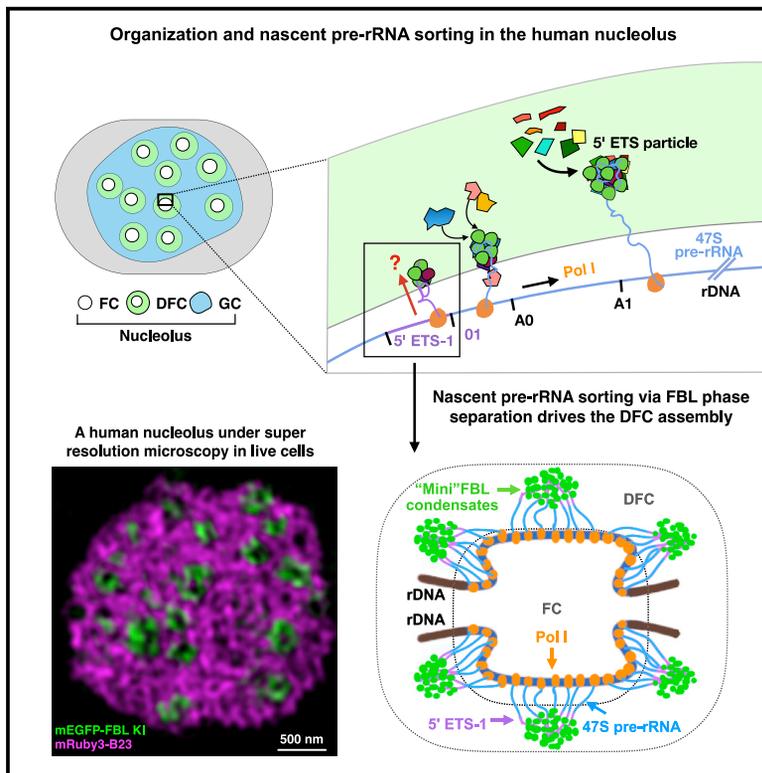


Molecular Cell

Nascent Pre-rRNA Sorting via Phase Separation Drives the Assembly of Dense Fibrillar Components in the Human Nucleolus

Graphical Abstract



Authors

Run-Wen Yao, Guang Xu, Ying Wang, ..., Yu-Hang Xing, Li Yang, Ling-Ling Chen

Correspondence

linglingchen@sibcb.ac.cn

In Brief

Yao et al. unveil the FC/DFC ultrastructure and rDNA arrangements in human nucleoli and show that a phase-separation mechanism promotes nascent pre-rRNA sorting and processing and the assembly of the DFC sub-nucleolar region.

Highlights

- Visualizing the ultrastructure of FC/DFC and rDNA arrangements in human nucleoli
- Processing factors, such as FBL, form protein clusters and then assemble into a DFC
- Self-association of GAR in FBL ensures sorting and processing of nascent 47S pre-rRNA
- Nascent pre-rRNA sorting via a phase-separation mechanism promotes DFC assembly

Nascent Pre-rRNA Sorting via Phase Separation Drives the Assembly of Dense Fibrillar Components in the Human Nucleolus

Run-Wen Yao,^{1,5} Guang Xu,^{1,5} Ying Wang,² Lin Shan,¹ Peng-Fei Luan,¹ Yang Wang,¹ Man Wu,¹ Liang-Zhong Yang,¹ Yu-Hang Xing,^{1,4} Li Yang,^{2,3} and Ling-Ling Chen^{1,3,6,*}

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, 320 Yueyang Road, Shanghai 200031, China

²CAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China

³School of Life Science and Technology, ShanghaiTech University, 100 Haik Road, Shanghai 201210, China

⁴Present address: Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

⁵These authors contributed equally

⁶Lead Contact

*Correspondence: linglingchen@sibcb.ac.cn

<https://doi.org/10.1016/j.molcel.2019.08.014>

SUMMARY

Fibrillar centers (FCs) and dense fibrillar components (DFCs) are essential morphologically distinct sub-regions of mammalian cell nucleoli for rDNA transcription and pre-rRNA processing. Here, we report that a human nucleolus consists of several dozen FC/DFC units, each containing 2–3 transcriptionally active rDNAs at the FC/DFC border. Pre-rRNA processing factors, such as fibrillarin (FBL), form 18–24 clusters that further assemble into the DFC surrounding the FC. Mechanistically, the 5' end of nascent 47S pre-rRNA binds co-transcriptionally to the RNA-binding domain of FBL. FBL diffuses to the DFC, where local self-association via its glycine- and arginine-rich (GAR) domain forms phase-separated clusters to immobilize FBL-interacting pre-rRNA, thus promoting directional traffic of nascent pre-rRNA while facilitating pre-rRNA processing and DFC formation. These results unveil FC/DFC ultrastructures in nucleoli and suggest a conceptual framework for considering nascent RNA sorting using multivalent interactions of their binding proteins.

INTRODUCTION

Like proteins, functions of RNAs are associated with unique sorting pathways and subcellular localization. Due to their single-stranded and flexible nature, newly synthesized precursor RNAs (pre-RNAs) must interact with RNA binding proteins (RBPs) both during and after transcription (Dreyfuss et al., 2002). Forming specific pre-RNA-protein (RNP) complexes may facilitate nascent pre-RNA directional movement toward

ultimate cellular destinations and may also prevent individual pre-RNAs from unproductive intra- and inter-molecular interactions. To date, the understanding of how nascent RNAs move with directionality in cells has been limited.

The mammalian nucleolus is assembled around nucleolar organizer regions (NORs), which are comprised of three morphologically distinct sub-regions named the fibrillar center (FC), dense fibrillar component (DFC), and granular component (GC) (Boisvert et al., 2007; McStay and Grummt, 2008). It is believed that such nucleolar ultrastructures are products of the functions they perform, which allow continuous RNA polymerase I (Pol I) transcription within the FC and the subsequent radial flux of rRNAs through the DFC into the GC and finally into the nucleoplasm (Boisvert et al., 2007; Farley et al., 2015; McStay, 2016). This system provides an attractive model to address how directional RNA sorting is achieved in cells. However, the compact and electron-dense nature of the nucleolus and the limitation of electron microscopy approaches in determination of individual protein localization have allowed only limited features of the sub-nucleolar organization to be revealed.

The assembly of non-membrane-bound cellular compartments requires condensation of both RBPs and RNAs involved in liquid-liquid phase separation (Feric et al., 2016; Shin and Brangwynne, 2017). Intrinsically disordered regions (IDRs) of RBPs often play important roles in RBP-RNA and protein networks (Calabretta and Richard, 2015), and RNAs were shown to promote phase separation for the formation of membrane-less bodies in a concentration- or structure-dependent manner (Langdon et al., 2018; Maharana et al., 2018). How RBPs and RNAs act in coordination to nucleate the assembly of membrane-less cellular bodies remains to be explored. Here we dissect sub-nucleolar organization in detail. By addressing the mechanism of nascent pre-rRNA sorting, we show that a phase separation mechanism-controlled process plays an important role in driving DFC assembly.

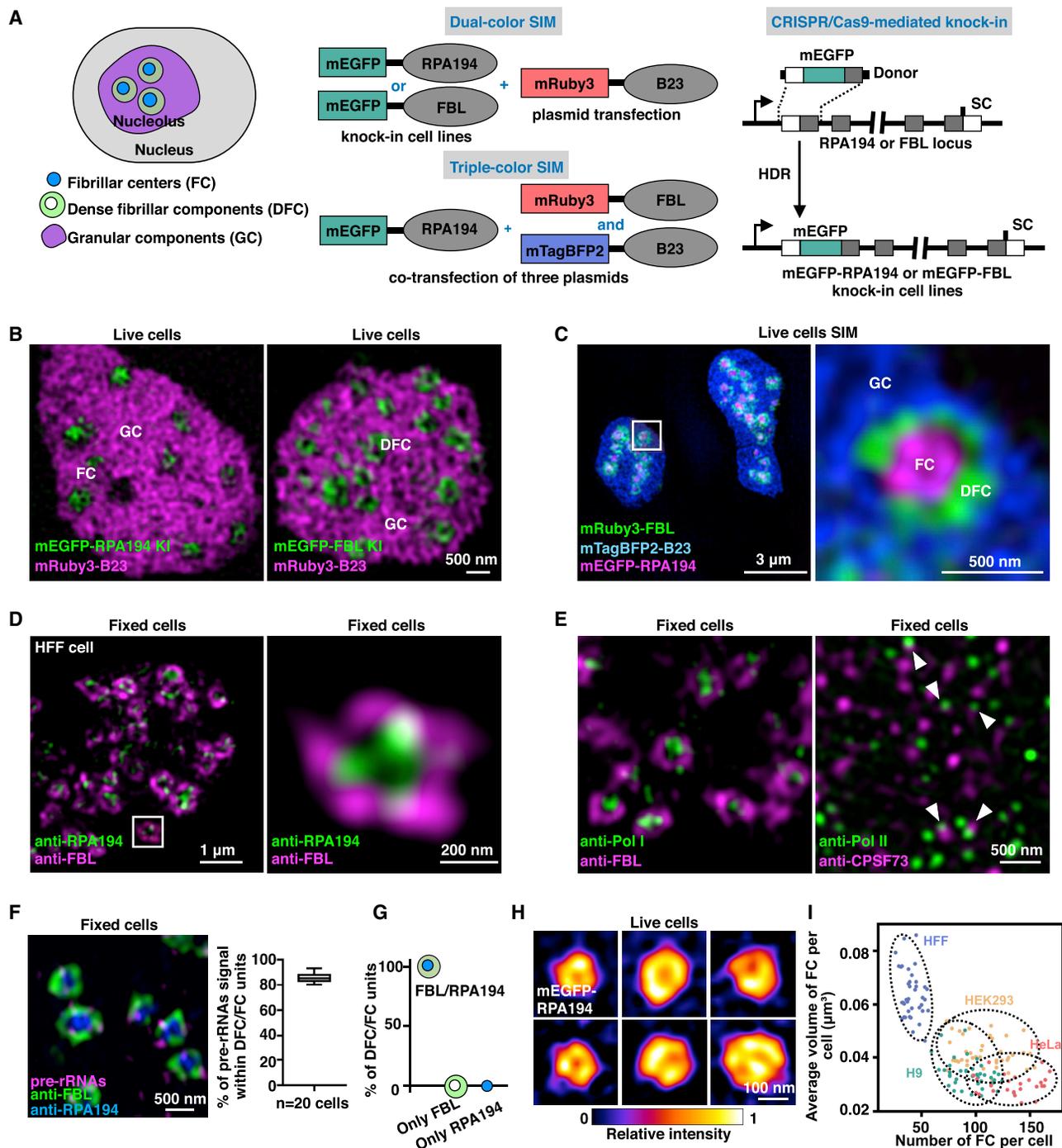


Figure 1. Visualization of FC/DFC Units in the Human Nucleolus

(A) Workflow of live-cell imaging to visualize nucleolar subdomains. Left, a mammalian nucleolus is comprised of three sub-regions. FCs, RPA194, blue; DFCs, FBL, green; GC, B23, purple. Right, plasmids and cell lines used to visualize nucleolar ultrastructure by SIM. The RPA194 or FBL locus was tagged with mEGFP via CRISPR/Cas9-mediated KI. White boxes, 5' and 3' untranslated regions (UTRs); dark gray boxes, exons; SC, stop codon; HDR, homology-directed repair. See also Figure S1B.

(B) Representative SIM images of nucleoli in live cells.

(C) Representative live HeLa cell SIM images of three nucleolar sub-regions. See also Figure S1C for wide field images.

(D) Representative SIM images of one FC/DFC in fixed HFF cells. See also Figure S1F.

(E) Distinct patterns of Pol I (left) and Pol II (right) factors for RNA transcription and processing. See also Figure S1G.

(legend continued on next page)

RESULTS

Visualization of FC/DFC Units in Human Cells

Structured illumination microscopy (SIM) provides multicolor imaging at ~ 120 nm resolution on the x-y scale (Fei et al., 2017) (Figure S1A). We expressed or knocked in fluorescently tagged proteins to visualize FC, DFC, and GC sub-nucleolar regions in HeLa cells under SIM (Figures 1A and S1B). RPA194 is a Pol I component in FCs; Fibrillarin (FBL) is a key factor involved in 47S pre-rRNA processing in DFCs; and B23, also called NPM1, is associated with the later stage of pre-rRNA processing in GCs (Boisvert et al., 2007). SIM observations showed significantly enhanced resolution of all examined sub-nucleolar regions compared to that of wide field imaging (Figures 1B, 1C, S1C, and S1D).

Consistent with electron microscopy (EM) observations in fixed cells (Boisvert et al., 2007; McStay and Grummt, 2008), we found that all FCs or DFCs were imbedded within GCs in live cells (Figures 1B and 1C). All detected RPA194 in FCs were surrounded by DFC-localized FBL in examined HeLa, HEK293, human foreskin fibroblast (HFF), and human embryonic stem cells (Figures 1C, 1D, S1E, and S1F). Each FC and DFC for Pol I transcription and pre-rRNA processing is distinct from the sparse co-localization pattern between the Pol II subunit POLR2A and examined pre-mRNA processing factors (PFs) (Figures 1E and S1G). Moreover, SIM images showed that 85% of pre-rRNA signals were localized within FC/DFCs (Figure 1F) and that all FC/DFCs were associated with pre-rRNAs (Figures 1F and 1G), confirming that each FC/DFC is transcriptionally active and forms a pre-rRNA transcription and processing unit.

Increased spatial resolution in both fixed and live cells (Figure S1D) revealed previously unknown FC ultrastructure. First, Pol I complex subunits RPA194 and RPA49 (Figures 1H, S1H, and S1I) are enriched at the FC border. These findings are in contrast to previous models in which Pol I complexes were thought to be distributed throughout the FC region (Cheutin et al., 2002). Second, human cells have variable numbers of FC/DFC units, from several dozen to over 100 (Figure 1I); these numbers, detected in live cells, are higher than previously observed under EM (Boisvert et al., 2007; McStay and Grummt, 2008). Third, numbers and sizes of FC regions are disparate among different cell types but are relatively fixed in the same type of cells (Figure 1I), suggesting that the morphology of FCs can be used as an indicator of cell type in addition to variations in nucleolar number and size (Farley et al., 2015).

Organization of Active and Inactive rDNA in NORs

About half of the several hundred rDNA repeats are transcriptionally active owing to protein synthesis requirements of the cell (Boisvert et al., 2007). We examined the arrangement of active

and inactive rDNAs in detail by fluorescence *in situ* hybridization (FISH) (van Sluis et al., 2016) under SIM (Figure 2A). We found that rDNAs are located both in the nucleolus and in the nucleoplasm (Figure 2B), the former pattern of rDNAs being known to be silent NORs (McStay and Grummt, 2008). Not all examined HeLa cells exhibited silent NORs, and, strikingly, not all rDNAs in NORs were transcriptionally active (Figure 2B).

On one hand, a significant fraction of rDNA clusters (Figures 2C and 2D) was located between the FC and DFC regions (Figure 2C), confirming that active Pol I transcription occurs at the border of FC/DFC (Cheutin et al., 2002; Koberna et al., 2002; Puvion-Dutilleul et al., 1991). On the other hand, a large portion of rDNAs were visualized as concentrated signals in active NORs (Figure 2C) and $\sim 71\%$ of rDNAs outside of FC/DFC units were co-localized with heterochromatin protein HP1 (Figure 2E), a well-established marker for inactive rDNAs in mammalian cells (Grummt, 2007; Li et al., 2013; Santoro et al., 2002; Yuan et al., 2007). As a control, the actively transcribed *SPA1* locus associated with Prader-Willi syndrome (Wu et al., 2016) was hardly coated by HP1 (Figure S2A), excluding staining artifacts of HP1. Together, these images suggest that active NORs contain transcriptionally inert rDNAs (Figures 2C–2E), supporting the earlier observation of discontinuously transcribed rDNA clusters in rDNA spreads (McKnight and Miller, 1976).

To address how many copies of active rDNAs are wrapped within each FC/DFC unit, we first applied stochastic optical reconstruction microscopy (STORM) to visualize RPA194 distribution within the FC/DFC units as one way to indicate active rDNAs. At 20 nm resolution by STORM (Rust et al., 2006), RPA194 was found to form 2 or 3 fiber-like patterns in each FC (Figure 2F). Such a Pol I arrangement is consistent with the previous reconstruction of EM images (Cheutin et al., 2002).

To confirm the STORM observation, we calculated active copies of rDNAs per cell. Quantification of rDNAs in all FC/DFC units by their fluorescent intensity (Figure S2B) and the copy number of rDNAs per cell (Figure S2C) revealed that, on average, each HeLa cell contains 821 rDNA copies, of which 42% are active (Figure 2G). Taking the average number of 136 FC/DFC units per HeLa cell (Figure 1I) into consideration, one FC/DFC unit contains ~ 2.5 copies of rDNA (345/136) for transcription (Figure 2G).

PFs Are Assembled into Clusters for Pre-rRNA Processing in the DFC

Pol I transcription is rapid and continuous (Albert et al., 2012; Coimai, 1999), and newly transcribed pre-rRNAs are moved into the DFC for processing (Watkins and Bohnsack, 2012). Presumably, pre-rRNAs must be efficiently processed by PFs in the DFC to coordinate with continuous Pol I transcription at the FC/DFC border. Thus, we asked how PFs are distributed. There are two possible PF distribution patterns: uniform or highly organized.

