Organelle proteome variation among different cell types: lessons from nuclear membrane proteins

Organelle Proteome Variation Among Different Cell Types: Lessons from Nuclear Membrane Proteins

Deirdre M. Kavanagh, William E. Powell, Poonam Malik, Vassiliki Lazou, and Eric C. Schirmer

Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, UK

Correspondence to:
Eric C. Schirmer, Ph.D.
Institute of Cell Biology and
Wellcome Trust Centre for Cell Biology
University of Edinburgh
Kings Building Campus
Swann 4.19, Mayfield Road
Edinburgh EH9 3JR
UK

Tel: +44 131 650 7090
Fax: +44 131 650 7360
e.schirmer@ed.ac.uk

Running title: Nuclear membrane proteome variation

Abbreviations: ER, endoplasmic reticulum; gp, glycoprotein; GFP, green fluorescent protein; HA, hemagglutinin; Ig, immunoglobulin; KASH, klarsicht ANC -1 and syne homology; INM, inner nuclear membrane; LAP1, lamina – associated polypeptide 1; LAP2, lamina – associated polypeptide 2; LBR, lamin B receptor; LEM, LAP2, emerin and MAN1 domain; MALDI, matrix-associated laser desorption/ionization; MMs, microsomal membranes; MudPIT, multi-dimensional protein identification technology; NE, nuclear envelope; Nesprins, nuclear envelope spectrin repeat; NET, nuclear envelope transmembrane protein; NPC, nuclear pore complex; Nurim, nuclear rim protein; ONM, outer nuclear membrane; PoM, pore membrane; Rb, retinoblastoma; Syne, synaptic nuclear envelope; SUN, sad -1 and UNC domain.

Table of Contents

Abstract

1. INTRODUCTION
2. THE HISTORY OF NETS
2.1 The Genomics Era
3. SUBTRACTIVE PROTEOMICS
3.1. Biochemical Fractionation
3.2. MudPIT Proteomics
3.3. Subtraction
4. CONFIRMATION OF NUCLEAR ENVELOPE RESIDENCE
4.1. Determination of Subcellular Localization
4.2. Determination of Membrane Insertion
4.3. Additional Complications
5. VARIATION IN THE NUCLEAR ENVELOPE PROTEOME
5.1. Tissue Variation: The Transcriptome Database
5.2. Evolutionary Variation
6. NETS IN MULTIPLE CELLULAR LOCATIONS
7. NUCLEAR ENVELOPE DISEASES AND FUNCTION
8. CONCLUSIONS
ACKNOWLEDGEMENTS
LEGENDS TO FIGURES AND TABLES
REFERENCES

Abstract

Most subcellular organelles are expected to be similar among different cell types; however, a recent study suggests a surprising amount of variation in the protein composition at the nuclear envelope. Therefore, to comprehensively identify proteins in subcellular organelles proteomics datasets may need to be generated from multiple cell types. In this chapter we describe a proteomics study that expanded the number of nuclear membrane proteins by 5-fold using a "subtractive" methodology in which a subcellular organelle is partially purified biochemically and partially purified *in silico*. The biochemical fraction of interest and a separate fraction enriched in proteins known to contaminate it, in this case nuclear envelopes and microsomes respectively, are first isolated and separately analyzed by mass spectrometry. For *in silico* purification, proteins appearing in both fractions are subtracted from the dataset in order to identify proteins that are unique to the organelle being investigated. This approach identified sixty-seven novel putative nuclear envelope transmembrane proteins in rodent liver. Further analysis of their expression levels in other tissues indicates that several are preferentially expressed in liver cell types, which in turn predicts considerable variation in the nuclear envelope proteome among different cell types. Finally, we discuss several issues associated with confirming that these peptide-based identifications represent proteins that truly localize to the nuclear envelope. These studies have complicated rather than simplified our view of the nuclear envelope, but proteomics has set the stage for beginning to understand this highly complex subnuclear organelle.

1. INTRODUCTION

It is clear that each cell type expresses only a fraction of the proteins encoded by the genome, with some unique and some ubiquitous proteins. It is less clear how these unique and ubiquitous proteins are distributed at a subcellular level. Cell-type specific proteins have long been demonstrated in the nucleus where distinctive transcription factors regulate which proteins are made, at the plasma membrane where many signalling cascades are instigated, and in the cytoplasm where many unique metabolic events occur. In contrast, the view is widely held that proteins in most other organelles such as mitochrondria, endoplasmic reticulum (ER), vacuoles, and nucleoli would be largely similar from one cell type to another. For some organelles, recent studies have shown that protein variation is greater than previously thought. For example, cell-type specific differences in the nucleus go well beyond transcription factors with certain histone H1 variants exchanged at specific stages in development and in differentiated cell types, presumably to faciliate unique patterns of gene expression (Brown, 2001; Lee et al., 2004). Historically the nuclear subcompartment termed the nuclear envelope (NE) was viewed as having only "structural" roles, serving as a barrier between the nucleus and cytoplasm; however analysis of a recent proteomics dataset suggests that its organization is quite intricate, with an unexpected degree of variation in the NE proteins present in different cell types (Schirmer et al., 2005; Schirmer et al., 2003; Schirmer and Gerace, 2005). Moreover, this variation in the NE proteome may provide a resolution to the conundrum of how mutations in a subset of ubiquitous NE proteins can cause tissue-specific diseases (Mounkes and Stewart, 2004; Worman and Courvalin, 2002).

The NE is an elaborate structure that can be divided into several distinct subdomains: the nuclear pore complexes (NPCs), the lamin polymer, and a double membrane system consisting of the outer nuclear membrane (ONM), inner nuclear membrane (INM), lumen, and pore membrane (PoM) together with their integral proteins (Figure 1). The ONM is not only continuous with the ER, but is also studded with ribosomes indicating that in addition to being the outermost layer of the nucleus it is also a subcompartment of the ER. How much of its complement of integral membrane proteins is unique from more distal ER remains unclear, but it has now been clearly demonstrated that proteins involved in tethering the nucleus to the cytoskeleton are distinctive components of the ONM (Crisp et al., 2006; Starr and Han, 2002; Wilhelmsen et al., 2005). The lumenal space between the ONM and INM is largely

unexplored territory, but recent studies suggest some proteins may make contacts across the lumen in order to maintain its highly regular spacing (Crisp et al., 2006). Such proteins may sterically interfere with lateral diffusion of membrane proteins with large lumenal domains and, indeed, transport from the ER to the INM of transmembrane protein chimeras is inhibited when they carry a large lumenal domain (Ohba et al., 2004). The ONM and INM are joined where NPCs perforate the membrane. The PoM effetively wraps around these > 1 MegaDalton macromolecular complexes. The INM has a more clearly defined set of integral membrane proteins compared to the ONM because many are physically associated with the nuclear lamin polymer that lies underneath it, and these proteins remain associated with lamins in a biochemical fraction when the lipid of the membranes is extracted with detergent. Lamins are type V intermediate filament proteins that maintain very stable homodimers through the formation of a 52 nm coiled-coil rod by their central alpha-helical domain. Flanking the rod is a short amino-terminal head domain, and a large globular tail that contains an Ig fold (Dhe-Paganon et al., 2002; Krimm et al., 2002; Stuurman et al., 1998). While the membrane forms a barrier to diffusion of soluble macromolecules between the nucleus and cytoplasm, the lamins provide mechanical stability to the structure (Lammerding et al., 2004; Liu et al., 2000; Schirmer et al., 2001).

2. THE HISTORY OF NETS

Before the application of proteomics to the NE, proteins of this structure were identified through a wide variety of approaches. Lamins were the first NE proteins characterized (roughly 30 years ago) because their solubility characteristics as intermediate filament proteins and their abundance facilitated easy biochemical enrichment (Aaronson and Blobel, 1975; Gerace et al., 1978). It is estimated that there are ~ 3 million lamin molecules in the average mammalian nucleus (Gerace and Burke, 1988).

Individually, the integral membrane proteins are much less abundant; thus, it was over a decade later that the first NE transmembrane proteins (NETs) were discovered. The lamin B receptor (LBR) was identified by its binding to lamin B1 (Worman et al., 1988), while lamina-associated polypeptides 1 and 2 (LAP1 and LAP2) were identified in a screen for monoclonal antibodies using the large number of proteins in a biochemically isolated NE fraction (Foisner and Gerace, 1993; Senior and Gerace, 1988). Consistent with their identification by this means, the two LAP proteins are likely to be among the most abundant of the NETs. This is further supported by the high relative abundance of LAP peptides recovered by proteomics (Schirmer et al., 2003). Two NETs were identified as part of the NPC, gp210 and POM121 (pore membrane 121) (Greber et al., 1990; Hallberg et al., 1993). Some NETs were identified by microscopy studies: MAN1 was identified fortuitously from an autoimmune serum that produced the nuclear "rim" staining characteristic of lamins (Paulin-Levasseur et al., 1996) while nurim (for nuclear rim) was identified by the screening of a GFP-cDNA fusion library for proteins that yielded a similar localization (Rolls et al., 1999). Human genetics identified emerin as the gene responsible for Emery-Dreifuss Muscular Dystrophy (Bione et al., 1994) and subsequent work to characterize its protein product determined that it was also a NET (Manilal et al., 1996).

Characterization of a protein identified in a 2-hybrid screen for partners of a kinase of the muscle postsynaptic membrane provided the first evidence of a new family of proteins at the NE (Apel et al., 2000). Originally named Syne for synaptic nuclear envelope (Apel et al., 2000), this family is becoming increasingly referred to as Nesprins for nuclear envelope spectrin repeat because each member contains multiple copies of this motif (Zhang et al., 2001). Two separate genes were originally identified (Apel et al., 2000) that share a common sequence motif termed the KASH domain (for Klarsicht, ANC-1, and Syne-1 homology after the founding members from Drosophila (Mosley-Bishop et al., 1999), *Caenorhabditis elegans* (Starr and Han, 2002), and mouse (Apel et al., 2000). A second protein family, denoted SUN (for Sad-1 and UNC domain), was originally identified at the NE in *C. elegans* (Unc84; Malone et al., 1999). Its first mammalian homolog, Unc84A/ SUN1, was discovered in the

first proteomics study for NETs (see below; Dreger et al., 2001). SUN2 was identified in an effort to isolate genes with differential expression in heart tissues (Sun et al., 2002).

That UNCL was found to target to the NE was surprising because it was uncovered in a study to identify chaperones that function in assembly of nicotinic acetylcholine receptors (Fitzgerald et al., 2000). Although little note was taken of the fact at the time, its discovery provided the first indication that NET composition differs among different cell types: the study was predicated on the observation that these receptors do not properly assemble in many mammalian cell lines and, indeed, UNCL is absent from most cell types.

2.1 The Genomics Era

This wide range of approaches has been replaced by proteomics in the post-genomics era. Interestingly, genomics itself was inadequate for the identification of NETs. For example, a weak homology "LEM" domain was identified when MAN1 was cloned and sequenced (Lin et al., 2000) because this domain was also found to occur in LAP2 and emerin (hence LEM for LAP2, emerin, and MAN1). However, not all proteins containing LEM domains are concentrated at the NE (Dechat et al., 1998; Raju et al., 2003) and the presence of a LEM domain is not always maintained within a protein family *e.g.* the Drosophila Syne/ Nesprin homolog (Klarsicht) contains a LEM domain while the mammalian members of this family do not (Wagner et al., 2004).

There were three initial proteomics studies of NE proteins. Two of these focussed on the NPCs and will not be detailed here other than to mention that the approaches were geared towards isolating its structural core (Cronshaw et al., 2002; Rout et al., 2000). The third study focussed on identifying novel INM proteins, using a comparative proteomic approach where different NE subfractions isolated from cultured neuroblastoma cells were compared based on the characteristics of previously known NETs (Dreger et al., 2001). Integral proteins are enriched after extraction with chaotropes (in this case 4 M urea, 0.1 M Na_2CO_3); however NEs so extracted yield a pellet containing membrane proteins not only of the INM, but also of the ONM and ER. Since known INM proteins remain pelletable after detergent extraction (in this case 1% Triton X-100), they considered proteins of the chaotrope pellet that were also found in the detergent pellet to be good candidates for novel INM proteins. Each subfraction was separated on 2-D gels, the protein spots were excised, and each spot was analyzed by MALDI mass spectrometry. This analysis identified most expected previously characterized INM proteins in both the chaotrope and detergent fractions; however LAP1 was not identified, nor were the later-identified SUN2 and nesprins. (UNCL was not identified as expected due to its limited tissue expression.) Moreover, LBR and emerin, though identified, did not behave as expected: LBR only appeared in the detergent-resistant fraction and emerin only appeared in the chaotrope-resistant fraction. Nonetheless, the approach was successful in identifying four novel proteins that appeared in both fractions: two additional splice variants of LAP2 (previously predicted from mRNA; Berger et al., 1996), the first mammalian SUN family protein (Unc84A/ SUN1), and a completely novel protein with no predicted functions that was named LUMA (Dreger et al., 2001). Several additional putative novel NETs were identified that only occurred in one or the other fraction.

Though both creative and productive, in retrospect there were certain limitations to this approach. The assumption that the membrane extraction characteristics of the proteins would be conserved was in practice not absolute. Moreover, the inherent problems of working with gel-extracted bands were exemplified by the inability to make identifications for roughly 25% of the gel spots (Dreger et al., 2001).

3. SUBTRACTIVE PROTEOMICS

Like the Dreger et al. study, the fourth NE proteomics study took the view that it is impossible to truly purify the NE biochemically because of its many connections to ER, cytoplasmic filament systems, and chromatin. It also assumed that any chemical extraction

chosen to improve purity might remove true components as well as contaminants. The design of this fourth approach renders it much better for identification of NE-specific transmembrane proteins than for soluble NE-specific components. The principle is that transmembrane proteins in isolated NEs should derive only from the INM, PoM, ONM, or ER. In contrast, microsomal membranes (MMs) are rich in ER transmembrane proteins, but are free of INM, PoM, or ONM because the large dense nuclei are readily pelletable away from the smaller membrane vesicles obtained by disruption of ER and other membrane systems. Therefore, while isolated nuclei will have contamination from ER, MMs will be free of nuclear contamination: by separately analyzing NEs and MMs and subtracting proteins identified in both fractions, NE-specific transmembrane proteins can be identified (Figure 2). In contrast, soluble contaminants could derive from "sticky" chromatin proteins released during NE purification, collapsed cytoplasmic filaments and associated proteins, or proteins caught in transit through the NPC at the time of purification. Although well-characterized contaminants such as cytoplasmic filaments can easily be identified and discarded, it is difficult to distinguish a separate fraction to subtract in order to identify which other soluble proteins are normally enriched at the NE.

3.1. Biochemical Fractionation

To purify NEs (Blobel and Potter, 1966; Dwyer and Blobel, 1976), cells are first lysed by dounce homogenization to release nuclei. Specific protocols for this step vary considerably from one cell type to another based on such criteria as nuclear:cytoplasmic volume ratio, types and abundance of cytoplasmic filament systems, and amount of connective tissue. For example, soft tissues with low nuclear:cytoplasmic ratios such as liver can be directly homogenized in a Potter-Elvehjem homogenizer (Blobel and Potter, 1966), while lymphocytes that have much higher nuclear:cytoplasmic ratios must be first hypotonically swollen and then dounce homogenized (Fields et al., 1988). Some protocols call for treatment with filament destabilizing drugs (such as cytochalasin B) because cytoskeletal filaments have been shown to collapse on the nuclear surface when cells are homogenized (Staufenbiel and Deppert, 1982), but the tightly banded filaments in muscle are impervious to such treatments. To remove these filaments, the muscle must be first minced or treated with a blade homogenizer (such as a polytron) prior to dounce homogenization. Care must be taken to break the muscle fibers into pieces that are much larger than nuclei because pieces of similar density will be difficult to separate from nuclei in subsequent steps (Held et al., 1977; Kuehl, 1977). It is noteworthy that many protocols for isolation of nuclei use detergent in buffers, but this must be avoided as detergents also remove most of the NE.

Because the intact nucleus is larger and denser than any other organelle in most cells, it is readily separated from other organelles by pelleting centrifugation. The speed and duration of this pelleting step should be appropriately modified for different tissues in order to separate the nuclei as much as possible from other cellular structures before subsequent steps (*e.g.* an initial step at slow speed to pellet large debris without nuclei, followed by a step at intermediate speed to pellet nuclei away from smaller cell debris). Many other membranes and contaminants that co-sediment with the nuclei in this first step are separated by floating on sucrose cushions that the intact nuclei are sufficiently dense to penetrate. Many membranes will float in 1.9 M sucrose; so sucrose is mixed with the homogenate to this density and underlayed with a sucrose cushion of higher density. If too much cell debris is loaded onto the sucrose cushions, it will accumulate at the interface and can prevent penetration by the nuclei. The size, shape, and density of nuclei vary considerably among different cell types and thus the density of sucrose cushions must be accordingly varied for effective recovery. It is also possible to isolate nuclei using other density gradient systems such as percoll (Hahn and Covault, 1990).

Isolated nuclei are then treated with DNase and RNase to digest DNA, which is then washed out of the nucleus along with a significant percentage of the nucleoplasmic content using high salt buffers (~half molar). This released chromatin material is then separated from NEs by floating on another sucrose cushion. Through this disruptive process NEs may

fracture, but their structure is largely maintained because the membranes and their integral protein complement are connected to the salt-resistant lamin polymer. The resulting nuclear envelopes are thus largely, but not completely, washed clean of both cytoplasmic and nucleoplasmic contaminants. Besides ER membranes (observed as single membrane vesicles by EM in NE preparations; see supplemental material in Schirmer et al., 2003), expected contaminants include mitochondria, cytoplasmic filaments, and some highly charged chromatin proteins that are very "sticky" and resistant to extraction with high salt.

To prepare MMs (Gerace et al., 1982; Scheele, 1983), nuclei are first removed from a cell homogenate by low-speed centrifugation. Due to the fact that many mitochondria will saturate the subsequent sucrose cushions, the post-nuclear supernatant is subjected to an intermediate-speed centrifugation step to pellet most mitochondria. The resulting supernatant is mixed with sucrose and layered to float membranes during the next centrifugation step. Different types of membranes will separate into the different density layers of sucrose: the MM fraction is recovered, the sucrose diluted, and the membranes pelleted at high speed.

3.2. MudPIT Proteomics

Multi-dimensional Protein Identification Technology (MudPIT) incorporates several liquid chromatography steps with tandem mass spectrometry to identify proteins in a complex biochemical fraction (Washburn et al., 2001; Wolters et al., 2001). Alternating reverse-phase and cation exchange steps provides separation of the enormous complexity of peptides generated by direct digestion of the fraction without prior separation of individual proteins. Among the advantages that this provides is that proteins which migrate too closely on gels or other isolation platforms can still be identified. Moreover, proteins of low abundance are more likely to be distinguished. These benefits are especially important for membrane proteins, which often have extreme isoelectric points that limit effective resolution on 2D gels (Santoni et al., 2000). As peptides are not linked to their corresponding protein band when using MudPIT, protein identifications must be made for each individual peptide. To accomplish this, tandem mass spectrometry is utilized in which each analyzed peptide is subsequently fragmented and the masses of the smaller peptides generated are also measured. In many cases this combined footprint can be used to match all the fragments within the full-length peptide with a unique sequence in the genome databases, thus enabling protein identification from a single peptide. Since a significant number of peptides analyzed do not produce identifications, the method is not quantitative alone; however, combining this method with analysis of metabolically labeled fractions can add information on abundance (MacCoss et al., 2005).

3.3. Subtraction

Rodent liver NEs and MMs were extracted with 0.1 M NaOH to enrich for membrane proteins. In the NE fraction 566 proteins were identified, while 652 proteins were identified in the MMs (Schirmer et al., 2003). 41% of proteins in the NE fraction were also observed in the MM fraction; therefore, according to the subtractive methodology, these proteins were disregarded. Because the well-characterized NETs observed in the earlier proteomics study did not all fractionate in extractions as expected (Dreger et al., 2001), a separate NE fraction was extracted with salt and detergent to enrich for proteins associated with the lamin polymer. This identified 1830 proteins, and again any also appearing in the MM fraction were disregarded. Proteins remaining in the NE dataset were further analyzed for probability of transmembrane spanning segments by computer algorithm. Most algorithms used for predicting transmembrane spanning segments measure alpha helices with a hydrophobic character. However, membrane-spanning segments could also be simple hydrophobic stretches, multimers of beta barrels (Tamm et al., 2004), or hydrophobic "domes" (McKinney and Cravatt, 2004). Algorithms for such predictions have evolved considerably in the last few years. At the time of the study, the dataset was only interrogated for hydrophobic helices and the available algorithms yielded different results. Because of this, an algorithm was

selected that compares proteins to a database of well-characterized membrane-spanning segments (TMPred: www.ch.embnet.org/software/TMPRED_form.html) and parameters were selected that resulted in correct predictions for the previously characterized NETs.

After eliminating previously characterized proteins, 67 putative new NETs were identified between the NaOH extracted and salt/ detergent extracted NE fractions. These included several novel proteins related to previously characterized NETs: the third member of the Syne/ Nesprin family, NET53 (later called Nesprin 3; Wilhelmsen et al., 2005); a homolog of LAP1, NET9 (later called LULL1; Goodchild and Dauer, 2005); and two new LEM domain proteins, NET25 (later called Lem2; Brachner et al., 2005; Lee and Wilson, 2004) and NET66.

4. CONFIRMATION OF NUCLEAR ENVELOPE RESIDENCE

For simplicity we will refer to these proteins collectively as NETs, although until they are confirmed at the NE they are in fact "putative" NETs. However, determination of targeting to the NE is not always trivial.

4.1. Determination of Subcellular Localization

The simplest method to test for NE residence is to transfect the NET as a fusion with an epitope tag into tissue culture cells and visualize by fluorescence microscopy. Concentration at the nuclear periphery would be considered confirmation of NE targeting. However, overexpression tends to result in accumulation in the ER as well as the NE. This is the case even for well-characterized NETs for which antibodies recognize only the nuclear rim, and is presumably due to saturation of binding sites at the NE (Ellenberg et al., 1997; Soullam and Worman, 1993). If a putative NET is normally expressed at a low level and requires a limiting NE binding partner for targeting, then distinguishing the staining at the nuclear periphery against the high background in the ER may be difficult.

Many characterized INM proteins have been shown to interact with lamins (Ye et al., 1998). For the subset of putative NETs that share this characteristic, a more definite result can be achieved. As previously discussed, the lamin polymer is insoluble to extraction with detergent and salt; therefore retention of NETs in cells extracted with detergent (*e.g.* 0.5% Triton X-100) prior to fixation for microscopy confirms their direct, or indirect, association with the lamin polymer and NE localization. However, loss of the protein to a pre-extraction with detergent would occur for proteins not tethered to the lamin polymer whether they are normally localized to the INM, ONM, or ER. All eight of the putative NETs originally tested targeted to the NE, but only five remained at the NE after the detergent pre-extraction (Schirmer et al., 2003).

Epitope tags produce false-negatives when fusion to the tag blocks access to sequences required for proper targeting. Many NE proteins have large nucleoplasmic regions that bind to both chromatin proteins and lamins and this binding appears to be responsible for their accumulation in the NE. Thus fusion of a large tag in a nucleoplasmic region could block access to these binding partners and cause mistargeting. This appears to be the case for LAP1 and LAP2, which both have large nucleoplasmic domains: chimeric proteins failed to target properly when GFP was fused to the amino-terminus of either protein, whilst GFP fusions to the carboxyl-terminus were correctly localized to the NE.

Although epitope tags may produce artifacts, so may antibodies and several NE proteins have been reported to have significant problems from epitope masking. In a study using three different lamin B1 antibodies on heart and hippocampus sections, one stained just cardiomyoctyes, another stained just hippocampus, and the third stained both (Tunnah et al., 2005). As each antibody recognized a different part of the protein, this argues that different epitopes are masked in different tissues. In the case of emerin, six different monoclonal antibodies that recognize the protein by immunofluorescence in other cell types failed to recognize emerin in spleen, even though it was clearly present by Western blot analysis

(Tunnah et al., 2005). Such extreme epitope masking for emerin may be understandable in light of the many partners that have been mapped to bind throughout its length (Bengtsson and Wilson, 2004).

4.2. Determination of Membrane Insertion

Soluble binding partners of NETs or lamins could also concentrate at the NE. Therefore full confirmation of these proteins as NETs requires also clear determination of membrane insertion. As mentioned above (sect. 3.3), prediction algorithms have improved considerably since this computational analysis was undertaken, yet different algorithms still produce contradictory predictions. Thus, some of the putative NETs may not be integral membrane proteins and all should be directly tested. A simple test is to parallel extraction characteristics with well-defined integral membrane proteins; however this proved problematic in the comparative proteomics study (Sect. 2.1; Dreger et al., 2001). Moreover, the NETs identified in the subtractive liver study had differing extraction characteristics: of the eight new NETs for which targeting to the NE was initially determined, only one appeared in both NaOH and salt/detergent extracted fractions, and, although the NaOH-extraction is considered better for isolation of transmembrane proteins, all three proteins that appeared in only the salt/detergent extracted fraction targeted to the NE (Schirmer et al., 2003). The best test for membrane insertion is the protection provided by membranes against proteases *e.g.* protein sequences in the lumen are resistant to proteolytic digestion. To test this, however, requires either multiple epitope tags or a panel of antibodies to the protein, which may not be readily available for the large set of proteins identified in a proteomics dataset.

Although the predictions generated by computer algorithm must be confirmed by direct testing for individual NETs, in aggregate they can provide important insights into a large dataset. For example, it is noteworthy that eight of the 67 putative NETs were predicted to have signal peptides, yet nuclear retention signals are frequently predicted to be signal peptides (for more information see http://www.cbs.dtu.dk/services/SignalP/). Three of these eight were tested for targeting with a small N-terminal hemagglutinin (HA) epitope tag: all three targeted to the NE, indicating that the predicted signal peptides were not cleaved and are instead NE retention signals or regular transmembrane segments. Moreover, all three resisted a pre-extraction with detergent, indicating that they occur in the INM. This finding is particularly interesting as it suggests most or all NETs are type II transmembrane proteins.

Another interesting aspect of this dataset comes from the topology predictions of the transmembrane algorithms. With the exception of gp210, the principal mass of the originally characterized NETs occurs in the nucleoplasm. However, many among this original set of NETs were identified on the basis of their interactions with the lamina; so a concentration in the nucleoplasm is not surprising. Within the set of novel putative NETs, where identification was based separately on membrane (NaOH-extracted fraction) and on lamina association (salt/ detergent extracted fraction), the topological predictions indicate that there are as many NETs with their principal mass in the lumen as in the nucleoplasm (Figure 3).

4.3. Additional Complications

Several other issues may further complicate determination of NE targeting. One of these is the existence of splice variants for many NETs. For example, the LAP2 gene produces at least seven separate mRNAs (Berger et al., 1996): most have a transmembrane segment, but in what cell types each is produced and whether all target to the NE has not been clarified. The alpha splice variant, which is soluble, concentrates both at the nuclear rim and in specific locations in the nuclear interior (Dechat et al., 1998). The Syne/ Nesprin family also has many splice variants, some of which do not target to the NE (Gough et al., 2003). Evidence of multiple splice isoforms for 18 out of the 67 novel putative NETs is present in the cDNA and EST databases. Some splice variants have different numbers of predicted membrane-spanning segments and large sequence blocks from other exons, which could alter

accessibility to or affinity for targeting sites in the NE. Thus, some splice variants could target to other cellular locations.

NE retention in most studied cases has been found to require a lamin-binding site (Soullam and Worman, 1993). However, NETs such as emerin have been shown to bind other NETs (Bengtsson and Wilson, 2004) and in *C. elegans* the NET Unc83 is dependent upon the NET Unc84 for its targeting to the NE (Malone et al., 1999). Unc84 itself is dependent upon lamin A (Lee et al., 2002). Thus if a particular partner protein is absent from a specific cell type, a NET could target to the NE in one cell type but not another.

5. VARIATION IN THE NUCLEAR ENVELOPE PROTEOME

Nearly all of the NETs identified prior to the subtractive study appeared to be ubiquitously expressed: the only exception was UNCL (Fitzgerald et al., 2000). As no message for UNCL was detected in liver by Northern analysis, it was not surprising that UNCL did not appear in the liver proteomics dataset (Schirmer et al., 2003; Schirmer and Gerace, 2005). The surprising number of NETs identified in the subtractive proteomics study raises the possibility that some might be cell-type specific because liver is a composite of several different cell types including hepatocytes, kuppfer cells, lipocytes, and endothelial cells contributed by its extensive vasculature.

5.1. Tissue Variation: The Transcriptome Database

The relative expression levels of NET mRNAs in different tissues were obtained from the transcriptome database of the Genome Institute of the Novartis Research Foundation. This DNA array dataset was generated using mRNAs isolated from 72 human and 61 mouse tissues and cell types (Su et al., 2002). Experiments were standardized within this large study so that the level of expression in one tissue could be compared to another through the numerical values given for each message in their respective tissues. When the highest numerical value given for a particular NET across the range of tissues was divided by the lowest numerical value given for that NET, a low ratio reflects ubiquitous expression of the NET at similar levels: a high ratio either indicates a wide range of expression levels or its absence in a subset of tissues. Of the 53 putative NETs that were encoded on the Affymetrix chips used, less than one fifth were uniformly expressed with lower than 6-fold variation among different tissues (Figure 4A).

When the highest value was instead divided by the median value, high ratios were also observed (Figure 4B). Here a high ratio is inferred to indicate a significant preference for expression in a subset of tissues or cell types. One third of the NETs yielded ratios greater than 5; thus, the tissues exhibiting highest expression have levels at least five-fold higher than the majority of tissues. The cell types where the highest expression levels occurred for NETs from the rodent liver dataset tended to be those present in the starting material *i.e.* liver, adipocytes (similar to lipoctyes), and blood cells (Schirmer et al., 2005; Schirmer and Gerace, 2005). These observations support the idea that a subset of NE proteins will be ubiquitous while another subset will be unique to each tissue investigated: for example, if each major cell type in liver contributed 5 unique NETs, there would be 50 core NETs and 30 tissue-specific NETs among the 80 (67 novel + 13 previously characterized) identified in rodent liver. This may also help to explain why the comparative study that used a single cell line (Dreger et al., 2001) identified a much smaller number of NETs compared to the subtractive liver study (Schirmer et al., 2003).

Surprisingly, the transcriptome data also indicates considerable variation in expression levels among different cell types for several of the previously characterized NETs (Figure 4C and D). Nonetheless these data suggest that the original NETs are more ubiquitous in their expression than the new NETs because their average levels of expression are much higher: using numbers from the relative values given for expression in the transcriptome data the average lowest levels of expression were roughly 2-fold higher in the

original NETs compared to the new NETs (161 and 83 respectively), while the average highest levels were more than 4-fold higher (4,797 *vs* 1,110). This is also consistent with the idea that their earlier identification was facilitated by their relative abundance. Although apparently ubiquitously expressed, LBR was so highly expressed in hematopoietic cells that its ratio for variation from highest to lowest expressing tissue was in excess of 100.

5.2. Evolutionary Variation

In addition to this variation between different tissues within an organism, the NE proteome appears to be varied between organisms. Blast searches were performed on all 80 NETs, and those with detectable homologs in several completely sequenced genomes are shown in Figure 5. Although the frequent annotations to the genome databases can cause fluxuations in the results for individual NETs, in aggregate the results are striking. Eight NETs appeared to be universally conserved from mammals to yeast. In contrast, seven other NETs were not conserved even between mouse, rat, and human (though this may reflect mistakes in the databases or non-equivalent prediction of hypothetical orfs). Moreover, there is no particular phylogenetic conservation for different NETs as some had yeast homologs, but neither plant nor fly homologs. Some observations were nonetheless striking and may reflect some phylogenetic grouping. For example, 17 NETs were conserved between mammals and flies, but were not found in worms; yet only 6 NETs conserved between mammals and worms did not appear in flies. NET variation between organisms was first highlighted because two proteins identified in Drosophila, Otefin (Ashery-Padan et al., 1997) and YA (Goldberg et al., 1998), have no mammalian homologs and most of the 13 previously characterized mammalian NETs do not have homologs in *Drosophila* (Gruenbaum et al., 2003).

6. NETS IN MULTIPLE CELLULAR LOCATIONS

The central limitation of the subtractive approach is that it identifies only proteins that are unique to the NE: it will disregard proteins that normally occur in both of the fractions analyzed. One such protein is Sec13, an ER trafficking protein that also has a demonstrated function in NPC mediated transport (Siniossoglou et al., 1996). Sec13 thus appeared in both datasets and was accordingly discarded from the NE-unique *in silico* purified protein set. TorsinA also appeared in both datasets and was discarded, but it has since been shown to shuttle between the INM and the ER (Goodchild and Dauer, 2004). As 41% of the 566 proteins identified in the membrane-enriched NE fraction were similarly discarded, there may be many other proteins listed in the pre-subtraction dataset that are truly NE proteins, though they are not unique to the NE.

NETs might also share cellular locations besides ER. The subtractive study focussed on identifying novel NETs, but several transmembrane proteins were identified that had been previously characterized in other cellular compartments (see supplemental material in Schirmer et al., 2003). These range from a plasma membrane receptor (Klein et al., 2004) to an ion transport pump (Arteaga et al., 2004) to mediators of signalling cascades such as the inositol triphosphate receptor (Cruttwell et al., 2005) and a cytosolic phospholipase A2 variant that relocates from the ER to the NE under certain conditions. Several of these have been shown by various means (immuno-EM, immunofluorescence microscopy, Western blotting analysis of subcellular fractions) to target to the NE. Although all of those shown in Table 1 have been tested for localization, typically only one antibody was used so the possibility of cross-reactivity cannot be ruled out. That not all the proteins included in the table were found in the subtractive liver NE study may be explained if they only occur at the NE in certain cell types. This was clearly shown in the case of the NCX1 ion transporter which was detected in brain and lymphocytes, but not in other tissues (Xie et al., 2004). The appearance of these proteins in the NE, though unexpected, makes sense. For example, in the case of ion transporters, the nucleus needs to maintain its internal ion concentration as much

as the cytoplasm, and the critical ion concentrations needed for chromatin-modifying enzymes are not the same as those for cytoplasmic enzymes. It is also logical for the NE to have additional mechanisms for signalling between the nucleus and cytoplasm besides the NPC. Thus clarification that other proteins in this proteomics dataset share multiple localizations may add more varied functions to the NE milieu.

7. NUCLEAR ENVELOPE DISEASES AND FUNCTION

Nearly 20 diverse inherited diseases and syndromes are now clearly linked to the NE (Mounkes and Stewart, 2004; Worman and Courvalin, 2002). These include muscular dystrophies (Bione et al., 1994; Bonne et al., 1999; Hanisch et al., 2002; Muchir et al., 2000; Raffaele Di Barletta et al., 2000), lipodystrophies (Cao and Hegele, 2000; Shackleton et al., 2000), neuropathy (De Sandre-Giovannoli et al., 2002), cardiomyopathies (Fatkin et al., 1999), dermopathies (Navarro et al., 2004), bone diseases (Hellemans et al., 2004; Waterham et al., 2003), and the aging disease progeria (De Sandre-Giovannoli et al., 2003; Eriksson et al., 2003). NE proteins mutated in these diseases include the intermediate filament lamin A/C protein and multiple NETs. In addition to diseases caused by proteins unique to the INM, mutations in torsinA which shuttles between the ER and INM cause the disease early-onset torsion dystonia and result in its accumulation in the NE (Goodchild and Dauer, 2004).

The combination of these diverse functions and the wide range of diseases associated with the NE suggest that fundamental regulatory mechanisms reside at the NE, the perturbation of which can have profound consequences for the cell. There are several hypotheses explaining how NE proteins could cause these diseases based on known functions for NE proteins. 1) Misregulation of gene expression could result in pathology because both lamins and NETs have been shown to influence transcription (Ellis et al., 1997; Moir et al., 2000; Nili et al., 2001; Spann et al., 1997). 2) Cell cycle misregulation could produce pathology in several ways. Both withdrawal from the cell cycle of a regenerating stem cell population and activation of the cell cycle in a differentiated population could produce pathology over time. Lamin interactions with Rb (Ozaki et al., 1994), DNA replication (Kennedy et al., 2000; Martins et al., 2003), and signal cascades (Steen et al., 2000) as well as NET influences on centriole positioning (Malone et al., 1999) could all affect the cell cycle. 3) Loss of mechanical/ structural stability in the lamina could weaken cells and produce pathology. Lamins and NETs have both been shown to impact on nuclear morphology (Hoffmann et al., 2002; Liu et al., 2000; Schirmer et al., 2001) and loss of lamin A has been shown to affect the mechanical stability of the lamina *in vivo* (Lammerding et al., 2004), *in vitro* (Schirmer and Gerace, 2004), and in disease (Fidzianska et al., 1998). None of these hypotheses can satisfactorily explain the wide range of tissue specificities for NE diseases. For example, while the mechanical stability hypothesis makes sense for muscular dystrophy, it is hard to see how it could apply to diseases affecting neurons or fat cells. These hypotheses also fail to explain how two proximally located mutations in the same protein can cause in one case neuropathy and in another muscular dystrophy.

The data presented above indicating that many NETs are preferentially expessed in certain tissues suggests an alternative hypothesis: the tissue specificity of NE diseases derives from tissue-specific NETs having specifically altered interactions with the ubiquitous NE proteins mutated in disease. This would result in selective dysfunction of multiprotein complexes in those tissues, which could themselves have functions in replication, gene expression, cell cycle, etc. The case of the NETs Unc83 and Unc84 (Sect. 4.3) provides an example of this. Unc83 is specific to *C. elegans* tissues where nuclear migration occurs, and it must interact with the widely expressed Unc84 to accumulate in the NE (Starr et al., 2001). Mutations in either interacting protein result in the same uncoordinated (*unc*) phenotype. Furthermore, Unc84 requires lamins for its targeting to the NE. This resembles human lamin A and emerin in that emerin requires lamin A to accumulate at the NE and mutations in either protein cause variants of Emery-Dreifuss muscular dystrophy. However both proteins are widely expressed and the relevant tissue-specific partner (such as Unc83) has not yet been

clearly identified, though new binding partners of emerin are being rapidly discovered (Sect. 4.1). The recent explosion in the number of proteins assigned to the NE opens a wide field of possible candidate partners, some of which may themselves cause additional NE diseases.

8. CONCLUSIONS

Many more proteomics datasets of subcellular organelles and compartments generated from different cell types will be required before the full extent of variability for each organelle can be determined. The unexpectedly high degree of variation indicated in the liver NE suggests that other organelles may also exhibit greater variation than previously thought. Currently our laboratory is in the process of analyzing nuclear envelopes from other cell types in an effort to address the former question. The latter question will require analysis of other organelles from multiple cell types.

ACKNOWLEDGEMENTS

We thank Dr. Gavin Wilkie for critical reading of the manuscript. This work was supported by a Dorothy Hodgkin Royal Society Fellowship to P. Malik and a Senior Research Fellowship to E. Schirmer from the Wellcome Trust.

REFERENCES

Aaronson, R. P., and Blobel, G. (1975) Isolation of nuclear pore complexes in association with a lamina. Proc Natl Acad Sci U S A 72: 1007-1011.

Apel, E. D., Lewis, R. M., Grady, R. M., and Sanes, J. R. (2000) Syne-1, a dystrophin- and Klarsicht-related protein associated with synaptic nuclei at the neuromuscular junction. J Biol Chem 275: 31986-31995.

Arteaga, M. F., Gutierrez, R., Avila, J., Mobasheri, A., Diaz-Flores, L., and Martin-Vasallo, P. (2004) Regeneration influences expression of the Na+, K+-atpase subunit isoforms in the rat peripheral nervous system. Neuroscience 129: 691-702.

Ashery-Padan, R., Weiss, A. M., Feinstein, N., and Gruenbaum, Y. (1997) Distinct regions specify the targeting of otefin to the nucleoplasmic side of the nuclear envelope. J Biol Chem 272: 2493-2499.

Bengtsson, L., and Wilson, K. L. (2004) Multiple and surprising new functions for emerin, a nuclear membrane protein. Curr Opin Cell Biol 16: 73-79.
— This review provides a comprehensive view of the many binding partners identified to date for emerin and supports the idea that some NETs may serve more as assembly points for functional complexes rather than having any inherent enzymatic activity.

Berger, R., Theodor, L., Shoham, J., Gokkel, E., Brok-Simoni, F., Avraham, K. B., Copeland, N. G., Jenkins, N. A., Rechavi, G., and Simon, A. J. (1996) The characterization and localization of the mouse thymopoietin/lamina- associated polypeptide 2 gene and its alternatively spliced products. Genome Res 6: 361-370.

Bione, S., Maestrini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G., and Toniolo, D. (1994) Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. Nat Genet 8: 323-327.
— This was the first link between a nuclear envelope protein and disease, although it was not until Manilal et al., in 1996 that emerin was recognized to be a nuclear envelope protein.

Bkaily, G., Nader, M., Avedanian, L., Jacques, D., Perrault, C., Abdel-Samad, D., D'Orleans-Juste, P., Gobeil, F., and Hazzouri, K. M. (2004) Immunofluorescence revealed the presence of NHE-1 in the nuclear membranes of rat cardiomyocytes and isolated nuclei of human, rabbit, and rat aortic and liver tissues. Can J Physiol Pharmacol 82: 805-811.

Blobel, G., and Potter, V. R. (1966) Nuclei from rat liver: isolation method that combines purity with high yield. Science 154: 1662-1665.

Bonne, G., Di Barletta, M. R., Varnous, S., Becane, H. M., Hammouda, E. H., Merlini, L., Muntoni, F., Greenberg, C. R., Gary, F., Urtizberea, J. A., *et al.* (1999) Mutations in the gene encoding lamin A/C cause autosomal dominant Emery- Dreifuss muscular dystrophy. Nat Genet 21: 285-288.
— This second association of a nuclear envelope protein with disease opened the door to an entirely new appreciation of the nuclear envelope and its functions. It provided the basis for the "guilt by association" approach, which served to identify many subsequent nuclear envelope diseases.

Brachner, A., Reipert, S., Foisner, R., and Gotzmann, J. (2005) LEM2 is a novel MAN1-related inner nuclear membrane protein associated with A-type lamins. J Cell Sci 118: 5797-5810.

Brown, D. T. (2001) Histone variants: are they functionally heterogeneous? Genome Biol 2: REVIEWS0006.

Cao, H., and Hegele, R. A. (2000) Nuclear lamin A/C R482Q mutation in Canadian kindreds with Dunnigan- type familial partial lipodystrophy. Hum Mol Genet 9: 109-112.

Crisp, M., Liu, Q., Roux, K., Rattner, J. B., Shanahan, C., Burke, B., Stahl, P. D., and Hodzic, D. (2006) Coupling of the nucleus and cytoplasm: role of the LINC complex. J Cell Biol 172: 41-53.
— Here data is presented suggesting that there are linkages across the lumen of the nuclear envelope that help to maintain the distance between the outer and inner nuclear membrane.

Cronshaw, J., Krutchinsky, A., Zhang, W., Chait, B., and Matunis, M. (2002) Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol 158: 915-927.

Cruttwell, C., Bernard, J., Hilly, M., Nicolas, V., Tunwell, R. E., and Mauger, J. P. (2005) Dynamics of the inositol 1,4,5-trisphosphate receptor during polarisation of MDCK cells. Biol Cell.

De Sandre-Giovannoli, A., Bernard, R., Cau, P., Navarro, C., Amiel, J., Boccaccio, I., Lyonnet, S., Stewart, C., Munnich, A., Le Merrer, M., and Levy, N. (2003) Lamin A Truncation in Hutchinson-Gilford Progeria. Science.

De Sandre-Giovannoli, A., Chaouch, M., Kozlov, S., Vallat, J., Tazir, M., Kassouri, N., Szepetowski, P., Hammadouche, T., Vandenberghe, A., Stewart, C., *et al.* (2002) Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope proteins, cause autosomal recessive axonal neuropathy in human (Charcot-Marie-Tooth disorder type 2) and mouse. Am J Hum Genet 70: 726-736.

Dechat, T., Gotzmann, J., Stockinger, A., Harris, C. A., Talle, M. A., Siekierka, J. J., and Foisner, R. (1998) Detergent-salt resistance of LAP2alpha in interphase nuclei and phosphorylation-dependent association with chromosomes early in nuclear assembly implies functions in nuclear structure dynamics. Embo J 17: 4887-4902.

Dhe-Paganon, S., Werner, E. D., Chi, Y.-I., and Shoelson, S. E. (2002) Structure of the globular tail of nuclear lamin. J Biol Chem 277: 17381-17384.

Dreger, M., Bengtsson, L., Schoneberg, T., Otto, H., and Hucho, F. (2001) Nuclear envelope proteomics: novel integral membrane proteins of the inner nuclear membrane. Proc Natl Acad Sci U S A 98: 11943-11948.
— The first use of proteomics to identify novel inner nuclear membrane proteins identified most of the previously characterized proteins as well as two additional splice variants of LAP2, and two novel NETs Unc84A and LUMA.

Dwyer, N., and Blobel, G. (1976) A modified procedure for the isolation of a pore complex-lamina fraction from rat liver nuclei. J Cell Biol 70: 581-591.

Ellenberg, J., Siggia, E. D., Moreira, J. E., Smith, C. L., Presley, J. F., Worman, H. J., and Lippincott-Schwartz, J. (1997) Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. J Cell Biol 138: 1193-1206.

Ellis, D. J., Jenkins, H., Whitfield, W. G., and Hutchison, C. J. (1997) GST-lamin fusion proteins act as dominant negative mutants in Xenopus egg extract and reveal the function of the lamina in DNA replication. J Cell Sci 110: 2507-2518.

Erickson, E. S., Mooren, O. L., Moore-Nichols, D., and Dunn, R. C. (2004) Activation of ryanodine receptors in the nuclear envelope alters the conformation of the nuclear pore complex. Biophys Chem 112: 1-7.

Eriksson, M., Brown, W. T., Gordon, L. B., Glynn, M. W., Singer, J., Scott, L., Erdos, M. R., Robbins, C. M., Moses, T. Y., Berglund, P., *et al.* (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. Nature 423: 293-298.

Fatkin, D., MacRae, C., Sasaki, T., Wolff, M., Porcu, M., Frenneaux, M., Atherton, J., Vidaillet, H. J., Spudich, S., De Girolami, U., *et al.* (1999) Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. N Engl J Med 341: 1715-1724.

Fidzianska, A., Toniolo, D., and Hausmanowa-Petrusewicz, I. (1998) Ultrastructural abnormality of sarcolemmal nuclei in Emery-Dreifuss muscular dystrophy (EDMD). J Neurol Sci 159: 88-93.
— This first demonstrated that peripheral chromatin distribution is altered in cells derived from patients with nuclear envelope diseases. Chromatin effects are a favored molecular mechanism for how pathology developes.

Fields, A. P., Pettit, G. R., and May, W. S. (1988) Phosphorylation of lamin B at the nuclear membrane by activated protein kinase C. J Biol Chem 263: 8253-8260.

Fitzgerald, J., Kennedy, D., Viseshakul, N., Cohen, B. N., Mattick, J., Bateman, J. F., and Forsayeth, J. R. (2000) UNCL, the mammalian homologue of UNC-50, is an inner nuclear membrane RNA-binding protein. Brain Res 877: 110-123.
— The tissue-restricted expression of this NET was the first indication that there is considerable variability in the nuclear envelope proteome.

Foisner, R., and Gerace, L. (1993) Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. Cell 73: 1267-1279.

Gerace, L., Blum, A., and Blobel, G. (1978) Immunocytochemical localization of the major polypeptides of the nuclear pore complex-lamina fraction. Interphase and mitotic distribution. J Cell Biol 79: 546-566.

Gerace, L., and Burke, B. (1988) Functional organization of the nuclear envelope. Annu Rev Cell Biol 4: 335-374.

Gerace, L., Ottaviano, Y., and Kondor-Koch, C. (1982) Identification of a major polypeptide of the nuclear pore complex. J Cell Biol 95: 826-837.

Goldberg, M., Lu, H., Stuurman, N., Ashery-Padan, R., Weiss, A. M., Yu, J., Bhattacharyya, D., Fisher, P. A., Gruenbaum, Y., and Wolfner, M. F. (1998) Interactions among Drosophila nuclear envelope proteins lamin, otefin, and YA. Mol Cell Biol 18: 4315-4323.
Goodchild, R. E., and Dauer, W. T. (2004) Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. Proc Natl Acad Sci U S A 101: 847-852.
Goodchild, R. E., and Dauer, W. T. (2005) The AAA+ protein torsinA interacts with a conserved domain present in LAP1 and a novel ER protein. J Cell Biol 168: 855-862.
Gough, L. L., Fan, J., Chu, S., Winnick, S., and Beck, K. A. (2003) Golgi localization of Syne-1. Mol Biol Cell 14: 2410-2424.
Greber, U. F., Senior, A., and Gerace, L. (1990) A major glycoprotein of the nuclear pore complex is a membrane-spanning polypeptide with a large lumenal domain and a small cytoplasmic tail. Embo J 9: 1495-1502.
Grewal, S., Herbert, S. P., Ponnambalam, S., and Walker, J. H. (2005) Cytosolic phospholipase A2-alpha and cyclooxygenase-2 localize to intracellular membranes of EA.hy.926 endothelial cells that are distinct from the endoplasmic reticulum and the Golgi apparatus. Febs J 272: 1278-1290.
Gruenbaum, Y., Goldman, R. D., Meyuhas, R., Mills, E., Margalit, A., Fridkin, A., Dayani, Y., Prokocimer, M., and Enosh, A. (2003) The nuclear lamina and its functions in the nucleus. Int Rev Cytol 226: 1-62.
Hahn, C. G., and Covault, J. (1990) Isolation of transcriptionally active nuclei from striated muscle using Percoll density gradients. Anal Biochem 190: 193-197.
Hallberg, E., Wozniak, R. W., and Blobel, G. (1993) An integral membrane protein of the pore membrane domain of the nuclear envelope contains a nucleoporin-like region. J Cell Biol 122: 513-521.
Hanisch, F., Neudecker, S., Wehnert, M., and Zierz, S. (2002) Hauptmann-Thannhauser muscular dystrophy and differential diagnosis of myopathies associated with contractures. Nervenarzt 73: 1004-1011.
Held, I. R., Rodrigo, R. T., Yeoh, H. C., and Tonaki, H. (1977) Isolation of nuclei from red and white skeletal muscles of the adult rat. Exp Cell Res 105: 191-197.
Hellemans, J., Preobrazhenska, O., Willaert, A., Debeer, P., Verdonk, P. C., Costa, T., Janssens, K., Menten, B., Van Roy, N., Vermeulen, S. J., *et al.* (2004) Loss-of-function mutations in LEMD3 result in osteopoikilosis, Buschke-Ollendorff syndrome and melorheostosis. Nat Genet 36: 1213-1218.
Hnasko, R., and Ben-Jonathan, N. (2005) Developmental regulation of PV-1 in rat lung: association with the nuclear envelope and limited colocalization with Cav-1. Am J Physiol Lung Cell Mol Physiol 288: L275-284.
Hoffmann, K., Dreger, C., Olins, A., Olins, D., Shultz, L., Lucke, B., Karl, H., Kaps, R., Muller, D., Vaya, A., *et al.* (2002) Mutations in the gene encoding the lamin B receptor produce an altered nuclear morphology in granulocytes (Pelger-Huet anomaly. Nat Genet 31: 410-414.
Kennedy, B. K., Barbie, D. A., Classon, M., Dyson, N., and Harlow, E. (2000) Nuclear organization of DNA replication in primary mammalian cells. Genes Dev 14: 2855-2868.
Klein, C., Gensburger, C., Freyermuth, S., Nair, B. C., Labourdette, G., and Malviya, A. N. (2004) A 120 kDa nuclear phospholipase Cgamma1 protein fragment is stimulated in vivo by EGF signal phosphorylating nuclear membrane EGFR. Biochemistry 43: 15873-15883.
Krimm, I., Ostlund, C., Gilquin, B., Couprie, J., Hossenlopp, P., Mornon, J., Bonne, G., Courvalin, J., Worman, H., and Zinn-Justin, S. (2002) The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy. Structure (Camb) 10: 811-823.
Kuehl, L. (1977) Isolation of skeletal muscle nuclei. Methods Cell Biol 15: 79-88.
Lammerding, J., Schulze, P., Takahashi, T., Kozlov, S., Sullivan, T., Kamm, R., Stewart, C., and Lee, R. (2004) Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. J Clin Invest 113: 370-378.
Lee, H., Habas, R., and Abate-Shen, C. (2004) MSX1 cooperates with histone H1b for inhibition of transcription and myogenesis. Science 304: 1675-1678.
Lee, K., Starr, D., Cohen, M., Liu, J., Han, M., Wilson, K., and Gruenbaum, Y. (2002) Lamin-dependent localization of UNC-84, a protein required for nuclear migration in Caenorhabditis elegans. Mol Biol Cell 13: 892-901.
Lee, K. K., and Wilson, K. L. (2004) All in the family: evidence for four new LEM-domain proteins Lem2 (NET-25), Lem3, Lem4 and Lem5 in the human genome. Symp Soc Exp Biol: 329-339.

Lin, F., Blake, D. L., Callebaut, I., Skerjanc, I. S., Holmer, L., McBurney, M. W., Paulin-Levasseur, M., and Worman, H. J. (2000) MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. J Biol Chem 275: 4840-4847.

Liu, J., Ben-Shahar, T., Riemer, D., Treinin, M., Spann, P., Weber, K., Fire, A., and Gruenbaum, Y. (2000) Essential roles for Caenorhabditis elegans lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. Mol Biol Cell 11: 3937-3947.

MacCoss, M. J., Wu, C. C., Matthews, D. E., and Yates, J. R., 3rd (2005) Measurement of the isotope enrichment of stable isotope-labeled proteins using high-resolution mass spectra of peptides. Anal Chem 77: 7646-7653.

Malone, C. J., Fixsen, W. D., Horvitz, H. R., and Han, M. (1999) UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development. Development 126: 3171-3181.

Manilal, S., Nguyen, T. M., Sewry, C. A., and Morris, G. E. (1996) The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. Hum Mol Genet 5: 801-808.

Mansharamani, M., Hewetson, A., and Chilton, B. S. (2001) Cloning and characterization of an atypical Type IV P-type ATPase that binds to the RING motif of RUSH transcription factors. J Biol Chem 276: 3641-3649.

Martins, S., Eikvar, S., Furukawa, K., and Collas, P. (2003) HA95 and LAP2 beta mediate a novel chromatin-nuclear envelope interaction implicated in initiation of DNA replication. J Cell Biol 160: 177-188.

McKinney, M. K., and Cravatt, B. F. (2004) Structure and Function of Fatty Acid Amide Hydrolase. Annu Rev Biochem.

Moir, R., Spann, T., Herrmann, H., and Goldman, R. (2000) Disruption of nuclear lamin organization blocks the elongation phase of DNA replication. J Cell Biol 149: 1179-1192.

Mosley-Bishop, K. L., Li, Q., Patterson, L., and Fischer, J. A. (1999) Molecular analysis of the klarsicht gene and its role in nuclear migration within differentiating cells of the Drosophila eye. Curr Biol 9: 1211-1220.

Mounkes, L., and Stewart, C. L. (2004) Structural organization and functions of the nucleus in development, aging, and disease. Curr Top Dev Biol 61: 191-228.

Muchir, A., Bonne, G., van der Kooi, A. J., van Meegen, M., Baas, F., Bolhuis, P. A., de Visser, M., and Schwartz, K. (2000) Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). Hum Mol Genet 9: 1453-1459.

Navarro, C. L., De Sandre-Giovannoli, A., Bernard, R., Boccaccio, I., Boyer, A., Genevieve, D., Hadj-Rabia, S., Gaudy-Marqueste, C., Smitt, H. S., Vabres, P., *et al.* (2004) Lamin A and ZMPSTE24 (FACE-1) defects cause nuclear disorganization and identify restrictive dermopathy as a lethal neonatal laminopathy. Hum Mol Genet 13: 2493-2503.

Nili, E., Cojocaru, G. S., Kalma, Y., Ginsberg, D., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Berger, R., Shaklai, S., Amariglio, N., *et al.* (2001) Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). J Cell Sci 114: 3297-3307.
— The finding that a NET can bind a transcriptional repressor provides a more detailed molecular mechanism for how the nuclear envelope could influence genome organization.

Ohba, T., Schirmer, E. C., Nishimoto, T., and Gerace, L. (2004) Energy- and temperature-dependent transport of integral proteins to the inner nuclear membrane via the nuclear pore. J Cell Biol 167: 1051-1062.
— The energy requirement observed indicates that lateral diffusion of integral proteins is regulated by a gated channel, and the inhibition of transport associated with increases in lumenal domains suggests that proteins cross this space to create a steric block.

Ozaki, T., Saijo, M., Murakami, K., Enomoto, H., Taya, Y., and Sakiyama, S. (1994) Complex formation between lamin A and the retinoblastoma gene product: identification of the domain on lamin A required for its interaction. Oncogene 9: 2649-2653.

Paulin-Levasseur, M., Blake, D. L., Julien, M., and Rouleau, L. (1996) The MAN antigens are non-lamin constituents of the nuclear lamina in vertebrate cells. Chromosoma 104: 367-379.

Raffaele Di Barletta, M., Ricci, E., Galluzzi, G., Tonali, P., Mora, M., Morandi, L., Romorini, A., Voit, T., Orstavik, K. H., Merlini, L., *et al.* (2000) Different Mutations in the LMNA Gene Cause Autosomal Dominant and Autosomal Recessive Emery-Dreifuss Muscular Dystrophy. Am J Hum Genet 66: 1407-1412.

Raju, G. P., Dimova, N., Klein, P. S., and Huang, H. C. (2003) SANE, a novel LEM domain protein, regulates bone morphogenetic protein signaling through interaction with Smad1. J Biol Chem 278: 428-437.

Rolls, M. M., Stein, P. A., Taylor, S. S., Ha, E., McKeon, F., and Rapoport, T. A. (1999) A visual screen of a GFP-fusion library identifies a new type of nuclear envelope membrane protein. J Cell Biol 146: 29-44.

Ronaldson, P. T., Bendayan, M., Gingras, D., Piquette-Miller, M., and Bendayan, R. (2004) Cellular localization and functional expression of P-glycoprotein in rat astrocyte cultures. J Neurochem 89: 788-800.

Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol 148: 635-651.

Santoni, V., Molloy, M., and Rabilloud, T. (2000) Membrane proteins and proteomics: un amour impossible? Electrophoresis 21: 1054-1070.

Scheele, G. (1983) Methods for the study of protein translocation across the RER membrane using the reticulocyte lysate translation system and canine pancreatic microsomal membranes. Methods Enzymol 96: 94-111.

Schirmer, E. C., Florens, L., Guan, T., Yates, J. R., 3rd, and Gerace, L. (2005). Identification of novel integral membrane proteins of the nuclear envelope with potential disease links using subtractive proteomics, In Novartis Found Symp No. 264 Nuclear Organization in Development and Disease, D. J. Chadwick, and J. A. Goode, eds. (Chichester and New York: John Wiley & Sons Ltd.), pp. 63-76; discussion 76-80, 227-230.

Schirmer, E. C., Florens, L., Guan, T., Yates, J. R. r., and Gerace, L. (2003) Nuclear membrane proteins with potential disease links found by subtractive proteomics. Science 301: 1380-1382.

— The use of a new subtractive proteomics methodology to identify 67 new putative NETs potentially expands the number of NETs by five times.

Schirmer, E. C., and Gerace, L. (2004) The stability of the nuclear lamina polymer changes with the composition of lamin subtypes according to their individual binding strengths. J Biol Chem.

Schirmer, E. C., and Gerace, L. (2005) The nuclear membrane proteome: extending the envelope. Trends Biochem Sci 30: 551-558.

— Further analysis of the subtractive proteomics nuclear envelope dataset, suggests that there is considerable variation in the NETs present in different tissues.

Schirmer, E. C., Guan, T., and Gerace, L. (2001) Involvement of the lamin rod domain in heterotypic lamin interactions important for nuclear organization. J Cell Biol 153: 479-489.

Senior, A., and Gerace, L. (1988) Integral membrane proteins specific to the inner nuclear membrane and associated with the nuclear lamina. J Cell Biol 107: 2029-2036.

Shackleton, S., Lloyd, D. J., Jackson, S. N., Evans, R., Niermeijer, M. F., Singh, B. M., Schmidt, H., Brabant, G., Kumar, S., Durrington, P. N., *et al.* (2000) LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. Nat Genet 24: 153-156.

Siniossoglou, S., Wimmer, C., Rieger, M., Doye, V., Tekotte, H., Weise, C., Emig, S., Segref, A., and Hurt, E. C. (1996) A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. Cell 84: 265-275.

Soullam, B., and Worman, H. J. (1993) The amino-terminal domain of the lamin B receptor is a nuclear envelope targeting signal. J Cell Biol 120: 1093-1100.

— This introduced the lateral diffusion hypothesis that NETs can freely diffuse between the ER and nuclear envelope, and indicated that they are retained in the nuclear envelope by binding to the lamin polymer.

Spann, T. P., Moir, R. D., Goldman, A. E., Stick, R., and Goldman, R. D. (1997) Disruption of nuclear lamin organization alters the distribution of replication factors and inhibits DNA synthesis. J Cell Biol 136: 1201-1212.

Starr, D. A., and Han, M. (2002) Role of ANC-1 in tethering nuclei to the actin cytoskeleton. Science 298: 406-409.

Starr, D. A., Hermann, G. J., Malone, C. J., Fixsen, W., Priess, J. R., Horvitz, H. R., and Han, M. (2001) unc-83 encodes a novel component of the nuclear envelope and is essential for proper nuclear migration. Development 128: 5039-5050.

Staufenbiel, M., and Deppert, W. (1982) Intermediate filament systems are collapsed onto the nuclear surface after isolation of nuclei from tissue culture cells. Exp Cell Res 138: 207-214.

Steen, R. L., Martins, S. B., Tasken, K., and Collas, P. (2000) Recruitment of protein phosphatase 1 to the nuclear envelope by A-kinase anchoring protein AKAP149 is a prerequisite for nuclear lamina assembly. J Cell Biol 150: 1251-1262.

Stuurman, N., Heins, S., and Aebi, U. (1998) Nuclear lamins: their structure, assembly, and interactions. J Struct Biol 122: 42-66.
— One of the most comprehensive reviews on the structural properties of lamins.

Su, A. I., Cooke, M. P., Ching, K. A., Hakak, Y., Walker, J. R., Wiltshire, T., Orth, A. P., Vega, R. G., Sapinoso, L. M., Moqrich, A., *et al.* (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 99: 4465-4470.
— This enormous undertaking has provided a map of expression levels for tens of thousands of genes in a wide range of tissues and cell types.

Sun, G., Yuen Chan, S., Yuan, Y., Wang Chan, K., Qiu, G., Sun, K., and Ping Leung, M. (2002) Isolation of differentially expressed genes in human heart tissues. Biochim Biophys Acta 1588: 241-246.

Tamm, L. K., Hong, H., and Liang, B. (2004) Folding and assembly of beta-barrel membrane proteins. Biochim Biophys Acta 1666: 250-263.

Tunnah, D., Sewry, C. A., Vaux, D., Schirmer, E. C., and Morris, G. E. (2005) The apparent absence of lamin B1 and emerin in many tissue nuclei is due to epitope masking. J Mol Histol 36: 337-344.
— The finding that multiple nuclear envelope proteins are differently epitope masked in different tissues argues that tissue-specific complexes form on these proteins.

Wagner, N., Schmitt, J., and Krohne, G. (2004) Two novel LEM-domain proteins are splice products of the annotated Drosophila melanogaster gene CG9424 (Bocksbeutel). Eur J Cell Biol 82: 605-616.

Washburn, M. P., Wolters, D., and Yates, J. R. r. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19: 242-247.

Waterham, H., Koster, J., Mooyer, P., Noort, G. G., Kelley, R., Wilcox, W., Wanders, R., Hennekam, R., and Oosterwijk, J. (2003) Autosomal Recessive HEM/Greenberg Skeletal Dysplasia Is Caused by 3beta-Hydroxysterol Delta14-Reductase Deficiency Due to Mutations in the Lamin B Receptor Gene. Am J Hum Genet 72: 1013-1017.

Wilhelmsen, K., Litjens, S. H., Kuikman, I., Tshimbalanga, N., Janssen, H., van den Bout, I., Raymond, K., and Sonnenberg, A. (2005) Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. J Cell Biol 171: 799-810.

Wolters, D. A., Washburn, M. P., and Yates, J. R. r. (2001) An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 73: 5683-5690.
— This was the first full description of the MudPIT proteomics approach.

Worman, H. J., and Courvalin, J. C. (2002) The nuclear lamina and inherited disease. Trends Cell Biol 12: 591-598.

Worman, H. J., Yuan, J., Blobel, G., and Georgatos, S. D. (1988) A lamin B receptor in the nuclear envelope. Proc Natl Acad Sci U S A 85: 8531-8534.

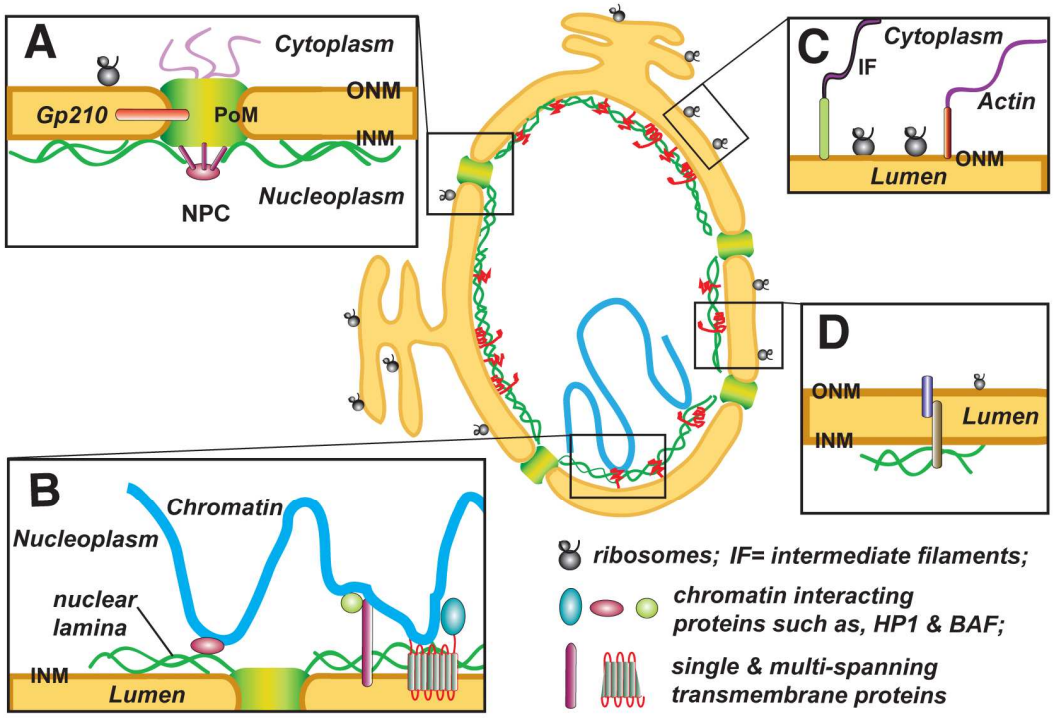
Xie, X., Wu, G., Lu, Z. H., Rohowsky-Kochan, C., and Ledeen, R. W. (2004) Presence of sodium-calcium exchanger/GM1 complex in the nuclear envelope of non-neural cells: nature of exchanger-GM1 interaction. Neurochem Res 29: 2135-2146.

Ye, Q., Barton, R. M., and Worman, H. J. (1998) Nuclear lamin-binding proteins. Subcell Biochem 31: 587-610.

Zhang, Q., Skepper, J. N., Yang, F., Davies, J. D., Hegyi, L., Roberts, R. G., Weissberg, P. L., Ellis, J. A., and Shanahan, C. M. (2001) Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. J Cell Sci 114: 4485-4498.

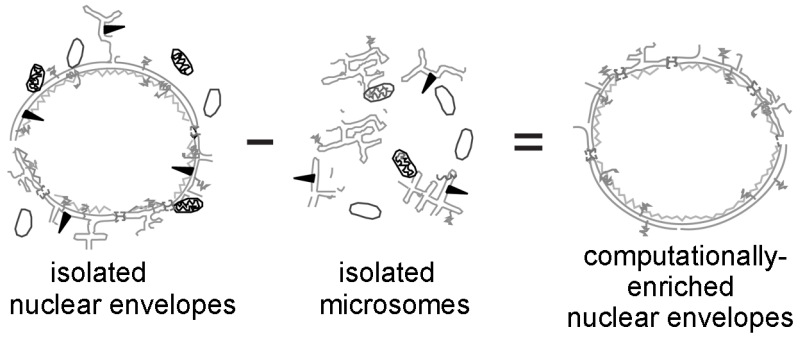
LEGENDS TO FIGURES AND TABLES

Figure 1.



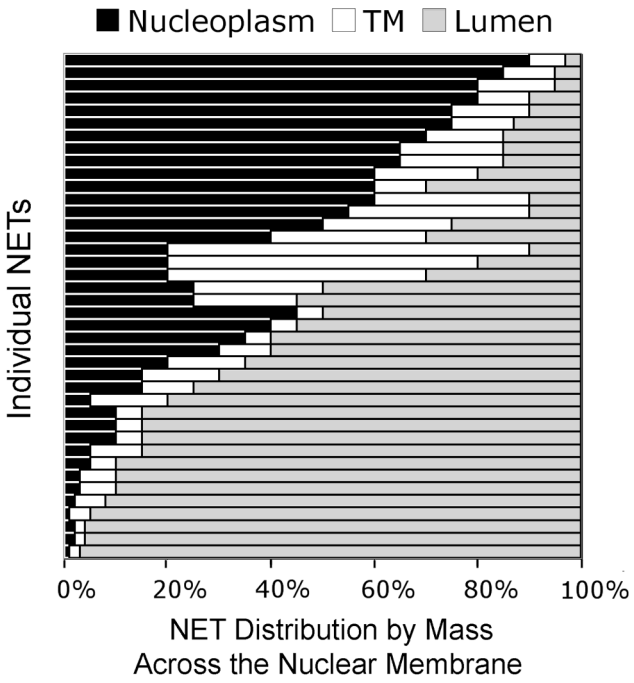
The nuclear envelope and its connections to other cellular compartments. (A.) The nuclear envelope is made of outer and inner nuclear membranes (ONM and INM), which are connected at the pore membrane (PoM) that encircles the nuclear pore complexes (NPCs). The NPCs regulate directional trafficking of soluble macromolecules in and out of the nucleus. On the cytoplasmic and nuclear side of the pore respectively are the cytoplasmic filaments and nuclear basket of the NPC. Nup153 of the nuclear basket interacts with lamins. NPCs are tethered to the PoM by two integral proteins, gp210 and POM121. The ONM is continuous with the endoplasmic reticulum (ER) and studded with ribosomes. Thus proximal ER is a major contaminant of NEs. (B.) Underneath the INM is the nuclear lamin polymer. Together with associated integral membrane proteins of the INM this polymer is called the nuclear lamina and some unique integral single or multi-spanning transmembrane proteins of the INM interact with chromatin and provide attachment points for the lamina. There are also integral proteins of the INM, PoM, and ONM that are not associated with the lamin polymer. Chromatin and soluble nucleoplasmic proteins may have specific interactions at the NE, but it is hard to distinguish these from contaminants because of the particular “stickiness” of many highly charged DNA and RNA-binding proteins. (C.) There are also integral proteins of the ONM that are not associated with the lamin polymer and some ONM-specific proteins bind to cytoplasmic filament systems, providing another major contaminant of NEs. Nesprin proteins appear to be able to link the NE to cytoplasmic intermediate filaments, while SUN proteins link the NE to cytoplasmic actin. (D.) Finally, the lumenal space between the ONM and INM, a soluble compartment that is also continuous with the ER is largely unexplored territory but some unidentified proteins may make contacts across the lumen in order to maintain its highly regular spacing.

Figure 2.



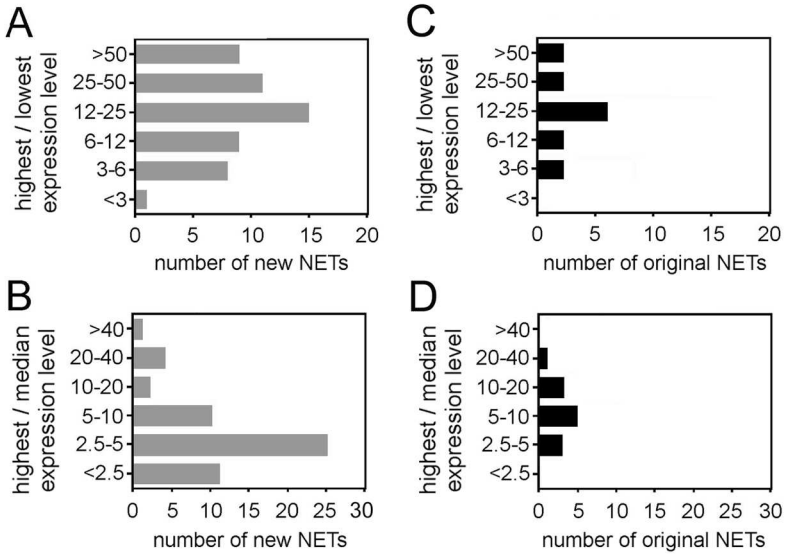
Subtractive proteomics. It is impossible to purify NEs to homogeneity because of the many connections to both the nucleoplasm and the cytoplasm. Thus, biochemically "purified" NEs are expected to be contaminated with chromatin and cytoskeletal proteins and with vesicles from organelles such as mitochodria and ER. In contrast, some of these expected contaminants can be purified free of NE contamination. One such contaminant is ER, which can be isolated as microsomes. Another is mitochondria, which has a well characterized protein complement. Therefore NE and microsomal membrane fractions are separately isolated and analyzed for protein content by MudPIT. All proteins appearing in both fractions are removed from the NE dataset because they could be due to ER vesicles sticking to the isolated nuclear NEs. Similarly, known mitochondrial proteins are removed. Because ER and mitochondria are the only expected membrane contaminants of NEs, all remaining integral membrane proteins in the NE fraction should be NE-specific in theory. After prediction by computer algorithm for membrane-spanning segments, an *in silico* purified NE transmembrane protein list is obtained. A limitation of this approach is that it discounts any proteins that are found both within the ER and the NE membranes (*e.g.* solid black triangles).

Figure 3.



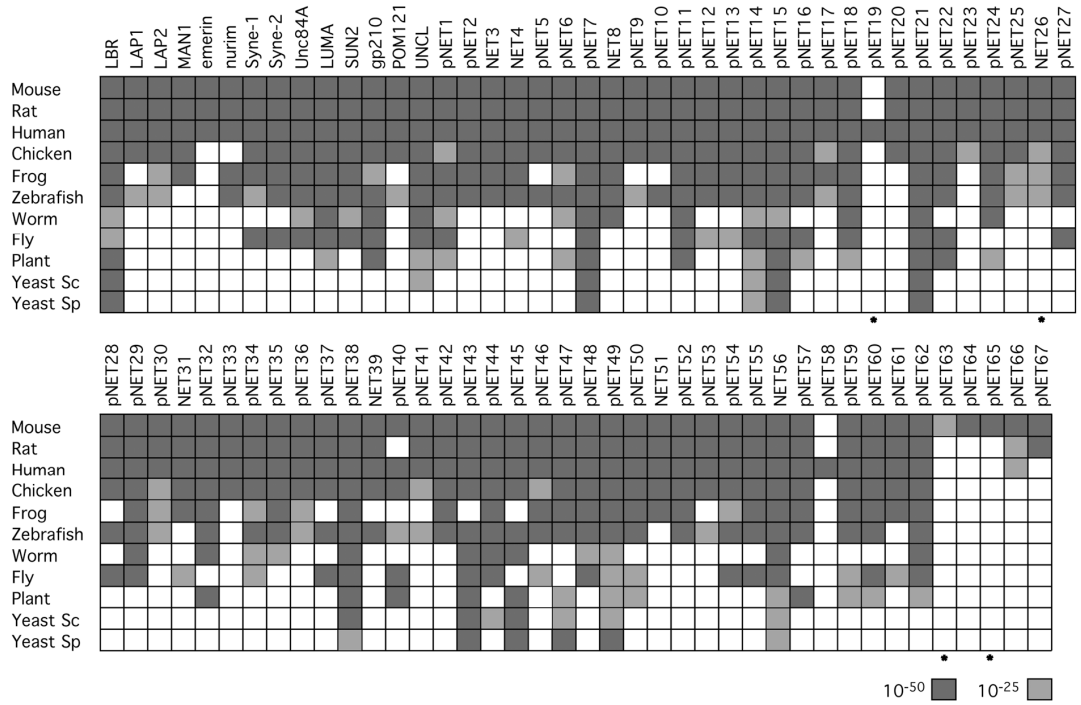
Predicted nuclear membrane topologies of new NETs. The TMHMM algorithm (http://www.cbs.dtu.dk/services/TMHMM-2.0/) was used to predict transmembrane helices for the NETs (plus splice variants for 5 NETs). This algorithm does not predict transmembrane segments for all the NETs that had been predicted by TMPred (Sect. 3.3); nonetheless, it was chosen for this analysis because it is considered to be more stringent. For those NETs with high probability predictions, the assigned topologies were used to calculate the percentage of the mass of each NET that should occur in the lumen, membrane, or nucleoplasm (nucleoplasm could be cytoplasm depending on if they reside in the INM or ONM). The percentage of the mass of each protein in each of the three domains is depicted graphically. Intriguingly, there were roughly as many proteins with the majority of their mass in the lumen as in the nucleoplasm. Several NETs also had a majority of their mass in the membrane.

Figure 4.



Histogram of NET expression variation. Expression data was generated by the laboratory of John Hogenesch at the Genomics Institute of the Novartis Research Foundation as part of their "transcriptome" database (available online at http://symatlas.gnf.org/SymAtlas/; Su et al., 2002). Maximum, median, and minimum values for NETs among the different tissues were determined and their ratios calculated. (A.) The ratio of the highest expression level to the lowest is presented for the new NETs identified in the subtractive study (Schirmer et al., 2003). If a NET is absent from some tissues and thus the lowest value is at the level of background, a high ratio would be produced. Very few NETs had low ratios. (B.) The ratio of the highest expression level to the median expression level for the new NETs. In this case a high ratio indicates a preference in expression for a subset of cell types. Though the distribution changes compared to A, there are still several NETs with very high ratios. (C and D.) The ratio of the highest expression level to the lowest level, C, and to the median level, D, is presented for the previously characterized NETs. Surprisingly, these also exhibited considerable variation.

Figure 5.



Evolutionary conservation of NETs. Strong homologs do not exist in the Genbank sequence database for many NETs in ten wide-ranging eukaryotic organisms that have completely sequenced genomes: Mouse, Mus musculus; Rat, Rattus norvegicus; Human, *Homo sapiens*; Chicken, *Gallus gallus*; Frog, *Xenopus laevis*; Zebrafish, *Danio rerio*; Worm, *Caenorhabditis elegans*; Fly, *Drosophila melanogaster*; Plant, *Arabidopsis thaliana*; Yeast *Sc*, *Saccharomyces cerevisiae*; Yeast *Sp Schizosaccharomyces pombe*. Dark boxes indicate probability scores lower than 10^{-50} and lighter shaded boxes indicate scores between 10^{-50} and 10^{-25}. White boxes indicate no homologs. The data is based on blast searches using NET protein sequences at different times, most recently updated on 27 April 2006. In cases where database sequences obtained from an earlier blast search (June 2005) had been removed pending sequence corrections, the homologs are maintained in this list. Some NETs that do not appear in all mammalian genomes may be absent because of sequence errors or failure to predict hypothetical orfs in those organisms; however, this seems unlikely for the set of original NETs which have been the subject of extensive searches over the years yet failed to identify homologs in flies (Gruenbaum et al., 2003). Asterisks indicate NETs that are included, yet had low probability scores (below the listed cut-offs) because the annotated sequence had been removed but the exons were still present in the genome sequence.

Table 1. Proteins from other cellular compartments also reported at the NE

Protein	Other Locations	System Tested	Cells /Tissue Tested	Detected at NE in Subtractive Study	Reference
TorsinA	ER	Exogenous expression of tagged protein	Gli36, BHK cells	Yes	(Goodchild and Dauer, 2004; Schirmer et al., 2003)
P-glycoprotein (multi-drug resistance-1 ATP-binding cassette subfamily B)	Plasma membrane, caveolae, and coated vesicles	Immunogold antibody labeling by EM.	Astrocytes	Yes	(Ronaldson et al., 2004; Schirmer et al., 2003)
CPLA2-a (cytosolic phospholipase A_2-a, 14-3-3 zeta)	ER	Determined with polyclonal antibodies	A23187 HeLa and A549 cells,	Yes	(Grewal et al., 2005; Schirmer et al., 2003)
EGFR (epidermal growth factor receptor)	Plasma membrane	At nuclear membrane by biochemical fractionization	*M. Musculus* liver	Yes	(Klein et al., 2004; Schirmer et al., 2003)
InsP 3R (inositol 1, 4, 5- triphosphate receptor type 2)	ER and cell periphery	Exogenous expression of tagged protein	MDCK cells	Yes	(Cruttwell et al., 2005; Schirmer et al., 2003)
Na+/K+ ATPase ß1 subunit	Plasma membrane	Subtype specific antibody labeling	Neurons, satellite cells in *R. norvegicus*	Yes	(Arteaga et al., 2004; Schirmer et al., 2003)
Ryanodine receptor	Various membranes	Determined by functional effect	NPCs in *X. laevis*	No	(Erickson et al., 2004)
NCX1 (Na+/ Ca++ exchanger membrane-1, solute carrier family 8)	Plasma membrane	Immunoblottting using various antibodies	*H. sapiens* brain and lymphocytes, not at NE in other tissues	No	(Xie et al., 2004)
PV-1 (plasmalemma vesicle protein-1)	ER	Exogenously expressed protein	Bovine aortic endothelial cells	No	(Hnasko and Ben-Jonathan, 2005)
NHE-1 (sodium/	Depends on cell	Fluorescent labeling, found at	Liver and	No	(Bkaily et al., 2004)

hydrogen exchanger)	type and species	nuclear membrane	cardiomyocytes in *R. norvegicus*		
Atypical type IV P-type ATPase	PM, post-mitochondrial supernatant	Immunoelectron microscopy and biochemical fractionization	*O. cuniculus* endometrium	Other family members were identified, but not this variant	(Mansharamani et al., 2001)