Molecular Biology of the Cell
Vol. 16, 260–269, January 2005

Proteomic Analysis of the *Arabidopsis* Nucleolus Suggests Novel Nucleolar Functions[D]

Alison F. Pendle,* Gillian P. Clark,† Reinier Boon,* Dominika Lewandowska,† Yun Wah Lam,‡ Jens Andersen,§ Matthias Mann,§ Angus I. Lamond,‡ John W. S. Brown,† and Peter J. Shaw*‖

*John Innes Centre, Norwich NR4 7UH, United Kingdom; †Scottish Crop Research Institute, Dundee DD2 5DA, Scotland; ‡University of Dundee, Dundee DD1 5HN, Scotland; and §University of Southern Denmark, DK-5230 Odense, Denmark

Submitted September 9, 2004; Accepted October 13, 2004
Monitoring Editor: Joseph Gall

The eukaryotic nucleolus is involved in ribosome biogenesis and a wide range of other RNA metabolism and cellular functions. An important step in the functional analysis of the nucleolus is to determine the complement of proteins of this nuclear compartment. Here, we describe the first proteomic analysis of plant (*Arabidopsis thaliana*) nucleoli, in which we have identified 217 proteins. This allows a direct comparison of the proteomes of an important nuclear structure between two widely divergent species: human and *Arabidopsis*. The comparison identified many common proteins, plant-specific proteins, proteins of unknown function found in both proteomes, and proteins that were nucleolar in plants but nonnucleolar in human. Seventy-two proteins were expressed as GFP fusions and 87% showed nucleolar or nucleolar-associated localization. In a striking and unexpected finding, we have identified six components of the postsplicing exon-junction complex (EJC) involved in mRNA export and nonsense-mediated decay (NMD)/mRNA surveillance. This association was confirmed by GFP-fusion protein localization. These results raise the possibility that in plants, nucleoli may have additional functions in mRNA export or surveillance.

INTRODUCTION

The eukaryotic nucleus is a complex, highly organized organelle containing a range of domains and nuclear bodies involved in DNA replication and gene expression (Lamond and Earnshaw, 1998). The most prominent nuclear subcompartment is the nucleolus, which is the site of transcription and processing of ribosomal RNAs (rRNAs), and their assembly into ribosomal subunits before export to the cytoplasm (Olsen *et al.*, 2000; Lafontaine and Tollervey, 2001; Fatica and Tollervey, 2002; Scheer and Hock, 1999). The nucleolus therefore contains the RNA polymerase I transcription machinery and the full range of small nucleolar ribonucleoprotein particles (snoRNPs) required to modify specific bases and to cleave and process sites in the precursor rRNA (pre-rRNA; Filipowicz and Pogacic, 2002; Kiss, 2002). It also contains ribosomal proteins (r-proteins), ribosome biogenesis factors required for rRNA folding and the step-wise pathway of ribosome biogenesis, and factors required for the export of the ribosomal subunits through the nuclear pore complex to the cytoplasm (Fatica and Tollervey, 2002).

The nucleolus is also implicated in a wider range of other activities and processes, such as maturation, assembly and export of RNP particles, telomerase activity, the cell cycle, and cell growth and aging (Lamond and Earnshaw, 1998; Pedersen, 1998; Cockell and Gasser, 1999; Scheer and Hock, 1999; Carmo-Fonseca *et al.*, 2000; Olsen *et al.*, 2000; Visintin and Amon, 2000; Shaw, 2004). Furthermore, it has recently been suggested that the nucleolus acts as a sensor for cellular stress, coordinating the cell's response after nucleolar disruption (Rubbi and Milner, 2003; Horn and Vousden, 2004) and thereby highlighting the importance of the nucleolus to the functional integrity of the cell.

Besides the nucleolus, animal nuclei contain a variety of other nuclear bodies or domains including Cajal bodies, gems, speckles, paraspeckles and PML bodies (Lamond and Earnshaw, 1998; Matera, 1999; Fox *et al.*, 2002). These nuclear subcompartments are involved in different aspects of transcription, RNA processing, maturation of RNP complexes and particles, and transport of RNAs and RNPs. The nucleolus is often closely associated with Cajal bodies, which function in the maturation and recycling of snRNPs; assembly, processing or trafficking to the nucleolus of snoRNPs; 3′ end processing of some histone mRNAs, and modification of snRNAs by small Cajal body-specific RNAs (scaRNAs; Sleeman and Lamond, 1999; Jady and Kiss, 2001; Darzacq *et al.*, 2002; Jady *et al.*, 2003). Nuclear proteins and RNPs are in dynamic interaction with chromatin and the various nuclear bodies and diffuse through the nucleoplasm in interchromatin regions. Accumulation of specific components in different nuclear structures occurs to carry out different functions and the composition of the bodies changes in response to cellular conditions and environmental cues (Sleeman and Lamond, 1999; Phair and Misteli, 2000; Chen and Huang, 2001; Leung and Lamond, 2002; Lamond and Spector, 2003).

Although the fundamental processes of ribosome biogenesis are very similar in all eukaryotes, there are significant

Article published online ahead of print. Mol. Biol. Cell 10.1091/mbc.E04–09–0791. Article and publication date are available at www.molbiolcell.org/cgi/doi/10.1091/mbc.E04–09–0791.

[D] The online version of this article contains supplemental material at MBC Online (http://www.molbiolcell.org).

‖ Corresponding author. E-mail address: peter.shaw@bbsrc.ac.uk.

differences in nucleolar organization between plants and animals, and even between different cell types within a single species. This suggests that there is more to nucleolar organization than making ribosomes; these differences may originate in other aspects of nucleolar function. In animal nucleoli, transcription occurs in the dense fibrillar component (DFC; Koberna *et al.*, 2002), often at the surface of the closely associated fibrillar centers (FC). The early processing of rRNAs occurs in the DFC and ribosomal subunits are assembled in the granular component (GC; Shaw and Jordan, 1995; Scheer and Hock, 1999; Carmo-Fonseca *et al.*, 2000). In plant nucleoli, the active rDNA transcription units are well dispersed throughout a region of the nucleolus broadly corresponding to the DFC, which constitutes up to 70% of the nucleolar volume in plants and which is surrounded by the GC; association of transcription with FCs is not necessary (Shaw and Jordan, 1995; Gonzalez-Melendi *et al.*, 2001). The rDNA transcription units have been imaged in plants as compacted "Christmas Trees" ~300 nm in length (Gonzalez-Melendi *et al.*, 2001). Recent studies of human HeLa cell nucleoli have reached a similar conclusion that transcription units are in the form of compacted, convoluted structures within the comparatively smaller DFC (Koberna *et al.*, 2002). The localization of a variety of proteins, rRNA and snoRNA species to subdomains within plant nucleoli correlated well with early and late events in rRNA processing (Beven *et al.*, 1996). From this, Brown and Shaw (1998) proposed a "vectorial" model of organization, in which successive biochemical steps in rRNA maturation and ribosome biogenesis occur in concentric layers enveloping the transcription sites. Finally, a distinctive feature of most plant nucleoli is a central region called the nucleolar cavity. Its function is unknown, but time-lapse microscopy has shown the cavity apparently emptying its contents into the nucleoplasm (Gunning, URL: www.plantcellbiologyonCD.com) and Beven *et al.* (1995, 1996) showed accumulation of snoRNAs and the presence of spliceosomal snRNAs in the nucleolar cavity.

Although many components of the nucleolus and other nuclear bodies have been identified, a full understanding of their functions will require far greater knowledge of their component proteins and other macromolecules, their interactions, and their responses to changes in cellular activity. The availability of complete genome sequences, together with rapid advances in mass spectrometry methods for analyzing complex polypeptide mixtures, means that it is now possible to begin to provide a complete identification of the proteins for cellular organelles or substructures of the size and complexity of the nucleolus (Rappsilber and Mann, 2002; Aebersold and Mann, 2003; Taylor *et al.*, 2003). Recent studies have provided proteomic analyses of the human nucleolus (Andersen *et al.*, 2002; Scherl *et al.*, 2002), leading ultimately to the identification of 692 different proteins (Lamond and Mann, unpublished data).

