Nucleolar structure evaluation and manipulation

Lafontaine, Denis; De Vleeschouwer, Christophe; Nicolas, Emilien; Parisot, Pascaline

Publication date:
2017

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
09 November 2017 (09.11.2017)

W!P O ¦ PCT

(10) International Publication Number
WO 2017/191187 A2

(51) International Patent Classification:
G06T 7/00 (2017.01) *G01N 33/50* (2006.01)

(21) International Application Number:
PCT/EP2017/060532

(22) International Filing Date:
03 May 2017 (03.05.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16168087.1 03 May 2016 (03.05.2016) EP

(71) Applicants: **UNIVERSITÉ LIBRE DE BRUXELLES** [BE/BE]; Avenue Franklin Roosevelt 50, 1050 Bruxelles (BE). **UNIVERSITÉ CATHOLIQUE DE LOUVAIN** [BE/BE]; 1, Place de l'Universite, 1348 Louvain-la-Neuve (BE).

(72) Inventors: **LAFONTAINE, Denis;** Avenue Leon Fournet, 9-204, 1342 Limelette (BE). **DE VLEESCHOUWER, Christophe;** Esplanade du Bon Air, 10, 5020 Vedrin (BE). **NICOLAS, Emilien;** Rue du Moulin, 70B, boite 3, 6740 Etalle (BE). **PARISOT, Pascaline;** Place Verte 60/201, 1348 Louvain-la-Neuve (BE).

(74) Agent: **DE CLERCQ & PARTNERS;** Edgard Gevaertdreef 10a, 9830 Sint-Martens-Latem (BE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: NUCLEOLAR STRUCTURE EVALUATION AND MANIPULATION

(57) Abstract: The nucleolus is a potent disease biomarker and a target in for instance cancer therapy. The present invention relates to how depletion of each of the eighty human r-proteins affects nucleolar structure, pre-rRNA processing, mature rRNA accumulation and p53 steady-state level. The present invention further relates to an image -processing program for qualitative and quantitative discrimination of normal from altered nucleolar morphology.

WO 2017/191187 A2

Nucleolar structure evaluation and manipulation

Field of the invention

The present invention relates to means and methods for evaluating and/or manipulating nucleolar morphology.

Background to the invention

Within the nucleus, the nucleolus is a specialized functional domain essential to gene expression (1). It is the site where the initial steps of ribosome biogenesis take place (2). Ribosomes are ribonucleoprotein (RNP) nanomachines converting the genetic information encoded in mRNAs into proteins. The human ribosome contains four ribosomal RNA (rRNAs) and eighty r-proteins organized in two subunits, each performing specialized functions in translation (3, 4). The small subunit (SSU), which consists of a single rRNA (18S) and thirty-three r-proteins, decodes the mRNA, while the large subunit (LSU), comprising three rRNAs (5S, 5.8S, and 28S) and forty-seven r-proteins, bears the peptidyl transferase center where amino acids are joined together into proteins. In the nucleolus, the 18S, 5.8S, and 28S rRNAs are synthesized by RNA polymerase I (Pol I) as long precursors, pre-rRNAs are modified, folded, and processed, and most r-proteins are assembled to form ribosomal subunits (2).

r-proteins are not involved in ribosome-mediated catalysis of peptide bond formation (3,5). Nonetheless, r-proteins play essential roles in shaping and maintaining the overall structure of the ribosomal subunits, and mutations in r-proteins are frequently associated with developmental disorders and human diseases (6). Notably, ribosomopathies are cancer predisposition syndromes caused by ribosome biogenesis dysfunction (7), due to mutations in r-proteins or ribosomal assembly factors. r-proteins are intimately linked to tumorigenesis, being directly involved in regulating the steady-state level of the anti-tumor protein p53 (8). This occurs via activation of specific anti-tumor surveillance pathways, through direct binding of specific r-proteins to the p53 regulator Hdm2 (see below and (9)).

The nucleolus is not limited by a lipid membrane. This makes it a highly dynamic structure that responds promptly, sometimes by profound morphological and compositional alterations, to cell stresses such as viral infections, DNA damage, and drug treatments (10,11). During interphase, the nucleoli of amniotic eukaryotes display three morphologically distinct layers (12, 13), which can be drastically re-organized under stress (14). During mitosis, the nucleolus undergoes a dramatic cycle of disassembly/reassembly that parallels Pol I activity controlled by specific phosphorylations (15,16). The number of nucleoli per cell nucleus and the shape and size of the nucleoli also vary greatly in proliferative diseases such as cancers (17). Cancer cells are more sensitive than non-cancer cells to inhibition of ribosome synthesis, and are killed selectively by treatment with Pol I inhibitors (18, 19). Despite the importance of the nucleolus as a cell stress sensor (20), disease biomarker and target for cancer therapy (21), how its structural integrity is maintained remains totally unclear.

While the principles of assembly and maintenance of the nucleolus are far from being understood (14), r-proteins which are very abundant, very basic, and which assemble mostly in the nucleolus onto pre-rRNAs to form ribosomal subunit precursors are likely to play an important role. The assembly of r-proteins is not random but follows a precise sequence of events. Groups of r-proteins have been defined on the basis of their assembly at early, intermediate, or late stages of ribosomal subunit biogenesis (22). Compromising the timely association of r-proteins with rRNA can indeed lead to severe pre-rRNA processing inhibitions, ribosomal subunit synthesis abortion, and sometimes to nucleolar structural alterations visible at the microscopic level (23). To date, no attempt has been made to systematically address the involvement of r-proteins in nucleolar structure maintenance or to grade their involvement in this process.

While the nucleolus is a long-known cancer biomarker (17) and a recently demonstrated therapeutic target (19), it is not widely used by pathologists, however, for lack of reliable clinical assays. Accordingly, there is a need in the art to improve qualitative as well as quantitative nucleolus assessment, in order to aid for instance in diagnostics, such as stratification of nucleolar morphology, but as well in the development of model systems for instance for use in drug development or evaluation, in particular associated with or otherwise dependent on nucleolar phenotypes, including morphology, such as nucleolar disruption or nucleolar integrity.

Summary of the invention

An object of the present invention is to provide methods and systems for discriminating populations of cells (e.g. normal/healthy versus abnormal/disease), for instance based on the analysis of the light captured by a sensor, when observing each population through an imaging device (e.g. a microscope).

To characterize nucleolar morphology defects both qualitatively and quantitatively, we developed a specific image-processing algorithm, of which a particular embodiment is detailed below. Briefly, we first segmented the observed nuclei on the basis of shape- and size-consistent adaptive thresholding of a nuclear stain (DAPI signal). Then, within each nuclear mass, GFP signal thresholding and mathematical morphology (see Methods) were applied to segment nucleoli into connected components. In order to optimize discrimination of nucleoli of cells depleted of an r-protein from those of SCR-treated control cells, five shape and textural features (or morphometric characteristics) were extracted from the largest connected components (LCC) of each nucleolus. These five features, selected from a set of eleven as the most discriminant ones, were: area, elliptical regularity, percentage of pixels below an optimized intensity threshold, smallest intensity, and number of local minima. For each of the five features, a d_k value corresponding to a statistically significant distance between the feature distribution in cells depleted of an r-protein and control cells was computed. Each population of cells was thus characterized by five d_k values. Principal component analysis (PCA) was used to reduce these five dimensions to two, allowing ready visualization of the data in a scatter plot (Fig. lb) where each dot corresponds to a population of cells treated with one siRNA.

Here we have depleted human cells systematically of each of the eighty r-proteins and investigated the consequences on nucleolar structural integrity, pre-rRNA processing, accumulation of mature rRNAs, and p53 steady-state level (see experimental strategy in Fig. la).

In summary, we show here that depletion of the vast majority of the eighty human r-proteins does not impact nucleolar structure. This notably applies to nearly all SSU r-proteins (Fig. 1). In striking contrast, about a third of the LSU proteins appear essential to maintaining normal nucleolar structure. This marked dichotomy is in line with the notion that pre-40S subunits are exported to the cytoplasm more rapidly than pre-60S subunits, whose production is more complex and requires numerous additional nuclear maturation steps. Among the strongest contributors to nucleolar structure are uL5 and uL18, known to form with 5S rRNA an Hdm2 trap and p53 stabilizer (25,26).

Most r-proteins assemble with pre-rRNAs within the nucleolus, quite early in the subunit assembly process. Notable exceptions are the acidic proteins uLlO (formerly P0), PI, and P2, which form the P-stalk on pre-60S subunits only after reaching the cytoplasm in yeast (47). Accordingly, we found the acidic r-proteins to have no impact on nucleolar structure (Supplementary Fig. 11).

While only a few r-proteins are required for nucleolar structure maintenance, most of them are essential to pre-rRNA processing (Fig. 4). The processing steps in which r-proteins are involved are primarily those which lead to synthesis of the rRNAs constituting the subunit to which they belong. The r-proteins whose depletion has the strongest impact on nucleolar structure are required for late processing reactions in the pathway of large subunit synthesis (Fig. 2a,b).

Pre-rRNA processing is an excellent proxy of ribosome assembly. Hence, by establishing the precise involvement of each r-protein in processing, we incidentally extend the conclusion (23,28,30,31,48) that the sequence of r-protein incorporation into maturing ribosomal subunits, and thus of ribosomal landmark formation, has been extremely well conserved throughout evolution from bacteria, to yeast and man (Supplementary Fig. 10). Importantly, we reveal this to be the case for human large ribosomal subunit assembly. This is quite remarkable, considering the tremendous differences in cell organization, in gene expression strategies and the increased complexity in ribosomal assembly machineries between prokaryotes and eukaryotes.

On mature 60S, the r-proteins whose depletion has the strongest impact on nucleolar structure form specific landmarks: the CP (uL5 and uL18), the L1-stalk, and a region directly below the L1-stalk (Fig. 2). These are late-assembling subunit structures. Furthermore, nucleolar structure is disrupted when uL5 or uL18 incorporation, and hence formation of the CP, is prevented by depletion of specific CP assembly factors (Fig. 5). We speculate that the importance of these r-proteins in maintaining the integrity of nucleolar structure reflects the emergence, during evolution, of checkpoints important to cell homeostasis, ensuring that the late steps of large subunit assembly, and particularly CP formation, occur properly.

Why should CP formation be monitored? Firstly, in the mature ribosome, the CP is involved in intersubunit interactions beneficial to translation(49,50). Furthermore, CP formation might be tightly coupled to maturation of essential ribosomal landmarks on the large subunit. This is plausible, given what is known about CP formation. The 5S RNP is incorporated into maturing 60S subunits as a pre-assembled

block (41), a step aided by the conserved assembly factors Rpf2(yeast)/BXDC1 (human) and Rrsl/RRSl (26,41,51) (see above). In precursor 60S, however, the 5S RNP does not adopt its final conformation until it undergoes a 180° rotation (51,52). This rotation seems to act as a power stroke promoting a cascade of subunit maturation events and the long-range transmission of mechano-chemical remodeling energy throughout the maturing 60S precursors (52,53). Structures whose formation may be strictly linked to that of the CP include the conserved A-site finger (ASF) helix 38, which is part of an intersubunit bridge that monitors the A-site tRNA throughout the decoding process (54), the peptidyl transferase center itself, where amino-acids are joined together, and possibly the phospho-stalk (52,53).

It is now well established that several r-proteins are essential to regulating the p53 level (9,43). In principle, depletion of any r-protein is expected to trigger a ribotoxic stress response leading to accumulation of unassembled ribosomal components, including uL5, uL18, and the 5S rRNA, and to sequestration of Hdm2. Therefore, it was expected that most r-proteins would be involved, in one way or another, in regulating the p53 level. In fact, setting a five-fold increase as significance threshold, we reveal that depletion of any one among two-thirds of the human r-proteins has no significant impact on p53 accumulation (Fig. 6). Nonetheless, we show that twenty-four r-proteins out of eighty are very important for p53 homeostasis, their depletion giving rise to a 5-to-10-fold increase in the p53 level. In all of these cases, the increase in p53 accumulation requires the presence of uL5 and uL18 (Supplementary Fig. 9). This implies activation of the anti-tumor nucleolar surveillance regulatory loop described above. Identification of the r-proteins whose depletion affects p53 accumulation provides essential insights into the etiology of ribosomopathies, which are cancer predisposition syndromes caused by mutations in r-proteins or by ribosome assembly defects (7).

Up to now, it has been unclear whether the activation of nucleolar surveillance, leading to p53 stabilization, systematically involves disruption of nucleolar structure or simply inhibition of nucleolar function. Induction of p53 in response to ribosome biogenesis inhibition has indeed been attributed to nucleolar disruption (55), but studies have also shown that a rise in the cellular level of p53 can occur after r-protein depletion, independently of gross nucleolar disruption (44,56). This is notably the case after uL30 (formerly: RpL7) depletion (44). We have confirmed this latter observation, showing that it applies, in fact, to a large group of twenty-one r-proteins (Table 5).

The nucleolus is a long-known cancer biomarker (17) and a recently demonstrated therapeutic target (19). It is not widely used by pathologists, however, for lack of reliable clinical assays. The image-processing algorithm and index of nucleolar disruption (iNo) developed here are robust and versatile tools for characterizing nucleolar morphological alterations both qualitatively and quantitatively. We have used them consistently in multiple cell lines, in time course analyses, and with either the dense fibrillar component or the granular component of the nucleolus (Fig. 3, Supplementary Figs S3,S4). We believe they hold great diagnostic and prognostic potential in cancer biology and research on ribosomopathies, several of which involve marked disruption of nucleolar integrity due to r-protein loss or mutation.

The complete dataset and additional information are accessible in a fully searchable information-rich database at www.RibosomalProteins.com as of the filing date of the present application.

The present invention is in particular captured by the appended claims, which are hereby also explicitly incorporated by reference.

Figure legends

FIG. 1: illustrates the systematic screening of human r-proteins, which reveals that uL5 (RPL11) and uL18 (RPL5) are the strongest contributors to nucleolar structure maintenance. Panel a) Experimental strategy: All eighty r-proteins were depleted one by one in human cells by use of specific siRNAs. The nucleolar structure (fluorescence microscopy), the accumulation of mature 18S and 28S rRNAs (electropherograms), pre-rRNA processing (high resolution northern blotting), and steady-state accumulation of p53 (fluorescent western blotting) were monitored. Panel b) Principal component analysis (PCA) showing a classification of r-proteins according to their requirement for nucleolar structure maintenance. Each r-protein was depleted in three knockdown experiments, each performed with a different siRNA. The image-processing algorithm that we designed for this analysis involves selecting five discriminant shape and textural features (or morphometric characteristics), computing five dk values, and reducing the five dimensions to two by PCA. In the resulting plot, each dot represents one population of cells treated with one siRNA. Dot intensity is indicative of the targeted protein: black shapes for small subunit (SSU) r-proteins and white shapes for large subunit (LSU) r-proteins. The mean of three populations of cells treated with a non-targeting control siRNA (SCR) is indicated with a line connected to the corresponding image. The grey symbols represent the six calibration controls (FBL, GFP, NCL, NPM, MOCK, and TIFIA, see Supplementary Fig. 1). Insets show images of the nuclei of cells depleted of representative proteins with the DNA stained brightly highlighted and the nucleoli appearing as black shapes (fibrillarin). For a few representative examples, a specific symbol is used (e.g., a diamond for uL5). RPS, r-proteins of the SSU; RPL, r-proteins of the LSU. Panel c) r-proteins and calibration controls classified according to the severity of nucleolar disruption caused by their absence. The iNo or index of nucleolar disruption was defined as the sum of the dk values of the five most discriminant shape and textural features identified in this work (LI -norm of the discrepancy vectors computed from the images presented in our experimental screens of 80 r-proteins (see Example 2). Higher iNo correspond to more severe disruption. Color-coding as in panel b; i.e., black for for small subunit (SSU) r-proteins and white for large subunit (LSU) r-proteins. The gray dots are the means of three individual experiments. Note: The r-proteins are named according to a recently revised nomenclature, where the "e" prefix stands for eukaryote-specific and "u" for universal (present in bacteria, archaea, and eukaryotes).

FIG. 2: depicts 3D models of human ribosomal subunits based on PDB entries 3J3D, 3J3A, 3J3F, and 3J3B. Late-assembling r-proteins of the large subunit are the strongest contributors to nucleolar structure maintenance and p53 homeostasis. The r-proteins are color-coded according to the impact of their depletion on nucleolar structure (iNo values) (a), pre-rRNA processing (b), or the p53 steady-state level (c). Left, subunit interface views; right, solvent-exposed views. The aminoacyl (A), peptidyl (P), and exit

(E) tRNA sites are indicated. Morphological features of the subunits are highlighted. On the LSU: the Ll-stalk, central protuberance (CP), and phospho-stalk (P-stalk). On the SSU, the beak (Be), head (H), platform (Pt), body (Bd), left foot (Lf), and right foot (Rf).

FIG. 3: illustrates the quantitative monitoring of nucleolar morphology in different human cell lines, based on detection of endogenous PES1. The data show, for a selection of eight representative r-proteins, that the r-proteins contributing weakly or strongly to nucleolar structure maintenance are largely the same in multiple cell lines.

(a) The indicated r-proteins were depleted with an siRNA for 3 days in two cervical carcinoma cell lines (HeLa-GFP-FBL, engineered to express fluorescent fibrillarin, and genetically unmodified HeLa), one colon carcinoma cell line (HCT116), and two lung carcinoma cell lines (A549 and H1944). Endogenous PES1 was detected by immunostaining with a specific antibody (see Example 1). As a control, cells were treated with a non-targeting control siRNA (SCR) and depleted of NPM (see also Supplementary Fig. 1). (b) Values of the nucleolar disruption index (iNo) obtained after 3 days of siRNA-mediated depletion of the indicated r-protein as calculated on the basis of the endogenous PES 1 signal.

FIG. 4: depicts schematically the involvement of human r-proteins in pre-rRNA processing. (a) The 28S/18S ratio calculated from Agilent bioanalyzer electropherograms. Data are shown for the two different siRNAs used (siRNA # 1 and #2). (b) Major pre-rRNA intermediates and probes used in this work. Three of the four rRNAs are produced by RNA Pol I as a long 47S primary transcript. The 18S, 5.8S, and 28S rRNAs are separated by noncoding external (ETS) and internal (ITS) transcribed spacers. Probes a, b, and c are the oligonucleotides LD1844, LD1827, and LD1828, respectively (see Example 1). (c) Pre-rRNA processing inhibitions upon depletion of SSU r-proteins. On the northern blots (see Supplementary Figs S5,S6, Sll), all RNA species were quantified with a Phosphorimager, normalized with respect to the non-targeting control (SCR), and their abundances represented on a heatmap using the color code indicated. The heatmap profiles were clustered with "R" and the corresponding proteins grouped in classes of r-proteins affecting the same or similar processing steps. The different siRNAs used are indicated (#). Asterisks refer to r-proteins assigned to two groups according to the siRNA used. (d) As in panel c for LSU r-proteins.

FIG. 5: reveals the central protuberance assembly factors BXDC1 and RRS1 are required for nucleolar structure integrity. (a) Cells expressing fibrillarin fused to the fluorescent protein were treated for 3 days with an siRNA targeting transcripts encoding the indicated protein. Two independent siRNAs (#1 and #2) were used in each case. Cells treated with a non-targeting (SCR) siRNA control are shown for reference. (b) For each depletion, the nucleolar disruption index (iNo) was calculated (see Fig. 1 and Example 1).

FIG. 6: illustrates the involvement of human r-proteins in p53 homeostasis. (a) Steady-state level of p53 determined by quantitative fluorescent western blotting. Western blots analysis are shown for representative r-proteins, with the p53 level indicated underneath as a mean of biological triplicates obtained after treatment of cells with the same siRNA (i, ii, and iii). The siRNA used was selected on the basis of its proven efficacy in the processing and nucleolar screens (Figs 1 and 4). The p53 signal corrected for loading (using β-actin as reference) was expressed with respect to the level observed in cells

treated with a non-targeting siRNA control (p53 +/+). The upper signals correspond with p53; the bottom signals with β-actin. A complete data set for all eighty r-proteins is available in Supplementary Fig. 11. As loading control we used HCT116 p53+/+ cells transfected with a non-targeting siRNA (p53 +/+) providing the basal level of p53 or with an antisense oligonucleotide suppressing the activity of the box C/D snoRNA U8 (#U8), thereby stimulating p53 accumulation up to 6-fold (J.-L.L, E.N. and D.L.J.L., submitted). As background control, we used a matched isogenic HCT116 cell line that does not express p53 (HCT116 p537-, ref.39) treated with a non-targeting siRNA (p53 -/-).

(b) r-proteins classified according to their impact on the p53 steady-state level. The non-targeting (SCR) control is shown in black, the SSU r-proteins in white, and the LSU r-proteins in grey. The histogram bars are the means of triplicates with SD. r-proteins whose depletion leads to a 5-fold increase in p53 level are highlighted in a gray box.

Supplementary FIG. 1: The nucleolar calibration set. The calibration set used consisted of four control proteins whose depletion, we established, strongly disrupts nucleolar structure. These control proteins are: the RNA polymerase I (Pol I) transcription factor TIF1A, nucleolin (NCL), and nucleophosmin (NPM). As further standardization controls, we used mock-treated (MOCK) cells, cells treated with an siRNA targeting FBL or GFP, and cells treated with a non-targeting siRNA (SCR). White signal, DNA stain (DAPI); black signal within white, GFP. Left column, images captured at 20x magnification in widefield mode. Right column, images captured at 40x magnification in confocal mode. To the right of these images, schematics depicting the effect of siRNA-mediated depletion on nucleolar structure and signal intensity.

Supplementary FIG. 2: Benchmarking the automated unsupervised classification of nucleolar disruption phenotypes. To benchmark our novel classification algorithm, we compared the automated classification described in the text with a manual one. The manual classification was based on the fine visual inspection of representative images obtained after depletion with each siRNA used and on the assignment of nucleolar disruption phenotypes to three arbitrarily defined classes corresponding to weak, intermediate, and strong disruption. Superimposition of the automatically and manually obtained classifications made us highly confident that assignment to phenotypic classes on the basis of our automated procedure is robust. (a) PCA (see Fig. lb). Each circle represents a population of cells treated with one siRNA specific to one r-protein. The circles are intensity-coded according to a manual classification based on the fine visual inspection of the microscopic images. The non-targeting control (Scramble, SCR) is indicated with an arrow. Strong contributors to nucleolar structure maintenance are black; intermediate contributors are in grey; weak contributors are in white. (b) Representative images illustrating the manual classification of r-proteins as weak, intermediate, and strong contributors to nucleolar structure maintenance. White signal, DNA stain; black signal, fibrillarin.

Supplementary FIG. 3: Kinetics of nucleolar disruption after siRNA-mediated r-protein depletion. The data show that the computed iNo values reliably reflect phenotype severity, and that nucleolar disruption is best scored in cells having undergone two successive rounds of nucleolar breakdown/nucleolar genesis, corresponding to two cell divisions (~72h). (a) Values of the nucleolar disruption index (iNo) obtained

after siRNA-mediated depletion of the indicated r- protein for 24 h, 48 h, and 72 h, which are white, grey, and black, respectively. (b) Representative images for thirteen r-proteins tested at each time point. White signal, DNA stain; Black signal, fibrillarin. #1 refers to the siRNA used (see Example 1).

Supplementary FIG. 4: Quantitative monitoring of nucleolar morphology based on detection of the endogenous granular component marker PES1. The data show that the iNo values can be computed equally on the basis of DFC (fibrilarin, FBL) or GC (PES1) antigen detection. (a) Values of the nucleolar disruption index (iNo) obtained after 3 days of siRNA-mediated depletion of the indicated r-protein as calculated on the basis of the fibrillarin signal (white bars) or the PES1 signal (black bars). (b) Representative images for the thirteen r-proteins tested. In each case the siRNA n°l (#1) was used (see Example 1). White signal, DNA stain; black signal, fibrillarin (top rows) or PES1 (bottom rows).

Supplementary FIG. 5: Involvement of small subunit r-proteins in pre-rRNA processing. (a) Representative examples of northern blots for each of the three classes of SSU r-proteins defined in this work (for a full dataset see Supplementary Fig. 11). A calibration set consisting of mock-treated cells and cells treated with a non-targeting siRNA (SCR) or a siRNA targeting UTP18 or NOL9 was used systematically (see ref.l). Schematics of the RNA intermediates detected are shown on the left. Ratios of 28S to 18S mature rRNA were calculated from bioanalyzer electropherograms. (b) Expanded version of Fig. 4c, showing all RNA intermediates detected and quantified.

Supplementary FIG. 6: Involvement of large subunit r-proteins in pre-rRNA processing. (a) Representative examples of northern blots for each of the four classes of LSU r-proteins defined in this work (for a full dataset see Supplementary Fig. 11). A calibration set consisting of mock-treated cells and cells treated with a non-targeting siRNA (SCR) or a siRNA targeting UTP18 or NOL9 was used systematically (see ref.l). Schematics of the RNA intermediates detected are shown on the left. 28S/18S mature rRNA ratios were calculated from bioanalyzer electropherograms. (b) Expanded version of Fig. 4d, showing all RNA intermediates detected and quantified.

Supplementary FIG. 7: Comparison of our classification of r-proteins according to their involvement in pre-rRNA processing with previous studies. The figure shows that our work either confirms (small subunit r-proteins) or considerably complements (large subunit r-proteins) the literature. 3-D models of human ribosomal subunits based on PDB entries 3J3D, 3J3A, 3J3F, and 3J3B. Left, subunit interface views; right, solvent-exposed views. The aminoacyl (A), peptidyl (P), and exit (E) tRNA sites are indicated. Morphological features of the subunits are highlighted. On the LSU: the Ll-stalk, central protuberance (CP), and phospho-stalk (P-stalk). On the SSU, the beak (Be), head (H), platform (Pt), body (Bd), left foot (Li), and right foot (Rf). (a) Previous studies: conducted on cervix cancer cells where p53 expression is disrupted by HPV integration (HeLa cells). The SSU r-proteins were tested in ref. 2; six out of the forty-seven LSU r-proteins were tested in ref. 3. (b) This work: conducted on colon carcinoma cells expressing p53 normally (HCT116 p53+/+)(based on Fig. 4 and Supplementary Figs. S5,S6,S11). All r-proteins were tested.

Supplementary FIG. 8: Efficiency of r-protein depletion established at the mRNA level by RTqPCR. For forty-eight r-proteins whose depletion did not significantly affect p53 accumulation (see Fig. 6b), the

residual level of mRNA was established by RTqPCR and found to be below 20% for forty of them, and to range between 20% and 45% for the remaining eight. Total RNA was extracted from HCT116 cells treated for 2 days with an siRNA specific to transcripts encoding the indicated r-proteins. Residual levels of mRNA were established by RTqPCR and normalized to those observed in cells treated with a non-targeting siRNA control (SCR). Each experiment was performed in triplicate.

Supplementary FIG. 9: Depletion of twenty-four r-proteins out of eighty leads to a significant five-fold-increased level of p53, and this increase requires the presence of uL5 and uL18. HCT116 p53+/+ cells were depleted for 2 days with an siRNA specific to transcripts encoding the indicated protein. Each protein was depleted by itself, or in combination with uL18 or uL5 depletion. As control, cells were treated with a non-targeting siRNA (SCR). Total protein was extracted and processed by western blotting with antibodies targeting p53, p21 (a transcriptional target of p53), and, as loading control, β-actin. Bands were revealed by luminescence. (a) Effect of individually depleting each of the 24 r-proteins. (b) Effect of depleting the central protuberance assembly factor BXDC1 or RRS1.

Supplementary FIG. 10: The sequence of assembly of r-proteins onto assembling ribosomal subunits has been remarkably conserved throughout evolution. (a) 3-D models of bacterial ribosomal subunits based on PDB entries 2AVY and 2AW4. The r-proteins are intensity-coded according to their early, intermediate, or late order of assembly, as established in vitro (refs 4-8). (b) 3-D models of budding yeast ribosomal subunits based on PDB entries 3U5B, 3U5C, 3U5D, and 3U5E. The r-proteins are labeled on the basis of their being required for specific early, intermediate, or late pre-rRNA processing steps in yeast cells (refs 9, 10). This largely corresponds to their kinetics of assembly established in vivo (ref. 11). (c) 3-D models of human ribosomal subunits based on PDB entries 3J3D, 3J3A, 3J3F, and 3J3B. The r-proteins are classified with respect to their impact on specific processing reactions in human cells according to this work (based on Fig. 4, Supplementary Figs. S5,S6,S11). Left, subunit interface views; right, solvent-exposed views. The aminoacyl (A), peptidyl (P), and exit (E) tRNA sites are indicated. Morphological features of the subunits are highlighted. On the LSU: the Ll-stalk, central protuberance (CP), and phospho-stalk (P-stalk). On the SSU, the beak (Be), head (H), platform (Pt), body (Bd), left foot (Lf), and right foot (Rf).

Supplementary FIG. 11: An example of a complete data sheet for one r-protein. The datasheets for all the eighty human r-proteins is available on the companion website at www.ribosomalProteins.com. The data sheet shows the position of the r-protein on mature subunit (A), its impacts on pre-rRNA processing and mature rRNA accumulation (B), on nucleolar structure (C), and on the p53 steady-state level (D).

(A) 3-D models of human ribosomal subunits based on PDB entries 3J3D and 3J3A, for SSU r-proteins, and 3J3F and 3J3B, for LSU r-proteins. The positions of individual r-proteins on the mature subunits are highlighted. On small subunits, the 18S rRNA is shown in gray; on large subunits, the 5S, 5.8S, and 28S rRNAs are shown in red, blue, and gray, respectively. Left, interface view; right, solvent view. The main ribosomal features are indicated (see Supplementary Fig. 10). (B) Effects of r-protein depletion on pre-rRNA processing and mature rRNA accumulation: northern blots and ethidium-bromide-stained denaturing agarose gels showing all the pre-rRNA intermediates and mature rRNAs detected. The

28S/18S rRNA ratio was calculated from electropherograms. A calibration set (described in Supplementary Figs S5,S6) is included for reference. Schematics representing the RNAs detected are shown to the left. Quantifications are available in Fig 4. and Supplementary Figs S5,S6. (C) Effect of r-protein depletion on nucleolar structure: representative microscopic images (blue signal, DAPI; green signal, fibrillarin) obtained after treatment in duplicate screens (i and ii) performed with three different siRNAs (#1, #2, and #3). iNo values for each screen are indicated to the right and on a scaled bar at the bottom. The iNo value ranges between 0 (unperturbed nucleolus) and 0.2 (severely disrupted structure). (D) Effect of r-protein depletion on the p53 steady-state level: fluorescent quantitative western blotting was performed in triplicate (i, ii, iii). The p53 signal was corrected for loading, using D-actin detection as a reference, and expressed with respect to the signal obtained in cells treated with a non-targeting (Scr) control (lane 2). The p53 signal was expressed as a mean of three independent experiments (see Fig. 6 for details). A calibration set (described in Fig. 6) is included for reference.

Supplementary FIG. 12: Examples of uncropped Northern blots. An example of a high resolution denaturing agarose gel is shown. (a) Ethidium bromide staining reveals the mature 18S and 28S rRNAs. (b,c,d) Northern blotting with specific probes reveals the pre-rRNAs. All RNA species were identified by differential hybridization with specific probes, and by reference to a calibration set consisting of two proteins (UTP18 and NOL9) whose depletion leads to well-characterized pre-rRNA processing inhibitions, as described in ref. 1. The probes used (ITSl in b, ITS2 in c, and 5'-ETS in d) are described in Example 1. Lane 1, mock; lane 2, SCR; lane 3, UTP18; lane 4, NOL9; lane 5, uS3; lane 6, uS17; lane 7, uL30#l; lane 8, uL30#2; lane 9, uL13#l; lane 10, uL13#2; lane 11, eL14#l; lane 12, eL14#2; lane 13, ul22#l; lane 14, ul22#2; lane 15, eL18#l; lane 16, eL18#2; lane 17, eL22#l; lane 18, eL22#2; lane 19, uL14#l; lane 20, uL14#2; lane 21, uL24#l; lane 22, uL24#2. #1 and #2 refer to the siRNAs used (see Example 1).

Supplementary FIG. 13: Examples of uncropped Western blots. All western blot hybridizations performed in this work used well-characterized commercially available antibodies (see Example 1). On fluorescent (a) and luminescent (b) detections, p53 migrated between the 35 kDa and the 55 kDa molecular weight bands, as expected (p53 has a molecular weight of 43.7 kDa). The p53 signal was increased upon nucleolar stress activation (U8 depletion in a, and uLl depletion in b, by comparison to the signals observed in cells treated with a non-targeting Scr control). The p53 level was severely reduced upon codepletion of uLl with uL5 or uL18 (b). On the fluorescent screening gels (representative examples shown in c), the β-actin loading control (bottom rows) was detected as a single band immediately below the p53 signal, as expected (β-actin has a molecular weight of 42 kDa). The p53 band (top row) was consistently detected at low levels in the duplicated HCT116 p53 +/+ cells treated with a non-targeting siRNA (p53 +/+ lanes), this served as a baseline control. The p53 levels increased substantially upon nucleolar stress caused by U8 snoRNA depletion (#U8 lanes), this served as a positive control. The p53 band was never detected in the negative control provided by the HCT116 p53 -/- isogenic cell line (p53 -/-). The stars denote nonspecific bands.

Supplementary FIG. 14: Nuclei aggregates are eliminated on the basis of a shape convexity analysis. Two segmented juxtaposed nuclei are illustrated. In the direction of the principal axis X and Y, if 8out, δï iη, and 82in are above a threshold on one of the lines parallel to the principal axis, the shape is considered to include more than one nucleus.

Supplementary FIG. 15: Results of segmentation of nuclei for highly contrasted (a) and weakly contrasted (b) DAPI images. The connected components are detected based on the proposed hierarchical thresholding. Irregular regions are rejected because of their concavity. Some regions are rejected because the DAPI/GFP intensity analysis indicates that cells are probably not in interphase. Only remaining nuclei circled are maintained in our analysis, and for the subsequent analysis, we consider the segmentation, obtained after morphological operations: 13x13 dilation, followed by 3x3 erosion.

Supplementary FIG. 16: Samples of nucleoli from SCR-treated control cells (a), from cells with a range of high level of nucleolar disruption (b) and from cells depleted of a specific protein of interest (c). Cell nuclei in each column were treated with the same siRNA. Images were normalized by percentile 99.9%.

Supplementary FIG. 17: Fibrillarin-GFP intensity profile of nucleoli in a SCR-treated control cell (a) or in a cell depleted for the protein uL5 (b). (a) corresponds to the nucleus in row 1, column 2 in Supplementary Fig. 16, panel a. (b) corresponds to the nucleus in row 1, column 5 in Supplementary Fig. 16, panel c.

Supplementary FIG. 18: Illustration of the shape factors associated to a connected component C (inner shape black contour). (a) The elongation factor reflects the ratio between the principal axes second order moments, while the elliptical regularity measures the area ratio of the smallest external ellipsoid $C_{ellipse}$ (outer shape in grey contour) to the connected component. (b) The concavity ratio measures the area ratio between the convex hull $C_{\infty\ n_y huii}$ (outer shape in grey contour) and the connected component.

Supplementary FIG. 19: Distribution of local maxima (white dots) and local minima (black dots) in the images shown in Supplementary Fig. 16. It is apparent that the density of local minima is higher in nucleoli from cells with high level of nucleolar disruption (panel b) or from cells depleted of a protein of interest (panel c) than in nucleoli from SCR-treated control cells (panel a).

Supplementary FIG. 20: Fisher's optimization criterion as a function of three features (a, b, and c) parameters.

Detailed description

Before the present system and method of the invention are described, it is to be understood that this invention is not limited to particular systems and methods or combinations described, since such systems and methods and combinations may, of course, vary. It is also to be understood that the terminology used herein is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

The terms "comprising", "comprises" and "comprised of" as used herein are synonymous with "including", "includes" or "containing", "contains", and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps. It will be appreciated that the terms "comprising", "comprises" and "comprised of" as used herein comprise the terms "consisting of, "consists" and "consists of".

The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

The term "about" or "approximately" as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, is meant to encompass variations of +/-10% or less, preferably *+1-5%* or less, more preferably +/-!% or less, and still more preferably +/-0.1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. It is to be understood that the value to which the modifier "about" or "approximately" refers is itself also specifically, and preferably, disclosed.

Whereas the terms "one or more" or "at least one", such as one or more or at least one member(s) of a group of members, is clear *per se*, by means of further exemplification, the term encompasses *inter alia* a reference to any one of said members, or to any two or more of said members, such as, *e.g.*, any >3, >4, >5, ≥6 or ≥7 etc. of said members, and up to all said members.

All references cited in the present specification are hereby incorporated by reference in their entirety. In particular, the teachings of all references herein specifically referred to are incorporated by reference.

Unless otherwise defined, all terms used in disclosing the invention, including technical and scientific terms, have the meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. By means of further guidance, term definitions are included to better appreciate the teaching of the present invention.

In the following passages, different aspects of the invention are defined in more detail. Each aspect so defined may be combined with any other aspect or aspects unless clearly indicated to the contrary. In particular, any feature indicated as being preferred or advantageous may be combined with any other feature or features indicated as being preferred or advantageous.

Reference throughout this specification to "one embodiment" or "an embodiment" means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, appearances of the phrases "in one embodiment" or "in an embodiment" in various places throughout this specification are not necessarily all referring to the same embodiment, but may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to a person skilled in the art from this disclosure, in one or more embodiments. Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the invention, and form different embodiments, as would be understood by those in the art. For example, in the appended claims, any of the claimed embodiments can be used in any combination.

In the present description of the invention, reference is made to the accompanying drawings that form a part hereof, and in which are shown by way of illustration only of specific embodiments in which the invention may be practiced. Parenthesized or emboldened reference numerals affixed to respective elements merely exemplify the elements by way of example, with which it is not intended to limit the respective elements. It is to be understood that other embodiments may be utilised and structural or logical changes may be made without departing from the scope of the present invention. The following detailed description, therefore, is not to be taken in a limiting sense, and the scope of the present invention is defined by the appended claims.

In an aspect, the invention relates to a method for characterizing nucleolar morphology, architecture, or integrity, comprising measuring or determining one or more morphometric characteristics of the nucleolar support area and the nucleolar intensity pattern. In certain embodiments, the morphometric characteristics are defined by parametric functions. In certain embodiments, the method comprises characterizing, in a measurable and quantifiable manner, the visual appearance of the nucleolus in an image, including morphology or form (encompassing size, shape, and number of subcomponents) and inner pattern structure (encompassing spatial organization and spatial variation of pixel intensity).

For example, individual cell nuclei and/or cell nucleoli may be segmented based on light intensity captures. Light intensity also reveals the nucleolar masses composing each cell nucleolus. Each population of cells may be described by the stochastic distributions of a set of discriminant nucleolar features (i.e. morphometric characteristics). A nucleolar feature is a scalar value that reflects one particular aspect of the spatial organization of nucleolar masses within one individual cell nucleus.

In certain embodiments, a feature discriminates between two populations of cells when the statistical distribution of the feature values measured over the two populations have a limited overlap, as for example measured by a large difference of averages compared to their mean standard deviation.

A set of nucleolar features that discriminate between two types of cell populations may in an aspect or an embodiment be derived in the following steps:

1) A set of arbitrary parametric features is defined, e.g. by an expert in the field, to capture the main salient differences observed between representative nucleoli samples of each population;

2) The parameters associated to each feature are then selected to minimize the overlap between pairs of exemplars of the populations to discriminate;

3) Only the features, and their associated parameters, leading to a large separation between populations are kept as discriminant features.

In an embodiment, such method comprises the steps of:

a) defining a set of arbitrary parametric morphometric characteristics to capture the most salient morphometric differences observed between representative nucleoli samples of two populations of interest;

b) selecting the parameters associated to each morphometric characteristic to maximize their discriminating power and/or to minimize the overlap between the morphometric characteristic probabilistic distributions derived from each population, preferably according to Fisher's criterium; and

c) selecting the morphometric characteristics and their associated optimized parameters leading to a large separation between populations to discriminate;

d) computing the distance between the two cells or cell populations based on a combination, e.g. using a weighted summation, of the distances measured in terms of the selected morphometric characteristics.

Accordingly, in an aspect, the invention relates to the use of the above method for identifying features that discriminate between two types of cell populations, such as identifying features that discriminate between two types of cell populations with different nucleolar morphology or architecture.

As an example to distinguish populations of cells with regard their nucleolar morphology, the present invention provides in certain embodiments a set of selected parametric features, their optimization, and combination. In certain embodiments, the combination results in the (weighed) sum of the absolute value of a normalized distance between the feature distributions (e.g. the Fisher score) associated to the population of cells we want to compare. Other features and combinations are possible.

The features that are envisioned to capture the main salient differences between the light intensity patterns revealing the nucleolar structure of cell samples are in certain embodiments:

- Features that characterize the size and/or shape of the spatial support of the nucleolar masses. Those features preferably count the number of distinct connected components contributing to the nucleoli, and/or the number of pixels that are part of each of those components. They may also include any quantitative metric that reflects the elongation and/or elliptical regularity and/or concavity of the nucleolar masses. All those features are preferably parametrized by the level of light intensity above which a pixel is considered to be part of the nucleolar masses.
- Features that characterize the spatial variation of the nucleolar signal within the nucleolar connected components. Spatial variations may be measured inside and/or on the boundary of connected components. In the neighbourhood of the boundary of a connected component, the sharpness of transition of nucleolus-related light intensity between the nucleolar mass and the background image may be measured. Inside a connected component, the features may be defined to capture the spatial variations of light intensities. In practice, the magnitude and number of local minima and local maxima of nucleolus-related light intensity values may be considered, as well as in addition or in the alternative the number of pixels having a relatively small intensity value and/or the number and shape of high intensity sub-regions that are separated by small intensity valleys inside a connected component. All those features may be parametrized by the thresholds adopted to define the notions of small/high intensity, and/or by the structuring elements considered to define the notions of inside/outside a component, preferably based on mathematical morphology.

As used herein, the terms "morphology" and "architecture" may generally refer to the overall appearance of nucleoli, such as for instance identified on images of cells or tissues. "nucleolar integrity" and the related "nucleolar disruption" may generally refer to the similarity or dissimilarity of the nucleolar morphology or architecture in a sample cell (population) or tissue compared to a reference nucleolar morphology or architecture. Preferably, such reference nucleolar morphology or architecture is a

"normal" nucleolar morphology or architecture, e.g. the nucleolar morphology or architecture of non-diseased cells, or cells known not to have an abnormal nucleolar morphology or architecture (and as such may also include diseased cells, but which do not have an abnormal nucleolar morphology or architecture). In certain embodiments, a standard or reference nucleolar morphology may be cell type specifically determined (e.g. for neuronal cells, epithelial cells, etc.). As used herein, the term "morphometric characteristic" may generally refer to any feature capable of describing the visual appearance, including morphology or form (encompassing size and shape) and inner pattern structure (encompassing spatial organization and variation), and which is measurable and quantifiable.

The term "nucleolar support area" may generally refer to the physical location of the nucleolus, or the area enclosing a nucleolus, e.g. on an image, as defined herein elsewhere. The term "nucleolar intensity pattern" may generally refer to specific intensity-associated morphometric features and their distributions, as described herein elsewhere.

In certain embodiments, the morphometric characteristics are determined on images of cells, cell populations, or tissues. In certain embodiments, the morphometric characteristics are determined on images of tissue sections. Cells or tissues may be prepared by techniques known in the art, and which allow cell or tissue imaging. In certain embodiments, the images are or include fluorescent components, e.g. DAPI, for nuclear staining, or fluorescent labels (e.g. antibodies) for specific cell structures or molecules (e.g. nucleolus-specific components, structures, or molecules). The present invention is however not limited to the use of fluorescent imaging. The skilled person can readily envisage alternative imaging methods and techniques. Also in vivo imaging methods are possible in certain embodiments according to the invention (see for instance US 2015/157254).

In certain embodiments, the method according to the invention as described herein is used to quantitatively score nucleolar morphology, architecture, integrity, or disruption.

In certain embodiments, the method according to the invention as described herein is used to qualitatively or quantitatively classify, score, diagnose, and/or grade disorders underlying and/or characterized by nucleolar disruption, or abnormal nucleolar morphology or architecture. In certain embodiments, such disorder is cancer, autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems).

In certain embodiments, the according to the invention as described herein is used to classify, score, diagnose, and/or grade a cell (population) or tissue, based on the severity of nucleolar disruption, or the deviation or difference of nucleolar integrity compared to a standard or reference.

In certain embodiments, the characterization of nucleolar morphology according to the invention as described herein is associated with or underlies a particular prognosis.

In certain embodiments, the characterization of nucleolar morphology according to the invention as described herein is associated with or underlies a particular treatment plan.