(F) FC/DFC units are associated with 47S pre-rRNAs. $>85\%$ of 5' ETS-1 signals localize within FC/DFC units that are representative of 47S pre-rRNAs (right, $n = 20$ cells, center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data).

(G) All detectable RPA194 is surrounded by FBL ($n = 20$ cells); a representative image is shown in (F).

(H) RPA194 is enriched at the edge of each FC, as shown by mEGFP-RPA194 KI live HeLa cells under SIM. See also Figures S1H and S1I.

(I) Different cells contain various sizes of FC regions and numbers of FC/DFC units. Volumes of individual FC regions and numbers of their numbers were measured and plotted. Each dot represents one cell, $n > 25$ cells for each cell type.

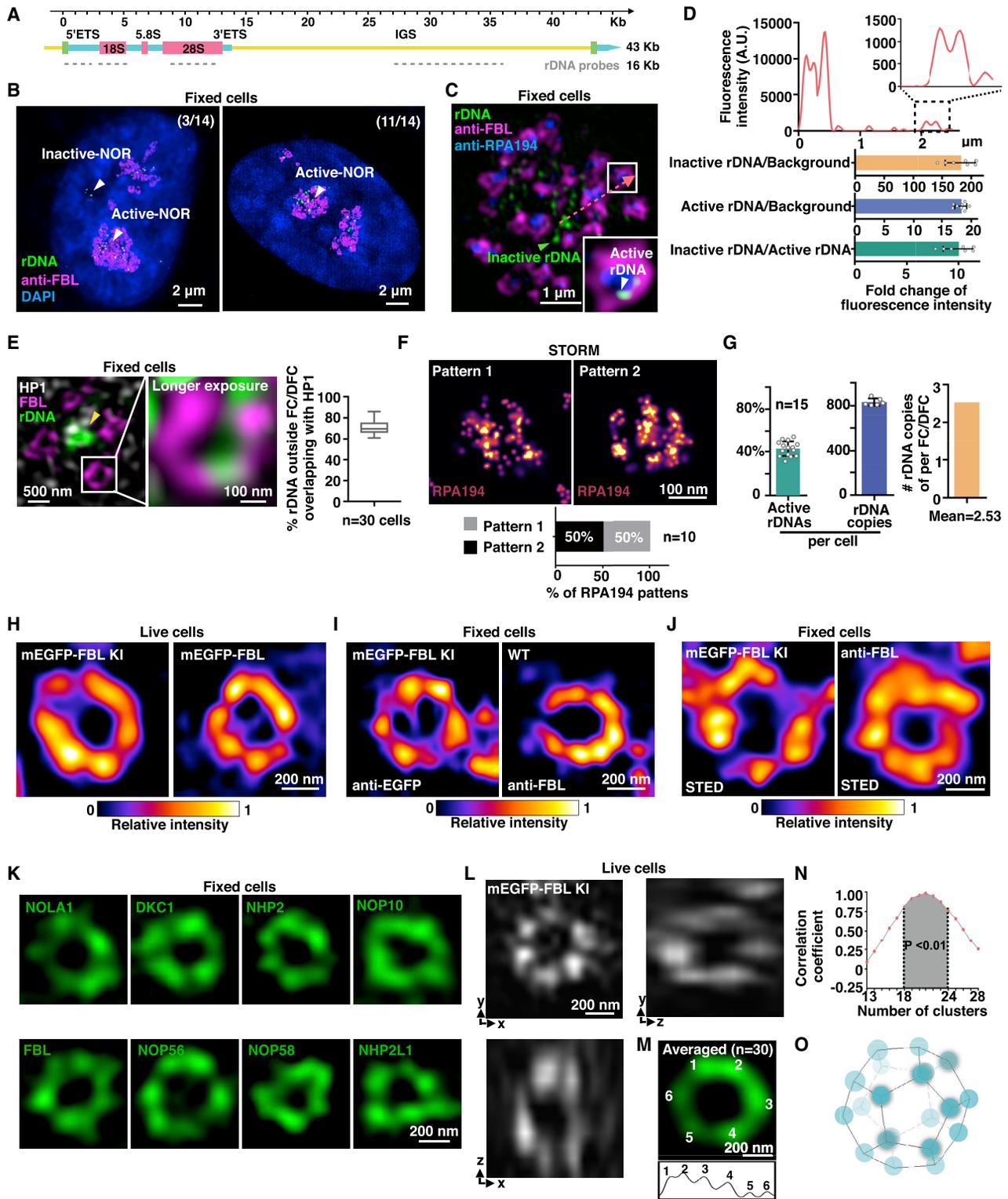


Figure 2. Organization of rDNAs at the FC/DFC Border and Pre-rRNA PFs in the DFC

(A) A schematic to illustrate rDNA probes that recognize human rDNAs used in DNA FISH.

(B) Organization of rDNA clusters in active and inactive NORs. Representative z-projected SIM images of rDNA and FBL HeLa cells are shown. 3 out of 14 examined cells displayed inactive NORs (left).

(legend continued on next page)

Live cells with mEGFP knocked in to the FBL N terminus (Figure 1A) or expression of mEGFP-FBL both revealed that FBL exhibited cluster-like patterns in the DFC on the x-y scale by SIM (Figure 2H). These observations were confirmed using direct antibody staining and stimulated emission depletion (STED) images (Figures 2I and 2J). Seven additional PFs (NOP56, NOP58, NHP2L1, NOLA1, DKC1, NHP2, and NOP10), which are key components of small nucleolar RNPs (snoRNPs), also exhibited similar patterns in the DFC (Figure 2K). As expected, FBL clusters were at least partially co-localized with NHP2L1 and DKC1 in dual-color fluorescent protein knockin (KI) cell lines (Figures S2D and S2E).

To gain more spatial details of PFs in the DFC, we analyzed the three-dimensional (3D) localization of FBL and found that FBL exhibited cluster-like patterns in the DFC (Figure 2L). The max-cross section presented 6 clusters on the x-y scale by both manual counting (Figure S2F) and average calculation of these max-cross section samples computationally (Guizar-Sicairos et al., 2008; Shi et al., 2017) (Figure S2G). The detached distance between two adjacent PF clusters was ~ 180 nm (Figures 2M and S2H).

The axial resolution is much lower than the horizontal in microscopy (Figures S1A and S2I). To solve the z-scale limitation, we constructed multiple hollow sphere models *in silico* with randomly distributed 12–30 small clusters and mimicked the SIM imaging process *in silico* (Figure S3A), followed by analysis of the likelihood of FBL cluster assembly in 3D between real SIM data and *in silico* SIM (Figure S3B). These analyses revealed a model with 18–24 clusters of PFs in the DFC surrounding the FC (Figures 2N, 2O, and S3C), among which an assembly pattern with 20–22 clusters exhibited the highest probability and the lowest error compared to SIM images (Figures S3C and S3D).

Why might it be important to form individual clusters of PFs at the DFC? One speculation is that forming clusters can increase PF local concentration to accelerate pre-rRNA processing. We sought to construct a simulation model to examine this hypothesis. Using parameters acquired by stoichiometric quantification (Figures S4A–S4E), we constructed a model based on the

Metropolis Hastings algorithm (Gršins and Mazets, 2014) (Figures S4F), and optimal efficiency was obtained when PFs were arranged into multiple clusters (Figure S4G; Video S1), consistent with observations in Figures 2L–2O and S3.

A human nucleolus consists of dozens of FC/DFC units that are formed around the rDNA tandem repeats (Figures 1 and S1). About 2–3 copies of active rDNAs are located at the border of each FC/DFC, where Pol I complexes are located and Pol I transcription occurs (Figures 2A–2G). Pre-rRNA PFs, including FBL, form 18–24 clusters that are further assembled into the DFC (Figures 2H–2O and S2–S4). On average, each spherical PF cluster is ~ 133 nm in diameter and each DFC region contains ~ 628 nm outer and ~ 362 nm inner diameters (Figure S2H).

Movement of Nascent 47S Pre-rRNA from the FC/DFC Border to DFC

The DFC is where co-transcriptional pre-rRNA processing takes place (Figure S5A) (Barandun et al., 2018). Although Pol I transcription occurs at the border of the FC/DFC (Figures 2B–2F), nascent pre-rRNAs were detected in the DFC (Figure 1F). Using two sets of single molecule RNA FISH (smFISH) probes, targeting upstream and downstream of the first cleavage site (O1 site) of pre-rRNA (Mullineux and Lafontaine, 2012) (Figure 3A) showed that the first 414 nucleotides (nt) of 47S pre-rRNA (labeled as 5' ETS-1) was mainly localized in the DFC (Figures 3B, 3C, S5B, and S5C), while the following 389 nt (498 nt to 977 nt) of 47S pre-rRNA (labeled as 5' ETS-2) was largely detected at the FC/DFC border (Figures 3B, 3C, S5B, and S5C). These observations suggest that, while pre-rRNA is still being transcribed, the terminus of its 5' ETS has been translocated to the DFC (Figures 3B and 3C). What mechanism accounts for such co-transcriptional nascent pre-rRNA sorting?

FBL Regulates the Directional Sorting of Nascent Pre-rRNA in FC/DFCs

We speculated that depletion of factors involved in pre-rRNA sorting would result in trapped 5' sequences of 47S pre-rRNAs at the FC/DFC border, shown as increased signals with FC

(C) Active NORs contain both active and inactive rDNAs. Green arrow, inactive rDNAs; white arrows, active rDNAs at the border of the FC/DFC and shown in the zoomed-in image with longer exposure time; pink line, used to plot the signal-to-background ratio (SBR) in (D).

(D) SBR of DNA FISH. The non-FC/DFC unit-localized rDNAs are highly concentrated compared to those within the FC/DFC units. Mean \pm SD are shown. SBR of active and inactive rDNAs are shown by plotting the pink line in (C).

(E) An active NOR contains both active and inactive rDNAs. Yellow arrow, co-localization of rDNAs with HP1; magenta, DFCs are shown by FBL. The zoomed-in image with longer exposure shows the FC/DFC engaged with active rDNAs. $\sim 71\%$ of rDNAs outside of the FC/DFCs were co-localized with HP1 (right, center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data). See also Figure S2A.

(F) A representative single layer STORM image shows that one FC/DFC contains two patterns of RPA194 clusters (top). Statistical results were calculated from 10 randomly selected STORM images in HeLa cells (bottom).

(G) Measurement of active rDNA copies per FC/DFC unit per HeLa cell. Mean \pm SD are shown. See also Figures S2B and S2C.

(H) FBL exhibits a clustered distribution pattern in the DFC. Representative SIM images of mEGFP-FBL KI cells (left) and exogenously expressed mEGFP-FBL in live HeLa cells (right).

(I) FBL exhibits clusters in the DFC. FBL patterns were confirmed with anti-mEGFP in fixed mEGFP KI cells and anti-FBL in fixed wild-type cells.

(J) FBL clusters shown by STED images. The clustered FBL patterns were confirmed with fixed mEGFP KI cells and anti-FBL in fixed wild-type cells.

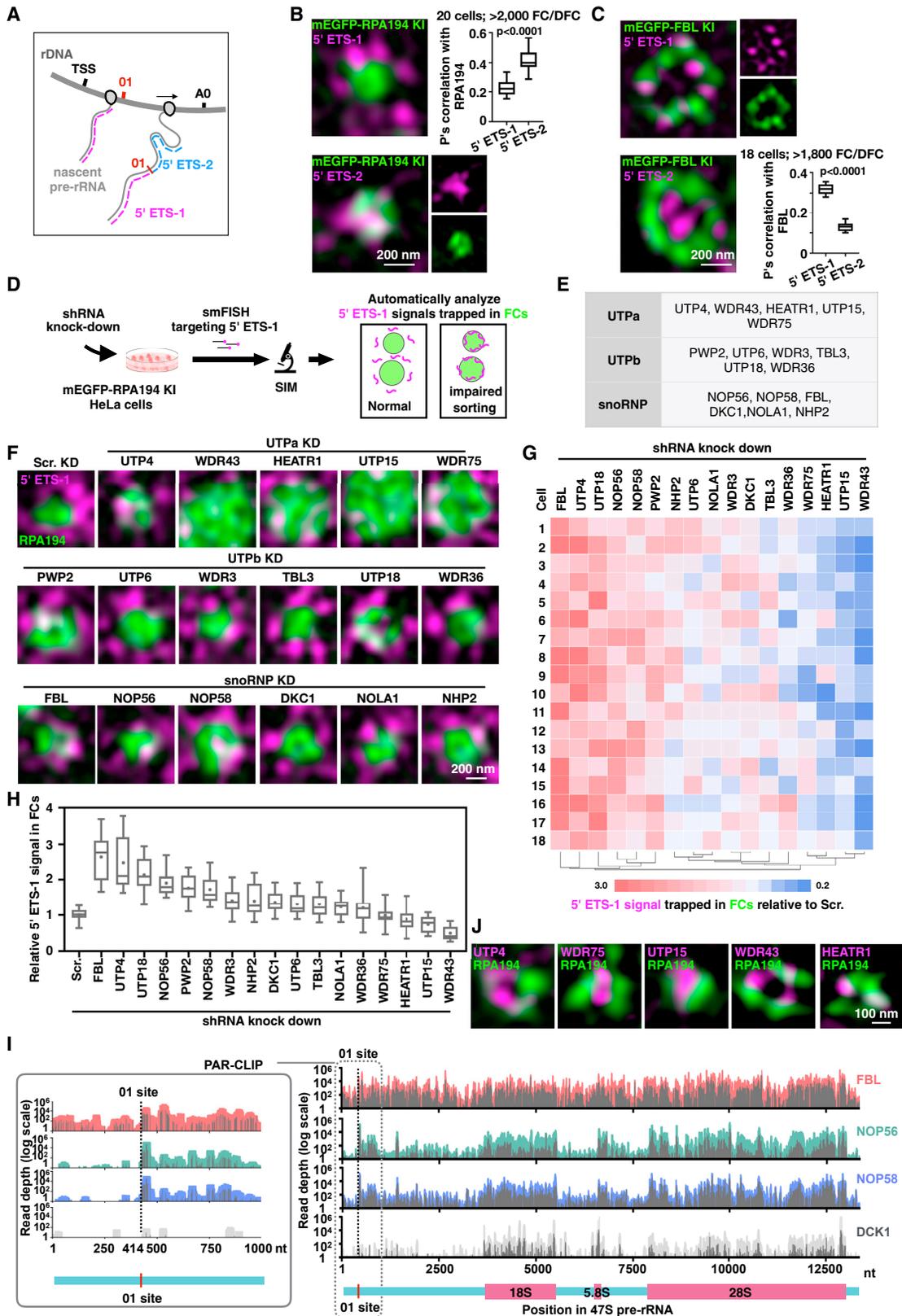
(K) Box H/ACA and Box C/D snoRNPs exhibit cluster-like distribution patterns in the DFC in HeLa cells. See also Figures S2D and S2E.

(L) FBL proteins are assembled into cluster-like distribution in 3D in the DFC, shown by representative live cell 3D-SIM images of mEGFP-FBL KI HeLa cells.

(M) Cross-correlation of aligned and averaged images shows that the max-cross sections of DFCs contain six FBL clusters. The intensity distribution of FBL was measured (bottom). See also Figures S2F–S2H.

(N) 18–24 clusters of PFs are presented in the DFC surrounding the FC. The p value is significant when the cluster numbers are 18–24. See also Figures S2I, S3A–S3D, and S4G.

(O) 3D model of the clustered FBL distribution in the DFC.



(legend on next page)

(Figure 3D). To identify such proteins, we performed screening assays by knocking down human homologs of yeast small subunit (SSU) processomes, including UTPa, UTPb, and snoRNP complexes (Figure 3E), which are involved in the early steps of eukaryotic ribosome biogenesis (Barandun et al., 2018) (Figure S5A), followed by evaluating the colocalization between 5' ETS-1 and RPA194 using effective short hairpin RNA (shRNA; Figure S5D) by quantitative SIM imaging (Figures 3D–3H; see also STAR Methods and Figure S6 for computational analysis). This screening showed that FBL knockdown (KD) resulted in the highest co-localization signal between RPA194 and 5' ETS-1, followed by the loss of UTP4; however, KD of other UTPb and U3 snoRNP components had only modest effects (Figures 3F–3H). Interestingly, KD of many UTPa proteins resulted in abnormal distribution of RPA194 (Figure 3F, top row), consistent with the notion that human orthologs of the yeast UTPa complexes can affect rDNA transcription (Prieto and McStay, 2007).

The strong effect on the nascent 47S pre-rRNA localization upon FBL depletion (Figures 3F–3H) prompted us to examine the binding preference between FBL and the 5' end of 47S pre-rRNA. Photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) assays (Kishore et al., 2013) revealed that FBL exhibited the strongest association to 5' ETS-1 of 47S pre-rRNA, compared to NOP56, NOP58, and DKC1 (Figure 3I), indicating an additional function of FBL in nascent pre-rRNA sorting beyond its classical role in U3 snoRNP.

Although KD of UTP4 resulted in mis-colocalization of 5' ETS-1 (Figures 3F–3H), all examined UTPa proteins were localized within FCs or the border of FC/DFCs (Figure 3J), thus unlikely accompanying nascent 5' ETS-1 because it travels to the DFC. Collectively, these data lead us to conclude that FBL plays a key role in mediating nascent 47S pre-rRNA transit from the border of FC/DFCs to DFCs.

