

**Manuscript Title:** Nucleolar enrichment of brain proteins with critical roles in human neurodevelopment

**Manuscript No:** MCP/2015/051920 [R1]

**Manuscript Type:** Research Article

**Date Submitted by the Author:** 30 Sep 2015

**Complete List of Authors:** Lukasz Slomnicki, Agata Malinowska, Michal Kistowski, Antoni Palusinski, Jing-Juan Zheng, Mari Sepp, Tonis Timmusk, Michal Dadlez, and Michal Hetman

**Keywords:** iTRAQ; Neurobiology\*; Ribosomes\*; Transcription\*; Translation\*; Chromatin ; Neurodevelopment; Nucleolus; Nucleus; Ribosome biogenesis

## **Nucleolar enrichment of brain proteins with critical roles in human neurodevelopment**

Lukasz Slomnicki<sup>1,\*\*</sup>, Agata Malinowska<sup>3,\*\*</sup>, Michal Kistowski<sup>3</sup>, Antoni Palusinski<sup>4</sup>, Jing-Juan Zheng<sup>1</sup>, Mari Sepp<sup>5</sup>, Tonis Timmusk<sup>5</sup>, Michal Dadlez<sup>3</sup>, Michal Hetman<sup>1,2,¶</sup>

<sup>1</sup>Kentucky Spinal Cord Injury Research Center and the Departments of Neurological Surgery and

<sup>2</sup>Pharmacology&Toxicology, University of Louisville, Louisville, Kentucky, USA, <sup>3</sup>Mass

Spectrometry Laboratory, Institute of Biochemistry and Biophysics, Polish Academy of

Sciences, Warsaw, Poland, <sup>4</sup>Department of Systems Biology, Faculty of Biology, University of

Warsaw, Warsaw, Poland, <sup>5</sup>Department of Gene Technology, Tallinn University of Technology,

Tallinn, Estonia

\*\* These authors contributed equally to this work.

¶ To whom correspondence should be addressed: Dr. Michal Hetman, Kentucky Spinal Cord Injury Research Center, University of Louisville, 511 S. Floyd St., MDR616, Louisville, KY40292, USA. Tel.: 502-852-3619; Fax: 502-852-5148

E-mail: [michal.hetman@louisville.edu](mailto:michal.hetman@louisville.edu)

Running Title: Nucleolar proteome of the developing brain

**Abbreviations:** aCSF, artificial cerebrospinal fluid; ActD, Actinomycin D; AS, Alazami Syndrome; bHLH, basic helix-loop-helix motif; BCS, Bowen-Conradi Syndrome; BP, biological function; CC, cellular compartment, CDLS, Cornelia De Lange syndrome; DFC, dense fibrillar component; DIV7, day *in vitro* 7; ECM, extracellular matrix; ER, endoplasmic reticulum; FASP, Filter-Aided Sample Preparation; GC, granular component; GFP, green fluorescent protein; GO, Gene Ontology; HP, human phenotype; iTRAQ, isobaric Tags for Relative and Absolute Quantitation; LARP7, La-related protein 7; MF, molecular function; NeuN, neuronal nuclei; No, nucleolus; Npm, nucleophosmin/B23, Nu, nucleus; OPP, O-propargyl-puromycin; P7, postnatal day 7, Pol1, RNA polymerase I; PTHS, Pitt-Hopkins syndrome; rDNA, ribosomal DNA; RFP, red fluorescent protein; RP, ribosomal protein; TCF4, transcription factor 4; TIF1A, transcription initiation factor 1A; UBF, upstream binding factor; UPLC, Ultra Performance Liquid Chromatography; wt, wild type;  $\beta$ -gal,  $\beta$ -galactosidase;

**Summary.** To study nucleolar involvement in brain development, the nuclear and nucleolar proteomes from the rat cerebral cortex at postnatal day 7 were analyzed using LC-MS/iTRAQ methodology. Data of the analysis are available via ProteomeXchange with identifier PXD002188. Among 504 candidate nucleolar proteins, the overrepresented gene ontology terms included such cellular compartment categories as “nucleolus”, “ribosome” and “chromatin”. Consistent with such classification, the most overrepresented functional gene ontology terms were related to RNA metabolism/ribosomal biogenesis, translation, and chromatin organization. Sixteen putative nucleolar proteins were associated with neurodevelopmental phenotypes in humans. Microcephaly and/or cognitive impairment were the most common phenotypic manifestations. While several such proteins have links to ribosomal biogenesis and/or genomic stability/chromatin structure (*e.g.* EMG1, RPL10, DKC1, EIF4A3, FLNA, SMC1, ATRX, MCM4, NSD1, LMNA or CUL4B), others including ADAR, LARP7, GTF2I or TCF4 have no such connections known. Although neither the Alazami syndrome-associated LARP7 nor the Pitt-Hopkins syndrome (PTHS)-associated TCF4 were reported in nucleoli of non-neural cells, in neurons, their nucleolar localization was confirmed by immunostaining. In cultured rat hippocampal neurons, knockdown of LARP7 reduced both perikaryal ribosome content and general protein synthesis. Similar anti-ribosomal/anti-translation effects were observed after knockdown of the ribosomal biogenesis factor EMG1 whose deficiency underlies Bowen-Conradi syndrome. Finally, moderate reduction of ribosome content and general protein synthesis followed overexpression of two PTHS mutant variants of TCF4. Therefore, dysregulation of ribosomal biogenesis and/or other functions of the nucleolus may disrupt neurodevelopment resulting in such phenotypes as microcephaly and/or cognitive impairment.

## **Introduction**

The nucleolus is the largest subnuclear structure and is the center of ribosomal biogenesis (1, 2). The RNA-polymerase-1 (Pol1)-mediated transcription of nucleolar rRNA genes is the first step of ribosome production that is also essential for establishment and maintenance of the nucleolus. In addition, pre-rRNA is processed in the nucleolus into mature rRNA species and assembled with tens of ribosomal proteins to form pre-ribosomal subunits (3, 4). This complex process requires hundreds of trans-acting factors including proteins and small RNAs. Because ribosome production requires many resources, it is highly regulated (2).

Ribosomes are relatively stable structures with an estimated half life time of several days (5). Hence, it is not surprising that a highly active nucleolus is often observed in intensely dividing cells which require many new ribosomes (1, 2). Conversely, non-dividing cells usually display small nucleoli. However, prominent nucleoli are present in various types of mature neurons where initial observations of that structure were reported over 150 years ago (6). Moreover, neuronal nucleoli appear to increase in size as the nervous system develops (7-9). Given the importance of protein synthesis for neuron biology, including neuritic morphogenesis, neurite maintenance, and synaptic plasticity, ribosomal biogenesis may be required to maintain and regulate the adequate translation apparatus of a neuron. In support of such a possibility, brain derived neurotrophic factor (BDNF), which promotes neuronal morphogenesis, survival, and synaptic plasticity, stimulates Pol1-driven transcription (10). Therefore, despite their post-mitotic status, neurons may maintain large nucleoli to rapidly produce additional ribosomes in response to neurotrophic stimulation.

Alternatively, the neuronal nucleolus may be involved in processes that are unrelated to ribosome biogenesis. For instance, as proficient ribosome biogenesis is required to suppress the

p53 pathway, the nucleolus serves as a cellular stress sensor whose dysfunction triggers the p53-mediated cell cycle arrest and/or apoptosis (11). The stress sensing ability of the nucleolus has been also demonstrated in neurons (12, 13). In addition, the nucleolus is a site of non-ribosomal ribonucleoprotein complex biogenesis, including the signal recognition particle or components of the spliceosome (14). Finally, export and degradation of nuclear proteins, regulated protein sequestration, as well as establishment and maintenance of chromatin structure add to the expanding list of non-ribosomal functions of the nucleolus (14-17).

Proteomic analysis of nucleoli from various types of proliferating cells supports a notion of the plurifunctional nucleolus. Initial reports revealed that ribosomal biogenesis factors represent about 40-50% of roughly 300 proteins that were found in isolated nucleoli (18-20). Improved depth of peptide identification expanded the nucleolar proteome to nearly 700 proteins with 25% of them being initially implicated in ribosomal biogenesis (20, 21). However, an RNAi screen whose design was based on such extended proteomic data sets uncovered tens of new ribosomal biogenesis factors expanding their nucleolar proteome contribution to more than 40% (22). Processing of tRNA, mRNA metabolism including splicing, DNA repair, DNA replication, chromatin organization, and cell cycle regulation are among other functional categories that are abundantly represented in the nucleolar proteome. Hence, analysis of the nucleolar proteome can be used to identify potentially novel functions of the nucleolus as well as novel regulators of its well established roles such as ribosomal biogenesis.

The qualitative proteomic analysis of several human cell lines that were grown under diverse conditions suggested that 4500 proteins may be located in human cell nucleoli in various cellular contexts (23). However, as nucleolar preparations may also be contaminated by other cellular compartments, at least some of these proteins may represent false positive hits (20).

Instead, quantitative analysis of nucleolar protein enrichment as compared to other cellular compartments provides a more specific approach to identify proteins with a localization preference for the nucleolus (24).

The functional organization of brain cell nucleoli remains mostly unexplored. In the nervous system, neither components nor regulators of the ribosomal biogenesis pathway have been systematically investigated with the exception of the Pol1 co-factor TIF1A whose knockdown or knockout blocks rRNA transcription in various neuronal populations (12, 13, 25, 26). Likewise, there is little information on non-ribosomal functions of brain cell nucleoli beyond their involvement in the p53 stress response pathway. Therefore, the current study was initiated to identify nucleolar proteins of the developing rat brain using quantitative iTRAQ proteomics.

## **Experimental Procedures**

*Animals.* Sprague–Dawley pregnant female rats were purchased from Harlan (Indianapolis, IN, USA). For proteomic analysis, three separate litters of neonate rats were used at postnatal day 7 (P7). The cortical tissue was collected from littermates of both sexes. All animal experiments strictly followed the protocols that were approved by the University of Louisville Institutional Animal Care and Use and Institutional Biosafety Committees.

*Materials.* All reagents were obtained from Sigma (St. Louis, MO), VWR (Radnor, PA), or Life Technologies-Invitrogen (Grand Island, NY) unless stated otherwise.