In an aspect, the invention relates to a method for scoring, quantifying, or classifying nucleolar disruption and/or for determining and/or prognosing the health status of an individual or of a cell or tissue or for diagnosis, comprising performing the method according to the invention as described herein. In certain

embodiments, the iNo value as described herein elsewhere is indicative of the severity of said nucleolar disruption, or the severity or grade of said health status. In certain embodiments, said health status or diagnosis is indicative of cancer/tumor/malignancy/proliferative disease or disorder (or pre-cancerous/pre-neoplastic/pre-malignant disorders or stages, such as hyperplasia, dysplasia, or metaplasia), autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems).

In certain embodiments, said method comprises obtaining or providing a cell or tissue sample of a subject or patient.

In an embodiment, the invention relates to a method according to the invention as described herein, comprising the steps:

a) obtaining or providing an image of one or more eukaryotic cell, or eukaryotic tissue, or preferably of a population of cells ;

b) optionally segmenting said image thereby obtaining individual cells or cell nuclei;

c) segmenting said image, thereby obtaining one or more nucleolar masses, preferably comprising a plurality of connected pixels, which connected pixels constitute or represent a nucleolar connected component;

d) optionally assigning by colocalization nucleolar masses to individual cells or cell nuclei and thus to individual cells;

e) extracting one or more feature (i.e. morphometric characteristic) characterizing the support area, and the intensity pattern of connected components;

f) computing probabilistic distributions of said features (e.g. for cell populations); and

g) comparing said probabilistic distributions with a reference model distribution and/or determining the overlap of said probabilistic distributions to a reference model distribution.

In the above method, b) and/or d) may both be optional. The method above may thus comprise b); b) and d); d); or none of b) and d).

As used herein, the term "connected pixels" refers to adjacent or neighbouring pixels on a graph or image, i.e. pixels which are next to each other (in any direction). According to certain embodiments, the set of connected pixels (and hence by extension the segment or component or connected segment or component) are or comprise or consist of adjacent pixels each having an intensity above a given threshold. Each pixel is a node, and the graph or image edges are directly derived from the definition of a neighborhood system, which typically assigns 4 (top, down, left, right) or 8 neighbors (4 + oblique directions) to each pixel. A connected component (or just component) of an undirected graph or image is a subgraph or subimage in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the supergraph or superimage. A vertex with no incident edges is itself a connected component. Hence, a nucleolar connected component is a set of nucleolar pixels in which any two pixels can be connected through a path (a sequence of pixel adjacency relationships) that only include nucleolar pixels.

In certain embodiments, such method further comprises treating a subject if the nucleolar morphology is determined to be affected, abnormal, or disrupted.

In an embodiment, as an initial step to delimit cell nucleoli from individual cells, the cell nucleus is segmented. Each cell has a single nucleus. The nucleoli are specialized subnuclear domains, which, by definition, are all contained within the nucleus. Segmenting the nucleus, which is dense, compact and easy to score, is thus a mean to delimit the cellular volume that contains nucleoli in each individual cell.

Nuclei may in certain embodiments be stained with a DNA stain, such as for example, and without limitation with DAPI. The nuclei contours may in certain embodiments be extracted from the DAPI channel (or other stain or means of visualization of the nucleus) in two steps. Firstly (step **1),** large connected components of relatively high intensity are identified using a thresholding method, such as for instance an adaptive thresholding method as described in Example 2. Optionally secondly (step 2), the connected components that show significant concavity, indicating they likely correspond to aggregated nuclei, may be rejected in certain embodiments.

Step **1** may in certain embodiments consist in a stepwise thresholding of the DAPI images (or other stain or means of visualization of the nucleus) such that nuclei corresponding to sufficiently large connected components of pixels lying above an intensity threshold are selected. To address stochastic variation, in certain embodiments, a hierarchical thresholding strategy may be adopted that progressively refines the segmentation by considering a sequence of increasing thresholds, while exploiting prior knowledge about the size range of human cell nuclei.

In step 2, the rejection of the connected components that likely correspond to multiple nuclei is considered. For this, in certain embodiments, the convexity of each connected component is analyzed. By means of example, and without limitation, in certain embodiments, a number of lines are drawn in parallel to the two principal axis of the connected component. When the connected component is convex, either one or zero segment lies inside the contour, for all parallel lines. In contrast, when the connected component represents aggregated nuclei, it presents a strong concavity and there exist parallel lines that include two or more segments lying inside the contour. If one parallel line supports two sufficiently long inner segments that are separated by a sufficiently long outer segment, the connected component is rejected. In certain embodiments, a threshold, such as a threshold of a particular number of pixels is set based on which is determined the length of inner and outer segments in order to include or exclude particular components. Those lengths in certain embodiments have to be larger than a threshold of for instance at least **1,** 2, 3, 4, or preferably at least 5 pixels to reject the component. The threshold value may in certain embodiments be set empirically to drastically reduce the number of multiple nuclei while keeping most of the single nuclei, compared to a manually generated ground truth.

In certain embodiments, only cells in interphase are analyzed. To identify them, in certain embodiments, the DAPI image (or other stain or means of visualization of the nucleus) and/or the distribution of nucleolus-specific stain (or other means of visualization of the nucleolus) in the segmented component are considered. A cell in certain embodiments is considered to be in interphase if the DAPI (or other stain or means of visualization of the nucleus) is sufficiently dense and spread (if for example at least 50% of

the pixels in the nucleus have a normalized DAPI value larger than 0.3, preferably larger than 0.4, more preferably larger than 0.47, with the DAPI image (or other stain or means of visualization of the nucleus) being normalized by its maximal value), and/or if the nucleolus specific stain (or other means of visualization of the nucleolus) is sufficiently localized (if at least 50% of the pixels in the nucleus have a normalized nucleolus specific stain (or other means of visualization of the nucleolus) value lower than 0.7, preferably lower than 0.6, more preferably lower than 0.53, with the nucleolus specific stain (or other means of visualization of the nucleolus) image being normalized by its maximal value). The thresholds in certain embodiments are set on the basis of a manual ground truth annotation of the images.

Finally, as a post-processing step, in certain embodiments, to ascertain all nucleoli of a cell are contained in each segmented regions, a basic mathematical morphology to close and enlarge each connected regions is applied. In certain embodiments, a dilation by a 13 x 13 structuring element is followed by a 3 x 3 erosion.

In certain embodiments, the morphometric characteristics as defined herein are independent of the actual position of the nucleoli within the nucleus. In certain embodiments, the morphometric characteristics as defined herein while (still) reflect the spatial spreading of the nucleolar masses. In certain embodiments, those characteristics are defined based on a manually-selected set of image samples, depicting typical normal and abnormal nucleolus patterns. The manual extraction of representative samples in certain embodiments allows to derive features that are relevant to make sure that the set of investigated features are able to discriminate among the variety of nucleolus appearances. In certain embodiments, each characteristic is systematically defined as a parametric function (advantageously so that its parameters can be optimized over the entire database to make the feature can differentiate between normal and gene-depleted cells images).

In certain embodiments a set of manually-selected cells nucleoli images that are representative of the appearance diversity in reference control cells is selected (cf. Supplementary Fig. 16). In addition, or in the alternative, in certain embodiments, a 2-D or 3-D graph representing nucleolar intensity may be use to visualize nucleolar appearance diversity.

In certain embodiments, for each nucleus, all morphometric characteristics are defined with respect to the segmentation of the nucleolus masses into a set of disjoint connected components. In certain embodiments, this segmentation is obtained by nucleolar image thresholding, which means that a pixel is considered to be part of the nucleolus if its intensity lies above the threshold (advantageously obtained with nucleolar specific staining or other nucleolar visualization methods or techniques). In certain embodiments, for each characteristic, the segmentation threshold parameter is defined automatically so as to maximize the separation between the distributions of the characteristics for control cells and for test cells, such as cells having disrupted nucleolar morphology or architecture. In certain embodiments, the segmentation threshold is a morphometric characteristic parameter, and might vary from one morphometric characteristic to the other.

In certain embodiments, the area of support of the nucleolus is characterized. In certain embodiments, the size and number of connected components obtained after thresholding with a so-called area segmentation

threshold τ_{α} is considered or determined. In certain embodiments, AAlcc measures the area of the largest connected component in the thresholded image. In certain embodiments, ANcc denotes the number of connected components in the nucleus. In certain embodiments, the sharpness index AS characterizes the sharpness of the intensity gradient along the frontier delimiting the nucleolar masses, and measures the ratio of the nucleus pixels that respectively lie above two thresholds τ_{i1} and ¾, with τ_{i1} > ¾ .

In certain embodiments, to characterize the shape and the texture of the nucleolus, only the largest connected component obtained after segmentation is considered.

In certain embodiments, to quantify the nucleolus shape regularity, we a shape segmentation threshold τ_s is adopted. In certain embodiments, the shape of the largest connected component in the segmented image is characterized. In certain embodiments, each characteristic describes the shape independently of its size. In certain embodiments, the elongation shape factor SElcc, is defined as the square root of the ratio of the two second order moments, λï and λ2, of the connected component c around its principal axes. In certain embodiments, the elliptical regularity factor SRlcc, is defined as the ratio between the area of the connected component, and the area of the smallest ellipse lying outside the connected component, and having the same center, the same principal axes, and the same elongation than the connected component. In certain embodiments, the concavity factor SClcc, is defined as the ratio between the area of the connected component and the area of its convex hull.

In certain embodiments, to characterize the nucleolar texture pattern scalar metrics to reflect the distribution of intensities inside one or more, preferably the largest connected component is considered. In certain embodiments, the distribution of pixel intensities is determined, preferably following cell-wise dynamic range normalization. In certain embodiments, the connected component is segmented based on a texture segmentation threshold τ. In certain embodiments, the texture histogram low tail index THlcc measures the percentage of pixels that lie below some intensity threshold a, preferably while being located inside the erosion of the connected component by a 3 x 3 pixels structuring element. An erosion may in certain embodiments be applied to the shape to get rid of the low intensity pixels lying on the border of the shape. In certain embodiments, the texture uplands index TUlcc is defined to be the number of connected regions lying above a threshold β, while being inside the connected component. In certain embodiments, the texture peaks index TPlcc is defined to be the number of local maxima in the connected component. In certain embodiments, the texture valleys index TVlcc is defined to be the number of local minima in the connected component. In certain embodiments, the texture local minimum TLMlcc is defined as the intensity of the smallest local minimum in the connected component.

In certain embodiments, selection of the morphometric characteristics is based on the Fisher's criterion introduced by the popular Linear Discriminant Analysis (22). In certain embodiments, those parameters or characteristics are selected so as to maximize the separation between the features distributions that we want to discriminate (see also Example 2). The skilled person will understand that alternative selection criteria may be applied, and which equally maximize the separation between the features to be

discriminated. In certain embodiments, a discriminant feature or characteristic is one for which the class-means are well separated, measured relative to the (sum of the) variances of the data assigned to a particular class.

In certain embodiments, considering a feature, parameterized by a vector p lying in a parameter space P, it is assumed that the feature distributions are known as a function of p, for the two classes of observations that we want to best discriminate. The Fishers' optimization criterion indicates that the vector p* that maximizes the separation between the class distributions is defined as:

$$\boldsymbol{p}^* = \underset{\boldsymbol{p} \in P}{\operatorname{argmax}} \frac{[\mu_1(\boldsymbol{p}) - \mu_2(\rho)]^2}{\sigma_1(\boldsymbol{p})^2 + \sigma_2(\boldsymbol{p})^2} \ ,$$

where μï(ρ), μ2(ρ) and σï(ρ), σ2(ρ) respectively denote the means and standard deviations of the distributions of the feature of interest, measured with parameter p for the two classes. This Fisher's criterion is equivalent to the Welch's adaptation of the t-test (23), widely used in image-based morphometry (24).

In certain embodiments, the parameters/characteristics are selected so as to maximize the sum of the separation measured between each pair of distributions extracted from control cells, and from test cells, such as cells having a disrupted nucleolar morphology.

In certain embodiments, a discrepancy vector is defined so as to summarize how the morphometric characteristic distributions associated to a set of nucleoli differ from their corresponding reference distributions. In certain embodiments, for a given image feature, the discrepancy d is defined between a test set and a reference set, preferably to be the ratio of the difference of the mean feature values on each set to the sum of their variance, preferably according to the following formula:

$$\mathbf{d(i)``} = \frac{\mu(i) - \mu_r(i)}{\sqrt{\sigma^2(i) + \sigma_r^2(i)}}$$

with μ(i), μr(i) and σ(i), or(i) denoting the means and standard deviations of the characteristic/feature; with d preferably being a discrepancy vector component.

In certain embodiments, to quantify the nucleolar disruption level, an index of nucleolar disruption, or iNo is defined

In certain embodiments the iNo is defined according to the following formula:

$$iNo = \sum_{i=1}^{N} a(i) * |d_k(i)|$$

wherein N is the number of features; wherein a(i) is a weighing factor for the ïth feature, wherein preferably 0 < a(i) < **1**; and wherein d(i) is the discrepancy vector component for the ïth feature. In certain embodiments, the iNo is the Ll-norm of the discrepancy vector (i.e. all a(i) are equal).

In certain embodiments, a lower dimensional representation of the data is obtained. In certain embodiments, such representation may be obtained by Principal components analysis (PCA). In certain embodiments, the first two principal components are considered. In certain embodiments, the PCA is applied to the discrepancy vectors that capture the average trends associated with nucleolar morphology,

architecture, or integrity. In certain embodiments, the 3, preferably 4, more preferably 5 features that have the largest Fisher's score, i.e. which best discriminates normal and altered nucleoli are considered. In certain embodiments, those features are as listed in the first column of Table 7.

In certain embodiments, the methods as described herein are computer implemented methods, in particular computer implemented methods for characterizing nucleolar morphology according to the invention as described herein. Accordingly, in certain embodiments, the invention relates to a method according to the invention as described herein executed on a computer. In related aspects, the invention provides a computer-readable medium comprising computer-readable instructions which, when loaded on the internal memory of a computer, cause the computer to execute a method according to the invention as described herein. In a further aspect, the invention relates to a data-stream comprising instructions for carrying out a method according to the invention as described herein. In a further aspect, the invention relates to a computer configured for executing a method according to the invention as described herein. In a further aspect, the invention relates to a data-carrier comprising computer-readable instructions for carrying out a method according to the invention as described herein. In a further aspect, the invention relates to a data-carrier comprising computer-readable instructions which, when loaded on the internal memory of a computer, cause the computer to execute a method according to the invention as described herein.

In an aspect, the invention relates to a method for predicting, diagnosing, prognosing, or determining nucleolar disruption or integrity, comprising determining the protein and/or mRNA expression level of one or more (ribosomal) protein. In certain embodiments, said method is an in vitro method, an in vivo, method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In certain embodiments, said (ribosomal) protein is selected as one or more, such as two or more, such as three or more, such as all comprised in of any of lists (a) to (pp):

(a) eL33, eS24, eL8, uS4, eS28, eS27, uS7, uS9, uS19, eL36, uSll, uS8, uL24, eL18, eL32, eS27L, eL29, uL16, uS13, eL24, eL15, uSIO, eSlO, eL42, uLlO, eL13, eS31, uLll, eL40, uL3, eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(b) eL33, eL8, eS27, eL36, uL24, eL18, eL32, eL29, uL16, eL24, eL15, eL42, uLlO, eL13, uLll, eL40, uL3, uL6, eL34, eL39, eL37, uL4, eL14, eL27, uL15, eL20, eL6, eL31, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(c) eL33, eL8, eL13, eS31, eL36, eSlO, eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(d) eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, uL4, eL14, eL27, uL15, eL20, eL6, eL31, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(e) eL13, eS31, eL36, eSlO, eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(f) eL13, eL36, uL6, eL34, eL39, eL37, uL4, eL14, eL27, uL15, eL20, eL6, eL31, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(g) eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14, uLlO, uS13, eS26, uL24, uL18, uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(h) eL29, eL15, eL31, eS30, eL27L, eL28, uLlO, uL24, uL18, eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(i) eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14, uLlO, uS13, eS26, uL24, uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(j) eL29, eL15, eL31, eS30, eL27L, eL28, uLlO, uL24, eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(k) uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(1) eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(m) eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(n) eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(o) eL33, eL8, eL13, eL34, eL39, eL14, uL22, uL13, eL38, eL21, uL2, anduLl;

(p) eL33, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(q) eSlO, eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(r) eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(s) uL6, eL34, eL39, eL37, uL4, eL14, eL27, uL15, eL20, eL6, eL31, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, anduL18;

(t) eSlO, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(u) uL4, eL14, eL27, uL15, eL20, eL6, eL31, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(v) eSlO, eL20, eL6, eL31, uS14, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(w) eL20, eL6, eL31, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(x) uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(y) uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, anduL18;

(z) uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, anduL18;

(aa) uL13, eL38, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(bb) eL38, eL19, eL21, eL43, uL2, uLl, uL5, anduL18;

(cc) eSlO, eL19, eL21, eL43, uL2, uLl, uL5, and uL18;

(dd) eL19, eL21, eL43, uL2, uLl, uL5, anduL18;

(ee) eSlO, eL21, eL43, uL2, uLl, uL5, anduL18;

(ff) eL21, eL43, uL2, uLl, uL5, and uL18;

(gg) eSlO, eL43, uL2, uLl, uL5, anduL18;

(hh) eL43, uL2, uLl, uL5, and uL18;

(ii) uL2, uLl, uL5, and uL18;

(jj) uLl, uL5, and uL18;

(kk) uL5, anduL18;

(11) uL18;

(mm) uL5;

(nn) BXDC1 and RRS1, or any protein involved in ribosomal central protuberance formation;

(oo) any one or more ribosomal protein selected from the above lists, except uL24, uL14, eS7, eL8, eS6, eS19, uSll, eL37, eS26, eS31, eS27L, uS12, anduSlO; or

(pp) any one or more ribosomal protein selected from the above lists, except uL5 and/or uL18.

In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, said use is an in vitro method, an in vivo, method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In a related aspect, the invention provides in the use of means for detecting the mRNA and/or protein expression level of one or more, such as two or more, such as three or more, such as all (ribosomal) protein for predicting, diagnosing, prognosing, or determining nucleolar disruption or integrity. In certain embodiments, such use involves the use of means for detecting the mRNA and/or protein expression level of any one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. Means for detecting mRNA or protein expression are known in the art, and include by means of further guidance, and without limitation primers, probes, and antibodies. In certain

embodiments, said use is an in vitro use, an in vivo use, or an ex vivo use. In certain embodiments, said use is not an in vivo use in a human or animal.

By extension, in an aspect, the invention relates to a method for predicting, diagnosing, prognosing, or determining a disease or disorder characterized by or associated with nucleolar disruption or integrity, comprising determining the protein and/or mRNA expression level of one or more, such as two or more, such as three or more, such as all (ribosomal) protein. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, said method is an in vitro method, an in vivo method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In a related aspect, the invention provides in the use of means for detecting the mRNA and/or protein expression level, such as an inducer or inhibitor of mRNA and/or protein expression, of one or more, such as two or more, such as three or more, such as all (ribosomal) protein for predicting, diagnosing, prognosing, or determining a disease or disorder characterized by or associated with nucleolar disruption or integrity nucleolar disruption or integrity. In certain embodiments, such use involves the use of means for detecting the mRNA and/or protein expression level of any one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, such disease or disorder is cancer/tumor/malignancy/proliferative disease or disorder (or pre-cancerous/pre-neoplastic/pre-malignant disorders or stages, such as hyperplasia, dysplasia, or metaplasia), autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems). In certain embodiments, said use is an in vitro use, an in vivo use, or an ex vivo use. In certain embodiments, said use is not an in vivo use in a human or animal.

In an aspect, the invention relates to a method for treating, alleviating or improving a disease or disorder characterized by or associated with nucleolar disruption or integrity, comprising (specifically) decreasing or increasing the mRNA or protein level or activity of one or more, such as two or more, such as three or more, such as all (ribosomal) protein, such as with an inducer or inhibitor of mRNA and/or protein expression. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, such disease or disorder is cancer/tumor/malignancy/proliferative disease or disorder (or pre-cancerous/pre-neoplastic/pre-malignant disorders or stages, such as hyperplasia, dysplasia, or metaplasia), autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems). In certain embodiments, said use is an in vitro method, an in vivo method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal. In certain embodiments, such method involves administering a therapeutically effective amount of said means for increasing or decreasing expression levels or activity.

In an aspect, the invention relates to means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for use in treating, alleviating, or improving a disease or disorder characterized by or associated with nucleolar disruption or integrity. In certain embodiments, said (ribosomal) protein is one

or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, such disease or disorder is cancer/tumor/malignancy/proliferative disease or disorder (or pre-cancerous/pre-neoplastic/pre-malignant disorders or stages, such as hyperplasia, dysplasia, or metaplasia), autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems).

In an aspect, the invention relates to the use of means for (specifically) increasing or decreasing the mRNA and/or protein level or activity of a (ribosomal) protein for the manufacture of a medicament for treating, alleviating, or improving a disease or disorder characterized by or associated with nucleolar disruption or integrity. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, such disease or disorder is cancer/tumor/malignancy/proliferative disease or disorder (or pre-cancerous/pre-neoplastic/pre-malignant disorders or stages, such as hyperplasia, dysplasia, or metaplasia), autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems).

In an aspect, the invention relates to the use of means for (specifically) increasing or decreasing the mRNA and/or protein level or activity of a (ribosomal) protein for treating, alleviating, or improving a disease or disorder characterized by or associated with nucleolar disruption or integrity. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, such disease or disorder is cancer/tumor/malignancy/proliferative disease or disorder (or pre-cancerous/pre-neoplastic/pre-malignant disorders or stages, such as hyperplasia, dysplasia, or metaplasia), autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems). In certain embodiments, said use is an in vitro use, an in vivo use, or an ex vivo use. In certain embodiments, said use is not an in vivo use in a human or animal.

In an aspect, the invention relates to a method for modifying or altering nucleolar morphology, architecture, or integrity, comprising (specifically) decreasing or increasing the mRNA or protein level or activity, such as with and inducer or inhibitor of mRNA and/or protein expression, of one or more, such as two or more, such as three or more, such as all (ribosomal) protein. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, said method is an in vitro method, an in vivo, method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal. In certain embodiments, said use is an in vitro method, an in vivo, method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In an aspect, the invention relates to means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for use in modifying or altering nucleolar morphology, architecture, or integrity, comprising (specifically) decreasing or increasing the mRNA or protein level or activity of one or more, such as two or more, such as three or more, such as all (ribosomal) protein. In certain embodiments, said

(ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above.

In an aspect, the invention relates to the use of means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for the manufacture of a medicament for modifying or altering nucleolar morphology, architecture, or integrity. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above.

In an aspect, the invention relates to the use of means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for use in modifying or altering nucleolar morphology, architecture, or integrity, comprising (specifically) decreasing or increasing the mRNA or protein level or activity of one or more, such as two or more, such as three or more, such as all (ribosomal) protein. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of any of (a) to (pp) above. In certain embodiments, said use is an in vitro method, an in vivo, method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In an aspect, the invention relates to a method for increasing or decreasing the protein level or activity of p53 in a cell, comprising respectively decreasing or increasing the mRNA or protein level or activity, such as with inducer or inhibitor of mRNA and/or protein expression, of one or more, such as two or more, such as three or more, such as all ribosomal proteins selected as one or more, such as two or more, such as three or more, such as all comprised in of any of lists (1) to (9):

1. eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14, uLlO, uS13, eS26, uL24, uL18, uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;
2. eL29, eL15, eL31, eS30, eL27L, eL28, uLlO, uL24, uL18, eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;
3. eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14, uLlO, uS13, eS26, uL24, uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;
4. eL29, eL15, eL31, eS30, eL27L, eL28, uLlO, uL24, eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

5. uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;
6. eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;
7. eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;
8. eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, anduLl;
9. eL33, eL8, eL13, eL34, eL39, eL14, uL22, uL13, eL38, eL21, uL2, anduLl;

In certain embodiments, said method is an in vitro method, an in vivo, method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In an aspect, the invention relates to means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for use in decreasing or increasing respectively the protein level or activity of p53. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of (1) to (9) above.

In an aspect, the invention relates to the use of means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for the manufacture of a medicament for decreasing or increasing respectively the p53 protein level or activity. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of (1) to (9) above.

In an aspect, the invention relates to the use of means for (specifically) increasing or decreasing the mRNA and/or protein level or activity, such as an inducer or inhibitor of mRNA and/or protein expression, of a (ribosomal) protein for decreasing or increasing respectively the p53 protein level or activity. In certain embodiments, said (ribosomal) protein is one or more, such as two or more, such as three or more, such as all (ribosomal) protein of (1) to (9) above. In certain embodiments, said use is an in vitro use, an in vivo use, or an ex vivo use. In certain embodiments, said use is not an in vivo use in a human or animal.

In certain embodiments, in the treatment methods and uses according to the invention as described herein (including manipulation of nucleolar morphology and/or p53 protein level), if a decreased mRNA and/or protein expression level of and one or more, such as two or more, such as three or more, such as all of the (ribosomal) proteins of (pp) above is detected or determined, said treatment method or use may involve decreasing or increasing the mRNA and/or protein level or activity of uL5 and/or uL18.

In an aspect, the invention relates to the use of any of the methods according to the invention as described herein for identifying compounds affecting or altering nucleolar morphology, architecture, integrity, or disruption or for identifying compounds affecting or altering (increasing or decreasing) p53 mRNA or

protein expression level, activity, stability and/or steady state levels. In certain embodiments, said use is an in vitro use, an in vivo use, or an ex vivo use. In certain embodiments, said use is not an in vivo use in a human or animal.

In an aspect, the invention relates to a screening method for identifying compounds affecting or altering nucleolar morphology, architecture, integrity, or disruption or for identifying compounds affecting or altering (increasing or decreasing) p53 mRNA or protein expression level, activity, stability and/or steady state levels, with any of the methods according to the invention as described herein. In certain embodiments, said method is an in vitro method, an in vivo method, or an ex vivo method. In certain embodiments, said method is not an in vivo method in a human or animal.

In any of these methods, the detection or induction or decrease of the recited (ribosomal) proteins may be done on (isolated) cells, in vitro, ex vivo, or in vivo in cells, tissues or organisms.

Methods for increasing or decreasing gene or protein expression or affecting protein stability are known in the art. Such methods include but are not limited to for instance (conditional or inducible) knock-out, (conditional or inducible) knock-down (e.g. siRNA), use of neutralizing compounds (e.g. small molecules, antibodies, etc.), (conditional or inducible) overexpression, transcriptional or translational activation, etc. Standard reference works setting forth the general principles of recombinant DNA technology include Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001; Current Protocols in Molecular Biology, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York, 1996 (with periodic updates) ("Ausubel et al. 1996"); Innis et al., PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990. General principles of microbiology are set forth, for example, in Davis, B. D. et al., Microbiology, 3rd edition, Harper & Row, publishers, Philadelphia, Pa. (1980).

Except when explicitly noted, "subject" or "patient" are used interchangeably and refer to animals or humans. Animals are preferably vertebrates, more preferably mammals such as veterinary animals, horse, rabbit, mouse, rat, pig, sheep, cow or dog, etc.. The term subject specifically includes human patients.

The terms "diagnosing" or "diagnosis" generally refer to the process or act of recognising, deciding on or concluding on a disease or condition in a subject on the basis of symptoms and signs and/or from results of various diagnostic procedures (such as, for example, from knowing the presence, absence and/or quantity of one or more biomarkers characteristic of the diagnosed disease or condition).

The terms "prognosticating" or "prognosis" generally refer to an anticipation on the progression of a disease or condition and the prospect (e.g., the probability, duration, and/or extent) of recovery.

A good prognosis of cancer may generally encompass anticipation of a satisfactory partial or complete recovery from cancer, preferably within an acceptable time period. A good prognosis of cancer may more commonly encompass anticipation of not further worsening or aggravating of the heart failure condition, preferably within a given time period.

A poor prognosis of cancer may generally encompass anticipation of a substandard recovery and/or unsatisfactorily slow recovery, or to substantially no recovery or even further worsening of cancer.

The terms "sample" or "biological sample" as used herein include any biological specimen obtained from a subject. Samples may include, without limitation, whole blood, plasma, serum, red blood cells, white blood cells (e.g., peripheral blood mononuclear cells), saliva, urine, stool (i.e., faeces), tears, sweat, sebum, nipple aspirate, ductal lavage, tumour exudates, synovial fluid, cerebrospinal fluid, lymph, fine needle aspirate, amniotic fluid, any other bodily fluid, cell lysates, cellular secretion products, inflammation fluid, semen and vaginal secretions. Samples may also include tissue samples and biopsies, tissue homogenates and the like.

The one or more nucleolar markers in the form of proteins, polypeptides, peptides, nucleic acids, or mRNA molecules are "measured" in a sample when the presence or absence and/or quantity of said marker is detected or determined in the sample, preferably substantially to the exclusion of other molecules and analytes.

The terms "quantity", "amount" and "level" are synonymous and generally well-understood in the art. The terms as used herein may particularly refer to an absolute quantification of a marker in a sample, or to a relative quantification of a marker in a sample, i.e., relative to another value such as relative to a reference value as taught herein, or to a range of values indicating a base-line expression of the marker. These values or ranges can be obtained from a single patient or from a group of patients.

An absolute quantity of a marker in a sample may be advantageously expressed as weight or as molar amount, or more commonly as a concentration, e.g., weight per volume or mol per volume.

A relative quantity of a marker in a sample may be advantageously expressed as an increase or decrease or as a fold-increase or fold-decrease relative to said another value, such as relative to a reference value from a certain type of tumour or a healthy sample.

The term "treating" as used herein includes treating any one or more of the diseases or disorders as described herein. Treatment of for instance cancer means administration of a medicament in the form of a compound or pharmaceutical composition of the invention with the result that cancer is stabilized, reduced or the patient is cured.

As used herein, the terms "treat" or "treatment" refer to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) an undesired physiological change or disorder, such as the development or spread of disease, e.g., cancer. Beneficial or desired clinical results include, but are not limited to, alleviation of symptoms, diminishment of extent of disease, stabilised (i.e., not worsening) state of disease, delay or slowing of disease progression, amelioration or palliation of the disease state, and remission (whether partial or total), whether detectable or undetectable. "Treatment" can also mean prolonging survival as compared to expected survival if not receiving treatment.

As used herein, a phrase such as "a subject in need of treatment" includes subjects, such as mammalian subjects, that would benefit from treatment of a given disease or disorder, such as, e.g., cancer.

Such subjects will typically include, without limitation, those that have been diagnosed with the condition, preferably a disease or disorder as defined herein elsewhere, e.g., cancer, those prone to have or develop the said condition and/or those in whom the condition is to be prevented.

The term "therapeutically effective amount" refers to an amount of a compound or pharmaceutical composition of the invention effective to treat a disease or disorder in a subject, i.e., to obtain a desired local or systemic effect and performance. By means of example and not limitation, in the case of proliferative disease, e.g., cancer, therapeutically effective amount of a drug may reduce the number of cancer cells; reduce the tumour size; inhibit (i.e., slow to some extent and preferably stop) cancer cell infiltration into peripheral organs; inhibit (i.e., slow to some extent and preferably stop) tumour metastasis; inhibit, to some extent, tumour growth; enhance efficacy of another cancer therapy; and/or relieve to some extent one or more of the symptoms associated with the cancer. To the extent the drug may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic. For cancer therapy, efficacy can, for example, be measured by assessing the time to disease progression (TTP) and/or determining the response rate (RR). The term thus refers to the quantity of compound or pharmaceutical composition that elicits the biological or medicinal response in a tissue, system, animal, or human that is being sought by a researcher, veterinarian, medical doctor or other clinician, which includes alleviation of the symptoms of the cancer being treated. In particular, these terms refer to the quantity of compound or pharmaceutical composition according to the invention which is necessary to prevent, cure, ameliorate, or at least minimize the clinical impairment, symptoms, or complications associated with cancer in either a single or multiple doses.

The treatment methods and uses according to the invention as described herein may in certain embodiments involve the administration of a composition comprising the means for inducing or reducing protein or mRNA expression level or activity. In certain embodiments, such composition is a pharmaceutical composition. In certain embodiments, such composition further comprises one or more pharmaceutically acceptable excipients.

The term "pharmaceutically acceptable" as used herein is consistent with the art and means compatible with the other ingredients of a pharmaceutical composition and not deleterious to the recipient thereof.

The term "pharmaceutically acceptable salts" as used herein means an inorganic acid addition salt such as hydrochloride, sulfate, and phosphate, or an organic acid addition salt such as acetate, maleate, fumarate, tartrate, and citrate. Examples of pharmaceutically acceptable metal salts are alkali metal salts such as sodium salt and potassium salt, alkaline earth metal salts such as magnesium salt and calcium salt, aluminum salt, and zinc salt. Examples of pharmaceutically acceptable ammonium salts are ammonium salt and tetramethylammonium salt. Examples of pharmaceutically acceptable organic amine addition salts are salts with morpholine and piperidine. Examples of pharmaceutically acceptable amino acid addition salts are salts with lysine, glycine, and phenylalanine.

Examples

Example 1: Materials and methods

Nucleolar screens

The nucleolar screens were performed on an automated high throughput platform. For each r-protein, three different siRNAs were used, and for each siRNA, 2,000 cells imaged. For consistency, the entire screen was duplicated. The efficiency of siRNA-mediated depletion was assessed in a random shotgun RTqPCR assay. A calibration set consisting of four control proteins whose depletion, we established, affects strongly nucleolar structure was used (Supplementary Fig. 1).

Cell lines

The cell lines used in this study are listed in Table 1. All cell lines were cultured at 37°C under 5% C02. Culture media were supplemented with 10% fetal bovine serum (FBS; Sigma) and 1% penicillin-streptomycin (Pen-Strep, Gibco). For consistency, the experiments were performed on cells grown for 10 to 15 passages. The nucleolar screens were conducted in cervical cancer (HeLa) cells stably expressing FBL in fusion with GFP (FIB364). All cell lines were purchased from the ATCC repository and regularly tested for contamination with the LookOut mycoplasma PCR detection kit (Sigma-Aldrich, MP0035).

Table 1

Name	Origin	Growth medium
HeLa-GFP-FBL	Epithelial (cervix)	DMEM / FBS / Pen-Strep
HeLa	Epithelial (cervix)	DMEM / FBS / Pen-Strep
HCT116 *p53+/+*	Epithelial (colon)	McCoy / FBS / Pen-Strep
HCT116 *p53-/-*	Epithelial (colon)	McCoy / FBS / Pen-Strep
A549	Epithelial (lung)	F12 / FBS / Pen-Strep
H1944	Epithelial (lung)	RPMI-1640 / FBS / Pen-Strep

siRNA depletion

The FIB364 cell line was transfected with either of 3 distinct siRNAs targeting each r-protein according to the protocol described below (Table 2 for siRNA sequences). The entire screening procedure was duplicated. Depletions were performed in 96-well plates (Porvair Sciences). A transfection reagent mix (0.125 µï of Interferin and 20 µï of Optimem) was added to each plate well and left to set for 10 min at RT. siRNA (10 µï of 100 nM stock) were added to this mix and left to set for another 30 min at RT. Cells (70 µï of 100,000 cells/ml) were added to each well and the plates were incubated for 3 days. For each individual plate, a set of 7 wells was used for negative and positive controls. Our calibration set consists of mock-treated cell (cells with the transfection reagent mix only) and cells treated with a non-targeting siRNA (Scramble, SCR), or with siRNA specific to GFP, fibrillarin (FBL), nucleophosmin (NPM), nucleolin (NCL), or TIFIA. Cells were fixed in 2% formaldehyde, washed in PBS, incubated 10 min in the presence of DAPI (1:20,000 of 5 mg/ml in PBS, Sigma), washed again and stored in PBS before imaging. The depletion of the central protuberance assembly factors RRS1 and BXDC1 were performed according to the same protocol.

Table 2

Ribosomal proteins (RPs)						
Gene name	**New nomenclature**	**GenBank accession number**	**siRNA number**	**siRNA ID**	**siRNA sequence (5' to 3')**	**Source**
RpSA	*uS2*	XM_939738	#1	s60434	CUCCUGGAACC UUCACUAAtt	Ambion, Life Technologies
			#2	s51977	GCACCAAUCU UGACUUCCAtt	Ambion, Life Technologies
			#3	s51978	GCAUCUAUAU CAUAAAUCUtt	Ambion, Life Technologies
RpS2	*uS5*	NM_002952	#1	s12252	CCAAGUCUCCC UAUCAGGAtt	Ambion, Life Technologies
			#2	s12253	UCUCCGCACCU GUGCCUAAtt	Ambion, Life Technologies
			#3	s12254	CCUAAGAAGC UGCUCAUGAtt	Ambion, Life Technologies
RpS3	*uS3*	NM_001005	#1	s12255	UGACUGCUGU AGUUCAGAAtt	Ambion, Life Technologies
			#2	s12256	GCGGAGACCC UGUUAACUAtt	Ambion, Life Technologies
			#3	s12257	AAGCUGAACU GAAUGAGUUtt	Ambion, Life Technologies
RpS3A	*eSl*	NM_001006	#1	sl2258	GCAACAAUCA GAUACGGAAtt	Ambion, Life Technologies
			#2	s12259	AAUUCAAGCU GAUUACUGAtt	Ambion, Life Technologies
			#3	s12260	AGAUUGGUAU GAUGUGAAAtt	Ambion, Life Technologies
RpS4X	*eS4*	NM_001007	#1	s12261	GCGGUUCAUU AAAAUCGAUtt	Ambion, Life Technologies
			#2	s12262	GCAUGCAGCG GUUCAUUAAtt	Ambion, Life Technologies
			#3	s12263	AGACUUAAGU AUGCCCUGAtt	Ambion, Life Technologies
RpS5	*uS7*	NM_001009	#1	s12267	CCGGAACAUU AAGACCAUUtt	Ambion, Life Technologies
			#2	sl2268	ACAUUGCAGU GAAGGAGAAtt	Ambion, Life Technologies
			#3	s12269	GCAUGCCUUC GAGAUCAUAtt	Ambion, Life Technologies
RpS6	*eS6*	NM_001010	#1	s12270	CCUUAAAUAA AGAAGGUAAtt	Ambion, Life Technologies
			#2	s12271	GGAACAAAUU GCGAAGAGAtt	Ambion, Life Technologies
			#3	s12272	CAGCGUACCA AGAAAAAUAtt	Ambion, Life Technologies
RpS7	*eS7*	NM_001011	#1	sl2288	CAUAAUCUUU GUUCCCGUUtt	Ambion, Life Technologies
			#2	sl2289	GAGUUUCAAU UGUAAACAAtt	Ambion, Life Technologies
			#3	s12290	GUACGCGAAU UGGAGAAAAtt	Ambion, Life Technologies

RpS8	*eS8*	NM__001012	#1	s12291	GCUAUGUGCU AGAGGGCAAtt	Ambion, Life Technologies
			#2	s12292	GAAAUAUGAU GAAAGGAAAtt	Ambion, Life Technologies
			#3	s12293	AGAGUGUUGU ACUCGUAAAtt	Ambion, Life Technologies
RpS9	*uS4*	NM__001013	#1	s12294	AGAUGAAGCU GGAUUACAUtt	Ambion, Life Technologies
			#2	s12295	AGAUAGAGGA UUUCUUAGAtt	Ambion, Life Technologies
			#3	s12296	AGCUGAAGCU GAUCGGCGAtt	Ambion, Life Technologies
RpSlO	*eSlO*	NM__001014	#1	s12297	GCAACCGAAU UCCAGUUUAtt	Ambion, Life Technologies
			#2	s12298	CGACCUGCGA GACUCACAAtt	Ambion, Life Technologies
			#3	s12299	CAUUUCUACU GGUACCUUAtt	Ambion, Life Technologies
RpSll	*uS17*	NM__001015	#1	s12300	AGAACAUGUC UGUACACCUtt	Ambion, Life Technologies
			#2	S200170	CUACAUCCGCA AGUACAACtt	Ambion, Life Technologies
			#3	S200171	CCAGAUCGGU GACAUCGUCtt	Ambion, Life Technologies
RpS12	*eS12*	NM__001016	#1	s52989	CCUUUGUGCU GAACACCAAtt	Ambion, Life Technologies
			#2	sl94763	GGCCUUUGUA AAAUUGACAtt	Ambion, Life Technologies
			#3	si94764	ACUGUGAUGA GCCUAUGUAtt	Ambion, Life Technologies
RpS13	*uS15*	NM__001017	#1	s12304	GAAAGGAUAA GGAUGCUAAtt	Ambion, Life Technologies
			#2	s12305	GGCCUUACUCC UUCACAGAtt	Ambion, Life Technologies
			#3	sl2306	GCUCGAUAUU AUAAGACCAtt	Ambion, Life Technologies
RpS14	*uSll*	NM__001025 070	#1	s12307	AGACCGAGAU GAAUCCUCAtt	Ambion, Life Technologies
			#2	sl2308	AGGUAAAGGC AGACCGAGAtt	Ambion, Life Technologies
			#3	s12309	CAUCCUUCAA UGACACUUUtt	Ambion, Life Technologies
RpS15	*uS19*	NM__001018	#1	sl2310	CCUACAAGCCC GUAAAGCAtt	Ambion, Life Technologies
			#2	si94765	CUUCAUCCCUC UCAAGUAAtt	Ambion, Life Technologies
			#3	sl94766	GCAUGGUGGG CGUCUACAAtt	Ambion, Life Technologies
RpS15 A	*uS8*	NM__001019	#1	sl2311	CCAAAGUCAU CGUCCGGUUtt	Ambion, Life Technologies
			#2	sl2312	AGAUUUGACG UGCAACUCAtt	Ambion, Life Technologies
			#3	sl2313	GAGUAUCAAC AAUGCCGAAtt	Ambion, Life Technologies
RpS16	*uS9*	NM__001020	#1	sl2314	AGAUUUAUGC	Ambion, Life

					UAUCCGUCAtt	Technologies
			#2	sl2315	GCAAUGGUCU CAUCAAGGUtt	Ambion, Life Technologies
			#3	sl2316	GGAGCGAUUU GCUGGUGUAtt	Ambion, Life Technologies
RpS17	*eS17*	NM__001021	#1	sl2317	GCCCGGGUCA UCAUAGAAAtt	Ambion, Life Technologies
			#2	sl2318	GGAGAUUAUU GAAGUAGAUtt	Ambion, Life Technologies
			#3	sl2319	GGAUCAGGAG AUUAUUGAAtt	Ambion, Life Technologies
RpS18	*uS13*	NM_ 022551	#1	s12320	UGCGAGUACU CAACACCAAtt	Ambion, Life Technologies
			#2	sl2321	UGAUCACCAU UAUGCAGAAtt	Ambion, Life Technologies
			#3	s12322	CGAUGGGCGG CGGAAAAUAtt	Ambion, Life Technologies
RpS19	*eS19*	NM__001022	#1	sl2323	GCCGCAAACU GACACCUCAtt	Ambion, Life Technologies
			#2	si94767	AAACCAUGCU GGGUUAAUAtt	Ambion, Life Technologies
			#3	sl94768	AGCUGGCCAA GCACAAAGAtt	Ambion, Life Technologies
RpS20	*uSIO*	NM_ 001023	#1	s12324	GGACCAGUUC GAAUGCCUAtt	Ambion, Life Technologies
			#2	s12325	GCCGCAACGU AAAAUCCUUtt	Ambion, Life Technologies
			#3	sl2326	CUAACAAGCC GCAACGUAAtt	Ambion, Life Technologies
RpS21	*eS21*	NM_001024	#1	s12327	AGUUUAAAAC UUAUGCUAUtt	Ambion, Life Technologies
			#2	S226974	GUAGGAUGGG UGAGUCAGAtt	Ambion, Life Technologies
			#3	si94769	GCAUCAUCGG UGCCAAGGAtt	Ambion, Life Technologies
RpS23	*uS12*	NM_ 001025	#1	sl2328	CCAAUGACGG UUGCUUGAAtt	Ambion, Life Technologies
			#2	s12329	GAAGCUCCGU AGUCACCGAtt	Ambion, Life Technologies
			#3	sl2330	UGACGGUUGC UUGAACUUUtt	Ambion, Life Technologies
RpS24	*eS24*	NM_ 001026	#1	sl2331	CCGACUACUUC AGAGGAAAtt	Ambion, Life Technologies
			#2	sl2332	AGACAGAAAU UCGGGAAAAtt	Ambion, Life Technologies
			#3	sl2333	CCUGGAUUAU GCAAAGAAAtt	Ambion, Life Technologies
RpS25	*eS25*	NM_ 001028	#1	sl2334	GCACAGAGCU CAAGUAAUUtt	Ambion, Life Technologies
			#2	sl2335	GGAGCUCCUU AGUAAAGGAtt	Ambion, Life Technologies
			#3	sl2336	GGAAGUUCCC AACUAUAAAtt	Ambion, Life Technologies
RpS26	*eS26*	NM_ 001093 731	#1	s229467	ACAGCAAAGU AGUCAGGAAtt	Ambion, Life Technologies