Both GAR and MD of FBL Are Required for Nascent Pre-rRNA Sorting and Processing in the FC/DFC

How does FBL regulate the 47S pre-rRNA sorting in FC/DFC units? The N-terminal glycine- and arginine-rich (GAR) domain

of FBL contains IDRs; its C-terminal region contains a methyltransferase domain (MD) (Figure 4A). Examination of pre-rRNA processing showed aberrant accumulation of 47S and 34S pre-rRNAs, accompanied by reduced 28S and 18S rRNAs in FBL KD cells (Figures 4B, S5E, and S5F). Of note, 34S pre-rRNA intermediates are typically formed upon FBL depletion (Tafforeau et al., 2013). RNA FISH showed that 47S pre-rRNAs mainly surrounded the RPA194-labeled FC in normal cells, while the 5' ETS-1 probe-labeled 47S and 34S pre-rRNAs were either associated with RPA194 or exhibited dispersed nucleolar patterns in FBL KD cells (Figure 4C).

The pre-rRNA processing mediated by FBL is unlikely to require the methyltransferase activity of FBL. To support this, the production of 34S, 28S, and 18S rRNAs could be almost fully rescued by either wild-type FBL or the methyltransferase null (Aittaleb et al., 2004; Defrasnes et al., 2016) mutants (Figures 4B and S5F). Correspondingly, aberrant 47S pre-rRNA localization in the nucleoli was also rescued with methyltransferase null FBL mutants (Figure 4C). In contrast, introducing either MD or GAR domain to FBL KD cells could not rescue 34S pre-rRNA production (Figure 4B), suggesting important roles of both domains of FBL in pre-rRNA processing.

The MD Is Required for Nascent Pre-rRNA Interaction with FBL in the FC/DFC

We next examined the roles of MD and GAR domains in pre-rRNA sorting and processing. Intriguingly, the MD, but not the GAR domain, appeared to be sufficient to constrain both 47S and 34S pre-rRNAs within FC/DFC units (Figure 4C, panels 7 and 8), indicating that binding to 5' ETS-1 of pre-rRNA by the MD of FBL could ensure pre-rRNA localization to FC/DFCs (Figure 4C, panel 7), although processing cannot be fully executed there (Figure 4B). Consistently, introducing GAR into cells led to dispersed localization throughout cells, whereas MD and FBL-full-length (FL) showed DFC localization (Figure S5G). Since GAR could not bind to pre-rRNA, such dispersed localization of GAR reflects the requirement of nascent pre-rRNA in maintaining FBL sub-nucleolar localization. Detailed imaging revealed that MD was no longer localized to DFCs, but it was found inside

Figure 3. Identification of FBL in Regulating Nascent Pre-rRNA Sorting in the FC/DFC

- (A) A schematic of RNA smFISH probes to detect transcribing pre-rRNAs. See also Figures S5A–S5C.
- (B) The 5' ETS-1 probe-detected 47S pre-rRNAs (top) are largely distributed outside of the FC, while the 5' ETS-2 probe-detected pre-rRNAs (bottom) are mainly located at the FC/DFC border, shown by SIM.
- (C) The 5' ETS-1 probe-detected 47S pre-rRNAs (top) are largely distributed in the DFC, while the 5' ETS-2 probe-detected pre-rRNAs (bottom) are mainly located at the FC/DFC border, shown by SIM.
- (B and C) The histogram is the Pearson correlation coefficient (P's) from 20 cells (center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data). See also Figures S5B and S5C. See also Figure S5B and S5C.
- (D) A screening schematic of factors involved in 47S pre-rRNA sorting. See also Figures S5A and S6 and STAR Methods.
- (E) List of candidates in screening assays.
- (F) Representative SIM images of 5' ETS-1 (magenta) and RPA194 (green) from cells that were treated with shRNAs individually targeted to proteins in (E). See also Figures S5D and S6.
- (G) Heatmap showing the impaired nascent pre-rRNA sorting after KD of individual factors shown in (E) and (F). The intensity of 5' ETS-1 signals trapped in the FCs of each KD relative to the scramble group is labeled underneath. Each colored square represents one cell. 18 cells of each KD condition were imaged and analyzed.
- (H) A boxplot shows statistic results presented in (G). Center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data; dots, mean.
- (I) FBL, NOP56, NOP58, and DKC1 PAR-CLIP assays show that the 1~414 nt of 5' ETS of 47S pre-rRNAs prefer to interact with FBL. The dark gray bar represents the T-to-C transition.
- (J) Representative SIM images of UTPa proteins (magenta) and RPA194 (green) in fixed HeLa cells.

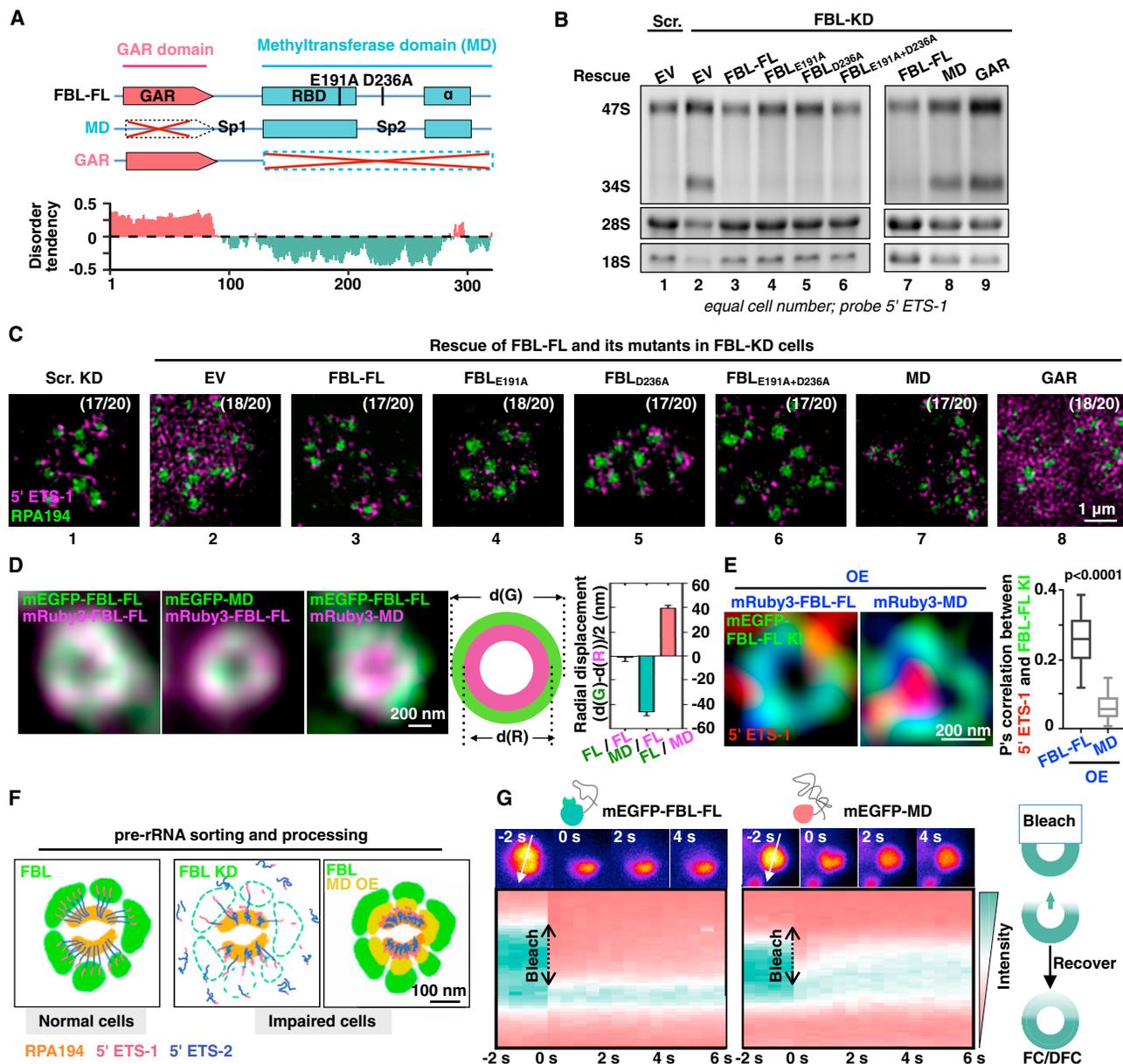


Figure 4. Both GAR and MD Domains of FBL Are Required for Nascent Pre-rRNA Sorting and Processing in the FC/DFC

(A) FBL truncations, methyltransferase null point mutations, and the disorder tendency of FBL amino acids calculated by IUPred2A (Mészáros et al., 2018) are shown.

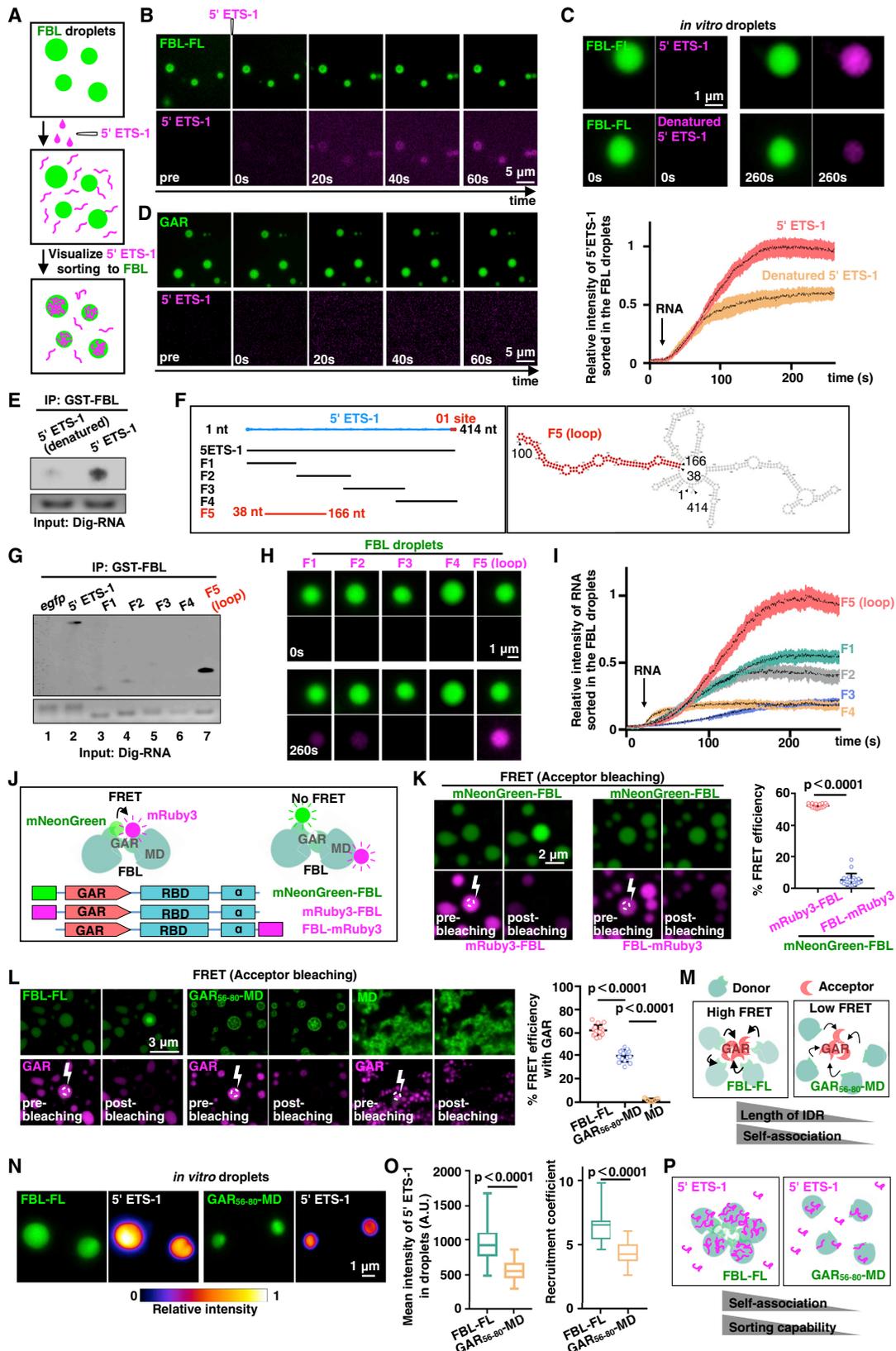
(B) Examination of 47S pre-rRNA processing defects by northern blots (NB) in FBL KD cells rescued with different FBL mutants. EV, empty vector. See also Figure S5F for longer exposure.

(C) FBL is required for nascent 47S pre-rRNA sorting to the DFC. Representative SIM images and statistics (the representative pattern/total number of imaged cells) are shown in the upper right for each condition.

(D) The MD is located within the FBL-FL-formed DFC. The radial displacement results were from 40 FC/DFC units under each condition. Mean \pm SEM are shown. (E) The 5' ETS-1 of 47S pre-rRNAs (red) are trapped at the border of the FC/DFC in MD overexpression (OE) HeLa cells (middle panel), compared to the DFC localization in FBL-FL OE cells (left panel). Right, statistical results of P's between 5' ETS-1 and mEGFP-FBL-FL ($n > 30$ FC/DFC units of each condition, center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data).

(F) A schematic showing that FBL plays a key role in 47S pre-rRNA sorting.

(G) The mobility of FBL-FL and the MD in one FC/DFC unit examined by FRAP (illustrated on right). Kymograph was recorded along the white arrow in the upper left panels. One representative of FC/DFC unit was shown. >10 cells were measured.



(legend on next page)

FBL-FL (Figure 4D) with an ~ 40 nm smaller radius than regular DFCs (Figure 4D). As expected, increased 47S pre-rRNA accumulation was observed at the FC/DFC border in mEGFP-FBL-FL KI cells with overexpressed MD (Figure 4E).

The observation that the MD itself fails to direct pre-rRNA migration to the DFC (Figures 3I, and 4C–4F) suggested that the GAR domain might be associated with the sorting of nascent pre-rRNAs toward the DFC for processing. To gain insight into this possibility, we examined the kinetics of FBL-FL and MD by fluorescence recovery after photobleaching (FRAP). Strikingly, MD exhibited much higher liquid-like dynamic movement (with $t_{1/2} \sim 1$ s) than FBL-FL (Figure 4G), indicating that the disordered GAR domain could increase the solid-like feature of FBL and the FBL-bound 5' ETS of 47S pre-rRNA to form clusters at DFCs (Figures 2H–2O).

GAR Domain-Mediated FBL Self-Association Promotes Pre-rRNA Sorting

Multivalent interactions between IDRs can promote self-association of IDR proteins, driving phase-separated droplet assembly (Banani et al., 2017; Shin and Brangwynne, 2017; Tatarakis et al., 2017). One way to promote the transport of MD-interacting 47S pre-rRNA from the FC/DFC border to the DFC (Figures 3 and 4) would be via self-association of FBL GAR domains. We set up assays to test this possibility (Figure 5A). Consistent with previous studies (Feric et al., 2016), purified FBL (Figure S7A) formed phase-separated droplets *in vitro* (Figure 5B). The addition of structured 1–414 nt 5' ETS-1 sequences led to 5' ETS-1 sorting to FBL droplets within 10 s (Figure 5B; Video S2), and saturated signals appeared at ~ 200 s (Figure 5C; Video S3). As controls, the same 5' ETS-1 barely moved into GAR droplets lacking RNA binding capability (Figure 5D; Video S4). Remarkably, this sorting required correctly folded RNAs because we observed that unstructured 5' ETS-1 could be hardly captured by FBL droplets (Figure 5C; Video S3). This finding was consistent with *in vitro* binding assays that FBL did not bind to unstructured 5'

ETS-1 (Figure 5E). Mapping between FBL and fragments of 5' ETS-1 (Figure 5F, left panel) revealed that FBL strongly interacted with a stem-loop region formed by nt 38–166 (Figure 5F, right panel, and Figure 5G, lane 7), compared to other examined fragments (Figure 5G, lanes 1 and 3–6). Importantly, this stem-loop structure had the highest efficiency of sorting to FBL droplets *in vitro* (Figures 5H and 5I; Video S5).

How is 5' ETS-1 specifically sorted to FBL droplets? One hypothesis is that FBL-bound 5' ETS-1 is locally concentrated via self-association of FBL GAR domains. If so, a positive correlation between FBL self-association and 5' ETS-1 sorting capability should be observed. We performed fluorescence resonance energy transfer (FRET) assays to examine whether FBL self-association is dependent on its GAR domain *in vitro* (Figure 5J). As shown in Figures 5J and 5K, FRET only appeared between mNeonGreen-FBL and mRuby3-FBL, which were both fused to the N-terminal GAR domain of FBL, but not between mNeonGreen-FBL and FBL-mRuby3 (C-terminal fusion). These analyses revealed that the self-association of FBL is achieved via its GAR domain, which leads to FBL droplet formation *in vitro*.

