DECIPHERING THE HUMAN NUCLEOLAR PROTEOME*

Yohann Couté,[1]* Jennifer A. Burgess,[1] Jean-Jacques Diaz,[2] Christine Chichester,[3] Frédérique Lisacek,[3,4,5] Anna Greco,[2] and Jean-Charles Sanchez[1]

[1]*Biomedical Proteomics Research Group, Département de Biologie Structurale et Bioinformatique, Centre Médical Universitaire, 1 Rue Michel Servet, 1211 Geneva 14, Switzerland*
[2]*Centre de Génétique Moléculaire et Cellulaire, CNRS-UMR 5534, Domaine Scientifique de la DOUA, Bât. Gregor Mendel, 43 Bd du 11 Novembre 1918, 69622 Villeurbanne Cedex, France*
[3]*Geneva Bioinformatics (GeneBio), 25 Av. de Champel, 1206 Geneva, Switzerland*
[4]*Swiss Institute of Bioinformatics, Proteome Informatics Group, 1 Rue Michel Servet, 1211 Geneva 14, Switzerland*
[5]*Génome & Informatique, Tour Evry 2, 91034 Evry Cedex, France*

Received 10 March 2005; received (revised) 6 June 2005; accepted 10 June 2005

Published online 4 October 2005 in Wiley InterScience (www.interscience.wiley.com) DOI 10.1002/mas.20067

Nucleoli are plurifunctional nuclear domains involved in the regulation of several major cellular processes such as ribosome biogenesis, the biogenesis of non-ribosomal ribonucleoprotein complexes, cell cycle, and cellular aging. Until recently, the protein content of nucleoli was poorly described. Several proteomic analyses have been undertaken to discover the molecular bases of the biological roles fulfilled by nucleoli. These studies have led to the identification of more than 700 proteins. Extensive bibliographic and bioinformatic analyses allowed the classification of the identified proteins into functional groups and suggested potential functions of 150 human proteins previously uncharacterized. The combination of improvements in mass spectrometry technologies, the characterization of protein complexes, and data mining will assist in furthering our understanding of the role of nucleoli in different physiological and pathological cell states. © 2005 Wiley Periodicals, Inc., Mass Spec Rev 25:215–234, 2006

Keywords: *sub-cellular proteomics; nucleolus; mass spectrometry; ribosome biogenesis; gene expression regulation*

I. INTRODUCTION: NUCLEOLI, PLURIFUNCTIONAL DOMAINS OF THE NUCLEUS

A. The Cell Nucleus: An Organized and Dynamic Cell Organelle

Microscopists first described the nucleus of eukaryotic cells about 200 years ago. It is physically separated from the cytoplasm by a lipid bilayer that acts as a regulator of import and export of molecules through nuclear pores. The nucleus has a high degree of organization. It contains a number of highly dynamic domains, in relation to DNA replication, RNA metabolism and transport (for reviews see Park & De Boni, 1999; Misteli, 2001). The organization of the nucleus depends on the species, cell type, and physiological or pathological states of the cell (Carter et al., 1993; Lewis & Tollervey, 2000; Dundr & Misteli, 2001). During interphase, the nucleus is compartmentalized into domains that support the major biological nuclear activities. Chromosomes are arranged as discrete entities called chromosome territories. Nascent pre-mRNAs localize at the surface of these chromosome territories. The mRNAs are synthesized from active genes within perichromatin fibrils where co-transcriptional splicing also occurs. In the interchromatin space, various molecular species constantly associate and dissociate to give rise to membrane-less nuclear domains whose composition and function have not yet been fully elucidated (Zimber, Nguyen, & Gespach, 2004). At this point in time, approximately 30 different nuclear domains have been characterized including nucleoli, the splicing factor compartments, the Cajal bodies (CBs), and the promyelocytic leukemia (PML) bodies (Dundr & Misteli, 2001).

Recently, the high dynamic of the cell nucleus has been demonstrated, notably by use of high resolution fluorescence microscopy and photobleaching techniques (Phair & Misteli, 2000). It was demonstrated that nuclear molecules freely diffuse through the nuclear space and that they are constantly exchanged from their accumulation domain with the nucleoplasm. This was shown for proteins (reviewed in Misteli, 2001) but also for RNAs (Politz et al., 1998, 1999). For example, it was calculated that the mean residence time of fibrillarin within nucleoli was less than 40 sec (Phair & Misteli, 2000). These observations allowed the proposal that nuclear molecules freely diffuse within nucleoplasm until they find a binding partner with which they interact for a certain time before being released and diffuse again in the nucleoplasm (Misteli, 2001). However, in spite of this continual exchange of molecules between nuclear compartments and the nucleoplasm, nuclear domains are maintained because, at a given time, the number of proteins within a domain is largely superior to the number of proteins that are released from this domain. For instance, it was calculated that approximately 12,000 molecules of fibrillarin leave nucleoli every second out of about 500,000 molecules (Phair & Misteli, 2000).

The supplementary materials referred to in this article can be found at http://www.interscience.wiley.com/jpages/0277-7037/suppmat/.

**Correspondence to:* Yohann Couté, BPRG, Département de Biologie Structurale et Bioinformatique, 1 Rue Michel Servet, 1211 Geneva 14, Switzerland. E-mail: yohann.coute@medecine.unige.ch

Many nuclear proteins are known to interact dynamically with one or several nuclear domains. For example, fluorescent tagged-Sm proteins are first imported into the nucleus, then accumulate in CBs, pass through nucleoli and finally reach the splicing-factor compartments (Sleeman & Lamond, 1999). Other proteins appear to act as molecular links between nuclear domains. This is the case of Nopp140 which shuttles between nucleoli and CBs (Isaac, Yang, & Meier, 1998). Finally, selected nuclear domains such as CBs and PML bodies move within the nucleus (Muratani et al., 2002; Platani et al., 2002). All these movements are potentially linked to biological activities.

The nucleolus is the most prominent and more extensively studied nuclear domain. In the 1960s, it was demonstrated that nucleoli are the site of ribosome biogenesis. Within nucleoli, genes coding for three of the four ribosomal RNAs (rRNAs) are transcribed, processed and assembled into ribosomal subunits. The sites containing rDNA genes are called nucleolar organizer regions (NORs) and are located on chromosomes 13, 14, 15, 21, and 22 in humans. Each NOR comprises 30–50 repeated rDNA genes in a tandem array separated by non-transcribed spacer DNA (Henderson, Warburton, & Atwood, 1972). Ribosome biogenesis gives rise to the characteristic nucleolar ultra structural organization, which consists of three distinct regions, visible by electron microscopy: the fibrillar center (FC) surrounded by the dense fibrillar component (DFC), which is encircled by the granular component (GC) (Granboulan & Granboulan, 1965; Hadjiolov, 1985; Melese & Xue, 1995; Olson, Dundr, & Szebeni, 2000). The FC contains RNA polymerase I (*Pol* I) and specific factors such as DNA topoisomerase I and the upstream binding factor (UBF). The DFC contains transcription factors, newly synthesized pre-rRNAs and pre-rRNA processing factors. Recent data suggest that the DFC is the primary site of active pre-rRNA synthesis (Huang, 2002). The GC, which is the site of late processing events, contains ribosomal proteins, factors involved in the assembly process, and nearly completed pre-ribosomal subunits. Transcription of rDNA repeats generates 47S pre-rRNAs that are cleaved and processed into 28S, 18S, and 5.8S rRNAs, and concomitantly assembled into large and small ribosomal subunits together with the 5S rRNA. During ribosome biogenesis, rRNAs are subjected to extensive modifications probably concomitantly to pre-rRNA synthesis within nucleoli. The two main modifications are pseudouridylation and 2′-*O*-methylation. In vertebrates some 95 pseudouridines and more than 100 2′-*O*-methyl groups have been observed at highly conserved, functionally important regions of rRNAs (Maden, 1990). The function(s) of these modifications is not yet clear although it appears that they are crucial for ribosome biogenesis and probably for the function of ribosomes in protein synthesis (Maden & Hughes, 1997). These modifications are made by small nucleolar ribonucleoprotein complexes (snoRNPs) composed of snoRNAs of between 60 and 250 nucleotides that belong to the box C/D snoRNAs or the box H/ACA snoRNAs, and of several different proteins. In yeast and human cells, box C/D snoRNAs are associated with at least the nucleolar proteins Snu13p, Nop56p, Nop5p/Nop58p, and Nop1p (fibrillarin) whereas box H/ACA snoRNAs are associated with Gar1p, Nhp2p, Nop10p, and Cbf5p (dyskerin). The RNA part of the snoRNP is required for recognition of target RNAs by base pairing and the modification itself is performed by the proteins of the snoRNP. Box C/D snoRNPs are principally responsible for the conversion of uridine into pseudo-uridine while box H/ACA snoRNPs are principally responsible for the 2′-*O*-methylation. Within the snoRNPs, dyskerin is probably required for conversion of uridine into pseudouridine by a mechanism that remains to be defined and fibrillarin catalyzes the methyl transfer reaction (Decatur & Fournier, 2003). Finally, during the maturation process, the pre-ribosomal particles move from FCs to GCs, and into the nucleoplasm (for a review see Scheer & Hock, 1999).

During the past decade, a number of experiments have demonstrated that nucleoli also sequester regulatory complexes and are involved in other biological functions such as tRNA maturation, maturation, and assembly of non-ribosomal RNPs, cell cycle regulation, cellular ageing, and mRNA transport, which has led to the notion of a plurifunctional nucleolus (Pederson, 1998; Pederson & Politz, 2000; Visintin & Amon, 2000; Andersen et al., 2002, 2005; Scherl et al., 2002; Pendle et al., 2005).

B. The Nucleolar Structure and Function Prior to Proteomics: Historical Overview of the Characterization of Nucleolar Proteins

From the time of the first observations of nucleoli using light microscopy in 1770 by Fontana (Busch & Smetana, 1970), numerous experiments have been undertaken in order to determine whether these nuclear structures truly exist within cells or whether they are artifacts resulting from cell preparation for microscopy. During the 19th century, the existence of nucleoli was unambiguously established using a large variety of basic dyes. The ubiquitous presence of nucleoli in every nucleus-containing cell type examined, together with the fact that nucleoli display a tremendous heterogeneity of morphology and number depending on cell type and status, led early biologists to propose that these structures play a determinant role in cell life.

Initial experimental strategies developed to decipher the function of nucleoli were dedicated to the identification of their composition. *In situ* staining of nucleoli provided evidence that nucleoli contain nucleic acids and proteins. Due to the abundance and the relatively low complexity of their RNA content it was concluded at the beginning of the 1960s that nucleoli contained rRNA, suggesting their involvement in ribosome biogenesis (Scherrer, Latham, & Darnell, 1963). Since that first discovery, it has been well demonstrated that nucleoli are the sites of rRNA synthesis and that they contain many other types of RNAs, notably snoRNAs, which are involved in rRNA maturation (Nazar, 2004). In contrast to RNA, although investigation into the protein content of nucleoli was undertaken at the same period, it is only recently that an extensive list of proteins potentially present within nucleoli at different times of the cell life has emerged (Andersen et al., 2002, 2005; Scherl et al., 2002; Pendle et al., 2005).

Long before these high throughput nucleolar proteome analyses, the first experiments dedicated to characterizing nucleolar proteins involved histochemical staining techniques,

purification of nucleoli followed by diverse analytical techniques such as saline- and acidic-soluble extractions, liquid chromatography, amino-acid composition determination, 1D- and 2D- electrophoresis (2-DE) and, more recently, MS. Except for the analyses involving the identification of proteins by MS, the other global analyses determined that nucleoli contain histones, ribosomal proteins, a few enzymes involved in RNA metabolism and non-ribosomal proteins called accessory proteins that are involved in several steps of ribosome biogenesis.

The 2-DE separation of acid-extracted nucleolar proteins, pre-ribosomal and ribosomal particles from different cell lines and tissues resulted in reproducible maps containing approximately 100 proteins in which ribosomal proteins were present together with several non-ribosomal proteins (Orrick, Olson, & Busch, 1973; Olson et al., 1974; Prestayko et al., 1974). Extensive studies have been performed on these so-called accessory proteins, but only a few of the more abundant nucleolar proteins, such as nucleolin (also called C23), B23 (also known as nucleophosmin, numatrin, NO38, NPM), fibrillarin and dyskerin, were identified and functionally characterized.

Prior to nucleolar proteomics, the protein content of nucleoli was poorly characterized, with no more than 120 different nucleolar proteins reported in the literature (Leung et al., 2003) and less than 35 human proteins annotated as nucleolar in the Swiss-Prot and TrEMBL databases (Couté et al., 2003). Proteomic analyses have therefore been launched in order to identify the proteins present within human nucleoli. From global proteomic analyses of the human nucleolus more than 700 different proteins have been identified, many of which are encoded by novel or uncharacterized genes. The presence within nucleoli of components that have been associated with functions unrelated to ribosome biogenesis has increased the list of the non-traditional roles of the nucleolus. Furthermore, the use of sensitive sequence analysis techniques and the high number of nucleolar proteins identified allowed a phylogenetic evolution study based on the known and the novel conserved motifs of the nucleolar proteins (Staub et al., 2004). The proposed model predicts that, in the different phases of the late and continuous evolution of the nucleolus in early eukaryotes, protein domains with different biochemical functions were recruited to the nucleolus. This model argues against an endosymbiotic origin of the nucleolus and the nucleus.

II. WHEN MASS SPECTROMETRY MET NUCLEOLI

The establishment of the nucleolar proteome using MS-based techniques relies first on a highly efficient purification of nucleoli. In the three published research articles describing the characterization of the human nucleolar proteome (Andersen et al., 2002, 2005; Scherl et al., 2002), similar procedures were used, adapted from Busch and Muramatsu's method established nearly 40 years ago (Busch et al., 1963; Muramatsu & Onishi, 1978). This protocol is based on an efficient purification of nuclei followed by their disruption by sonication, which releases the nucleoli. Nucleoli, which have a high density, can then be isolated from the other nuclear components by sedimentation through a sucrose cushion. The efficiency of this fractionation procedure has been validated by several tests, including light and electron microscopic observations, Western-blots and functional assays.

To date, proteomic analyses of purified nucleoli have followed classical gel-based approaches for the identification of proteins using MS (Andersen et al., 2002, 2005; Scherl et al., 2002). Following the separation of nucleolar proteins by 1- or 2-DE, the peptides extracted after in-gel digestion were analyzed by matrix-assisted laser desorption/ionization (MALDI), nanoelectrospray (nanoES), or liquid chromatography (LC) coupled to MS. These workflows have proved to be robust and relatively simple to integrate in laboratories with access to MS and tandem MS (MS/MS) instruments. The excellent sensitivity of such approaches has been confirmed by the types of proteins that have been identified in these proteomic analyses of nucleoli preparations. Proteins of low abundance and proteins not previously described to be associated with nucleoli were identified.

For the analysis of proteins from 2-DE gels, MALDI-TOF MS and peptide mass fingerprinting (PMF) has been widely employed and was used for the analysis of 2-DE-separated proteins in initial analyses of nucleolar preparations (Andersen et al., 2002; Scherl et al., 2002). With an increase in sample complexity, however, methods permitting MS/MS analysis of peptides have become central to proteomics (reviewed in Aebersold & Mann, 2003). In the three proteomic studies of the nucleolus that have been published to date, MS/MS analysis was used to obtain additional information on peptides that were extracted from 1-DE gel bands. In the first study by Andersen et al. (2002), nanoES was used for MS/MS of peptides from an unseparated mixture that had been extracted from gel bands. The MS/MS analysis was performed using a quadrupole-time-of-flight (quad-ToF) mass spectrometer. Scherl et al. (2002) employed ESI-LC-MS/MS of peptides extracted from 1-DE gel bands on the same type of mass spectrometer. In the most recent report, Andersen et al. (2005) also used LC-MS/MS on two different ESI instruments, the quad-ToF instrument used in the first study and an ion trap-Fourier transform ion cyclotron resonance (ion trap-FT-ICR) mass spectrometer.

A. ToF Instruments

Resurgence in ToF mass analysis followed the implementation of new (and rediscovered) tools for improving the resolution and mass accuracy and the concomitant advances in electronics. These improvements included the use of delayed extraction ToF analyzers for MALDI and the introduction of orthogonal acceleration (oa-ToF) (for a review see Guilhaus, Selby, & Mlynski, 2000).

ToF mass spectrometers have been widely used for proteomics analyses. MALDI-ToF MS has been the method of choice for PMF analyses of gel-separated proteins. Hybrid mass spectrometers employing a ToF as the final stage of mass analysis have been widely used with nanoES ionization (Wilm & Mann, 1994, 1996) or LC-ESI-MS/MS for the analysis of more complex samples. These types of instruments include the EBE-ToF configuration (magnetic sector-ToF), ion trap-ToF configurations, and the quadrupole-ToF instrument used in the analyses

of the nucleolar proteome (Andersen et al., 2002, 2005; Scherl et al., 2002).

B. Trapping Instruments

The 3-D ion trap (or Paul trap) has been used in proteomic analyses for more than a decade. In this instrument the ions are trapped within three electrodes (a ring electrode and two endcap electrodes) and subsequently scanned out of the trap to a detector (for reviews, see McLuckey et al., 1994; Jonscher & Yates, 1997). The 3-D ion trap is a very robust, bench top instrument that is sensitive and has a fast scan rate. It does, however, suffer from low trapping efficiencies, low resolution, and poor mass accuracy.

More recently 2-D ion traps have become commercially available (Hager, 2002; Schwartz, Senko, & Syka, 2002). The 2-D quadrupole ion traps have an increased ion capacity compared to 3-D ion traps and a higher trapping efficiency leads to enhanced sensitivity (Schwartz, Senko, & Syka, 2002). The 2-D ion trap is also capable of much faster scan speeds than the 3-D ion trap. Both types of ion trap have limited mass accuracy and resolution. Slowing the scan speed can improve the resolution obtained but these scan types must be performed over a smaller mass range to limit the space charging effects in the trap (Yates, 2004).

Another tandem mass spectrometer that has been recently commercialized and has important implications in the field of proteomics is the linear ion trap-FT-ICR, as used by Andersen et al. (2005) for the most recent proteomic analysis of the nucleolus. This instrument couples the 2-D linear quadrupole ion trap described above with an FT-ICR (Syka et al., 2004). FT-ICR-MS provides unsurpassed resolution and mass accuracy. The utility of the linear ion trap-FT is exemplified in the data from the most recent analysis of the nucleolar proteome. Over 11,000 unique peptide sequences were unambiguously matched. The combination of LC-MS/MS with the high scan rate of the linear ion trap and high mass accuracy of the FT-ICR has resulted in an increased number of identified proteins and the capability of applying more stringent mass errors in database searching.

C. Database Searching Strategies

Several methods have been developed for the identification of proteins from databases using information derived from MS analysis (for a review see Sadygov, Cociorva, & Yates, 2004). PMF uses the masses of peptides from an enzymatic digest to identify proteins from an *in silico* digest of proteins in a database. The method relies on the enzymatic cleavage of separated proteins and is not amenable to the analysis of complex mixtures. The other two techniques rely on peptide sequence information produced by MS/MS of peptides from an enzymatic digest. The most widely applied approach uses software algorithms to compare the experimental MS/MS data against proteins from a database. The software retrieves a "match" from the database that most closely corresponds to the experimental data. Alternatively, the MS/MS data can be processed in a more manual fashion to generate a potential sequence or partial sequence of the peptide and this sequence can then be searched against the database. This manual approach, sometimes referred to as sequence-only searching, requires considerable user input and is therefore limited to the analysis of small numbers of peptides.

For the analysis of the nucleolar proteome, a combination of the above methods has been applied. In the initial studies (Andersen et al., 2002; Scherl et al., 2002) MALDI-TOF data from *in gel* digests of 2-DE gel spots were used to identify proteins using PMF. In both of these studies, 1-DE gel bands were analyzed by MS/MS. Scherl et al. submitted LC-MS/MS data to a software algorithm, Mascot (Perkins et al., 1999), to search the data against Swiss-Prot and TrEMBL databases. The retrieved matches were manually validated as required. Manual sequence-only searching of interpreted spectra was used to check uncertain matches and to search high quality MS/MS spectra that did not result in any matches. In the first publication of the nucleolar proteome, Andersen et al. (2002) used nanoES, and database searching was performed immediately whilst data acquisition continued. Peptide sequence tags (Mann & Wilm, 1994) were generated and searched against databases to retrieve protein matches. A minimum of two peptides were selected and fragmented for each protein and then additional peptides from the same proteins could be excluded based on theoretical masses from the database. Andersen et al. (2002) reported that this method was very useful for the identification of novel proteins. The second nucleolar proteome publication from Andersen et al. (2005) employed LC-MS/MS for the analysis of gel-separated proteins using two different mass spectrometers. The data from both instruments was submitted to Mascot (Perkins et al., 1999) for automated searching of the data against different databases. The accurate data obtained from the linear ion trap-ICR instrument permitted a narrower mass error range in the database search (3 ppm).

A major challenge that has arisen with the analysis of larger and larger volumes of data is how to distinguish between true "positive" identifications and "false positives." Furthermore, how can we reduce the number of "false positive" results without increasing the number of "false negatives," i.e., identifications that are true matches but have scores that fall below the criteria for acceptance? The development of the field of proteomics has also led to the use of MS data by researchers who have little or no experience in MS and there is therefore a temptation to accept protein identifications regardless of the quality of the data and search criteria. These problems have been addressed by a number of researchers (Keller et al., 2002; Peng et al., 2003a; Cargile, Bundy, & Stephenson, 2004) and the importance of publishing full details of the criteria used for the acceptance of protein identifications is becoming more widely understood. A consortium has also been established to develop general criteria for the publication and exchange of MS data and database search results (Kaiser, 2002) and guidelines are now in place to assist researchers in the publication of protein identification from MS data (Carr et al., 2004). A method to estimate the rate of false positive identifications using a "reverse database" (Peng et al., 2003a) has provided a technique for reporting the stringency of the search parameters (Cargile, Bundy, & Stephenson, 2004). This approach was applied to the analysis of nucleolar proteins in the most recent publication of the nucleolar proteome (Andersen et al., 2005).

III. RESULTS OF HUMAN NUCLEOLAR PROTEOMICS

A. Establishment of the Human Nucleolar Proteome and Functional Classification of Proteins Identified within Purified Nucleoli

The two first published proteomic analyses of human nucleoli purified from HeLa cells identified 257 proteins (Andersen et al., 2002) and 210 proteins (Scherl et al., 2002). The combination of these complementary results provided a set of approximately 350 different proteins that form the basis for the understanding of nucleolar functions. Recently, a new study identified 667 proteins within nucleoli prepared from HeLa cells (Andersen et al., 2005). The results of these three analyses are available in online databases at respectively www.dundee.ac.uk/lifesciences/lamonddatabase/, www.expasy.org/ch2d/, and lamondlab.com/nopdb/. The comparison of the results obtained from the three independent analyses using BLASTp searches resulted in a list of 713 individual proteins (Supplemental Table). Recently, a proteomic analysis of nucleoli purified from *Arabidopsis* cells led to the identification of 217 different proteins among which approximately 70% were homologues to proteins identified in the human nucleolar proteome (Pendle et al., 2005). The nucleolar proteome of baker's yeast has been investigated using GFP-fusion proteins (Huh et al., 2003). This study led to the identification of 164 nucleolar proteins. Compiling these results with the 219 proteins annotated as nucleolar in the *Saccharomyces* Genome Database (www.yeastgenome.org) provides a set of 249 yeast nucleolar proteins. Among them, 163 (65%) have a direct counterpart in the human nucleolar proteome (Supplemental Table). These results suggest that the basic protein content of nucleoli is highly conserved throughout the eukaryotic kingdom.

Compiling a list of the proteins identified from the nucleolar fraction does not provide an understanding of the biological roles of nucleoli. Extensive data and text mining is required to assign biological functions to the identified proteins and group them into functional classes. Using classifications adapted from (Scherl et al., 2002), the 713 different proteins identified within purified nucleoli from HeLa cells can be divided into 9 major functional classes: ribosomal proteins, ribosome biogenesis, chromatin structure, mRNA metabolism, translation, chaperones, fibrous proteins, others, and unpredictable function (Fig. 1 and Supplemental Table). The class "mRNA metabolism" can be further subdivided into the following subgroups: transcription, splicing, 3′ end cleavage, and polyadenylation, editing, stability/degradation, export/trafficking, and plurifunctional factors; and "others" can be subclassified into DNA replication, DNA repair, mitosis/cytokinesis/cell cycle regulation, ubiquitination/protein degradation, sumoylation, nucleocytoplasmic transport, cytoskeleton organization/transport, kinases/phosphatases, other enzymes, and others. This classification tends to reinforce the notion of plurifunctional nucleoli. Further experiments are needed to fully understand the role of nucleoli in each of these biological processes. Nevertheless, nucleolar proteomics provides molecular bases for such an understanding.

The functional classification of the proteins identified in individual proteomic studies of human nucleoli reveals the complementary nature of the studies published to date. Nevertheless, some differences in the relative importance of some protein classes were observed between the studies. In the two first studies (Andersen et al., 2002; Scherl et al., 2002), ribosomal proteins and proteins involved in ribosome biogenesis together represented more than one half of the identified proteins (Fig. 1A,B) whereas in the third study these protein categories represented only one third of the identified proteins (Andersen et al., 2005) (Fig. 1C). In contrast, the proportion of proteins involved in mRNA metabolism and grouped in the class called "others" was much more highly represented in the most recent study (more than one third of the identified proteins) (Fig. 1C). This appears to be directly linked to the total number of proteins identified in the most recent study, in which approximately twice as many proteins were identified compared to the first two studies. This indicates that technological improvements facilitated a more extensive survey of the nucleolar proteome. These improvements allowed the identification of proteins that are weakly concentrated within nucleoli, but probably also led to the identification of more contaminants inherent in any subcellular proteomics approach, as described in "Drawbacks and Limitations."

B. Exploring the Nucleolar Proteome to Unravel its Biological Functions

1. Disclosure of the Human Proteins Involved in Ribosome Biogenesis

Ribosome biogenesis is a complex process that involves the synthesis and assembly of four different mature rRNA molecules (in eukaryotes, 28S/25S, 5.8S, 5S for the large ribosomal subunit 60S, and 18S for the small subunit (SSU) 40S) and about 80 different proteins. The key steps of ribosome biogenesis take place within nucleoli and involve more than 100 different proteins that are not part of the ribosome itself (Kressler, Linder, & de La Cruz, 1999; Venema & Tollervey, 1999). This process is very efficient, since 14,000 ribosomal subunits can be synthesized every minute in an exponentially growing cell (Gorlich & Mattaj, 1996). Genomic-based high-throughput studies have identified numerous novel proteins directly or indirectly involved in rRNA processing (Peng et al., 2003b; Krogan et al., 2004). Recently, several groups have used functional proteomic approaches to identify the complexes and factors involved in ribosome biogenesis in *Saccharomyces cerevisiae* (Bassler et al., 2001; Harnpicharnchai et al., 2001; Dragon et al., 2002; Fatica et al., 2002; Grandi et al., 2002; Nissan et al., 2002; Saveanu et al., 2003; Schafer et al., 2003; Horsey et al., 2004). For this, specific proteins have been TAP-tagged, the complexes in which they are present purified and the proteins contained within the complexes identified by MS. This led to the identification of new *trans*-acting factors involved in ribosome biogenesis. Furthermore, the clustering of the proteins identified in the different complexes according to the intracellular localization of the tagged proteins and to the maturation stages of the containing pre-rRNAs revealed the dynamics of ribosome biogenesis in yeast (Table 1). These studies demonstrated that ribosome biogenesis is a highly dynamic and coordinated process giving rise to ribosomal sub-

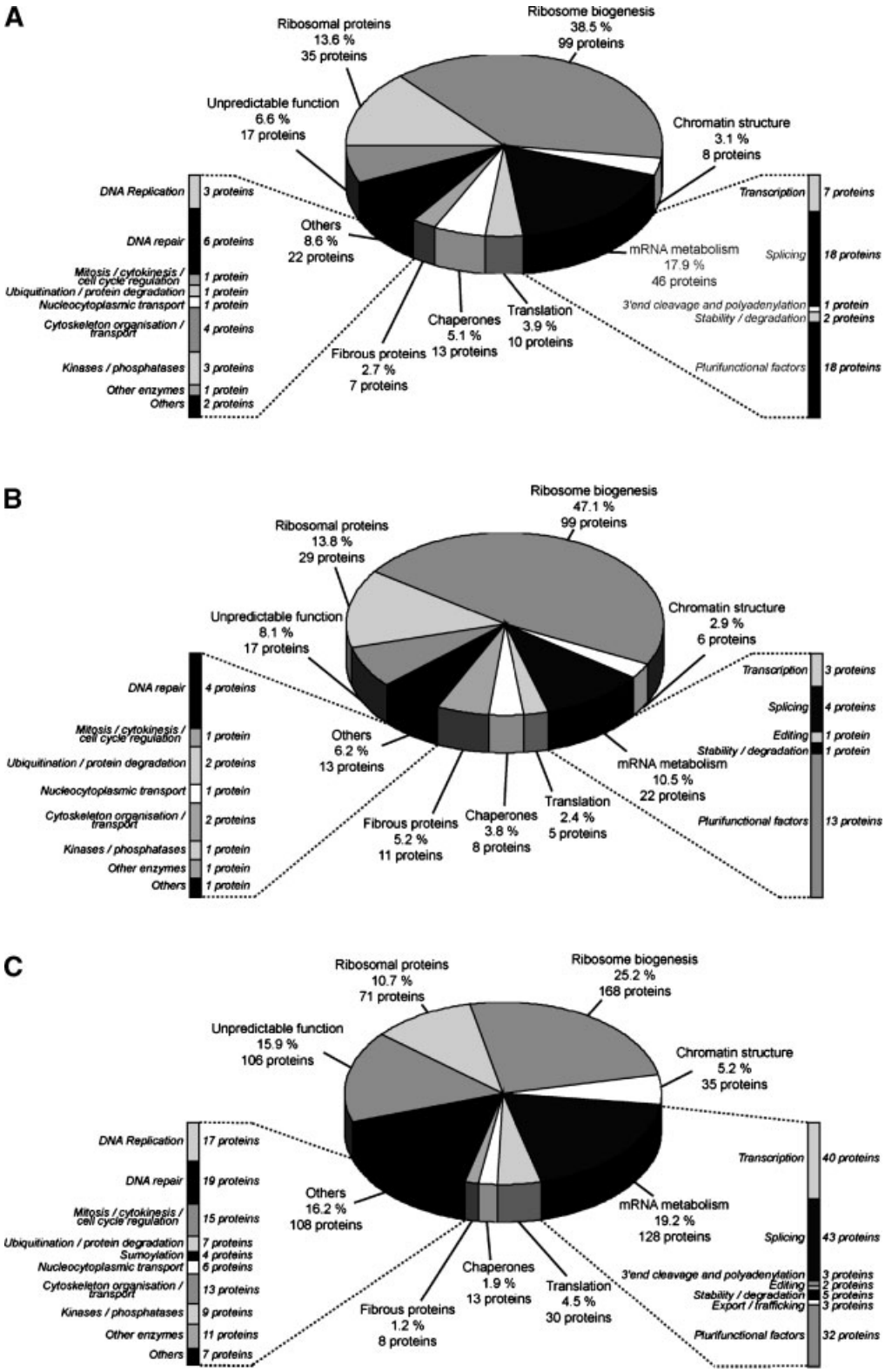

A
Ribosome biogenesis
38.5 %
99 proteins
Ribosomal proteins
13.6 %
35 proteins
Unpredictable function
6.6 %
17 proteins
Chromatin structure
3.1 %
8 proteins
Others
8.6 %
22 proteins
mRNA metabolism
17.9 %
46 proteins
Chaperones
5.1 %
13 proteins
Translation
3.9 %
10 proteins
Fibrous proteins
2.7 %
7 proteins
DNA Replication 3 proteins
DNA repair 6 proteins
Mitosis / cytokinesis / cell cycle regulation 1 protein
Ubiquitination / protein degradation 1 protein
Nucleocytoplasmic transport 1 protein
Cytoskeleton organisation / transport 4 proteins
Kinases / phosphatases 3 proteins
Other enzymes 1 protein
Others 2 proteins
Transcription 7 proteins
Splicing 18 proteins
3'end cleavage and polyadenylation 1 protein
Stability / degradation 2 proteins
Plurifunctional factors 18 proteins
B
Ribosome biogenesis
47.1 %
99 proteins
Ribosomal proteins
13.8 %
29 proteins
Unpredictable function
8.1 %
17 proteins
Chromatin structure
2.9 %
6 proteins
Others
6.2 %
13 proteins
mRNA metabolism
10.5 %
22 proteins
Translation
2.4 %
5 proteins
Chaperones
3.8 %
8 proteins
Fibrous proteins
5.2 %
11 proteins
DNA repair 4 proteins
Mitosis / cytokinesis / cell cycle regulation 1 protein
Ubiquitination / protein degradation 2 proteins
Nucleocytoplasmic transport 1 protein
Cytoskeleton organisation / transport 2 proteins
Kinases / phosphatases 1 protein
Other enzymes 1 protein
Others 1 protein
Transcription 3 proteins
Splicing 4 proteins
Editing 1 protein
Stability / degradation 1 protein
Plurifunctional factors 13 proteins
C
Ribosomal proteins
10.7 %
71 proteins
Ribosome biogenesis
25.2 %
168 proteins
Unpredictable function
15.9 %
106 proteins
Chromatin structure
5.2 %
35 proteins
Others
16.2 %
108 proteins
mRNA metabolism
19.2 %
128 proteins
Chaperones
1.9 %
13 proteins
Translation
4.5 %
30 proteins
Fibrous proteins
1.2 %
8 proteins
DNA Replication 17 proteins
DNA repair 19 proteins
Mitosis / cytokinesis / cell cycle regulation 15 proteins
Ubiquitination / protein degradation 7 proteins
Sumoylation 4 proteins
Nucleocytoplasmic transport 6 proteins
Cytoskeleton organisation / transport 13 proteins
Kinases / phosphatases 9 proteins
Other enzymes 11 proteins
Others 7 proteins
Transcription 40 proteins
Splicing 43 proteins
3'end cleavage and polyadenylation 3 proteins
Editing 2 proteins
Stability / degradation 5 proteins
Export / trafficking 3 proteins
Plurifunctional factors 32 proteins

FIGURE 1.

units 60S and 40S after numerous sequential events of rRNA maturation and protein assembly and disassembly (reviewed in Fatica & Tollervey, 2002; Fromont-Racine et al., 2003; Tschochner & Hurt, 2003).

The majority of the human factors involved in ribosome biogenesis were unknown until the first publications of nucleolar proteomics. One of the most striking results of human nucleolar proteomics is the uncovering of 170 different proteins that are potentially involved in ribosome biogenesis. For the vast majority of these proteins (~70%), no data were available from the literature and they have been grouped in this functional class based upon their homology with other factors, which are for most part of yeast origin. Sequence homology searches were carried out using the BLASTp network service (www.expasy.org/tools/blast) between each of the proteins without any known function and all the proteins present in the Swiss-Prot and TrEMBL databases. The obtained data suggest that a large proportion of the proteins involved in rRNA processing have been conserved through evolution, notably the 5′–3′ exonuclease RAT1, the exosome bearing a 3′–5′ exonucleasic activity, and the MRP RNase involved in the separation of the precursors of 5.8S and 28S rRNAs.

According to the sequential events described in yeast (Table 1), it is now possible to propose a model of ribosome biogenesis in human (Fig. 2). Concomitantly to transcription of the 47S pre-rRNA by RNA *Pol* I, a 90S pre-ribosomal particle containing the pre-rRNA, a number of ribosomal proteins and processing and assembly factors is formed. One of the most abundant proteins of human nucleoli, nucleolin, may play a major role in the coupling of pre-rRNA transcription and pre-ribosomal particle formation (Roger et al., 2003). The 90S pre-ribosomal particle contains processing factors that are mainly involved in the maturation of the small ribosomal subunit and notably contains the huge U3 snoRNP now known as SSU processome (Dragon et al., 2002; Bernstein et al., 2004). After separation of the small and large pre-ribosomal subunits, numerous factors associate and dissociate sequentially from the pre-40S and pre-60S particles to process the pre-rRNAs as well as to assemble the subunits before leaving the nucleolus. The transport of the large ribosomal subunit precursor from the nucleolus to the nucleoplasm is dependent on the exchange of the NOC1–NOC2 complex by the NOC3–NOC2 complex (Milkereit et al., 2001), whereas for the SSU precursor, the NOC4–NOP14 heterodimer may play a major role (Milkereit et al., 2003). In the nucleoplasm, the pre-ribosomal subunits undergo some re-arrangements that allow their subsequent export to the cytoplasm. Several factors able to induce such a remodeling of the pre-60S particle in yeast have been identified: RIX7 (Gadal et al., 2001a), REA1 and the RIX1 complex (composed of RIX1, IPI1 and IPI3) (Nissan et al., 2004). The nucleo-cytoplasmic export of both pre-ribosomal subunits is dependent on several nucleoporins and karyopherins as well as the Ran GTPase (Hurt et al., 1999; Moy & Silver, 1999; Stage-Zimmermann, Schmidt, & Silver, 2000). In yeast, Nmd3p associates with the ribosomal protein Rpl10p and acts as an adaptor to link the pre-60S particle to the exportin Crm1p (Ho, Kallstrom, & Johnson, 2000; Gadal et al., 2001b). In human, the same mechanism is probably used and NMD3 seems to associate with the pre-60S particle in the nucleolus (Trotta et al., 2003). Recently, it has been demonstrated in yeast that RRP12 is required for export of both ribosomal subunits, probably through its binding to several nucleoporins and to the Ran GTPase (Oeffinger, Dlakic, & Tollervey, 2004).

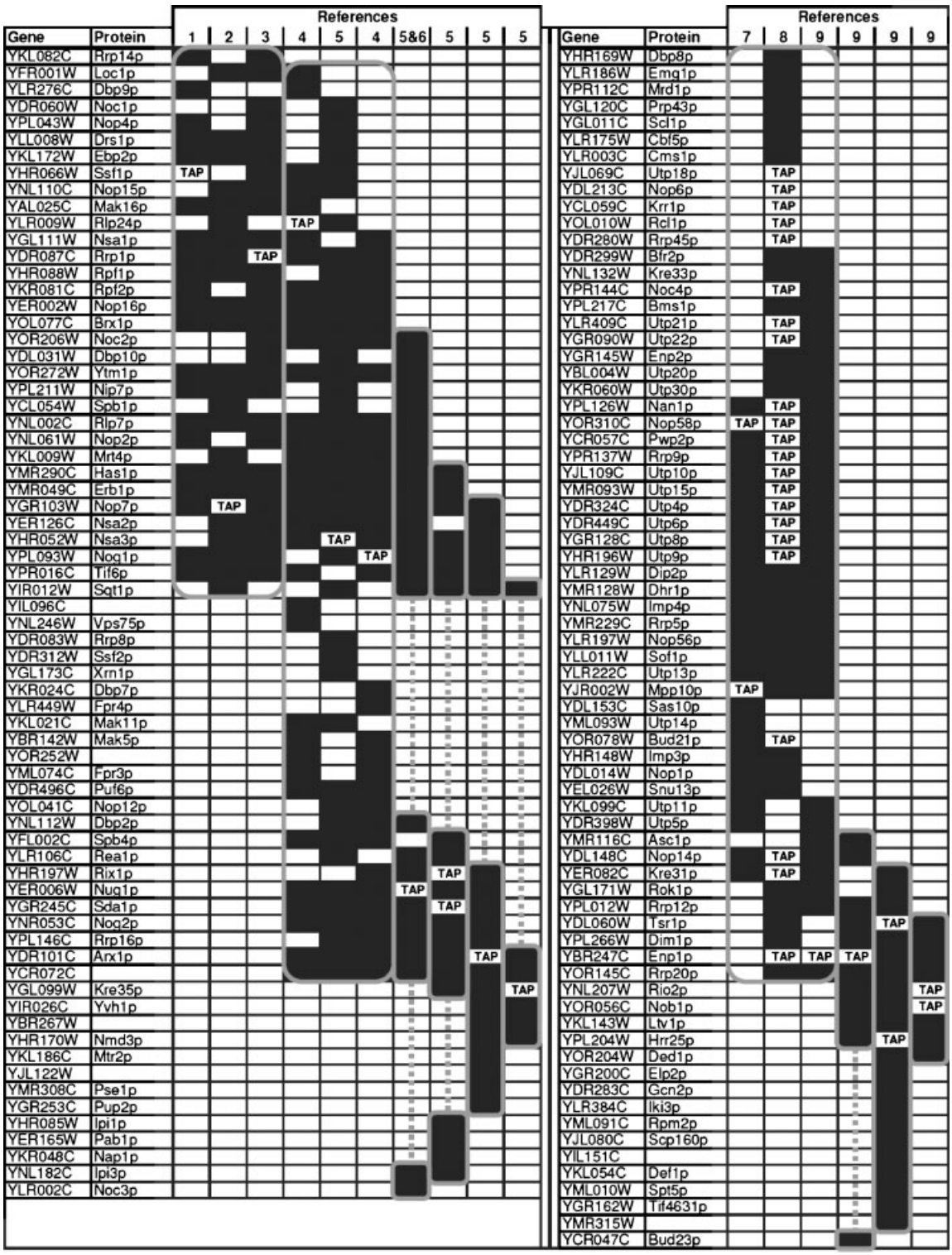
In addition to the 170 proteins most likely involved in ribosome biogenesis (Supplemental Table, Fig. 2), several other factors grouped within other functional classes might play a role in this process in humans. For example, the DNA–PK complex is able to regulate rRNA synthesis (Kuhn et al., 1995). Ubiquitin has been shown to be coupled to some ribosomal proteins, facilitating ribosome biogenesis (Finley, Bartel, & Varshavsky, 1989) and, in yeast, an ubiquitin ligase seems to be involved in rRNA processing and export (Neumann et al., 2003). The yeast homolog of p68 and p72 helicases, Dbp2p, as well as the eIF6 homolog, Tif6p, are involved in the 60S subunit processing (Basu et al., 2001; Bond et al., 2001). In *E. coli*, some chaperones have been shown to participate in ribosome biogenesis (El Hage, Sbai, & Alix, 2001; Maki, Schnobrich, & Culver, 2002). Finally, several proteins identified in nucleolar proteomic analyses have also been found associated with human pre-ribosomal particles (Yanagida et al., 2001, 2004; Fujiyama et al., 2002; Hayano et al., 2003) (see Supplemental Table) and may also be involved in ribosome biogenesis in humans.

2. *Nucleoli in mRNA Metabolism*

Another striking result obtained from the proteomic analyses of nucleoli is the unexpected proportion of identified proteins (~20%) that are involved in at least one step of mRNA metabolism. Several of these proteins may play a role in ribosome biogenesis as described above. It is difficult to conceive, however, that all these proteins could be *trans*-acting factors of ribosomal subunits processing and assembly. Therefore, even if the nucleolar localization of most of these identified proteins needs to be confirmed, several direct and indirect roles of nucleoli in mRNA metabolism have already been reported. Various snRNPs that are involved in mRNA splicing transiently localize within nucleoli where their snRNAs are modified by nucleolar enzymes (Gerbi, Borovjagin, & Lange, 2003). It is now commonly accepted that mRNAs leave their synthesis and maturation sites

FIGURE 1. Functional classification of the proteins identified in the three published nucleolar proteomics studies. The known or predicted biological functions of the proteins identified and listed in the Supplemental Table allow their grouping into nine functional classes. **A**: Functional classification of the 257 proteins identified in Andersen et al. (2002). **B**: Functional classification of the 210 proteins identified in Scherl et al. (2002). **C**: Functional classification of the 667 proteins identified in Andersen et al. (2005). In these three independent studies, the name, the percentage of total proteins, and the number of proteins are given for each class. The "mRNA metabolism" and "others" classes have been divided into sub-classes and the number of proteins present in each sub-class is given.

TABLE 1. Clustering of the proteins identified in the pre-ribosomal particles isolated in *Saccharomyces cerevisiae*



Pre-ribosomal particles have been isolated in yeast by several teams. Techniques used for isolation of complexes varied but identification of their protein content always used MS identification of the binding partners of TAP-tagged proteins. The names of genes and corresponding proteins that have been identified are shown in the two first columns. These data have been collected from the following references: 1. Fatica et al. (2002); 2. Harnpicharnchai et al. (2001); 3. Horsey et al. (2004); 4. Saveanu et al. (2003); 5. Nissan et al. (2002); 6. Bassler et al. (2001); 7. Dragon et al. (2002); 8. Grandi et al. (2002); 9. Schafer et al. (2003). The presence of a protein in isolated pre-ribosomal particles is indicated with a black rectangle. Baits are designated as "TAP." Pre-60S particles are encircled on the left side of the table and pre-40S particles are encircled on the right side.

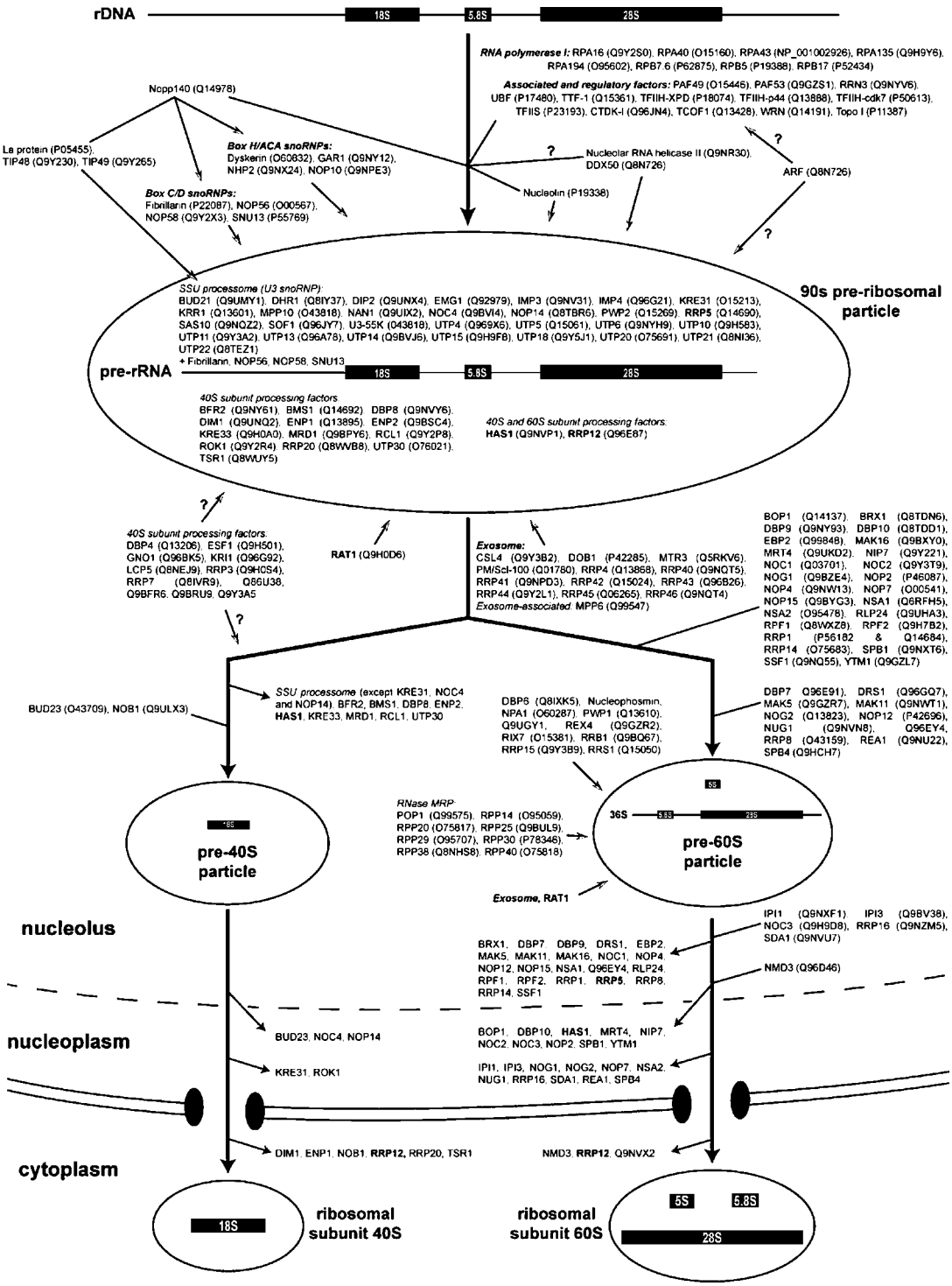


FIGURE 2. Model of ribosome biogenesis dynamics in humans. Bibliographic and bioinformatic searches resulted in a set of 170 different proteins identified by nucleolar proteomics and potentially involved in ribosome biogenesis. By comparison to the model built for yeast, a human ribosome biogenesis pathway can be proposed. Each protein is represented by its usual name (or the usual name of its homolog allowing its classification) and its Swiss-Prot or TrEMBL accession number is given in brackets. Proteins and complexes thought to be involved in maturation of both ribosomal subunits are written in bold.

in speckles to reach, by passive diffusion, the nuclear pores from where they will be exported into the cytoplasm. However, several studies have revealed the nucleolar localization of mRNAs and a role of nucleoli in mRNA export was therefore proposed (Ideue et al., 2004 and references herein). Translation factors have been identified in nucleoli and it has also been demonstrated that other molecules involved in mRNA translation, such as the signal recognition particle (SRP) and tRNAs, are localized within nucleoli (reviewed in Pederson & Politz, 2000). Taken together these observations highlight the possibility that the translation machinery could be pre-assembled within these nuclear domains prior to their export into the cytoplasm. Alternatively, the translation machinery could be used directly within nuclei, coupling transcription and translation, a mechanism recently proposed (Iborra, Jackson, & Cook, 2001; Brogna, Sato, & Rosbash, 2002) but still debated (Dahlberg, Lund, & Goodwin, 2003). On the other hand, it may suggest a role in mRNA quality control within the nucleus (Wilkinson & Shyu, 2002).

3. *Other Nucleolar Functions*

Nucleoli, known to be a retention site of regulatory proteins, are involved in cell cycle regulation (Visintin & Amon, 2000). In yeast, the Cdc14 phosphatase is maintained within nucleoli throughout its interaction with Net1p during the G_1, S, and G_2 phases of the cell cycle. At the end of anaphase, Cdc14p is released from nucleoli and dephosphorylates its targets, resulting in the inactivation of mitotic cyclin-dependent kinases and the exit of mitosis. Cdc14p is then sequestered within nucleoli until the next anaphase (Bachant & Elledge, 1999; Geymonat, Jensen, & Johnston, 2002). The same type of events seems to regulate the functions of Cms1p, Lsr4p (Rabitsch et al., 2003), and Pch2p (San-Segundo & Roeder, 1999), which are involved in meiosis. In mammals, this regulation mode may well be used for NuSAP, a nucleolar protein that re-localizes on microtubules during mitosis and that seems to play a crucial role in mitotic spindle organization (Raemaekers et al., 2003). The use of growing, non-synchronized cells for nucleolar proteomic analyses has not addressed the role of nucleoli in cell cycle regulation. Several proteins involved in this process have, however, been identified within nucleoli purified from HeLa cells and it would be of great interest to determine if their nucleolar localization regulates their functions.

A link between aging and nucleoli has also been described in yeast (reviewed in Sinclair, Mills, & Guarente, 1998). In mammals, mutations that inactivate the nucleolar WRN helicase are responsible for Werner's syndrome, a disease characterized notably by premature aging. However, the model used for nucleolar proteomics did not address the role of nucleoli in cellular aging.

Finally, nucleoli may be involved in the regulation of apoptosis (Horky et al., 2002). The presence within nucleoli of proteins such as PARP-1 and DEDD supports this hypothesis. Furthermore, it has been shown that nucleolar localization of the tumor suppressor ING1 facilitates apoptosis following UV stress (Scott et al., 2001). Several proteins identified in the nucleolar proteomics studies are known to be involved in DNA repair, which reinforces the hypothesis that nucleoli could be major cellular stress sensors (Olson, 2004).

IV. DRAWBACKS AND LIMITATIONS

A. Sub-Cellular Fractionation and Organelle Cross-Contamination

Sub-cellular fractionation generally consists of two major steps: disruption of the cell assembly and organization (homogenization step) and separation of the different functional complexes of the cell (fractionation step) (Huber, Pfaller, & Vietor, 2003). The cell fractionation technique used for nucleoli purification relies first on an efficient separation of nuclei from other cell components. Nuclei are then ruptured by sonication and nucleoli purified by differential centrifugation owing to their high density. A high degree of enrichment is crucial and has been demonstrated by several tests but perfect purity is almost impossible. Some sub-cellular compartments or artefactual complexes formed during cell disruption may share a similar density and co-fractionate to some extent with nucleoli and thus be a potential source of contamination. Moreover, since nucleoli are membrane-less structures, some components may "stick" to nucleoli and co-purify with them. Distinguishing contaminant proteins from "bonafide" nucleolar proteins is therefore a crucial issue. As already described above, the nucleolar proteome has been widely investigated, resulting in the identification of several hundred proteins. The global proteomics analyses of nucleoli have identified a number of unexpected proteins, which raises the question as to whether they are contaminants, or whether they interact transiently with nucleolar elements. These questions can partially be answered by localization validation. When antibodies are available, the cellular immuno-localization can be studied. Alternatively, fusion proteins containing either an immunoreactive or fluorescent tag can potentially validate the nucleolar localization but achieving the tagging of several hundreds of proteins represents an enormous effort. Furthermore, the presence of a tag could potentially lead to the mislocalization of particular proteins. More recently, Ranish et al. (2003) proposed an ICAT labeling strategy followed by quantitative MS to potentially distinguish between cellular non-specific contaminants and the sub-proteome of interest. Proteins originating in the sub-fraction of interest were cysteine labeled with the heavy stable isotope and the proteins from the starting sample with light stable isotope. Proteins were combined and enzymatically cleaved. Peptides were purified and analyzed by MS/MS. The relative abundance of labeled peptide pairs was determined. Specific "bonafide" proteins were detected by their increased abundance in the sub-fraction sample compared to the cellular contaminants. This strategy can potentially be used with any labeling technology including iTRAQ (Ross et al., 2004) and SILAC (Ong et al., 2002). Theoretically, these strategies should exclude any cellular contaminants. However, only the low affinity non-specific proteins are ruled out. The stable contaminants will still be detected with increased abundance and thus included in the sub-cellular proteome of interest. This strategy involves some risk as it may give misleading confidence in contaminating proteins.

A number of well-known nucleolar proteins remain to be detected using current proteomics strategies. This is the case for subunits of the transcription factor SL1 (Learned, Cordes, & Tjian, 1985), the telomerase reverse transcriptase (Yang et al.,

2002), and the SRP68 protein (Politz et al., 2000). Several reasons could explain their absence. Being part of a membraneless sub-organelle, some of the nucleolar proteins that have an association of low stability with nucleolar components are probably lost during the multi-step fractionation procedures. Sample preparation and solubilization prior to separation is a key question especially for poorly soluble proteins. Despite the advances in the last 20 years there are still technical challenges in efficient protein separation. Separation of very acidic, basic, small or large proteins can cause major problems with classical 1-DE, 2-DE, and LC approaches. Other proteins are present in only minute abundance, which prevents their identification using current MS technology. An alternative reason for the absence of known nucleolar proteins could be due to the highly dynamic state of the nucleolar proteome. Photobleaching techniques have very nicely shown the brief occurrence and continuous exchange of proteins between nucleoli and the surrounding nucleoplasm (Phair & Misteli, 2000; Chen & Huang, 2001). The existence of certain proteins in nucleoli is also highly dependent on the synthesis of ribosome subunits and the presence of other components (Andersen et al., 2005).

B. Global Versus Directed Analysis: Deciphering the Whole Nucleolar Proteome Versus Subset of Protein Complexes

Nucleolar proteins are organized into multi-functional protein complexes that control and execute biological processes. Many biochemical and catalytic activities are regulated and modulated through protein–protein interactions. Most of the spatio-temporal information of these multi protein complexes is lost when the whole nucleolar proteome is analyzed. An understanding of the interactions between proteins within the nucleolus is crucial to gain a functional understanding of nucleolar proteins. Projects directed at revealing protein interactions within the nucleolus may require different methodologies than projects designed to examine differences in protein composition. In proteomics strategies the analysis of protein complexes and interactions usually implies the use of ''affinity'' techniques for the isolation and/or enrichment of the complexes prior to MS analysis. A variety of affinity-based purification techniques are available including co-immunoprecipitation with specific antibodies and pull-down experiments using epitope-tagged proteins. Potentially extremely powerful, affinity techniques allow, in principal, the purification and detection of many direct and indirect protein partners and are suitable for any cell type or organism. Compared with the yeast two-hybrid system, the affinity techniques do not suffer from the limited binary interactions. On the other hand, the yeast two-hybrid system might be more efficient to detect transient or unstable interactions.

Four studies from Takahashi's group have described small-scale efforts to detect the interacting partners of four nucleolar bait proteins, nucleolin, parvulin, hNOP56 and fibrillarin, using pull-down or co-immunoprecipitation assays (Yanagida et al., 2001, 2004; Fujiyama et al., 2002; Hayano et al., 2003). More than 100 similar proteins (50% ribosomal and 50% non-ribosomal) were reported to potentially associate with each bait. Most of the interactions were RNA dependant and very few direct partners were found. Moreover, the four proteins are known to interact with rRNAs and pre-ribosomal complexes. Owing to these known associations with additional molecules, these experiments have not been able to define the precise involvement of the baits in ribosome biogenesis but give a basis for a better understanding of the interactions between molecules within human nucleoli.

Specific nucleolar protein–protein interactions are highly dynamic. A major challenge for the future will be to unequivocally delineate the precise roles of each protein in each biological process. Proteomics can help in this task by comprehensively and rapidly comparing and analyzing the protein profiles of complexes in different functional states or isolated using different ionic strengths. However, there is no magic, global, and unique proteomics strategy to fully characterize the nucleolar proteome and sub-nucleolar complexes. Any experiment is merely a snapshot that represents a static version of the nucleolar machinery. Comparative proteomics needs to be repeated nearly an infinite number of times to comprehensively cover any cell state and the dynamic protein content of nucleoli. Integration of these data with orthogonal functional genomic datasets may facilitate the interpretation of functional diversity and cellular processes.

V. PERSPECTIVES: IMPROVING THE CHARACTERIZATION OF THE NUCLEOLAR PROTEOME

A. Mass Spectrometry Techniques

As mentioned above, the published studies on the nucleolar proteome have relied on a limited range of experimental workflows, each involving 1- or 2-DE and MS analysis using MALDI-ToF for PMF or ESI-MS/MS for peptide sequence information. There are, however, many available workflows to choose from when attempting to establish the proteome of a tissue, body-fluid, cell-type, sub-cellular organelle, or sub-cellular structure. The well-known limitations of 2-DE (i.e., limited number of proteins, the inability to analyze hydrophobic proteins, very basic proteins, and very large proteins and a limited dynamic range) have shifted the emphasis of proteomics to include different types of separations, including 1-DE and liquid-based separations (Rabilloud, 2002). Many laboratories have begun to focus on methods that circumvent the use of gels in the separation of proteins. Liquid-based separations may result in the identification of groups of proteins that are not well represented in gel-based approaches. Figure 3 shows a range of alternative strategies available to researchers when undertaking a proteome analysis. The left side of the figure shows the gel-based approaches that have been used to date for the analysis of the nucleolar proteome. The right-hand side shows potential workflows using liquid-based methods. A combination of the approaches can also be envisaged, for example shotgun proteomics using isoelectric focusing (IEF) gels as presented in Figure 3 and published by Cargile, Talley, and Stephenson (2004). Future experiments for the characterization of the nucleolar proteome could call on a

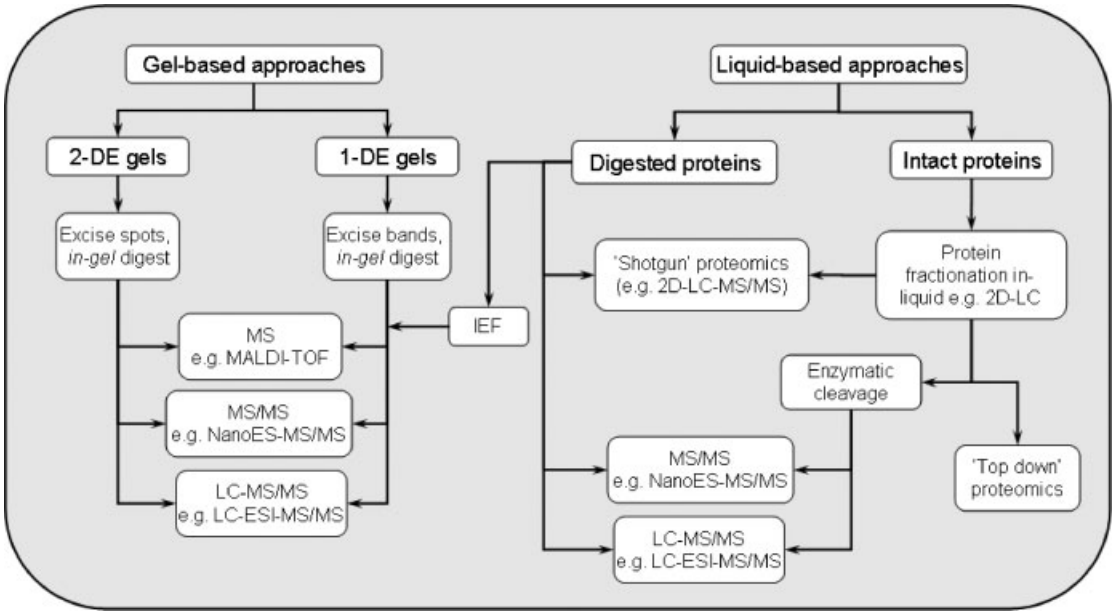


FIGURE 3. Summary of various proteomics workflows.

combination of the above workflows. Furthermore, the use of MALDI-MS/MS, which has shown to give complementary results to ESI-MS/MS (Bodnar et al., 2003) may also provide additional information.

1. Liquid-Based Proteomics Workflows

In the multidimensional protein identification technology (MuDPIT) or shotgun proteomics approach, the entire protein sample is digested, producing a highly complex mixture of peptides that is subsequently analyzed by 2D-LC-MS/MS (Link et al., 1999; Yates, 2004). A combination of the presence of highly abundant proteins, a large number of co-eluting peptides and ion suppression effects in the MS source limit the resolving capacity of the system. Some of these drawbacks can be minimized by prefractionation of the proteins prior to MuDPIT analysis (Wienkoop et al., 2004).

Another approach that is becoming more widely used is 2-D chromatography for the separation of proteins followed either by ''top down'' proteomics or in-solution digestion and LC-MS/MS analysis (Opiteck et al., 1997, 1998; Wang & Hanash, 2005). Capillary electrophoresis has also been used in multidimensional separations of proteins prior to MS (Zhou & Johnston, 2004).

2. Choice of Ionization Technique

ESI and MALDI ionization processes provide complementary results and the choice of whether MALDI or ESI is used will often depend on access to specific tandem MS instruments.

With the recent introduction of MALDI tandem mass spectrometers (Laiko, Baldwin, & Burlingame, 2000; Medzihradszky et al., 2000; Huang et al., 2002), the use of LC-MALDI techniques has substantially risen. In this technique the eluent is spotted onto a MALDI sample plate and analyzed at a later point in time, whereas with LC-ESI techniques the eluent is sprayed directly into the ESI source and data acquired in ''real-time.''

LC-MALDI techniques do not suffer from the time constraints imposed by the transient presence of peptides eluting from a column. The acquisition of data is limited by sample consumption from the MALDI process rather than time. However, because LC-MALDI is a decoupled technique, the data is not available immediately, as is the case with ESI analyses. This has implications in the detection of chromatographic problems. A further disadvantage of LC-MALDI is that the resolution of the separation is dependent on the time between each fraction deposited on the plate. To overcome this drawback, systems have been developed that use continuous deposition of the column eluent (Wall et al., 2002).

The complementary nature of MALDI and ESI MS/MS has been demonstrated using a post-column flow split for concurrent LC-ESI-MS and LC-MALDI sample deposition (Wattenberg et al., 2002; Bodnar et al., 2003). MALDI-MS/MS spectra may contain additional structural information that is not present in the equivalent ESI-MS/MS spectra, however, the multiply charged ions generated by ESI often produce MS/MS spectra that are cleaner and potentially easier to interpret (Bodnar et al., 2003).

3. The Potential of FT-ICR

The recent publication of Andersen et al. (2005) showed the utility of a hybrid tandem mass spectrometer employing FT-ICR as the last stage of mass analysis. FT-ICR-MS is unique in that it provides unsurpassed resolution and mass accuracy. The development of hybrid mass spectrometers employing FT-ICR has created new opportunities for MS analysis in the field of proteomics and further developments in this technology hold much promise for proteome research.

FT-ICR-MS has also been applied to proteomics analyzes without the use of MS/MS (reviewed in this special issue by Zimmer and co-workers). The approach uses the high mass accuracy and resolution of the FT-ICR to obtain very accurate molecular masses of the peptides (accurate mass tags), which are then searched against masses obtained from an *in silico* digest of proteins from a database (Smith et al., 2002). The accurate mass tags approach combined with refined developments over the last decade in nanoscale capillary LC have led to an accurate mass and time (AMT) tag approach (described in Zimmer and co-workers) that benefits from high dynamic range, mass resolution, and mass accuracy.

A further application of FT-ICR has been in ''top down'' proteomics, where proteins are not cleaved into peptides but directly analyzed by MS and MS/MS. A major challenge lies in the separation of intact proteins in sufficient concentrations for this type of analysis. The ''top down'' approach for proteomics has a major advantage in that information on the intact protein (such as PTMs and alternative splice forms) is retained (for a review see Reid & McLuckey, 2002).

B. Revealing Nucleolar Post-Translational Modifications

Chemical modifications that occur following translation of a protein are critical in the regulation of enzyme activity, protein transport and localization, protein interactions, and protein metabolism. The identification and characterization of these modifications is by no means a trivial task. Over 300 different types of PTMs are known (Jensen, 2004) and often the modifications are only present transiently and in low levels. Little is known about the PTMs of nucleolar proteins and the functional role of such PTMs (Leary & Huang, 2001). The binding of nucleolin to rDNA requires phosphorylation and the same enzyme that phosphorylates nucleolin is known to phosphorylate other nucleolar proteins. Some of the results of these phosphorylation events have been noted although other effects are not clear. It is known, however, that phosphorylation plays a role in both the positive and negative regulation of rRNA synthesis. It has also been suggested that the translocation of pre-ribosomal particles from the nucleolus to the nucleoplasm is at least partially mediated by phosphorylation. Acetylation and methylation are also known to be involved in the synthesis of rRNAs. Further studies into the role of PTMs in nucleoli are sure to reveal additional PTMs that are important and may help improve our understanding of the functional roles of those PTMs that have already been observed in nucleoli.

The analysis of PTMs using MS has been extensively reviewed (Schweppe et al., 2003; Cantin & Yates, 2004; Jensen, 2004; Seo & Lee, 2004; Meng et al., 2005) and therefore readers are directed to the literature for additional information. Briefly, five different approaches have been used to identify PTMs using MS. These approaches are (i) gel-based methods, (ii) affinity enrichment methods, (iii) chemical derivatization methods, (iv) quantitative *in vivo* labeling, and (v) MS-specific methods.

Gel-based methods for the identification of PTMs use specific staining of proteins that have been separated by 1-DE or 2-DE. Staining can be performed once the proteins have been separated in the gel or after blotting the gel-separated proteins on to a membrane. Alternatively, the proteins can be radio-labeled (^{32}P) prior to separation, and the proteins that have incorporated the label detected following separation. The proteins can then be excised, *in gel* digested, and identified using MS.

Affinity chromatographic techniques have been extensively investigated for the enrichment of proteins with specific PTMs. The goal is to increase the concentration of the modified proteins or peptides to facilitate their detection and analysis by MS. This category includes immunoprecipitation of phosphoproteins using antibodies, enrichment of phosphopeptides using immobilized metal ion affinity chromatography (IMAC) and the use of lectin columns for the enrichment of glycoproteins and glycopeptides.

The third approach relies on the conversion of the PTM group into another moiety that can be more easily detected using MS methods. This chemical modification may be designed to be less labile than the original modification, to have ideal properties for affinity selection or MS analysis, or to provide a specific cleavage site for an enzyme. These approaches have been used for the analysis of phosphoproteins and glycoproteins (Oda, Nagasu, & Chait, 2001; Rusnak, Zhou, & Hathaway, 2002). A drawback of labeling strategies and enrichment procedures is the number of processing steps. It is advantageous to minimize the number of sample handling steps required in order to minimize protein losses.

Recently, Pandey and co-workers published the use of SILAC to examine changes in protein phosphorylation (Ibarrola et al., 2003). The principal uses metabolic labeling and immunoprecipitation of a protein of interest. MS analysis and examination of the peptide ratios reveals two different distributions of peptide pairs—unmodified peptides pairs show similar intensities and peptides that have undergone modification reveal a difference in the peptide ratio (Ibarrola et al., 2003). The method has been used to identify and quantify phosphorylation sites (Ibarrola et al., 2003, 2004; Blagoev et al., 2004; Gruhler et al., 2005) and methylation sites (Ong, Mittler, & Mann, 2004).

Finally, capabilities of certain mass spectrometers can be used to detect and characterize PTMs. The fragmentation of some PTMs results in the generation of specific fragment ions or neutral losses that are indicative of the PTM. Precursor ion scans and neutral loss scans can be used to detect peptides that fragment to produce these diagnostic ions. In reality, the technology is not yet sufficiently sensitive to attain the low concentration levels required for the analysis of biologically relevant levels of PTMs. Advances in instrumentation and sample preparation techniques will be required before these methods can be used for the analysis of PTMs in the nucleolar proteome.

One advance in MS that has important implications in the analysis of PTMs is the use of FT-ICR-MS (Meng et al., 2005) and specifically the fragmentation of peptides using electron capture dissociation (ECD). ECD, unlike collision-induced dissociation (CID), results in fragmentation that retains the modified groups on the peptide chain (Cooper, Hakansson, & Marshall, 2005), making it much easier to identify the sites of modification. Further advances in this field of MS could potentially lead to more fruitful characterization of nucleolar proteins' PTMs.

C. Data and Text Mining to Improve the Functional Classification of Proteomics Outcomes

Another area that is crucial for the advancement of our understanding of proteins identified in nucleolar proteomic analyses is data mining.

1. Data Mining and Proteomics

As pointed out earlier, assigning a biological function to proteins identified by MS entails extensive bioinformatics and bibliographic analyses. Indeed, a large collection of bioinformatics resources is available to help correlate protein identity with functional characteristics through the interpretation of sequence similarities, domain composition, or pathway information. The uncertainty or the ambiguity of these interpretations can be reduced *via* further literature searches.

2. What to Expect from Databases

For many years, data resources have promoted rapid increases in the number of database entries, reflecting an incessant production of data. Efforts are often guided by criteria such as maximizing the number of topics or species, aiming towards an exhaustive coverage of the many different aspects of proteins. Consequently, in many large databases, information is often redundant or alternatively, the data is averaged and specificity is lost. The criterion of supplying comprehensive information has resulted in ever-increasing database size, and not necessarily an increase in database quality. In effect, biological data were often collected because they were available rather than being selected for their utility. Resulting trends can be identified in databases. For instance, most protein family databases are biased towards enzyme-related domains, given the traditional tight link between structural domains, and enzymatic activity in protein studies (see for instance Mulder et al., 2003).

The recent release of genome data has modified this view. An exhaustive dataset is now accepted to include only a small number of entities. The number of entries in a proteome database is not expected to grow but information related to each entry should be further explored. Now, properties of objects are expected to be as comprehensive as possible. Database expansion has shifted from breadth to depth and from objects to properties.

As a result, the first step in interpreting proteomics data is a fact-based assessment of what is known about the particular set of proteins and the bioinformatics resources (databases and tools) associated with the generation of these facts. Irregular data production spread over many years implies that a lot can be known about some proteins and very little about others.

3. Text-Based Information

Reviewing research literature is a problem that has already received a great deal of attention. Much of the work has concentrated on integrating information from as many documents as possible, by automatic means. A number of applications have been developed to identify gene or protein names and symbols in biomedical texts (Proux et al., 1998; Krauthammer et al., 2000; Tanabe & Wilbur, 2002; Yu et al., 2002). Once assembled as networks, these enormous collections of data allow hypotheses to be generated and tested using pre-existing data.

4. Text and Sequence Information

The determination of the biological function of a new protein is usually approached as a sequence-matching problem. Sequence similarity is interpreted as functional equivalence so that results with similar sequences from database searches are used to transfer functional features from the annotated sequences to the query sequence. Two types of problems may arise with this method, under- or over-prediction of functional assignments.

Over-prediction of functional assignments based on sequence similarities is a typical error caused by automatic attribution of a function based on a high scoring sequence alignment. A high score can be due to sequence similarity in a region other than that credited to that specific function. This error is apparent in cases where high scoring sequence homologues do not share the same catalytic activity because their active sites are not conserved. It may happen that lower scoring homologues have perfect matches within the conserved catalytic site.

Under-prediction may occur when a sequence contains multiple functional domains. Each of these functional units may have a different level of sequence similarity. In these cases the functional properties may only be partially attributed to the sequence because the annotation is only inferred from the best scoring domain match.

Despite the possibility of incorrect predictions, homology searching is perhaps the most widely used bioinformatics tool. Thus in an attempt to complement and improve these results, text mining of functional links based on document similarity can be employed. Many text-mining methods require a corpus of articles relevant to all of the proteins being studied and the generation of an index associating the article to the appropriate genes or proteins.

Strategies for knowledge discovery, based on data mining combined with literature analysis, have produced promising results. In most cases, sequence homology and/or gene expression profiling together with manual literature analysis help in the interpretation of results (Blaschke, Oliveros, & Valencia, 2001; Chaussabel & Sher, 2002; Raychaudhuri & Altman, 2003). To date no publication has described similar achievements with protein expression data.

5. Literature Provides the Link between Knowledge and Biological Data

As pointed out elsewhere (Lisacek et al., 2004), efforts are geared to federate rather than integrate data. A major challenge in bioinformatics is to initiate the shift from linking the data sources independently to incorporating the textual knowledge into data mining procedures. In fact, information from the literature can not only be part of the data mining process, but also provide guidelines for cross-linking biological data and potentially discover new interactions.

Improvements in data and text mining strategies will therefore be important in the future for understanding the roles

played by nucleoli and nucleolar proteins in normal cells as well as in pathological states. This section cannot exhaustively survey the numerous bioinformatics resources that are commonly used to further characterize proteins. We have addressed only those approaches that were used for analyzing nucleolar data. However, popular methods often rely on exploiting information stored in Gene Ontology (www.geneontology.org) as exemplified by Jensen et al. (2003). Other important initiatives focus on the study of protein–protein interactions (for a review, see Salwinski & Eisenberg, 2003). Even more promising are the so-designated ''systems biology'' studies which attempt to integrate multiple source data into network models (for an example, see Yan et al., 2004).

VI. CONCLUSION

Improvements in fractionation techniques and MS technologies will increase our knowledge of the proteins localized within nucleoli. This will be achieved both by improving the sensitivity of detection and by distinguishing between true nucleolar proteins and contaminating components. Together with developments in data and text mining procedures, the role of nucleoli and nucleolar proteins in normal and aberrant cellular processes will be further clarified. There is increasing evidence suggesting that nucleoli and nucleolar proteins could be involved in various pathologies including cancer (Trere, 2000; Ruggero & Pandolfi, 2003; Horn & Vousden, 2004; Pandolfi, 2004) and viral infections (Hiscox, 2002). Functional nucleolar proteomics using well-established cellular models will be a powerful tool to better understand the fundamental functions performed by nucleoli and nucleolar proteins in these pathologies.

REFERENCES

Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. Nature 422:198–207.

Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, Mann M, Lamond AI. 2002. Directed proteomic analysis of the human nucleolus. Curr Biol 12:1–11.

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M. 2005. Nucleolar proteome dynamics. Nature 433:77–83.

Bachant JB, Elledge SJ. 1999. Mitotic treasures in the nucleolus. Nature 398:757–758.

Bassler J, Grandi P, Gadal O, Lessmann T, Petfalski E, Tollervey D, Lechner J, Hurt E. 2001. Identification of a 60S preribosomal particle that is closely linked to nuclear export. Mol Cell 8:517–529.

Basu U, Si K, Warner JR, Maitra U. 2001. The *Saccharomyces cerevisiae TIF6* gene encoding translation initiation factor 6 is required for 60S ribosomal subunit biogenesis. Mol Cell Biol 21:1453–1462.

Bernstein KA, Gallagher JE, Mitchell BM, Granneman S, Baserga SJ. 2004. The small-subunit processome is a ribosome assembly intermediate. Eukaryot Cell 3:1619–1626.

Blagoev B, Ong SE, Kratchmarova I, Mann M. 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol 22:1139–1145.

Blaschke C, Oliveros JC, Valencia A. 2001. Mining functional information associated with expression arrays. Funct Integr Genomics 1:256–268.

Bodnar WM, Blackburn RK, Krise JM, Moseley MA. 2003. Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. J Am Soc Mass Spectrom 14:971–979.

Bond AT, Mangus DA, He F, Jacobson A. 2001. Absence of Dbp2p alters both nonsense-mediated mRNA decay and rRNA processing. Mol Cell Biol 21:7366–7379.

Brogna S, Sato TA, Rosbash M. 2002. Ribosome components are associated with sites of transcription. Mol Cell 10:93–104.

Busch H, Smetana K. 1970. Nucleoli of tumor cells. In: Busch H, Smetana K, editor. The nucleolus. New York: Academic Press. pp 448–471.

Busch H, Muramatsu M, Adams H, Steele WJ, Liau MC, Smetana K. 1963. Isolation of nucleoli. Exp Cell Res 24(Suppl 9):150–163.

Cantin GT, Yates JR III. 2004. Strategies for shotgun identification of post-translational modifications by mass spectrometry. J Chromatogr A 1053:7–14.

Cargile BJ, Bundy JL, Stephenson JL, Jr. 2004. Potential for false positive identifications from large databases through tandem mass spectrometry. J Proteome Res 3:1082–1085.

Cargile BJ, Talley DL, Stephenson JL, Jr. 2004. Immobilized pH gradients as a first dimension in shotgun proteomics and analysis of the accuracy of pI predictability of peptides. Electrophoresis 25:936–945.

Carr S, Aebersold R, Baldwin M, Burlingame A, Clauser K, Nesvizhskii A. 2004. The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. Mol Cell Proteomics 3:531–533.

Carter KC, Bowman D, Carrington W, Fogarty K, McNeil JA, Fay FS, Lawrence JB. 1993. A three-dimensional view of precursor messenger RNA metabolism within the mammalian nucleus. Science 259:1330–1335.

Chaussabel D, Sher A. 2002. Mining microarray expression data by literature profiling. Genome Biol 3:RESEARCH0055.

Chen D, Huang S. 2001. Nucleolar components involved in ribosome biogenesis cycle between the nucleolus and nucleoplasm in interphase cells. J Cell Biol 153:169–176.

Cooper HJ, Hakansson K, Marshall AG. 2005. The role of electron capture dissociation in biomolecular analysis. Mass Spectrom Rev 24:201–222.

Couté Y, Scherl A, Dieckmann R, Callé A, Hoogland C, Kindbeiter K, Déon C, Greco A, Sanchez J-C, Diaz J-J, Hochstrasser DH. 2003. Nucleolar proteomics. In: Cooper DN, editor. Nature encyclopedia of the human genome. London: Macmillan Publishers Ltd., Nature Publishing Group. pp 403–408.

Dahlberg JE, Lund E, Goodwin EB. 2003. Nuclear translation: What is the evidence? RNA 9:1–8.

Decatur WA, Fournier MJ. 2003. RNA-guided nucleotide modification of ribosomal and other RNAs. J Biol Chem 278:695–698.

Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y, Beyer AL, Hunt DF, Baserga SJ. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature 417:967–970.

Dundr M, Misteli T. 2001. Functional architecture in the cell nucleus. Biochem J 356:297–310.

El Hage A, Sbai M, Alix JH. 2001. The chaperonin GroEL and other heat-shock proteins, besides DnaK, participate in ribosome biogenesis in *Escherichia coli*. Mol Gen Genet 264:796–808.

Fatica A, Tollervey D. 2002. Making ribosomes. Curr Opin Cell Biol 14:313–318.

Fatica A, Cronshaw AD, Dlakic M, Tollervey D. 2002. Ssf1p prevents premature processing of an early pre-60S ribosomal particle. Mol Cell 9:341–351.

Finley D, Bartel B, Varshavsky A. 1989. The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. Nature 338:394–401.

Fromont-Racine M, Senger B, Saveanu C, Fasiolo F. 2003. Ribosome assembly in eukaryotes. Gene 313:17–42.

Fujiyama S, Yanagida M, Hayano T, Miura Y, Isobe T, Fujimori F, Uchida T, Takahashi N. 2002. Isolation and proteomic characterization of human Parvulin-associating preribosomal ribonucleoprotein complexes. J Biol Chem 277:23773–23780.

Gadal O, Strauss D, Braspenning J, Hoepfner D, Petfalski E, Philippsen P, Tollervey D, Hurt E. 2001a. A nuclear AAA-type ATPase (Rix7p) is required for biogenesis and nuclear export of 60S ribosomal subunits. EMBO J 20:3695–3704.

Gadal O, Strauss D, Kessl J, Trumpower B, Tollervey D, Hurt E. 2001b. Nuclear export of 60s 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.

Gerbi SA, Borovjagin AV, Lange TS. 2003. The nucleolus: A site of ribonucleoprotein maturation. Curr Opin Cell Biol 15:318–325.

Geymonat M, Jensen S, Johnston LH. 2002. Mitotic exit: The Cdc14 double cross. Curr Biol 12:R482–R484.

Gorlich D, Mattaj IW. 1996. Nucleocytoplasmic transport. Science 271: 1513–1518.

Granboulan N, Granboulan P. 1965. Ultrastructure cytochemistry of the nucleolus II. Study of the sites of RNA synthesis in the nucleolus and the nucleus. Exp Cell Res 38:604–619.

Grandi P, Rybin V, Bassler J, Petfalski E, Strauss D, Marzioch M, Schafer T, Kuster B, Tschochner H, Tollervey D, Gavin AC, Hurt E. 2002. 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. Mol Cell 10:105–115.

Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NF, Mann M, Jensen ON. 2005. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol Cell Proteomics 4:310–327.

Guilhaus M, Selby D, Mlynski V. 2000. Orthogonal acceleration time-of-flight mass spectrometry. Mass Spectrom Rev 19:65–107.

Hadjiolov AA. 1985. The nucleolus and ribosome biogenesis. In: Beermann AM, Goldstein L, Porter KR, Sitte P, editors. Cell biology monographs. Wien/New York: Springer Verlag. pp 1–263.

Hager JW. 2002. A new linear ion trap mass spectrometer. Rapid Commun Mass Spectrom 16:512–526.

Harnpicharnchai P, Jakovljevic J, Horsey E, Miles T, Roman J, Rout M, Meagher D, Imai B, Guo Y, Brame CJ, Shabanowitz J, Hunt DF, Woolford JL, Jr. 2001. Composition and functional characterization of yeast 66S ribosome assembly intermediates. Mol Cell 8:505–515.

Hayano T, Yanagida M, Yamauchi Y, Shinkawa T, Isobe T, Takahashi N. 2003. Proteomic analysis of human Nop56p-associated pre-ribosomal ribonucleoprotein complexes: Possible link between Nop56p and the nucleolar protein treacle responsible for treacher Collins syndrome. J Biol Chem 278:34309–34319.

Henderson AS, Warburton D, Atwood KC. 1972. Location of ribosomal DNA in the human chromosome complement. Proc Natl Acad Sci USA 69:3394–3398.

Hiscox JA. 2002. The nucleolus—A gateway to viral infection? Arch Virol 147:1077–1089.

Ho JH, Kallstrom G, Johnson AW. 2000. Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. J Cell Biol 151:1057–1066.

Horky M, Kotala V, Anton M, Wesierska-Gadek J. 2002. Nucleolus and apoptosis. Ann NY Acad Sci 973:258–264.

Horn HF, Vousden KH. 2004. Cancer: Guarding the guardian? Nature 427:110–111.

Horsey EW, Jakovljevic J, Miles TD, Harnpicharnchai P, Woolford JL, Jr. 2004. Role of the yeast Rrp1 protein in the dynamics of pre-ribosome maturation. RNA 10:813–827.

Huang S. 2002. Building an efficient factory: Where is pre-rRNA synthesized in the nucleolus? J Cell Biol 157:739–741.

Huang L, Baldwin MA, Maltby DA, Medzihradszky KF, Baker PR, Allen N, Rexach M, Edmondson RD, Campbell J, Juhasz P, Martin SA, Vestal ML, Burlingame AL. 2002. The identification of protein–protein interactions of the nuclear pore complex of *Saccharomyces cerevisiae* using high throughput matrix-assisted laser desorption ionization time-of-flight tandem mass spectrometry. Mol Cell Proteomics 1:434–450.

Huber LA, Pfaller K, Vietor I. 2003. Organelle proteomics: Implications for subcellular fractionation in proteomics. Circ Res 92:962–968.

Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. Nature 425:686–691.

Hurt E, Hannus S, Schmelzl B, Lau D, Tollervey D, Simos G. 1999. A novel *in vivo* assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants. J Cell Biol 144:389–401.

Ibarrola N, Kalume DE, Gronborg M, Iwahori A, Pandey A. 2003. A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture. Anal Chem 75:6043–6049.

Ibarrola N, Molina H, Iwahori A, Pandey A. 2004. A novel proteomic approach for specific identification of tyrosine kinase substrates using [13C]tyrosine. J Biol Chem 279:15805–15813.

Iborra FJ, Jackson DA, Cook PR. 2001. Coupled transcription and translation within nuclei of mammalian cells. Science 293:1139–1142.

Ideue T, Azad AK, Yoshida J, Matsusaka T, Yanagida M, Ohshima Y, Tani T. 2004. The nucleolus is involved in mRNA export from the nucleus in fission yeast. J Cell Sci 117:2887–2895.

Isaac C, Yang Y, Meier UT. 1998. Nopp140 functions as a molecular link between the nucleolus and the coiled bodies. J Cell Biol 142:319–329.

Jensen ON. 2004. Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. Curr Opin Chem Biol 8:33–41.

Jensen LJ, Gupta R, Staerfeldt HH, Brunak S. 2003. Prediction of human protein function according to Gene Ontology categories. Bioinformatics 19:635–642.

Jonscher KR, Yates JR III. 1997. The quadrupole ion trap mass spectrometer—A small solution to a big challenge. Anal Biochem 244:1–15.

Kaiser J. 2002. PROTEOMICS: Public-private group maps out initiatives. Science 296:827.

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74:5383–5392.

Krauthammer M, Rzhetsky A, Morozov P, Friedman C. 2000. Using BLAST for identifying gene and protein names in journal articles. Gene 259:245–252.

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

Krogan NJ, Peng WT, Cagney G, Robinson MD, Haw R, Zhong G, Guo X, Zhang X, Canadien V, Richards DP, Beattie BK, Lalev A, Zhang W, Davierwala AP, Mnaimneh S, Starostine A, Tikuisis AP, Grigull J, Datta N, Bray JE, Hughes TR, Emili A, Greenblatt JF. 2004. High-definition macromolecular composition of yeast RNA-processing complexes. Mol Cell 13:225–239.

Kuhn A, Gottlieb TM, Jackson SP, Grummt I. 1995. DNA-dependent protein kinase: A potent inhibitor of transcription by RNA polymerase I. Genes Dev 9:193–203.

Laiko VV, Baldwin MA, Burlingame AL. 2000. Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry. Anal Chem 72:652–657.

Learned RM, Cordes S, Tjian R. 1985. Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I. Mol Cell Biol 5:1358–1369.

Leary DJ, Huang S. 2001. Regulation of ribosome biogenesis within the nucleolus. FEBS Lett 509:145–150.

Leung AK, Andersen JS, Mann M, Lamond AI. 2003. Bioinformatic analysis of the nucleolus. Biochem J 376:553–569.

Lewis JD, Tollervey D. 2000. Like attracts like: Getting RNA processing together in the nucleus. Science 288:1385–1389.

Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR III. 1999. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17:676–682.

Lisacek F, Chichester C, Gonnet P, Jaillet O, Kappus S, Nikitin F, Roland P, Rossier G, Truong L, Appel RD. 2004. Shaping biological knowledge: Applications in proteomics. Comp Funct Genom 5:190–195.

Maden BE. 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. Prog Nucleic Acid Res Mol Biol 39:241–303. Write to the Help Desk NCBI/ NLM/NIH Department of Health & Human Services Privacy Statement/Freedom of Information Act/Disclaimer.

Maden BE, Hughes JM. 1997. Eukaryotic ribosomal RNA: The recent excitement in the nucleotide modification problem. Chromosoma 105:391–400. Write to the Help Desk NCBI/NLM/NIH Department of Health & Human Services Privacy Statement/Freedom of Information Act/Disclaimer.

Maki JA, Schnobrich DJ, Culver GM. 2002. The DnaK chaperone system facilitates 30S ribosomal subunit assembly. Mol Cell 10:129–138.

Mann M, Wilm M. 1994. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem 66:4390–4399.

McLuckey SA, Vanberkel GJ, Goeringer DE, Glish GL. 1994. Ion trap mass spectrometry of externally generated ions. Anal Chem 66:A689–A696.

Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL. 2000. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. Anal Chem 72:552–558.

Melese T, Xue Z. 1995. The nucleolus: An organelle formed by the act of building a ribosome. Curr Opin Cell Biol 7:319–324.

Meng F, Forbes AJ, Miller LM, Kelleher NL. 2005. Detection and localization of protein modifications by high resolution tandem mass spectrometry. Mass Spectrom Rev 24:126–134.

Milkereit P, Gadal O, Podtelejnikov A, Trumtel S, Gas N, Petfalski E, Tollervey D, Mann M, Hurt E, Tschochner H. 2001. Maturation and intranuclear transport of pre-ribosomes requires Noc proteins. Cell 105:499–509.

Milkereit P, Strauss D, Bassler J, Gadal O, Kuhn H, Schutz S, Gas N, Lechner J, Hurt E, Tschochner H. 2003. A Noc complex specifically involved in the formation and nuclear export of ribosomal 40 S subunits. J Biol Chem 278:4072–4081.

Misteli T. 2001. Protein dynamics: Implications for nuclear architecture and gene expression. Science 291:843–847.

Moy TI, Silver PA. 1999. Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins. Genes Dev 13:2118–2133.

Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Barrell D, Bateman A, Binns D, Biswas M, Bradley P, Bork P, Bucher P, Copley RR, Courcelle E, Das U, Durbin R, Falquet L, Fleischmann W, Griffiths-Jones S, Haft D, Harte N, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lopez R, Letunic I, Lonsdale D, Silventoinen V, Orchard SE, Pagni M, Peyruc D, Ponting CP, Selengut JD, Servant F, Sigrist CJ, Vaughan R, Zdobnov EM. 2003. The InterPro Database, 2003 brings increased coverage and new features. Nucleic Acids Res 31:315–318.

Muramatsu M, Onishi T. 1978. Isolation and purification of nucleoli and nucleolar chromatin from mammalian cells. Methods Cell Biol 17: 141–161.

Muratani M, Gerlich D, Janicki SM, Gebhard M, Eils R, Spector DL. 2002. Metabolic-energy-dependent movement of PML bodies within the mammalian cell nucleus. Nat Cell Biol 4:106–110.

Nazar RN. 2004. Ribosomal RNA processing and ribosome biogenesis in eukaryotes. IUBMB Life 56:457–465.

Neumann S, Petfalski E, Brugger B, Grosshans H, Wieland F, Tollervey D, Hurt E. 2003. Formation and nuclear export of tRNA, rRNA, and mRNA is regulated by the ubiquitin ligase Rsp5p. EMBO Rep 4:1156–1162.

Nissan TA, Bassler J, Petfalski E, Tollervey D, Hurt E. 2002. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO J 21:5539–5547.

Nissan TA, Galani K, Maco B, Tollervey D, Aebi U, Hurt E. 2004. A pre-ribosome with a tadpole-like structure functions in ATP-dependent maturation of 60S subunits. Mol Cell 15:295–301.

Oda Y, Nagasu T, Chait BT. 2001. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol 19:379–382.

Oeffinger M, Dlakic M, Tollervey D. 2004. A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits. Genes Dev 18:196–209.

Olson MO. 2004. Sensing cellular stress: Another new function for the nucleolus? Sci STKE 2004:pe10.

Olson MO, Dundr M, Szebeni A. 2000. The nucleolus: An old factory with unexpected capabilities. Trends Cell Biol 10:189–196.

Olson MO, Prestayko AW, Jones CF, Busch H. 1974. Phosphorylation of proteins of ribosomes and nucleolar preribosomal particles from Novikoff hepatmoa ascites cells. J Mol Biol 90:161–168.

Ong S-E, Mittler G, Mann M. 2004. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. Nat Methods 1:119–126.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386.

Opiteck GJ, Lewis KC, Jorgenson JW, Anderegg RJ. 1997. Comprehensive on-line LC/LC/MS of proteins. Anal Chem 69:1518–1524.

Opiteck GJ, Ramirez SM, Jorgenson JW, Moseley MA III. 1998. Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping. Anal Biochem 258:349–361.

Orrick LR, Olson MO, Busch H. 1973. Comparison of nucleolar proteins of normal rat liver and Novikoff hepatoma ascites cells by two-dimensional polyacrylamide gel electrophoresis. Proc Natl Acad Sci USA 70:1316–1320.

Pandolfi PP. 2004. Aberrant mRNA translation in cancer pathogenesis: An old concept revisited comes finally of age. Oncogene 23:3134–3137.

Park PC, De Boni U. 1999. Dynamics of structure–function relationships in interphase nuclei. Life Sci 64:1703–1718.

Pederson T. 1998. The plurifunctional nucleolus. Nucleic Acids Res 26: 3871–3876.

Pederson T, Politz JC. 2000. The nucleolus and the four ribonucleoproteins of translation. J Cell Biol 148:1091–1095.

Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, Mann M, Lamond AI, Brown JW, Shaw PJ. 2005. Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. Mol Biol Cell 16:260–269.

Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. 2003a. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. J Proteome Res 2:43–50.

Peng WT, Robinson MD, Mnaimneh S, Krogan NJ, Cagney G, Morris Q, Davierwala AP, Grigull J, Yang X, Zhang W, Mitsakakis N, Ryan OW, Datta N, Jojic V, Pal C, Canadien V, Richards D, Beattie B, Wu LF, Altschuler SJ, Roweis S, Frey BJ, Emili A, Greenblatt JF, Hughes TR.

2003b. A panoramic view of yeast noncoding RNA processing. Cell 113:919–933.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567.

Phair RD, Misteli T. 2000. High mobility of proteins in the mammalian cell nucleus. Nature 404:604–609.

Platani M, Goldberg I, Lamond AI, Swedlow JR. 2002. Cajal body dynamics and association with chromatin are ATP-dependent. Nat Cell Biol 4:502–508.

Politz JC, Browne ES, Wolf DE, Pederson T. 1998. Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. Proc Natl Acad Sci USA 95: 6043–6048.

Politz JC, Tuft RA, Pederson T, Singer RH. 1999. Movement of nuclear poly(A) RNA throughout the interchromatin space in living cells. Curr Biol 9:285–291.

Politz JC, Yarovoi S, Kilroy SM, Gowda K, Zwieb C, Pederson T. 2000. Signal recognition particle components in the nucleolus. Proc Natl Acad Sci USA 97:55–60.

Prestayko AW, Klomp GR, Schmoll DJ, Busch H. 1974. Comparison of proteins of ribosomal subunits and nucleolar preribosomal particles from Novikoff hepatoma ascites cells by two-dimensional polyacrylamide gel electrophoresis. Biochemistry 13:1945–1951.

Proux D, Rechenmann F, Julliard L, Pillet VV, Jacq B. 1998. Detecting gene symbols and names in biological texts: A first step toward pertinent information extraction. Genome Inform Ser Workshop Genome Inform 9:72–80.

Rabilloud T. 2002. Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains. Proteomics 2: 3–10.

Rabitsch KP, Petronczki M, Javerzat JP, Genier S, Chwalla B, Schleiffer A, Tanaka TU, Nasmyth K. 2003. Kinetochore recruitment of two nucleolar proteins is required for homolog segregation in meiosis I. Dev Cell 4:535–548.

Raemaekers T, Ribbeck K, Beaudouin J, Annaert W, Van Camp M, Stockmans I, Smets N, Bouillon R, Ellenberg J, Carmeliet G. 2003. NuSAP, a novel microtubule-associated protein involved in mitotic spindle organization. J Cell Biol 162:1017–1029.

Ranish JA, Yi EC, Leslie DM, Purvine SO, Goodlett DR, Eng J, Aebersold R. 2003. The study of macromolecular complexes by quantitative proteomics. Nat Genet 33:349–355.

Raychaudhuri S, Altman RB. 2003. A literature-based method for assessing the functional coherence of a gene group. Bioinformatics 19:396–401.

Reid GE, McLuckey SA. 2002. 'Top down' protein characterization via tandem mass spectrometry. J Mass Spectrom 37:663–675.

Roger B, Moisand A, Amalric F, Bouvet P. 2003. Nucleolin provides a link between RNA polymerase I transcription and pre-ribosome assembly. Chromosoma 111:399–407.

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169.

Ruggero D, Pandolfi PP. 2003. Does the ribosome translate cancer? Nat Rev Cancer 3:179–192.

Rusnak F, Zhou J, Hathaway G. 2002. Identification of phosphorylated and glycosylated sites in peptides by chemically targeted proteolysis. J Biomol Tech 13:228–237.

Sadygov RG, Cociorva D, Yates JR III. 2004. Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. Nat Methods 1:195–202.

Salwinski L, Eisenberg D. 2003. Computational methods of analysis of protein–protein interactions. Curr Opin Struct Biol 13:377–382.

San-Segundo PA, Roeder GS. 1999. Pch2 links chromatin silencing to meiotic checkpoint control. Cell 97:313–324.

Saveanu C, Namane A, Gleizes PE, Lebreton A, Rousselle JC, Noaillac-Depeyre J, Gas N, Jacquier A, Fromont-Racine M. 2003. Sequential protein association with Nascent 60S ribosomal particles. Mol Cell Biol 23:4449–4460.

Schafer T, Strauss D, Petfalski E, Tollervey D, Hurt E. 2003. The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO J 22:1370–1380.

Scheer U, Hock R. 1999. Structure and function of the nucleolus. Curr Opin Cell Biol 11:385–390.

Scherl A, Coute Y, Deon C, Calle A, Kindbeiter K, Sanchez JC, Greco A, Hochstrasser D, Diaz JJ. 2002. Functional proteomic analysis of human nucleolus. Mol Biol Cell 13:4100–4109.

Scherrer K, Latham H, Darnell JE. 1963. Demonstration of an unstable RNA and of a precursor to ribosomal RNA in HeLa cells. Proc Natl Acad Sci USA 49:240–248.

Schwartz JC, Senko MW, Syka JE. 2002. A two-dimensional quadrupole ion trap mass spectrometer. J Am Soc Mass Spectrom 13 :659–669.

Schweppe RE, Haydon CE, Lewis TS, Resing KA, Ahn NG. 2003. The characterization of protein post-translational modifications by mass spectrometry. Acc Chem Res 36:453–461.

Scott M, Boisvert FM, Vieyra D, Johnston RN, Bazett-Jones DP, Riabowol K. 2001. UV induces nucleolar translocation of ING1 through two distinct nucleolar targeting sequences. Nucleic Acids Res 29:2052–2058.

Seo J, Lee KJ. 2004. Post-translational modifications and their biological functions: Proteomic analysis and systematic approaches. J Biochem Mol Biol 37:35–44.

Sinclair DA, Mills K, Guarente L. 1998. Molecular mechanisms of yeast aging. Trends Biochem Sci 23:131–134.

Sleeman JE, Lamond AI. 1999. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. Curr Biol 9:1065–1074.

Smith RD, Anderson GA, Lipton MS, Pasa-Tolic L, Shen Y, Conrads TP, Veenstra TD, Udseth HR. 2002. An accurate mass tag strategy for quantitative and high-throughput proteome measurements. Proteomics 2:513–523.

Stage-Zimmermann T, Schmidt U, Silver PA. 2000. Factors affecting nuclear export of the 60S ribosomal subunit *in vivo*. Mol Biol Cell 11:3777–3789.

Staub E, Fiziev P, Rosenthal A, Hinzmann B. 2004. Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire. Bioessays 26:567–581.

Syka JE, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt DF. 2004. Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. J Proteome Res 3:621–626.

Tanabe L, Wilbur WJ. 2002. Tagging gene and protein names in biomedical text. Bioinformatics 18:1124–1132.

Trere D. 2000. AgNOR staining and quantification. Micron 31:127–131.

Trotta CR, Lund E, Kahan L, Johnson AW, Dahlberg JE. 2003. Coordinated nuclear export of 60S ribosomal subunits and NMD3 in vertebrates. EMBO J 22:2841–2851.

Tschochner H, Hurt E. 2003. Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol 13:255–263.

Venema J, Tollervey D. 1999. Ribosome synthesis in *Saccharomyces cerevisiae*. Annu Rev Genet 33:261–311.

Visintin R, Amon A. 2000. The nucleolus: The magician's hat for cell cycle tricks. Curr Opin Cell Biol 12:372–377.

Wall DB, Berger SJ, Finch JW, Cohen SA, Richardson K, Chapman R, Drabble D, Brown J, Gostick D. 2002. Continuous sample deposition from reversed-phase liquid chromatography to tracks on a matrix-assisted laser desorption/ionization precoated target for the analysis of protein digests. Electrophoresis 23:3193–3204.

Wang H, Hanash S. 2005. Intact-protein based sample preparation strategies for proteome analysis in combination with mass spectrometry. Mass Spectrom Rev 24:413–426.

Wattenberg A, Organ AJ, Schneider K, Tyldesley R, Bordoli R, Bateman RH. 2002. Sequence dependent fragmentation of peptides generated by MALDI quadrupole time-of-flight (MALDI Q-TOF) mass spectrometry and its implications for protein identification. J Am Soc Mass Spectrom 13:772–783.

Wienkoop S, Glinski M, Tanaka N, Tolstikov V, Fiehn O, Weckwerth W. 2004. Linking protein fractionation with multidimensional monolithic reversed-phase peptide chromatography/mass spectrometry enhances protein identification from complex mixtures even in the presence of abundant proteins. Rapid Commun Mass Spectrom 18: 643–650.

Wilkinson MF, Shyu AB. 2002. RNA surveillance by nuclear scanning? Nat Cell Biol 4:E144–E147.

Wilm MS, Mann M. 1994. Electrospray and Taylor-cone theory, Doles beam of macromolecules at last. Int J Mass Spectrom Ion Process 136:167–180.

Wilm M, Mann M. 1996. Analytical properties of the nanoelectrospray ion source. Anal Chem 68:1–8.

Yan W, Lee H, Yi EC, Reiss D, Shannon P, Kwieciszewski BK, Coito C, Li XJ, Keller A, Eng J, Galitski T, Goodlett DR, Aebersold R, Katze MG. 2004. System-based proteomic analysis of the interferon response in human liver cells. Genome Biol 5:R54.

Yanagida M, Shimamoto A, Nishikawa K, Furuichi Y, Isobe T, Takahashi N. 2001. Isolation and proteomic characterization of the major proteins of the nucleolin-binding ribonucleoprotein complexes. Proteomics 1:1390–1404.

Yanagida M, Hayano T, Yamauchi Y, Shinkawa T, Natsume T, Isobe T, Takahashi N. 2004. Human fibrillarin forms a sub-complex with splicing factor 2-associated p32, protein arginine methyltransferases, and tubulins alpha 3 and beta 1 that is independent of its association with preribosomal ribonucleoprotein complexes. J Biol Chem 279:1607–1614.

Yang Y, Chen Y, Zhang C, Huang H, Weissman SM. 2002. Nucleolar localization of hTERT protein is associated with telomerase function. Exp Cell Res 277:201–209.

Yates JR III. 2004. Mass spectral analysis in proteomics. Annu Rev Biophys Biomol Struct 33:297–316.

Yu H, Hatzivassiloglou V, Friedman C, Rzhetsky A, Wilbur WJ. 2002. Automatic extraction of gene and protein synonyms from MEDLINE and journal articles. Proc AMIA Symp 919–923.

Zhou F, Johnston MV. 2004. Protein characterization by on-line capillary isoelectric focusing, reversed-phase liquid chromatography, and mass spectrometry. Anal Chem 76:2734–2740.

Zimber A, Nguyen QD, Gespach C. 2004. Nuclear bodies and compartments: Functional roles and cellular signalling in health and disease. Cell Signal 16:1085–1104.

Yohann Couté received his Ph.D. from the University of Lyon, France in December 2003, where he worked on the characterization of the nucleolar proteome of human cells, notably using proteomic strategies. He is currently a post-doctoral research fellow in the Biomedical Proteomics Research Group at the Medical Faculty of Geneva. His interests include the discovery of new molecular targets for the treatment of colorectal cancer dissemination using sub-cellular proteomics strategies.

Jennifer A. Burgess graduated with a degree in Medicinal Chemistry from the University of Wollongong, Australia in 1998 and gained her Ph.D. in Chemistry from the same University in 2001. Her thesis concerned the use of mass spectrometry for the analysis of peptides and proteins and included work from an 8 month scholarship to study at UMIST in Manchester, U.K. Following her Ph.D., she spent 6 months as a post-doc at the Australian Cataract Research Foundation, where she carried out research on modified lens proteins. From 2002 to April 2005 Jennifer worked at the University of Geneva as a post-doctoral researcher in mass spectrometry. There her research focused on the separation and analysis of proteins from plasma, urine, and CSF. She currently works in the U.K. for Waters Ltd. as an M.S. specialist.

Jean-Jacques Diaz received his Ph.D. from Claude Bernard Lyon 1 University in 1988 after studying the role of ribosomal protein phosphorylation in translational regulation. He moved to Kansas State University (August 1988–January 1991) to develop a new saturated mutagenesis technology in order to study the structure–function relationships of several ribosomal proteins. In 1991, he joined the Medical University of Lyon to develop a research program to study the host cell viral interactions using herpes simplex virus type 1. In 1993, he became a member of the Institut National de la Santé Et de la Recherche Medicale. Since 1999 he is the group leader of a team whose main objective is to identify the structure and the function of several nuclear domains at the Centre de Génétique Moléculaire et Cellulaire in Lyon. His recent research interest is also to develop new kinds of anti-herpes molecules.

Christine Chichester received her Ph.D. from the University of California, Davis in 1993 after studying mechanism cytochrome P450 dependent cytotoxicity by developing techniques for analytical measurements using fluorescence microscopy and image analysis, and HPLC and MS. During a post-doctoral fellowship at the Stanford University, Palo Alto, California she continued to study cytochrome P450 from a molecular biological perspective. The focus on bioinformatics mining of genomic and genetic information for cytochrome P450 lead to an interest representation of biological data, information retrieval and the design of new knowledge engineering processes. She continues to pursue these interests in the private sector in an effort to drive knowledge discovery from publicly available scientific literature.

Frédérique Lisacek received a Ph.D. in Computer Science (Artificial Intelligence) from the University Pierre & Marie Curie, Paris, France in 1984. From then on she has held research positions in bioinformatics in France, Japan, and Australia working on knowledge representation and sequence analysis in various fields of biology. Since 2000, she shares research and teaching activities in proteomics between the Proteome Informatics Group of the Swiss Institute of Bioinformatics in Geneva and Geneva Bioinformatics (GeneBio) S.A., while being on temporary leave from a tenured position in the Genome & Informatics Laboratory in Evry, France.

Anna Greco received her Ph.D. at the University Claude Bernard of Lyon in 1980 after studying the role of tRNAs in the optimization of translation. Since 1982 her research focuses on herpes viruses. She became a member of the Institut National de la Santé Et de la Recherche Médicale in 1983. She identified the protein responsible of the reactivation of Epstein Barr virus, and she has shown that herpes simplex virus type 1 induces a translational control of gene expression. Her recent interests center on the role of nucleolar proteins in the outcome of viral infection, and on the development of new kinds of anti-viral compounds based on cellular protein targets.

Jean-Charles Sanchez has been working since 1989 in the field of proteomics. He contributed in the development of two-dimensional polyacrylamide gel electrophoresis and related procedures and their application in biomarker discovery research. He obtained his Ph.D. in Biochemistry at the Buckingham University (U.K.) in the field of proteomics and diabetes. In 1997, he obtained a Business Administration diploma from the Officer School of Lausanne. Since 1995, he has been the head of the Biomedical Proteomics Group at the BPRG at the Geneva University Hospital and director of the Proteomics Facility of the Faculty of Medicine. He is a founder of the Swiss Proteomics Society, Swiss-2Dservice, and Phenomics Sciences Consulting. The activities in his group cover the discovery of new disease markers and drug targets in multi-factorial disorders including cancer, alzheimer, CJD, stoke, obesity, and type II diabetes.