			#2	s53606	CAAGCUGUAU GUGAAGCUAtt	Ambion, Life Technologies
			#3	s229468	AUUCGUCAUU CGAAACAUAtt	Ambion, Life Technologies
RpS27	*eS27*	NM_ 001030	#1	sl2337	GGAGGAAAAG CAAGGCUUAtt	Ambion, Life Technologies
			#2	sl2338	AUGCACAAAC GGUAGUUUUtt	Ambion, Life Technologies
			#3	si94771	GGAGGAAGCA GCACUAAAAtt	Ambion, Life Technologies
RpS27 L	*eS27L*	NM_ 015920	#1	s27333	AAGUCCAAAU UCUUACUUUtt	Ambion, Life Technologies
			#2	s27334	AAAUGUCCAG GUUGCUACAtt	Ambion, Life Technologies
			#3	s27335	UGUCCAGGUU GCUACAAGAtt	Ambion, Life Technologies
RpS27 A	*eS31*	NM_ 002954	#1	s12340	GUACUUUGUC UGACUACAAtt	Ambion, Life Technologies
			#2	sl2341	GGACGUACUU UGUCUGACUtt	Ambion, Life Technologies
			#3	s226976	GAAGAAGUCU UACACCACUtt	Ambion, Life Technologies
RpS28	*eS28*	NM_ 001031	#1	s12342	CCAUCAUCCGC AAUGUAAAtt	Ambion, Life Technologies
			#2	sl2343	GCCGAUCCAUC AUCCGCAAtt	Ambion, Life Technologies
			#3	S226977	GCGUGGAAUU CAUGGACGAtt	Ambion, Life Technologies
RpS29	*uS14*	NM_ 001030 001	#1	S227039	GCACGGUCUG AUCCGGAAAtt	Ambion, Life Technologies
			#2	si94772	CGUCAGUACG CGAAGGAUAtt	Ambion, Life Technologies
			#3	sl96844	GGAAAUAUGG CCUCAAUAUtt	Ambion, Life Technologies
RpS30	*eS30*	NM_ 001997	#1	s5039	UGAGAGGUCA GACUCCUAAtt	Ambion, Life Technologies
			#2	s5040	CCAGAUCAAG GCUCAUGUAtt	Ambion, Life Technologies
			#3	si94404	UUCUGGCUUU CUCUAAUAAtt	Ambion, Life Technologies
Rackl	*Rackl*	NM_ 006098	#1	s20340	CAGGCUAUCU GAACACGGUtt	Ambion, Life Technologies
			#2	s20341	CAAACACCUU UACACGCUAtt	Ambion, Life Technologies
			#3	s20342	GGAUGAGACC AACUAUGGAtt	Ambion, Life Technologies
RpL3	*uL3*	NM_ 000967	#1	sl2142	GGCAUAAAUC UAAGAAGAAtt	Ambion, Life Technologies
			#2	sl2143	ACGGCAAGCU GAUCAAGAAtt	Ambion, Life Technologies
			#3	sl2144	GACUAUGACC UAUCUGACAtt	Ambion, Life Technologies
RpL4	*uL4*	NM_ 000968	#1	sl2148	GACCAGAUAU UGUGAACUUtt	Ambion, Life Technologies
			#2	sl2149	GGCCGAAUGU	Ambion, Life

					UUGCACCAAtt	Technologies
			#3	sl2150	CCGCUUCCCUC AAGAGUAAtt	Ambion, Life Technologies
RpL5	*uL18*	NM_ 000969	#1	sl2151	GACGAGAGGG UAAAACUGAtt	Ambion, Life Technologies
			#2	sl2152	GGAGGAGAUG UAUAAGAAAtt	Ambion, Life Technologies
			#3	sl2153	GAAGUACAUC GGAAGCACAtt	Ambion, Life Technologies
RpL6	*eL6*	NM_ 000970	#1	sl2154	AGCGCAAGAU UGAUCAGAAtt	Ambion, Life Technologies
			#2	sl2155	CCAAAAUCGA UAUCAGCAAtt	Ambion, Life Technologies
			#3	sl2156	CCCAAAACAUC UUACUGAUtt	Ambion, Life Technologies
RpL7	*uL30*	NM_ 000971	#1	s352	CUCGAUCUCU UGGUAAAUAtt	Ambion, Life Technologies
			#2	s353	CACUAUCACA AGGAAUAUAtt	Ambion, Life Technologies
			#3	s354	CGUCAAAUCU UCAAUGGAAtt	Ambion, Life Technologies
RpL7A	*eL8*	NM_000972	#1	sl2157	GGUGAACUCG GAAGACAAAtt	Ambion, Life Technologies
			#2	sl2158	CAGCUGUCGU GAAGAAGCAtt	Ambion, Life Technologies
			#3	sl2159	AGUACAGACC AGAGACAAAtt	Ambion, Life Technologies
RpL8	*uL2*	NM_ 000973	#1	sl2160	GCCGAAUUGA CAAACCCAUtt	Ambion, Life Technologies
			#2	sl2161	CCAUGCCUGA GGGUACAAUtt	Ambion, Life Technologies
			#3	sl2162	CGGGAUCCGU AUCGGUUUAtt	Ambion, Life Technologies
RpL9	*uL6*	NM_ 000661	#1	sl2163	AUGUCGACAU UACUCUGAAtt	Ambion, Life Technologies
			#2	sl2164	GAAAGAUGAA UUAAUCCUUtt	Ambion, Life Technologies
			#3	sl2165	CCAGAAAAUG UCGACAUUAtt	Ambion, Life Technologies
RpLlO	*uL16*	NM_ 006013	#1	sl2166	GAAACAGGUU GACAACUCAtt	Ambion, Life Technologies
			#2	sl2167	CUAUGUCUUU GUAUCUACAtt	Ambion, Life Technologies
			#3	si96275	CCCGAAUUUG UGCCAAUAAtt	Ambion, Life Technologies
RpLlO A	*uLl*	NM_ 007104	#1	s9421	GUCCGGGCCU UAUAUAUCAtt	Ambion, Life Technologies
			#2	s9422	CAGAGUCUCU GAUCAAGCAtt	Ambion, Life Technologies
			#3	s9423	UGAAGAAGGU GUUAUGUCUtt	Ambion, Life Technologies
RpLll	*uL5*	NM_ 000975	#1	sl2168	GGUGCGGGAG UAUGAGUUAtt	Ambion, Life Technologies
			#2	sl2169	CAACUUCUCA GAUACUGGAtt	Ambion, Life Technologies

			#3	sl2170	GGAACUUCGC AUCCGCAAAtt	Ambion, Life Technologies
RpL12	*uLll*	NM_ 000976	#1	si94741	CCCAGUCAGU GGGCUGUAAtt	Ambion, Life Technologies
			#2	si94742	CCAGCCAGUU AAGCACAAAtt	Ambion, Life Technologies
			#3	si94743	AGCCCUCAAG GAACCACCAtt	Ambion, Life Technologies
RpL13	*eL13*	NM_ 000977	#1	sl2171	GGAAGAGAAG AAUUUCAAAtt	Ambion, Life Technologies
			#2	sl2172	AGAAGAAUUU CAAAGCCUUtt	Ambion, Life Technologies
			#3	sl2173	CGGCAUACGG GCAAAAAGAtt	Ambion, Life Technologies
RpL13 A	*uL13*	NM_ 012423	#1	s58865	AGCUCAUGAG GCUACGGAAtt	Ambion, Life Technologies
			#2	s23995	GGUGUUUGAC GGCAUCCCAtt	Ambion, Life Technologies
			#3	s23996	GGAAACAGGC CGAGAAGAAtt	Ambion, Life Technologies
RpL14	*eL14*	NM_ 001034 996	#1	sl7238	GACAGAUUUU GAUCGUUUUtt	Ambion, Life Technologies
			#2	sl7239	CCAGAAGUAU GUCCGACAAtt	Ambion, Life Technologies
			#3	s17240	AGAUCACCGCC GCGAGUAAtt	Ambion, Life Technologies
RpL15	*eL15*	NM_ 002948	#1	sl2174	CCUUUCAAGU GUGAGCUUAtt	Ambion, Life Technologies
			#2	sl2175	CCUUAAGAUU GGUAAGCUAtt	Ambion, Life Technologies
			#3	sl2176	GGUGUAGACU UUUUAAGUUtt	Ambion, Life Technologies
RpL17	*uL22*	NM_ 000985	#1	sl2177	CACGAAAUCA UGCAAAUCAtt	Ambion, Life Technologies
			#2	sl2178	CAAUCUUCGU GUUCACUUUtt	Ambion, Life Technologies
			#3	sl2179	GCACAUGCUU AAAAACGCAtt	Ambion, Life Technologies
RpL18	*eL18*	NM_ 000979	#1	si94744	GGCUGUUGGU CAAGUUAUAtt	Ambion, Life Technologies
			#2	si94745	AAACUAACCC UGGAUCCUAtt	Ambion, Life Technologies
			#3	si94746	CGGAUGAUCC GGAAGAUGAtt	Ambion, Life Technologies
RpL18 A	*eL20*	NM_ 000980	#1	sl2180	GUACUUUGUA UCUCAGUUAtt	Ambion, Life Technologies
			#2	sl2181	CCCACAACAUG UACCGGGAtt	Ambion, Life Technologies
			#3	sl2182	ACUCCAUUCA GAUCAUGAAtt	Ambion, Life Technologies
RpL19	*eL19*	NM_ 000981	#1	sl2183	AGACCAAGGA AGCACGCAAtt	Ambion, Life Technologies
			#2	sl2184	AGAAGAUACC GUGAAUCUAtt	Ambion, Life Technologies
			#3	sl2185	GCUCAGAAGA	Ambion, Life

					UACCGUGAAtt	Technologies
RpL21	*eL21*	NM_ 000982	#1	sl2186	CACUCUAAGA GCCGAGAUAtt	Ambion, Life Technologies
			#2	sl2187	GGAAGAGUCU ACAAUGUUAtt	Ambion, Life Technologies
			#3	sl2188	GGUGAUAUUG UAGACAUCAtt	Ambion, Life Technologies
RpL22	*eL22*	NM_ 000983	#1	sl2189	GAGAGUUACG AAUUACGUUtt	Ambion, Life Technologies
			#2	sl2190	CAAAGAGAGU UACGAAUUAtt	Ambion, Life Technologies
			#3	sl2191	GGUUGCGCGU AGUUGCUAAtt	Ambion, Life Technologies
RpL23	*uL14*	NM_ 000978	#1	sl7871	GCAGGAGUCA UAGUGAACAtt	Ambion, Life Technologies
			#2	sl7872	CAAUAAAGGC GAGAUGAAAtt	Ambion, Life Technologies
			#3	sl7873	CAACGAAAGU CAUACCGUAtt	Ambion, Life Technologies
RpL23 A	*uL23*	NM_000984	#1	sl2192	GAGUUAGUGU CCUAGGAAAtt	Ambion, Life Technologies
			#2	sl2193	AGAGAUCUUU GUGACUAGAtt	Ambion, Life Technologies
			#3	sl2194	CUGUUCUACU UAUCCUUUUtt	Ambion, Life Technologies
RpL24	*eL24*	NM_ 000986	#1	sl2198	CGAGCAGUCA AAUUCCAGAtt	Ambion, Life Technologies
			#2	sl2199	AGAAACCUGA AGUUAGAAAtt	Ambion, Life Technologies
			#3	s12200	CUACAAAGGC AGCACCUAAtt	Ambion, Life Technologies
RpL26	*uL24*	NM_ 000987	#1	sl2201	CCACAUUCGA AGGAAGAUUtt	Ambion, Life Technologies
			#2	s12202	GGUUGUACGU GGACACUAUtt	Ambion, Life Technologies
			#3	s12203	GAAAUAUGUU AUCUACAUUtt	Ambion, Life Technologies
RpL27	*eL27*	NM_ 000988	#1	s12204	GAGAGAUACA AGACAGGCAtt	Ambion, Life Technologies
			#2	s12205	ACAAAACUGU CGUCAAUAAtt	Ambion, Life Technologies
			#3	s12206	AAACUGUCGU CAAUAAGGAtt	Ambion, Life Technologies
RpL27 A	*uL15*	NM_ 000990	#1	sl2210	GGAUCAACUU CGACAAAUAtt	Ambion, Life Technologies
			#2	sl2211	GCCCAACUGUC AACCUUGAtt	Ambion, Life Technologies
			#3	sl2212	CACUUAAAGA GGAACCAGAtt	Ambion, Life Technologies
RpL28	*eL28*	NM_ 000991	#1	sl2213	CACCGUCUCUA AAAUAAAAtt	Ambion, Life Technologies
			#2	sl2214	GCACUGAGCCC AAUAACUUtt	Ambion, Life Technologies
			#3	sl2215	UGAUCAAGAG GAAUAAGCAtt	Ambion, Life Technologies

RpL29	*eL29*	NM_ 000992	#1	sl2216	CCUGUGCUAU UUGUACAAAtt	Ambion, Life Technologies
			#2	sl2217	CGAUCACAAA GAUACGAAUtt	Ambion, Life Technologies
			#3	si94747	GGAACAUGCG CUUUGCCAAtt	Ambion, Life Technologies
RpL30	*eL30*	NM_ 000989	#1	s12207	GGCUAUCAUU GAUCCAGGUtt	Ambion, Life Technologies
			#2	s12208	GGCUCCAACUC GUUAUGAAtt	Ambion, Life Technologies
			#3	s12209	GGAAAUCUGA AAUAGAGUAtt	Ambion, Life Technologies
RpL31	*eL31*	NM_ 000993	#1	sl2218	CUCCAGAUGU GCGCAUUGAtt	Ambion, Life Technologies
			#2	si94749	GGAAUGUGCC AUACCGAAUtt	Ambion, Life Technologies
			#3	s231333	CCAAAUAAGC UAUAUACUUtt	Ambion, Life Technologies
RpL32	*eL32*	NM_000994	#1	s12220	UGGUUAUGGA AGCAACAAAtt	Ambion, Life Technologies
			#2	sl2221	GACUGGUACU CAGAAUUUAtt	Ambion, Life Technologies
			#3	S227220	ACAGGGUUCG UAGAAGAUUtt	Ambion, Life Technologies
RpL34	*eL34*	NM_ 000995	#1	s12222	AUAGAAUUGU UUACCUUUAtt	Ambion, Life Technologies
			#2	s12223	AAAUCGUUGU GAAAGUGUUtt	Ambion, Life Technologies
			#3	s12224	CAAUACAGCC UCUAACAAAtt	Ambion, Life Technologies
RpL35	*uL29*	NM_ 007209	#1	s22151	CAAGCUCUCU AAGAUCCGAtt	Ambion, Life Technologies
			#2	s22152	GGAAAUCCAU UGCCCGUGUtt	Ambion, Life Technologies
			#3	s22153	UAACCAGACU CAGAAAGAAtt	Ambion, Life Technologies
RpL35 A	*eL33*	NM_ 000996	#1	s12225	GUAUAUAAAG CAAAGAACAtt	Ambion, Life Technologies
			#2	s12226	ACAGAAUUCU AUUUGGGCAtt	Ambion, Life Technologies
			#3	s12227	GAGAUGCGCU UAUGUAUAUtt	Ambion, Life Technologies
RpL36	*eL36*	NM_ 015414	#1	s24653	CCCUCAAAUU UAUCAAGAAtt	Ambion, Life Technologies
			#2	s24654	AAACGGGCCC UCAAAUUUAtt	Ambion, Life Technologies
			#3	s226046	GCGCCAUGGA GUUACUGAAtt	Ambion, Life Technologies
RpL36 A	*elA2*	NM_ 021029	#1	sl99106	AGUGUCAUCU UUUAUUAUGtt	Ambion, Life Technologies
			#2	s231826	GACUUUCUGU AAGAAGUGUtt	Ambion, Life Technologies
			#3	s231827	CUAAGCCGAU UUUCCGGAAtt	Ambion, Life Technologies
RpL37	*eL37*	NM_ 000997	#1	s12230	GAACAACACC	Ambion, Life

					UAAACCCAAtt	Technologies
			#2	s12231	GGUCGAAUGA GGCACCUAAtt	Ambion, Life Technologies
			#3	s12232	GCUAAAAGAC GAAAUACCAtt	Ambion, Life Technologies
RpL37 A	*elA3*	NM_000998	#1	sl2233	CACUUCCGCUG UCACGGUAtt	Ambion, Life Technologies
			#2	si94751	CGCCAAGUAC ACUUGCUCUtt	Ambion, Life Technologies
			#3	si94752	GGUAAAGUCC GCCAUCAGAtt	Ambion, Life Technologies
RpL38	*eL38*	NM_000999	#1	s12234	AGGACAACGU GAAGUUUAAtt	Ambion, Life Technologies
			#2	S194753	GGACAACGUG AAGUUUAAAtt	Ambion, Life Technologies
			#3	si94754	CCGACGAAAG GAUGCCAAAtt	Ambion, Life Technologies
RpL39	*eL39*	NM_001000	#1	s12235	UCACGAUCAU GUUACCAUAtt	Ambion, Life Technologies
			#2	si94755	GGAGAUUUCG ACGUGUUUUtt	Ambion, Life Technologies
			#3	sl94756	CCACUAUCUG GAGAUUUCGtt	Ambion, Life Technologies
RplA0	*elA0*	NM_001033930	#1	sl4556	AGACAAGGAG GGUAUCCCAtt	Ambion, Life Technologies
			#2	sl4557	AAAUACAACU GCGACAAGAtt	Ambion, Life Technologies
			#3	sl4558	GUCUGAUAUU UGCCGGCAAtt	Ambion, Life Technologies
RpIAl	*elAl*	NM_001035267	#1	sl2236	CAAUGGAUCU AGAACUUCAtt	Ambion, Life Technologies
			#2	s12237	CCACCUUGCUC AUAAACAAtt	Ambion, Life Technologies
			#3	sl2238	GUAACAACCA UAUAAUAAAtt	Ambion, Life Technologies
P0	*uLlO*	NM_001002	#1	s226965	AGAUCAUCCA ACUAUUGGAtt	Ambion, Life Technologies
			#2	s226966	CGAGGGCACC UGGAAAACAtt	Ambion, Life Technologies
			#3	s785	GUUUCAUUGU GGGAGCAGAtt	Ambion, Life Technologies
PI	*PI*	NM_001003	#1	s12239	GCACGACGAU GAGGUGACAtt	Ambion, Life Technologies
			#2	si94758	GAGCCUCAUC UGCAAUGUAtt	Ambion, Life Technologies
			#3	si94759	ACAUGGGCUU UGGUCUUUUtt	Ambion, Life Technologies
P2	*P2*	NM_001004	#1	si94760	CCAGGGUAUU GGCAAGCUUtt	Ambion, Life Technologies
			#2	si94761	GACAUGGGAU UUGGCCUUUtt	Ambion, Life Technologies
			#3	si94762	GACCGGCUCA ACAAGGUUAtt	Ambion, Life Technologies
e, eukaryotic; u, universal						

CP assembly factors	GenBank accession number	siRNA number	siRNA ID	siRNA sequence (5' to 3')	Source
BXDC1	NM_032194	SEQ ID NO:l	HSS.RNAI.N032194.12.1	rUrCrUrUrCrUrArUrArArUrCrUrUrCrUrGrUrUrArCrArUrCrGrArA	IDT
		SEQ ID NO:2	HSS.RNAI.N032194.12.2	TArArUrArCrCrUrArArUrUrCrArArUrCrArUrArUrCrCrArGrCrArC	IDT
		SEQ ID NO:3	HSS.RNAI.N032194.12.3	rUrCrUrUrGrCrUrUrCrUrGrCrArUrArUrGrArArUrCrCrUrUrCrCrA	IDT
RRS1	NM_015169	SEQ ID NO:4	HSS.RNAI.N015169.12.1	rUrCrCrUrCrCrUrCrCrCrUrCrArUrCrUrGrCrUrUrArUrUrGrGrUrG	IDT
		SEQ ID NO:5	HSS.RNAI.N015169.12.4	rArUrCrArGrCrCrArUrUrCrUrUrUrGrGrUrGrUrCrGrUrCrCrCrGrG	IDT
		SEQ ID NO:6	HSS.RNAI.N015169.12.6	rCrUrUrUrCrCrUrUrCrUrUrGrGrCrCrUrGrArArUrCrCrGrCrUrUrG	IDT

Calibration set Gene name	GenBank accession number	siRNA ID	siRNA or snoRNA sequence (5' to 3')	Source
Scramble (SCR)	/	s4390844	not available	Ambion, Life Technologies
GFP	/	AM4626	not available	Ambion, Life Technologies
Fibrillarin (FBL)	NM_0014363	s4821	UGACAUCGUUGGUCCGGAUtt	Ambion, Life Technologies
Nucleophosmin (NPM)	NM_002520	s9677	GCAAUGAAUUACGAAGGCAtt	Ambion, Life Technologies
Nucleolin (NCL)	NM_005381	s9312	GGAAAUGGCCAAACAGAAAtt	Ambion, Life Technologies
TIF1A	NM_018427	s29325	CGAUGUUGCUCUCCAUGGAtt	Ambion, Life Technologies
UTP18	NM_016001	s27417	CAUUCUCUUCAGAUAGUAAtt	Ambion, Life Technologies
NOL9	NM_024654	s36145	GCCCUGGAAGAGUUAGUCAtt	Ambion, Life Technologies
U8	NR_033294	SEQ ID NO:7	mGmGmAmUmUATCCCACCTGmAmCmGmAmU	IDT

Imaging

Imaging was performed on a Zeiss Axio Observer.Zl microscope with a motorized stage, driven by MetaMorph (MDS Analytical Technologies, Canada). Images were captured in widefield mode with a 20x objective (Plan NeoFluar, Zeiss), a LED illumination (CoolLed pE-2) and a CoolSnap HQ2 camera. Sixteen independent fields of view were captured automatically for each well. The correct focal plane was maintained by using the built-in autofocus module of MetaMorph. High-resolution images were captured

in confocal mode using a Yokogawa spindisk head and the HQ2 camera with a laser from Roper (405 nm 100 mW Vortran, 491 nm 50 mW Cobolt Calypso, and 561 nm 50 mW Cobolt Jive) and a 40x objective (Plan NeoFluar, Zeiss).

PES1 detection by indirect immunofluorescence

After 3 days siRNA-mediated depletion, cells were fixed in 2% formaldehyde, washed in PBS and blocked in PBS supplemented with 5% BSA and 0.3% Triton X-100 during 1 hour at RT. Anti-PESl antibody (anti-rat, 1:1,000; courtesy from E. Kremmer) was diluted in PBS supplemented with 1% BSA, 0.3% Triton X-100 and incubated with the cells overnight at 4°C. Cells were washed in PBS and incubated with a secondary Alexa Fluor 594 anti-rat antibody (1:1,000; Invitrogen) in PBS, 1% BSA, 0.3% Triton X-100 during 1 hour at RT. Cells were finally washed in PBS, treated with DAPI and imaged with the Zeiss microscope as described above.

Image processing and iNo index

Example 2 presents the methodology for distinguishing populations of normal and altered nucleoli, based on statistical morphometric information. Shape and textural features were first derived to characterize nucleolar morphology in individual cell nuclei, so as to distinguish normal from altered nucleoli morphology in FIB-GFP images. Each feature was systematically defined as a parametric function, so that its parameters could be optimized over the entire database to maximize Fisher's criterion computed between the distributions of the features observed in r-protein-depleted cells and SCR-treated control cells. Given these features, we then performed a quantitative analysis of differences between their statistical distributions in a population of r-protein-depleted cells, compared to their distributions in a reference population. For this, we introduced a so-called discrepancy vector, each component of this vector being associated with a specific feature, and measured the distance between the distribution observed for a population of cells depleted of a given r-protein and that observed for a reference population of cells (SCR-treated cells). We then defined the index of nucleolar disruption, iNo, as the Ll-norm of the discrepancy vector. This index reflects the degree of severity of nucleolar disruption, i.e. it ranks the r-proteins according to their impact on nucleolar structure. Additionally, Principal Component Analysis (PCA) of the discrepancy vectors was used to extract and visualize the major trends affecting the morphology of the nucleolus upon gene product depletion. PCA assumes linear embedding for dimensionality reduction and allows unsupervised clustering of the nucleolar disruption phenotypes. The computer code is described in the Supplementary information section and available upon request.

Pre-rRNA processing analysis

For the pre-rRNA analysis, we used a colon carcinoma cell line (HCT116) expressing normally p53 (39). Northern-blot analyses were performed essentially as described in (38) and www.RibosomeSynthesis.com. Briefly, HCT116 cells were transfected with one siRNA specific to transcripts encoding each r-protein in 6-well plates and incubated for 2 days prior to total RNA extraction and Northern-blot analysis. The probes used are described in Table 3. Two distinct siRNAs were used in two independent experiments. The "R" software was used to generate and cluster the heatmaps. These heatmaps are a visual representation of the logarithm of the ratio of the pre-rRNA level in the knockdown

condition respective to its level in the non-targeting (Scramble, SCR) control. The calibration set used in the pre-rRNA processing analysis consists of mock-treated cells, and cells treated with a non-targeting siRNA (Scramble, SCR), or with siRNAs specific to UTP18 or NOL9 (see Table 2 and (38) and www.RibosomeSynthesis.com). In our clustering analysis, we did not average the processing data obtained with the two different siRNAs used in this work for each r-protein, but rather, we considered them as individual experiments. In most cases, the two independent processing datasets obtained for any particular r-protein are highly clustered, demonstrating the robustness of our screens. In a few cases (denoted with a star in Fig. 4, and observed only for two SSU and six LSU r-proteins), the heatmaps do not belong to the same class, reflecting the inherent variation in depletion efficiency from one individual siRNA to another. Note that all RNA species detected were used to cluster the heatmaps shown in Fig. 4c and Fig. 4d but only those directly relevant to synthesis of the small (in panel c), or large (in panel d) subunit are shown for simplicity. The clusters with all the RNA species are shown in Supplementary Fig. 5,6. The 28S/18S rRNA ratios were calculated from Agilent bioanalyzer electropherograms according to the manufacturer's instructions. Examples of uncropped Northern blots are show in Supplementary Fig. 12.

Table 3

Oligo probe name	**Northern blot probes**	
LD1844 (5'-ETS)	CGGAGGCCCAACCTCTCCGACGACAGGTCGCCAGAGGA CAGCGTGTCAGC	SEQ ID NO:8
LD1827 (ITS1)	CCTCGCCCTCCGGGCTCCGGGCTCCGTTAATGATC	SEQ ID NO:9
LD1828 (ITS2°	CTGCGAGGGAACCCCCAGCCGCGCA	SEQ ID NO:10

Oligo probe name	**RTqPCR primers**	
LD1818 (GAPDH Forward)	TGCACCACCAACTGCTTAG	SEQ ID NO: 11
LD1819 (GAPDH Reverse)	GTTCAGCTCAGGGATGACC	SEQ ID NO: 12
LD3657 (eL22 Forward)	TTGGTGCTCTGTGGATTGAG	SEQ ID NO: 13
LD3658 (eL22 Reverse)	TTGCCAAAGAGCAACTGATG	SEQ ID NO: 14
LD3659 (eS24 Forward)	AACTGGCTTTGGCATGATTT	SEQ ID NO: 15
LD3660 (eS24 Reverse)	GCAGCACCTTTACTCCTTCG	SEQ ID NO: 16
LD3661 (uS12 Forward)	CCCACTTTCCGTAGGATCAA	SEQ ID NO: 17
LD3662 (uS12 Reverse)	GTCCTGAGGCTGGATATGGA	SEQ ID NO: 18
LD3663 (eS 10 Forward)	CGCAGAGATGTTGATGCCTA	SEQ ID NO: 19
LD3664 (eS 10 Reverse)	TCCAGGCAAACTGTTCCTTC	SEQ ID NO:20
LD2855 (eS6 Forward)	CTAGGACCAAAGCACCCAAG	SEQ ID NO:21
LD2856 (eS6 Reverse)	GGAAAGTCTGCGTCTCTTCG	SEQ ID NO:22
LD3665 (RACK1 Forward)	CAAGCTGAAGACCAACCACA	SEQ ID NO:23
LD3666 (RACK1 Reverse)	CACACAGCCAGTAGCGGTTA	SEQ ID NO:24
LD3667 (uL23 Forward)	AACTTGCCTCCATGGTTGAG	SEQ ID NO:25
LD3668 (uL23 Reverse)	TTGTCCTTTATTGGGCAAGG	SEQ ID NO:26
LD3669 (eS7 Forward)	ACGGCAGCTAAGGAAATTGA	SEQ ID NO:27

LD3670 (eS7 Reverse)	GTACGGCTTTTTCGAGTTGG	SEQ ID NO:28
LD3671 (eL41 Forward)	CTCGGCACTTAGCATCATCA	SEQ ID NO:29
LD3672 (eL41 Reverse)	TCCATGTTTCTGCTCCTGTG	SEQ ID NO:30
LD3576 (eSl_Forward)	CCGGAAGAAGATGATGGAAA	SEQ ID NO:31
LD3577 (eSl Reverse)	TCCATGAGCTTTCCCAATTC	SEQ ID NO:32
LD3673 (uS7 Forward)	GTGAACGCCATCATCAACAG	SEQ ID NO:33
LD3674 (uS7 Reverse)	AGGCACTCAGCAATGGTCTT	SEQ ID NO:34
LD2859 (eS12 Forward)	GAAGCTGCCAAAGCCTTAGA	SEQ ID NO:35
LD2860 (eS12 Reverse)	AACCACTTTACGGGGTTTCC	SEQ ID NO:36
LD3675 (uS19 Forward)	ACAACGGCAAGACCTTCAAC	SEQ ID NO:37
LD3676 (uS19 Reverse)	GGAGTCATGTGCGCCTTTAT	SEQ ID NO:38
LD3677 (uS15 Forward)	GGACTTGCTCCTGATCTTCCT	SEQ ID NO:39
LD3678 (uS15 Reverse)	AGGGCAGAGGCTGTAGATGA	SEQ ID NO:40
LD3679 (eL42 Forward)	TGATTGCTCCTACCGACTCC	SEQ ID NO:41
LD3680 (eL42 Reverse)	GGCGTACAGAGAATCCTTGC	SEQ ID NO:42
LD3431 (uL5 Forward)	GCATCCGGAGAAATGAAAAG	SEQ ID NO:43
LD3432 (uL5 Reverse)	GTCCAGGCCGTAGATACCAA	SEQ ID NO:44
LD3681 (eS8 Forward)	TGAAGAATTGCATCGTGCTC	SEQ ID NO:45
LD3682 (eS8 Reverse)	TCCTCCAGGAGACTGCTGAT	SEQ ID NO:46
LD3683 (eL29 Forward)	ACATGGCCAAGTCCAAGAAC	SEQ ID NO:47
LD3684 (eL29 Reverse)	ATTGTTGGCCTGCATCTTCT	SEQ ID NO:48
LD3785 (uS2 Forward)	AGACGGCTGTGCTGAAGTTT	SEQ ID NO:49
LD3786 (uS2 Reverse)	CAGCGCAATGGTAGGTAGGT	SEQ ID NO:50
LD3685 (eL15 Forward)	AGCTCTCTGCTCTCCACAGG	SEQ ID NO:51
LD3686 (eL15 Reverse)	TGAAGGCTTCGAGCAAACTT	SEQ ID NO:52
LD3687 (eS4 Forward)	CCTGGATCTTTTGACGTGGT	SEQ ID NO:53
LD3688 (eS4 Reverse)	TTCACCCACTGCTCTGTTTG	SEQ ID NO:54
LD3689 (uS5 Forward)	TTATGCCAGTGCAGAAGCAG	SEQ ID NO:55
LD3690 (uS5 Reverse)	ATCTTGTTCCCCCAGTAGCC	SEQ ID NO:56
LD3691 (P2 Forward)	AGCTTGCCAGTGTACCTGCT	SEQ ID NO:57
LD3692 (P2 Reverse)	GGGGAGCAGGAATTTAATCA	SEQ ID NO:58
LD3693 (eL31 Forward)	TGCCATCAACGAAGTGGTAA	SEQ ID NO:59
LD3694 (eL31 Reverse)	TATTCCTTTGGCCCAGACAG	SEQ ID NO:60
LD3695 (eS30 Forward)	TCGCTTCTTCCTCTTTCTCG	SEQ ID NO:61
LD3696 (eS30 Reverse)	GGAGCACGACTTGATCTTCC	SEQ ID NO:62
LD2861 (eS27L Forward)	GTCTGGTAGGGCTGAGCTTG	SEQ ID NO:63
LD2862 (eS27L Reverse)	ACTGTCTGAGCATGGCTGAA	SEQ ID NO:64
LD3697 (uS3 Forward)	CTGGGCATCAAGGTGAAGAT	SEQ ID NO:65
LD3698 (uS3 Reverse)	CCTGTTATGCTGTGGGGACT	SEQ ID NO:66
LD3699 (uS9 Forward)	GGCAATGGTCTCATCAAGGT	SEQ ID NO:67
LD3700 (uS9 Reverse)	GGCTTTGGAGATGGACTGAC	SEQ ID NO:68
LD3701 (eS17 Forward)	CCTGTGCTTCCTGTTTCCTC	SEQ ID NO:69
LD3702 (eS17 Reverse)	GCTATCTTGTTGCGGAGCTT	SEQ ID NO:70
LD3703 (uSll Forward)	CTTTCAGGGAGGAGCTTGTG	SEQ ID NO:71

LD3704 (uSll Reverse)	ACCTTCATCCCACCAGTCAC	SEQ ID NO:72
LD2863 (eS28 Forward)	TCCATCATCCGCAATGTAAA	SEQ ID NO:73
LD2864 (eS28 Reverse)	AGTTACGTGTGGCGGACAAA	SEQ ID NO:74
LD3705 (eL28 Forward)	ACAGACATCACGGGAGGAAG	SEQ ID NO:75
LD3706 (eL28 Reverse)	GGACAATGCTAAGGCTGCTC	SEQ ID NO:76
LD3707 (uS 14 Forward)	AACCAGAGACCCTGGCTTTT	SEQ ID NO:77
LD3708 (uS 14 Reverse)	ACTTGGGAGGCTGAGACAGA	SEQ ID NO:78
LD3709 (uLlO Forward)	TCGACAATGGCAGCATCTAC	SEQ ID NO:79
LD3710 (uL 10 Reverse)	ATCCGTCTCCACAGACAAGG	SEQ ID NO: 80
LD2045 (uS13 Forward)	GCAGCCATGTCTCTAGTGATCC	SEQ ID NO: 81
LD2046 (uS13 Reverse)	GGATCTTGTACTGGCGTGGA	SEQ ID NO: 82
LD3711 (eS26 Forward)	GAACGCATTTCCACCCTAGA	SEQ ID NO: 83
LD3712 (eS26 Reverse)	GCACGACCATTGTTCCTTCT	SEQ ID NO: 84
LD3713 (uL24 Forward)	AGGCATTTCAATGCACCTTC	SEQ ID NO: 85
LD3714 (uL24 Reverse)	CTGCACCCGTTCAATGTAGA	SEQ ID NO: 86
LD2841 (uL 18 Forward)	GCAGGATGGGGTTTGTTAAA	SEQ ID NO: 87
LD2842 (uL 18 Reverse)	ACGGGCATAAGCAATCTGAC	SEQ ID NO:88
LD3715 (uS8 Forward)	GACCTGGAAAAATGGCAGAA	SEQ ID NO: 89
LD3716 (uS8 Reverse)	CCAGAGTCCATGAGGCATTT	SEQ ID NO:90
LD2005 (uS17 Forward)	AAGATGGCGGACATTCAGAC	SEQ ID NO:91
LD2006 (uS17 Reverse)	TACCAGTGAAGGGGCATTTC	SEQ ID NO:92
LD3717 (PI Forward)	GGTCCTTCCGAGGAAGCTAA	SEQ ID NO:93
LD3718 (PI Reverse)	AACATTTACACCGGCTGCTT	SEQ ID NO:94
LD3719 (eS17 Forward)	ACAAGAAGAAACGCCTGGTG	SEQ ID NO:95
LD3720 (eS 17 Reverse)	AGTGCTGCTTCCTCCTGAAG	SEQ ID NO:96
LD3721 (eL24 Forward)	CAAATTCCAGAGGGCCATTA	SEQ ID NO:97
LD3722 (eL24 Reverse)	TTTGCTTAGGTGCTGCCTTT	SEQ ID NO:98
LD2857 (uS4 Forward)	GATTACATCCTGGGCCTGAA	SEQ ID NO:99
LD2858 (uS4 Reverse)	CGCAGAGAGAAGTCGATGTG	SEQ ID NO: 100
LD3723 (uL14 Forward)	CTGACAACACAGGAGCCAAA	SEQ ID NO: 101
LD3724 (uL14 Reverse)	ACACGCCATCTTTTCTACGG	SEQ ID NO: 102
LD3725 (eS 19 Forward)	TCTCCACCACTGTTCCTTCC	SEQ ID NO: 103
LD3726 (eS 19 Reverse)	GGTGTCTAGTGAGGGGTGGA	SEQ ID NO: 104
LD2845 (uL16 Forward)	AGCAAGGGTAGGTGTGCATC	SEQ ID NO: 105
LD2846 (uL16 Reverse)	AGACCTTTGGTCAGGTGGTG	SEQ ID NO: 106
LD3727 (eS21 Forward)	GCTGCTTCCTTTCTCTCTCG	SEQ ID NO: 107
LD3728 (eS21 Reverse)	GCCTGTGACCTTGTCAACCT	SEQ ID NO: 108

p53 steady-state level analysis

For p53 steady-state analysis, we used a colon carcinoma cell line (HCT116) expressing p53 (39). For quantitative Western-blot analysis, HCT116 cells were depleted three times independently with one siRNA specific to transcripts encoding each r-protein. The transfection protocol used was similar to the one described above in the rRNA processing analysis section. For total protein extractions, cells from 6-

well plates were first detached with 300 µl of trypsin-EDTA (ATCC) and pelleted at 100 g for 10 min at RT. Cells were washed in 1 ml of cold PBS and pelleted again at 100 g for 10 min at RT. Cells were then lysed in 30 µl of lysis buffer (Tris-HCl pH 8.0, 20 mM; NP40, 0.5%; NaCl, 150 mM; EDTA, 1 mM, protease inhibitor-Roche) during 15 min on ice. Lysed cells were then centrifuged at 20,000 g for 10 min at 4°C and supernatants were recovered from the pellet of cellular debris. As controls, we used the non-targeting scramble siRNA and an antisense oligonucleotide targeting the U8 snoRNA (IDT) (Table 2). Forty µg of total protein were separated on a 4-12 % polyacrylamide gel (Novex, Life Technologies, Bolt Bis-Tris Plus) and transferred on low-fluorescence PVDF membrane (Immobilon-FL, Millipore) according to the manufacturer protocol. The membranes were blocked in Odyssey blocking buffer (Li-Cor) for 1 hour at RT. Primary antibodies (1:4,000 anti-β-actin, Santa Cruz, SC69879; and 1:1,000 anti-p53, Bethyl Laboratories, A300-247A) were added to the Odyssey blocking buffer supplemented with 0.2% Tween-20 (Sigma) and membranes were incubated overnight at 4°C with agitation. Membranes were washed three times in TBS supplemented with 0.1% Tween-20 (TBS-T). Secondary antibodies carrying fluorescent dyes (1:2,000 DyLight 550 anti-mouse, Thermo Scientific, 84540; and 1:2,000 IRDye 680 anti-rabbit, Westburg, 926-68071) were added to Odyssey blocking buffer supplemented with 0.1% SDS and 0.2% Tween-20 and membranes were incubated 1 hour at RT with agitation. Membranes were washed three times in TBS-T before imaging of the fluorescent signals with the Chemidoc (Biorad). Cellular p53 steady-state level was assessed by calculating a ratio between the red fluorescent signal (corresponding to p53) and the green fluorescent signal (corresponding to β-actin). For each experiment, two independent lanes corresponding to HCT116 cells treated with the SCR siRNA were loaded on the gel, and the results from these two lanes were averaged to determine the level of p53 in this control condition. All data were then harmonized to this averaged value in order to determine the variation in the p53 steady-state level under this reference condition. Examples of uncropped Western-blots are shown in Supplementary Fig. 13. In Supplementary Fig. 9, the Western-blots were performed according to the same protocol, except that the gels were transferred onto PVDF (Amersham Hybond-P, RPN303F), and revealed with an HRP-conjugated secondary antibody (Santa Cruz) and the Supersignal WestPico chemiluminescent ECL substrate (Thermo Scientific). The anti-p21 antibody was purchase from Cell signalling (2947S).

For forty-eight r-proteins whose depletion did not affect the p53 level, the residual mRNA level was established by RTqPCR (see Supplementary Fig. 8). Reverse transcription was performed with the qScript cDNA supermix (Quanta Biosciences). qPCR was performed on a StepOne Plus Real-Time PCR machine (ThermoFisher Scientific) with specific primer pairs (Supplementary Table 3) and the perfecta SYBR green supermix (Quanta Biosciences). Each reaction was performed in triplicate. The residual level of mRNA was normalized to that of GAPDH and expressed with respect to that observed in cells treated with a non-targeting siRNA control (SCR).

Example 2: Image processing analysis

In this Example, we describe our procedure to extract qualitative and quantitative morphometric information from nucleoli using a low dimensional feature vector to provide a statistically validated tool to discriminate between populations of normal and altered nucleoli.

Overview

First, we describe how individual cell nuclei are segmented within each image of the database, in order to localize individual cells nucleoli. Then, we present the methodology adopted to derive a small set of shape and textural features that characterize the nucleolar morphology in each individual cell nucleus. Finally, we present a quantitative analysis of the differences observed between the distributions of these features in populations of cells depleted for specific gene products and the ones of a reference population. This leads us to use dimensionality reduction techniques, to stratify and rank the r-proteins according to their impact on the nucleolar structure. The stratification is based on a Principal Components Analysis (PCA) of the five distances (dk, k<5) measured between the distributions observed for a population of cells depleted for a given r-protein and those observed for a reference population of cells (SCR-treated cells), each distribution being associated to a specific shape/texture feature. For the ranking, we introduce an index of nucleolar disruption, or iNo, corresponding to the sum of the five absolute Idkl values.

1. Cell nuclei segmentation

As an initial step to delimit cell nucleoli from individual cells we segmented the cell nucleus. Each cell has a single nucleus. The nucleoli are specialized subnuclear domains, which, by definition, are all contained within the nucleus. Segmenting the nucleus, which is dense, compact and easy to score, is thus a mean to delimit the cellular volume that contains nucleoli in each individual cell.

Nuclei are stained with the DNA stain DAPI. The nuclei contours are extracted from the DAPI channel (blue) in two steps. Firstly, large connected components of relatively high intensity are identified using an original adaptive thresholding method. Secondly, the connected components that show significant concavity, indicating they likely correspond to aggregated nuclei, are rejected (Supplementary Fig. 14).

Step 1 consists in a stepwise thresholding of the DAPI images such that nuclei corresponding to sufficiently large connected components of pixels lying above an intensity threshold are selected. Although all parameters of samples preparation (cell seeding, transfection procedure, DAPI staining, cell fixation, etc.) and of image capture (illumination, exposure time, etc.) are fully standardized and automatized, we observed an inherent variability in the DAPI signal intensity of individual nuclei. Supplementary Fig. 15 compares two extreme cases of such variability in panels (a) and (b). To address this we adopt a hierarchical thresholding strategy that progressively refines the segmentation by considering a sequence of K increasing thresholds, while exploiting prior knowledge about the size range of human cell nuclei. Detailed nuclei segmentation pseudo-code is provided below.

Inputs:

I: image

$\lambda_1 < \lambda_2 < \cdots < \lambda_k < \cdots < \lambda_K$: thresholds on intensity

S: minimal area threshold for eligible connected component

S_{max} : maximal nucleus area

dl : boundary of the image I (first and last rows and columns)

Output:

C: set of connected components associated to segmented nuclei

$I_1 \leftarrow I > \lambda_1$

$D \leftarrow connected\_components(I\{)$

$C_1 \leftarrow \{d \in D \mid area(d) > S)$

For $k = 2 \cdots K$

$I_k \leftarrow I > \lambda_k$

$D \leftarrow connected\_components(I_k)$

$C_k \leftarrow \emptyset$

For all $c \in C_{k-1}$

$L(c) = \{d \in D \mid d \subset c\}$ % connected components in c

$L_s(c) = \{d \in L(c) \mid areaid) > S)$ % components in $L(c)$ larger than S

If $(((|«c)| == 1) \wedge (|_W\ c)| == 1) \wedge (\sim \frac{t\epsilon \quad (c))}{intensity(c)} > 1.15)) \vee (fe\ (c)| \geq 2))$

% one big connected component with important gain in intensity

% OR at least two big connected components

$C_k \leftarrow C_k \cup L_s(c)$

Else % keep c

$C_\epsilon \leftarrow C_\epsilon \cup \{c\}$

End if

End for all

End for

% Remove the connected components that are too big or that touch the image boundary

$C \leftarrow \{c \in C_K \mid (area(c) < S_{max})\ A\ is\_empty(c \cap 3/))$

The image processing code was programmed in MatLab. The 3-D models of ribosomal subunits were generated with Pymol v1.5.0.3, the images for microscopy illustrations produced with ImageJ (http://imagej.nih.gov/ij/), and the graphs generated and analyzed with Prism.

