EMBO reports

review

Organellar proteomics: turning inventories into insights

Jens S. Andersen[1] *& Matthias Mann*[2]

[1]University of Southern Denmark, Odense, Denmark, and [2]Max-Planck Institute for Biochemistry, Martinsried, Germany

Subcellular organization is yielding to large-scale analysis. Researchers are now applying robust mass-spectrometry-based proteomics methods to obtain an inventory of biochemically isolated organelles that contain hundreds of proteins. High-resolution methods allow accurate protein identification, and novel algorithms can distinguish genuine from co-purifying components. Organellar proteomes have been analysed by bioinformatic methods and integrated with other large-scale data sets. The dynamics of organelles can also be studied by quantitative proteomics, which offers powerful methods that are complementary to fluorescence-based microscopy. Here, we review the emerging trends in this rapidly expanding area and discuss the role of organellar proteomics in the context of functional genomics and systems biology.

Keywords: functional genomics; mass spectrometry; protein dynamics; protein identification; protein localization

EMBO reports (2006) **7,** 874–879. doi:10.1038/sj.embor.7400780

Introduction

The 23,000 or so human genes give rise to a much larger number of active proteins owing to processes such as alternative splicing and post-translational modification. A further level of complexity is introduced by the precisely controlled temporal and spatial organization of these proteins. Microarray experiments provide information on the expression of transcripts, but do not address the important question of where in the cell proteins are directed to and where they are active. To understand cellular function, we need to know the cast of characters at each time point and location in the cell.

A simple hierarchy of cellular organization would include cellular compartments, such as the cytosol and nucleus, membrane-enclosed organelles, and large and small multiprotein complexes. There is some overlap between these levels, and the definitions are partly arbitrary (Fig 1). Organelles have classically been studied by biochemists with the help of separation and enrichment methods, and by cell biologists using microscopy. The first proteome-wide localization study using green fluorescent protein (GFP) fusion proteins in yeast has been published (Huh *et al*, 2003). In mammalian systems, however, such methods face numerous challenges as fusion proteins are usually overexpressed, and tagging itself is difficult and can lead to artefacts. In principle, mass-spectrometry-based proteomics (Aebersold & Mann, 2003) allows all the proteins in a purified complex to be identified. The first multiprotein complexes were characterized shortly after modern proteomics methods were introduced (Neubauer *et al*, 1997, 1998; Rout *et al*, 2000), and the proposal to systematically map multiprotein complexes dates back 10 years (Lamond & Mann, 1997). Here, we review mass-spectrometry-based proteomics methods that generate organellar proteomes and focus on approaches that go beyond simple protein lists to provide functional proteomics data. More than 250 organellar proteomics studies have been published and several specialized reviews can be found (see, for example, Brunet *et al*, 2003; Dundr & Misteli, 2001; Handwerger & Gall, 2006; Taylor *et al*, 2003; Yates *et al*, 2005).

Mass spectrometry to obtain protein inventories

After an organelle has been biochemically purified, it can be analysed by mass spectrometry, which has now almost completely replaced two-dimensional (2D) gel electrophoresis as a method for proteome analysis. Many mass-spectrometry-based methods have been used in organellar proteomics, the most robust and powerful of which involve one-dimensional (1D) gel electrophoresis, enzymatic degradation of the proteins, chromatographic separation of the resulting peptides and electrospray tandem mass spectrometry (MS^2; Fig 2). Advantages of this workflow are that 1D gel electrophoresis is an almost universal protein-separation method, and the use of sodium dodecyl sulphate (SDS) during this step allows the efficient denaturation of proteins and the removal of buffer constituents that are detrimental to mass spectrometry. However, direct 'in-solution' digestion followed by liquid-chromatography MS^2 is attractive for proteomes of medium complexity, particularly when measurements need to be repeated many times and a relatively small amount of sample is available. 2D peptide chromatography can be highly automated and can also characterize complex proteomes (Link *et al*, 1999).

