The Complexity of Human Ribosome Biogenesis

Revealed by Systematic Nucleolar Screening

of Pre-rRNA Processing Factors

Lionel Tafforeau, Christiane Zorbas, Jean-Louis Langhendries, Sahra-Taylor Mullineux, Vassiliki Stamatopoulou, Romain Mullier, Ludivine Wacheul, and Denis L.J. Lafontaine
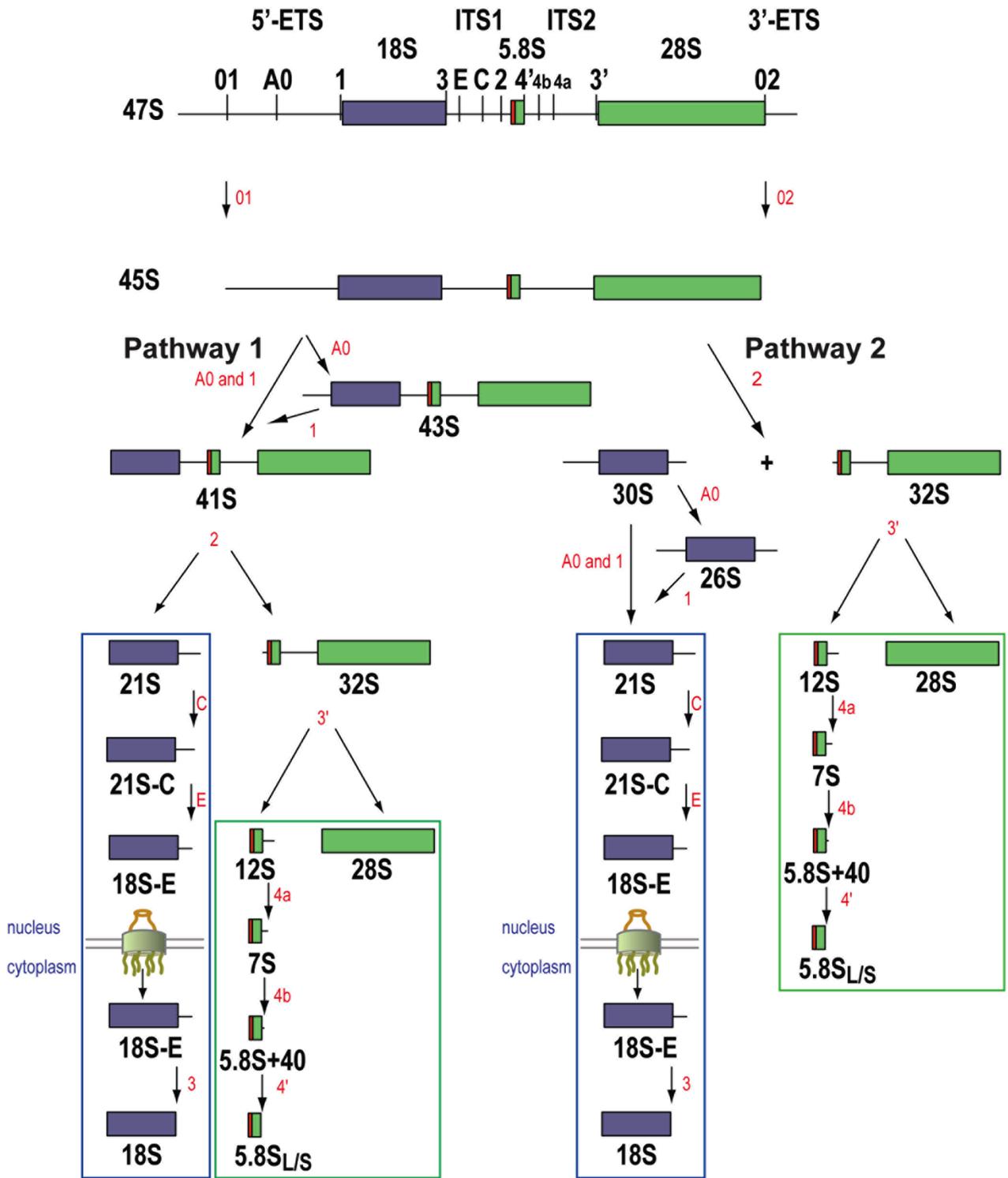
Figure S1. Pre-rRNA Processing Pathway in Cultured Human Cells, Related to Figure 1

