Current Biology, Vol. 12, 1–11, January 8, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0960-9822(01)00650-9

Directed Proteomic Analysis of the Human Nucleolus

Jens S. Andersen,[1] Carol E. Lyon,[2] Archa H. Fox,[2] Anthony K.L. Leung,[2] Yun Wah Lam,[2] Hanno Steen,[1] Matthias Mann,[1,3,4] and Angus I. Lamond[2,4]
[1]Department of Biochemistry and Molecular Biology
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark
[2]Wellcome Trust Biocentre
MSI/WTB Complex
University of Dundee
Dundee DD1 4HN
United Kingdom
[3]MDS Proteomics
Staermosegaadsvej 6
5230 Odense M
Denmark

Summary

Background: The nucleolus is a subnuclear organelle containing the ribosomal RNA gene clusters and ribosome biogenesis factors. Recent studies suggest it may also have roles in RNA transport, RNA modification, and cell cycle regulation. Despite over 150 years of research into nucleoli, many aspects of their structure and function remain uncharacterized.

Results: We report a proteomic analysis of human nucleoli. Using a combination of mass spectrometry (MS) and sequence database searches, including online analysis of the draft human genome sequence, 271 proteins were identified. Over 30% of the nucleolar proteins were encoded by novel or uncharacterized genes, while the known proteins included several unexpected factors with no previously known nucleolar functions. MS analysis of nucleoli isolated from HeLa cells in which transcription had been inhibited showed that a subset of proteins was enriched. These data highlight the dynamic nature of the nucleolar proteome and show that proteins can either associate with nucleoli transiently or accumulate only under specific metabolic conditions.

Conclusions: This extensive proteomic analysis shows that nucleoli have a surprisingly large protein complexity. The many novel factors and separate classes of proteins identified support the view that the nucleolus may perform additional functions beyond its known role in ribosome subunit biogenesis. The data also show that the protein composition of nucleoli is not static and can alter significantly in response to the metabolic state of the cell.

[4]Correspondence: a.i.lamond@dundee.ac.uk (A.I.L.), mann@bmb.sdu.dk (M.M.)

Introduction

Many studies have shown that the cell nucleus contains distinct and often dynamic compartments [1–4]. Chromosome territories are separated by the interchromatin space, which contains discrete nuclear bodies, including nucleoli, Cajal bodies, speckles, gems, and promyelocytic leukemia (PML) bodies that have mainly been characterized at the morphological level [5]. Nuclear bodies contain complexes of proteins and/or RNPs but are not separated by membranes from the surrounding nucleoplasm. In vivo experiments using fluorescence photobleaching (FRAP and FLIP) techniques have shown that their protein components move by passive diffusion and are in continuous flux between nuclear compartments [6, 7]. Self-organization has been suggested as a mechanism for nuclear body formation and maintenance [8]. In addition, nuclear bodies themselves can move in the nucleoplasm. For example, the mobility of Cajal bodies [9] may also involve diffusion, although active and directed mechanisms could also play a role [10].

In most cases, the biological roles of nuclear bodies are still not well understood, and their molecular constituents have not been comprehensively identified. Recent advances in MS techniques, coupled with the accumulation of genomic and EST DNA sequence databases, have facilitated the global characterization of protein complexes and organelles [11–14]. In the case of the nucleus, a recent study has applied MS to analyze the protein composition of fractions from mouse liver cells enriched in interchromatin granule clusters (IGCs) [15]. Previous work from our own laboratories and from other groups has employed proteomic methods to characterize the protein composition of a variety of multiprotein complexes and subcellular organelles from yeast to human, including U1 snRNPs [16], spliceosomes [17], and nuclear pore complexes [18].

The best-studied example of a membrane-free nuclear subdomain is the nucleolus, a "cellular factory" in which 28S, 18S, and 5.8S ribosomal RNAs (rRNAs) are transcribed and together with 5S rRNA are processed and assembled into the ribosome subunits. The nucleolus is a dynamic structure that disassembles and reforms during each cell cycle around the rRNA gene clusters [19]. Within the nucleolus, three distinct subcompartments have been described based on their morphology in the electron microscope. These correspond to the fibrillar centers (FC), dense fibrillar components (DFC), and granular components (GC). The current consensus view is that rDNA transcription is restricted to the periphery of the FC, while transient accumulation, modification, and processing of primary rRNA transcripts occurs in the DFC, and later processing and rRNA assembly into ribosomal subunits occurs in the GC [20–23].

While it is clear that the major role of the nucleolus is in ribosome subunit biogenesis, it is interesting that recent studies suggest that there may be additional functions for the nucleolus [24–26]. For example, it may also be a site for the biogenesis and/or maturation of

other ribonucleoprotein machines, including the signal recognition particle [27], the spliceosomal small nuclear RNPs [28], and telomerase [29]. The nucleolus may also participate in processing or export of some mRNAs and tRNAs [26, 30] and can control the activities of specific regulatory factors by a sequestration mechanism [31]. Moreover, association between the nucleolus and other nuclear bodies, such as the perinucleolar compartment (PNC) [32] and the Cajal body [33], also raises the possibility that nucleoli have additional functions.

Here we present an analysis of the proteome of purified human nucleoli, using MS and computer searches in sequence databases, including the draft human genome. Over 30% of the 271 nucleolar proteins identified are encoded by novel human genes, and a subset of factors are shown to accumulate in nucleoli specifically when transcription is inhibited.

Results

Isolation of Nucleoli

It was essential to have highly purified preparations of nucleoli as a source material for direct protein analysis. The inherent density of nucleoli [34] facilitated their isolation from cultured cells, using a combination of sonication and sucrose density centrifugation (see Experimental Procedures). Thus, nucleoli were isolated from HeLa cell nuclei, and their purity and integrity were analyzed using both light and electron microscopy and protein blotting (Figure 1). As judged by light microscopy, over 95% of the particles detected also stained positive with the RNA-specific dye Pyronin Y and were immunolabeled with antibodies to nucleolar proteins (Figure 1A). The isolated nucleoli were embedded in resin, sectioned, and examined in the electron microscope (Figure 1B), showing that they had typical nucleolar morphology, as seen in sections through intact nuclei (Figure 1B, compare nucleoli in situ [1Bi] with isolated nucleoli [panels 1Bii–1Biv]). At the highest magnification, the internal nucleolar substructures, corresponding to the FC, DFC, and GC, are clearly visible in sections of the isolated nucleoli (Figure 1Biv). To address the purity of the isolated nucleoli using a biochemical assay, protein samples from (1) unfractionated nuclear extract, (2) nucleoplasmic fraction, and (3) purified nucleoli were separated by SDS PAGE, transferred to a nitrocellulose membrane, and probed with antibodies specific for either nucleolin, fibrillarin, Lamin B, or NUP 62 (Figure 1C). This shows that both nucleolar proteins tested, i.e., nucleolin and fibrillarin, are highly enriched in the nucleolar fraction (compare lanes 2 and 3). In contrast, two nuclear proteins that do not accumulate in nucleoli, i.e., lamin B and NUP62, are not detected in the purified nucleolar fraction.

To evaluate the suitability of these purified samples for large-scale MS characterization, a preliminary MS analysis was conducted following separation of the purified nucleolar proteins by one-dimensional (1D) SDS PAGE. Gel lanes were completely sliced, in-gel digested with trypsin, and the resulting peptides analyzed automatically by high mass-accuracy peptide-mass mapping using MALDI-MS (Figure 2 and data not shown). These data were screened against a human-specific nonredundant protein database (nrdb). The resulting list of 80 proteins included many known nucleolar proteins without obvious major protein contaminants (see below). We conclude that the isolation procedure yields intact preparations of HeLa cell nucleoli that are sufficiently pure for use in proteomic analysis.

Identification and Characterization of Nucleolar Proteins

The purified nucleoli were used in a large-scale MS analysis, involving the identification of proteins separated by both 1D and 2D SDS PAGE (Figure 2; Table S1 [Table S1 is included as a special reprint of this article and as Supplementary Material available online]). A variety of different gel/buffer systems, which each have advantages for resolving different types of proteins (for example, compare Figures 2A and 2B), were used to ensure maximal coverage of nucleolar proteins. However, it was not necessary to resolve every individual protein component prior to MS analysis, because nanoelectrospray (nanoES) tandem MS could identify separate proteins even in complex mixtures, including up to eight different factors (Figure 3A). The majority of nucleolar proteins could therefore be identified by analyzing proteins separated using the different 1D SDS PAGE systems followed by nanoES sequencing of proteins in individual gel slices.