In plants, progress has been made in determining the proteomes of organelles such as chloroplasts, mitochondria, and peroxisomes (Mo *et al.*, 2003). Here we describe the first proteomic analysis of plant nucleoli. We have carried out a proteomic analysis of the *Arabidopsis* nucleolus in order 1) to address the function of the nucleolus in plants; 2) to provide tools for identification of other nuclear bodies and to examine nuclear dynamics of proteins in the plant nucleus; and 3) to identify proteins that are common between or specific to plant and human nucleoli. We have identified 217 proteins from purified *Arabidopsis* nucleoli. Over 70 such proteins have been expressed as GFP fusions in *Arabidopsis* cells and 87% show labeling of, or associated with, the nucleolus, attesting to the quality of the preparation. The protein profile resembles that from human nucleoli, containing many expected nucleolar and ribosomal proteins, but in addition identifies conserved proteins of unknown function, plant-specific proteins and proteins with different localization between the two organisms. Intriguingly, we have identified a number of proteins known to be part of the postsplicing exon-junction complex (EJC), as nucleolar proteins.

MATERIALS AND METHODS

Purification of Nucleoli from Arabidopsis Culture Cells

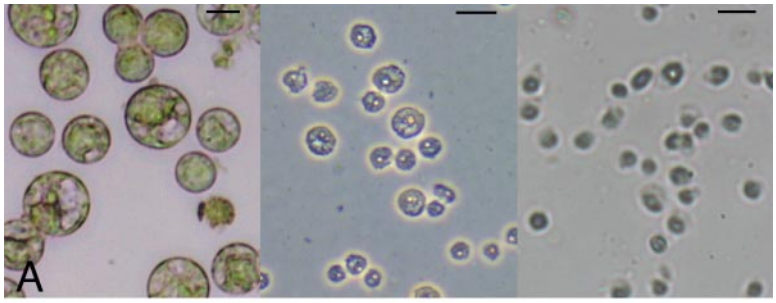
An *Arabidopsis* suspension culture derived from a Landsberg erecta line was used as a source of material for nucleolar purification. Cultures were grown at 25°C in the light with constant shaking on an orbital incubator at 200 rpm in AT medium (4.4% Murashige & Skoog salts, 3% sucrose, 0.05 mg/L kinetin, and 0.5 mg/L NAA, pH 5.8) and were harvested 3 d after subculturing by gentle centrifugation, and then protoplasts were generated essentially as described by Saxena *et al.* (1985). Cells were treated with cell wall degrading enzymes (2% cellulase R10; Onozuka, Tokyo, Japan, 0.025% pectolyase Y23, Seishin Corp, Tokyo, Japan, made up in 0.5 M sorbitol, 10 mM MES (2-*N*-morpholino-ethane-sulfonic acid)-KOH, pH 5.5, 1 mM $CaCl_2$) for ~3 h at room temperature (RT), until smooth, round protoplasts were obtained, as judged by light microscopy. The protoplasts were spun at 300 × *g* for 5 min and resuspended in 60% percoll buffer (60% percoll in 0.44 M sorbitol, 10 mM MES-KOH, pH 5.5, 1 mM $CaCl_2$), overlaid with a stepwise percoll gradient (45, 35, and 0% percoll in 0.44 M sorbitol, 10 mM MES-KOH, pH 5.5, 1 mM $CaCl_2$), and spun at 300 × *g* for 5 min. The intact protoplasts floated to the 0–35% interface and were collected by pipette. They were washed twice in 0.5 M sorbitol, 10 mM MES-KOH, pH 5.5, 1 mM $CaCl_2$ and quantified using a hemocytometer. Finally they were resuspended in Nuclear Isolation Buffer (NIB) at a concentration of no more than 10^6 protoplasts/ml (NIB: 10 mM MES-KOH, pH 5.5, 0.2 M sucrose, 2.5 mM EDTA, 2.5 mM dithiothreitol, 0.1 mM spermine, 0.1 mM spermidine, 10 mM NaCl, and 10 mM KCl) and were ruptured by 1–3 strokes in a stainless steel homogenizer. In some cases nuclei were produced first and then resuspended in NIB and further homogenized to produce nucleoli. The nucleoli were gently pelleted, then resuspended in 0.35 M sucrose, 0.5 mM $MgCl_2$, and frozen in aliquots at −80°C. Control of the $[Mg^{2+}]$ was important; if Mg^{2+} was added to the NIB, the nucleoli could not be separated from the network of nuclear chromatin fibers, whereas without Mg^{2+} in the buffer the nucleoli began to show signs of disintegration after 1–2 h. Therefore Mg^{2+} was added to the storage buffer, within 30 min of nuclear breakage. 1 liter of cell culture generated 4.8×10^8 nuclei and subsequently 3.2×10^8 nucleoli. All stages were monitored by phase contrast light microscopy.

Proteomic Analysis

Mass spectrometric analysis of isolated *Arabidopsis* nucleoli was done as previously described (Andersen *et al.*, 2002). Briefly, *Arabidopsis* nucleoli from 20 μl of the suspension (4×10^5 nucleoli) were solubilized in 500 μl of 8 M urea, 50 mM Tris-HCl, pH 8.0, and then reduced in 10 μg/ml DTT at RT for 1 h. Iodoactamide, 25 μg, and 1 μg of endoproteinase Lys-C (Roche) were added, and the mixture was incubated at RT for 3 h. The mixture was then diluted 5× with 25 mM Na_2CO_3 and incubated with 25 μg of trypsin (Promega, Madison, WI) at RT for an additional 8 h. The resulting peptide mixture was desalted and concentrated and then pressure-loaded onto a pulled fused silica capillary with an internal diameter of 100 μm and a tip opening of 8 μm (New Objective, Woburn, MA) filled with Vydac 218MSB3 3 μm reverse-phase material. Peptides were eluted directly into a quadrupole time-of flight mass spectrometer (MDS Sciex, Toronto, Ont., Canada) with a 90-min linear gradient of 5–60% buffer B (80% acetonitrile, 0.5% acetic acid, 0.02% heptafluorobutyric acid) from an HPLC system (Agilent, Palo Alto, CA). Combined peak lists were searched in the International Protein Index (IPI) database (http://www.ebi.ac.uk/IPI/IPIhelp.html) using the Mascot program (Matrix Science, Boston, MA).

Sequence Analysis

Arabidopsis protein sequences identified in the proteomic analysis were compared with proteins encoded in the human and yeast (*Saccharomyces cerevisiae*) genome and to the core set of 692 proteins identified in human nucleoli (Andersen *et al.*, 2002, Lamond and Mann, unpublished data) using BLAST (Altschul *et al.*, 1997). Alignments of all comparisons with an E value of more than $1.0e^{-20}$ were examined visually to determine their significance. To confirm the similarity of best matches between *Arabidopsis* and human proteins and to identify other related *Arabidopsis* genes, a reciprocal BLAST of the human protein with the highest score was carried out against the *Arabidopsis* protein database. In the majority of cases the human protein identified the original *Arabidopsis* protein or a related protein. The reciprocal BLAST analysis also identified homologous genes, providing information on multigene



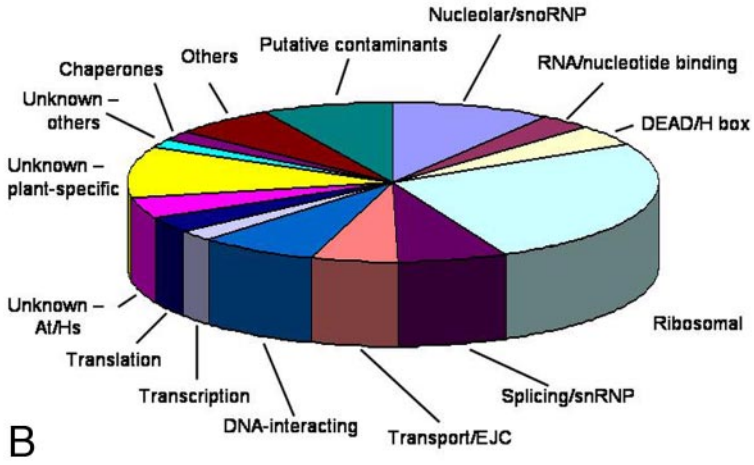


Figure 1. Preparation of *Arabidopsis* nucleoli and classification of proteins identified by mass spectrometry. (A) Nucleoli were purified from protoplasts and visualized using phase contrast light microscopy. Left: purified protoplasts (bar, 10 μm); center: purified nuclei (bar, 5 μm); right: purified nucleoli (bar, 2 μm). (B) Distribution of identified proteins among different protein classes

families. In a small number of cases, the human protein identified a different *Arabidopsis* protein than the original *Arabidopsis* protein. This suggests that either the homology between the proteins is low or limited to particular domains, and/or *Arabidopsis* contains proteins not identified in the proteomic analysis but which have stronger matches to human proteins. This may reflect different functions of the related proteins or that the limitations of the proteomic analysis have not yet identified all proteins in the plant nucleolus.

