanied by the sorting of nascent 47S pre-rRNA from the FC–DFC border to the DFC⁴, we speculated that 3' ETS processing takes place towards the PDFC (Fig. 4). Visualization of 3' ETS-attached intermediates by single-molecule fluorescence *in situ* hybridization (smFISH), together with mTagBFP2–FBL and eGFP–URB1 knock-in revealed that these 3' ETS-attached pre-rRNAs were located mainly in a ring-shaped region around the DFC (Fig. 5a and Extended Data Fig. 9a). Similarly, U8 snoRNAs also surrounded the DFC (Extended Data Fig. 9b), consistent with its role in 3' ETS removal²⁵ and our hypothesis that the DFC–PDFC boundary is where 3' ETS processing takes place.

Upon URB1 depletion, these 3' ETS-attached intermediates (Fig. 4d) and U8 snoRNAs were no longer located within the DFC–PDFC boundary, and were dispersed in the GC (Fig. 5b and Extended Data Fig. 9c,d), as shown by reduced colocalization between 3' ETS-attached pre-rRNA intermediates and FBL in URB1-depleted cells compared with control cells (Fig. 5b). Together, these analyses showed that pre-rRNA 3' ETS processing occurs at the DFC–PDFC boundary (Fig. 5a,b) and that the layered nucleolar proteins can facilitate efficient pre-rRNA processing.

Loss of URB1 dampens nucleolar function

Consistent with the reduced amount of 28S rRNA upon URB1 depletion, in both HeLa and 293FT cells, the biogenesis of 60S ribosomal subunits was markedly impaired, and the numbers of monosomes and polysomes were correspondingly reduced (Extended Data Fig. 9e). The impaired ribosome biogenesis was accompanied by visible shrinkage of the nucleolus and a reduced GC volume in URB1-KD cells (Extended Data Fig. 9f,g). Although the number of DFCs per cell remained unchanged, the distance between adjacent DFC units decreased (Extended Data

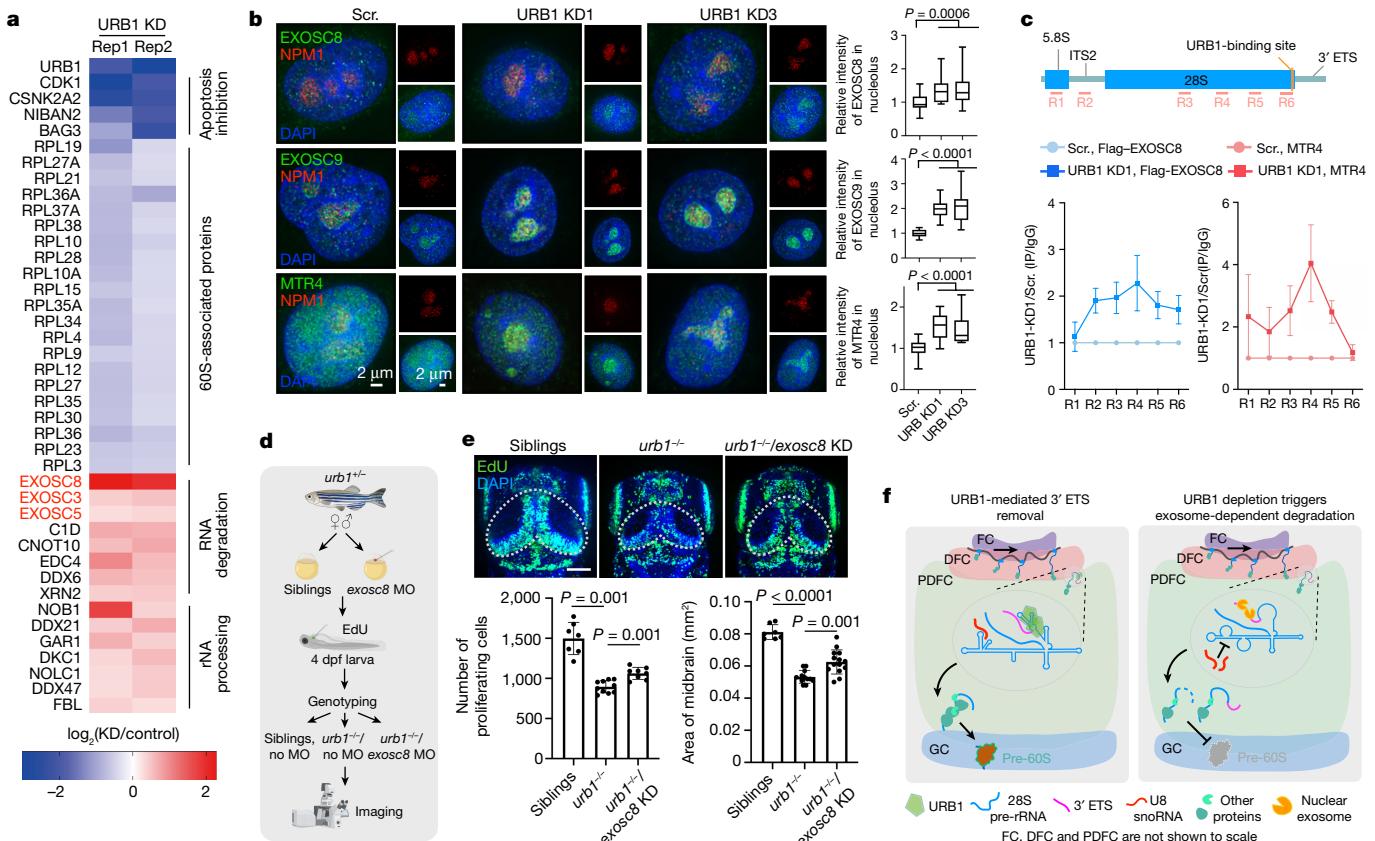


Fig. 6 | URB1 depletion activates exosome-dependent pre-rRNA surveillance, impairing embryonic development. **a**, Heat map of proteins altered upon URB1 knockdown in HeLa cells shows reduced expression of ribosome components and increased expression of exosome components. The experiment was performed twice (rep1 and rep2). **b**, Representative images of EXOSC8, EXOSC9 and MTR4 (left), and their relative intensities (right) in scramble shRNA-treated and URB1-KD HeLa cells are shown. Exosome components accumulate in nucleoli upon URB1 knockdown. $n = 77$ (scramble), 70 (URB1 KD1) and 65 (URB1 KD3) cells over 3 independent experiments. Bars show the range of values, boxes delineate the middle quartiles and the centre line represents the median. **c**, RNA immunoprecipitation (IP) assays followed by RT-qPCR with primers targeting 32S pre-rRNA shows increased EXOSC8 and MTR4 binding with 28S pre-rRNA in URB1-KD HeLa cells. Signals of anti-MTR4

and anti-Flag-EXOSC8 are normalized to those of anti-IgG of the same set of primers. $n = 3$. **d**, The workflow of the experiments shown in **e**. Created with BioRender.com. **e**, EXOSC8 knockdown (KD) partially rescues phenotypes of *urb1^{-/-}* zebrafish. Top, representative EdU-positive proliferating cells in control siblings, *urb1^{-/-}* and *urb1^{-/-}* exosc8-KD larvae at 3 dpf. Bottom, number of proliferating cells and midbrain area. $n = 25$ fish. **f**, PDFC-localized URB1 mediates 3' ETS removal from pre-rRNA, which is conserved and essential for cell proliferation and embryonic development. URB1 depletion results in weakened binding between U8 snoRNA and pre-rRNA. Consequently, the accumulation of 3' ETS-attached aberrant pre-rRNA molecules triggers the activation of nucleolar surveillance by exosome complexes. **c,e**, Data are mean \pm s.e.m. **b**, Two-sided Mann–Whitney tests. **e**, Two-tailed Student's *t*-test.

Fig. 9g), and mTagBFP2–NPM1 in URB1-KD cells showed a reduced mobility (Extended Data Fig. 9h). Owing to the reduced GC volume, some DFC units could not be fully embedded within the GC or the nucleolus, consistent with the impaired pre-rRNA processing and ribosome biogenesis (Fig. 4). We observed a corresponding reduction of cell proliferation upon URB1 knockdown (Extended Data Fig. 9i).

These nucleolar functional and morphological changes accompanying URB1 depletion highlighted a close functional connection between nucleolar substructure and ribosome biogenesis that is linked to cell proliferation (Extended Data Fig. 9i) and developmental defects (Fig. 3). Consistently, 3' ETS-attached pre-rRNAs in zebrafish ZF4 cells were also localized to the periphery of the merged FC–DFC region, where the zebrafish U8 snoRNA was localized (Extended Data Fig. 9j). Moreover, *urb1^{-/-}* mutants displayed similar phenotypes—with less-defined midbrain-hindbrain boundary and disorganized trunk vasculature—as observed in U8-knockout zebrafish²⁶ (Extended Data Fig. 9k), indicating that both URB1 and U8 are involved in the same 3' ETS removal pathway. Indeed, the level of 3' ETS-attached pre-rRNAs in the head, especially in the midbrain, was increased in *urb1^{-/-}* zebrafish embryos, accompanying the delayed midbrain growth (Fig. 5c,d).

Furthermore, smFISH revealed a diffuse distribution of 3' ETS-attached pre-rRNA intermediates to GC in Urb1-KD NIH 3T3 cells and *Urb1*-knockout mouse E4.0 embryos (Extended Data Fig. 9l,m), similar to what was observed in URB1-KD HeLa cells.

URB1 loss activates exosome surveillance

Dysregulated rRNA processing and 60S ribosomal subunit assembly have been linked to impaired protein translation^{31,32}. The reduced levels of rRNA and 60S ribosomal subunits upon URB1 loss prompted us to identify proteins with altered abundance upon URB1 depletion and the specific pathway associated with URB1-controlled 3' ETS processing. We performed stable isotope labelling using amino acids in cell culture (SILAC) followed by mass spectrometry in scramble shRNA-treated and URB1-KD cells (Fig. 6a). The abundance of URB1 proteins was decreased to 25% of wild-type levels in URB1-KD HeLa cells, validating the efficiency of URB1 depletion and the accuracy of SILAC labelling results (Extended Data Fig. 10a). In total, 4,688 proteins were detected in replicates after URB1 knockdown, among which 274 proteins were upregulated more than twofold and 170 proteins were downregulated

more than twofold (Extended Data Fig. 10b,c). The expression of 38 large ribosomal proteins (RPLs) associated with the 60S complex was significantly reduced (Extended Data Fig. 10d), consistent with the reduction in 28S rRNA (Fig. 4b,c) and 60S ribosomal subunit (Fig. 5d and Extended Data Fig. 9h) in URB1-KD cells. Notably, the expression of several components of the exosome complex involved in RNA degradation and nucleolar factors involved in pre-rRNA processing was increased (Fig. 6a and Extended Data Fig. 10d,e).

Both core components of the nuclear exosome complex (EXOSC8 and EXOSC9) and the cofactor MTR4 accumulated in nucleoli after URB1 knockdown (Fig. 6b), although the total abundance of EXOSC9 and MTR4 were not altered (Extended Data Fig. 10e). In addition, EXOSC9 and MTR4 were co-localized with the aberrantly processed 3' ETS-containing pre-rRNA intermediates (Extended Data Fig. 10f). RNA immunoprecipitation assays showed increased binding of EXOSC8 and MTR4 to pre-28S rRNA transcripts upstream of the URB1-binding site (Fig. 6c). Of note, in nucleoli, MTR4 is involved in the processing of rRNA precursors and the decay of processing by-products³³. These results suggest that the augmented exosome components in the nucleolus could contribute to pre-rRNA surveillance by degrading the aberrantly processed transcripts from their 3' end upon URB1 depletion (Fig. 4a–d and Extended Data Fig. 10e). To explore this possibility, we assessed whether loss of exosome components in URB1-KD HeLa cells affected 32S pre-rRNA abundance. Knockdown of most exosome components (EXOSC1, EXOSC7, EXOSC8, EXOSC9 and EXOSC10) or co-factors (ERI1 and MTR4) could at least partially rescue the misprocessing and reduction of 32S pre-rRNA, whereas their knockdown in wild-type cells did not affect normal 32S pre-rRNA processing (Extended Data Fig. 10g–i). Consistently, these extended 3' ETS-containing pre-rRNAs carry short poly(A) tails known to target RNAs for exosome-mediated degradation³⁴, as shown by 3' rapid amplification of cDNA ends (RACE) and oligo-dT pull-down upon URB1 depletion (Extended Data Fig. 10j,k).

These observations support the view that the exosome components cooperate with URB1 in the surveillance of 3' ETS processing of pre-rRNAs (Extended Data Fig. 10h), a process that is essential for development (Fig. 3). Consistently, further knockdown of the exosome component gene *exosc8* in *urb1*^{-/-} zebrafish embryos (Fig. 6d) alleviated the cell proliferation defect and head deformity (Fig. 6e).

Discussion

Here we have identified PDFC as a unique nucleolar sub-region that contains at least 12 proteins, including URB1 (Fig. 1). URB1 supports PDFC integrity (Fig. 2) and ensures 3' ETS removal by acting as an immobile protein to anchor the 3' end of pre-rRNA (Fig. 4e,f), thereby providing a sufficient residence time for pre-rRNA folding and U8 recognition (Fig. 4g–j) required for pre-rRNA processing²⁶ at the DFC–PDFC boundary (Fig. 5). Such a highly coordinated processing of pre-rRNA in the nucleolus with a 3D radial flux model can facilitate pre-rRNA radial transport along these layered nucleolar structures. The requirement of the static URB1 for the appropriate pre-rRNA folding and processing amongst phase-separated proteins in the nucleolus indicates a general principle that some immobile proteins can facilitate the folding of single-stranded RNAs in complex condensates.

The rapid transcription of hundreds of rDNA repeats and the massive pre-rRNA processing make ribosome production a multi-step, error-prone process. By nature, it requires hundreds of nucleolar factors to coordinately function in specialized sub-nucleolar domains. URB1 at the PDFC provides an essential platform to ensure 3' ETS folding and processing (Figs. 4 and 5), preventing the activation of nucleolar exosome surveillance by the aberrant 3' ETS-attached pre-rRNAs (Fig. 6a–c and Extended Data Fig. 10j,k). This erroneous processing in the absence of URB1 leads to defects in cell proliferation (Extended Data Fig. 9i) and failure in development (Fig. 3).

The activation of exosome-dependent surveillance upon URB1 loss indicates that the 3' ETS-attached pre-rRNAs can recruit exosomes for degradation, similar to the exosome-dependent surveillance pathway that has been found to eliminate defective 60S pre-ribosomal subunits^{35,36}. However, there is also the possibility that URB1 binding per se blocks exosome recruitment. Further studies are warranted to clarify the underlying mechanism. Nonetheless, the static URB1-mediated 3' ETS removal in the PDFC and the coupled, exosome-pre-rRNA surveillance pathway are essential and evolutionarily conserved (Fig. 6f).

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-023-05767-5>.

1. Lafontaine, D. L. J., Riback, J. A., Bascetin, R. & Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **22**, 165–182 (2021).
2. Thul, P. J. et al. A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017).
3. Huh, W. K. et al. Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
4. Yao, R. W. et al. Nascent pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus. *Mol. Cell* **76**, 767–783.e711 (2019).
5. Wu, M. et al. lncRNA SLERT controls phase separation of FC/DFCs to facilitate Pol I transcription. *Science* **373**, 547–555 (2021).
6. Boisvert, F. M., van Koningsbruggen, S., Navascues, J. & Lamond, A. I. The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* **8**, 574–585 (2007).
7. Andersen, J. S. et al. Nucleolar proteome dynamics. *Nature* **433**, 77–83 (2005).
8. Andersen, J. S. et al. Directed proteomic analysis of the human nucleolus. *Curr. Biol.* **12**, 1–11 (2002).
9. Frottin, F. et al. The nucleolus functions as a phase-separated protein quality control compartment. *Science* **365**, 342–347 (2019).
10. Scherl, A. et al. Functional proteomic analysis of human nucleolus. *Mol. Biol. Cell* **13**, 4100–4109 (2002).
11. Hu, R. et al. ZNF668 functions as a tumor suppressor by regulating p53 stability and function in breast cancer. *Cancer Res.* **71**, 6524–6534 (2011).
12. Guo, D. C. et al. Loss-of-function mutations in YY1AP1 lead to grange syndrome and a fibromuscular dysplasia-like vascular disease. *Am. J. Hum. Genet.* **100**, 21–30 (2017).
13. Stenstrom, L. et al. Mapping the nucleolar proteome reveals a spatiotemporal organization related to intrinsic protein disorder. *Mol. Syst. Biol.* **16**, e9469 (2020).
14. Cuyley, S. et al. Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature* **535**, 308–312 (2016).
15. Meszaros, B., Erdos, G. & Dosztanyi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **46**, W329–W337 (2018).
16. Feric, M. et al. Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 (2016).
17. Grob, A., Colleran, C. & McStay, B. Construction of synthetic nucleoli in human cells reveals how a major functional nuclear domain is formed and propagated through cell division. *Genes Dev.* **28**, 220–230 (2014).
18. Drygin, D. et al. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res.* **69**, 7653–7661 (2009).
19. Granick, D. Nucleolar necklaces in chick embryo fibroblast cells. I. Formation of necklaces by dichlororibobenzimidazole and other adenosine analogues that decrease RNA synthesis and degrade preribosomes. *J. Cell Biol.* **65**, 398–417 (1975).
20. Mnaimneh, S. et al. Exploration of essential gene functions via titratable promoter alleles. *Cell* **118**, 31–44 (2004).
21. Dez, C. et al. Npa1p, a component of very early pre-60S ribosomal particles, associates with a subset of small nucleolar RNPs required for peptidyl transferase center modification. *Mol. Cell. Biol.* **24**, 6324–6337 (2004).
22. Narla, A. & Ebert, B. L. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**, 3196–3205 (2010).
23. Tafforeau, L. et al. The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol. Cell* **51**, 539–551 (2013).
24. Zubradt, M. et al. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* **14**, 75–82 (2017).
25. Peculis, B. A. & Steitz, J. A. Sequence and structural elements critical for U8 snRNP function in *Xenopus* oocytes are evolutionarily conserved. *Genes Dev.* **8**, 2241–2255 (1994).
26. Badrock, A. P. et al. Analysis of U8 snoRNA variants in zebrafish reveals how bi-allelic variants cause leukoencephalopathy with calcifications and cysts. *Am. J. Hum. Genet.* **106**, 694–706 (2020).
27. Leone, S., Bar, D., Slabber, C. F., Dalcher, D. & Santoro, R. The RNA helicase DHX9 establishes nucleolar heterochromatin, and this activity is required for embryonic stem cell differentiation. *EMBO Rep.* **18**, 1248–1262 (2017).
28. Calo, E. et al. RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* **518**, 249–253 (2015).

29. Bond, A. T., Mangus, D. A., He, F. & Jacobson, A. Absence of Dbp2p alters both nonsense-mediated mRNA decay and rRNA processing. *Mol. Cell. Biol.* **21**, 7366–7379 (2001).
30. Martin, R., Straub, A. U., Doebele, C. & Bohnsack, M. T. DExD/H-box RNA helicases in ribosome biogenesis. *RNA Biol.* **10**, 4–18 (2013).
31. Thomson, E., Ferreira-Cerca, S. & Hurt, E. Eukaryotic ribosome biogenesis at a glance. *J. Cell Sci.* **126**, 4815–4821 (2013).
32. Farley-Barnes, K. I. et al. Diverse regulators of human ribosome biogenesis discovered by changes in nucleolar number. *Cell Rep.* **22**, 1923–1934 (2018).
33. Schmid, M. & Jensen, T. H. The nuclear RNA exosome and its cofactors. *Adv. Exp. Med. Biol.* **1203**, 113–132 (2019).
34. Wasmuth, E. V., Januszyk, K. & Lima, C. D. Structure of an Rrp6–RNA exosome complex bound to poly(A) RNA. *Nature* **511**, 435–439 (2014).
35. Dez, C., Houseley, J. & Tollervey, D. Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of *Saccharomyces cerevisiae*. *EMBO J.* **25**, 2662–2662 (2006).
36. Lafontaine, D. L. J. A. ‘Garbage can’ for ribosomes: how eukaryotes degrade their ribosomes. *Trends Biochem. Sci.* **35**, 267–277 (2010).
37. Lu, Z. et al. RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Methods

Cell culture and cell transfection

Human HeLa and HEK293FT cells were purchased from the American Type Culture Collection (ATCC; <http://www.atcc.org>) and were originally authenticated using STR profiling. Human U2OS cells were purchased from the National Collection of Authenticated Cell Cultures. Human K562 cells were a gift from the laboratory of F.-L. Meng. Human H9 cell lines were purchased from WiCell. Zebrafish ZF4 cell lines were purchased from China Zebrafish Resource Center. HeLa, HEK293FT and U2OS cells were maintained in DMEM, which was supplemented with 10% fetal bovine serum (FBS). Human K562 cells were maintained in RPMI1640, which was supplemented with 10% FBS. H9 cells were maintained in DMEM/F-12 supplemented with 20% KnockOut Serum Replacement, 0.1 mM Glutamax, 0.1 mM non-essential amino acids and 0.1 mM mercaptoethanol and 4 ng ml⁻¹ b-FGF and cultured with irradiated mouse embryonic fibroblast feeder cells with daily changed cultured medium and passaged weekly. Zebrafish ZF4 cells were maintained in DMEM/F-12 supplemented with 10% FBS. We maintained human cell lines and zebrafish cell lines at 37 °C and 28 °C respectively in a 5% CO₂ cell culture incubator and tested all cell lines routinely to exclude mycoplasma contamination. Transfection of plasmids was carried out with Lipofectamine 3000 Transfection Reagent (Invitrogen) according to the manufacturer's protocol.

Lentivirus production, infection and generation of stable cell lines

To produce lentiviral particles, HEK293FT cells (60%–70% confluence) in a 10 cm dish were co-transfected with 10 µg pHAGE or pLKO.1-TRC construct, 7.5 µg of psPAX2 and 3 µg pMD2.G. The supernatant containing viral particles was collected twice at 48 and 72 h after transfection, filtered through Millex-GP Filter Unit (0.45 µm pore size, Millipore), and stored at -80 °C till use. To infect HeLa cells with lentivirus, cells were cultured in the medium containing lentivirus and 10 µg ml⁻¹ polybrene (Sigma). To increase the efficiency, infected cells were cultured under 2 days of 500 ng ml⁻¹ puromycin. The cells were recovered at least 7 days before performing subsequent experiments.

Plasmid construction and transfection

To construct the tet-on shRNA vectors, shRNA sequences and a scrambled sequence were individually cloned into tet-on pLKO.1-TRC vector between the AgeI and EcoRI sites to obtain all shRNA constructs. Lentiviral particles were produced in HEK293FT cells and stable inducible knockdown HeLa cell lines were generated as described above. To knock down protein candidates, individual cell lines were seeded and induced with 1 µg ml⁻¹ doxycycline for 6 days.

To construct expression plasmids for eGFP-nucleolar proteins (200 proteins were tested in this way; Extended Data Fig. 2 and Supplementary Table 2), the nucleolar protein candidates were individually amplified from HeLa cDNAs, and inserted into pcGFP-C1 vector (with a CMV promoter) using HiFiClone Plus One Step Cloning Kit (Yeasen, 10911ES20) or the enzyme digestion and ligation method. Plasmids transformation was performed in T1 *Escherichia coli* (Trans1-T1 Phage Resistant Chemically Competent Cell, Transgen). The transfection was performed as described above.

Primers and shRNA sequences for plasmids construction used are listed in Supplementary Table 1.

Knock-in fluorescent protein via the CRISPR–Cas9 system

To construct the plasmids used in fluorescent protein knock-in, DNA sequences for left homology arm and right homology arm of targeted genes (DKCI, NPM1, URB1, URB2, TCOF1 and MKI67) were amplified from genomic DNA using the primer pairs listed in Supplementary Table 1. To generate the donor plasmids, we introduced silent mutations within

the Cas9 nuclease binding region of the left/right homology arms. The coding sequence of mRuby3, eGFP, mNeonGreen or mTagBFP2 was amplified by PCR using the primers listed in Supplementary Table 1. The PCR products were purified and cloned into a pCRII plasmid that was digested with BamHI/EcoRI using the one-step clone method. Design of the guide RNAs was carried out using the Benchling CRISPR tool (<http://www.benchling.com>) to minimize potential off-target effects. Oligonucleotide pairs (Supplementary Table 1) were cloned into the vector pS338. The final bicistronic vector encoded the genomic RNA and the Cas9 nuclease. To obtain the knock-in HeLa cell lines, 1 × 10⁶ cells per well were seeded in a 6-well plate with DMEM supplemented with 10% FBS at 37 °C, 5% CO₂. The following day, transfection was carried out using the bicistronic nuclease plasmid with the corresponding donor plasmid at the ratio of 2 to 1 and a total 2.5 µg plasmid were transfected as described above. One day later, puromycin (1 mg ml⁻¹) was added to the cells to increase the knock-in efficiency. Three days later, the fluorescent positive cells were sorted and collected using FACSaria (BD Biosciences).

The fluorescent protein knock-in cell lines were then observed using an epifluorescence microscope (DeltaVision Elite) and further confirmed by genotyping PCR. In brief, genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Tiangen) from a confluent well of a 12-well plate. In total, 100 ng genomic DNA was used as a template for an out-out PCR with the primers listed in Supplementary Table 1 and the PCR products were analysed by Sanger sequencing.

Colony formation assay

Three thousand scramble shRNA-treated or URB1-KD H9 cells were counted with a Countess II FL Automated Cell Counter (Therm Fisher) and seeded on irradiated mouse embryonic fibroblast feeder cells with 10 mM Y-27632 (Tocris) on the first day, then cultured for 7 days by changing the culture medium daily. Surviving colonies were stained with 0.1% crystal violet at room temperature for 2 min followed by washing with Dulbecco's PBS (DPBS) twice. Pictures were taken and the average number of colonies per well were measured. The colony formation ability was presented by the ratio of survival colonies to the total seeded cells on the first day.

RNA isolation, RT–qPCR and northern blots

Total RNAs from the equal number of cultured cells were extracted with Trizol Reagent (Invitrogen) according to the manufacturer's protocol. For RT–qPCR, the cDNA synthesis was carried out using PrimeScript RT reagent Kit (TaKaRa) with oligo(dT) or random hexamers. Quantitative PCR was performed using SYBR Green Realtime PCR Master Mix (TOYOBO) and a StepOnePlus real-time PCR system (Applied Biosystems). Primer sequences for RT–qPCR used are listed in Supplementary Table 1.

To examine pre-rRNAs, northern blots were carried out according to the manufacturer's protocol (DIG Northern Starter Kit, Roche). RNA was loaded on agarose gels and the Dig-labelled antisense probes of 5' ETS-1, ITS1, ITS2, 5.8S, 28S-3' ETS and 3' ETS as previously reported²³ were used. Probe sequences for northern blots are listed in Supplementary Table 1.

Western blotting

Cells were collected after treatments and resuspended in lysis buffer (1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 50 mM Tris and 1× protease inhibitor cocktail, pH 8.0) for 10 min. After centrifugation, supernatants containing soluble proteins were resolved on polyacrylamide gel with 10% SDS and analysed by western blot with anti-Flag (Sigma, 1:1,000 dilution), anti-URB1 (Proteintech, 1:500 dilution), anti-EXOSC8 (Proteintech, 1:500 dilution), anti-MTR4 (Proteintech, 1:500 dilution), anti-RPL23 (Proteintech, 1:500 dilution), anti-RPL10A (Abcam, 1:500 dilution) or anti-ACTB (Sigma, 1:5,000 dilution) antibodies.

Cell proliferation assay

To validate the effect of URB1 on cell proliferation, the tet-on URB1 shRNA stably expressed cell lines were induced with 1 $\mu\text{g ml}^{-1}$ doxycycline for 5 days. Cell proliferation was detected by the MTT assay that measures OD490 according to CellTiter 96 AQueous One Solution Cell Proliferation Assay (Promega) and Envision Multimode Plate Reader (PerkinElmer) that measures the occupied surface area of cells to indicate confluence. Five thousand HeLa cells were seeded into each well of 96-well dishes, then cultured at 37 °C in a 5% CO₂ cell culture incubator for indicated time points. Cell proliferation at day 0 was detected after cells were seeded for 5 h to allow their attachment to the plate. Live cells at day 1, day 2 and day 3 were detected by MTT and Envision Multimode Plate Reader at corresponding time points. Data were normalized to the value at day 0.

Ribosome fractionation

About 2 $\times 10^7$ scramble shRNA-treated and URB1-KD HeLa or HEK293FT cells were firstly treated with cycloheximide (100 $\mu\text{g ml}^{-1}$) for 15 min and then lysed with 500 μl lysis buffer (pH 7.4 20 mM Tris, 15 mM MgCl₂, 200 mM KCl, 1 mM dithiothreitol, 100 $\mu\text{g ml}^{-1}$ cycloheximide, 1 mg ml^{-1} heparin, 40 U ml^{-1} RNasin (Promega) and 1% Triton X-100). Supernatants were collected at 4 °C by centrifugation with 12,000g for 10 min, and then loaded onto 5%–50% sucrose gradients for fractionation by ultracentrifugation. Linear sucrose gradients were prepared with a Gradient Master (Biocomp) according to the manufacturer's suggestion, and the ultracentrifugation was performed with a SW41 rotor (Beckman) at 4 °C with 36,000g for 2.5 h. Ribosome distribution on the sucrose gradients was recorded at 254 nm by BIOCOMP Piston Gradient Fractinator equipped with Bio-Rad ECONO UV Monitor.

Protein visualization

To detect protein localization by immunofluorescence in fixed cells, cells were seeded on High Performance No. 1.5 18 \times 18 mm glass coverslips and were fixed with 4% PFA for 15 min, followed by permeabilization with 0.5% Triton X-100 for 5 min. Then cells were blocked with 1% BSA for 1 h at room temperature. Primary antibodies were diluted with 1% BSA (anti-GFP 1:100, anti-FBL 1:400, anti-EXOSC8 1:50, anti-EXOSC9 1:50, anti-MTR4 1:50, anti-NPM1 1:50) and incubated for 1 h at room temperature. After washing with 1× DPBS 3 times, fluorescent secondary antibodies (goat anti-mouse secondary antibody–Alexa Fluor 555 and goat anti-rabbit secondary antibody–Alexa Fluor 488) were diluted 1:1,000 in 1% BSA and incubated for 1 h at room temperature. Samples were mounted in Vectashield antifade mounting medium (Vector Lab).

For live-cell imaging, cells were seeded on 35 mm No. 1.5 glass bottom dishes (Labtide, M03-0601) one day prior to imaging. Cells were washed once with PBS and the medium was replaced by FluoroBrite DMEM (GIBCO) supplemented with 10% FBS and placed back in the incubator for 1 h. All images were obtained at 37 °C with 5% CO₂ condition.

To detect mNeonGreen or mRuby3-tagged proteins in fixed cells, cells were seeded on High Performance No. 1.5 18 \times 18 mm glass coverslips and were fixed with 3% PFA/0.1% glutaraldehyde in DPBS for 10 min at room temperature. Then we performed blocking and smFISH (single-molecule RNA fluorescent in situ hybridization) as described below. Samples were mounted with Vectashield antifade mounting medium (Vector Lab) or SlowFade Diamond (Invitrogen).

Distribution analysis of POIs

Each POI expression plasmid was transfected and co-imaged with mRuby3–DKC1 and mTagBFP2–NPM1 by epifluorescence microscope (DeltaVision Elite). The 3D stacks were imported into Fiji/ImageJ and analysed by an in-house ImageJ script. In brief, (1) the reconstructed images were split into single channels; (2) the POI, DKC1 and NPM1 were marked using optimum threshold, respectively; (3) the masks of the POI, DKC1 and NPM1 were generated from step 2 and the background

outside the masks was set to 0; (4) for the colocalization analysis of POI signal, DKC1 and NPM1 was calculated by comparing the corresponding masks modulated in step 3 using MOC³⁸ as the percentage of above-background pixels for POIs that overlap above-background pixels for DKC1 ($M_{\text{nu}/\text{DKC1}}$), the percentage of above-background pixels for DKC1 that overlap above-background pixels for POIs ($M_{\text{DKC1}/\text{nu}}$), the percentage of above-background pixels for POIs that overlap above-background pixels for NPM1 ($M_{\text{nu}/\text{NPM1}}$), the percentage of above-background pixels for NPM1 that overlap above-background pixels for POIs ($M_{\text{NPM1}/\text{nu}}$); (5) based on the distinct radio flux structure of POIs of proteins obtained in an inner localization had a smaller volume compared to those in outer localization, the localization of POIs could be simplified as Score_{DKC} = $M_{\text{nu}/\text{DKC1}} - M_{\text{DKC1}/\text{nu}}$ and Score_{NPM1} = $M_{\text{nu}/\text{NPM1}} - M_{\text{NPM1}/\text{nu}}$.

smFISH

All smFISH probes were designed via Stellaris Probe Designer and labelled with Cy3 on the 3' ends (Supplementary Table 1). In brief, cells were fixed with 4% PFA for 15 min, followed by permeabilization with 0.5% Triton X-100 for 5 min. Cells were incubated in 10% formamide/2× SSC for 10 min at room temperature followed by hybridization at 37 °C for 16 h. After hybridization, the cells were blocked and incubated with antibodies as described above to visualize proteins. Samples were mounted in Vectashield antifade mounting medium (Vector Lab).

Wide-field microscopy

All wide-field microscopy was performed on a DeltaVision Elite imaging system equipped with a 60×1.42 NA Plan Apo oil-immersion objective, or a 100×1.40 NA Plan Apo oil-immersion objective (Olympus), as well as the CoolSnap HQ2 camera (Photometrics) equipped with the live-cell imaging environment control system (Live Cell Instrument). Raw data of all presented figures were deconvoluted by softWoRx 7.0 using the enhanced ratio method.

SIM procedure

All SIM experiments were performed on a DeltaVision OMX V4 system (GE Healthcare) equipped with a 60×1.42 NA Plan Apo oil-immersion objective (Olympus) and six laser beams (405, 445, 488, 514, 568 and 642 nm; 100 mW) or a DeltaVision OMX SR system (GE Healthcare) equipped with a 60×1.42 NA Plan Apo oil-immersion objective (Olympus) and four laser beams (405, 488, 568 and 642 nm; 100 mW). The microscope was routinely calibrated with a special image registration slide and algorithm provided by GE healthcare. To obtain optimal images, immersion oil with refractive indices of 1.516 was used at 25 room temperature and 1.520 for 37 °C. SIM image stacks were captured with a z-distance of 0.125 μm and with 5 phases, 3 angles, 15 raw images per plane. The raw data were reconstructed with channel specific OTFs and a Wiener filter was set to optimum value by using softWoRx 6.5 package (GE Healthcare). Images were registered with alignment parameters obtained from calibration measurements with 100 nm diameter TetraSpeck Microspheres with four colours (Molecular Probes).

STED procedure

STED was performed using the Inspector Software (16.3.14287-w2129) (Abberior) and the Abberior STED (Infinity Line) with a 775 nm STED-Laser from Abberior Instruments. In brief, the immunolabelled samples with the fluorophore Alexa Fluor 555 was excited at 568 nm, and Abberior STAR RED was excited at 640 nm. STED was performed with a 775 nm depletion laser. Images were recorded in the 2D STED mode. For visualization of URB1, URB2, DDX21 and FBL, eGFP–URB1 or eGFP–URB2 knock-in HeLa cells and wild-type HeLa cells were seeded and fixed as described above. The primary antibody targeting GFP, DDX21 and FBL were diluted at 1:100 or 1:400. The fluorescent (Alexa Fluor 555, Invitrogen; Abberior STAR RED, Abberior) secondary antibodies were diluted at 1:200 or 1:1,000 with 1% BSA. Samples were mounted with SlowFade

Diamond (Invitrogen). A UPlanSApo 100× oil-immersion objective was used (Olympus Corporation). All STED images were deconvolved using Huygens software (Scientific Volume Imaging).

Spinning disk confocal superresolution microscopy

Spinning disk confocal superresolution microscope images were acquired with the Olympus IXplore SpinSR microscope equipped with Yokogawa CSU-W1 SoRa (50 μm and SoRa disks) and operated with cellSens Dimension software. For visualization of URB1, URB2, DDX21 and FBL, HeLa eGFP–URB1 or eGFP–URB2 knock-in cells and HeLa wild-type cells were seeded and fixed as described above. The primary antibody targeting GFP, DDX21 and FBL were diluted at 1:100 or 1:400. The fluorescent (Alexa Fluor 555) secondary antibodies were diluted at 1:200 or 1:1,000 with 1% BSA. Samples were mounted with SlowFade Diamond (Invitrogen).

CLEM procedure

Live-cell dishes with photo-etched gridded coverslips were used to culture mNeoGreen–URB1-knock-in HeLa cells. Cells were fixed with 4% PFA first and then a DeltaVision Elite imaging system was used to collect 1–2 difference interference contrast and wide-field images to document the arrangement of the cells at different magnifications. Then, the cells were fixed with 2.5% glutaraldehyde for 2 h at room temperature. Samples were then dehydrated with a graded ethanol series (50%, 70%, 80%, 90% and 100%) for 2 min each. Samples were infiltrated and embedded in PON812 resin. After polymerizing for 48 h at 60 °C, 70-nm-thick ultrathin sections were cut using a diamond knife, and then picked up with slot grids. The sections were double stained with uranyl acetate and lead citrate. After air drying, samples were examined with a transmission electron microscope H-7800 at an acceleration voltage of 80 kV.

FRAP assay

Cells were cultured on 35 mm no.1.5 glass-bottomed dishes (Cellvis). FRAP assays were performed on a DeltaVision OMX SR system (GE Healthcare) equipped with a 60 ×/1.42 NA Plan Apo oil-immersion objective (Olympus) and four laser beams (405, 488, 568 and 642 nm; 100 mW). The region of interest was photobleached and the recovery of fluorescence intensity within the region of interest was obtained for each experiment. Photokinetic Data Analysis was used to analyse FRAP data.

iCLIP-seq

iCLIP-seq and data analyses were performed as described³⁹ with slight modifications. HeLa cells expressing Flag-tagged URB1 were crosslinked under UV irradiation, resuspended in lysis buffer and sonicated 6 times for 20 s each on 25% duty. Supernatants were incubated with anti-Flag antibodies (Shanghai Genomics) or control IgG-conjugated beads for 2 h at 4 °C. Beads with protein–RNA complexes were treated with 5 Kunz U ml⁻¹ of micrococcal nuclease in the presence of 5 μg yeast tRNAs for 5 min at 37 °C, followed by dephosphorylation of RNA 3' ends with PNK. Then pre-adenylated linker L3 (3' modification with biotin) was linked to 3' end of RNA using T4 RNA ligase 1 at 16 °C overnight. The 5' end of RNA was labelled with [γ -³²P]ATP using PNK. Protein–RNA complexes were eluted with 1× Nupage loading buffer (Invitrogen) at 70 °C for 10 min, resolved on 4%–12% NuPAGE Bis-Tris gel (Invitrogen), transferred to a nitrocellulose membrane and exposed to a Fuji film at -80 °C. Bands corresponding to specific protein–RNA complexes were excised and then digested with proteinase K. RNAs were purified and cDNAs were made using Superscript III (Invitrogen) with Rclip primers. cDNAs were separated on 6% denaturing urea polyacrylamide gel, then three bands at 120–200 nt (high), 85–120 nt (medium) and 70–85 nt (low) were excised, purified and circularized by Circligase II (Epicenter). Single-stranded cDNAs were hybridized with Cut-Oligo and then digested by BamHI. Linearized cDNAs were PCR-amplified and

purified. The obtained iCLIP libraries were subjected to high throughput deep sequencing.

In vitro RNA transcription and purification

Linear RNA was in vitro transcribed from T7 expression vector prepared by RiboMax large RNA production system (Promega) according to the manufacturer's protocol with slight modifications. In brief, 1 μg PCR-amplified T7-DNA fragments were incubated with 2 μl T7 RNA polymerase enzyme and 0.5 mM Biotin-labelled dNTPs (Roache). In vitro transcription was carried out for 3 h at 37 °C, followed by DNase I treatment for 30 min at 37 °C to remove DNA templates. Transcribed RNA was precipitated with ethanol and washed with 75% ethanol and resuspended in RNase-free water. For purification, the transcribed RNA was resolved on denaturing urea polyacrylamide gel and visualized by ethidium bromide staining. The corresponding bands on denaturing urea polyacrylamide gel were excised for further purification.

28S-3' ETS RNA–U8 snoRNA in vitro binding assay

Purified biotin-labelled 28S-3' ETS RNA fragments (4 pmol) and Dig-labelled U8 snoRNAs (4 pmol) were denatured for 5 min at 65 °C in RNA structure buffer (10 mM HEPES pH 7.0, 10 mM MgCl₂) and slowly cooled down to room temperature and then incubated together with the cell lysates collected from scramble shRNA-treated or URB1-KD HeLa cells, respectively. Of note, cells were collected 4 days after URB1 knockdown, when cells were still alive upon URB1 depletion. All components were incubated with rotation at room temperature for 30 min. U8 snoRNAs bound to 28S-3' ETS RNA were purified by incubating with 50 μl Streptavidin beads at room temperature for 20 min. Beads were washed twice with 500 μl washing buffer (100 mM Tris pH 7.4, 10 mM EDTA, 1 M NaCl, 0.1% Tween 20), extracted with Trizol Reagent (Invitrogen) according to the manufacturer's protocol, and proceeded for anti-Dig northern blot detection.

Nascent pre-rRNA metabolic chasing and fluorescence labelling assay

HeLa cells were cultured in the medium supplemented with 1 mM 5-EU (Click Chemistry Tools 1261) for 10 min. Then the medium was replaced with the fresh 5-EU-free medium. Next, cells were fixed at indicated time points with 4% PFA for 15 min, followed by permeabilization with 0.5% Triton X-100 for 15 minutes. The click reaction for nascent RNA labelling was carried out as described in the manufacturer's protocol (Thermo Fisher C10642) with the following modifications: (1) the free copper (component C) and copper protectant (component D) level in the reaction mix was medium (1:1); (2) the Alexa Fluor picolyl azide concentration was 1 μM; (3) cells were incubated with the reaction cocktail for 30 min and were washed 3 times with DPBS. Finally, samples were mounted in Vectashield antifade mounting medium (Vector Lab H-1900) for imaging.

The rate of pre-rRNA radial flux movement was measured by calculating the diameter of nascent pre-rRNA fluorescent signals in each nucleolus at different time points.

SILAC followed by mass spectrometry sequencing

To differentially label HeLa cells with light and heavy amino acids, we replaced the following components in the culture medium: DMEM formulated without lysine and arginine instead of DMEM, dialysed serum instead of regular FBS, and lysine and arginine added separately to light (regular L-lysine and L-arginine) and heavy ([¹³C₆, ¹⁵N₂]L-lysine and [¹³C₆, ¹⁵N₄]L-arginine) media. All SILAC reagents were purchased from Thermo Fisher Scientific. To quantitatively identify changes in protein levels upon URB1 knockdown, tet-on shRNA plasmids were constructed and the Tet-inducible stable knockdown cell lines were generated as described above. Individual cell lines were seeded and induced with 1 μg ml⁻¹ doxycycline for 6 days. After washing with ice-cold PBS, the control and URB1-KD cells were collected and counted. An equivalent

Article

number of SILAC-labelled cells (4×10^6) was lysed and sonicated in RIPA buffer (20 mM Tris-HCl, pH 7.5; 150 mM NaCl; 1 mM EDTA; 1 mM EGTA; 1% NP-40; 1% sodium deoxycholate; and protease inhibitor cocktail) on ice. After centrifugation at 14,000 rpm for 10 min at 4 °C, proteins of equal number of cells were combined and the concentrations were estimated using BCA. The proteins were boiled with SDS loading buffer and subjected to SDS-PAGE followed by colloidal blue staining. Gel fragments were excised, washed, and sent for liquid chromatography-mass spectrometry.

Zebrafish embryology

Injection of morpholinos was performed on one-cell-stage zebrafish embryos at the indicated concentrations (see Supplementary Table 1 for morpholino sequences) (Gene Tools). Injected embryos were incubated at 28 °C. For cartilage staining with alcian blue, embryos were collected 3 or 4 dpf. Alcian blue staining was performed as described⁴⁰.

Single guide RNAs (sgRNAs) targeting to *urb1* were designed using Benchling (<https://www.benchling.com>) and synthesized as previously described⁴¹. (See Supplementary Table 1 for the sgRNA sequences). sgRNAs were combined with Cas9 protein (5 µg) and co-injected (-1 nl) into the one-cell stage wild-type embryos. All injected embryos were raised to adulthood, crossed to wild-type zebrafishes and screened for passing the mutant allele to the next generation. Founder fish with desirable mutations were selected.

EdU-incorporation assay

For the EdU-incorporation assay, the BeyoClick EdU Cell Proliferation Kit (Beyotime, C0071S) was used. In brief, zebrafish transgenic embryos aged 4 dpf were intravenously injected with EdU (10 mM) and incubated in embryo buffer for 3 h. After fixation in 4% PFA for 4 h at room temperature, embryos were dehydrated with methanol and kept at -20 °C for over 2 h. After rehydration, permeabilization and re-fixation, the embryos were incubated with 488-dye for 1.5 h at room temperature, according to the standard protocol. Subsequently, embryos were rinsed two times with PBST and photographed in 1% agarose gel using confocal microscopy (Olympus SpinSR, 20X).

WISH

WISH was performed as previously described⁴². In brief, sense control and antisense probes were used to detect *urb1* expression (see Supplementary Table 1 for probe sequences). The embryos were observed and photographed with stereomicroscope (Nikon SMZ18).

Whole-mount smFISH in zebrafish larvae

Whole-mount smFISH for zebrafish larvae was developed by adapting smFISH for cultured cells and standard WISH for zebrafish embryos⁴³. In brief, larvae were fixed in 4% PFA at room temperature for 4 h. Larvae were then briefly washed with PBST once, dehydrated with methanol and kept at -20 °C for over 2 h. After rehydration, permeabilization and re-fixation, the larvae were incubated in prehybridization buffer (10% formamide, 2× SSC, 0.1% Triton X-100) at 30 °C for 5 min. The probe stock solution (25 µM) was diluted to 1:10 (Supplementary Table 1). The larvae were soaked in 100 µl hybridization mix in 1.5 ml microcentrifuge tube and incubated at 30 °C overnight in dark. On the following day, larvae were washed twice with prehybridization buffer at 30 °C for 30 min each and then briefly washed with PBST. Larvae were mounted with 1% agarose with or without DAPI (Thermo Scientific) and photographed using confocal microscopy (Olympus SpinSR, 20X, 60X).

Mouse embryology

Eight-week-old female B6D2F1 mice (C57BL/6 × DBA/2) were superovulated and mated with male B6D2F1 mice. Fertilized embryos were then collected from oviducts after 21–23 h. By using a FemtoJet microinjector (Eppendorf) with constant flow settings, pre-mixed Cas9 mRNAs (100 ng µl⁻¹) and guide RNA targeting *Urb1* (100 ng µl⁻¹) were injected

into the cytoplasm of fertilized embryos in a droplet of HEPES-CZB medium containing 5 µg µl⁻¹ cytochalasin B (Sigma). After injection, fertilized embryos were cultured in KSOM medium with amino acids until they reached 2-cell stage and were transferred into the oviducts of E0.5jm mice (pseudopregnant mice mated with vasectomized male ICR mice one day before).

In-cell SHAPE-MaP by NAI

In-cell SHAPE-MaP was performed in HeLa cells as previously described³⁹ with modifications. In brief, the Tet-inducible stable URB1-KD HeLa cell lines were seeded and induced with 1 µg ml⁻¹ doxycycline on 6 cm dishes for 6 days. After washed with DPBS, cells were incubated with 270 µl of cell culture medium and 30 µl of 10× SHAPE Chemical in DMSO with the final concentration of NAI (EMD Millipore) at 100 mM for 10 min at 37 °C. After reaction medium was removed, RNAs were isolated with 1 ml Trizol reagent (Invitrogen) according to the manufacturer's protocol. The same procedure was also performed in parallel for the untreated control samples, but with the addition of only DMSO. In the denaturing control reaction, RNAs were suspended in a denaturing buffer containing 50% formamide and were incubated at 95 °C before modification with SHAPE reagents.

Isolated RNAs were treated with DNase I (Ambion, DNA-freeTM kit) to remove possible DNA contamination. About 50–100 ng of RNAs were obtained under each treatment and were then used for SHAPE-MaP reverse transcription by adding 1 mL (200 U ml⁻¹) of SuperScript II (Invitrogen), 6 mM Mn²⁺ and gene-specific primers for 3'ETS. Second-strand synthesis was performed with Taq DNA polymerase. The resulting PCR products were further isolated. Primers for SHAPE-MaP reverse transcription and second-strand synthesis (1st round PCR and nested PCR reactions) were listed in Supplementary Table 1.

SHAPE-MaP libraries were prepared from 1 ng of DNAs, and size-selected with AmpureXP beads (Agencourt) with a 1:1 (bead to sample) ratio to obtain library DNA products spanning 100–400 bp in length. Final libraries were quantified using Agilent Bioanalyzer 2100 and QuBit high-sensitivity dsDNA assay. The 150-nt paired-end deep sequencing was performed by Illumina NextSeq 500 at Shanghai Institute Nutrition and Health, Big Data Center Omics Core. About 15–25 million mapped sequencing reads were obtained for each sample, with 88% of bases at or above Q30.

DMS-MaPseq

In brief, the Tet-inducible stable URB1-KD HeLa cell lines were seeded and induced with 1 µg ml⁻¹ doxycycline on 6 cm dishes for 6 days. After washed with DPBS, cells in 6 cm dishes with 1 ml cell culture medium were treated with the addition of 15 µl DMS and incubation at 37 °C for 5 min^{24,44}. Cell culture media with DMS was slowly decanted. Then cells were washed twice in 30% v/v 2-mercaptoethanol (diluted in DPBS) and collected into a 15 ml centrifuge tube using a scraper. Cells were centrifuged at 1,000g at 4 °C for 3 min and the BME solution. Cells were washed with 1 ml DPBS twice and centrifuged again.

Isolated RNAs were treated with DNase I (Ambion, DNA-free kit) to remove possible DNA contamination. About 50–100 ng of RNAs were obtained under each treatment and were then used for reverse transcription by adding 1 mL (200 U ml⁻¹) of SuperScript II (Invitrogen), 6 mM Mn²⁺ and gene-specific primers for 3'ETS. Second-strand synthesis was performed with TAQ DNA polymerase. The resulting PCR products were further isolated. Primers for reverse transcription and second-strand synthesis (1st round PCR and nested PCR reactions) were listed in Supplementary Table 1.

Sequencing libraries were prepared from 1 ng of DNAs, and size-selected with AmpureXP beads (Agencourt) with a 1:1 (bead to sample) ratio to obtain library DNA products spanning 100–400 bp in length. Final libraries were quantified using Agilent Bioanalyzer 2100 and QuBit high-sensitivity dsDNA assay. The 150-nt paired-end deep sequencing was performed by Illumina NextSeq 500 at Shanghai Institute Nutrition

and Health, Big Data Center Omics Core, Shanghai, China. About 15–25 million mapped sequencing reads were obtained for each sample, with 88% of bases at or above Q30.

SHAPE reactivity calculation and RNA secondary structure modelling

Low quality reads from SHAPE-MaP and DMS-MaP sequencing data were trimmed by Trimmomatic (v.038), then clean data were analysed by ShapeMapper (V2.1.3) to obtain the ‘reactivity value’ (named ‘shape-score’ for SHAPE-MaP and ‘DMS-score’ for DMS-MaP) in SHAPE profiles. For SHAPE-MaP, the NAI-modified, untreated and denatured data were used for mutations calling and reactivity value calculation in the ShapeMapper, and the DMS-modified data were used in DMS-MaP. In the software, reactivity value was only computed for nucleotides with sequencing depths above 1,000 in all the treatments by set parameter “--min-depth 1000”, and further normalized by a model-free box-plot approach⁴⁵. Then, the SHAPE profiles were used to model RNA secondary structures by RNAfold with parameters (-p-d2 --shapeMethod=D). Finally, the structures were redrawn by StructureEditor.

Delta shape-score calculation between the scramble shRNA-treated and shRNA-knockdown samples

The changes of RNA SHAPE value were calculated by URB1-KD samples minus scramble samples using deltaSHAPE (v1.0) with the parameters (--all--colorfill--noshow--pdf-p5-z0-s0). Positive changed regions were labelled in green and negative changed regions were labelled in purple in Fig. 4h.

Native RNA–protein complex immunoprecipitation

Scramble and URB1-KD HeLa cells (1×10^7) were collected, resuspended in 1 ml lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.05% Igepal, 0.5% NP-40, 0.5 mM PMSF, 2 mM ribonucleoside vanadyl complex (RVC), protease inhibitor cocktail (MedChemExpress)) followed by 4 \times 10 s sonication with an interval of 1 min on ice. After centrifuging at 13,000 rpm for 10 min at 4 °C, the supernatant was pre-cleared with 20 μ l Dynabeads Protein G (Invitrogen). The pre-cleared lysates were further incubated with 2 μ g anti-MTR4 (Proteintech) or anti-Flag (Sigma) antibody for 2 h at 4 °C. Then 20 μ l Dynabeads Protein G beads (blocked with 1% BSA and 20 μ g ml⁻¹ yeast tRNA) were added to the mixture and incubated for another 1 h at 4 °C followed by washing with wash buffer (50 mM Tris pH 7.4, 300 mM NaCl, 0.05% sodium deoxycholate, 0.5% NP-40, 0.5 mM PMSF, 2 mM RVC, protease inhibitor cocktail (MedChemExpress)). The RNA–protein complex was eluted with elution buffer (100 mM Tris pH 6.8, 4% SDS) at room temperature for 15 min. For RT-qPCR, each RNA sample was treated with DNase I (Ambion, DNA-freeTM kit) and then reverse transcription was performed with SuperScript III Reverse Transcriptase (Invitrogen) with random hexamers followed by RT-qPCR analysis. Primers were listed in Supplementary Table 1.

3' RACE

3' RACE was performed using the 3'-RACE Kit (Sangon, B605101-0010). RNA was extracted from scramble shRNA-treated and URB1-targeted shRNA-treated HeLa cells, and RACE was performed according to the standard manufacturer's protocol.

Animal use and care

All animal procedures were performed under the ethical guidelines of Center for Excellence in Molecular Cell Science, CAS. Animal ages are stated in the text and figure legends.

Statistics and reproducibility

Histogram and line chart were plotted with GraphPad Prism 9. Microscopy imaging, northern blotting and western blotting were repeated

independently at least three times with similar results. For the statistical significance and sample size of all graphs, please see figure legends and methods for details.

Materials availability

Plasmids and cell lines generated in this research are available from the Chen laboratory under a material transfer agreement with the Center for Excellence in Molecular Cell Science, CAS. Requests should be submitted by email to L.-L.C.

Reporting summary

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

Data availability

All data supporting the findings of this study are available in the manuscript, Source Data, and at <https://data.mendeley.com/datasets/8fmmx5vpkt/draft?a=c1a207f4-eda5-48be-99a9-92863496c868>. SHAPE-MaP data with NAI and DMS, and iCLIP-seq datasets of Flag-URB1 are available at the Gene Expression Omnibus under accessions GSE194413 and GSE196625. Source data are provided with this paper.

Code availability

Custom code used in SHAPE-MaP and iCLIP-seq analysis and the ImageJ script used in image analysis in this study are available from <https://github.com/YangLab/Nucleolar-URB1-ensures-3-ETS-rRNA-removal-to-prevent-exosome--surveillance>.

38. Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. Cell Physiol.* **300**, C723–C742 (2011).
39. Liu, C. X. et al. Structure and degradation of circular RNAs regulate PKR activation in innate immunity. *Cell* **177**, 865–880.e821 (2019).
40. Walker, M. B. & Kimmel, C. B. A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech. Histochem.* **82**, 23–28 (2007).
41. Shen, B. et al. Efficient genome modification by CRISPR–Cas9 nickase with minimal off-target effects. *Nat. Methods* **11**, 399–402 (2014).
42. Thissen, C. & Thissen, B. High-resolution *in situ* hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69 (2008).
43. Oka, Y. & Sato, T. N. Whole-mount single molecule FISH method for zebrafish embryo. *Sci. Rep.* **5**, 8571 (2015).
44. Cordero, P., Kladwang, W., VanLang, C. C. & Das, R. Quantitative dimethyl sulfate mapping for automated RNA secondary structure inference. *Biochemistry* **51**, 7037–7039 (2012).
45. Deigan, K. E., Li, T. W., Mathews, D. H. & Weeks, K. M. Accurate SHAPE-directed RNA structure determination. *Proc. Natl Acad. Sci. USA* **106**, 97–102 (2009).

Acknowledgements We thank members of the Chen laboratory for discussions; J. Zhang, B.-Y. Zou and J.-G. Li for supporting our experiments; and the core facilities of Zhejiang University Medical Center and Liangzhu Laboratory, and Optofem Technology for technical support with STED. This study was supported by the CAS Project for Young Scientists in Basic Research (YSBR-009), the National Key R&D Program of China (2021YFA1100203), the Shanghai Municipal Commission for Science and Technology (20JC1410300), the National Natural Science Foundation of China (NSFC) (31830108, 31821004, and 31725009), the Center for Excellence in Molecular Cell Science (CEMCS) (2020DF03), the HHMI International Program (55008728) and the Xplorer Prize to L.-L.C.

Author contributions L.-L.C. supervised and conceived the project. L.-L.C., L.S. and G.X. designed experiments. Y.L. performed CLEM, supervised by L. Yu. L.Z. generated *Urb1*^{-/-} mice, supervised by J.L. S.-X.G. performed SILAC, supervised by C.C.L.W. P.-H.Z. performed computational analyses, supervised by L. Yang. L.S., G.X., Y.H., R.-W.Y., P.-F.L., Y.-H.P., X.G., S.L., L.-Z.Y., S.-M.C., Z.-H.Y. and Y.W. performed all other experiments and analyses, supervised by L.-L.C. L.S., G.X. and L.-L.C. drafted the manuscript. L.-L.C. edited the manuscript.

Competing interests The authors declare no competing interests.

Additional information

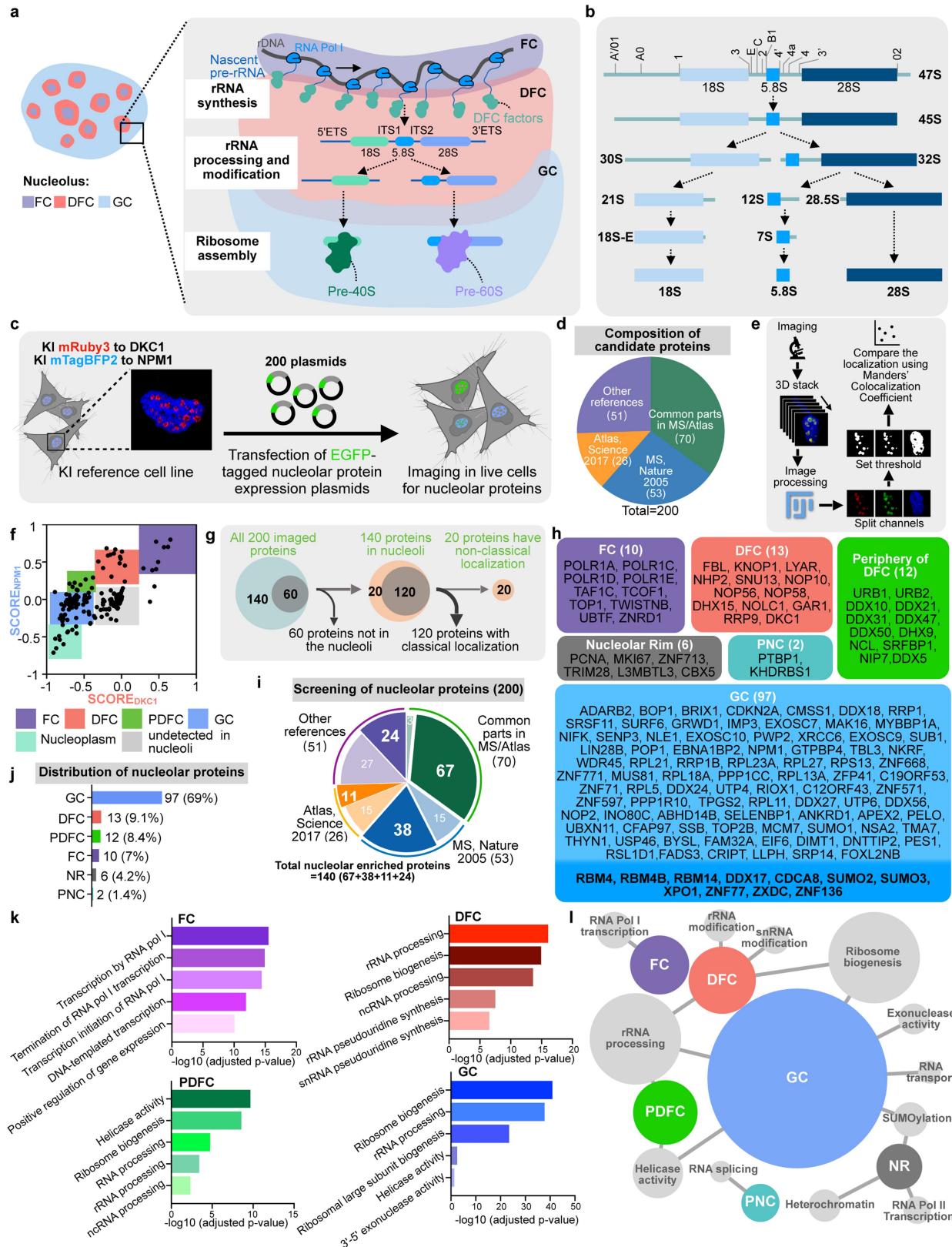
Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-023-05767-5>.

Correspondence and requests for materials should be addressed to Ling-Ling Chen.

Peer review information *Nature* thanks Jean-Denis Beaudoin, Prasanth Kannanganattu and Denis Lafontaine for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Article



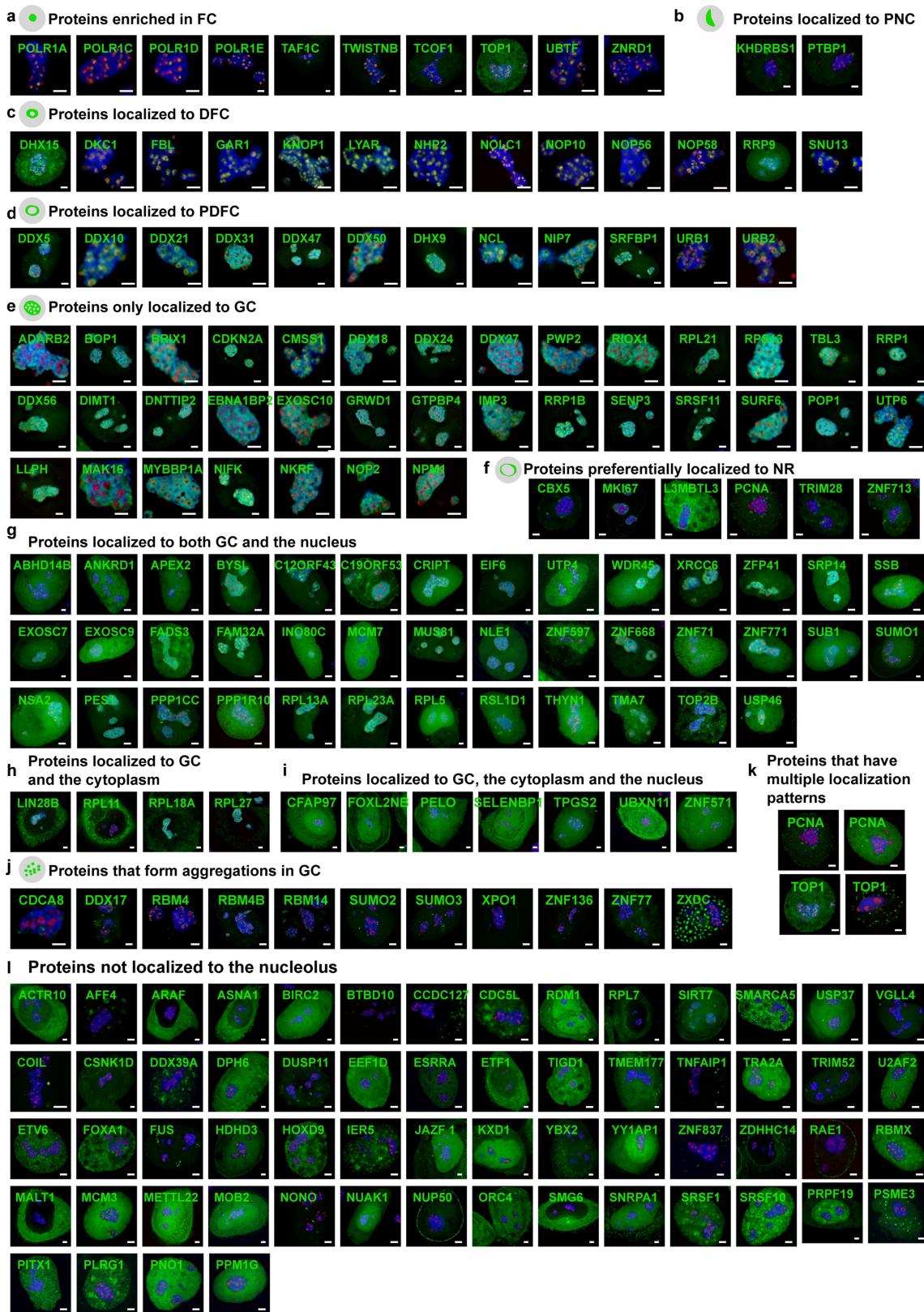
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | The schematic of pre-rRNA processing and experiment procedure of image-based nucleolar protein identification.

a. Coordination of the sub-nucleolar compartment with its function in ribosome biogenesis. b. The schematic of pre-rRNA processing in human cells. c. The workflow of live-cell imaging-based mapping of nucleolar proteins. d. Composition of nucleolar protein candidates in this paper^{2,7}. e. Workflow of acquiring, processing and analyzing image data by Fiji/ImageJ. f. Cluster distribution of candidate nucleolar proteins. SCORE_{NPML} and SCORE_{DKCI} (see Methods for details) are shown to distinguish different nucleolar sub-regions.

g. Distribution analysis procedure of candidate nucleolar proteins. h. The list of 140 proteins in different nucleolar sub-regions. i. Composition of 140 nucleolar enriched proteins from different reports^{2,7,10}. j. Proportion of different nucleolar sub-regions (FC, DFC, PDFC, NR, GC and PNC) of the 140 nucleolar enriched proteins. k. Biological process GO-enrichment of nucleolar proteins enriched in FC, DFC, PDFC and GC. P-values are calculated by two-sided Mann-Whitney tests. l. Biological process enrichments in FC, DFC, PDFC, NR, GC and PNC. The circle areas are scaled by the number of proteins in each category.

Article

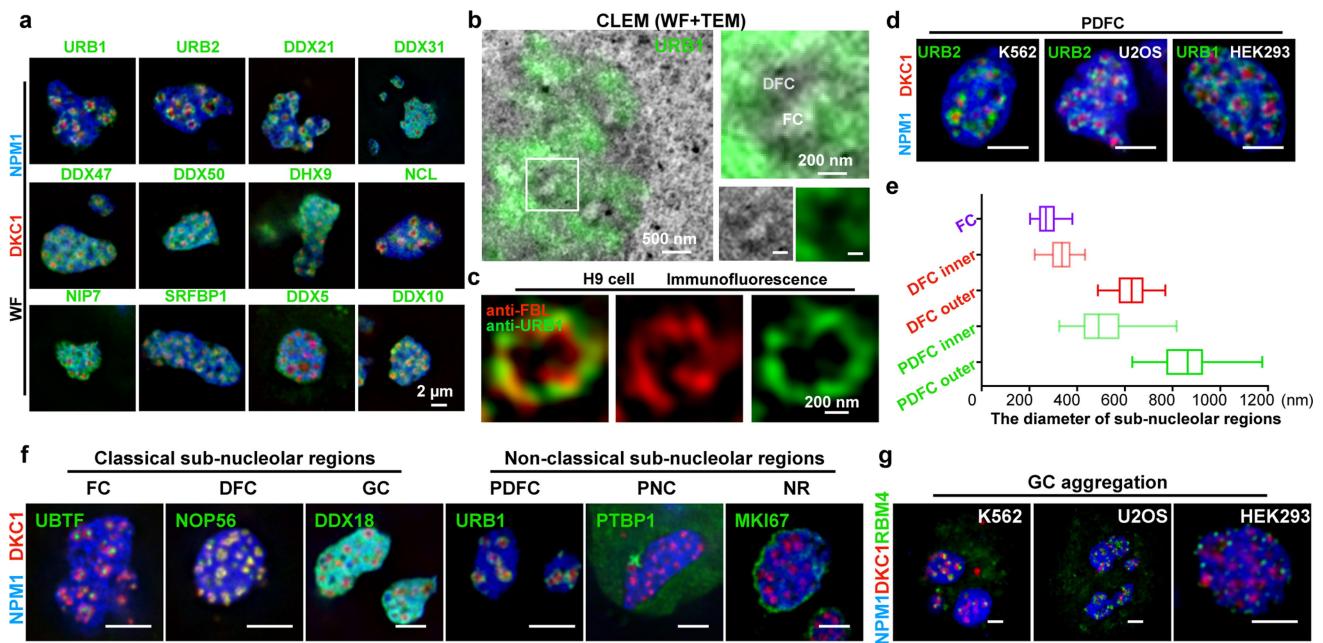


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Gallery of nucleolar protein candidates. Representative images of all screened 200 candidate proteins by exogenous expression of EGFP tagged candidate protein in the dual-color reference knock-in cells shown in Extended Data Fig. 1c. Data was collected from n = 4 cells. Scale bar 2 μm.
a. Representative images of nucleolar proteins enriched in FC. b. Representative images of nucleolar proteins enriched in PNC. c. Representative images of nucleolar proteins enriched in DFC. d. Representative images of nucleolar proteins enriched in PDFC. e. Representative images of nucleolar proteins

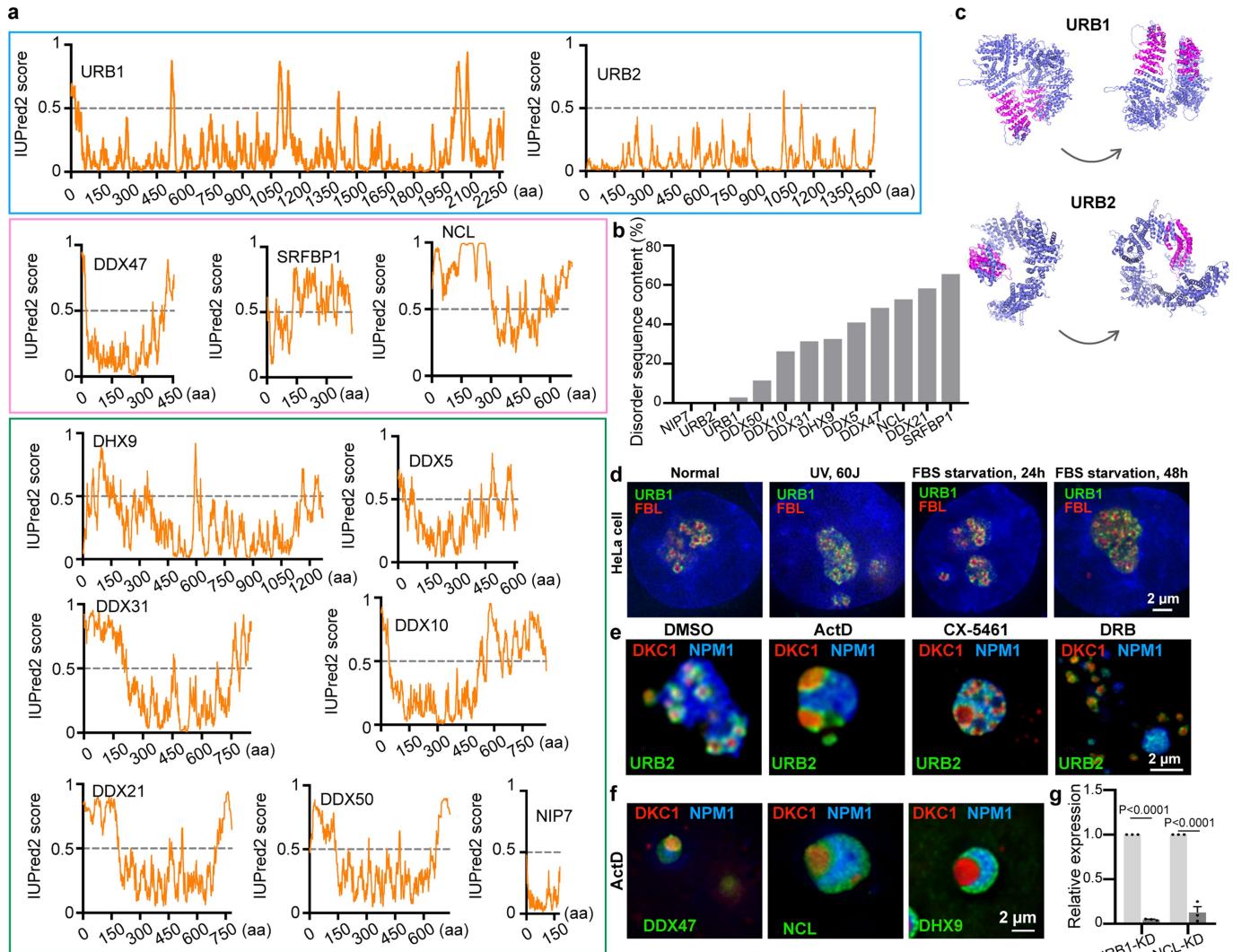
enriched only in GC. f. Representative images of nucleolar proteins enriched in NR. g. Representative images of nucleolar proteins enriched in both GC and the nucleus. h. Representative images of nucleolar proteins enriched in both GC and the cytoplasm. i. Representative images of nucleolar proteins enriched in GC, the nucleus and the cytoplasm. j. Representative images of nucleolar proteins formed aggregations in GC. k. Representative images of protein candidates that obtained multiple distribution patterns. l. Representative images of protein candidates that did not show obvious signals in nucleoli.

Article



Extended Data Fig. 3 | Validation of nucleolar sub-regions. a. Representative wide-field images of each individual PDFC protein in fixed HeLa cells with DKC1/NPM1 as markers of DFC/GC, respectively. Data was collected from $n = 4$ cells. Scale bar 2 μ m. b. CLEM images of EGFP-KI URB1. Left, the merged image of EGFP-URB1 acquired by light microscopy and correlative nucleolus acquired by TEM. Right, the zoomed-in image from the white square. c. Validation of PDFC in H9 cells. Representative SIM image of PDFC (URB1) and DFC (FBL) by direct anti-URB1 and anti-FBL IF in H9 cells. d. Validation of PDFC in different cell types. Representative wide-field images of PDFC marker proteins URB1 and URB2 in K562, U2OS and HEK293 cells by exogenous expression of EGFP-tagged individual candidate proteins, mRuby3-DKC1 and mTagBFP2-NPM1. Data was collected from $n = 4$ cells. Scale bar 2 μ m. e. Diameters of different nucleolar

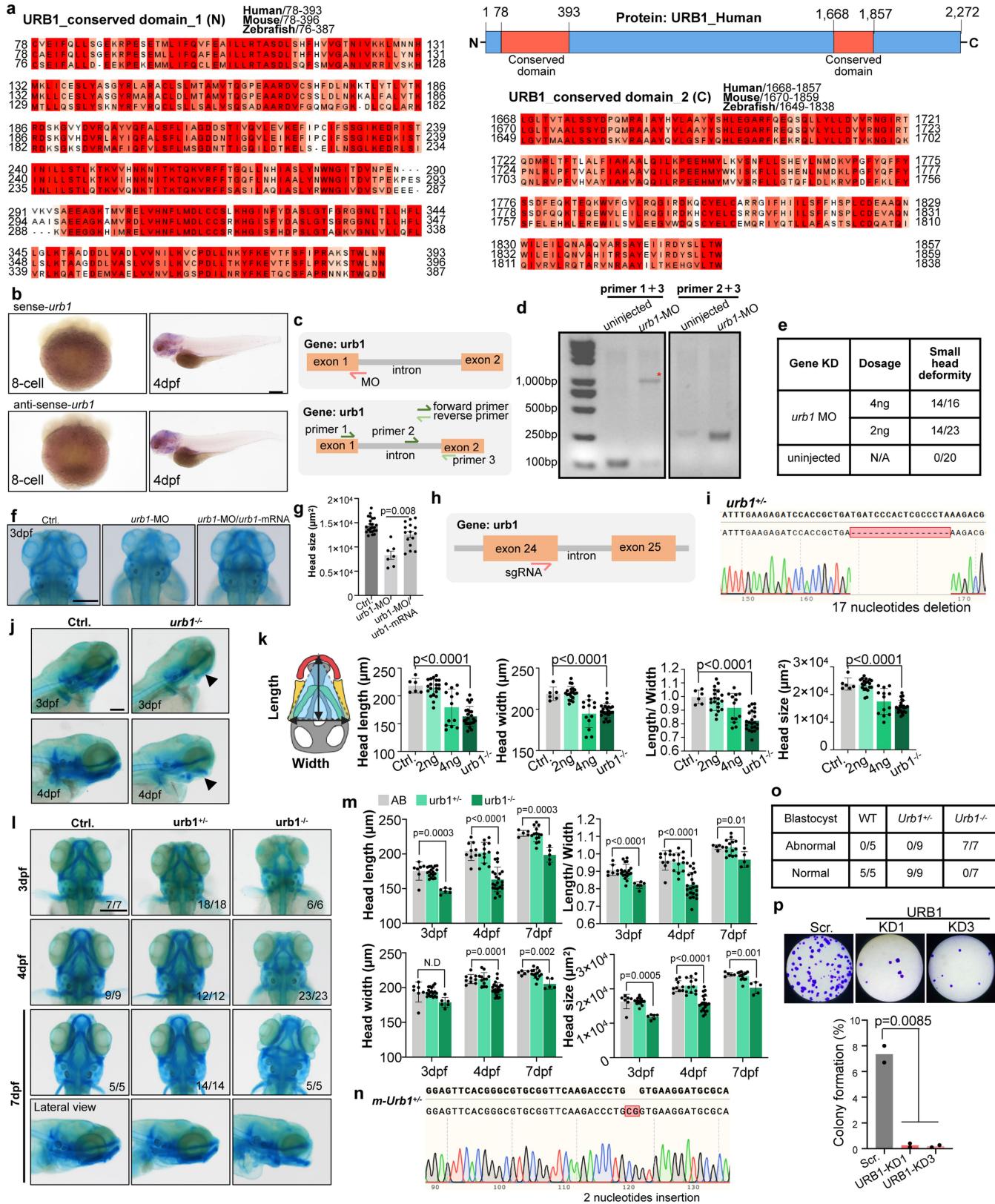
compartments. RPA194, DKC1, URB1, and NPM1 were used as FC, DFC, PDFC and GC marker proteins, respectively, $n = 60$. f. Representative wide-field images of classical and non-classical nucleolar sub-regions. Left panel, classical nucleolar sub-regions are represented by exogenous expression of UBTF (FC marker), NOP56 (DFC marker) and DDX18 (GC marker) in DKC1/NPM1-KI dual-color cells as described in Extended Data Fig. 1c. Non-classical nucleolar sub-regions are marked by exogenous expression of URB1 (PDFC marker), PTBP1 (PNC marker) and MKI67 (NR marker). Data was collected from $n = 4$ cells. Scale bar 2 μ m. g. Validation of GC condensates in different cell types. Representative wide-field images of GC condensates in K562, U2OS and HEK293 cells by exogenous expression of EGFP-RBM4, mRuby3-DKC1 and mTagBFP2-NPM1. Data was collected from $n = 4$ cells. Scale bar 2 μ m.



Extended Data Fig. 4 | Characterization of nucleolar PDFC proteins. **a.** The disorder tendency of 12 PDFC proteins calculated by IUPred2A. **b.** The rank of PDFC proteins about the disorder sequence content. URB1 and URB2 obtain the lowest disorder sequence content among all PDFC proteins. **c.** Two different perspectives of the URB1 and URB2 protein structures predicted by AlphaFold2. The predicted RNA binding domains are shown in magenta. **d.** Representative images of PDFC (URB1) and DFC (FBL) in HeLa cells treated with FBS starvation (24 h and 48 h), or UV radiation (60J). Data was collected from $n = 4$ cells. Scale bar 2 μ m. **e.** URB2 surrounds DFC under different treatments. Representative images of DKC1 (DFC), URB2 (PDFC), and NPM1 (GC) proteins under DMSO

(control), ActD, cx-5461 and DRB treatments in live HeLa cells. Of note, mRuby3-DKC1/mTagBFP2-NPM1 knock-in HeLa cells were introduced with EGFP-URB2 expression, followed by different treatments (ActD: 10 ng/mL, 3 h; DRB: 50 μ M, 1.5 h; CX-5461: 2 μ M, 1 h). **f.** PDFC proteins surrounds DFC under Pol I inhibition. Representative images of DFC, PDFC (DDX47, NCL and DHX9) and GC proteins under ActD treatment in live HeLa cells. Data was collected from $n = 4$ cells. Scale bar 2 μ m. **g.** RT-qPCR validation of the shRNA knockdown efficiency in URB1 KD and NCL KD cells. Data are presented as mean values \pm SEM. P-values are calculated by two-tailed Student's t-test. $n = 3$ independent experiments.

Article



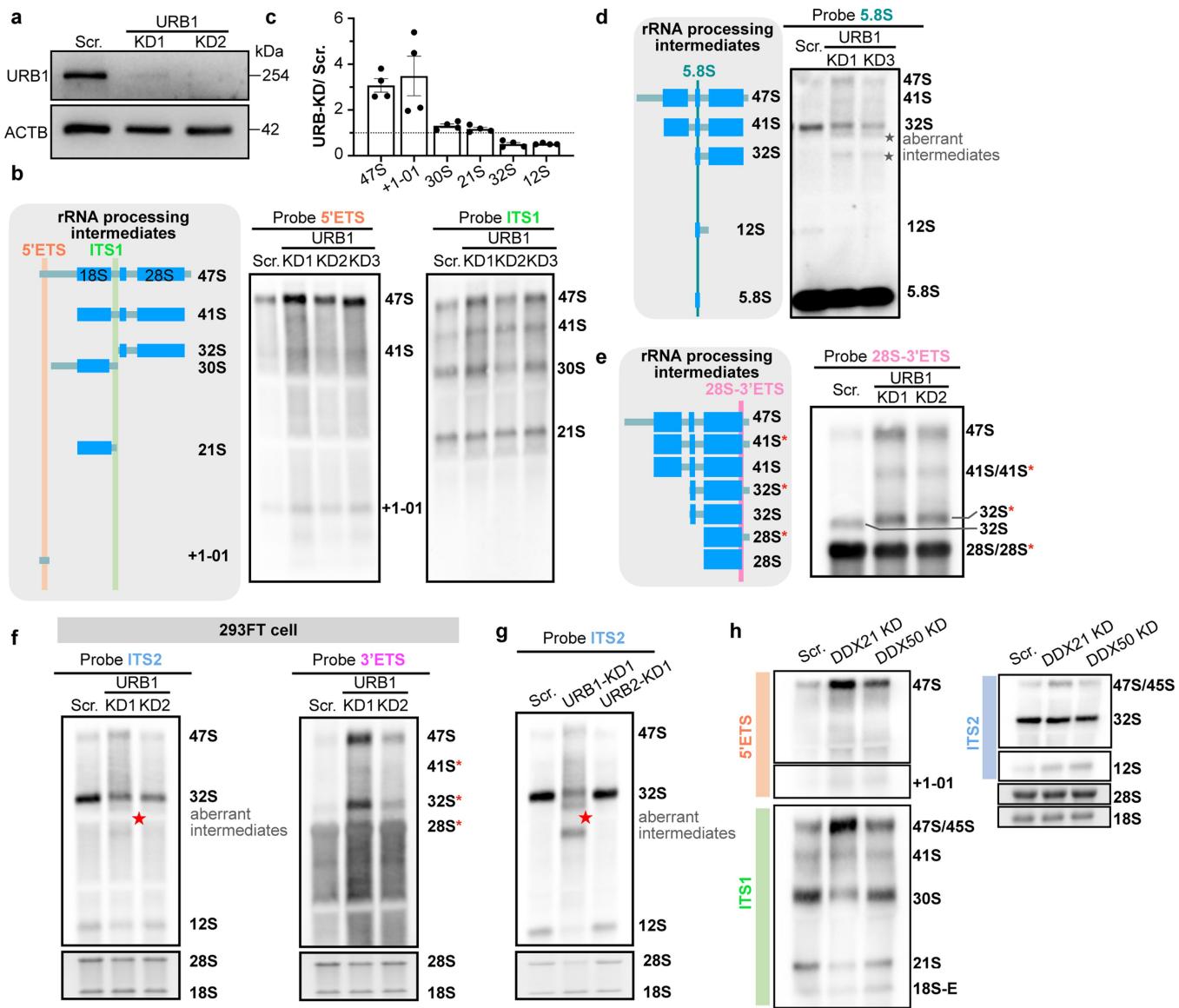
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Generation of URB1 depleted zebrafish and mouse models.

a. The conservation analysis of URB1 between human, mouse and zebrafish. URB1 has two conserved domains in N- and C-terminus. The conserved amino acids are shown in red according to conservative degrees. b. *urb1* is barely expressed in zebrafish embryos at 8-cell and 4dpf stages, as shown by WISH with either the sense (control, upper panels) or the antisense *urb1* (bottom panels) probe. Scale bar 100 μm. c. The schematic of the MO targeting site and KD efficiency detecting primers in zebrafish *urb1* gene locus. d. RT-PCR validation of the MO efficiency targeting *urb1*. The agarose gel shows the MO KD efficiency detected by PCR using detecting primers in un-injected and *urb1*-MO zebrafish, n = 15. The red asterisk indicated the splicing-blocked mRNA of *urb1* after MO injection. e. Statistics of small head deformity with different dosages of *urb1* MO in zebrafish. f. Introducing *urb1* mRNA rescues the *urb1* MO phenotypes in zebrafish. Representative images of zebrafish cartilages stained with alcian blue after injection of *urb1* MO alone, or *urb1* MO and *in vitro* transcribed zebrafish *urb1* mRNA. Scale bar 100 μm. g. Statistics of the head size in (f). n = 45 biologically independent animals. Data are presented as mean values +/- SEM. h. The schematic of the URB1-sgRNA targeting site in zebrafish *urb1* gene locus. i. The sanger sequencing result of *urb1*^{−/−} zebrafish confirmed the *urb1*^{−/−} zebrafish has a 17-nucleotides deletion at exon 24, which

resulted in a frameshift in the following *urb1*^{−/−} embryos. j. URB1 depletion results in zebrafish cranial cartilage hypoplasia. Lateral view of alcian blue stained control and *urb1*^{−/−} larvae at 3dpf and 4dpf. n = 6 animals. Scale bar 100 μm. k. URB1 depletion results in zebrafish cranial cartilage malformations. Statistics of width, length and head area (illustrated on left) in control, *urb1* MO-injected and *urb1*^{−/−} larvae at 4dpf. n = 62 biologically independent animals. Error bars, mean ± SEM. P-values are calculated by two-tailed Student's t-test. l. Dorsal and lateral view of alcian blue stained control, *urb1*^{−/−} and *urb1*^{−/−} larvae. Ratio of observed phenotypes is labeled lower right of the panel. Scale bar 100 μm. m. Statistics of the head length, head width, length/width and head size in (j). *urb1*^{−/−} larvae have a retardation in various aspects detected in head development compared to control and *urb1*^{−/−} larvae. n = 102 independent animals. Data are presented as mean values +/- SEM. P-values are calculated by two-tailed Student's t-test. n. Sanger sequencing confirmed that the *Urb1*^{−/−} mouse received a 2-nucleotides insertion at exon 1, which resulted in a frameshift in *Urb1*^{−/−} embryos. o. Statistics of blastocysts (E4.5) isolated from *Urb1*^{+/−} intercrosses. p. URB1 depletion impaired the colony formation ability of H9 cells. Data are presented as mean values +/- SEM. P-values are calculated by two-tailed Student's t-test.

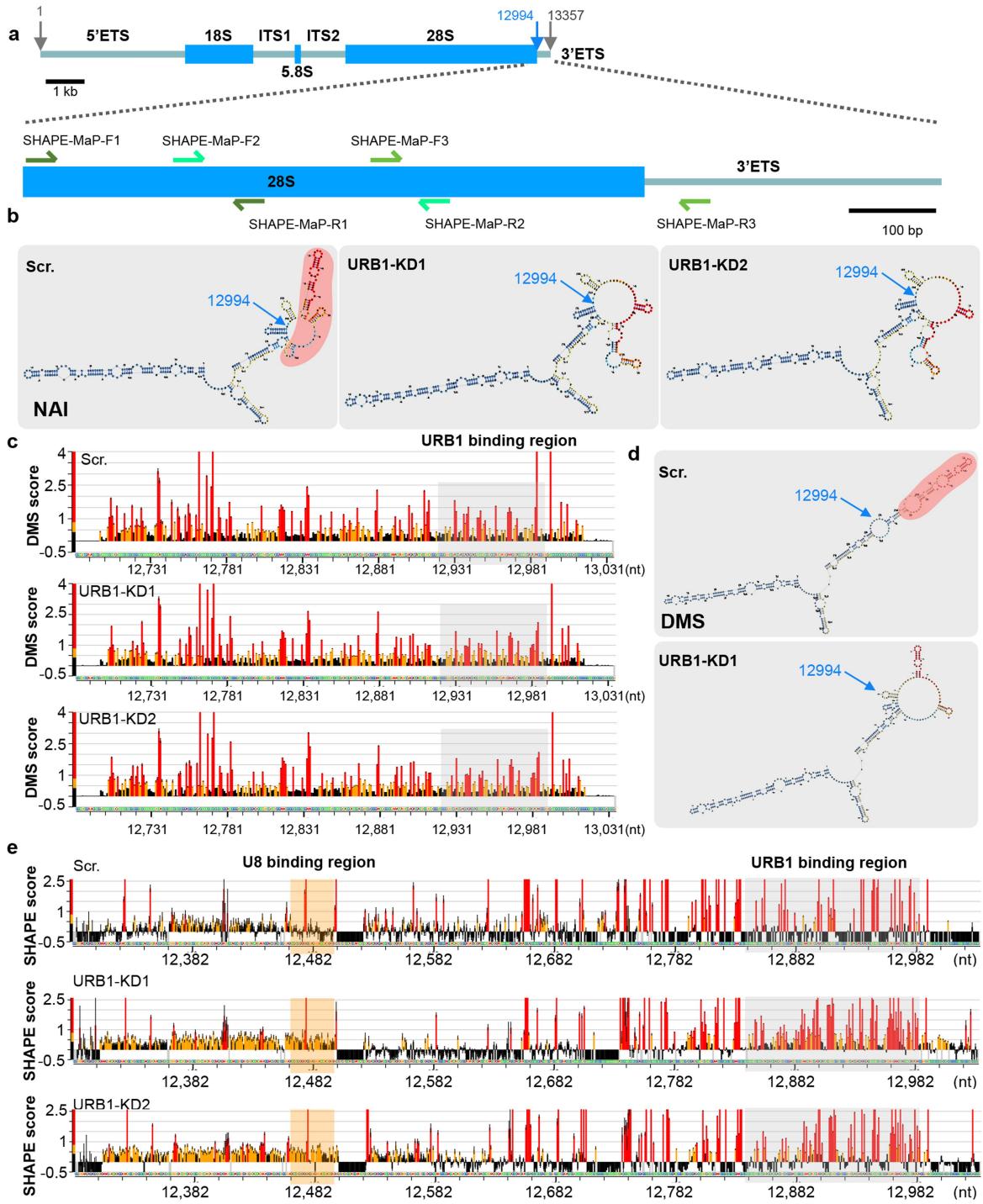
Article



Extended Data Fig. 6 | URB1 depletion impairs 32S pre-rRNA biogenesis.

a. WB validation of URB1 depletion efficiency in URB1 KD cell lines. b. URB1 depletion affects pre-rRNA biogenesis. Left, schematic of human pre-rRNA processing and positions of probes used in Northern Blots (NB). Right, NB results of pre-rRNA intermediates in scramble shRNA-treated and URB1-KD HeLa cells, with probes recognizing 5'ETS and ITS1, respectively. c. Quantification of pre-rRNA intermediates detected by different NB probes shown in (b). Each indicated rRNA intermediate is normalized to scramble shRNA-treated cells and shown as the relative intensities. n = 4 independent experiments. Data are presented as mean values +/- SEM. d. The aberrant 32S pre-rRNA intermediates contain the 5.8S rRNA region. Left, schematic representation of human 5.8S probe related rRNA intermediates in NB. Right, NB analyses of 5.8S probe related rRNA intermediates in scramble and URB1 KD HeLa cells. e. Aberrant

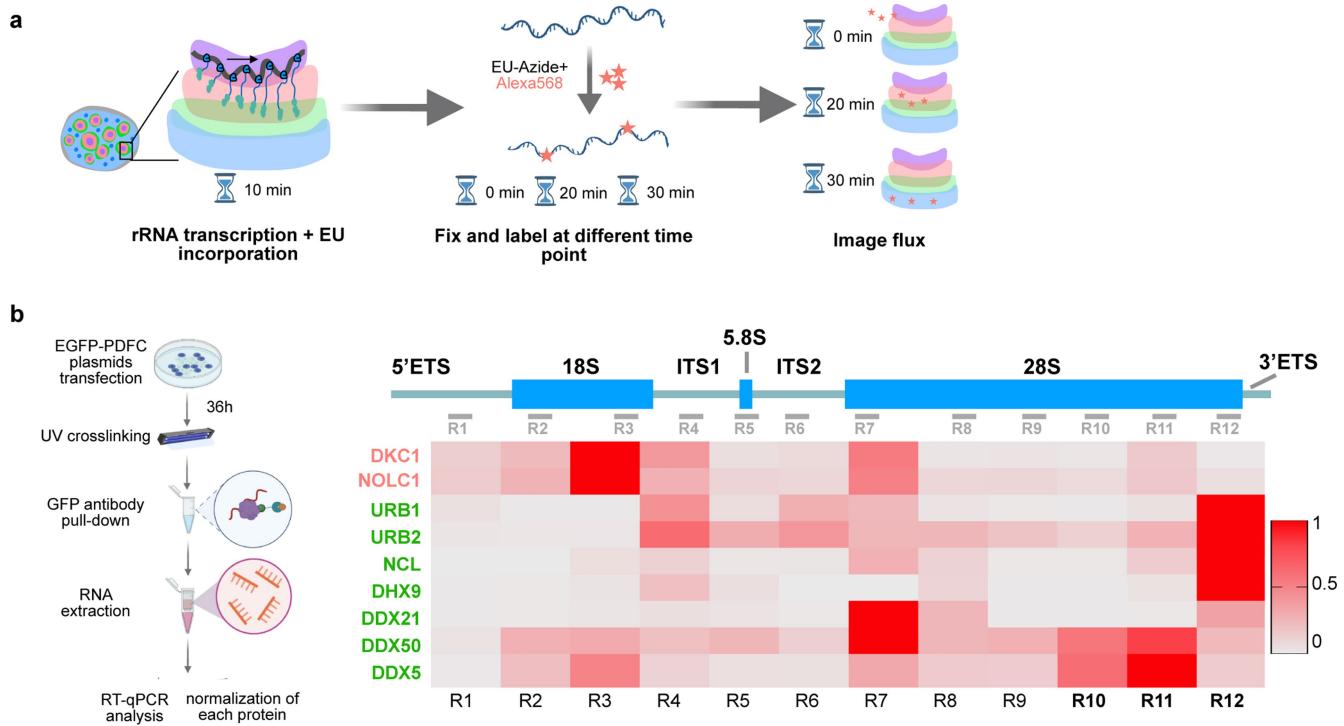
3' 32S pre-rRNA takes up a majority of total 32S pre-rRNA after URB1 depletion. Left, schematic of human 28S-3' ETS related rRNA intermediates in NB. Right, NB of 28S-3' ETS probe related rRNA intermediates in scramble shRNA-treated and URB1 KD HeLa cells. f. URB1 depletion disrupts 32S pre-rRNA biogenesis in 293FT cells. NB analyses of ITS2 and 3'ETS probe related rRNA intermediates in scramble shRNA-treated and URB1 KD 293FT cells. Red pentacle, aberrant intermediates. g. NB analyses of ITS2 probe related rRNA intermediates show no obvious reduction in 32S pre-rRNA and no accumulation of aberrant intermediates after URB2 KD compared to URB1 KD. Red pentacle, aberrant intermediates. h. NB analyses of rRNA intermediates in scramble shRNA-treated, DDX21 KD and DDX50 KD HeLa cells, respectively with 5'ETS, ITS1 and ITS2 probe shows no obvious or subtle reduction in 32S pre-rRNA compared to URB1 KD.



Extended Data Fig. 7 | The depletion of URB1 leads to an increase in the SHAPE value of binding region and its upstream revealed by SHAPE-MaP.
a. Schematic of primers used in SHAPE-MaP experiments. **b.** Predicated secondary structural modes of 28S-3'ETS region in scramble shRNA-treated and URB1 KD HeLa cells labeled by NAI labeling. Red region, URB1 binding region.

c. The DMS scores of 28S-3'ETS region in scramble shRNA-treated and URB1 KD HeLa cells labeled by DMS. **d.** Predicated secondary structure modes in (c). Red region, URB1 binding region. **e.** The SHAPE scores of 28S-3'ETS region and U8-snoRNA binding region in scramble shRNA-treated and URB1 KD HeLa cells labeled by NAI.

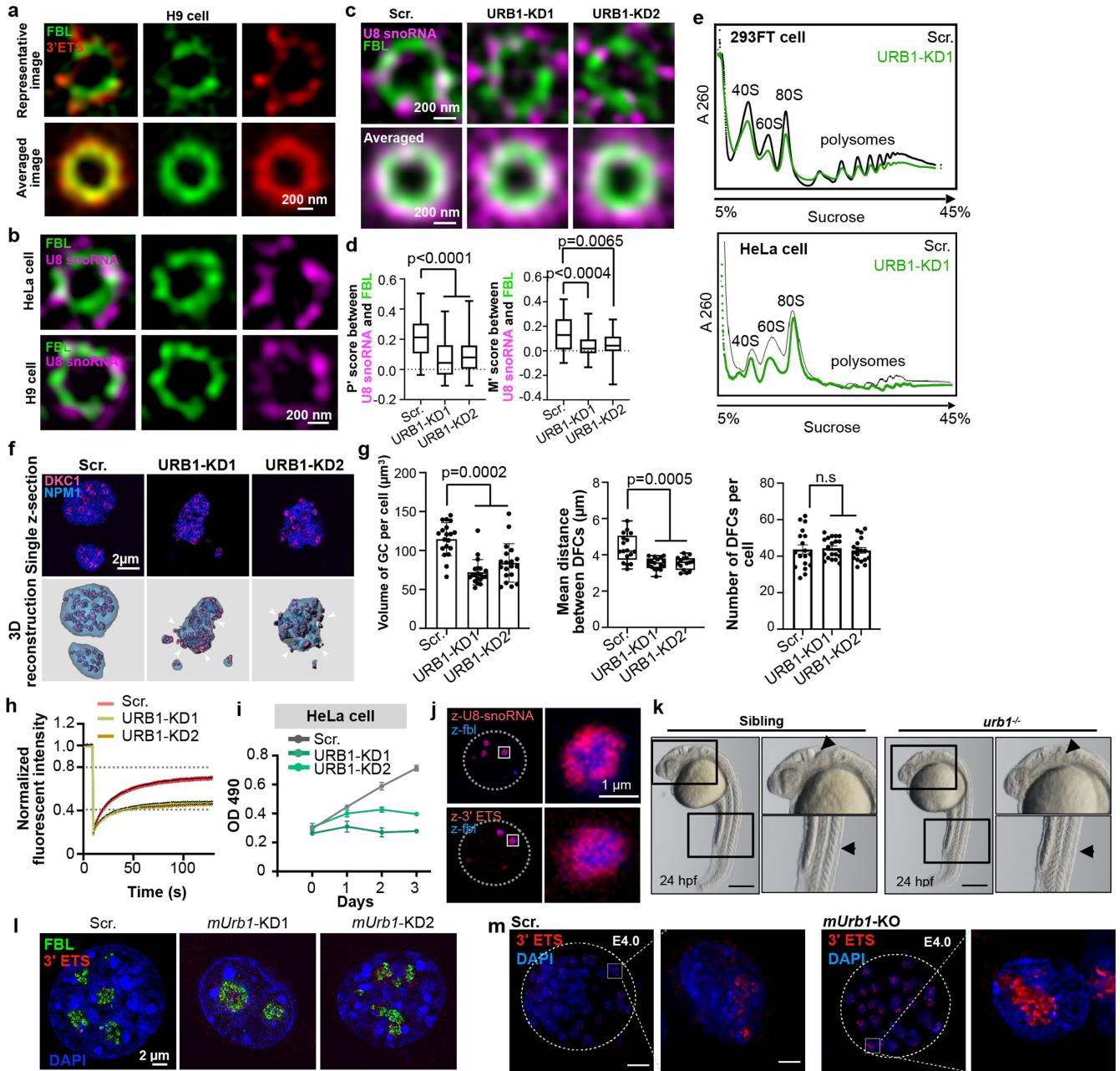
Article



Extended Data Fig. 8 | Examined PDFC proteins prefer to bind the 28S and 3'

a. Schematic of pulse-chase EU labelling experiment.
b. Several examined PDFC proteins preferentially bind to the 3' end of 28S pre-rRNA analyzed by RNA Immunoprecipitation (RIP) in HeLa cells. Left, the

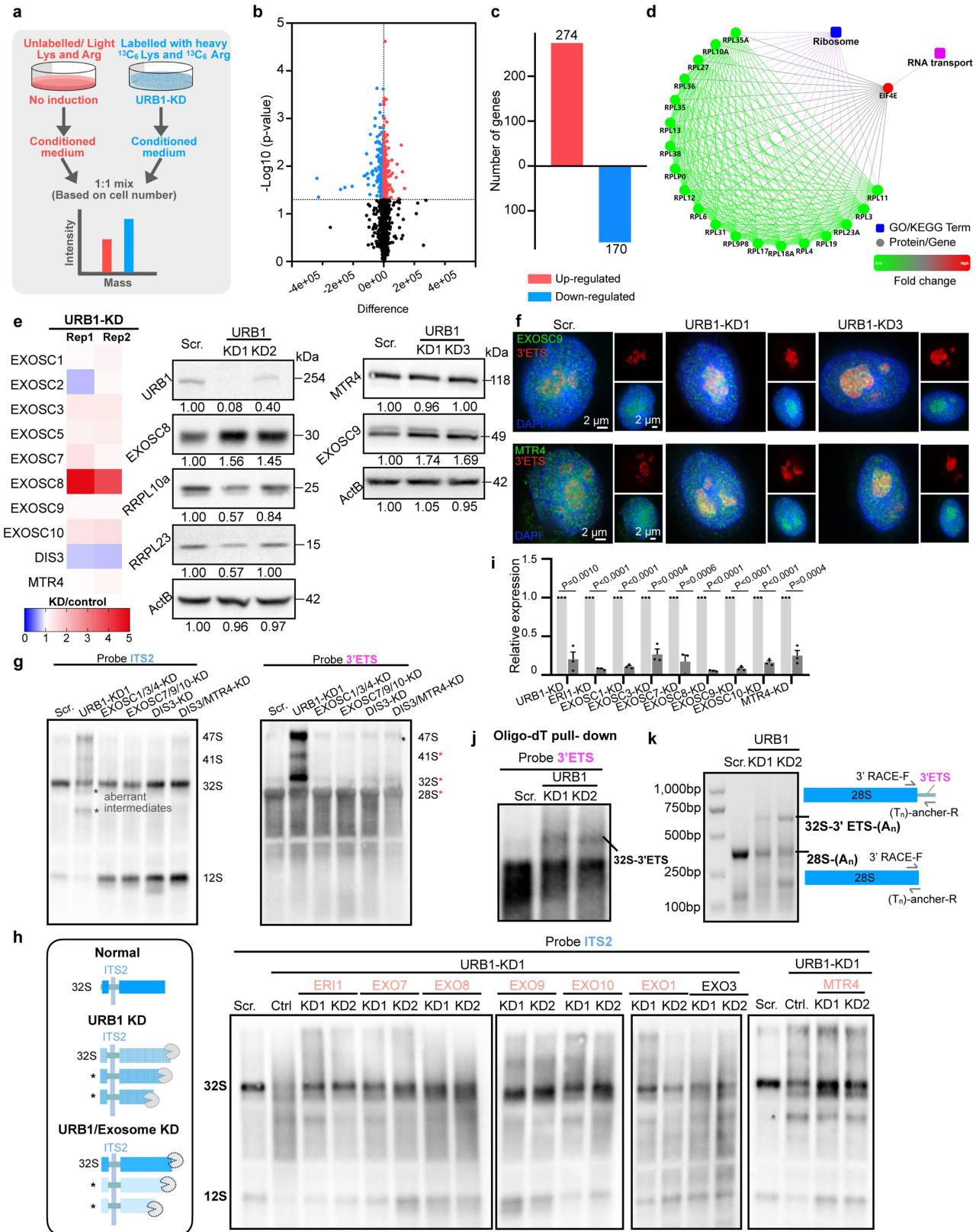
schematic of RIP experiments. Right, the heatmap of PDFC proteins binding to the 47S pre-rRNA. Data was collected from three independent experiments; heatmap values were generated by averaging IP/Input of each examined PDFC protein, and normalized to the highest binding capability, respectively.



Extended Data Fig. 9 | URB1 depletion alters the distribution of both 3' ETS-attached pre-rRNAs and U8 snoRNAs. a. 3' ETS-attached pre-rRNAs localize at the periphery of DFC in H9 cells. Top, representative single-z-section SIM images of 3' ETS-attached pre-rRNAs (red) and DFC (FBL, green). Bottom, representative averaged images. b. U8 snoRNAs localize at the periphery of DFC. Top, representative single-z-section SIM images of U8 snoRNAs (magenta) and DFC (FBL, green) in HeLa cells. Bottom, representative single-z-section SIM images of U8 snoRNA and DFC (FBL) in H9 cells. c. URB1 KD alters the distribution of U8 snoRNAs. U8 snoRNAs (magenta) detected by smFISH were aberrantly dispersed to GC after URB1 knockdown in HeLa cells. Top, representative single-z-section SIM images of U8 snoRNAs and DFC in scramble and URB1 KD cells. Bottom, averaged images shown above, n = 20. d. Statistics of the altered distributions of U8 snoRNAs in (c). The floating bar shows the range of values, and the center line represents the median. P-values are calculated by two-sided Mann-Whitney tests. e. URB1 depletion led to reduced ribosome production in 293FT cells and HeLa cells, as shown by ribosome fractionations. The position of URB1 KD is noted by the green line and the scramble is noted by the black line. f. URB1 KD leads to reduced GC volume. Representative single z-section (upper panels) and 3D-reconstructed (bottom panels) SIM images of DKC1 (magenta) and NPM1 (blue) in scramble shRNA-treated and URB1-KD HeLa cells. White arrows show that some DFCs are extruded outside of GC. g. Statistics of the GC

volume, the mean distance between DFCs and DFC numbers per cell in (f), n = 15. The floating bar shows the range of values. Data are presented as mean values +/- SEM. P-values are calculated by two-tailed Student's t-test. (NS, not significant). h. URB1 KD led to decreased mobility of NPM1 in HeLa cells. FRAP analysis of NPM1 in scramble shRNA-treated and URB1 KD HeLa cells, n = 10. i. URB1 KD led to retarded cell proliferation in HeLa cells, revealed by MTT cell proliferation assays, n = 3. Data are presented as mean values +/- SEM. j. 3' ETS-attached pre-rRNAs and U8 snoRNAs localize at the periphery of merged FC/DFC in zebrafish embryo. U8-snoRNAs (top, red) and 3' ETS-attached pre-rRNAs (bottom, red) detected by smFISH surround z-fbl (blue) in zebrafish 8 hpf embryos, n = 3. Data are presented as mean values +/- SEM. k. *urb1*^{-/-} larvae exhibit a less-defined midbrain-hindbrain boundary and disorganized trunk vasculature (arrowheads) compared to siblings at 24 hpf. Scale bar 100 μm. l. *mUrb1*-KD impairs the radial flux processing of pre-rRNAs. 3' ETS-attached pre-rRNAs (red) detected by 3' ETS smFISH probes were diffused to the outside of DFC (FBL detected by IF) after *mUrb1* knockdown in mouse NIH-3T3 cells, n = 20. m. *mUrb1* depletion impairs the radial flux processing of pre-rRNAs in blastocyst (E4.0). Representative images of 3' ETS-attached pre-rRNAs in control and *mUrb1*-KO (knockout) blastocyst by smFISH showed a diffused pattern of pre-rRNAs into GC region after *mUrb1* depletion, n = 3. Scale bar 10 μm.

Article



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Loss of URB1 activates the exosome surveillance in the nucleolus. a. The workflow of stable isotope labeling using amino acids in cell culture (SILAC) followed by mass spectrometry (MS) in scramble shRNA-treated and URB1 KD HeLa cells. b. Volcano plot of the changes in protein expression upon URB1 KD. 4,688 proteins were overlapped in independent replicates. P-values are calculated by two-sided Mann-Whitney tests. c. The number of up-regulated and down-regulated proteins in URB1 KD HeLa cells. 274 proteins are up-regulated >2-fold and 170 proteins are downregulated >2-fold. d. GO/KEGG analyses of expression changed proteins upon URB1 KD in HeLa cells. Briefly, the expression levels of ribosome components (especially LSU) are down regulated and EIF4E (RNA transport factor) is up regulated. e. WB validation of the related changes in protein expression upon URB1 KD. EXOSC8 is up regulated, RPL10 and RPL23 are down regulated. f. Exosome components accumulate with 3' ETS-attached pre-rRNAs in nucleoli upon URB1 KD. Representative images of EXOSC8, EXOSC9 and MTR4 with 3' ETS-attached pre-rRNAs in scramble shRNA-treated and URB1 KD HeLa cells are shown by smFISH and IF. g. NB analyses of rRNA intermediates in scramble shRNA-treated,

URB1, EXOSC1/3/4, EXOSC7/9/10, DIS3 and DIS3/MTR4 KD HeLa cells, respectively with ITS2 or 3' ETS probes. Left, depletion of different exosome components led to a reduction or no change in 32S pre-rRNA expression, and no obvious aberrant intermediates were observed. Right, depletion of different exosome components did not affect 3' ETS processing. h. Most exosome components KD rescues the abnormal exosome-dependent 32S pre-rRNA degradation. Left, the schematic of URB1 and exosome components that affect the 32S pre-rRNA degradation. Right, NB detection of the ITS2-containing pre-rRNA intermediates in scramble shRNA-treated and URB1 KD cells that were further depleted exosome components. i. RT-qPCR validation of the shRNA knockdown efficiency in cells KD of URB1 and exosome components. P-values are calculated by two-tailed Student's t-test. j. URB1 depletion generated 3' ETS-attached pre-rRNAs carry short poly(A) tails. NB detection using the 3' ETS probes after oligo-dT pull-down in scramble shRNA-treated and URB1 KD HeLa cells. k. 3'RACE analysis of pre-rRNA intermediates in scramble shRNA-treated and URB1 KD HeLa cells shows short poly(A) tails in 3' ETS-attached pre-rRNAs.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

WF, SIM and FRAP images were acquired with softWoRx 7.0 (GE). Spin disk images were acquired with cellSens (3.1). Stereoscopic images were acquired with Nikon NIS Elements (9.0). STED images were acquired by Abberior Imspector Software (16.3.14287-w2129).

Data analysis

For imaging analyses, Fiji Image J (ImageJ 1.53q) and Imaris (9.0) were used. GraphPad Prism9, R (v9.0.0) and Microsoft office 365 were used for statistical analysis and graphing. For iCLIP-seq data, raw read quality was evaluated by FastQC (v0.11.9). Deep sequencing datasets were mapped with Bowtie2 (Bowtie2 v2.3.5, parameters: -q --sensitive -a -p 8 --no-mixed --reorder --no-unal) and aligned to 45SrRNA sequence in human reference genome; then, reads introduced by PCR duplicates were removed via a custom script. We obtained the read-coverage on 45SrRNA using genomeCoverageBed (genomeCoverageBed v2.29.0, parameters: -split), and calculated the fraction of reads overlapping each position in 45S rDNA. For SHAPE-MaP data, raw read quality was evaluated by FastQC (v0.11.9). Trimmed data was mapped with ShapeMapper2 (v 2.1.3, parameters: --verbose --serial --nproc 16 --min-depth 1000). The RNA secondary structure from SHAPE-MaP was predicted by RNAfold (v 2.4.14, parameters: -p -d2 --shapeMethod=D). The changes of RNA SHAPE value were calculated by URB1 KD samples minus scramble samples using deltaSHAPE (v1.0) with the parameters (--all --colorfill --noshow --pdf -p 5 -z 0 -s 0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data supporting the findings of this study are available in the manuscript, Source Data, and Mendeley Data <https://data.mendeley.com/datasets/8fmmx5vpkt/>

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Preliminary experiments were performed to determine the sample size. Sample size sufficiency were determined by previous experiments from our laboratories. No statistical methods were used to predetermine the sample size.
Data exclusions	No data was excluded.
Replication	Embryos were collected from at least 6 pairs of fish/mouse breedings. Data come from at least two independent experiments. All replication experiments gave similar results. All samples were pulled together during analysis.
Randomization	For the mouse and zebrafish experiments: embryos were allocated into experimental groups based on their genotypes. For zebrafish MO injections, embryos were allocated into experimental groups based on different treatments. For cell experiments: cells were allocated into experimental groups based on different shRNA treatments.
Blinding	Not applicable due to the lack of randomizable treatments/interventions in the experimental plan.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

Antibodies

Antibodies used	<ol style="list-style-type: none">1) anti-Flag (Sigma; F1804 and Shanghai Genomics; SG4110-16)2) anti-URB1 (Proteintech; 20023-1-AP)3) anti-Digoxigenin-AP, Fab fragments (Roche; 11093274910)4) anti-EXOSC8 (Proteintech; 11979-1-AP)5) anti-MTR4 (Proteintech; 12719-2-AP)6) anti-RPL23 (Proteintech; 16086-1-AP)7) anti-RPL10A (Abcam; ab174318)8) anti-ACTB (Sigma; A3854)9) anti-EXOSC9 (Proteintech; 24470-1-AP)10) anti-NPM1 (Santa Cruz Biotechnology; sc-53175)11) Goat anti-Mouse Secondary Antibody, Alexa Fluor 555 (Invitrogen; A-21424)12) Goat anti-Rabbit Secondary Antibody, Alexa Fluor 488 (Invitrogen; A-11034)13) anti-URB1 (Invitrogen; PA5-53787)14) anti-GFP (Abcam; ab6556)15) anti-FBL (Abcam; ab5821)16) anti-FBL (Santa Cruz Biotechnology; sc-374022)17) Goat anti-Mouse Secondary Antibody, Abberior STAR RED (Abberior Cat# 2-0002-011-2, RRID:AB_2810982)18) Goat anti-Rabbit Secondary Antibody, Abberior STAR RED (Abberior Cat# STRED-1002-500UG, RRID:AB_2833015)
-----------------	--

Validation

All antibodies against endogenous proteins had non-expressing tissues along the expression domains within images, which exhibited background level homogeneous and non-specific staining. The background staining was corrected during image analysis. EXOSC8 rabbit polyclonal antibody (Proteintech; 11979-1-AP) is validated for Immunofluorescent analysis in 4T1 cells at 1:50 dilution and for WB analysis of HeLa cell lysates at 1:500 dilution (manufacture's website). SKIV2L2 rabbit polyclonal antibody (Proteintech; 12719-2-AP) is validated for WB analysis of HepG2 cells at 1:2000 dilution (manufacture's website). RPL23 rabbit polyclonal antibody (Proteintech; 16086-1-AP) is validated for WB analysis of HeLa cells at 1:200 dilution (Nature Cell Biology vol. 19, 1 (2017): 60-67.). RPL10A rabbit monoclonal antibody (Abcam; ab174318) is validated for WB analysis of HeLa cells at 1:10000 dilution (manufacture's website). Anti-b-Actin-HRP conjugated primary antibody (Sigma; A3854) is validated for WB analysis at 1:10000 dilution (Cancer research, 72(10), 2645-2656 (2012-04-12); Cell reports, 28(7), 1703-1716 (2019-08-15)). EXOSC9 rabbit polyclonal antibody (Proteintech; 24470-1-AP) is validated for Immunofluorescent analysis in HeLa cells at 1:25 dilution and for WB analysis at 1:1000 dilution (manufacture's website). NPM1 mouse monoclonal antibody (Santa Cruz Biotechnology; sc-53175) is validated for Immunofluorescent analysis in HeLa cells (The EMBO Journal (2015) 34:2758-2774). URB1 rabbit polyclonal antibody (Invitrogen; PA5-53787) is validated for Immunofluorescent analysis in human A-431 cells (manufacture's website). FBL rabbit polyclonal antibody (Abcam; ab5821) and mouse monoclonal antibody (Santa Cruz Biotechnology; sc-374022) are validated for Immunofluorescent analysis by previous experiments from our laboratories.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

Cell lines are from commercial sources.
HeLa and HEK293FT: ATCC
U2OS: National Collection of Authenticated Cell Cultures
K562: National Collection of Authenticated Cell Cultures
Human H9 cell lines were purchased from WiCell.
Zebrafish ZF4 cell lines were purchased from China Zebrafish Resource Center (CZRC).

Authentication

Vendor of the commercially-obtained cell lines (ATCC/ National Collection of Authenticated Cell Cultures/ WiCell/ CZRC) provide further information on the generation, characteristics and authentication of the cell line in its website.

Mycoplasma contamination

All cells were tested for free of mycoplasm contamination.

Commonly misidentified lines
(See [ICLAC](#) register)

No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Zebrafish embryos: sex of the embryos was not determined. To obtain genetic mutants, a pair of female and male adult zebrafish (8 months-1year old) were mated and the resulting embryos were developed for an additional 3 days for cartilage stainings or 4 days for whole-mount smFISH. Injections were performed at the one cell stage.
Mouse embryos: sex of the embryos was not determined. To obtain genetic mutants, a pair of female and male adult mice were mated and the resulting embryos were developed to E4.5 and collected for phenotype analysis and smFISH/IF staining. Mice were given free access to food and water, and were maintained under a 12/12 h light-dark cycle with controlled temperature (20-25°C) and humidity (50 ± 10%).

Wild animals

No wild animals were used in the study.

Field-collected samples

No field-collected samples were used in the study.

Ethics oversight

All animal procedures were performed under the ethical guidelines of the CAS Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences (CAS).

Note that full information on the approval of the study protocol must also be provided in the manuscript.