One potential source of error in obtaining a proteomic inventory is the erroneous identification of proteins, and many published proteomes contain significant numbers of these 'false positives' (Steen & Mann, 2004). Fortunately, the development of high-accuracy mass spectrometers (Syka *et al*, 2004) and multiple fragmentation methods now allow the generation of high-quality data in high-throughput experiments (Olsen & Mann, 2004; Olsen *et al*, 2004). Furthermore, proper statistical interpretation of identification

[1]Centre for Experimental Bioinformatics (CEBI), University of Southern Denmark, Campusvej 55, DK-5230 Odense, Denmark
Tel: +45 6550 2365; Fax: +45 6593 3018;
E-mail: jens.andersen@bmb.sdu.dk
[2]Department of Proteomics and Signal Transduction, Max-Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany
Tel: +49 89 8578 2557; Fax: +49 89 8578 2219;
E-mail: mmann@biochem.mpg.de

Submitted 8 May 2006; accepted 7 July 2006

©2006 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

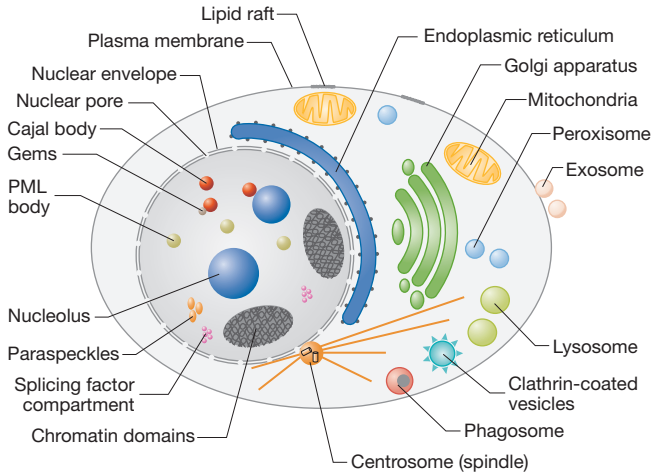


Fig 1 | What is an organelle? The figure depicts classical membrane-enclosed organelles, nuclear structures without membranes, and large multiprotein complexes or 'molecular machines' with distinct morphology and/or function, which can all be considered organelles. They are often highly dynamic: for example, some change their proteome in response to metabolic state and some exist only in certain phases of the cell cycle. PML, promyelocytic leukaemia.