*Nucleolar isolation.* Rat neocortical nuclei and nucleoli were purified using a previously published procedure with modifications (18) (and <http://www.lamondlab.com/pdf/noprotoocol.pdf>, Fig. 1A). For preparation of each set of samples 5-8 pups from a single litter were used. Pulling at least 5 animals/preparation was necessary to obtain sufficient amount of proteins in the nucleolar sample. Three independent preparations/sample sets from three different litters were used for the subsequent proteomic analysis. All steps were performed on ice or at 4°C. The S2 solution as well as RIPA lysis buffer (Santa Cruz Biotechnology, Dallas, TX, #24948; 1xPBS (pH 7.4), 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 0.04% sodium azide) were supplemented with protease inhibitors. Dissected cerebral cortices were placed in the ice-cold artificial cerebral spinal fluid (aCSF, 126 mM NaCl, 25 mM NaHCO<sub>3</sub>, 25 mM Glucose, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>-H<sub>2</sub>O, 2.5 mM KCl, 2 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, saturated with 5% CO<sub>2</sub>) and gently dissociated in a Dounce tissue homogenizer by turning the pestle to avoid generation of negative pressure. Cells were pelleted at 218 x g, resuspended in aCSF, and applied onto 40 µm nylon Cell Strainer (BD Falcon, Franklin Lakes, NJ) to separate tissue chunks and blood vessels. The filtrate was again



centrifuged at 218 x g. The cell pellet was re-suspended in 5 ml of hypotonic Buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.5 mM DTT) and incubated on ice for 5 min. The cell suspension was transferred directly to a pre-cooled Dounce tissue homogenizer and homogenized by 15 pestle strokes using a tight pestle until approximately 90 % of cells burst. After centrifugation at 218 x g for 5 min, whole nuclei were collected from the pellet in 5 ml of ice-cold S2 solution (0.35 M Sucrose, 0.5 mM MgCl<sub>2</sub>). One tenth of the sample was lysed in 1 ml RIPA buffer and used as a nuclear fraction for further analyses. The rest of the sample was sonicated at 40% power in 4 bursts, 10s each with 10s intervals to disrupt approximately 90% of nuclei. The suspension was ultracentrifuged at 1430 x g for 5 min. The resulting nucleolar pellet was re-suspended in 5 ml ice-cold S2 followed by ultracentrifugation and lysis in 1 ml RIPA. The lysates were frozen on dry ice and stored at -80°C.

*Western blotting.* Western blotting was performed using standard procedures. The primary antibodies were as follows: anti-NeuN (1:1000, Millipore, Billerica, MA ABN78), anti-nucleophosmin/B23 (1:1000, Sigma #B0556), anti-fibrillarin (1:1000, Abcam, Cambridge, MA, AB4566), anti-UBF (1:1000, Santa Cruz Biotechnology, SC-13125), anti-actin (1:1000, Sigma, #A4700), anti-β-tubulin (1:1000, Sigma, T5293), anti-LARP7 (1:1000, Abcam, AB105683), anti-EMG1 (1:1000, Proteintech Group, 11965-1-AP), anti-GFP (MBL, 1:1000), anti-β-galactosidase (MP Biomedicals, 1:1000), anti-Tom20 (1:1500, a gift of Dr. Brian Wattenberg, University of Louisville, Louisville, KY), anti-calreticulin (1:1000, Enzo); secondary antibodies were horseradish peroxidase conjugated.

*Dot blot analysis of protein concentration.* Two μl of each sample were spotted onto chromatography paper filter (Whatman). After all spots dried (about an hour at RT), the membrane was incubated for at least 30 min at RT in fixing solution (40% methanol, 10% glacial

acetic acid) on a platform shaker. Next, fixing solution was replaced by Coomassie G250 staining buffer (0,025% Coomassie, 10% glacial acetic acid) and incubated for at least 3 h. The membrane was then incubated with destaining buffer (25% methanol, 10% glacial acetic acid). Standard curve spots containing variable amount of BSA were used to estimate protein concentration of each sample.

*Mass spectrometry.* All the samples were brought to the following reagent concentrations: 2% of SDS and 100 mM DTT. Protein concentration was estimated with dot-blot. and 100 µg of protein from each sample was taken for further preparation. Proteins were denatured and reduced at 95°C for 5 minutes, alkylated with iodacetamide and subsequently digested with trypsin (Promega) in filter devices (10 MWCO, Sartorius Stedim) using the Filter-Aided Sample Preparation (FASP) method (27). The digestion was performed for 16h at enzyme/protein ratio of 1/100. For better compatibility with iTRAQ, we used triethylammonium bicarbonate buffer (TEAB, Sigma) instead of ammonium bicarbonate. . Further on, all samples were lyophilized and incubated with iTRAQ mass tags as recommended in the manufacturer's protocol (Agilent, Santa Clara, CA). iTRAQ 8-plex was used and the labeling pattern was as follows: nuclear fractions in channels 113, 115 and 117, nucleolar fractions in channels 114, 116 and 118. After 1 hour the labeling reaction was stopped by diluting each sample with an appropriate amount of sample buffer (8 M urea, 0.2% IPG buffer pH 3–11 NL (GE Healthcare), 0.002% bromophenol blue in 50 mM Tris–HCl, pH 8.0). Labeled samples were then pooled, concentrated in a Speedvac and fractionated on the IEF strip. The strip was then cut into 18 fractions and from each fraction peptides were extracted. Peptide mixtures were analyzed with an LC-MS system, composed of a UPLC chromatograph (nanoAcquity, Waters) directly coupled to a Q Exactive mass spectrometer (Thermo), working in the data-dependent acquisition mode.

*Proteomic data analysis.* The acquired MS/MS data were pre-processed with Mascot Distiller software (v.2.4.3, MatrixScience, Boston, MA) and a search was performed with the Mascot Search Engine (MatrixScience, Mascot Server 2.4.1) against the UniProt mouse protein database (UniProt 2013\_11; 43,236 sequences). To reduce mass errors, the peptide and fragment mass tolerance settings were established separately for individual LC-MS/MS runs after a measured mass recalibration, as described previously (28). Subsequently, data from individual MS runs were merged into a single Mascot search. The parameters were as follows: enzyme - Trypsin, missed cleavages - 1, fixed modifications - Carbamidomethyl(C), variable modifications - Oxidation (M), quantitation - iTRAQ 8plex, taxonomy - unrestricted, instrument - HCD, parent MMD - 5 ppm, fragment MMD - 0.08 Da, Decoy option - active. A statistical assessment of the confidence of peptide assignments was based on the target/decoy database search strategy (29). This procedure provided q-value estimates for each peptide spectrum match in the dataset. All queries with q-values >0.01 were removed from further analysis. A protein was regarded as identified confidently when at least two peptides of this protein were found. Proteins identified by a subset of peptides from another protein were excluded from the analysis. Proteins that exactly matched the same set of peptides were combined into a single group (family). The mass calibration and data filtering described above were carried out with MScan software, developed in-house (<http://proteom.ibb.waw.pl/mscan/>). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (<http://www.ebi.ac.uk/pride/archive/>) with the dataset identifier PXD002188. In addition, detailed peptide/protein identification data are presented as a supplementary material (DiffprotOutput.txt and Table S5).

*Nucleolar classifier.* iTRAQ results were also incorporated into features used to create a classifier for prediction of nucleolar localization of identified proteins. The Multilayer Perceptron neural network-based classifier was constructed using WEKA suite (30). It included the following features: iTRAQ protein ratio (each sample pair separately), number of peptides, nucleolar localization sequences predicted by the Nod software (31), number of various types of amino acids present in protein sequence, length of amino acid sequence, gene ontology and interaction data (from the IntAct database) and presence of various protein domains. Out of them, the following were selected by the software to perform the classification: iTRAQ protein ratio, interaction data, presence of lysine-rich domain (PS50318). The training set – which included proteins present in nucleus but not nucleolus and known nucleolar proteins - was constructed using published data (32). The classifier achieved 93% of correctly assigned instances and precision of 0.88 in cross-validation. Two hundred two proteins identified in all preparations were assigned nucleolar localization by the classifier, 88 of which acquired q-value below 0.05 in the statistical analysis. The remaining 114 proteins had on average less peptides identified (2-5, compared to median of 9 peptides for significantly enriched proteins) but higher No/Nu enrichment ratios.

*Gene ontology analysis.* Gene ontology analysis was performed with the NCBI's DAVID analysis tool. GO\_CC\_FAT, GO\_BP\_FAT, and GO\_MF\_FAT algorithms were used (33)(<http://david.abcc.ncifcrf.gov/>). The default parameters of the GO enrichment searches were used with a count threshold set at 2 and the EASE (*p*-value) threshold set at 0.1; fold enrichment as well as FDR values were displayed in search results. Human phenotype analysis was performed with the ToppGene using the ToppFun algorithm (34)(<https://toppgene.cchmc.org/>).

*Plasmid DNA.* The following plasmids were previously described: pCDNA3.1-based expression vectors for human wild type TCF4 B- isoform and its mutants (R580W or R582P) (35); chicken  $\beta$ -actin promoter-driven-RFP or GFP (36) and EF1 $\alpha$ -LacZ (EF1 $\alpha$  promoter-driven  $\beta$ -galactosidase) (37). To prepare Emg1-EGFP expressing plasmid vector 720 bp-long cDNA encoding rat Emg1 was cloned into pEGFP-N1 plasmid with XhoI/BamHI restriction enzymes.

*Cell culture and transfections.* Cortical and hippocampal neurons were prepared from newborn Sprague–Dawley rats at postnatal day 0 as described (38). Cells were plated onto poly-D-lysine- and laminin-coated 12 mm diameter plastic coverslips that were produced in the lab from electron microscopy-grade mylar masks (Electron Microscopy Sciences, Hatfield, PA, #50425). Cells were grown in Neurobasal medium (supplemented with the B-27 supplement (Invitrogen), 1 mM L-glutamine, 100 units/ml penicillin, and 0.1 mg/ml streptomycin) half of which was changed every other day and transfected on day in vitro 3 (DIV3) or DIV6 using Lipofectamine 2000 (Invitrogen) as previously described (39). Cerebrocortical astrocyte-enriched glial cultures were prepared from neonate rats as reported elsewhere (40).

*Immunostaining.* All stainings of cultured neurons were performed following 20 min fixation in 4% formaldehyde at room temperature followed by permeabilization by 0.5% NP-40 in PBS for 15 minutes. The following primary antibodies were used: rabbit anti-NeuN (1:1000, Millipore, ABN78), mouse anti-nucleophosmin/B23 (1:750, Sigma #B0556), mouse anti-fibrillarin (1:750, Abcam AB4566), mouse anti-UBF (1:750, St. Cruz, sc-13125), rabbit anti-LARP7 (1:100, Abcam, AB105682) and rabbit anti-TCF4 (1:100, Abcam AB185736). The following secondary antibodies were used Alexa Fluor 488 goat anti-rabbit IgG (1:300, Invitrogen, A11034), Alexa Fluor 594 anti-mouse IgG (1:300, Invitrogen, A11005) and Alexa Fluor 555 anti-rabbit IgG (1:300, Invitrogen, A21429). Brain tissue was obtained from P7 rats; immersion-fixed in 4%

paraformaldehyde for 24h and 20  $\mu$ m sections were prepared using as previously described (41). Antigen recovery was performed by incubating brain sections with 10 mM sodium citrate, 0.05% Tween 20, pH 6.0 at 90-95°C for 20 min followed by standard immunostaining protocol.

*Generation of shRNA expression constructs.* To generate shRNAs against rat *Larp7* and *Emg1* each mRNA sequence was analyzed using shRNA design software (rnaidesigner.lifetechnologies.com) followed by off-target exclusion using Blastn (NCBI).