The 18S, 5.8S, and 28S rRNA sequences are flanked by the 5' and 3' ETS, ITS1 and ITS2. The primary transcript (47S) is cleaved at sites 01 and 02, generating the 45S pre-rRNA, which is processed by two alternative pathways. In a minor pathway, site A0, in the 5' ETS, and site 1, at the 5' end of 18S rRNA, are cleaved, yielding the 41S pre-rRNA. The 41S is digested at site 2 within ITS1, separating the RNA precursors destined to become the small and large subunits, the 21S and 32S species, respectively. The 21S is cleaved at site E, producing the 18S-E intermediate, which is then processed at site 3 into the mature 18S rRNA. Maturation of the 21S is a sequential process, and a 21S-C intermediate is detected upon depletion of specific assembly factors (e.g., see RPS3 depletion, Fig 1C). Processing of the 32S within ITS2 generates the 12S and the 28S mature rRNA. The 12S pre-rRNA is successively trimmed to produce the 5.8S rRNA by a series of exoribonucleolytic digestions involving specific subunits of the RNA exosome (Fig 6). In the alternative major pathway, the 45S pre-rRNA is directly cleaved at site 2 in ITS1, prior to cleavage at sites A0 and 1, generating the 30S and 32S species. Processing of the 30S pre-RNA at sites A0 and 1 produces, respectively, the 26S and 21S. In human cells, as in budding yeast, there are two forms, short and long, of 5.8S rRNA (red extension). Synthesis of the 26S results from partial uncoupling of cleavages at sites A0 and 1; when such uncoupling occurs the 43S pre-rRNA can also be observed. Similar branches in the two major pathways are boxed in blue and green.
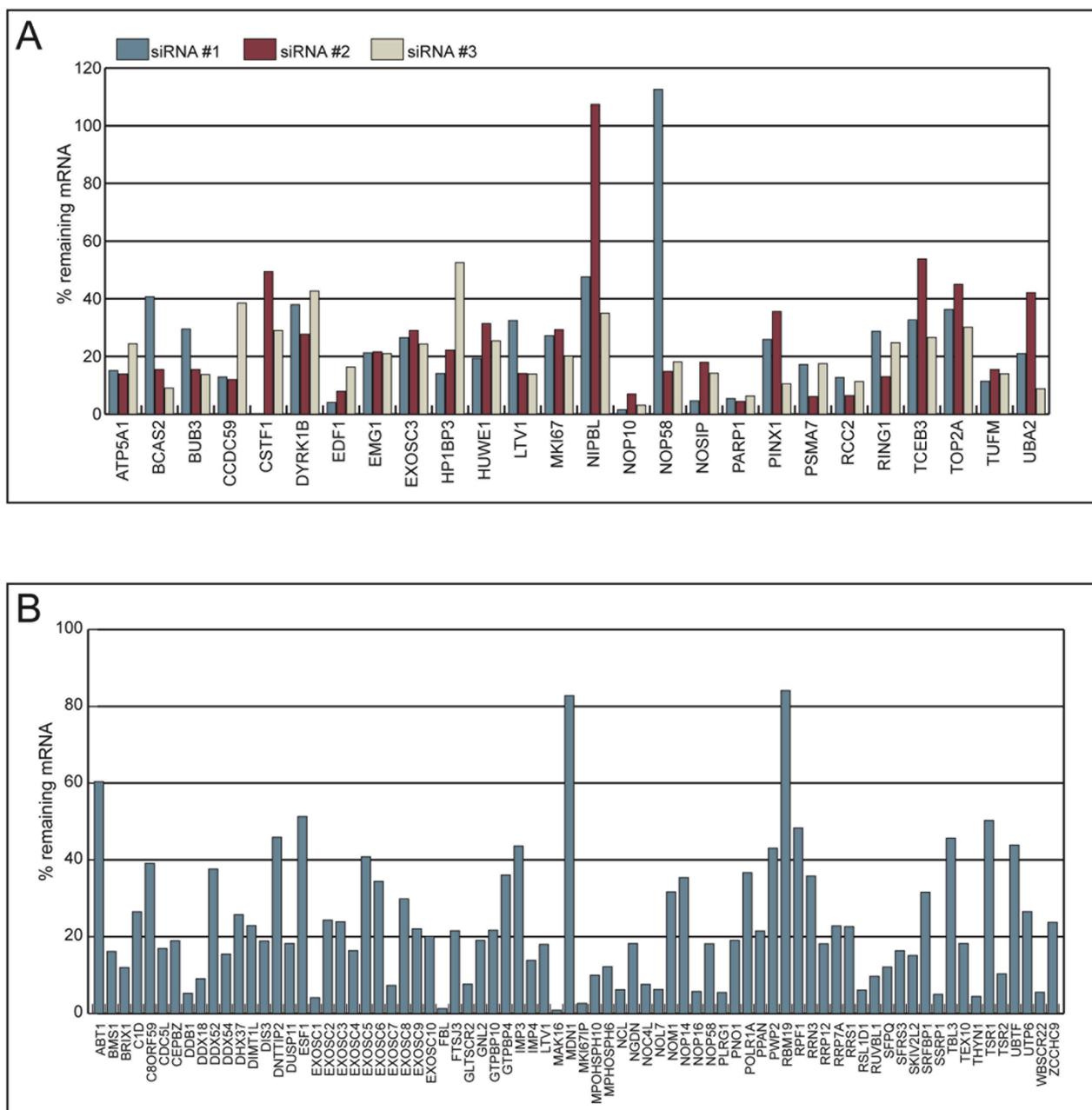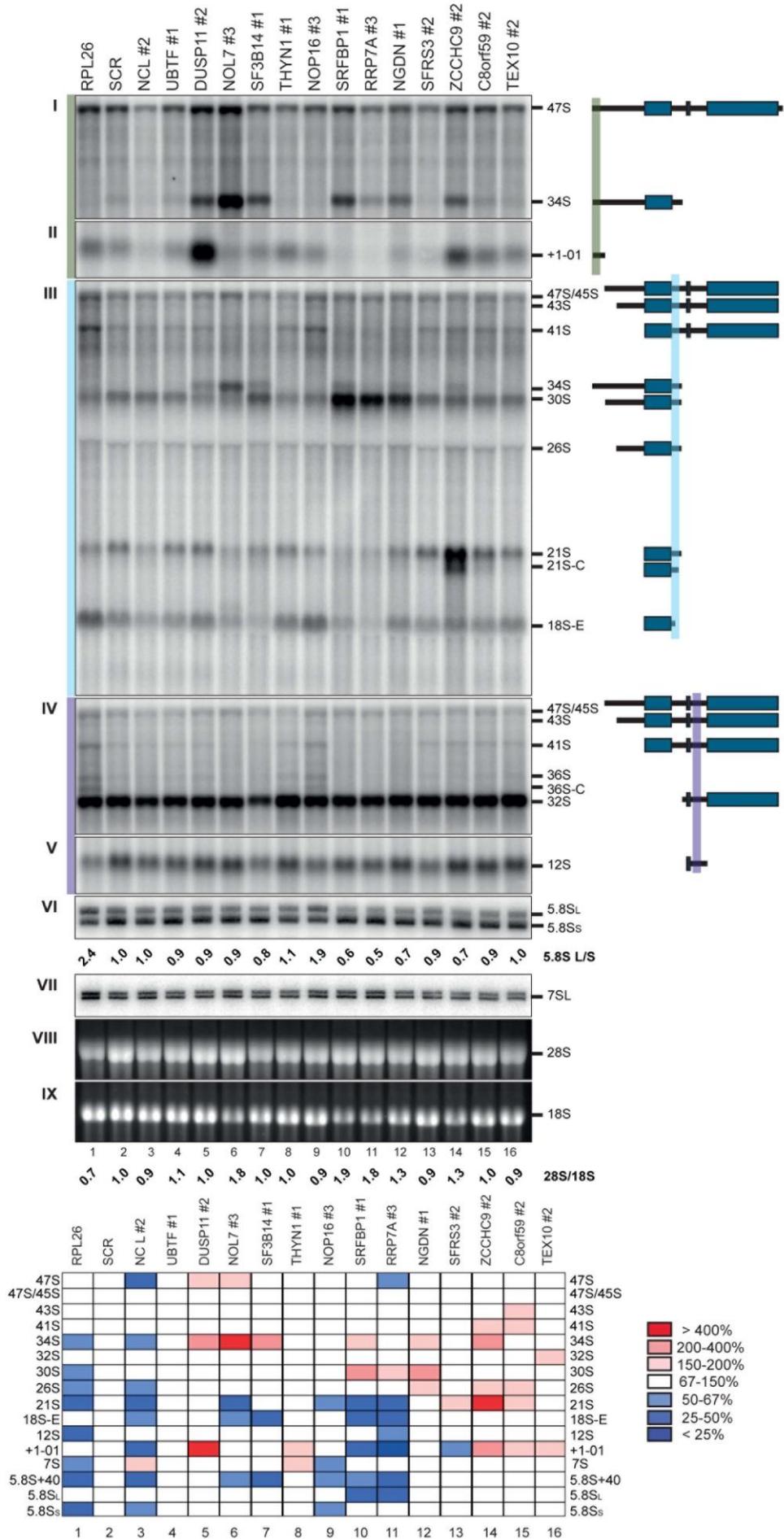
Figure S2. Shotgun qPCR Validation of mRNA Depletion for Randomly Selected Genes, Related to Figure 2