To facilitate the analysis of the complex range of nucleolar proteins, a layered analytical strategy [35] was refined to allow online interrogation of complex peptide mixtures and to identify proteins directly in genome sequences (Figure 3). Peptide sequence tags [36] assigned from fragmentation spectra of peptide signals were searched against various sequence databases in "real time" during data acquisition. The sequences of other expected peptides in the mixture were then identified on the basis of already identified proteins. These peptides were labeled and excluded from further sequencing, thereby sequentially reducing the complexity of the peptide mixture. This approach, termed "directed sequencing," proved to be highly efficient at allocating sequencing time to the identification of the largest number of proteins from each sample (Figure 3B). Unambiguous sequence identification was established by directed sequencing of a minimum of two peptides for each protein, followed by matching expected and measured peptide fragment ions. Directed sequencing of key peptides was particularly useful in (1) the identification of low abundance proteins, (2) the identification of proteins represented by limited sequence information, e.g., EST's, and (3) to differentiate between protein isoforms, for example, PSP1 where sequencing of the peptide ALESAVGESE PAAAAAMALALAGEPAPPAPAPPEDHPDEEMGFTI DIK was required to confirm the N-terminal extended form [37].

The 1D PAGE analyses showed that the majority of nucleolar proteins were in the size range ~25–100 kDa. 2D gel electrophoresis was used to expand the separation of proteins in this size range (Figure 2C). Individual spots from the 2D gels were detected as described in Experimental Procedures, excised, and in-gel digested with trypsin prior to MS analysis. These data confirmed

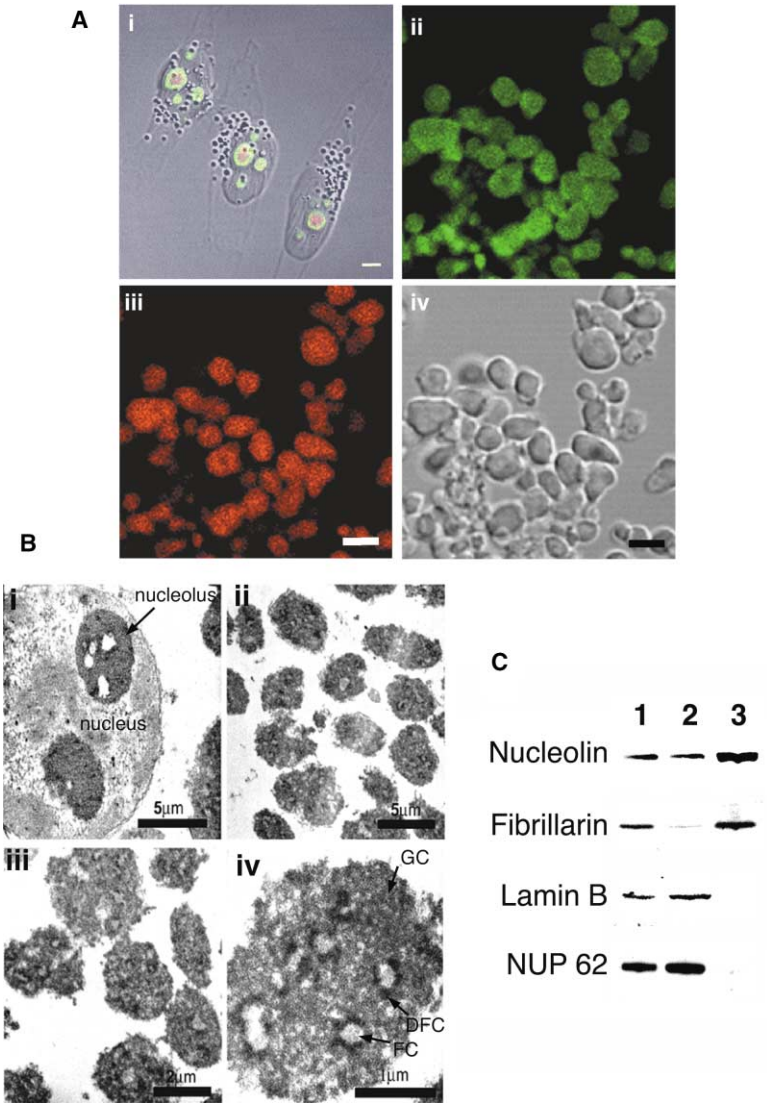

Figure 1. Nucleoli Can Be Isolated from HeLa Cells in Large Quantities and with High Purity

The purity of the isolated nucleolar preparations was monitored using several methods. (A) Light microscopy can demonstrate the purity of purified HeLa nucleoli. Panel (Ai) shows an intact HeLa cell stained with two nucleolar-specific labels: Pyronin Y (red) and nucleolin antibody (green). These reagents were also used to stain purified nucleoli (nucleolin in panel [Aii] and Pyronin Y in panel [Aiii]). DIC imaging of the purified material (Aiv) also shows that it does not contain significant amounts of nonnucleolar particles. (B) TEM images of thin sections of a HeLa cell nucleus with two nucleoli in situ, marked by arrows (Bi). Panels (Bii)–(Biv) show TEM images of sections of isolated nucleoli in increasing magnification. These clearly show that the isolated nucleoli are morphologically intact; they contain clearly defined granular component (GC), dense fibrillar component (DFC), and fibrillar centers (FC). (C) Proteins from preparations of sonicated whole nuclei, nucleoplasm, and nucleoli were separated by gel electrophoresis, transfered to nitrocellulose, and immunolabeled with anti-nucleolin, anti-fibrillarin, anti-lamin B, and anti-Nup 62. The nucleolar fraction is enriched in nucleolin and fibrillarin but not in the nonnucleolar proteins.

the identification of many of the factors detected in the 1D gel systems and identified 20 additional proteins (Table S1). However, many basic proteins identified by 1D PAGE analyses were not resolved in the 2D gel system.

The peptide sequence tag search algorithm [36] was adapted to allow protein identification directly in the draft genome sequence [38]. Protein identity was confirmed by detecting two or more peptides matching within either a single exon or within different exons confined to the same genomic region. Examples of direct MS gene identification in the unannotated human genome are shown in Figures 3C–3E. In the first example, a peptide sequence tag (underlined in Figure 3D) matched exclusively a novel exon located in the genome sequence within chromosome 16. Directed sequencing of three additional peptides (underlined in Figure 3D), predicted from the same exon, confirmed the identification. The coding sequence of this unusual large exon, encoding a putative exoribonuclease, had a calculated M_r of 25.993 kDa, consistent with the observed M_r of 26 kDa for the cognate gel band. The second example illustrates genome searching combined with gene prediction (Figure 3E). In this case, three peptides that were identified all mapped to a short region of the genome, spanning ~100 kb within chromosome 12. This region was therefore used as a constraint for further analysis. Sequencing of additional peptides that either lay within or spanned exons from this region allowed us to refine the predicted gene structure at this locus, including the identification of an additional exon. Importantly, this MS analysis also identified the N-acetylated N-terminal peptide that includes the initiating methionine. A total of 14 novel proteins that copurify with nucleoli were identified directly by analysis of the human genome sequence (Table S1). In other cases, peptides identified and matched to the genome sequence allowed us to revise the structures of predicted genes retrieved from a nonredundant database (for example, GCN1L1, SAZD, and BING4). The above examples show how the MS data, combined with direct analysis of the human genome sequence, facilitate the identification of novel nucleolar

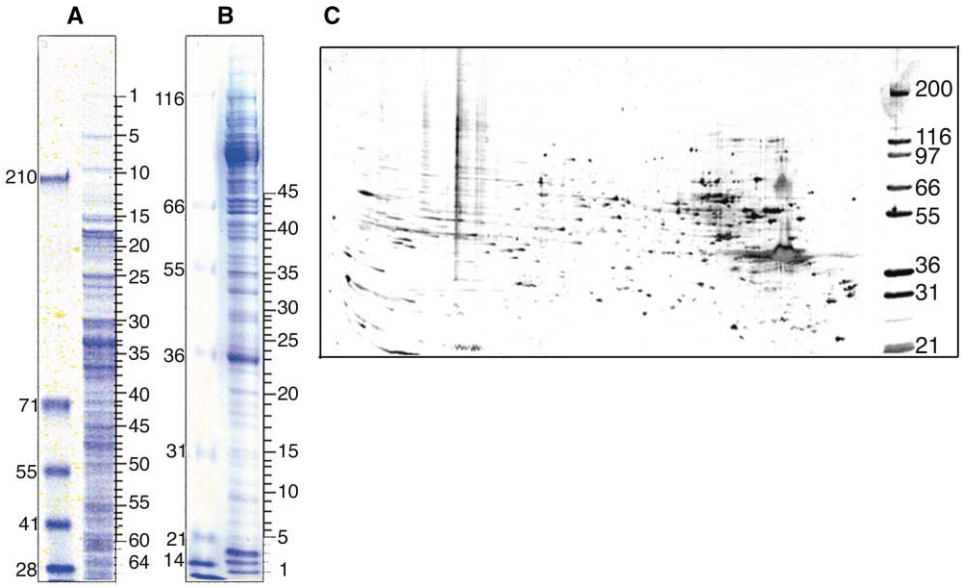

Figure 2. Proteins from Isolated Nucleoli Were Resolved by Both 1D and 2D SDS/PAGE to Maximize the Coverage of Nucleolar-Associated Proteins

(A and B) Representative Coomassie-stained 1D SDS/PAGE of nucleolar proteins separated utilizing either a tris-glycine (A) or a tris-acetate buffer system (B), respectively. Molecular weight size markers (kDa) are shown to the left of each gel, and rulers used for excising gel slices are shown at right.