In short, let λ_k denote the k^{th} intensity threshold, with $0 < k \leq K$, and $\lambda_k < \lambda_{k+1}$, $\forall k < K$. Let I_k denote the thresholded binary image, i.e. Ik (x) = **1** if I (x) > λk, and 0 elsewhere, with $x \in [1,H] \times [1,W]$, H and W denoting the height and width of the image, respectively. We also introduce C_k to denote the set of sufficiently large (compared to a size threshold S) connected components at step k. At initialization, C_1 includes the connected components in image L that are larger than S. The set $C_{k+}i$ of connected components at step **k+1** is then derived iteratively from C_k and $I_{k+}i$, as follows. $C_{k+}i$ is initialized to the empty set. For each connected component $c \in C_k$, we considered the list L(c) of connected components in $I_{k+}i$ that are included in c. If L(c) includes at least two connected components with sizes larger than S,

then those connected components are added to C_{k+i}. If the list L(c) includes a single connected component c' larger than S, then either c' or c are added to C_{k+i}, depending on whether the gain in mean intensity between c and c' is larger or smaller than 15%, respectively. Note that a higher gain in mean intensity reveals a more accurate segmentation, and that the value of 15% was set empirically. If the list L(c) does not include any component larger than S, then c is added to C_{k+i}. The connected components in C_K that are smaller than a threshold S_{max} are expected to reasonably segment the nuclei. Practically, the threshold S and S_{max} have been set to 500 and 5000 pixels, so as to include most of the size range of human cell nuclei. Regarding $\Box_k$, we considered a sequence of thresholds increasing from 250 to 450 by steps of 20, and from 500 to 1300 by steps of 100. The resulting nuclei segmentation appeared to be relatively independent of the actual sequence used, as long as its range and granularity were sufficient.

As a second step, we consider the rejection of the connected components that likely correspond to multiple nuclei in C_K. For this, we analyzed the convexity of each connected component (Supplementary Fig. 14). Specifically, a number of lines are drawn in parallel to the two principal axis of the connected component. When the connected component is convex, either one or zero segment lies inside the contour, for all parallel lines. In contrast, when the connected component represents aggregated nuclei, it presents a strong concavity and there exist parallel lines that include two or more segments lying inside the contour. If one parallel line supports two sufficiently long inner segments that are separated by a sufficiently long outer segment, the connected component is rejected. In Supplementary Fig. 14, the length of the outer and inner segments respectively correspond to δ_{0Ut}, δ^1_{in}, and δ^2_{in}. Those lengths have to be larger than a threshold of 5 pixels to reject the component. The threshold value has been set empirically to drastically reduce the number of multiple nuclei while keeping most of the single nuclei, compared to a manually generated ground truth. C denotes the subset of C_K that includes all and only all non-rejected components.

Among the nuclei segmented in C, we were only really interested to analyze further those of cells in interphase. To identify them, the DAPI image and the distribution of FIB-GFP (green channel) in the segmented component are considered. A cell is considered to be in interphase if the DAPI is sufficiently dense and spread (if at least 50% of the pixels in the nucleus have a normalized DAPI value larger than 0.47, with the DAPI image being normalized by its maximal value.), and if the FIB-GFP is sufficiently localized (if at least 50% of the pixels in the nucleus have a normalized FIB-GFP value lower than 0.53, with the FIB-GFP image being normalized by its maximal value). The thresholds were set on the basis of a manual ground truth annotation of the images. These thresholds are quite stringent and their use results in the loss of a small fraction of cells in interphase, however our aim to only consider cells in interphase for further analysis is successfully achieved.

Finally, as a post-processing step, to ascertain all nucleoli of a cell are contained in each segmented regions, we apply basic mathematical morphology to close and enlarge each connected regions. Specifically, a dilation by a 13 x 13 structuring element was followed by a 3 x 3 erosion.

As depicted in Supplementary Fig. 15, our proposed method segments nuclei effectively, both in highly (panel a) and weakly (panel b) contrasted DAPI images. In this figure, the set of connected components

that segment interphase nuclei are depicted in red. A reasonable detection rate was achieved, in conjunction with a very small false positive rate. This result is well suited to our needs since we are not interested in detecting all nuclei but rather in collecting a sufficient number of representative nucleoli patterns samples from each FIB-GFP image.

2. Nucleoli image features

This section introduces the image features that we consider to discriminate normal and altered nucleoli morphology in FIB-GFP images. For each segmented nucleus, its FIB-GFP signal is normalized by a percentile of 99.9% and all the features presented in this section are computed on this normalized signal.

Section 2.1 introduces a number of original parameterized image features to measure the most significant visual differences observed in a set of representative nucleoli. Section 2.2 optimizes the parameters of those features, so as to maximize the discrimination between the distributions of the features in r-proteins-depleted cells and SCR-treated control cells.

2.1 Discriminant nucleolar morphometric features

The distribution of the nucleolar masses within a cell nucleus soon appeared to be subject to important stochastic variability. This is well illustrated in Supplementary Fig. 16, showing digitally resected nucleoli from control cells (panel a) and from cells depleted of specific proteins of interest (panels b, c). In these, the spatial organization of the nucleolar masses with respect to the nucleus center or to its principal axis fluctuates substantially across the images of a given panel, and do not help in differentiating the images from each panel. Hence, any features that would measure how the nucleolar topology is defined in terms of the absolute and normalized position of its components are not relevant to our problem. For example, the popular object recognition approaches that define the object appearance in terms of the intensities and gradients observed on small patches defined by their size and location in a normalized image do not help (12-14). The same holds true regarding transport-based features (15-17), since they measure the discrepancy compared to a reference distribution of masses, which is not available here.

We therefore consider the definition of a set of ad-hoc features that are independent of the actual position of the nucleoli within the nucleus, while still reflecting the spatial spreading of the nucleolar masses. Those features have been defined based on a manually-selected set of image samples, depicting typical normal and abnormal nucleolus patterns (see Supplementary Fig. 16). The manual extraction of representative samples is required to derive features that are relevant to the problem at hand, i.e. to make sure that the set of investigated features are able to discriminate among the variety of nucleolus appearances. However, to avoid (over)fitting our investigated features to those manually annotated samples, in the rest of the section, each feature is systematically defined as a parametric function, so that its parameters can be optimized over the entire database to make the feature can differentiate between normal and gene-depleted cells images (see Section 2.2).

To derive our set of parametric features, Supplementary Fig. 16 presents a set of manually-selected cells nucleoli images that are representative of the appearance diversity in reference control cells, and in cells depleted of proteins of interest. Panel c (resp. b), shows that the nucleoli from cells depleted of specific

proteins of interest (resp. nucleoli with very high level of disruption) are generally spread over large and often irregular shapes, which contrasts with the rather circular spot distribution observed in control cells (panel a).

In addition, the 3-D graphs depicted in Supplementary Fig. 17 reveal that the distribution of FIB-GFP intensity of normal nucleoli is smoother and less peaky than in nucleoli of cells depleted of a protein of interest. Those observations motivated us to use features that characterize: (i) the area of support, (ii) the shape regularity, and (iii) the variations of intensities, i.e. the texture, of the nucleolar GFP signal.

Practically, for each nucleus, all our proposed features are defined with respect to the segmentation of the nucleolus masses into a set of disjoint connected components. This segmentation is obtained by FIB-GFP image thresholding, which means that a pixel is considered to be part of the nucleolus if its intensity lies above the threshold. For each feature, the segmentation threshold parameter is defined automatically according to the method proposed in the next section, so as to maximize the separation between the distributions of the features for SCR-treated control cells and for cells depleted of proteins of interest. Hence, the segmentation threshold is a feature parameter, and might vary from one feature to the other. Other feature parameters are optimized similarly, and are thus defined automatically, as described in the next section.

Area of support:

To characterize the area of support of the nucleolus, we first consider the size and number of connected components obtained after thresholding with a so-called area segmentation threshold τ_a. Specifically,

- AA_{cc} measures the area of the largest connected component in the thresholded image, and
- AN_{cc} denotes the number of connected components in the nucleus.

In addition, to characterize the sharpness of the intensity gradient along the frontier delimiting the nucleolar masses, we introduced a sharpness index AS that measures the ratio of the nucleus pixels that respectively lie above two thresholds ¾ and ¾, with $\tau_{il} >$ ¾.

Shape and texture:

To characterize the shape and the texture of the nucleolus, we only consider the largest connected component obtained after segmentation, because the small-sized components naturally tend to reduce to single circular peaks, making the largest connected component more representative with respect to shape and texture.

Shape:

To quantify the nucleolus shape regularity, we adopt a shape segmentation threshold τ_S, and consider three distinct shape factors to characterize the shape of the largest connected component in the segmented image. Each factor describes the shape independently of its size.

They are illustrated in Supplementary Fig. 18, and correspond to:

- The elongation shape factor SE_{cc}, which is defined as the square root of the ratio of the two second order moments, λ_1 and λ_2, of the connected component c around its principal axes;

- The elliptical regularity factor SRi_{cc}, which is defined as the ratio between the area of the connected component, and the area of the smallest ellipse lying outside the connected component, and having the same center, the same principal axes, and the same elongation than the connected component.
- The concavity factor SCi_{cc}, which is defined as the ratio between the area of the connected component and the area of its convex hull.

Texture:

To characterize the nucleolar texture pattern, after having investigated without any success (data not shown) some conventional texture descriptors such as the local binary patterns (18), the region covariance (19), or the grey level aura matrices (20), we introduced a number of original scalar metrics to reflect the distribution of intensities inside the largest connected component segmented based on a texture segmentation threshold τ. Those metrics are:

- The texture histogram low tail index THi_{cc} , which measures the percentage of pixels that lie below some intensity threshold a, while being located inside the erosion of the connected component by a 3 x 3 pixels structuring element21. An erosion is applied to the shape to get rid of the low intensity pixels lying on the border of the shape;
- The texture uplands index TUi_{cc}, which is defined to be the number of connected regions lying above a threshold β, while being inside the connected component;
- The texture peaks index TPi_{cc}, which is defined to be the number of local maxima in the connected component;
- The texture valleys index TVi_{cc} , which is defined to be the number of local minima in the connected component;
- The texture local minimum $TLM_{|cc}$, which is defined as the intensity of the smallest local minimum in the connected component;

As can be observed in Table 6 and Supplementary Fig. 19, those scalar features have reasonably distinct values for the representative images depicted in Supplementary Fig. 16. Supplementary Fig. 19 illustrates that the number of local minima (red dots) is generally more important in morphologically disrupted nucleoli (panel b and c) than in control nucleoli (panel a). We thus expect that they are appropriate to differentiate normal from altered nucleoli (see below).

Table 6. Histogram low tail index (x_t=100, a=150), number of uplands region (x_il=100,x_i2=200), and smallest local minimum (x_a=100) (∞ indicates the absence of a local minimum) in the largest connected component of the texture segmented images derived from Supplementary Fig. 16. (a) nucleoli from 12 SCR-treated control cells, (b) nucleoli with high level of nucleolar disruption, (c) nucleoli from cells depleted of a protein of interest (see Supplementary Fig. 16).

(a)	1.22% 2 200	12.57% 1 ∞	2.83% 1 ∞	1.20% 1 216	1.40% 1 240

	6.02% 2 ∞	4.05% 1 ∞	0% 1 ∞	9.12% 3 173	5.05% 1 ∞
(b)	31.51% 1 104	27.72% 8 101	51.30% 5 112	29.93% 7 142	4.93% 6 190
	29.58% 5 ∞	11.17% 2 ∞	8.57% 2 ∞	20.83% 2 ∞	7.78% 7 140
(c)	22.42% 6 132	0% 1 ∞	2.18% 1 165	6.83% 3 162	11.21% 10 118
	22.51% 5 111	1.77% 2 ∞	4.78% 4 159	6.17% 1 149	14.55% 4 160

2.2. Supervised optimization of features parameters

This section explains how the parameters involved in the definition of the above features are selected to best discriminate between normal and altered nucleoli.

Following the Fisher's criterion introduced by the popular Linear Discriminant Analysis (LDA) (22), we propose to select those parameters so as to maximize the separation between the features distributions that we want to discriminate.

In short, the Fisher's optimization criterion considers the problem of estimating whether a feature can discriminate between two classes of data, knowing the feature values for a representative set of data samples from each class. A natural step to answer this question consists in looking at the average (or the mean) of the feature values from each class. Intuitively, the closer the means are, the less discriminant the feature is. This is because a large distance between the means implies that the gap between the classes is expected to be large in the corresponding feature space. However, before drawing a conclusion about this gap, we also need to account for the spreading of the features around their respective mean, so that we can decide whether a given distance between the means is significant or not. Based on this reasoning, Fisher has defined the separation between the distributions associated to two classes of observations to be the ratio of the squared distance between the means to the sum of the variance within each class (22). We adopted the same criterion to optimize the parameters of our features. In other words, a discriminant feature is one for which the class-means are well separated, measured relative to the (sum of the) variances of the data assigned to a particular class.

Formally, considering a feature, parameterized by a vector p lying in a parameter space P, we assume that the feature distributions are known as a function of p, for the two classes of observations that we want to best discriminate. Then, the Fishers' optimization criterion informs us that the vector p* that maximizes the separation between the class distributions is defined as:

$$\boldsymbol{p}^* = \underset{\boldsymbol{p} \in P}{\operatorname{argmax}} \frac{[\mu_1(\boldsymbol{p}) - \mu_2(\rho)]^2}{\sigma_1(\boldsymbol{p})^2 + \sigma_2(\boldsymbol{p})^2} \quad ,$$

where $\mu_1(\rho)$, $\mu_2(\rho)$ and $\sigma_1(\rho)$, $\sigma_2(\rho)$ respectively denote the means and standard deviations of the distributions of the feature of interest, measured with parameter p for the two classes. This Fisher's criterion is equivalent to the Welch's adaptation of the t-test (23), widely used in image-based morphometry (24).

Since it relies on the distributions of features that are observed for the two classes to discriminate, our proposed parameter optimization method has to be supervised. In our case, we know by design of our experimental set up which cells images correspond to control and (gene-depleted) test cells. Hence, we can readily identify pairs of distributions that should be discriminated one from the other. Practically, we selected the parameters so as to maximize the sum of the separation measured between each pair of distributions extracted from SCR-treated control cells, and from cells depleted of a protein of interest. Supplementary Fig. 20 presents the Fisher's optimization criterion for three different features, as a function of their associated parameter. We observe that the parameter selection significantly impact the discriminative power of the feature.

It is worth noting here that our methodology has been defined to limit the impact of supervision on the outcome of the data mining process presented in Section 3. Specifically, supervision is deliberately restricted to the independent selection of individual feature parameters, without being involved in how the resulting features will be combined in the next section, based on a strictly unsupervised approach. Moreover, by defining the features parameters to differentiate distributions of samples extracted from the same culture, we avoid biasing the selection of features induced by the exploitation of a class containing different kind of deviations compared to the reference class.

3. Nucleolar features distribution analysis

This section analyzes how the distributions of the features of nucleoli observed in SCR-treated control cells compare to those of nucleoli of cells depleted of a protein of interest. As a primary objective, we aimed at quantifying the degree of nucleolar disruption associated with the depletion of a specific protein, based on the analysis of the distribution of the features of the associated nucleoli. Therefore, we introduced a discrepancy vector. Each component of this vector is associated to a specific feature, and measures the separation between the reference distribution and the gene-depleted distribution of interest. We then defined the index of nucleolar disruption, or iNo, to be the L1-norm of the discrepancy vector. This allowed us to rank the degree of severity of nucleolar disruption.

As a second and complementary outcome, we analyzed the principal components among the set of discrepancy vectors, assuming linear embedding for dimensionality reduction. It allows for extracting and visualizing the major trends affecting the morphology of the nucleolus when it is subject to gene-depletion. This allowed us to regroup nucleolar disruption phenotypes in classes in an unsupervised fashion.

3.1. Discrepancy-based distribution characterization

A discrepancy vector is defined so as to summarize how the features distributions associated to a set of nucleoli differ from their corresponding reference distributions.

Formally, let S_r and S_k denote two sets of nucleoli images, respectively obtained from normal reference cells and from cells that have been subject to the k^{th} gene depletion process, i.e. to the k^{th} silencer. For a given image feature f, we define the discrepancy d_k between set S_k and the reference S_r to be the ratio of the difference of the mean feature values on each set to the sum of their variance. The definition naturally extends to N features fi, 0<i<N, and the i^{th} component of the discrepancy vector d_k associated to the set S_k writes

$$p^* = \underset{p \in F}{\operatorname{argmax}} \frac{[\mu_1(\rho) - \mu_2(\rho)]^2}{\sigma_1(\rho)^2 + \sigma_2(\rho)^2} ,$$

with $\mu_k(i)$, $\mu_r(i)$ and $\sigma_k(i)$, $\sigma_r(i)$ denoting the means and standard deviations of the i^{th} feature over sets S_k and S_r, respectively.

3.2. L I norm of discrepancy vectors

To quantify the disruption level associated with the depletion of a specific protein of interest, we defined an index of nucleolar disruption, or iNo, as the L I-norm of the discrepancy vector computed over the set of nucleoli images obtained from cells depleted for that given protein of interest (Fig. lc).

Letting S_k denote the set of nucleoli obtained upon treatment with silencer k, and N be the 11 features defined in Section 2.1, the nucleolus disruption index Ak is measured as:

$$\Delta_k = \sum_{i=1}^{N} |d_k(i)|$$

3.3. Principal component analysis of discrepancy vectors

Principal components analysis (PCA) is an unsupervised method for dimensionality reduction. PCA is used to visualize the most important phenotypic classes observed in our work. PCA searches for directions in the data that have the largest variance, and subsequently projects the data onto it. Following such an approach, we obtained a lower dimensional representation of the data, which removes some of the 'noisy', supposedly less meaningful, directions.

Since we are interested to score the nucleolar disruption associated with the depletion of particular proteins of interest, we applied the PCA to the discrepancy vectors that capture the average trends associated with nucleoli in cells depleted with a specific silencer, and not to individual nucleolar feature vectors.

To facilitate the interpretation of the eigenvectors associated to the principal components, we only consider the 5 features that have the largest Fisher's score, i.e. which best discriminates normal and altered nucleoli. Those features are listed in the first column of Supplementary Table 7.

Table 7. Two most significant PCA vectors.

Feature	Principal component	Second principal component
AAi_{cc}	0.4808	-0.4864
$S_{R\,icc}$	-0.5503	-0.6191
TH_{icc}	0.3772	0.3789
TLM_{icc}	-0.4470	0.2151
TV_{icc}	0.3521	-0.4363

Fig. 1b presents a PCA scatter plot depicting the 2-D points obtained by projecting each discrepancy vector on the two most significant PCA components. We observed that the PCA analysis has successfully found linear combinations of the proposed features that separate out the ground truth clusters, corresponding to different levels of disruption and phenotypic classes. We indeed observe visually that the nature of the disruption changes depending on the position in the scatter plot.

Supplementary Table 7 presents the first two principal components, i.e. the two directions of maximal variability of the projected discrepancy vectors. It reveals the main trends in the disruption process.

From the signs of the components in the first vector, we learned that the dominant disruption process increases the area of the nucleolus (AAlcc), and reduces its elliptical regularity (SRlcc). It also increases the number of low intensity pixels in the segmented nucleoli (THlcc), as well as the deepness (TLMlcc) and number of local minima (TVlcc), which reflects the scattered nature of the nucleoli spread.

The second vector induces an opposite trend compared to the one induced by the first vector, except for the elliptical regularity (SRlcc) and for the histogram low tail index (THlcc). Hence, a positive second PCA coefficient tends to foster the decrease of elliptical regularity, while a negative second PCA coefficient mitigates it, compared to what would result from the single vector only. This is reflected in Fig. 1b by more regular and circular shapes of the nucleolus masses in case of negative second PCA coefficient. In contrast, a positive second PCA coefficient corresponds to a more severe disruption, with less regular shape than the one observed for nucleoli of similar size in absence of second PCA component.

It is well established in the literature that inhibition of RNA polymerase I (Pol I) leads to a very specific nucleolar morphology alterations referred to as "nucleolar segregation" or "nucleolar caps" (25). Such caps are for example observed when cells are treated with low doses of actinomycin D, or in our experimental set up, in cells depleted of the Pol I transcription factor TIF1A (Fig. 1b, caps are seen as tiny bright dots).

As an illustration that the PCA analysis is a powerful method to classify in an unsupervised fashion distinct nucleolar disruption phenotypes, nucleoli of cells depleted of TIF1A, which in agreement with the literature (25) have a markedly different nucleolar disruption phenotype by comparison to the other control or test cells, correspond to two magenta triangles totally isolated in the upper left corner of the graph and characterized by a negative first PCA component value, which is in contrast to most other magenta triangles corresponding to cells depleted for other test genes.

References specific for Example 2 and Supplementary Figures

1. Tafforeau, L. *et al.* The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Molecular cell* **51,** 539-551 (2013).
2. O'Donohue, M.F., Choesmel, V., Faubladier, M., Fichant, G. & Gleizes, P.E. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *The Journal of cell biology* **190,** 853-866 (2010).
3. Robledo, S. *et al.* The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *Rna* **14,** 1918-1929 (2008).

4. Chen, S.S. & Williamson, J.R. Characterization of the ribosome biogenesis landscape in E. coli using quantitative mass spectrometry. *Journal of molecular biology* **425,** 767-779 (2013).
5. Sykes, M.T. & Williamson, J.R. A complex assembly landscape for the 30S ribosomal subunit. *Annual review of biophysics* **38,** 197-215 (2009).
6. Talkington, M.W., Siuzdak, G. & Williamson, J.R. An assembly landscape for the 30S ribosomal subunit. *Nature* **438,** 628-632 (2005).
7. Mulder, A.M. *et al.* Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science* **330,** 673-677 (2010).
8. Jomaa, A. *et al.* Functional domains of the 50S subunit mature late in the assembly process. *Nucleic acids research* **42,** 3419-3435 (2014).
9. Ferreira-Cerca, S. *et al.* Analysis of the in vivo assembly pathway of eukaryotic 40S ribosomal proteins. *Molecular cell* **28,** 446-457 (2007).
10. Ferreira-Cerca, S., Poll, G., Gleizes, P.E., Tschochner, H. & Milkereit, P. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Molecular cell* **20,** 263-275 (2005).
11. Gamalinda, M. *et al.* A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. *Genes & development* **28,** 198-210 (2014).
12. Viola, P. & Jones, M. Robust real-time object detection. *IJCV* **57,** 137-154 (2004).
13. Dalai, N. & Triggs, B. Histograms of oriented gradients for human detection. *IEEE International Conference on Computer Vision and Pattern Recognition* (2005).
14. Felzenszwalb, P., Girshick, R.B., McAllester, D. & Ramanan, D. Object detection with discriminatively trained part-based models. *IEEE Trans on PAMI32,* 1627-1645 (2010).
15. Gangbo, W. & McCann, R. The geometry of optimal transportation. *Acta Math* **177,** 113-161 (1996).
16. Haker, S., Angenent, S. & Tannenbaum, A. Minimizing flows for the monge-kantorovich problem. *SIAM J. Math Analysis* **35** (2003).
17. Wang, W., Slepcev, D., Basu, S., Ozolek, J.A. & Rohde, G.K. A linear optimal transportation framework for quantifying and visualizing variations in sets of images. *Int. J. Computer Vision* **101,** 254-269 (2013).
18. Ojala, T., Pietikainen, M. & Harwood, D. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* **29,** 51-59 (1996).
19. Tuzel, O., Porikli, F. & Meer, P. Region covariance: a fast descriptor for detection and classification. *ECCV* (2006).
20. Qin, X. & Yang, Y.-H. Similarity measure and learning with gray level aura matrices (GLAM) for texture image retrieval. *CVPR* (2004).
21. Serra, J. *Image analysis and mathematical morphology.* (1982).
22. Fisher, R.A. The use of multiple measurmements in taxonomic problems. *Annals of Eugenics* **7,** 179-188 (1936).
23. Welch, B.L. The generalization of "Student's" problem when several different population variances are involved. *Biometrika* **34,** 28-35 (1947).
24. Wolfe, P. *et al.* Using nuclear morphometry to discriminate the tumorigenic potential of cells: a comparison of statistical methods. *Cancer Epidmiol Biomarkers Prev* **13,** 976-988 (2004).
25. Hernandez-Verdun, D., Roussel, P., Thiry, M., Sirri, V. & Lafontaine, D.L.J. The nucleolus: structure/function relationship in RNA metabolism. *Wiley interdisciplinary reviews. RNA* **1,** 415-431 (2010).

Example 3: Effects of r-protein depletion on nucleolar structure

Human cells stably producing the nucleolar methyltransferase fibrillarin (FBL) fused to a green fluorescent protein (GFP) were transfected with siRNAs targeting the appropriate transcripts, incubated for 3 days, and imaged by fluorescence microscopy (see Methods). Each r-protein was depleted in three experiments, a different siRNA being used in each experiment. The entire screen was duplicated. A non-targeting siRNA (SCR), mock-treated cells (MOCK), and a calibration set were included (see Methods

and Supplementary Fig. 1). The calibration set consisted of proteins whose depletion leads to moderate to severe nucleolar disruption, formation of nucleolar "caps" (see below), or a reduction in fluorescence intensity (Supplementary Fig. 1).

To characterize nucleolar morphology defects both qualitatively and quantitatively, we developed a specific image-processing algorithm (see Example 2). Briefly, we first segmented the observed nuclei on the basis of shape- and size-consistent adaptive thresholding of a nuclear stain (DAPI signal). Then, within each nuclear mass, GFP signal thresholding and mathematical morphology (see Example 1) were applied to segment nucleoli into connected components. In order to optimize discrimination of nucleoli of cells depleted of an r-protein from those of SCR-treated control cells, five shape and textural features were extracted from the largest connected components (LCC) of each nucleolus. These five features, selected from a set of eleven as the most discriminant ones, were: area, elliptical regularity, percentage of pixels below an optimized intensity threshold, smallest intensity, and number of local minima (see Methods). For each of the five features, a d_k value corresponding to a statistically significant distance between the feature distribution in cells depleted of an r-protein and control cells was computed. Each population of cells was thus characterized by five d_k values. Principal component analysis (PCA) was used to reduce these five dimensions to two, allowing ready visualization of the data in a scatter plot (Fig. lb) where each dot corresponds to a population of cells treated with one siRNA.

The PCA revealed groups of proteins whose depletion leads to similar nucleolar morphological phenotypes (Fig. lb). Four major groups emerged. The largest one, indicated by a gray ellipse containing the SCR control (shown as a red dot), comprises r-proteins whose depletion had no significant impact on nucleolar structure. Importantly, most of the r-proteins are in this group, i.e. nearly all of the SSU proteins (shown in green) and roughly two-thirds of the LSU proteins (in magenta). A second group, beneath the SCR control, comprises proteins whose depletion did not alter the nucleolar structure but reduced the fluorescence intensity (e.g. control cells treated with siRNAs against GFP or FBL, see also Supplementary Fig. 1). Cells depleted of the RNA Pol I transcription factor TIF1A formed distinctive "nucleolar caps", in keeping with the known effects of RNA Pol I inhibitions (14), and appeared isolated in the upper left part of the graph. The fourth group comprises the few r-proteins whose depletion was found to impact nucleolar structure very severely, remarkably they are almost exclusively LSU proteins. This cluster forms a tail in the right part of the graph. In cells depleted of these major contributors to normal nucleolar structure, the nucleoli were detected as "unfolded beaded necklaces" (Fig. lb). Our automated classification was benchmarked with a manual one and found to be extremely robust (Supplementary Fig. 2).

To stratify the r-proteins according to the severity of nucleolar disruption caused by their absence, we defined an index of nucleolar disruption (iNo) as the sum of the absolute values of the five d_k distances (see Methods). For each r-protein, an average iNo, based on the values obtained with the three different siRNAs used, was calculated and plotted. In the resulting graph, the r-proteins are listed from top to bottom in the order of increasing impact on nucleolar structure (Fig. lc). As concluded from the PCA, depletion of most r-proteins appears to have no significant impact on nucleolar structure (iNo <0.05), and

the proteins whose depletion has the greatest effect belong to the LSU (magenta). Unexpectedly, the r-proteins uL5 (formerly RPL11, (24)) and uL18 (formerly RPL5) appear among the strongest contributors to maintenance of nucleolar structural integrity (Fig. lc). These are precisely the proteins which, together with the 5S rRNA, form a small ribonucleoprotein complex, the 5S RNP, which acts as an HDM2 trap and controls the steady-state level of p53 in a regulatory circuit known as p53-dependent anti-tumor nucleolar surveillance (25, 26). Briefly, in unstressed cells, p53 is constitutively targeted for proteosomal degradation by Hdm2-mediated ubiquitination. In the event of a nucleolar stress, such as a ribosome biogenesis dysfunction, unassembled ribosomal components accumulate. These include the 5S RNP, which interacts with Hdm2, sequestering it away from p53. As a result, p53 is stabilized and induces cell cycle arrest and cell death (9). In mature 60S subunits, the 5S RNP constitutes the central protuberance (CP), a late-assembling structure (see below).

Ribosomal subunit assembly is a sequential process involving progressive binding of r-proteins to nascent rRNAs and gradual formation of ribosomal landmarks (23, 27-31). We wondered if the r-proteins important for nucleolar structure might map to particular areas on mature ribosomal subunits. Color-coding of the r-proteins according to their iNo values, on a 3-D model based on the crystal structure of the human ribosome32 (Fig. 2a), made it obvious that the strongest contributors to nucleolar structure maintenance belong to the LSU and are not randomly distributed over it: rather, they are preferentially located at the subunit interface in areas corresponding to the CP, the L1-stalk, and a region directly below the Ll-stalk (Fig. 2a). All of these are late-forming structures (see below).

The nucleolus is a highly dynamic structure capable of responding through profound morphological alterations to cellular stresses such as drug treatment or viral infection (20). In interphase, however, it is quite stable. It is disassembled at the onset of mitosis and reassembled at the end of this process (14). In our nucleolar screens, cells were imaged after three days of r-protein depletion, as we reasoned that cells might have to undergo at least two cycles of nucleolar breakdown/nucleolar genesis for nucleolar alterations to become readily detectable. This assumption was confirmed when we established the time course of the appearance of nucleolar morphological defects (Supplementary Fig. 3). Focusing on thirteen representative r-proteins, and monitoring changes at 24-h intervals over a 3-day depletion period, we indeed found nucleolar disruption to increase steadily (Supplementary Fig. 3b), in parallel with an increase in iNo values. Nucleolar disruption became obvious only after 72 h of depletion (Supplementary Fig. 3a).

The nucleoli of cells of amniotic organisms have three nucleolar subcompartments (12,13,33). In our original screens, we used FBL, a dense fibrillar component (DFC) marker, to assess nucleolar morphology. To extend our conclusions, we examined whether nucleolar structural defects due to r-protein depletion might be equally observable with a marker of a different nucleolar subcompartment. We chose to monitor by immunofluorescence a granular component (GC) marker, the PESl antigen, in depletion experiments focusing on thirteen representative r-proteins (Supplementary Fig. 4). As expected for a GC protein, PESl staining was peripheral to the FBL signal (Supplementary Fig. 4b). Remarkably, we observed extreme closeness between the iNo scores computed from the FBL and PESl signals, and

the ranking of r-proteins according to phenotype severity was largely similar (Supplementary Fig. 4a). We conclude that the nucleolar structural defects due to r-protein depletion can be monitored similarly with a DFC or a GC antigen.

We conducted our nucleolar screens in HeLa cells because of the large size of their nucleus, which makes them ideal for use in high-throughput screens with visual readouts (e.g. 34, 35). To see how general the effects observed in HeLa-GFP-FBL cells might be, we tested five cell lines: two cervical carcinoma cell lines (HeLa-GFP-FBL and HeLa), one colon carcinoma cell line (HCT116), and two lung cancer cell lines (A549 and H1944). We selected eight representative r-proteins, depleted them for three days in each of the five cell lines, and monitored nucleolar structure by immunostaining of endogenous PESl and iNo score computation (Fig. 3). For the r-proteins tested, we found the weak and strong contributors to nucleolar structure maintenance to be largely the same in all five cell lines (Fig. 3), with uL5 and uL18 playing an important role in each case.

Example 4: Effects of r-protein depletion on pre-rRNA processing

In an attempt to correlate the effects of r-protein depletion on nucleolar structure with defects in ribosome biogenesis, we determined which r-proteins are essential to pre-rRNA processing (Fig 4). Mature rRNAs are produced from long precursor molecules. They are embedded in non-coding spacers and require extensive processing to be generated (2, 36). Pre-rRNA processing analysis is a good proxy for ribosomal assembly analysis, because failure of an r-protein to bind to nascent ribosomes leads to ribosome biogenesis blockade, pre-rRNA processing inhibitions, and subunit biogenesis abortion (23, 30, 31, 37).

The synthesis of each of the eighty r-proteins was knocked down for 2 days in HCT116 cells with an appropriate siRNA. Total RNA was then extracted, run on a bioanalyzer, and analyzed by high-resolution quantitative northern blotting. Two different siRNAs were used for each r-protein and yielded largely similar results (Fig 4). As controls, we used UTP18 and NOL9 because their depletion leads to well-established pre-rRNA processing defects (Supplementary Figs 5,6,11; see (38)). As further controls, non-targeting siRNA (SCR) and mock-treated cells were used (Supplementary Figs 5,6). HCT116 and HeLa cells are both of epithelial origin and, as shown above, their nucleolar structure is similarly affected by r-protein depletion (Fig. 3). We performed our RNA processing work and p53 steady-state accumulation analysis (see below) on HCT116 cells because, unlike HeLa cells, they express p53 normally39. For the RNA analysis, cells were depleted for only 2 days, as we had established beforehand, precisely in HCT116 cells, that bona fide pre-rRNA processing inhibitions are early defects preceding cell-cycle arrest and apoptosis and are best captured at this time point (discussed in (38)).

The ratio of 28S to 18S mature rRNA was extracted from bioanalyzer electropherograms (Fig. 4a). The accumulation of small subunit 18S rRNA was strongly decreased, and the 28S/18S ratio accordingly increased, by SSU r-protein depletion (Fig. 4a). Reciprocally, LSU r-protein depletion led to decreased accumulation of the large subunit 28S rRNA and to a reduced 28S/18S ratio (Fig. 4a). Northern blots were probed with specific radioactively labeled oligonucleotides, revealing all major known pre-rRNA intermediates (Fig. 4b, Supplementary Figs. 5,6,11). Each band detected was quantified with a

phosphoimager and normalized with respect to the SCR control. The signals were represented on heatmaps (Fig. 4c for SSU r-proteins, Fig. 4d for LSU r-proteins and Supplementary Fig. 11; see also (38)). The heatmaps were clustered with the software "R", revealing functionally related groups of r-proteins whose depletion affects similar processing steps (Fig. 4c,d, and Supplementary Figs S5,6 and Supplementary Fig. 11). For the SSU r-proteins, three groups emerged: proteins whose depletion affects early processing (class 1), late processing (class 3), or has no significant effect on processing (class 2)(Fig. 4c, see representative examples in Supplementary Fig. 5 and Supplementary Fig. 11 for a full dataset). Our classification of the SSU r-proteins corresponds largely to that previously established in HeLa cells (23). We identified four classes of LSU r-proteins (Fig. 4d and Supplementary Figs. 6,11): those whose depletion affects early cleavage steps (class 1), intermediate cleavage steps (class 3), late cleavage steps (class 4), or has no substantial impact on processing (class 2)(Fig. 4d and Supplementary Fig. 6). Importantly, no such classification of LSU r-proteins has been reported previously. Our classification of r-proteins' involvement in pre-rRNA processing thus confirms and largely extends previous work (Supplementary Fig. 7).

The r-proteins were mapped on a 3-D model of the human ribosome according to their involvement in processing (Fig. 2b). This revealed, on both subunits, a strikingly asymmetric distribution. On the mature small subunit, the r-proteins required for early processing steps are those forming the body and platform (Fig. 2b), both of which are known as early-assembling subunit structures (27,28,40). The r-proteins affecting late cleavage steps, in contrast, correspond to the head and beak (Fig. 2b), which are late-forming structures (27,28,40). On the large subunit, the r-proteins important for early processing are mainly exposed on the solvent side of the ribosome (in blue on the right-hand-side cartoon Fig. 2b), while those required for intermediate cleavages are at the interface side (in orange on the left-hand-side cartoon), below the LI stalk, and those important for late processing correspond largely to the CP and Ll-stalk (in red). Remarkably, this is precisely the order in which these structures have been shown to form in budding yeast (30).

A comparison of our nucleolar structure and rRNA processing data reveals that the r-proteins whose depletion has the greatest effect on nucleolar structure (in red in Fig. 2a) are largely those required for intermediate or late processing steps in the formation of the large ribosomal subunit (in orange and red in Fig. 2b). Within this subunit, they belong mostly to late-assembling structures, including the Ll-stalk and CP (uL5 and uL18). In conclusion, while practically all r-proteins appear important for pre-rRNA processing, most of them have no incidence on the structural integrity of the nucleolus.

Several trans-acting factors, including BXDC1 and RRS1, are required for CP assembly (26,41). This function is conserved between yeast and human (26,41). Considering the strong effect of uL5 or uL18 depletion on nucleolar structure, and because both of these proteins are CP components, we predicted that depletion of factors involved in CP formation should also cause profound nucleolar structure alterations. This proved to be true: we found depletion of BXDC1, RRS1, or both to affect nucleolar structure severely (Fig. 5a), almost as strongly as does uL5 or uL18 depletion (Fig. 5b).

Example 5: Effects of r-protein depletion on p53 steady-state levels

Given the numerous connections between p53 and ribosomal component synthesis on the one hand (42), and between the functional integrity of the nucleolus and p53 metabolic stability on the other (43), we examined systematically how depletion of each individual r-protein might affect the steady-state level of p53. Colon carcinoma cells expressing p53 (HCT116 p53+/+, (39)) were transfected with one siRNA targeting each r-protein transcript and incubated for 2 days. In this analysis we used a single siRNA, selected on the basis of its proven efficacy in the nucleolar and processing screens, and carried out depletion for 2 days to allow a direct comparison with the RNA analysis. Total protein was then extracted and analyzed by quantitative fluorescent western blotting (Fig. 6a and Supplementary Fig. 11).

The p53 steady-state level increase observed ranged from 0 to 10-fold (Fig. 6b). About a third of the r-proteins (24/80) were found to affect p53 level at least 5-fold (Fig. 6b, gray box), the cut-off we adopted arbitrarily as significance threshold. As observed for the effects on nucleolar structure, we found nearly all of these r-proteins to belong to the LSU, the sole exception being eS31. Interestingly, depletion of uL5 or uL18, involved in p53-dependent nucleolar surveillance (discussed above), had no significant impact on the p53 steady-state level, in keeping with previous reports (25,26,44) and in contrast to the role of these proteins in forming an Hdm2 trap when they accumulate in cells (44-46). As an additional control we established by RTqPCR, for forty-eight r-proteins whose depletion did not significantly affect p53 accumulation (see Fig. 6b, from eL22 to eS21), the efficiency of r-protein depletion at the mRNA level (Supplementary Fig. 8). We found depletion to be effective for all the r-proteins tested, the residual mRNA level for most of them (40 out of 48) being below 20%. Note that 31 of the 48 candidates tested showed a marked processing defect upon depletion (see Fig. 4, Supplementary Figs S5,S6,S11), a further indication that depletion was efficient.

In view of the model of nucleolar stress above, we wondered if the significant increase in p53 observed upon depletion of 24 r-proteins might involve uL5 and uL18. We found this to be the case: co-depletion of any one of the 24 r-proteins and either uL5 or uL18 led to normal levels of both p53 and its transcriptional target p21 (Supplementary Fig. 9a). The effects of BXDC1 and RRS1 were also investigated. As expected from the role of these proteins as ribosome (CP) assembly factors, their depletion also caused p53 and p21 to increase, and this rise was dependent on uL5 and uL18 (Supplementary Fig. 9b).

Figure 2c shows the distribution of the r-proteins on mature subunits according to their impact on p53 expression. This shows that the significant contributors to p53 homeostasis all correspond to late-assembling structures on the subunits (Fig. 2b,c).

Table 5. Comparison between the effects of r-protein depletion on nucleolar structure integrity and p53 homeostasis

SSU	iNo	p53	LSU	iNo	p53
eS1 (Rps3a)	0.02/0.05	0.8	**uL1 (RpL10a)**	0.17/0.19	7.2

uS2 (RpSA)	0.02/0.05	1.0	**uL2 (RpL8)**	0.16/0.19	7.1
uS3 (RpS3)	0.01/0.04	1.3	**uL3 (RpL3)**	0.04/0.06	4.2
uS4 (RpS9)	0.02/0.04	2.2	**uL4 (RpL4)**	0.04/0.05	4.2
eS4 (RpS4)	0.02/0.03	1.1	**uL5 (RpL11)**	0.19/0.20	1.0
uS5 (RpS2)	0.02/0.03	1.1	**uL6 (RpL9)**	0.04/0.04	6.3
eS6 (RpS6)	0.02/0.04	0.6	**eL6 (RpL6)**	0.05/0.05	4.8
uS7 (RpS5)	0.03/0.04	0.8	**eL8 (RpL7a)**	0.02/0.03	7.2
eS7 (RpS7)	0.02/0.05	0.8	**uL10 (P0)**	0.03/0.04	1.7
uS8 (RpS15a)	0.03/0.05	1.9	**uL11 (RpL12)**	0.04/0.03	3.3
eS8 (RpS8)	0.02/0.03	1.0	**uL13 (RpL13a)**	0.07/0.07	8.5
uS9 (RpS16)	0.03/0.05	1.4	**eL13 (RpL13)**	0.03/0.05	8.0
uS10 (RpS20)	0.03/0.05	4.6	**uL14 (RpL23)**	0.02/0.04	2.2
eS10 (RpS10)	0.03/0.07	0.5	**eL14 (RpL14)**	0.04/0.04	10.0
uS11 (RpS14)	0.03/0.09	1.4	**uL15 (RpL27a)**	0.05/0.12	5.2
uS12 (RpS23)	0.02/0.03	0.5	**eL15 (RpL15)**	0.03/0.02	1.1
eS12 (RpS12)	0.01/0.04	0.9	**uL16 (RpL10)**	0.03/0.03	2.5
uS13 (Rps18)	0.03/0.06	1.7	**uL18 (RpL5)**	0.19/0.20	1.9
uS14 (RpS29)	0.05/0.06	1.6	**eL18 (RpL18)**	0.03/0.05	3.5
uS15 (RpS13)	0.02/0.11	0.9	**eL19 (RpL19)**	0.08/0.1	5.4
uS17 (Rps11)	0.02/0.04	2.0	**eL20 (RpL18a)**	0.05/0.05	6.2
eS17 (RpS17)	0.01/0.04	1.4	**eL21 (RpL21)**	0.11/0.16	8.3
uS19 (RpS15)	0.03/0.05	0.9	**uL22 (RpL17)**	0.05/0.06	8.2
eS19 (RpS19)	0.04/0.06	2.4	**eL22 (RpL22)**	0.02/0.02	0.5
eS21 (RpS21)	0.02/0.04	2.9	**uL23 (RpL23a)**	0.02/0.04	0.7
eS24 (RpS24)	0.02/0.04	0.5	**uL24 (RpL26)**	0.03/0.06	1.8
eS25 (RpS25)	0.02/0.04	3.2	**eL24 (RpL24)**	0.03/0.07	2.2
eS26 (RpS26)	0.01/0.05	1.8	**eL27 (RpL27)**	0.05/0.11	6.3
eS27 (RpS27)	0.03/0.03	2.1	**eL28 (RpL28)**	0.02/0.03	1.5
eS27L (RpS27L)	0.03/0.04	1.3	**uL29 (RpL35)**	0.06/0.04	5.0
eS28 (RpS28)	0.02/0.03	1.5	**eL29 (RpL29)**	0.03/0.03	1.0
eS30 (RpS30)	0.05/0.05	1.3	**uL30 (RpL7)**	0.06/0.06	5.1
eS31 (RpS27a)	0.03/0.03	6.9	**eL30 (RpL30)**	0.07/0.15	5.7
RACK1	0.02/0.02	0.6	**eL31 (RpL31)**	0.05/0.09	1.1
			eL32 (RpL32)	0.03/0.05	4.8
			eL33 (RpL35a)	0.02/0.04	8.9
			eL34 (RpL34)	0.04/0.07	8.3
			eL36 (RpL36)	0.03/0.04	5.5
			eL37 (RpL37)	0.04/0.10	6.9
			eL38 (RpL38)	0.08/0.11	8.9
			eL39 (RpL39)	0.04/0.09	10.4

eL40 (RpL40)	0.04/0.08	3.2
eL41 (RpL41)	0.02/0.02	0.8
eL42 (RpL36a)	0.03/0.05	0.9
eL43 (RpL37a)	0.13/0.20	5.7
P1	0.02/0.04	2.1
P2	0.02/0.03	1.1

The r-proteins are listed according to a recently revised nomenclature (Ban et al, 2014, PMID 24524803). The former nomenclature is provided in parentheses. The index of nucleolar disruption (iNo score) ranges from 0 (normal nucleolus) to 0.2 (disrupted nucleolus) (see Fig 1C). The iNo values were computed from two independent screens (screen I/screen II). The p53 level is expressed in fold increase (from 0 to 10.4, see Fig 4B). Arbitrarily set significance cut-off: p53 fold increase ≥ 5 ; iNo ≥ 0.1. **Dark grey,** above cut-off; medium grey, below cut-off; white, curated "near cut-off cases. SSU, small subunit r-protein; LSU, large subunit r-protein; iNo, index of nucleolar disruption. The iNo scores and p53 levels were established in epithelial human cells. The iNo scores were determined in HeLa-FBL-GFP cells after r-protein depletion for 3 days; the p53 levels in HCT1 16 p53 +/+ cells after r-protein depletion for 2 days.