The histograms illustrate the residual level of targeted mRNA (normalized to GAPDH), as assayed by qPCR after a 3 d siRNA-mediated depletion. **A** the efficiency of the three siRNAs was tested for 26 randomly selected genes. **B** the efficiency of one siRNA was tested for 76 randomly selected genes. See also Supplemental Table S9.

A

B

	SCR	POLR1A (Rrpa190) #2	RRN3 (Rrn3) #1	PWP2 (Utp1) #1	MKI67IP (Nop15) #2	MPHOSPH10 (Mpp10) #1	BMS1 (Bms1) #1	DDX52 (Rok1) #2	LTV1 (Ltv1) #2	PNO1 (Dim2) #1	TSR2 (Tsr2) #1	RRP12 (Rrp12) #3	DHX37 (Dhr1) #1	WBSCR22 (Bud23) #1	DIMT1L (Dim1) #3	TSR1 (Tsr1) #1	RRS1 (Rrs1) #2	DDX54 (Dbp10) #3	GNL2 (Nog2) #2	MDN1 (Mdn1) #3	MAK16 (Mak16) #2	CEBPZ (Mak21) #3	
28S/18S	1	0.9	0.8	1.3	0.6	1.4	0.6	1.6	1.5	1.4	1.1	1.3	1.7	1.33	1.0	1.2	0.7	0.7	0.6	0.9	0.6	0.8	

Heatmap legend:
- > 400%
- 200-400%
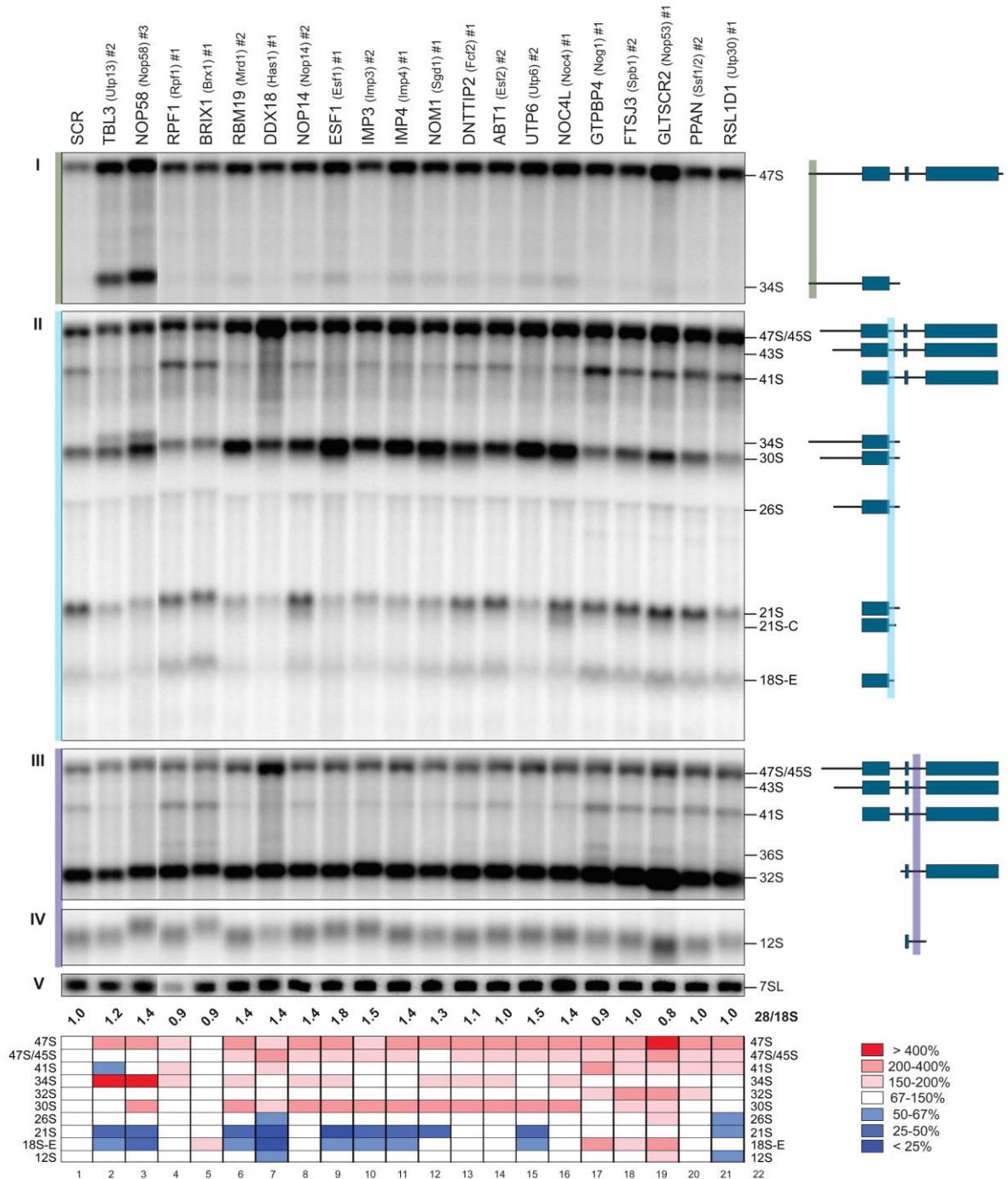- 150-200%
- 67-150%
- 50-67%
- 25-50%
- < 25%

Figure S3. Involvement in Pre-rRNA Processing of 11 Human Genes without a Yeast Homolog and 39 Human Genes with a Yeast Ribosome Assembly Factor Homolog, Related to Figure 3

Analysis of pre-rRNA processing by Northern blotting. Total RNA was extracted from HeLa cells treated for 3 d with a siRNA specific to the target gene, 5 μg was separated by denaturing gel electrophoresis and hybridized with probes (see Fig 1B). As a control, a non-targeting siRNA (SCR) was used. RNAs were resolved on agarose gels with the exception of the 7SL, which was analyzed on acrylamide and provides a loading control. A schematic of the pre-rRNAs detected is provided. siRNA sequences (#1 to #3) are listed in Table S7. The 28S/18S

ratio was calculated from Agilent electropherograms. Bands were quantitated with a Phosphorimager (Fuji FLA-7000) and converted into a heatmap. The signal was standardized to SCR and corrected for loading with the 7SL.

A 11 human genes. RPL26, NCL, UBTF and TEX10 samples were loaded as controls.

Panels I and II, 5' ETS probe; panel III, ITS1 probe; panels IV and V, ITS2 probe; panel VI, probe LD2132; panel VII, probe LD2133. Panels VIII and IX show the large mature rRNAs on an ethidium bromide stained gel.

B 21 human genes with a yeast homolog. The yeast homolog is indicated in brackets. POLR1A and RRN3, both primarily involved in RNA synthesis, were loaded as controls.

Panels I, 5' ETS probe; panel II, ITS1 probe; panels III and IV, ITS2 probe; panel V, probe LD2133.

C 20 human genes with a yeast homolog. Hybridizations as in panel B.

A Human proteins without a yeast homolog

Proteins involved in early processing steps:

Cells depleted for DUSP11 and NOL7 strongly accumulate the 47S primary transcript and 34S RNA, indicating these proteins act at cleavage sites 01, A0 and 1 (panels I and III, lanes 5-6). DUSP11-depleted cells further accumulated a 5'-ETS spacer fragment (+1-01, panel II, lane 5). An increase in 34S was also detected upon the depletion of SF3B14, SRFBP1, NGDN and ZCCHC9 (panel I, lanes 7, 10, 12 and 14). However, in these cases the primary transcript did not accumulate significantly; rather, for SRFBP1 and NGDN, it was the 30S pre-rRNA that became more abundant upon depletion, indicating that cleavages at A0 and 1 are preferentially affected (panel III, lanes 10 and 12). This was also the case for RRP7A (panel III, lane 11). Cells depleted for NOP16 were more abundant in 41S pre-rRNA, indicative of impaired cleavage at site 2 (panel III, lane 9).

Proteins involved in 18S rRNA maturation:

In addition to its effect on 34S accumulation, depletion of ZCCHC9 led to a striking increase in the amount of 21S (panel III, lanes 14), indicating that ZCCHC9 is involved in 18S rRNA maturation. Depletion of SFRS3 and C8orf59 also led to the accumulation of the 21S. For SFRS3, ZCCHC9 and C8orf59, the shorter 21S intermediate (21S-C), typically detected in RPS3-depleted cells (see Fig 1C), was also detected (panel III, lanes 13-15; best seen for ZCCHC9. See RNA exosome role in ITS1 processing. Reduced levels of NOP16 also led to the accumulation of the 18S-E pre-rRNA, a form of 18S rRNA extended by 24 nt at the 3'-end.

Proteins involved in 28S rRNA maturation:

The depletion of THYN1 led to a moderate accumulation of 32S pre-rRNA, similar to what is seen in cells depleted for the ITS2 processing factor TEX10 (panel IV, lanes 8 and 16 and Castle et al., 2012).

B, C Human genes with a yeast homolog

Proteins involved in early processing steps:

Most of the 39 proteins described are involved in early pre-rRNA nucleolar cleavages, as shown by the accumulation of the 47S pre-rRNA. This was particularly evident for GLTSCR2/Nop53 (C, panel I, lane 19). Depletion of PWP2/Utp1, BMS1/Bms1, DDX52/Rok1, TBL3/Utp13 and NOP58/Nop58 and, to a lesser extent, that of four other gene products, led to the accumulation of the aberrant 34S RNA, indicative of cleavage inhibition at 01, A0 and 1 (B, panels I and II, lanes 4, 7-8 and C, same panels, lanes 2-3). MKI67IP/Nop15 and MAK16/Mak16 appeared to be primarily required for cleavage at site 2, as shown by the accumulation of the 41S pre-rRNA (B, lanes 5 and 21). Seven other genes products were also notably involved in processing at site 2: DHX37/Dhr1, TSR1/Tsr1, RRS1/Rrs1, DDX54/Dbp10, GNL2/Nog2, CEBPZ/Mak21 and GTBP4/Nog1. The depletion of BMS1/Bms1 and, to a lesser extent, that of 14 other gene products, led to the accumulation of the 30S (B, panel II, lane 7), indicating a requirement for cleavages at sites A0 and 1. The 26S, resulting from uncoupling at cleavage sites A0 and 1, was particularly abundant in cells deprived of PNO1/Dim2, as well as DIMT1L/Dim1 (B, panel II, lanes 10 and 15).

Proteins involved in 18S rRNA maturation:

Cells lacking LTV1/Ltv1, RRP12/Rrp12 or DHX37/Dhr1 strongly accumulated the 21S (B, panel II, lanes 9, 12 and 13). For LTV1/Ltv1, RRP12/Rrp12, 21S-C was also detected. The 21S-C was also accumulated upon depletion of NOP14/Nop14 and NOC4L/Noc4 (C, panel II, lanes 8 and 16). Depletion of WBSCR22/Bud23, MAK16/Mak16, TSR2/Tsr2, GTPBP4/Nog1 or GLTSCR2/Nop53 led to the accumulation of the late 18S rRNA intermediate, the 18S-E (B, panel II, lanes 11, 14, 21 and C, panel II, lanes 17, 19).