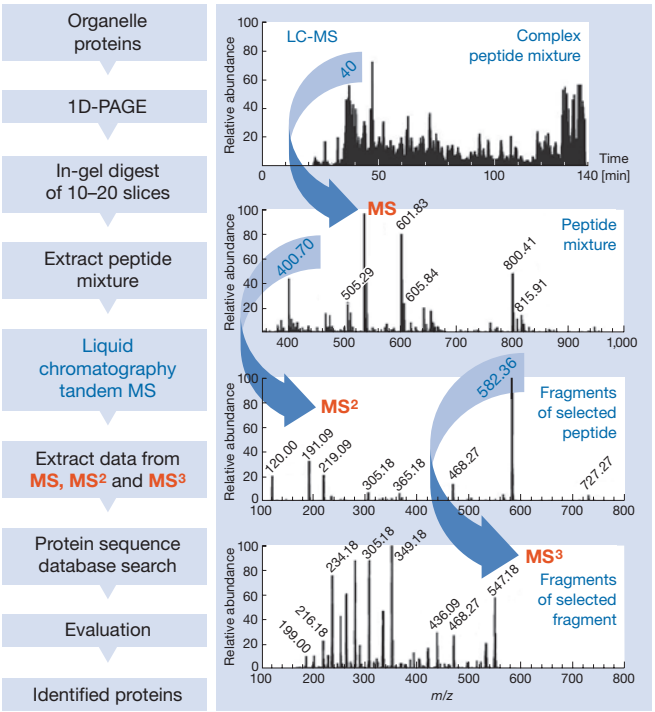


Fig 2 | Proteomics workflow. A protein mixture enriched in the organelle of interest is separated by 1D-PAGE. The entire gel band is cut into 10–20 slices, which are trypsin digested. The resulting peptide mixtures are separated by liquid chromatography and peptides are on-line ionized by electrospray mass spectrometry (MS). The panels show the summed ion signals of all the peptides eluting during the chromatographic separation, an example of a mass spectrum of eluting peptides, and a tandem (MS^2) spectrum obtained by isolating and fragmenting one of the eluting peptides. The mass and fragmentation information is matched against a sequence database by a search algorithm, resulting in a list of reported protein identifications for the organelle. Peptide sequences can be confirmed with MS^3 spectra obtained by isolating and fragmenting the most intense fragment in the MS^2 spectra (see Steen & Mann (2004) for an introduction to peptide sequencing by MS). 1D-PAGE, one-dimensional polyacrylamide gel electrophoresis.

results (Elias *et al*, 2005; Keller *et al*, 2002; Nesvizhskii *et al*, 2003) and the data requirements of journals (Carr *et al*, 2004) are helping to make results more transparent. Implementation of these technical, algorithmic and reporting standards could lead to the complete elimination of false-positive protein identifications. This is important if proteomic inventories are to lead to useful annotations in protein databases. Such annotations are difficult to reverse at a later date, and even a small error rate in each individual report on the same organelle will eventually lead to an annotated proteome consisting mostly of incorrect proteins.

Validation and integration

Proteins can be identified correctly but still might not be genuine components of an organelle. The high sensitivity and sequencing speed of modern mass spectrometers makes it inevitable that many co-purifying proteins are also identified, even in high-purity preparations. Therefore, additional data are needed to validate the proteins. 'Subtractive proteomics' compares the identified constituents of the complex of interest and a related 'background' complex. For example, the proteins identified in the 'H-complex' of the human spliceosome were subtracted from the spliceosomal inventory (Neubauer *et al*, 1998; Zhou *et al*, 2002). Schirmer and colleagues identified the proteins in a nuclear-envelope preparation and subtracted those that were also found in a microsomal membrane fraction (Schirmer *et al*, 2003). The remaining 67 previously unknown candidate integral nuclear-envelope proteins were then investigated in the literature for any disease association. Among these, 23 proteins were located in large chromosome regions associated with dystrophies and are being investigated further (Wilkie & Schirmer, 2006). Subtractive proteomics is limited by the 'sampling problem': in a complex proteome, not all peptides are sequenced and successive runs of the same sample do not overlap completely. To compare two related proteomes, it is preferable to use stable isotope labelling or to directly compare the ion current of identified peptides, even if they were sequenced in only one of the experiments (Foster *et al*, 2005). In addition, the composition of ill-defined compartments, such as lipid rafts, is difficult to study using proteomics. Using the stable isotope labelling by amino acids in cell culture (SILAC) technique (Ong *et al*, 2002) for comparative proteomics (Fig 3), Foster and colleagues isolated detergent-resistant lipid rafts from labelled and unlabelled cells, with and without depletion of cholesterol, respectively (Foster *et al*, 2003). In this way, any protein with a cholesterol-dependent ratio is, by definition, a member of the cholesterol-dependent lipid rafts. The quantitative data revealed an enrichment of signalling molecules in lipid rafts, which provides biochemical evidence for a functional connection between rafts and receptor-mediated signalling. In protein correlation profiling (PCP), the information inherent in the fractionation profile distinguishes between