Construction of GFP Fusion Proteins of Nucleolar Components

The list of putative nucleolar proteins was used to search the SSP Orfeome database (http://signal.salk.edu/SSP/). Approximately half of the cDNAs were available as trimmed U clones (cDNAs trimmed to the annotated start and stop codons in the Cre-Lox pUNI vector). The SSP clones were transferred to a binary plant expression vector, designed to express N-terminal GFP fusions of the proteins (i.e., GFP::protein fusions) and expressed in *Arabidopsis* suspension culture cells as described by Koroleva *et al.* (2004). Briefly, the coding regions from the SSP pUNI U clones were amplified by high-fidelity PCR using standard primers for the adjacent vector sequence and then inserted by recombinase cloning into a standard Gateway entry vector (pDONr207; www.invitrogen.com). The entry vector clones were then transferred by recombinase cloning, en masse in 96-well format, into Gateway binary plant expression vector GFP-N-bin (a gift from Ben Trewaskis, Canberra, ACT, Australia), designed to express N-terminal GFP fusions of the proteins (i.e., GFP::protein fusions). The binary vector clones were then used to transform Agrobacterium. The transformed Agrobacterium lines were used in transient expression experiments with *Arabidopsis* suspensions as described by Koroleva *et al.* (2004).

Microscopy and Image Analysis

For microscopy, a small volume of suspension was placed on a slide and covered with a coverglass, which was held in place by nail polish. Initial surveys of GFP expression were carried out on a Nikon E600 epifluorescence microscope (Nikon UK, Kingston upon Thames, United Kingdom). Detailed imaging was carried out either using a Leica SP laser scanning confocal microscope (Leica Microsystems, Heidelberg, Germany), or a wide-field CCD camera microscope (Orca ER, CCD camera, Prior filter wheels and X,Y,Z translation stage, controlled by Metamorph operating software; Universal Imaging, West Chester, PA). CCD images were deconvolved using a maximum likelihood blind restoration algorithm (Autodeblur, AutoQuant, Imaging, Troy, NY). The use of a high-sensitivity CCD camera enabled us to image faintly labeled cells, thus minimizing the chance of mis-localization due to overexpression. For each cell a single section through the middle of the nucleolus was selected from the 3D data stacks. Image montages were prepared using ImageJ, a public domain image processing program written by Wayne Rasband (available from http://rsb.info.nih.gov/ij/), and Adobe Photoshop (San Jose, CA).

RESULTS

Purification and Mass Spectrometric Analysis of Arabidopsis Nucleoli

Nucleoli were purified from nuclei prepared from *Arabidopsis* cell culture protoplasts. The nucleolar preparation was characterized by phase contrast optical microscopy and also by fluorescence microscopy after staining with 7 μM pyronin Y (a nucleolus-specific fluorescent stain) and showed small discrete nucleoli, essentially free of contaminating nuclear debris and chromatin (Figure 1A). The nucleoli were also characterized by immunofluorescence labeling with an antibody to the nucleolar protein, SSB1 (unpublished data). Tandem mass spectrometric analysis of complex mixtures of peptides, generated by tryptic digestion of aliquots of the nucleolar preparation, identified 217 different proteins (Figure 1B; Supplementary Table 1).

Localization of GFP Fusion Proteins of a Subset of Identified Proteins

The inserts from 76 cDNA clones were transferred into a plant Gateway-compatible binary vector, designed to express N-terminal GFP-protein fusions from the Cauliflower Mosaic Virus (CaMV) 35S promoter (Koroleva *et al.*, 2005). For localization we used transient expression in *Arabidopsis* cell cultures using Agrobacterium infection (Koroleva *et al.*, 2005), which gave strong, reproducible expression. GFP fluorescence was successfully detected in 72 clones and expres-

Table 1. Localization of expressed GFP fusion proteins in transient assay

	No. of fusions[a]
GFP fusion location	
Nucleolar/nucleolus-associated structures	42 (58)
Weak nucleolar labelling	21 (29)
Nuclear excluded from nucleolus	2 (3)
Nuclear body only	1 (1)
Cytoplasmic	6 (8)
Total GFP fusions surveyed	72
No detectable expression	4
Nuclear bodies	20 (28)

[a] Values in parentheses are percentages.

sion patterns were imaged using either confocal microscopy or wide-field CCD imaging followed by image deconvolution. Although it is possible that some of the localizations observed may be affected by the addition of GFP or by overexpression in the transient system, the range of different labeling patterns observed, the expected localization of proteins of known function (e.g., ribosomal and snoRNP proteins), and the differential localization of similar proteins (e.g., the ALY family) argue against the observed patterns being artifacts. In addition, use of the CCD camera allowed the imaging of cells which fluoresced weakly to minimize the possibility of artifacts due to overexpression. For each fusion, a range of cells with relatively low levels of fluorescence were examined. The localization of each construct was annotated with a controlled vocabulary (nucleolus, nucleolar-associated structure, nuclear bodies, nucleoplasm and extranuclear). This annotation is included in Supplementary Table 1.

Overall statistics of the localization are shown in Table 1, and a selection of different GFP fusions is shown in Figure 2. Different localization patterns were observed with various combinations of labeling of the nucleolus, nucleolar-associated structures, nuclear bodies, nucleoplasm, or extranuclear labeling. In some cases, the fusion proteins were largely restricted to the nucleolus (e.g., Figure 2, A and L), whereas in others, the nucleoplasm (e.g., Figure 2, B, G, H, and I) and/or other nuclear structures or regions were also labeled. Many fusion proteins (28%) labeled small nuclear bodies in addition to the nucleolus, some of which were reminiscent of Cajal bodies or speckles, (e.g., Figure 2, D–F, H, R, and S). Some fusions showed clear labeling of nucleolar-associated structures (Figure 2, P and Q) or subdomains (Figure 2, C and L). From the 72 GFP fusions surveyed, 42 (58%) showed strong labeling of the nucleolus and 21 (29%) showed weak labeling (e.g., Figure 2, G and I) such that 87% showed some degree of labeling of the nucleolus and thus would be expected to be present in a purified nucleolar fraction. On the other hand, two GFP-fusions labeled the nucleoplasm and appeared to be excluded from the nucleolus (unpublished data). Only 6 (9%) showed only extranuclear/cytoplasmic labeling and of the 217 proteins, and only 18 (8%) are putative cytoplasmic, chloroplast, or mitochondrial proteins. Because of the greater sensitivity of detection by mass spectrometry, the results from GFP fluorescence may be an underestimate of the number of proteins in the nucleolus. Lack of accumulation in the nucleolus of some GFP fusions may be due to the protein being either present in low abundance or only transiently associated with the nucleolus at particular cell stages or under certain conditions (Fox *et al.*, 2002; Leung *et al.*, 2003). Overall, these results provide the most practical assessment and verification that the majority of identified proteins are bona fide nucleolar components.

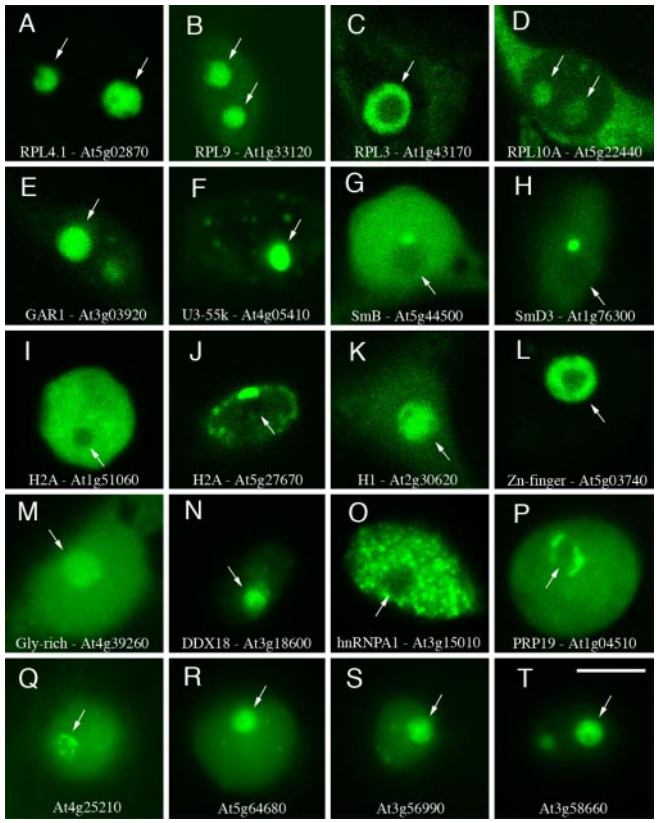


Figure 2. GFP localizations of a selection of proteins identified in the *Arabidopsis* nucleolar proteomic analysis. Confocal GFP images of single cells are shown and nucleoli indicated by arrows. (A–D) Ribosomal proteins: RPL4.1 (At5g02870), RPL9 (At1g33120), RPL3 (At1g43170), and RPL10A (At5g22440). Note that some ribosomal protein fusions label the cytoplasm strongly, others weakly; (E–H snoRNP and snRNP proteins: GAR1 (At3g03920), U3–55k (At4g05410), SmB (At5g44500), and SmD3 (At1g76300); (I–L) DNA-interacting proteins: H2A (At1g51060), H2A (At5g27670), H1 (At2g30620), and Zinc-finger protein (At5g03740). Note that I and J are both H2A histones, but the homologue in I is widely present throughout the chromatin, whereas that in J is apparently concentrated in heterochromatin; (M–P) RNA-interacting proteins: glycine-rich RNA-binding protein (At4g39260), DDX18 (At3g18600), hnRNPA1 (At3g15010), and PRP19 (At1g04510); (Q–T) proteins with unknown function: plant-specific proteins (At4g25210 and At5g64680), and conserved proteins in both plant and human proteomes (At3g56990 and At3g58660). Bar, 5 μm.

Comparison of the Arabidopsis and Human Nucleolar Proteomes

Of the *Arabidopsis* nucleolar set of 217 proteins, 186 (86%) and 145 (67%) had homologous proteins encoded in the human and yeast genomes, respectively (Supplementary Table 1). We also compared the *Arabidopsis* proteins to the core set of 692 proteins identified in the human nucleolus (Lamond and Mann, unpublished data), which extends the initial, published human nucleolar proteome of 271 proteins (Andersen *et al.*, 2002). Just over two-thirds (69%) of the *Arabidopsis* proteins had a direct counterpart in the human

nucleolar proteome. The profile of the *Arabidopsis* nucleolar proteome resembled that of human in containing many expected nucleolar proteins (e.g., nucleolin, B23, NOL1, NOLC1, Nopp34, etc.), ribosome biogenesis proteins (e.g., RRS1 and RRP5), and ribosomal proteins. In addition, there were many components of transcription, splicing, and translation, DEAD box proteins and RNA/DNA interacting/modifying proteins, nucleotide-binding proteins, and chaperones. Thirty-seven proteins were of unknown function (Figure 1B and Supplementary Table 1).

All eight core proteins of the box C/D and box H/ACA snoRNPs (fibrillarin, NOP5/NOP58, NOP56, p15.5 (Snu13p homologue), and GAR1, DKC1 (Cbf5p/NAP57), NOP10 and NHP2) were present, along with U3snoRNP-specific proteins (U3–55k and Mpp10), and nonsnoRNP proteins involved in snoRNP assembly or nucleolar localization (TIP48, TIP49, and Nopp140/NOLC1; Pluk *et al.*, 1998; Westendorf *et al.*, 1998; Watkins *et al.*, 2000, 2002; Wang and Meier, 2004). *Arabidopsis* contains three fibrillarin genes but only two, AtFib1 and AtFib2, are expressed (Barneche *et al.*, 2000), and both are detected in the proteomic analysis. In addition, among a range of RNA-interacting proteins were nine DEAD/H box putative RNA helicase proteins, a class of proteins with important functions in RNA metabolism (Rocak and Linder, 2004). These included the nucleolar RNA helicase II (DDX21), two splicing helicases, including UAP56, and eIF-4AIII, recently identified as an exon junction complex protein (see below). In addition, five other DEAD/H box proteins identified (DDX3, DDX9, DDX10, DDX15, and DDX18) were also found in the human proteome.

Nearly 60 different ribosomal proteins were identified making up 27% of the proteins detected in the *Arabidopsis* proteomic analysis. The multigenic nature of many plant and *Arabidopsis* genes is clearly demonstrated by ribosomal protein genes where, to date, 249 genes have been identified in the *Arabidopsis* genome (Barakat *et al.*, 2001). The 249 genes represent variants of 80 different ribosomal protein genes, organized into gene families with two to seven members (Barakat *et al.*, 2001), and contrasts greatly with human and yeast where most ribosomal proteins are encoded by a single gene. Where gene families encode identical or very similar proteins (e.g., RPS18, RPL29, RPL36a, RPL38, and RPL40), the different isoforms may not be distinguished by the MS analysis. In contrast, for other gene families which encode proteins with more variable sequences (e.g., L7Ae, L22, and L23), different isoforms were identified.

The direct comparison between the protein profiles of *Arabidopsis* and human nucleoli has given rise to a number of novel observations. First, 23 proteins of unknown function were plant specific, possibly reflecting the differences in structure of plant and animal nucleoli (Shaw and Jordan, 1995), differences in rRNA gene organization and transcription and the presence of a nucleolar cavity. For example, five of the plant-specific proteins are potential DNA-interacting proteins, which may reflect differences in plant rRNA transcription (Gonzalez-Melendi *et al.*, 2001). Second, 10 proteins of unknown function were conserved proteins found in both plant and animal nucleoli. One such protein (DKFZP564M182–At3g58660) was also present in a proteomic analysis of human Nop56p-containing pre-rRNP complexes (Hayano *et al.*, 2003), demonstrating its close association with ribosome biogenesis. Similarly, the WD-repeat–containing protein of unknown function encoded by At3g21540 is a homologue of Utp12p found in the yeast SSU processome, along with U3snoRNP proteins, Rrp5p and Rrp9p (Dragon *et al.*, 2002). Some of the other human/plant conserved proteins contained recognized domains involved in RNA-protein and protein-protein interactions: WD40 repeat, GTPase- and RNA-binding domains. Of 12 plant-specific or plant-human conserved proteins for which GFP fusions were constructed, 10 labeled the nucleolus strongly, with some also labeling nuclear bodies (e.g., Figure 2, Q–T). Third, many known proteins, not expected to be nucleolar proteins, were either found in both proteomes or strongly associated only with the plant nucleolus. For example, our results show the presence in the nucleolus of many proteins involved in pre-mRNA splicing and translation. This confirms similar unexpected findings in the human nucleolar proteome (Andersen *et al.*, 2002; Scherl *et al.*, 2002). In *Arabidopsis*, the majority of splicing proteins detected were snRNP core proteins: SmD1, SmD2, SmD3, SmF, and SmG, and snRNP-specific proteins: U2A′, and the U5 116k and 200k proteins. This observation is consistent with previous evidence for the presence of U1, U2, and U6 snRNAs in the nucleolar cavity in plant nucleoli (Beven *et al.*, 1995). In addition, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), a glycolytic enzyme, was found in both the *Arabidopsis* and human nucleolar proteomes. The *Arabidopsis* GAPDH:GFP fusion protein showed clear labeling of the nucleoplasm and fainter labeling of the nucleolus.

Thirty-nine (18%) of the *Arabidopsis* proteins have human homologues that, however, are not present in the more extensive human nucleolar proteome. This number is too great to represent contamination of the plant nucleolar preparation and indeed only five of these proteins are probable contaminants, being organellar, or associated with ER or peroxisomes. The remaining proteins potentially reflect either divergent functions or different localization between organisms and include, for example, exon junction complex (EJC) proteins (see below) and *S*-adenosylmethionine (SAM) synthetase.

Identification of Exon Junction Complex Proteins in the Nucleolus

A surprising finding was the presence in the nucleolar proteome of six proteins known from animal studies to be part of the exon junction complex (EJC): ALY/REF, Mago, Y14, UAP56, RNPS1, and the eukaryotic translation initiation factor eIF4A-III (eIF4A-III; Dreyfuss *et al.*, 2002; Chan *et al.*, 2004; Maquat, 2004; Palacios *et al.*, 2004). The EJC is a multiprotein complex that marks splice junctions in mRNAs after mRNA splicing, linking transcription and processing to mRNA export and NMD (Dreyfuss *et al.*, 2002; Maquat, 2004). The EJC contains spliceosome-associated proteins (ALY, UAP56, RNPS1, SRm160, and DEK); Y14 and Mago, which interact with mRNP export adaptor proteins (p15 and TAP); and proteins involved in nonsense-mediated decay (Upf2 and Upf3; Zhou *et al.*, 2000; Le Hir *et al.*, 2001a; Lykke-Andersen *et al.*, 2001; Dreyfuss *et al.*, 2002; Reed and Hurt, 2002; Stutz and Izzuralde, 2003; Maquat, 2004). The *Arabidopsis* proteins AtUAP56–2, AtY14, AtRNPS1, AtALY-4, and AteIF4A-III all showed clear localization to the nucleolus as well as labeling of the nucleoplasm (Figure 3). AtY14 labels the nucleolus much more strongly than the nucleoplasm and, in particular, the peripheral nucleolar region. In addition it is located in foci in the nucleolar cavity (Beven *et al.*, 1995; Brown and Shaw, 1998; Figure 3A). AtY14 is also concentrated in numerous small nucleoplasmic bodies compared with the faint, diffuse labeling of the nucleoplasm. In contrast, AtMago shows strong nucleoplasmic labeling with a number of small bodies and fainter punctuate nucleolar labeling (Figure 3B). AtRNPS1 shows intense foci in the

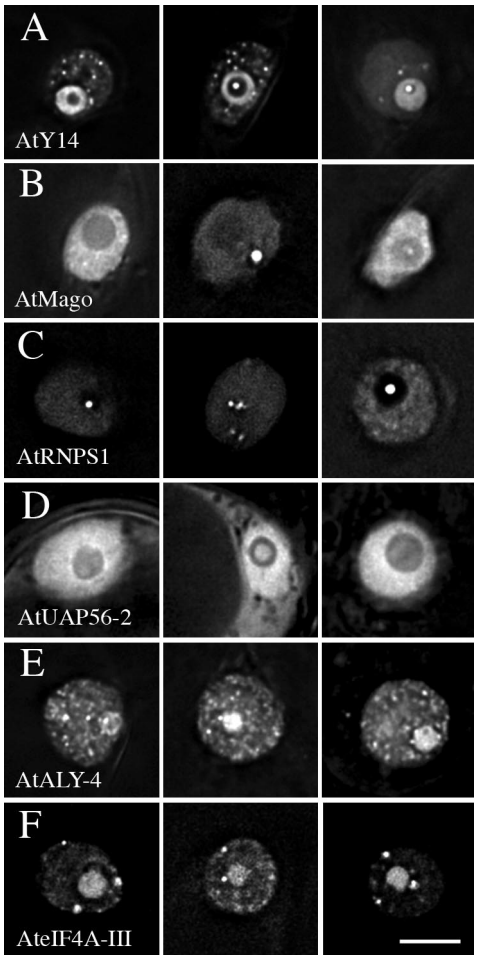


Figure 3. GFP localizations of EJC proteins identified in the *Arabidopsis* nucleolar proteomic analysis. Images of three different cells are shown for each of the Y14, Mago, RNPS1, UAP56–2, ALY-4, and eIF4A-III GFP fusions. A single *z* plane passing through the center of a nucleolus from a deconvoluted CCD image is shown in each case. Bar, 5 μm.

nucleolar cavity with diffuse, nonuniform nucleoplasmic labeling (Figure 3C). AtUAP56–2 also shows accumulation in the nucleolar cavity but is more diffuse than AtY14 or AtRNPS1 and also shows strong nucleoplasmic labeling. AtALY-4 labels the nucleolus strongly and shows a more discrete nuclear speckled labeling (Figure 3E). Finally, AteIF4A-III labels the nucleolus and nuclear bodies strongly with fainter nucleoplasmic labeling (Figure 3F). Thus, all six EJC proteins found in the nucleolar proteomic preparation showed localization in the nucleolus, suggesting that the nucleolus may be a site of storage, assembly, or function of the EJC components or complexes or involved in other RNA metabolism processes.

We have also identified *Arabidopsis* homologues of all known EJC components (Table 2). AtALY, AtUAP56, AtDEK, and Atp15 comprise small multigene families (Table 2), which is a feature of many plant genes and may reflect the prevalence of hybridization and polyploidy in plant evolution (Wendel, 2000). The remaining EJC components had a single homologue, apart from TAP, for which no plant homologue has yet been identified. To investigate whether there are potential functional differences between homologues, full-length cDNAs and GFP-fusions were prepared

Table 2. Putative *Arabidopsis* orthologues of EJC components

Protein	Locus	BLAST e-value[a]
AtY14	At1g51510	$2.2e^{-33}$
AtMago Nashi	At1g02140	$4.9e^{-61}$
AtUAP56–1	At5g11200	$1.0e^{-157}$
AtUAP56–2	At5g11170	$1.0e^{-156}$
AtALY-1	At5g59950.1	$8.1e^{-27}$
AtALY-1	At5g59950.2(splicevariant)	$1.0e^{-26}$
AtALY-2	At5g02530	$3.4e^{-26}$
AtALY-3	At5g66260	$2.5e^{-22}$
AtALY-4	At1g37720	$1.3e^{-21}$
AtRNPS1	At1g16610	$6.5e^{-13}$
AtSRm160	At2g29210	$8.1e^{-40}$
AteIF-4AIII	At3g19760	$1.0e^{-174}$
AtDEK-1	At3g48710	$1.2e^{-20}$
AtDEK-2	At5g63550	$8.6e^{v15}$
AtDEK-3	At4g26630	$4.4e^{-10}$
AtDEK-4	At5g55660	$1.3e^{-08}$
AtP15–1	At1g11570	$1.4e^{-08}$
AtP15–2	At1g27970	$1.8e^{-08}$
AtP15–3	At1g27310	$5.5e^{-07}$
AtUPF2	At2g39260	$4.6e^{-148}$
AtUPF3	At1g33980	$6.1e^{-19}$
TAP	No hits	No hits

[a] BLAST against human genome.

for the three remaining ALY genes (At*ALY-1*, At*ALY-2*, and At*ALY-3*). The labeling pattern of each of the four AtALY proteins was distinct, showing differential nuclear and subnuclear localizations. AtALY-2 labeled the nucleoplasm strongly in a speckled network including bright foci, with only very faint labeling of the nucleolus (Figure 4B). Like AtALY-4, both AtALY-1 and AtALY-3 labeled the nucleolus

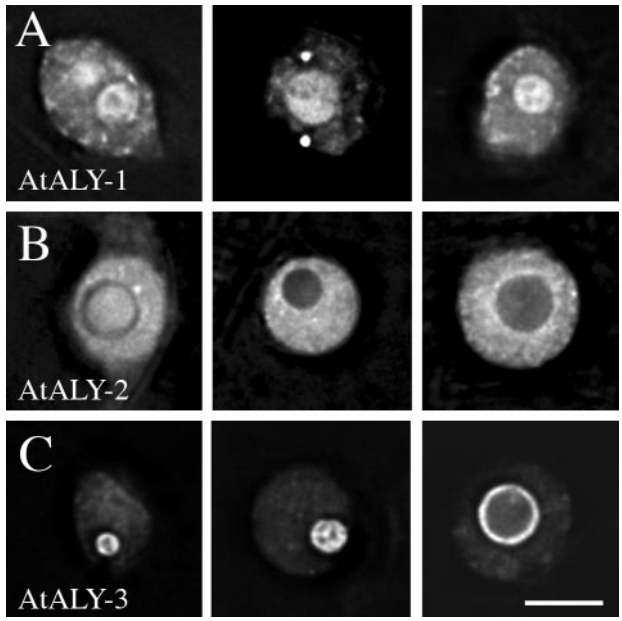


Figure 4. GFP localizations of the three ALY homologues not detected in the *Arabidopsis* nucleolar proteomic analysis. Images of three different cells are shown for ALY-1, ALY-2, and ALY-3 GFP fusions. A single *z* plane passing through the center of a nucleolus from a deconvoluted CCD image is shown in each case. Bar, 5 μm.

strongly (Figure 4, A and C) but nucleoplasmic labeling with AtALY-3 was less intense than with the other ALY homologues. AtALY-1 showed a small number of intensely staining bodies, in contrast to AtALY-4, which had a speckled pattern. AtALY-3 consistently labeled a peripheral region of the nucleolus.

DISCUSSION

On the basis of the high incidence of known or expected nucleolar proteins and the low occurrence of, for example, cytoplasmic or organellar proteins, the *Arabidopsis* nucleolar preparation used in the analysis is of good purity and quality. Furthermore, the high percentage (87%) of the 72 protein:GFP fusions that showed labeling of nucleoli when expressed in *Arabidopsis* cells, provides complementary evidence for the overall quality of the original nucleolar preparation. The high throughput method for cloning of GFP fusions makes such verification by localization a feasible adjunct to proteomic analyses. Compared with the most recent proteomic analysis of human nucleoli (Lamond and Mann, unpublished data), it is likely that this proteomic analysis represents only an initial estimate of plant nucleolar protein complexity. For example, only three-quarters of the *Arabidopsis* ribosomal proteins were detected in this analysis, and only the largest subunit of RNA polymerase I was detected. The most recent human proteomic analysis was carried out using more sensitive mass spectrometric techniques than were available when this proteomic study was performed. We anticipate that we will identify a similar number of plant nucleolar proteins when the most sensitive current methods are applied.

The analyses of the proteome of the human nucleolus have demonstrated the complexity of this nuclear body in terms of the number and nature of protein factors (692 proteins) and further underpins the multitude of functions in which the nucleolus may have a role (Pedersen, 1998; Andersen *et al.*, 2002; Scherl *et al.*, 2002; Leung *et al.*, 2003). Furthermore, the level and localization of subsets of proteins can vary under different conditions (e.g., inhibition of transcription) and in different cell types or cell cycle stages reflecting transient association with the nucleolus or dynamic interactions of components of the nucleolus, with other nuclear bodies and the nucleoplasm (Andersen *et al.*, 2002). Little is known about nuclear architecture in plant cells, and to date, the best characterized nuclear bodies or domains are the nucleolus, Cajal bodies, and SR protein-containing speckles (Beven *et al.*, 1995; Shaw and Jordan, 1995; Lorkovic *et al.*, 2004a). However, other plant proteins involved in RNA metabolism or signaling have been recently shown to accumulate in nuclear speckle-like bodies (Kircher *et al.*, 2002; Chen *et al.*, 2003; Han *et al.*, 2004; Lorkovic *et al.*, 2004b; Shaw and Brown, 2004), further demonstrating the expected complexity of plant subnuclear organization. Many of the nucleolar protein fusions analyzed here also labeled nuclear bodies. Although some appear similar to Cajal body-like structures or SR-protein containing speckles, it is possible that they label other novel nuclear domains. Confirmation of the nature of these domains will require colocalization analyses with marker proteins for the known bodies and may lead to the identification of other plant nuclear bodies.

The value of comparative proteomic analyses between widely different species is demonstrated here by the identification of known proteins, which were not expected to be nucleolar but which were either found in both proteomes or associated only with the plant nucleolus, and proteins of unknown function that were either plant-specific or conserved in both plant and animal nucleoli. For example, the presence of many proteins involved in pre-mRNA splicing and translation in the *Arabidopsis* nucleolus confirms the findings of the human nucleolar proteome analyses (Andersen *et al.*, 2002; Scherl *et al.*, 2002). Although it is not known whether these proteins are associated in RNPs or with mRNAs, the presence of the many snRNP proteins in the *Arabidopsis* nucleolar proteome, taken with evidence of snRNAs being present in the nucleolar cavity (Beven *et al.*, 1995), suggests that mature snRNPs exist in the nucleolus.

GAPDH was also detected in both the *Arabidopsis* and human nucleolar proteomes. Recent evidence has shown that besides its glycolytic activity, GAPDH is a multifunctional protein with both cytoplasmic and nuclear functions, one of which is as an essential component of a transcriptional activator complex regulating mammalian histone H2B expression (Mazzola and Sirover, 2002; Zheng *et al.*, 2003). The dual functionality of cytoplasmic enzymes in the activation or repression of transcription or in modulating splicing or stability of mRNAs illustrates potential links between gene regulation and the metabolic activity of the cell (Macknight, 2003). The finding of GAPDH in the nucleolus of both plants and mammals may reflect one of the functions of the nucleolus in sequestration of regulatory components (Pedersen, 1998; Cockell and Gasser, 1999; Olsen *et al.*, 2000; Visintin and Amon, 2000).

We showed the presence of two different SAM synthetase proteins in the plant nucleolar preparation, whereas this protein was not detected in the human nucleolar proteome. This enzyme catalyzes the conversion of L-methionine and ATP to *S*-adenosyl-L-methionine, which is a methyl group donor. Even considering the extensive methylation of rRNAs by fibrillarin in the nucleolus, the presence of SAM synthetase is unexpected. However, a recent proteomic analysis of the nuclear matrix of *Arabidopsis* (Calikowski *et al.*, 2003) identified 36 proteins of which many were nucleolar (expected because of the integral association of the nucleolar matrix with such preparations) and two were SAM synthetases, one of which (At3g17390) was also present in the nucleolar proteome. *Arabidopsis* contains five SAM synthetase genes and our results suggest that at least two SAM synthetases are found in the nucleolus, perhaps for rapid production of SAM for rRNA methylation.

Finally, the *Arabidopsis* nucleolar proteomic analysis has identified an association of six EJC components with the nucleolus, confirmed by GFP localization. In contrast, human and *Drosophila* EJC proteins localize to the nucleoplasm and nuclear speckles but are excluded from the nucleolus (Kataoka *et al.*, 2001; Le Hir *et al.*, 2001b; Custódio *et al.*, 2004; Palacios *et al.*, 2004). In the original published human nucleolar proteome of 271 proteins, the only EJC component to be identified was ALY/Ref (Andersen *et al.*, 2002). In light of the clear association of plant EJC proteins with the nucleolus, we searched the most up-to-date human nucleolar proteome of 692 proteins and found six EJC components (ALY/Ref, UAP56, Y14, DEK, eIF4A-III, and UPF3X; Lamond and Mann, unpublished data). This suggests that at least some EJC components are also associated with the nucleolus in animal cells even though the animal proteins may be present in much lower concentrations as evidenced by fluorescence microscopy.

The exon junction complex couples transcription, RNA processing and, in particular, splicing, to export and other downstream processes such as nonsense-mediated decay (NMD) or mRNA surveillance (Dreyfuss *et al.*, 2002; Maquat, 2004). After splicing, the exon junction complex is deposited

around 20–24 nt upstream of the exon:exon junction. The complex contains proteins involved in spliceosome assembly and export adapter proteins, which interact with export receptors to contact the nuclear pore complex and effect mRNP export (Le Hir *et al.*, 2001a; Dreyfuss *et al.*, 2002; Maquat, 2004). The EJC also recruits Upf proteins involved in NMD, which degrade mRNAs containing premature stop codons, produced by mutation or by aberrant expression and RNA processing. The selective degradation of such transcripts avoids the production of truncated proteins that could have dominant negative effects and be deleterious to the cell.

NMD is a translation-dependent process and the majority of NMD is nuclear, suggesting a pioneer round of translation of mRNAs bound by the cap-binding complex at the 5′ cap (Ishigaki *et al.*, 2001; Lejeune *et al.*, 2002). Although the site of nuclear NMD is unknown, current models propose translation by cytoplasmic ribosomes as mRNAs pass through the nuclear pore complex or nuclear translation (Dreyfuss *et al.*, 2002; Wilkinson and Shyu, 2002; Maquat, 2004). The human EJC proteins ALY, UAP56, SRm160, RNPS1, Y14, and Magoh have recently been shown in mammalian cells to accumulate at sites of transcription/splicing and they therefore bind pre-mRNAs/mRNAs cotranscriptionally (Custódio *et al.*, 2004). In contrast, the mRNA export adaptor proteins p15 and TAP are not recruited to these sites and must interact later in the export pathway. Similarly, the presence of nonsense codons does not affect splicing of pre-mRNAs suggesting that nuclear mRNA surveillance is unlikely to occur cotranscriptionally (Lytle and Steitz, 2004).

If the detection of ALY, UAP56, RNPS1, Y14, and Mago in the plant nucleolar preparation reflects the presence of mRNAs bound by EJC complexes, then such mRNPs may be targeted to the nucleolus after release from transcription sites. Therefore, our results suggest that the plant nucleolus functions in storage or assembly of EJC subcomplexes, mRNA export or in NMD. Both human and plant nucleolar proteomes contained a variety of translation initiation and elongation factors (including the EJC component, eIF4A-III, and EF-Tu, a mitochondrial translation elongation factor; Andersen *et al.*, 2002; Scherl *et al.*, 2002; this article; Lamond and Mann, unpublished data). Translation factors were also identified in plant nuclear matrix preparations (Calikowski *et al.*, 2003). Although the nuclear/nucleolar functions of these proteins are unknown, it is possible that they have multiple and novel functions perhaps related to mRNA surveillance.

A major difference between some of the EJC components of plants and animals is the multigene nature of *ALY*, *DEK*, *p15*, and *UAP56* genes. Organization of genes into multigene families reflects the prevalence of hybridization and polyploidy, which has occurred in plant evolution (up to 70% of plant species may have undergone such genome amplification events; Wendel, 2000). Duplicated genes may diverge and either generate complementary expression patterns or develop differential functions, ensuring continued selection of both genes. Thus it is possible that some plant EJC homologues are functionally distinct, forming different complexes by specific association with different members of the EJC or other proteins. They may therefore bind or associate with different subsets of mRNAs or in some way identify and target mRNAs for transport to and from the nucleolus, export from the nucleus or NMD. The different localization of all four ALY homologues may reflect such differential activities. Detailed cell biological and biochemical analyses of all *Arabidopsis* EJC proteins, complexes, and associated mRNAs is required to address these major questions in posttranscriptional gene regulation. Nevertheless, the importance of the comparative proteomics approach between *Arabidopsis* and human nucleoli is evident from the identification of species-specific proteins, proteins with differential localization, and potentially novel functions for the eukaryotic nucleolus.

Note added in proof: The Arabidopsis nucleolus proteome information is available on a database at: http://bioinf.scri.sari.ac.uk/cgi-bin/atnopdb/home

ACKNOWLEDGMENTS

We thank Olga Koroleva and Matt Tomlinson, JIC, for help with high-throughput GFP expression, and Ben Trewaskis, MPI, Germany for the gift of the GFP-N-bin vector. We thank Dave Marshall and Paul Shaw of the Computational Biology Programme, Scottish Crop Research Institute, for bioinformatics advice and help. This research was supported by funding from the Scottish Executive Environment and Rural Affairs Department (SEERAD) to SCRI and the Biotechnology and Biological Sciences Research Council of the United Kingdom (BBSRC) to the John Innes Centre. D.L. was funded by a European Union Marie Curie fellowship.

REFERENCES

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Andersen, J. S., Lyon, C. E., Fox, A. H., Leung, A.K.L., Lam, Y. W., Steen, H., Mann, M., and Lamond, A. I. (2002). Directed proteomic analysis of the human nucleolus. Curr. Biol. *12*, 1–11.

Barakat, A., Szick-Miranda, K., Chang, I.-F., Guyot, R., Blanc, G., Cooke, R., Delseny, M., and Bailey-Serres, J. (2001). The organization of cytoplasmic ribosomal protein genes in the *Arabidopsis* genome. Plant Physiol. *127*, 398–415.

Barneche, F., Steinmetz, F., and Echeverria, M. (2000). Fibrillarin genes encode both a conserved nucleolar protein and a novel small nucleolar RNA involved in ribosomal RNA methylation in *Arabidopsis thaliana.* J. Biol. Chem. *275*, 27212–27220.

Beven, A. F., Simpson, G. G., Brown, J.W.S., and Shaw, P. J. (1995). The organization of spliceosomal components in the nuclei of higher plants. J. Cell Sci. *108*, 509–518.

Beven, A. F., Lee, R., Razaz, M., Leader, D. J., Brown, J.W.S., and Shaw, P. J. (1996). The organization of ribosomal RNA processing correlates with the distribution of nucleolar snRNAs. J. Cell Sci. *109*, 1241–1251.

Brown, J.W.S., and Shaw, P. J. (1998). Small nucleolar RNAs and pre-rRNA processing in plants. Plant Cell *10*, 649–657.

Calikowski, T. T., Meulia, T., and Meier, I. (2003). A proteomic study of the *Arabidopsis* nuclear matrix. J. Cell. Biochem. *90*, 361–378.

Carmo-Fonseca, M., Mendez-Soares, L., and Campos, I. (2000). To be or not to be in the nucleolus. Nat. Cell Biol. *2*, E107–E112.

Chan, C. C., Dostie, J., Diem, M. D., Feng, W., Mann, M., Rappsilber, J., and Dreyfuss, G. (2004). eIF4A3 is a novel component of the exon junction complex. RNA *10*, 200–209.

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

Chen, M., Schwab, R., and Chory, J. (2003). Characterization of the requirements for localization of phytochrome B to nuclear bodies. Proc. Natl. Acad. Sci. USA *100*, 14493–14498.

Cockell, M. M., and Gasser, S. M. (1999). Nucleolar space for RENT. Curr. Biol. *9*, R575–R576.

Custódio, N., Carvalho, C., Condado, I., Antoniou, M., Blencowe, B. J., and Carmo-Fonseca, M. (2004). In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei. RNA *10*, 622–633.

Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. EMBO J. *21*, 2746–2756.

Dragon, F. *et al.* (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature *417*, 967–970.

Dreyfuss, G., Kim, V. N., and Kataoka, N. (2002). messenger-RNA-binding proteins and the messages they carry. Nat. Rev. Mol. Cell. Biol. *3*, 195–205.

Fatica, A., and Tollervey, D. (2002). Making ribosomes. Curr. Opin. Cell Biol. *14*, 313–318.

Filipowicz, W., and Pogacic, V. (2002). Biogenesis of small nucleolar ribonucleoproteins. Curr. Opin. Cell Biol. *14*, 319–327.

Fox, A. H., Lam, Y. W., Leung, A.K.L., Lyon, C. E., Andersen, J., Mann, M., and Lamond, A. I. (2002). Paraspeckles: a novel nuclear domain. Curr. Biol. *12*, 13–25.

Gonzalez-Melendi, P., Wells, B., Beven, A. F., and Shaw, P. J. (2001). Single ribosomal transcription units are linear, compacted Christmas trees in plant nucleoli. Plant J. *27*, 223–233.

Gunning, B.E.S. Plant Cell Biology on CD—information for students and a resource for teachers, Part 1. World Wide Web URL: www.plantcellbiologyonCD.com. Accessed November 18, 2004.

Han, M. H., Goud, S., Song, L., and Fedoroff, N. (2004). The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. Proc. Natl. Acad. Sci. USA *101*, 1093–1098.

Hayano, T., Yanagida, M., Yamauchi, Y., Shinkawa, T., Isobe, T., and Takahashi, N. (2003). Proteomic analysis of human Nop56p-associated pre-ribosomal ribonucleoprotein complexes. J. Biol. Chem. *278*, 34309–34319.

Horn, H. F., and Vousden, K. H. (2004). Guarding the guardian. Nature *427*, 110–111.

Ishigaki, Y., Li., X. Serin, G., and Maquat, L. E. (2001). Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. Cell *106*, 607–617.

Jady, B. E., Darzacq, X., Tucker, K. E., Matera, A. G., Bertrand, E., and Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. EMBO J. *22*, 1878–1888.

Jady, B., and Kiss, T. (2001). A small nucleolar guide RNA functions in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J. *20*, 541–551.

Kataoka, N., Diem, M. D., Kim, V. N., Yong, J., and Dreyfuss, G. (2001). Magoh, a human homologue of *Drosophila* mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. EMBO J. *20*, 6424–6433.

Kircher, S., Gil, P., Kozma-Bognar, L., Fejes, E., Speth, V., Husselstein-Muller, T., Bauer, D., Adam, E., Schafer, E. and Nagy, F. (2002). Nucleocytoplasmic partitioning of the plant photoreceptors phytochrome A, B, C, D, and E is regulated differentially by light and exhibits a diurnal rhythm. Plant Cell, *14*, 1541–1555.

Kiss, T. (2002). Small nucleolar RNAs: an abundant group of non-coding RNAs with diverse cellular functions. Cell *109*, 145–148.

Koberna, K., Malinsky, J., Pliss, A., Masata, M., Vecerova, J., Fialova, M., Bednar, J., and Raska, I. (2002). Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of "Christmas trees" in situ. J. Cell Biol. *157*, 743–748.

Koroleva, O., Tomlinson, M., Leader, D., Shaw, P., and Doonan, J. (2005). High-throughput protein localization in *Arabidopsis* using Agrobacterium-mediated transient expression of GFP-ORF fusions. Plant J. (*in press*).

Lafontaine, D.L.J., and Tollervey, D. (2001). The function and synthesis of ribosomes. Nat. Rev. Mol. Cell. Biol. *2*, 514–520.

Lamond, A. I., and Earnshaw, W. C. (1998). Structure and function in the nucleolus. Science *291*, 843–847.

Lamond, A. I. , and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. Nat. Rev. Mol. Cell. Biol. *4*, 605–612.

Le Hir, H., Gatfield, D., Izauralde, E., and Moore, M. J. (2001a). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J. *20*, 4987–4997.

Le Hir, H., Gatfield, D., Braun, I. C., Forler, D., and Izzuralde, E. (2001b). The protein Mago provides a link between splicing and mRNA localization. EMBO Rep. *2*, 1119–1124.

Lejeune, F., Ishigaki, Y., Li, X., and Maquat, L. E. (2002). The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodelling. EMBO J. *21*, 3536–3545.

Leung, A.K.L., and Lamond, A. I. (2002). In vivo analysis of NHPX reveals a novel nucleolar localization pathway involving a transient accumulation in splicing speckles. J. Cell Biol. *157*, 615–629.

Leung, A.K.L., Andersen, J. S., Mann, M., and Lamond, A. I. (2003). Bioinformatics analysis of the nucleolus. Biochem. J. *376*, 553–569.

Lorkovic, Z. J., Hilscher, J., and Barta, A. (2004a). Use of fluorescent protein tags to study nuclear organization of the spliceosomal machinery in transiently transformed living plant cells. Mol. Biol. Cell *15*, 3233–3243.

Lorkovic, Z. J., Lopato, S., Pexa, M., Lehner, R., and Barta, A. (2004b). Interactions of arabidopsis RS domain containing cyclophilins with SR proteins and U1 and U11 snRNP-specific proteins suggest their involvement in pre-mRNA splicing. J. Biol. Chem. *279*, 33890–33898.

Lykke-Andersen, J., Shu, M. D., and Steitz, J. A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science *293*, 1836–1839.

Lytle, J. R., and Stietz, J. A. (2004). Premature termination codons do not affect the rate of splicing of neighbouring introns. RNA *10*, 657–668.

Macknight, S. (2003). Gene switching by metabolic enzymes—how did you get on the invitation list? Cell *114*, 150–152.

Maquat, L. E. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNA dynamics. Nat. Rev. Mol. Cell. Biol. *5*, 89–99.

Matera, A. G. (1999). Nuclear bodies: multifaceted sub-domains of the interchromatin space. Trends Cell Biol. *9*, 302–309.

Mazzola, J. L., and Sirover, M. A. (2002). Alteration of intracellular structure and function of glyceraldehydes-3-phosphate dehydrogenase: a common phenotype of neurodegenerative disorders? Neurotoxicology *23*, 603–609.

Mo, B., Tse, Y. C., and Jiang, L. (2003). Organelle identification and proteomics in plant cells. Trends Biotechnol. *21*, 331–332.

Olsen, M. O., Dundr, M., and Szebeni, A. (2000). The nucleolus: an old factory with unexpected capabilities. Trends Cell Biol. *10*, 189–196.

Palacios, I. M., Gatfield, G., St. Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature *427*, 753–757.

Pedersen, T. (1998). The plurifunctional nucleolus. Nucleic Acids Res. *26*, 3871–3876.

Phair, R. D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. Nature *404*, 604–609.

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

Rappsilber, J., and Mann, M. (2002). What does it mean to identify a protein in proteomics? Trends Biochem. Sci. *27*, 74–78.

Reed, R., and Hurt, E. A. (2002). A conserved mRNA export machinery coupled to pre-mRNA splicing. Cell *108*, 523–531.

Rocak, S., and Linder, P. (2004). DEAD-box proteins: the driving forces behind RNA metabolism. Nat. Rev. Mol. Cell. Biol. *5*, 232–241.

Rubbi, C. P., and Milner, J. (2003). Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. EMBO J. *22*, 6068–6077.

Saxena, P.K., Fowke, L. C., and King, J. (1985). An efficient procedure for isolation of nuclei from plant protoplasts. Protoplasma *128*, 184–189.

Scherl, A., Couté, Y., Déon, C., Callé, A., Kindbeiter, K., Sanchez, J.-C., Greco, A., Hochstrasse, D., and Diaz, J.-J. (2002). Functional proteomics analysis of human nucleolus. Mol. Biol. Cell *13*, 4100–4109.

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

Shaw, P. J. (2004). The nucleolus. In: Nature Encyclopaedia of Life Sciences/ www.els.net. Edited by: Nature Publishing Group.

Shaw, P. J. and Brown, J.W.S. (2004). Plant nuclear bodies. Curr. Opin. Plant Biol. *7*, 614–620.

Shaw, P. J., and Jordan, E. G. (1995). The nucleolus. Annu. Rev. Cell Dev. Biol. *11*, 93–121.

Sleeman, J. E., and Lamond, A. I. (1999). Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. Curr. Biol. *9*, 1065–1074.

Stutz, F., and Izauralde, E. (2003). The interplay of nuclear mRNP assembly, mRNA surveillance and export. Trends Cell Biol. *13*, 319–327.

Taylor, S. W., Fahy, E., and Ghosh, S. S. (2003). Global organellar proteomics. Trends Biotechnol. *21*, 82–87.

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

Wang, C., and Meier, U. T. (2004). Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. EMBO J. *23*, 1857–1867.

Watkins, N. J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Lührmann, R. (2000). A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. Cell *103*, 457–466.

Watkins, N. J., Dickmanns, A., and Lührmann, R. (2002). Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. Mol. Cell. Biol. *22*, 8342–8352.

Wendel, J. F. (2000). Genome evolution in polyploids. Plant Mol. Biol. *42*, 225–249.

Wilkinson, M. F., and Shyu, A.-B. (2002). RNA surveillance by nuclear scanning? Nat. Cell Biol. *4*, E144–E147.

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

Zheng, L., Roeder, R. G., and Luo, Y. (2003). S phase activation of the histone promoter OCA-S, a coactivator complex that contains GAPDH as a key component. Cell *114*, 255–266.

Zhou, Z., Luo, M. J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger RNA splicing to nuclear export in metazoans. Nature *407*, 401–405.