No effect on nucleolar structure & no effect on p53 level: 50 r-proteins

No effect on nucleolar structure & p53 induction: 2 1 r-proteins

Effect on nucleolar structure & no p53 induction: 2 r-proteins (uL5 and uL18)

Effect on nucleolar structure & p53 induction: 8 r-proteins

uL5 and uL18 are unique in that, remarkably, they are the only two r-proteins, out of eighty, whose depletion impact strongly nucleolar structure without inducing p53.

References

1. Boisvert, F.M., van Koningsbruggen, S., Navascues, J. & Lamond, A.I. The multifunctional nucleolus. *Nature reviews. Molecular cell biology* **8,** 574-585 (2007).
2. Lafontaine, D.L.J. Noncoding RNAs in eukaryotic ribosome synthesis and function. *Nature Structural and Molecular Biology* **22,** 11-19 (2015).
3. Steitz, T.A. A structural understanding of the dynamic ribosome machine. *Nature reviews. Molecular cell biology* **9,** 242-253 (2008).
4. Melnikov, S. *et al.* One core, two shells: bacterial and eukaryotic ribosomes. *Nature structural & molecular biology* **19,** 560-567 (2012).
5. Schmeing, T.M. & Ramakrishnan, V. What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461,** 1234-1242 (2009).
6. Danilova, N. & Gazda, H.T. Ribosomopathies: how a common root can cause a tree of pathologies. *Dis Model Mech* **8,** 1013-1026 (2015).
7. Narla, A. & Ebert, B.L. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115,** 3196-3205 (2010).
8. Zhang, Y. & Lu, H. Signaling to p53: ribosomal proteins find their way. *Cancer cell* **16,** 369-377 (2009).
9. Chakraborty, A., Uechi, T. & Kenmochi, N. Guarding the 'translation apparatus': defective ribosome biogenesis and the p53 signaling pathway. *Wiley interdisciplinary reviews. RNA* **2,** 507-522 (2011).

10. Lam, Y.W., Evans, V.C., Heesom, K.J., Lamond, A.I. & Matthews, D.A. Proteomics analysis of the nucleolus in adenovirus-infected cells. *Molecular & cellular proteomics : MCP* **9,** 117-130 (2010).
11. Moore, H.M. *et al.* Quantitative proteomics and dynamic imaging of the nucleolus reveal distinct responses to UV and ionizing radiation. *Molecular & cellular proteomics : MCP* **10,** Mill 009241 (2011).
12. Thiry, M. & Lafontaine, D.L.J. Birth of a nucleolus: the evolution of nucleolar compartments. *Trends in cell biology* **15,** 194-199 (2005).
13. Lamaye, F., Galliot, S., Alibardi, L., Lafontaine, D.L.J. & Thiry, M. Nucleolar structure across evolution: the transition between bi- and tri-compartmentalized nucleoli lies within the class Reptilia. *Journal of structural biology* **174,** 352-359 (2011).
14. Hernandez-Verdun, D., Roussel, P., Thiry, M., Sirri, V. & Lafontaine, D.L.J. The nucleolus: structure/function relationship in RNA metabolism. *Wiley interdisciplinary reviews. RNA* **1,** 415-431 (2010).
15. Sirri, V., Hernandez-Verdun, D. & Roussel, P. Cyclin-dependent kinases govern formation and maintenance of the nucleolus. *The Journal of cell biology* **156,** 969-981 (2002).
16. Leung, A.K. *et al.* Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. *The Journal of cell biology* **166,** 787-800 (2004).
17. Derenzini, M., Montanaro, L. & Trere, D. What the nucleolus says to a tumour pathologist. *Histopathology* **54,** 753-762 (2009).
18. Bywater, M.J. *et al.* Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. *Cancer cell* **22,** 51-65 (2012).
19. Peltonen, K. *et al.* A targeting modality for destruction of RNA polymerase I that possesses anticancer activity. *Cancer cell* **25,** 77-90 (2014).
20. Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.M. & Lamond, A.I. The nucleolus under stress. *Molecular cell* **40,** 216-227 (2010).
21. Hein, N., Hannan, K.M., George, A.J., Sanij, E. & Hannan, R.D. The nucleolus: an emerging target for cancer therapy. *Trends in molecular medicine* **19,** 643-654 (2013).
22. de la Cruz, J., Karbstein, K. & Woolford Jr, J.L. Functions of Ribosomal Proteins in Assembly of Eukaryotic Ribosomes In Vivo. *Annual review of biochemistry* **84,** 93-129 (2015).
23. O'Donohue, M.F., Choesmel, V., Faubladier, M., Fichant, G. & Gleizes, P.E. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *The Journal of cell biology* **190,** 853-866 (2010).
24. Ban, N. *et al.* A new system for naming ribosomal proteins. *Curr Opin Struct Biol* **24,** 165-169 (2014).
25. Donati, G., Peddigari, S., Mercer, C.A. & Thomas, G. 5S ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2-p53 checkpoint. *Cell reports* **4,** 87-98 (2013).
26. Sloan, K.E., Bohnsack, M.T. & Watkins, N.J. The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress. *Cell reports* **5,** 237-247 (2013).
27. Chen, S.S. & Williamson, J.R. Characterization of the ribosome biogenesis landscape in E. coli using quantitative mass spectrometry. *Journal of molecular biology* **425,** 767-779 (2013).
28. Talkington, M.W., Siuzdak, G. & Williamson, J.R. An assembly landscape for the 30S ribosomal subunit. *Nature* **438,** 628-632 (2005).
29. Sashital, D.G. *et al.* A combined quantitative mass spectrometry and electron microscopy analysis of ribosomal 30S subunit assembly in E. coli. *eLife* **3** (2014).
30. Gamalinda, M. *et al.* A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. *Genes & development* **28,** 198-210 (2014).
31. Ferreira-Cerca, S. *et al.* Analysis of the in vivo assembly pathway of eukaryotic 40S ribosomal proteins. *Molecular cell* **28,** 446-457 (2007).
32. Anger, A.M. *et al.* Structures of the human and Drosophila 80S ribosome. *Nature* **497,** 80-85 (2013).
33. Thiry, M., Lamaye, F. & Lafontaine, D.L.J. The nucleolus: when 2 became 3. *Nucleus* **2,** 289-293 (2011).
34. Schmitz, M.H. *et al.* Live-cell imaging RNAi screen identifies PP2A-B55alpha and importin-betal as key mitotic exit regulators in human cells. *Nature cell biology* **12,** 886-893 (2010).

35. Wild, T. *et al.* A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS biology* **8,** el000522 (2010).
36. Mullineux, S.T. & Lafontaine, D.L.J. Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? *Biochimie* **94,** 1521-1532 (2012).
37. Robledo, S. *et al.* The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *Rna* **14,** 1918-1929 (2008).
38. Tafforeau, L. *et al.* The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Molecular cell* **51,** 539-551 (2013).
39. Bunz, F. *et al.* Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* **282,** 1497-1501 (1998).
40. Sykes, M.T. & Williamson, J.R. A complex assembly landscape for the 30S ribosomal subunit. *Annual review of biophysics* **38,** 197-215 (2009).
41. Zhang, J. *et al.* Assembly factors Rpf2 and Rrsl recruit 5S rRNA and ribosomal proteins rpL5 and rpLl 1 into nascent ribosomes. *Genes & development* **21,** 2580-2592 (2007).
42. Ruggero, D. & Pandolfi, P.P. Does the ribosome translate cancer? *Nature reviews. Cancer* **3,** 179-192 (2003).
43. James, A., Wang, Y., Raje, H., Rosby, R. & DiMario, P. Nucleolar stress with and without p53. *Nucleus* **5** (2014).
44. Fumagalli, S., Ivanenkov, V.V., Teng, T. & Thomas, G. Suprainduction of p53 by disruption of 40S and 60S ribosome biogenesis leads to the activation of a novel G2/M checkpoint. *Genes & development* **26,** 1028-1040 (2012).
45. Sun, X.X., Wang, Y.G., Xirodimas, D.P. & Dai, M.S. Perturbation of 60 S ribosomal biogenesis results in ribosomal protein L5- and Lll-dependent p53 activation. *The Journal of biological chemistry* **285,** 25812-25821 (2010).
46. Bursac, S. *et al.* Mutual protection of ribosomal proteins L5 and Lll from degradation is essential for p53 activation upon ribosomal biogenesis stress. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 20467-20472 (2012).
47. Lo, K.Y. *et al.* Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. *Molecular cell* **39,** 196-208 (2010).
48. Ferreira-Cerca, S., Poll, G., Gleizes, P.E., Tschochner, H. & Milkereit, P. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Molecular cell* **20,** 263-275 (2005).
49. Ben-Shem, A. *et al.* The structure of the eukaryotic ribosome at 3.0 A resolution. *Science* **334,** 1524-1529 (2011).
50. Khatter, H., Myasnikov, A.G., Natchiar, S.K. & Klaholz, B.P. Structure of the human 80S ribosome. *Nature* **520,** 640-645 (2015).
51. Madru, C. *et al.* Chaperoning 5S RNA assembly. *Genes & development* **29,** 1432-1446 (2015).
52. Leidig, C. *et al.* 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. *Nature communications* **5,** 3491 (2014).
53. Bassler, J. *et al.* A network of assembly factors is involved in remodeling rRNA elements during preribosome maturation. *The Journal of cell biology* **207,** 481-498 (2014).
54. Yusupov, M.M. *et al.* Crystal structure of the ribosome at 5.5 A resolution. *Science* **292,** 883-896 (2001).
55. Rubbi, CP. & Milner, J. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *The EMBO journal* **22,** 6068-6077 (2003).
56. Fumagalli, S. *et al.* Absence of nucleolar disruption after impairment of 40S ribosome biogenesis reveals an rpLl 1-translation-dependent mechanism of p53 induction. *Nature cell biology* **11,** 501-508 (2009).

Claims

1. A method for predicting, diagnosing, or determining nucleolar disruption or integrity, comprising determining the protein or mRNA expression level of one or more, such as two or more, such as three or more, such as all (ribosomal) protein selected from the group comprising or consisting of any of (a) to (PP):

(a) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eL13, and eL33 ;

(b) uLl, uL2, eL21, eL38, uL13, uL22, eL14, eL39, eL34, eL13, eL8, and eL33;

(c) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eL13, eL8, and eL33;

(d) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eS31, eL13, eL8, and eL33;

(e) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eL13, eL8, eL33, eL32, eL6, uL4, uL3, eL18, uLll, eL40, uL16, uL14, and eL24;

(f) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eS31, eL13, eL8, eL33, eL32, eL6, uSIO, uL4, uL3, eL18, uLll, eS25, eL40, eS21, uL16, eS19, uL14, uS4, eL24, eS27, uS17, and uS8;

(g) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eL13, eL8, eL33, eL32, eL6, uL4, uL3, eL18, uLll, eL40, uL16, uL14, eL24, uLlO, eL28, eL27L, eS30, eL31, eL15, and eL29;

(h) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eS31, eL13, eL8, eL33, eL32, eL6, uSIO, uL4, uL3, eL18, uLll, eS25, eL40, eS21, uL16, eS19, uL14, uS4, eL24, eS27, uS17, uS8, uL24, eS26, uS13, uLlO, eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14;

(i) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL31, eL6, eL20, uL15, eL27, eL14, uL4, eL37, eL39, eL34, uL6 , eL13, and eL36;

(j) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, uS14, eL31, eL6, eL20, uL15, eL27, eS30, eL14, uL4, eL37, eL39, eL34, uL6, eS19, eSlO, eL13, eS31, and eL36;

(k) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL31, eL6, eL20, uL15, eL27, eL14, uL4, eL37, eL39, eL34, uL6 , eL13, eL36, eL33, and eL8;

(1) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, uS14, eL33, eL8, eL13, eS31, eL36, eSlO, eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, and eL31;

(m) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL33, eL8, eS27, eL36, uL24, eL18, eL32, eL29, uL16, eL24, eL15, eL42, uLlO, eL13, uLll, eL40, uL3, uL6, eL34, eL39, eL37, uL4, eL14, eL27, uL15, eL20, eL6, and eL31;

(n) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL33, eS24, eL8, uS4, eS28, eS27, uS7, uS9, uS19, eL36, uSll, uS8, uL24, eL18, eL32, eS27L, eL29, uL16, uS13, eL24, eL15, uSIO, eSlO, eL42, uLlO, eL13, eS31, uLll, eL40, uL3, eS19, uL6, eL34, eL39, eL37, uL4, eL14, eS30, eL27, uL15, eL20, eL6, eL31, and uS14;

(o) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eS31, eL13, eL8, eL33, eL32, eL6, uSIO, uL4, uL3, eL18, uLll, eS25, eL40, eS21, uL16, eS19, uL14, uS4, eL24, eS27, uS17, uS8, uL18, uL24, eS26, uS13, uLlO, eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, and uS14;

(p) uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL20, uL15, eL27, eL14, eL37, eL39, eL34, uL6, eL36, eL13, eL8, eL33, eL32, eL6, uL4, uL3, eL18, uLll, eL40, uL16, uL14, eL24, uL24, uL18, uLlO, eL28, eL27L, eS30, eL31, eL15, and eL29;

(q) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, uS14, eL31, eL6, eL20, uL15, eL27, eS30, eL14, uL4, eL37, eL39, eL34, uL6, eS19, and eSlO;

(r) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, uS14, eL31, eL6, eL20, uL15, eL27, eS30, eL14, uL4, eL37, eL39, eL34, uL6, and eS19;

(s) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL31, eL6, eL20, uL15, eL27, eL14, uL4, eL37, eL39, eL34, and uL6;

(t) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, uS14, eL31, eL6, eL20, uL15, eL27, eS30, eL14, uL4, and eSlO;

(u) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL31, eL6, eL20, uL15, eL27, eL14, and uL4;

(v) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, uS14, eL31, eL6, eL20, and eSlO;

(w) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, uL22, eL31, eL6, and eL20;

(x) uL18, uL5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, uL30, and uL22;

(y) ull8, ul5, uLl, uL2, eL43, eL21, eL19, eL38, uL13, eL30, uL29, and uL30;

(z) ul18, ul5, uLl , uL2, eL43, eL21, eL19, eL38, uL13, uL29, and eL30;

(aa) ul18, ul5, uLl , uL2, eL43, eL21, eL19, eL38, and uL13;

(bb) ul18, ul5, uLl , uL2, eL43, eL21, eL19, and eL38;

(cc) ul18, ul5, uLl , uL2, eL43, eL21, eL19, and eSlO;

(dd) ul18, ul5, uLl , uL2, eL43, eL21, and eL19;

(ee) ul18, ul5, uLl , uL2, eL43, eL21, and eSlO;

(ff) ul18, ul5, uLl , uL2, eL43, and eL21;

(gg) ul18, ul5, uLl , uL2, eL43, and eSlO;

(hh) ul18, ul5, uLl , uL2, and eL43;

(ii) ul18, ul5, uLl , and uL2;

(jj) uL18, uL5, and uLl;

(kk) uL5, anduL18;

(11) uL18;

(mm) uL5;

(nn) BXDC1 and RRS1, or any protein involved in ribosomal central protuberance formation;

(00) any one or more ribosomal protein selected from the above lists, except uL24, uL14, eS7, eL8, eS6, eS19, uSll, eL37, eS26, eS31, eS27L, uS12, and uSIO; or

(pp) any one or more ribosomal protein selected from the above lists, except uL5 and/or uL18.

2. A method for increasing or decreasing the protein level of p53 in a cell, comprising respectively decreasing or increasing the mRNA or protein level of one or more, such as two or more, such as three or more, such as all ribosomal proteins selected from the group comprising or consisting of any of (a) to (i):

(a) eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14, uLlO, uS13, eS26, uL24, uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, anduLl;

(b) eL29, eL15, eL31, eS30, eL27L, eL28, uLlO, uL24, eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(c) uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(d) eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(e) eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(f) eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(g) eL33, eL8, eL13, eL34, eL39, eL14, uL22, uL13, eL38, eL21, uL2, anduLl;

(h) eL29, uS2, eL15, eS4, uS5, eL31, eS30, eL27L, uS3, uS9, eS17, uSll, eS28, eL28, uS14, uLlO, uS13, eS26, uL24, uL18, uS8, uS17, eS27, eL24, uS4, uL14, eS19, uL16, eS21, eL40, eS25, uLll, eL18, uL3, uL4, uSIO, eL6, eL32, eL33, eL8, eL13, eS31, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, and uLl;

(1) eL29, eL15, eL31, eS30, eL27L, eL28, uLlO, uL24, uL18, eL24, uL14, uL16, eL40, uLll, eL18, uL3, uL4, eL6, eL32, eL33, eL8, eL13, eL36, uL6, eL34, eL39, eL37, eL14, eL27, uL15, eL20, uL22, uL30, uL29, eL30, uL13, eL38, eL19, eL21, eL43, uL2, anduLl;

3. An inhibitor or inducer of one or more ribosomal proteins selected from the group comprising or consisting of any of (a) to (pp) as defined in claim 1 or any of (a) to (i) as defined in claim 2 for use in

respectively increasing or decreasing the protein level of p53 in a cell, or for use in therapy, or for use in altering nucleolar morphology, architecture, or integrity.

4. A method for characterizing nucleolar morphology, architecture, or integrity, comprising measuring or determining one or more morphometric characteristics of the nucleolar support area and the nucleolar intensity pattern;

optionally wherein the morphometric characteristics are defined by parametric functions, preferably to characterize, in a measurable and quantifiable manner, the visual appearance of the nucleolus in an image, including morphology or form (encompassing size, shape, and number of subcomponents) and inner pattern structure (encompassing spatial organization and spatial variation of pixel intensity).

5. The method according to claim 4, wherein nucleolar morphology, architecture, or integrity is characterized based on eukaryotic cell or tissue images, preferably based on a population of cells.

6. The method according to claim 4 or 5, which is a computer implemented method and/or an automated method.

7. The method according to any of claims 4 to 6, comprising the steps:

a) obtaining or providing an image of one or more eukaryotic cell, or eukaryotic tissue, or preferably of a population of cells ;

b) optionally segmenting said image thereby obtaining individual cells or cell nuclei;

c) segmenting said image, thereby obtaining one or more nucleolar masses, preferably comprising or consisting of a plurality of connected pixels, which connected pixels constitute or represent a nucleolar connected component;

d) optionally assigning by colocalization nucleolar masses to individual cells or cell nuclei and thus to individual cells;

e) extracting for each cell or nucleus one or more morphometric characteristic characterizing the support area, and/or the intensity pattern of the nucleolar connected components optionally assigned to the cell, preferably the largest one;

f) computing probabilistic distributions (over the set of cells) of said morphometric characteristics; and

g) comparing said probabilistic distributions with a reference model distribution and/or determining the overlap of said probabilistic distributions to a reference model distribution.

8. The method according to any of claims 4 to 7, comprising defining a set of discriminant nucleolar morphometric characteristics, preferably derived by:

a) optionally providing or obtaining cell (population) or tissue image samples from cell(s) (populations) or tissues that have to be differentiated;

b) defining a set of arbitrary parametric nucleolar morphometric characteristics to capture the main salient differences observed between representative cell nucleolus samples of two populations of interest;

c) selecting the parameters associated to each nucleolar morphometric characteristic to maximize its discriminating power, i.e. to minimize the overlap between the morphometric characteristic probabilistic distributions derived from the cell(s) (populations) to differentiate;

d) selecting as discriminant morphometric characteristics the morphometric characteristics and their associated optimized parameters leading to a large separation between cell(s) (populations) to discriminate.

9. The method according to any of claims 4 to 8,

wherein said nucleolar support area is characterized by morphometric characteristics that determine the size and shape of the spatial support of the nucleolar masses, preferably wherein said morphometric characteristics count the number of distinct nucleoli or distinct connected components contributing to the nucleoli, and/or the number of pixels that are part of each of those components, and/or wherein said morphometric characteristics include any quantitative metric that reflects the elongation, (e.g. aspect ratio), and/or elliptical regularity, and/or concavity of the nucleolar masses; preferably wherein all morphometric characteristics are parametrized by the level of (pixel) intensity above which a pixel is considered to be part of the nucleolar masses; and/or

wherein said nucleolar intensity pattern is characterized by morphometric characteristics that determine the nucleolar intensity patterns and/or nucleolar connected components intensity patterns, preferably the spatial variation of the nucleolar signal within the nucleoli or nucleolar connected components, preferably the largest one; wherein spatial variations are preferably measured inside and on the boundary of the nucleoli or nucleolar connected components, and wherein preferably in the neighborhood of the boundary of a nucleolus or nucleolar connected component, the sharpness of transition of nucleolus-related pixel intensity between the nucleolar mass and the background image is measured; and wherein preferably inside a nucleolus or nucleolar connected component, the morphometric characteristics are defined to capture the spatial variations of pixel intensities; optionally wherein the magnitude and/or number of local minima and/or local maxima of nucleolus-related pixel intensity values are measured, and/or the number of pixels having a relatively small intensity value and/or the number and/or shape of high intensity sub-regions that are separated by small intensity valleys inside a nucleolus or nucleolar connected component, preferably wherein all morphometric characteristics are parametrized by thresholds adopted to define the notions of small/high intensity, and/or by the structuring elements considered to define the notions of inside/outside a nucleolus or nucleolar connected component, preferably based on mathematical morphology.

10. The method according to any of claims 7 to 9, wherein segmenting involves determining whether or not the intensity of a pixel lies above a threshold, and preferably wherein the pixel is considered to be part of the physical element of interest, such as nucleus or nucleolus, when the intensity of the pixel lies above the threshold; wherein the threshold is preferably determined automatically;

optionally wherein the threshold considered for segmentation is defined using an automatic and locally adaptive method that selects the threshold within a pre-defined set of values, preferably as the

largest value for which the area of the connected set of pixels above the threshold is larger than a lower bound of the size expected for the physical element such as a nucleus or nucleolus to segment; and/or

optionally wherein the threshold considered for segmentation is defined to maximize the discriminant power of said discriminant nucleolar morphometric characteristics.

11. The method according to any of claims 4 to 10,

wherein said nucleolar support area is determined by measuring the nucleolus area, preferably of the largest nucleolus or largest nucleolar connected component in a cell, the number of nucleoli or nucleolar connected components in a cell, and/or the sharpness index, preferably of the largest nucleolus or largest nucleolar connected component in a cell;

optionally, wherein the sharpness of transition of nucleolus-related pixel intensity between the nucleolus of nucleolar connected component and the background image is measured as the ratio of the nucleolar pixels that respectively lie above two pixel intensity thresholds x_i and x_2, with $x_i > x_2$.

12. The method according to any of claims 4 to 11, wherein said nucleolar support area is determined by measuring the nucleolar support shape regularity, preferably by measuring the elongation shape factor, the elliptical regularity factor, and/or the concavity factor preferably of the largest nucleolus or largest nucleolar connected component in a cell; preferably by measuring the elliptical regularity factor;

optionally, wherein said elongation shape factor is determined based on the two second order moments λ_1 and λ_2 of the nucleolus or nucleolar connected component around its principal axes, preferable as the square root of the ratio of the two second order moments λ_1 and λ_2 of the nucleolus or nucleolar connected component around its principal axes,

optionally, wherein the elliptical regularity factor is determined as the ratio between the area of the nucleolus or nucleolar connected component and the area of the smallest ellipse lying outside the nucleolus or nucleolar connected component and having the same center, the same principal axes, and the same elongation as the nucleolus or nucleolar connected component, and/or

optionally, wherein the concavity factor is determined as a decreasing function of the ratio between the area of the nucleolus or the nucleolar connected component and the area of its convex hull.

13. The method according to any of claims 4 to 12, wherein said nucleolar intensity pattern is determined by measuring the distribution of pixel intensities preferably in the largest nucleolus or largest nucleolar connected component in a cell; preferably following cell-wise dynamic range normalization; preferably by measuring the texture histogram low tail index, the texture uplands index, the texture peaks index, the texture valleys index, and/or the texture local minimum, preferably of the largest nucleolus or largest nucleolar connected component in a cell; preferably by measuring the texture histogram low tail index, the texture valleys index, and the texture local minimum;

optionally, wherein the texture histogram low tail index is defined as the percentage of pixels that lie below a pixel intensity threshold a,

optionally wherein the texture uplands index is defined as the number of connected nucleolar pixels having their intensity above a threshold β,

optionally wherein the texture peaks index is defined as the number of local maxima in the nucleolus or nucleolar connected component,

optionally wherein the texture valleys index is defined as the number of local minima in the nucleolus or nucleolar connected component, and/or

optionally wherein the texture local minimum is defined as the intensity of the smallest local minimum in the nucleolus or nucleolar connected component.

14. The method according to any of claims 4 to 13, comprising determining:

- the area of the largest nucleolus in a cell;
- the elliptical regularity factor;
- the texture histogram low tail index;
- the texture valleys index; and
- the texture local minimum.

15. The method according to any of claims 4 to 14, wherein the nucleolus is identified based on image pixel intensity thresholding, preferably in an image obtained by identifying or quantifying a nucleolus-specific feature (e.g. density, optical height) or marker (e.g. nucleolus specific or enriched molecule), such as FBL or PES1; preferably as a connected component in a thresholded image;

optionally wherein the nucleolar support area is determined based on an area segmentation threshold, wherein the nucleolar shape regularity is determined based on a shape segmentation threshold, and/or the nucleolar texture pattern is determined based on a texture segmentation threshold, each threshold being preferably defined to maximize the discriminant power of corresponding morphometric characteristics.

16. The method according to any of claims 4 to 15, wherein the nucleolar morphology, architecture, or integrity is compared to a standard or reference nucleolar morphology, architecture, or integrity, and preferably wherein the nucleolar morphology, architecture, or integrity is scored based on the discrepancy with respect to the standard or reference nucleolar morphology, architecture, or integrity, preferably wherein said standard or reference nucleolar morphology, architecture, or integrity is a normal nucleolar morphology, architecture, or integrity.

17. The method according to claim 16, wherein said discrepancy is determined as a discrepancy vector based on the distances between the distributions of each morphometric characteristic of a population of said nucleoli or nucleolar connected components and the distribution of the corresponding morphometric characteristic of a population of standard or reference nucleoli or nucleolar connected components, wherein the discrepancy vector component associated with a given morphometric characteristic (i) for a particular nucleolus or nucleolar connected component population d is preferably determined according to the following formula representing a discrepancy vector component:

$$\mathbf{d}(i) = \frac{\mu(i) - \mu_r(i)}{\sqrt{\sigma^2(i) + \sigma_r^2(i)}}$$

wherein $\mu(i)$ and $\mu_r(i)$ are the mean of morphometric characteristic (i) of said nucleoli or nucleolar connected components and said standard or reference nucleoli or nucleolar connected components, respectively; and wherein $\sigma(i)$ and $\sigma_r(i)$ are the standard deviation of morphometric characteristic (i) of said nucleoli or nucleolar connected components and said standard or reference nucleoli or nucleolar connected components, respectively.

18. The method according to any of claims 4 to 17, wherein nucleolar morphology, architecture, or integrity is characterized based on the index of nucleolar disruption (iNo) which is derived from an increasing function of the absolute value of the morphometric characteristic discrepancy vector components, preferably as the Ll-norm of the discrepancy vector, wherein said iNo is determined according to the following formula:

$$iNo = \sum_{i=1}^{N} a(i) * |d(i)|)$$

wherein N is the number of morphometric characteristics; wherein a(i) is a weighing factor for the i^{th} morphometric characteristic, wherein preferably $0 < a(i) < 1$; and wherein d(i) is the discrepancy vector component for the i^{th} morphometric characteristic.

19. The method according to any of claims 7 to 18, further comprising regrouping nucleolar disruption phenotypes in classes and/or presenting them in lower dimension by means of unsupervised clustering and/or principal components analysis.

20. A method for scoring, quantifying, or classifying nucleolar disruption and/or for determining the health status of an individual or of a cell or tissue or for diagnosis, comprising performing the method according to claim 18 or 19, wherein the iNo value is indicative of the severity of said nucleolar disruption, or the severity or grade of said health status;

optionally wherein said health status or diagnosis is indicative of cancer, autoimmune disease, viral infection, neurodegenerative disorder, or ribosomopathy (such as cancer predisposition, skeletal problems, or hematological problems).

21. A method for quantifying nucleolar integrity or nucleolar disruption, comprising the steps of:

a) identifying one or more nucleolar morphometric characteristic capable of discriminating between different nucleolar morphologies or architectures, preferably capable of discriminating between normal and disrupted nucleolar morphology, architecture, or integrity, preferably according to Fisher' s criterium;

b) assigning a numerical value to each of said one or more morphometric characteristic, wherein said value correlates with the distance between said morphometric characteristic of said normal and disrupted nucleolar morphologies or architectures; and

c) integrating all numerical values to obtain a global score which quantifies nucleolar integrity or nucleolar disruption.

22. A method for identifying a set of one or more nucleolar morphometric characteristics that discriminate between two cells or cell populations, comprising the steps of:

a) defining a set of arbitrary parametric morphometric characteristics to capture the most salient morphometric differences observed between representative nucleoli samples of two populations of interest;

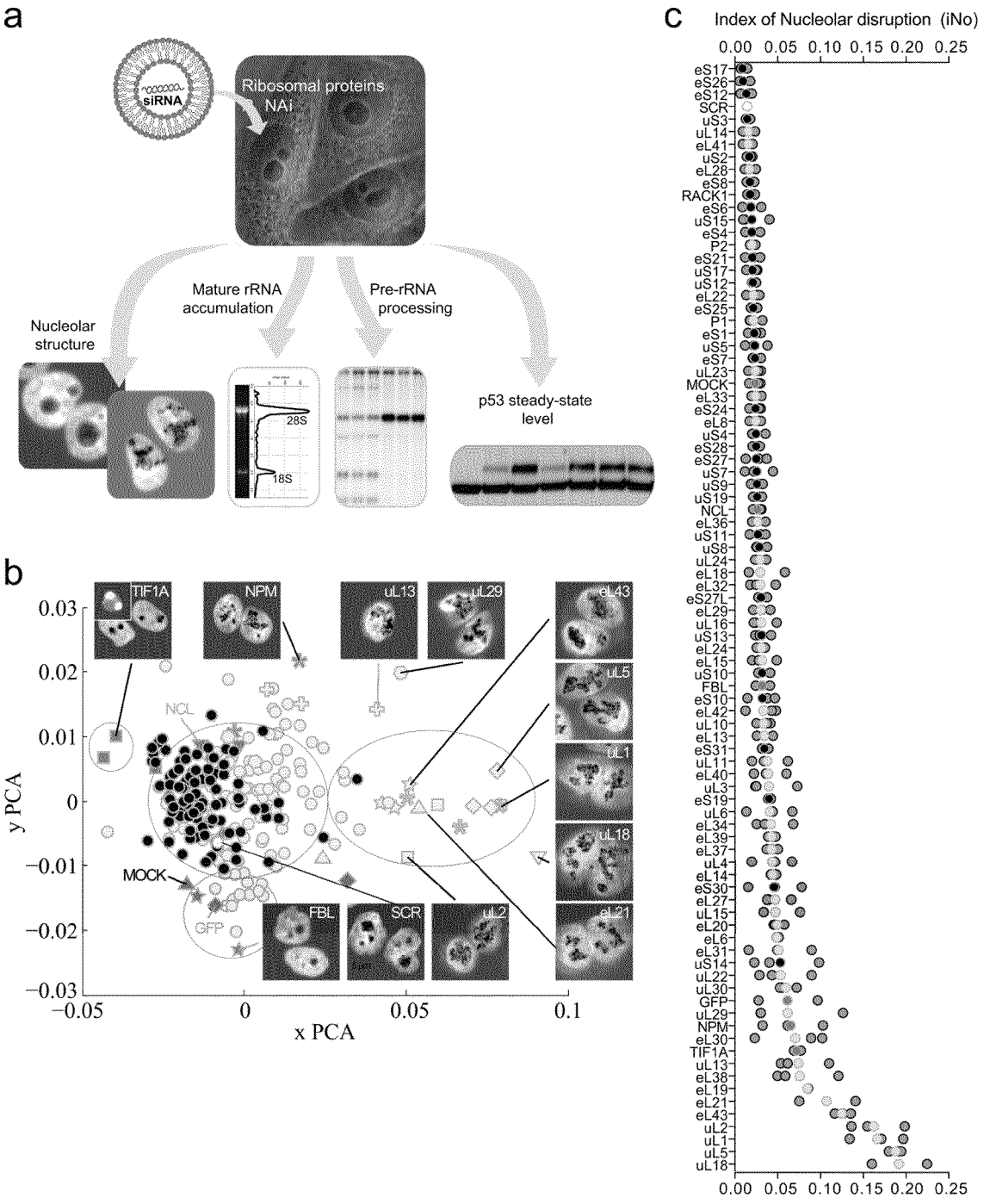
b) selecting the parameters associated to each morphometric characteristic to maximize their discriminating power and/or to minimize the overlap between the morphometric characteristic probabilistic distributions derived from each population, preferably according to Fisher's criterium; and

c) selecting the morphometric characteristics and their associated optimized parameters leading to a large separation between populations to discriminate;

d) computing the distance between the two cells or cell populations based on a combination, e.g. using a weighted summation, of the distances measured in terms of the selected morphometric characteristics.

23. Use of a method according to any of claims 1,2 or 4 to 20 for identifying compounds affecting nucleolar morphology, architecture, integrity, or disruption or for identifying compounds affecting p53 stability and/or steady state levels.

FIG. 1


a
siRNA
Ribosomal proteins NAi
Mature rRNA accumulation
Pre-rRNA processing
Nucleolar structure
28S
18S
p53 steady-state level
b
TIF1A
NPM
uL13
uL29
eL43
uL5
uL1
uL18
eL21
NCL
MOCK
GFP
FBL
SCR
uL2
y PCA
x PCA
c
Index of Nucleolar disruption (iNo)
0.00 0.05 0.10 0.15 0.20 0.25

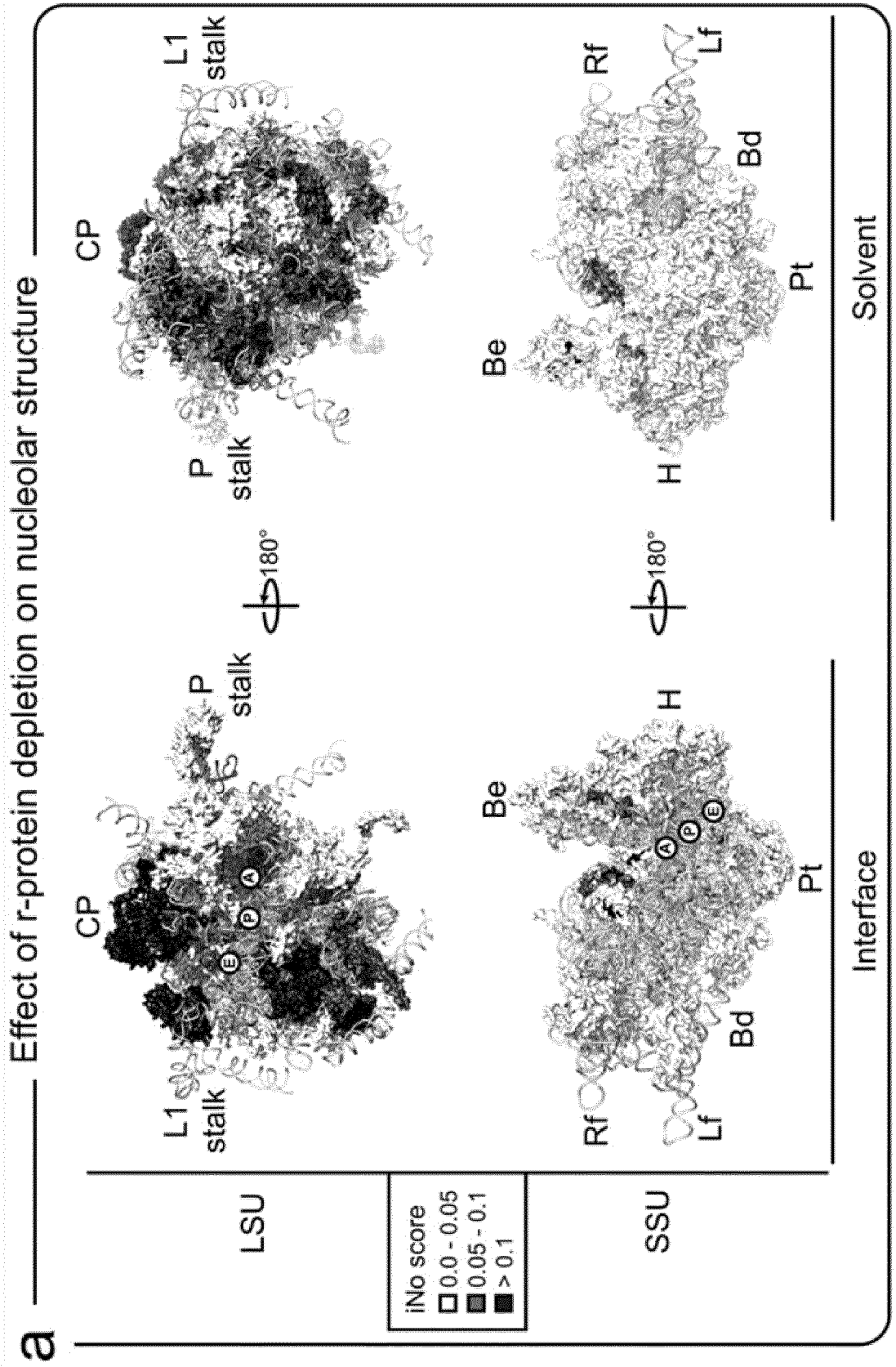

a
Effect of r-protein depletion on nucleolar structure
LSU
SSU
iNo score
0.0 - 0.05
0.05 - 0.1
> 0.1
CP
L1 stalk
P stalk
180°
Be
Rf
Lf
Bd
Pt
H
Interface
Solvent

Fig.2

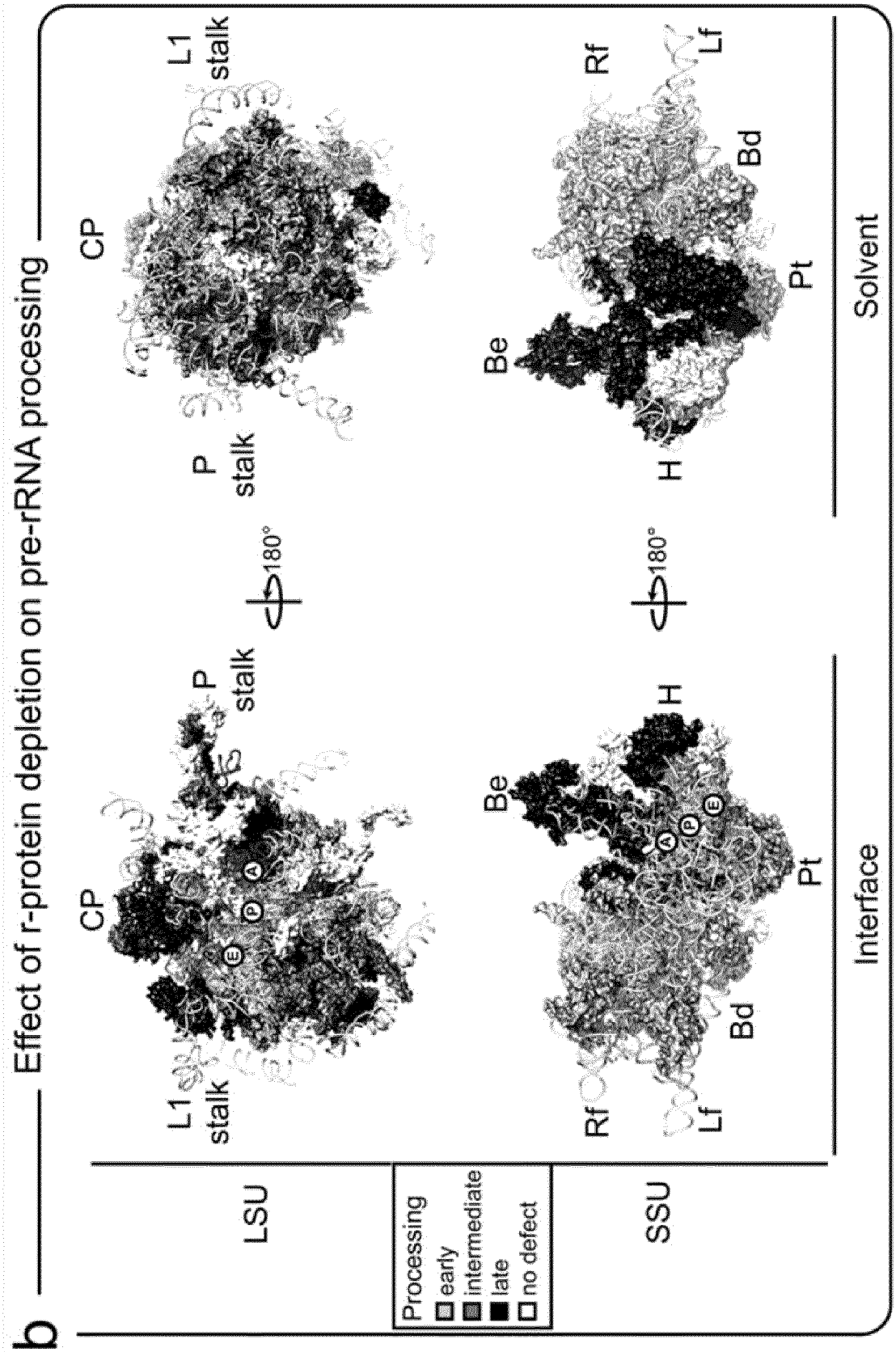

b
Effect of r-protein depletion on pre-rRNA processing
LSU
SSU
Processing
early
intermediate
late
no defect
CP
L1 stalk
P stalk
E
P
A
180°
Be
H
Rf
Lf
Bd
Pt
Interface
Solvent