©2006 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

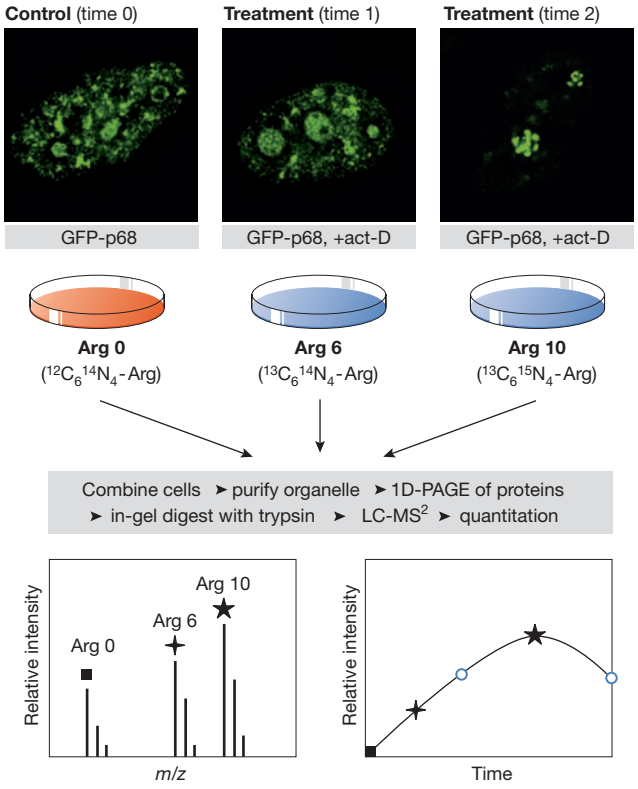


Fig 3 | Organellar dynamics using SILAC and fluorescent microscopy. Three cell populations are labelled to completion by isotope-encoded essential amino acids before the experiment. In the example shown here, transcription is inhibited by actinomycin D (act-D) for three different time periods. Cells are mixed and the nucleoli are isolated. Each peptide occurs as a triplet in mass spectrometry (MS) and peak heights reflect the relative amounts of the protein in the nucleolus at each time point. Multiplexing of the experiment results in a kinetic curve of protein recruitment to the nucleolus, in this case for p68. The top panels depict fluorescence microscopy for green fluorescent protein (GFP)-p68. 1D-PAGE, one-dimensional polyacrylamide gel electrophoresis; LC-MS, liquid chromatography MS; SILAC, stable isotope labelling by amino acids in cell culture.

genuine and background proteins (Fig 4). Andersen and colleagues used PCP to eliminate about 90% of proteins in a centrosomal preparation, leading to the high-confidence identification of 114 true centrosomal proteins (Andersen *et al*, 2003). PCP can also be performed with internal standards (Fig 4) to increase the quantification accuracy, which has resulted in the identification of additional centrosomal proteins (J.S.A., unpublished data). Foster and colleagues extended the PCP analysis to all membrane-enclosed organelles in the mouse liver (Foster *et al*, 2006). In total, 1,900 proteins were quantified and 1,404 localized to 10 subcellular compartments. This large-scale PCP data set suggests error rates in published organellar data sets of between 3 and 64% owing to co-purifying proteins. Furthermore, the global nature of this study revealed multiple locations for 39% of all organellar proteins.

The candidate proteomic inventory can be analysed with a variety of bioinformatic methods, and several web-based organellar databases and prediction algorithms (Donnes & Hoglund, 2004) can independently support parts of the inventory. Bioinformatic predictions are usually based on sequence features, such as signal peptides or nuclear-localization sequences, and are most useful for membrane-bound organelles. By contrast, the components of subnuclear domains cannot be predicted accurately. In this case, an increasingly useful tool for any proteome analysis is the Gene Ontology project (Ashburner *et al*, 2000), which classifies proteins not only by localization but also by biological process and molecular function. Even though the categories are broad, they give a useful overall assessment of the data.

All these analytical tools implicitly assume that organellar proteomes are the same, regardless of tissue or cell type. However, a study of the mitochondrial proteome showed substantial variations depending on the tissue from which it was derived (Mootha *et al*, 2003a). Emili and co-workers recently separated cell lysates into cytosolic, nuclear, mitochondrial and membrane fractions in six different mouse tissues (Kislinger *et al*, 2006). This study provided tissue expression and compartment information for more than 5,000 proteins and showed that organelle protein composition varied between tissues. There might also be differences in subcellular organization between cells in culture and primary cells or tissues. Nevertheless, bioinformatics analyses suggest a high degree of evolutionary conservation of organellar proteomes. For example, human homologues of proteins localizing to the yeast nucleolus also had an 89% probability of being nucleolar (Andersen *et al*, 2005). A similar result was found in a large-scale analysis of membrane-bound organelles (Foster *et al*, 2006).