(C) Silver-stained 2D gel of nucleolar samples focused in the first dimension using a 3-10 strip then subsequently electrophoresed on a 12.5% SDS gel. Molecular weight size markers are shown at right (kDa).

proteins and simultaneously improve the annotation of the draft human genome sequence. Altogether, a total of 271 separate gene products were identified from the nucleolar preparations (Table S1).

Characterization of Identified Nucleolar Proteins

Many of the genes identified above encoded known proteins that had previously been shown to localize to nucleoli in human cells or else had close homologs that had been shown to localize to nucleoli in other species (Table S1). However, over 30% of the genes encoded novel or uncharacterized proteins whose localization was unknown. In addition, some of the known gene products identified had not been reported as nucleolar factors. To address whether these classes of proteins are genuine nucleolar components in vivo, we adopted the strategy, as in our previous study of the spliceosome [17], of tagging them with fluorescent proteins and examining their localization in the fluorescence microscope following transient expression in cultured cells (Figure 4). Due to the large number of proteins involved, it was beyond the scope of this study to isolate and tag full-length cDNA clones for every gene identified. Therefore, we selected for analysis 18 examples of proteins that were either novel, or not known to be nucleolar, isolated cDNA clones and fused them to yellow fluorescent protein (YFP) at their amino termini in the pEYFPC1 expression vector (see Experimental Procedures). The proteins selected were chosen to include a range of sizes, pI values, and motifs. Following transient transfection and expression in HeLa cells, 15 of the fusion proteins localized to nucleoli, showing a variety of different patterns (Figure 4 and data not shown). Some of the proteins localized exclusively to nucleoli, whereas others were detected in nucleoli but also accumulated at other nuclear and/or cytoplasmic structures. For example, the human Pescadillo homolog (PES1) localized specifically to nucleoli (Figure 4A), whereas the PWP1 protein showed extensive cytoplasmic localization as well as accumulating in nucleoli (Figure 4D). Some proteins were detected within distinct subregions of the nucleolus, such as the novel protein SAZD (Figure 4E). The variety of localization patterns observed with the YFP-fusion proteins is consistent with the wide spectrum of nucleolar labeling patterns detected previously for known nucleolar proteins. During the course of our analyses, several recent studies have independently reported the nucleolar localization of some of the proteins we have identified by MS (for example, NOH61 [39] and Bop1 [40]). As discussed below, the YFP-tagging data support the view that the majority of the novel proteins detected here by MS can interact with nucleoli in vivo. Some or all of the tagged proteins that could not be confirmed here as nucleolar may also interact with nucleoli in vivo (see Discussion).

An overview of the separate classes of proteins identified in the purified nucleoli reported in Table S1 is presented in the form of a pie chart in Figure 5. The largest single category (~31%) represents the group of novel and previously uncharacterized factors, including those whose function cannot be readily inferred based upon recognizable homology to known proteins. The next largest category (24%) corresponds to proteins that bind to nucleic acid and/or to nucleotides, consistent with the known role of the nucleolus in ribosome biogenesis and the presence in nucleoli of large amounts of rRNA and the rRNA repeat gene clusters. Other classes listed

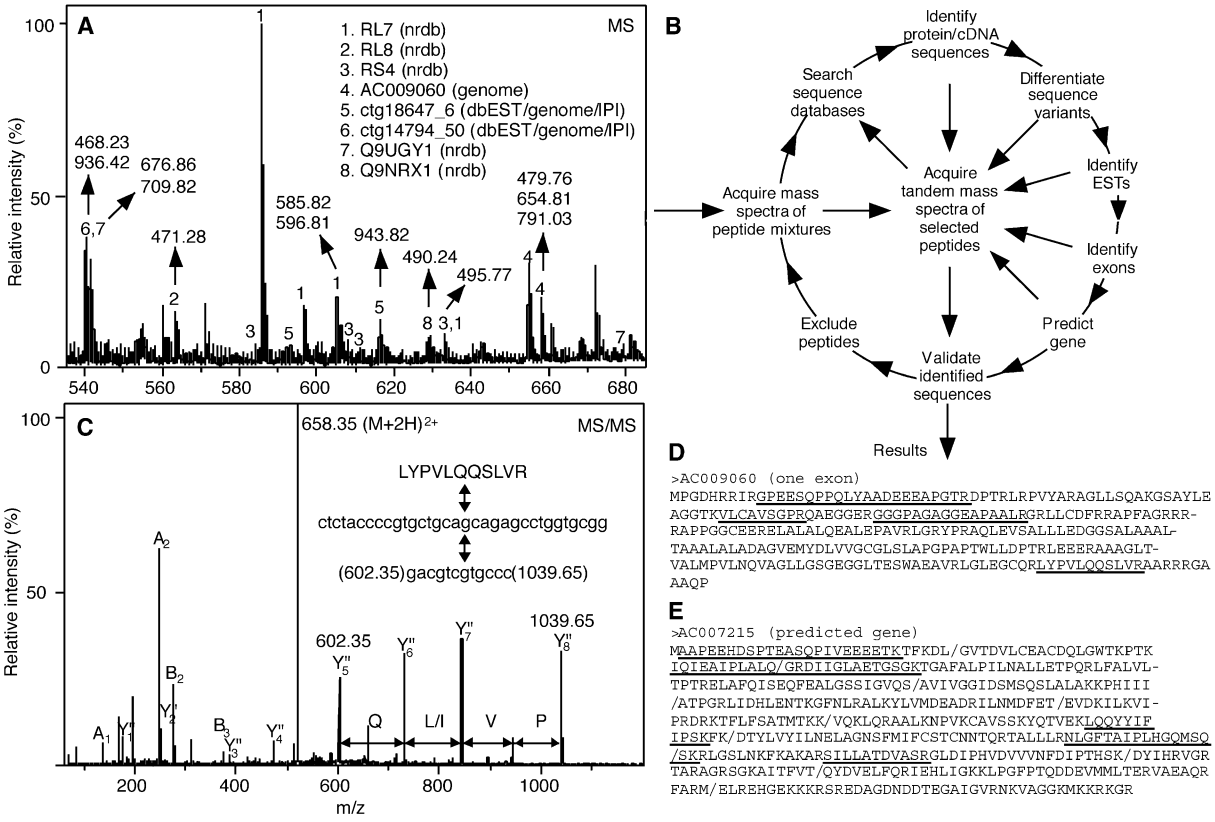

Figure 3. Identification of Proteins from Isolated Nucleoli

The gel lanes/spots from the 1D and 2D gels of nucleolar proteins (see Figure 2) were cut into slices, enzymatically digested, and the resulting peptide mixtures analyzed by MALDI and nanoES MS. (A) Part of nanoES mass spectrum of tryptic peptides from gel slice 18P. Marked peaks (1) were fragmented and validated identification of RL7 by MALDI-MS peptide mapping. Tandem MS of unexplained peptide signals, followed by peptide sequence tag database searches, resulted in identification of seven additional proteins. Database searches in real time allowed unambiguous identification by directed sequencing of at least two peptides for each of the retrieved sequences. (B) Flow chart illustrating the directed sequencing approach to identify and validate a maximum number of proteins from a peptide mixture. (C–E) Examples of identifying proteins in the human genome. (C) Fragment ion mass spectrum of the peak at m/z 558.35. A peptide sequence tag was assembled from the series of C-terminal fragment ions. The amino acid sequence of the search string was translated into the corresponding degenerated nucleotide sequence. Potential hits in the forward or reverse direction of the genome data were checked as to whether they coded for the amino acid sequence defined by the peptide sequence tag. Additional N- and C-terminal fragment ions (B and Y″ ions) matched the retrieved peptide LYPVLQQSLVR. (D) Amino acid sequence translated from the identified exon. Underlined peptides were fragmented to confirm identification and to narrow down the probable terminal splice sites. (E) Amino acid sequence translated from a gene predicted from genomic data using the peptides IGLAETGSGK, LQQYYIFIPSK, and ILLATDVASR identified in slice 36P as coding constraints. Underlined peptides were sequenced, including the N-acetylated N-terminal peptide and peptides spanning exon boundaries that partially confirmed the gene prediction.

also include families of RNA or nucleotide binding proteins whose functions are more clearly defined. For example, the DEAD box motif that is characteristic of the superfamily of RNA-dependent ATPases and helicases [41] is found in 8% of the total nucleolar proteins, including 3% of the novel factors. The ribosomal proteins and other translation factors comprise a similar fraction of nucleolar proteins as do the RNA modifying enzymes (including snoRNP proteins) and chaperones. The category listed as “others” includes proteins with a wide range of known motifs and disparate functions. Collectively, these data illustrate the surprisingly high complexity of the nucleolar proteome.

Actinomycin D Treatment Changes the Nucleolar Proteome

We analyzed whether the nucleolar proteome was affected by treatment of HeLa cells with Actinomycin D, which is known to affect the structure of the nucleus, the localization pattern of nuclear proteins, and the integrity of the nucleolus [42]. The efficacy of Actinomycin D treatment was confirmed by double labeling treated and control HeLa cells with both pyronin Y and anti-coilin antibodies (Figures 6A and 6B). As expected [43, 44], Actinomycin D-treated cells showed a relocalization of coilin to form caps at the nucleolar periphery (cf. panels 6A and 6B, green label shows coilin and red foci indicate nucleoli). Additional experiments to monitor the incorporation of Br-UTP before and after Actinomycin D treatment indicated that both RNA polymerase I and II activities had been inhibited by the drug treatment (data not shown). Therefore, we isolated quantitative preparations of nucleoli from Actinomycin-treated cells to allow for direct MS analysis of their protein composition (Figure 6). Double labeling of the nucleoli isolated from Actinomycin-treated cells with anti-coilin antibodies and pyro-

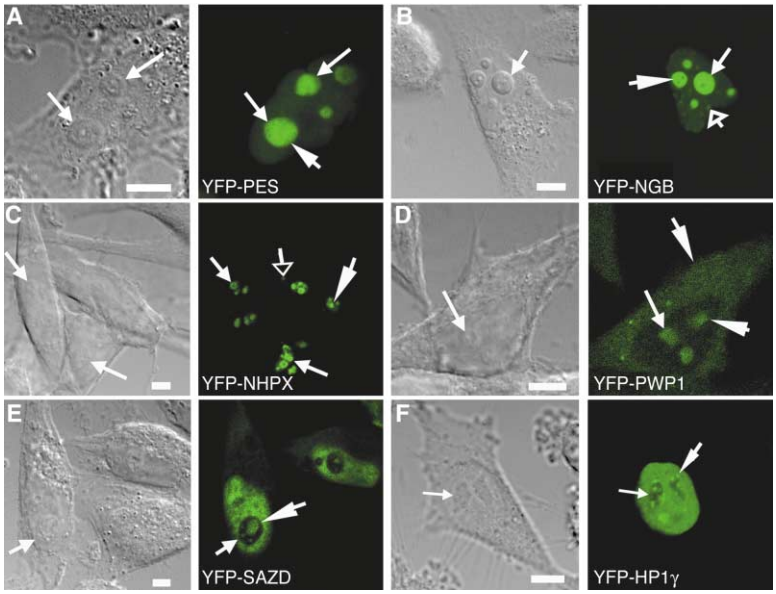

Figure 4. Tagged Nucleolar Candidate Proteins, Identified by MS, Accumulate within Nucleoli in HeLa Cells, Albeit with Varying Patterns

Each panel shows a confocal fluorescence micrograph of HeLa cells expressing YFP fusion proteins, with the corresponding Nomarski image shown on the left. Constructs expressed were as follows (names and accession numbers in brackets): (A) YFP-PES1 (Pescadillo, Hs.13501), (B) YFP-NGB (GTP binding protein, Hs.215766), (C) YFP-NHPX (Hs.182255), (D) YFP-PWP1 (nuclear phosphoprotein similar to *S. cerevisiae* PWP1, Hs.172589), (E) YFP-SAZD (Hs.114416), and (F) YFP-HP1γ (Heterochromatin Protein 1 γ isoform, Hs.278554). Small arrows indicate nucleoli, large arrowheads indicate the localization of the fusion proteins within the nucleoli, and open arrowheads highlight small nuclear bodies also labeled by the fusion proteins (in the case of YFP-NGB and YFP-NHPX). Scale bars, 5 μm.

nin Y confirmed that they had an identical morphology to the nucleoli in intact Actinomycin D-treated cells (cf. panels 6C and 6D with 6B, green label shows coilin, and red foci indicate nucleoli). The purity and morphology of these Actinomycin-treated nucleolar preparations was also confirmed and analyzed in the electron microscope and will be presented in detail elsewhere (C.E.L. et al., unpublished data).

Proteins from nucleoli isolated from both control and Actinomycin D-treated HeLa cells were separated by 1D SDS PAGE and stained with Coomassie dye (Figure 6G). The total proteomes are similar for both control and Actinomycin D-treated nucleoli as seen by SDS-PAGE (Figure 6G) and by additional MS analysis (data not shown). However, there were some protein bands whose intensity in the Actinomycin-treated nucleoli was increased relative to the control nucleoli (Figure 6G, cf. Lanes 1 and 2, arrows). These bands were excised, in-gel digested with trypsin, and analyzed by both MALDI peptide mapping and nanoES MS. This identified 11 proteins as major candidates for factors whose abundance in the nucleolar proteome was increased following Actinomycin D treatment (Figure 6G, proteins listed beside arrows). This included the protein p80 coilin (running at ~66 kDa in both extracts in this gel system), which was shown by immunocytochemical analysis to have increased association with nucleoli following Actinomycin D treatment (see also Figures 6A–6D). Interestingly, the other ten proteins identified were members of three separate protein families, i.e., DEAD box proteins (DDX9, p72, and p68), hnRNP proteins (hnRNPs K, G, and A2/B1), and a group of related RNA binding proteins (PSF, PSP2/CoAA, PSP1, and p54/nrb). We therefore selected a number of proteins from two of these families (i.e., p68, p72, p54/nrb, PSP1, and PSP2/CoAA), tagged them with YFP, and compared their localization in both control and Actinomycin-treated cells (Figures 6E and 6F show one example where the green signal shows YFP-p68 and the red channel Pyronin Y; see also [37] and other data not shown). Surprisingly, all the proteins analyzed were predominantly nucleoplasmic but relocated to the nucleolar periphery following Actinomycin D treatment (a detailed analysis of the localization and behavior of PSP1 is reported in [37]). Furthermore, consistent with this MS identification, a recent study has reported that GFP-PSF also relocalizes to the nucleolar periphery following Actinomycin D treatment [45].

In summary, we conclude that metabolic perturbations, as induced here by Actinomycin D treatment, can induce changes in the relative abundance of a subset of the nucleolar proteome.

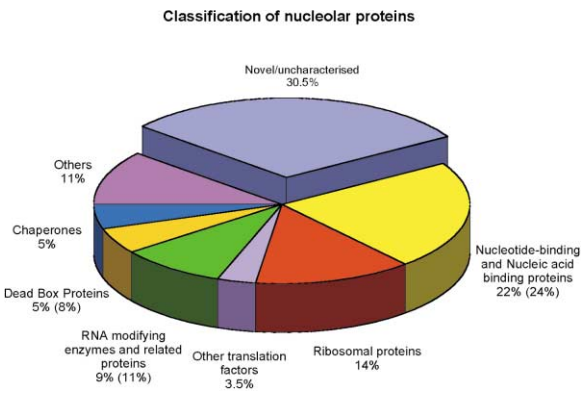

Figure 5. Distribution of Conserved Motifs and Putative Functional Categories of the Identified Proteins

The proteins from Table S1 were divided into groups as indicated above. The percentages in brackets indicate the number of proteins, including those classed as novel/uncharacterized, that contain the motif in question.

Discussion

We have reported here a characterization of the proteome of nucleoli isolated from cultured human cells. This study represents the largest proteomic analysis

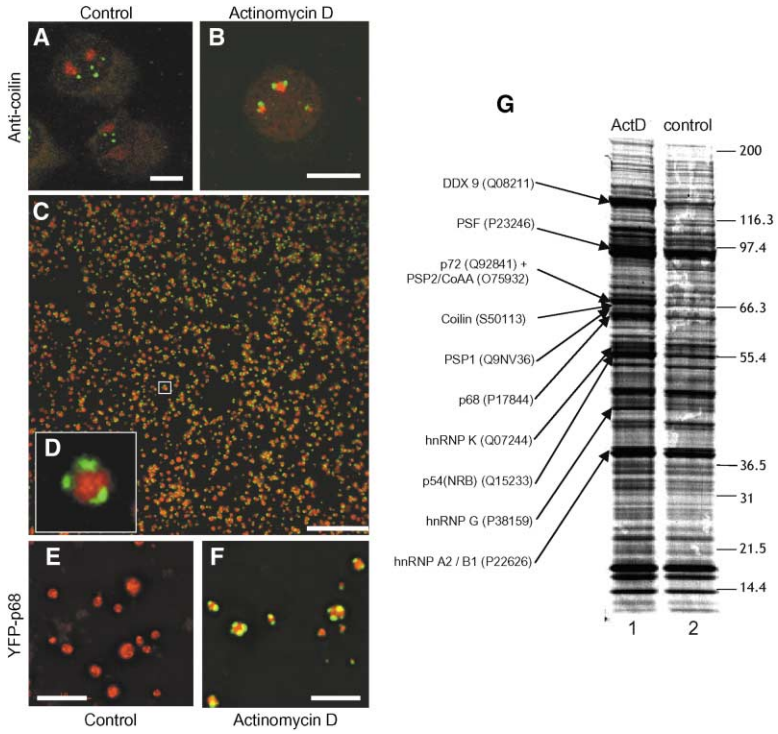

Figure 6. Changes in Nucleolar Structure and Composition after Actinomycin D Treatment of HeLa Cells

Following incubation of HeLa cells with Actinomycin D (see Experimental Procedures), coilin (as detected with anti-coilin antiserum, green) relocalized from Cajal bodies, as seen in untreated cells (A), to caps, labeled by Pyronin Y (red, [B]). These perinucleolar coilin caps remained detectable in nucleoli isolated from Actinomycin D-treated cells (C). (D) A magnified view of one of the isolated nucleoli. (G) Proteins from nucleoli isolated from Actinomycin D-treated cells (lane 1) and the untreated control nucleoli (lane 2) were separated by SDS-PAGE and stained with Coomassie blue. Proteins that are more abundant in lane 1 were analyzed by MS and identified as labeled (numbers in brackets are SwissProt accession numbers). One of the identified proteins, p68, was tagged with YFP and expressed in HeLa cells. YFP-p68 (green signal) was not significantly detectable in nucleoli isolated from the YFP-p68 stable cell line (red signal is Pyronin Y staining) (E) but was found at caps following Actinomycin D treatment (F). Scale bars, 10 μm (A and B); 25 μm (C); and 15 μm (E and F).

reported so far for a single organelle and identifies 80 novel human genes that encode putative nucleolar proteins. Importantly, the proteomic analysis of nucleoli isolated from cells in which transcription had been inhibited revealed the dynamic nature of the nucleolar proteome. A subset of 11 nucleolar proteins was shown to increase their association with nucleoli following treatment of cells with Actinomycin D. This analysis provides new insights into the complexity of proteins in the nucleolus and identifies many nucleolar factors that are conserved in evolution from budding yeast to humans. The data indicate that the nucleolus is an organelle with a high degree of functional complexity.

The large-scale characterization and identification of nucleolar proteins required the application of a variety of parallel approaches for both the effective separation of different classes of proteins and their subsequent analysis by both MALDI TOF and nanoES MS. The identities of the proteins included in the nucleolar database were each confirmed through the detection of at least two independent peptides. Furthermore, in many cases, genes were identified from the MS data through direct online searching of the draft human genome sequence. The strategy used, involving real-time database searching during data acquisition, facilitated the identification of proteins in complex peptide mixtures. This provided major practical advantages in dealing with such a large-scale analysis as that involved in determining the nucleolar proteome. Previously, we demonstrated that 18 novel human proteins analyzed in a large-scale proteomic study of the spliceosome could be identified by screening EST databases [17]. Now, with access to the human genome sequence, we have demonstrated that searches directly in the human genome can be advantageous when identifying novel proteins from peptide mixtures. In these searches, different peptides were immediately linked by mapping to a confined region in the genome. In addition, the resulting information, which defines detailed gene structures, can play an important role in both extending and refining the annotation of the human genome. In this way, the predicted coding products of the genome can be compared with the actual primary structures of the cognate proteins expressed in vivo.

For over half of the known proteins identified in the purified nucleoli, evidence exists in the literature to support their nucleolar association. We are confident that the majority of the additional proteins identified, including those encoded by novel genes, are also bona fide nucleolar factors. Few known contaminants were detected in the nucleolar preparations, and among the novel proteins we observed low scores for signal peptide sequences that could be indicative of contaminating cytoplasmic and ER proteins. The surprisingly high proportion of novel proteins also argue for highly enriched preparations. Furthermore, transient transfection analysis of 18 selected genes from the list in Table S1, each tagged with YFP, showed that 15 associated with nucleoli, as judged by fluorescence microscopy. Extrapolating from this YFP-tagging analysis, we estimate that at least 80% of the unverified genes are likely to encode nucleolar proteins. In fact, this figure may underestimate the fraction of genes encoding genuine nucleolar proteins for several reasons. First, the presence of the YFP tag in the fusion protein may interfere with nucleolar localization. Second, some proteins are known only to localize to nucleoli at specific stages of the cell cycle [46, 31] or under specific metabolic conditions (this study). Third, because of the high sensitivity of protein detection by MS, our analysis may have identified proteins where

only a small fraction of the total protein is localized to nucleoli at one specific time and therefore may not be readily detected by fluorescence microscopy (for example, see [37]). Therefore, at least some of the proteins analyzed here that did not show obvious accumulation in nucleoli when fused to YFP may nonetheless correspond to endogenous nucleolar factors. Correspondingly, we anticipate that the great majority, if not all, of the proteins listed in Table S1 are nucleolar proteins.

The nucleolar proteome includes a wide range of different proteins (see Figure 5 and Table S1). The most common motifs found in these proteins are nucleic acid and nucleotide binding domains. The DEAD-box helicase motifs characteristic of the superfamily of RNA-dependent ATPases were present in 22 proteins, suggesting that the control of RNA base pairing interactions may be an important feature of nucleolar function. This conclusion is supported by the recent work of Bickmore and colleagues who found using a gene-trap method that the DEAD box motif was the most common motif found in the pool of nucleolar proteins detected [47]. A more detailed bioinformatic analysis of the nucleolar proteome is currently in progress and will be presented elsewhere. However, preliminary analysis indicates that there are no simple targeting motifs shared by all the nucleolar proteins, suggesting that their localization to the nucleolus involves a variety of mechanisms.

The major function known for the nucleolus is the transcription and processing of rRNAs and their subsequent assembly into ribosomal subunits. Consistent with this, the nucleolar proteome includes many ribosomal proteins, processing factors, and components required for transcription of the rRNA gene clusters, as well as homologs of genes involved in these processes in other organisms. This analysis has also identified novel proteins that may be involved in these processes, such as NNP8, a protein related to the yeast pre-rRNA processing factor Ski6p, and the putative exoribonucleases NNP15 and NNP11. However, not all the proteins appear to have functions associated with the known roles of the nucleolus. For example, we have also identified several protein translation factors, including eIF4A, eIF5A, eIF6, and ETF1 (peptide chain release factor subunit 1), which bind to active ribosomes in the cytoplasm but were not known to preassemble in the nucleolus with the individual ribosomal subunits [48]. Interestingly, eIF6 was recently reported to localize to nucleoli in mast cells [49], consistent with the observation reported here for HeLa cells. The presence of these translation factors raises the interesting possibility of nucleolar/nuclear translation activity [50, 51].

The high protein complexity of the nucleolus revealed in Table S1 implies that either the biogenesis of ribosomes is a surprisingly complex process and/or that the nucleolus carries out additional functions, consistent with the theory of a plurifunctional nucleolus [24, 26]. One of these additional roles may be in regulating the localization of nuclear factors in a cell cycle-dependent manner and thereby controlling their access to interaction partners [22, 31]. The purified nucleoli studied in this project were isolated from unsynchronized cells and therefore the proteins listed in Table S1 will include factors that may associate with nucleoli only at specific stages of the cell cycle. For example, among the proteins identified, the DEAD box protein p68 associates with nucleoli specifically at telophase [52], while the Bloom's Syndrome factor (BLM) specifically accumulates in nucleoli only during S phase [53]. It will be interesting in future to determine whether any of the novel protein factors show similar properties.

In addition to cell cycle-specific associations, other nuclear factors can show facultative interactions with the nucleolus that depend upon the metabolic state of the cell. This was illustrated in this study by the 11 proteins that associate with nucleoli after treatment of cells with the transcription inhibitor Actinomycin D. In the accompanying article by Fox et al. (this issue of *Current Biology*), we describe the detailed characterization of one such protein identified here, the novel factor PSP1 [37]. PSP1 shows a nucleolar relationship affected by the transcription state of the cell. Photobleaching analyses have revealed that PSP1 constantly traffics through the nucleolus yet has a steady-state accumulation within novel nucleoplasmic structures called "paraspeckles." However, when transcription is inhibited, PSP1 rapidly relocalizes to the nucleolar periphery where it accumulates within perinucleolar caps. This is consistent with the MS data reported here showing PSP1 is enriched in preparations of purified nucleoli from Actinomycin D-treated cells (Figure 6). These data are also in agreement with recent evidence showing other nuclear proteins are dynamic and can traffic rapidly between different nuclear compartments [8]. It is interesting that the 11 proteins found here enriched in the preparations of nucleoli purified from Actinomycin D-treated cells have obvious features in common. Many of the proteins have RNA binding motifs, and several have been implicated in similar modes of transcriptional control. For example, p72 and p68 can act as estrogen receptor coactivators [54, 55]; PSF [56] and p54nrb can act as corepressors interacting with the DNA binding domains of nuclear hormone receptors; and PSP2 (also called CoAA) coactivates transcription and interacts with a thyroid hormone receptor binding protein [57]. These data suggest a possible novel role for the nucleolus, involving the cycling of transcription factors, and it will be interesting to investigate this further in the future. It is likely that further analysis will also identify additional factors whose association with nucleoli is either increased or decreased by Actinomycin D treatment.

The proteins listed in Table S1 represent a core of major nucleolar proteins, most of which were detected multiple times in independent preparations of nucleoli. Clearly, it does not include every human protein that may associate or interact with nucleoli in vivo, and some previously reported nucleolar proteins were not detected. Weakly associated factors may be lost during the purification protocol, while other proteins may interact with nucleoli only under specific metabolic conditions that were not sampled in our study. Very low abundance proteins or factors with unusual structures or modifications may have escaped detection with the methods we have used so far. Our analysis of HeLa cell nucleoli will also exclude some cell type-specific nucleolar proteins. We intend to extend our coverage of nucleolar proteins by further proteomic analysis of nucleoli

purified from a variety of sources, including primary cells and cell lines derived from different tissues. In addition, it is anticipated that future improvements in the sensitivity of protein detection and analysis methods will enable the identification of additional nucleolar proteins.

In conclusion, although more work remains to be done, we believe that the nucleolar proteome detailed here already represents a significant advance toward defining a comprehensive inventory of nucleolar proteins. These data should be of value for future studies on the range of biological roles performed by the nucleolus as well as the mechanisms involved in its assembly and function.

Experimental Procedures

Nucleolar Isolation

Nucleoli were prepared from HeLa cell nuclei (Computer Cell Culture Centre, Belgium), using a method based on that first described by Muramatsu and coworkers in 1963 [58]. Aliquots (250 μl) containing ~1 × 10^8 nuclei were washed three times with PBS, resuspended in 5 ml buffer A (10 mM HEPES-KOH [pH 7.9], 1.5 mM $MgCl_2$, 10 mM KCL, 0.5 mM DTT), and dounce homogenized ten times using a tight pestle. Dounced nuclei were centrifuged at 228 × *g* for 5 min at 4°C. The nuclear pellet was resuspended in 3 ml 0.25 M sucrose, 10 mM $MgCl_2$, and layered over 3 ml 0.35 M sucrose, 0.5 mM $MgCl_2$, and centrifuged at 1430 × *g* for 5 min at 4°C. The clean, pelleted nuclei were resuspended in 3 ml 0.35 M sucrose, 0.5 mM $MgCl_2$, and sonicated for 6 × 10 s using a microtip probe and a Misonix XL 2020 sonicator at power setting 5. The sonicate was checked using phase contrast microscopy, ensuring that there were no intact cells and that the nucleoli were readily observed as dense, refractile bodies. The sonicated sample was then layered over 3 ml 0.88 M sucrose, 0.5 mM $MgCl_2$ and centrifuged at 2800 × *g* for 10 min at 4°C. The pellet contained the nucleoli, while the supernatant consisted of the nucleoplasmic fraction. The nucleoli were then washed by resuspension in 500 μl of 0.35 M sucrose, 0.5 mM $MgCl_2$, followed by centrifugation at 2000 × *g* for 2 min at 4°C.

Microscopy and Immunostaining of Samples

Whole HeLa cells were grown on coverslips, washed in PBS, then fixed for 5 min with paraformaldehyde (4% in 10 mM pipes [pH 6.8], 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, and 2 mM EDTA) and immunolabeled. Permeabilization was performed with 1% Triton X-100 in PBS for 15 min at room temperature. Immunofluorescence labeling was carried out on fixed permeabilized cells by 1 hr incubation in 100 μl primary antibody (diluted in PBST [PBS plus 0.05% Tween 20]), followed by three washes in PBST and incubation for 1 hr in 100 μl secondary antibody (diluted in PBST) followed by three washes in PBST. Antibodies used were anti-p80 coilin monoclonal 5P10 (dilution 1:10) and anti-nucleolin 7G2 (1:1000, kindly provided by G. Dreyfuss) and anti-mouse FITC-conjugated secondary antibody (Jackson Lab, 1:250). Prior to mounting, samples were stained with 0.66 mM Pyronin Y (Sigma) for 2 s, then mounted using DABCO containing Mowiol. Purified nucleoli were immobilized on poly-L-lysine slides (BDH) and air dried. After rehydration with PBS (5 min), the nucleoli were incubated with monoclonal anti-nucleolin antibody 7G2 (1:100) or monoclonal anti-coilin antibody 5P10 (1:5) for 30 min. The slides were washed with PBS (3 × 5 min) and incubated with FITC-conjugated secondary antibody (Jackson Lab, 1:250) for 30 min. The nucleoli were counterstained with 0.66 mM Pyronin Y (Sigma) for 1 min. After washing as above, the nucleoli were embedded in DABCO containing Mowiol. Images were obtained using a Zeiss LSM 410 confocal laser scanning microscope.

Isolated nucleoli were processed for transmission electron microscopy using standard methods. In brief, nucleoli were centrifuged at 5000 rpm for 2 min. The pelleted nucleoli were washed briefly in PBS, fixed in 3.7% paraformaldehyde in PBS for 20 min at room temperature, washed three times in PBS, washed in water, postfixed in 1% osmium tetroxide in water for 15 min at room temperature, washed three times in water, dehydrated in 70% ethanol for 2 × 5 min, contrasted with 3% uranyl acetate in 70% ethanol for 15 min at room temperature, washed in 70% ethanol further dehydrated through 90% ethanol, 1 × 5 min, 100% ethanol, 3 × 5 min, propylene oxide, 2 × 5 min, 50:50 propylene oxide:epoxy resin mix for 2 hr at room temperature before embedding in epoxy resin. Sections were cut using a Reichart ultracut ultramicrotome and visualized in a JOEL 1200EX TEM.

Electrophoresis and Immunoblotting

For 1D SDS/PAGE, purified nucleoli were dissolved in 1× LDS sample buffer (Novex) + 100 mM DTT and heated at 70°C for 10 min. Nucleolar proteins were then separated on 3%–8% SDS Tris-Acetate gels and 4%–12% Bis-Tris gels (Novex) and stained with Coomassie colloidal blue according to the manufacturer's instructions (Novex). For 2D SDS/PAGE, IPG strips (Pharmacia) were in-gel rehydrated with samples dissolved in 2 M thiourea, 6 M urea, 1% CHAPS, 0.4% DTT, and 0.5% Pharmalyte (either 3-10 or 4-7). Proteins were focused for 40000 V in the first dimension, separated on vertical 12.5% SDS gels in the second dimension, and stained with silver. Protein slices/spots were excised, deposited in 96-well plates, in-gel reduced, alkylated with iodoacetamide, and digested with trypsin as previously described [59]. The resulting peptides were analyzed by MALDI-MS and nanoES tandem MS (see below).

For immunoblotting, protein samples separated by 1D SDS/PAGE (see above) were subsequently transferred onto nitrocellulose membrane using a submarine system (Novex) and buffer containing 12 mM Tris, 100 mM glycine, and 20% methanol. Following blocking with 5% milk powder in PBS + 0.05% Tween 20, the membranes were incubated with one of the following antibodies: rabbit anti-fibrillarin, Fib 42 (1:10000, gift from F. Fuller-Pace), mouse monoclonal anti-nucleolin, 7G2 (1:1000, G. Dreyfuss), mouse monoclonal anti-Lamin B, LN43.2, (gift from E.B. Lane), and nucleoporin NUP62 was detected using mouse monoclonal 414 (1:5000, Babco), then bound antibody was probed using anti-rabbit HRP conjugate (1:2000 dilution) or anti-mouse HRP conjugate (1:5000 dilution) (Pierce Chemical Co.) in PBS containing 5% milk powder and 0.05% Tween 20, and detected via chemiluminescence with ECL (Amersham Pharmacia Biotech).

Mass Spectrometry

MALDI mass spectra were acquired automatically on a Bruker REFLEX III reflectron time of flight (TOF) mass spectrometer (Bruker-Franzen, Bremen, Germany) as previously described [60, 61]. Peptides were also analyzed by nanoES tandem MS on a quadrupole time-of-flight mass spectrometer (QSTAR Pulsar, PE Sciex, Toronto, Canada) equipped with a nanoES ion source (MDS-Proteomics, Denmark). The peptide mixtures from in-gel digests were purified and concentrated prior to nanoESMS as described [62], except that the remaining supernatant were loaded on columns in parallel from 96-well plates.

Databases and Searching

Nonredundant (nrdb), predicted (IPI.1), expressed sequence tags and finished and unfinished human genome sequences (phases 0-3) were downloaded (ftp://ncbi.nlm.nih.gov/genbank and ftp://ftp.ensembl.org/IPI/) and converted into FASTA formatted sequence index files accepted by the PepSea database search software system (MDS proteomics, Denmark). Peptide masses measured by MALDI-MS were searched with 40 ppm mass accuracy, and "used" peptide masses were subtracted prior to repeated rounds of database searches. Peptide sequence tags [36] were assigned from tandem mass spectra assisted by the Inspector software (MDS Proteomics, Denmark) and used for database searching in real time against the respective databases using PepSea software. Calculated peptide ion and fragment ion masses of retrieved sequences were displayed in the mass spectra to establish unambiguous identification and to identify additional proteins by selection of unexplained peaks for MS/MS. Genome searching were performed as recently described [38].

Gene Prediction

Two WWW-based gene prediction programs were employed for further characterization of identified coding regions of the human

genome: GENSCAN, at the Massachusetts Institute of Technology (MIT, Boston, http://genes.mit.edu/GENSCAN.html), and HMMgene, at the Centre or Biological Sequence Analysis (CBS, The Technical University of Denmark, Lyngby, Denmark, http://www.cbs.dtu.dk/services/HMMgene). Nucleotide sequences corresponding to identified peptides were defined as coding and used as constraints in HMMgene predictions. The GPMAW software (Lighthouse Data, Denmark) were modified to calculate m/z values of exon-exon spanning peptides.

Cloning Nucleolar Proteins, Tagging with Fluorescent Protein, and Expression in HeLa Cells

Full-length cDNAs encoding selected nucleolar candidate proteins were amplified by PCR using specific ESTs (UK HGMP Resource Centre) or HeLa cDNA library (Clontech) as templates. Each construct was cloned into the pEYFP-C1 vector (Clontech), using restriction sites engineered onto the 5′ ends of each amplification primer. Details of ESTs, primer sequences, and restriction sites are available on request.

HeLa cells were grown in DMEM supplemented with 10% fetal calf serum and 1% penicillin streptomycin (Life Technologies). Fluorescent protein fusion constructs were transfected into HeLa cells seeded onto coverslips using Effectene reagent (Qiagen). At 16–18 hr posttransfection, the cells were washed in PBS and fixed in paraformaldehyde as above. Coverslips were mounted in Mowiol/Dabco onto slides for imaging. Images were obtained using a Zeiss LSM 410 confocal laser scanning microscope.

Actinomycin D Treatment of HeLa Cells

To study the changes in nucleolar composition caused by inhibition of transcription, we incubated HeLa cells in 5 μg/ml Actinomycin D (Sigma) diluted in DMEM (Gibco) with 10% fetal calf serum (Gibco). After 3 hr, the cells were harvested by trypsinization, washed three times with cold PBS, then resuspended at $\sim 5 \times 10^7$/ml with 10 mM HEPES-KOH (pH 7.9), 10 mM KCl, 1.5 mM MgCl2, 0.5 mM DTT. The cell suspension was incubated on ice for 5 min and then disrupted by dounce homogenation (ten strokes) on ice. The nuclei released by homogenation were pelleted by centrifugation at $1000 \times g$ for 5 min. Nucleoli were isolated from these nuclei as described above.

Supplementary Material

Supplementary Material including a table of proteins identified from nucleoli isolated from cultured human cells can be found at http://images.cellpress.com/supmat/supmatin.htm.

Acknowledgments

A.I.L. is a Wellcome Trust Principal Research Fellow and is funded by a Wellcome Trust Programme grant. The authors thank Peter Højrup for modifying GPMAW. Work in M.M.'s laboratory is funded by a Danish National Research Foundation grant to the Centre for Experimental BioInformatics. A.K.L.L. is funded by a Croucher studentship; Y.W.L. is funded by a Croucher postdoctoral fellowship; A.H.F. is funded by a Wellcome Trust International Travelling Fellowship; and C.E.L. is funded by the Wellcome Trust.

Received: October 15, 2001
Revised: November 14, 2001
Accepted: November 14, 2001
Published: January 8, 2002

References

1. Misteli, T., and Spector, D.L. (1998). The cellular organization of gene expression. Curr. Opin. Cell Biol. *10*, 323–331.
2. Lamond, A.I., and Earnshaw, W.C. (1998). Structure and function in the nucleus. Science *280*, 547–553.
3. Lewis, J.D., and Tollervey, D. (2000). Like attracts like: getting RNA processing together in the nucleus. Science *288*, 1385–1389.
4. Dundr, M., and Misteli, T. (2001). Functional architecture in the cell nucleus. Biochem. J. *356*, 297–310.
5. Matera, A.G. (1999). Nuclear bodies: multifaceted subdomains of the interchromatin space. Trends Cell Biol. *9*, 302–309.
6. Phair, R.D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. Nature *404*, 604–609.
7. Kruhlak, M.J., Lever, M.A., Fischle, W., Verdin, E., Bazett-Jones, D.P., and Hendzel, M.J. (2000). Reduced mobility of the alternate splicing factor (ASF) through the nucleoplasm and steady state speckle compartments. J. Cell Biol. *150*, 41–51.
8. Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. Science *291*, 843–847.
9. Platani, M., Goldberg, I., Swedlow, J.R., and Lamond, A.I. (2000). In vivo analysis of Cajal body movement, separation, and joining in live human cells. J. Cell Biol. *151*, 1561–1574.
10. Swedlow, J.R., and Lamond, A.I. (2001). Nuclear dynamics: where genes are and how they got there. Genome Biol. *2*. Published online March 9, 2001. reviews0002.1-0002.7.
11. Lamond, A., and Mann, M. (1997). Cell biology and the genome projects—a concerted strategy for characterizing multiprotein complexes by using mass spectrometry. Trends Cell Biol. *7*, 139–142.
12. Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. Nature *405*, 837–846.
13. Blackstock, W.P., and Weir, M.P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. Trends Biotechnol. *17*, 121–127.
14. Aebersold, R., and Goodlett, D.R. (2001). Mass spectrometry in Proteomics. Chem. Rev. *101*, 269–295.
15. Mintz, P.J., Patterson, S.D., Neuwald, A.F., Spahr, C.S., and Spector, D.L. (1999). Purification and biochemical characterization of interchromatin granule clusters. EMBO J. *18*, 4308–4320.
16. Neubauer, G., Gottschalk, A., Fabrizio, P., Seraphin, B., Luhrmann, R., and Mann, M. (1997). Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc. Natl. Acad. Sci. USA *94*, 385–390.
17. Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A., and Mann, M. (1998). Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. Nat. Genet. *20*, 46–50.
18. Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. J. Cell Biol. *148*, 635–651.
19. Shaw, P.J., and Jordan, E.G. (1995). The nucleolus. Annu. Rev. Cell Dev. Biol. *11*, 93–121.
20. Scheer, U., and Hock, R. (1999). Structure and function of the nucleolus. Curr. Opin. Cell Biol. *11*, 385–390.
21. Venema, J., and Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. Annu. Rev. Genet. *33*, 261–311.
22. Carmo-Fonseca, M., Mendes-Soares, L., and Campos, I. (2000). To be or not to be in the nucleolus. Nat. Cell Biol. *2*, E107–112.
23. Dundr, M., Misteli, T., and Olson, M.O. (2000). The dynamics of postmitotic reassembly of the nucleolus. J. Cell Biol. *150*, 433–446.
24. Pederson, T. (1998). The plurifunctional nucleolus. Nucleic Acids Res. *26*, 3871–3876.
25. Olson, M.O., Dundr, M., and Szebeni, A. (2000). The nucleolus: an old factory with unexpected capabilities. Trends Cell Biol. *10*, 189–196.
26. Schneiter, R., Kadowaki, T., and Tartakoff, A.M. (1995). mRNA transport in yeast: time to reinvestigate the functions of the nucleolus. Mol. Biol. Cell *6*, 357–370.
27. Politz, J.C., Yarovoi, S., Kilroy, S.M., Gowda, K., Zwieb, C., and Pederson, T. (2000). Signal recognition particle components in the nucleolus. Proc. Natl. Acad. Sci. USA *97*, 55–60.
28. Lange, T.S., and Gerbi, S.A. (2000). Transient nucleolar localization of U6 small nuclear RNA in Xenopus Laevis oocytes. Mol. Biol. Cell *11*, 2419–2428.
29. Mitchell, J.R., Wood, E., and Collins, K. (1999). A telomerase component is defective in the human disease dyskeratosis congenita. Nature *402*, 551–555.
30. Bertrand, E., Houser-Scott, F., Kendall, A., Singer, R.H., and Engelke, D.R. (1998). Nucleolar localization of early tRNA processing. Genes Dev. *12*, 2463–2468.

31. Visintin, R., and Amon, A. (2000). The nucleolus: the magician's hat for cell cycle tricks. Curr. Opin. Cell Biol. *12*, 752.
32. Huang, S. (2000). Review: perinucleolar structures. J. Struct. Biol. *129*, 233–240.
33. Gall, J.G. (2000). Cajal bodies: the first 100 years. Annu. Rev. Cell Dev. Biol. *16*, 273–300.
34. Vincent, W.S. (1955). Structure and chemistry of nucleoli. Int. Rev. Cytol. *4*, 269–298.
35. Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Boucherie, H., and Mann, M. (1996). Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. USA *93*, 14440–14445.
36. Mann, M., and Wilm, M. (1994). Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal. Chem. *66*, 4390–4399.
37. Fox, A.H., Lam, Y.W., Leung, A., Lyon, C.E., Andersen, J.S., Mann, M., and Lamond, A.I. (2002). Paraspeckles: A novel nuclear domain. Curr. Biol. *12*, this issue, 13–25.
38. Kuster, B., Mortensen, P., Andersen, J.S., and Mann, M. (2001). Mass spectrometry allows direct identification of proteins in large genomes. Proteomics *5*, 641–650.
39. Zirwes, R.F., Eilbracht, J., Kneissel, S., and Schmidt-Zachmann, M.S. (2000). A novel helicase-type protein in the nucleolus: protein NOH61. Mol. Biol. Cell *11*, 1153–1167.
40. Strezoska, Z., Pestov, D.G., and Lau, L.F. (2000). Bop1 is a mouse WD40 repeat nucleolar protein involved in 28S and 5. 8S RRNA processing and 60S ribosome biogenesis. Mol. Cell. Biol. *20*, 5516–5528.
41. Tanner, N.K., and Linder, P. (2001). DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol. Cell *8*, 251–262.
42. Ghosh, S. (1976). The nucleolar structure. Int. Rev. Cytol. *44*, 1–28.
43. Raska, I., Ochs, R.L., Andrade, L.E., Chan, E.K., Burlingame, R., Peebles, C., Gruol, D., and Tan, E.M. (1990). Association between the nucleolus and the coiled body. J. Struct. Biol. *104*, 120–127.
44. Carmo-Fonseca, M., Pepperkok, R., Carvalho, M.T., and Lamond, A.I. (1992). Transcription-dependent colocalization of the U1, U2, U4/U6, and U5 snRNPs in coiled bodies. J. Cell Biol. *117*, 1–14.
45. Dye, B.T., and Patton, J.G. (2001). An RNA recognition motif (RRM) is required for the localization of PTB-associated splicing factor (PSF) to subnuclear speckles. Exp. Cell Res. *263*, 131–144.
46. Xirodimas, D., Saville, M.K., Edling, C., Lane, D.P., and Lain, S. (2001). Different effects of p14ARF on the levels of ubiquitinated p53 and Mdm2 in vivo. Oncogene *20*, 4972–4983.
47. 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.
48. Pestova, T.V., Kolupaeva, V.G., Lomakin, I.B., Pilipenko, E.V., Shatsky, I.N., Agol, V.I., and Hellen, C.U. (2001). Molecular mechanisms of translation initiation in eukaryotes. Proc. Natl. Acad. Sci. USA *98*, 7029–7036.
49. Oh, C.K., Filler, S.G., and Cho, S.H. (2001). Eukaryotic translation initiation factor-6 enhances histamine and IL-2 production in mast cells. J. Immunol. *166*, 3606–3611.
50. Pederson, T., and Politz, J.C. (2000). The nucleolus and the four ribonucleoproteins of translation. J. Cell Biol. *148*, 1091–1095.
51. Iborra, F.J., Jackson, D.A., and Cook, P.R. (2001). Coupled transcription and translation within nuclei of mammalian cells. Science *293*, 1139–1142.
52. Nicol, S.M., Causevic, M., Prescott, A.R., and Fuller-Pace, F.V. (2000). The nuclear DEAD box RNA helicase p68 interacts with the nucleolar protein fibrillarin and colocalizes specifically in nascent nucleoli during telophase. Exp. Cell Res. *257*, 272–280.
53. Sanz, M.M., Proytcheva, M., Ellis, N.A., Holloman, W.K., and German, J. (2000). BLM, the Bloom's syndrome protein, varies during the cell cycle in its amount, distribution, and co-localization with other nuclear proteins. Cytogenet. Cell Genet. *91*, 217–223.
54. Endoh, H., Maruyama, K., Masuhiro, Y., Kobayashi, Y., Goto, M., Tai, H., Yanagisawa, J., Metzger, D., Hashimoto, S., and Kato, S. (1999). Purification and identification of p68 RNA helicase acting as a transcriptional coactivator specific for the activation function 1 of human estrogen receptor alpha. Mol. Cell. Biol. *19*, 5363–5372.
55. Watanabe, M., Yanagisawa, J., Kitagawa, H., Takeyama, K., Ogawa, S., Arao, Y., Suzawa, M., Kobayashi, Y., Yano, T., Yoshikawa, H., et al. (2001). A subfamily of RNA-binding DEAD-box proteins acts as an estrogen receptor alpha coactivator through the N-terminal activation domain (AF-1) with an RNA coactivator, SRA. EMBO J. *20*, 1341–1352.
56. Mathur, M., Tucker, P.W., and Samuels, H.H. (2001). PSF is a novel corepressor that mediates its effect through Sin3A and the DNA binding domain of nuclear hormone receptors. Mol. Cell. Biol. *21*, 2298–2311.
57. Iwasaki, T., Chin, W.W., and Ko, L. (2001). Identification and characterization of rrm-containing coactivator activator (coaa) as trbp-interacting protein, and its splice variant as a coactivator modulator (coam). J. Biol. Chem. *276*, 33375–33383.
58. Muramatsu, M., Smetana, K., and Busch, H. (1963). Quantitative aspects of isolation of nucleoli of the Walker carcinosarcoma and liver of the rat. Cancer Res. *25*, 693–697.
59. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. *68*, 850–858.
60. Vorm, O., Roepstorff, P., and Mann, M. (1994). Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. Anal. Chem. *66*, 3281–3287.
61. Jensen, O.N., Mortensen, P., Vorm, O., and Mann, M. (1997). Automation of matrix-assisted laser desorption/ionization mass spectrometry using fuzzy logic feedback control. Anal. Chem. *69*, 1706–1714.
62. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano- electrospray mass spectrometry. Nature *379*, 466–469.