Oligonucleotides were designed:

Larp7-1: 5'gatccccgccagtcagcacattcgattcaagagaatcgaatgtgctgactggctttta3',

Larp7-2: 5'gatccccgctaatacaciaagctgagttcaagagactcagctttggtgattagctttta3',

Larp7-3: 5'gatccccggccaaagctaaagaagagttcaagagactcttcttagctttggcctttta3',

Emg1-1: 5'gatcccccttacgagctactcaactgttcaagagacagttgagtagctcgaagttttta3',

Emg1-2: 5'gatccccgaatgtgctcattgaagtgttcaagagacacttcaatgagcacattctttta3' together with their complementary counterparts, annealed and subcloned into the pSuper vector (Oligoengine) digested with BglIII and HindIII.

*Quantitative Reverse Transcriptase PCR (qRT-PCR).* Tri Reagent-based (Sigma) RNA isolation, cDNA synthesis and quantitative PCR were performed as described previously (12). The following oligonucleotide primers were used: bacterial  $\beta$ -galactosidase forward 5'tactgacgaaacgcctgccagtat3' and reverse 5'gctgtgccgaaatgtccatcaaa3'; human TCF4-specific forward: 5'acttgcttcgatctattctccag3' and reverse 5'cataattaggagacgatgaggcc3'. The RNA levels were analyzed using  $\Delta\Delta$ Act method. Expression values obtained from triplicate runs of each cDNA sample were calculated relative to the triplicate value for the bacterial  $\beta$ -galactosidase from the same cDNA preparation.

*Quantification of ribosomes.* Cultured hippocampal neurons were fixed with 4% formaldehyde for 20 min. Ribosomes were visualized with the NeuroTrace 500/525 Green Fluorescent Nissl Stain (Invitrogen, N21480, 1:500); DNA was counterstained with 2.5 µg/mL Hoechst 33258. To control for the specificity of NeuroTrace staining, some coverslips with attached cells were pre-treated with RNase A (Qiagen, 20 u in PBS, 1h at 37°C). Cells were imaged using a 40x objective lens and a Zeiss Axioobserver inverted microscope system. Digitalized pictures were captured using the AxioVision software (Zeiss) followed by conversion to gray scale TIFF files. Identical exposure times were used for all pictures that were used for comparative analyses. Quantification of the fluorescence intensity was prepared by using the „Integrated Density” parameter in the ImageJ software with background correction.

*Protein synthesis assay.* To visualize protein synthesis in individual neurons O-propargyl-puromycin (OPP) labeling method was used followed by Click-iT chemistry detection of nascent peptides (42). A commercial kit was employed according to manufacturer’s recommendations (Click-iT Plus OPP Alexa Fluor 594 Protein Synthesis Assay Kit, Molecular Probes, C10457). To control for the specificity of OPP staining, some coverslips with attached cells were pre-treated with cycloheximide (CHX, Sigma, 50 µM, 1h at 37°C). Quantification of newly synthesized, OPP-labeled peptides was performed as described for ribosome content analysis.

*Statistical analysis.* Statistical analysis of quantitative proteomic results was performed with in-house software (Diffprot) using the following parameters: data normalization – LOWESS, clustering of peptide sets with over 80% similarity, statistics – Wilcoxon test (28). To ensure that individual variability between biological replicates is not a major contributor to quantitative differences between nucleolar- and nuclear data sets, boxplot analysis of interquartile range and principal component analysis of quantitative data from individual samples were performed (Fig.

S1). In addition, the Diffprot algorithm penalizes quantitative inconsistencies in control samples, which further reduces the risk of obtaining artifacts. Statistical analysis of ribosome content and protein synthesis activity was performed using one-way ANOVA that was followed by Fisher's LSD posthoc tests.

## **Results**

*Isolation of brain nucleolar proteins.* To obtain the nucleolus-enriched protein fraction, whole nuclei were first isolated from the neocortex of P7 rats (Fig. 1). Then, nucleoli were obtained from the whole nuclei by sonication and ultracentrifugation in sucrose solution (Fig. 1). Relative purity of the isolated whole nuclei- and nucleoli was verified by phase contrast microscopy (Fig. 2A-B). Immunofluorescence for the nucleolar markers fibrillarin and nucleophosmin/B23 (Npm) confirmed successful isolation of the nucleoli (Fig. 2C). Sixty one % of the nuclei from which nucleoli were isolated were of neuronal origin as determined by the presence of the neuronal nuclear marker NeuN (Fig. 2D-E). Importantly, as in P7 cerebral cortex nucleolar presence of Npm is disrupted by stressors that interfere with rRNA transcription (41), maintenance of the nucleolar Npm staining throughout the nucleolar isolation procedure suggest well preserved integrity of the isolated nucleoli (Fig. 2C, E).

Next, protein lysates were prepared and relative nucleolar enrichment was evaluated by western blotting for several nucleolar marker proteins including Npm, fibrillarin, and upstream binding factor (UBF) (Fig. 3A-B). As compared to the nuclear lysates, all the nucleolar markers were enriched in the nucleolar fractions. In addition, relative purity of the nucleolar fraction was further verified by western blotting for markers of the cytoskeleton (actin, tubulin), the



mitochondrion (Tom20), and, the endoplasmic reticulum (calreticulin). In nucleolar fractions, their fibrillar-in-normalized signals were 14-50 times lower than in whole cortical lysates (Fig. 3C). Thus, enriched nucleolar protein samples can be obtained from the cerebral cortex of neonate rats.

*Proteomic analysis of nucleoli from the developing cerebral cortex.* Quantitative proteomics using the iTRAQ labeling method was applied to compare samples of intact nuclei and isolated nucleoli (Fig. 4A). Three sets of such samples, each prepared independently from different litters of neonate rats, were evaluated. Altogether, 3997 proteins were identified (Tables S1 and S5). Six hundred forty three proteins were significantly enriched in nucleoli and 487 proteins were significantly depleted from this compartment (Table S2, Fig. 4B-C). The gene ontology (GO) term “nucleolus” was associated with 12.6% of all nucleolus-enriched proteins (Fig. 4D). That fraction increased to 22.9 and 38.3% when the analysis was limited to nucleolar proteins with at least 1.5- and 2-fold enrichment, respectively (Fig. 4D).

Next an artificial neural network-based classifier algorithm was built to investigate whether apparent but non-significant nucleolar enrichment/depletion of some proteins may represent real differences in their subnuclear distribution (see Experimental Procedures for more details). Using such an approach, an additional 114 proteins were identified as likely candidates for differential nucleolar distribution (Table S3). This set included 102 proteins with apparent Nucleolus-enrichment; 75 of them were enriched at least 1.3 fold.

*Gene ontology analysis of nucleolar proteome from the developing cerebral cortex.* For Gene Ontology (GO) analysis, a list of candidate nucleolar proteins with a minimum 1.3 fold nucleolar-enrichment was prepared by combining significantly enriched proteins (437 of 643) with those whose nucleolar enrichment trend was supported by the classifier (75 of 102). For simplicity, proteins from this list will be further referred to as nucleolar.

GO terms associated with the cellular compartment (GO\_CC) were assigned to 433 nucleolar proteins with several GO\_CCs being significantly enriched (FDR  $p < 0.05$ , Table 1, Fig. 4E). The GO\_CC “nucleolus” was assigned to 17% proteins while 9% was associated with “ribosome”/“ribosomal subunit”. In addition, 7% proteins were assigned to chromatin-related GO\_CCs and 3% were splicosomal. Of note, with more than 22% coverage, the GO\_CC “mitochondrion” had the largest presence among nucleolar proteins of all enriched GO\_CCs that were organelle/structure specific. Large representation of mitochondrial proteins can be also found in other nucleolar proteome data sets that were obtained using similar fractionation methodology (21). Indeed, this mitochondrial enrichment appears to be a consequence of contamination whose origins are in the whole nuclei fraction (43). In support of that notion, similar fold enrichment of mitochondrial proteins was present in nucleolar- and nuclear samples from the brain (2.47-, and 1.89 fold, respectively, Table 1). Contamination may also account for a smaller, yet significantly enriched group of extracellular matrix/basal membrane components (3%, Table 1). Other enriched GO\_CC terms were unrelated to a specific subcellular organelle/structure and largely overlapped with the aforementioned specific GO\_CCs (*e.g.* “ribonucleoprotein complex”, “non-membrane-bounded organelle”, data not shown). Of note, 30% nucleolar proteins did not associate with any of the enriched GO\_CC terms.

Gene ontology terms for Biological Process (GO\_BP) were assigned to 411 nucleolar proteins. The greatest enrichment was observed for categories related to RNA metabolism including “ribosome biogenesis”, “ribonucleoprotein complex biogenesis”, “RNA processing” and several other similar terms which all together spanned 17% of the entire nucleolar protein list (Table 2, Fig. 4F). “Translation” was another highly enriched GO\_BP term followed by several GO\_BPs that were associated with chromatin organization (*e.g.* “chromosome organization”, “chromatin assembly or disassembly”, “chromatin organization”, “DNA packaging” “nucleosome assembly”, “nucleosome organization”). Their contribution to the nucleolar proteome was 8 and 9%, respectively. Ten % of the identified proteins were associated with mitochondrial GO\_BPs (*e.g.* “cellular respiration”, “generation of precursor metabolites and energy”, “electron transport chain” “oxidation reduction” “acetyl-CoA metabolic process”, “tricarboxylic acid cycle” and other related terms, Fig. 3C and data not shown). Finally, 53% of nucleolar proteins were not associated with any significantly enriched GO\_BP.

As expected for the center of ribosomal biogenesis, the most Nucleolus-enriched gene ontology terms of molecular function (GO\_MF) were associated with rRNA binding and ribosome structure (Table 3). In addition other enriched GO\_MFs included helicases, RNA binding proteins, ATPases and nucleotide binding proteins. Taken together, GO analysis supports functional specialization of brain cell nucleoli confirming their involvement in ribosome synthesis.

*Ribosomal proteins of the developing cerebral cortex.* Ribosome diversity has been previously proposed as a possible mechanism of translational regulation that may be active in the nervous system (44). Such a ribosomal diversity may originate from differential use of ribosomal proteins

(RPs), as demonstrated for RPL38 (45). In addition, mutations of some ribosomal proteins including S3, S4, S15A and L10 have relatively selective impact on forebrain development (46, 47). In the current data set, 75 nuclear ribosomal proteins were detected (Fig. S2). Of them, 51 were Nucleolus-enriched (Fig. S2). The abundant nucleolar presence of S3, S4, S15A and L10 suggests that these RPs are included into newly made brain ribosomes. Interestingly, two pairs of RP paralogs were also detected (L7/L711 and L22/L2211). Thus, differential use of these paralogs as well as inclusion of S3, S4, S15 and L10 may support ribosomal specialization in the developing cerebral cortex.