Fluorescence-based microscopy is another powerful technique for studying protein subcellular localization. However, microscopy is based on single cells, whereas analysing the constituents of organelles that have been biochemically enriched averages over millions of cells. Furthermore, microscopy is usually performed either with antibodies, which potentially raise specificity issues, or with tagged proteins, which need to be shown to behave similarly to endogenous proteins. A combination of MS-based proteomics and microscopy is therefore an attractive validation method for both, as illustrated here with examples from mitosis-related structures. In the centrosome study mentioned above, a subset of 23 novel centrosomal proteins with correct PCP profiles was amino-terminally and carboxy-terminally tagged, and shown to localize to the centrosome (Andersen *et al*, 2003). Sauer and colleagues investigated the components of the spindle pole body and reported 151 proteins, 17 of which were tagged and 6 localized to the spindle (Sauer *et al*, 2005). Skop and co-workers reported an inventory of the midbody in Chinese hamster ovary cells (Skop *et al*, 2004). An RNA interference (RNAi) follow-up of homologues in *Caenorhabditis elegans* revealed a furrow-formation defect in 100 of the 172 candidates, and *in vivo* localization validated 10 of these as novel midbody proteins. The use of orthogonal methods to validate subcellular localization is important not only to establish subcellular assignment of the identified proteins but also to explore the functional value of proteomics experiments. In a study of proteins binding to major components of focal-adhesion complexes, de Hoog and colleagues identified many RNA-binding proteins. On the basis of quantitative data, these proteins could not easily be ignored as contaminants. Imaging experiments subsequently led to the discovery of a new structure that is involved in cell adhesion (de Hoog *et al*, 2004).

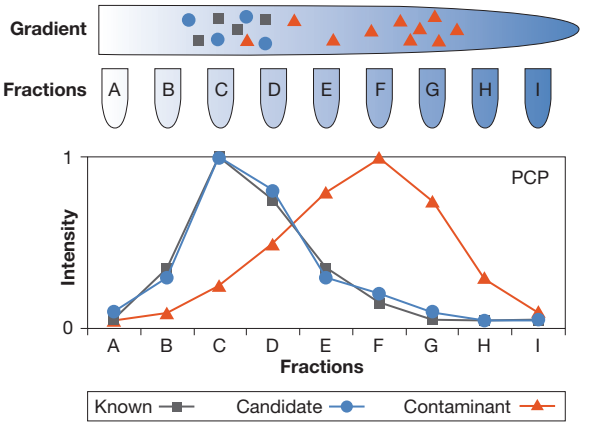
©2006 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

Dynamics of organelles

The protein constituents of most organelles are in constant exchange with the rest of the cell. Several organelles—particularly nuclear subdomains—contain both resident and transient proteins. For example, during their maturation, proteins and protein complexes might pass through several subdomains (Swedlow & Lamond, 2001). Furthermore, perturbation of the cell might change the proteome of an organelle. So far, these processes have generally been studied using microscopy, which, despite its tremendous power, remains a candidate-based approach and only allows a few factors to be followed in a single experiment. By contrast, proteomic investigation can follow changes in the proteome in an unbiased way for all the components of an organelle. Using the SILAC technology, Andersen and colleagues have performed the first comprehensive study to follow the temporal change of an organelle after perturbation (Andersen *et al*, 2005). Transcription was inhibited in SILAC-labelled cells for different lengths of time, and changes in the proteome were quantified at up to nine time points (Fig 3). The experiments were performed using stable cell lines expressing GFP-tagged nucleolar proteins to compare the MS data directly with the corresponding quantitative fluorescence measurements in the same cell cultures. Many, but not all, nucleolar proteins left the nucleolus after transcription inhibition, whereas a small group of proteins was recruited. Functionally or physically interacting proteins, such as the exosome subunits and the polymerase I subunits, had similar kinetic profiles. The response of the nucleolar proteome was also followed for several other perturbations, with interesting results. For example, proteasome inhibition resulted in largely opposite recruitment patterns for the ribosomal proteins compared with transcription inhibition, which suggests the involvement of the ubiquitin proteasome system in ribosome biogenesis. Furthermore, through a variation of this experimental strategy, protein synthesis and maturation can be studied directly (Doherty & Beynon, 2006). Importantly, organellar dynamics can be assessed for a large variety of perturbations once the organellar proteome has been defined and the method has been established. In this way, quantitative mass spectrometry can be used as a functional assay similar to others that are routinely used in cell biology.