Proteins involved in 28S rRNA maturation:

ITS2 processing was most obviously affected by the depletion of GNL2/Nog1, leading to the accumulation of 32S (B, panel III, lane 19); 10 other gene products were also involved in this processing. The involvement of GNL2/Nog1 in ITS2 processing was also manifested by the accumulation of the 12S, a species that also accumulated upon depletion of MDN1/Mdn1 and GLTSCR2/Nop53 (B, panel IV, lanes 19-20 and C, same panel, lane 19). Finally, several non-canonical species were detected. For example, a truncated 18S rRNA fragment in cells depleted for DHX37 (B, panel II, lane 13, see *) and extended versions of the 32S pre-rRNA (36S and 36S-C) typically detected in RPL26-defective DBA patients for MKI67IP and MAK16 (B, panel III, lanes 5 and 21 and see Figs 3 and 4).
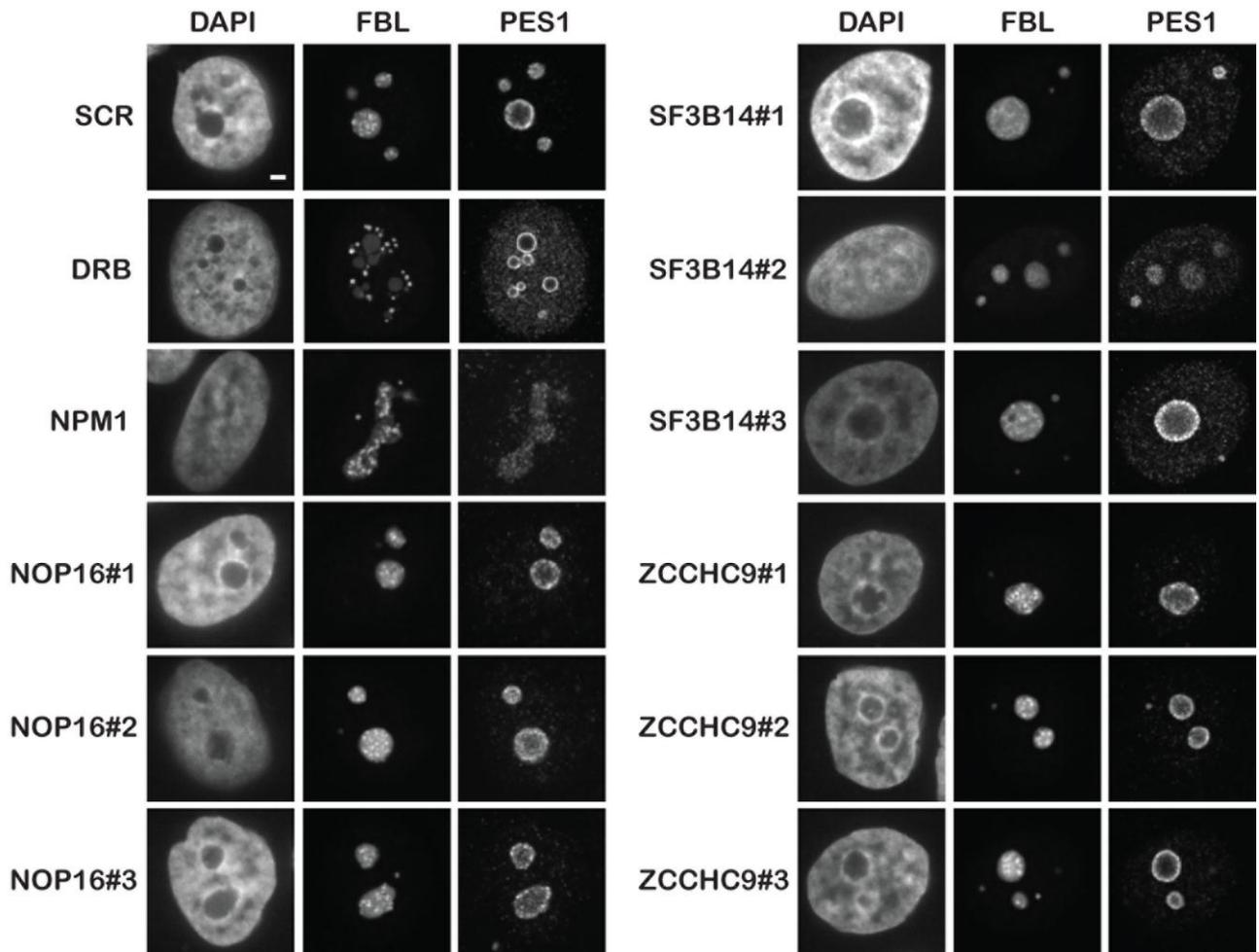
Figure S4. Human Genes without a Yeast Homolog Do Not Grossly Affect Nucleolar Structure, Related to Figure 4

To test for potential effects on nucleolar structure, the 11 proteins were depleted in HeLa cells stably expressing the dense fibrillar component (DFC) marker Fibrillarin (FBL) fused to GFP. Following a 3 d depletion, cells were fixed and immunolabeled with an antibody specific to Pescadillo (PES1), to detect the granular component (GC) of the nucleolus, and imaged by spin disc confocal microscopy. As a control for the disruption of nucleolar structure, the antitumor agent 5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) was added to cells. DRB treatment led to the formation of small fibrillarin foci extruding from the main nucleolar volume (counterstained with DAPI), while the GC component PES1 was partially relocated to the nucleoplasm and formed striking perinucleolar rings. Depletion of the control gene nucleophosmin (NPM1) also led to profound nucleolar reorganization with the formation of beaded necklaces. None of the tested proteins showed any gross alteration in nucleolar structure (shown for 3 genes and data not shown). In some instances, PES1 was partially relocated to the nucleoplasm, as seen upon SF3B14 depletion, and to a lesser extent with other genes (data not shown). Scale bar, 2 μm.
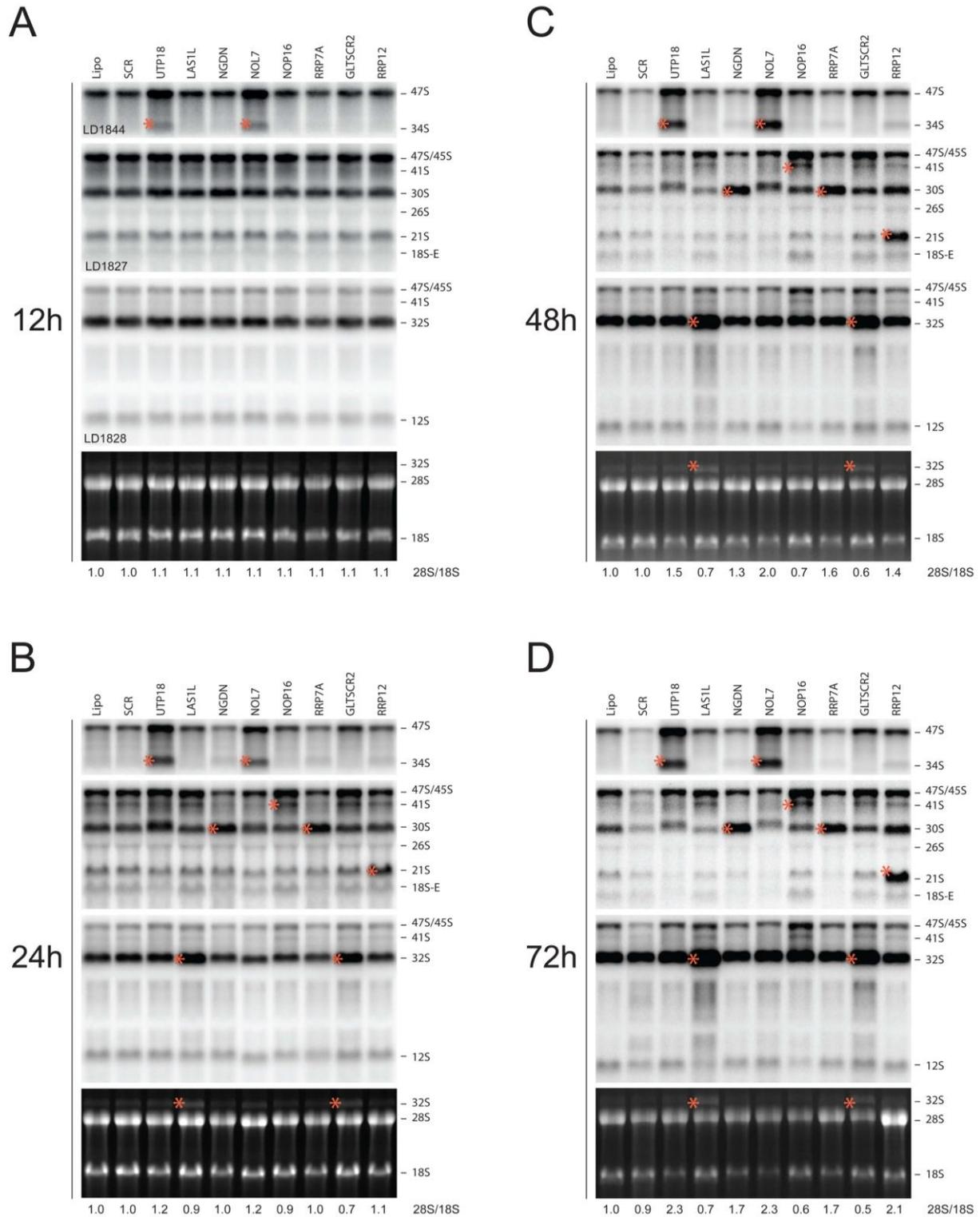
Figure S5. Pre-rRNA Processing Inhibitions Are Early Defects Detected as Soon as 24 hr, or Even 12 hr, after Depletion, Related to Figure 5

HCT116 p53 +/+ cells were depleted for the target genes in a time course analysis. Total RNA was extracted at the indicated time points and analyzed by Northern blotting. The mature rRNAs are shown on an ethidium bromide stained gel. Untreated (Lipo, lipofectamine) cells and cells treated with SCR siRNA were loaded as

controls. On relevant panels, a star highlights the RNA species diagnostic of the defect. 28S/18S ratio determined from Agilent electropherograms. **A** 12 h; **B** 24 h; **C** 48 h; **D** 72 h.
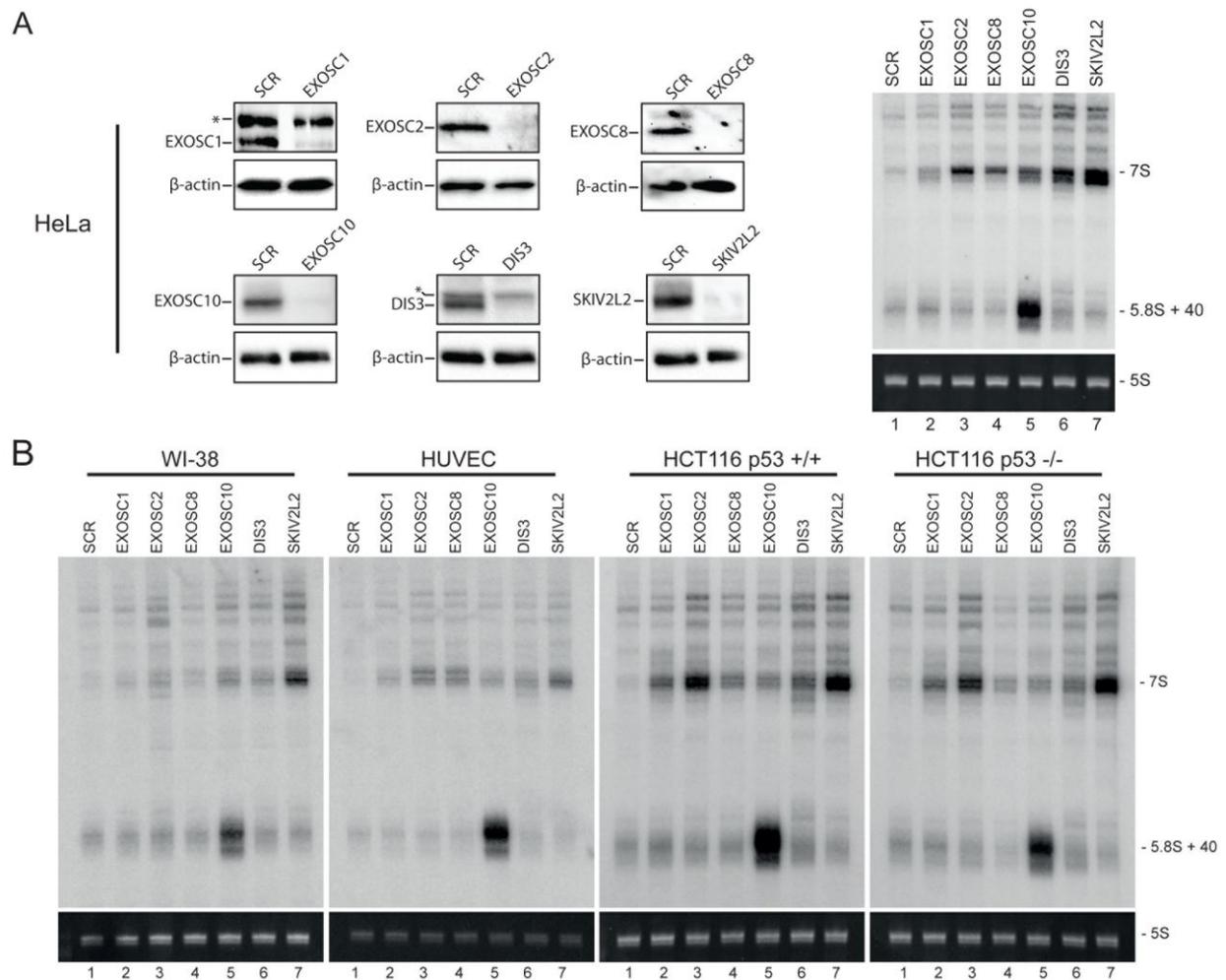

Figure S6. The Requirement for Exosome Subunits and Cofactors in ITS2 Processing Is Conserved in Different Cell Types, Related to Figure 6

A HeLa cells. Select exosome subunits and cofactors were depleted for 3 d in HeLa cells. Total protein was analyzed by Western blotting (left panels) with antibodies specific to the subunits and β-actin as loading control. In the EXOSC1 and DIS3 panels, a star denotes an unspecific cross-reacting band. Total RNA (5 μg) was analyzed by Northern blotting and hybridized with probe LD2079 (right panels). The 5S rRNA detected by ethidium bromide staining provides a loading control.

B WI-38, HUVEC, HCT116 p53 +/+ and HCT116 p53 -/- cells. Indicated cell types were depleted as in panel A. For all cell types, but HUVEC, 3 μg total RNA was loaded; 1 μg was loaded for HUVEC.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Phosphorimager Quantification

Northern blots were exposed to Fuji imaging plates (Fujifilm) and signals were acquired with a Phosphorimager (FLA-7000, Fujifilm). A calibration set (Fig 1) was loaded on each gel. The following rRNA fragments were quantified using MultiGauge (v3.1): the 47S pre-rRNA and aberrant 34S (with the 5' ETS probe LD1844); the 47S/45S, 41S, 30S, 26S, 21S and 18-E pre-rRNAs (ITS1 probe LD1827); and the 32S and 12S (ITS2 probe LD1828). Each gel lane and RNA band were identified manually. The background was set automatically using a polygonal lane with settings H ratio= 10 and V ratio= 75 and was then manually curated. The dataset was assembled in Excel and integrated by hierarchical clustering with "R" (http://www.r-project.org/index.html).

Mature rRNA Quantification

The 28S/18S ratio was determined using the Agilent RNA 6000 nano kit (REF 5067-1511) on a BioAnalyzer 2100 (Agilent).

Western Blotting

For immunoblotting, protein lysates corresponding to 6.0×10^4 cells were prepared as described in Castle et al., 2010. Protein concentration was evaluated with the Bradford assay (Bio-Rad). Total protein was separated by SDS–PAGE (Nu-Page, Invitrogen), transferred to PVDF (GE Healthcare) and analyzed by Western blotting (ECL, ThermoScientific). The following antibodies were used: ExoSC2 (Bethyl Laboratories), ExoSC8 (Abcam), ExoSC10 (Sigma-Aldrich), DIS3 (Bethyl Laboratories), β-ACTIN and p53 (Santa Cruz biotechnology). The ExoSC1 and SKIV2L2 antibodies were kindly provided by Dr Nicholas J. Watkins (Sloan et al., 2013). Images were acquired and processed with a ChemiDoc MP (Bio-Rad).

Shotgun qPCR

The efficiency of siRNA-mediated knockdown was determined by qPCR using a shotgun approach (8.2% of siRNA depletions tested altogether, see Fig S2), using GAPDH as an endogeneous control. Reactions were performed as previously described (Schillewaert et al., 2012). Primer sequences are listed in Table S9. Triplicate reactions were performed for each sample. Data were analyzed with the software StepOne (version 2.1) from Applied Biosystems (Life Technologies). The comparative threshold cycle method ($\Delta\Delta CT = \Delta C_T target - \Delta C_T scrambled$) was used to determine the relative abundance of mRNA transcripts.

Selection of 625 Targets

The list of candidate proteins was generated from the human nucleolar databases (Andersen et al., 2002; Scherl et al., 2002) and expanded using the bovine and plant *Arabidospis thaliana* nucleolar proteomes (Patel et al., 2010; Pendle et al., 2005), with selected entries from the Gene Ontology database (GO0042254 and GO0006364). A compilation of the yeast ribosome processing factors was generated mainly from Fromont-Racine et al., 2003; Kressler et al., 1999; Strunk and Karbstein, 2009. Human homologs were retrieved using a blastp search (expected p value < 10^{-15}). Note that a few yeast proteins have multiple human counterparts (especially helicases) and that several (30 proteins) do not have human homologs. From this assembled list of ~800 human proteins with putative functions in pre-rRNA processing we excluded all ribosomal proteins, as they were already described (O'Donohue et al., 2010), or are currently studied elsewhere, several histones and other proteins with conflicting subcellular localization. The final list comprised 625 candidates.

Generation of Heatmaps and Functional Cluster

The dataset was normalized as follows: each rRNA species was normalized to its corresponding fragment in the SCR sample loaded on the same gel. In the functional clustering analysis, all the calibration sets clustered consistently, providing validation of the grouping approach. The dataset was assembled into a matrix in Excel (see www.ribogenesis.com) and hierarchical clustering was done with the R freeware using the Gplots and plotrix packages (www.r-project.org/). The pre-RNA profiles obtained for each gene with all three independent siRNAs were systematically inspected in the functional cluster (see www.ribogenesis.com). Pre-rRNA processing profiles were considered similar when they were within a functional neighborhood of 10 genes.

Microscopy

A stable HeLa cell line, expressing a GFP-FBL fusion was used (a kind gift from Prof. D Hernandez-Verdun, Institut Jacques Monod, Paris). siRNA-mediated depletion was done as described above. As control, cells were treated with DRB (0.1 mM, 2 h). After 72h of depletion, cells were fixed with 2% formaldehyde for 15 min, and permeabilized with 2% BSA/0.2% Triton X-100 for 10 min. Unspecific binding was blocked with PBS/2% BSA for 30 min. PES1 was detected with a 1:1000 dilution of anti-Pes1 antibody (Ascenion), incubated for 2h at RT in a humidified chamber. After washing with PBS/2% BSA, the Alexa fluor 568 anti-rat (Invitrogen, A11077) secondary antibody was incubated in a dilution of 1:1000 for 2h at RT in a humidified chamber. Cells were washed with PBS and DNA was stained with 100 ng/mL DAPI (Sigma) for 10 min. Images were acquired using MetaMorph and a spin disc microscope (Yokogawa CSU-X1) with a HQ^2 Coolsnap Roper camera and a Plan-

Apochromat 100x, 1.46 numerical aperture objective lens. Exposure for DAPI, GFP and TxRed was 60 ms, 30 ms and 2 s, respectively.

Time Course Analysis

siRNA transfection: HCT116 cells were reverse-transfected as described for HeLa in 6-well plate but the number of cells seeded per well were modified according to the depletion time ($5x10^5$ cells for 12h depletion, $4x10^5$ cells for 24h, $3x10^5$ cells for 48h and $1.5x10^5$ cells for 72h). RNA were extracted using TRI reagent (Life Technologies). For Annexin V and caspase 3/7 assays (apoptosis assays), cells were transfected as above but in a 96-well format ($30x10^3$ cells for 12-h depletion, $20x10^3$ cells for 24-h, $10 x10^3$ cells for 48-h and $5 x10^3$ cells for 72-h). As a control, cells were treated with 1 μM camptothecin (Sigma). Immunodetection: for protein extraction, cells were washed in PBS 1x and lysed in 300 μl RIPA buffer (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% NP-40, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, supplemented by a protease inhibitor cocktail (Complete, Roche)) for 15 minutes on ice. Lysates were cleared by centrifugation. Proteins were separated by SDS-PAGE on precast gels (criterion TGX 4-20%, Bio-Rad) and transferred to a PVDF membrane (GE healthcare). β-actin and p53 (Santa Cruz biotechnology) antibodies were used to detect both proteins. Signals generated by chemiluminescence (Thermo Scientific) were detected using a ChemiDoc MP system (Bio-Rad) and quantified using ImageLab (Bio-Rad). Apoptosis: three assays were used to estimate the level of apoptosis. In the TUNEL assay, cells were treated according to the manufacturer's instructions (Apo-BrdU TUNEL assay, Life Technologies), stained by propidium iodide to follow the DNA and analyzed by flow cytometry (FACScantoII, BD). Cell-cycle was analyzed using FlowJo (TreeStar). For the Annexin V assay, we used Annexin-V-FLUOS (Roche) and analyzed differentially stained cells by microscopy and manual counting. Images were captured with a wide field fluorescent microscope (Observer.Z1, CoolLED excitation system) and an EC Plan-Neofluar 20x, 0.5 numerical aperture objective lens. For the caspase 3/7 assay, we used the ApoTox-Glo kit (Promega) and a Tecan luminometer plate reader.

SUPPLEMENTAL REFERENCES

Andersen, J.S., Lyon, C.E., Fox, A.H., Leung, A.K., Lam, Y.W., Steen, H., Mann, M., and Lamond, A.I. (2002). Directed proteomic analysis of the human nucleolus. Curr Biol *12*, 1-11.

Castle, C.D., Cassimere, E.K., and Denicourt, C. (2012). LAS1L interacts with the mammalian Rix1 complex to regulate ribosome biogenesis. Mol Biol Cell *23*, 716-728.

Castle, C.D., Cassimere, E.K., Lee, J., and Denicourt, C. (2010). Las1L is a nucleolar protein required for cell proliferation and ribosome biogenesis. Mol Cell Biol *30*, 4404-4414.

Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. Gene *313*, 17-42.

Kressler, D., Linder, P., and de La Cruz, J. (1999). Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae. Mol Cell Biol *19*, 7897-7912.

O'Donohue, M.F., Choesmel, V., Faubladier, M., Fichant, G., and Gleizes, P.E. (2010). Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. J Cell Biol *190*, 853-866.

Patel, A.K., Olson, D., and Tikoo, S.K. (2010). Proteomic analysis of bovine nucleolus. Genomics Proteomics Bioinformatics *8*, 145-158.

Pendle, A.F., Clark, G.P., Boon, R., Lewandowska, D., Lam, Y.W., Andersen, J., Mann, M., Lamond, A.I., Brown, J.W., and Shaw, P.J. (2005). Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. Mol Biol Cell *16*, 260-269.

Scherl, A., Coute, Y., Deon, C., Calle, A., Kindbeiter, K., Sanchez, J.C., Greco, A., Hochstrasser, D., and Diaz, J.J. (2002). Functional proteomic analysis of human nucleolus. Mol Biol Cell *13*, 4100-4109.

Schillewaert, S., Wacheul, L., Lhomme, F., and Lafontaine, D.L.J. (2012). The Evolutionarily Conserved Protein LAS1 Is Required for Pre-rRNA Processing at Both Ends of ITS2. Mol Cell Biol *32*, 430-444.

Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J., and Watkins, N.J. (2013). Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. J Cell Biol *200*, 577-588.

Strunk, B.S., and Karbstein, K. (2009). Powering through ribosome assembly. RNA *15*, 2083-2104.