*Human neurological phenotypes associated with nucleolar proteins of the cerebral cortex.*

ToppFun analysis of proteins with a minimum 1.3 fold nucleolus-enrichment revealed significant over-representation of such overlapping human phenotype (HP) terms as “Abnormality of the cerebellum”, “abnormality of hindbrain morphology”, “abnormality of brain morphology” (FDR  $p < 0.05$ ). Of 69 proteins that were in these categories, 48 were associated with GO\_CCs “mitochondrion” and “ECM part” suggesting that they represent contaminations rather than *bona fide* nucleolar proteins. Three additional potential contaminants were two lysosomal hydrolases (NAGA, GLB1) and an ER glycosylase (STT3A). The remaining proteins included 14 whose mutations result in neurodevelopmental syndromes (Table 4). Two additional nucleolar proteins that are linked to neurodevelopmental disorders were identified by analyzing literature (RPL10 and EIF4A3, Table 4). Adult onset neurodegeneration was associated with mutations of three candidate nucleolar proteins (Table 5).

Microcephaly and/or cognitive impairment were the most common elements of the neurodevelopmental phenotypes (14 of 16). The target proteins for these phenotypes included a ribosomal protein (RPL10), four established ribosomal biogenesis factors (EMG1, DKC1, EIF4A3, FLNA), and seven chromatin organization/genome maintenance regulators (SMC1A, MCM4, CUL4B, PRKDC, ATRX, NSD1, LMNA). Moreover, neurodevelopmental phenotypes were associated with several nucleolus-enriched proteins that did not have any demonstrated connections to ribosomal biogenesis and/or chromatin structure. That group included transcription regulators (LARP7, TCF4, and GTF2I, Table 4) and an RNA editing enzyme (ADAR, Table 4).

*Verification of nucleolar enrichment of LARP7 and TCF4.* Previously published proteomic studies of human cell lines of various origins did not report LARP7 or TCF4 to be nucleolar (24, 48, 49). Moreover, immunofluorescence analysis of several other human cell lines did not support their nucleolar localization either (<http://www.proteinatlas.org/>). Therefore, in the developing rat brain, the apparent nucleolar enrichment of LARP7 and TCF4 were further examined using specific antibodies. First, western blot analysis of the nucleolar- and the whole nuclear lysates of the P7 cerebral cortex confirmed nucleolar enrichment of either protein (Fig. 4G). In addition, nucleolar enrichment of the established ribosomal biogenesis factor EMG1 was also confirmed by western blotting (Fig. 4G).

Second, immunofluorescence analysis of P7 rat brain sections revealed a strong nucleolar-like signal that was obtained with the anti-LARP7 antibody (Fig. 5A-F). In the neocortex and the hippocampus, nucleolar presence of LARP7 was observed in neurons but not in non-neuronal cells as determined by co-staining for the neuronal marker NeuN (Fig. 5A, E,

data not shown). Its nucleolar character was also confirmed by co-stainings for the nucleolar markers Npm and UBF (Fig. 5B, C, F). In addition, a diffused LARP7 signal was observed in the nucleoplasm (Fig. 5A-F). A similar pattern of LARP7 immunofluorescence was found in other analyzed structures including the thalamus (data not shown).

Also, cultured rat neurons displayed nucleolar enrichment of LARP7 (Fig. 5G,H). However, as compared to cells from brain sections, stronger nuclear signal was observed in cultured cells. Conversely, in cultured rat primary astrocytes, no nucleolar enrichment was observed despite a strong nuclear signal of LARP7 (Fig. 5I). Therefore, nucleolar accumulation of LARP7 appears to be neuron-specific and may be sensitive to culture conditions.

When cortical neurons were imaged at high magnification, diffused- and fine granular staining of LARP7 was visible throughout the nucleus (Fig. 5J). Such a distribution is consistent with a previously proposed role of LARP7 in regulation of RNA-Polymerase-2 (50-52). In the nucleolus, LARP7 signal only partially overlapped with that of Npm, suggesting its absence from the Npm-positive granular component of the nucleolus (GC) (Fig. 5J) (1). Conversely, LARP7 appeared to co-localize with fibrillarin which marks the dense fibrillar component (DFC) of the nucleolus where rRNA genes are transcribed and where their primary transcripts undergoes initial processing (Fig. 5J) (1).

The transcriptional inhibitor Actinomycin D (ActD) led to nucleoplasmic release of Npm but not LARP7 suggesting that rRNA-protein interactions are dispensable for its nucleolar localization (Fig. 5H). In ActD-treated neurons, LARP and fibrillarin separated into apparently overlapping structures that were highly similar to nucleolar caps (Fig. 5J) (53). Hence, the

LARP7-containing caps resemble the previously reported light nucleolar caps (LNC) which are also known to sequester several snoRNAs (53).

Unfortunately, in brain section, immunofluorescence analysis with the anti-TCF4 antibody was not conclusive due to a very bright staining of the blood vessels and the extracellular matrix (data not shown). These signals which likely represent non-specific staining prevented conclusive observations of any staining from inside the cells. However, a specific signal was obtained in cultured primary rat brain cells suggesting perikaryal and nuclear presence of TCF4 (Fig. 6). In cortical- and hippocampal neurons as well as astrocytes, the nuclear signal was concentrated in the nucleolus (Fig. 6A-C). Besides nucleolus, both a diffuse- and a fine granular signal appeared throughout the nucleoplasm and in the perikarya (Fig. 6A-D). Hence, only a relatively minor fraction of cellular TCF4 may reside in the nucleolus.

In nucleoli of cortical neurons, TCF4 overlapped nearly completely with fibrillarin but only partially with Npm or UBF (Fig. 6D). As in the case of LARP7, ActD did not trigger nucleoplasmic release of TCF4 which, instead, appeared sequestered in nucleolar cap-like structures (Fig. 6D). However, the TCF4-containing caps were fibrillarin-negative but UBF-positive suggesting they were the fibrillar caps that sequester various component of the Pol1 machinery including UBF (53). Thus, at least in rat cortical neurons, LARP7 and TCF4 appear to be concentrated in the DFC compartment of the nucleolus and undergo nucleolar cap sequestration after inhibition of Pol1. However, they segregate into distinct types of caps which may suggest different roles at the early stages of ribosomal biogenesis.

*Requirement of LARP7 for ribosomal biogenesis.* To test the role of LARP7 in ribosomal biogenesis, three different shRNAs against rat LARP7 were designed. When overexpressed in hippocampal neurons, these constructs reduced expression of endogenous LARP7 (Fig. 7A). Importantly, in transfected neurons, shRNAs against LARP7 prevented nucleolar staining with the LARP7 antibody further confirming specificity of the nucleolar signal (Fig 7A). Moreover, each of these constructs reduced perikaryal ribosome content by 22-40% as compared to a control shRNA (Fig. 7B-C). As reduction in ribosome content is expected to compromise protein synthesis activity, perikaryal accumulation of nascent peptides was monitored in LARP7-deficient neurons using co-translational labeling with puromycin (Fig. 7D-E). Indeed, shLAR7 reduced general protein synthesis. However, the observed declines by 50-65% were greater than those in ribosome content (Fig. 7D-E). The effects of shLARP7 were compared to those of knocking down EMG1. EMG1 deficiency is a cause of Bowen-Conradi syndrome that combines neurodevelopmental delay and microcephaly. Two different shRNAs against EMG1 reduced ribosome content and protein synthesis by 32-40% and 30-36%, respectively (Fig. 8). Thus, knockdowns of LARP7 and EMG1 produce similar anti-ribosomal effects. However, translational inhibition was stronger with shRNAs against LARP7. Therefore, at least in hippocampal neurons LARP7 is required for ribosomal biogenesis. It may also have additional functions in regulation of general protein synthesis.

*Potential role of TCF4 in ribosomal biogenesis.* To achieve loss-of-function of TCF4, two mutant forms of human TCF4 were overexpressed in cultured cortical neurons (Fig. 9A). These mutants (R580W and R582P ) disrupt the bHLH domain producing dominant-negative effects against wtTCF4 and the related members of the E-box family (35). Moreover, these mutants



cause a neurodevelopmental disease, PTHS which may also result from TCF4 haploinsufficiency. After overexpressing the mutants for 3 days, perikaryal ribosome content was reduced by 20% or 15% as compared to empty vector (EV) or wtTCF4, respectively (Fig. 9B-C). Similar reductions of general protein synthesis were also observed (Fig. 9D-E). These findings suggest that in hippocampal neurons TCF4 and/or its relatives from the E-box family of transcription factors play a role in ribosomal biogenesis.

## **Discussion**

Quantitative proteomic analysis uncovered several hundred proteins that are enriched in nucleoli of the cerebral cortex of neonate rats. Due to limitations of the applied methodology, it is likely that in many cases the nucleolar enrichment was underestimated. Thus, protein and peptide ratios derived from the multiplexed quantification (iTRAQ or TMT) of highly complicated mixtures are known to be distorted due to the interference from the co-fragmenting ions. High fold changes tend to be substantially underestimated, and, in consequence, the distribution of peptide/protein ratios is artificially narrowed (54). This effect might be dealt with using several methods. They are, however, either moderately effective (computing isolation specificity, fractionation) or costly in terms of protein identification efficiency (12-22% reduction of non-redundant protein identifications when isolation window is narrowed or additional MS3 fragmentation is employed). Increasing the accuracy of protein ratio at the expense of proteome coverage might be in some cases a prudent solution. However, this study was more focused on the enrichment/depletion status of proteins than on actual protein ratio and was also aiming at possibly comprehensive protein sequencing. Therefore the iTRAQ ratio compression was considered acceptable, and the assessment of accurate protein ratios was left

for the follow-up experiments. However, it is possible that the currently used proteomic approach missed a significant number of proteins that are enriched in brain cell nucleoli.

Despite these limitations, the emerging nucleolar proteome landscape of the cerebral cortex of neonate rats suggests that ribosomal biogenesis remains active during postnatal stages of brain development. In addition, the importance of nucleolar contributions to neurodevelopment is suggested by human phenotypes that are associated with 16 proteins that are enriched in cerebro-cortical nucleoli including the ribosomal biogenesis factor EMG1, the RNA binding protein LARP7, and, the transcription factor TCF4.

Surprisingly, only 106 of 286 human ribosomal biogenesis factors were represented by rodent homologs among all Nucleolus-enriched proteins from rat brain (including 643 proteins with significant Nucleolus-enrichment and 102 proteins with Nucleolus-enrichment trends that were supported by the Nucleolus-classifier). Such observations suggest a unique organization of the ribosomal biogenesis pathway in brain cells, including neurons. That concept is consistent with a report that in mouse brain there are different species of rRNA processing intermediates than in mouse kidney (55). In addition, as brain ribosomal production is relatively rapidly down-regulated during the maturation phase of neurodevelopment (56-58) (Slomnicki and Hetman, unpublished results), lesser number of ribosomal biogenesis factors may be required to satisfy reduced needs for new ribosomes. In contrast, in proliferating cells, many additional enzymes/regulators may be necessary to maximize rate of ribosome synthesis.

Overall, the overlap between nucleolar proteomes of the brain and proliferating cells was relatively limited. For instance, when the list of putative brain nucleolar proteins with at least 1.3-fold No/Nu enrichment was compared to 344 such proteins from human HCT116 colon carcinoma cells (24), only 114 proteins were found in both data sets. On that short list, the overrepresented GO\_BPs were mostly related to ribosomal biogenesis and RNA metabolism.

In addition, the nucleolus-enriched brain proteome was compared to non-quantitative data sets of 726 or 873 nucleolar proteins from human cervical carcinoma HeLa- or human T-cell lymphoma Jurkat cells, respectively. The overlaps of 192 or 189 proteins were observed for each comparison (48, 49). The most highly enriched GO\_BPs among the shared proteins were related to ribosomal biogenesis and RNA metabolism. Enrichment was also found for GO\_BPs that are related to chromatin organization. When 263 brain-unique components of the nucleolar proteome (*e.g.* not present in either HeLa or Jurkat data sets) were analyzed for GO\_BP enrichment, chromatin organization- but not ribosomal biogenesis remained overrepresented. The GO\_BP term “chromosome organization” was associated with 21 proteins that were found in brain nucleoli but not in any other of the three nucleolar data sets from human cell lines (Table S4). Overall, it appears that nucleoli of mostly post-mitotic brain cells and proliferating cells share various proteins of the ribosomal biogenesis pathway and chromatin organization machinery. However, in brain cell nucleoli, relative representation of ribosomal biogenesis decreases. Conversely, the contribution of chromatin organization increases, including the appearance of brain nucleolus-specific chromatin components and modifiers. Therefore, brain cell nucleoli may participate not only in producing ribosomes but also in regulating chromatin structure. Alternatively, nucleolar chromatin structure may be more regulated and/or have

relatively greater modulatory effect on brain cell ribosome production. Finally, the previously demonstrated nucleolar association of specifically regulated chromatin domains may be greater in brain cells increasing co-purification of chromatin organization proteins together with the nucleoli (59, 60).

Several putative nucleolar proteins from the developing brain are associated with human neurodevelopmental diseases including three established ribosomal biogenesis factors and a ribosomal protein (Table 4). A question that emerges from such an association is whether the nucleolus contributes to neurodevelopment? Rapidly proliferating cells are considered to be most sensitive to deficiencies in ribosome biogenesis (61, 62). Hence, neuroprogenitor proliferation could be impaired if ribosome production declines. In addition, the nucleolar/ribosomal stress response could initiate p53-dependent apoptosis eliminating neuroprogenitors with suboptimal ribosome production (13). Such a defective neurogenesis could account for microcephaly that is observed in 9 out of 16 syndromes that are associated with mutations of the putative nucleolar brain proteins (Table 4).

In addition, neuronal maturation is associated with intense neurite outgrowth that involves ribosomal biogenesis (10)(Slomnicki and Hetman, unpublished results). In fact, on postnatal day 7, when the nucleolar fractions of the cerebral cortex were obtained, there is neuritic growth and maturation but little neurogenesis in that region of the rodent brain (63). Therefore, it is tempting to speculate that ribosome depletion and consequent dysregulation of neuronal maturation including reduced neuritic morphogenesis contribute to at least some neurodevelopmental diseases that are associated with the nucleolar proteins.

The newly identified components of neuronal nucleolar proteome include the RNA binding protein LARP7 and the E-box transcription factor TCF4 whose deficiencies in humans are associated with cognitive impairment and microcephaly (64, 65). By binding and stabilizing a small non-coding RNA, 7SK RNA, LARP7 negatively regulates the general Pol2 transcription elongation factor P-TEFb as well as the transcription factor HMGA1 (50-52). However, it is unclear whether dysregulation of these transcription factors contributes to pathology of Alazami syndrome that besides severe cognitive impairment also includes dwarfism (64). In fact, such a pathogenic contribution appears counterintuitive as excessive activity of P-TEFb or HMGA1 may underlie increased proliferative growth following transformation (see discussion in (64)). Our findings that neuronal LARP7 is required for ribosomal biogenesis suggest that, at least, neurological symptoms of Alazami syndrome are caused by translational deficits and/or nucleolar stress.

Similarly to Alazami syndrome, little is known about pathogenic mechanisms that underlie PTHS. The disease is caused by heterozygous deletions or point mutations disrupting the function of TCF4 by a loss of one active allele, generation of hypomorphic alleles or dominant negative effects (35). Our data indicate that the PTHS-associated dominant-negative mutants of TCF4 moderately reduce ribosomal biogenesis and general protein synthesis. While future experiments are needed to clarify role of TCF4 in neuronal ribosome biogenesis, it is noteworthy that ribosomal proteins as well as many trans-acting factors of ribosome synthesis and nucleolus-associated proteins were shown to have their gene regulatory regions bound to TCF4 (66). In addition, the GO\_BP term “ribosomal biogenesis” appeared to be overrepresented

among genes that were down-regulated after TCF4 knockdown in human neuroblastoma cells (67). In proliferating cells, many of these genes, as well as rDNA, are regulated by E-boxes that interact with c-Myc (68, 69). Conversely, TCF4 and/or a TCF4-related transcription factors may regulate ribosomal biogenesis by engaging the same E-boxes in postmitotic and/or differentiating cells.

Putative proteins of the nucleolar brain proteome include 7 chromatin organization factors that are also associated with neurodevelopmental diseases. At least for one of such factors, the cohesin SMC1A, deficient rRNA transcription has been proposed as a mechanism that could explain cohesinopathy pathology in non-dividing cells including neurons (70, 71). Of note, two other cohesinopathy-associated proteins including NIPBL and SMC3 were also enriched in cerebro-cortical nucleoli, albeit less than 1.3 fold (OMIM #122470 and #610759, Table S2). Critical role of chromatin organization factors in rRNA transcription is not limited to cohesins but has been also reported for the chromatin remodeling complexes B-WICH and FACT whose various components are enriched in brain cell nucleoli(72, 73) (Table S2). Future studies are needed to investigate the role of nucleolus-enriched chromatin organizers in ribosome synthesis in the brain.

Taken together, the presented proteomic characterization of the nucleoli from the cerebral cortex of neonate rats supports the notion that the nucleolus is critical for neurodevelopment. Ribosomal biogenesis appears as top candidate function of the nucleolus through which it may affect formation and maturation of the nervous system. Disruption of ribosomal biogenesis by genetic and/or environmental factors may contribute to pathogenesis of various neurodevelopmental disorders. Finally, the current description of the nucleolar proteome offers a

starting point to examine brain cell specific regulation of ribosomal synthesis, brain ribosome diversity, and nucleolar impact on neuronal chromatin organization and genomic maintenance.

## References

1. Sirri, V., Urcuqui-Inchima, S., Roussel, P., and Hernandez-Verdun, D. (2008) Nucleolus: the fascinating nuclear body. *Histochem Cell Biol* 129, 13-31
2. Drygin, D., Rice, W. G., and Grummt, I. (2010) The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu Rev Pharmacol Toxicol* 50, 131-156
3. Freed, E. F., Bleichert, F., Dutca, L. M., and Baserga, S. J. (2010) When ribosomes go bad: diseases of ribosome biogenesis. *Mol Biosyst* 6, 481-493
4. Kressler, D., Hurt, E., and Bassler, J. (2010) Driving ribosome assembly. *Biochimica et biophysica acta* 1803, 673-683
5. Stoykova, A. S., Dudov, K. P., Dabeva, M. D., and Hadjiolov, A. A. (1983) Different rates of synthesis and turnover of ribosomal RNA in rat brain and liver. *J Neurochem.* 41, 942-949.
6. Hetman, M., and Pietrzak, M. (2012) Emerging roles of the neuronal nucleolus. *Trends Neurosci* 35, 305-314
7. Lafarga, M., Villegas, J., and Crespo, D. (1985) Changes in nucleolar morphology and volume of the supraoptic nucleus neurons during postnatal development of the rat. *Brain research* 354, 310-313

8. Clark, P., Jones, K. J., and LaVelle, A. (1990) Ultrastructural and morphometric analysis of nucleolar and nuclear changes during the early growth period in hamster facial neurons. *J Comp Neurol* 302, 749-760
9. Berciano, M. T., Novell, M., Villagra, N. T., Casafont, I., Bengoechea, R., Val-Bernal, J. F., and Lafarga, M. (2007) Cajal body number and nucleolar size correlate with the cell body mass in human sensory ganglia neurons. *J Struct Biol.* 158, 410-420. Epub 2006 Dec 2028.
10. Gomes, C., Smith, S. C., Youssef, M. N., Zheng, J. J., Hagg, T., and Hetman, M. (2011) RNA polymerase 1-driven transcription as a mediator of BDNF-induced neurite outgrowth. *J Biol Chem* 286, 4357-4363
11. Boulon, S., Westman, B. J., Hutten, S., Boisvert, F. M., and Lamond, A. I. (2010) The nucleolus under stress. *Molecular cell* 40, 216-227
12. Kalita, K., Makonchuk, D., Gomes, C., Zheng, J. J., and Hetman, M. (2008) Inhibition of nucleolar transcription as a trigger for neuronal apoptosis. *Journal of neurochemistry* 105, 2286-2299
13. Parlato, R., Kreiner, G., Erdmann, G., Rieker, C., Stotz, S., Savenkova, E., Berger, S., Grummt, I., and Schutz, G. (2008) Activation of an endogenous suicide response after perturbation of rRNA synthesis leads to neurodegeneration in mice. *J Neurosci.* 28, 12759-12764.
14. Pederson, T., and Tsai, R. Y. (2009) In search of nonribosomal nucleolar protein function and regulation. *The Journal of cell biology* 184, 771-776
15. Boyd, M. T., Vlatkovic, N., and Rubbi, C. P. (2011) The nucleolus directly regulates p53 export and degradation. *The Journal of cell biology* 194, 689-703



16. Latonen, L., Moore, H. M., Bai, B., Jaamaa, S., and Laiho, M. (2011) Proteasome inhibitors induce nucleolar aggregation of proteasome target proteins and polyadenylated RNA by altering ubiquitin availability. *Oncogene* 30, 790-805
17. Guetg, C., and Santoro, R. (2012) Formation of nuclear heterochromatin: the nucleolar point of view. *Epigenetics* 7, 811-814
18. Andersen, J. S., Lyon, C. E., Fox, A. H., Leung, A. K., Lam, Y. W., Steen, H., Mann, M., and Lamond, A. I. (2002) Directed proteomic analysis of the human nucleolus. *Curr Biol* 12, 1-11
19. Scherl, A., Coute, Y., Deon, C., Calle, A., Kindbeiter, K., Sanchez, J. C., Greco, A., Hochstrasser, D., and Diaz, J. J. (2002) Functional proteomic analysis of human nucleolus. *Mol Biol Cell* 13, 4100-4109
20. Coute, Y., Burgess, J. A., Diaz, J. J., Chichester, C., Lisacek, F., Greco, A., and Sanchez, J. C. (2006) Deciphering the human nucleolar proteome. *Mass spectrometry reviews* 25, 215-234
21. Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005) Nucleolar proteome dynamics. *Nature*. 433, 77-83.
22. Tafforeau, L., Zorbas, C., Langhendries, J. L., Mullineux, S. T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D. L. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Molecular cell* 51, 539-551
23. Ahmad, Y., Boisvert, F. M., Gregor, P., Cobley, A., and Lamond, A. I. (2009) NOPdb: Nucleolar Proteome Database--2008 update. *Nucleic Acids Res* 37, D181-184

24. Boisvert, F. M., Lam, Y. W., Lamont, D., and Lamond, A. I. (2010) A quantitative proteomics analysis of subcellular proteome localization and changes induced by DNA damage. *Mol Cell Proteomics* 9, 457-470
25. Rieker, C., Engblom, D., Kreiner, G., Domanskyi, A., Schober, A., Stotz, S., Neumann, M., Yuan, X., Grummt, I., Schutz, G., and Parlato, R. (2011) Nucleolar disruption in dopaminergic neurons leads to oxidative damage and parkinsonism through repression of mammalian target of rapamycin signaling. *J Neurosci* 31, 453-460
26. Kreiner, G., Bierhoff, H., Armentano, M., Rodriguez-Parkitna, J., Sowodniok, K., Naranjo, J. R., Bonfanti, L., Liss, B., Schutz, G., Grummt, I., and Parlato, R. (2013) A neuroprotective phase precedes striatal degeneration upon nucleolar stress. *Cell Death Differ* 20, 1455-1464
27. Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nature methods* 6, 359-362
28. Malinowska, A., Kistowski, M., Bakun, M., Rubel, T., Tkaczyk, M., Mierzejewska, J., and Dadlez, M. (2012) Diffprot - software for non-parametric statistical analysis of differential proteomics data. *Journal of proteomics* 75, 4062-4073
29. Elias, J. E., Haas, W., Faherty, B. K., and Gygi, S. P. (2005) Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nature methods* 2, 667-675
30. Hall, M., Frank, E., Holmes, G., Pfaringer, B., Reutermann, P., and Witten, I. H. (2009) The WEKA Data Mining Software: An Update. *SIGKDD Explorations* 11
31. Scott, M. S., Troshin, P. V., and Barton, G. J. (2011) NoD: a Nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC bioinformatics* 12, 317

32. Sutherland, H. G., Mumford, G. K., Newton, K., Ford, L. V., Farrall, R., Dellaire, G., Caceres, J. F., and Bickmore, W. A. (2001) Large-scale identification of mammalian proteins localized to nuclear sub-compartments. *Hum Mol Genet* 10, 1995-2011
33. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37, 1-13
34. Chen, J., Bardes, E. E., Aronow, B. J., and Jegga, A. G. (2009) ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 37, W305-311
35. Sepp, M., Pruunsild, P., and Timmusk, T. (2012) Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum Mol Genet* 21, 2873-2888
36. Jaworski, J., Spangler, S., Seeburg, D. P., Hoogenraad, C. C., and Sheng, M. (2005) Control of dendritic arborization by the phosphoinositide-3'-kinase-Akt-mammalian target of rapamycin pathway. *J Neurosci.* 25, 11300-11312.
37. Kalita, K., Kharebava, G., Zheng, J. J., and Hetman, M. (2006) Role of megakaryoblastic acute leukemia-1 in ERK1/2-dependent stimulation of serum response factor-driven transcription by BDNF or increased synaptic activity. *J Neurosci.* 26, 10020-10032.
38. Habas, A., Kharebava, G., Szatmari, E., and Hetman, M. (2006) NMDA neuroprotection against a phosphatidylinositol-3 kinase inhibitor, LY294002 by NR2B-mediated suppression of glycogen synthase kinase-3beta-induced apoptosis. *Journal of neurochemistry* 96, 335-348
39. Hetman, M., Hsuan, S. L., Habas, A., Higgins, M. J., and Xia, Z. (2002) ERK1/2 Antagonizes Glycogen Synthase Kinase-3beta -induced Apoptosis in Cortical Neurons. *J Biol Chem* 277, 49577-49584.

40. Yang, P., Arnold, S. A., Habas, A., Hetman, M., and Hagg, T. (2008) Ciliary neurotrophic factor mediates dopamine D2 receptor-induced CNS neurogenesis in adult mice. *J Neurosci* 28, 2231-2241
41. Pietrzak, M., Smith, S. C., Geraldts, J. T., Hagg, T., Gomes, C., and Hetman, M. (2011) Nucleolar disruption and apoptosis are distinct neuronal responses to etoposide-induced DNA damage. *Journal of neurochemistry* 117, 1033-1046
42. Liu, J., Xu, Y., Stoleru, D., and Salic, A. (2012) Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proceedings of the National Academy of Sciences of the United States of America* 109, 413-418
43. Liang, Y. M., Wang, X., Ramalingam, R., So, K. Y., Lam, Y. W., and Li, Z. F. (2012) Novel nucleolar isolation method reveals rapid response of human nucleolar proteomes to serum stimulation. *Journal of proteomics* 77, 521-530
44. Holt, C. E., and Schuman, E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* 80, 648-657
45. Kondrashov, N., Pusic, A., Stumpf, C. R., Shimizu, K., Hsieh, A. C., Xue, S., Ishijima, J., Shiroishi, T., and Barna, M. (2011) Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* 145, 383-397
46. Uechi, T., Nakajima, Y., Nakao, A., Torihara, H., Chakraborty, A., Inoue, K., and Kenmochi, N. (2006) Ribosomal protein gene knockdown causes developmental defects in zebrafish. *PLoS One* 1, e37
47. Brooks, S. S., Wall, A. L., Golzio, C., Reid, D. W., Kondyles, A., Willer, J. R., Botti, C., Nicchitta, C. V., Katsanis, N., and Davis, E. E. (2014) A novel ribosomopathy caused by

dysfunction of RPL10 disrupts neurodevelopment and causes X-linked microcephaly in humans. *Genetics* 198, 723-733

48. Leung, A. K., Trinkle-Mulcahy, L., Lam, Y. W., Andersen, J. S., Mann, M., and Lamond, A. I. (2006) NOPdb: Nucleolar Proteome Database. *Nucleic Acids Res.* 34, D218-220.

49. Jarboui, M. A., Wynne, K., Elia, G., Hall, W. W., and Gautier, V. W. (2011) Proteomic profiling of the human T-cell nucleolus. *Mol Immunol* 49, 441-452

50. Krueger, B. J., Jeronimo, C., Roy, B. B., Bouchard, A., Barrandon, C., Byers, S. A., Searcey, C. E., Cooper, J. J., Bensaude, O., Cohen, E. A., Coulombe, B., and Price, D. H. (2008) LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res* 36, 2219-2229

51. He, N., Jahchan, N. S., Hong, E., Li, Q., Bayfield, M. A., Maraia, R. J., Luo, K., and Zhou, Q. (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Molecular cell* 29, 588-599

52. Eilebrecht, S., Brysbaert, G., Wegert, T., Urlaub, H., Benecke, B. J., and Benecke, A. (2011) 7SK small nuclear RNA directly affects HMGA1 function in transcription regulation. *Nucleic Acids Res* 39, 2057-2072

53. Shav-Tal, Y., Blechman, J., Darzacq, X., Montagna, C., Dye, B. T., Patton, J. G., Singer, R. H., and Zipori, D. (2005) Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. *Mol Biol Cell* 16, 2395-2413

54. Ting, L., Rad, R., Gygi, S. P., and Haas, W. (2011) MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature methods* 8, 937-940

55. Akhmanova, A., Verkerk, T., Langeveld, A., Grosveld, F., and Galjart, N. (2000) Characterisation of transcriptionally active and inactive chromatin domains in neurons. *Journal of cell science* 113 Pt 24, 4463-4474
56. Adams, D. H. (1966) The relationship between cellular nucleic acids in the developing rat cerebral cortex. *Biochem J* 98, 636-640
57. Johnson, T. C. (1967) The effects of maturation on in vitro RNA synthesis by mouse brain cells. *Journal of neurochemistry* 14, 1075-1081
58. Berthold, W., and Lim, L. (1976) Nucleo-cytoplasmic relationships of high-molecular-weight ribonucleic acid, including polyadenylated species, in the developing rat brain. *Biochem J* 154, 529-539
59. van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G. J., Ariyurek, Y., den Dunnen, J. T., and Lamond, A. I. (2010) High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol Biol Cell* 21, 3735-3748
60. Nemeth, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Peterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Langst, G. (2010) Initial genomics of the human nucleolus. *PLoS Genet* 6, e1000889
61. Trainor, P. A., and Merrill, A. E. (2014) Ribosome biogenesis in skeletal development and the pathogenesis of skeletal disorders. *Biochim Biophys Acta* 1842, 769-778
62. Ellis, S. R. (2014) Nucleolar stress in Diamond Blackfan anemia pathophysiology. *Biochim Biophys Acta* 1842, 765-768
63. Bayer, S. A., and Altman, J. (1987) Directions in neurogenetic gradients and patterns of anatomical connections in the telencephalon. *Prog Neurobiol* 29, 57-106

64. Alazami, A. M., Al-Owain, M., Alzahrani, F., Shuaib, T., Al-Shamrani, H., Al-Falki, Y. H., Al-Qahtani, S. M., Alsheddi, T., Colak, D., and Alkuraya, F. S. (2012) Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability, and primordial dwarfism. *Hum Mutat* 33, 1429-1434
65. Sweatt, J. D. (2013) Pitt-Hopkins Syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. *Experimental & molecular medicine* 45, e21
66. Ghosh, H. S., Cisse, B., Bunin, A., Lewis, K. L., and Reizis, B. (2010) Continuous expression of the transcription factor e2-2 maintains the cell fate of mature plasmacytoid dendritic cells. *Immunity* 33, 905-916
67. Forrest, M. P., Waite, A. J., Martin-Rendon, E., and Blake, D. J. (2013) Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS One* 8, e73169
68. Grandori, C., Gomez-Roman, N., Felton-Edkins, Z. A., Ngouenet, C., Galloway, D. A., Eisenman, R. N., and White, R. J. (2005) c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nature cell biology* 7, 311-318
69. Poortinga, G., Wall, M., Sanij, E., Siwicki, K., Ellul, J., Brown, D., Holloway, T. P., Hannan, R. D., and McArthur, G. A. (2011) c-MYC coordinately regulates ribosomal gene chromatin remodeling and Pol I availability during granulocyte differentiation. *Nucleic Acids Res* 39, 3267-3281
70. Xu, B., Lee, K. K., Zhang, L., and Gerton, J. L. (2013) Stimulation of mTORC1 with L-leucine rescues defects associated with Roberts syndrome. *PLoS Genet* 9, e1003857

71. Xu, B., Sowa, N., Cardenas, M. E., and Gerton, J. L. (2014) l-leucine partially rescues translational and developmental defects associated with zebrafish models of Cornelia de Lange syndrome. *Hum Mol Genet*
72. Tsai, Y. C., Greco, T. M., Boonmee, A., Miteva, Y., and Cristea, I. M. (2012) Functional proteomics establishes the interaction of SIRT7 with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription. *Mol Cell Proteomics* 11, 60-76
73. Tessarz, P., Santos-Rosa, H., Robson, S. C., Sylvestersen, K. B., Nelson, C. J., Nielsen, M. L., and Kouzarides, T. (2014) Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* 505, 564-568



## **Footnotes**

*Acknowledgments.* The authors declare no competing financial interests. This work was supported by NIH (R01 NS073584 and P30 GM103507 to MH), NSF (IOS1021860 to MH), the Commonwealth of Kentucky Challenge for Excellence Fund (MH), Estonian Research Council (IUT19-18 to TT) and Estonian Academy of Sciences (TT). The authors wish to thank Dr. Marta Tkaczyk for advice on design of the initial proteomic experiments, and Dr. Scott R. Whittemore for critical reading of the manuscript. The equipment used was sponsored in part by the Centre for Preclinical Research and Technology (CePT), a project co-sponsored by European Regional Development Fund and Innovative Economy, The National Cohesion Strategy of Poland.

## Figure Legends

**Figure 1. Preparation of nucleolar proteins from the cerebral cortex of neonate rats.** The schematic presents the nucleolar isolation procedure. First, whole nuclei were isolated from the neocortex of P7 rats. After sonication, the released nucleoli were purified by ultracentrifugation; some intact nuclei were still present in such preparations (see Materials and Methods for more details).

**Figure 2. Morphology of isolated nuclei and nucleoli from the cerebral cortex of neonate rats. A-B.** Phase contrast micrographs of unfixed nuclei (arrows) and nucleoli (arrowheads). **C-D,** Isolated nucleoli (C) or nuclei (D-E) were fixed and immunostained as indicated; DNA was visualized by counterstaining with Hoechst-33258. In isolated nucleoli (C), or, in intact nuclei (D-E), nucleolar markers fibrillarin and nucleophosmin (Npm) remained nucleolar suggesting good preservation of the nucleolus throughout the fractionation procedure. In D, the isolated whole nuclei were co-immunostained for the neuronal nuclear marker NeuN and the Npm. The arrow points a NeuN-reactive (*i.e.* neuronal) nucleus; the arrowhead marks a non-neuronal nucleus. Note that in neuronal nuclei, nucleoli were usually larger than in non-neuronal nuclei.

**Figure 3. Enrichment of nucleolar marker proteins in isolated nucleoli from the cerebral cortex of neonate rats.** Enrichment of nucleolar proteins in nucleolar preparations (No) was verified by Western blotting. **A,** Note that nucleolar proteins including fibrillarin, Npm and UBF were more abundant in No lysates than in total nuclear lysate (Nu) as compared to a nuclear marker NeuN. **B,** Averages  $\pm$  SEM of such NeuN-normalized enrichments from 8 independent preparations are presented on the graph. **C,** Nucleolar enrichment was further evaluated by

western blotting for markers of other cellular components including microfilaments (actin), microtubules (tubulin), mitochondria (Tom20) and endoplasmic reticulum (calreticulin). When compared to total cortical- (TL), or, whole nuclear lysates (Nu), the fibrillar-in-normalized content of these potential contaminants was lowest in nucleolar lysates (No). Similar results were observed with lysates from two independent experiments/isolations. Band intensities were quantified by densitometry and normalized against intensity of NeuN-(B) or fibrillar-in (C) signals.

**Figure 4. Analysis of the nucleolar proteome from the developing rat cerebral cortex. A,** Three independent sets of total nuclear/nucleolar protein samples from the cerebral cortex of P7 rats were analyzed using quantitative iTRAQ proteomics. Filter Aided Sample Preparation (FASP) was applied to obtain peptides that were tagged using an 8-plex iTRAQ system for quantitative LC/MS analysis. **B,** The pie chart illustrates the numbers of proteins with differential abundance in the nucleolus (No). **C,** Distribution of relative No enrichments among 1,130 differentially localized proteins. Numbers over the bars correspond to numbers of proteins in each enrichment bin. **D,** Increased No enrichment correlates with greater fraction of proteins that are associated with the gene ontology (GO) term “nucleolus”. Numbers over the bars correspond to numbers of proteins that are passing each enrichment threshold. **E-F,** Gene ontology (GO) analysis of 504 mapped proteins whose nucleolar enrichment was at least 1.3 fold. The significance threshold for GO enrichment was  $FDR < 0.05$ . The pie charts present fractions of No proteins that were assigned to enriched GO\_CC- (cellular compartment, **E**) or GO\_BP (biological process, **F**) categories. Related terms were grouped together; overlapping GO assignments were avoided to accurately determine the fraction of unclassified proteins. More

details about GO enrichment are in Tables 1-3. **G**, Nucleolar enrichment of the candidate nucleolar proteins EMG1, LARP7 and TCF4 was verified by Western blotting. The established ribosomal biogenesis factor EMG1 was used as an additional control. Note that relative abundance of each of these three proteins is greater in nucleolar lysates as compared to whole nuclear lysates resembling nucleolar enrichment of fibrillarin. Similar trends were observed in two independent experiments/isolations.

**Figure 5. Nucleolar presence of the RNA-binding protein LARP7 in forebrain neurons. A-F**, LARP7 immunofluorescence in P7 rat brain sections that were cut through the dorsal hippocampus; co-stainings with neuronal- or nucleolar markers are shown as indicated; DNA was counterstained with Hoechst-33258; in D, LARP7 antibody was replaced by control rabbit IgG. Representative micrographs of cells from the neocortical layer V (A-D) and the hippocampal field CA1 (E-F, the pyramidal layer) are shown. Both nuclear and nucleolar signals of LARP7 are visible in neurons from each region. **G-J**, LARP7 Immunostaining in cultured primary rat brain cells including DIV6 cortical neurons (G, J), DIV7 hippocampal neurons (H) and DIV10 astrocytes (I). Note nucleolus-like staining in neurons but not astrocytes; both cell types display nuclear presence of LARP7. In **J**, representative nuclear profiles are shown following 6h treatment with 0.1% DMSO (vehicle) or 1  $\mu$ M ActinomycinD (ActD). In vehicle-treated cells, overlapping stainings of LARP7 and fibrillarin suggest LARP7 localization to the dense fibrillar component (DFC) of the nucleolus. After transcriptional inhibition, Npm but not LARP7 or fibrillarin were released into nucleoplasm; moreover, the persistent overlap of LARP7 and fibrillarin indicates co-localization to light nucleolar caps which sequester early rRNA processing factors including snoRNAs (see text for details).

**Figure 6. Nucleolar presence of the transcription factor TCF4 in primary cultures of forebrain neurons and astrocytes.** Cultured rat brain cells were immunostained for TCF4 and the nucleolar markers as indicated. **A**, Representative images depicting TCF4 immunofluorescence in DIV6 cortical neurons. Both diffused and fine granular signals were observed in the nucleus and the perikaryon. In addition, nucleolus-like structures were detected (arrows). **A'**, No signal was observed with a control IgG. Similar TCF4 staining pattern including nucleolar-like structures was found in DIV7 hippocampal neurons (**B**) and in DIV10 primary astrocytes (**C**). **D**, Representative nuclear profiles of DIV6 cortical neurons following 6h treatment with 0.1% DMSO (vehicle) or 1  $\mu$ M ActinomycinD (ActD). Note that the TCF4 signal shows the most extensive overlap with fibrillarin and only a partial overlap with Npm or UBF suggesting its localization to the transcriptionally active DFC. After transcriptional inhibition, TCF4, fibrillarin and UBF remained associated with the nucleolus. Note that fibrillarin and TCF4 segregated into distinct structures (indicated by a star) while UBF and TCF4 overlapped (indicated by a hash sign). Such a segregation pattern resembles fibrillar caps that sequester core components of the nucleolar transcription machinery (see text for details).

**Figure 7. Reduced neuronal ribosome content and inhibition of protein synthesis after knockdown of LARP7 in hippocampal neurons.** DIV4 (A) or DIV6 (B-E) rat hippocampal neurons were co-transfected with expression vectors for RFP and shRNAs against LARP7 or a control shRNA (shLuc, 0.2 + 0.6  $\mu$ g of plasmid DNAs/ $2 \times 10^5$  neurons, respectively); in D-E, 0.1  $\mu$ g EGFP expression vector was used instead of RFP. **A**, Representative images of transfected neurons (i.e. RFP-positive) that were fixed and immunostained for endogenous LARP7 at 72h

post-transfection (arrows). Note that shLARP7 constructs reduced expression of endogenous LARP7 including disruption of the LARP7 signal in the nucleoli. **B-C**, After 3 days, perikaryal ribosome content was evaluated in RFP-positive cells by measuring intensity of ribosomal staining with NeuroTrace Green. For quantification, the ribosomal signal was normalized against DNA content in the nucleus as determined by counterstaining with Hoechst-33258. **B**, Representative images of transfected (*i.e.* RFP-positive) neurons whose ribosomes were stained with NeuroTrace Green (arrows). **C**, Quantification of ribosomal content in transfected neurons. shRNAs reduced mean ribosomal content. Data represent means  $\pm$ SEM of at least 72 cells from three independent experiments; \*\*\*,  $p < 0.001$ . **D-E**, Three days after transfection, O-propargyl-puromycin (OPP) was added to the culture medium to label nascent peptides which were visualized after fixation by fluorescent Click iT chemistry that was followed by GFP immunostaining. **D**, Representative images of transfected (*i.e.* GFP-positive) neurons with nascent peptide accumulation (arrows). **E**, Nascent peptide accumulation in the perikarya was reduced by shLARP7 constructs suggesting global translation deficiency. Data represent means  $\pm$ SEM of at least 54 individual cells from two independent experiments; \*\*\*,  $p < 0.001$ . Signal specificity controls for NeuroTrace staining of ribosomes (loss of staining after RNase A treatment) and OPP labeling of nascent peptides (loss of signal in cycloheximide-treated neurons) are presented in Figure S3.

**Figure 8. Reduced neuronal ribosome content and inhibition of protein synthesis after knockdown of EMG1 in hippocampal neurons.** **A**, COS7 cells were co-transfected with expression vector for EMG1-EGFP together with the expression vector for  $\beta$ -galactosidase ( $\beta$ -gal) and shRNAs against EMG1 or control shRNA (shLuc, 1 + 1 + 2  $\mu$ g of plasmid DNA/60 mm

plate). Western blot for GFP revealed efficient knockdown of the target 48 h post transfection; the membrane was also probed for  $\beta$ -gal to ensure similar transfection efficiency. **B-E**, DIV6 rat hippocampal neurons were transfected and stained as described for experiments in Fig. 7B-E. **B-C**, Knockdown of EMG1 reduced mean ribosomal content in the perikarya. Data represent means  $\pm$ SEM of at least 58 cells from three independent experiments; \*\*\*,  $p < 0.001$ . **D-E**, Knockdown of EMG1 reduced general protein synthesis. Data represent means  $\pm$ SEM of at least 48 cells from two independent experiments, \*\*\*,  $p < 0.001$ .

**Figure 9. Reduced neuronal ribosome content and inhibited protein synthesis after overexpression of the PTHS-associated dominant-negative mutant form of TCF4.** DIV3 (A) or DIV6 (B-D) rat hippocampal neurons were co-transfected with expression vectors for human TCF4 including its wt- and two PTHS-associated mutant variants together with expression constructs for  $\beta$ -galactosidase ( $\beta$ -gal, A) or RFP (B,C) or EGFP (D,E) (0.4+0.2  $\mu$ g of plasmid DNAs/ $2 \times 10^5$  neurons, respectively). **A**, After 24 h, human *TCF4* mRNA levels were determined by qRT-PCR and normalized to  *$\beta$ -gal* mRNA. Human *TCF4* was detected in neurons that received TCF4 plasmids but not the empty vector control. **B-E**, After 3 days, perikaryal ribosome content (B-C) and general protein synthesis (D-E) were evaluated as described for Fig. 7. The mutants but not the wt reduced mean ribosomal content and general protein synthesis. Data represent means  $\pm$ SEM of at least 80 cells from five independent experiments (C) or 68 cells from three independent experiments (E); \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; NS,  $p > 0.05$ .

**Table 1. The FDR-verified enrichment of GO\_CC categories**

GOTERM_CC_FAT*	NUCLEOLUS <sup>^</sup>			NUCLEUS (all identified) <sup>~</sup>		
	Fold Enriched	% List	FDR	Fold Enriched	% List	FDR
GO:0030684~preribosome	17.77	1.59	1.10E-04	3.36	0.26	NS
GO:0005730~nucleolus	7.73	16.47	2.94E-47	1.93	3.58	1.67E-13
GO:0000786~nucleosome	7.45	3.17	2.70E-06	2.32	0.86	0.001155
GO:0005840~ribosome	7.22	9.52	3.96E-24	2.52	2.90	5.58E-22
GO:0033279~ribosomal subunit	6.56	2.98	5.76E-05	2.51	0.99	8.55E-06
GO:0000792~heterochromatin	6.23	2.18	0.010759	2.22	0.68	NS
GO:0005681~spliceosome	4.66	3.97	6.36E-05	2.92	2.17	1.04E-21
GO:0000785~chromatin	4.38	4.96	3.13E-06	1.85	1.83	6.27E-05
GO:0005694~chromosome	3.13	8.13	3.42E-07	1.49	3.37	8.18E-04
GO:0044427~chromosomal part	2.72	5.95	0.002514	1.47	2.79	0.01671
GO:0005604~basement membrane	6.72	3.37	3.71E-06	1.43	0.63	NS
GO:0044420~extracellular matrix part	5.34	3.37	1.21E-04	NE	NE	NE
GO:0005739~mitochondrion	2.47	22.42	3.12E-17	1.89	14.99	6.36E-66

\* only organelle/structure/complex-specific GO\_CC categories are not shown

<sup>^</sup> the analyzed list included 504 mapped IDs of proteins with at least 1.3 fold No enrichment; of them 433 had assigned GO\_CC terms

<sup>~</sup> the analyzed list included 3830 mapped IDs, of them 2865 had assigned GO\_CC terms

NS, p>0.05

NE, not enriched



**Table 2. The FDR-verified enrichment of GO\_BP categories**

GO_BP_FAT Term*	NUCLEOLUS ^			NUCLEUS (all identified)~		
	Fold Enriched	% list	FDR	Fold Enriched	% list	FDR
GO:0042254~ribosome biogenesis	13.85	8.73	1.90E-23	2.69	1.70	4.1078E-13
GO:0006396~RNA processing	12.99	13.89	2.18E-33	2.36	5.80	5.7144E-39
GO:0006364~rRNA processing	11.50	6.15	0.005802	2.76	1.15	1.5491E-08
GO:0016072~rRNA metabolic process	11.02	6.15	0.007979	2.72	1.15	2.8768E-08
GO:0006412~translation	10.86	9.92	2.61E-30	2.60	4.67	2.3355E-38
GO:0034470~ncRNA processing	9.80	6.75	0.018929	1.67	1.49	NS
GO:0034660~ncRNA metabolic process	9.60	6.94	0.004574	1.88	2.14	2.82E-06
GO:0051276~chromosome organization	9.45	8.53	0.024588	1.53	3.47	0.0001808
GO:0006397~mRNA processing	9.12	6.35	0.031575	2.76	4.07	2.5363E-37
GO:0006333~chromatin assembly or disassembly	8.97	4.17	2.39E-07	2.38	1.46	4.48E-08
GO:0006325~chromatin organization	7.25	6.94	5.93E-06	1.62	2.87	9.0115E-05
GO:0006323~DNA packaging	7.05	3.97	8.79E-06	2.25	1.28	1.062E-05
GO:0016071~mRNA metabolic process	6.96	6.55	1.07E-05	2.55	4.33	8.101E-34
GO:0008380~RNA splicing	6.96	5.16	1.07E-05	2.98	3.37	3.5164E-35
GO:0006334~nucleosome assembly	6.55	3.17	2.39E-07	2.48	1.02	1.759E-05
GO:0031497~chromatin assembly	6.37	3.17	1.25E-07	2.47	1.04	1.1704E-05
GO:0034728~nucleosome organization	5.40	3.17	3.74E-04	2.38	1.02	7.316E-05

\* all mitochondria related GO\_BP categories were excluded as a potential contamination

^ the analyzed list included 504 mapped IDs of proteins with at least 1.3 fold No enrichment; of them 411 had assigned GO\_BP terms

~ the analyzed list included 3830 mapped IDs, of them 2930 had assigned GO\_BP terms

NS, p>0.05

**Table 3. The FDR-verified enrichment of GO\_MF categories**

GOTERM_MF_FAT*	NUCLEOLUS^			NUCLEUS (all identified)~		
	Fold Enriched	% List	FDR	Fold Enriched	% List	FDR
GO:0019843~rRNA binding	16.69	2.38	3.8E-08	3.07	0.42	0.030584
GO:0003735~structural constituent of ribosome	9.95	8.93	1.6E-28	2.90	2.48	4.82E-24
GO:0070035~purine NTP-dependent helicase activity	6.60	3.37	6.2E-06	2.31	1.12	3.67E-05
GO:0008026~ATP-dependent helicase activity	6.60	3.37	6.2E-06	2.31	1.12	3.67E-05
GO:0004386~helicase activity	5.69	4.37	3.1E-07	2.18	1.59	4.60E-07
GO:0005198~structural molecule activity	4.53	12.10	5.6E-20	1.93	4.91	3.91E-19
GO:0003723~RNA binding	3.83	15.28	1.7E-21	2.13	8.12	1.27E-44
GO:0042623~ATPase activity, coupled	3.14	4.17	1.8E-02	2.03	2.56	2.73E-10
GO:0000166~nucleotide binding	1.54	20.04	5.8E-03	1.87	23.11	1.85E-104

\* mitochondrial protein-enriched categories are not shown

^ the analyzed list included 504 mapped IDs of proteins with at least 1.3 fold No enrichment; of them 398 had assigned GO\_MF terms

~ the analyzed list included 3830 mapped IDs, of them 2881 had assigned GO\_MF terms

**TABLE 4. Human counterparts of No-enriched rat brain proteins whose mutations cause neurodevelopmental diseases.**

<b>Protein</b>	<b>No/Nu</b>	<b>Function</b>	<b>Disease/neurological phenotype</b>	<b>OMIM#/ (Reference)</b>
<b>EMG1</b>	1.66	Ribosomal biogenesis	Bowen-Conradi Syndrome/microcephaly, developmental delay	211180
<b>RPL10</b>	1.43	Structural protein of the ribosome	X-linked microcephaly, growth retardation, and seizures	(44)
<b>DKC1</b>	2.07	Ribosomal biogenesis/telomere maintenance	Dyskeratosis Congenita, X-Linked/learning difficulties, microcephaly, cerebellar hypoplasia	305000
<b>EIF4A3</b>	1.36	Ribosomal biogenesis, mRNA splicing and decay	Richieri-Costa-Pereira Syndrome/developmental delay, learning disability, language delay/disorder	268305
<b>FLNA</b>	1.63	Actin regulation, negative regulation of PI3K signaling, negative regulation of Pol1, pre-rRNA processing	Heterotopia, Periventricular, X-Linked Dominant/ seizures, cortical dysplasia	300049
<b>SMC1A</b>	1.42	Chromatid cohesion, general transcription regulation	Cornelia De Lange Syndrome 2/microcephaly, mental retardation, dilated ventricles, seizures	300590
<b>MCM4</b>	1.93*	Prevents DNA re-replication, maintains genomic stability	Natural Killer Cell And Glucocorticoid Deficiency With Dna Repair Defect/microcephaly	609981
<b>NSD1</b>	1.3	Transcription factor and chromatin modifier	Sotos Syndrome 1/ macrocephaly, ventriculomegaly, seizures, developmental delay, mental retardation, poor coordination	117550
<b>CUL4B</b>	1.31	DNA damage response	Mental Retardation, X-Linked, With Short Stature	300354
<b>PRKDC</b>	2.2*	DNA damage response	Immunodeficiency 26 With Or Without Neurologic Abnormalities/ Microcephaly, mental retardation)	615966
<b>ATRX</b>	1.43	Chromatin remodeling	Alpha-Thalassemia and/or Mental Retardation Syndromes, X-Linked; Linked/mental retardation, microcephaly	301040; 309580
<b>LMNA</b>	1.54	Nuclear envelop protein	Charcot-Marie-Tooth Disease, Axonal, Type 2b1 (LMNA loss of function)/axonal atrophy in peripheral nerves, distal limb muscle weakness/atrophy due to peripheral neuropathy; Malouf Syndrome (heterozygous point mutation in LMNA, possible gain of function)/mental retardation	605588; 212112
<b>LARP7</b>	2.05	Stabilizes 7SK ncRNA that inhibits the p-TEFb transcription factor	Alazami Syndrome/ intellectual disability and microcephaly	615071
<b>TCF4</b>	2.23*	Transcription factor	Pitt-Hopkins Syndrome/postnatal microcephaly, cerebral atrophy, mental retardation, poor or absent speech development, seizures, hypotonia	610954
<b>GTF2I</b>	1.32	General transcription factor	Williams-Beuren Syndrome/mental retardation, poor visual-motor integration, hypotonia, poor balance and coordination	194050
<b>ADAR</b>	1.47	RNA editing	Aicardi-Goutieres Syndrome 6/atrophy of striatum, leukodystrophy, dystonia, tremor, loss of motor skills	615010

Shading: dark gray- ribosomal biogenesis-related proteins, light gray- chromatin organization/ genome maintenance-related protein,

\* Non-significant enrichment that was supported by the classifier

**TABLE 5. Human counterparts of No-enriched rat brain proteins whose mutations cause neurodegenerative diseases.**

<b>Protein</b>	<b>No/Nu</b>	<b>Function</b>	<b>Disease/neurological phenotype</b>	<b>OMIM#/ (Reference)</b>
<b>EWSR1</b>	1.9	RNA binding protein, RNA splicing, ribosomal biogenesis	Amyotrophic Lateral Sclerosis	(45)
<b>TARDBP</b>	1.45	DNA/RNA binding protein, RNA splicing	Amyotrophic Lateral Sclerosis 10 With Or Without Frontotemporal Dementia; Als10	612069
<b>LMNB1</b>	1.47	nuclear envelop protein	Leukodystrophy, Demyelinating, Adult-Onset, Autosomal Dominant (LMNB1 duplication); ADLD; cerebellar ataxia, cognitive decline, personality changes due to degeneration of the CNS white matter	169500



