Functional genomics with subcellular resolution

The location of a protein is an important independent parameter that can be used in interpreting the results of any large-scale experiment. For example, Mootha and colleagues used the mitochondrial proteome to identify the gene mutated in a form of Leigh syndrome (Mootha *et al*, 2003b). This disorder shows clinical and biochemical features pointing to the involvement of mitochondria, and genetic analysis of affected families has implicated a specific genomic region in the development of the disease. Only a single gene in this region encodes a protein that is a member of the mitochondrial proteome data set and is co-regulated at the messenger RNA level with other mitochondrial genes. Data from the same mitochondrial proteome mapping experiment were also used in a sophisticated bioinformatic analysis to identify transcription-factor-binding sites upstream from mitochondrial genes. Furthermore, a gene-expression neighbourhood index was defined to capture genes the transcripts of which showed significant co-regulation with mitochondrial genes. Several transcription factors and low-abundance factors that were not sequenced in the proteome experiments were retrieved by this analysis. Foster and co-workers have recently reported a similar analysis in their whole-cell study of organellar structures by PCP (Foster *et al*, 2006). These analyses would not be possible with the entire proteome, but derive their statistical power from the fact that a single organelle is queried. Organellar data can generally be layered on top of any other large-scale data. Thus, as more protein–protein interaction data become available, it will be useful to interpret them in the light of protein co-localization in organelles (Hinsby *et al*, 2006). Large-scale mutagenesis screens and live-cell RNAi experiments (Pepperkok & Ellenberg, 2006) combined with proteomics data are also expected to provide insight into the function of organelles. Through a genetic screen in zebrafish looking for



Protein correlation profiling

Normalized intensity of same peptide *m/z* in A, B, C, D, E
Compare profiles with profile for known organelle protein

Localization of organelle proteins by isotope tagging (LOPIT)

Ratio of peptides: $^{light}B/^{heavy}F$, $^{light}C/^{heavy}G$, $^{light}D/^{heavy}H$
Compare set of ratios with ratios for known organelle protein

PCP with matching internal standard

Ratio of peptides: $^{light}A/X$, $^{light}B/X$, $^{light}C/X$, $^{light}D/X$, $^{light}E/X$
where $X=(^{heavy}A + ^{heavy}B + ^{heavy}C + ^{heavy}D + ^{heavy}E)/5$
Compare set of ratios with ratios for known organelle protein

Subtractive proteomics

Proteins in G subtracted from proteins in C

Comparative proteomics

Ratio of peptides: $^{light}C(control)/^{heavy}C(case)$

Fig 4 | Protein correlation profiling. Fractions adjacent to the peak fraction of an organelle are also measured, and signals from the same peptides are correlated with each other. A quantitative curve is obtained for each protein. Marker proteins for an organelle define a consensus curve and deviation of each protein profile from the consensus curve is a measure of its likelihood of being a genuine member of the organelle. Proteins can be quantified by integrating peptide ion currents (PCP), by stable isotope-based quantification of adjacent fractions (Dunkley *et al*, 2004) or, most accurately, by using stable isotope labelling by amino acids in cell culture (SILAC) to provide the same internal standard for each fraction. Alternative approaches include subtractive proteomics or comparative proteomics (see main text). PCP, protein correlation profiling.

©2006 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

kidney cyst defects, Sun and colleagues identified several genes with homologues in flagella that were linked to human kidney disease (Sun *et al*, 2004). Supporting this association between cilia and cystogenesis, three proteomic investigations into the eukaryotic flagellum also found several homologues of human ciliary diseases (Broadhead *et al*, 2006; Keller *et al*, 2005; Pazour *et al*, 2005). Proteomic characterization of the human centrosome resulted in the identification of several disease-associated genes, including nephrocystin 6, which is another cystic kidney disease gene that is mutated in Joubert syndrome (Sayer *et al*, 2006).