Fig.2

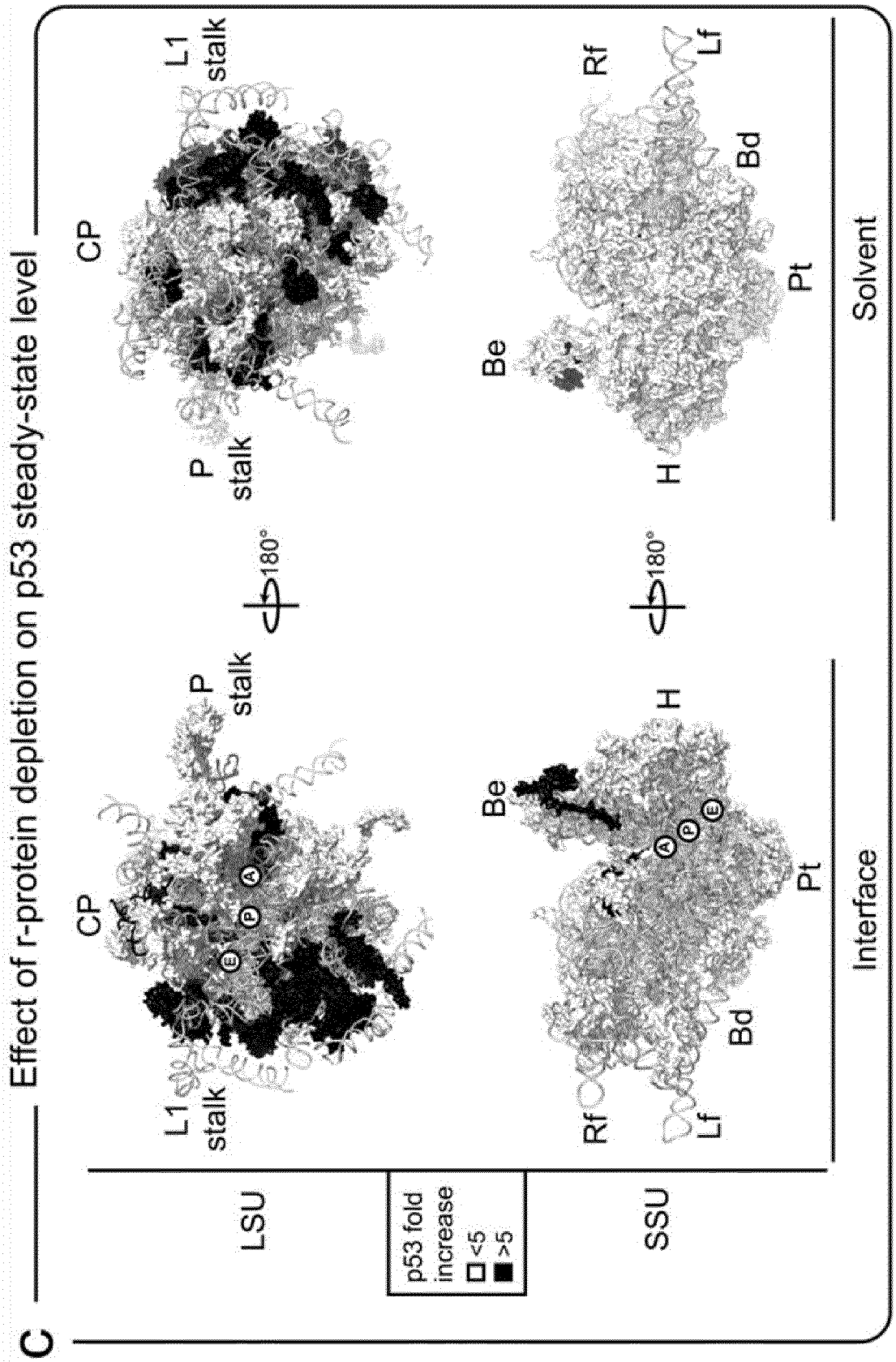

C
Effect of r-protein depletion on p53 steady-state level
LSU
SSU
p53 fold increase
<5
>5
CP
L1 stalk
P stalk
A
P
E
180°
Be
H
Rf
Lf
Bd
Pt
Interface
Solvent

Fig.2

FIG. 3

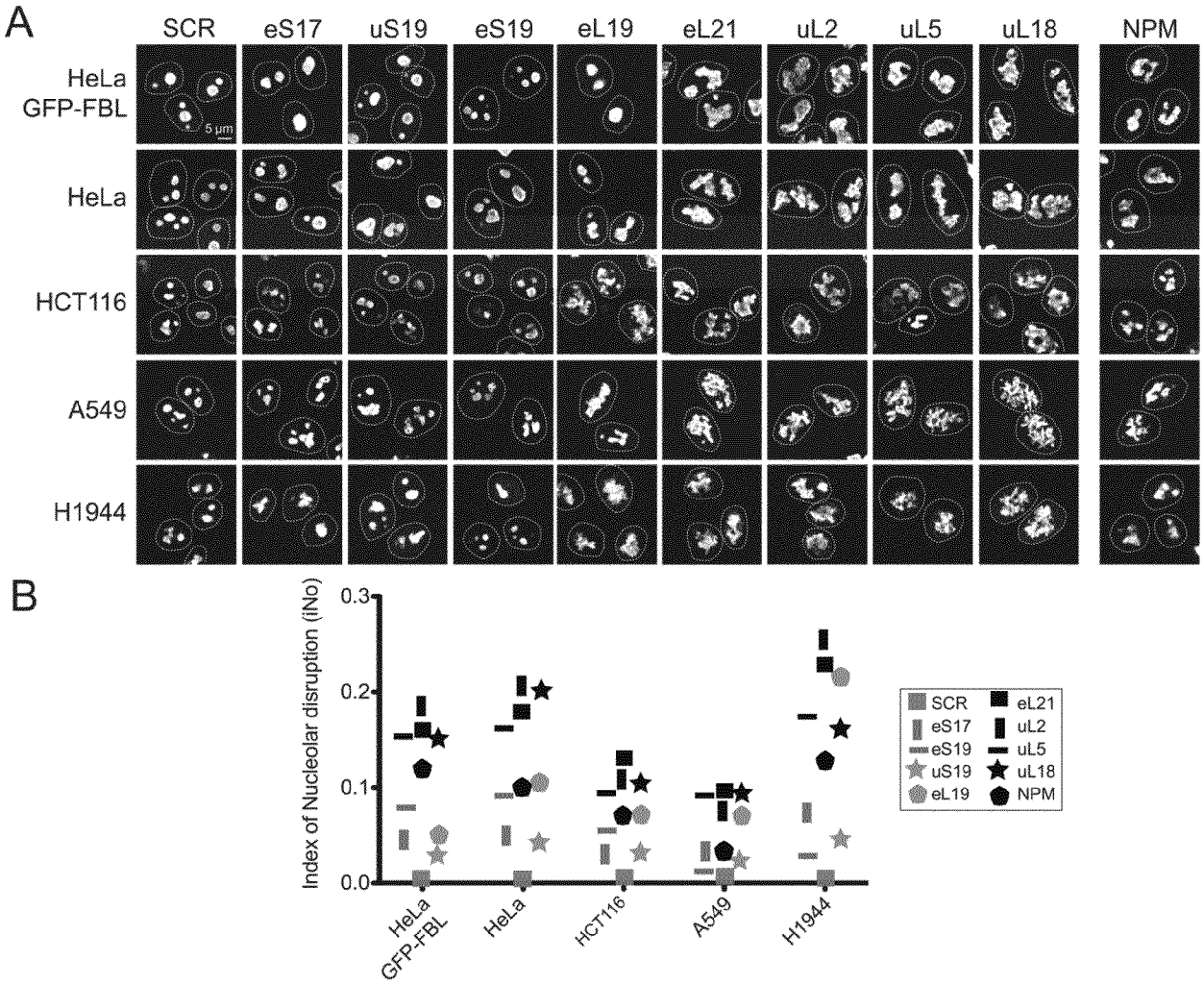

A
SCR
eS17
uS19
eS19
eL19
eL21
uL2
uL5
uL18
NPM
HeLa GFP-FBL
5 µm
HeLa
HCT116
A549
H1944
B
Index of Nucleolar disruption (iNo)
0.3
0.2
0.1
0.0
HeLa GFP-FBL
HeLa
HCT116
A549
H1944
SCR
eS17
eS19
uS19
eL19
eL21
uL2
uL5
uL18
NPM

FIG. 4

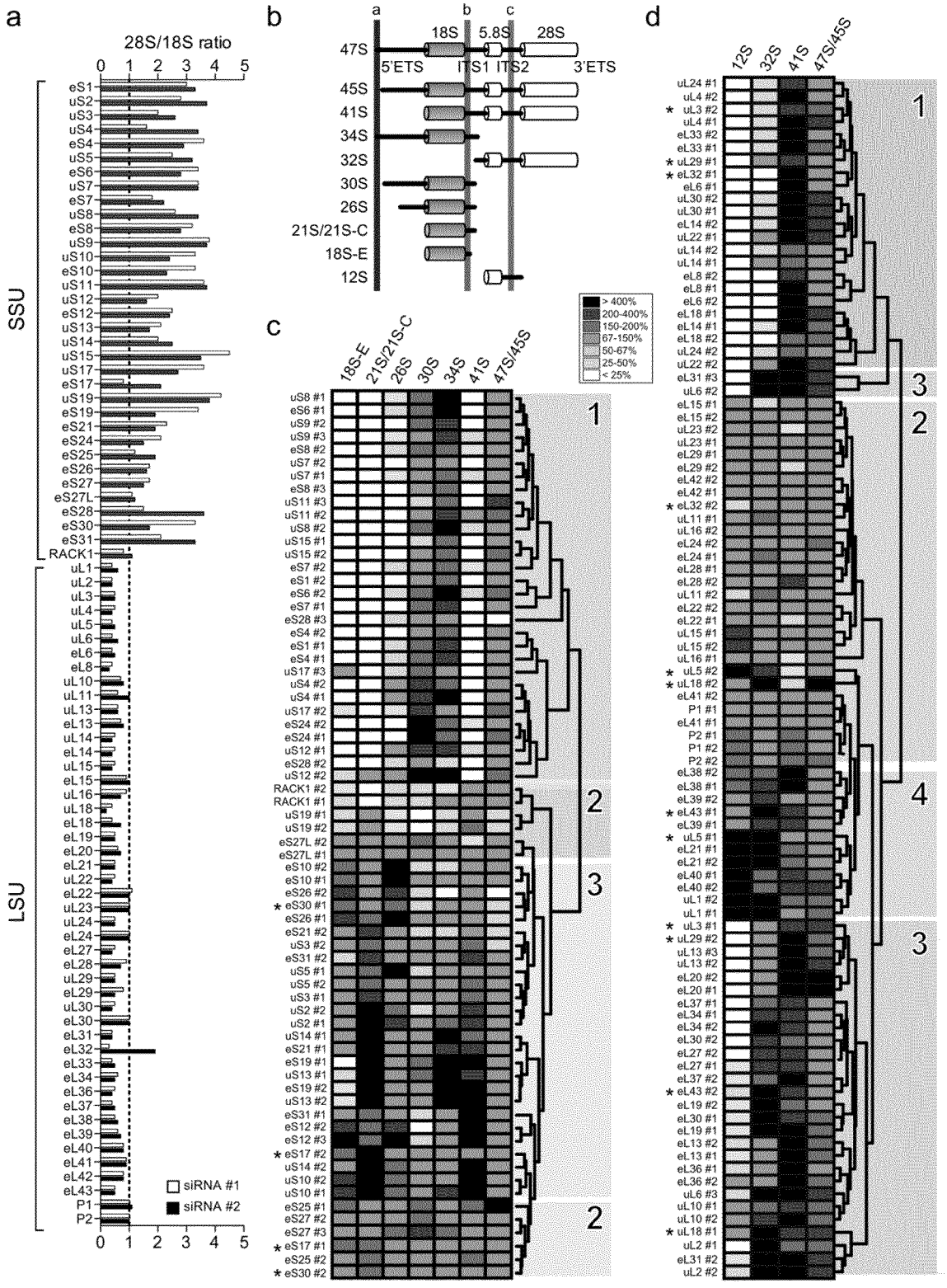

a
28S/18S ratio
0 1 2 3 4 5
SSU
LSU
eS1
uS2
uS3
uS4
eS4
uS5
eS6
uS7
eS7
uS8
eS8
uS9
uS10
eS10
uS11
uS12
eS12
uS13
uS14
uS15
uS17
eS17
uS19
eS19
eS21
eS24
eS25
eS26
eS27
eS27L
eS28
eS30
eS31
RACK1
uL1
uL2
uL3
uL4
uL5
uL6
eL6
eL8
uL10
uL11
uL13
eL13
uL14
eL14
uL15
eL15
uL16
uL18
eL18
eL19
eL20
eL21
uL22
eL22
uL23
uL24
eL24
eL27
eL28
uL29
eL29
uL30
eL30
eL31
eL32
eL33
eL34
eL36
eL37
eL38
eL39
eL40
eL41
eL42
eL43
P1
P2
siRNA #1
siRNA #2
b
a b c
18S 5.8S 28S
47S
5'ETS ITS1 ITS2 3'ETS
45S
41S
34S
32S
30S
26S
21S/21S-C
18S-E
12S
c
18S-E 21S/21S-C 26S 30S 34S 41S 47S/45S
> 400%
200-400%
150-200%
67-150%
50-67%
25-50%
< 25%
d
12S 32S 41S 47S/45S
1
2
3
4

FIG. 5

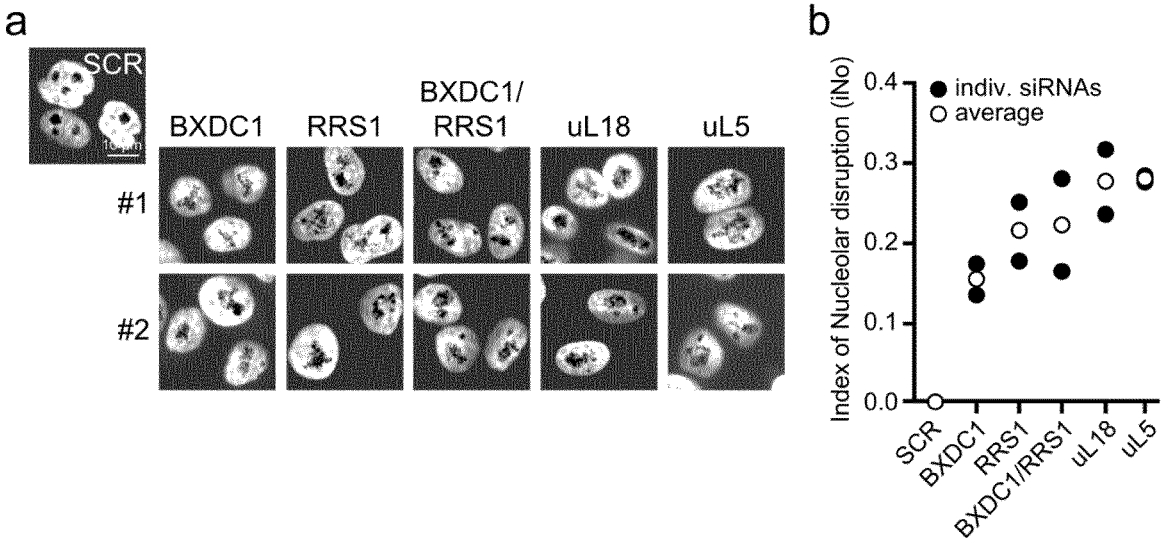

a
SCR
BXDC1
RRS1
BXDC1/
RRS1
uL18
uL5
#1
#2
b
Index of Nucleolar disruption (iNo)
0.4
0.3
0.2
0.1
0.0
indiv. siRNAs
average
SCR
BXDC1
RRS1
BXDC1/RRS1
uL18
uL5

FIG. 6

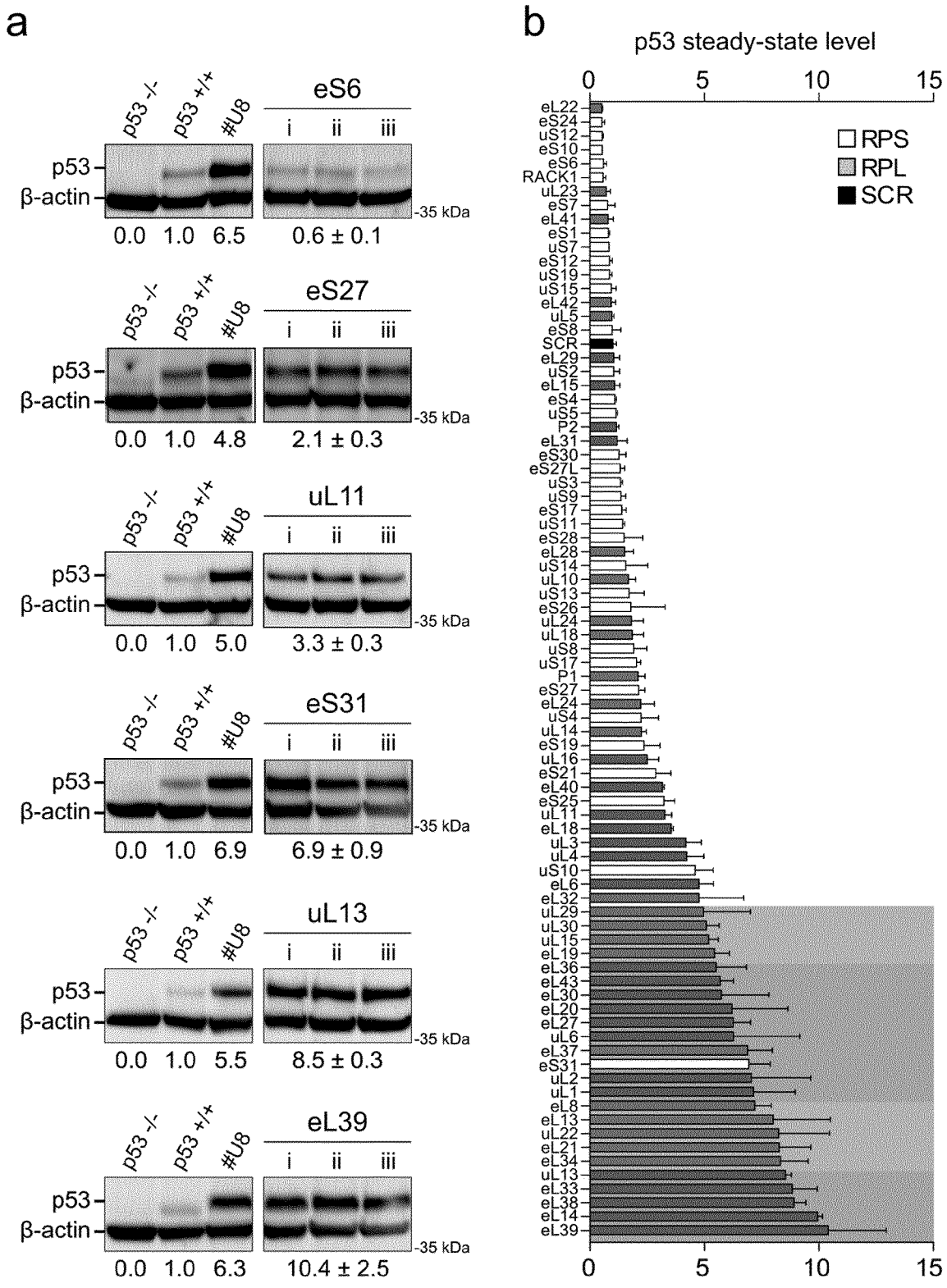

a
p53 -/-
p53 +/+
#U8
eS6
i ii iii
p53
β-actin
-35 kDa
0.0 1.0 6.5
0.6 ± 0.1
eS27
0.0 1.0 4.8
2.1 ± 0.3
uL11
0.0 1.0 5.0
3.3 ± 0.3
eS31
0.0 1.0 6.9
6.9 ± 0.9
uL13
0.0 1.0 5.5
8.5 ± 0.3
eL39
0.0 1.0 6.3
10.4 ± 2.5
b
p53 steady-state level
0 5 10 15
RPS
RPL
SCR
eL22
eS24
uS12
eS10
eS6
RACK1
uL23
eS7
eL41
eS1
uS7
eS12
uS19
uS15
eL42
uL5
eS8
SCR
eL29
uS2
eL15
eS4
uS5
P2
eL31
eS30
eS27L
uS3
uS9
eS17
uS11
eS28
eL28
uS14
uL10
uS13
eS26
uL24
uL18
uS8
uS17
P1
eS27
eL24
uS4
uL14
eS19
uL16
eS21
eL40
eS25
uL11
eL18
uL3
uL4
uS10
eL6
eL32
uL29
uL30
uL15
eL19
eL36
eL43
eL30
eL20
eL27
uL6
eL37
eS31
uL2
uL1
eL8
eL13
uL22
eL21
eL34
uL13
eL33
eL38
eL14
eL39

FIG. S1

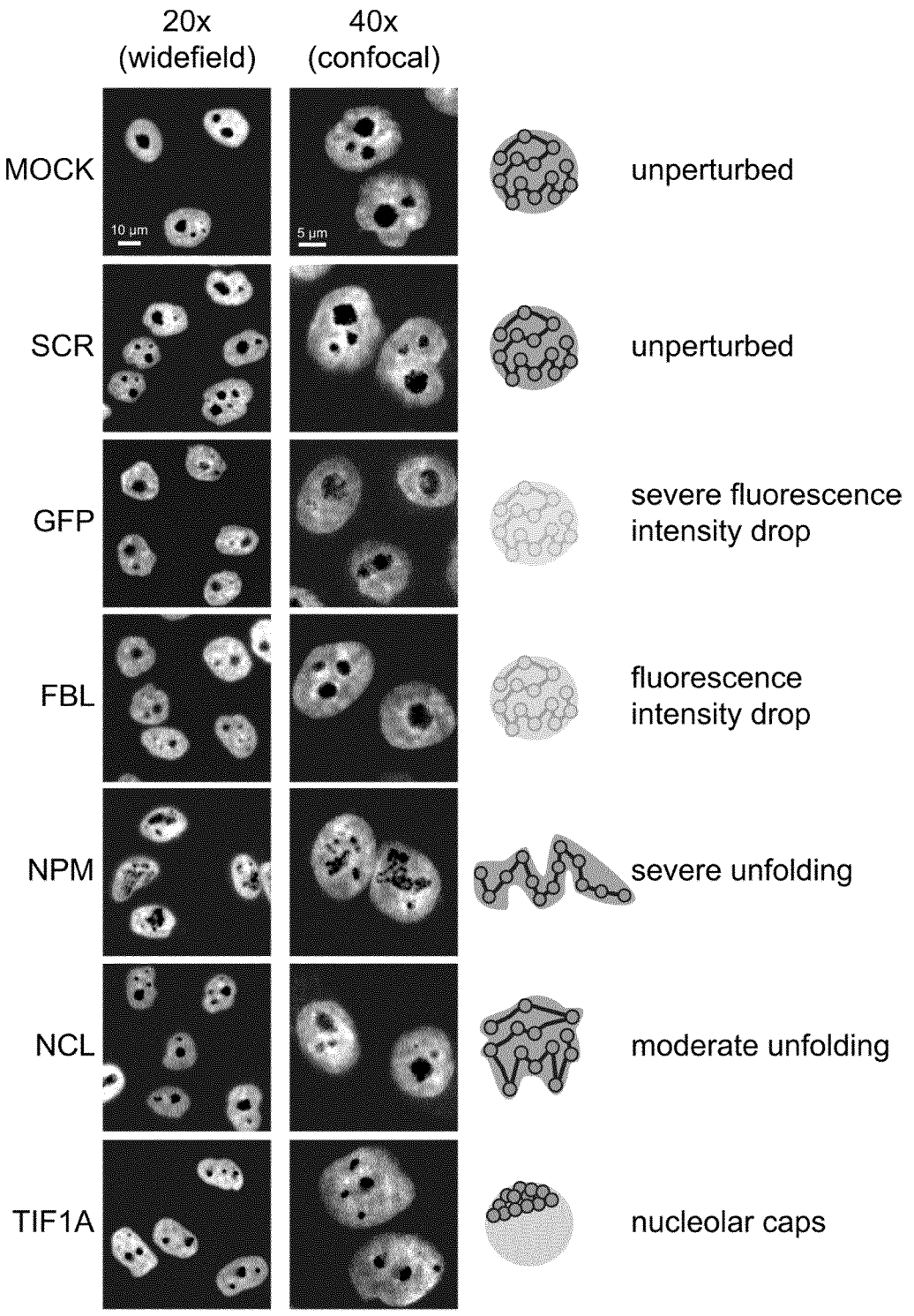

20x
(widefield)
40x
(confocal)
MOCK
10 µm
5 µm
unperturbed
SCR
unperturbed
GFP
severe fluorescence
intensity drop
FBL
fluorescence
intensity drop
NPM
severe unfolding
NCL
moderate unfolding
TIF1A
nucleolar caps

FIG. S2

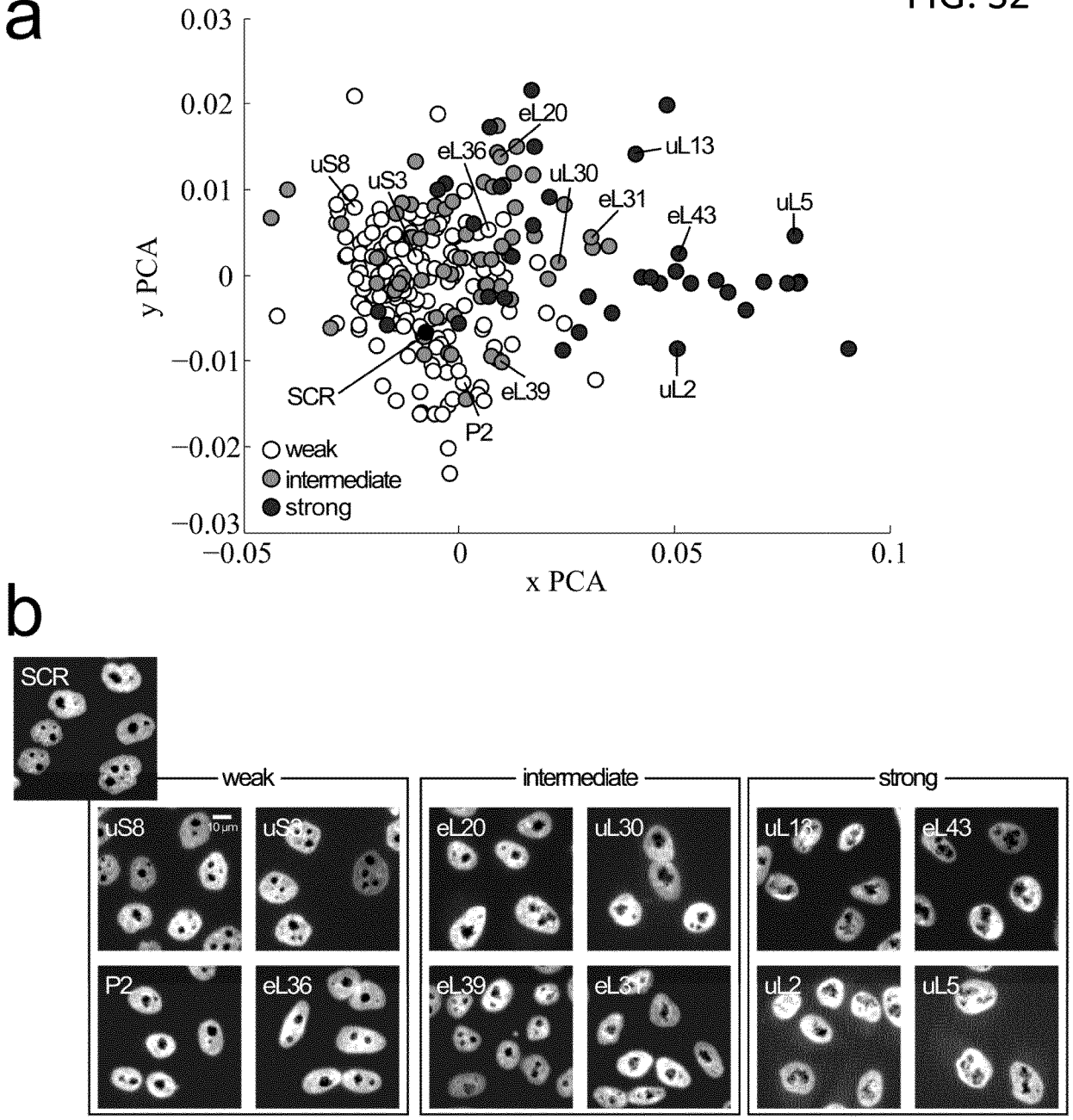

a
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
y PCA
−0.05
0
0.05
0.1
x PCA
eL20
uL13
eL36
uS8
uS3
uL30
eL31
eL43
uL5
SCR
eL39
uL2
P2
weak
intermediate
strong
b
SCR
weak
intermediate
strong
uS8
10 µm
uS3
eL20
uL30
uL13
eL43
P2
eL36
eL39
eL31
uL2
uL5

FIG. S3

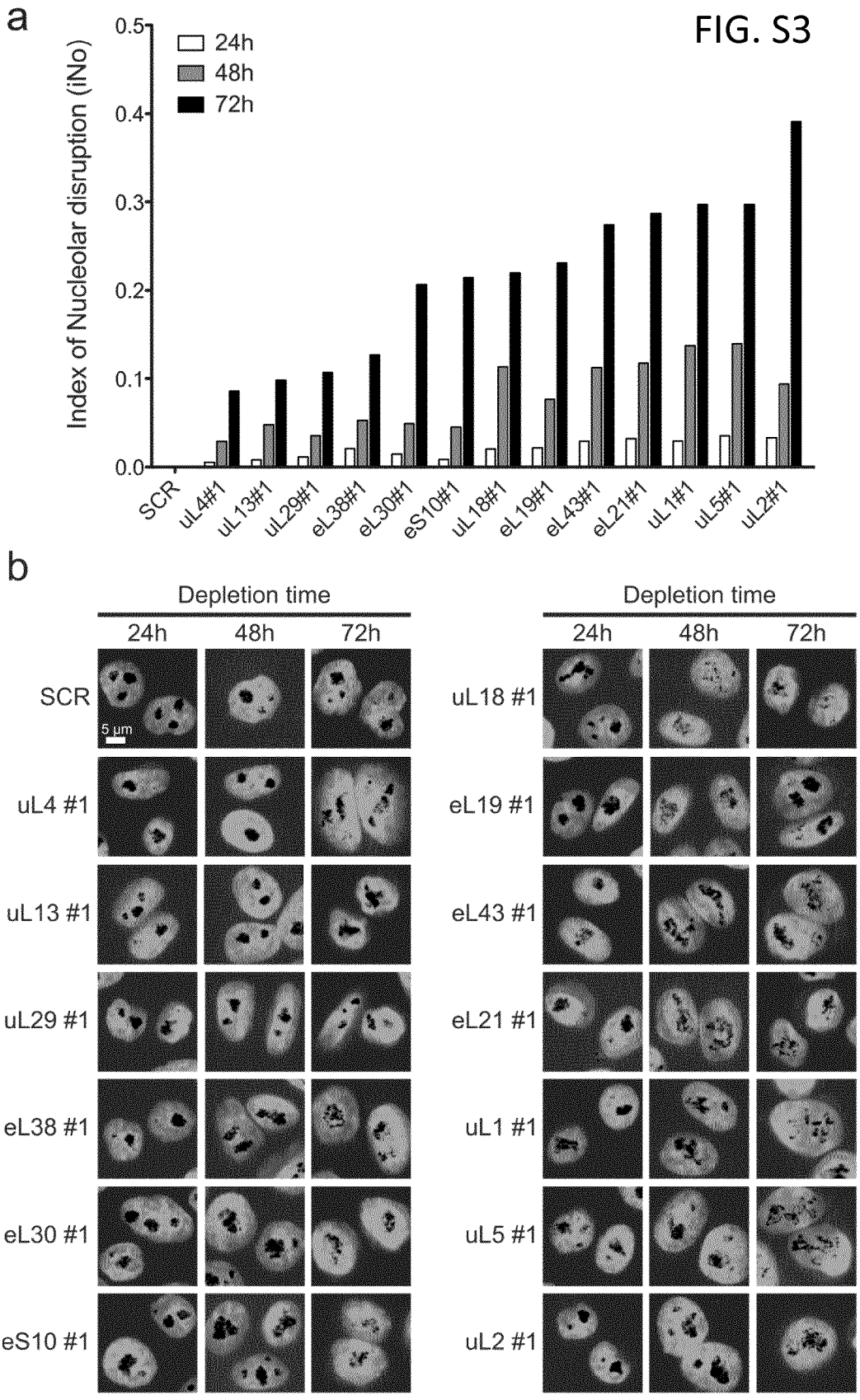

a
Index of Nucleolar disruption (iNo)
0.5
0.4
0.3
0.2
0.1
0.0
24h
48h
72h
SCR
uL4#1
uL13#1
uL29#1
eL38#1
eL30#1
eS10#1
uL18#1
eL19#1
eL43#1
eL21#1
uL1#1
uL5#1
uL2#1
b
Depletion time
24h
48h
72h
SCR
5 µm
uL4 #1
uL13 #1
uL29 #1
eL38 #1
eL30 #1
eS10 #1
Depletion time
24h
48h
72h
uL18 #1
eL19 #1
eL43 #1
eL21 #1
uL1 #1
uL5 #1
uL2 #1

FIG. S4

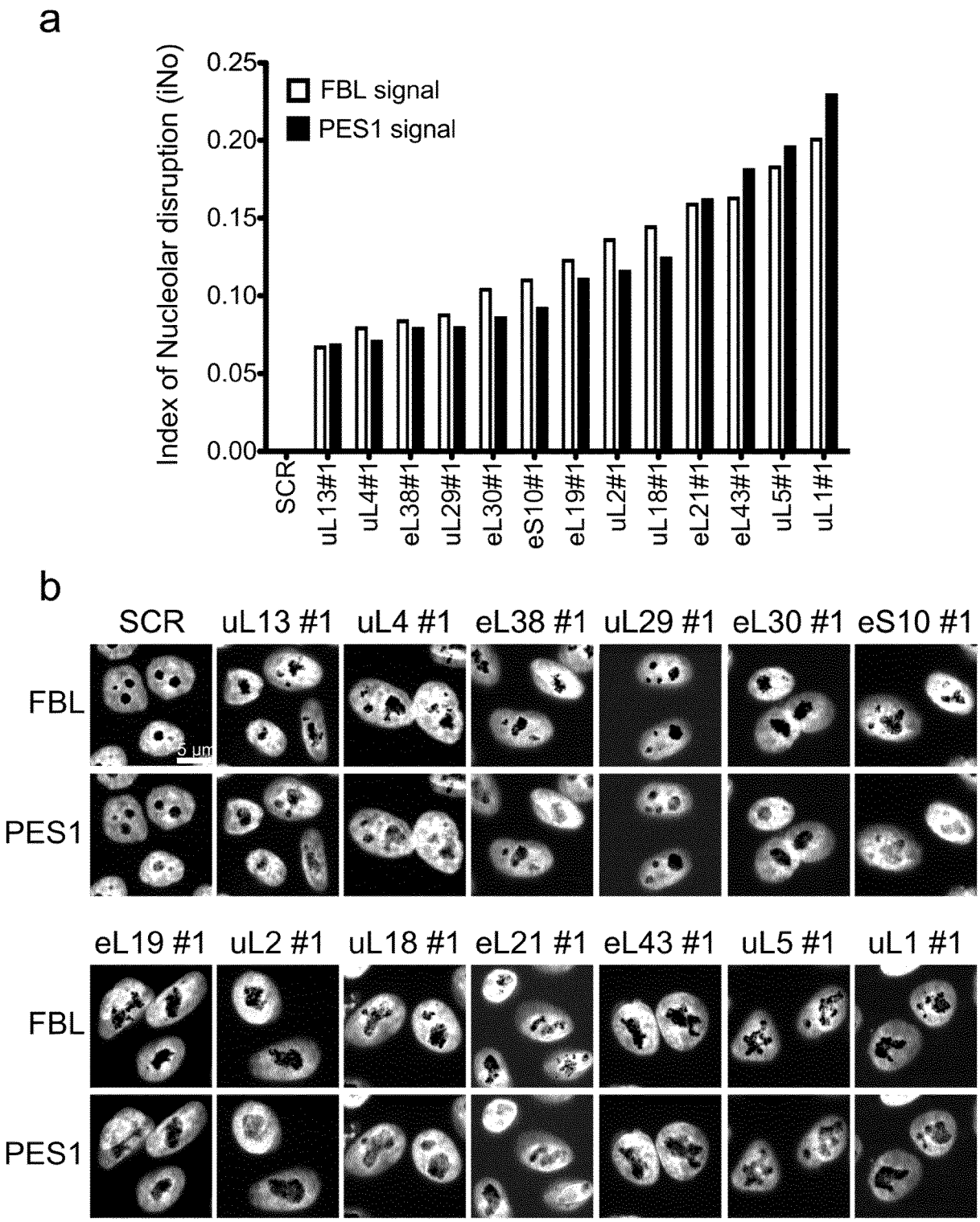
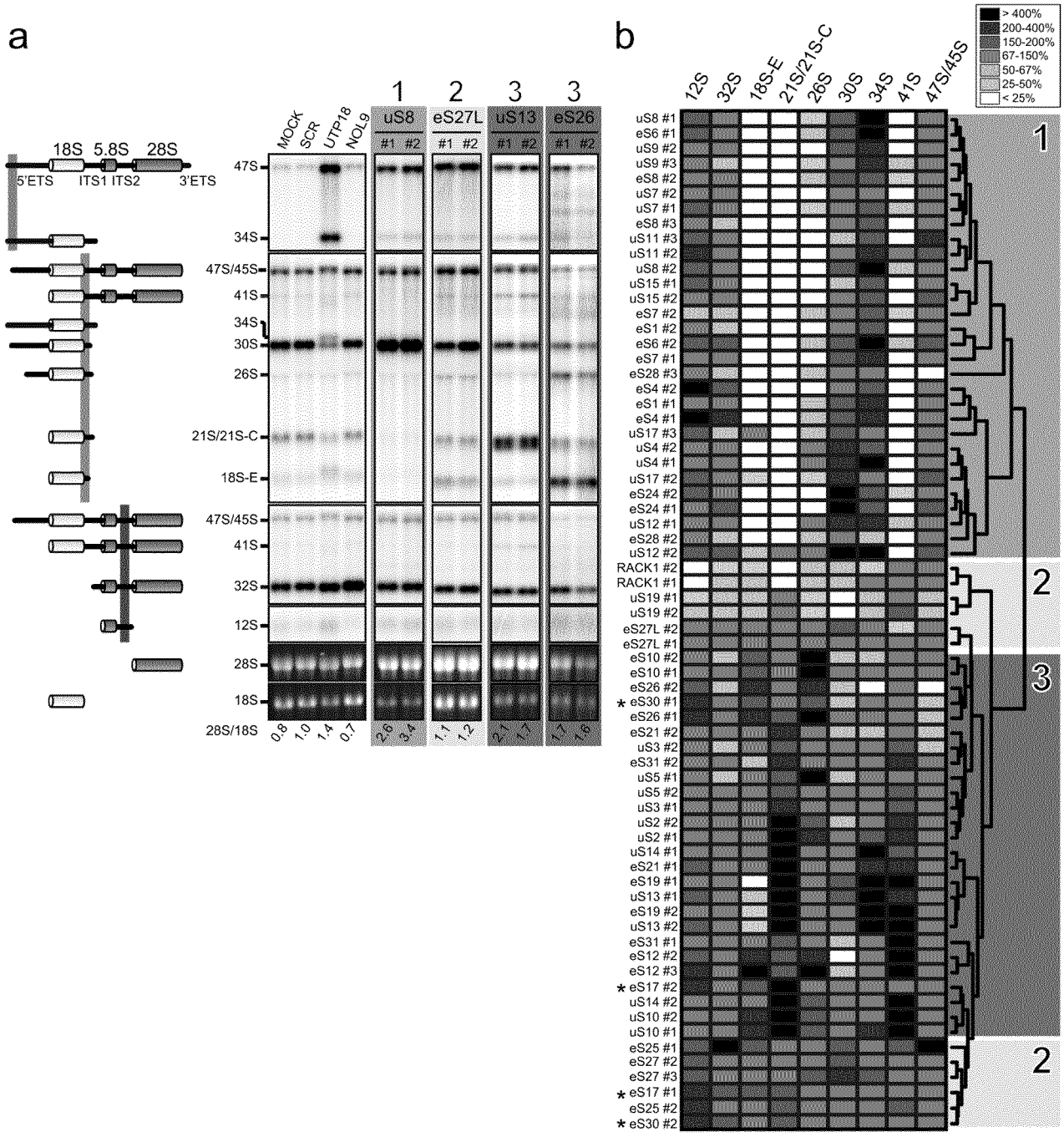

a
Index of Nucleolar disruption (iNo)
0.25
0.20
0.15
0.10
0.05
0.00
FBL signal
PES1 signal
SCR
uL13#1
uL4#1
eL38#1
uL29#1
eL30#1
eS10#1
eL19#1
uL2#1
uL18#1
eL21#1
eL43#1
uL5#1
uL1#1
b
SCR
uL13 #1
uL4 #1
eL38 #1
uL29 #1
eL30 #1
eS10 #1
FBL
PES1
5 µm
eL19 #1
uL2 #1
uL18 #1
eL21 #1
eL43 #1
uL5 #1
uL1 #1
FBL
PES1

FIG. S5


a
18S 5.8S 28S
5'ETS ITS1 ITS2 3'ETS
MOCK SCR UTP18 NOL9
1 2 3 3
uS8 eS27L uS13 eS26
#1 #2 #1 #2 #1 #2 #1 #2
47S
34S
47S/45S
41S
34S
30S
26S
21S/21S-C
18S-E
47S/45S
41S
32S
12S
28S
18S
28S/18S
0.8 1.0 1.4 0.7
2.6 3.4 1.1 1.2 2.1 1.7 1.7 1.6
b
12S 32S 18S-E 21S/21S-C 26S 30S 34S 41S 47S/45S
> 400%
200-400%
150-200%
67-150%
50-67%
25-50%
< 25%
uS8 #1
eS6 #1
uS9 #2
uS9 #3
eS8 #2
uS7 #2
uS7 #1
eS8 #3
uS11 #3
uS11 #2
uS8 #2
uS15 #1
uS15 #2
eS7 #2
eS1 #2
eS6 #2
eS7 #1
eS28 #3
eS4 #2
eS1 #1
eS4 #1
uS17 #3
uS4 #2
uS4 #1
uS17 #2
eS24 #2
eS24 #1
uS12 #1
eS28 #2
uS12 #2
RACK1 #2
RACK1 #1
uS19 #1
uS19 #2
eS27L #2
eS27L #1
eS10 #2
eS10 #1
eS26 #2
* eS30 #1
eS26 #1
eS21 #2
uS3 #2
eS31 #2
uS5 #1
uS5 #2
uS3 #1
uS2 #2
uS2 #1
uS14 #1
eS21 #1
eS19 #1
uS13 #1
eS19 #2
uS13 #2
eS31 #1
eS12 #2
eS12 #3
* eS17 #2
uS14 #2
uS10 #2
uS10 #1
eS25 #1
eS27 #2
eS27 #3
* eS17 #1
eS25 #2
* eS30 #2
1
2
3
2

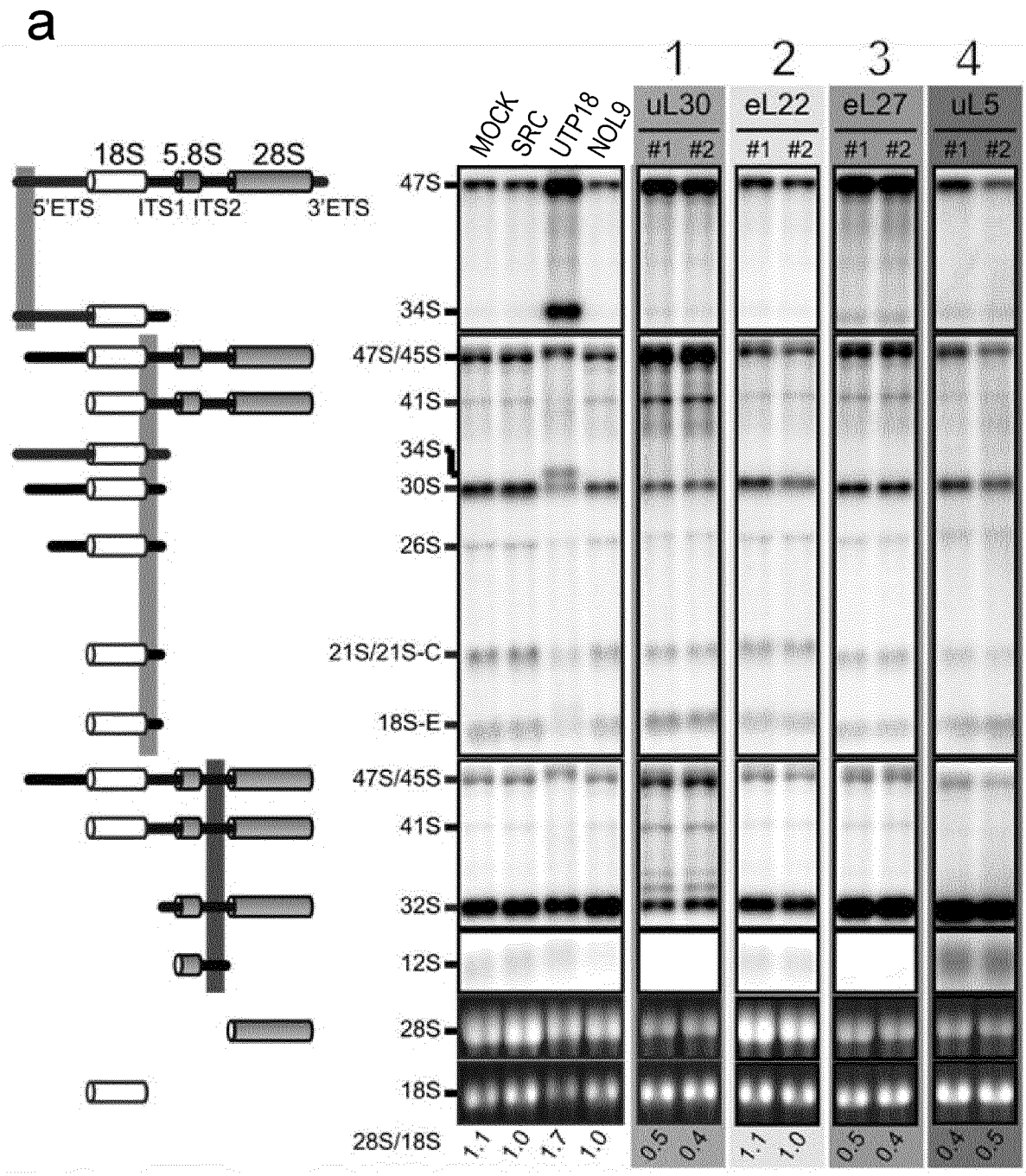

a
18S 5.8S 28S
5'ETS ITS1 ITS2 3'ETS
MOCK
SRC
UTP18
NOL9
1
2
3
4
uL30
eL22
eL27
uL5
#1 #2
47S
34S
47S/45S
41S
34S
30S
26S
21S/21S-C
18S-E
47S/45S
41S
32S
12S
28S
18S
28S/18S
1.1 1.0 1.7 1.0
0.5 0.4
1.1 1.0
0.5 0.4
0.4 0.5

Fig. S6

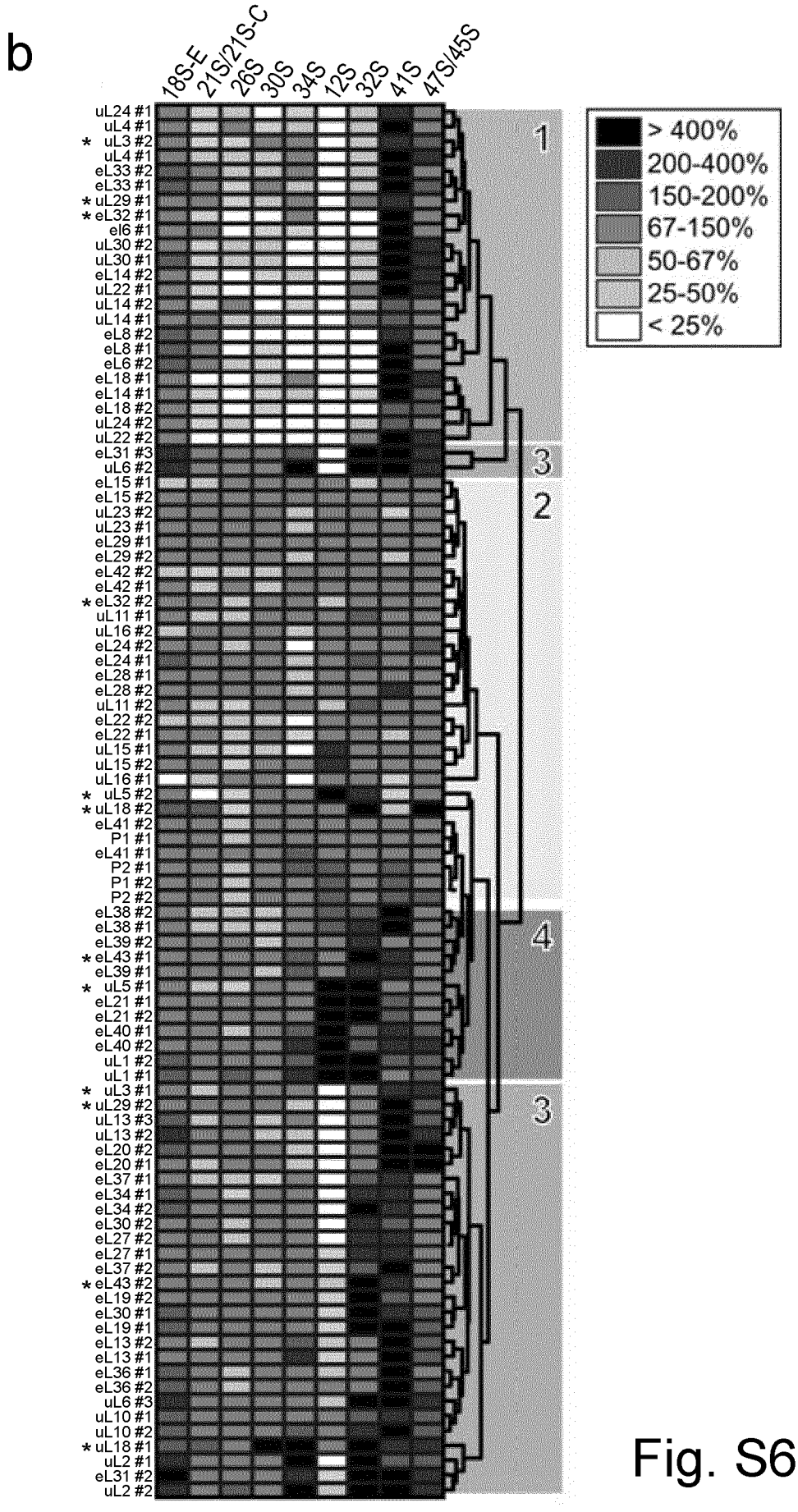

b
18S-E
21S/21S-C
26S
30S
34S
12S
32S
41S
47S/45S
uL24 #1
uL4 #1
* uL3 #2
uL4 #1
eL33 #2
eL33 #1
* uL29 #1
* eL32 #1
eI6 #1
uL30 #2
uL30 #1
eL14 #2
uL22 #1
uL14 #2
uL14 #1
eL8 #2
eL8 #1
eL6 #2
eL18 #1
eL14 #1
eL18 #2
uL24 #2
uL22 #2
eL31 #3
uL6 #2
eL15 #1
eL15 #2
uL23 #2
uL23 #1
eL29 #1
eL29 #2
eL42 #2
eL42 #1
* eL32 #2
uL11 #1
uL16 #2
eL24 #2
eL24 #1
eL28 #1
eL28 #2
uL11 #2
eL22 #2
eL22 #1
uL15 #1
uL15 #2
uL16 #1
* uL5 #2
* uL18 #2
eL41 #2
P1 #1
eL41 #1
P2 #1
P1 #2
P2 #2
eL38 #2
eL38 #1
eL39 #2
* eL43 #1
eL39 #1
* uL5 #1
eL21 #1
eL21 #2
eL40 #1
eL40 #2
uL1 #2
uL1 #1
* uL3 #1
* uL29 #2
uL13 #3
uL13 #2
eL20 #2
eL20 #1
eL37 #1
eL34 #1
eL34 #2
eL30 #2
eL27 #2
eL27 #1
eL37 #2
* eL43 #2
eL19 #2
eL30 #1
eL19 #1
eL13 #2
eL13 #1
eL36 #1
eL36 #2
uL6 #3
uL10 #1
uL10 #2
* uL18 #1
uL2 #1
eL31 #2
uL2 #2
1
3
2
4
3
> 400%
200-400%
150-200%
67-150%
50-67%
25-50%
< 25%

Fig. S6

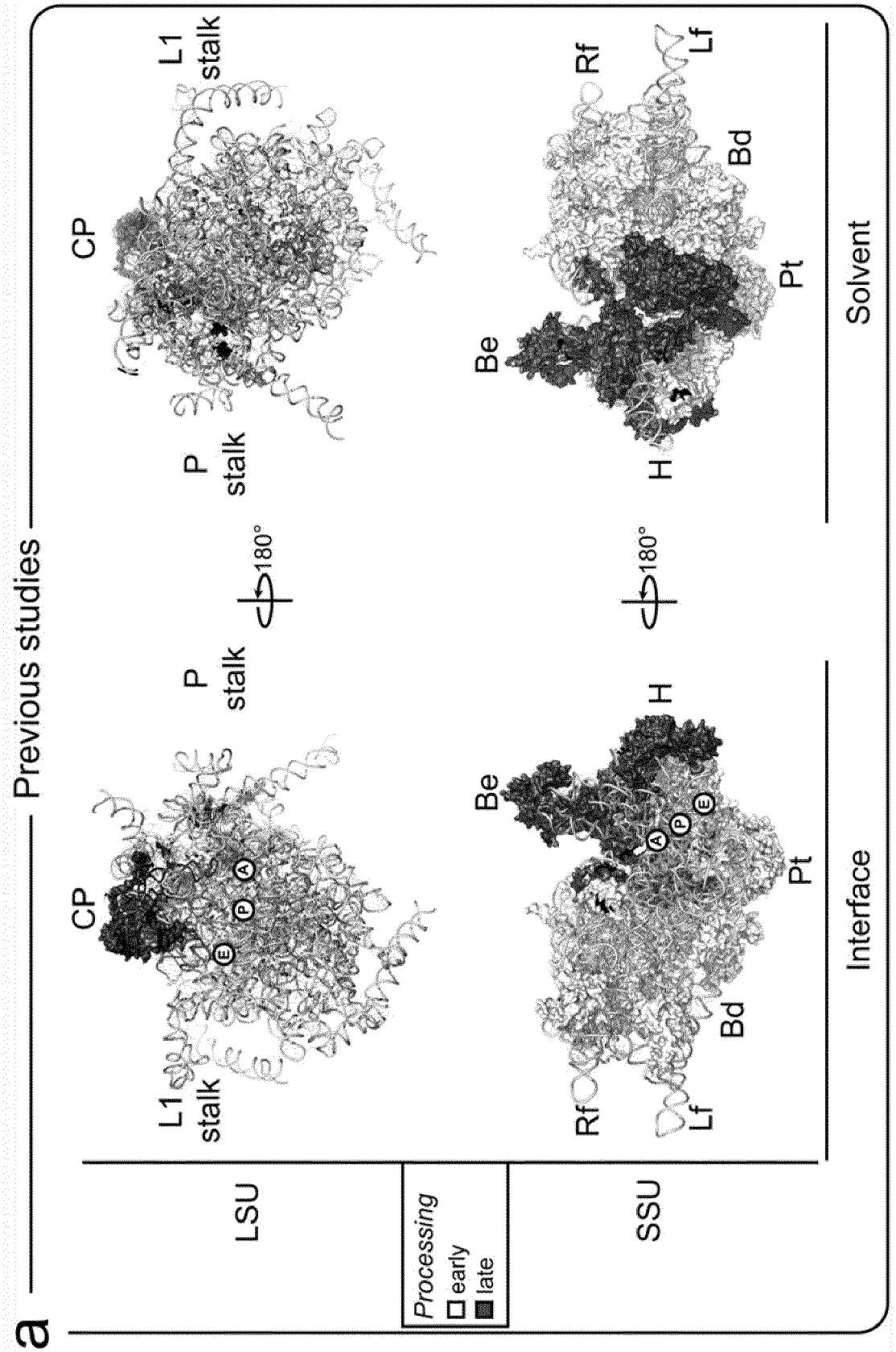

a
Previous studies
LSU
SSU
Processing
early
late
CP
L1 stalk
P stalk
180°
Be
Rf
Lf
Bd
Pt
H
Interface
Solvent
E
P
A

Fig. S7

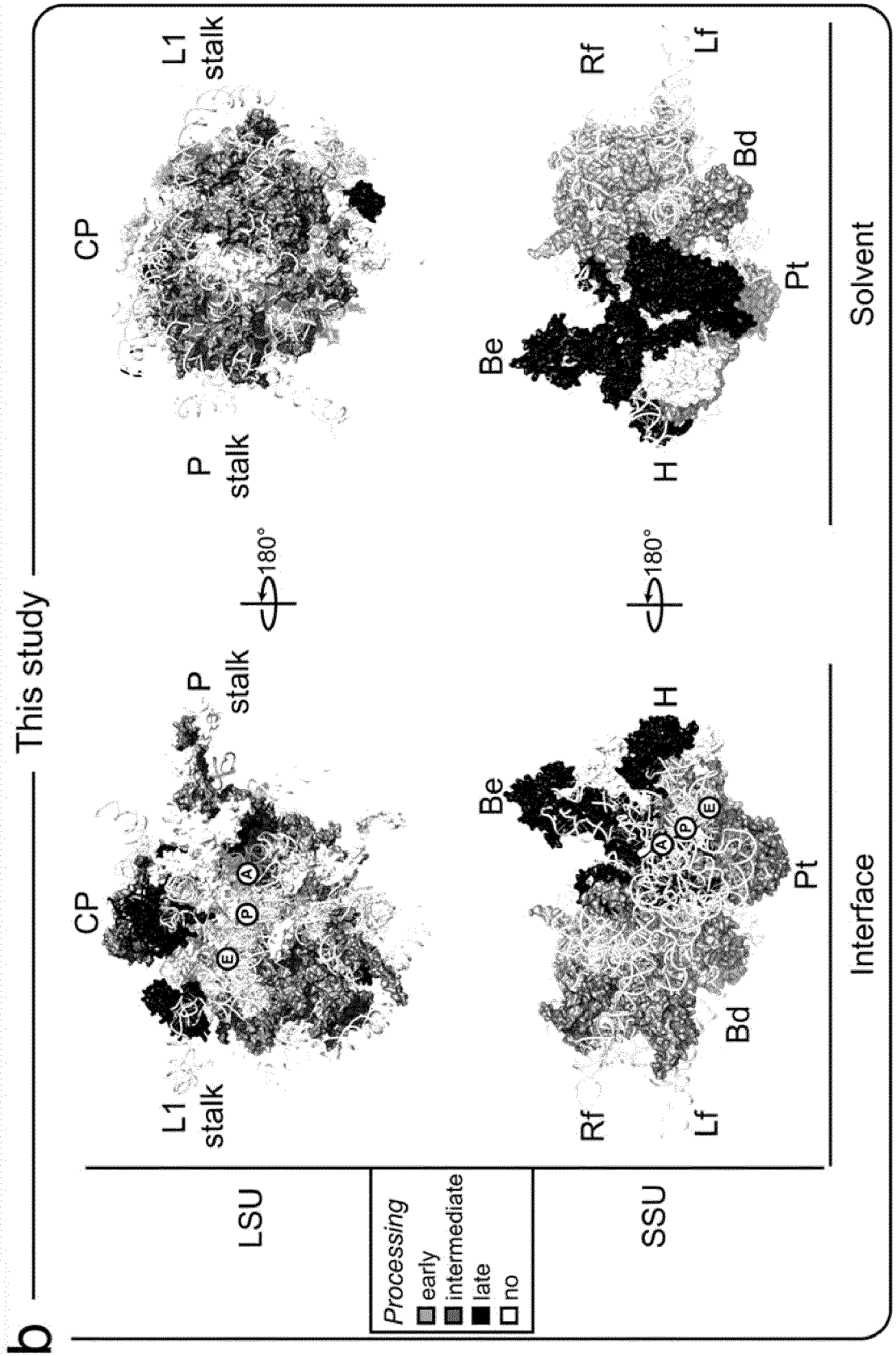

b
This study
LSU
SSU
Processing
early
intermediate
late
no
CP
L1 stalk
P stalk
E
P
A
180°
Be
Rf
Lf
Bd
Pt
H
Interface
Solvent

Fig. S7

FIG. S8

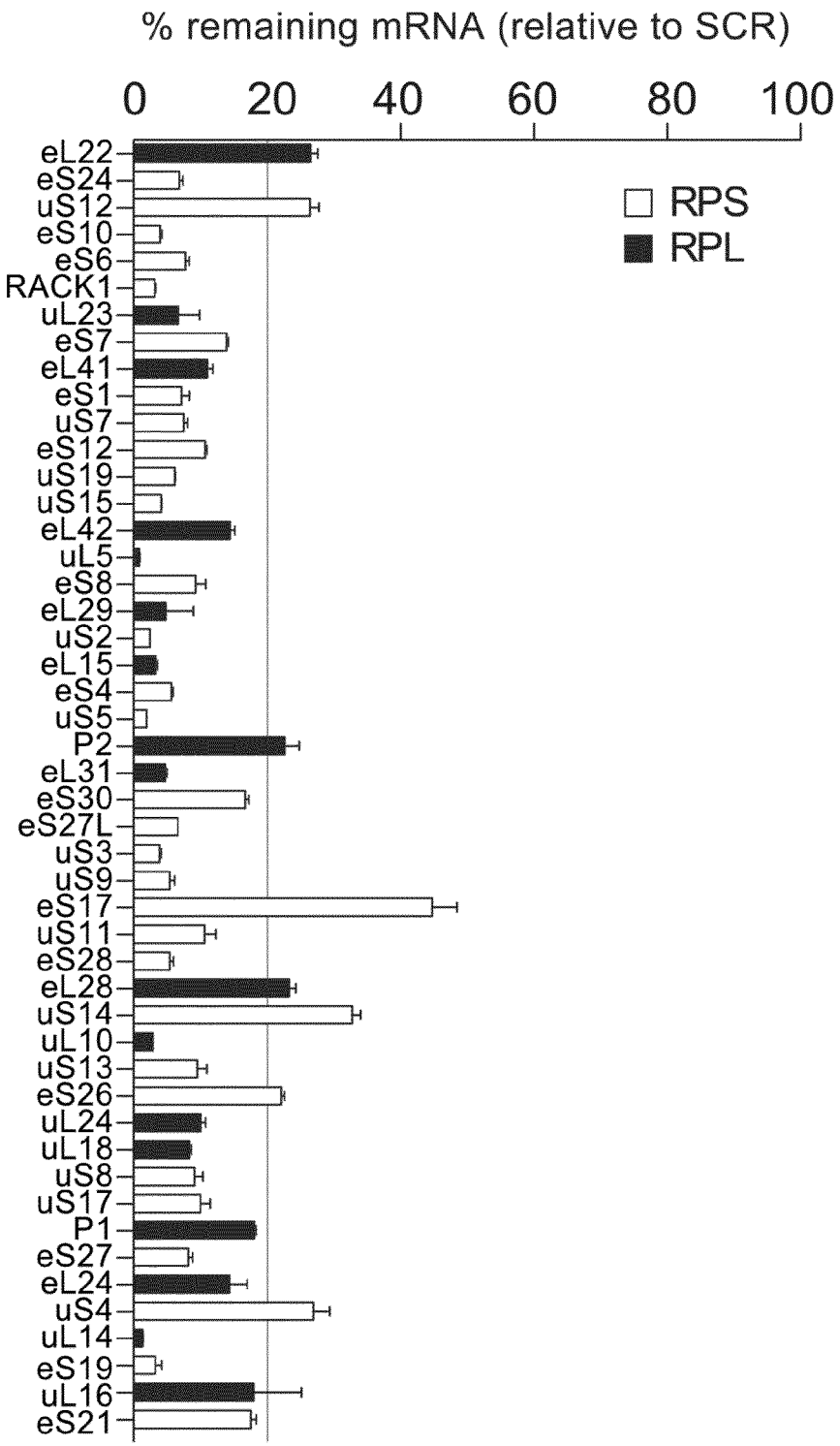
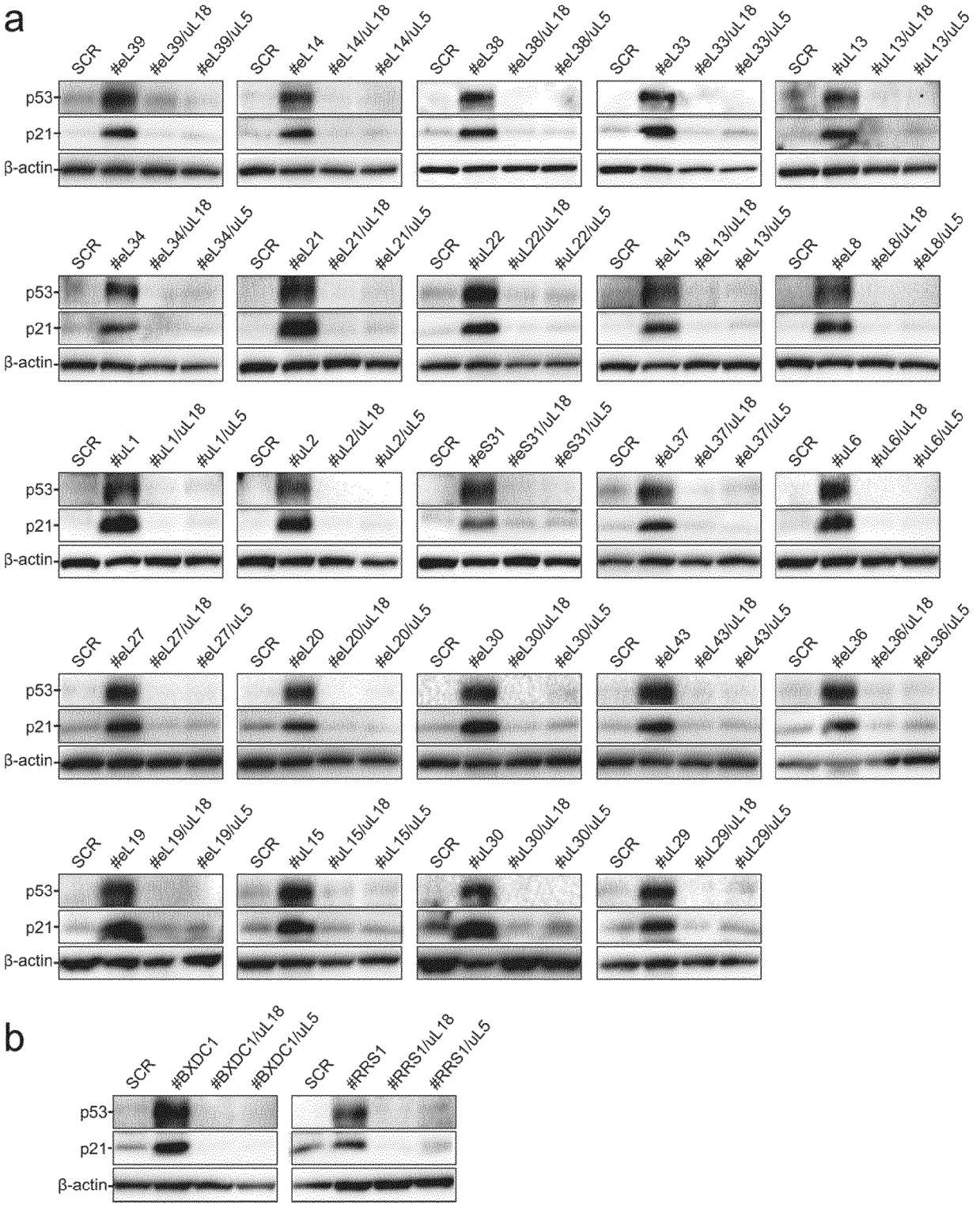

% remaining mRNA (relative to SCR)
0
20
40
60
80
100
RPS
RPL
eL22
eS24
uS12
eS10
eS6
RACK1
uL23
eS7
eL41
eS1
uS7
eS12
uS19
uS15
eL42
uL5
eS8
eL29
uS2
eL15
eS4
uS5
P2
eL31
eS30
eS27L
uS3
uS9
eS17
uS11
eS28
eL28
uS14
uL10
uS13
eS26
uL24
uL18
uS8
uS17
P1
eS27
eL24
uS4
uL14
eS19
uL16
eS21

FIG. S9


a
SCR
#eL39
#eL39/uL18
#eL39/uL5
#eL14
#eL14/uL18
#eL14/uL5
#eL38
#eL38/uL18
#eL38/uL5
#eL33
#eL33/uL18
#eL33/uL5
#uL13
#uL13/uL18
#uL13/uL5
#eL34
#eL34/uL18
#eL34/uL5
#eL21
#eL21/uL18
#eL21/uL5
#uL22
#uL22/uL18
#uL22/uL5
#eL13
#eL13/uL18
#eL13/uL5
#eL8
#eL8/uL18
#eL8/uL5
#uL1
#uL1/uL18
#uL1/uL5
#uL2
#uL2/uL18
#uL2/uL5
#eS31
#eS31/uL18
#eS31/uL5
#eL37
#eL37/uL18
#eL37/uL5
#uL6
#uL6/uL18
#uL6/uL5
#eL27
#eL27/uL18
#eL27/uL5
#eL20
#eL20/uL18
#eL20/uL5
#eL30
#eL30/uL18
#eL30/uL5
#eL43
#eL43/uL18
#eL43/uL5
#eL36
#eL36/uL18
#eL36/uL5
#eL19
#eL19/uL18
#eL19/uL5
#uL15
#uL15/uL18
#uL15/uL5
#uL30
#uL30/uL18
#uL30/uL5
#uL29
#uL29/uL18
#uL29/uL5
p53
p21
β-actin
b
#BXDC1
#BXDC1/uL18
#BXDC1/uL5
#RRS1
#RRS1/uL18
#RRS1/uL5

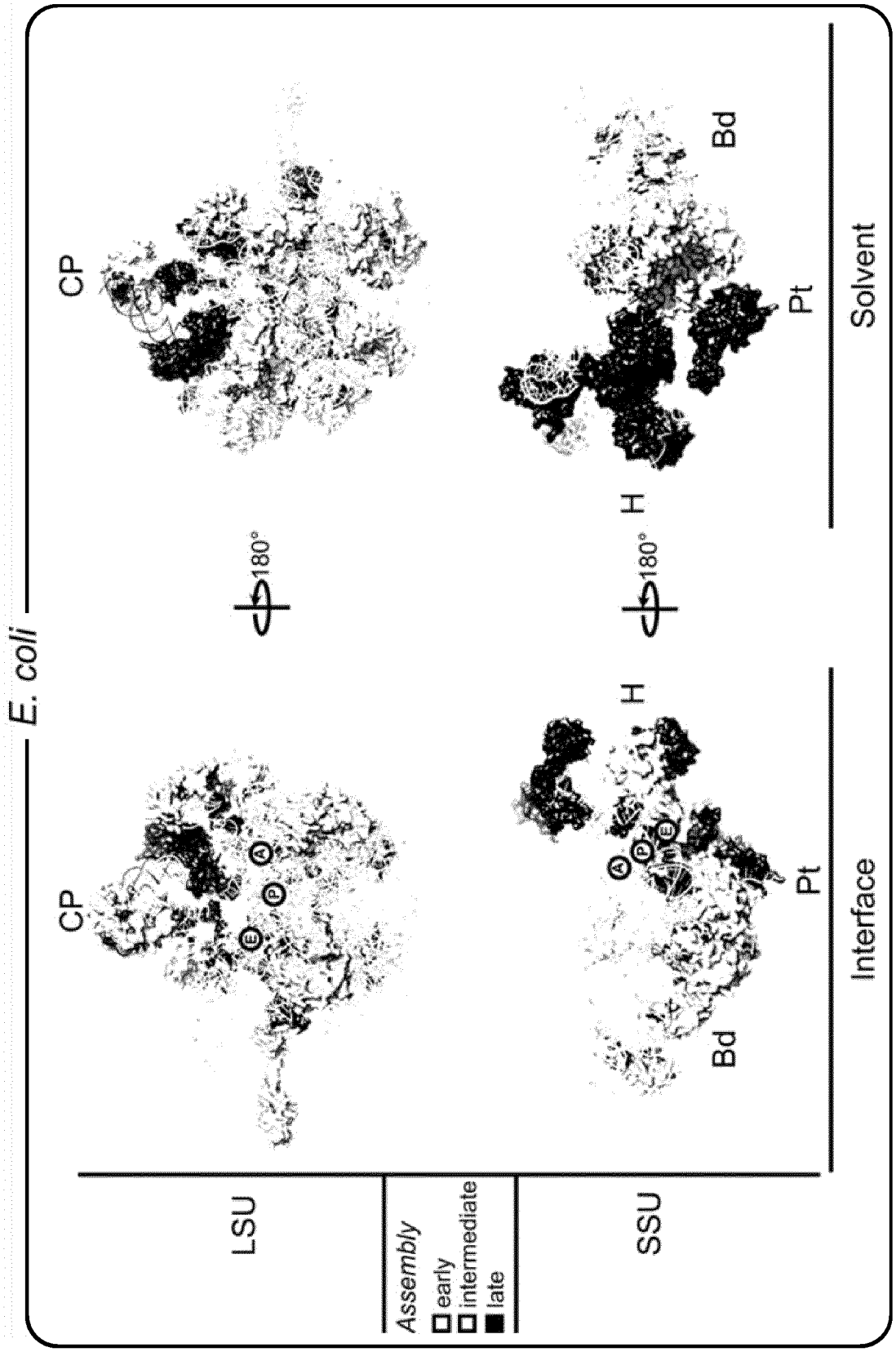

E. coli
LSU
SSU
Assembly
early
intermediate
late
CP
CP
180°
A
P
E
H
Bd
Pt
180°
H
Bd
Pt
Interface
Solvent

Fig. S10

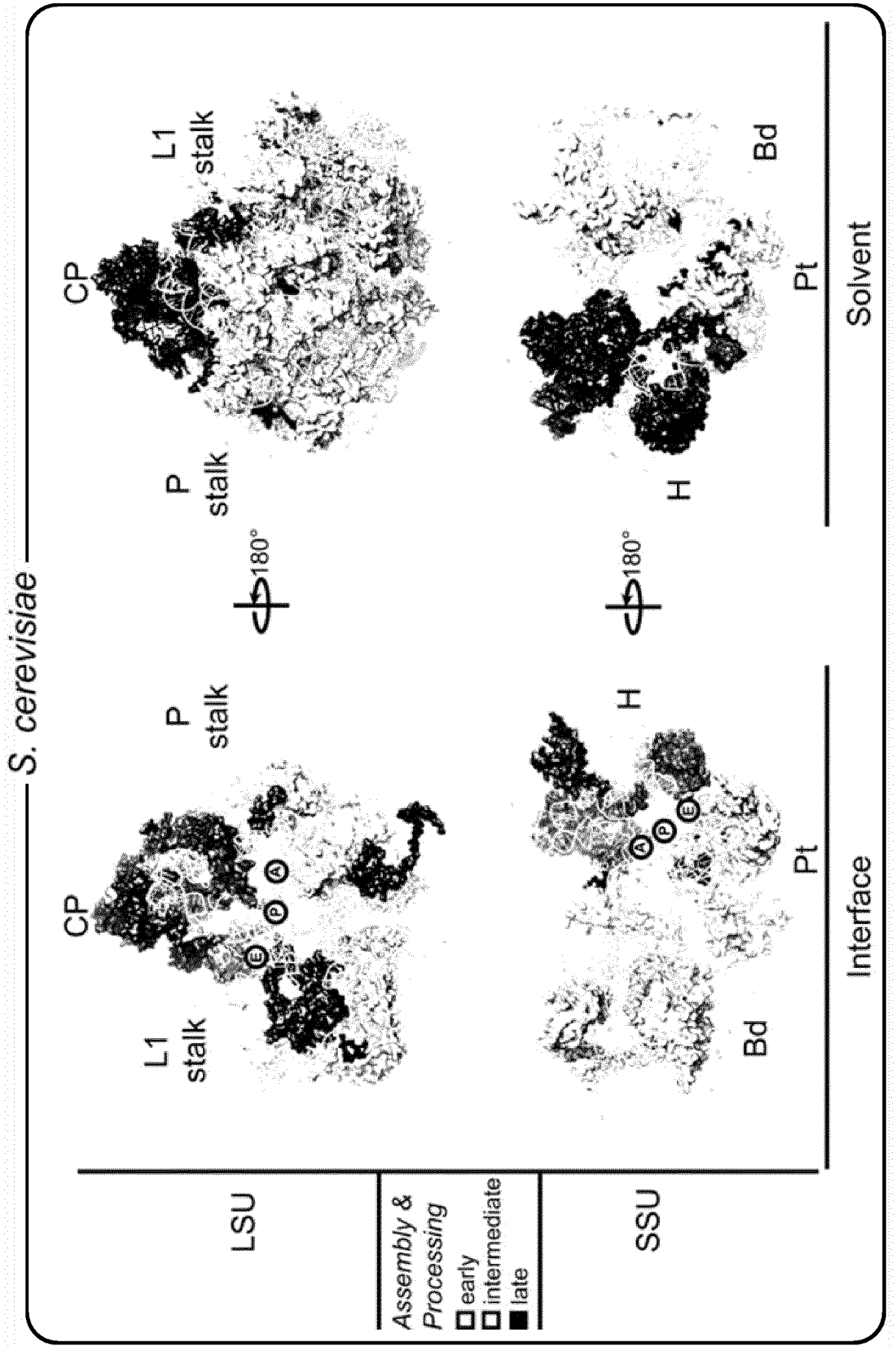


Fig. S10 continued

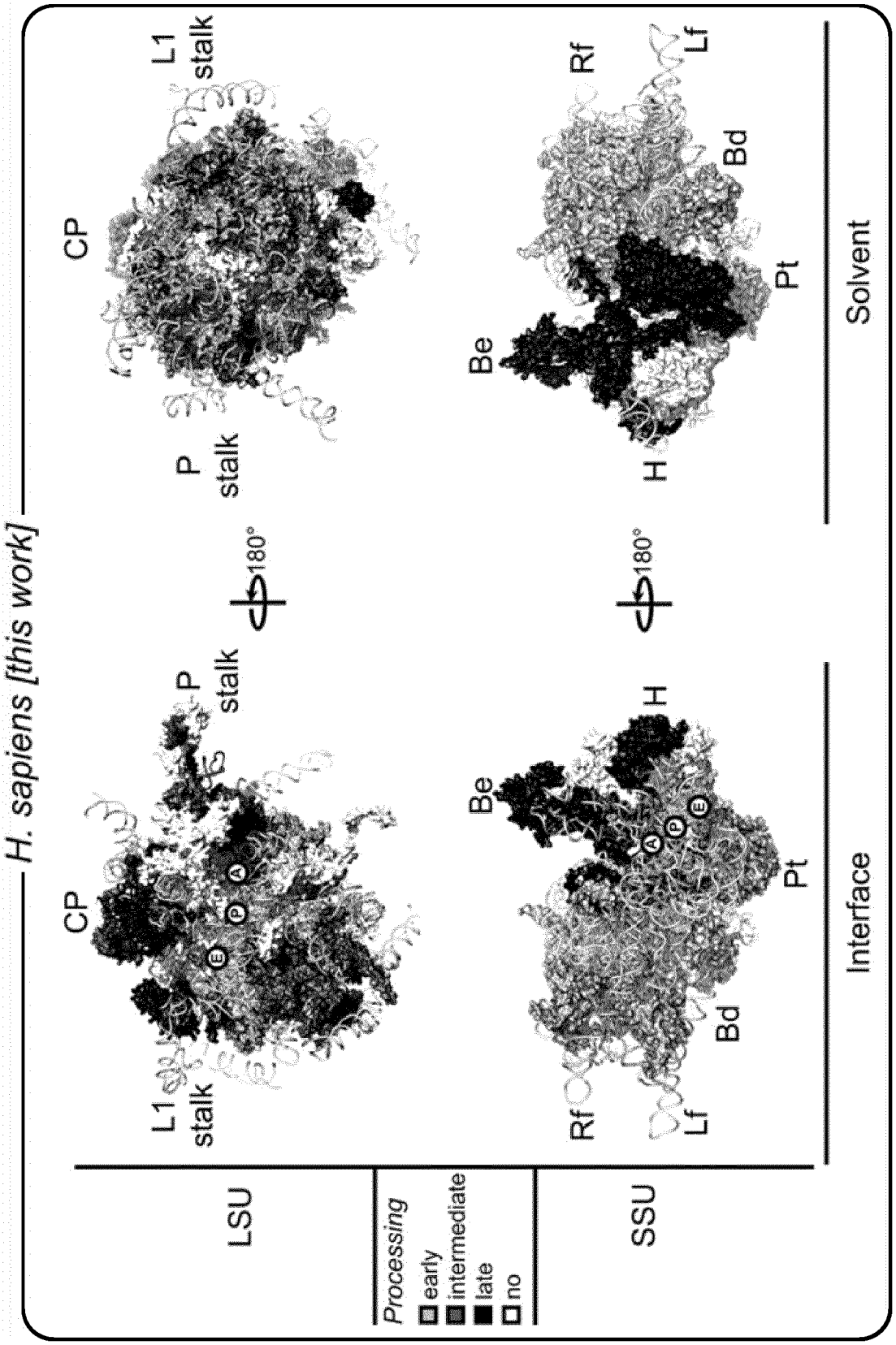
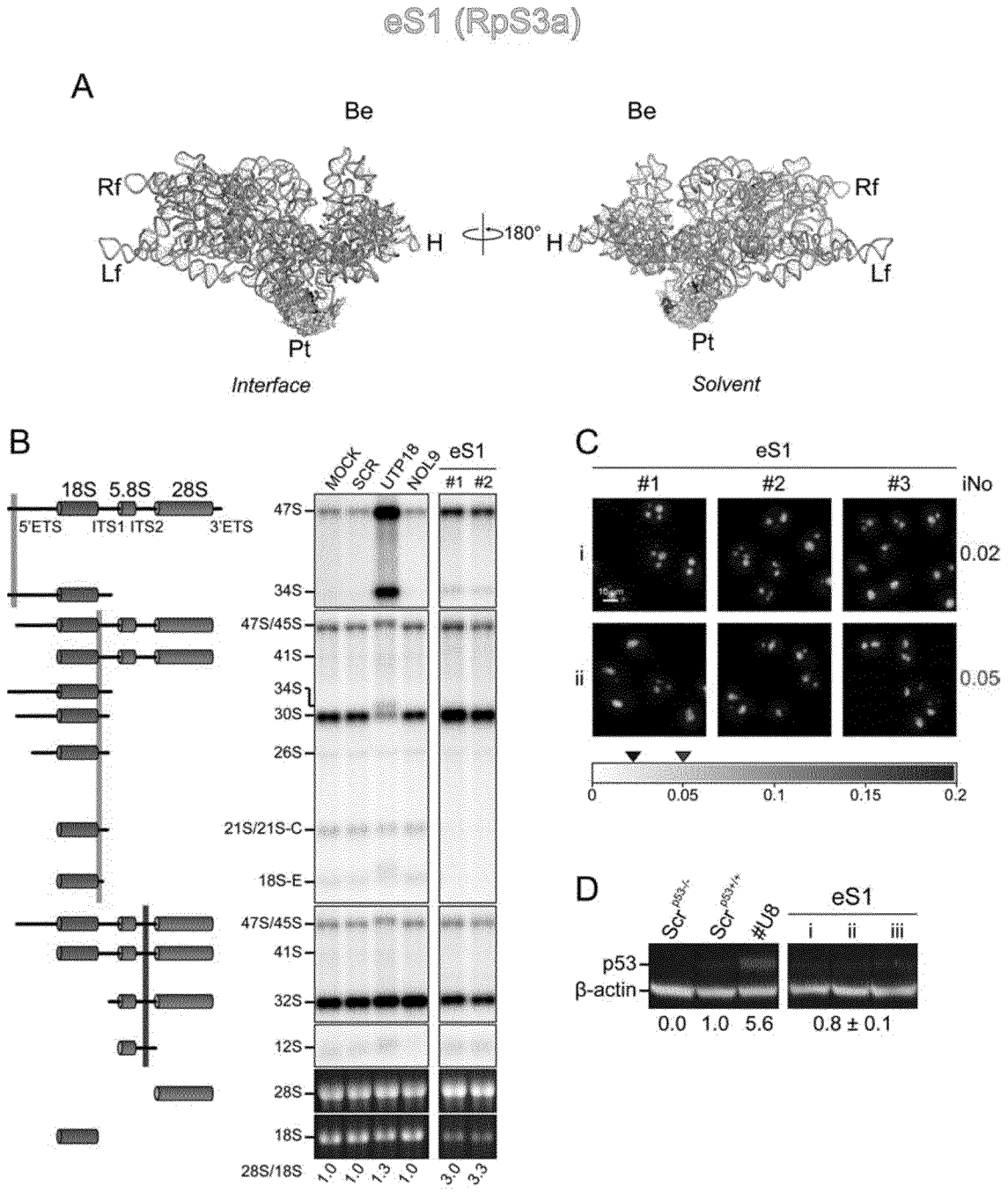


Fig. S10 continued

FIG. S11


eS1 (RpS3a)
A
Be
Be
Rf
Rf
H
180°
H
Lf
Lf
Pt
Pt
Interface
Solvent
B
18S 5.8S 28S
5'ETS ITS1 ITS2 3'ETS
MOCK SCR UTP18 NOL9 eS1
#1 #2
47S
34S
47S/45S
41S
34S
30S
26S
21S/21S-C
18S-E
47S/45S
41S
32S
12S
28S
18S
28S/18S
1.0 1.0 1.3 1.0 3.0 3.3
C
eS1
#1 #2 #3 iNo
i
0.02
ii
0.05
0 0.05 0.1 0.15 0.2
D
Scr p53-/- Scr p53+/+ #UB eS1
i ii iii
p53
β-actin
0.0 1.0 5.6 0.8 ± 0.1

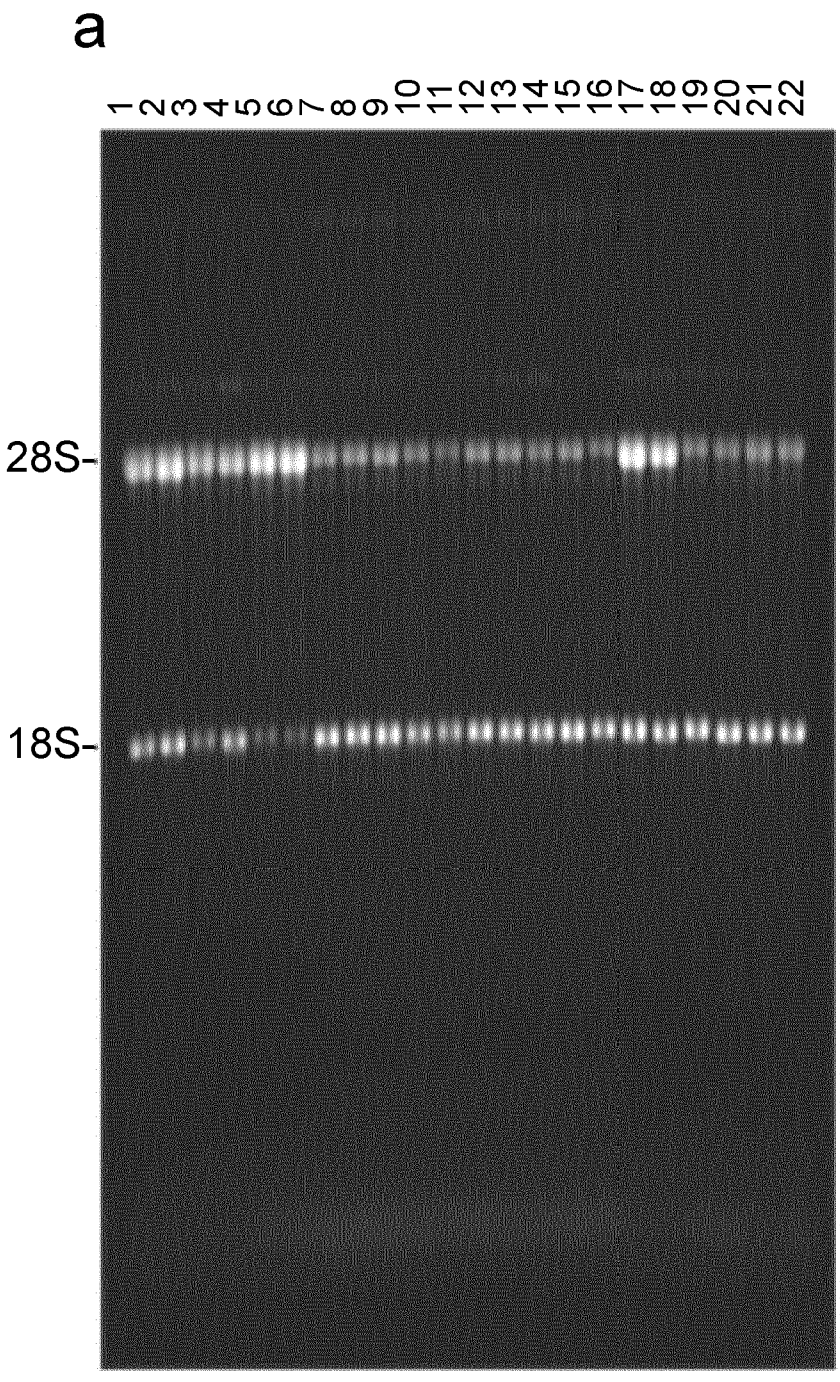

a
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
28S-
18S-
Ethidium Bromide

Fig. S12

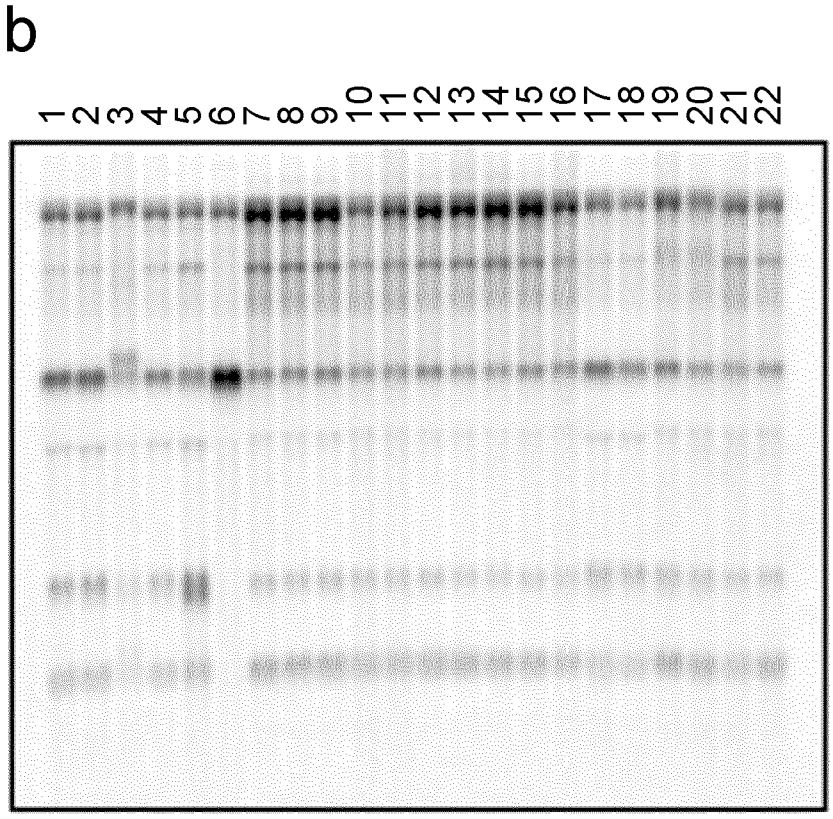
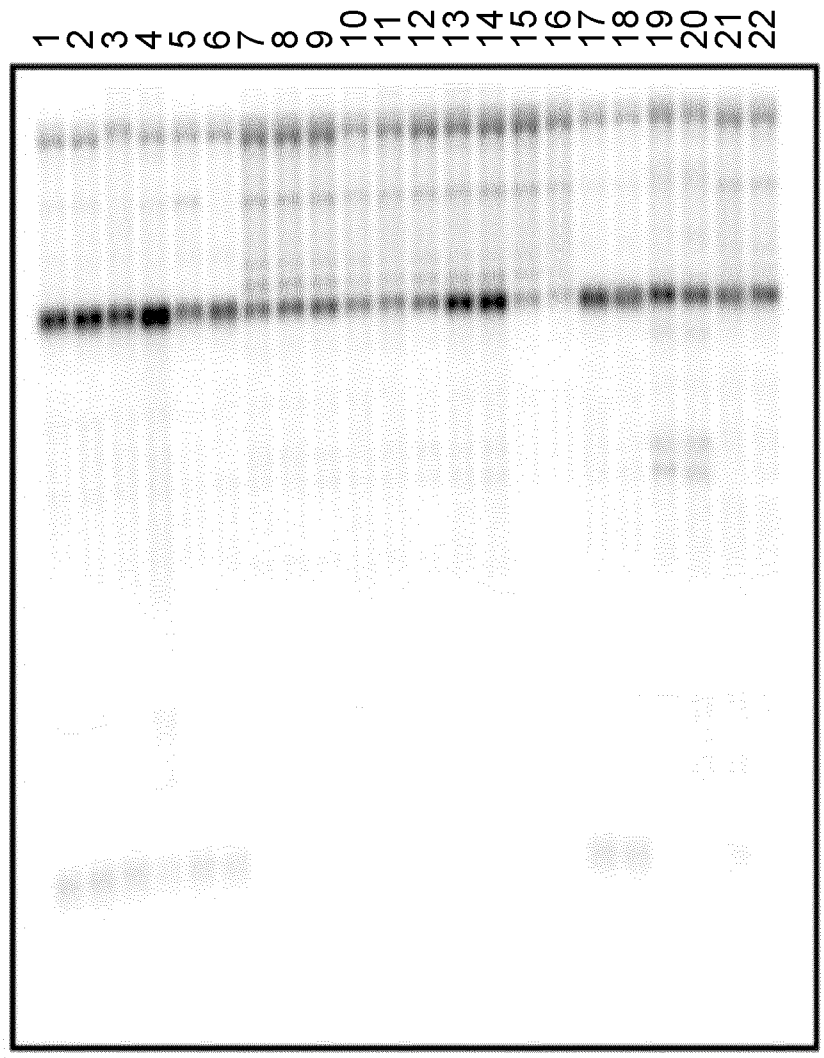


Fig. S12

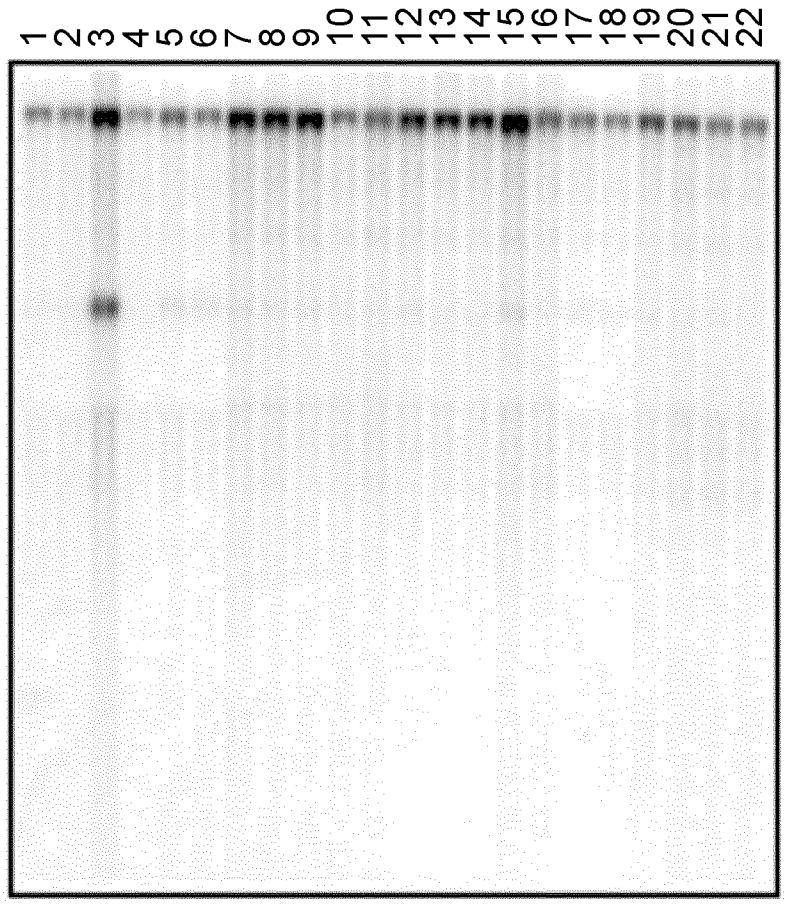
C



ITS2

Fig. S12

d



5'-ETS

Fig. S12

FIG. S13

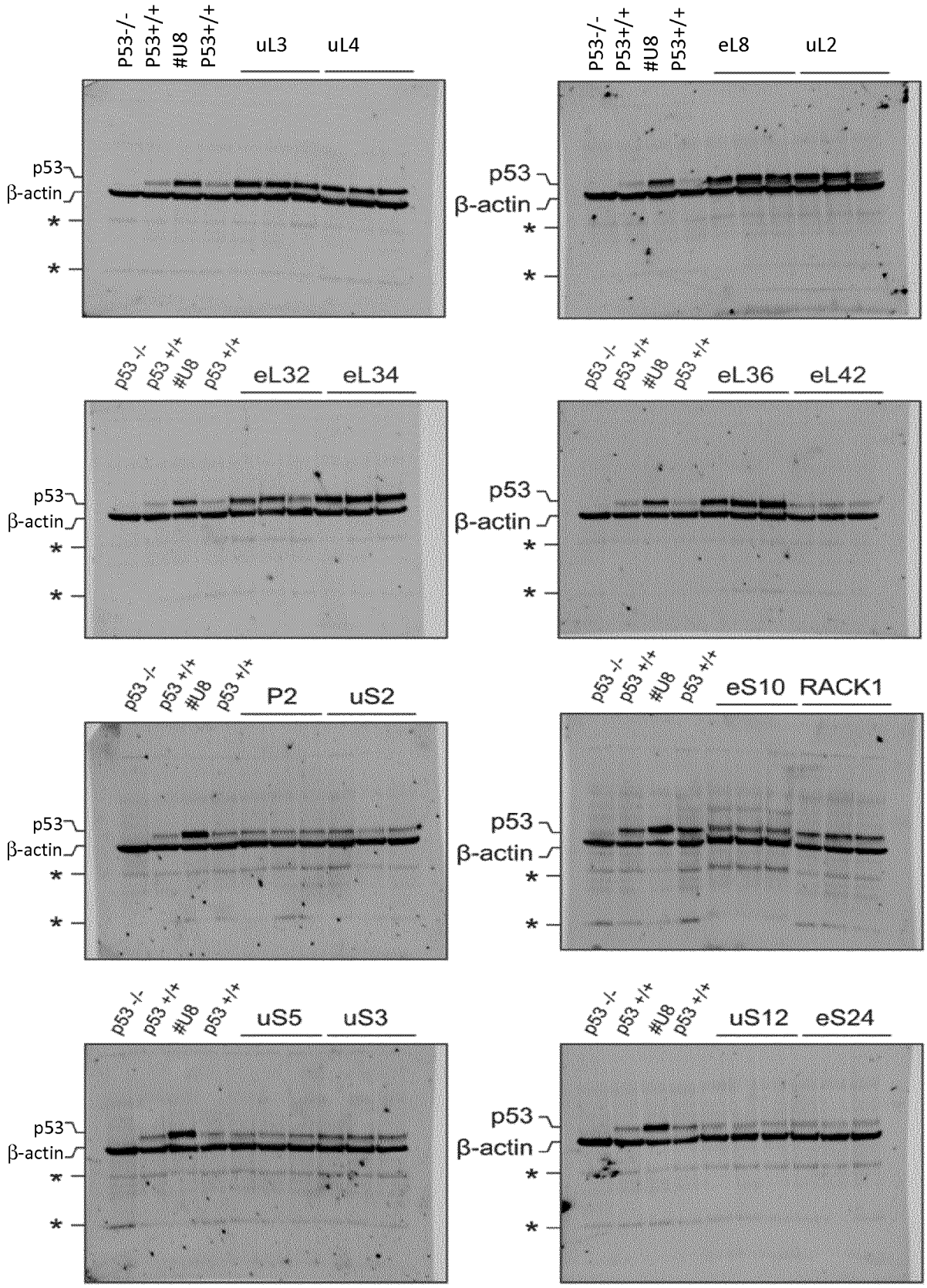

P53-/-
P53+/+
#U8
P53+/+
uL3
uL4
p53
β-actin
*
*
P53-/-
P53+/+
#U8
P53+/+
eL8
uL2
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eL32
eL34
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eL36
eL42
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
P2
uS2
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eS10
RACK1
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
uS5
uS3
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
uS12
eS24
p53
β-actin
*
*

FIG. S13 cont.

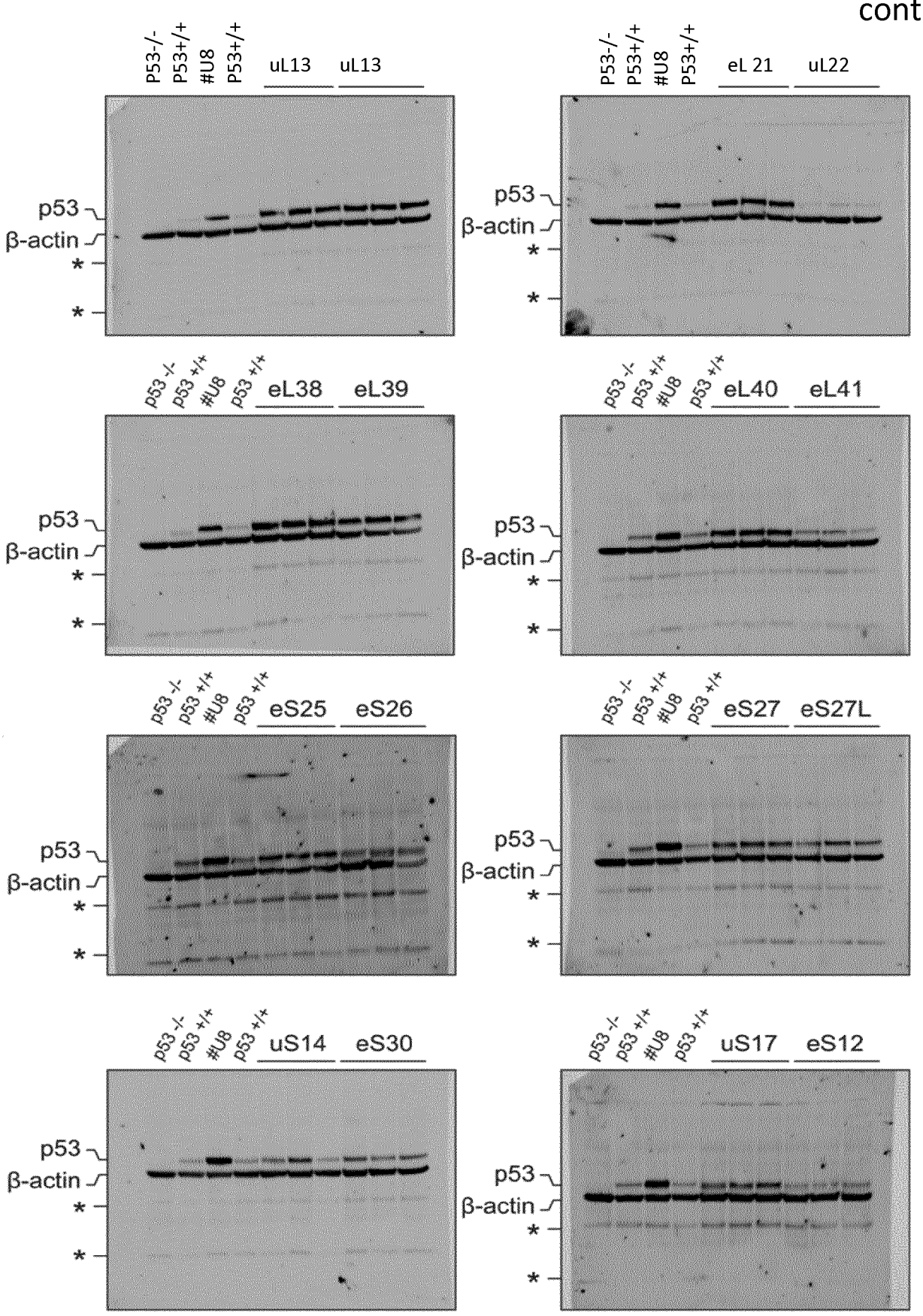

P53-/-
P53+/+
#U8
P53+/+
uL13
uL13
p53
β-actin
*
*
P53-/-
P53+/+
#U8
P53+/+
eL 21
uL22
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eL38
eL39
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eL40
eL41
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eS25
eS26
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
eS27
eS27L
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
uS14
eS30
p53
β-actin
*
*
p53 -/-
p53 +/+
#U8
p53 +/+
uS17
eS12
p53
β-actin
*
*

FIG. S14

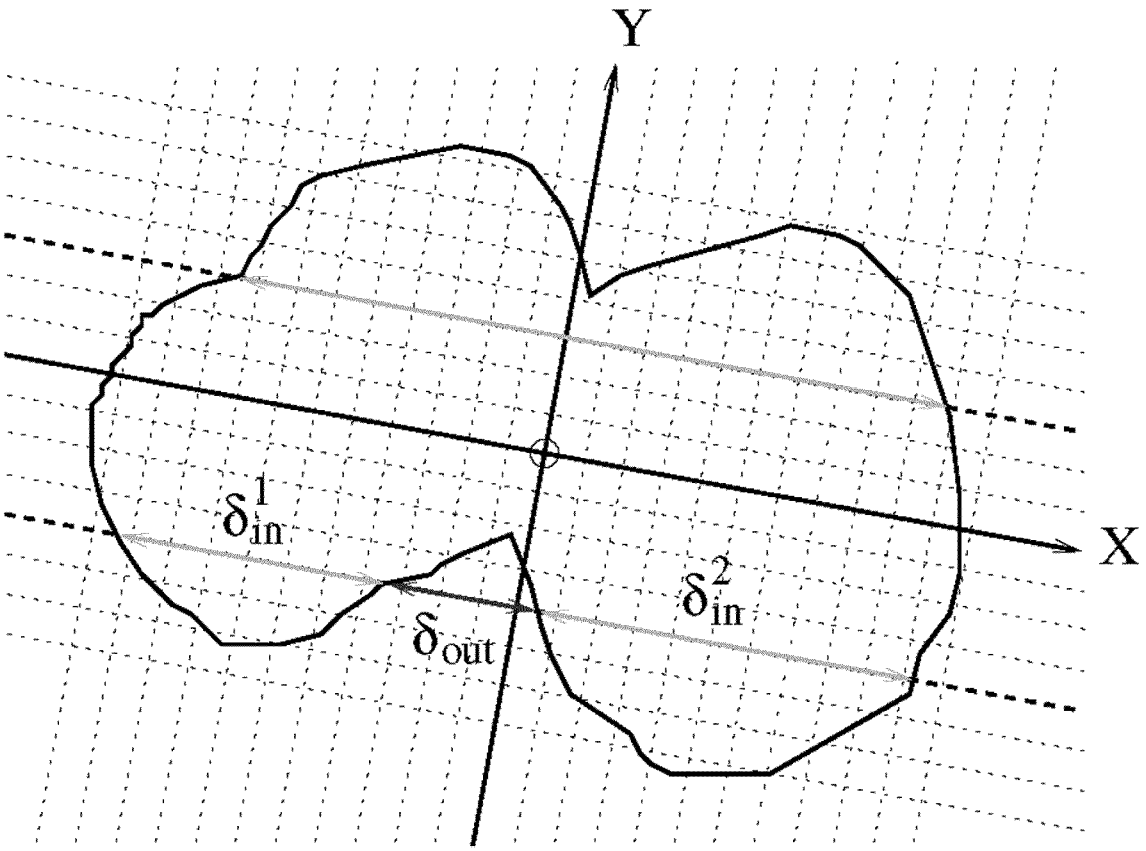


FIG. S15

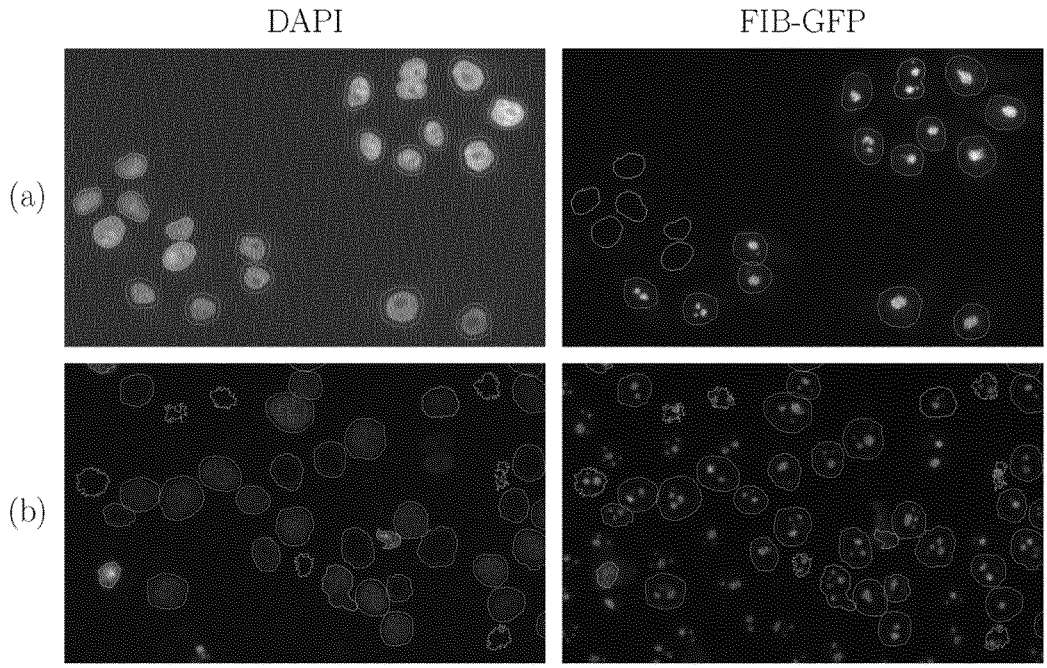
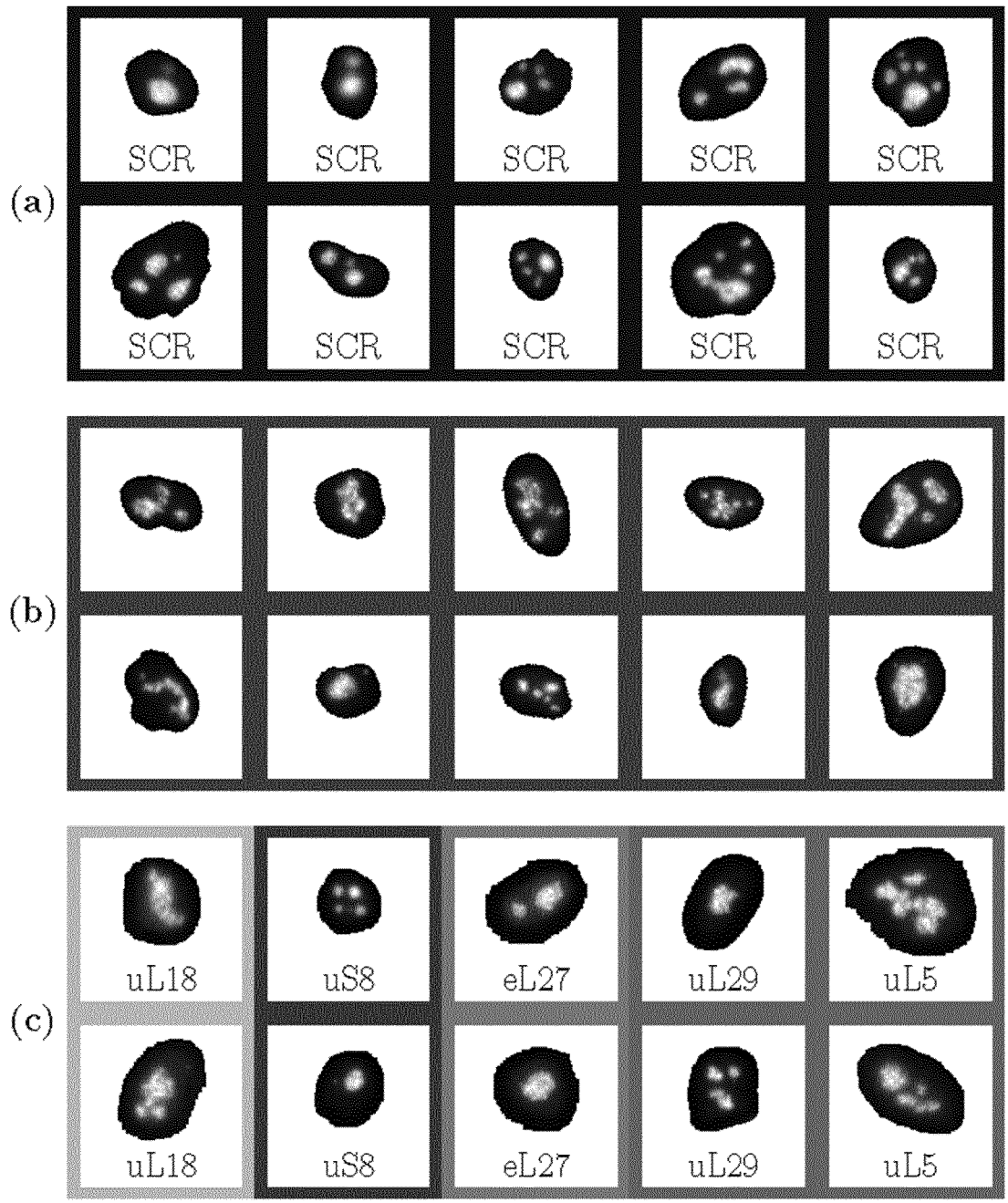


FIG. S16


(a)
SCR
SCR
SCR
SCR
SCR
SCR
SCR
SCR
SCR
SCR
(b)
(c)
uL18
uS8
eL27
uL29
uL5
uL18
uS8
eL27
uL29
uL5

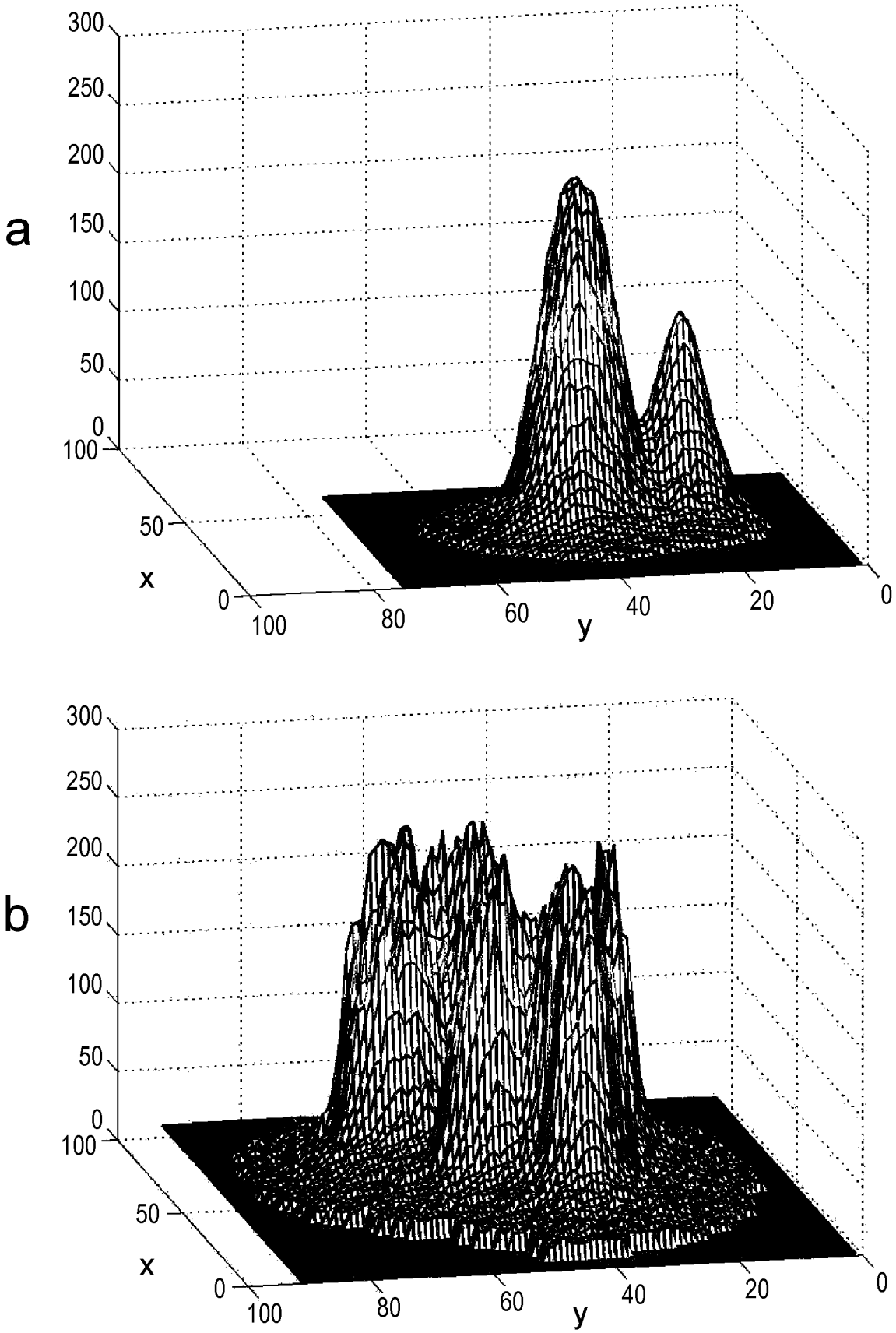

a
300
250
200
150
100
50
0
100
50
x
0
100
80
60
40
20
0
y
b
300
250
200
150
100
50
0
100
50
x
0
100
80
60
40
20
0
y

Fig. S17

FIG. S18

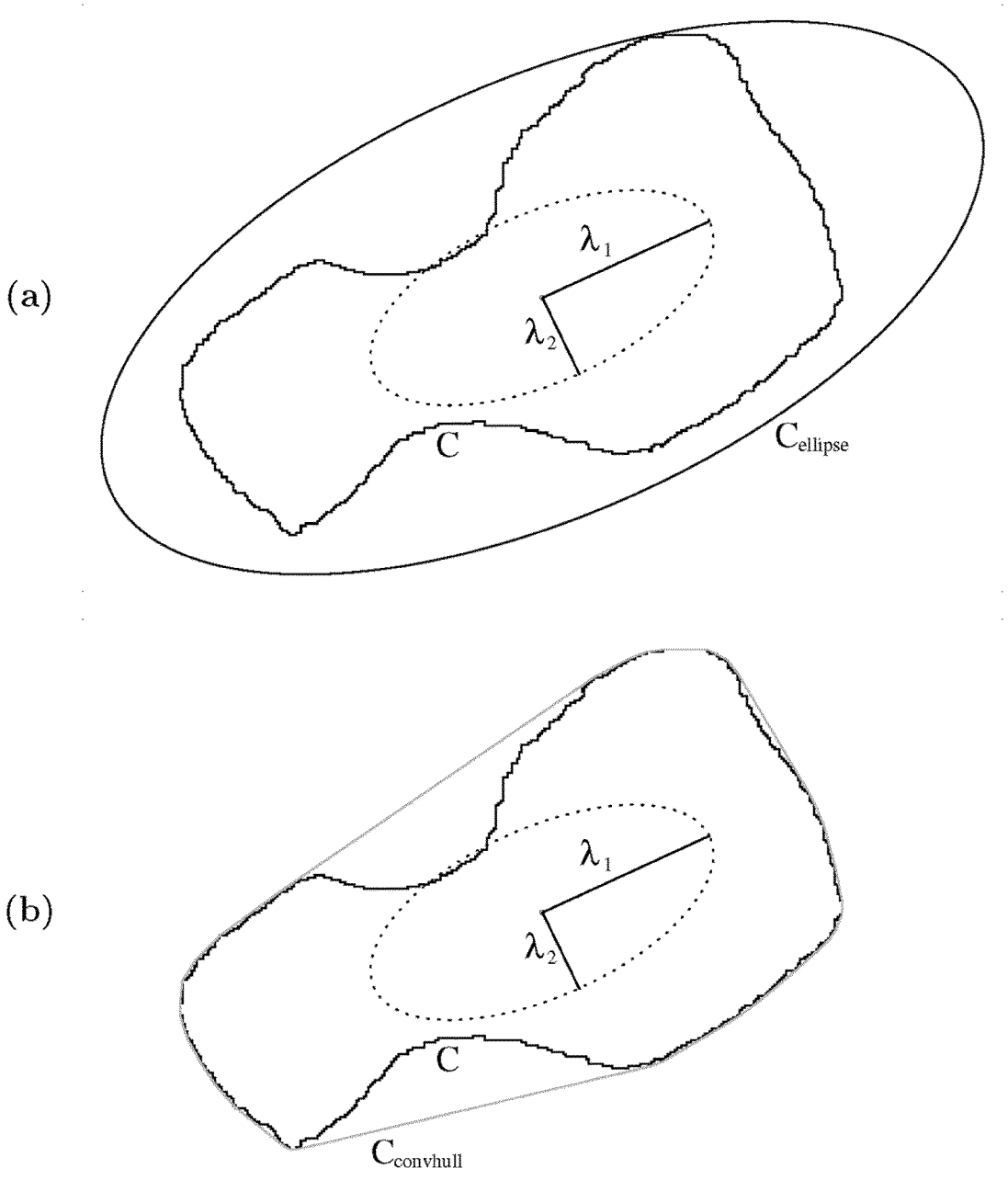
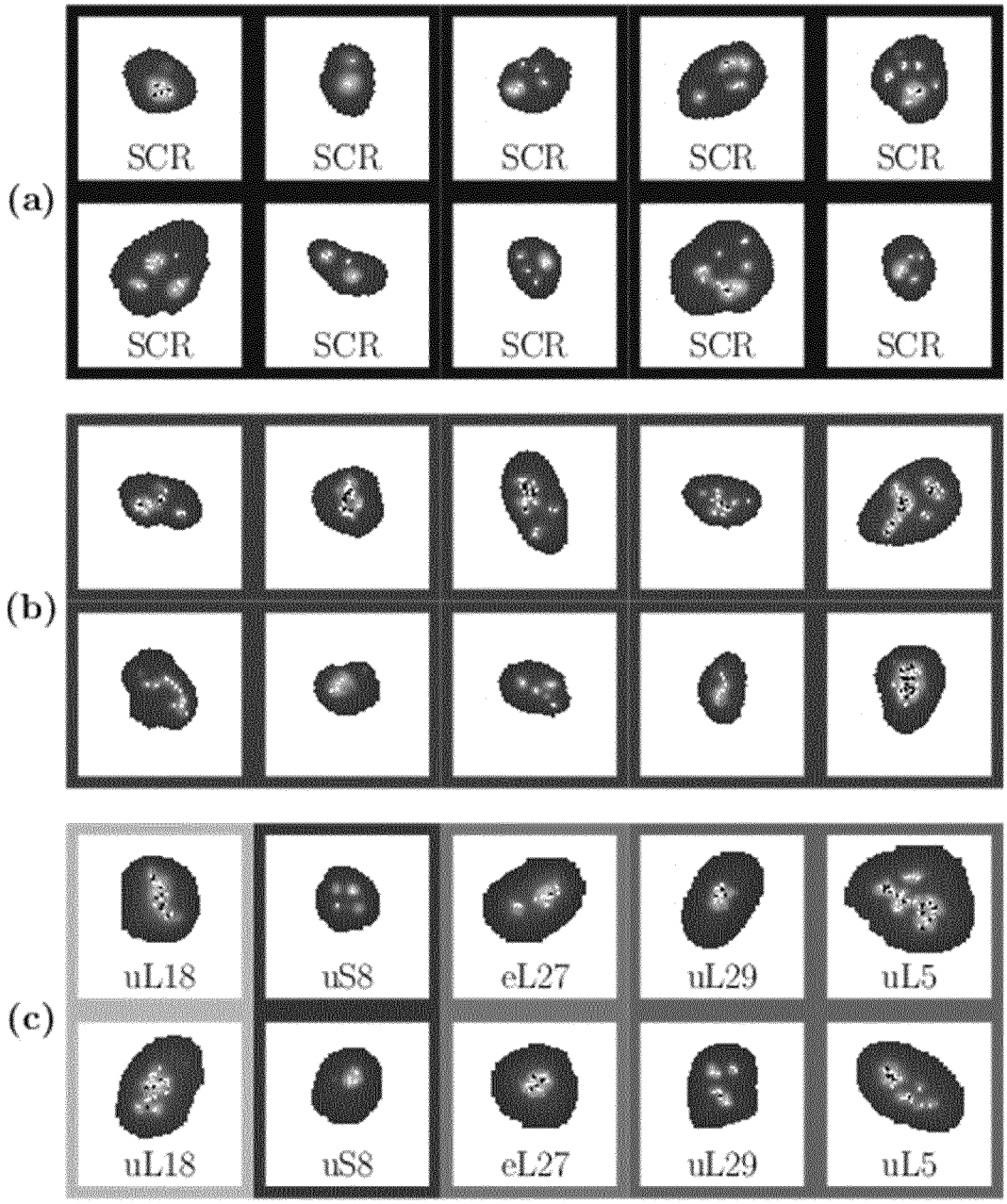
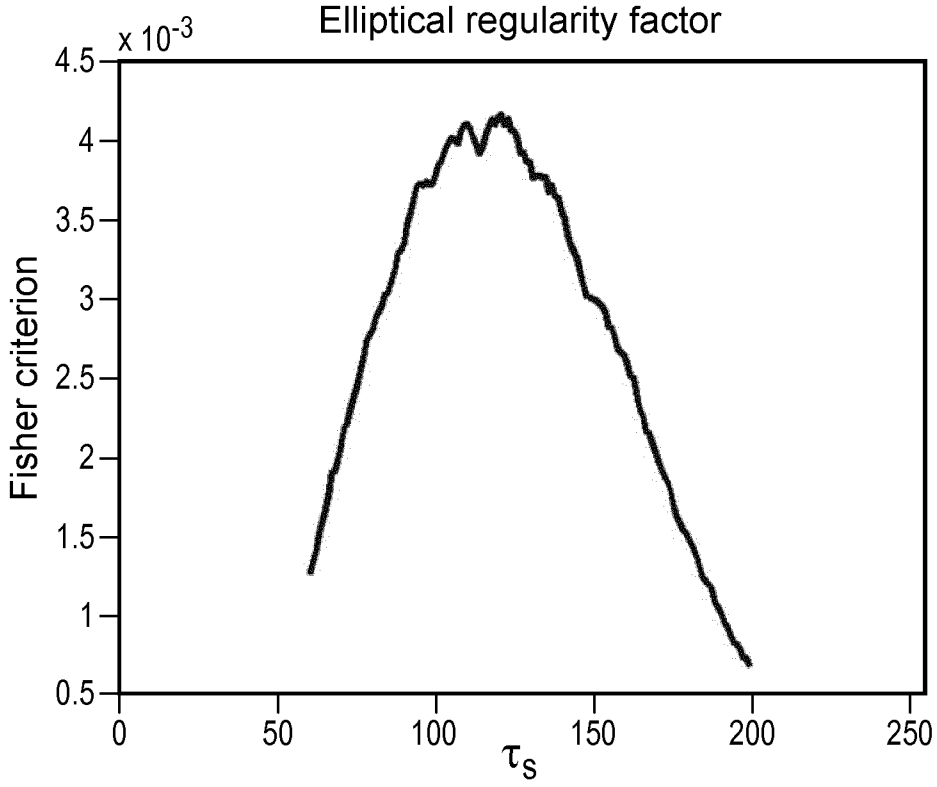


FIG. S19


(a)
SCR
SCR
SCR
SCR
SCR
SCR
SCR
SCR
SCR
SCR
(b)
(c)
uL18
uS8
eL27
uL29
uL5
uL18
uS8
eL27
uL29
uL5

a


Elliptical regularity factor
x 10-3
4.5
4
3.5
3
2.5
2
1.5
1
0.5
Fisher criterion
0
50
100
150
200
250
τs

b

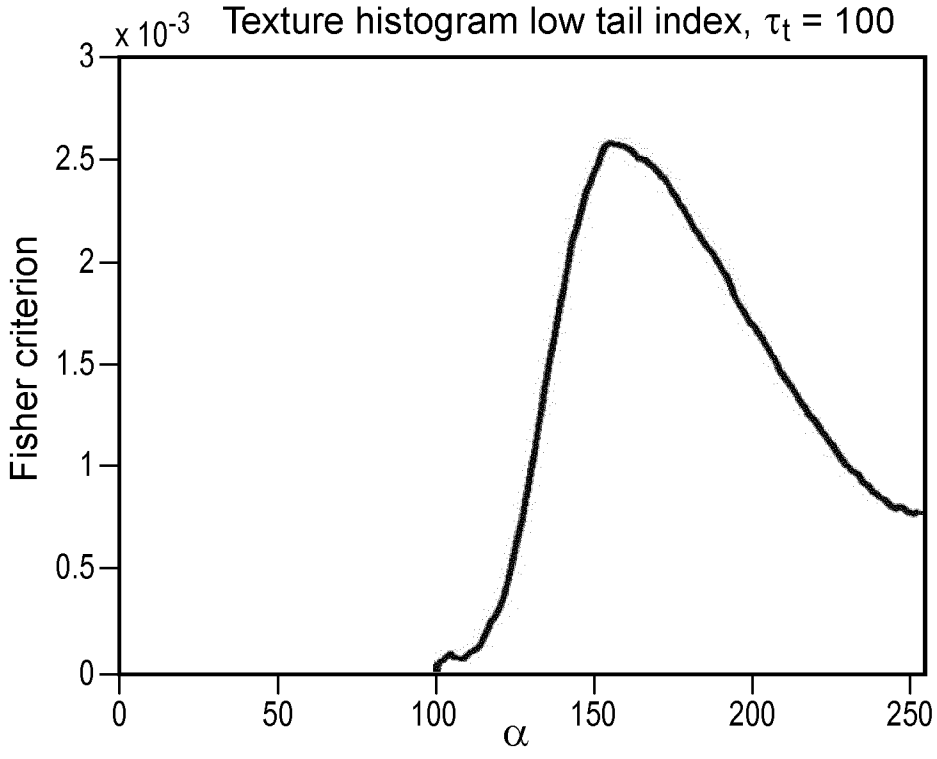

Texture histogram low tail index, τt = 100
x 10-3
3
2.5
2
1.5
1
0.5
0
Fisher criterion
0
50
100
150
200
250
α

Fig. S20

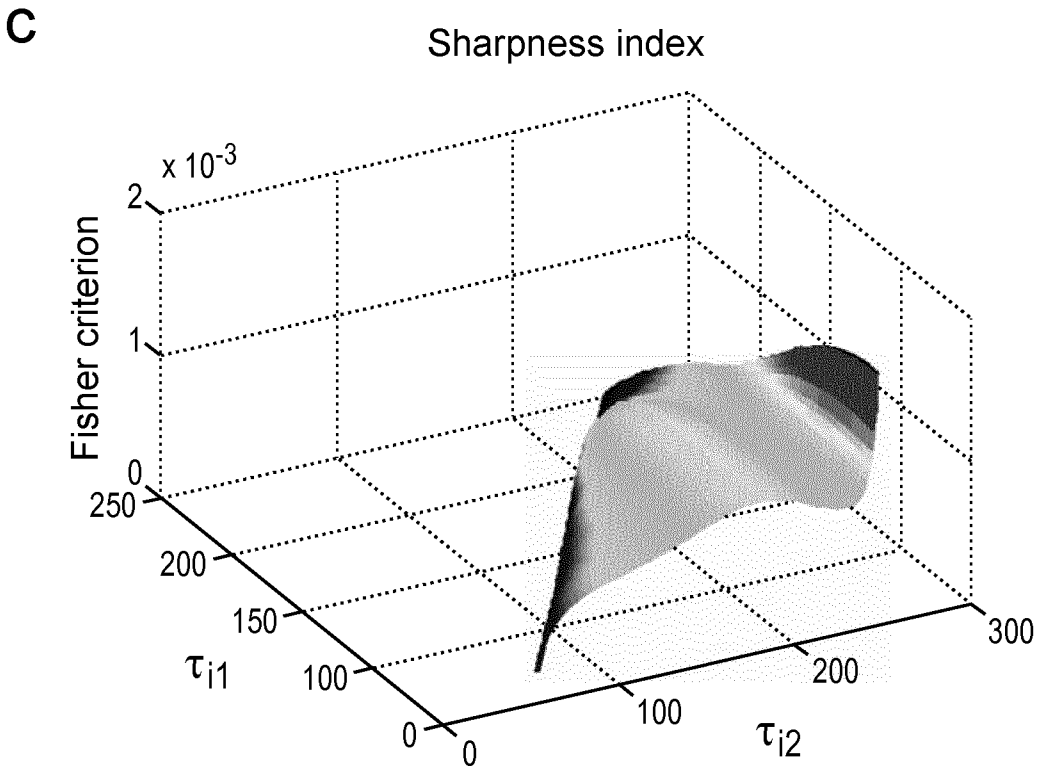

C
Sharpness index
x 10-3
2
1
0
Fisher criterion
250
200
150
100
0
τi1
0
100
200
300
τi2

Fig. S20