Next, we asked how the GAR domain modulates the strength of FBL self-association. As the GAR domain contains IDRs, one possibility was that different lengths of IDRs would determine FBL self-association. Using the same FRET strategy (Figure 5L), we compared the self-association capability of the purified GAR domain (Figure S7A) with FBL, a truncated FBL that contains only one third the length of the GAR (GAR_{56–80}-MD), or the MD (Figure S7B). The FRET value between the GAR and FBL was much stronger than that between the GAR and GAR_{56–80}-MD (Figure 5L). As controls, no FRET value could be detected between the purified GAR domain and purified MD (Figure 5L). Together, these results (Figures 5J–5L) suggest that the GAR domain determines FBL self-association, which is modulated by its IDR length (Figure 5M).

Finally, we asked whether the addition of 5' ETS-1 to droplets formed by FBL or GAR_{56–80}-MD with different self-association

Figure 5. FBL Self-Association via GAR Correlates with Pre-rRNA Sorting Capability

(A) *In vitro* RNA sorting assays to visualize 5' ETS-1 sorting to FBL.

(B) Cy3-labeled 5' ETS-1 (magenta) is sorted to the mNeonGreen-FBL-FL droplets (green) *in vitro*. A representative time series of micrographs are shown. See also Video S2.

(C) The 5' ETS-1 sorting to mNeonGreen-FBL-FL droplets requires folded RNAs (top), but not denatured RNAs (bottom). $n > 10$ droplets, mean \pm SD are shown. See also Video S3.

(D) Cy3-labeled 5' ETS-1 (magenta) cannot be captured by the droplets formed by mNeonGreen-GAR (green) *in vitro*. See also Video S4.

(E) FBL specifically interacts with structured (lane 2) but not denatured 5' ETS-1 (lane 1).

(F) A schematic to show 5' ETS-1 fragments used in *in vitro* binding assays and *in vitro* sorting assays. The structure of 5' ETS-1 is predicted by RNAfold (Gruber et al., 2008).

(G) FBL specifically interacts with 38–166 nt of 5' ETS-1. An equal molar amount of Dig-labeled RNA was used in each binding assay.

(H) Stem-loop structure formed by 38–166 nt of 5' ETS-1 is crucial for its sorting to FBL-FL droplets *in vitro*. See also Video S5.

(I) Kinetics of (H). $n > 10$ droplets, mean \pm SD are shown.

(J) A schematic to illustrate detection of GAR domain-mediated FBL self-association by FRET assays.

(K) FBL self-association is controlled by GAR domain. $n > 15$ individual FRET assays for each examined pair (Student's *t* test). Mean \pm SD are shown.

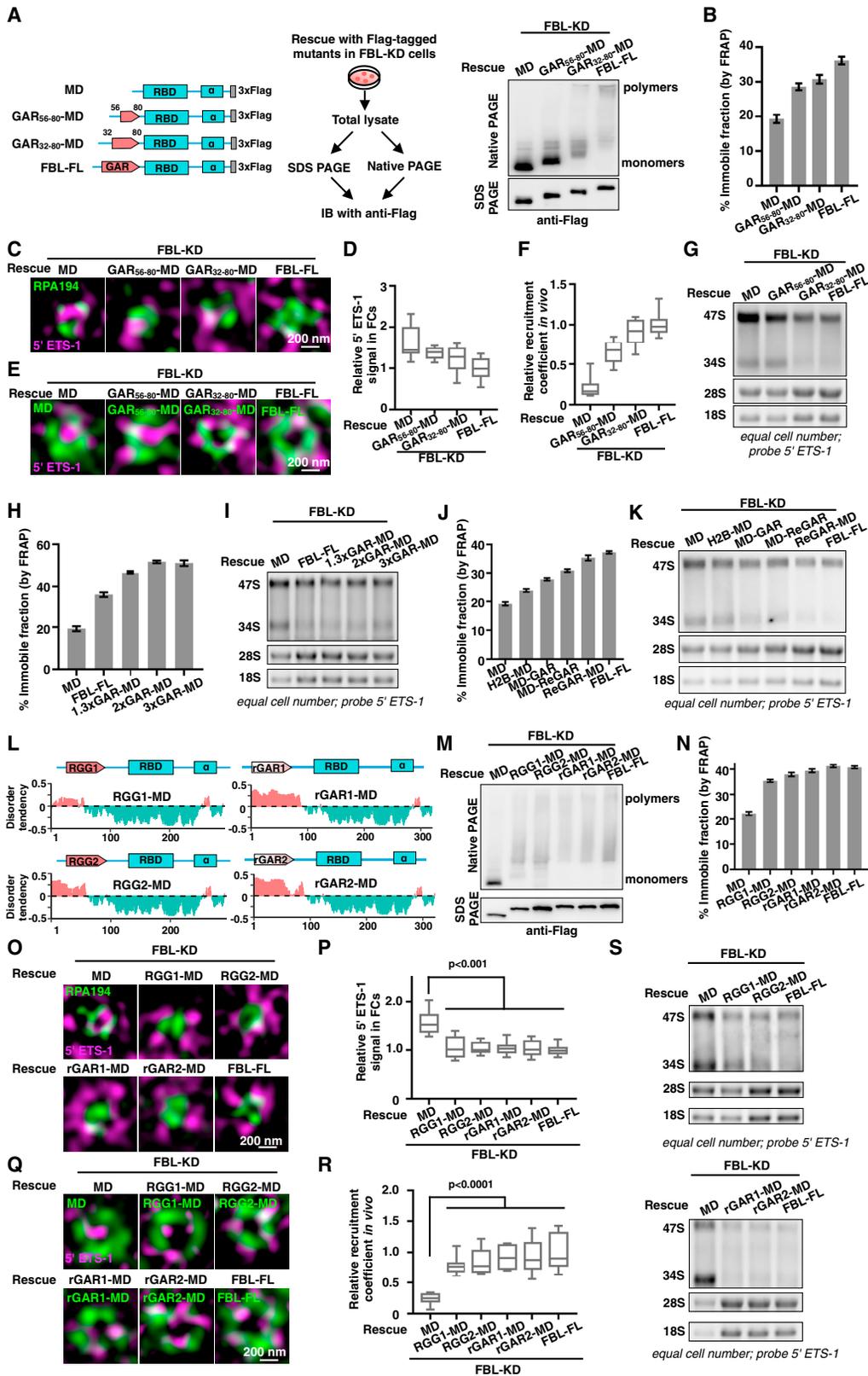
(L) The length of the GAR domain correlates with FBL self-association strength, shown by FRET assays. $n > 15$ individual FRET assays for each examined pair (Student's *t* test). Mean \pm SD are shown. See also Figure S7B.

(M) An illustration summarizes results obtained from (J)–(L) that the IDR length positively correlates with the strength of FBL self-association.

(N) Reduced length of GAR domain in FBL results in the decreased sorting capability of 5' ETS-1. Representative micrographs are shown. See also Figure S7C.

(O) Statistics of the mean intensity recruitment coefficients of 5' ETS-1 in droplets are shown. $n = 50$ droplets ($n = 50$ droplets, center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data). See also Figure S7D.

(P) An illustration summarizes results obtained from (N) and (O) that the strength of FBL self-association positively correlates with the 5' ETS-1 sorting capability to *in vitro* FBL droplets.



(legend on next page)

capabilities (Figures 5J–5M) could result in a sorting difference. The average fluorescence intensity and recruitment coefficient of 5' ETS-1 were markedly reduced in GAR_{56–80}-MD droplets compared to those in FBL droplets (Figures 5N and 5O). A smaller recruitment coefficient indicated less concentrated RNA in preformed droplets (Langdon et al., 2018), ruling out any influence from the size of FBL droplets. Considering the average size of FBL droplets was >2-fold larger than that of GAR_{56–80}-MD droplets (Figure S7C), the total intensity of 5' ETS-1 in FBL droplets was even higher than that in GAR_{56–80}-MD droplets (Figure S7D). Collectively, these assays support the view that FBL self-association via its GAR domain promotes pre-rRNA sorting *in vitro* (Figure 5P).

The IDR in the GAR Domain Immobilizes FBL for 47S Pre-rRNA Sorting to the DFC

To examine the importance of GAR domain-mediated pre-rRNA sorting *in vivo*, we introduced FBL mutants with different lengths of IDRs to FBL KD cells, followed by native PAGE to detect self-associated proteins as polymers (Figure 6A). Consistent with *in vitro* findings (Figures 5J–5M), increased IDR length in the GAR promoted FBL-self-association in cells, revealed by enhanced FBL self-association as polymers on native gels with increased IDR length in FBL mutants (Figure 6A).

Next, we used FRAP assays to examine the association kinetics (Crosby et al., 2013) of FBL and its mutants (Figure 6B and S7E) in cells. The highest association kinetics in individual FC/DFC units was FBL, followed by GAR_{32–80}-MD and GAR_{56–80}-MD. The MD alone exhibited the lowest association kinetics (Figures 6B and S7E), confirming that reduced IDR length leads to decreased self-association in cells.

We employed a number of assays to determine whether the reduced self-association of FBL mutants would cause impaired nascent 47S pre-rRNA sorting and processing in FC/DFCs. First, the reduced self-association capability of FBL mutants led to increased 5' ETS-1 accumulation on RPA194, with the highest

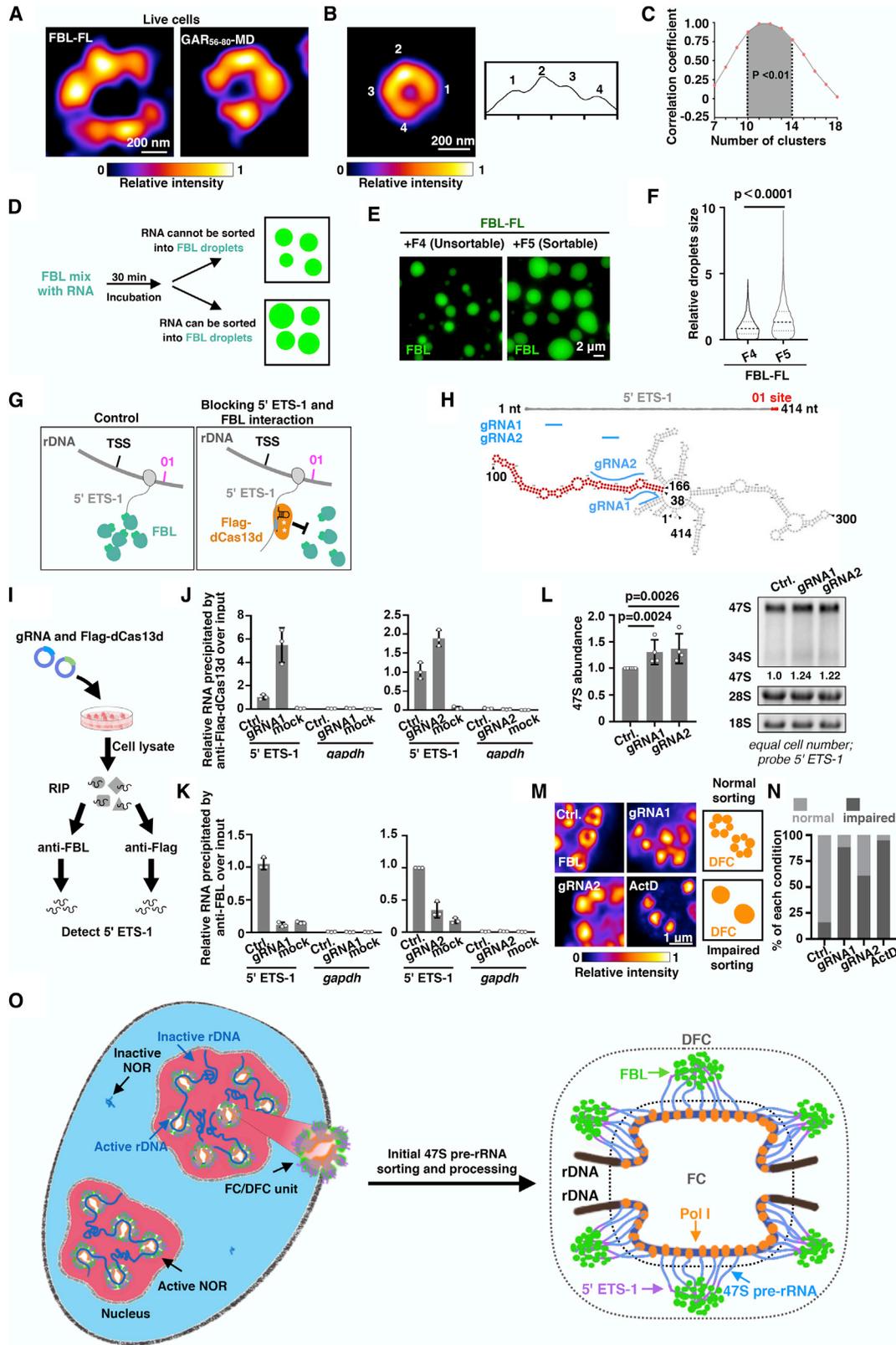
accumulation by MD alone (Figures 6C and 6D). Second, the reduced self-association capability of FBL mutants led to decreased 5' ETS-1 signals in different FBL mutant-formed “DFC” regions (Figures 6E and 6F). Third, increased 5' ETS-1 sorting to the DFC (Figures 6C–6F) by augmented GAR self-association (Figures 6A and 6B) led to increased rescue of 47S pre-rRNA processing (Figure 6G). Since all mutants contain the intact MD required for pre-rRNA binding, the abnormal 47S pre-rRNA accumulation was likely due to sorting defects by the impaired self-association of GAR domains (Figures 6A–6G). We also examined the effects of FBL mutants with extended GAR domains. Interestingly, increased GAR domains (Figure S7F) led to augmented self-association of FBL as expected, but it reached a plateau at 2×GAR (Figures 6H and S7G). FBL mutants with extended GAR domains did not further promote nascent pre-rRNA processing (Figure 6I), indicating that the natural length of the GAR domain in FBL is optimal for pre-rRNA sorting and processing.

We also constructed several chimeric FBL mutants that contained reversed GAR sequences or a reversed order of GAR fused to the MD (Figures S7H and S7I) and introduced them individually to FBL KD cells. As expected, pre-rRNAs were sequestered within FC/DFC units in all conditions, because each mutant contains the MD (Figure S7J). Compared to the MD alone, mutants containing multiple types of GAR domains, which individually exhibited a low-sequence complexity (Figure S7I) and high self-association (Figures 6J and S7K), led to increased pre-rRNA processing in FBL KD cells (Figure 6K). Thus, the disordered property, rather than the orientation of GAR per se is important for FBL self-association-mediated pre-rRNA sorting. In contrast, a mutant containing a high complexity sequence from H2B (Figure S7I) with low self-association (Figures 6J and S7K) resulted in the same high level of aberrant 34S pre-rRNA accumulation as by expression of the MD alone (Figure 6K).

Finally, to confirm that FBL phase separation via GAR domain self-association, rather than by specific sequence and/or

Figure 6. The Immobility of FBL Correlates with 47S Pre-rRNA Sorting and Processing

- (A) The IDR length in the GAR domain promotes FBL self-association in cells. Left, a workflow shows the determination of FBL self-association *in vivo*. Right, increased GAR domain length in FBL mutants led to augmented FBL self-association shown by increased multimerization.
- (B) Increased IDR length in the GAR domain immobilized FBL mutants in cells. See also Figure S7E.
- (C and D) Increased IDR length in the GAR domain of FBL mutants reduced 5' ETS-1 signal trapped in FC (C). 20 cells were analyzed under each condition by boxplot (D).
- (E and F) Increased IDR length in the GAR domain of FBL mutants promotes sorting ability of pre-rRNAs (E). 20 cells were analyzed under each condition by boxplot (F).
- (G) Increased IDR length in the GAR domain of FBL positively correlates with proper 47S pre-rRNA processing, shown by NB. See Figure 4B for detail.
- (H) FBL mutants with the extended length of GARs (Figure S7F) display different immobility in FC/DFC units. See also Figure S7G.
- (I) FBL mutants with the extended length of GARs in FBL mutants did not further promote pre-rRNA processing, shown by NB. See Figure 4B for detail.
- (J) FBL mutants and FBL-FL have different immobility in FC/DFC units. ReGAR, reversed amino acid sequence of GAR. See also Figures S7H–S7K.
- (K) FBL mutants with increased immobility positively correlates with 47S pre-rRNA processing, shown by NB. See Figure 4B for detail.
- (L) The disorder tendency of different IDR mutants (RGG1-MD, RGG2-MD, rGAR1-MD, and rGAR2-MD) in the chimeric FBL mutants is shown in pink, whereas that of the MD is shown in green. See also Figure S7L.
- (M) FBL mutants shown in (L) and FBL-FL have similar self-association strength in cells.
- (N) FBL mutants shown in (L) and FBL-FL have similar immobility in FC/DFC units. See also Figure S7M.
- (O and P) FBL mutants shown in (L) and FBL-FL have a similar capability of pre-rRNA sorting (O). 20 cells were analyzed under each condition by boxplot (P).
- (Q and R) FBL mutants shown in (L) and FBL-FL have similar capability of pre-rRNA sorting (Q). 20 cells were analyzed under each condition by boxplot (R).
- (S) FBL mutants shown in (L) and FBL-FL largely rescue aberrant 47S pre-rRNA processing, shown by NB. See Figure 4B for detail.
- In (D), (F), (P), and (R), center line, median; box limits, upper and lower quartiles; whiskers, maximum or minimum of the data shown in each boxplot. In (B), (H), (J) and (N), relative immobile fraction was measured from the plateau value using one-phase association non-linear regression to regress the FRAP results in related supplemental figures; mean ± 95% confidence intervals are shown.