Large-scale proteomics of post-translational modifications can also be performed with subcellular resolution (Jensen, 2006). For example, Nousiainen and co-workers defined 736 phosphorylation sites on mitotic spindle-associated proteins (Nousiainen *et al*, 2006), which constitute a promising starting point for functional studies. Temporal or comparative analysis of site-specific modifications by quantitative proteomics helps to distinguish basal sites from regulatory sites. It will be interesting to use such approaches to directly study signalling to organelles.

In the cell nucleus, organelles are difficult to define and purify. In several proteomic analyses of nuclear structures (Andersen *et al*, 2005; Jurica & Moore, 2003; Saitoh *et al*, 2004), unexpected proteins are commonly observed. Their presence in certain compartments possibly reflects non-specific associations with proteins roaming the nuclear space (Gorski & Misteli, 2005). Dynamic organellar proteomics might help to address this issue and to determine how these structures are formed and maintained. A systems biology approach to the temporal and spatial organization of the nucleus, including chromatin states, would be of great value in understanding gene expression and other fundamental molecular events in the cell (Gorski & Misteli, 2005).

Conclusion

The increasing power of mass-spectrometry-based proteomics now makes it possible to characterize organelles with more than 1,000 proteins and with a dynamic range in protein abundance of several orders of magnitude. In fact, obtaining mass-spectrometry data has already ceased to be the limiting step in organellar proteomics. Instead, the main challenges are purifying the organelle and removing the background proteins. As shown by several examples, integrating organellar data sets with existing functional genomics data can be fruitful. Conversely, large-scale organellar data enhance other types of data, giving them 'subcellular resolution'. In the next few years, we expect mass-spectrometry-based proteomics to define a basic set of inventories for the main cellular organelles. These inventories will then allow even more interesting questions to be asked about the effect of perturbations on organellar proteomes and what can be revealed by the dynamic interchange of proteins between compartments.

REFERENCES

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

Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, Mann M (2003) Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426:** 570–574

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

Ashburner M *et al* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25:** 25–29

Broadhead R *et al* (2006) Flagellar motility is required for the viability of the bloodstream trypanosome. *Nature* **440:** 224–227

Brunet S, Thibault P, Gagnon E, Kearney P, Bergeron JJ, Desjardins M (2003) Organelle proteomics: looking at less to see more. *Trends Cell Biol* **13:** 629–638

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

de Hoog CL, Foster LJ, Mann M (2004) RNA and RNA binding proteins participate in early stages of cell spreading through spreading initiation centers. *Cell* **117:** 649–662

Doherty MK, Beynon RJ (2006) Protein turnover on the scale of the proteome. *Expert Rev Proteomics* **3:** 97–110

Donnes P, Hoglund A (2004) Predicting protein subcellular localization: past, present, and future. *Genomics Proteomics Bioinformatics* **2:** 209–215

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

Dunkley TP, Watson R, Griffin JL, Dupree P, Lilley KS (2004) Localization of organelle proteins by isotope tagging (LOPIT). *Mol Cell Proteomics* **3:** 1128–1134

Elias JE, Haas W, Faherty BK, Gygi SP (2005) Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat Methods* **2:** 667–675

Foster LJ, de Hoog CL, Mann M (2003) Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc Natl Acad Sci USA* **100:** 5813–5818

Foster LJ, Zeemann PA, Li C, Mann M, Jensen ON, Kassem M (2005) Differential expression profiling of membrane proteins by quantitative proteomics in a human mesenchymal stem cell line undergoing osteoblast differentiation. *Stem Cells* **23:** 1367–1377

Foster LJ, de Hoog CL, Zhang Y, Zhang Y, Xie X, Mootha VK, Mann M (2006) A mammalian organelle map by protein correlation profiling. *Cell* **125:** 1–13

Gorski S, Misteli T (2005) Systems biology in the cell nucleus. *J Cell Sci* **118:** 4083–4092

