multimedia

REVIEW ARTICLE

Bioinformatic analysis of the nucleolus

Anthony K. L. LEUNG*[1], Jens S. ANDERSEN†, Matthias MANN† and Angus I. LAMOND*

*Division of Gene Regulation and Expression, School of Life Sciences, Wellcome Trust Biocentre, University of Dundee, Dundee DD1 5EH, Scotland, U.K., and †Center for Experimental BioInformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230, Odense M, Denmark

The nucleolus is a plurifunctional, nuclear organelle, which is responsible for ribosome biogenesis and many other functions in eukaryotes, including RNA processing, viral replication and tumour suppression. Our knowledge of the human nucleolar proteome has been expanded dramatically by the two recent MS studies on isolated nucleoli from HeLa cells [Andersen, Lyon, Fox, Leung, Lam, Steen, Mann and Lamond (2002) Curr. Biol. **12**, 1–11; Scherl, Coute, Deon, Calle, Kindbeiter, Sanchez, Greco, Hochstrasser and Diaz (2002) Mol. Biol. Cell **13**, 4100–4109]. Nearly 400 proteins were identified within the nucleolar proteome so far in humans. Approx. 12 % of the identified proteins were previously shown to be nucleolar in human cells and, as expected, nearly all of the known housekeeping proteins required for ribosome biogenesis were identified in these analyses. Surprisingly, approx. 30 % represented either novel or uncharacterized proteins. This review focuses on how to apply the derived knowledge of this newly recognized nucleolar proteome, such as their amino acid/peptide composition and their homologies across species, to explore the function and dynamics of the nucleolus, and suggests ways to identify, *in silico*, possible functions of the novel/uncharacterized proteins and potential interaction networks within the human nucleolus, or between the nucleolus and other nuclear organelles, by drawing resources from the public domain.

Key words: amino acid composition, bioinformatics, nucleolus, profiling, proteomics, sequence homology.

BACKGROUND

A eukaryotic cell, by definition, segregates its DNA content within an internal compartment known as the nucleus. As early as in the 1800s, it was already obvious from light-microscopy studies that the interior of the nucleus was not uniform, notably with the discovery of the nucleolus. As noted by Gabriel Gustav Valentin, who first described the presence of the nucleolus in somatic cells [1,2], "*In every cell, without exception, there exists a somewhat darker-appearing and compact nucleus of a round or nearly round shape. Mostly it is located in the centre of a specific cell, composed of finely granular material and containing in its interior an exactly spheroidal body which in this way forms a kind of secondary nucleus within the first one*."

Because the nucleolus is dense and highly refractile, it was the first subnuclear organelle intensely studied by light microscopists (Figure 1), and both its variability in size and number and the dynamic aspects of this organelle during mitosis were noted within 50 years of its discovery [3]. It was, however, not until the 1960s that the role of the nucleolus was established as the 'ribosome factory', where the rDNAs are kept and transcribed and the rRNA transcripts are processed and assembled with ribosomal proteins to form ribosomes. Although all rDNA repeats are housed within nucleoli, only half of the rDNA repeat copies are utilized even at the maximal transcriptional output [4]. Yet, surprisingly, this expends about 50–80 % of the total transcriptional activity in transcribing the rDNA and making other small nucleolar RNAs required for rRNA processing in an actively cycling eukaryotic cell [5]. However, an exclusive role of the nucleolus as a ribosome factory may not explain the recent discovery of viral, cell-cycle regulatory and tumour-related proteins within this structure [6–8], where the nucleolus acts as a molecular 'safe' or 'sink' to retain proteins and thus prevent them from interacting with their potential downstream partners until a specific cell-cycle stage, or metabolic state. This view is consistent with analyses showing that many human nucleolar proteins, identified independently in two recent organelle-directed proteomic studies, are not obviously related to ribosomal biogenesis [9,10].

Owing to their inherent high density, human nucleoli can be isolated by sucrose-gradient centrifugation following the disruption of nuclei by sonication. The recent advances in protein-separation strategies and MS allow high-throughput, reproducible identification of the human nucleolar proteome [9,10]. In these studies, the proteins in purified samples of human nucleoli were separated by either one-dimensional (1D) or two-dimensional (2D) SDS/PAGE, followed by in-gel trypsinization, and the peptides were analysed by tandem mass spectrometry (MS/MS) to identify not only the peptide masses, but also their amino acid sequences [11,12]. Alternatively, the isolated nucleoli were solubilized, then trypsinized in solution and the resulting peptides were separated by liquid chromatography (LC) before analysing by MS/MS. The identified peptides from both 1D/2D gel–MS/MS and LC–MS/MS were then used to interrogate the NCBI (National Center for Biotechnology Information) database and, in some cases, the human-genome database to identify the corresponding proteins [13].

Nearly 300 nucleolar proteins were identified using MALDI–TOF (matrix-assisted laser-desorption ionization–time-of-flight) and nanoelectrospray MS on proteins separated by a combination of 1D and 2D gel analyses [9]. Recently, the coverage of nucleolar proteins was extended to over 400 members by performing LC–MS/MS ([10]; J. S. Andersen, Y. W. Lam, A. K. Leung, A. I. Lamond and M. Mann, unpublished work). More than 90 % of the proteins were identified multiple times by three MS studies of the human nucleolar proteome, suggesting that the current data are reproducible and rigorous ([9,10]; J. S. Andersen, Y. W. Lam, A. K. Leung, A. I. Lamond and M. Mann, unpublished work). Approx. 12 % of the identified proteins were known

Abbreviations used: 1D, one-dimensional; 2D, two-dimensional; LC, liquid chromatography; MS/MS, tandem mass spectrometry; PP1, protein phosphatase-1; PSP1, paraspeckle protein 1; RNP, ribonucleoprotein particle; RRM, RNA-recognition motif; snRNP, small nuclear RNP.

[1] To whom correspondence should be addressed (e-mail a.k.l.leung@dundee.ac.uk).

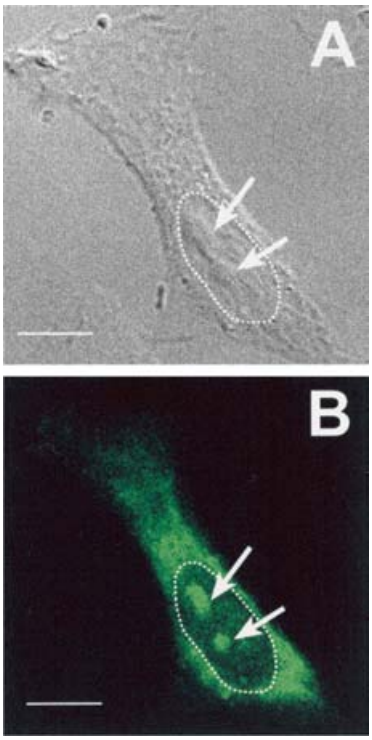


Figure 1 The nucleolus

(**A**) Because nucleoli are denser and more refractile than the nucleoplasm, they appear as darker bodies under the light microscope. (**B**) A fluorescent micrograph depicting the same cells expressing a GFP–ribosomal protein which localizes both in nucleoli and endoplasmic reticulum in the cytoplasm. Arrows indicate nucleoli and dotted lines indicate the boundaries of nuclei. Scale bar, 5 μm.

previously to be nucleolar in human cells (see the Expansion of the known human nucleolar proteome section). Surprisingly, approx. 30 % represented either novel or uncharacterized proteins (see the Bioinformatic analyses section). As expected, nearly all of the known housekeeping proteins required for ribosomal biogenesis were identified in these analyses. So, the immediate question is what are the remaining 88 % of proteins doing in the nucleolus? The present review focuses on our current understanding of the nucleolar proteome and how to apply bioinformatic analyses of this proteome database to ask questions and, more importantly, design experiments that explore the function and dynamics of the nucleolus [6,15–19].

EXPANSION OF THE KNOWN HUMAN NUCLEOLAR PROTEOME

Before 2002, 121 human proteins had been reported to localize in the nucleolus in the published literature (Table 1). In all these cases, the proteins were first identified biochemically and then their subcellular localizations were shown subsequently to be nucleolar by antibody staining and/or tagging with fluorescent proteins. This is in stark contrast with the recent proteomic approach, in which nucleolar protein components were identified en masse from highly purified samples of nucleoli. Examining Table 1, we note that more than 90 % of the human proteins that were previously reported to be involved in ribosomal biogenesis were identified multiple times, both from the 1D/2D gel and the LC approaches (Tables 1a and 1d). In general, the LC–MS/MS approach appears more sensitive than the 1D/2D gel approach in detecting peptides and thus gave rise to a higher coverage of nucleolar proteins (indicated by ticks in the LCMS/MBC columns rather than the CB02 column). Approx. 70 % of the proteins identified by 1D/2D gel–MS/MS analyses were also detected by the LC–MS/MS studies.

Even so, only less than 10 % of proteins that were previously reported to show facultative nucleolar localization, i.e. to be localized in nucleoli during only part of the cell cycle, or under certain metabolic conditions, were identified in any of the proteomic analyses (Tables 1b and 1c). This may be because the nucleoli analysed were isolated from unsynchronized HeLa cells and hence such facultative nucleolar proteins may only constitute a minor fraction of the proteins isolated. Although LC–MS/MS has already increased the sensitivity of detection, it may still not be sensitive enough to detect very low abundance factors. Future analyses will therefore focus on nucleoli isolated at specific cell-cycle stages and under particular metabolic conditions (see the Dynamic proteome and Proteome profiling sections). This should help to increase the total coverage of the human nucleolar proteome.

BIOINFORMATIC ANALYSES

The large amount of data acquired from these proteomic studies requires a systematic way to analyse and integrate it with the information already deposited in the databases that are publicly available. To facilitate these analyses, several databases were downloaded to our Unix server and were interrogated with software such as stand-alone BLAST [20,21] and in-house customized Perl scripts. The resulting information was presented in a searchable database to provide a one-stop, user-friendly gateway to other relevant databases for non-specialists where each protein was given a summary of up-to-date, annotated information (Figure 2; http://www.dundee.ac.uk/lifesciences/lamonddatabase/). The most abundant motifs within the nucleolar proteome were the RNA-binding RRM (RNA-recognition motif) domain, DEAD/H (Asp-Glu-Ala-Asp/His) box helicase domain and the protein–protein interaction WD (Trp-Asp) domain (Table 2). These motifs support the role of the nucleolus being an organelle of multiple interlinked complexes acting in both ribosomal maturation and other RNA-processing pathways [22].

An overview of the separate classes of proteins identified in the purified nucleoli is presented in Figure 3. Surprisingly, the subsequent addition of more than 100 newly identified nucleolar proteins from the LC–MS/MS data resulted in minimal change to the distribution of proteins between categories, as previously assigned based on the analysis of 271 proteins [9]. DNA- and RNA-binding proteins that are involved in DNA repair, transcription, the unwinding of nucleic acid and RNA modification, including splicing, constitute roughly a quarter of the nucleolar proteome. Approx. 16 % of the human nucleolar proteome is accounted for by translation-related factors, i.e. ribonucleoprotein particles/RNAs related to translation, such as SRP (signal recognition particle), tRNA and 5 S RNA, and nearly all the ribosomal proteins [9,10,17,23,24]. However, there is no evidence that this is related to the controversial idea of mRNA translation in the nucleus [25,26]. Yet, it should be noted that, during the 1960s, data were published reporting that amino acids could be incorporated into peptides within isolated nucleoli [27–29]. The high protein complexity of the nucleolus 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 [30,31].

Surprisingly, approx. 30 % of the nucleolar proteins are either uncharacterized or else are the products of novel open reading frames. This partly reflects the currently limited status of the human genome annotation and also presents a challenge for us to unravel the functional complexity of the nucleolus [9]. However,

if we pool information from different sources, including gene-expression activity and information on homologues in other species, this approach should help speed up our understanding of the human nucleolar proteome, as discussed in the following sections.

PATHWAY PREDICTION

The ribosomal biogenesis pathway was so far studied mainly in yeast, while its characterization in other eukaryotes, especially humans, lags far behind [32,33]. The accumulated knowledge

Table 1 Known nucleolar proteome

In order to compare our nucleolar proteome with the list of previously published nucleolar proteins, it was first necessary to define this comprehensive list. Known nucleolar proteins were identified and extracted from PubMed (1978–2002) using the MeSH (medical subject heading) key words 'amino acid sequences' and word 'nucleol*' in the abstract field. In total, 123 nucleolar proteins were thus identified; however, two of these proteins were not annotated with amino acid sequences and hence were not compared with our MS results. Part (a) shows nucleolar proteins identified in more than one MS analysis. Proteins known to be localized in nucleoli during only part of the cell cycle and/or under certain metabolic conditions are shown in parts (b) and (c) respectively. Other known proteins localized in nucleoli are shown in part (d). The column 'Accession' shows the corresponding NCBI GI (GenInfo Identifier) number of the protein. A tick (√) in the column 'CB02', 'LCMS' or 'MBC' signifies the identification of the protein by the 1D/2D gel–MS/MS approach [9], by our LC–MS/MS approach (J. S. Andersen, Y. W. Lam, A. K. Leung, A. I. Lamond and M. Mann, unpublished work) or by the MS/MS approach described by Scherl et al. [10]. The literature regarding the nucleolar localization of each protein is documented under the column 'Literature'. Note that some proteins fall into multiple categories.

(a) Known nucleolar proteins found in more than one MS analysis

Protein	Name	Accession	CB02	LCMS	MBC	Literature
BLM	Bloom syndrome	4557365	√	√		[64], Table 1(b)
BRIX	BRIX	19311012	√		√	[88]
CGI-94	CGI-94 protein	7705809	√	√		[89]
DDX21	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 21	13787209	√	√	√	[90]
DDX5	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 5 (RNA helicase, 68 kDa)	4758138	√	√	√	[91], Table 1(b)
DKC1	Dyskeratosis congenita 1, dyskerin	4503337	√	√	√	[92]
FBL	Fibrillarin	12056465	√	√	√	[93]
H2BFE	H2B histone family, member E	4504263	√	√	√	[94]
hPOP1/KIAA0061	KIAA0061 protein	23097292	√	√		[95]
HSPA1B	Heat shock 70 kDa protein 1B	5123454	√		√	[96], Table 1(c)
HUMAUANTIG	Nucleolar GTPase	3334276	√	√	√	[97]
Ku70/G22P1	Thyroid autoantigen 70 kDa (Ku antigen)	4503841	√	√	√	[98]
MKI67	Antigen identified by monoclonal antibody Ki-67	19923217	√	√		[99]
MKI67IP	MKI67 (FHA domain) interacting nucleolar phosphoprotein	21314753	√	√	√	[100]
MPHOSPH10	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)	2230873	√	√	√	[101]
NCL	Nucleolin	4885511	√	√	√	[102]
NHP2L1	NHP2 non-histone chromosome protein 2-like 1 (*S. cerevisiae*)	4826860	√	√	√	[103]
NOH61	Putative nucleolar RNA helicase	9506931	√	√	√	[104]
NOL1	Nucleolar protein 1 (120 kDa)	5453792	√	√	√	[105]
NOL5A	Nucleolar protein 5A (56 kDa with KKE/D repeat)	5453794	√	√	√	[106]
NOLC1	Nucleolar and coiled-body phosphoprotein 1	4758860	√	√		[107]
NOP5/NOP58	Nucleolar protein NOP5/NOP58	7706254	√	√	√	[108]
Nop52/D21S2056E	DNA segment on chromosome 21 (unique) 2056 expressed sequence	4503247	√	√	√	[109]
NPM1	Nucleophosmin (nucleolar phosphoprotein B23, numatrin)	10835063	√	√	√	[110]
Nrap	Nucleolar RNA-associated protein	18644728	√	√	√	[111]
PMSCL1	Polymyositis/scleroderma autoantigen 1 (75 kDa)	4826922	√	√		[112]
PMSCL2	Polymyositis/scleroderma autoantigen 2 (100 kDa)	4505917	√	√	√	[113]
PPP1CC	Protein phosphatase 1, catalytic subunit, γ isoform	4506007	√	√		[114]
PTBP1	Polypyrimidine tract-binding protein 1	4506243	√	√	√	[115]
RPL22	Ribosomal protein L22	4506613	√	√	√	[116]
RPL5	Ribosomal protein L5	14591909	√	√		[117]
RPL6	Ribosomal protein L6	16753227	√	√	√	[118]
RPL9	Ribosomal protein L9	15431303	√	√	√	[119]
RPS6	Ribosomal protein S6	17158044	√	√		[120]
SURF6	Surfeit 6	19557702	√	√		[121]
TCOF1	Treacher Collins–Franceschetti syndrome 1	4507411	√	√		[122]
TOP2B	Topoisomerase (DNA) IIβ (180 kDa)	19913408	√	√		[123]
U3-55K	U3 snoRNP (small nucleolar RNP)-associated 55 kDa protein	4759276	√	√	√	[124]
UBTF	Upstream binding transcription factor, RNA polymerase I	7657671	√	√		[125]

(b) Known proteins localized in nucleoli during part of the cell cycle

Protein	Name	Accession	CB02	LCMS	MBC	Literature
S phase						
BLM	Bloom syndrome	4557365	√	√		[64]
PCNA	Proliferating cell nuclear antigen	4505641				[93]
G_2/M phase						
CENPC1	Centromere protein C1	4502779				[126]
STK18	Serine/threonine kinase 18	21361433				[127]
Telophase						
DDX5	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 5 (RNA helicase, 68 kDa)	4758138	√	√	√	[91]

Table continues on following page.

Table 1 (contd.)

(c) Known proteins localized in nucleoli under certain metabolic conditions

Protein	Name	Accession	CB02	LCMS	MBC	Literature
Apoptosis						
DEDD	Death-effector-domain-containing	14670394				[128]
VIL2	Villin 2 (ezrin)	31283				[129]
Growth-related						
ANG	Angiogenin, ribonuclease, RNase A family, 5	4557313				[130]
AREG	Amphiregulin (schwannoma-derived growth factor)	4502199				[131]
FGF2	Fibroblast growth factor 2 (basic)	15451898				[132]
FGF3	Fibroblast growth factor 3 [murine mammary tumour virus integration site (v-int-2) oncogene homologue]	4885233				[133]
PTHLH	Parathyroid-hormone-like hormone	4506269				[134]
Heat shock						
HSP105B	Heat shock 105 kDa	2495344				[135]
HSPA1B	Heat shock 70 kDa protein 1B	5123454	✓		✓	[96]
HSPB1	Heat shock 27 kDa protein 1	4504517				[136]
PPID	Peptidylprolyl isomerase D (cyclophilin D)	4826932				[137]
Interferon induction						
PRKR	Protein kinase, interferon-inducible double-stranded RNA dependent	4506103				[138]
SP110a	SP110 nuclear body protein isoform a	17986254				[139]
SP110b	SP110 nuclear body protein isoform b	17986256				[139]
SP110c	SP110 nuclear body protein isoform c	17986252				[139]
UV induction						
ING1	Inhibitor of growth family, member 1	19923771				[140]
RAD17	RAD17 homologue (*Schizosaccharomyces pombe*)	4506383				[103]
Others						
ARL4	ARF-like protein 4	5031603				[141]
CALR	Calreticulin	4757900				[142]
CBFA2T3	Core-binding factor, runt domain, α subunit 2; translocated to, 3	20127532				[143]
DNAJB1	DnaJ (Hsp40) homologue, subfamily B, member 1	1706473				[144]
GAS41	Glioma-amplified sequence-41	5729838				[145]
MDM2	Mdm2, transformed 3T3 cell double minute 2, p53 binding protein (mouse)	4505137				[146]
OXR1	Oxidation resistance 1	8922241				[147]
p14ARF	p14 ARF	17738294				[148]
PNMA1	Paraneoplastic antigen MA1	14719430				[149]
SE20-4	Cutaneous T-cell lymphoma-associated tumour antigen se20-4; differentially expressed nucleolar TGF-β1 target protein (DENTT)	14861044				[150]
TP53	Tumour protein p53 (Li–Fraumeni syndrome)	8400738				[151]
YY1	YY1 transcription factor	4507955				[152]
ZFP37	Zinc-finger protein 37 homologue (mouse)	4507963				[153]
ZNF259	Zinc-finger protein 259	4508021				[154]

(d) Other known proteins localized in nucleoli

Protein	Name	Accession	CB02	LCMS	MBC	Literature
Ribosomal proteins						
MRPL3	Mitochondrial ribosomal protein L3	6005862				[155]
RPS7	Ribosomal protein S7	4506741		✓		[156]
RNA-processing factors						
POP4	POP4 (processing of precursor, *S. cerevisiae*) homologue	5729986		✓		[95]
RNASE3	Ribonuclease, RNase A family, 3 (eosinophil cationic protein)	4506551				[157]
RPP14	Ribonuclease P (14 kDa)	5902066				[158]
RPP38	Ribonuclease P (38 kDa)	5454026				[158]
SKIV2L	Superkiller viralicidic activity 2-like (*S. cerevisiae*)	20631987				[159]
Rpp21/FLJ22638	Hypothetical protein FLJ22638	15080755				[160]
FTSJ2	FtsJ homologue 2 (*Escherichia coli*)	7019377				[161]
GEMIN4	Gem (nuclear organelle)-associated protein 4	7657122				[162]
Translation-related factors						
SRP19	Signal-recognition particle 19 kDa	4507213				[163]
SRP68	Signal-recognition particle 68 kDa	7657617				[163]
SRP72	Signal-recognition particle 72 kDa	5902124				[163]
DEAD box proteins						
DDX11	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homologue, *S. cerevisiae*)	13787200				[164]
DDX12	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homologue, *S. cerevisiae*)	1517818				[164]
Nucleotide or nucleic-acid-binding proteins						
hCAP-C/SMC4L1	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	21361252		✓		[165]
hCAP-H/KIAA0074	KIAA0074 protein	559715				[165]

Table continues on following page.

Table 1 (contd.)

Protein	Name	Accession	CB02	LCMS	MBC	Literature
SMC2L1	SMC2 structural maintenance of chromosomes 2-like 1 (yeast)	5453591				[165]
ERCC5	Excision repair cross-complementing rodent repair deficiency, complementation group 5 (xeroderma pigmentosum, complementation group G (Cockayne syndrome)	4503601				[166]
FMR1	Fragile X mental retardation 1	4503765				[167]
FXR1	Fragile X mental retardation, autosomal homologue 1	4826736				[168]
FXR2	Fragile X mental retardation, autosomal homologue 2	4758410				[168]
WRN	Werner syndrome	5739524				[169]
ANP32B	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member B	5454088				[170]
JAZ	Double-stranded RNA-binding zinc-finger protein JAZ	6912440				[171]
KHDRBS3	KH-domain-containing, RNA-binding, signal-transduction-associated 3	5730073		√		[172]
SLC29A2	Solute-carrier family 29 (nucleoside transporters), member 2	10835019				[173]
STAU	Staufen, RNA-binding protein (*Drosophila*)	4759176				[116]
TARBP2	TAR (HIV) RNA-binding protein 2	19743840				[174]
TERT	Telomerase reverse transcriptase	4507439				[175]
TOP3A	Topoisomerase (DNA) IIIα	10835218				[176]
ZNF274	Zinc-finger protein 274	7706507				[177]
RB1	Retinoblastoma 1 (including osteosarcoma)	4506435				[178]
Others						
CDC2	Cell-division cycle 2, G_1 to S and G_2 to M	4502709				[179]
PPM1D	Protein phosphatase 1D magnesium-dependent, δ isoform	4505997				[180]
PPP1R10	Protein phosphatase 1, regulatory subunit 10	4506009				[181]
ASE-1	CD3-ε-associated protein; antisense to ERCC-1	6912246		√		[182]
ASNA1	arsA arsenite transporter, ATP-binding, homologue 1 (bacterial)	4757796		√		[38]
COIL	Coilin	4758024	√			[183]
HIC p40	Human I-mfa-domain-containing protein (HIC p40)	14739599				[184]
HSA6591	Nucleolar cysteine-rich protein	7657196				[185]
MCRS1	Microspherule protein 1	5453694				[186]
NOL3	Nucleolar protein 3 (apoptosis repressor with CARD domain)	4505419				[187]
NOL4	Nucleolar protein 4	4505421				[188]
SC65	Nucleolar autoantigen (55 kDa) similar to rat synaptonemal complex protein	3183090				[189]
SENP3	Sentrin/SUMO-specific protease 3	11245811		√		[190]
SSB	Sjogren syndrome antigen B (autoantigen La)	10835067				[191]
SSRP1	Structure-specific recognition protein 1	4507241		√		[192]

during the last decade of genetic defects in the yeast ribosome-synthesis pathway, as well as the recent success in the identification of pre-ribosomes using MS, has expanded our understanding of the ribosome-synthesis pathway dramatically (http://www.proteome.com; http://www.pre-ribosome.de). In particular, the large-scale systematic purification of protein complexes using tandemly arranged, high-affinity tags has led to tremendous progress in dissecting how the pre-60 S particles are formed. This advance is particularly impressive given that the half-lives of these transient structures are only approx. 1 min *in vivo* [34–37]. To define better the pathway of 60 S subunit synthesis, a series of particles was pulled out using tagged versions of several previously identified components from the yeast pre-60 S particles, revealing a series of 'snapshots' of the pre-60 S ribosomes as they move from the nucleolus to the cytoplasm.

Drawing on published data, as well as identifying homologues from the yeast ribosome-biogenesis pathway, Figure 4 shows all the possible 90 human homologues in the nucleolar proteome that could potentially be involved in ribosomal biogenesis, e.g. POP4, FIB and NOP52, organized here in Figure 4(A) according to how these proteins assist along the pathway. In this analysis, nearly 40 novel/uncharacterized proteins, constituting a third of this category in the nucleolar proteome, could potentially have functions related to ribosomal biogenesis based on the function of their yeast homologues. For other human proteins with limited characterization, e.g. many DEAD/H box helicases, we can now pinpoint more precisely their possible functions and ask whether they perform a similar function in the human system as their yeast counterparts. This illustrates the potential for using the information from the proteomic studies to elucidate human ribosomal biogenesis further.

To survive, a cell requires energy, and nucleotides such as ATP and GTP are commonly utilized as the major energy currency for nuclear export, transcription and other enzymic activities, e.g. unwinding DNA/RNA. So far one putative ATP transporter has been identified within the human nuclear proteome, namely ASNA1, a human homologue of a bacterial arsenite ATP-binding transporter arsA ([38]; J. S. Andersen, Y. W. Lam, A. K. Leung, A. I. Lamond and M. Mann, unpublished work). In addition, putative GTPases (NNP47, NNP51 and NNP57/NGB) and AAA-type ATPases (NNP86 and MDN1) can be identified and we propose that they may be involved in human ribosome biogenesis, especially in the export of pre-ribosomal particles. However, there are proteins, such as nucleophosmin (NPM/B23), that have been shown previously to be involved in ribosomal biogenesis, which are human- or higher metazoan-specific and thus were unavoidably excluded in this approach.

AMINO ACID COMPOSITION ANALYSES

Taking advantage of the recent nucleolar proteome expansion based on these organelle-directed MS studies, we analysed whether or not certain amino acids are enriched, as compared with either total-nuclear or total-cellular protein sequences (Figure 5).

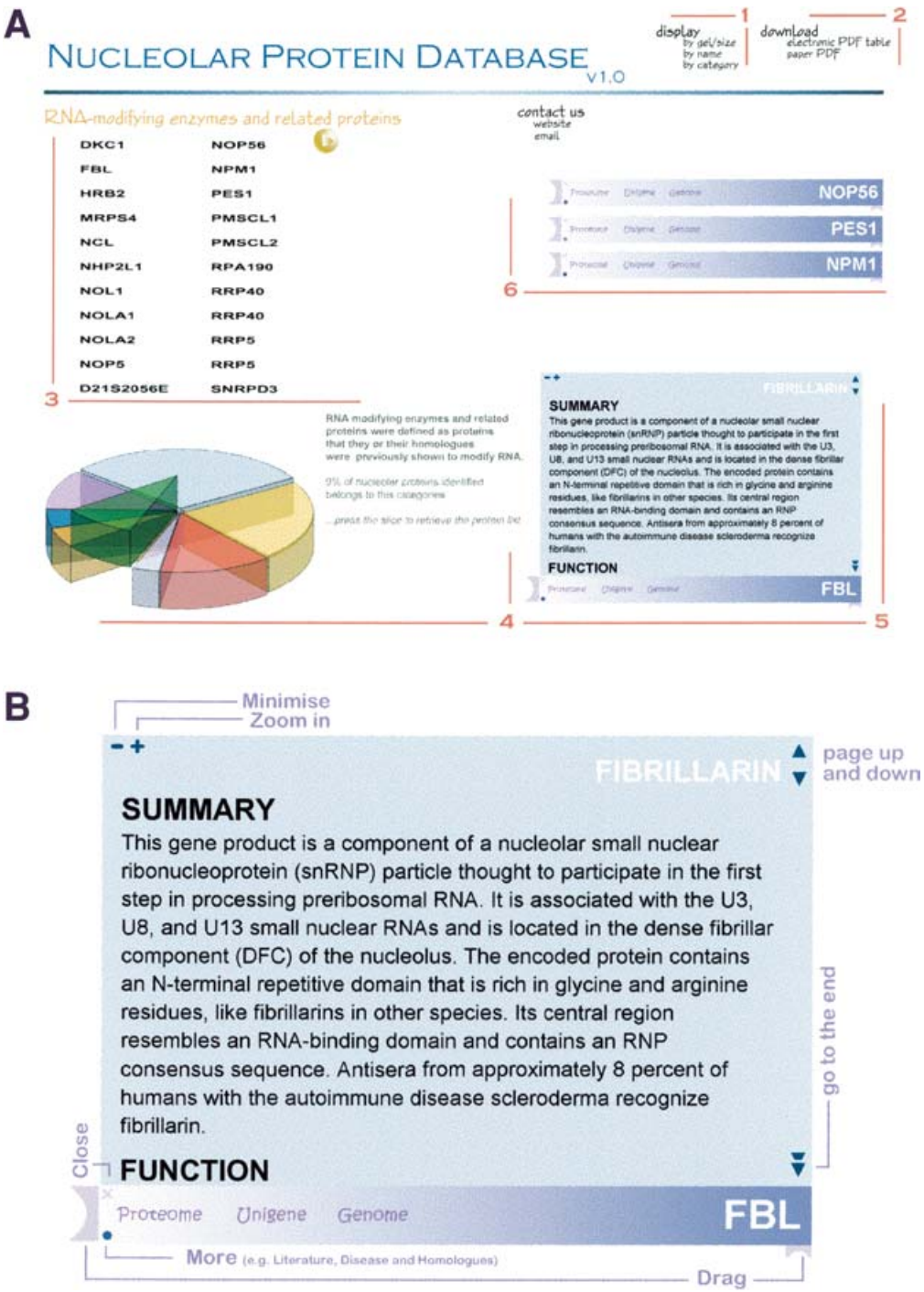


Figure 2 Nucleolar Protein Database

(**A**) A snapshot of the Nucleolar Protein Database (http://www.dundee.ac.uk/lifesciences/lamonddatabase/) showing the proteome by category, e.g. RNA-modifying enzymes and related proteins. (1) The database provides three options to display, i.e. by gel/size, by name or by category; (2) the complete dataset can be downloaded electronically; (3) the current set of proteins under a particular category (4); (5) the information of a particular protein is provided in details in a 'console window' on the lower, right-hand corner by clicking the protein name, where different protein console windows can be minimized as demonstrated in (6). (**B**) A protein 'console window' showing the information of a particular protein, e.g. fibrillarin. The construction of the Nucleolar Protein Database is summarized as follows: the peptide sequences obtained from MS were assigned to a particular protein sequence either by BLAST or, more commonly, by Mascot (http://www.matrixscience.com). All the sequences were then retrieved by Batch Entrez from the NCBI database. The protein sequences were then assigned to their corresponding gene by BLAST analysis against the UniGene database (http://www.ncbi.nlm.nih.gov/UniGene/). The UniGene entry provides a starting point leading to various other relevant databases, such as LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink) and OMIM (Online Mendelian Inheritance in Man; http://www.ncbi.nlm.nih.gov/OMIM/), to provide the cognate genomic and disease-linked information of that protein respectively. Other information is provided, including the data regarding corresponding homologues and the published literature and linking to other proteome databases.

Table 2 The most abundant motifs in the human nucleolar proteome

The frequency of the appearance of the 'motif' is indicated as *a*/*b*, where *a* indicates the number of proteins containing that particular motif and *b* indicates the total count of the motif within the human nucleolar proteome. 'Proteins' shows the corresponding human nucleolar proteins containing such a motif.

Motif	Frequency	Proteins
RNA-recognition motif (RRM)	40/51	NNP18, NNP59, NNP6, NNP64, NNP84, NNP88, NNP90, GRSF1, HNRPA0, HNRPA1, HNRPA2B1, HNRPA3, HNRPAB, HNRPC, HNRPDL, HNRPH1, HNRPH3, HNRPL, HNRPM, HNRPR, NCL, NONO, nPTB, PSP1, RBM14, RBM3, RBM4, RBMX, SFPQ/PSF, SFRS1, SFRS10, SFRS2, SFRS3, SFRS5, SFRS7, SFRS9, TARDBP, TASR1, U2AF1, U2AF65
DEAD/DEAH box helicase	23/25	BLM, CHD4, DDX10, DDX15, DDX17/P72, DDX18, DDX21, DDX24, DDX31, DDX33, DDX5, DDX9, DDXL, GU2, NOH61, NNP33, NNP40, NNP41, NNP48, NNP75, NNP121, EIF4A1, IF4N
Helicase conserved C-terminal domain	20/22	BLM, CHD4, DDX10, DDX17/P72, DDX18, DDX21, DDX24, DDX31, DDX33, DDX5, DDX9, DDXL, GU2, NOH61, NNP33, NNP41, NNP48, NNP75, EIF4A1, IF4N
WD domain	12/13	NNP116, NNP24/BING4, NNP30, NNP36, NNP42, NNP43/SAZD, NNP62, NNP63, NNP34, CSTF1, NMP200, U3-55K
Intermediate filament proteins	9/9	RAD50, SMC4L1, DSP, KRT17, KRT18, KRT8, LMNA, LMNB1, VIM
Myosin tail	9/10	CSPG6, RAD50, SMC1L1, SMC4L1, DSP, KRT17, KRT18, KRT8, LMNB1, NNP19
Elongation factor Tu GTP-binding domain	6/6	NNP69, EEF1A1, EEF1A2, EEF2, EIF2S3, TUFM

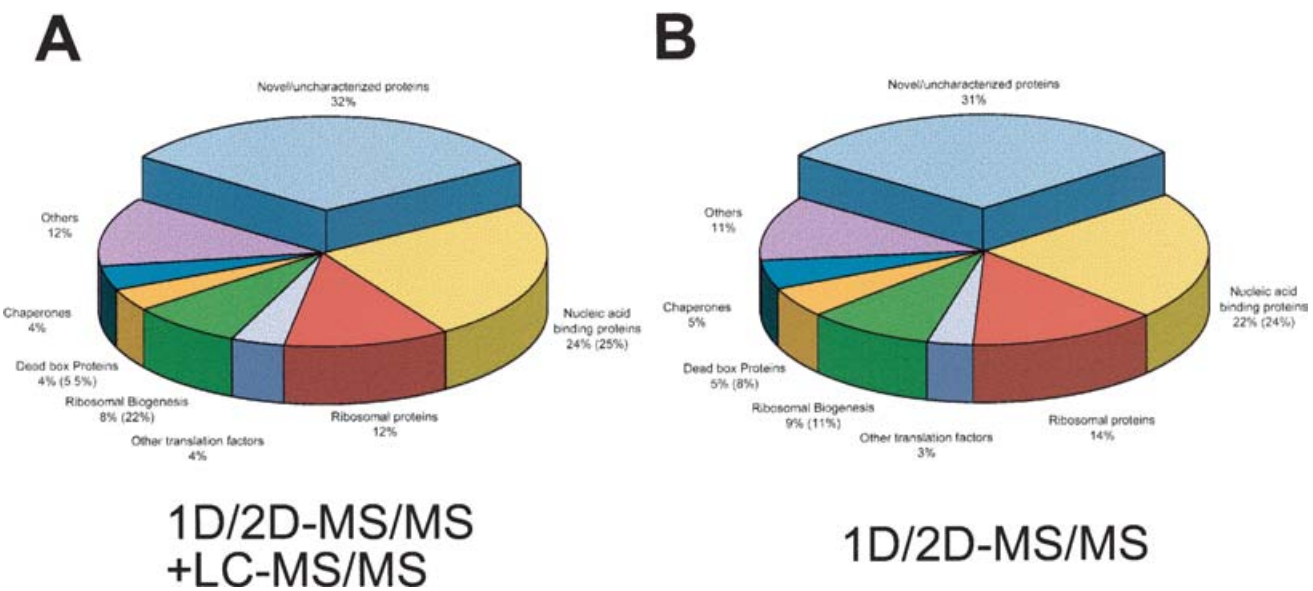


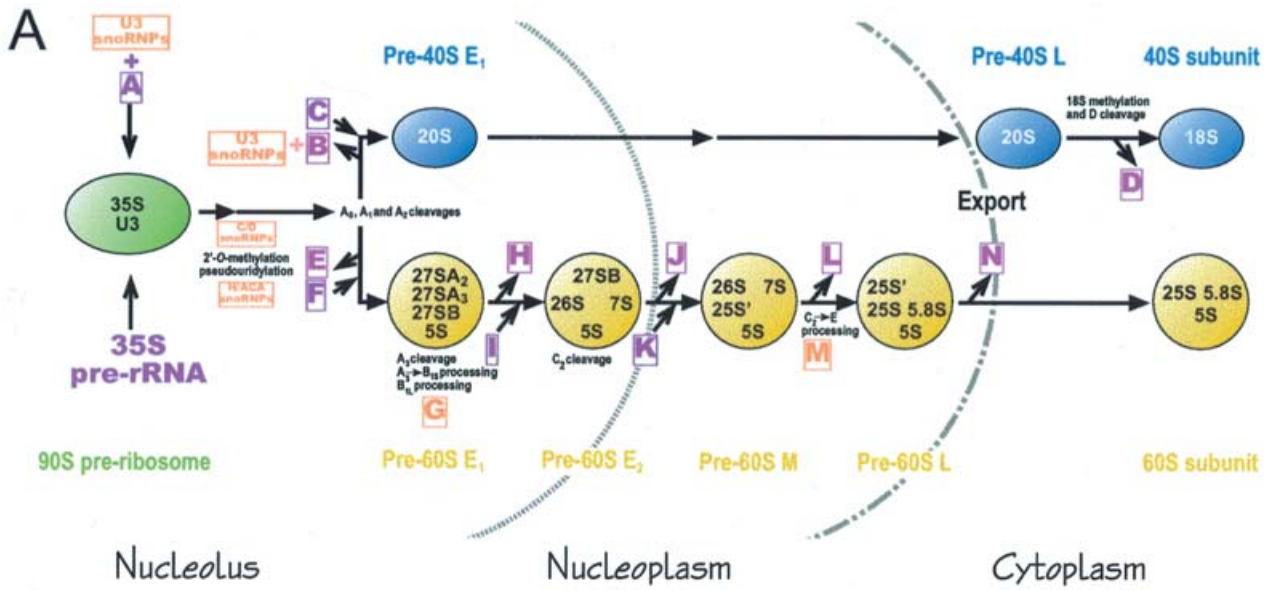
Figure 3 Distribution of conserved motifs and putative functional categories of the identified proteins in the nucleolar proteome

(**A**) Current analysis and (**B**) 1D/2D gel analysis data [9]. The percentages in brackets indicate the maximum number of proteins that could possibly be classified in the category in question. For example, 22 % of the proteins may be related to ribosomal biogenesis from the current analysis (**A**), due to their yeast orthologues being present in the ribosomal synthesis pathway.

Apparently, charged amino acids, such as glutamate, aspartate, lysine and arginine, are more favoured with respect to nuclear-specific, or total-cellular, protein sequences, whereas neutral amino acids, such as proline and cysteine, are highly disfavoured. Surprisingly, the polar amino acid serine, which is an important target for phosphorylation, was also not common within the nucleolar protein sequences. These apparent amino acid biases may reflect nucleolar targeting motifs, although these remain poorly characterized [39]. The initial search for such targeting motifs using software, such as MEME (Multiple Expectation Maximization for Motif Elicitation; http://meme.sdsc.edu/meme/website/intro.html; [40]), was only sensitive enough to retrieve motifs, such as RRM, or WD domains, as mentioned in the Bioinformatic analyses section.

Although it was not possible to define nucleolar-specific motifs shared by all nucleolar proteins, it was possible to determine whether or not some short peptide motifs show any specific enrichment in the human nucleolar proteome. For example, the protein phosphatase-1 (PP1)-binding motif, (Lys/Arg)-Val-Xaa-Phe, constitutes 0.032 % of all tetrapeptide sequences present in the nucleolar proteome. This is approx. 1.4-fold more common than in the total nuclear proteome, suggesting that a protein containing one or more PP1-binding site is more common in the nucleolus than in other parts of the nucleus. In contrast, the SUMO (small ubiquitin-related modifier)-conjugation site, (Ile/Val/Leu/Met/Phe)-Lys-Xaa-(Glu/Asp), constitutes approx. 0.2 % of tetrapeptides in nucleolar protein sequences, a value similar to its frequency in the total nuclear proteome.

Unlike arginine, glycine does not show a particular preference within the nucleolar proteome (Figure 5); however, the tripeptides GRG and RGG are enriched in the nucleolar proteome at least 2- and 3-fold respectively, as compared with either total cellular, or nuclear-specific, proteins (Figure 6A) and roughly a third of the human nucleolar proteome contains the RGG motif at least once



B

Gene Name	Yeast GN	In	Out	Sc	Sp	Ce	Dm	Remarks
NNP59	BMS1	A	B					GTPase
DDX37	DHR1	A	B					DEAD/H box Helicase
NNP53	DIP2	A	B					
NNP7	EMG1	A	B					
HRB2	KRR1	A	B					
NNP57	NOP14	A	B					
NNP52	PWP2	A	B					
RNAC	RCL1	A	B					
NNP41	ROK1	A	B					DEAD/H box Helicase
RRP5	RRP5	A	B					
NNP90	RRP7	A	B					
U3-55K	RRP9	A	B					
NNP13	UTP11	A	B					
NNP43/SAZD	UTP13	A	B					
NNP120	UTP15	A	B					
NNP117	UTP19	A	B					
NNP49	UTP21	A	B					
NNP27	UTP22	A	B					
NNP114	UTP4	A	B					
NNP105	UTP6	A	B					
NNP24/BING4	UTP7	A	B					
NNP52	YGR145W	A	B					
NNP46	YKR056W	A	B					
NNP66	KRE33	A	B, E					
NNP71	UTP10	A	B, E					
BYSL	ENP1	A	D					
NNP108	YOR145C	C	D					

C

Gene Name	Yeast GN	In	Out	Sc	Sp	Ce	Dm	Remarks
NNP66	KRE33	A	B, E					
NNP71	UTP10	A	B, E					
NNP68	RRP12	A	E					
NOH61	DBP9	F	H					DEAD/H box Helicase
NNP59	NOP4	F	H					
NNP125	RPF2	F	H					
NNP37	SSF1	F	H					
NNP81	YKL082C	F	H					
NNP3	BRX1	F	J					
NNP48	DRS1	F	J					DEAD/H box Helicase
NNP17	EBP2	F	J					
NNP78	MAK16	F	J					
NNP23	NSA1	F	J					
NNP22	RPF1	F	J					
ITGB4BP	TIF6	F	J					
NNP30	YTM1	F	J					
NNP54/BOP1	ERB1	F	L					
DDX18	HAS1	F	L					DEAD/H box Helicase
NNP98	NIP7	F	L					
NOL1	NOP2	F	L					SAM-dependent methyltransferase
PES1	NOP7	F	L					Ribosomal subunit export
NNP51	NOG1	F	N					GTPase
NOP52	RRP1	F						
NNP18	NOP15	I	J					
NNP12S/RPS0	MRT4	I	L					
NNP12	NSA2	I	L					
NNP86	RIX1	I	L					Ribosomal subunit export
FTSJ3	SPB1	I	L					SAM-dependent methyltransferase
NNP107	RLP24	I	N					
NNP76	DBP10	K	L					DEAD/H box Helicase
MDN1	MDN1	K	L					
NNP56	NOC2	K	L					Ribosomal subunit export
NNP110	NOC3	K	L					Ribosomal subunit export
NNP47	NUG1	K	L					GTPase, Ribosomal subunit export
NNP50	SDA1	K	L					
NNP116	YCR072C	K	L					
NNP38	YPL148C	K	L					
NNP57/NGB	NOG2	K	N					GTPase, Ribosomal subunit export

D

Gene Name	Yeast GN	Complex	Sc	Sp	Ce	Dm	Remarks
XRN2	RAT1	C					
NNP61	POP1	C, RNase MRP					
POP4	POP4	C, RNase MRP					
NNP55	DIS3	M, Exosome					DEAD/H box Helicase
NNP67	MTR4	M, Exosome					
NNP10	RRP4	M, Exosome					
RRP40	RRP40	M, Exosome					
PMSCL1	RRP45	M, Exosome					
RRP46	RRP46	M, Exosome					
PMSCL2	RRP6	M, Exosome					
NNP5	SKI6	M, Exosome					
DKC1	CBF5	H/ACAsnoRNP					
NOLA1	GAR1	H/ACAsnoRNP					
NOLA2	NHP2	H/ACAsnoRNP					
NOLA3	NOP10	H/ACAsnoRNP					
MRPS4	IMP3	U3 snoRNP					
NNP14	IMP4	U3 snoRNP					
NNP96	LCP5	U3 snoRNP					
MPHOPSH10	MPP10	U3 snoRNP					
NNP34	SOF1	U3 snoRNP					

Gene Name	Yeast GN	Complex	Sc	Sp	Ce	Dm	Remarks
FBL	NOP1	U3 snoRNP, C/D snoRNP					SAM-dependent methyltransferase
NOP5/NOP58	NOP58	U3 snoRNP, C/D snoRNP					
NOP56	SIK1	U3 snoRNP, C/D snoRNP					
NHP2L1	SNU13	U3 snoRNP, C/D snoRNP					

Figure 4 For legend see facing page.

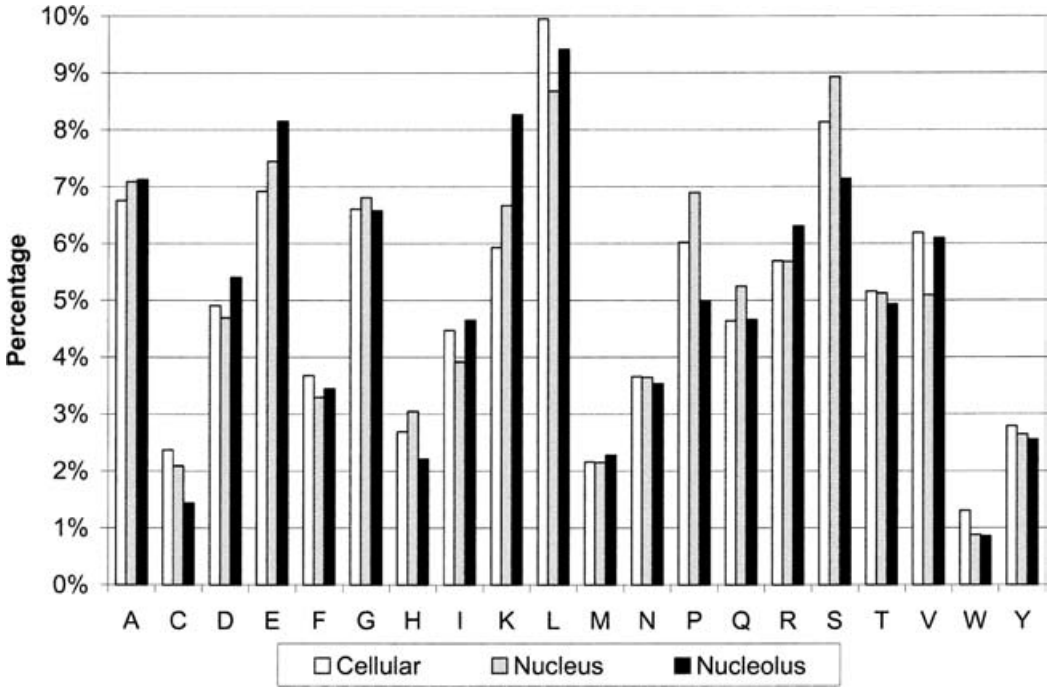


Figure 5 Abundance of individual amino acids in the human nucleolar proteome

For comparison, we generated sets of protein sequences extracted from either the human-specific total cellular proteomes or the human nuclear-specific proteomes and each set contains roughly the same number of sequences and amino acids (approx. 30 000) as the identified nucleolar proteome.

(Figure 6B). Previous studies have suggested that SMN (survival of motor neuron 1), a component of the nucleolar accessory bodies, recently renamed as Cajal bodies after the name of their discoverer [41,42], utilizes this RG-rich domain to interact with a range of proteins involved in pre-rRNA processing, ribosome production, pre-mRNA splicing, transcription and recruitment to Cajal bodies. The close association of the Cajal body with the nucleolus has been well established since its discovery. It is noted that Cajal body components containing RG-rich domains, such as Sm proteins, have been shown to traffic through the nucleolus [43]. It remains to be established whether such trafficking behaviour is due to their involvement with various RG-containing proteins within the nucleolus.

More surprisingly, the tetrapeptides FGGR and RGGF have over 10- and 20-fold enrichment, respectively, within the nucleolus, compared with either total-cellular or nuclear-specific proteins (Figure 6 and results not shown), and these two particular peptide sequences fall into the consensus motif, i.e. (G/F)-GGRGG(G/F), for asymmetric dimethylation of arginine residues [44]. This is in line with the recent result from a budding yeast study showing that several nucleolar factors are modified by arginine methyltransferase *in vivo* (Figure 6B; [45]). Although methylation does not change the overall charge on an arginine residue, addition of methyl groups increases steric hindrance and removes potential hydrogen bonds from the amino hydrogens and hence may affect potential protein–protein interactions [46]. The enrichment of such motifs possibly reflects unknown functional properties of the human nucleolus that are distinct from other subnuclear structures or, alternatively, may reflect a common mechanism for specific nucleolar localization. Consistent with the latter idea, a direct role for the asymmetric dimethylation of arginine residues within RG-rich domains was shown to be essential for nucleolar localization in the higher-molecular-mass form of basic fibroblast growth factor [47]. On the other hand, the change in interaction between proteins mediated by arginine methylation can macroscopically affect nuclear body formation, as recently documented in the case for p80 coilin, the marker protein of Cajal bodies [48,49].

Figure 4 Ribosome-biogenesis pathway in yeast

The 90 S pre-ribosomal complex is proposed to contain the 35 S rRNA and the U3 snoRNP (small nucleolar RNP). The early pre-rRNA cleavages at sites A_0 to A_2 lead to the separation of the pre-40 S and pre-60 S particles. In both pathways, a series of predicted intermediates are drawn, which are designated early (E), middle (M) and late (L), according to their positions on the proposed pathway [32]. The processing steps envisaged to be associated with each of these complexes are indicated in (**B**)–(**D**), as is the likely time of export from the nucleolus to the cytoplasm [33]. Nucleolar proteins found in this study that are putative ribosomal biogenesis factors based on homology with their yeast counterparts are represented in the following panels: (**B**) 40 S ribosomal subunit maturation pathway; (**C**) 60 S ribosomal subunit maturation pathway and (**D**) other processing complexes. Note that it is very probable that other pre-ribosomal complexes exist in addition to those shown in (**A**), and it is not clear in what order the components are gained and lost between the complexes. The colour indicated on the left of each entry corresponds to the classification of the nucleolar proteome as denoted in Figure 3. The pre-60 S pathway includes only proteins for which corroborating data exist, supporting a direct role in ribosome synthesis. 'Yeast GN' shows the closest matching yeast gene to the corresponding proteins. 'In' and 'Out' in (**B**) and (**C**) indicate the corresponding input and output complex during the maturation pathway as illustrated in (**A**). The shaded panels represent the results of BLASTp searches for each protein against the species-specific proteome database from EBI/SGD databases: *S. cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Drosophila melanogaster* (Dm) and *Caenorhabditis elegans* (Ce). The results are shaded such that a black panel represents the expectancy value (e-value) of 0.0 and thus indicates a nearly perfect match; 80 % grey for e-value $< 1e^{-100}$; 50 % for $1e^{-100} <$ e-value $< 1e^{-75}$; 40 % for $1e^{-75} <$ e-value $< 1e^{-50}$; 25 % for $1e^{-50} <$ e-value $< 1e^{-25}$ and a white panel for e-value $> 1e^{-25}$. It should be noted that the e-value between known yeast homologues of human protein can be as high as $1e^{-10}$.

A

Motif	Occurrence	No/ N	No/C
RRR	154	1.16	0.60
RRG	86	1.21	1.29
RGR	128	1.55	1.59
RGG	202	3.19	3.29
GRR	52	0.81	0.66
GRG	181	2.05	2.12
GGR	133	1.75	1.93
GGG	352	1.84	1.15

Motif	Occurrence	No/ N	No/C
RGGR	34	4.11	5.09
RGGG	58	3.82	5.34
RGGF	20	29.00	23.94
RRGG	11	1.60	1.65
GRGG	77	6.98	9.22
FRGG	12	8.70	4.79

Motif	Occurrence	No/N	No/C
GRGR	27	1.70	4.04
GRGG	77	6.98	9.22
GRGF	4	0.28	4.79
RGRG	53	2.96	3.96
GGRG	58	5.24	6.31
FGRG	8	5.80	4.79

B

Protein	Sequence
NNP6	MADKMDMSLDDIIKLNRSQRGGRGGGRNRRAGSQGGPAVGAGPARVNRGGGPIRAAIAGAAGGGGRN RPAPYSRPKQLPDKWQHDLFDSGFGGGAGVETGGKLLVSNLDFGVSDADIQELFAEFGTLKKAAVHYDR SGRSLGTADVHFERKADALKAMKQYNGVPLDGRPMNIQLVTSQIDAQRRPAQSVNRGGMTRNRGAGGF GGGGGTRRGTRGGARGRGRGAGRNSKQQLSAEELDAQLDAYNARV
FBL	MKPGFSPRGGGFGGRGGFGDRGGRGGRGGFGGGRGRGGGFRGRGRGGGGGGGGGGGGGRGGGGFHSGGNR
NOLA1	MSFRGGGRGGFNRGGGGGGFNRGGSSNHFRGGGGGGGGGNFRGGGRGGFGRGGGRGGFNK
NNP88	DRGGRGRGGSGGGGGGGGGGYNRSSGGYEPRGRGGGRGGRGGMGGSDRGGFNKFGGPRDQGSRHDSEQDN
HNRPA3	GSQRGRGGGSGNCMGHRGNFGGGGGNFGRDGNFGGRGGYGGGGGGSRGSYGGGDGGYNGLGGDGGNYGSG
HNRPU	RGGGHGRGRGGLNMRGGNFRGGAPGNRGGYNRRGNMPQRGGGGGGSGGIGYPYPRAPVFPGRGSYSNRGNY
SFRS10	PTYGSSRRRDYYDRGYDRGYDDRDYYSRSYRGGGGGGGGWRAAQDRDQIYRRRSPSPYYSRGGYRSRSRS
COILIN	GRPGPGLSSQTAGAAGWRRSGSNGGGQAPGASPSVSLPASLGRGWGREENLFSWKGAKGRGMRGRGRGRG
RBM3	SLDGRQIRVDHAGKSARGTRGGGFGAHGRGRSYSRGGGDQGYGSGRYYDSRPGGYGYGYGRSRDYNGRNQ
RBMX	GKSLDGKAIKVEQATKPSFESGRRGPPPPPRSRGPPRGLRGGRGGSGGTRGPPSRGGHMDDGGYSMNFNM
HNRPH3	RGGYYGAGRGSMYDRMRRGGDGYDGGYGGFDDYGGYNNYGYGNDGFDDRMRDGRGMGGHG
HNRPA0	FGFVYFQNHDAADKAAVVKFHPIQGHRVEVKKAVPKEDIYSGGGGGGSRSSRGGRGGRGRSGGRDQNGLS
HNRPK	DLISESPIKGRAQPYDPNFYDETYDYGGFTMMFDDRRGRPVGFPMRGRGGFDRMPPGRGGRPMPPSRRDY
NNP74	QAEGGERGGGPAGAGGEAPAALRGRLLCDFRRAPFAGRRRRAPPGGCEERELALALQEALEPAVRLGRYP
NNP37	RWEMDRGRGRLCDQKFPKTKDKSQGAQARRGPRGASRDGGRGRGRGRPGKRVA
SFPQ/PSF	MSRDRFRSRGGGGGGFHRRGGGGGRGGLHDFRSPPPGMGLNQNRGPMGPGPGQSGPKPPIPPPPPHQQQQ
HNRPDL	YTDEEPVKKLLESRYHQIGSGKCEIKVAQPKEVYRQQQQQQKGGRGAAAGGRGGTRGRGRGQGQNWNQGF
ILF2	MRGDRGRGRGGRFGSRGGPGGGFRPFVPHIPFDFYLCEMAFPRVKPAPDETSFSEALLKR
ADAR	LRPRFPVLLASSTRGRQVDIRGVPRGVHLGSQGLQRGFQHPSPRGRSLPQRGVDCLSSHFQELSIYQDQE
DDX3	SFGSRSDSRGKSSFFSDRGSGSRGRFDDRGRSDYDGIGSRGDRSGFGKFERGGNSRWCDKSDEDDWSKPL
H4F2	MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKVFLENVIRDA
HNRPR	AARQASRSTAYEDYYYHPPPRMPPPIRGRGRGGGRGGYGYPPDYYGYEDYYDDYYGYDYHDYRGGYEDPY
SNRPD3	LILPDMLKNAPMLKSMKNKNQGSGAGRGKAAILKAQVAARGRGRGMGRGNIFQKRR

Figure 6 Enrichment of short peptide sequences within the human nucleolar proteome

(**A**) The relative abundance of selected tripeptides and tetrapeptides containing arginine and glycine residues. (**B**) Examples of human nucleolar proteins containing at least one RGG motif, where RG/GR residues are coloured red and RGGF/FRGG tetrapeptides are purple. The database was searched for all the possible dipeptides, tripeptides, tetrapeptides and pentapeptides within the nucleolar proteome using customized written Perl scripts. For a protein sequence of n amino acids, it consists of $n-1$ dipeptides, $n-2$ tripeptides, $n-3$ tetrapeptides and $n-4$ pentapeptides. For example, a pentapeptide sequence of MALAL has two tetrapeptides (MALA and ALAL), three tripeptides (MAL, ALA and LAL) and three different dipeptides (MA, AL and LA) with dipeptide AL appearing twice. Therefore, from the frequency of every possible combination of dipeptides, tripeptides and tetrapeptides, we can identify whether certain short peptides show specific enrichments in the human nucleolar proteome.

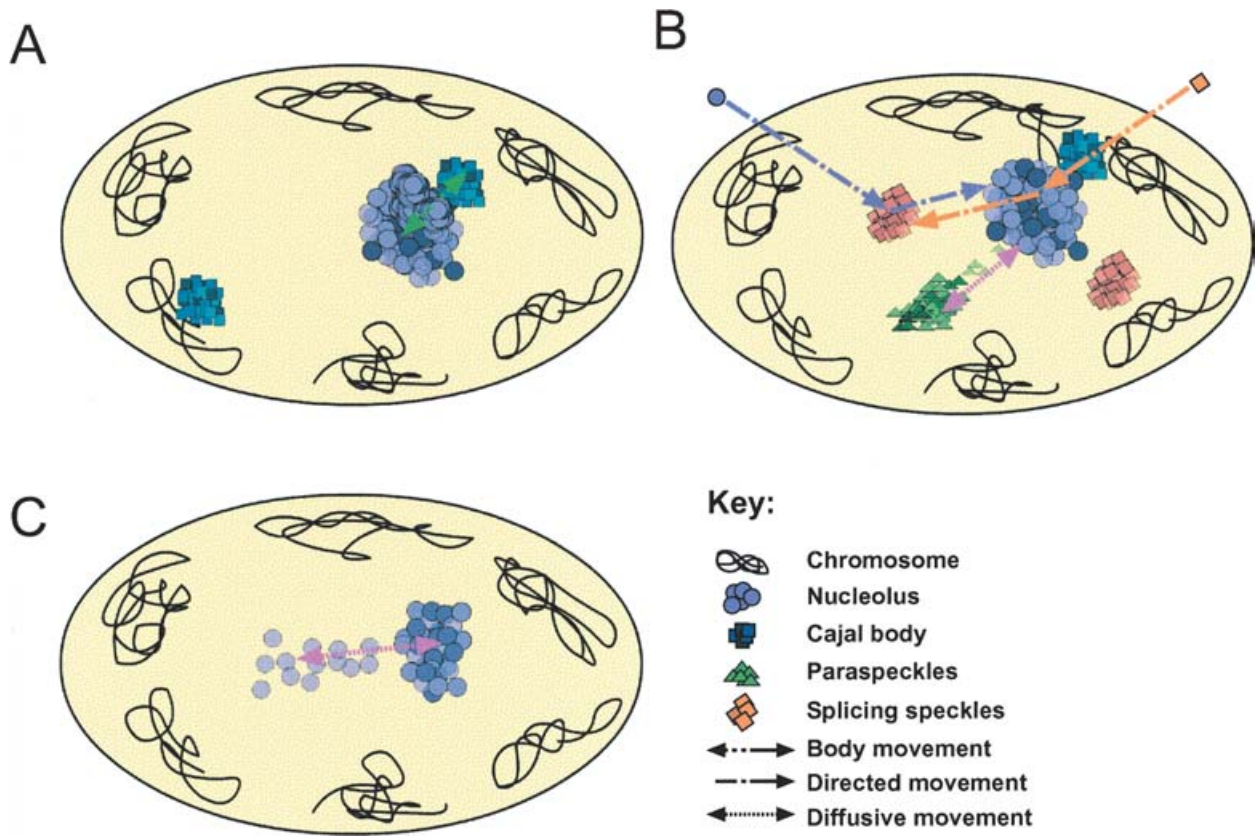


Figure 7 Dynamic nuclear proteome (animated versions to illustrate the dynamic behaviour can be found at http://www.BiochemJ.org/bj/376/bj3760553add.htm)

(**A**) The Cajal body (blue squares) has a close association with the nucleolus on the periphery and are even sometimes located within the nucleolus. (**B**) Splicing factor Sm (orange square in the top-right corner) travels via the nucleolus/Cajal body before accumulating in the splicing speckles, whereas nucleolar protein NHPX (purple circle in the top-left corner) transiently interacts with the splicing speckles before entry into the nucleolus. Marker protein of paraspeckles PSP1 (green triangle) continually exchanges between the nucleolus and the paraspeckles. (**C**) All the nucleolar proteins analysed so far continually exchange with the nucleoplasm, according to recent fluorescence recovery after photobleaching (FRAP) data.

DYNAMIC PROTEOME

An important aspect of subnuclear structures is their dynamic and regulated association with one another (Figure 7; [43,50,51]). This is especially evident for the nucleolus. Even at the light-microscopy level, a close association between nucleoli and Cajal bodies had been observed in the early twentieth century (Figure 7A; [41,42]). They are often associated with the nucleolar periphery, or even located inside the nucleolus [52–55]. Tracking of Cajal body movement in both plant and human cells showed that Cajal bodies can enter into nucleoli [56–58]. Recent studies have shown that mechanisms are operating to organize the temporal distribution of proteins and RNPs (ribonucleoprotein particles) to distinct subnuclear bodies, including nucleoli, Cajal bodies and splicing speckles, a nuclear domain where the splicing factors are accumulated [59–61]. For example, the Sm proteins of spliceosome subunits (snRNPs; small nuclear RNPs) localize to Cajal bodies and nucleoli before accumulating in splicing speckles (Figure 7B; orange squares [59,60]), whereas nucleolar protein NHPX travels transiently to the splicing speckles before accumulating in nucleoli (Figure 7B; purple circles [61]). This directed nucleolar entry may also be subject to cell-cycle regulation. For example, both the p68 DEAD box RNA helicase and the Bloom syndrome candidate gene product, BLM, accumulate within the nucleolus specifically at late telophase and S phase respectively [62–64].

So far only a few examples of directed intranuclear pathways have been uncovered, and the underlying mechanisms are still not known. In future, when more nuclear organelle-specific proteomes are available, comparative analyses may help to identify other proteins that undergo similar intricate traffic within the nucleus. It should be noted that both Sm proteins and NHPX have previously been identified in the organelle-specific proteomic studies on the nucleolus and splicing speckles/spliceosome ([9,10,65–69]; J. S. Andersen, Y. W. Lam, A. K. Leung, A. I. Lamond and M. Mann, unpublished work). The link between splicing factors and the nucleolus was reinforced further by two recent reports showing transient accumulation of more splicing factors within the nucleolus [70,71].

Recent advances in light microscopy and, in particular, the use of live cell fluorophores revealed the surprising result that many components within the nucleolus are not static, but rather are dynamic (Figure 7C; [19,72–74]). The dynamic behaviour of nucleolar proteins is also manifested in the nucleolar proteome. For example, a subset of proteins that were shown to accumulate within nucleoli on transcriptional inhibition [9,75], and several proteins normally localized in Cajal bodies, including snRNP proteins and the Cajal body marker protein p80 coilin, are

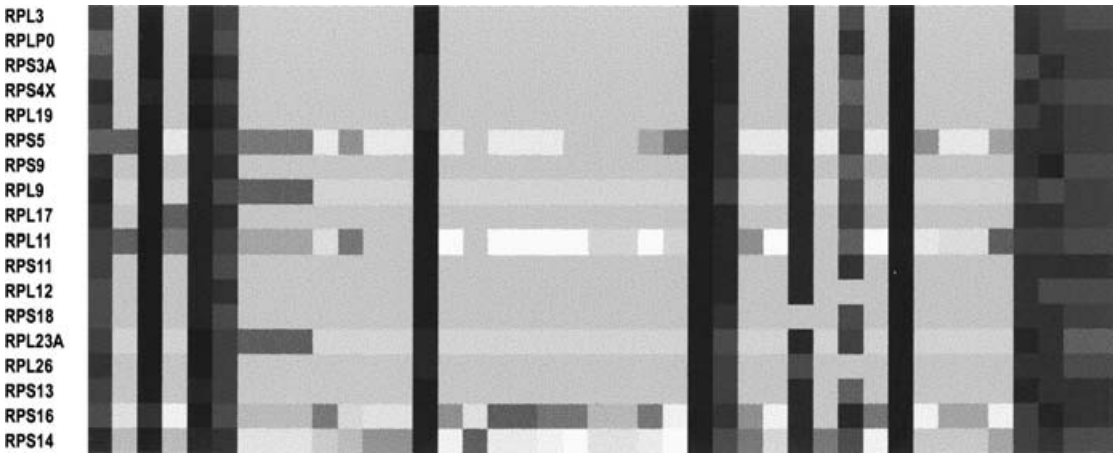


Figure 8 Nucleolar proteome profiling in terms of its homology with other model organisms

The data presented here represent a subset of 20 randomly chosen ribosomal proteins within the human nucleolar proteome. To analyse and present homologues from different completed genome sequences, we can view them as a profile, where each profile is a single row consisting of multiple elements that individually encode either the presence (100 % black) or absence (0 % black) of sequence homologues in a particular genome. Each profile is a representation of the homologues across different species; matching profiles identifies proteins with similar patterns of inheritance. When placing all the profiles parallel to one another, each row represents a single protein, whereas each column represents a single genome and the blackness of each element represents the degree of homology between the human gene and orthologue from that particular organism. Note that no homology in either primary sequence, or in tertiary structure, is required among the proteins with similar BLASTp profiles in this approach to reveal a possible functional relationship between two proteins.

accumulated in the nucleolus when dephosphorylation is inhibited [54]. During the identification of the human nucleolar proteome, a new subnuclear structure known as 'paraspeckles' was identified and three distinct paraspeckle proteins were all shown to associate with the nucleolus in a transcription-dependent manner (Figure 7B, green triangles; [75]). The marker protein, paraspeckle protein 1 (PSP1), is continuously exchanging between the nucleolus and paraspeckles, but PSP1 accumulates at the nucleolar periphery upon transcriptional inhibition. Further comparative analyses between specific organelle proteomes at different cell-cycle stages/metabolic conditions may therefore identify additional complexes that have either cell-cycle-specific or metabolically dependent interactions with the nucleolus.

PROTEOME PROFILING

Most bioinformatics analyses performed so far have analysed either individual proteins or genes and have determined whether certain motifs are present or whether the respective orthologues in other model organisms perform certain functions. In an alternative approach, proteins can be analysed either pairwise or in clusters, because proteins rarely operate alone, but usually function together in a network/complex. Proteins that form a common complex are often expressed co-ordinately in multiple organisms because many pathways or complexes are crippled by the loss of any of the individual components. For example, a set of nucleolar proteins, which are also found in the ribosome, is profiled in Figure 8, according to how they are inherited in different genomes. The similarity of the profile of each individual protein within this subset suggests a potential interdependency in different organisms, where they associate in a complex. Similar phenomena have been observed in yeast, where co-inherited proteins are usually functionally related [76]. Likewise, proteins are often expressed at either the same time, or place, as functionally related proteins. By either choosing different cell tissues [77], or by varying the growth conditions of the cells [78], enough variation in gene expression can be observed to identify co-expressing genes that may exist as complexes within the nucleolar proteome. By exercising these profiling techniques within the human nucleolar proteome, we may get a glimpse of the possible interaction networks within the human nucleolus. The recent expansion of protein–protein interaction data in yeast, both from large-scale yeast two-hybrid assays and from high-throughput MS, provides a basis for evaluating such potential associations identified from proteome profiling within the nucleolus [79,80].

As discussed in the Dynamic proteome section, the nucleolar proteome is highly dynamic and therefore the 400 human nucleolar proteins identified could facilitate an understanding of how the nucleolus responds to changes in metabolic activity. As recently demonstrated, the inhibition of transcription, e.g. by actinomycin D treatment, causes a set of 11 proteins to accumulate within the nucleolus [75]. Because the main advantage of MS is its high mass accuracy, it is possible to add a quantitative dimension to proteomic experiments by incorporating into proteins an isotope with a molecular mass that is different from the isotope used in the control experiment [81]. The chemically identical analytes with different stable isotope compositions can then be differentiated by their mass differences and the signal intensities between the two forms used to calculate their relative abundance ratio. This quantitative approach provides another dimension to our understanding of how the composition of the nucleolus changes under different metabolic states and at specific cell-cycle stages. It may also be possible to identify whether or not certain proteins move either into or out of the nucleolus as a complex using this type of proteome-profiling technique. For example, upon actinomycin D treatment, proteins that are normally localized in the paraspeckles accumulate into the nucleolus at roughly the same rate, suggesting that these proteins may potentially exist as a complex (J. S. Andersen, Y. W. Lam, A. K. Leung, A. I. Lamond and M. Mann, unpublished work).

PERSPECTIVES

In the present review, we have surveyed the previously established and recently identified nucleolar proteins known in human cells (c.f. Table 1; [9,10]). This includes over 500 proteins, which

are encoded by approx. 0.6–3 % of the human genome, assuming that the current estimate of 15 000–80 000 human genes is correct. Surprisingly, considering that the nucleolus has been studied for more than 150 years, approx. 30 % of the proteins are still defined as either novel, or previously uncharacterized, according to their genome annotations. By mining the databases that are publicly available, we could obtain interesting clues regarding the functions within the nucleolus. The nucleolus is known to be the place where ribosomal subunit biogenesis occurs. However, the pathway of ribosome synthesis in humans has still not been defined in detail, although this situation may be improved in the near future, especially now with the corresponding human homologues of budding-yeast factors being mapped in the nucleolus (see the Pathway prediction section). In comparison with the known ribosomal biogenesis pathway from yeast, mainly *Saccharomyces cerevisiae*, roughly one-third of the novel or uncharacterized human nucleolar proteins were annotated to be related to ribosomal biogenesis (c.f. Figure 4). The future analyses of these putative ribosomal biogenesis factors, which comprise approx. 20 % of the nucleolar proteome, should advance our understanding of how the ribosomal subunits form in human cells and how conserved this process is. The roles of these putative ribosome biogenesis factors can then be tested rigorously by biochemical methods.

As proteins rarely work alone within the cell, nucleolar proteins may work in a complex with one or more partner proteins to perform a specified function. The proteome-profiling technique, by incorporating a multitude of information deposited in public databases, provides a basis for studying biological process modularly, rather than individually (see the Proteome profiling section; [76,82,83]). The fortuitous discovery of nucleolar functions through studying a single protein requires a systematic way to analyse its interaction partners, either along a pathway or within a complex [84]. Even though the number of the human nucleolar proteins is of the order of hundreds, it is still relatively simple to use computational methods to map potential relationships of proteins with one another. The future mining of the wealth of existing and continually expanding information in databases, especially together with the forthcoming quantitative proteomic data, requires efforts to develop good database structures and user-friendly interfaces [85,86]. Good visualization tools are also important, especially to allow non-specialists to have ready access to the existing data. Moreover, standardization of protocols for archiving data is also crucial for transferring information effectively between laboratories and such procedures can be modelled on existing systems provided, for example, by NCBI (http://ncbi.nlm.nih.gov/) and EBI (European Bioinformatics Institute; http://www.ebi.ac.uk/) for genomic data and OME (Open Microscopy Environment; http://www.openmicroscopy.org/) for microscopic data [87].

Another approach for studying dynamic changes in the nucleolar proteome would be to carry out parallel studies using microarray techniques. Gene chips can be tailor-made to contain the 500 corresponding nucleolar cDNAs, allowing changes in the transcriptional levels of the cognate genes to be studied under different growth conditions. Correlative studies with the quantitative measurement of the relative protein abundance derived from proteomic data may provide insights into possible post-transcriptional regulation of the respective nucleolar genes. Moreover, stable cell lines that label multiple subcompartments within the nucleolus with fluorescent-protein-tagged marker proteins may provide a means to study changes in subcellular level under different growth conditions in live cells. Concurrent isolation of RNA and nucleoli from these cell lines in parallel allows quantitative comparisons of the localization pattern with the expression of genes at both transcription and translation levels.

FINAL WORDS

As described by Valentin back in 1836, the nucleolus is like ‘a secondary nucleus within the nucleus’. In fact, its complexity, as revealed by its subcompartmentalization, may provide a good model to study how the nucleus is organized. The number of nucleolar proteins (> 500) is ideal for identifying modular sets of proteins that perform specific functions by large-scale profiling of the proteins under many different conditions. The opportunity to isolate functionally intact nucleoli using a well-established protocol, including from established cell lines with fluorescently labelled subcompartments, provides a reproducible *in vitro* system to test models both biochemically and microscopically. The current understanding of the human nucleolar proteome and dynamics may be nothing but a teaser trailer, yet we are all welcome to the show.

We thank our colleagues for critically reading the manuscript. A. K. L. L. is a Croucher Foundation Scholar (Hong Kong) and ORS (Overseas Research Students Awards Scheme) awardee (U.K.). A. I. L. is a Wellcome Trust Principal Research Fellow. Work in the Lamond laboratory is supported by the Wellcome Trust and work in the Mann laboratory is funded by a Danish National Research Foundation grant to the Centre for Experimental BioInformatics.

REFERENCES

1 Valentin, G. (1836) Repertorium für Anatomie und Physiologie, vol. 1, Verlag von Veit und Comp, Berlin
2 Franke, W. W. (1988) Matthias Jacob Schleiden and the definition of the cell nucleus. Eur. J. Cell Biol. **47**, 145–156
3 Montgomery, T. H. (1898) Comparative cytological studies, with especial regard to the morphology of the nucleus. J. Morphol. **15**, 265–582
4 Moss, T. and Stefanovsky, V. Y. (2002) At the center of eukaryotic life. Cell **109**, 545–548
5 Warner, J. R. (1999) The economics of ribosome biosynthesis in yeast. Trends Biochem. Sci. **24**, 437–440
6 Visintin, R. and Amon, A. (2000) The nucleolus: the magician's hat for cell cycle tricks. Curr. Opin. Cell Biol. **12**, 372–377
7 Hiscox, J. A. (2002) The nucleolus – a gateway to viral infection? Arch. Virol. **147**, 1077–1089
8 Carmo-Fonseca, M., Mendes-Soares, L. and Campos, I. (2000) To be or not to be in the nucleolus. Nat. Cell Biol. **2**, E107–E112
9 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
10 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
11 Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. Nature (London) **422**, 198–207
12 Tyers, M. and Mann, M. (2003) From genomics to proteomics. Nature (London) **422**, 193–197
13 Andersen, J. S. and Mann, M. (2000) Functional genomics by mass spectrometry. FEBS Lett. **480**, 25–31
14 Reference deleted
15 Olson, M. O., Dundr, M. and Szebeni, A. (2000) The nucleolus: an old factory with unexpected capabilities. Trends Cell Biol. **10**, 189–196
16 Scheer, U. and Hock, R. (1999) Structure and function of the nucleolus. Curr. Opin. Cell Biol. **11**, 385–390
17 Pederson, T. and Politz, J. C. (2000) The nucleolus and the four ribonucleoproteins of translation. J. Cell Biol. **148**, 1091–1095
18 Carmo-Fonseca, M. (2002) The contribution of nuclear compartmentalization to gene regulation. Cell **108**, 513–521
19 Leung, A. K. and Lamond, A. I. (2003) The dynamics of the nucleolus. Crit. Rev. Eukaryotic Gene Expression **13**, 49–64
20 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. J. Mol. Biol. **215**, 403–410
21 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402

22 Gerbi, S. A., Borovjagin, A. V. and Lange, T. S. (2003) The nucleolus: a site of ribonucleoprotein maturation. Curr. Opin. Cell Biol. **15**, 318–325
23 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. U.S.A. **97**, 55–60
24 Politz, J. C., Lewandowski, L. B. and Pederson, T. (2002) Signal recognition particle RNA localization within the nucleolus differs from the classical sites of ribosome synthesis. J. Cell Biol. **159**, 411–418
25 Pederson, T. (2001) Is the nucleus in need of translation? Trends Cell Biol. **11**, 395–397
26 Iborra, F. J., Jackson, D. A. and Cook, P. R. (2001) Coupled transcription and translation within nuclei of mammalian cells. Science **293**, 1139–1142
27 Birnstiel, M. and Flamm, W. (1964) Intranuclear site of histone synthesis. Science **145**, 1435–1437
28 Birnstiel, M. and Hyde, B. (1963) Protein synthesis by isolated pea nucleoli. J. Cell Biol. **18**, 41–50
29 Maggio, R. (1966) Progress report on the characterization of nucleoli from guinea pig liver. Natl. Cancer Inst. Monogr. **23**, 213–222
30 Pederson, T. (1998) The plurifunctional nucleolus. Nucleic Acids Res. **26**, 3871–3876
31 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
32 Fatica, A. and Tollervey, D. (2002) Making ribosomes. Curr. Opin. Cell Biol. **14**, 313–318
33 Tschochner, H. and Hurt, E. (2003) Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol. **13**, 255–263
34 Nissan, T. A., Bassler, J., Petfalski, E., Tollervey, D. and Hurt, E. (2002) 60 S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO J. **21**, 5539–5547
35 Gadal, O., Strauss, D., Kessl, J., Trumpower, B., Tollervey, D. and Hurt, E. (2001) Nuclear export of 60 S ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. Mol. Cell. Biol. **21**, 3405–3415
36 Ho, J. H., Kallstrom, G. and Johnson, A. W. (2000) Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. J. Cell Biol. **151**, 1057–1066
37 Aitchison, J. D. and Rout, M. P. (2000) The road to ribosomes. Filling potholes in the export pathway. J. Cell Biol. **151**, F23–F26
38 Kurdi-Haidar, B., Hom, D. K., Flittner, D. E., Heath, D., Fink, L., Naredi, P. and Howell, S. B. (1998) Dual cytoplasmic and nuclear distribution of the novel arsenite-stimulated human ATPase (hASNA-I). J. Cell. Biochem. **71**, 1–10
39 Hatanaka, M. (1990) Discovery of the nucleolar targeting signal. BioEssays **12**, 143–148
40 Bailey, T. L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. **2**, 28–36
41 Gall, J. G. (2000) Cajal bodies: the first 100 years. Annu. Rev. Cell Dev. Biol. **16**, 273–300
42 Cajal, S. (1903) Un sencillo metodo de coloracion seletiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. Trab. Lab. Invest. Biol. (Madrid) **2**, 129–221
43 Sleeman, J. E. and Lamond, A. I. (1999) Nuclear organization of pre-mRNA splicing factors. Curr. Opin. Cell Biol. **11**, 372–377
44 Gary, J. D. and Clarke, S. (1998) RNA and protein interactions modulated by protein arginine methylation. Progr. Nucleic Acid Res. Mol. Biol. **61**, 65–131
45 Xu, C., Henry, P. A., Setya, A. and Henry, M. F. (2003) *In vivo* analysis of nucleolar proteins modified by the yeast arginine methyltransferase Hmt1/Rmt1p. RNA **9**, 746–759
46 McBride, A. E. and Silver, P. A. (2001) State of the Arg: protein methylation at arginine comes of age. Cell **106**, 5–8
47 Pintucci, G., Quarto, N. and Rifkin, D. B. (1996) Methylation of high molecular weight fibroblast growth factor-2 determines post-translational increases in molecular weight and affects its intracellular distribution. Mol. Biol. Cell **7**, 1249–1258
48 Hebert, M. D., Shpargel, K. B., Ospina, J. K., Tucker, K. E. and Matera, A. G. (2002) Coilin methylation regulates nuclear body formation. Dev. Cell **3**, 329–337
49 Boisvert, F. M., Cote, J., Boulanger, M. C., Cleroux, P., Bachand, F., Autexier, C. and Richard, S. (2002) Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. J. Cell Biol. **159**, 957–969
50 Janicki, S. M. and Spector, D. L. (2003) Nuclear choreography: interpretations from living cells. Curr. Opin. Cell Biol. **15**, 149–157
51 Carmo-Fonseca, M., Platani, M. and Swedlow, J. R. (2002) Macromolecular mobility inside the cell nucleus. Trends Cell Biol. **12**, 491–495
52 Malatesta, M., Zancanaro, C., Martin, T. E., Chan, E. K., Amalric, F., Luhrmann, R., Vogel, P. and Fakan, S. (1994) Is the coiled body involved in nucleolar functions? Exp. Cell Res. **211**, 415–419
53 Ochs, R. L., Stein, Jr, T. W. and Tan, E. M. (1994) Coiled bodies in the nucleolus of breast cancer cells. J. Cell Sci. **107**, 385–399
54 Lyon, C. E., Bohmann, K., Sleeman, J. and Lamond, A. I. (1997) Inhibition of protein dephosphorylation results in the accumulation of splicing snRNPs and coiled bodies within the nucleolus. Exp. Cell Res. **230**, 84–93
55 Sleeman, J., Lyon, C. E., Platani, M., Kreivi, J. P. and Lamond, A. I. (1998) Dynamic interactions between splicing snRNPs, coiled bodies and nucleoli revealed using snRNP protein fusions to the green fluorescent protein. Exp. Cell Res. **243**, 290–304
56 Snaar, S., Wiesmeijer, K., Jochemsen, A. G., Tanke, H. J. and Dirks, R. W. (2000) Mutational analysis of fibrillarin and its mobility in living human cells. J. Cell Biol. **151**, 653–662
57 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
58 Boudonck, K., Dolan, L. and Shaw, P. J. (1999) The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. Mol. Biol. Cell **10**, 2297–2307
59 Sleeman, J. E., Ajuh, P. and Lamond, A. I. (2001) snRNP protein expression enhances the formation of Cajal bodies containing p80-coilin and SMN. J. Cell Sci. **114**, 4407–4419
60 Sleeman, J. E. and Lamond, A. I. (1999) Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. Curr. Biol. **9**, 1065–1074
61 Leung, A. K. and Lamond, A. I. (2002) *In vivo* analysis of NHPX reveals a novel nucleolar localization pathway involving a transient accumulation in splicing speckles. J. Cell Biol. **157**, 615–629
62 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
63 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
64 Yankiwski, V., Marciniak, R. A., Guarente, L. and Neff, N. F. (2000) Nuclear structure in normal and Bloom syndrome cells. Proc. Natl. Acad. Sci. U.S.A. **97**, 5214–5219
65 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
66 Zhou, Z., Licklider, L. J., Gygi, S. P. and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. Nature (London) **419**, 182–185
67 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
68 Rappsilber, J., Ryder, U., Lamond, A. I. and Mann, M. (2002) Large-scale proteomic analysis of the human spliceosome. Genome Res. **12**, 1231–1245
69 Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M. and Luhrmann, R. (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. Science **298**, 2205–2208
70 Wagner, S., Chiosea, S. and Nickerson, J. A. (2003) The spatial targeting and nuclear matrix binding domains of SRm160. Proc. Natl. Acad. Sci. U.S.A. **100**, 3269–3274
71 Lai, M. C., Kuo, H. W., Chang, W. C. and Tarn, W. Y. (2003) A novel splicing regulator shares a nuclear import pathway with SR proteins. EMBO J. **22**, 1359–1369
72 Chen, D. and Huang, S. (2001) Nucleolar components involved in ribosome biogenesis cycle between the nucleolus and nucleoplasm in interphase cells. J. Cell Biol. **153**, 169–176
73 Reits, E. A. and Neefjes, J. J. (2001) From fixed to FRAP: measuring protein mobility and activity in living cells. Nat. Cell Biol. **3**, E145–E147
74 Misteli, T. (2001) Protein dynamics: implications for nuclear architecture and gene expression. Science **291**, 843–847
75 Fox, A. H., Lam, Y. W., Leung, A. K., Lyon, C. E., Andersen, J., Mann, M. and Lamond, A. I. (2002) Paraspeckles, a novel nuclear domain. Curr. Biol. **12**, 13–25
76 Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D. and Yeates, T. O. (1999) Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc. Natl. Acad. Sci. U.S.A. **96**, 4285–4288
77 Boguski, M. S. and Schuler, G. D. (1995) ESTablishing a human transcript map. Nat. Genet. **10**, 369–371
78 Lash, A. E., Tolstoshev, C. M., Wagner, L., Schuler, G. D., Strausberg, R. L., Riggins, G. J. and Altschul, S. F. (2000) SAGEmap: a public gene expression resource. Genome Res. **10**, 1051–1060
79 Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M. et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature (London) **415**, 141–147

80 Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K. et al. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature (London) **415**, 180–183

81 Flory, M., Griffin, T., Martin, D. and Aebersold, R. (2002) Advances in quantitative proteomics using stable isotope tag. Trends Biotechnol. **20**, S23–S29

82 Hartwell, L. H., Hopfield, J. J., Leibler, S. and Murray, A. W. (1999) From molecular to modular cell biology. Nature (London) **402**, C47–C52

83 Eisenberg, D., Marcotte, E. M., Xenarios, I. and Yeates, T. O. (2000) Protein function in the post-genomic era. Nature (London) **405**, 823–826

84 Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S. et al. (2002) A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science **295**, 321–324

85 Stein, L. (2002) Creating a bioinformatics nation. Nature (London) **417**, 119–120

86 Aebersold, R. and Watts, J. D. (2002) The need for national centers for proteomics. Nat. Biotechnol. **20**, 651

87 Swedlow, J. R., Goldberg, I., Brauner, E. and Sorger, P. K. (2003) Informatics and quantitative analysis in biological imaging. Science **300**, 100–102

88 Kaser, A., Bogengruber, E., Hallegger, M., Doppler, E., Lepperdinger, G., Jantsch, M., Breitenbach, M. and Kreil, G. (2001) Brix from *Xenopus laevis* and brx1p from yeast define a new family of proteins involved in the biogenesis of large ribosomal subunits. Biol. Chem. **382**, 1637–1647

89 Heese, K., Nakayama, T., Hata, R., Masumura, M., Akatsu, H., Li, F., Nagai, Y., Yamamoto, T., Kosaka, K., Suemoto, T. and Sawada, T. (2002) Characterizing CGI-94 (comparative gene identification-94) which is down-regulated in the hippocampus of early stage Alzheimer's disease brain. Eur. J. Neurosci. **15**, 79–86

90 Valdez, B. C., Henning, D., Busch, R. K., Woods, K., Flores-Rozas, H., Hurwitz, J., Perlaky, L. and Busch, H. (1996) A nucleolar RNA helicase recognized by autoimmune antibodies from a patient with watermelon stomach disease. Nucleic Acids Res. **24**, 1220–1224

91 Iggo, R. D., Jamieson, D. J., MacNeill, S. A., Southgate, J., McPheat, J. and Lane, D. P. (1991) p68 RNA helicase: identification of a nucleolar form and cloning of related genes containing a conserved intron in yeasts. Mol. Cell. Biol. **11**, 1326–1333

92 Heiss, N. S., Knight, S. W., Vulliamy, T. J., Klauck, S. M., Wiemann, S., Mason, P. J., Poustka, A. and Dokal, I. (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nat. Genet. **19**, 32–38

93 Busch, H., Busch, R. K., Black, A., Chan, P. K., Chatterjee, A., Durban, E., Freeman, J., Ochs, R., Reichlin, M., Tan, E. M. et al. (1987) Novel nucleolar antigens in autoimmune disease. J. Rheumatol. **14** (Suppl. 13), 70–77

94 di Padua Mathieu, D., Mura, C. V., Frado, L. L., Woodcock, C. L. and Stollar, B. D. (1981) Differing accessibility in chromatin of the antigenic sites of regions 1–58 and 63–125 of histone H2B. J. Cell Biol. **91**, 135–141

95 Pluk, H., van Eenennaam, H., Rutjes, S. A., Pruijn, G. J. and van Venrooij, W. J. (1999) RNA–protein interactions in the human RNase MRP ribonucleoprotein complex. RNA **5**, 512–524

96 Milarski, K. L., Welch, W. J. and Morimoto, R. I. (1989) Cell cycle-dependent association of HSP70 with specific cellular proteins. J. Cell Biol. **108**, 413–423

97 Racevskis, J., Dill, A., Stockert, R. and Fineberg, S. A. (1996) Cloning of a novel nucleolar guanosine 5′-triphosphate binding protein autoantigen from a breast tumor. Cell Growth Differ. **7**, 271–280

98 Francoeur, A. M., Peebles, C. L., Gompper, P. T. and Tan, E. M. (1986) Identification of Ki (Ku, p70/p80) autoantigens and analysis of anti-Ki autoantibody reactivity. J. Immunol. **136**, 1648–1653

99 Schonk, D. M., Kuijpers, H. J., van Drunen, E., van Dalen, C. H., Geurts van Kessel, A. H., Verheijen, R. and Ramaekers, F. C. (1989) Assignment of the gene(s) involved in the expression of the proliferation-related Ki-67 antigen to human chromosome 10. Hum. Genet. **83**, 297–299

100 Takagi, M., Sueishi, M., Saiwaki, T., Kametaka, A. and Yoneda, Y. (2001) A novel nucleolar protein, NIFK, interacts with the forkhead associated domain of Ki-67 antigen in mitosis. J. Biol. Chem. **276**, 25386–25391

101 Westendorf, J. M., Konstantinov, K. N., Wormsley, S., Shu, M. D., Matsumoto-Taniura, N., Pirollet, F., Klier, F. G., Gerace, L. and Baserga, S. J. (1998) M phase phosphoprotein 10 is a human U3 small nucleolar ribonucleoprotein component. Mol. Biol. Cell **9**, 437–449

102 Peter, M., Nakagawa, J., Doree, M., Labbe, J. C. and Nigg, E. A. (1990) Identification of major nucleolar proteins as candidate mitotic substrates of cdc2 kinase. Cell **60**, 791–801

103 Chang, M. S., Sasaki, H., Campbell, M. S., Kraeft, S. K., Sutherland, R., Yang, C. Y., Liu, Y., Auclair, D., Hao, L., Sonoda, H. et al. (1999) HRad17 colocalizes with NHP2L1 in the nucleolus and redistributes after UV irradiation. J. Biol. Chem. **274**, 36544–36549

104 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

105 Busch, H., Busch, R. K., Freeman, J. W. and Perlaky, L. (1991) Nucleolar protein P120 and its targeting for cancer chemotherapy. Boll. Soc. Ital. Biol. Sper. **67**, 739–750

106 Gautier, T., Berges, T., Tollervey, D. and Hurt, E. (1997) Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. Mol. Cell. Biol. **17**, 7088–7098

107 Pai, C. Y., Chen, H. K., Sheu, H. L. and Yeh, N. H. (1995) Cell-cycle-dependent alterations of a highly phosphorylated nucleolar protein p130 are associated with nucleologenesis. J. Cell Sci. **108**, 1911–1920

108 Lyman, S. K., Gerace, L. and Baserga, S. J. (1999) Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. RNA **5**, 1597–1604

109 Savino, T. M., Bastos, R., Jansen, E. and Hernandez-Verdun, D. (1999) The nucleolar antigen Nop52, the human homologue of the yeast ribosomal RNA processing RRP1, is recruited at late stages of nucleologenesis. J. Cell Sci. **112**, 1889–1900

110 Yung, B. Y., Bor, A. M. and Yang, Y. H. (1990) Immunolocalization of phosphoprotein B23 in proliferating and non-proliferating HeLa cells. Int. J. Cancer **46**, 272–275

111 Utama, B., Kennedy, D., Ru, K. and Mattick, J. S. (2002) Isolation and characterization of a new nucleolar protein, Nrap, that is conserved from yeast to humans. Genes Cells **7**, 115–132

112 Gelpi, C., Alguero, A., Angeles Martinez, M., Vidal, S., Juarez, C. and Rodriguez-Sanchez, J. L. (1990) Identification of protein components reactive with anti-PM/Scl autoantibodies. Clin. Exp. Immunol. **81**, 59–64

113 Bluthner, M. and Bautz, F. A. (1992) Cloning and characterization of the cDNA coding for a polymyositis-scleroderma overlap syndrome-related nucleolar 100-kD protein. J. Exp. Med. **176**, 973–980

114 Trinkle-Mulcahy, L., Sleeman, J. E. and Lamond, A. I. (2001) Dynamic targeting of protein phosphatase 1 within the nuclei of living mammalian cells. J. Cell Sci. **114**, 4219–4228

115 Huang, S., Deerinck, T. J., Ellisman, M. H. and Spector, D. L. (1997) The dynamic organization of the perinucleolar compartment in the cell nucleus. J. Cell Biol. **137**, 965–974

116 Le, S., Sternglanz, R. and Greider, C. W. (2000) Identification of two RNA-binding proteins associated with human telomerase RNA. Mol. Biol. Cell **11**, 999–1010

117 Michael, W. M. and Dreyfuss, G. (1996) Distinct domains in ribosomal protein L5 mediate 5 S rRNA binding and nucleolar localization. J. Biol. Chem. **271**, 11571–11574

118 Thomson, S. R. and Johnson, S. E. (2001) Isolation and characterization of chicken TaxREB107, a putative DNA binding protein abundantly expressed in muscle. Gene **278**, 81–88

119 Vater, C. A., Bartle, L. M., Leszyk, J. D., Lambert, J. M. and Goldmacher, V. S. (1995) Ricin A chain can be chemically cross-linked to the mammalian ribosomal proteins L9 and L10e. J. Biol. Chem. **270**, 12933–12940

120 Franco, R. and Rosenfeld, M. G. (1990) Hormonally inducible phosphorylation of a nuclear pool of ribosomal protein S6. J. Biol. Chem. **265**, 4321–4325

121 Magoulas, C. and Fried, M. (2000) Isolation and genomic analysis of the human surf-6 gene: a member of the Surfeit locus. Gene **243**, 115–123

122 Winokur, S. T. and Shiang, R. (1998) The Treacher Collins syndrome (TCOF1) gene product, treacle, is targeted to the nucleolus by signals in its C-terminus. Hum. Mol. Genet. **7**, 1947–1952

123 Zini, N., Santi, S., Ognibene, A., Bavelloni, A., Neri, L. M., Valmori, A., Mariani, E., Negri, C., Astaldi-Ricotti, G. C. and Maraldi, N. M. (1994) Discrete localization of different DNA topoisomerases in HeLa and K562 cell nuclei and subnuclear fractions. Exp. Cell Res. **210**, 336–348

124 Pluk, H., Soffner, J., Luhrmann, R. and van Venrooij, W. J. (1998) cDNA cloning and characterization of the human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. Mol. Cell. Biol. **18**, 488–498

125 Jantzen, H. M., Admon, A., Bell, S. P. and Tjian, R. (1990) Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature (London) **344**, 830–836

126 Pluta, A. F. and Earnshaw, W. C. (1996) Specific interaction between human kinetochore protein CENP-C and a nucleolar transcriptional regulator. J. Biol. Chem. **271**, 18767–18774

127 Hudson, J. W., Kozarova, A., Cheung, P., Macmillan, J. C., Swallow, C. J., Cross, J. C. and Dennis, J. W. (2001) Late mitotic failure in mice lacking Sak, a polo-like kinase. Curr. Biol. **11**, 441–446

128 Stegh, A. H., Schickling, O., Ehret, A., Scaffidi, C., Peterhansel, C., Hofmann, T. G., Grummt, I., Krammer, P. H. and Peter, M. E. (1998) DEDD, a novel death effector domain-containing protein, targeted to the nucleolus. EMBO J. **17**, 5974–5986

129 Kaul, S. C., Kawai, R., Nomura, H., Mitsui, Y., Reddel, R. R. and Wadhwa, R. (1999) Identification of a 55-kDa ezrin-related protein that induces cytoskeletal changes and localizes to the nucleolus. Exp. Cell Res. **250**, 51–61

130 Moroianu, J. and Riordan, J. F. (1994) Identification of the nucleolar targeting signal of human angiogenin. Biochem. Biophys. Res. Commun. **203**, 1765–1772

131 Modrell, B., McDonald, V. L. and Shoyab, M. (1992) The interaction of amphiregulin with nuclei and putative nuclear localization sequence binding proteins. Growth Factors **7**, 305–314

132 Gualandris, A., Arese, M., Shen, B. and Rifkin, D. B. (1999) Modulation of cell growth and transformation by doxycycline-regulated FGF-2 expression in NIH-3T3 cells. J. Cell. Physiol. **181**, 273–284

133 Kiefer, P. and Dickson, C. (1995) Nucleolar association of fibroblast growth factor 3 via specific sequence motifs has inhibitory effects on cell growth. Mol. Cell. Biol. **15**, 4364–4374

134 Henderson, J. E., Amizuka, N., Warshawsky, H., Biasotto, D., Lanske, B. M., Goltzman, D. and Karaplis, A. C. (1995) Nucleolar localization of parathyroid hormone-related peptide enhances survival of chondrocytes under conditions that promote apoptotic cell death. Mol. Cell. Biol. **15**, 4064–4075

135 Parsell, D. A., Sanchez, Y., Stitzel, J. D. and Lindquist, S. (1991) Hsp104 is a highly conserved protein with two essential nucleotide-binding sites. Nature (London) **353**, 270–273

136 Biggiogera, M., Tanguay, R. M., Marin, R., Wu, Y., Martin, T. E. and Fakan, S. (1996) Localization of heat shock proteins in mouse male germ cells: an immunoelectron microscopical study. Exp. Cell Res. **229**, 77–85

137 Owens-Grillo, J. K., Czar, M. J., Hutchison, K. A., Hoffmann, K., Perdew, G. H. and Pratt, W. B. (1996) A model of protein targeting mediated by immunophilins and other proteins that bind to hsp90 via tetratricopeptide repeat domains. J. Biol. Chem. **271**, 13468–13475

138 Jeffrey, I. W., Kadereit, S., Meurs, E. F., Metzger, T., Bachmann, M., Schwemmle, M., Hovanessian, A. G. and Clemens, M. J. (1995) Nuclear localization of the interferon-inducible protein kinase PKR in human cells and transfected mouse cells. Exp. Cell Res. **218**, 17–27

139 Welsh, G. I., Kadereit, S., Coccia, E. M., Hovanessian, A. G. and Meurs, E. F. (1999) Colocalization within the nucleolus of two highly related IFN-induced human nuclear phosphoproteins with nucleolin. Exp. Cell Res. **250**, 62–74

140 Scott, M., Boisvert, F. M., Vieyra, D., Johnston, R. N., Bazett-Jones, D. P. and Riabowol, K. (2001) UV induces nucleolar translocation of ING1 through two distinct nucleolar targeting sequences. Nucleic Acids Res. **29**, 2052–2058

141 Lin, C. Y., Huang, P. H., Liao, W. L., Cheng, H. J., Huang, C. F., Kuo, J. C., Patton, W. A., Massenburg, D., Moss, J. and Lee, F. J. (2000) ARL4, an ARF-like protein that is developmentally regulated and localized to nuclei and nucleoli. J. Biol. Chem. **275**, 37815–37823

142 Opas, M., Dziak, E., Fliegel, L. and Michalak, M. (1991) Regulation of expression and intracellular distribution of calreticulin, a major calcium binding protein of nonmuscle cells. J. Cell. Physiol. **149**, 160–171

143 Hoogeveen, A. T., Rossetti, S., Stoyanova, V., Schonkeren, J., Fenaroli, A., Schiaffonati, L., van Unen, L. and Sacchi, N. (2002) The transcriptional corepressor MTG16a contains a novel nucleolar targeting sequence deranged in t (16; 21)-positive myeloid malignancies. Oncogene **21**, 6703–6712

144 Hattori, H., Liu, Y. C., Tohnai, I., Ueda, M., Kaneda, T., Kobayashi, T., Tanabe, K. and Ohtsuka, K. (1992) Intracellular localization and partial amino acid sequence of a stress-inducible 40-kDa protein in HeLa cells. Cell Struct. Funct. **17**, 77–86

145 Munnia, A., Schutz, N., Romeike, B. F., Maldener, E., Glass, B., Maas, R., Nastainczyk, W., Feiden, W., Fischer, U. and Meese, E. (2001) Expression, cellular distribution and protein binding of the glioma amplified sequence (GAS41), a highly conserved putative transcription factor. Oncogene **20**, 4853–4863

146 Rizos, H., Darmanian, A. P., Mann, G. J. and Kefford, R. F. (2000) Two arginine rich domains in the p14ARF tumour suppressor mediate nucleolar localization. Oncogene **19**, 2978–2985

147 Fischer, H., Zhang, X. U., O'Brien, K. P., Kylsten, P. and Engvall, E. (2001) C7, a novel nucleolar protein, is the mouse homologue of the *Drosophila* late puff product L82 and an isoform of human OXR1. Biochem. Biophys. Res. Commun. **281**, 795–803

148 Lindstrom, M. S., Klangby, U., Inoue, R., Pisa, P., Wiman, K. G. and Asker, C. E. (2000) Immunolocalization of human p14(ARF) to the granular component of the interphase nucleolus. Exp. Cell Res. **256**, 400–410

149 Dalmau, J., Gultekin, S. H., Voltz, R., Hoard, R., DesChamps, T., Balmaceda, C., Batchelor, T., Gerstner, E., Eichen, J., Frennier, J., Posner, J. B. and Rosenfeld, M. R. (1999) Ma1, a novel neuron- and testis-specific protein, is recognized by the serum of patients with paraneoplastic neurological disorders. Brain **122**, 27–39

150 Chai, Z., Sarcevic, B., Mawson, A. and Toh, B. H. (2001) SET-related cell division autoantigen-1 (CDA1) arrests cell growth. J. Biol. Chem. **276**, 33665–33674

151 Benninghoff, J., Kartarius, S., Teleb, Z., Selter, H., Unteregger, G., Zwergel, T., Wullich, B. and Montenarh, M. (1999) Two different forms of p53 localized differently within cells of urogenital tumours. Cancer Lett. **144**, 55–64

152 McNeil, S., Guo, B., Stein, J. L., Lian, J. B., Bushmeyer, S., Seto, E., Atchison, M. L., Penman, S., van Wijnen, A. J. and Stein, G. S. (1998) Targeting of the YY1 transcription factor to the nucleolus and the nuclear matrix *in situ*: the C-terminus is a principal determinant for nuclear trafficking. J. Cell. Biochem. **68**, 500–510

153 Payen, E., Verkerk, T., Michalovich, D., Dreyer, S. D., Winterpacht, A., Lee, B., De Zeeuw, C. I., Grosveld, F. and Galjart, N. (1998) The centromeric/nucleolar chromatin protein ZFP-37 may function to specify neuronal nuclear domains. J. Biol. Chem. **273**, 9099–9109

154 Galcheva-Gargova, Z., Gangwani, L., Konstantinov, K. N., Mikrut, M., Theroux, S. J., Enoch, T. and Davis, R. J. (1998) The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus of proliferating cells. Mol. Biol. Cell **9**, 2963–2971

155 Ou, J. H., Yen, T. S., Wang, Y. F., Kam, W. K. and Rutter, W. J. (1987) Cloning and characterization of a human ribosomal protein gene with enhanced expression in fetal and neoplastic cells. Nucleic Acids Res. **15**, 8919–8934

156 Annilo, T., Karis, A., Hoth, S., Rikk, T., Kruppa, J. and Metspalu, A. (1998) Nuclear import and nucleolar accumulation of the human ribosomal protein S7 depends on both a minimal nuclear localization sequence and an adjacent basic region. Biochem. Biophys. Res. Commun. **249**, 759–766

157 Wu, H., Xu, H., Miraglia, L. J. and Crooke, S. T. (2000) Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J. Biol. Chem. **275**, 36957–36965

158 Jarrous, N., Wolenski, J. S., Wesolowski, D., Lee, C. and Altman, S. (1999) Localization in the nucleolus and coiled bodies of protein subunits of the ribonucleoprotein ribonuclease P. J. Cell Biol. **146**, 559–572

159 Lee, S. G., Lee, I., Park, S. H., Kang, C. and Song, K. (1995) Identification and characterization of a human cDNA homologous to yeast SKI2. Genomics **25**, 660–666

160 Jarrous, N., Reiner, R., Wesolowski, D., Mann, H., Guerrier-Takada, C. and Altman, S. (2001) Function and subnuclear distribution of Rpp21, a protein subunit of the human ribonucleoprotein ribonuclease P. RNA **7**, 1153–1164

161 Ching, Y. P., Zhou, H. J., Yuan, J. G., Qiang, B. Q., Kung, H. F. and Jin, D. Y. (2002) Identification and characterization of FTSJ2, a novel human nucleolar protein homologous to bacterial ribosomal RNA methyltransferase. Genomics **79**, 2–6

162 Charroux, B., Pellizzoni, L., Perkinson, R. A., Yong, J., Shevchenko, A., Mann, M. and Dreyfuss, G. (2000) Gemin4: a novel component of the SMN complex that is found in both gems and nucleoli. J. Cell Biol. **148**, 1177–1186

163 Jacobson, M. R. and Pederson, T. (1998) Localization of signal recognition particle RNA in the nucleolus of mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **95**, 7981–7986

164 Amann, J., Kidd, V. J. and Lahti, J. M. (1997) Characterization of putative human homologues of the yeast chromosome transmission fidelity gene, CHL1. J. Biol. Chem. **272**, 3823–3832

165 Cabello, O. A., Eliseeva, E., He, W. G., Youssoufian, H., Plon, S. E., Brinkley, B. R. and Belmont, J. W. (2001) Cell cycle-dependent expression and nucleolar localization of hCAP-H. Mol. Biol. Cell **12**, 3527–3537

166 Knauf, J. A., Pendergrass, S. H., Marrone, B. L., Strniste, G. F., MacInnes, M. A. and Park, M. S. (1996) Multiple nuclear localization signals in XPG nuclease. Mutat. Res. **363**, 67–75

167 Willemsen, R., Bontekoe, C., Tamanini, F., Galjaard, H., Hoogeveen, A. and Oostra, B. (1996) Association of FMRP with ribosomal precursor particles in the nucleolus. Biochem. Biophys. Res. Commun. **225**, 27–33

168 Tamanini, F., Kirkpatrick, L. L., Schonkeren, J., van Unen, L., Bontekoe, C., Bakker, C., Nelson, D. L., Galjaard, H., Oostra, B. A. and Hoogeveen, A. T. (2000) The fragile X-related proteins FXR1P and FXR2P contain a functional nucleolar-targeting signal equivalent to the HIV-1 regulatory proteins. Hum. Mol. Genet. **9**, 1487–1493

169 Marciniak, R. A., Lombard, D. B., Johnson, F. B. and Guarente, L. (1998) Nucleolar localization of the Werner syndrome protein in human cells. Proc. Natl. Acad. Sci. U.S.A. **95**, 6887–6892

170 Zhu, L., Perlaky, L., Henning, D. and Valdez, B. C. (1997) Cloning and characterization of a new silver-stainable protein SSP29, a member of the LRR family. Biochem. Mol. Biol. Int. **42**, 927–935

171 Yang, M., May, W. S. and Ito, T. (1999) JAZ requires the double-stranded RNA-binding zinc finger motifs for nuclear localization. J. Biol. Chem. **274**, 27399–27406

172 Venables, J. P., Vernet, C., Chew, S. L., Elliott, D. J., Cowmeadow, R. B., Wu, J., Cooke, H. J., Artzt, K. and Eperon, I. C. (1999) T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis. Hum. Mol. Genet. **8**, 959–969

173 Williams, J. B. and Lanahan, A. A. (1995) A mammalian delayed-early response gene encodes HNP36, a novel, conserved nucleolar protein. Biochem. Biophys. Res. Commun. **213**, 325–333

174 Eckmann, C. R. and Jantsch, M. F. (1997) Xlrbpa, a double-stranded RNA-binding protein associated with ribosomes and heterogeneous nuclear RNPs. J. Cell Biol. **138**, 239–253

175 Yang, Y., Chen, Y., Zhang, C., Huang, H. and Weissman, S. M. (2002) Nucleolar localization of hTERT protein is associated with telomerase function. Exp. Cell Res. **277**, 201–209
176 Lin, C. W., Darzynkiewicz, Z., Li, X., Traganos, F., Bedner, E. and Tse-Dinh, Y. C. (2000) Differential expression of human topoisomerase IIIα during the cell cycle progression in HL-60 leukemia cells and human peripheral blood lymphocytes. Exp. Cell Res. **256**, 225–236
177 Huang, Z., Philippin, B., O'Leary, E., Bonventre, J. V., Kriz, W. and Witzgall, R. (1999) Expression of the transcriptional repressor protein Kid-1 leads to the disintegration of the nucleolus. J. Biol. Chem. **274**, 7640–7648
178 Bukovsky, A., Caudle, M. R., Keenan, J. A., Wimalasena, J., Foster, J. S. and Van Meter, S. E. (1995) Quantitative evaluation of the cell cycle-related retinoblastoma protein and localization of Thy-1 differentiation protein and macrophages during follicular development and atresia, and in human corpora lutea. Biol. Reprod. **52**, 776–792
179 Ino, H., Mochizuki, T., Yanaihara, N. and Chiba, T. (1993) p34cdc2 homologue is located in nucleoli of the nervous and endocrine systems. Brain Res. **614**, 131–136
180 Kotani, H., Ito, M., Hamaguchi, T., Ichikawa, K., Nakano, T., Shima, H., Nagao, M., Ohta, N., Furuichi, Y., Takahashi, T. and Umekawa, H. (1998) The δ isoform of protein phosphatase type 1 is localized in nucleolus and dephosphorylates nucleolar phosphoproteins. Biochem. Biophys. Res. Commun. **249**, 292–296
181 Kreivi, J. P., Trinkle-Mulcahy, L., Lyon, C. E., Morrice, N. A., Cohen, P. and Lamond, A. I. (1997) Purification and characterisation of p99, a nuclear modulator of protein phosphatase 1 activity. FEBS Lett. **420**, 57–62
182 Whitehead, C. M., Winkfein, R. J., Fritzler, M. J. and Rattner, J. B. (1997) ASE-1: a novel protein of the fibrillar centres of the nucleolus and nucleolus organizer region of mitotic chromosomes. Chromosoma **106**, 493–502
183 Bohmann, K., Ferreira, J. A. and Lamond, A. I. (1995) Mutational analysis of p80 coilin indicates a functional interaction between coiled bodies and the nucleolus. J. Cell Biol. **131**, 817–831
184 Thebault, S., Basbous, J., Gay, B., Devaux, C. and Mesnard, J. M. (2000) Sequence requirement for the nucleolar localization of human I-mfa domain-containing protein (HIC p40). Eur. J. Cell Biol. **79**, 834–838
185 Bolivar, J., Diaz, I., Iglesias, C. and Valdivia, M. M. (1999) Molecular cloning of a zinc finger autoantigen transiently associated with interphase nucleolus and mitotic centromeres and midbodies: orthologous proteins with nine CXXC motifs highly conserved from nematodes to humans. J. Biol. Chem. **274**, 36456–36464
186 Ren, Y., Busch, R. K., Perlaky, L. and Busch, H. (1998) The 58-kDa microspherule protein (MSP58), a nucleolar protein, interacts with nucleolar protein p120. Eur. J. Biochem. **253**, 734–742
187 Stoss, O., Schwaiger, F. W., Cooper, T. A. and Stamm, S. (1999) Alternative splicing determines the intracellular localization of the novel nuclear protein Nop30 and its interaction with the splicing factor SRp30c. J. Biol. Chem. **274**, 10951–10962
188 Ueki, N., Kondo, M., Seki, N., Yano, K., Oda, T., Masuho, Y. and Muramatsu, M. (1998) NOLP: identification of a novel human nucleolar protein and determination of sequence requirements for its nucleolar localization. Biochem. Biophys. Res. Commun. **252**, 97–102
189 Ochs, R. L., Stein, Jr, T. W., Chan, E. K., Ruutu, M. and Tan, E. M. (1996) cDNA cloning and characterization of a novel nucleolar protein. Mol. Biol. Cell **7**, 1015–1024
190 Nishida, T., Tanaka, H. and Yasuda, H. (2000) A novel mammalian Smt3-specific isopeptidase 1 (SMT3IP1) localized in the nucleolus at interphase. Eur. J. Biochem. **267**, 6423–6427
191 Carmo-Fonseca, M., Pfeifer, K., Schroder, H. C., Vaz, M. F., Fonseca, J. E., Muller, W. E. and Bachmann, M. (1989) Identification of La ribonucleoproteins as a component of interchromatin granules. Exp. Cell Res. **185**, 73–85
192 Hsu, T., King, D. L., LaBonne, C. and Kafatos, F. C. (1993) *A Drosophila* single-strand DNA/RNA-binding factor contains a high-mobility-group box and is enriched in the nucleolus. Proc. Natl. Acad. Sci. U.S.A. **90**, 6488–6492

Received 1 August 2003; accepted 8 October 2003
Published as BJ Immediate Publication 8 October 2003, DOI 10.1042/BJ20031169