(legend on next page)

structure elements in GAR domains, causes pre-rRNA sorting and promotes processing, we replaced the GAR domain in FBL with two arginine-glycine-rich disordered domains (RGG1 and RGG2) from another protein, NOLA1, or with randomized GARs (rGAR1 and rGAR2) from the wild-type GAR of FBL (Figures 6L and S7L). These FBL mutants containing completely different IDRs could largely rescue FBL self-association (Figures 6M, 6N, and S7M) and 5' ETS-1 signal trapped in the FC (Figures 6O and 6P), promote nascent pre-rRNA sorting (Figures 6Q and 6R), and reduce aberrant pre-rRNA processing in FBL KD cells, just as wild-type FBL does (Figure 6S). As controls, fusion of two copies of mEGFP to FBL did not alter its mobility (Figure S7N), eliminating the concern that fluorescently tagging FBL might change its function. Collectively, these results reveal that GAR domain self-association confers the capability for FBL to form clusters, which can accompany nascent 47S pre-rRNAs in transit from the FC/DFC border to the DFC.

Nascent Pre-rRNA Sorting Is Required for DFC Assembly

It is known that Pol I transcription and the presence of pre-rRNAs are required for nucleolar formation (Falahati et al., 2016; Feric et al., 2016; Kopp et al., 2007). Interestingly, along with its reduced sorting capability (Figures 5K–5P and 6A–6G), GAR_{56–80}-MD formed fewer clusters in the DFC than FBL (Figures 7A–7C and S7O), indicating that the nascent pre-rRNA sorting process may facilitate DFC assembly.

Consistent with this idea, the addition of sortable 5' ETS-1 fragments (Figures 5F–5I) led to enlarged FBL droplets *in vitro* (Figures 7D–7F). We next applied the catalytically dead, RNA-guided, RNA-targeting clustered regularly interspaced short palindromic repeats (CRISPR) system Cas13d (dCas13d) and CRISPR guide RNAs (gRNAs) (Konermann et al., 2018) to block FBL binding to nascent pre-rRNA in cells (Figure 7G). Two gRNAs were designed to target the FBL-interacting loop region in 5' ETS-1 (Figure 7H). Transfection of Flag-dCas13d and gRNAs allowed Flag-dCas13d to target 5' ETS-1, shown by observations that Flag-dCas13d efficiently associated with 5'

ETS-1 (Figures 7I and 7J) and that FBL-associated 5' ETS-1 was reduced (Figure 7K), accompanied by measurable 47S pre-rRNA accumulation without processing being affected (Figure 7L). Remarkably, transfection of Flag-dCas13d and gRNAs led to retarded formation of individual DFC regions in these cells, similar to that mediated by treating cells with the Pol I inhibitor actinomycin D (Figures 7M and 7N). These results suggest that nascent pre-rRNA sorting is likely involved in DFC assembly.

DISCUSSION

Understanding how newly transcribed RNA sorting is achieved has remained challenging due to the low abundance of most nascent RNAs that usually undergo rapid processing. The relatively high abundance of nascent 47S pre-rRNA and its radial flux mode of processing in FC/DFCs (Boisvert et al., 2007; Farley et al., 2015) make the nucleolus an attractive model to address this fundamental question.

Recently developed super-resolution microscopy enables the observation and quantification of sub-cellular structures at the nanometer level (Schermelleh et al., 2019). Here, we describe nucleolar ultrastructures in live cells (Figures 1 and 2) and demonstrate that FBL self-association plays a key role in controlling the 5' terminus of nascent pre-rRNA sorting both *in vitro* and *in vivo* (Figures 3, 4, 5, and 6). Such a phase separation mechanism-controlled pre-rRNA sorting process is likely required for DFC nucleation (Figure 7). These findings represent a substantial advance in our understanding of how IDRs contribute to nascent pre-rRNA sorting and nucleolar spatial organization (Figure 7O).

Our data support the view that the initial step of ribosome biogenesis begins at the sorting step of newly transcribed 5' ends of 47S pre-rRNAs from the border of the FC/DFC to the DFC, which likely occurs prior to the processome assembly in the DFC (Figures 3 and S5) and is modulated by FBL (Figures 3, 4, 5, and 6). In addition to FBL, depletion of the UTPa complex also leads to impaired pre-rRNA sorting (Figures S5E–S5I). UTPa

Figure 7. Nascent Pre-rRNA Sorting Promotes DFC Formation

- (A) GAR_{56–80}-MD tends to form fewer clusters than FBL-FL in HeLa cells. Representative SIM images are shown.
- (B) Cross-correlation of aligned and averaged images shows that the max-cross sections of the GAR_{56–80}-MD contains 4 major clusters (left panel). The intensity distribution of GAR_{56–80}-MD was measured (right). See Figures S2G, 2M, and S7O for details.
- (C) Statistics of GAR_{56–80}-MD clusters in the max-cross section of DFCs. >75 DFCs were counted by SIM. See Figures S3 and 2N for details.
- (D) An assay to evaluate effects of 5' ETS-1 fragment sorting on sizes of FBL droplets.
- (E and F) The 38–166 nt of 5' ETS-1 (F5) enlarges FBL-FL droplets. Representative images are shown in (E) and statistic results of the relative droplet sizes of FBL-FL with unsortable (F4) and sortable (F5) fragments of 5' ETS-1 are shown in (F).
- (G) A schematic to illustrate that dCas13d is directed to 1–414 nt of 5' ETS (5' ETS-1) of 47S pre-rRNA by gRNAs (right), thereby presumably blocking the interaction between 5' ETS-1 and FBL (left).
- (H) Two gRNAs were designed to target the 38–166 nt loop in 5' ETS-1 pre-rRNA that specifically interacts with FBL. See also Figures 5F–5I.
- (I) A schematic to quantify gRNA/Flag-dCas13d-mediated FBL and 5' ETS-1 disruption by RNA immunoprecipitation.
- (J) Flag-dCas13d was efficiently associated with 5' ETS-1 shown by anti-Flag RNA immunoprecipitation (RIP). As a control, *Gapdh* was not enriched by anti-Flag.
- (K) FBL-associated 5' ETS-1 was reduced using Flag-dCas13d targeting 5' ETS-1, as shown by anti-FBL RIP. As a control, *Gapdh* mRNA was not enriched by anti-FBL.
- (L) Blocking FBL and 5' ETS-1 interaction by Flag-dCas13d/gRNAs led to an increased level of 47S pre-rRNAs. The abundance of 47S pre-rRNAs was detected by qRT-PCR and NB.
- (M) Blocking FBL and 5' ETS-1 interaction by Flag-dCas13d and gRNAs impaired the DFC formation. Representative normal and impaired DFC regions (shown by FBL) in HeLa cells treated with Flag-dCas13d/gRNAs or actinomycin D (ActD) (50 ng/mL) are shown.
- (N) Statistics of DFC patterns under each condition (>20 cells) in (M) are shown.
- (O) A model of FC/DFC organization and 47S pre-rRNA sorting in the human nucleolus. See text for details.
- In (J)–(L), all qRT-PCRs were performed from more than three biological replicates. Mean ± SD are shown. p values were calculated from Student's t test.

proteins are thought to bind to the very first part of the 35S pre-rRNA in yeast (Gallagher et al., 2004; Krogan et al., 2004; Pérez-Fernández et al., 2007), but the 5' ETS sequences between yeast and human are not conserved (Mullineux and Lafontaine, 2012). Thus, how UTPa proteins are involved in 47S pre-rRNA sorting and processing in human FC/DFCs warrants further study.

Different droplet surface tensions are important for nucleolus assembly (Feric et al., 2016), but how sub-nucleolar compartments are assembled in cells remains unexplored. The presence of PF clusters (Figures 2H–2J and S2H) suggests the co-existence of multiple phase-separated sub-droplets in the DFC (i.e., “multiple phases” within “one phase”) that, together, underlie the ultimate formation of individual DFCs. It has been shown that nucleoli in *Caenorhabditis elegans* embryos form only when the expression of FBL (and perhaps other key snoRNP components) is above a threshold concentration (Weber and Brangwynne, 2015). It is possible that higher concentrations of FBL and other pre-rRNA PFs are required to nucleate the one-phase DFC than those in the hierarchy assembly pattern of DFC that includes multiple phase-separated PF clusters (Figures 2H–2O and 7O). More broadly, multiple phase-separated sub-droplets are likely required for the assembly of other types of large, membrane-less cellular condensates.

Our data reveal that FBL plays a key role in directing the 5' end of 47S pre-rRNA to the DFC via an IDR-dependent mechanism (Figures 3, 4, 5, and 6). The binding and sorting of 47S pre-rRNA utilizes different FBL domains: the MD first binds to the 5' end of nascent pre-rRNA and the FBL-RNA complexes then move toward the DFC by GAR domain self-association (Figures 4, 5, and 6). Such pre-rRNA sorting strongly correlates with FBL self-association via IDRs (Figures 5N–5P and 6A–6S) and is required for pre-rRNA processing (Figures 6G, 6I, 6K, and 6S), thus revealing the important functional relevance of this sorting process. Finally, as many RBPs associated with nascent pre-mRNA processing events are known to contain structurally disordered regions (Banani et al., 2017; Gueroussov et al., 2017; Shin and Brangwynne, 2017; Ying et al., 2017), we speculate that a similar mechanism is likely used by the cell to keep other types of nascent RNAs from unnecessary or unwanted self-aggregation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Human Cell Lines
- METHOD DETAILS
 - Cell culture and cell transfection
 - Lentivirus infection and production and generation of stable cell lines
 - Plasmids construction
 - Protein expression and purification

- RNA isolation, RT-qPCR and Northern Blots
- Measurement of rDNA copy numbers
- Knock-in fluorescent protein by CRISPR/Cas9
- Protein visualization
- Single molecule RNA Fluorescent *in situ* Hybridization (smFISH)
- DNA Fluorescent *in situ* Hybridization (DNA FISH)
- Widefield Microscopy procedure
- Structured Illumination Microscopy (SIM) procedure
- Stimulated Emission Depletion Microscopy (STED) procedure
- Stochastic Optical Reconstruction Microscopy (STORM) procedure
- Screening factors involved in 47S pre-rRNA sorting
- Phase separation assay
- RNA-Protein *in vitro* binding assay
- *In vitro* RNA sorting assay
- Fluorescence Resonance Energy Transfer (FRET) assay
- Fluorescence Recovery After Photobleaching (FRAP) assay
- Native RNA Immunoprecipitation
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Measurement of FC number and size
 - Measurement of the proportion of active rDNA copies
 - Measurement of the proportion of nascent pre-rRNAs associated with FC/DFC units
 - Measurement of the 5' ETS-1 signal trapped in FCs
 - Measurement of the 5' ETS-1 signal sorted in FBL mutants
 - Quantification of RPA194 and FBL per FC/DFC unit
 - Measurement of RNA recruitment coefficient
 - Image alignment for FBL distribution in the max-cross x-y section
 - Digital simulation of FBL distribution in 3D
 - Metropolis Hastings strategy to search the optimal distribution of processing factors
 - Statistics and reproducibility
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.molcel.2019.08.014>.

ACKNOWLEDGMENTS

We are grateful to Gordon Carmichael and all Chen lab members for reading the manuscript. We acknowledge Yun-Xiao He from SIPPE, CAS; Fei Hao, Hai-Jia Yu, and Xin-Ping Zhu from GE Healthcare Life Sciences; You-Jun Chu from ShanghaiTech University; and Yan Wang from NCPP (Shanghai) for support on SIM; Yan Wang, Yi-Hui Tu, and Xue-Dong Wang from SIBCB, CAS for support on microscopy and FACS; Yuan-Yuan Su from Leica Microsystems (Shanghai) for support on Leica STED; Dr. Dan Li from Nikon instruments (Shanghai) for support on STORM; and Jun Chu from SIAT, CAS for providing mNeonGreen and mRuby3. This work was supported by the Ministry of Science and Technology of China (2016YFA0100701), the Chinese Academy of Sciences (XDB19020104), the National Natural Science Foundation of China (31861143025, 31830108, 31821004, 31725009, and 31730111), and the HHMI International Research Scholar Program (55008728).

AUTHOR CONTRIBUTIONS

Conceptualization, R.-W.Y. and L.-L.C.; Methodology, R.-W.Y., G.X., and L.-L.C.; Investigation, R.-W.Y., G.X., L.S., P.-F.L., L.-Z.Y., Yang Wang, M.W., and Y.-H.X.; Software, Ying Wang; Formal Analysis & Validation, R.-W.Y., G.X., and L.-L.C.; Writing – Original Draft, L.-L.C.; Writing – Review & Editing, R.-W.Y., G.X., L.Y., and L.-L.C.; Visualization, R.-W.Y., G.X., and L.-L.C.; Supervision, L.-L.C.; Funding Acquisition, L.Y. and L.-L.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 4, 2018

Revised: February 4, 2019

Accepted: August 13, 2019

Published: September 17, 2019

SUPPORTING CITATIONS

The following references appear in the Supplemental Information: Ball et al. (2015).

REFERENCES

- Aittaleb, M., Visone, T., Fenley, M.O., and Li, H. (2004). Structural and thermodynamic evidence for a stabilizing role of Nop5p in S-adenosyl-L-methionine binding to fibrillarin. *J. Biol. Chem.* *279*, 41822–41829.
- Albert, B., Perez-Fernandez, J., Léger-Silvestre, I., and Gadal, O. (2012). Regulation of ribosomal RNA production by RNA polymerase I: does elongation come first? *Genet. Res. Int.* *2012*, 276948.
- Bajar, B.T., Wang, E.S., Lam, A.J., Kim, B.B., Jacobs, C.L., Howe, E.S., Davidson, M.W., Lin, M.Z., and Chu, J. (2016). Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Sci. Rep.* *6*, 20889.
- Ball, G., Demmerle, J., Kaufmann, R., Davis, I., Dobbie, I.M., and Schermelleh, L. (2015). SIMcheck: a Toolbox for Successful Super-resolution Structured Illumination Microscopy. *Sci. Rep.* *5*, 15915.
- Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* *18*, 285–298.
- Barandun, J., Hunziker, M., and Klinge, S. (2018). Assembly and structure of the SSU processome—a nucleolar precursor of the small ribosomal subunit. *Curr. Opin. Struct. Biol.* *49*, 85–93.
- Boisvert, F.M., van Koningsbruggen, S., Navascués, J., and Lamond, A.I. (2007). The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* *8*, 574–585.
- Calabretta, S., and Richard, S. (2015). Emerging Roles of Disordered Sequences in RNA-Binding Proteins. *Trends Biochem. Sci.* *40*, 662–672.
- Cheutin, T., O'Donohue, M.F., Beorchia, A., Vandelaer, M., Kaplan, H., Deféver, B., Ploton, D., and Thiry, M. (2002). Three-dimensional organization of active rRNA genes within the nucleolus. *J. Cell Sci.* *115*, 3297–3307.
- Comai, L. (1999). The nucleolus: a paradigm for cell proliferation and aging. *Braz. J. Med. Biol. Res.* *32*, 1473–1478.
- Crosby, K.C., Postma, M., Hink, M.A., Zeelenberg, C.H., Adjobo-Hermans, M.J., and Gadella, T.W. (2013). Quantitative analysis of self-association and mobility of annexin A4 at the plasma membrane. *Biophys. J.* *104*, 1875–1885.
- Defrasnes, C., Marsh, G.A., Foo, C.H., Rootes, C.L., Gould, C.M., Grusovin, J., Monaghan, P., Lo, M.K., Tompkins, S.M., Adams, T.E., et al. (2016). Genome-wide siRNA Screening at Biosafety Level 4 Reveals a Crucial Role for Fibrillarin in Herpesvirus Infection. *PLoS Pathog.* *12*, e1005478.
- Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* *3*, 195–205.
- Falahati, H., Pelham-Webb, B., Blythe, S., and Wieschaus, E. (2016). Nucleation by rRNA Dictates the Precision of Nucleolus Assembly. *Curr. Biol.* *26*, 277–285.
- Farley, K.I., Surovtseva, Y., Merkel, J., and Baserga, S.J. (2015). Determinants of mammalian nucleolar architecture. *Chromosoma* *124*, 323–331.
- Fei, J., Jadalih, M., Harmon, T.S., Li, I.T.S., Hua, B., Hao, Q., Holehouse, A.S., Reyer, M., Sun, Q., Freier, S.M., et al. (2017). Quantitative analysis of multilayer organization of proteins and RNA in nuclear speckles at super resolution. *J. Cell Sci.* *130*, 4180–4192.
- Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* *165*, 1686–1697.
- Gallagher, J.E., Dunbar, D.A., Granneman, S., Mitchell, B.M., Osheim, Y., Beyer, A.L., and Baserga, S.J. (2004). RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev.* *18*, 2506–2517.
- Grišins, P., and Mazets, I.E. (2014). Metropolis-Hastings thermal state sampling for numerical simulations of Bose-Einstein condensates. *Comput. Phys. Commun.* *185*, 1926–1931.
- Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., and Hofacker, I.L. (2008). The Vienna RNA websuite. *Nucleic Acids Res.* *36*, W70–W74.
- Grummt, I. (2007). Different epigenetic layers engage in complex crosstalk to define the epigenetic state of mammalian rRNA genes. *Hum. Mol. Genet.* *16*, R21–R27.
- Guerousov, S., Weatheritt, R.J., O'Hanlon, D., Lin, Z.Y., Narula, A., Gingras, A.C., and Blencowe, B.J. (2017). Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. *Cell* *170*, 324–339.e323.
- Guizar-Sicairos, M., Thurman, S.T., and Fienup, J.R. (2008). Efficient subpixel image registration algorithms. *Opt. Lett.* *33*, 156–158.
- Hacker, I., Sander, B., Golas, M.M., Wolf, E., Karagoz, E., Kastner, B., Stark, H., Fabrizio, P., and Luhrmann, R. (2008). Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. *Nat Struct Mol Biol* *15*, 1206–1212.
- Kishore, S., Gruber, A.R., Jedlinski, D.J., Syed, A.P., Jorjani, H., and Zavolan, M. (2013). Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol.* *14*, R45.
- Koberna, K., Malinský, J., Pliss, A., Masata, M., Vecerova, J., Fialová, M., Bednár, J., and Raska, I. (2002). Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of “Christmas trees” in situ. *J. Cell Biol.* *157*, 743–748.
- Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., and Hsu, P.D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* *173*, 665–676.e614.
- Kopp, K., Gasiorowski, J.Z., Chen, D., Gilmore, R., Norton, J.T., Wang, C., Leary, D.J., Chan, E.K., Dean, D.A., and Huang, S. (2007). Pol I transcription and pre-rRNA processing are coordinated in a transcription-dependent manner in mammalian cells. *Mol. Biol. Cell* *18*, 394–403.
- Krogan, N.J., Peng, W.T., Cagney, G., Robinson, M.D., Haw, R., Zhong, G., Guo, X., Zhang, X., Canadien, V., Richards, D.P., et al. (2004). High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell* *13*, 225–239.
- Langdon, E.M., Qiu, Y., Ghanbari Niaki, A., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M., et al. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. *Science* *360*, 922–927.
- Li, T., Lu, Q., Wang, G., Xu, H., Huang, H., Cai, T., Kan, B., Ge, J., and Shao, F. (2013). SET-domain bacterial effectors target heterochromatin protein 1 to activate host rDNA transcription. *EMBO Rep.* *14*, 733–740.
- Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovskiy, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* *360*, 918–921.
- McKnight, S.L., and Miller, O.L., Jr. (1976). Ultrastructural patterns of RNA synthesis during early embryogenesis of *Drosophila melanogaster*. *Cell* *8*, 305–319.

- McStay, B., and Grummt, I. (2008). The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.* *24*, 131–157.
- McStay, B. (2016). Nucleolar organizer regions: genomic 'dark matter' requiring illumination. *Genes Dev* *30*, 1598–1610.
- Mészáros, B., Erdos, G., and Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* *46* (W1), W329–W337.
- Mullineux, S.T., and Lafontaine, D.L. (2012). Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? *Biochimie* *94*, 1521–1532.
- Oruganti, S., Zhang, Y., Li, H., Robinson, H., Terns, M.P., Terns, R.M., Yang, W., and Li, H. (2007). Alternative conformations of the archaeal Nop56/58-fibrillar complex imply flexibility in box C/D RNPs. *J. Mol. Biol.* *371*, 1141–1150.
- Pérez-Fernández, J., Román, A., De Las Rivas, J., Bustelo, X.R., and Dosil, M. (2007). The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol. Cell. Biol.* *27*, 5414–5429.
- Phair, R.D., Gorski, S.A., and Misteli, T. (2004). Measurement of dynamic protein binding to chromatin *in vivo*, using photobleaching microscopy. *Methods Enzymol* *375*, 393–414.
- Prieto, J.L., and McStay, B. (2007). Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells. *Genes Dev.* *21*, 2041–2054.
- Puvion-Dutilleul, F., Bachelier, J.P., and Puvion, E. (1991). Nucleolar organization of HeLa cells as studied by *in situ* hybridization. *Chromosoma* *100*, 395–409.
- Raj, A., and Tyagi, S. (2010). Detection of individual endogenous RNA transcripts *in situ* using multiple singly labeled probes. *Methods Enzymol* *472*, 365–386.
- Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* *3*, 793–795.
- Santoro, R., Li, J., and Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat. Genet.* *32*, 393–396.
- Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O., and Drummen, G.P.C. (2019). Super-resolution microscopy demystified. *Nat. Cell Biol.* *21*, 72–84.
- Shi, X., Garcia, G., 3rd, Van De Weghe, J.C., McGorty, R., Pazour, G.J., Doherty, D., Huang, B., and Reiter, J.F. (2017). Super-resolution microscopy reveals that disruption of ciliary transition-zone architecture causes Joubert syndrome. *Nat. Cell Biol.* *19*, 1178–1188.
- Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* *357*, eaaf4382.
- Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L. (2013). The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol. Cell* *51*, 539–551.
- Tatarakis, A., Behrouzi, R., and Moazed, D. (2017). Evolving Models of Heterochromatin: From Foci to Liquid Droplets. *Mol. Cell* *67*, 725–727.
- van Sluis, M., van Vuuren, C., and McStay, B. (2016). The Relationship Between Human Nucleolar Organizer Regions and Nucleoli, Probed by 3D-ImmunoFISH. *Methods Mol. Biol.* *1455*, 3–14.
- Watkins, N.J., and Bohnsack, M.T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* *3*, 397–414.
- Weber, S.C., and Brangwynne, C.P. (2015). Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr. Biol.* *25*, 641–646.
- Wu, H., Yin, Q.F., Luo, Z., Yao, R.W., Zheng, C.C., Zhang, J., Xiang, J.F., Yang, L., and Chen, L.L. (2016). Unusual processing generates SPA LncRNAs that sequester multiple RNA binding proteins. *Mol. Cell* *64*, 534.
- Ying, Y., Wang, X.J., Vuong, C.K., Lin, C.H., Damianov, A., and Black, D.L. (2017). Splicing Activation by Rbfox Requires Self-Aggregation through Its Tyrosine-Rich Domain. *Cell* *170*, 312–323.e310.
- Yuan, X., Feng, W., Imhof, A., Grummt, I., and Zhou, Y. (2007). Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a. *Mol. Cell* *27*, 585–595.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
RPA194 Antibody (C-1)	Santa Cruz	Cat#: sc-48385; RRID: AB_675814
Anti-Fibrillarin antibody	Abcam	Cat#: ab5821; RRID: AB_2105785
Goat anti-Rabbit Secondary Antibody, Alexa Fluor 488	Invitrogen	Cat#: A-11034; RRID: AB_2576217
Goat anti-Mouse Secondary Antibody, Alexa Fluor 488	Invitrogen	Cat#: A-11029; RRID: AB_2534088
Goat anti-Rabbit Secondary Antibody, Alexa Fluor 555	Invitrogen	Cat#: A-21428; RRID: AB_2535849
Goat anti-Mouse Secondary Antibody, Alexa Fluor 555	Invitrogen	Cat#: A-21424; RRID: AB_141780
Goat anti-Rabbit Secondary Antibody, Alexa Fluor 647	Invitrogen	Cat#: A-21244; RRID: AB_2535812
Goat anti-Mouse Secondary Antibody, Alexa Fluor 647	Invitrogen	Cat#: A-21236; RRID: AB_141725
Donkey Anti-Mouse IgG (H+L), Highly Cross-Adsorbed, CF647, Single Label for STORM	Biotium	Cat#: 20810
Monoclonal ANTI-FLAG M2 antibody produced in mouse	Sigma	Cat#: F1804; RRID: AB_262044
Flag tag Antibody	proteintech	Cat#: 20543-1-AP; RRID: AB_11232216
GFP Tag Antibody	proteintech	Cat#: 66002-1-Ig; RRID: AB_11182611
β-Actin	Sigma	Cat#: A3854; RRID: AB_262011
HP1 (E-6)	Santa Cruz	Cat#: sc-515341
Rabbit anti-CPSF73 Antibody	Bethyl Laboratories	Cat#: A301-091A; RRID: AB_2084528
CSTF64 Antibody	Bethyl Laboratories	Cat#: A301-092A; RRID: AB_873014
Anti-XRN2 antibody	Abcam	Cat#: ab72181; RRID: AB_2241927
Rabbit anti-CPSF30 Antibody	Bethyl Laboratories	Cat#: A301-584A; RRID: AB_1078872
Anti-RNA Polymerase II 8WG16	BioLegend	Cat#: MMS-126R; RRID: AB_10013665
Anti-Digoxigenin-AP, Fab fragments	Roche	Cat#: 11093274910; RRID: AB_514497
NHP2L1 Antibody	proteintech	Cat#: 15802-1-AP; RRID: AB_2251452
Anti-DKC1 antibody	Abcam	Cat#: ab93777; RRID: AB_2245868
HEATR1(B-11)	Santa Cruz	Cat#: sc-390445
Chemicals, Peptides, and Recombinant Proteins		
Ribonucleoside Vanadyl Complex	NEB	Cat#: S1402S
cOMplete ULTRA Tablets, Mini, EASYpack Protease Inhibitor Cocktail	Roche	Cat#:000 000005892970001
Lipofectamine 3000 Reagent	Thermo	Cat#: 21341
VECTASHIELD Antifade Mounting Medium	Vector Lab	Cat#: H-1000
ProLong Diamond Antifade Mountant	Invitrogen	Cat#: P36970
TetraSpeck Microspheres, 0.1 μm	Molecular Probes	Cat#: T7279
Cyanine 3-dUTP	Enzo Life	Cat#: ENZ-42501
Red 650 dUTP	Enzo Life	Cat#: ENZ-42522
Hoechst 33342	Sigma	Cat#: B2261-25MG
DAPI	Invitrogen	Cat#: D1306
Triton X-100	ABCONE	Cat#: X10010
Bovine Serum Albumin	ABCONE	Cat#: A23088
Actinomycin D	MedChemExpress	Cat#: HY-17559
DPBS	GIBCO	Cat#: 14190136
Paraformaldehyde	Sigma	Cat#: 158127-500G
Glutaraldehyde Solution	Sigma	Cat#: G6257-100ML
Ni Sepharose 6 Fast Flow	GE healthcare	Cat#: 17-5318-01
5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside	Sigma	Cat#: D1916-10MG

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
DNA-free™ kit	Ambion	Cat#: AM1907
DIG Northern Starter Kit	Roche	Cat#: 12039672910
Hieff Clone One Step Cloning Kit	Yeasen	Cat#: 10905ES25
2 × T5 Super PCR Mix	TSINGKE	Cat#: TSE005
PrimeScript RT reagent Kit (Perfect Real Time)	TaKaRa	Cat#: RR037A
RiboMAX Large Scale RNA Production System	Promega	Cat#: P1300
Nick Translation Kit	Abbott	Cat#: 07J00-001
StarPrep Gel Extraction Kit StarPrep	GenStar	Cat#: D205-04
HighGene Transfection Reagent	ABclonal	Cat#: RM09014
One-tube General Sample DNAup for PCR	Sangon Biotech	Cat#: B518401
TIANprep Mini Plasmid Kit	Tiagen	Cat#: DP103
Dynabeads Protein G	Invitrogen	Cat#: 1003D
Deposited Data		
Original uncropped data	This paper; Mendeley data	https://doi.org/10.17632/kbm9nww4xn.1
Original screening imaging data	This paper; Mendeley data	https://doi.org/10.17632/k2n56cvbwd.1
Experimental Models: Cell Lines		
HeLa	ATCC	ATCC CCL-2; RRID: CVCL_0030
HEK293	ATCC	ATCC CRL-1573; RRID: CVCL_0045
H9	WiCell Research Institute	N/A
HFF	Stem Cell Bank, Chinese Academy of Sciences	SCSP-109
Recombinant DNA		
pmEGFP-C1-RPA194	This paper	N/A
pmEGFP-C1-FBL	This paper	N/A
pmRuby3-C1-B23	This paper	N/A
pmRuby3-C1-FBL	This paper	N/A
pmTagBFP2-C1-B23	This paper	N/A
pCMV-VSV-G	Addgene	Cat#: 8454
pMD2.G	Addgene	Cat#: 12259
pHAGE-FBL-3xFLAG-IRES-NeoR	This paper	N/A
pHAGE-MD-3xFLAG-IRES-NeoR	This paper	N/A
pHAGE-GAR-3xFLAG-IRES-NeoR	This paper	N/A
pHAGE-ReGAR-MD-3xFLAG-IRES-NeoR	This paper	N/A
pHAGE-MD-GAR-3xFLAG-IRES-NeoR	This paper	N/A
pHAGE-MD-ReGAR-3xFLAG-IRES-NeoR	This paper	N/A
pHAGE-H2B-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-GAR ₅₆₋₈₀ -MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-GAR ₃₂₋₈₀ -MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-1.3xGAR-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-2xGAR-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-3xGAR-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-rGG1-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-rGG2-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-rGAR1-MD-3xFLAG-IRES-NeoR	This paper	N/A
PHAGE-rGAR2-MD-3xFLAG-IRES-NeoR	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pmEGFP-C1-FBL	This paper	N/A
pmEGFP-C1-MD	This paper	N/A
pmEGFP-C1-GAR	This paper	N/A
pmEGFP-C1-ReGAR-MD	This paper	N/A
pmEGFP-C1-MD-GAR	This paper	N/A
pmEGFP-C1-MD-ReGAR	This paper	N/A
pmEGFP-C1-H2B-MD	This paper	N/A
pmEGFP-C1-GAR ₅₆₋₈₀ -MD	This paper	N/A
pmEGFP-C1-GAR ₃₂₋₈₀ -MD	This paper	N/A
pmEGFP-C1-1.3xGAR-MD	This paper	N/A
pmEGFP-C1-2xGAR-MD	This paper	N/A
pmEGFP-C1-3xGAR-MD	This paper	N/A
pmEGFP-C1-mEGFP-GAR-MD	This paper	N/A
pmRuby3-C1-MD	This paper	N/A
pmRuby3-C1-GAR	This paper	N/A
pmRuby3-C1-Lifeact	This paper	N/A
tet-on pLKO.1 puro	Addgene	Cat#: 21915
px330-sgRNA-RPA194	This paper	N/A
px330-sgRNA-FBL	This paper	N/A
px330-sgRNA-NHP2L1	This paper	N/A
px330-sgRNA-DKC1	This paper	N/A
pCRII-TOPO	Invitrogen	Cat#: K461020
pCRII-TOPO-mEGFP-RPA194	This paper	N/A
pCRII-TOPO-mEGFP-FBL	This paper	N/A
pCRII-TOPO-mCherry-NHP2L1	This paper	N/A
pCRII-TOPO-mCherry-DKC1	This paper	N/A
pGEX-4T-1	GE Healthcare	Cat#: 28-9545-49
pGEX-4T-1-FBL	This paper	N/A
pET-28a	Novagen	Cat#: 69864-3
pET-28a-EGFP	This paper	N/A
pET-28a-mNeonGreen-FBL-FL	This paper	N/A
pET-28a-mNeonGreen-GAR ₅₆₋₈₀ -MD	This paper	N/A
pET-28a-mNeonGreen-MD	This paper	N/A
pET-28a-mNeonGreen-GAR	This paper	N/A
pET-28a-mRuby3-FBL-FL	This paper	N/A
pET-28a-FBL-FL-mRuby3	This paper	N/A
pET-28a-mRuby3-GAR	This paper	N/A
pmEGFP-C1-GAR1	This paper	N/A
pmEGFP-C1-NHP2L1	This paper	N/A
pmEGFP-C1-DKC1	This paper	N/A
pmEGFP-C1-NOP56	This paper	N/A
pmEGFP-C1-NOP58	This paper	N/A
pmEGFP-C1-NOP10	This paper	N/A
pmEGFP-C1-NHP2	This paper	N/A
pmEGFP-C1-UTP4	This paper	N/A
pmEGFP-C1-WDR75	This paper	N/A
pmEGFP-C1-UTP15	This paper	N/A
pmEGFP-C1-WDR43	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pHAGE-dCas13d	This paper	N/A
pdCas13d-gRNA1	This paper	N/A
pdCas13d-gRNA2	This paper	N/A
Software and Algorithms		
Fiji/ImageJ	Fiji/ImageJ	https://imagej.net/Fiji
GraphPad Prism	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
sgRNAs Design Tool	CRISPR DESIGN	http://zlab.bio/guide-design-resources
LAS X	Leica	https://www.leicabiosystems.com/
Huygens Professional	Scientific Volume Imaging	https://svi.nl/Huygens-Professional
softWoRx 6.5	GE Healthcare	N/A
NIS Elements AR	Nikon	https://www.microscope.healthcare.nikon.com/products/software
Office365	Microsoft	https://www.office.com
iWork	Apple	https://www.apple.com/cn/iwork/
MATLAB	MathWorks	https://www.mathworks.com
Sketchbook	Autodesk	https://www.sketchbook.com
IUPred2A	N/A	https://iupred2a.elte.hu/
RNAfold web server	N/A	http://ma.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Ling-Ling Chen (linglingchen@sibcb.ac.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Cell Lines

Human cell lines including HeLa and HEK293 cells were purchased from the American Type Culture Collection (ATCC; <http://www.atcc.org>). Human embryonic stem cell line H9 was from WiCell Research Institute. HFF cells were kindly provided by Stem Cell Bank, Chinese Academy of Sciences (<http://www.cellbank.org.cn/>).

METHOD DETAILS

Cell culture and cell transfection

Human HeLa and HEK293 cell lines were cultured using standard protocols from the American Type Culture Collection (ATCC; <http://www.atcc.org>). Human HFF was cultured using standard protocols from Stem Cell Bank, Chinese Academy of Sciences (<http://www.cellbank.org.cn/>). Human embryonic stem cell line H9 was maintained as previously described (Wu et al., 2016). Briefly, H9 cells were cultured on plates coated with growth-factor-depleted Matrigel (BD Biosciences, San Jose, CA) in either serum-free, defined mTeSR medium (StemCell Technologies, Inc, Vancouver, BC, Canada) or fibroblast-conditioned medium (CM) with irradiation inactivated mouse embryo fibroblasts supplemented with 4 ng/ml human Basic Fibroblast Growth Factor (bFGF) (Life Technologies, Rockville, MD). Cultured H9 cells were regularly evaluated for Oct3/4 expression every 3-4 weeks and passaged every 6-7 days.

Transfection of plasmid was carried out with Lipofectamine 3000 Transfection Reagent (Invitrogen) according to the manufacturer's protocol.

Lentivirus infection and production 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 μ g pHAGE or tet-on pLKO.1-TRC construct, 7.5 μ g of psPAX2 and 3 μ g pMD2. G. The supernatant containing viral particles were harvested twice at 48 and 72 hours after transfection, filtered through Millex-GP Filter Unit (0.22 μ m pore size, Millipore), and stored at -80° C till use. To infect HeLa cells with lentivirus, cells were cultured in medium containing lentivirus and 1 μ g/ml polybrene (Sigma). To increase the

efficiency, infected cells were under several days of puromycin or G418 selection. The cells were recovered at least 7 days before performing subsequent experiments.

Plasmids construction

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 KD HeLa cell lines were generated as described above.

To construct expression plasmids for mEGFP-FBL, mEGFP-RPA194, mEGFP-RPA49, mRuby3-FBL, mRuby3-B23, mTagBFP2-B23, mEGFP-GAR1, mEGFP-DKC1, mEGFP-NHP2, mEGFP-NOP10, mEGFP-NHP2L1, mEGFP-NOP58, mEGFP-NOP56 and the chimeric FBL mutants (mEGFP-MD, mEGFP-GAR, mEGFP-ReGAR-MD, mEGFP-MD-GAR, mEGFP-MD-ReGAR, mEGFP-H2B-MD, mRuby3-MD, mEGFP-GAR₅₆₋₈₀-MD, mEGFP-GAR₃₂₋₈₀-MD 2xmEGFP-GAR-MD), the full-length FBL, RAP194, RPA49, B23, GAR1, DKC1, NHP2, NOP10, NHP2L1, NOP58 and NOP56 were individually amplified from HeLa cDNAs, and inserted into pmEGFP-C1, pmRuby3-C1 or pmTagBFP2-C1 vector using one-step clone method. Plasmids transformation was performed in T1 *E. coli* (Trans1-T1 Phage Resistant Chemically Competent Cell, Transgen). The transfection was performed as described above.

To construct expression plasmids for chimeric FBL, sequences of FBL-FL, MD, GAR, ReGAR-MD, MD-GAR, MD-ReGAR, H2B-MD, GAR₅₆₋₈₀-MD and GAR₃₂₋₈₀-MD with 3xFlag were individually amplified and cloned into pHAGE-EF1 α -IRES-NeoR using one-step clone method.

To construct protein purification plasmids for chimeric FBL, sequences of mNeongreen-FBL-FL, mNeongreen-GAR₅₆₋₈₀-MD, mNeongreen-MD, mNeongreen-GAR, mRuby3-FBL-FL, FBL-FL-mRuby3 or mRuby3-GAR were inserted into pET-28a using one step clone method for protein purification from *E. coli*.

Primer and shRNAs sequences for plasmids construction used were listed in [Table S1](#).

Protein expression and purification

Expression plasmids for His-tagged mEGFP, mNeongreen-FBL-FL, mNeongreen-GAR₅₆₋₈₀-MD, mNeongreen-MD, mNeongreen-GAR, mRuby3-FBL-FL, FBL-FL-mRuby3 and mRuby3-GAR in pET-28a were individually transformed into *E. coli* expression strain BL21 [Transetta (DE3) chemically competent cell (Transgen Biotech, CD801)]. After transformation, a single colony was inoculated in 5 mL LB media supplemented with 100 μ g/L kanamycin at 250 rpm, 37°C. After overnight growth, the culture was diluted 100-fold into 500 mL LB medium supplemented with 100 μ g/L kanamycin. Absorbance was monitored at a wavelength of 600 nm, and upon reaching an optical density (OD₆₀₀) of 0.6 - 0.8, IPTG was added to LB medium at the concentration of 0.5 mM for the induction of protein expression. After overnight incubation at 180 rpm, 16°C, cell pellets were harvested by centrifugation (5,000 rpm, 10 min, 4°C), resuspended in lysis buffer (20 mM Tris-HCl pH 7.5-8.0, 500 mM NaCl, 12 mM β -mercaptoethanol, 0.5 mM PMSF) with 1 mg/ml lysozyme rotated at 4°C for 30 min, and fragmented by high-pressure homogenizer (Ultrahigh pressure cell crusher UH-06; Union-biotech) at 4°C. After centrifugation at 10,000 rpm for 30 min at 4°C, the supernatant cell lysates were filtered through a 0.45 filter and then incubated with Ni Sepharose (GE healthcare, 17-5318-01) for 2 hours at 4°C. The Sepharose beads were washed with washing buffer (20 mM Tris-HCl pH 7.5 - 8.0, 500 mM NaCl, 20 mM imidazole 0.5 mM PMSF), and bound proteins were eluted with elution buffer (20 mM Tris-HCl pH 7.5 - 8.0, 500 mM NaCl, 250 mM imidazole 0.5 mM PMSF) for twice. Then the protein further purified over the gel filtration chromatography (Superdex-200; GE Healthcare) equilibrated with storage buffer (20 mM Tris-HCl pH 7.5 - 8.0, 500 mM NaCl, 0.1 mM PMSF and 3% Glycerol). The concentration of purified protein was determined by using Modified Bradford Protein Assay Kit (Sangon Biotech, C503041) and checked by Coomassie blue staining.

Expression plasmids for GST-tagged full-length FBL in pGEX-4T-1 were individually transformed into *E. coli* expression strain BL21 [Transetta (DE3) chemically competent cell (TRANSGEN BIOTECH, CD801)]. After transformation, a single colony was inoculated in 5 mL LB media supplemented with 100 μ g/L kanamycin at 250 rpm, 37°C. After overnight growth, the culture was diluted 100-fold into 100 mL LB medium supplemented with 100 μ g/L kanamycin. Absorbance was monitored at a wavelength of 600 nm, and upon reaching an optical density (OD₆₀₀) of 0.6 - 0.8, IPTG was added to LB medium at the final concentration of 0.2 mM for the induction of protein expression. After overnight incubation at 250 rpm, 16°C, cell pellets were harvested by centrifugation (5,000 rpm, 10 min, 4°C), resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM PMSF,) with 1 mg/ml lysozyme on ice for 30 min, and sonicated for 10 min (5 s on/off) on ice. After centrifugation at 10,000 rpm for 30 min at 4°C, the supernatant cell lysates were incubated with Glutathione Sepharose (GE healthcare, 17-0756-01) for 2 hours at 4°C. The Sepharose was washed 4 times with lysis buffer, and bound protein was eluted with elution buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM PMSF, 20 mM Glutathione). Elution was repeated twice to gain maximum yield. Eluted protein was dialyzed in Dialysis buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1 mM PMSF, 10% Glycerol) overnight at 4°C. The concentration of purified protein was determined using Modified Bradford Protein Assay Kit (Sangon Biotech, C503041) and checked by Coomassie blue staining.

RNA isolation, RT-qPCR and Northern Blots

Total RNAs from equal number of cultured cells or cultured cells with different treatments 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 by using SYBR Green Realtime PCR Master Mix

(TOYOBO) and a StepOnePlus real-time PCR system (Applied Biosystems). Primer sequences for RT-qPCR used were listed in Table S1.

To examine intermediates of pre-rRNAs, Northern Blots (NB) were carried out according to the manufacturer's protocol (DIG Northern Starter Kit, Roche). RNA was loaded on agarose gels and the digoxigenin (Dig) labeled antisense probes of 5' ETS-1 as previously reported (Tafforeau et al., 2013) were used. Probe sequences for NB were listed in Table S1.

Measurement of rDNA copy numbers

A serial dilution of the linearized plasmids containing rDNA sequences were used to generate a standard curve for rDNA by qPCR. The copy numbers of the diluted plasmids were calculated by DNA/RNA Copy Number Calculator from website (<http://endmemo.com/bio/dnacopynum.php>). To measure the rDNA copy number in HeLa cells, total RNAs were extracted from 2×10^6 cells, then DNAs were extracted and diluted for qPCR analysis, and the copy number of rDNA could be quantitated from the standard curve. Primer sequences were listed in Table S1.

Knock-in fluorescent protein by CRISPR/Cas9

To construct the plasmid using in knock-in, DNA sequences for left homology arm and right homology arm of targeted genes were amplified from genomic DNA using the primer pairs listed in Table S1. 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 mEGFP or mCherry was PCR amplified using the primers listed in Table S1. Overlap PCR was used for generating the left homology-mEGFP/mCherry-right homology arm sequences. The PCR products were purified, digested with BamHI/NotI and cloned into a pCRII plasmid that was digested with BamHI/NotI by a standard two fragments ligation. Design of the guide RNAs was carried out using the CRISPR Design Tool (<http://zlab.bio/guide-design-resources>) to minimize potential off-target effects. Oligonucleotide pairs (Table S1) were cloned into the vector pS338. The final bicistronic vector encoded the gRNA and the Cas9 nuclease.

To obtain the mEGFP-KI or mCherry-KI HeLa cell lines, 1×10^6 cells per well were seeded in a 6-well plate with supplemented DMEM+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 μg/ml) was added to the cells to increase the KI efficiency. Three days later, the cells were inspected by fluorescence microscopy and mEGFP or mCherry positive cells were sorted into a new 10cm dish using FACSAria (BD Biosciences, Heidelberg, Germany). Within about two to three weeks after single-cell sorting, mEGFP or mCherry-positive single colonies were picked up and transferred into 24-well plates.

The mEGFP or mCherry -KI cell lines were then observed using an epifluorescence microscope (DeltaVision Elite). Cells that showed a correct fluorescence localization patterns were further confirmed using PCR. Briefly, genomic DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen) from a confluent well of a 12-well plate. 100 ng genomic DNA was used as a template for an out-out PCR with the primers listed in Table S1 and the PCR products were analysis by Sanger sequencing.

Protein visualization

To detect protein localization by immunofluorescence in fixed cells, cells were seeded on High Performance No.1.5 18 × 18mm 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 hour at room temperature. Primary antibodies were diluted with 1% BSA (FBL 1:400, RPA194 1:400, Pol II 1:200, CPSF73 1:100, CstF64 1:100, CPSF30 1:100, XRN2 1:100, HP1 1:50, EGFP 1:400, FLAG 1:400) and incubate for 1 hour at room temperature. After washing with 1xDPBS 3 times, fluorescent secondary antibodies were 1: 1,000 diluted in 1% BSA and incubated for 1 hour at room temperature. Samples were mounted in VECTASHIELD antifade mounting medium (Vector Lab).

To detect mEGFP/mRuby3-tagged proteins in fixed cells, cells were seeded on High Performance No.1.5 18 × 18mm glass coverslips and were fixed with 3% PFA/0.1% glutaraldehyde in DPBS for 10 min at room temperature. Then we performed blocking and single molecule RNA fluorescent *in situ* hybridization as describe below. Samples were mounted with VECTASHIELD antifade mounting medium (Vector Lab).

For live cell imaging, cells were seeded on 35 mm No.1.5 glass bottom dishes (Labtite, 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 hour. All images were obtained at 37°C with 5% CO₂ condition.

Single molecule RNA Fluorescent *in situ* Hybridization (smFISH)

All single molecule RNA FISH probes were designed via Stellaris Probe Designer and labeled with cy3 or cy5 on the 3' ends (Table S1). RNA FISH was carried as described before (Raj and Tyagi, 2010). Briefly, 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 10min at room temperature follow by hybridization at 37°C for 16 hours. 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). For samples labeled with Cy5, ProLong Diamond antifade reagent (Thermo Fisher) was used.

DNA Fluorescent *in situ* Hybridization (DNA FISH)

To detect rDNAs, cells were seeded on High Performance No.1.5 18 × 18mm glass coverslips and fixed with 4% PFA for 15 min, followed by permeabilization with 0.5% Triton X-100 for 5 min, incubated at 37°C for 1 hour with RNase A. Cells were then denatured at 80°C for 10 min in prewarmed 2 × SSC and 70% deionized formamide, pH 7.0. Next, cells were hybridized with the denatured DNA probes prepared from Nick Translation (Abbott) overnight. After hybridization, two washes of 10 min at 37°C with 50% formamide in 2 × SSC were performed, followed by two washes of 15 min at 37°C with 4 × SSC. To co-localize proteins, IF was performed as described above. Slides were mounted with ProLong Diamond antifade reagent (Thermo Fisher).

Widefield Microscopy procedure

All widefield microscopy images were 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 6.5 using the enhanced ratio method.

Structured Illumination Microscopy (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 642nm; 100mW) 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 642nm; 100mW). 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°C 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 colors (Molecular Probes).

Stimulated Emission Depletion Microscopy (STED) procedure

STED images were acquired with Leica TCS SP8 STED 3X microscope equipped with a 100 × 1.4 NA HC PL APO CS2 oil immersion objective and operated with the LAS-X imaging software. Excitation was with an argon laser and emission was detected with hybrid (HyD) detectors. For visualization of FBL, cells were seeded and fixed as described above. The primary antibody targeting FBL and the fluorescent (Alexa488) secondary antibodies were diluted at 1:50 and 1:200, respectively with 1% BSA. Samples were mounted in ProLong Diamond antifade reagent (Thermo Fisher SCIENTIFIC). The Alexa488 labeled FBL was excited by 488nm laser and the STED beam was generated by a 592 nm depletion beam. All STED images were deconvolved using Huygens software (Scientific Volume Imaging).

Stochastic Optical Reconstruction Microscopy (STORM) procedure

Super-resolution STORM imaging was performed on a Nikon N-STORM 5.0 microscope equipped with a motorized inverted microscope ECLIPSE Ti-E, an Achromat TIRF 100 × oil immersion lens with a N.A. of 1.49 (Nikon), a quad band filter composed of a quad line beam splitter (zt405/488/561/640rpc TIRF, Chroma Technology Corporation) and a quad line emission filter (brightline HC 446, 523, 600, 677, Semrock, Inc.). The focus was kept stable during acquisition using Nikon focus system. For the excitation of Alexa 647 or CF 647, the 647 nm continuous wave visible fiber laser was used, and the 405 nm diode laser (CUBE 405-100C, Coherent Inc.) was used for switching back the fluorophores from dark to the fluorescent state. Super-resolution images were reconstructed from a series of 20,000–25,000 frames using the N-STORM analysis module of NIS Elements AR.

Screening factors involved in 47S pre-rRNA sorting

To screening of factors involved in nascent 47S pre-rRNA sorting, tet-on shRNA plasmids were constructed and the Tet inducible stable KD cell lines were generated as described above. To perform the screening, individual cell lines were seeded and induced with 500ng/ml doxycycline for 72h followed by fixation, smFISH and SIM imaging as described above. The 5' ETS-1 intensity on RPA194 in each cell was analyzed using Fiji/ImageJ with in-house build macro automatically (See also [Figure S6](#)). The results obtained in individual KD cell lines were normalized by the scramble shRNA treated cells.

Phase separation assay

Phase separation assay was performed as described previously ([Feric et al., 2016](#)) with modifications. The purified mNeogreen or mRuby3 tagged FBL-FL, GAR₅₆₋₈₀-MD, GAR or MD were assembled by diluting the protein from a high salt-containing storage buffer to a physiological buffer (20 mM Tris PH7.5, 150 mM NaCl, 0.1 mM PMSF) or physiological buffer with 10% Dextran. Samples (7 μM) were prepared on High Performance No.1.5 18 × 18 mm glass coverslips (Schott) and were imaged within 30 mins after drop assembly with a DeltaVision Elite imaging system.

RNA-Protein *in vitro* binding assay

300 ng per reaction of synthetic RNAs labeled with Digoxin (Dig) 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. 40 μL per reaction of Glutathione Sepharose (GE Healthcare) was washed with 300 μL Binding buffer [50 mM HEPES pH 7.0, 150 mM NaCl, 10 mM MgCl₂, 0.5 mM PMSF, 2 mM RVC, protease inhibitor cocktail (Roche)] for twice and suspended in 300 μL Binding buffer. 300 ng per reaction of recombinant proteins (GST-FBL) were added to Glutathione Sepharose and incubated at 4°C for 1 hour. Then RNA was renatured and 2 μL RNasin was added to the Glutathione Sepharose. The binding reaction incubated 2 hours at 4°C, followed by 3 times washing with 500 μL binding buffer. Then RNAs were extracted and analyzed by PAGE gel using anti-Dig antibody.

In vitro RNA sorting assay

For *in vitro* RNA sorting assay, *in vitro* transcribed (Promega) 5' ETS-1 fragments RNA were labeled with Cy3 at the 3' ends and were diluted to concentrations of 500 nM in 10 mM Tris pH 7.0, 10 mM MgCl₂, 25 mM NaCl buffer. RNA was denatured at 95°C for 3 min and refolded by cooling down at 1–4°C per min to 37°C final temperature in a thermocycler. To obtain denatured RNA, RNA was denatured at 95°C for 3 min and cooled down on ice for 5 min. The mNeogreen tagged proteins were phase separated as described above. Samples were prepared on High Performance No.1.5 18 × 18 mm glass coverslips (Schott). RNA was added into droplets to final concentrations of 80 nM and observed with a DeltaVision Elite imaging system with 500ms/frame imaging speed for needed time.

Fluorescence Resonance Energy Transfer (FRET) assay

FRET efficiency was measured with acceptor bleaching approach. Proteins were phase separated as described above and the mNeonGreen-mRuby3 FRET pair was used (Bajar et al., 2016). After droplets formation, imaging was performed on a Leica TCS SP8 X WLL confocal laser scanning microscope (Leica) with LAS X FRAP ab module. In brief, zooming into droplet of interest, highlighting an ROI (region of interest) in which the photo-destruction of the acceptor (mRuby3) occurred and began the program. For photo-destruction of the acceptor, cells were photo-bleached with 558 nm laser line (set at 100% intensity, 2 times). The images were captured in both channels before and after photo bleaching. In all experiments, about 20 droplets were measured, and FRET efficiency was calculated as $E = (1 - \text{Pre}/\text{Post}) \times 100\%$, where Pre and Post represent the intensity of donor (mNeogreen) fluorescence before and after photo bleaching. The final concentration of all proteins used in FRET assay was 7 μM.

Fluorescence Recovery After Photobleaching (FRAP) assay

Cells were cultured on 35 mm no.1.5 glass-bottomed dishes (Cellvis). All FRAP assays were performed on Leica TCS SP8 STED 3X microscope equipped with a 100 × 1.4 NA HC PL APO CS2 oil immersion objective and operated with the LAS-X imaging software. The region of interest was photobleached and the recovery of fluorescence intensity within the region of interest was obtained for each experiment. Intensity recovery curves were normalized and corrected for photobleaching (Phair et al., 2004). The recovery curves were fit to the following expression by GraphPad:

$$Y(t) = A \cdot (1 - e^{-\tau \cdot t})$$

Where A is the end-value of the recovered intensity, τ is the fitted parameter and t is the time after the bleaching pulse.

Native RNA Immunoprecipitation

HeLa cells (4×10^7) were rinsed twice with ice-cold PBS and suspended in 2 ml RIP buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.5% Igepal, 1 mM phenylmethyl sulfonyl fluoride (PMSF), 1 × protease inhibitor cocktail (Roche) and 2 mM ribonucleoside vanadyl complex (VRC, NEB)) followed by sonication. Cell lysates were centrifuged at 1,000 g for 10 min at 4 °C and the supernatants were pre-cleared with 20 μl Dynabeads Protein G (Invitrogen). The pre-cleared supernatants were then divided into four parts equally and incubated with 20 μl Dynabeads Protein G with antibodies for FBL, FLAG, rabbit IgG2b or mouse IgG2b for 2 h at 4 °C, followed by washing three times with high salt buffer (RIP buffer with 0.5 M NaCl, 0.5% sodium deoxycholate and 0.1% Igepal) and twice with RIP buffer. The beads were incubated with elution buffer (100 mM Tris pH 6.8, 4% SDS, 10 mM EDTA) at r.t. for 10 min. One-third of the eluted sample was used for western blot and the remainder was used for RNA extraction.

QUANTIFICATION AND STATISTICAL ANALYSIS

Measurement of FC number and size

For measurement of the FC number and size in different cell lines in Figure 11, HEK293, HFF, HeLa or H9 cells were transfected with FC marker pmEGFP-RPA194 as described above. At least 20 cells were imaged for each type of cell lines by DeltaVision Elite imaging system as described above. The images were then measured by Imaris software using surface building function. The average volume and number of FC labeled by mEGFP-RPA194 in each cell was calculated and plotted.

Measurement of the proportion of active rDNA copies

To measure the proportion of active rDNAs in HeLa cells, rDNA and FBL were co-stained and imaged by SIM as described above. The 3D SIM stacks were imported into Fiji/ImageJ and analyzed by an in-house ImageJ script. Briefly, (1) the reconstructed images were split into single channels; (2) the rDNAs and FC/DFC units (labeled by FBL) were marked using Moments threshold; (3) the masks of rDNAs and FC/DFC units were generated from step (2) and the background outside the masks was set to 0; (4) the integrated density of rDNA signals in the rDNA masks (I_{total}) and that in the FC/DFC unit masks (I_{active}) were counted. The proportion of active rDNAs was calculated as:

$$\text{Active rDNA\%} = \frac{I_{active}}{I_{total}} \cdot 100\%$$

Measurement of the proportion of nascent pre-rRNAs associated with FC/DFC units

Pre-rRNAs were co-stained and imaged with FBL by SIM as described above. The 3D SIM stacks were imported into Fiji/ImageJ and analyzed by an in-house ImageJ script. Briefly, (1) the reconstructed images were split into single channels; (2) the pre-rRNAs and FBL were marked using optimum threshold, respectively; (3) the masks of pre-rRNAs and FBL were generated from step (2) and the background outside the masks was set to 0; (4) the masks of FBL were dilated (three dimensional) and generated the FC/DFC units' masks; (5) the integrated density of pre-rRNAs signal in the pre-rRNAs masks (I_{pre}) and that in the FC/DFC unit masks ($I_{FC/DFC\ unit}$) were counted. The proportion of nascent pre-rRNAs associated with FC/DFC units was calculated as:

$$\text{pre-rRNAs in FC/DFC units \%} = \frac{I_{FC/DFC\ unit}}{I_{pre}} \cdot 100\%$$

Measurement of the 5' ETS-1 signal trapped in FCs

For each KD condition, pre-rRNAs were co-stained and imaged with RPA194 by SIM as described above. The 3D SIM stacks were imported into Fiji/ImageJ and analyzed by an in-house ImageJ script. Briefly, (1) the reconstructed images were split into single channels; (2) the pre-rRNAs and RPA194 were marked using optimum threshold, respectively; (3) the masks of pre-rRNAs and RPA194 were generated from step (2) and the background outside the masks was set to 0; (4) the masks of RPA194 were dilated (three dimensional) and generated the FC units' masks; (5) the integrated density of pre-rRNAs signal in the FC unit masks ($I_{FC\ unit}$) for each cells were counted. For each of KD conditions, the integrated density 5' ETS-1 signals in FCs ($I_{FC\ unit}$) for each cell were normalized by the average value of $I_{FC\ unit}$ in Scr. group. See also [Figure S6](#).

Measurement of the 5' ETS-1 signal sorted in FBL mutants

For each condition, pre-rRNAs were co-stained and imaged with FBL mutants by SIM as described above. The 3D SIM stacks were imported into Fiji/ImageJ and analyzed by an in-house ImageJ script. Briefly, (1) the reconstructed images were split into single channels; (2) the pre-rRNAs and FBL mutants were marked using optimum threshold, respectively; (3) the masks of pre-rRNAs and FBL mutants were generated from step (2) and the background outside the masks was set to 0; (4) the masks of FBL mutants were dilated (three dimensional) 40nm and filled holes to generate the FC/DFC units' masks; (5) the integrated density of pre-rRNAs signal in the FBL mutants masks ($I_{FBL\ mutants}$, 5' ETS-1 sorted in FBL mutants) and that in the FC/DFC units' masks (I_{pre} , total 5' ETS-1) for each cells were counted. For each cells in each KD conditions, the integrated density 5' ETS-1 signals in FCs for each cell were normalized by its own total 5' ETS-1 integrated density: $(I_{FBL\ mutants})/(I_{pre})$; and for each conditions, $(I_{FBL\ mutants})/(I_{pre})$ was normalized by the average value of $(I_{FBL\ mutants})/(I_{pre})$ in control group. See also [Figure S6](#).

Quantification of RPA194 and FBL per FC/DFC unit

Numbers of FC and DFC molecules per FC/DFC unit were quantified using a workflow shown in [Figure S4A](#). The CRISPR/Cas9 mediated mEGFP KI cell lines were generated as described above and the mEGFP were purified as described above. To generate the standard curve of the mEGFP fluorescent intensity and their concentrations, purified mEGFP at 5 different concentrations (0.2579 $\mu\text{g}/\mu\text{l}$, 0.1289 $\mu\text{g}/\mu\text{l}$, 0.08595 $\mu\text{g}/\mu\text{l}$, 0.0737 $\mu\text{g}/\mu\text{l}$, 0.0645 $\mu\text{g}/\mu\text{l}$) were imaged on Leica TCS SP8 STED 3X microscope using a 100 \times 1.4 NA HC PL APO CS2 oil immersion objective and hybrid (HyD) detectors in photon counting mode to minimize the background. For each concentration, six 4.9 $\mu\text{m} \times 4.9 \mu\text{m} \times 3.2 \mu\text{m}$ cubes were imaged at 1% 488 nm laser power with line average for 3 times. The images were then analyzed by Imaris software to calculate the mean intensity (I) and the standard curve was obtained using linear regression equation as:

$$I(c) = 29.84 \cdot c - 0.3028$$

Where c is the concentration of mEGFP, the R^2 of this regression is 0.9993.

With the same image acquisition parameters, at least 20 mEGFP-FBL or mEGFP-RPA194 KI HeLa cells were imaged and analyzed by Imaris software to measure the mean intensity (I) and mean volume (V) of FBL or RPA194 in FC/DFC unit. Using the equation above, the concentration of FBL or RPA194 can be calculated as:

$$c = (I + 0.3018)/29.84$$

Thus, the numbers (N) of RPA194 or FBL molecules per FC/DFC unit can be calculated as:

$$N = (c \cdot V) / M \cdot N_A$$

Where M is the molar mass of the RPA194 or FBL; N_A is Avogadro constant.

Measurement of RNA recruitment coefficient

To measure RNA recruitment coefficient *in vitro* (Figures 5N and 5O), proteins were phase separated and RNAs were sorted *in vitro* as described above. The recruitment coefficient is defined by the following equation as described (Langdon et al., 2018):

$$RC = \frac{(RNA)PD}{(RNA)O}$$

where RC is the recruitment coefficient, (RNA)PD is the intensity of added Cy3-labeled RNA within the preformed droplets, (RNA)O is the intensity of added Cy3-labeled outside the preformed droplets. A higher RC indicates that the RNA is more concentrated in preformed droplets. All image analysis was performed using Fiji/ImageJ with in-house build macro automatically.

Similarly, to measure RNA recruitment coefficient *in vivo* (Figures 6F and 6G), FBL KD cell rescued with mEGFP tagged FBL mutants were imaged by SIM as described above. The recruitment coefficient is defined by the equation above, where RC is the recruitment coefficient, (RNA)PD is the intensity of 5' ETS-1 signal within each FBL mutant, (RNA)O is the intensity of 5' ETS-1 signal outside of each FBL mutant. A higher RC indicates that the RNA is more effectively sorted to this FBL mutant. All image analyses were performed using Fiji/ImageJ with in-house build macro automatically.

Image alignment for FBL distribution in the max-cross x-y section

To view an average distribution of the max-cross section of FBL-FL and GAR₅₆₋₈₀-MD clusters in DFCs in the x-y scale, we used image alignment by Discrete Fourier Transform (DFT) and cross-correlation (Guizar-Sicairos et al., 2008; Shi et al., 2017). First, we randomly selected an image as the reference. Then, another image was rotated from 0° to 359° by 1° per rotation and the rotated images were stored after each rotation (total 360 rotated images). After applying fast Fourier transform to the reference image and rotated images, the cross-correlation peak was individually calculated between the reference image and each rotated image. A rotated image, which has the best cross-correlation peak with the reference image, was then refined in the x-y scale according to the reference image and further averaged with the original reference image to generate a new reference image for subsequent analysis. After applying these steps to 30 different images, an image with six FBL clusters was obtained to show the FBL distribution in the max-cross section in DFCs (Figure S2G) and an image with four GAR₅₆₋₈₀-MD clusters was obtained to show the distribution of the GAR₅₆₋₈₀-MD mutant in the max-cross section in DFCs (Figure 7B).

Digital simulation of FBL distribution in 3D

We performed computer simulation to build the 3D of FBL distribution. This simulation was performed in a Cartesian coordinate system *in silico*, in which one lattice site equals to 1 nm³. Each cluster was set as a solid sphere with the diameter of 133 nm for FBL-FL and 140nm for GAR₅₆₋₈₀-MD, which were both measured under SIM images. A given number (12-30) of FBL clusters were individually arranged around the original point in the shell of DFC sphere with the radius of 247.5 nm for FBL-FL and 170 nm for GAR₅₆₋₈₀-MD, which were both measured under SIM images. For each condition with a different number of FBL clusters, the position of clusters in the shell was randomly simulated for 100 times; and in each time, images of the max-cross section were collected from 200 different angles randomly selected (Figure S3A). Simulated FBL distributions were obtained in the max-cross section called from 200 different randomly rotated angles and the numbers of clusters (vertices) in a total of 100*200 = 20,000 max-cross sections of each polyhedral hollow sphere were counted. These simulated FBL distributions (Figure S3B, left) were then compared with the actual FBL distribution (Figure S3B, right) observed by SIM images. These analyses revealed that 18-24 clusters of FBL-FL and 10-14 clusters of GAR₅₆₋₈₀-MD in 3D in the DFC sphere matched the best with actual observations under SIM (Figures 2N and 7C).

Metropolis Hastings strategy to search the optimal distribution of processing factors

By developing a mathematical model (https://github.com/suduwoniu/nucleolus_model), we calculate the binding ability of processing factors, such as FBL, to pre-rRNAs. If FBL proteins are randomly distributed in the space between the FC/DFC border to the outside surface of the DFC, the binding ability of each FBL is equal, we then use $P(\text{binding}) = 1/N$ (N is the number of factors in this region) to represent the binding ability of each FBL (Figure S4F). In the condition that FBL proteins are not equally distributed between the FC/DFC border to the outside surface of the DFC (Figure S4F), we apply a formula to calculate binding ability of a given FBL:

$$P(x, y, z) = \frac{\rho(x, y, z)}{N} \quad (1)$$

where (x, y, z) is the coordinate of a given FBL factor, $\rho(x, y, z)$ is the concentration of FBL at the location) and $N = \int \rho(\Delta l) \Delta l$ (where l is the path from the FC/DFC border to the outside surface of the DFC (Figure S4F).

First, we applied a series of experiments to measure the numbers and sizes of processing factors in FC and DFC (Figure S4E). To simplify this model, we examined RPA194, the largest component of the Pol I complex, as the key nucleolar factor in the FC, and FBL as the key nucleolar factor in the DFC. The stoichiometric quantification showed that about 175 RPA194 and 3,600 FBL molecules together with 2 ~3 rDNA copies located in each FC/DFC unit (Figures 2F, 2G, and S4D). Accordingly, we set 1,800 FBL dimers (Oruganti et al., 2007) (3,600 FBL molecules) in the DFC and 175 Pol I complexes in the FC in this model. The simulation was performed in the lattice-based Cartesian system composed of 50×50×50 lattice sites, and the diameters of FC and DFC were shown in Figure S2G. Each processing factor (FBL dimer in this model) was estimated approximately to 10×10×10 nm³ in size (according to the size of LSm8 and Prp3 complex in the one arm of Yeast tri-RNP proteins in spliceosome with the similar molecular weight) (Hacker et al., 2008), which occupies about one lattice site (one grid of the Cartesian system). FC was set to be a sphere region filled with synthetic resources in the center of the system, the diameter of which is 13 grids (equal to 130 nm, Figure S2G). After being transcribed at the surface of the FC sphere, nascent pre-rRNAs were bound by FBL factors in DFC to be processed and translocated from the border of the FC/DFC to the DFC region.

Then, by using the formula (1) (Figure S4F) with the calculated parameters (Figure S4E) as input, we applied Metropolis-Hasting algorithm to simulate multivariate distributions (Grişins and Mazets, 2014) to achieve optimal distribution of FBL proteins. In the initial state, FBL proteins were set as randomly (and evenly) distributed. Metropolis-Hasting algorithm automatically calculated the best FBL distribution from this initial random distribution model. After performing 10,000,000 times of FBL disturbance, we obtained a final FBL distribution as multiple clusters that yielded the highest binding ability of FBL with pre-rRNA. The multiple-cluster distribution of FBL showed ~10 times higher binding ability than the initial random distribution (9.6 for the initial random distribution, and 100.1 for the multiple-cluster distribution).

Statistics and reproducibility

The data used in this study are presented as mean ± standard deviation (SD) or standard error of the mean (SEM) in triplicate experiments unless otherwise stated. Statistical analyses (two-tailed Student's t test, Mann-Whitney Test, linear regression, correlation analysis and so on) were performed using existing software (GraphPad Prism 7). $p < 0.05$ was considered significant. Representative pictures for microscopy imaging, Northern blotting and western blotting were obtained from at least two independent experiments. For the statistical significance and sample size of all graphs, please see figure legends and STAR Methods for details.

DATA AND CODE AVAILABILITY

Raw data of key experiments can be accessed on Mendeley data:

<https://doi.org/10.17632/kbm9nww4xn.1>

<https://doi.org/10.17632/k2n56cvbwd.1>