Handwerger KE, Gall JG (2006) Subnuclear organelles: new insights into form and function. *Trends Cell Biol* **16:** 19–26

Hinsby AM, Kiemer L, Andersen JS, Mann M, Brunak S (2006) Wiring the human nucleolus. *Mol Cell* **22:** 285–295

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

Jensen ON (2006) Interpreting the protein language using proteomics. *Nat Rev Mol Cell Biol* **7:** 391–403

Jurica MS, Moore MJ (2003) Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell* **12:** 5–14

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

Keller LC, Romijn EP, Zamora I, Yates JR, Marshall WF (2005) Proteomic analysis of isolated *Chlamydomonas* centrioles reveals orthologs of ciliary-disease genes. *Curr Biol* **15:** 1090–1098

Kislinger T *et al* (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* **125:** 173–186

Lamond AI, Mann M (1997) Cell biology and the genome projects a concerted strategy for characterizing multiprotein complexes by using mass spectrometry. *Trends Cell Biol* **7:** 168–173

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

Mootha VK *et al* (2003a) Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115:** 629–640

Mootha VK *et al* (2003b) Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. *Proc Natl Acad Sci USA* **100:** 605–610

Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **75:** 4646–4658

©2006 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION

Neubauer G, Gottschalk A, Fabrizio P, Seraphin B, Luhrmann R, Mann M (1997) Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc Natl Acad Sci USA* **94:** 385–390

Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A, Mann M (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat Genet* **20:** 46–50

Nousiainen M, Sillje HH, Sauer G, Nigg EA, Korner R (2006) Phosphoproteome analysis of the human mitotic spindle. *Proc Natl Acad Sci USA* **103:** 5391–5396

Olsen JV, Mann M (2004) Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proc Natl Acad Sci USA* **101:** 13417–13422

Olsen JV, Ong SE, Mann M (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* **3:** 608–614

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

Pazour GJ, Agrin N, Leszyk J, Witman GB (2005) Proteomic analysis of a eukaryotic cilium. *J Cell Biol* **170:** 103–113

Pepperkok R, Ellenberg J (2006) High-throughput fluorescence microscopy for systems biology. *Nature Reviews Mol Cell Bio* [published online ,19 Jul] doi:10.1038/nrm1979

Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol* **148:** 635–651

Saitoh N, Spahr CS, Patterson SD, Bubulya P, Neuwald AF, Spector DL (2004) Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell* **15:** 3876–3890

Sauer G, Korner R, Hanisch A, Ries A, Nigg EA, Sillje HH (2005) Proteome analysis of the human mitotic spindle. *Mol Cell Proteomics* **4:** 35–43

Sayer JA *et al* (2006) The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4. *Nat Genet* **38:** 674–681

Schirmer EC, Florens L, Guan T, Yates JR, Gerace L (2003) Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* **301:** 1380–1382

Skop AR, Liu H, Yates J, Meyer BJ, Heald R (2004) Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. *Science* **305:** 61–66

Steen H, Mann M (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5:** 699–711

Sun Z, Amsterdam A, Pazour GJ, Cole DG, Miller MS, Hopkins N (2004) A genetic screen in zebrafish identifies cilia genes as a principal cause of cystic kidney. *Development* **131:** 4085–4093

Swedlow JR, Lamond AI (2001) Nuclear dynamics: where genes are and how they got there. *Genome Biol* **2:** REVIEWS0002

Syka JE *et al* (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

Taylor SW, Fahy E, Ghosh SS (2003) Global organellar proteomics. *Trends Biotechnol* **21:** 82–88

Wilkie GS, Schirmer EC (2006) Guilt by association: the nuclear envelope proteome and disease. *Mol Cell Proteomics* 21 June [Epub ahead of print]

Yates JR, Gilchrist A, Howell KE, Bergeron JJ (2005) Proteomics of organelles and large cellular structures. *Nat Rev Mol Cell Biol* **6:** 702–714

Zhou Z, Licklider LJ, Gygi SP, Reed R (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature* **419:** 182–185

Jens S. Andersen

Matthias Mann

©2006 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION