THE UNIVERSITY *of* EDINBURGH

Nuclear envelope transmembrane proteins in differentiation systems

Dzmitry G Batrakou

Thesis presented for the degree of Doctor of Philosophy

Wellcome Trust Centre for Cell Biology

School of Biological Sciences

University of Edinburgh

August 2011

I hereby declare that this thesis has been composed by myself and that the work presented is my own and has not been submitted for any other degree, except where stated otherwise

Dzmitry Batrakou

Acknowledgements

My warmest gratitude goes to Darwin Trust of Edinburgh, one of the few funding bodies that continue providing scholarships to overseas students despite the ridiculous fees for those not from EU countries.

I am grateful to Dr Eric C Schirmer for beneficial training, countless hours of education and freedom to “play” with the new techniques. The last might have reduced my actual achievements however it allowed me to acquire a number of skills that I believe are more important for my future career. I would also like to thank Eric and The Wellcome Trust for my financial support during the 4th year of my PhD and summer projects that I have done in the laboratory before starting PhD.

I am thankful to Nadia, my friend and colleague, for long-term supervision and support during a few ups and many downs of my PhD.

I would also like to thank the open source software community for proving that free software can be better than commercial one. I thank developers of Ubuntu, OpenOffice, Bibus, GIMP, ImageJ, GENtle, R project, and especially Ubuntu One cloud project that kept the draft of this thesis safe.

Abbreviations

°C	degree Celsius
2D	two dimensional
3D	three dimensional
A	alanine
AP	affinity purified
aa	amino acid
Amp	ampicillin
ATP	adenosine-5'-triphosphate
BAC	bacterial artificial chromosome
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
CMV	cytomegalovirus
CNBr	cyanogen bromide
cRNA	complementary RNA
ctrl	control
D	aspartic acid
Da	dalton
DamID	DNA adenine methyltransferase identification
DMS	dexamethasone
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide
Dr	doctor
DSB	double strand break
DTBP	dimethyl 3,3'-dithiobispropionimidate
DTT	dithiothreitol
DUF	domain of unknown function
E	glutamic acid
E.coli	Escherichia coli
EDMD	Emery-Dreifuss muscular dystrophy
EDTA	ethylenediaminetetraacetic acid
E-GFP	enchanced green fluorescent protein
ER	endoplasmic reticulum
F	phenylalanine
FBS	fetal bovine serum
FISH	fluorescence in situ hybridisation
FITC	fluorescein isothiocyanate
FPLD	Dunnigan-type familial partial lipodystrophy
FT	flow-through
g	gram
h	hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HR	homologous recombination
HRP	horse radish peroxidase
IBMX	3-isobutyl-1-methylxanthine
IF	immunofluorescence
Ig	immunoglobulin
INM	inner nuclear membrane
IP	immunoprecipitation
IVCL	in vivo crosslinking
K	lysine
Kan	kanamycin
LC	liquid chromatography
LAD	lamina associated domain
LINC	links the nucleoskeleton and cytoskeleton (complex)
log	logarithm
MALDI TOF	matrix-assisted laser desorption/ionization time-of-flight
MEM	minimal essential medium
min	minute
mRFP	monomeric red fluorescent protein
mRNA	messenger RNA
MS	mass spectrometry
MudPIT	multidimensional protein identification technology
N/A	not applicable
NCBI	National Center for Biotechnology information
NE	nuclear envelope
NET	nuclear envelope transmembrane protein
NHEJ	non-homologous end joining
NIH	National Institutes of Health
NPC	nuclear pore complex
NTD	N-terminal domain
OAc	acetate group
ONM	outer nuclear membrane
PBS	phosphate buffered saline
PC	personal computer
PCR	polymerase chain reaction
pd	pulldown
pH	$-\log_{10}(a_H+)$

PHA	phytohaemagglutinine	**PTM**	post-translational modification
RNA	ribonucleic acid	**qPCR**	quantitative PCR
RT	reverse transcription	**g**	relative centrifugal force
s	second	**SSC**	saline sodium citrate
S	serine	**SV40**	simian virus 40
SBP	streptavidin binding peptide	**TCA**	trichloroacetic acid
SDS	sodium dodecyl sulphate	**TEN**	Tris, EDTA, NaCl buffer
SDS-PAGE	SDS polyacrylamide gel electrophoresis	**TM**	transmembrane
SHM	Sucrose, HEPES, Mg^{2+}-containing buffer	**TMEM**	transmembrane protein
shRNA	small hairpin RNA	**TRIzol**	total RNA isolation reagent
siRNA	short inhibitory RNA	**UV**	ultraviolet
PoM	pore membrane	**Y**	tyrosine
PVDF	polyvinylidene fluoride	**Y2H**	yeast two-hybrid

List of figures

List of tables

Table of contents

Abstract

Historically, our perception of the nuclear envelope has evolved from a simple barrier isolating the genome from the rest of a cell to a complex system that regulates functions including transcription, splicing, DNA replication and repair and development. Several recent proteomic studies uncovered a great variety of nuclear envelope transmembrane proteins (NETs). Diseases associated with several nuclear envelope proteins, mostly NETs, affect many tissues e.g. muscle, adipose tissue, skin, bones. Many NETs of the inner nuclear membrane have been shown to interact with chromatin, suggesting that their influencing gene expression might explain NET roles in disease.

This work is focused on finding novel interactions of NETs with chromatin. First, SUN2 post-translational modifications were analysed and the effect of phosphomimetic and phospho-null mutants on heterochromatin and the cytoskeleton was tested by overexpression. However, no obvious changes were found. Second, several tissue-preferential NETs were tested in an adipocyte differentiation system. NET29 changed chromosome 6 position in pre-adipocytes. This matched changes in chromosome positioning that occur during adipocyte differentiation when NET29 is normally induced. Post-translational modifications of NET29 are likely to play a vital role in this process because a phospho-null mutant dominantly blocked chromosome repositioning. The effect of over-expression and down-regulation of NET29 on transcription was tested and results suggest that NET29 negatively regulates expression of myogenic genes during adipogenesis.

This thesis is split into six chapters. Chapter I is an overview of the nuclear envelope, adipogenesis and chromatin remodelling, Chapter II is a detailed description of methods used in this study. Chapter III focuses on post-translational modifications of SUN2, as well as trials to identify novel partners of SUN2. Chapter IV and V deal with a novel nuclear envelope transmembrane protein and its role in adipogenesis. Finally, the last chapter includes a discussion and recommended future directions.

CHAPTER I: Introduction

1.1 Nuclear envelope

1.1.1 General overview

The nuclear envelope (NE) is a double membrane system that provides a barrier between cytoplasmic processes and those inside of the nucleus. The NE is comprised of an inner nuclear membrane (INM) and an outer nuclear membrane (ONM) (Fig 1.1.1). The space between them is around 50 nm and is called the lumen (Callan and Tomlin, 1950). The ONM is, basically, an extension of the rough endoplasmic reticulum (ER) and connects with the INM in areas adjacent to the nuclear pore complexes (NPCs). The INM and ONM connect at where NPCs are inserted in what is often referred to as the pore membrane (PoM) (reviewed in Newport and Forbes, 1987). Underlying the INM is a meshwork made of polymers of different lamin isoforms (Gerace *et al.*, 1984). A large body of evidence argues that the NE has many complex functions (reviewed in Akhtar and Gasser, 2007; Stewart *et al.*, 2007; Batrakou *et al.*, 2009).

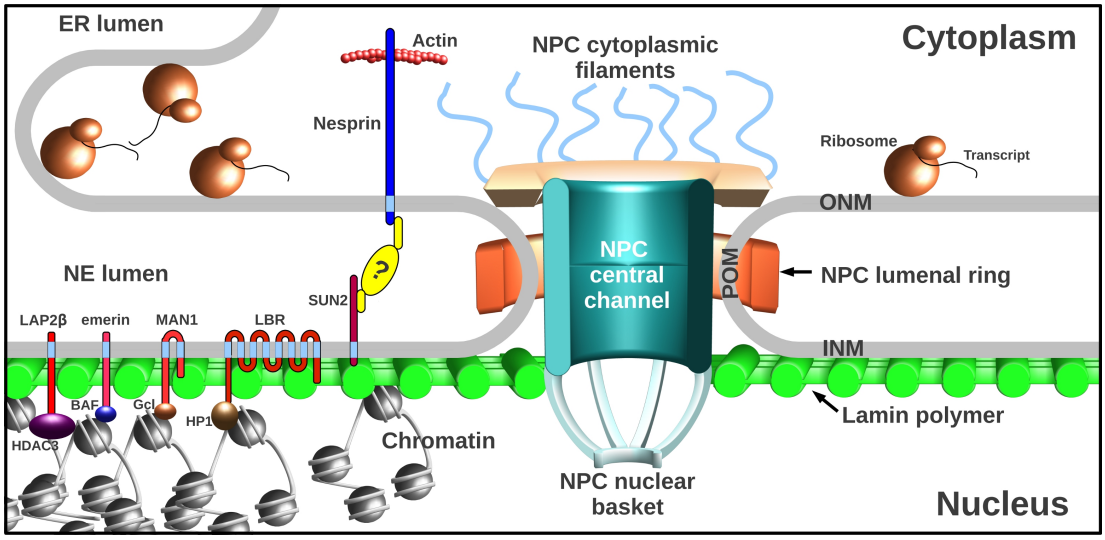


Fig 1.1.1: The nuclear envelope (modified from Batrakou et al., 2009)

The nuclear envelope is a double membrane system that separates the nucleus from the cytoplasm. The inner nuclear membrane (INM), the outer nuclear membrane (ONM) and the pore membrane (PoM) contain specific sets of transmembrane proteins. The ONM is also encrusted with ribosomes and the INM is underlined with the nuclear lamina – a meshwork of intermediate filaments. The nuclear pore complex (NPC) connects the cytoplasm with the nucleoplasm. The ONM located nesprins interact with the cytoskeleton and, possibly through a partner protein, with the INM located SUN proteins. SUN proteins interact with the lamina, thus the nesprin-SUN protein interaction bridges the cytoskeleton with the peripheral nuclear skeleton. Several proteins of the INM interact with various chromatin components (more detailed in Fig 1.1.3).

1.1.2 Outer Nuclear Membrane *vs* Inner Nuclear Membrane

Until recently, the ONM was considered to be simply an ER membrane extension with the same pool of proteins. Both the ER and the ONM are decorated with ribosomes engaged in protein synthesis. But it was shown that at least one protein family is distinct to the ONM – nesprins (Zhen *et al.*, 2002). Nesprins are giant spectrin repeat-containing transmembrane proteins that connect the cytoskeleton to the nucleus.

On the other hand, the INM was determined as a unique membrane long before the ONM. Between 1988 - 2001 eleven INM transmembrane proteins had been discovered and initially described: LBR – lamin B receptor (Worman *et al.*,

1988), LAP1 and LAP2 – lamina associated polypeptides 1 and 2 (Senior and Gerace, 1988), gp210 (Greber *et al.*, 1990), POM121 (Hallberg *et al.*, 1993), emerin (Manilal *et al.*, 1996), nurim (Rolls *et al.*, 1999), SUN1 (Malone *et al.*, 1999), SUN2 (Hoffenberg *et al.*, 2000), MAN1 (Lin *et al.*, 2000) and LUMA (Dreger *et al.*, 2001).

1.1.3 INM resident proteins and heterochromatin

There is a great potential for the NE to be involved in genome regulation because of its many connections to chromatin (Fig 1.1.3). Both lamins and several LEM domain-containing INM proteins have been shown to interact with core histones, transcription factors and transcriptional repressors and the DNA crosslinking factor BAF (reviewed in Mattout-Drubezki and Gruenbaum, 2003; Tsuchiya, 2008). SUN1 is functionally associated with the histone acetyltransferase hALP (NET43) and they are both required for chromosome de-condensation after mitosis (Chi *et al.*, 2007). In vitro, LBR binds to heterochromatin protein 1 (Ye and Worman, 1996), which associates with under-acetylated chromatin through its interaction with histone H3/H4 oligomers (Polioudaki *et al.*, 2001). Recent studies suggest a much greater variety of the integral proteins of the NE and their functions, respectively (Schirmer *et al.*, 2003; Malik *et al.*, 2010; Korfali *et al.*, 2010; Wilkie *et al.*, 2010).

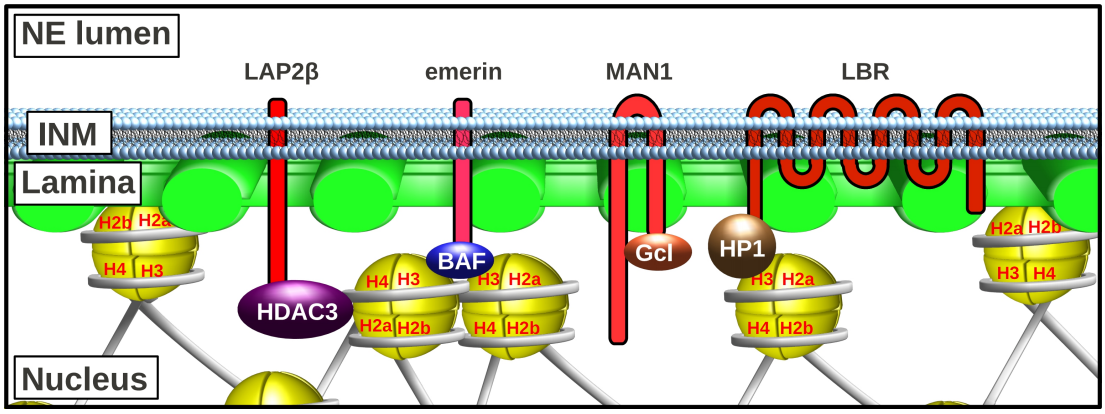


Fig 1.1.3: Proteins of the INM (modified from Batrakou et al., 2009)

Proteins that reside within the INM are connected with the chromatin. Barrier to autointegration factor (BAF) was shown to mediate this connection in the case of the LEM domain-containing proteins LAP2, emerin and MAN1 (Furukawa, 1999, Lee et al., 2001, Liu et al., 2003). LBR is associated with heterochromatin protein 1 (HP1) which binds to H3K9me3 marked chromatin and appears to enhance and maintain its inactive state (Lachner et al., 2001, Bannister et al., 2001). Germ cell-less (Gcl), a transcriptional repressor, also was shown to bind LAP2β, emerin and MAN1 (Nili et al., 2001; Holaska et al., 2003; Mansharamani and Wilson, 2005). LAP2β was also shown to bind histone deacetylase-3 HDAC3 associated with the silent chromatin (Somech et al., 2005).

1.1.4 The nuclear envelope and DNA double strand break repair

Recent studies linked the NE with DNA repair, specifically the repair of persistent double strand breaks (DSB).

DNA DSB repair is an important mechanism of cell defence against the serious threats to chromatin integrity by external (for example, gamma-rays) and internal (free radicals) factors. There are two main mechanisms to re-join two ends of DNA - homologous recombination (HR) and non-homologous end joining (NHEJ). The former mechanism prevails in S and G_2 phases of the cell cycle when an exact copy of the damaged strand is available in a sister chromatid. The intact copy is used then as a template to repair the damaged locus error-free.

However, when no homology to the damaged locus is found, NHEJ takes place. The Ku70/Ku80 heterodimer binds to the broken ends and recruits DNA-dependent protein kinase (DNA-PK), which binds to the DNA ends and

autophosphorylates itself. Upon autophosphorylation it changes conformation and is then released from the DSB ends, allowing DNA ligase to join the ends. The processing step of DNA ends is likely to introduce deletions into the re-joined locus.

A histone 2 variant, H2AX, seems to play an important role in DSB response. This histone variant is distributed throughout chromatin and is phosphorylated upon DNA damage by DNA-PK and ATM kinase that is involved in the HR repair (Burma *et al.*, 2001). The phosphorylated serine 139 form of the histone is called gamma-H2AX (γ-H2AX) and it marks several megabases of DNA flanking a DSB (Rogakou *et al.*, 1998).

Despite many sophisticated mechanisms of DNA repair, there are occasions when a DSB persists for a long time. Even a single DSB will lead to a cell cycle arrest at the G_2/M DNA damage checkpoint in yeast. However, the checkpoint was shown to be overridden after about 15 hours despite the continued presence of DSBs (Toczyski *et al.*, 1997). Another study showed that persistent DSBs are relocated to the nuclear periphery and a SUN domain-containing protein, Mps3p, is required for the peripheral localisation of DSBs in yeast (Oza *et al.*, 2009). Nagai *et al.* (2008) have also shown that DSBs are tethered to the NE and this requires the Nup84 subcomplex of nuclear pores and a SUMO-dependent ubiquitin ligase.

A more detailed mechanism of the DSB tethering to the NE in yeast was described in a recent study (Kalocsay *et al.*, 2009). Rad51 which is involved in HR was shown to spread outward of persistent DSBs and, thus, mark the entire pieces of the broken chromosome. H2AZ, another variant of the H2A family of histones, is also recruited to the DSB and its SUMOylation is important for the DSB translocation to the NE. The DNA damage checkpoint seems to be required for the translocation as well.

Thus, at least in yeast, a SUN domain-containing protein of the NE appears to play a role in the DNA damage response. It is not yet clear why the DSBs are tethered to the periphery. There could be a NE-associated backup mechanism of DSB repair. Alternatively, association of persistent DSBs with the NE might enable the

arrested cell to divide. But it is obvious that there is an active mechanism that recognises the broken chromosome and relocates either adjacent to the DSB domains or moves the whole chromosome to the nuclear periphery.

1.1.5 The nuclear envelope and disease

Mutations of NE proteins have been shown to be responsible for at least 15 heritable human diseases - the nuclear envelopathies (diseases associated with NE proteins). Several mutations of Emerin, a predominantly INM protein, are responsible for X-linked Emery-Dreifuss muscular dystrophy, EDMD (Bione *et al.*, 1994). A range of diseases with distinct tissue-specific pathologies are caused by mutations in NE proteins. These include lipodystrophies, neuropathy, dermopathy, dystonia and premature ageing syndromes (reviewed in Worman and Bonne, 2007).

Specifically familial partial lipodystrophy of the Dunnigan type (FPLD) is characterized by a lack of adipose tissue in the limbs, buttocks and trunk with fat accumulation in the neck and face. It is caused by heterozygous missense mutations in *LMNA* (Shackleton *et al.*, 2000) and possibly by disrupting the differentiation of adipocytes. This is supported by the interaction between the sterol response element binding factor 1 (SREBF1), an adipocyte differentiation factor, and lamin A, which is noticeably reduced by FPLD causing mutations (Lloyd *et al.*, 2002). It was also shown that overexpression of lamin A inhibits the *in* vitro 3T3-L1 pre-adipocyte differentiation to adipocytes by suppressing PPARγ and Slc2a4 (Boguslavsky *et al.*, 2006). The same study describes that murine embryonic fibroblasts derived from lamin A knockout mice accumulate more lipids and synthesize more triglycerides compared to wild-type fibroblasts.

1.2 LINC complex

1.2.1 Discovery

The term LINC complex was coined by Crisp *et al.* (2006) and it is defined as

a complex that **li**nks the **n**ucleoskeleton and **c**ytoskeleton. It consists of SUN proteins and the KASH domain-containing nesprins proteins and connects all major components of the cytoskeleton through the NE to the peripheral nuclear skeleton – the lamina (Fig 1.2.1). Crisp *et al.* (2006) showed a direct interaction of murine SUN1 and SUN2 with all types of lamins whilst the rest of the complex interactions were shown earlier in *C. elegans*. First, the SUN domain-containing protein UNC-84 and the KASH domain-containing protein Syne/ANC-1 have been shown to be important for nuclear migration and anchoring (Sulston and Horvitz, 1981; Hedgecock and Thomson, 1982). Decades later the *C. elegans* KASH domain-containing ANC-1 was cloned and shown to localise at the NE in an UNC84-dependent manner, as well as to interact with filamentous actin (Starr and Han, 2002).

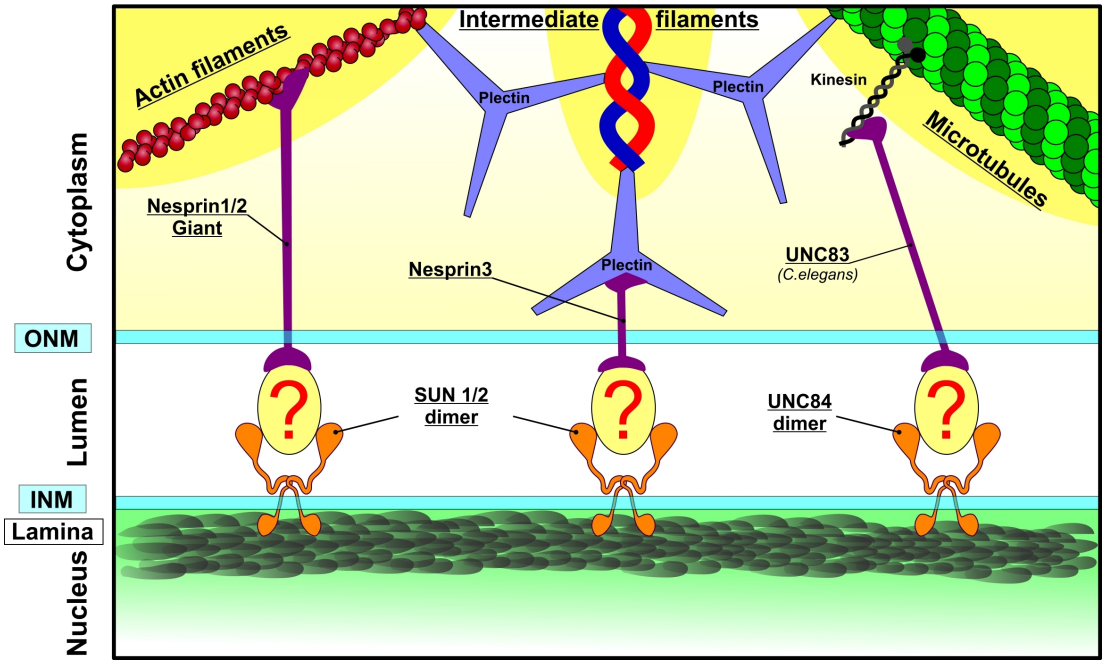


Fig 1.2.1: The LINC complex

The LINC complex connects the peripheral nucleoskeleton (lamina) to the cytoskeleton. Different forms of nesprins have been shown to interact directly or indirectly with all major cytoskeletal components: actin filaments (directly), microtubules (through kinesin) and intermediate filaments (through plectin, which also binds actin and tubulin). SUN domain-containing proteins, on the other hand, bind the lamina – a polymer of intermediate filaments that are underlying the inner nuclear membrane. It is possible that the interaction between SUN proteins and nesprins is mediated by another yet unidentified protein (question mark).

1.2.2 Cytoplasmic side of the LINC complex: KASH domain proteins

The KASH domain (for Klarsicht, Anc-1, Syne homology) is a small (around 60 aa) conserved domain that includes a transmembrane region and less than 35 C-terminal amino acids. *C. elegans* has three KASH domain containing proteins: ANC-1, UNC-83 and ZYG-12. Only two proteins of the family are found in *D. melanogaster*: Klarsicht and MSP300. Mammals have four KASH domain containing proteins: nesprin-1, nesprin-2, nesprin-3 and nesprin-4. Genes encoding nesprins, syne1-4 (for synaptic nuclei expressed), give rise to many isoforms (more than 12 for syne-1 and syne-2 each) (Zhang *et al.*, 2001; Zhen *et al.*, 2002).

Nesprins (Fig 1.2.1) are type II integral membrane proteins with a short lumenal KASH domain and a large (around 1MDa for Nesprin-1 and 0.8 MDa for Nesprin-2) cytoplasmic N-terminal domain, composed of spectrin repeats that are predicted to form a rod domain similar to that of dystrophin, a giant protein that connects the muscle fibre to the extracellular matrix (Zhang *et al.*, 2001). Apart from that, Nesprin-1 and Nesprin-2 have N-terminal calponin-homology domains that bind to actin, the building block of microfilaments and thin filaments of muscle (Starr and Han, 2002). Nesprin-3 does not have the domain but it was shown to bind plectin, a protein that links microfilaments, intermediate filaments and microtubules (Wilhelmsen *et al.*, 2005). Furthermore, UNC83, a KASH domain-containing protein from *C. elegans*, has been shown to interact with the kinesin-1 light chain, a motor protein that binds microtubules (Meyerzon *et al.*, 2009). As nesprins are anchored in the ONM and are linked to actin, plectin and, indirectly, to tubulin, they potentially connect the nucleus to the cytoskeleton that is comprised of microfilaments, intermediate filaments and microtubules. It should be mentioned that cytoplasmic localisation of nesprins is not exclusive. It was shown that some short isoforms of Nesprin-1 and Nesprin-2 are localised to the nucleus where they co-localise with heterochromatin and bind lamin A and emerin (Zhang *et al.*, 2001; Mislow *et al.*, 2002; Zhang *et al.*, 2005).

1.2.3 Nucleoplasmic side of the LINC complex: SUN domain proteins

The SUN domain (for Sad, Unc-84 homology) encompasses about 130 aa. Apart from the SUN domain, sequences of SUN proteins are quite diverse. After molecular cloning of *UNC-84* (Malone *et al.*, 1999) and establishing its role in connecting the peripheral nucleoskeleton to the cytoskeleton (Starr and Han, 2002), many laboratories have focused their research on SUN domain proteins. Studies of mammalian homologues of the protein revealed multiple functions of proteins of the family.

Mammalian SUN domain proteins include six members – SUN1, SUN2, SUN3, SUN4 (SPAG4), SUN5 (SPAG4L) and SPAG4L2. SUN1 and SUN2 proteins

are ubiquitously expressed whereas others are found mainly in testis.

SUN1 is a type 2 transmembrane protein with a nucleoplasmic N-terminus and a C-terminus present in the NE lumen. Its nucleoplasmic domain can bind lamin A and, to lesser degree, lamin B (Crisp *et al.*, 2006). The N-terminus of SUN1, lacking the transmembrane region, localises in the cytoplasm and forms dot-like structures, presumably at centrosomes (Wang *et al.*, 2006). Thus, transient interaction of SUN1 and the microtubule organising organelle is possible during the cell cycle when the NE is disassembled.

A recent report has argued that SUN1 has a functional association with NPCs – knocking down the protein caused NPCs to cluster (Liu *et al.*, 2007) suggesting that SUN1 is important for the uniform distribution of NPCs in the NE. Confocal microscopy showed a spatial pattern and partial co-localisation of SUN1 with NPCs (Lu *et al.*, 2008). Furthermore, the high immobility of the protein in the NE suggests that it is a part of a big complex.

SUN2 is also a type II transmembrane protein with a nucleoplasmic N-terminus and a lumenal C-terminus. As for SUN1, it has been shown to bind Lamin A and Lamin B (Crisp *et al.*, 2006). Interestingly, a fraction of endogenous SUN2 was shown to localise at the centrosomes in a cell cycle dependent manner (Wang *et al.*, 2006); however, no functional data was provided.

A recent paper revealed that SUN2 is involved in repositioning of telomeres during meiotic prophase I (Schmitt *et al.*, 2007). It was shown in rat testicular cells to be a part of a fibrillar complex that interconnects the cytoskeleton with telomeres. In budding yeast a SUN domain containing protein, Sad1, was shown to change localisation pattern with depletion of another NET Ima I, a homolog of NET5 (King *et al.*, 2008). Ima I was shown to indirectly bind to the heterochromatin of centromere regions of chromosomes.

Thus, SUN domain proteins are linked to the peripheral nucleoskeleton through their direct interaction with lamins and to chromosomes through as yet unidentified partners. Known partners of SUN2 are listed in Table 1.

Table 1: SUN2 interactors

#	Partner	SUN2 interacting domain	Method	Direct binding	Reference
1	Rab5	C-terminus after TM	Y2H, in vitro binding assays	Yes	Hoffenberg *et al.*, 2000
2	SUN2	C-terminus after TM, N-terminus	Co-IP of tagged expressed proteins	N/A	Wang *et al.*, 2006
3	SUN1	N/A	Co-IP of tagged expressed proteins	N/A	Wang *et al.*, 2006
3	Lamin A	N-terminus, first 165aa	Pull-down	Yes	Crisp *et al.*, 2006
4	PreLamin A	N-terminus, first 165aa	Pull-down, cotransfection/IF	Yes	Crisp *et al.*, 2006
5	Nesprin2G	N/A	Co-IP	N/A	Crisp *et al.*, 2006
6	Nesprin3α	N/A	Co-IP	N/A	Ketema *et al.*, 2007
7	Emerin	N-terminus	Co-IP, pull-down	N/A	Haque *et al.*, 2010

1.2.4 Connection between SUN and KASH domains

It is well established that SUN domains interact with KASH domains (Starr *et al.*, 2001; Crisp *et al.*, 2006; Haque *et al.*, 2010) in the NE lumen and bridge the INM and the ONM. Moreover, localisation of nesprins is dependent on both SUN1 and SUN2 proteins (Crisp *et al.*, 2006), or only SUN1 (Padmakumar *et al.*, 2005). The last two studies gave different results most likely due to different anti-Nesprin-2 antibodies used, one that recognises specifically Nesprin-2 Giant isoform (Padmakumar *et al.*, 2005) and the other that recognises many nesprin isoforms (Crisp *et al.*, 2006).

However, the length of the lumenal C-termini of SUN1 or SUN2 is about 500 aa and includes only a few coiled-coil domains; the rest is presumably globular. The KASH domain without the transmembrane region is about 35 aa. It is unlikely that these two structures cover about 50 nm of lumenal distance between INM and ONM (Fig 1.2.1). Moreover, when direct binding of SUN1 and Nesprin-2 was tested using the yeast two hybrid assay, the SUN domain showed only a weak interaction with the KASH domain (Padmakumar *et al.*, 2005). It is the part of SUN1 between the SUN

and coiled-coil domains that showed a strong interaction. Also, both SUN1 and SUN2 N-termini have been shown to interact with nesprins (Haque *et al.*, 2010). The last fact suggests that a fraction of both SUN and KASH domain-containing proteins are located within the same membrane, either the INM or the ONM or both. This could potentially explain results that were interpreted as a consequence of interaction between SUN and KASH domains. Further, the N-terminus of nesprins contains spectrin repeats and the C-terminus of SUN proteins contains coiled coil domains. Both motifs are known to interact non-specifically.

Based on the statements above it can be hypothesised that the interaction between KASH and SUN domains through the NE lumen is indirect. Hence, it is mediated by a partner(s) located in NE lumen and interacting with both SUN and KASH domains. One possible candidate for the mediator is torsin A, a protein that resides in the NE lumen. It was shown that the KASH domain of Nesprin 3 binds torsin A (Nery *et al.*, 2008). The NE localisation of Torsin A-ΔE, a mutant of torsin that causes torsion dystonia, is SUN1-dependent (Jungwirth *et al.*, 2011). Moreover, the protein was found to co-immunoprecipitate with SUN2, though it needs to be confirmed (present work, see Table 13).

1.3 Nuclear envelope proteome

1.3.1 Pre-proteomics studies

Xenopus laevis oocytes served as an excellent model system for studying NE components, especially NPC transport (reviewed in Peters, 2006). Extracts from the oocytes can assemble a NE upon addition of sperm chromatin (Lohka and Masui, 1983). Two distinct vesicle fractions of the oocytes extract were identified and named NEP-A and NEP-B, for **n**uclear **e**nvelope **p**recursor (Vigers and Lohka, 1991). Both fractions are involved in NPC formation (Salpingidou *et al.*, 2008).

The very first NET was identified through its interaction with lamin polymers

as the latter were already biochemically characterised. LBR was co-purified with lamin B1, hence the name - lamin B receptor (Worman *et al.*, 1988). LAP proteins appeared to be major antigens in a NE fraction from rat liver (Senior and Gerace, 1988). Gp210 was identified as a major transmembrane glycoprotein of the NPC (Greber *et al.*, 1990). POM121 was identified from a rat liver NE fraction through its binding to WGA (Hallberg *et al.*, 1993). Emerin was first shown to cause X-linked EDMD (Bione *et al.*, 1994) and later described as a NET (Manilal *et al.*, 1996). Nurim was found at the NE in a visual screen with a GFP-tagged library of proteins (Rolls *et al.*, 1999). SUN1 was identified due to its effect on nuclear migration and anchoring in *C. elegans* (Malone *et al.*, 1999). SUN2 was first described as a protein interacting with Rab5 (Hoffenberg *et al.*, 2000) and later was defined as a NET (Hodzic *et al.*, 2004). MAN1 was identified as a protein recognised by autoantibodies in a patient with a collagen vascular disease (Lin *et al.*, 2000). Nesprin-1 was identified through its binding to muscle kinase (Apel *et al.*, 2000).

Thus, only 12 NETs were identified before 2001, and only a few studies were aiming to identify novel NETs. Development of mass spectrometry and the NE fractionation techniques lead to a new stage of NETs identification - proteomic approaches.

1.3.2 Proteomic analysis of the NE

Whereas some of the NETs were and continue to be identified by traditional approaches, the development of mass spectrometry allowed more directed studies of NE proteomes. However, to obtain a clean NE fraction two major obstacles had to be overcome. First of all, the outer nuclear membrane is continuous with ER and any attempt to purify the NE will co-purify proteins from ER. Second of all, the lamina is tightly associated with chromatin that will also add contaminants into the NE fraction.

In order to overcome the problems mentioned above, the Otto group compared insoluble material from chaotropes (urea and sodium carbonate) and non-

ionic detergent (Triton X-100) extractions using 2D SDS-PAGE (Dreger *et al.*, 2001). While chaotropes should solubilise the lamina and most of the protein-protein interactions, leaving the membranes intact, detergent extraction should remove membranes and proteins that are weakly associated with the lamina. Hence, insoluble material in both extractions represents potential INM NETs that are strongly associated with the lamina. This approach identified two novel NETs - mammalian SUN1 and LUMA. Such a poor number of identified proteins could be explained by the fact that the purified fractions were resolved on gels, visualised with Coomassie, then digested in the gel and subjected to MALDI TOF mass spectrometry. Each of these steps has introduced limitations into the final number of identified proteins.

On the other hand, Schirmer *et al.* (2003) used a subtractive approach. Both crude NEs and, separately, microsomes were purified and analysed by multidimensional protein identification technology (MudPIT) that avoids gels and couples multiple chromatography steps with tandem mass spectrometry. Then the proteins that were identified in the microsomal fraction were subtracted from the proteins identified in the crude NE fraction. The presence of predicted transmembrane domains among the rest of the proteins defined a large (67) number of novel putative NETs. Subsequent studies identified in total over five hundred potential NETs in three different tissues (Korfali *et al.*, 2010; Wilkie *et al.*, 2010), with only

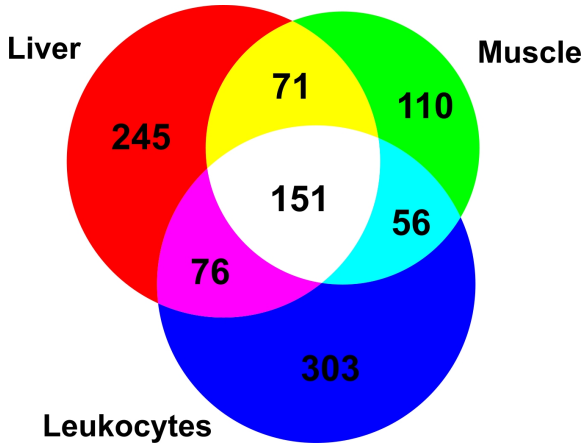


Fig 1.3.2: Distribution of putative NETs identified in three different tissues

Total number of proteins that were over-represented in the NE fractions compared to microsomal fractions and contain at least one strongly predicted transmembrane domain is compared between three different tissues. Only a third of putative NETs is shared between liver, muscles and blood, suggesting that some of the NETs are tissue-specific. Liver and muscle tissues are from rat and leukocytes are from human peripheral blood.

about a third being shared between the tissues (Fig 1.3.2, N Korfali, GS Wilkie, SK Swanson, V Srsen, J de las Heras, DG Batrakou, ARW Kerr, L Florens, and EC Schirmer, manuscript in preparation), suggesting tissue-specific or tissue preferential expression for the other two thirds of corresponding genes.

Existing transcriptome data supports this mass spectrometry based observation (refer to results, Fig 4.1.1 a). Interestingly, many known diseases associated with NE have tissue-specific phenotypes. This suggests that mutations in some of the newly identified putative (with about 70 of them being confirmed at the time of writing this manuscript) NETs could cause or contribute to the tissue-specific disease phenotypes. Any investigation of tissue-specific proteins should be done in cells of the corresponding tissue; moreover, the pre-differentiated state of the same cells could provide additional insights into role of the tissue preferential NETs. Among a few *in vitro* differentiation systems adipogenesis is considered to be both robust and reliable. More importantly, several NETs were found to be highly expressed in adipocytes (refer to results, Fig 4.1.1 a). Thus we took advantage of the murine 3T3-L1 pre-adipocyte cell line that can be induced to differentiate into adipocytes.

1.4 Adipogenesis

1.4.1 Adipose tissue

Adipocytes are specialised cells that store energy in the form of triacylglycerol in order to provide energy when the more preferential source of energy - glucose - is not available. Apart from acting as storage for energy, adipocytes also secrete adipokines and play a role in the control of metabolism (reviewed in Vázquez-Vela *et al.*, 2008).

There are two types of mammalian adipocytes, white and brown adipocytes. Brown adipocytes store less lipid, have more mitochondria and are able to dissipate energy as heat without the generation of ATP. Expression profiles of white and brown

adipocytes are similar, however brown adipocytes additionally express some distinct genes. In infancy, humans have large amounts of brown adipose tissue, but only small amounts persist in adults (reviewed in Koppen and Kalkhoven, 2010).

The adipose tissue, together with muscle and bone, develops from the mesenchymal stem cells derived from the mesodermal layer of the embryo (reviewed in Park *et al.*, 2008). Multipotent precursor cells in the vascular stroma of adipose tissue become restricted to the adipocyte lineage without actually expressing any markers of terminal expression. Multiple transcription factors are subsequently activated causing the cells to differentiate. This process is initiated by factors secreted by cells in the vascular stroma and/or adipocytes undergoing hypertrophy.

The 3T3-L1 cell line is a well-established murine preadipocyte cell line that has already been committed to the adipocyte lineage. The 3T3-L1 cells were selected for their ability to accumulate cytoplasmic triacylglycerol while in a resting state (Green and Kehinde, 1975). The cells cannot be distinguished from their precursor cells morphologically, but can be induced to terminally differentiate to adipocytes using a cocktail of insulin, 3-isobutyl-1-methylxanthine (IBMX) and dexamethasone (DMS) (Student *et al.*, 1980). 3T3-L1 cells offer a homogeneous cell population and a stable pre-differentiated state.

1.4.2 Transcriptional regulation of adipogenesis

The transcriptional networks in adipogenesis are highly regulated and involve many different factors, of which a part may still be unknown (Fig 1.4.2).

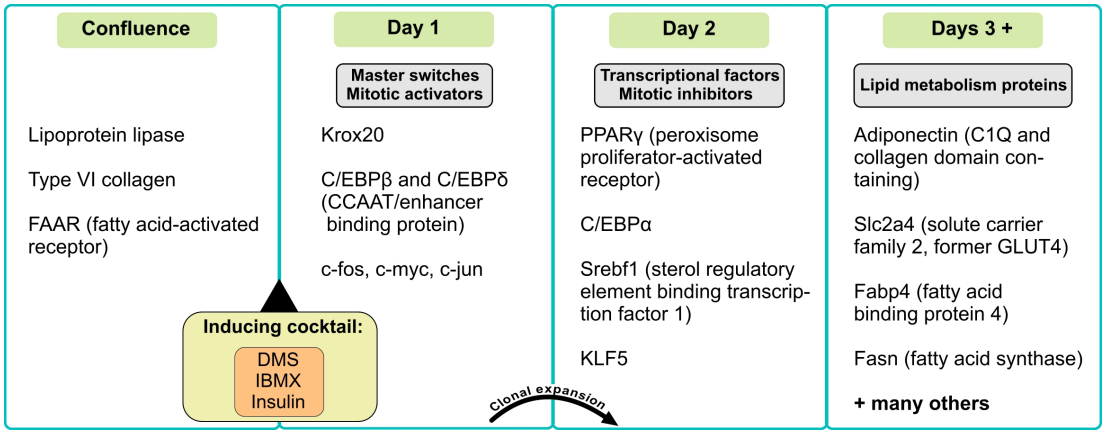


Fig 1.4.2: Overview of the 3T3-L1 differentiation cascade

Different sets of genes are expressed at different stages of adipogenesis. Most notably, Krox20 C/EBP family proteins are activated upon addition of a differentiation cocktail and activate expression of PPARG, SREBF1 KLF5. The latter, in their turn, activate expression of a large number of adipogenic genes, including Slc2a4 (GLUT4).

Different members of the C/EBP (CCAAT enhancer binding protein) family are temporally expressed during adipocyte differentiation. Early expression of C/EBPβ and C/EBPδ leads to induction of C/EBPα. C/EBPα is directly involved in the induction of some adipocyte genes and *in vivo* studies suggest it has an important role in the development of white adipose tissue (Chen *et al.*, 2000). Despite the important role of C/EBPs in adipocyte differentiation these factors are not able to function without peroxisome proliferator-activated receptor γ, PPARγ (Rosen *et al.*, 2002).

PPARγ belongs to the nuclear-receptor superfamily. It is absolutely required for adipogenesis and able to induce adipocyte differentiation in fibroblasts when overexpressed (Tontonoz *et al.*, 1994). The function of the most pro-adipogenic factors seems to be at least partially to activate PPAR expression. There are two protein-isoforms, PPARγ1 and PPARγ2, which are produced by alternative splicing and promoter usage. Both isoforms are induced during adipogenesis and PPARγ1 is also expressed in other cell types. PPARγ2 appears not to be absolutely required for adipogenesis and plays a role in insulin sensitivity (Medina-Gomez *et al.;* 2005,

Zhang *et al.*, 2004).

The expression of sterol response element binding factor (SREBF1) is regulated by C/EBP transcription factors that are expressed earlier in adipogenesis (Payne *et al.*, 2010). It induces PPARγ (Kim *et al.*, 1998; Kim and Spiegelman, 1996) and influences the induction of lipid biosynthesis by insulin in adipocytes (Kim *et al.*, 1998).

The roles of PPARγ, C/EBPs and SREBF1 in adipogenesis are now well defined, but there are over a hundred other factors expressed in adipocytes which all have a role in the differentiation process. Many transcription factors repress adipogenesis by promoting alternative cell fates and are down-regulated in adipocytes, which suggests that repression of these genes is an important task for pro-adipogenic factors. For example, addition of sonic hedgehog or its activated receptor diverges 3T3-L1 cells from differentiation into adipocytes and activates expression of osteogenic genes. The hedgehog pathway regulates gene expression through members of the GLI family and Smo (reviewed in Cousin *et al.*, 2007). Conversely, the over-expression of dominant negative GLI as well as the addition of Smo inhibitor promotes adipogenesis (Suh *et al.*, 2006).

The tight regulation of both positive and negative gene expression is crucial to effective adipogenesis and it is very likely that there are other mechanisms in place to fine tune the regulation.

1.4.3 Signalling cascades in adipogenesis

Different extracellular and intracellular factors activate different signalling pathways that in their turn activate downstream transcription factors that induce the expression of genes responsible for adipogenesis.

One of the most important regulators of adipogenesis is insulin. At the onset of adipogenesis it acts through insulin-like growth factor receptors, however with the progress of differentiation expression of a more specific insulin receptor increases (Smith *et al.*, 1988). Insulin has been shown to activate expression of SREBF1, a

transcription factor important for adipogenesis (Kim *et al.*, 1998). It also prevents activation of forkhead transcription factor FoxO1 early in adipogenesis. Active FoxO1 is able to block adipogenesis by inhibition of expression of adipogenic genes (Nakae *et al.*, 2003). Insulin also induces transcription of Krox20 through MED23 (Wang *et al.*, 2009). Krox20 is induced transiently after addition of the differentiation cocktail and its over-expression with C/EBPβ can induce full differentiation of adipocytes (Chen *et al.*, 2005).

Recent studies provide solid evidence that SMAD signalling plays an important role in adipogenesis. A Smad3 -/- mouse was shown to be protected from diet-induced obesity and diabetes (Yadav *et al.*, 2011). What's more, SMAD3-/- white adipocytes acquire the expression profile of brown adipocytes and, accordingly, contain an increased number of mitochondria. This suggests that SMAD signalling is important for switching the cell fate of pre-adipocytes towards differentiation into brown adipose tissue. Recent studies show that proteins of the inner nuclear membrane affect SMAD signalling. The NET MAN1 binds to receptor-regulated SMADs and reduces signalling by transforming growth factor-beta, activin and bone morphogenic protein (Pan *et al.*, 2005). The phosphorylation of SMADs also seems to be regulated by Lamin A and C (Van Berlo *et al.*, 2005). This shows that proteins within and associated with the inner nuclear membrane influence signalling pathways that regulate adipogenesis.

The Wnt signalling pathway is highly conserved and functions through secreted glycoproteins to influence cell fate and development. Wnt signalling was shown to inhibit adipocyte differentiation by blocking the expression of PPARγ and C/EBPα (Ross *et al.*, 2000). The Wnt signalling pathway is also affected by a NE protein. Emerin regulates the activity of a downstream target of the Wnt signalling pathway, β-catenin, by restricting its accumulation in the nucleus (Markiewicz *et al.*, 2006).

1.4.4 Chromatin rearrangement during adipogenesis

Adipogenesis is influenced by almost every important signalling pathway some of which exercise both positive and negative regulation. All of the events lead to the activation of genes involved in adipogenesis and the repression of those that are not needed, or are inhibitory to, adipogenesis. However, expression of transcription factors is not sufficient for adequate gene activation. The other important factor is accessibility of target genes and it is postulated that in every cell a part of the genome is tightly packed into heterochromatin and therefore silent. During gross changes of a cell's fate, such as differentiation, some of this heterochromatin must be made accessible, i.e. de-compacted to promote gene activation. Conversely, a part of the previously active chromatin must be assembled into higher order structures to inactivate undesirable genes and free some room inside of the nucleus.

A recent study used DNAse I hypersensitive site (DHS) analysis of genomic DNA of 3T3-L1 cells to gain information about changes in active chromatin during adipogenesis (Siersbæk *et al.*, 2011). Sensitivity to DNAse I treatment corresponds to accessibility of the genomic DNA as open chromatin is more readily subjected to DNAse I degradation. Interestingly, major chromatin rearrangements take place in the first hours after pharmacological induction (Fig 1.4.4). What's more, undifferentiated and differentiated cells have substantial similarity in the open chromatin content.

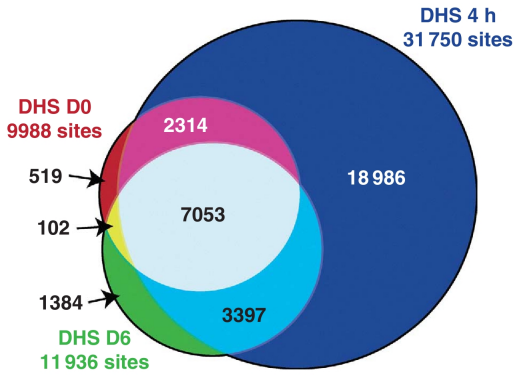


Fig 1.4.4: DNAse I hypersensitive site analysis of differentiating 3T3-L1 cells

Venn diagram representing the overlap between DHS sites in pre-adipocytes immediately before induction of differentiation (day 0; red), in adipocytes (day 6; green), and in cells stimulated for 4 h with the differentiation cocktail (blue). From Siersbæk et al., 2011.

These observations suggest that, shortly after induction a large number of chromatin

sites become open. One obvious reason for this would be activation of adipogenic genes and it was shown that sites open at 4 h correlate with induced and transiently induced genes. The latter suggests that some of the genes, despite the association with transcription factors, are not expressed at the time. This opening and association with transcription factors significantly preceding transcription could be explained by recording of an epigenetic program of differentiation. Post-translational modifications (PTMs) of histones or even histone variant exchanging could potentially mark the sites as active at certain times later in adipogenesis. In any case, opening of large areas of chromatin might be undesirable for many reasons, including overloading of transcriptional complexes and lack of intranuclear space. Thus growth-arrested 3T3-L1 cells re-enter the cell cycle, which highly condenses the chromosomes. After two rounds of mitosis, regions of the chromatin are activated, possibly in accordance with recorded earlier epigenetic marks. Hence, adipocyte differentiation involves the opening of a large number of chromatin sites followed by total condensation of chromatin. Subsequently, a different set of chromatin sites are then opened allowing the activation of alternative genes.

Some evidence exists to support involvement of the chromatin remodeling complex (RSC) in activation of PPARγ (Salma *et al.*, 2004) and SREBF1 (Lee *et al.*, 2007), but it is not clear whether opening of chromatin upon induction is achieved by this complex. Changes in gene and even whole chromosome positioning during adipogenesis have been described (Kuroda *et al.*, 2004, Szczerbal *et al.*, 2009). Thus it is clear that chromatin remodelling is important for differentiation and that the most likely outcome of the remodelling is altered gene expression.

1.5 Spatial genome organisation

1.5.1 Introduction

Considering the enormous size of the genome it could only fit into the nucleus if it is highly compacted. However it is not packed the same way in different cells or, even, in the same cell. Some of it is universally archived and represents the

constitutive heterochromatin. Some of it can be found either condensed or open, depending on a cell needs. That is so called facultative heterochromatin. The rest is euchromatin, a fraction that is always open. It is reasonable that the state of chromatin is linked to transcriptional activity of the genes located within. Housekeeping genes that are always expressed are located within euchromatin. Conversely, heterochromatin is associated with gene silencing, and constitutive chromatin represents gene poor areas and centromeric and pericentromeric regions.

Regulation of facultative heterochromatin appears to be crucial as it is the way for a cell to react to changes in conditions. What is astonishing is that every cell, despite its origin and differentiation, seems to possess the full spectrum of responses. This is derived from observations that the nucleus of a differentiated somatic cell from an adult organism, when transferred to an enucleated unfertilised egg, is able to program development of a new individual and fulfil what was programmed (Wilmut *et al.*, 1997).

Differentiation is, perhaps, the most challenging task out of all responses that a cell accomplishes upon receiving signals from outside. It involves gross changes in the gene expression profile and is often associated with changing cell morphology and function. Therefore cell systems that can be differentiated in a controlled way serve as good model systems to investigate chromatin dynamics.

1.5.2 Two distinct mechanisms of chromosome relocation

A study of localisation of two human chromosomes, 18 and 19, showed that they have distinct places inside of the nucleus (Croft *et al.*, 1999). In several cell lines, including primary lymphocytes and fibroblasts, gene-rich chromosome 19 was located preferably in the nuclear interior. On the other hand, gene-poor chromosome 18 was localised mostly at the nuclear periphery.

Chromosome 18 also appeared to occupy less volume than chromosome 19, despite containing more DNA. One suggested explanation was that because chromosome 19 is transcriptionally more active, it is located in the nuclear centre in

a more open state to allow transcription. This is in accordance with the observation that upon inhibition of transcription chromosome 19 seemed more compact. However, absence of transcription did not affect its localisation. Interestingly, chromosome 18 moves towards the centre of the nucleus when cells become quiescent or senescent and re-locates back to the nuclear periphery when the cell cycle in quiescent cells is induced, but the re-location seems to require a cell division (Bridger *et al.*, 2000). While restoration of chromosome 18 to the nuclear periphery depends upon the cell cycle, movement of the chromosome into the interior of the nucleus is rapid and complete after 15 minutes of serum starvation (Mehta *et al.*, 2010). What's more, it requires energy and nuclear myosin and, possibly, actin.

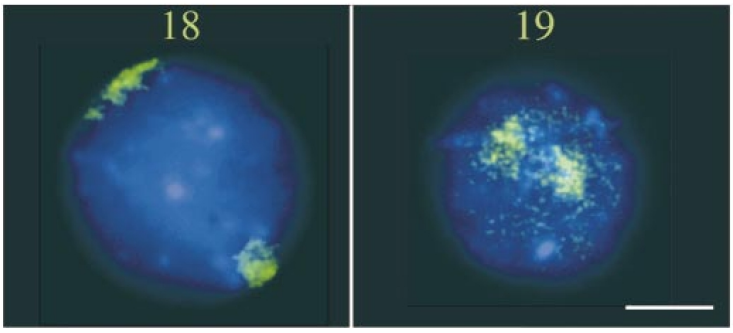


Fig 1.5.2: Human chromosome 18 and 19 position within the nucleus (Croft et al., 1999)

Human chromosomes 18 and 19 have similar size (78 Mb and 59 Mb, respectivey), yet their localisation within the interphase nucleus is completely different. Most of the chromosome 18 territory is at the nuclear periphery whereas chromosome 19 is centrally located. The cells depicted are primary lymphocytes. The scale bar is 2 µm.

Obviously, there seems to be at least two different mechanisms of chromosome relocations. One is a response to withdrawal from cell cycle. It is rapid and requires motor activity within the nucleus. The other one happens shortly after the first cell cycle of previously quiescent cells. The requirement of a cell division suggests that the INM associated proteins could play a role in this relocation mechanism. The NE breaks down during mitosis and it was shown that lamina-depleted NE membranes bind chromosomes (Foisner and Gerace, 1993). However, the described 2-4 h time window early in G1, when the movement takes place (Bridger *et al.*, 2000), seems to be contradictory to this hypothesis as affinity between a chromosome and a NET should bring the chromosome to NE upon the NE reformation in telophase. On the other hand, the delay could be explained by

the NE invaginations at the NET-chromosome attachment sites. This would suggest that the cytoskeleton, LINC complex and lamina might be involved in this type of chromosome movement. These structures are responsible for positioning of the nucleus and, therefore, are capable of transmitting forces significant for chromosome movement.

Interestingly, chromosome positioning of proliferating fibroblasts from patients with Hutchinson-Gilford Progeria Syndrome (HGPS) resembles that of quiescent control fibroblasts (Mehta *et al.*, 2011). HGPS is a rare disease that is caused by point mutations in the *LMNA* gene and leads to translation of truncated lamin A. Therefore lamin A is required for normal chromosome distribution. Moreover, the nuclear lamina has been shown to be in proximity to multiple chromosomal loci (Guelen *et al.*, 2008).

1.5.3 Gene repositioning and its effect on transcription

In uninduced murine erythroleukemia cells both alleles of the β-globin locus are associated with peripheral and centromeric heterochromatin (Francastel *et al.*, 2001). Upon terminal differentiation these loci move to euchromatic regions of the nucleus which correlates with activation of β-globin expression. Hence upon induction to differentiate one of the allele moves away from the nuclear periphery and becomes transcriptionally active. A more recent study corroborated this finding and showed that transcription of β-globin starts even before its translocation into the nuclear interior (Ragoczy *et al.*, 2006).

Zink *et al.* (2004) investigated localisation of a human chromosome 7 locus, containing three genes, in different cell lines. The localisation was compared to expression level of these genes in corresponding cells. Although transcription levels of the genes were individual, localisation of the whole locus seemed to be correlated with the sum of transcription levels of all three genes. Interestingly, a synteny locus on mouse chromosome 6 did not display a similar correlation in a follow up study (Sadoni *et al.*, 2008). On the other hand human chromosomes in mouse hybrid cells

behave similarly to their syntenic mouse chromosomes (Meaburn *et al.*, 2008; Sengupta *et al.*, 2008). This suggests that chromosome positioning is one of the features that changed dramatically during evolution of mammals.

A number of adipogenic genes are translocated towards the nuclear centre in porcine adipogenesis which correlates with their becoming more actively transcribed (Fig 1.5.3; Szczerbal *et al.*, 2009). Some of these genes were not co-localised with their chromosome territories in differentiated cells. This is in agreement with other observations that transcriptionally active genes may loop away from the main body of the chromosome they belong to. This also explains why more transcriptionally active chromosomes take up more intranuclear volume compared to gene-poor chromosomes of similar size.

Therefore, emerging evidence argues that both chromosome and gene repositioning during differentiation are able to influence gene expression in differentiation (for extended review see Schöfer and Weipoltshammer, 2008). The general idea is that less active chromosomes are located at the nuclear periphery and are more compact. Conversely, gene-rich and highly transcribed chromosomes reside closer to the centre of the nucleus and are spread over a bigger volume due to massive gene looping. However just as gene looping from peripheral chromosomes may allow their expression, accessible genes of centrally located chromosomes may not be transcribed due to lack of corresponding transcription factors or presence of repressors.

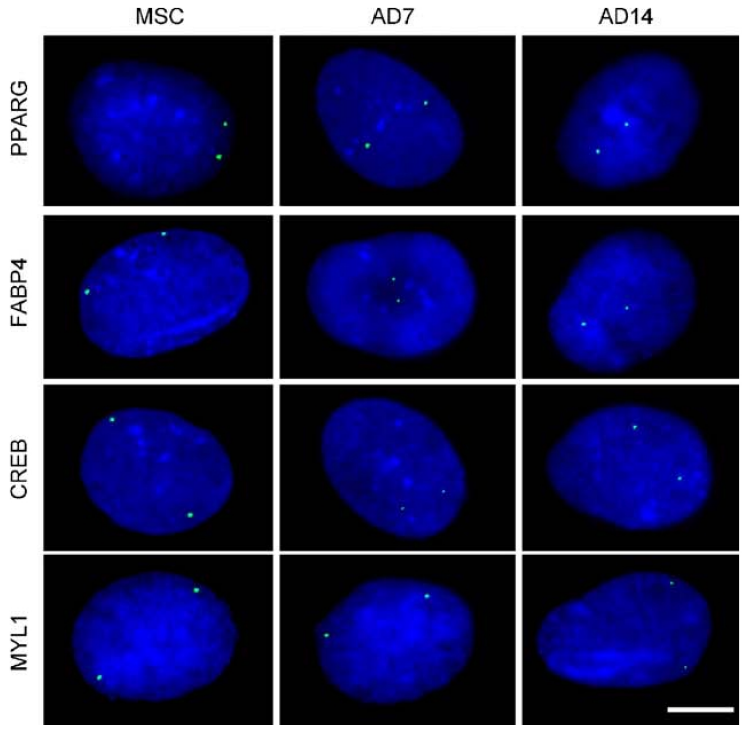


Fig 1.5.3: Localisation of adipogenic gene loci during porcine adipogenesis (from Szczerbal et al., 2009)

In blue – DNA visualised with DAPI. In green: genes visualised with specific BAC probes. MCS – mesenchymal stem cells, AD7 – preadipocytes, AD14 – adipocytes. PPARG, FABP4 and CREB – adipogenic markers, MYL1 – control, myosin light chain. Adipogenic gene loci move from the nuclear periphery during adipogenesis, which correlates with activation of their transcription. The scale bar is 10 µm.

1.6 Summary

It is clear that the NE is involved in many processes within the cell, mainly owing to the proteins that are anchored in its membranes. Some of the 13 originally identified NETs have been linked to cell division (Liu *et al.*, 2003), signalling pathways (Pan *et* al., 2005; Markiewicz *et al.*, 2006), transcription (Nili *et al.*, 2001; Holaska *et al.*, 2003), mechanical stability of the nucleus (Crisp *et al.*, 2006), DNA replication (Gant *et al.*, 1999) and DNA repair (Oza *et al.*, 2009). Proteomics identified hundreds of putative novel NETs, many of which have been confirmed to reside at the NE (Schirmer *et al.*, 2003; Malik *et al.*, 2010; Korfali *et al.*, 2010;

Wilkie *et al.*, 2010). It is likely that some of the novel NETs have new functions that link the NE to yet more vital cellular processes.

One predictable function for the INM proteins is involvement in chromatin organisation. The goal of the present work is to find novel interactions between NETs and chromatin. First, PTMs of SUN2 were identified and their effect on chromatin distribution was tested. Second, a novel protein NET29 and its effect on chromosome positioning was studied.

CHAPTER II: Materials and Methods

2.1 Materials

Chemicals were purchased from Sigma Aldrich, Calbiochem, Invitrogen and VWR. Restriction enzymes, T4 DNA ligase, T4 Polynucleotide kinase, ThermoScript™ RNAse H- Reverse Transcriptase were from New England Biolabs, Fermentas and Invitrogen. Phusion® High-Fidelity DNA Polymerase was from Finnzymes. Oligodeoxynucleotide primers were synthesised by Eurogentec and Sigma-Genosys. SybrGreen Master Mix was from Roche.

Cell culture media was from Lonza. Foetal bovine serum was bought in large stocks from HyClone. Penicillin-Streptomycin solution, MEM non-essential amino acids and sodium pyruvate were from Invitrogen.

Plasmids and their relevant features are listed in Table 2. Fugene 6 and Fugene HD were from Roche, JetPRIME reagent was from Polyplus transfection ™.

Table 2: Plasmids used in this study

Name	Type	Promoter	Features	Source
pSC-B	Intermediate cloning		Amp, Kan, lacZ	Stratagene
pEGFP	Mammalian transient	CMV early	E-GFP, Kan	Clontech
pLKO.1-puro	Mammalian transient/stable	SV40, U6	Amp	Sigma
pET28a	Bacterial expression	T7	Kan	Novagen

Isoelectrofocusing gels DryStrip and IPG buffers were obtained from GE Healthcare. Immobilon PVDF membrane was from Millipore, Odyssey nitrocellulose membrane was from Li-Cor.

Primary and secondary antibodies are listed in Table 3 and Table 4.

Table 3: Primary antibodies used in this study

Antigen	Source species	WB dilution	IF dilution	Source
β-actin	mouse	1:5000	1:500	Sigma
α-tubulin	mouse	1:5000	1:1000	Sigma
SUN2	rabbit	1:1000	1:200	Didier Hodzic
GFP	rabbit	1:1000	1:250	this study
NET29	rabbit	1:1000	N/A	Millipore
γ-H2AX	mouse	1:2000	1:800	Millipore
vimentin	mouse	1:5000	1:500	Sigma
SUN2 (new)	rabbit	1:1000	1:250	Millipore
Lamin A/C	rabbit	1:5000	1:1000	Larry Gerace

Table 4: Secondary antibodies used in this study

Antigen	Conjugate	Source species	WB dilution	IF dilution	Source
Rabbit IgG	AlexaFluor 488	goat	NA	1:1000	Invitrogen
Rabbit IgG	AlexaFluor 594	donkey	NA	1:1000	Invitrogen
Rabbit IgG	IRDye 800CW	donkey	1:5000	NA	LI-COR
Rabbit IgG	HRP	donkey	1:3000	NA	GE Healthcare
Mouse IgG	AlexaFluor 488	goat	1:1000	NA	Invitrogen
Mouse IgG	AlexaFluor 568	goat	1:1000	NA	Invitrogen
Mouse IgG	IRDye 680CW	donkey	1:5000	NA	LI-COR
Mouse IgG	HRP		1:3000	NA	Promega
Biotin	AlexaFluor 488	NA	NA	1:1000	Invitrogen
Biotin	AlexaFluor 594	NA	NA	1:1000	Invitrogen

Matrix support for affinity medium included Protein-A Agarose beads, CNBr-activated Sepharose™ 4B (GE Healthcare), streptavidin coated magnetic beads (Roche), chitin beads (New England Biolabs) and His-Select™ Nickel beads (Sigma). Affinity columns were from BioRad and Pierce.

Commercial kits included QiaFilter Plasmid Midi Prep kit, Zymoclean™ Gel DNA Recovery Kit and Illumina® TotalPrep RNA Amplification Kit.

2.2 Tissue culture techniques

2.2.1 Cell culture and transfection

Leukocytes

Human buffy coats from healthy donors were obtained from the Scottish National Blood Transfusion service. After diluting 4 times with PBS the blood was gently overlayed on top of an equal amount of HISTOPAQUE™-1077 media (Sigma) and centrifuged at 400x g for 30 min at room temperature. The white middle layer containing leukocytes was collected and washed 3 times with PBS spinning the cells down at 250 x g for 10 min at room temperature. Then the cells were resuspended in RPMI medium supplemented with 10% FBS and 100 units/mL penicillin, 100 µg/mL streptamycin to a final cell concentration of 1-5 x 10^6/ml.

HT1080, HCT116

Human fibrosarcoma cell line HT1080 and human colorectal carcinoma cell line HCT116 were maintained in Dulbecco's Modified Eagle Medium (Lonza) containing 4.5 g/L glucose and L-Glutamine and supplemented with 10% FBS and 1% penicillin/streptamycin.

Transfection of HT1080 and HCT116 cells was done using Fugene6 transfection reagent (Roche) following the manufacturer's recommendations.

3T3-L1

Mouse 3T3-L1 cells were cultured in Dulbecco's Modified Eagle Medium (Lonza) containing 4.5 g/L glucose and L-Glutamine and supplemented with 10% FBS, 0.1 M non-essential amino acids, 1 mM sodium pyruvate and 1% penicillin/streptamycin. Cells were incubated at 37°C in a humid atmosphere with 5% CO_2, split at 80-90% confluency and diluted 1:4 to 1:6 for plating.

Transfection of 30-40% confluent uninduced 3T3-L1 was done using QiaFilter™ Plasmid Midi Prep kit purified DNA. For each 10 cm plate 27µl of FugeneHD and 7.5 µg of DNA were used, for 1 well of a 6 well plate, 4 µL of

FugeneHD and 1 µg of DNA were used. The cells were allowed to express exogenous proteins for 3 days without reaching confluency. Transfection of siRNA oligos was done using Polyplus jetPRIME™ transfection reagent following the manufacturer's recommendations.

2.2.2 Pharmacological induction

Lymphocyte activation

On day 7 after lymphocyte isolation phytohaemagglutinin PHA (Sigma L1668) was added to the cells to 1% final concentration.

Double strand break induction

HCT116 cells were treated with 20 µM etoposide (Sigma). Alternatively, a short time period (5 min) of UV radiation from sterile hood built-in UV lamps was used on cells with the lid off the culture dish.

A*dipocyte differentiation*

To differentiate into adipocytes, 3T3-L1 cells were left for two days after reaching confluency, and fresh medium containing 10 µg/ml insulin, 0.5 mM IBMX and 1 µM dexamethasone was added and replaced after two days. After four days, medium containing 10 µg/ml of insulin was added and replaced every two days onward.

2.2.3 Generation of stable cell lines

Low passage undifferentiated 3T3-L1 cells were transfected using FugeneHD with linearised plasmid DNA containing the puromycin resistance marker. After two days, puromycin was added to a final concentration of 5 µg/ml and after overnight selection the medium was replaced. After that the cells were maintained in medium supplemented with 2.5 µg/ml of puromycin.

2.3 Nucleic acid techniques

2.3.1 Total RNA/protein isolation

Total RNA and proteins were isolated using TRIzol reagent following the manufacturers protocol. Briefly, cells were washed once with PBS on plates and 1 mL of TRIzol per confluent 10 cm plate was added. Then the plate was rotated for a minute or until the cells were lysed (for differentiated 3T3-L1 a cell scraper was used) and 1 mL of the lysate was transferred into a 1.5 mL microcentrifuge tube. 200 µL of chloroform were added and the tube was vigorously shaken for 30 s and phases were allowed to separate for 7 min. The lysates were centrifuged for 15 min at 12000x g, 4°C and the top half of the aqueous phase (about 300 µl) was transferred into a fresh tube. 0.5 mL of isopropanol was added, incubated for 7 min and RNA was pelleted by centrifugation at 12000x g for 10 min, 4°C. The pellet was washed once with 1 mL of 70% ethanol, air-dried and resuspended in 50 µL of water. The quality and concentration of the RNA were accessed using NanoDrop 2000c spectrophotometer. The RNA was stored at -80°C.

To recover proteins from the same lysates as total RNA samples, the rest of the aqueous phase and genomic DNA (white intermediate layer) were discarded. 1 mL of isopropanol was added to the organic bottom phase, followed by brief mixing and 5 min incubation. The proteins were pelleted by centrifugation at 12000x g for 10 min, 4°C. The pellet was washed once with 0.3 M guanidine·HCl in 95% ethanol and stored in 80% acetone at -20°C. The protein pellet was centrifuged for 10 min at 18000x g, 4°C. The supernatant was removed and the pellet was air-dried for 5 min. 200 µL of reducing buffer was added and the pellet was sonicated until it was dissolved. The resuspended proteins were stored at -20°C.

2.3.2 Reverse transcription

10 µg of total RNA was brought to 11.5 µL with water. 2 µL of reverse primers mix (6.66 µM each), 4µl of Thermoscript RNAse H– 5x reaction buffer, 0.5 µL of RNAsin, 0.5 µL of 0.1 M DTT and 0.5 µL of dNTPs mix (25 mM each) were

added and mixed by pipetting. The mix was divided into two halves, 9.5 µL each. To one half, 0.5 µL of Thermoscript RNAse H– Reverse Transcriptase was added; to the other half, 0.5 µL of water was added. The tubes were vortexed, centrifuged briefly, and incubated at 51°C in a thermocycler with heated lid at 80°C for 2 h. Upon completion of the incubation, 1 µL of 7 mg/ml RNAse A was added to each tube, and incubated at 37°C in a thermocycler with heated lid at 80°C for 30 min. 90 µL of water were added to each tube and the reactions were stored at -20°C.

2.3.3 Real-time polymerase chain reaction

Primers for real-time PCR were designed using RealTime PCR web tool (http://eu.idtdna.com/Scitools/Applications/RealTimePCR) with default parameters except for amplicon size, which was set to 100 bp minimum, 150 bp optimum and 200 bp maximum.

20 µL reactions containing 8.4 µL of diluted cDNA (serial dilutions of 5 fold), 800 nM each of forward and reverse primers and 1x LightCycler® 480 SYBR Green I Master were carried out in a LightCycler® 480 Multiwell Plate 96. The plates were briefly vortexed and centrifuged. Polymerase chain reactions were carried out in a LightCycler® 480 using the program detailed in Table 5. Primers used are listed in Table 6.

Expression data was analysed using LightCycler® 480 Software v1.5.0.39.

Table 5: LightCycler® 480 program

Target (°C)	**Acquisition mode**	**Hold (mm:ss)**	**Ramp rate (°C/s)**	**Acquisitions (per °C)**
Pre-incubation, 1 cycle				
95		05:00	4.4	
Amplification, 40 cycles				
95		00:10	4.4	
56		00:01	2.2	
51		00:15	1	
72	Single	00:21	4.4	
Melting Curve, 1 cycle				
95		00:05	4.4	

Target (°C)	Acquisition mode	Hold (mm:ss)	Ramp rate (°C/s)	Acquisitions (per °C)
65		01:00	2.2	
97	Continuous		0.19	3
Cooling, 1 cycle				
40		00:30	2.2	

Table 6: Primers used for quantitative real time PCR

Gene Name	Forward primer	Reverse primer
Adipoq	tgtctgtacgattgtcagtgg	agtaacgtcatcttcggcatg
C14orf1 (NET51)	tctacgagaagctctacactgg	tggatgtcaatggcacagag
Fabp4	caccgagatttccttcaaactg	cacgcctttcataacacattcc
Fasn	ctcaagatgaaggtggcagag	ggtcggtggctgtgtattc
Insr	ggaagctacatctgattcgagg	tgagtgatggtgaggttgtg
Lmna	cttatgcctccagtgtcacag	ggcaggtcccagattacatg
Pparg	tcacaagagctgacccaatg	atgctttatccccacagactc
Scara5 (NET33)	agcttcaagggactttctgg	tcaaagatggagccgttgtc
Slc2a4 (GLUT4)	gtaacttcattgtcggcatgg	tgctctaaaagggaaggtgtc
Srebf1	gaacctgaccctacgaagtg	tttcatgccctccatagacac
Tmem120a (NET29)	gaaaaccagatgaaagagcgc	gtgaaggagatgacgatgagg
Tmem53 (NET4)	aagaacagacaaggtgggaag	acgcgaagtgaagggatg

2.3.4 Processing of RNA samples for microarray hybridisation

RNA was labelled with biotin using the Illumina® TotalPrep RNA Amplification Kit following the manufacturer's instructions. Briefly, 500 ng of total RNA was brought to 11 µL with water and 9 µL of reverse transcription master mix, consisting of T7 oligo primer, 1x First Strand Buffer, dNTP mix, RNase Inhibitor and ArrayScript, were added. The reactions were incubated for 2 h at 42 °C in a thermocycler with 50 °C heated lid. After the incubation, reactions were placed on ice and 80 µL of Second Strand Master Mix, consisting of 1x Second Strand Buffer, dNTP mix, DNA Polymerase and RNase H, were added. The reactions were incubated for 2 h at 16 °C in a thermocycler with the heated lid off. After the incubation double-stranded cDNA was purified on a filter cartridge and eluted with 20 µL of water. The *in vitro* translation reaction was carried out in a thermocycler at

37°C with 105°C heated lid for 13-14 h. The resulting cRNA was purified on a filter cartridge and eluted with 100 µL of water. The eluate was stored at -80 °C in aliquotes.

2.3.5 Microarray data processing and analysis

Microarray data were quantile normalised and analysed in the R environment using the Bioconductor package Limma (Wettenhall and Smyth, 2004). A low intensity filter was applied, eliminating transcripts that do not reach a log2(signal) threshold of at least 6.5 in one of the samples, and we chose a low fold-difference cut-off of 1.4 (abs(log2 NET/ref)=0.5) to remove transcripts exhibiting small differences.

2.3.6 Site-directed mutagenesis

Site-directed mutagenesis was performed using the Phusion® Site-Directed Mutagenesis Kit from NEB following the manufacturer's recommendations, with the following modifications. Briefly, non-phosphorylated reverse-phase cartridge purified primers (Table 7) were used in the PCR reaction with Phusion® High-Fidelity DNA Polymerase (NEB, F-530S), the PCR product was gel purified using Zymoclean™ Gel DNA Recovery Kit according to the manufacturer's instructions, followed by simultaneous phosphorylation and ligation by T4 Polynucleotide Kinase (Fermentas, EK0031) and T4 DNA Ligase (NEB, M0202S) in T4 DNA Ligase buffer at room temperature for 2 h. Chemically competent DH5α cells were transformed with the resulting ligation mix. Selected clones were verified by Sanger dideoxy sequencing.

Table 7: Mutagenic primers

Name	Sequence	Name	Sequence
sun2.S12A.f	gcccggggtgacgatgac	sun2.S12E.f	gagcggggtgacgatgacgg
sun2.S12.r	gtagcgcgtgaggcgc	sun2.S63A.f	gcagcagatgcacacacctcctactacagt
sun2.S63E.f	gaggaggatgcacacacctcctactacagt	sun2.S63.r	cgggcccagctgtgg

Name	Sequence	Name	Sequence
sun2.S295A.f	gcagcaaactggcagaaggaggcca	sun2.S295E.f	gaggagaactggcagaaggaggcca
sun2.S295.r	aaattcagcagcaagagcttccagac	sun2.S641A.f	gcagcagcccccaaggacttcgc
sun2.S641E.f	gaggaggcccccaaggacttcgc	sun2.S641.r	gatagtgctgttgggtgacaaggc
sun2.S68A.f	gcatactacagtgagtcgctggtcca	sun2.S68E.f	gagtactacagtgagtcgctggtcca
sun2.S68.r	ggtgtgtgcatcagaggacgg		
net29Hs.PhM.f	cgaggaggagaagttcaagctctacctcac	net29Hs.PhM.r	tccttctcggcaaacttagcctgcttgc
net29Hs.Ph0.f	cgagttcgagaagttcaagctctacctcac	net29Hs.Ph0.r	tccttgaaggcaaacttagcctgcttgc

2.4 Protein techniques

2.4.1 Western blotting

The proteins were transferred onto PVDF membranes with a pore size 0.45 µm (Millipore, IPFL00010) using a Trans-Blot SD (BioRad, 170-3848) at 15 V for 25 min. The membrane was blocked with 6% BSA in PBS for 1 day, following incubation with anti-Phosphoserine mouse antibody in 6% BSA (PBS) for 1 day, washing with 6% BSA (PBS) and incubation with anti-mouse HRP conjugated antibody. In-house made ECL reagents were mixed prior to adding to the membrane. For other antibodies 6% non fat milk in PBST (0.1% Tween 20) was used as a blocking solution and incubation times were shortened. Signals were detected either with HRP or InfraRed Dye conjugated secondary antibody. HRP was detected with ECL solutions and processed using an automatic X-Ray developer. InfraRed Dye visualised bands were scanned on a LiCor scanner connected to a PC running Odyssey software.

2.4.2 Immunoprecipitation

To immobilize antibodies, ProteinA Sepharose CL-4B beads (Pharmacia Biotech AB, 17-0780-01) were used. The beads were swollen in PBS and washed extensively first with PBS and the last three washes with 1x RIPA buffer (50 mM Tris-HCl pH 7.4, 1% NP-40, 0.25% Na-deoxycholate, 150 mM NaCl, 1 mM EDTA,

1mM Na_3VO_4, 1mM NaF). Antibody was added to the beads in 1x RIPA buffer and incubated at room temperature for 1 h. The supernatant was removed and the beads were washed three times with 1x RIPA buffer.

To covalently crosslink antibody CNBr-activated Sepharose 4B (GE Healthcare) was used. 30 mg of the beads were swollen in 0.01M HCl and washed extensively to remove additives. Antibody was added to the beads in 1 ml of 0.1M $NaHCO_3$, 0.5M NaCl and incubated at room temperature for 3-5 h. Then the supernatant was removed and the beads were washed three times with 0.1M $NaHCO_3$, 0.5M NaCl. After that the beads were incubated with 0.1M Tris pH 8 at 4°C overnight. After altered washes, 3 times with each buffer – 0.1M Na acetate, 0.5M NaCl pH 4 and 0.1M Tris, 0.5M NaCl pH 8.0 – the beads were used or stored at 4°C in 0.1M Tris, 0.5M NaCl pH 8, 1mM EDTA, 0.02% NaN_3.

Isolated from buffy coats lymphocytes were either directly lysed or nuclei were isolated to perform the crosslinking assay. For lysis, the cells were spun down and resuspended in a protease inhibitor cocktail supplemented with the serine protease inhibitor isocoumarin. Then 3x RIPA buffer was added and the cells were incubated at room temperature for 20 min. Water was added dilute the RIPA buffer to 1x and the lysate was centrifuged at 3000x g for 3 min to remove insoluble material. The supernatant was applied to the immobilized or covalently crosslinked antibody. To minimise crosslinking of contaminants from the cytoplasm, isolation of nuclei was performed prior to crosslinking. Leukocytes were swollen in hypotonic buffer and burst in a loose dounce homogenizer by 10 vigorous strokes. 2.2 M SHKM and 1 M KCl were added immediately to the buffer to make it isotonic. The nuclei were pelleted at 2000x g in a swinging bucket rotor for 30 min at 4°C. To digest the chromatin, the nuclei were resuspended in 10% SHM and incubated at room temperature for 20 min with 20 units/ml of DNAse I and 10 ug/ml of RNAse A.

For the crosslinking of proteins, the isolated nuclei were incubated for 20 min at 30°C with 2 mM $CuSO_4$ and 4 mM 1,10-orthophenanthroline. To stop the reaction EDTA to 5mM was added.

Immunoprecipitation was carried out on a rotating wheel either for 1 hour at room temperature or overnight at 4°C. The beads were washed 3 times in 1x RIPA buffer and immunocomplexes were eluted by adding either Laemmli protein sample buffer (50 mM Tris pH 6.8, 2% SDS, 5% Glycerol, 0.05% Bromophenol Blue) or 3% SDS solution, supplemented with 50 mM DTT.

2.4.3 2D SDS-PAGE

The lymphocyte concentration was measured by counting cells using a haemocytometer and 40 million of resting and activated cells were incubated with 100 mM of the potent phosphatase inhibitor calyculin A (NEB) for 40 min at 37°C, 5% CO2 in humid atmosphere. The cells were then centrifuged at 250x g, 4ºC for 10 min. The supernatant was discarded and the cells were washed with 50 ml of ice cold PBS and resuspended in 50 µl of freshly prepared inhibitor cocktail (5 µl protease inhibitor cocktail Sigma D-7910, 1 mM 3,4-Dichloroisocoumarin, 20 mM activated sodium orthovanadate, 200 mM naphtyl acid phosphate in PBS). 450 µL of 1.1x lysis buffer (8.8M urea, 2.2% CHAPS, freshly added 55 mM DTT) was added and the lysate was vortexed for 1 min, briefly sonicated and stored at -80ºC. One day before the first dimension run the lysates were thawed on ice and appropriate IPG buffer was added to 1% final concentration. 250 µL of lysate was used to re-swell the first dimension gel (DryStrip pH 3-11 NL, GE Healthcare) overnight under mineral oil. The next day the gels were mounted onto a Multiphor II (Pharmacia Biotech) electrophoresis unit equipped with a water cooling system. Settings of the first dimension run are listed in Table 8.

Table 8: Isoelectric focusing program

Step	V	mA	W	Time (h:mm)	Mode
1	300	1	5	00:01	Gradient
2	300	1	5	06:00	Gradient
3	3500	1	5	05:00	Gradient
4	3500	1	5	06:00	Gradient

After the first dimension the strips were frozen and stored at -80°C. Before the second dimension run, the gels were treated with DTT and iodoacetamide to reduce cysteins and SDS to denature and negatively charge proteins. Separation of proteins based on mass was performed on an SE 400 Sturdier Vertical Electrophoresis Unit (Hoefer, SE400-15-1.5) in Tris/Glycine buffer at 40 mA for 3-5 h.

2.4.4 Silver staining

After SDS-PAGE, the gel was briefly rinsed in water and fixed for 1 hour in 40% ethanol, 10% acetic acid. Then the gel was washed twice in 30% ethanol and once in water, 20 min each wash. The gel was sensitized in 0.02% $Na_2S_2O_3$ for 1 min, washed 3 times in water, 20 s each wash. Staining was done at 4°C for 20 min in 0.2% $AgNO_3$, 0.02% formaldehyde. After staining, the gel was washed 3 times in water, 20 s each wash and moved to a different chamber. Developing was done in 3% Na_2CO_3, 0.05% formaldehyde, changing the solution when its colour was turning yellow. When developing was complete, the gel was briefly washed in water and incubated for a few minutes in 5% acetic acid. The stained gel was stored in 1% acetic acid at 4ºC. Digital imaging of the stained gels was performed using the film mode of an HP scanner.

2.4.5 TCA precipitation

To precipitate proteins, sodium deoxycholate was added protein solutions to 0.1% and mixed. 100% w/v TCA was then added to a final concentration of 25% and vortexed. The solution was placed on ice for 30 min and centrifuged at 20800x g, 4°C, for 30 min. Ice-cold acetone was added to the pellet, incubated at -20°C for 30 min and pelleted by centrifugation at 20800x g, 4°C for 10 min. The pellet was air-dried at room temperature.

2.5 Fluorescence microscopy

2.5.1 Immunofluorescence microscopy

After 30 h cells were either directly fixed for 7 min in 3.7% formaldehyde or washed with PBS, then extracted for 1 min with 20 mM HEPES pH 7.4, 110 mM KOAc, 2 mM $Mg(OAc)_2$, 1 mM EDTA, 2 mM DTT, 0.2% Triton X-100, washed again with PBS and then fixed with formaldehyde. For antibody staining, cells that were not pre-extracted were permeabilised for 6 min in 0.2% Triton X-100 after fixation. Cells were then blocked with 5% FBS in PBS and incubated for 1 h at RT with relevant antibodies. DNA was visualised with Hoechst 33342 and coverslips mounted in Fluoromount-G (Southern Biotech).

2.5.2 2D Fluorescence *in situ* Hybridization

Cells plated on 13 mm coverslips were washed with PBS and fixed with -20°C cold methanol/acetic acid (3:1) for 30 min at room temperature. The fixation step was repeated with -20°C cold methanol/acetic acid for 10 min at room temperature. Coverslips were left to dry at room temperature, incubated with 0.7 µg/ml RNAse A in PBS for 30 min at 37°C, washed 3 times with 2x SSC buffer, and the cells were dehydrated by incubation with 70%, 90% and 100% ethanol, 2 min each. Consequently, coverslips were left to dry at room temperature and then incubated for 5 min at 70°C. DNA was denatured with 70% formamide (Sigma F7503) in 2x SSC at 80°C for 7-15 min, followed by dehydration in 70%, 90% and 100% ethanol (with 70% being at -20°C). Drying at room temperature and incubation at 70°C were repeated, then the coverslips were incubated at 80°C for 5 min and moved to 42°C. 1.5µl of biotin-labelled concentrated chromosome probes (Cambio 1187-"Chromosome number"MB-02) were resuspended in 6 µL of supplied hybridization buffer containing 7 µg of mouse C_0t1 DNA and 4 µg of salmon sperm DNA, denatured at 80°C for 5 min, pre-annealed at 42°C for 3 min and spotted onto pre-heated at 42°C slides, then covered with the processed coverslips and sealed with

rubber cement. The slides were incubated in a humid chamber overnight at 37°C. After the incubation the coverslips were washed 4 times with 2x SSC buffer at 45°C, 4 times with 0.1x SSC buffer at 60°C and 1 time with 4x SSC, 0.1% Tween 20 at room temperature, 3 min each wash. Cells were blocked with 5% non-fat milk in 2x SSC for 10 min, followed by incubation with 594 Alexa Fluor conjugated streptavidin and Hoechst 33342 in blocking buffer for 1 h. After the incubation was complete, the coverslips were washed 3 times with 4x SSC, 0.1% Tween 20 for 2 min, and mounted on slides with Fluoromount-G (Southern Biothech).

2.5.3 3D Fluorescence *in situ* Hybridization

The cells on coverslips were washed once with PBS and fixed with 3.7% freshly diluted formaldehyde in PBS for 7-10 min. After fixing, the cells were rinsed with PBS once and, optionally, left for up to 7 days at 4°C with addition of 0.05% of sodium azide. Cell membranes were permebilised with 0.5% Saponin, 0.5% Triton X-100 in PBS for 10 min, followed by 3 washes in PBS. Non-specific binding was blocked with 2% BSA, fraction V, in PBS for 20 min, and anti-GFP antibody (Invitrogen) was added, followed by incubation for 1 h. The coverslips were washed 3 times with PBS, re-fixed with 1% formaldehyde in PBS for 45 s and washed twice with PBS, briefly, and once with 0.1 M Tris·HCl pH 7.4 for 10 min. After, the coverslips were incubated in 20% glycerol in 0.8x PBS and subjected to three freeze/thaw cycles, freezing in liquid nitrogen and thawing in 20% glycerol in PBS. After rinsing with PBS, the coverslips were incubated with 0.1 M HCl for 10 min, followed by incubation in 0.5% saponin, 0.5% Triton X-100 in PBS. The coverslips were rinsed with PBS and incubated in 70% formamide in 2x SSC for 15 min at 80°C. 1.5 µl of biotin-labelled concentrated chromosome probes were resuspended in 6 µL of supplied hybridization buffer, containing 15 µg of mouse C_0t1 DNA and 7 µg of salmon sperm DNA, denatured at 80°C for 5 min, pre-annealed at 42°C for 3 min and spotted onto pre-heated at 42°C slides. Coverslips incubated in formamide were mounted onto the slides, and sealed with rubber cement. Hybridization was carried out overnight in a humid chamber at 37°C. The next day, the coverslips were washed

4 times with 2x SSC buffer at 45°C, 4 times with 0.1x SSC buffer at 60°C and 1 time with 4x SSC, 0.1% Tween 20 at room temperature, 3 min each wash. Blocking was done in 2x SSC containing 0.1% Tween 20 for 5 min. After that, coverslips were incubated with AlexaFluor 488 conjugated secondary antibody, Alexa Fluor 594 conjugated streptavidin and Hoechst 33342 in blocking buffer, followed by 3 washes in 4x SSC containing 0.1% Tween 20, 2 min each wash, rinsed in PBS and mounted on slides in Fluoromount-G (Southern Biotech). Modified from Solovei *et al.*, 2002.

2.5.4 Image acquisition and analysis

Microscope pictures were obtained using a Nikon TE-2000 microscope equipped with a 1.45 numerical aperture 100x objective, Sedat quad filter set and CoolSnapHQ High Speed Monochrome charge-coupled device camera (Photometrics) connected to a PC running Metamorph software. The erosion macro for Image-Pro Plus was adopted from (Croft *et al.*, 1999) by David Kelly.

2.6 Software

2.6.1 Standalone applications

Software	Description	Source
GeneDoc	Multiple sequence alignment	http://www.nrbsc.org/gfx/genedoc/
GENtle	Multipurpose, molecular biology	http://gentle.magnusmanske.de/
Vector NTI®	Multipurpose, molecular biology	Invitrogen
JalView	Sequence alignment editor	http://www.jalview.org/
Metamorph	Image acquisition	Molecular Devices
ImageJ	Image processor	http://rsbweb.nih.gov/ij/
Image-Pro Plus	Image processor	Media Cybernetics

2.6.2 Web-based services

Service	Description	URL
NetPhos 2.0 Server	Prediction of phosphorylation	http://www.cbs.dtu.dk/services/NetPhos/
TMHMM Server v. 2.0	Prediction of transmembrane domains	http://www.cbs.dtu.dk/services/TMHMM/
Finnzymes TM Calculator	Annealing temperature determination for PCR	http://www.finnzymes.fi/tm_determination.html
RealTime PCR	Design of primers for qPCR	http://eu.idtdna.com/scitools/Applications/RealTimePCR/

CHAPTER III: Testing for SUN2 interactions with chromatin and the cytoskeleton

A distinct feature of diseases associated with NE proteins is remarkable tissue specificity. For example, mutations of Lamin A have been shown to cause lipodystrophy, dermopathy, cardiomyopathy, muscular dystrophy, nerve atrophy and progeria (reviewed in Worman and Bonne, 2007). One of the hypotheses that explains such diverse effects of mutations in an ubiquitously expressed protein postulates the presence of tissue specific interacting partners. Thus tissues lacking those partners do not show a phenotype, but in the affected tissue, dissociation of normally interacting proteins leads to severe malfunctioning. The fact that mutations in a few identified Lamin A interactors can cause the same diseases as mutations of Lamin A itself serves as a strong supporting evidence of the hypothesis (reviewed in Maraldi *et al.*, 2011).

Given nuclear localisation of the putative Lamin A binding partners and, hence, proximity to the chromatin, it is possible to speculate that at least some of the partners may alter gene expression by direct or indirect modification of chromatin. It has been shown that simple attraction of chromatin to the nuclear periphery may alter gene transcription (Ruault *et al.*, 2008; Finlan *et al.*, 2008).

The aim of the present work is to identify NETs that cause chromatin rearrangements and investigate the mechanism and consequences of this.

3.1 SUN2 PTMs

3.1.1 Human lymphocytes as a model system for chromatin rearrangement

Human peripheral blood lymphocytes provide an excellent model system for

chromatin rearrangement studies. These cells have the advantage of having two states - resting and inducible activated state. The resting state is characterised by a minimal gene expression profile and small cell size. Upon stimulation with an antigen the resting lymphocytes become activated and this induces the expression of many genes and cell growth. What's more, abundant peripheral heterochromatin of resting cells dissipates upon activation (Fig 3.1.1, A and B). Intriguingly, disruption of peripheral heterochromatin is a hallmark of many diseases associated with NE proteins (Fig 3.1.1, A). The cytoskeleton has also been shown to undergo drastic changes after lymphocyte activation (Samstag *et al.*, 2003).

Apart from this specific feature, lymphocytes are useful in research as primary cells that are easy to isolate from donor blood obtained from local transfusion centres. The cells can be easily activated *in vitro* by addition of phytohaemagglutinin (PHA), a lectin from plants. PHA binds to the T lymphocyte antigen receptor (Chilson *et al.*, 1984) and stimulates activation of cells.

However, there are also some drawbacks of using lymphocytes as a model system. First of all, purification of white blood cells from peripheral blood yields a rather heterogeneous population of cells, lymphocytes being ~70% of the total number of cells (Fig 3.1.1, C). Second, only about 20% of lymphocytes are activated after a week of stimulation with PHA (Fig 3.1.1, D). Third, primary cells are very difficult to transfect and sufficient expression of exogenous proteins is difficult to achieve.

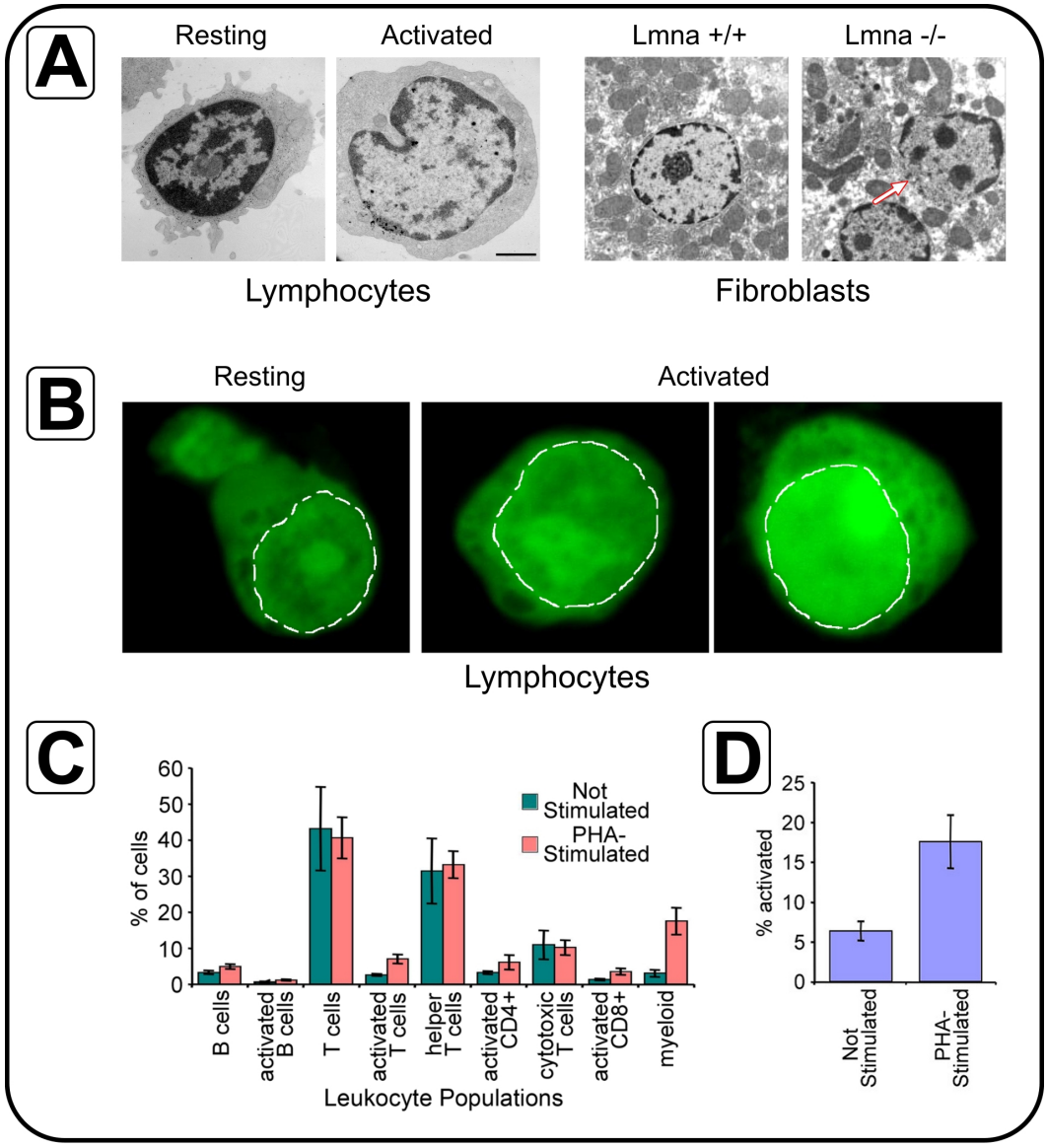


Fig 3.1.1: Lymphocytes as a model system for investigation of changes in peripheral heterochromatin distribution

A *-* ***Electron micrographs****. Upon activation of lymphocytes, dense heterochromatin dissipates from the nuclear periphery. Scale bar represents 1 µm. Fibroblasts with LMNA gene knockout also appear to have peripheral heterochromatin disturbed (arrow).* ***B*** *-* ***Acridine orange staining of lymphocytes****. Acridine orange stains single stranded RNA and DNA. White lines outline the nucleus. More intense staining in activated lymphocytes indicates chromatin decompaction.* ***C*** *-* ***FACS analysis of white cells from peripheral blood****. The leukocyte population was determined by flow cytometry after staining for specific antigens. Activated CD4+ and CD8+ cells are increased in PHA-stimulated cells.* ***D*** *-* ***Content of CD4+ and CD8+ cells in resting and PHA-activated peripheral blood cells****. Upon addition of PHA, number of activated cells increases three-fold.*

Leukocyte and lymphocyte data is from Korfali et al., 2010, fibroblast data is from Sullivan et al., 1999.

3.1.2 Differences in PTMs of some NETs between states of the leukocytes indicated by peptide recovery[1]

One of the focuses of our laboratory was the question whether NETs might be responsible for major changes in chromatin and cytoskeletal organisation during lymphocyte activation. An optimised protocol (Korfali *et al.*, 2009) was used to determine the proteome of NEs and two datasets of proteins identified in NE fractions of resting and activated leukocytes were generated and analysed (Korfali *et al.*, 2010). Out of a total 3174 identified proteins, 423 proteins were predicted to contain at least one transmembrane domain by TMHMM v2.0 (Krogh *et al.*, 2001) or a membrane anchor by SignalP v 3.0 (Nielsen *et al.*, 1997). About three quarters of the proteins were common for resting and PHA-activated leukocytes, but the rest of the proteins that were distinct to either of the states had a low peptide recovery and thus were not accounted for. Low peptide recovery might be explained by sequence specific reasons (excessive fragmentation, hydrophobicity, charges), low abundance of the protein in the sample or heavy PTMs. In either case it means that the assignment of poorly recovered proteins to one state but not the other is of low confidence. Thus our next hypothesis was that post-translational modifications of abundant proteins might explain changes in the heterochromatin distribution and/or cytoskeleton re-organisation. The process of post-translational modification of existing proteins would also ensure rapid response upon activation.

Due to the complexity of the NE proteome, initial peptide mass fingerprinting was done without any predictions of PTMs. As a consequence, modified peptides were not assigned to the proteins they belong to in the initial analysis. Taking this into account, identified peptides of most abundant NETs were mapped and aligned with serines, predicted to be phosphorylated by NetPhos v2.0 (Blom *et al.*, 1999). As a rule, there tended to be fewer or no peptides found in the regions with many predicted phosphorylation sites (Fig 3.1.2). The absence of peptides in these regions thus strongly suggests that predicted phosphorylation occurs at these sites *in vivo*. More significantly, in some of these regions with predicted modification sites,

1 Fig 3.1.2 and data associated with it is part of my Diploma thesis

peptides were found only in one of the two states of the leukocytes, either resting or activated. This suggested that in those places the phosphorylation is specific to the state where peptides were not recovered. Based on this hypothesis, prediction of putative sites of phosphorylation specific to one state of leukocytes was made. In activated leukocytes these included S49, S52, S57 and S58 of emerin, S66 and S67 of LAP2β, S267, S324 and S337 of SUN2. In resting leukocytes putative phosphorylated residues were assumed as S141, S142 and S143 of emerin and S407 of SUN2[2].

2 Number of phosphorylation sites depicted in Fig 3.1.2 was expanded at the time of writing. However, prediction of state-specific PTMs was made based on the information available to time.

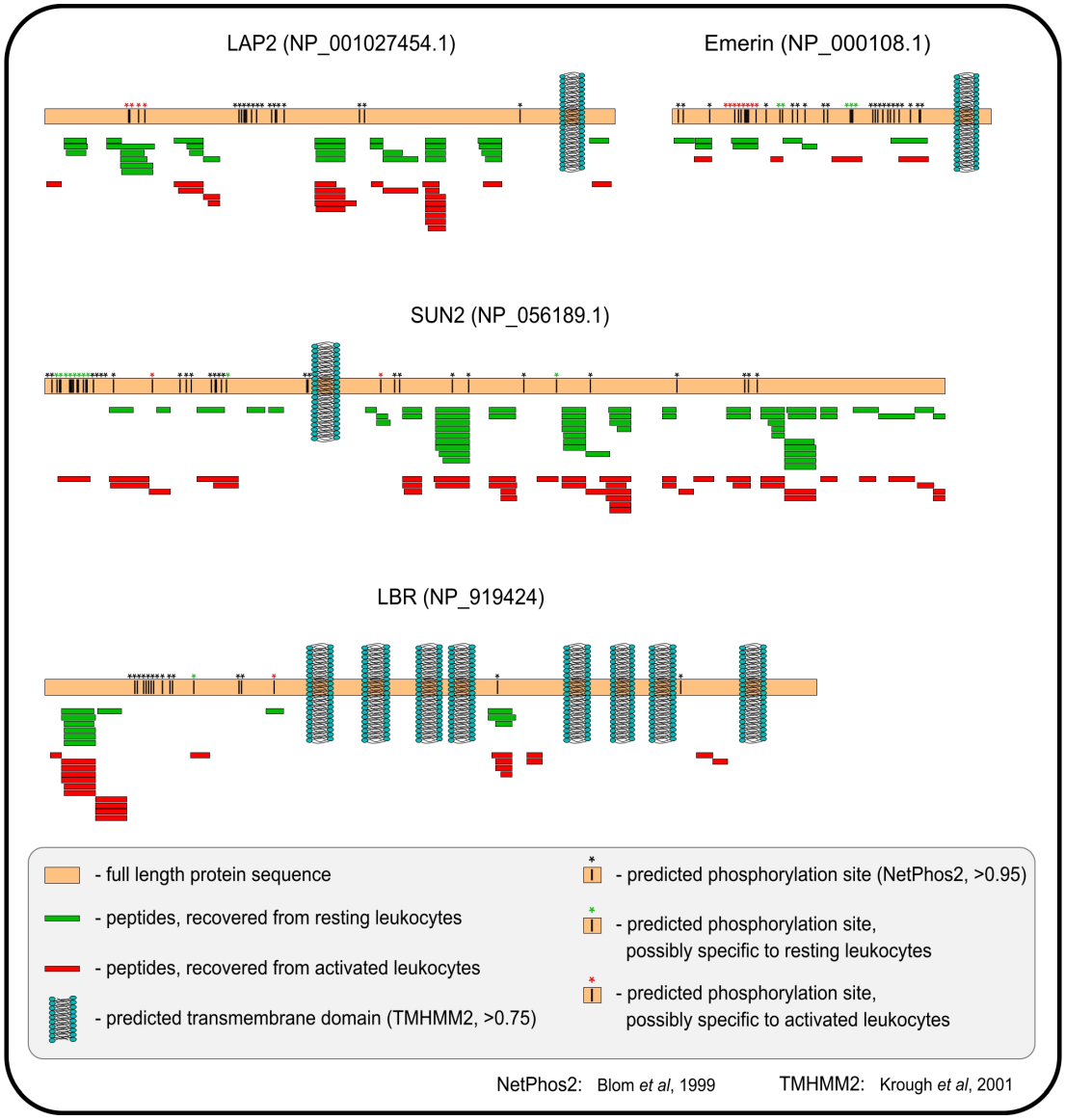


Fig 3.1.2: Recovery of peptides suggests state-specific PTMs of NETs in lymphocytes

Recovered peptides from resting and activated lymphocytes were aligned to the full-length proteins to relative scale. Serines predicted to be phosphorylated with NetPhos are depicted as stars. As a rule, there are no or few peptides recovered in the places with many serines that are predicted to be phosphorylated. However, in some of these places peptides are still found in either the resting or activated state, suggesting absence of PTMs in this state and presence in the other state of leukocytes.

3.1.3 Search for PTMs of four NETs in the NE databases[3].

As explained earlier, it is nearly impossible to identify the modifications of all peptide masses in a complex mixture of proteins such as a NE fraction. However, when a clear protein identification has been made from unmodified peptides, it is possible to look for modified peptides that belong to these identified proteins among previously unassigned masses. The whole NE MS/MS dataset was first searched in parallel for modifications against a database containing only the sequences for the proteins of interest; then MS/MS spectra matching modified peptides were searched for PTMs against the complete sequence database to make sure that the best possible matches were obtained (Fig 3.1.3, A).

First, raw files were searched without specifying differential modifications against a database combining non-redundant human, mouse and rat protein sequences downloaded from NCBI, sequences for NET proteins, and 177 sequences for usual contaminants (keratins, proteases, IgG *etc.*). In addition, to estimate false discovery rates, each non-redundant protein entry was randomized and the resulting sequences were added to the database for a final search space of 175024 amino acid sequences (Table 9). This SEQUEST "ASAP" (All Spectra against All Proteins) step allowed establishing the list of proteins detected in the samples. To account for carboxamidomethylation, +57 Da were added statically to cysteine residues for all the searches.

Table 9: Amino acid sequence database used to search the MS/MS dataset

Database	Sequences	Contaminant	Shuffled
Homo sapiens NCBI (05/09/2006)	29890	0	0
Mus musculus NCBI (22/06/2007)	29998	0	0
Rattus norvegicus NCBI (10/07/2006)	28400	0	0
NETs annotated (17/07/2007)	313	0	0
Contaminants (15/06/2005)	177	177	0
Combined (including shuffled)	175024	177	87512

3 The analysis of PTMs described in this chapter is done by Dr Laurence Florens, Stowers Institute, Kansas, USA

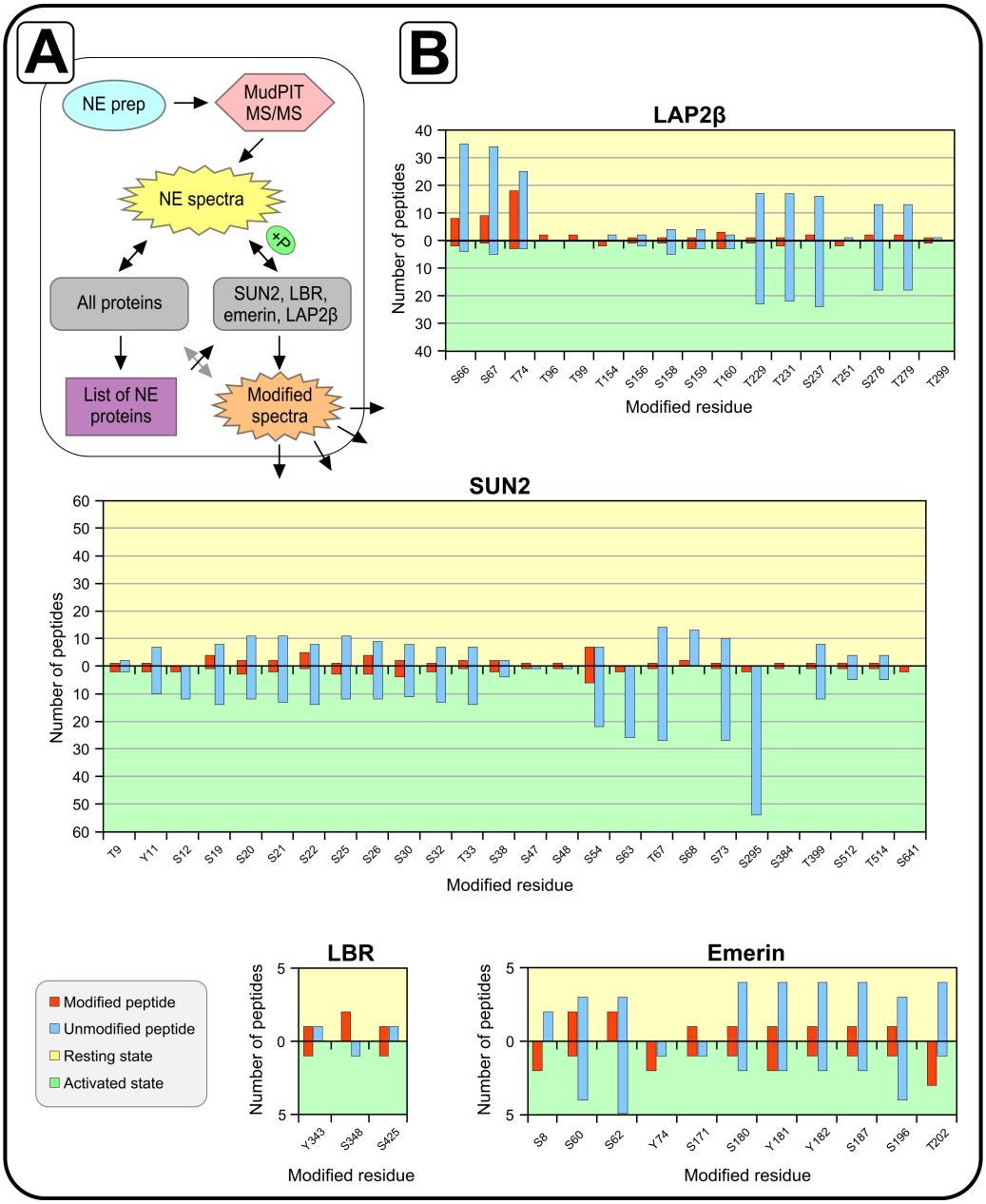


Fig 3.1.3: Analysis of phosphorylation of selected NETs

A *-* ***Phosphorylation analysis scheme****. The NE proteome was processed by MudPIT followed by tandem mass spectrometry (LC-LC MS-MS). Resulting spectra were searched against a combined human, rat and mouse database of proteins to gain the list of identified NE proteins. Out of this list, a few abundant proteins were selected and total NE spectra were searched for masses, corresponding to peptides modified by phosphorylation of the selected NETs. To eliminate false identification, spectra of the modified peptides were searched against the combined database of proteins and matches were removed.* ***B*** *-* ***Phosphorylation levels of selected NETs****. Modified and unmodified peptides of LAP2β, SUN2, LBR and emerin, recovered from resting and activated leukocytes, are plotted.*

For the third round of searches, only spectra matching modified peptides were selected. The resulting database contained a few hundreds of MS/MS spectra that could be subjected to the differential search against the complete database without having to pay too high of a computer-hour price (SEQUEST "MSAP": Modified Spectra against All Proteins). This step allowed the spectra that matched modified peptides from the NETs sequences to be checked that they did not find a better match against the larger protein database.

As a result of these three searches, a number of spectra corresponding to peptides modified by phosphorylation from SUN2, LBR, emerin and LAP2β, were identified. Data about phosphorylated residues is presented in tables 10 and 11 and depicted in Fig 3.1.3, B.

Table 10: Phosphorylated sites detected on LBR, emerin and LAP2β

Residue position	NetPhos2 Probability	Resting			Activated		
		Modified spectra count	Unmodified spectra count	Repeats (out of 7)	Modified spectra count	Unmodified spectra count	Repeats (out of 7)
LBR							
Y343	0.03	1	1	1	1	0	1
S348	0.14	2	0	1	0	1	0
S425	0	1	1	1	1	0	1
Emerin							
S8	0.9	0	2	0	2	0	1
S60	0.9	2	3	2	1	4	1
S62	0.11	2	3	2	0	5	0
Y74	0.67	0	0	0	2	1	2
S171	0.99	1	0	1	1	1	1
S180	0.98	1	4	1	1	2	1
Y181	0.03	1	4	1	2	2	2
Y182	0.07	1	4	1	1	2	1
S187	0.25	1	4	1	1	2	1
S196	0.96	1	3	1	1	4	1
T202	0.43	0	4	0	3	1	2
Lap2 β							
S66	0.98	8	35	1	2	4	1
S67	0.99	9	34	1	1	5	1
T74	0.97	18	25	2	3	3	1
T96	0.86	2	0	1	0	0	0
T99	0.5	2	0	1	0	0	0
T154	0.57	0	2	0	2	0	1
S156	0.98	1	2	1	1	2	1

Residue position	NetPhos2 Probability	Resting			Activated		
		Modified spectra count	Unmodified spectra count	Repeats (out of 7)	Modified spectra count	Unmodified spectra count	Repeats (out of 7)
S158	0.9	1	4	1	1	5	1
S159	0.99	1	4	1	3	3	2
T160	0.96	3	2	3	3	3	2
T229	0.13	1	17	1	1	23	1
T231	0.05	1	17	1	2	22	1
S237	0.9	2	16	1	0	24	0
T251	0.61	0	1	0	2	0	1
S278	0.11	2	13	1	0	18	0
T279	0.85	2	13	1	0	18	0
T299	0.02	1	1	1	1	0	1

Twenty-six phosphorylation sites were confidently identified (i.e. detected with more than one spectrum) on SUN2, twenty of which are located within nucleoplasmic N-terminus (Table 11). Two of these sites, S12 and S54, have been reported in the literature as part of a large-scale screen for phosphorylated peptides (Grønborg *et al.*, 2002), the remaining 24 sites are novel. 5 of the sites appear to be lymphocyte state-specific: S12, S63, S295 and S641 phosphorylations were detected only in activated lymphocytes, whereas phosphorylation on S68 was detected only in resting lymphocytes (Table 11, in grey).

Table 11: Phosphorylated sites detected on SUN2

Residue position	NetPhos2 Probability	Resting			Activated		
		Modified spectra count	Unmodified spectra count	Repeats (out of 7)	Modified spectra count	Unmodified spectra count	Repeats (out of 7)
T9	0.43	1	2	1	2	2	1
Y11	0.21	1	7	1	2	10	2
S12	0.99	0	0	0	2	12	2
S19	0.8	4	8	1	1	14	1
S20	0.94	2	11	1	3	12	2
S21	0.94	2	11	1	2	13	2
S22	0.96	5	8	2	1	14	1
S25	0.81	1	11	1	3	12	2
S26	0.58	4	9	1	3	12	2
S30	0.84	2	8	1	4	11	3
S32	0.01	1	7	1	2	13	1
T33	0.46	2	7	1	1	14	1
S38	0.99	2	2	1	2	4	2
S47	0.99	1	0	1	1	1	1

Residue position	NetPhos2 Probability	Resting			Activated		
		Modified spectra count	Unmodified spectra count	Repeats (out of 7)	Modified spectra count	Unmodified spectra count	Repeats (out of 7)
S48	0.5	1	0	1	1	1	1
S54	0.99	7	7	3	6	22	2
S63	0.12	0	0	0	2	26	2
T67	0.03	1	14	1	1	27	1
S68	0.81	2	13	1	0	0	0
S73	0.02	1	10	1	1	27	1
S295	0.02	0	0	0	2	54	2
S384	0.84	1	0	1	1	0	1
T399	0.2	1	8	1	1	12	1
S512	0.07	1	4	1	1	5	1
T514	0.03	1	4	1	1	5	1
S641	0.02	0	0	0	2	0	1

A fair number of phosphorylated sites were reproducibly detected in the NE leukocyte dataset on SUN2, LBR, emerin and LAP2β. Some of these modifications might be regulated by the cell state: SUN2 S12/S63/S295/S641, emerin S8/Y74/T202, and LAP2 T154/T251 are found to be phosphorylated only in the active leukocytes, while SUN2 S68, emerin S62, LAP2 T96/T99/S237, and LBR S348 seem specific to the resting cells. It should be mentioned that the analysed NE samples were prepared without phosphatase inhibitors as there was no intention to do phosphorylation analysis at the time.

Only LAP2β S66 and S67 out of these originally predicted phosphorylation sites have been confirmed by re-analysis of PTMs of the four NETs. However, there is no evidence that these residues are phosphorylated exclusively in activated cells. There are multiple reasons why the described approach may have failed to correctly identify PTMs. First of all, it used a prediction algorithm, and only phosphorylation has been accounted for. Failure of a peptide's identification, while it is identified in a similar sample, suggests modification of its mass, however, it cannot distinguish between many possible modifications. Second of all, good enrichment of the proteins of interest is necessary to obtain a large number of recovered peptides. Few peptides in one state but not the other could represent sample variation. And third of all, to ensure optimal comparison, normalization between two different samples should be

done, possibly based on total spectra.

3.1.4 2D SDS-PAGE analysis of SUN2 from resting or activated lymphocytes

We focused on SUN2 because of its abundance in lymphocytes and its known function in the cell as a part of the LINC complex. On the one hand, its N-terminus is nucleoplasmic (Hodzic *et al.*, 2004), and could potentially cause changes to heterochromatin. On the other hand, its C-terminal SUN domain interacts with the KASH domains of Nesprins. As different Nesprins have been shown to bind actin, plectin and tubulin, the SUN/KASH interaction links the cytoskeleton to the nucleoskeleton and so it is possible for PTMs of SUN2 to affect the cytoskeleton.

To confirm differences in the phosphorylation of SUN2 between the resting and activated cells, two-dimensional polyacrylamide gel electrophoresis was applied followed by Western blotting. Similar amounts of total cell lysate were first separated based on pI followed by molecular mass. Anti-SUN2 serum was used for detection of endogenous SUN2 (from Didier Hodzic). Differential distribution of detected spots supported the considerable changes in PTMs of SUN2 with activation of the leukocytes (Fig 3.1.4 a).

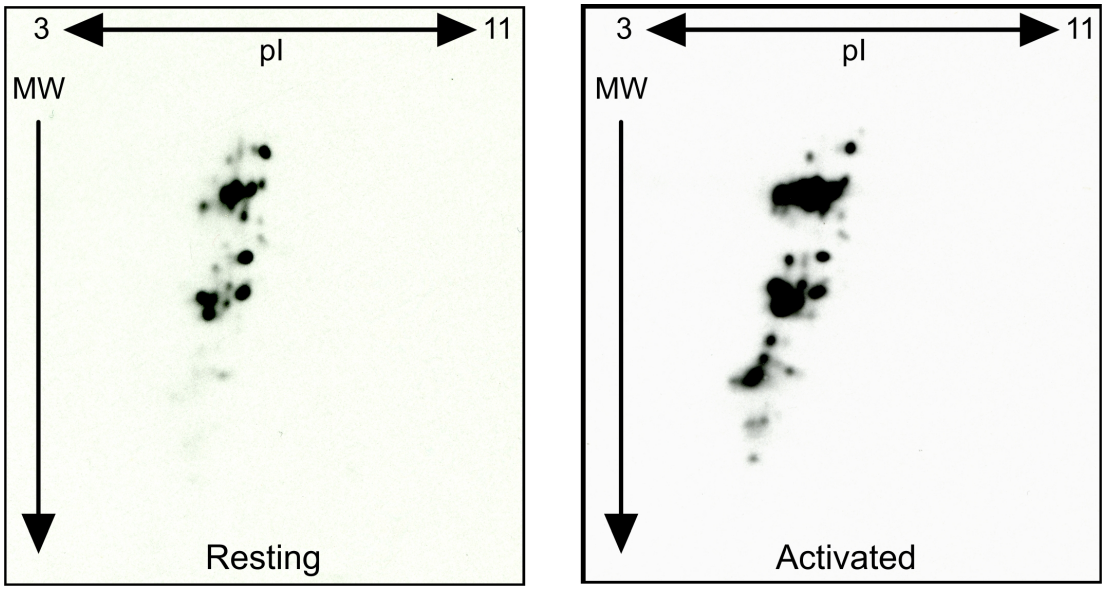


Fig 3.1.4 a: 2D SDS-PAGE analysis suggests multiple changes in SUN2 PTMs upon activation of leukocytes.

Similar amounts of resting and activated leukocytes were lysed and separated by protein pI and molecular weight, followed by Western blotting using an anti-SUN2 antibody from Didier Hodzic. The different pattern of the signals suggests multiple PTM differences between resting and activated cells.

However, even the affinity-purified antibody obtained from the serum under standard Western blotting conditions recognises two lower MW species apart from the full length SUN2 (Hodzic *et al.*, 2004). Neither of those were immunoprecipitated under more stringent conditions (Fig 3.1.4 b).

To clarify that the different phosphorylated species were not due to cross-reactivity of the antibody, the experiment was repeated using a new peptide antibody against SUN2 that was raised in collaboration with Millipore. The antibody from

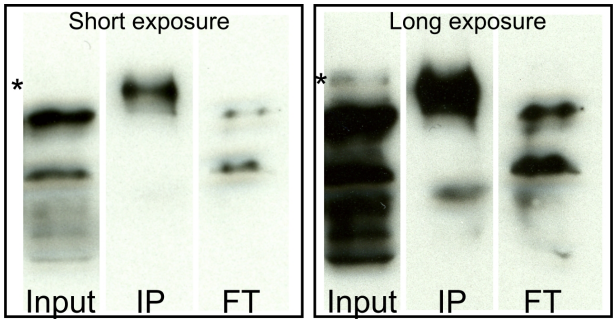


Fig 3.1.4 b: SUN2 antibody precipitates only the full length SUN2

Two exposures of the same blot are shown. Asterisk represents the expected size of SUN2. Input – cleared lysate, IP – immunoprecipitated material, FT – flow-through.

rabbit 11905 recognises an antigen that migrates according to the full-length SUN2 (Fig 3.1.4 c).

2D SDS-PAGE analysis followed by WB with the Millipore peptide anti-SUN2 antibody was repeated for resting lymphocytes only (Fig 3.1.4 d). The membrane was also probed with an anti-phosphoserine antibody (P3430, Sigma) but the signal did not co-localize with any of the SUN2 dots. However, known limitations of phosphoserine antibodies and/or lack of specific phosphatase inhibitors could explain this (only Calyculin A, activated sodium orthovanadate and naphtyl acid phosphate were used).

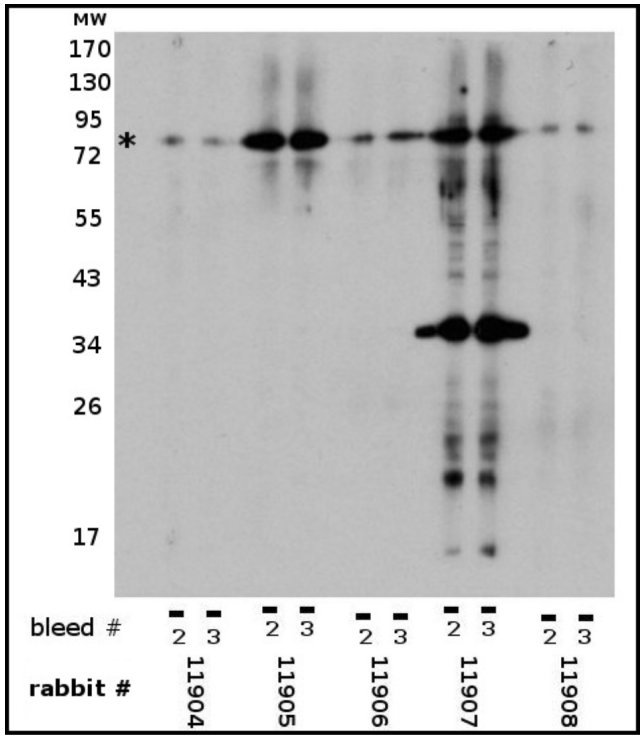


Fig 3.1.4 c: New anti-SUN2 antibody recognises a single antigen of expected molecular weight

5 rabbits were injected with a synthesised peptide corresponding to aa 1-17 of human SUN2. Rabbit 11905 developed a specific antibody.

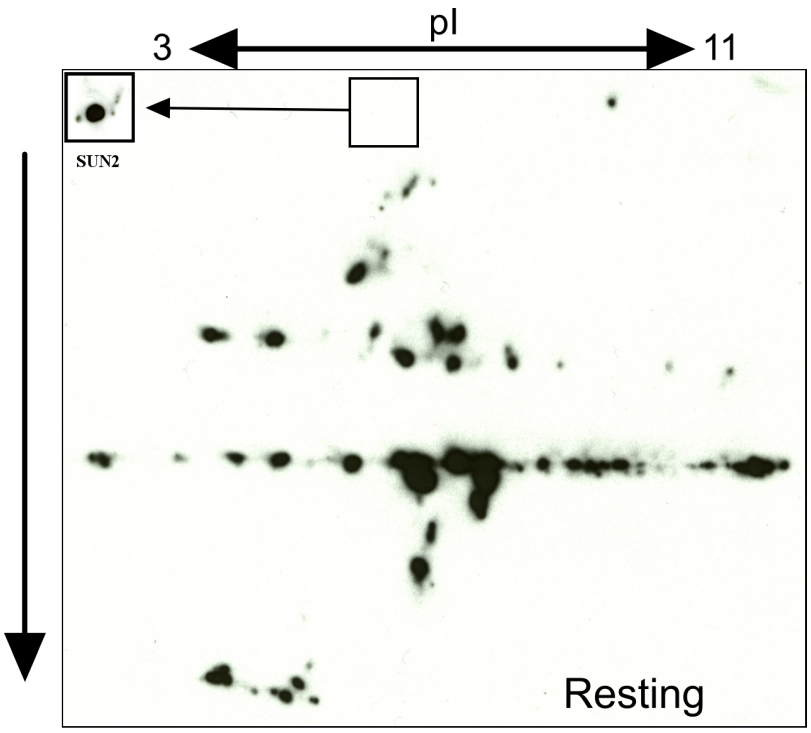


Fig 3.1.4 d: An anti-phosphoserine antibody does not recognise SUN2

The resting leukocytes lysate was resolved by pI and MW, followed by Western blotting using an anti-phosphoserine antibody. The same membrane was then probed with the new anti-SUN2 antibody (insert box). No co-localisation of the signals from the two antibodies is found.

3.1.5 Analysis of phospho-null and phosphomimetic mutants of SUN2

In order to determine whether identified phosphorylation sites are responsible for any of the changes in heterochromatin or the cytoskeleton during lymphocyte activation, site-directed mutagenesis was performed on SUN2 cDNA (from D. Hodzic) to replace serines of interest to either alanine (phospho-null mutants) or glutamic acid (phosphomimetic mutant), mimicking either of the lymphocyte states. Thus, the cDNA mimicking SUN2 from the resting state of lymphocytes contained combined mutations S12A, S63A, S68E, S295A and S641A; the cDNA mimicking SUN2 from activated lymphocytes contained combined mutations S12E, S63E, S68A, S295E and S641E.

HT1080 cells were transiently transfected with either wild type SUN2 fused

with GFP, or one of the two mutants of SUN2. The cells were allowed to express exogenous proteins for two days after transfection and then fixed with formaldehyde and stained with an antibody against histone 3 tri-methylated on lysine 9 (a mark of transcriptionally repressed chromatin), histone 3 methylated on lysine 27 (a heterochromatin mark) (Fig 3.1.5 a), vimentin (an intermediate filament) and γ-tubulin (microtubules) (Fig 3.1.5 b). In the case of heterochromatin we anticipated changes in peripheral nuclear staining similar to changes in lymphocytes upon activation. Thus, the resting state is characterised by dense peripheral heterochromatin, whereas activation leads to its dissipation. If the effect was present, it could have been quantified by a macro that is used for quantification of FISH-stained chromosomes (Croft *et al.*, 1999). In the case of the cytoskeleton, we anticipated structural changes that would lead to reduced cell stiffness. Wound healing assays would be used to quantify effect of the SUN2 mutations. However, no apparent changes in the heterochromatin or cytoskeleton stainings were noticed between cells over-expressing wild type SUN2 or its mutants.

As endogenous SUN2 was present in the cells over-expressing SUN2 mutants, only dominant negative mutant would potentially show an effect. Thus to remove endogenous SUN2 we decided to construct a plasmid that would simultaneously express an shRNA against endogenous SUN2 and either wild type or mutated SUN2 fused to GFP, modified to be resistant to the knockdown. A short hairpin RNA-encoding oligo designed based on efficient siRNA oligo (Didier Hodzic, personal communication) was cloned into a modified pLKO.1-puro plasmid (TRC1, Sigma) under the control of U6 promoter. Next, a GFP expression cassette including the CMV early promoter and SV40 poly-adenylation sequence was subcloned from pEGFP-C3 vector (Promega). However, transient over-expression of this construct in HT1080 cells and staining with anti-SUN2 antibody did not show efficient knock-down of endogenous SUN2 (Fig 3.1.5 c).

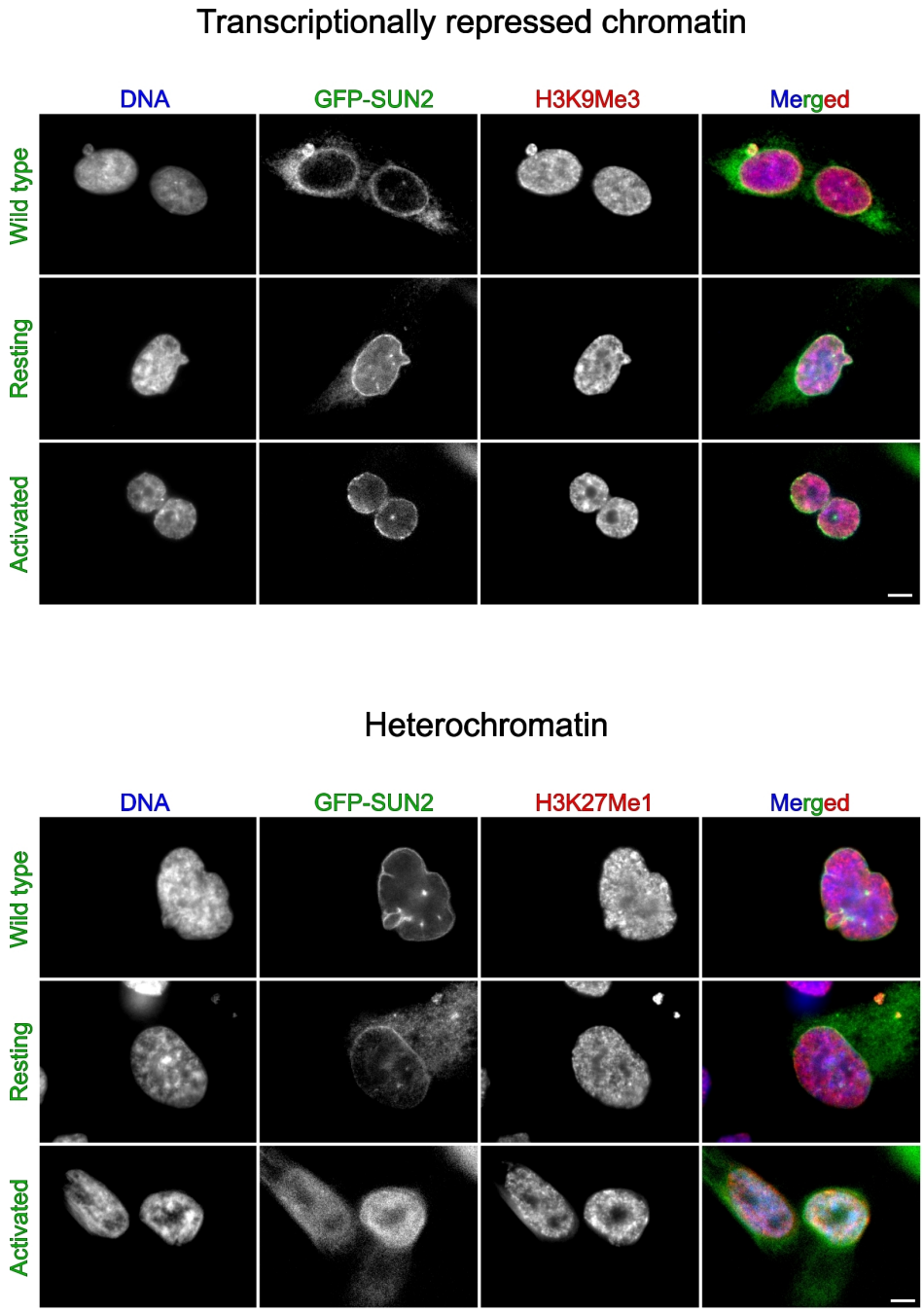


Fig 3.1.5 a: Heterochromatin is not changed upon overexpression of SUN2 mutants

HT1080 cells were transfected with either wild type SUN2-GFP fusion or SUN2 mutants (Resting – combined S12A, S63A, S68E, S295A, S641A, Activated – combined S12E, S63E, S68A, S295E, S641E) and stained with antibodies against epigenetic marks of transcriptionally repressed chromatin (H3K9Me3) and heterochromatin (H3K27Me1). No changes in distribution of the signals were observed upon overexpression of SUN2 mutants. Scale bars represent 10 µm.

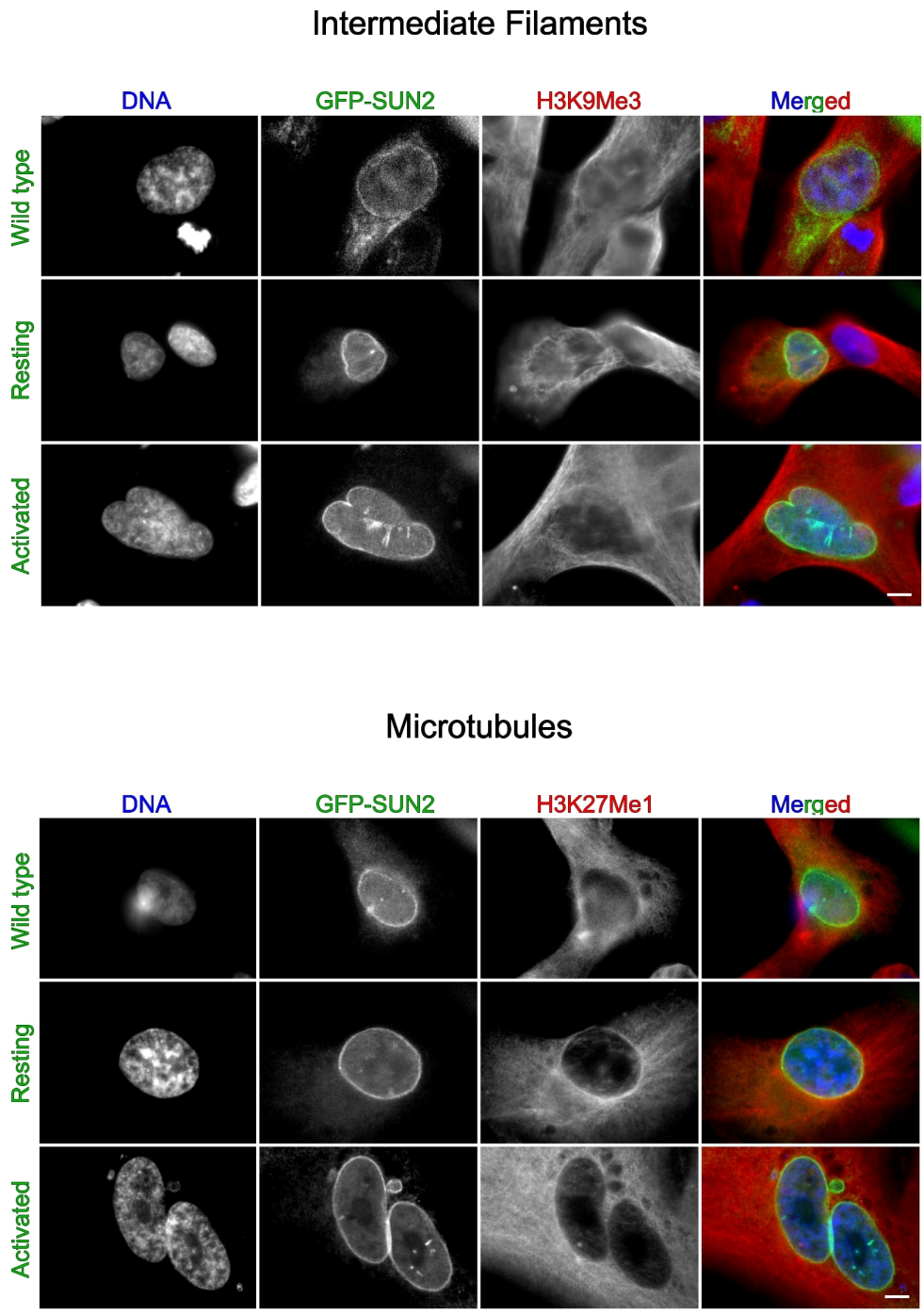


Fig 3.1.5 b: Cytoskeleton is not affected by overexpression of SUN2 mutants

HT1080 cells were transfected with either the wild type SUN2-GFP fusion or SUN2 mutants (Resting – combined S12A, S63A, S68E, S295A, S641A, Activated – combined S12E, S63E, S68A, S295E, S641E) and stained with antibodies against different components of the cytoskeleton: intermediate filaments (vimentin) and microtubules (α-tubulin). No changes in distribution of the signals were observed upon overexpression of SUN2 mutants. Scale bars represent 10 µm.

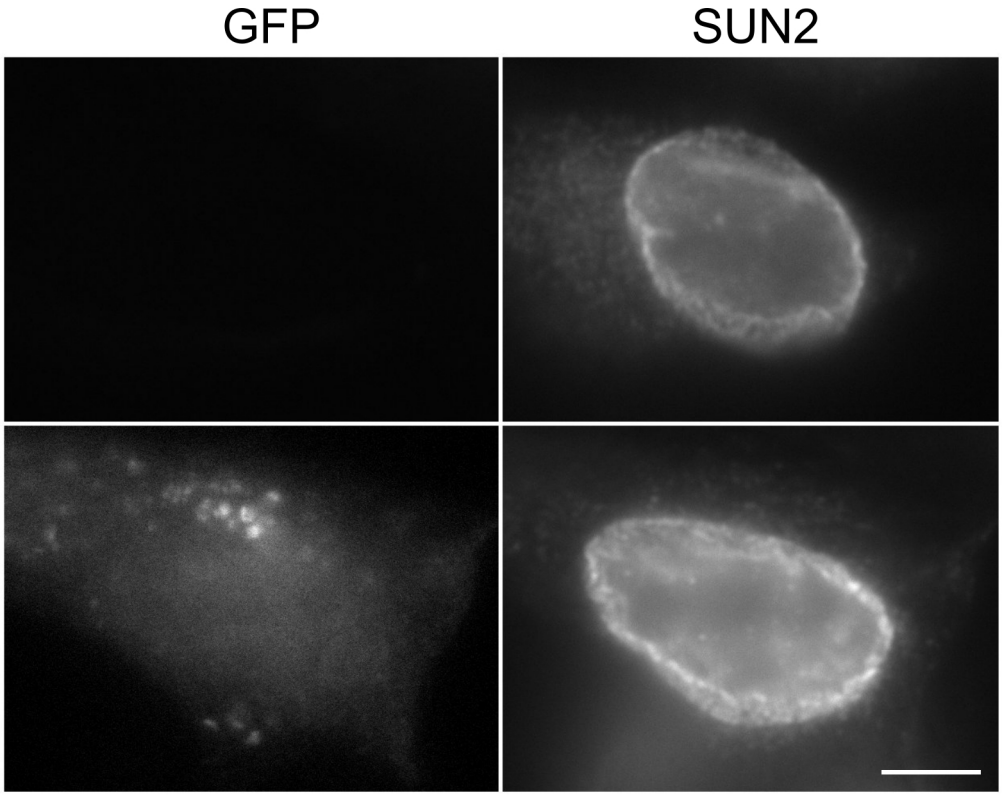


Fig 3.1.5 c: SUN2 shRNA failed to down-regulate SUN2

HT1080 cells were transfected with plasmid DNA containing an shRNA against SUN2 and GFP expression cassette. After 72 hours after transfection the cells were fixed and stained with anti-SUN2 antibody. The signal from SUN2 was similar between transfected and non-transfected cells. Scale bar represents 10 µm.

Although siRNA against SUN2 could provide sufficient knockdown, the project was not continued after this point due to multiple reasons (see below).

3.2 Potential interactors of SUN2

3.2.1 Immunoprecipitation of SUN2[4]

To identify additional binding partners of SUN2 and, possibly, identify partners that differ between the resting and activated lymphocytes, we have immunoprecipitated the protein from resting and activated lymphocyte cell lysates.

4 Mass-spectrometry in this chapter was done by Dr Laurence Florens, Stowers Institute, Kansas, USA

IPs were performed both with regular lysates and also from isolated lymphocyte nuclei with reversibly covalently cross-linked cysteines. This procedure was developed in our laboratory because many binding partners of NETs are likely to be lost with the extraction conditions required to remove the proteins from the membrane and their associations with the insoluble lamin polymer. The protein mixture that was eluted from the anti SUN2 antibody after the IP was analysed by tandem mass spectrometry. Identified proteins were sorted by number of times detected in 6 different immunoprecipitations, and only proteins that were detected at least 4 times were analysed further (the threshold was chosen based on a low number of proteins identified in immunoprecipitations with CnBr coupled antibody). Next, proteins were ranked according to enrichment in resting lymphocytes, CnBr-crosslinked antibody, crosslinked cysteines immunoprecipitation over control immunoprecipitation. The proteins that were enriched 2-fold and more are listed in Table 8.

Table 12: Highly ranked proteins from SUN2 IP

Protein	**NCBI accession number**	**ProteinA-Abs**		**CnBr-Abs**		**CnBr-Abs, Cys X-link**		**Con-trol**
Description		**Rest**	**Activ**	**Rest**	**Activ**	**Rest**	**Activ**	**Rest**
unc-84 homolog B	NP056189	275	46	65	65	538	266	32
S100 calcium-binding protein A9	NP002956	31	7	0	0	157	112	72
S100 calcium-binding protein A8	NP002955	89	11	0	0	143	606	69
H2B histone family, member S	NP059141	3	7	0	0	75	28	24
ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit precursor	NP001001937	85	16	1	3	57	102	26
histone 1, H2ad	NP066409	8	5	2	13	52	143	18
lamin B receptor	NP919424	14	3	0	0	15	16	1
profilin 1	NP005013	9	13	0	6	14	37	1
heterogeneous nuclear ribonucleoprotein H1	NP005511	6	8	0	11	11	20	4
heterogeneous nuclear ribonucleoprotein F	NP004957	4	4	0	5	11	22	4
adenylyl cyclase-associated protein	NP006358	20	4	0	0	10	32	4

Protein	**NCBI accession number**	**ProteinA-Abs**		**CnBr-Abs**		**CnBr-Abs, Cys X-link**		**Con-trol**
Description		**Rest**	**Activ**	**Rest**	**Activ**	**Rest**	**Activ**	**Rest**
desmoplakin isoform I	NP004406	0	17	1	30	7	10	2
heat shock 70kDa protein 9B precursor	NP004125	1	2	0	0	7	29	3
heterogeneous nuclear ribonucleoprotein C isoform a	NP112604	4	2	0	18	5	15	2
SAM domain- and HD domain-containing protein 1	NP056289	111	8	0	0	3	17	1
hypothetical protein LOC345651	NP001017992	75	24	28	11	28	19	21
F-actin capping protein alpha-1 subunit	NP006126	165	38	0	0	25	9	17
WD repeat-containing protein 1 isoform 1	NP059830	27	1	0	0	8	33	6
filamin 1 (actin-binding protein-280)	NP001447	639	25	0	0	189	128	409
actinin, alpha 1	NP001093	448	72	0	1	120	41	137
gelsolin isoform b	NP937895	409	15	0	9	42	40	79
PREDICTED: similar to Prostate, ovary, testis expressed protein on chromosome 2	XP934799	312	30	31	26	31	70	44
junction plakoglobin	NP068831	1	10	0	10	2	5	4

In order to analyse only immunoprecipitations with reversible cysteine crosslinking, proteins were sorted by the number on spectra identified in resting lymphocytes, CnBr-crosslinked antibody, crosslinked cysteines immunoprecipitation and proteins that were not detected in this IP were removed from analysis.

Table 13: Analysis of IP after cysteine crosslinking

Protein	**NCBI accession number**	**ProteinA-Abs**		**CnBr-Abs**		**CnBr-Abs, Cys X-link**		**Con-trol**
Description		**Rest**	**Activ**	**Rest**	**Activ**	**Rest**	**Activ**	**Rest**
lamin B receptor	NP919424	14	3	0	0	15	16	1
profilin 1	NP005013	9	13	0	6	14	37	1
PREDICTED: similar to peptidylprolyl isomerase A isoform 1	XP945701	23	0	0	0	37	25	3

Protein	**NCBI accession number**	**ProteinA-Abs**		**CnBr-Abs**		**CnBr-Abs, Cys X-link**		**Con-trol**
Description		**Rest**	**Activ**	**Rest**	**Activ**	**Rest**	**Activ**	**Rest**
H2A histone family, member X	NP006017	8	0	0	0	37	123	5
peptidoglycan recognition protein 1	NP005082	1	0	0	0	15	18	3
hexokinase 1 isoform HKI-R	NP277031	14	1	0	0	5	23	1
Fc fragment of IgG, high affinity Ia, receptor (CD64)	NP000557	2	3	0	1	8	4	0
cytochrome c oxidase subunit VIIc precursor	NP001858	0	0	0	0	6	0	0
HIG1 domain family, member 1A	NP054775	0	0	0	0	5	0	0
guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	NP006487	11	5	0	0	5	13	0
myosin, heavy polypeptide 14	NP079005	16	7	0	1	5	1	0
PREDICTED: similar to filaggrin 2	XP001127446	0	0	0	1	4	2	0
surfeit 4	NP149351	9	0	0	0	4	8	0
UV excision repair protein RAD23 homolog B	NP002865	0	0	0	0	4	9	0
solute carrier family 2 (facilitated glucose transporter), member 3	NP008862	0	0	0	0	4	12	0
hypothetical protein LOC79887	NP079105	0	0	0	0	4	0	0
GDP dissociation inhibitor 1	NP001484	0	1	0	0	3	15	0
grancalcin, EF-hand calcium binding protein	NP036330	34	0	0	0	3	2	0
translocator protein (18kDa) isoform PBR	NP000705	13	0	0	0	3	12	0
cell division cycle 10 isoform 1		4	0	0	0	3	11	0
torsin A	NP000104	4	0	0	0	3	1	0
granzyme A precursor	NP006135	0	0	0	0	3	77	0
progesterone membrane binding protein	NP006311	0	0	0	0	3	10	0
bleomycin hydrolase	NP000377	0	0	0	0	3	2	0
cytochrome c oxidase subunit VIa polypeptide 1 precursor	NP004364	0	0	0	0	3	0	0

It is obvious that too many proteins were identified as potential interactors in the SUN2 immunoprecipitations. Although SUN2 was identified with the highest

number of spectra in every sample, it was also present in the negative control (Table 12). This can be explained by its association with the cytoskeleton, and the abundant components of the latter are common contaminants in pull-down assays. As the negative control was processed in a similar way and from the same amount of cells, it is most comparable with the SUN2 immunoprecipitation from resting leukocytes with reversible crosslinking. The number of SUN2 spectra is about 15 times less in the negative control than in the IP sample. A similar spectral ratio was observed for LBR, profilin 1 and histone H2AX (Tables 12 and 13).

LBR is an interesting candidate for a SUN2 binding partner. Given its multiple membrane-spaning domains and INM localisation, it could anchor the LINC complex in the NE and, therefore, transmit pulling or pushing forces from the cytoskeleton laterally in the NE. In the support of this idea, mutations of LBR causes huclear hyposegmentation of neutrophils called Pelger-Huët anomaly (Hoffmann *et al.*, 2002). Healthy neutrophils have segmented nuclei, however homozygous mutations in the LBR gene can lead to the nuclei being round.

Profilin 1, on the other hand, is an actin monomer-binding protein that can regulate actin polymerisation and cellular motility (Bae *et al., 2010*). The protein localises in the cytoplasm and the direct binding to SUN2 could be explained only by dual localisation of SUN2 to both the INM and the ONM. However, the link between the LINC complex and the cell motility is obvious as the nucleus of a migrating cell also migrates (Meindl *et al.*, 1994).

The significance of the potential interaction between SUN2 and histone H2AX is discussed in chapter 3.3.

3.3 DNA double strand break repair and SUN2

A recent study revealed that a *S. cerevisiae* SUN domain-containing protein Mps3 is required for peripheral localisation of slowly repaired DNA double strand breaks (Oza *et al.*, 2009). However, the molecular mechanism of this is still unclear. On the other hand, mass spectrometry analysis of SUN2 immunoprecipitation

involving cysteine cross-linking identified H2AX as a potential binding partner of SUN2 (Table 13). Apart from cross-linking of cysteines, the last SUN2 immunoprecipitation protocol included isolation of nuclei and digestion of chromatin with DNAse I, which generated a large number of non-repairable DNA double strand breaks prior to cross-linking. Hence it was hypothesised that it is the phosphorylated form of H2AX, one of the proteins that directly bind DNA near DSB, which interacts with SUN2 and brings DSBs to the nuclear periphery. In support of the hypothesis, a number of proteins involved in DSB repair were detected in the immunoprecipitations (Table 14). Moreover, the peptide of H2AX containing phosphorylated upon DNA DSB serine 140 was not identified, suggesting it might have been modified post-translationally (Fig 3.3 a, A).

Table 14: Co-immunoprecipitated proteins involved in DNA double strand break repair

Lymphocytes state	**NCBI accession number**	**Resting**		**Activated**		**Ctrl**
Name \ DNAseI		-	+	-	+	+
SUN2		275	538	46	266	32
H2A.X	NP006017	8	37	0	123	5
RAD23 homolog B	NP002865	0	4	0	9	0
FK506 binding protein 12-rapamycin associated protein 1	NP004949	8	0	0	0	0
branched chain ketoacid dehydrogenase kinase	NP005872	2	0	0	0	0
flap structure-specific endonuclease 1	NP004102	2	0	0	0	0
transformation/transcription domain-associated protein	NP003487	5	0	0	0	0
ATP-dependent DNA helicase II, 70 kDa subunit	NP001460	3	3	1	20	9
ATP-dependent DNA helicase II	NP066964	4	1	0	22	8
APEX nuclease	NP542380	0	1	0	9	4

In order to probe the hypothesis, first a mammalian system with inducible DNA DSB had to be tested. Human colorectal carcinoma-derived HCT116 cells were treated either with etoposide, a topoisomerase II inhibitor, or UV, fixed with

formaldehyde and stained with a commercial anti-γ-H2AX antibody. After the treatments there was clearly an increase of the γ-H2AX signal (3.3 a, B) that seemed to be accumulate at the nuclear periphery.

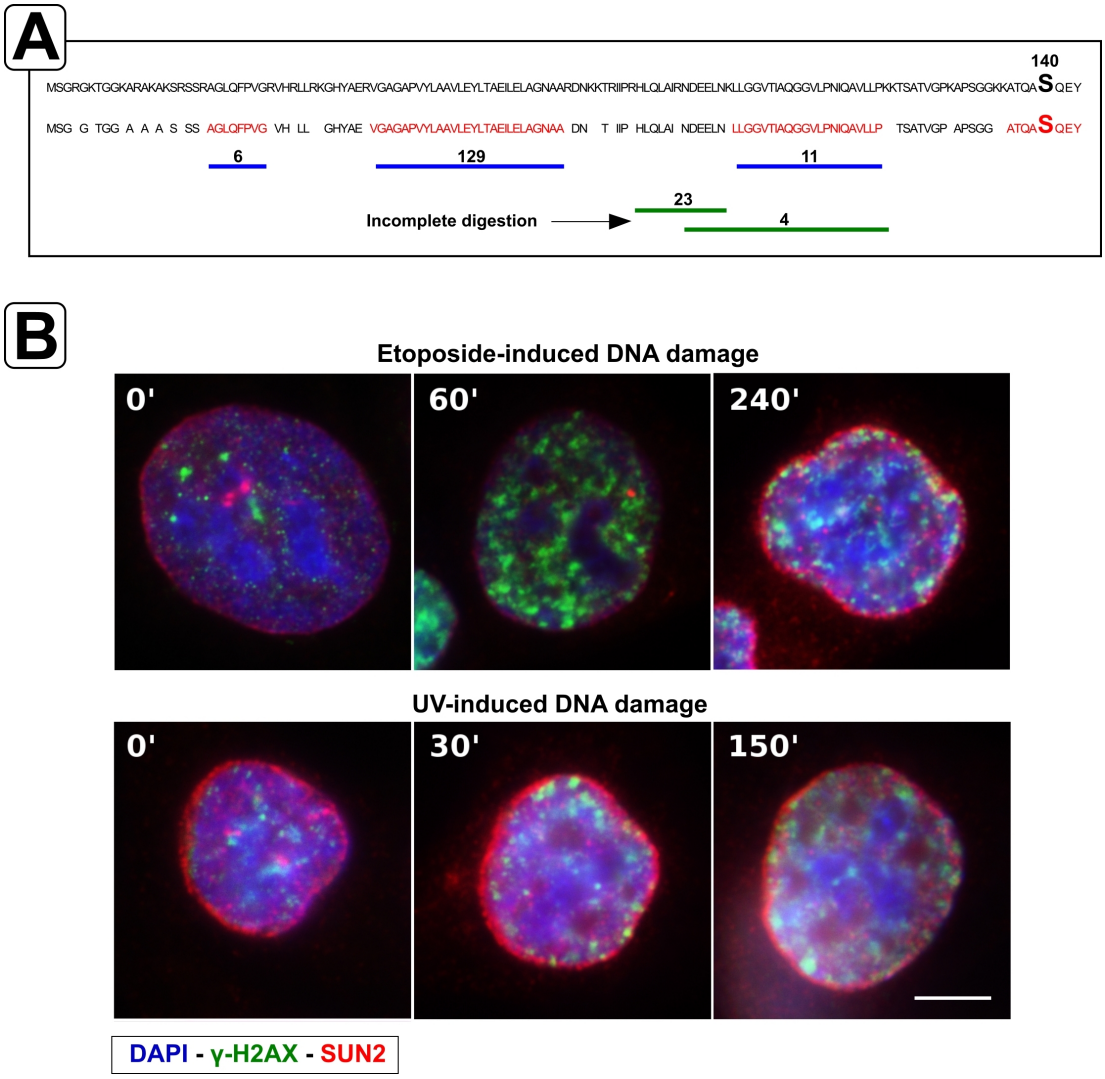


Fig 3.3 a: DNA double strand breaks are found at the nuclear periphery

***A** - **Analysis of peptides recovered from H2AX co-immunoprecipitated with SUN2.** First line - H2AX sequence, serine 140 is in bigger font size. Second line - products of complete trypsin digestion of H2AX, in red are unique peptides that may be identified by mass spectrometry. Blue lines represent recovered H2AX peptides with corresponding number of spectra, green lines - identified products of incomplete trypsin digestion. **B** - **Etoposide- and UV-induced DNA DSB can be found at the nuclear periphery.** Both etoposide and UV induced DSB are accumulated at the NE after prolonged periods of time after their induction. It is consistent with the hypothesis that unrepairable DSB are translocated to the NE. Scale bar represents 10 μm.*

Second, the distribution of the γ-H2AX-positive dots was to be tested upon SUN2 knock-down. However, as the cloned shRNA against SUN2 did not reduce the level of SUN2 significantly, it was decided instead over-express soluble N-terminus of SUN2 and look for its effect on the distribution of the γ-H2AX. SUN2 was reported to be a type II transmembrane protein with N-terminus in the nucleus and C-terminus in the nuclear envelope lumen (Hodzic *et al.*, 2004). In order to predict transmembrane regions two different tools were used: TMHMM v2.0 (Krogh *et al.*, 2001) and TMpred (Hofmann and Stoffel, 1993). While TMHMM predicted strongly one transmembrane domain (aa 213-325), TMpred strongly predicted 3 transmembrane regions (aa 159-177, 183-204, 212-234). As both variants could be applicable for a type II transmembrane protein, three SUN2 deletion mutants were generated - amino acids 1 to 158, 1 to 182 and 1 to 213 of SUN2 fused to GFP at the C-terminus. The resulting constructs were transfected into HCT116 cells and allowed to express for one day. The cells were fixed with formaldehyde and GFP signal was monitored. While both SUN2 1-182 and SUN2 1-213 constructs showed rim staining, the SUN2 1-158 construct localised mainly in the nucleus (Fig 3.3 b). This suggests that SUN2 has 3 transmembrane domains.

The soluble N-terminal fragment of SUN2 was then tested for co-localisation with γ-H2AX-positive dots with or without etoposide treatment. However, no correlation between γ-H2AX and SUN2 signals was found (Fig 3.3 b). This suggests that the N-terminus of SUN2 does not interact with γ-H2AX. Furthermore, it was reported that the DNA ends that appeared as a result of a double strand break do not move significantly in mammalian cells (Soutoglou *et al.*, 2007). Therefore it seemed less likely that the yeast mechanism of DSB tethering to the NE is functional in higher organisms.

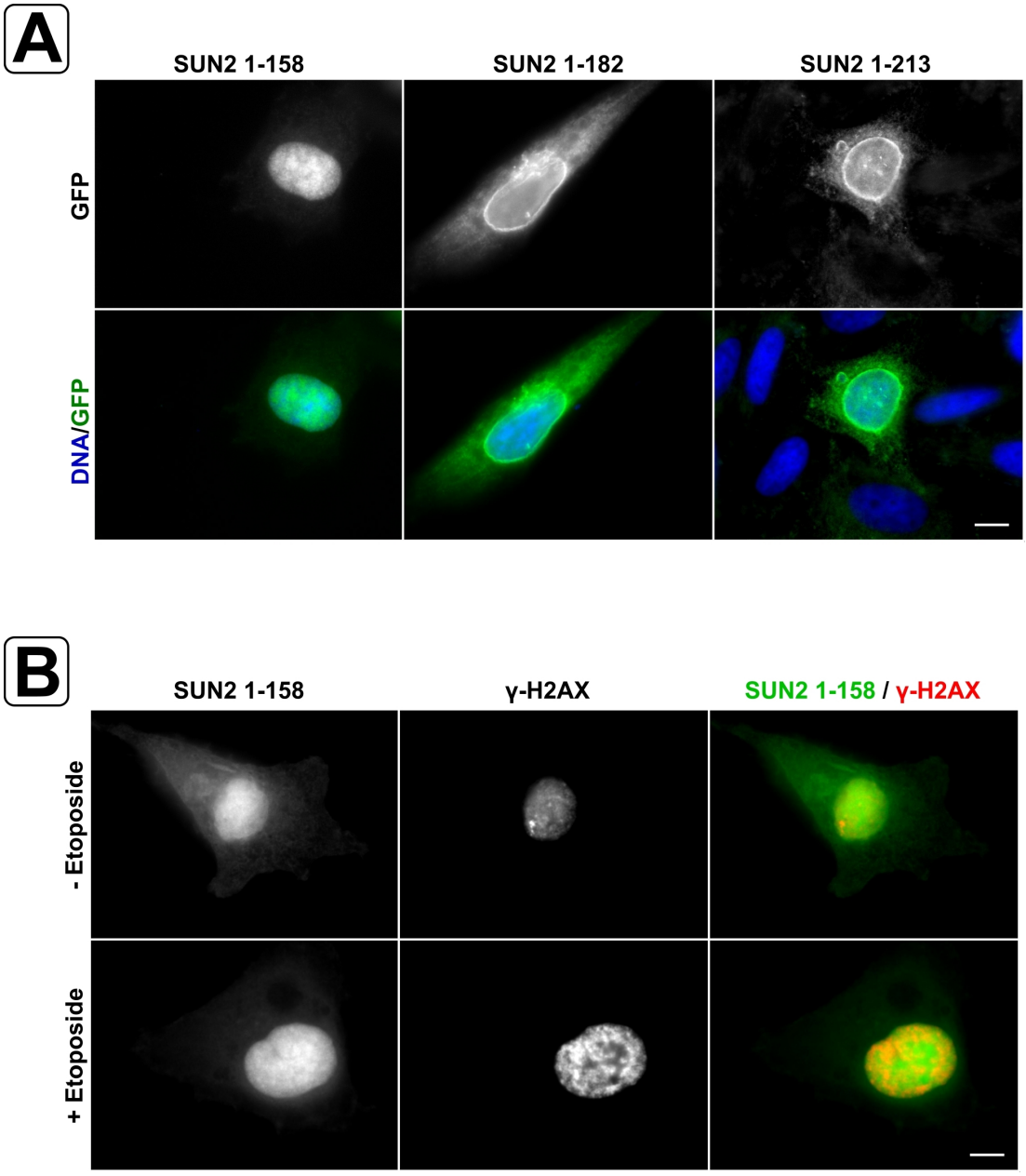


Fig 3.3 b: SUN2 NTD is not enriched within γ-H2AX-positive dots

***A** - **N-terminus of SUN2 has mostly nuclear localisation.** SUN2 is most likely to have three transmembrane domains, as predicted by Tmpred. **B** - **No apparent correlation found between localisation of N-terminal fragment of SUN2 and γ-H2AX.** Etoposide-induced DNA double strand breaks are marked by phosphorylated H2AX. GFP fusion of the soluble nucleoplasmic domain of SUN2 does not seem to co-localise with γ-H2AX. Scale bars represent 10µm.*

3.4 Summary and conclusions

The SUN2 phosphorylation data presented in this chapter is unclear and contradictory. On the one hand, the search within the whole NE datasets identified masses corresponding to the phosphorylated peptides of SUN2. Two of the 26 identified phospho-sites were identified previously in a high-throughput analysis of phospho-peptides that were captured using phospho-specific antibody (Grønborg *et al.*, 2002). On the other hand, Western blot analysis of SUN2 in resting leukocytes showed no co-staining with anti-phosphoserine antibody. Furthermore, mass-spectrometry analysis of SUN2 immunoprecipitations did not confirm any phospho-sites on the protein. The immunoprecipitation samples are much less complex than the NE fractions and, therefore, the data obtained from the former should be more confident. Thousands of proteins were identified in NE preparations, and the vast majority of masses recorded (about 90%) could not be assigned to the identified proteins due to multiple reasons, including non-unique peptides, modifications of peptides and inaccurate mass identification due to charge or hydrophobicity issues. Moreover, while SUN2 IPs were done in the presence of the activated sodium orthovanadate and alpha-naphthyl acid phosphate, general phosphatase inhibitors, the isolation of NEs was done without phosphatase inhibitors in most of the repeats. However, differences in the buffers used for immunoprecipitations and the NE isolations could also explain the contradiction. Also, it was impossible to control for similarity between the donor blood used for leukocyte isolation as the samples were anonymous.

Mass spectrometry analysis of SUN2 immunoprecipitations identified too many proteins as potential binding partners. The covalent cross-linking approach relied on cysteines from SUN2 and its potential binding partner to lie within a short distance, and the occurrence of cysteines is low in an average protein. There are only 4 cysteines in SUN2, one of which is most probably within the transmembrane domain, and none in the nucleoplasmic N-terminal fragment. Hence, the probability of successful crosslinking is quite low. Although the hypothesis of SUN2 mediated

tethering of DNA double strand breaks to the NE is not likely to apply to mammalian cells, H2AX is still a potential SUN2 partner. However a lack of a sufficient amount of anti-SUN2 antibody and further evidence of the interaction stopped us from a more careful investigation.

At this point it was inconclusive whether the project was viable. Further work would include SUN2 immunoprecipitation in denaturating conditions with subsequent mass spectrometry-based analysis of phosphorylation. Different crosslinking reagents would be tested in order to identify SUN2 partners more conclusively.

On the other hand, a screen in our laboratory identified several novel NETs that affected the localisation of a LacO array integrated into human chromosome 5 (N Zuleger, S Boyle, DA Kelly, J de Las Heras, DG Batrakou, V Lazou, GR Otti, N Korfali, KN Randles, GE Morris, DJ Harrison, WA Bickmore and EC Schirmer, paper in submission). Thus, NET29 tethered the LacO array to the nuclear periphery upon overexpression. Moreover, transcriptome data (Wu *et al.*, 2009) suggested a tissue-specific expression of NET29, raising the possibility of tissue-specific chromatin organisation orchestrated by the NE. After consideration, I decided to focus my research on NET29 and its link with chromatin. The goal of this project is to uncover the significance of NET-chromatin interactions in cell differentiation.

CHAPTER IV: NET-chromatin interactions in adipogenesis

4.1 Several NETs are upregulated in adipogenesis

4.1.1 Tissue specificity of novel NETs

One of the hypotheses to explain tissue-specific phenotypes of the nuclear envelopathies postulates that there are tissue-specific NETs. In order to test this NE proteomes from 3 different tissues were purified and analysed by MudPIT followed by tandem mass spectrometry (Schirmer *et al.*, 2003; Korfali *et al.*, 2010; Wilkie *et al.*, 2010). Supporting this initial hypothesis, only about a third of putative NETs was shared between the tissues analysed.

To corroborate further that novel NETs might be tissue-specific or tissue-preferential, transcriptome data was extracted from the public database bioGPS (Wu *et al.*, 2009). As was anticipated, some of the NETs have tissue-preferential expression based on mRNA expression profiles (Fig 4.1.1 a). For example, NET4, NET29 and NET33 are expressed highly in adipose tissue. As fat is one of the tissues affected by mutations in Lamin A, it was decided first of all to confirm upregulation of the three NETs in adipocytes and, second of all, investigate their role in fat cells.

To confirm upregulation of NET4, NET29 and NET33 in adipocytes, the 3T3-L1 inducible system was chosen. 3T3-L1 is a mouse cell line established from NIH 3T3 cells (Green and Kehinde, 1975). The cells have fibroblast-like morphology and can be rapidly differentiated into adipocytes by addition of a drug cocktail con-

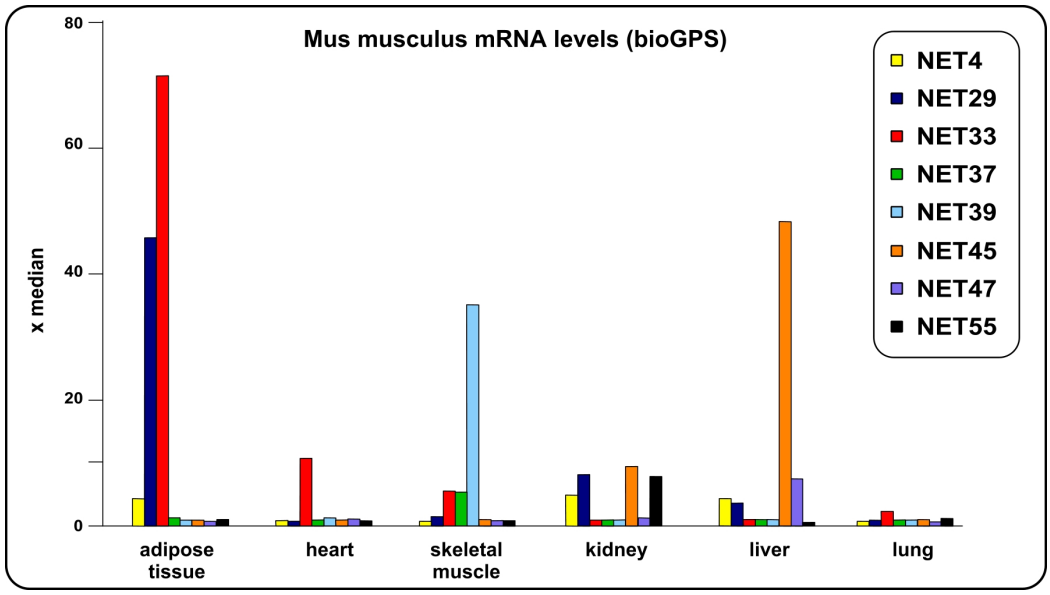


Fig 4.1.1 a: Transcription profiling of some novel NETs

Data extracted from BioGPS database (Wu et al., 2009) and plotted in medians of 78 different mouse tissues and cells. Median is a middle value of a list of values. Representative tissues were selected.

sisting of a glucocorticoid dexamethasone (DMS), a phosphodiesterase inhibitor 3-isobutyl-1-methylxanthine (IBMX) and hormone insulin (Rubin *et al.*, 1978). Some of the key players of differentiation have been identified and the order of their upregulation is well described (reviewed in Fajas *et al.*, 1998). Differentiated cells change their shape and accumulate lipid droplets that can be visualised with lipid dyes (Fig 4.1.1 b).

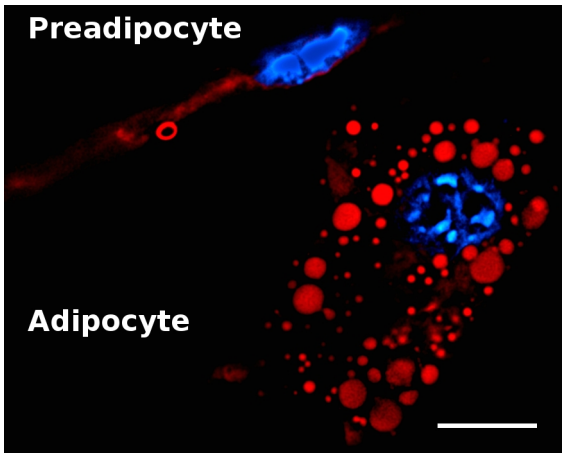


Fig 4.1.1 b: Undifferentiated and differentiated 3T3-L1 cells

Differentiated adipocytes accumulate lipid droplets that can be visualised with lipophilic dyes (here – Nile Red). Scale bar represents 10 µm.

4.1.2 Upregulation of adipogenic markers

In order to confirm the reproducibility of results it was decided to first assess expression of adipogenic markers by reverse transcription followed by quantitative PCR. RNA from undifferentiated sub-confluent 3T3-L1 cells or cells differentiated for 1, 2, 3, 4, 6 and 8 days was isolated and reverse transcribed using gene-specific reverse primers. First, primer pairs for different genes have been calibrated on cDNA obtained from reverse transcribed RNA from 3T3-L1 cells differentiated for 6 days. RNA from differentiated cells was used for calibration as some of the adipogenic markers are not expressed in preadipocytes and therefore could not be amplified from undifferentiated cells. Primer efficiency was between 1.879 for *Pparg* and 2.047 for *Srebf1* (Fig 4.1.2, A). To ensure specificity of the amplified product, melting curves also have been measured at the end of all reactions and analysed (Fig 4.1.2, B).

The amount of synthesised cDNA was measured by SYBR Green I based qPCR on LightCycler 480 apparatus. *LMNA* cDNA was used as a reference; the primer set was specific to lamin A and did not amplify lamin C, which is a splicing isoform of the *LMNA* gene.

A good correlation between published data on adipogenic markers and our qPCR results was obtained (4.1.2, C). Two master regulators of adipogenesis, peroxisome proliferator-activated receptor gamma (*Pparg*) and sterol regulatory element binding transcription factor 1 (*Srebf1*) are upregulated on the second day after addition of differentiation cocktail. The rest of the tested genes, expressing proteins of lipid metabolism such as fatty acid synthase (*Fasn*), fatty acid binding protein 4 (*Fabp4*), solute carrier family 2 (facilitated glucose transporter), member 4 (*Slc2a4*, G*lut4*) and adiponectin (*Adipoq*), were upregulated later in adipogenesis.

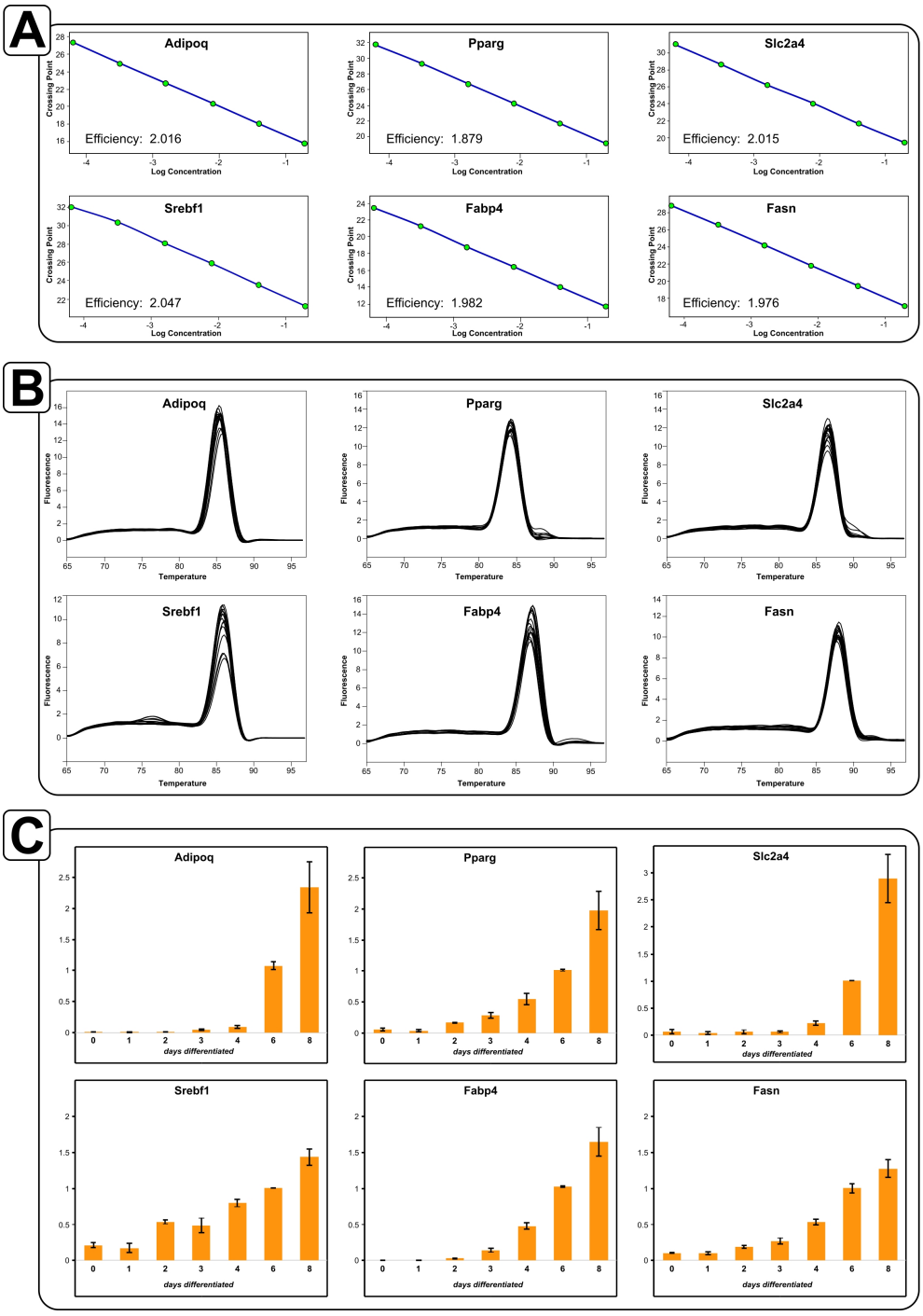


Fig 4.1.2: qRT-PCR analysis of adipogenic markers in adipocyte differentiation

A - Callibration curves of primer pairs used for PCR. *Callibration curves are built to determine efficiency of the primer pairs. Efficiency should be close to 2 with an acceptable range of 1.8-2.1.* ***B - Melting curves of amplified products.*** *Specificity of amplified products is assessed upon reaction completion. Single melting peaks represent high specificity of primer pairs.* ***C - Relative mRNA level of corresponding adipogenic markers.*** *Relative levels of adipogenic markers mRNA were detected with SYBR green and ratio to LMNA mRNA level is plotted.*

4.1.3 NETs 4, 29 and 33 are upregulated in adipogenesis

Having established the model system, expression of NETs 4, 29 and 33 was assessed the same way. First, primer pairs were calibrated and their efficiency was calculated, ranging from 1.979 for NET33 to 1.99 for NET4 (Fig 4.1.3, A). Single peaks with melting curve genotyping were obtained (Fig 4.1.3, B) supporting specificity of the amplified products.

After that, cDNA levels of the three NETs have been measured in reverse transcribed RNA isolated from subconfluent undifferentiated 3T3-L1 cells and cells differentiated for 1, 2 ,3, 4, 6 and 8 days. cDNA levels are plotted relative to lamin A in Fig 4.1.3, C. All three NETs are upregulated during adipogenesis, but each one has a different pattern.

NET4 is upregulated on day 6, similar to Adipoq and Slc2a4 (Glut4). NET4, or transmembrane protein 53 (*Tmem53*), is a 277 aa protein with one transmembrane domain strongly predicted by TMHMM2 and two other hydrophobic regions. It also contains a domain of unknown function DUF829. It has been shown by overexpression and downregulation with siRNA to alter cell cycle progression and change the level of phosphorylated Rb (Korfali *et al.*, 2011).

NET29 is upregulated on day 2, similar to *Pparg* and *Srebf1*, a master regulator of adipogenesis and a key transcription factor. NET29 is transmembrane protein 120a (*Tmem120a*), also known as transmembrane protein induced by tumor necrosis factor α (*TMPIT*), however no evidence of it being regulated by TNFα was published. It is 343 aa long and has 5 transmembrane regions predicted by TMHMM2. Interestingly, this protein has been identified to alter the chromosome 5 position in HT1080 cells with a LacO array integrated into chromosome 5 (N Zuleger, S Boyle, DA Kelly, J de Las Heras, DG Batrakou, V Lazou, GR Otti, N Korfali, KN Randles, GE Morris, DJ Harrison, WA Bickmore and EC Schirmer, paper in submission).

NET33 is upregulated already on day 1 of adipogenesis, however changes in its expression are less than that of NET4 and NET29. NET33 is known as a

(putative) scavenger receptor class A, member 5 (*Scara5*). It is 495 aa long and has one strongly predicted transmembrane domain. It also has been found to have an effect on the chromosome 5 position in HT1080 cells, albeit weaker than that of NET29.

After consideration, I have decided to focus on NET29 for a number of reasons. First of all, it is upregulated at the same time as important adipogenesis factors as *Pparg* and *Srebf1* (Fig 4.1.3, D). Second of all, the chromosome repositioning effect it possesses might explain its role in adipogenesis. And third of all, we have an affinity purified antibody generated against a NET29 peptide that recognises a protein of the correct molecular weight. It was used to confirm that NET29 protein level is increased during adipogenesis (Fig 4.1.3, E).

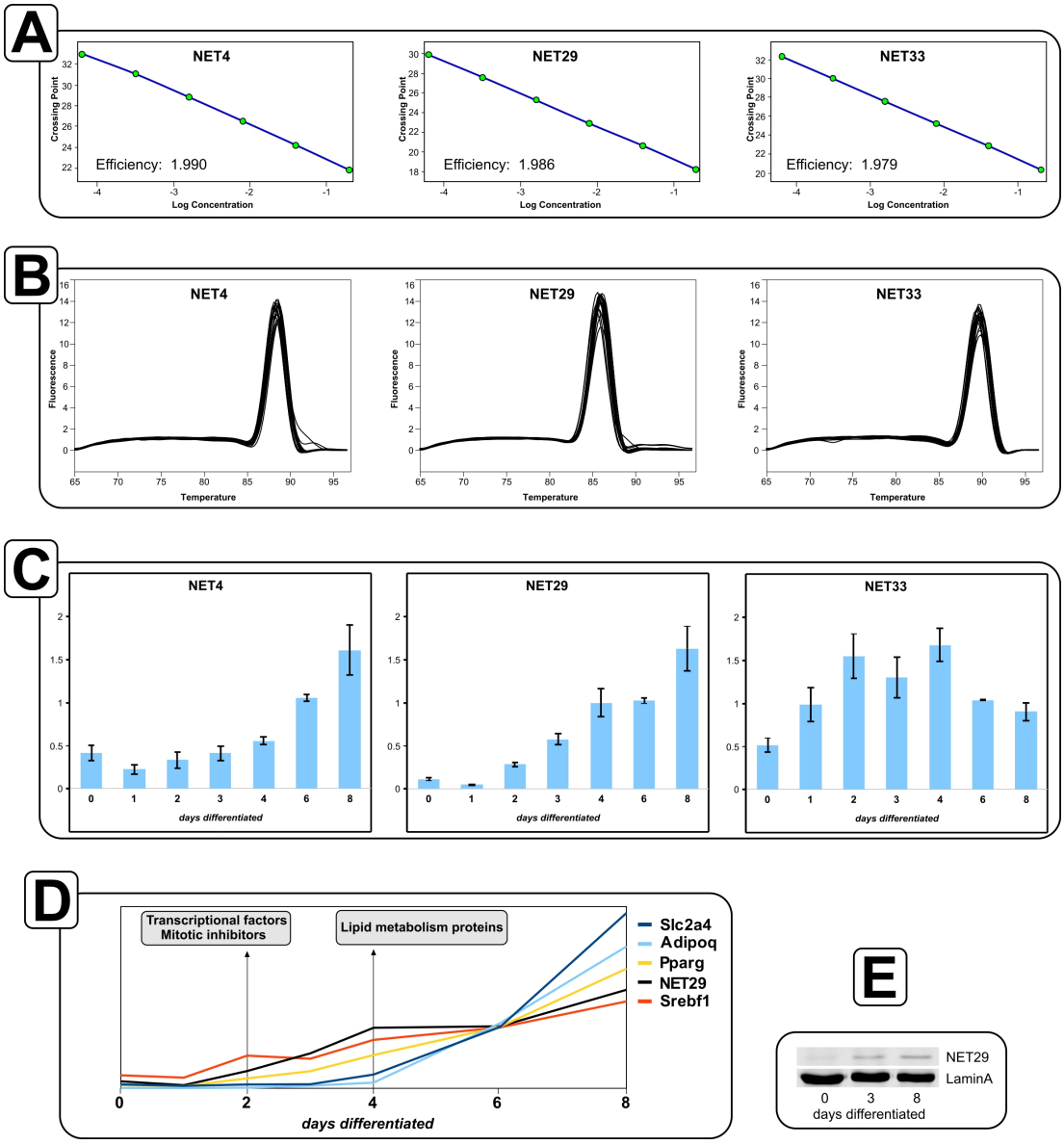


Fig 4.1.3: NET29 is upregulated early in adipogenesis

A - Callibration curves of primer pairs used for PCR. *Callibration curves are built to determine efficiency of the primer pairs. Efficiency should be close to 2 with an acceptable range 1.8-2.1.* ***B - Melting curves of amplified products.*** *Specificity of amplified products is assessed upon reaction completion. Single melting peaks represent high specificity of primer pairs.* ***C - NETs 4, 29 and 33 are upregulated in adipogenesis.*** *NET33 is upregulated on the first day of differentiation, NET29 - on the second day, and NET4 is upregulated late in adipogenesis.* ***D - NET29 follows the pattern of transcription factors.*** *Normalized to day 6 of differentiation mRNA levels of two transcription factors, Pparg and Srebf1, two lipid metabolism proteins, Slc2a4 and Adipoq, and NET29 are compared. Upregulation on the second day suggests important role in adipogenesis for NET29.* ***E - NET29 upregulation on protein level.*** *Protein extracts from undifferentiated or differentiated for 3 and 8 days 3T3-L1 cells were probed with AP NET29 antibody. Lamin A protein level was used as a loading control.*

4.2 NET29 effect on chromatin and its consequences

4.2.1 Chromosome repositioning effect of NET29[5]

As mentioned before, NET29 has a chromosome repositioning effect in human HT1080 cells where it alters the position of chromosome 5. To investigate if it has a similar role in 3T3-L1 mouse cells we took advantage of a fluorescence *in situ* hybridization (FISH) approach upon overexpression of human NET29 cDNA fused to an encoding GFP sequence. Chromosomes 6 and 16 were chosen for testing because they both bear adipogenic markers: *PPARG* and *LEP* on chromosome 6 and *ADIPOQ* on chromosome 16. Chromosome 7 was chosen to be used as a control.

Undifferentiated subconfluent 3T3-L1 cells were transfected either with a NET29-GFP fusion or GFP alone and were stained for chromosomes 6, 7 and 16. In order to quantify the effect, a macro for Image-Pro Plus, adapted by David Kelly from Croft *et al.* (1999), was applied. The macro divides the nucleus (based on DAPI signal) into 5 concentric rings, equal in area (Fig 4.2.1 a) and counts area and pixel intensity of whole chromosome painting intersected with different rings. Hence, when a chromosome is located in the central area of the nucleus, higher proportion of its signal will be in central rings. Similarly, if a chromosome is located close to the nuclear periphery, most of its signal will belong to the peripheral rings. However, the nucleus is a three-dimensional structure, and for the analysis only one focal plane that contains most of the chromosomes [6] was used. This underestimates the number of peripherally located chromosomes, as peripheral chromosomes, located at the top or at the bottom of the nucleus (relative to the slide) will appear centrally located. This underestimation, though, should be similar among all the samples.

5 Gerlinde Otti has performed FISH experiments and macro analysis in this chapter.

6 3T3-L1 cells, similar to many other cell lines, do not have a stable karyotype and often are polyploid. However this should not affect the analysis as the signal that is quantified as relative to the total chromosomal signal of a cell.

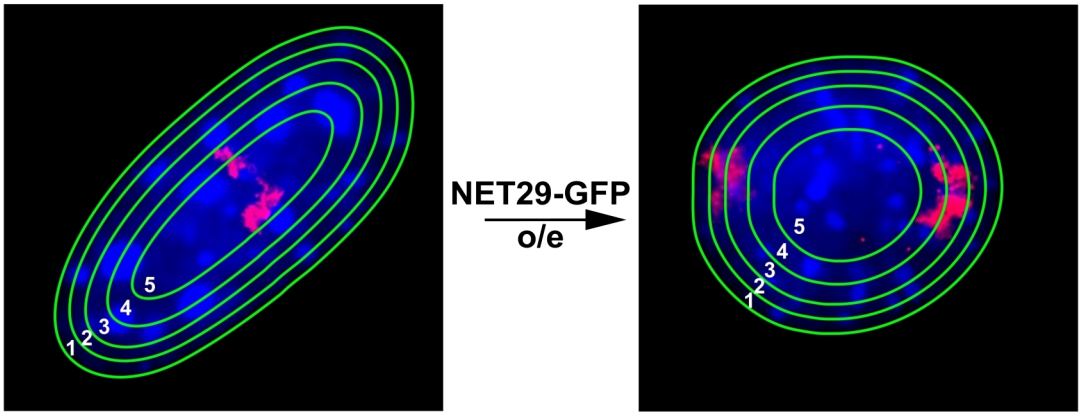


Fig 4.2.1 a: A macro divides the nucleus into 5 concentric rings of the same area

Rings contain 20% of nuclear area (in blue) each and numbered 1-5 from the periphery to the centre. Upon overexpression of NET29-GFP fused protein, chromosome 6 (in red) is found more often at the nuclear periphery.

As the changes in chromosome positioning are not observed in every transfected cell, analysis of a large population of cells had to be performed. 100 cells for each chromosome/protein combination have been analysed using the Image-Pro Plus macro. Relative intensities (percent of total signal from chromosomes in each cell) have been calculated for each ring. As chromosomes are rather large structures, it was decided to sum signals from rings 1 and 2 (peripheral signal) and 4 and 5 (central signal). The resulting values are plotted (Fig 4.2.1 b). It is obvious, that NET29-GFP overexpression has no effect on chromosomes 7 and little effect on chromosome 16. However, chromosome 6 tends to localize more often at the periphery in cells overexpressing NET29-GFP, but not GFP alone.

As NET29 is upregulated during adipogenesis, it was decided to test whether chromosome 6 is relocated to the periphery during adipocyte differentiation. However, fluorescence *in situ* hybridization failed to produce any specific signal in differentiated adipocytes fixed with formaldehyde. To overcome this problem, methanol/acetic acid fixation has been employed. This type of fixation, along with an ethanol dehydration series also used in this procedure, flatten nuclei and thus destroys its three-dimensional structure. On the other hand, methanol/acetic acid

fixation also destroys GFP fluorescence and epitopes, making it impossible to distinguish expressing exogenous protein cells from not expressing ones. As transient transfection of 3T3-L1 cells in our hands was at best around 60%, this would diminish the chromosome repositioning effect measured if NET29-GFP expressing preadipocytes had been processed the same way. However despite the fact that overexpressing NET29-GFP preadipocytes have been processed differently from this experiment, they have also been analysed using the two-dimensional macro. Thus, the results of both series of experiments should be qualitatively comparable.

4 and 7 days differentiated 3T3-L1 cells have been tested, and chromosome 6, indeed, has been found at the nuclear periphery more often in 7 days differentiated adipocytes (Fig 4.2.1 b). This suggests that upregulation of endogenous NET29 during adipogenesis leads to chromosome 6 repositioning. However, it is not clear whether only NET29 is responsible for the effect. Knock-down of endogenous NET29 expression must be done along with a rescue experiment in order to confirm this suggestion. If cells with downregulated endogenous protein will not exhibit chromosome 6 repositioning in adipogenesis, and introduction of exogenous NET29 will be able to rescue the phenotype, then NET29 is, indeed, the only factor of chromosome 6 repositioning in differentiating into adipocytes cells.

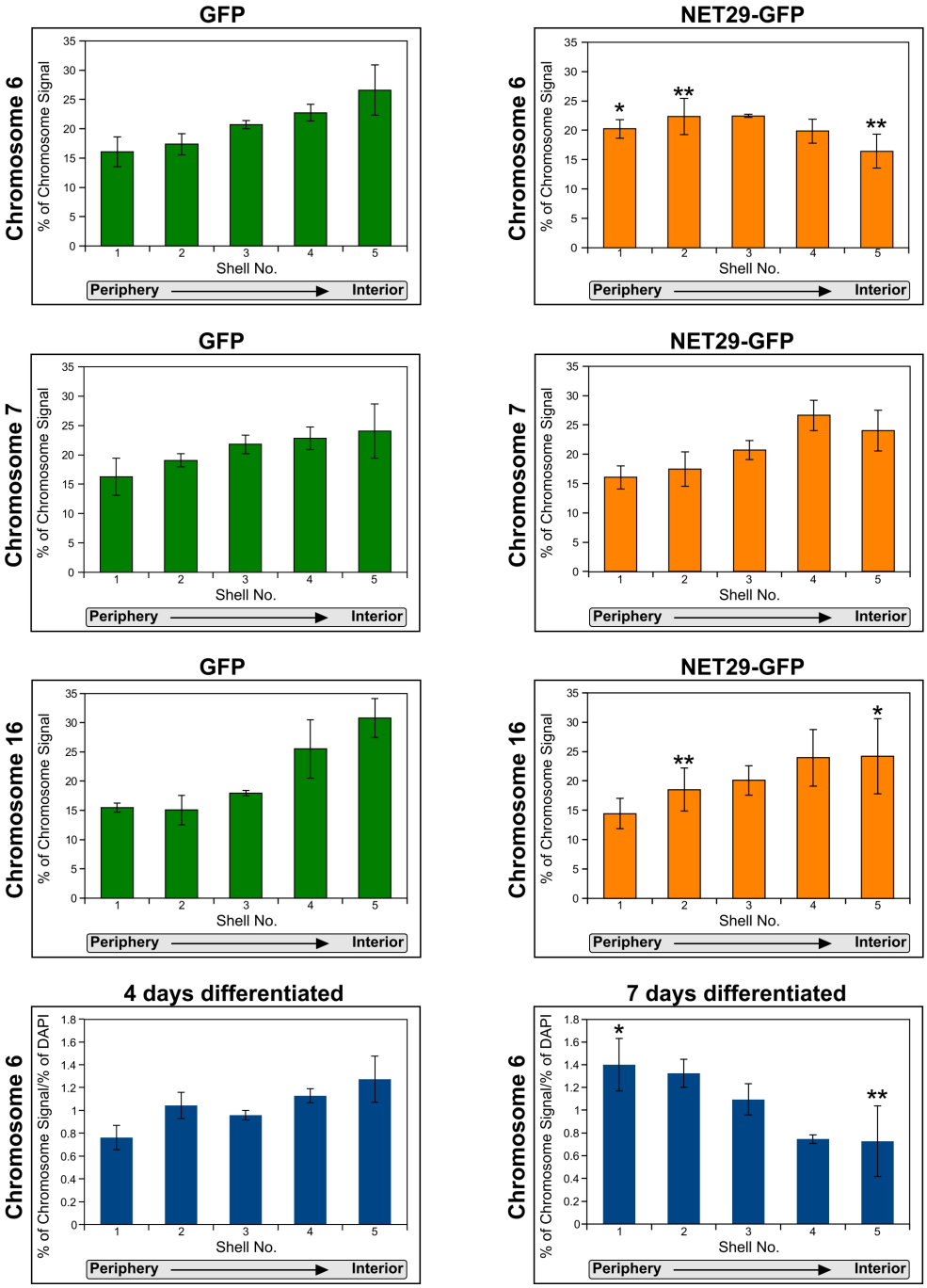


Fig 4.2.1 b: Chromosome 6 positioning is altered by NET29 overexpression and in adipogenesis

The graphs display the distribution of chromosome signal in each of the five shells. Shell 1 is the nuclear periphery, and shell 5 is the centre of the nucleus. 100 cells per each chromosome/protein combination have been analysed. 4 and 7 days differentiated samples were normaslised by DAPI signal. The asterisks represent significant differences (- $p < 0.05$, ** - $p < 0.01$, Kolmogorov -Smirnov test) for the shell when compared to the corresponding shell of GFP overexpressing cells (for NET29 overexpression) or 4 days differentiated cells (for 7 days differentiation).*

4.2.2 Transcription analysis of cells undergoing chromosome repositioning[7]

It has been already shown using artificial systems, that tethering of a chromosomal loci to the nuclear periphery may lead to both silencing or activation of gene expression (Finlan *et al.*, 2008; Ruault *et al.*, 2008). The presence of tissue-specific proteins that are able to rearrange chromosomes and tune expression of their genes could provide another way to regulate global changes that cells undergo during differentiation.

To identify consequences of chromosome 6 and, possibly, other untested chromosomes repositioning on gene expression, microarray analysis was carried out. Along with NET29, both NET4 and NET33 (the latter also had a chromosome repositioning effect in HT1080 cells) have also been tested. Total RNA was isolated from 3T3-L1 cells, transiently transfected with either mRFP alone, or C-terminal fusions of NETs 4, 29 and 33 with mRFP. Three independent RNA isolations have been done for all overexpressed proteins. The RNA have been reverse transcribed into cDNA, and double-stranded, biotin-labelled cRNA have been synthesized and amplified using the Illumina® TotalPrep RNA Amplification Kit following manufacturers instructions. Louise Evenden at Wellcome Trust Clinical Research Facility in Edinburgh has performed quality analysis of cRNA using an Agilent Bioanalyser and hybridization to Sentrix® BeadChip (6 samples/chip). BeadArray technology from Illumina (4.2.2 a, A) was chosen because of its reproducibility, precision (detects 1.3 fold difference with 95% probability) and economical effectiveness (Kuhn *et al.*, 2004).

First, the distribution of raw signals from controls (4.2.2 a, B) and all the samples (4.2.2 a, C) has been assessed. As array chips are scanned as 16 bit images, values of signals can only be within a 0 and 65535 interval, or, more conveniently, logarithms to base 2 of signal values can only be between 0 and 16. However,

7 Analysis of microarray data on transiently over-expressed NETs and, partially, stable cell lines over-expressing NET29 has been performed by Jose de las Heras

practically, logarithms to base 2 of background values are about six, so the dynamic range of expressed RNA signals is roughly between 6 and 16[8], where the lower threshold is calculated based on negative control values and the upper threshold is defined by scanner intensity, and should not be too close to 16 to prevent saturation. This way, up to a 1000-fold difference in expression could be observed.

Microarray controls (4.2.2 a, B) include negative controls (probes, that should not hybridize with any cRNA), and positive biotin controls (probes labelled with biotin). The former indicate background level in a particular chip and the latter should give high signals close to the upper limit of detection[9]. All of the samples showed good distribution of negative (below 7) and positive biotin (around 14) controls.

Distribution of all sample signals (4.2.2 a, C) shows that over 50% of the probes are within the background level (boxes are below 7). This could be due to different and not mutually exclusive reasons. First of all, Illumina arrays often have more than one probe per gene, and while some of them detect all splice variants, most of them detect single splice variants. So that if only one particular splice form is expressed, only two probes will detect it, but not the others. Second of all, uninduced 3T3-L1 cells are not fully differentiated and, hence, many cell type specific genes are not expressed. Another possibility is poor labelling of cRNA with biotin, but the fact that some of the probe values were as high as 15 disproves this possibility. Despite this, the level of biotin incorporation into cRNA has been compared to that of commercial biotinylated oligonucleotides by dot blot assay (as described in Dölken *et al.*, 2008) and was found to be at similar levels. Poor handling of samples during hybridization is excluded by good distribution of controls.

Signals of different samples have been normalized in R v2.11.1 (RTeam, 2010) using the quantile method of the Limma package (Wettenhall and Smyth, 2004). Normalized data has been averaged between 3 repeats of each sample and plotted as two-dimensional scatter plots, where on the vertical axis plotted probe va-

8 Here and later, when signals from microarray are discussed, the numbers are logarithms to base 2.

9 Due to differences in biotin incorporation and length between probes and hybridized RNA, highly expressed and, mostly, long RNA might give signals higher than biotin control.

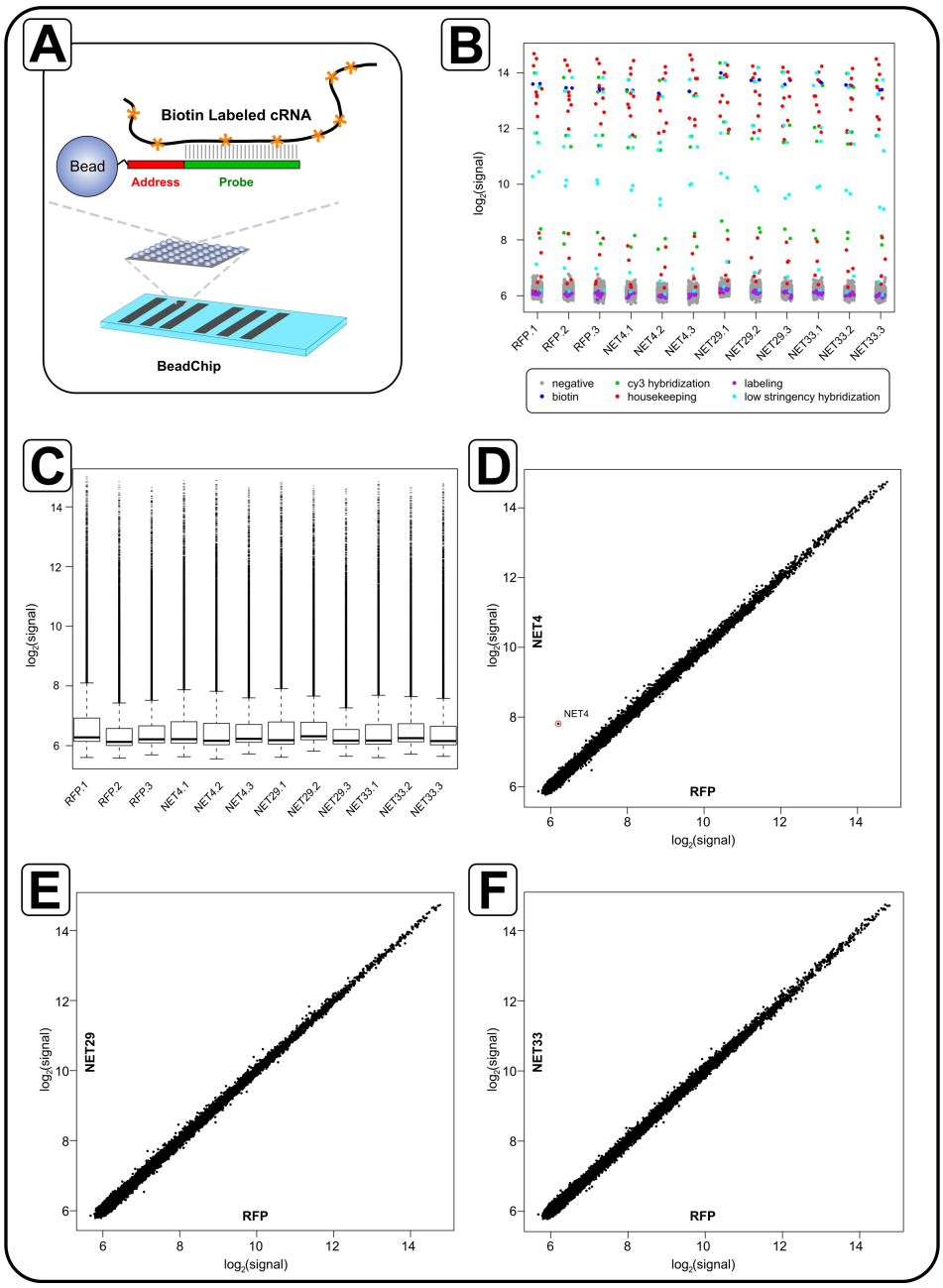


Fig 4.2.2 a: Transient overexpression of three NETs does not affect transcription

A *-* ***Schematic overview of Illumina Sentrix Bead Array technology.*** *Each array is in a format of a glass slide with imprinted hollows. Each hollow holds a single 3 µm silica bead with hundreds of thousands bar-coded oligonucleotides of the same sequence attached to it. Each array has about 20 beads with the same oligonucleotides attached to it.* ***B*** *-* ***Raw signal distribution of controls****. Included into every array detection controls look good. Uneven distribution between samples indicates the need of normalisation of signals.* ***C*** *-* ***Raw signal distribution of samples****. Over a half of signals are at the background level (discussed in text).* ***D-F*** *-* ***Normalized signals of NETs plotted against RFP****. The easiest way spot differences in gene expression between two samples is to plot each probe on a two-dimensional graph with values from one sample on x axis and values from another sample on y axis. Probes that are expressed differently between the samples will fall outside of the line.*

lues from NET-RFP fusion overexpressing sample, and on the horizontal axis - from RFP overexpressing sample (4.2.2 a, D-F). However, no significant differences between the sample pairs could be observed. As the exogenous sequences were human, upregulation was not detected in the mouse microarray, apart from one of the NET4 probes that is similar between the mouse and human sequence (4.2.2 a, D).

Absence of differences in expression between cells overexpressing RFP fusions of NETs and control cells (overexpressing RFP alone) could be due to low transfection efficiency and/or insufficient time between transfection and RNA isolation. Transfection of 3T3-L1 cells is a challenging task, out of different methods and reagents, including nucleofection, FugeneHD has performed best, transfecting up to 60% of the cells based on the RFP signal. However, the level of expression is greatly variable from a barely visible rim to bulks of overexpressed protein aggregates in the ER.

To overcome problems with transfection efficiency, 3T3-L1 cell lines stably overexpressing the NET29-GFP fusion and GFP alone were generated. The differentiation ability and the position of chromosome 6 in the cell lines has been assessed. Both cell lines have been able to differentiate into adipocytes and chromosome 6 has been found at the periphery more often in the cell line overexpressing NET29-GFP fusion, than in the overexpressing GFP cell line. 3 individual RNA isolations from both cell lines have been done and microarray analysis has been carried out. As a positive control of differential expression, RNA from undifferentiated and 9 days differentiated 3T3-L1 cells have also been used for microarray analysis, one repeat each.

Normalised averaged signals from NET29-GFP and GFP overexpressing cells have been analysed and plotted (Fig 4.2.2 b). Out of 11 highly scored genes that were expressed at different levels in NET29-GFP cells and GFP cells, GRB10 had the biggest difference. GRB10 is growth factor receptor bound protein 10 and it had been shown to be a negative regulator of insulin signalling (Mori *et al.*, 2005). Reverse transcription followed by qPCR has been used to confirm over 3-fold upregulation of

GRB10 in cells overexpressing the NET29-GFP fusion, compared to the cells overexpressing GFP alone (Fig 4.2.2 b, insertion box).

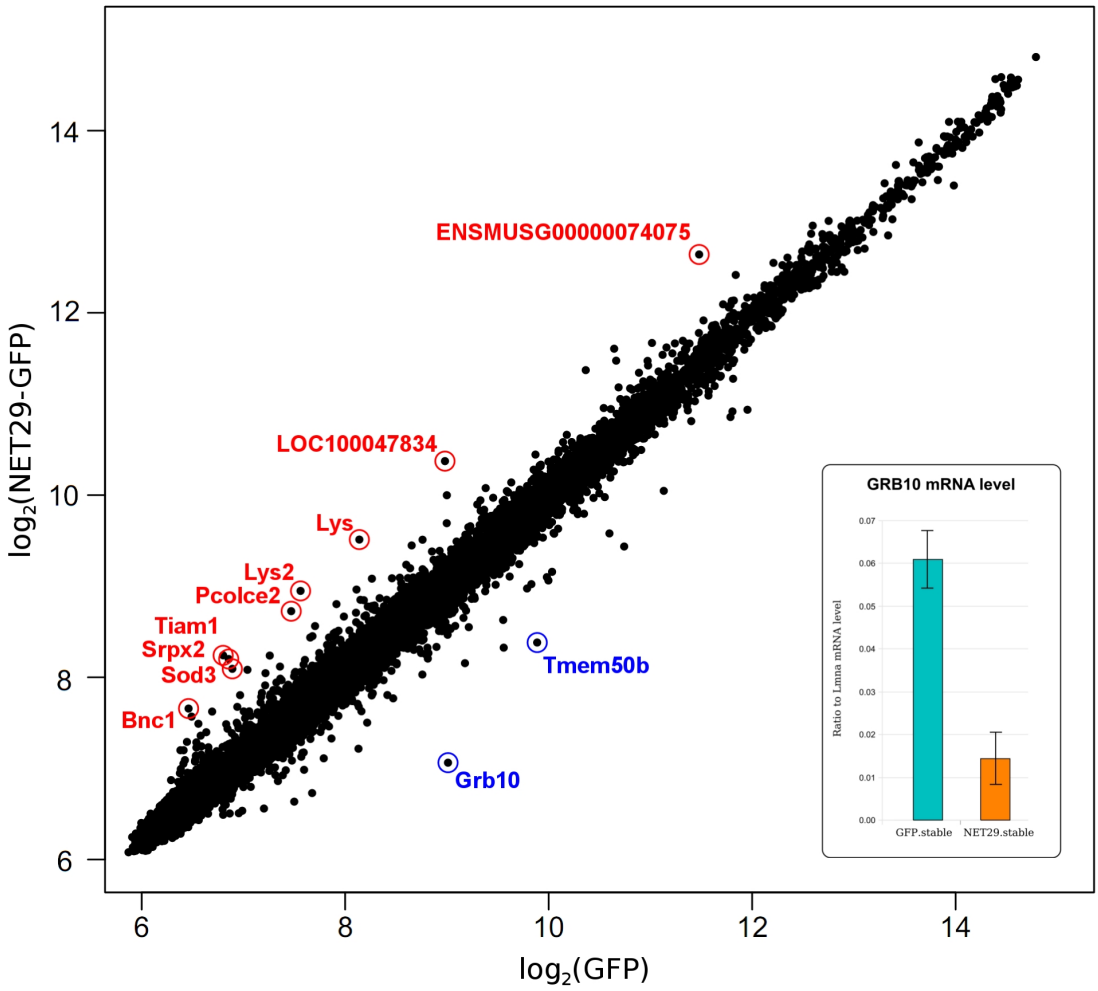


Fig 4.2.2 b: Expression of several genes is affected by NET29 overexpression

Microarray signals from the NET29-GFP overexpressing cell line were plotted against the GFP overexpressing cell line. Several genes are upregulated (in red) in the NET29-GFP cell line while two (in blue) are downregulated. Insertion box is RT-qPCR analysis of GRB10 expression in NET29-GFP and GFP cell lines.

However, this way of analysis cannot distinguish whether a particular gene is downregulated in the NET29-GFP cell line, or upregulated in the GFP cell line. In order to normalize data to the "initial" state of the cells, signals from undifferentiated untransfected cells have been subtracted from signals of NET29-GFP and GFP

overexpressing cells and plotted (Fig 4.2.2 c). This way, any signal that is uniform among all the cells will have coordinates roughly 0,0. The two most outstanding probes are EGFP and GRB10. EGFP is a probe against enchanced GFP, and since untransfected cells do not express this exogenous sequence, this probe is most differently expressed between the cells. As the probes coordinates are roughly equal (6,6), it suggests that overexpression of the NET29-GFP fusion and GFP alone is at a similar level, hence the two cell lines can be compared safely. The other probe, GRB10, has coordinates at about 3.5,1 which means that both stable cell lines have it upregulated, however the level of expression is higher in GFP cells and therefore it appeared downregulated in NET29-GFP cell lines in the previous analysis. Thus, it is not likely that the chromosome 6 repositioning is the cause for the difference in GRB10 expression.

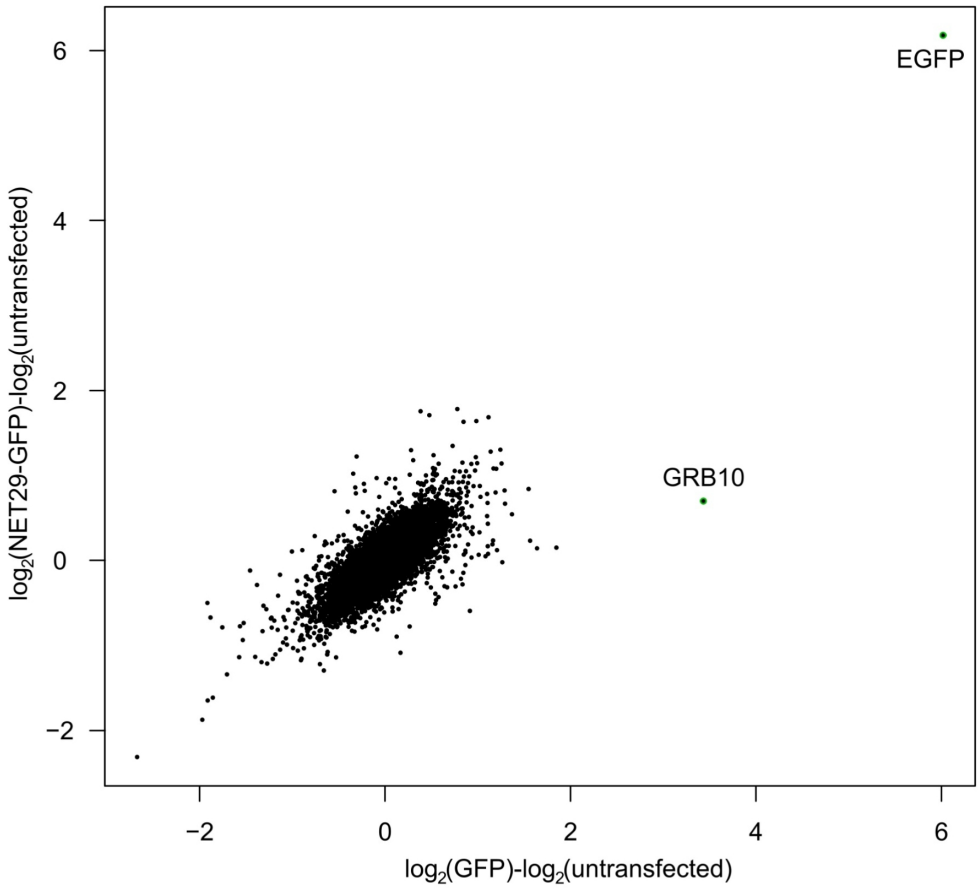


Fig 4.2.2 c: Expression of GRB10 is also affected in the control cell line

Signals from undifferentiated cells were subtracted from corresponding signals from NET29-GFP and GFP overexpressing cells.

To select for genes that are more likely to change expression due to NET29 overexpression two criteria have been applied: the difference in gene expression must not exceed 0.3 (about 1.2-fold difference) in GFP overexpressing cells and must be more than 0.75 (about 1.7-fold difference) in NET29-GFP overexpressing cells (Fig 4.2.2 d). The selected genes included 10 upregulated messages and 7 downregulated messages in the NET29-GFP cell line that are located on different chromosomes.

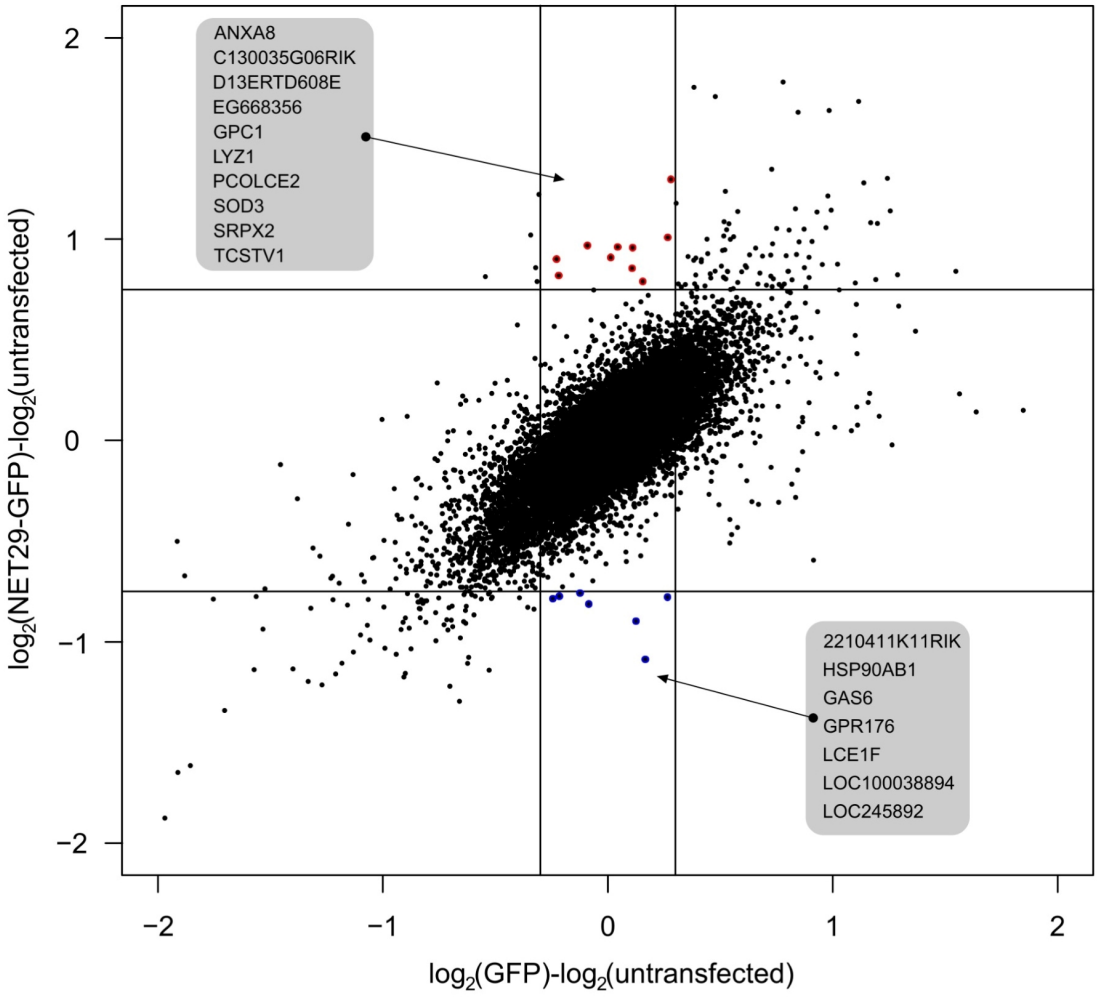


Fig 4.2.2 d: Contrast analysis identifies genes affected by NET29 overexpression

Signals from undifferentiated cells were subtracted from corresponding signals from NET29-GFP and GFP overexpressing cells. Genes expressed differently in NET29-GFP, but not GFP cell line are selected (upregulated genes are in red, downregulated - in blue).

Included into the microarray analysis controls, undifferentiated 3T3-L1 cells and 9-day differentiated cells, provide an opportunity to find more NETs upregulated or downregulated during adipogenesis. First, a list of genes with a two-fold and over difference in expression during differentiation have been built. Then the list was searched for matching gene names of NETs. A number of NETs appeared in the list of genes up- or downregulated during adipogenesis. Apart from the already tested

NETs 4, 29 and 33, upregulated NETs include lNETs 26 (*TMEM14C*), 50 (*DHRS7*) and 62 (*MCAT*) identified in liver dataset, and mNETs 20 (*MBOAT5*) and 22 (*TMEM66*) identified in muscle dataset. Liver NET3 (*TMEM48/Ndc1*), a core NPC protein, appears to be downregulated in adipogenesis.

In order to assess the quality of the RT-qPCR data, the behaviour of the *LMNA* transcript during adipogenesis was monitored. All six *LMNA* probes were downregulated about two fold on day 9 as compared to undifferentiated cells and therefore *LMNA* should not be used as a loading control for comparison of undifferentiated and differentiating cells. This does not seriously affect the qRT-PCR results showing upregulation of NETs 4, 29 and 33 as microarray data supports their upregulation. After searching for an appropriate loading control for qRT-PCR, *GAPDH* was chosen as a gene potentially unchanged during adipogenesis based on a relatively even level of transcript in both undifferentiated and differentiated cells.

4.2.3 Transcriptional analysis of differentiating 3T3-L1 cells with down-regulated NET29

As preadipocytes might lack factors needed for the chromosome repositioning to affect transcription, it was decided to down-regulate NET29 in differentiating cells and test their transcription. siGENOME SMART pool, combining 4 siRNA oligos against NET29 was purchased from Dharmacon (M-040281-00). Either the SMART pool or a non-target siRNA oligo (CGUACGCGGAAUACUUCGAdTdT) were transfected into undifferentiated 3T3-L1 to 100 nM final (total) concentration. RNA from undifferentiated cells was isolated two days after transfection, and RNA from cells differentiated for 2 days was isolated 4 days after transfection. The NET29 transcript level was measured by qRT-PCR using *LMNA* as a loading control (Fig 4.2.3, A). Although the NET29 mRNA level was significantly reduced in undifferentiated cells, 2 days differentiated cells showed no difference between SMART pool and non-target oligo transfected cells. This could be explained by two reasons - oligo degradation and NET29 upregulation.

It is well known, that differentiated 3T3-L1 cells are extremely difficult to transfect, hence multiple transfections of differentiating cells would not be efficient. In order to obtain a significant decrease in NET29 mRNA level in differentiating cells, a set of shRNA vectors was purchased from Sigma, including 5 vectors with different anti-NET29 oligo sequences (Table 15) and one with a non-target oligo sequence. 6 stable cell lines were generated and RNA and proteins from 5 days differentiated stable cells lines were isolated. NET29 transcript and protein levels were assessed by qRT-PCR and quantitative Western blot, accordingly. The cell line expressing oligo TRCN0000247877 (sh77) was the most successful in downregulating NET29 (Fig 4.2.3, B-D).

Table 15: NET29 shRNA vectors sense oligo sequences

shRNA vector	**sense sequence**
TRCN0000247873	TTCCTGCTGGTCTGGTATTAT
TRCN0000247874	TGTGCAATTCCTGCAGTATTA
TRCN0000247875	ACAATGGTTCCAGGATCAAAG
TRCN0000247876	CATCATTATGTGTCCACATTC
TRCN0000247877	CAAAGACGAGTACGAGAAATT

Total RNA from the sh77 cell line, cell line expressing non-target shRNA and non-transfected 3T3-L1 cells, all differentiated for 4 days, was isolated, labelled with biotin and microarray analysis was carried out. Raw signals have been plotted to control for large chip abnormalities and distribution of in-built controls has been monitored to ensure adequate hybridisation. After normalisation, subtractions of signals from either non-target shRNA expressing cell line or sh77 cell line and signals from non-transfected cells have been plotted (Fig 4.2.3, E). To select genes that are specifically affected by NET29 down-regulation two rules have been applied: the difference between non-target shRNA expressing cell line and non-transfected cells must not exceed 0.5 and the difference between sh77 and non-transfected cells must be not less than 0.75. Genes upregulated upon NET29 knock-down are listed in Table 16, downregulated genes are listed in Table 17.

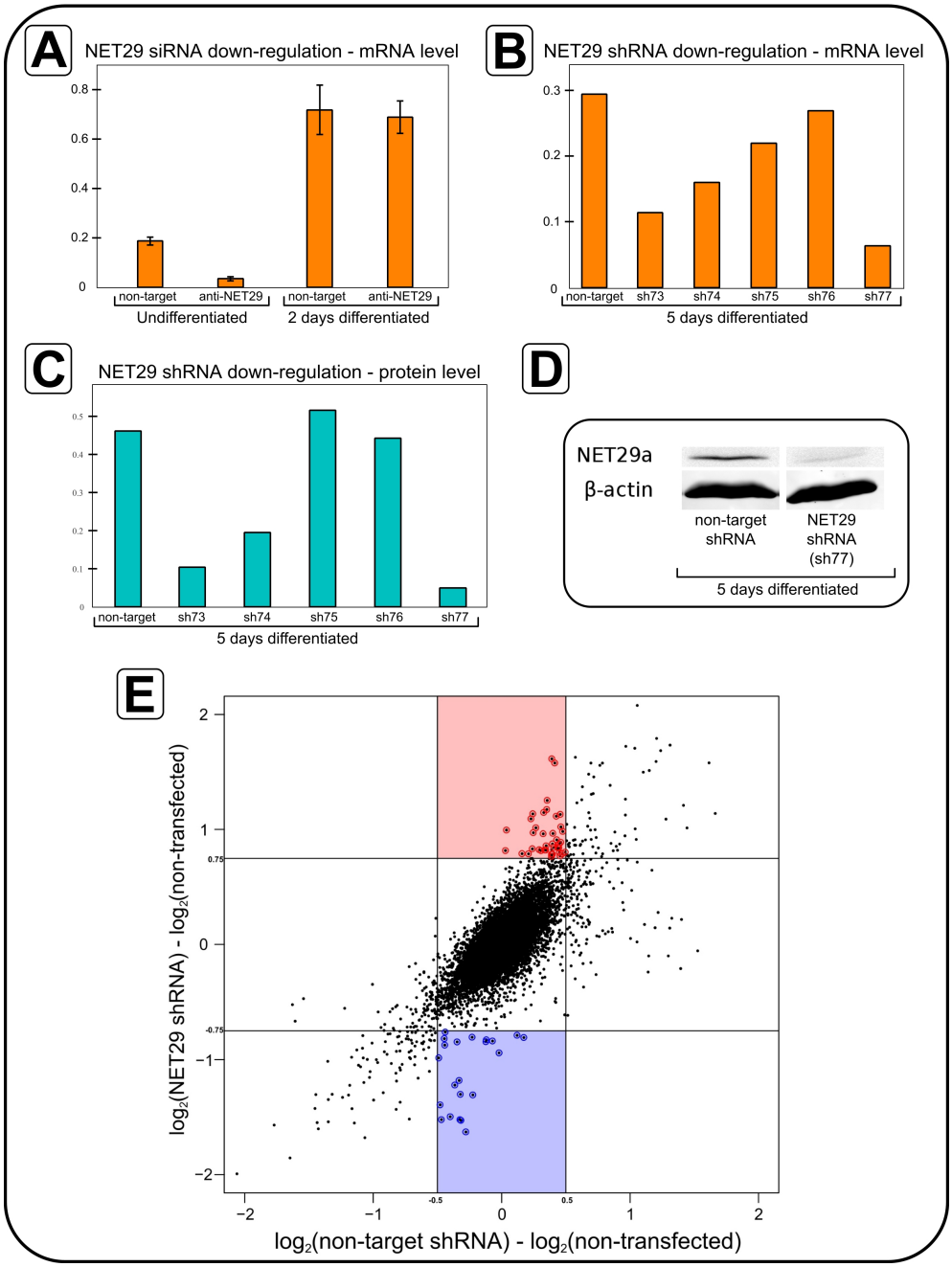


Fig 4.2.3: Downregulation of NET29 affects transcription in adipocytes

***A - siRNA downregulates NET29 in pre-adipocytes**. SMART pool oligos were transfected into undifferentiated 3T3-L1 cells and RNA was extracted either after two days post transfection or the cell were allowed to become confluent and induced to differentiate. Then RNA was extracted from two days differentiated cells (in total 4 days after transfection). **B-D - shRNA downregulates NET29 in adipocytes**. Stable cells lines expressing shRNA hairpin against NET29 were established and simultaneous RNA and protein isolation was done from 5 days differentiated cells. sh77 cell line shows best down-regulation of NET29 on mRNA and protein levels. **E - Microarray analysis of differentiating 3T3-L1 cells with downregulated NET29**. In text.*

Table 16: List of genes upregulated in sh77 cell line

Gene	Chromosome	Number of probes	Notes
coiled-coil domain containing 74A (2310015A05Rik)	16	1	
actinin alpha 3 (Actn3)	19	2	
brain-specific angiogenesis inhibitor 2 (Bai2)	4	1	
cDNA sequence BC011487 (BC011487)	**6**	1	
cDNA sequence BC064033 (BC064033)	**6**	1	
crystallin, alpha B (Cryab)	9	1	up in myogenesis
chemokine (C-X-C motif) ligand 12 (Cxcl12)	**6**	2	**induces myogenic differentiation**
eyes absent 1 homolog (Drosophila) (Eya1)	1	1	
glucuronyl C5-epimerase (Glce)	9	1	
hyaluronan and proteoglycan link protein 2 (Hapln2)	3	1	
histone cluster 1, H3e (Hist1h3e)	13	1	
histone cluster 1, H4a (Hist1h4a)	13	1	
histone cluster 1, H4h (Hist1h4h)	13	1	
histone cluster 1, H4j (Hist1h4j)	13	1	
histone cluster 2, H2aa1, (Hist2h2aa1)	3	1	
histone cluster 2, H2aa2 (Hist2h2aa2)	3	1	
histone cluster 2, H2be, (Hist2h2be)	3	1	
kallikrein 1-related peptidase b24 (Klk1b24)	7	1	
kallikrein 1-related pepidase b4 (Klk1b4)	7	1	
keratin 16 (Krt16)	11	1	
keratin 42 (Krt42)	11	3	
PREDICTED: similar to histone H4 (LOC100041230)	13	1	
PREDICTED: similar to clusterin (LOC100046120)	14	1	discontinued
NAD(P)H dehydrogenase, quinone 1 (Nqo1)	8	1	
procollagen C-endopeptidase enhancer 2 (Pcolce2)	9	1	down in myogenesis
phosphofructokinase, platelet (Pfkp)	13	1	
Src homology 2 domain containing F (Shf)	2	2	
solute carrier family 9 (sodium/hydrogen exchanger), member 3 regulator 1 (Slc9a3r1)	11	1	
transmembrane protein 132A (Tmem132a)	19	1	
tetraspanin 17 (Tspan17)	13	1	
PREDICTED: WAS/WASL interacting protein family, member 3 (Wipf3)	**6**	1	
WW domain containing E3 ubiquitin protein ligase 2 (Wwp2)	8	2	

Gene	Chromosome	Number of probes	Notes
zinc finger and BTB domain containing 7C (Zbtb7c)	18	1	

Table 17: List of genes downregulated in sh77 cell line

Gene	Chromosome	Number of probes	Notes
RIKEN cDNA 6330406I15 gene (6330406I15Rik)	5	2	
angiopoietin-like 7 (Angptl7)	4	1	
endoplasmic reticulum metallopeptidase 1 (Ermp1)	19	1	
fibulin 1 (Fbln1)	15	1	down in myogenesis
GATA binding protein 6 (Gata6)	18	1	
insulin-like growth factor binding protein 4 (Igfbp4)	11	2	up in myogenesis
insulin-like growth factor binding protein 7 (Igfbp7)	5	1	
mannosidase 2, alpha B1 (Man2b1)	8	1	
nucleotide-binding oligomerization domain containing 1 (Nod1)	**6**	1	
prolactin family 2, subfamily c, member 2 (Prl2c2)	13	2	**inhibits myogenesis differentiation**
prolactin family 2, subfamily c, member 3 (Prl2c3)	13	1	
prolactin family 2, subfamily c, member 4 (Prl2c4)	13	1	
resistin like alpha (Retnla)	16	1	
ribonuclease, RNase A family 4 (Rnase4), transcript variant 1	14	1	
sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A (Sema5a)	15	2	
sulfatase 2 (Sulf2)	2	1	
transmembrane protein 120A (Tmem120a)	5	1	NET29

Not surprisingly, the genes affected by NET29 are located on several chromosomes, including the tested chromosome 6. Most notably, chromosome 13 is highly affected by NET29 knock-down. It would be interesting to test whether the chromosome 13 nuclear localisation is also affected by NET29.

Interestingly, a number of genes affected by NET29 downregulation are also affected in myogenesis. Moreover, chemokine (C-X-C motif) ligand 12 was shown to

induce myogenic differentiation (Melchionna *et al.*, 2010) and is also upregulated upon NET29 depletion. On the other hand, prolactin family 2, subfamily c, member 2 was shown to inhibit myogenesis (Wilder and Linzer, 1989) and it is repressed in the NET29-depleted cell line. This suggests that NET29 upregulation in adipogenesis helps to inhibit an alternative differentiation pathway – myogenesis.

4.3 Summary and conclusions

Three NETs are shown to be upregulated during *in vitro* adipogenesis. Available transcriptomics data suggests that these NETs are highly expressed in mouse fat tissue. In a separate screen, two of the NETs affected chromatin distribution. This study is focused on NET29, a NET that is upregulated early in adipogenesis along with crucial adipogenic genes such as Peroxisome proliferator-activated receptor gamma and Sterol regulatory element-binding transcription factor 1. This suggests an important function of NET29 in adipogenesis.

We tested the effect of NET29 overexpression in preadipocytes on the radial distribution of several chromosomes. Most notably, chromosome 6 was relocated to the nuclear periphery by NET29 overexpression. Transcriptomic analysis of preadipocytes with upregulated NET29 revealed several genes that are affected by NET29. However, downregulation of NET29 in the differentiating cells had a more pronounced effect on transcription. Two genes important for myogenesis were affected among several others. Chemokine (C-X-C motif) ligand 12 (*CXCL12*) is upregulated upon NET29 knockdown, and it can induce myogenic differentiation (Melchionna *et al.*, 2010). Prolactin family 2, subfamily c, member 2 *(PRL2C2)* was shown to inhibit myogenesis (Wilder and Linzer, 1989) and it is downregulated along with NET29. Therefore it is possible that NET29 helps to shut alternative differentiation pathways down. It is not yet clear whether the effect on transcription is mediated by the NET29-mediated chromatin rearrangement.

CHAPTER V: A Molecular Mechanism of Chromosome Repositioning

This chapter is focused on uncovering the molecular mechanism of chromosome repositioning of NET29. It is mostly ongoing work with some significant and promising results. First, a dominant negative mutant of NET29 is described. Phosphorylation analysis of NET29 and attempt of identification of the protein's binding partners will also be described.

5.1 Dominant negative mutant of NET29

In order to facilitate identification of the region of NET29 that is involved in chromosome 6 repositioning, multiple protein sequence alignments of NET29 from diverse species have been done. Using OrthoMCL DB (Altschul *et al.*, 1997), NET29 orthologs have been found in Plants and various classes of Animals: Nematodes, Insects, Amphibia, Fish, Birds and Mammals. JalView v2.6.1 (Waterhouse *et al.*, 2009) was used to build the alignment of NET29 protein sequences from different animals (Fig 5.1 a). Along with this analysis, prediction of potential phosphorylation sites has been done using NetPhos v2.0 (Blom *et al.*, 1999), as well as prediction of potential transmembrane domains using TMHMM v2.0 (Krogh *et al.*, 2001).

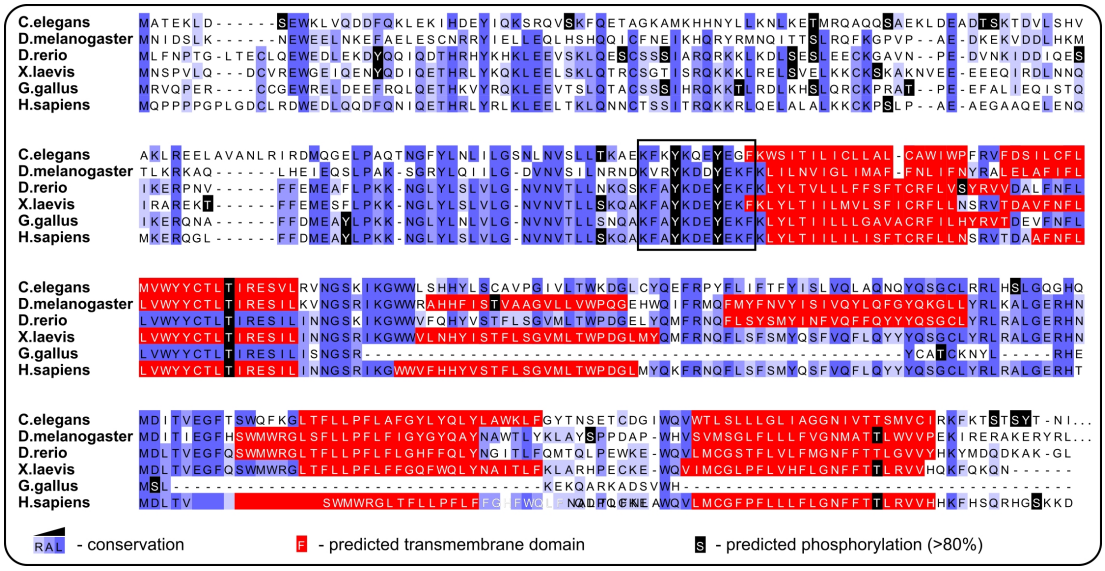


Fig 5.1 a: ClustalW alignment of NET29 from multiple species

NET29 sequences from Caenorhabditis elegans, Drosophila melanogaster, Danio rerio, Xenopus laevis, Gallus gallus, and Homo sapiens were aligned to highlight the most conserved regions. Box outlines a very conserved region that contains two potentially phosphorylated tyrosines.

There are three residues highly scored for phosphorylation that are located within highly conserved regions. However, the T176 residue resides within a hydrophobic sequence that is predicted to form a transmembrane helix. Both Y129 and Y133, on the other hand, are located on the nucleoplasmic N-terminus and have the potential to be involved in chromatin binding. Considering all this, residues Y129 and Y133 have been mutated to either phenylalanines (Y129F, Y133F - phospho-null mutations) or glutamic acids (Y129E, Y133E - phosphomimetic mutations). 3T3-L1 cells have been transiently transfected with plasmid DNA encoding either wild type or mutated NET29 sequences in frame with the GFP coding sequence. Intracellular localisation was monitored by fluorescence microscopy after direct fixation with PFA or with Triton X-100 pre-extraction. The last procedure washes away proteins that are weakly associated with the NE. Wild type NET29 as well as its mutants formed rim-like structures that are characteristic to NETs. They also were retained at the NE after pre-extraction further supporting similarity in the localisation between wild type

and mutant proteins (Fig 5.1 b).

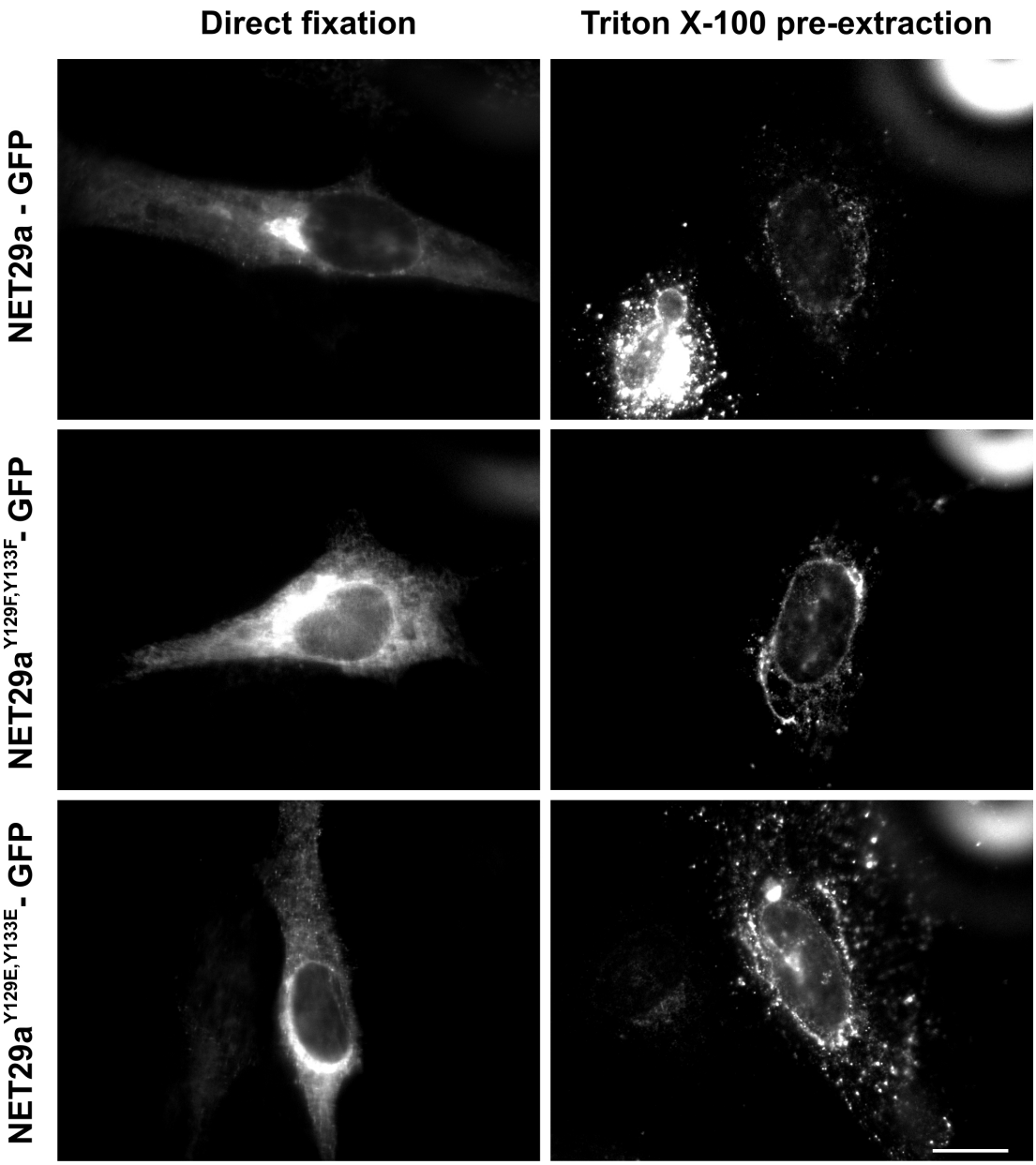


Fig 5.1 b: Mutants of NET29 have a similar cellular localisation as the wild type protein

3T3-L1 pre-adipocytes expressing wild type NET29 or its mutants fused to GFP were directly fixed with PFA or pre-extracted with Triton X-100 and then fixed. The mutant's localisation and resistance to Triton X-100 pre-extraction is similar to wild type NET29. Scale bar represents 10 µm.

3T3-L1 preadipocytes were allowed to express exogenous cDNA for three days and the chromosome 6 position was assessed using FISH[10]. Cells overexpressing the phosphomimetic mutant showed chromosome 6 distribution that is similar to the cells overexpressing wild type NET29 (Fig 5.1 c). However, cells overexpressing the phospho-null mutant had chromosome 6 in the centre of the nucleus more often than in cells overexpressing wild type NET29, or even non-transfected cells. This lead to two suggestions: first, the conserved region before the first predicted transmembrane domain interacts either directly or indirectly with chromatin. Second, this interaction most probably requires phosphorylation of either Y129 or Y133, or both of the residues.

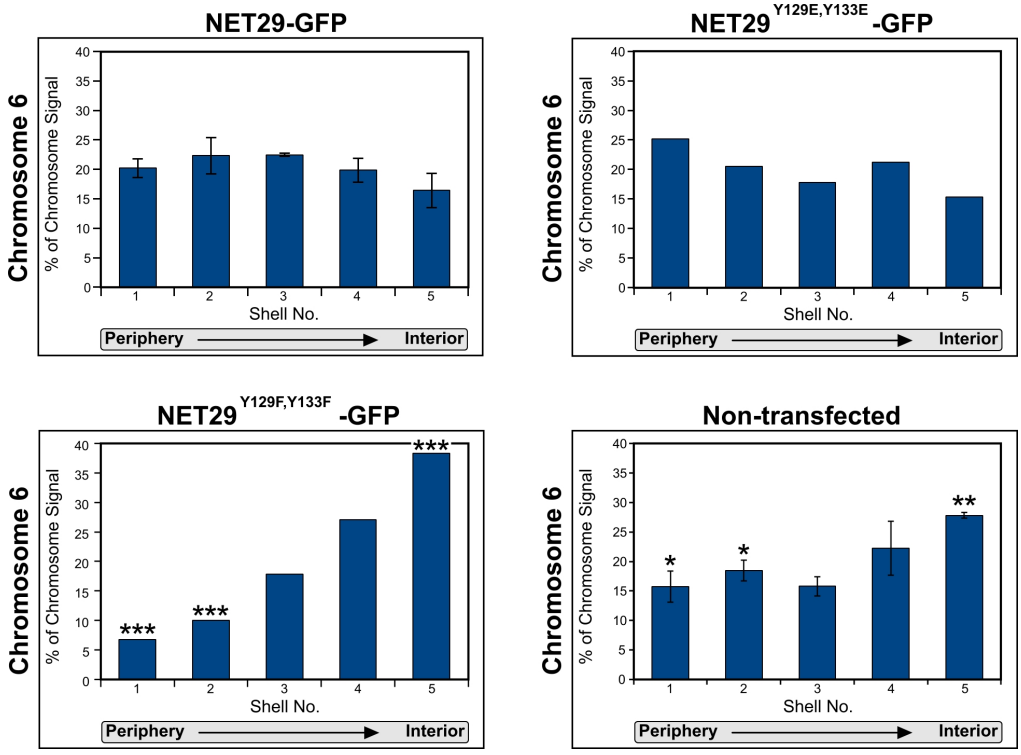


Fig 5.1 c: NET29 phospho-null mutant fails to tether chromosome 6

*hromosome 6 localisation was determined by FISH upon overexpression of NET29 mutants. The phosphomimetic mutant has a chromosome tethering effect similar to the wild type. However the phospho-null mutant reverts the effect of NET29 overexpression. The asterisks represent significant differences (***** - $p < 0.05$, ****** - $p < 0.01$, ******* - $p < 0.001$, Kolmogorov -Smirnov test) for the shell when compared to the corresponding shell of wild-type NET29 overexpressing cells.*

10 Initial FISH experiment has been done by Gerlinde Otti

5.2 Analysis of PTMs of NET29

In order to confirm the presence of phosphorylation at Y129 and Y133 of NET29 *in vivo*, immunoprecipitation of over-expressed NET29-GFP fusion protein from undifferentiated 3T3-L1 cells was carried out. To avoid contamination of eluted material, an anti-GFP antibody was covalently cross-linked to CnBr-Sepharose 4B agarose beads. A rabbit polyclonal antibody that recognizes GFP was generated against full-length GFP protein purified from *E.coli*. The affinity purified antibody specifically recognises GFP (Fig 5.2 a) and works well in immunoprecipitation assays.

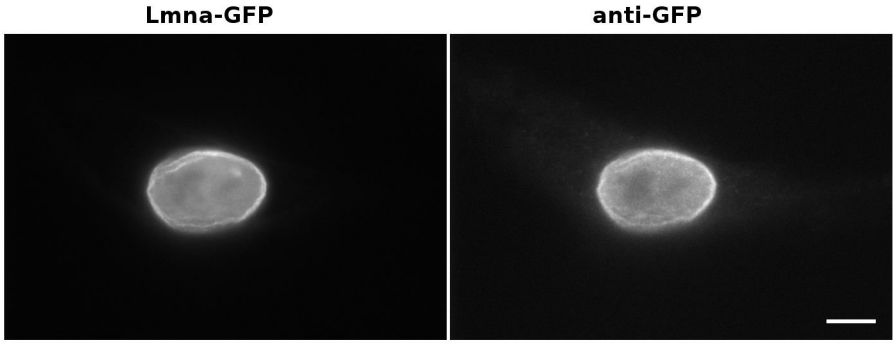


Fig 5.2 a: Affinity purified anti-GFP antibody staining

HT1080 cells stably expressing Lamin A - GFP fusion were fixed with PFA, stained with affinity purified anti-GFP antibody and labelled with anti-rabbit immunoglobulin Cy5 conjugated antibody. Signal of anti-GFP antibody corresponds to the signal of the GFP fluorescence thus showing the specificity of the antibody. Scale bar represents 10 µm.

Cell lysates are prone to de-phosphorylation by phosphatases present in every cell, hence phosphatase inhibitors were used for the immunoprecipitation assay. In order to obtain information about phosphatases that are highly expressed in differentiated 3T3-L1 cells, microarray data from 9 days differentiated cells was searched for annotated phosphatases or phosphatase subunits and results were scored by their expression level (Table 18). Based on these data, available specific and general phosphatase inhibitors were selected and used during the assay (Table 19).

Table 18: Phosphatase genes, highly expressed in 3T3-L1 adipocytes

adip.log2	Target ID	Definition
13.16	Ppp2ca	Protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform
13.1	Ppp1r3c	Protein phosphatase 1, regulatory (inhibitor) subunit 3C
12.32	Ppp1ca	Protein phosphatase 1, catalytic subunit, alpha isoform
11.92	Ppm1a	Protein phosphatase 1A, magnesium dependent, alpha isoform
11.82	Ppp2cb	Protein phosphatase 2 (formerly 2A), catalytic subunit, beta isoform
11.46	Dusp1	Dual specificity phosphatase 1
11.44	Ssu72	Ssu72 RNA polymerase II CTD phosphatase homolog
11.34	Ppp2r1a	Protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), alpha isoform
11.29	Ptplad1	Protein tyrosine phosphatase-like A domain containing 1
11.16	Ppp6c	Protein phosphatase 6, catalytic subunit
11.12	Loc100046393	PREDICTED: Similar to Protein phosphatase 2, regulatory subunit B (B56), alpha
11.04	Pptc7	PTC7 protein phosphatase homolog (S. cerevisiae)
10.99	Ptpn11	Protein tyrosine phosphatase, non-receptor type 11
10.71	Ppp4c	Protein phosphatase 4, catalytic subunit
10.69	Ublcp1	Ubiquitin-like domain containing CTD phosphatase 1
10.58	Ctdsp2	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 2
10.57	Ppm1b	PREDICTED: Protein phosphatase 1B, magnesium dependent, beta isoform
10.47	Dusp11	Dual specificity phosphatase 11 (RNA/RNP complex 1-interacting)
10.34	Phpt1	Phosphohistidine phosphatase 1
10.27	Ilkap	Integrin-linked kinase-associated serine/threonine phosphatase 2C
10.26	Ppp4r1	Protein phosphatase 4, regulatory subunit 1
10.2	Ppp1r10	Protein phosphatase 1, regulatory subunit 10
10.18	Ppm1g	Protein phosphatase 1G (formerly 2C), magnesium-dependent, gamma isoform
10.06	Ptp4a1	Protein tyrosine phosphatase 4a1
10.05	Ppp2r5d	Protein phosphatase 2, regulatory subunit B (B56), delta isoform

Table 19: Phosphatase inhibitors used for immunoprecipitation

Inhibitor	Concentration	Target
bpV(bipy)	1 µM	Protein tyrosine phosphatases

Inhibitor	Concentration	Target
4-methoxyphenacyl Br	500 µM	Protein tyrosine phosphatases
Phenylarsine oxide	50 µM	Protein tyrosine phosphatases
Sodium stibogluconate	100 µM	Protein tyrosine phosphatases
Chelerythrine chloride	50 µM	Dual specificity phosphatase 1
Ockadaic acid	50 nM	Protein phosphatases 2A, 1, 2B
Sodium fluoride	10 mM	Acid protein phosphatases
Imidazole	10 mM	Alkaline protein phosphatases
Beta-glycerophosphate	25 mM	Broad range
Alpha-naphthyl acid phosphate	10 mM	Broad range

Western blotting on eluted material with anti-NET29 affinity purified antibody showed a prominent band corresponding to the right molecular size of the NET29-GFP fusion (Fig 5.2 b, A). The rest of the eluted material was run on Bis-Tris 4-12% gradient gel (Invitrogen) and stained with colloidal coomassie blue. A faint band of similar size was visible in the anti-GFP immunoprecipitate but not in the IgG control eluate (Fig 5.2 b, B). Covalently cross-linked rabbit IgGs (Sigma) were used to pre-clear the lysate before incubation with the anti-GFP antibody. The 70 kDa band, as well as the 110 kDa, 60 kDa and 45 kDa bands, were cut out of the gel and mass spectrometry analysis of the thermolysin-digested gel pieces is ongoing in collaboration with Dr Thierry Le Bihan at The University of Edinburgh.

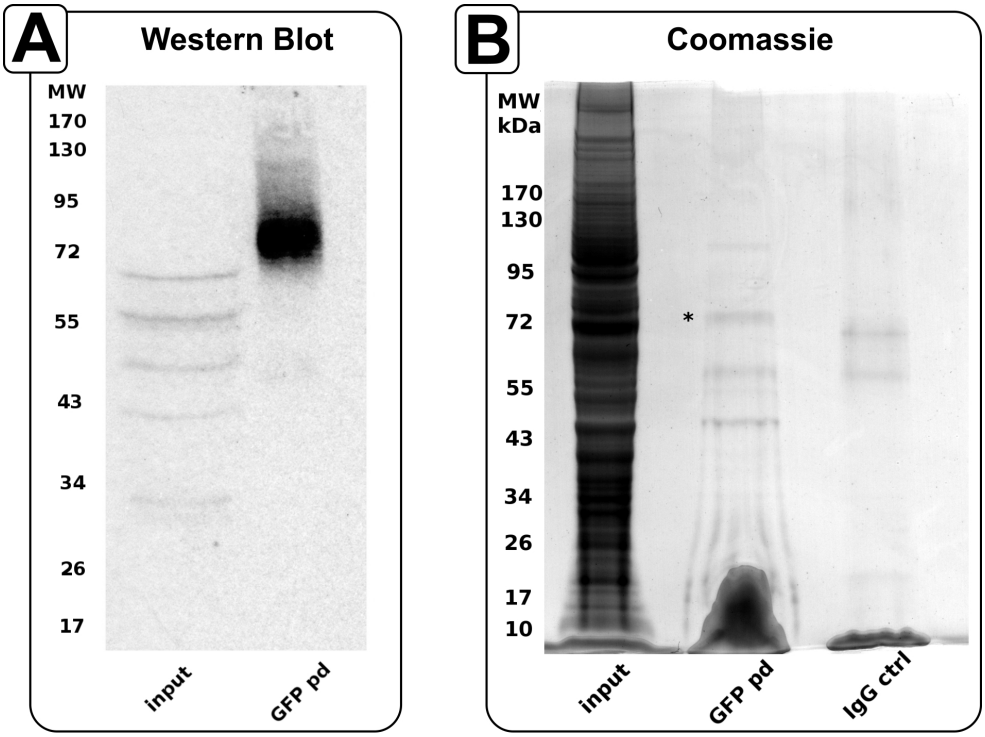


Fig 5.2 b: NET29-GFP immunoprecipitation

A - Western blot with anti-NET29 antibody shows enrichment of NET29 fused to GFP after immunoprecipitation with anti-GFP antibody. *3T3-L1 pre-adipocytes were transiently transfected with plasmid DNA encoding murine NET29 fused to GFP and immunoprecipitation using anti-GFP antibody was carried out.*

B - Coomassie staining of NET29-GFP immunoprecipitate. *Eluted material, as well as control-bound (rabbit IgG) material were run on SDS-PAGE and stained with colloidal Coomassie. Apart from the 70 kDa band that corresponds to NET29-GFP fusion size (asterisk), three other bands are visible and might represent co-precipitated partners of NET29.*

5.3 Identification of NET29 partners mediating chromosome repositioning effect

5.3.1 Development of *in vivo* cross-linking (IVCL) approach for identifying NET29 partner proteins

In order to obtain information about proteins, interacting with NET29 in general and, more specifically, interactors responsible for chromosome relocation to

the nuclear periphery, an *in vivo* cross-linking (IVCL) approach was developed and applied.

Transmembrane proteins have always been considered difficult to work with, but transmembrane proteins of the NE are particularly hard to solubilise and require strict conditions, such as high salt and detergent concentration and excessive sonication. However under harsh conditions, most if not all of the interactors are likely to get washed away. Alternatively, exogenous soluble fragments of transmembrane proteins could be used. This also incorporates a few potential drawbacks, including protein misfolding, mislocalisation to an unnatural compartment and lack of PTMs if the fragment is expressed in a different organism. Moreover, soluble fragments of NETs tend to auto-activate β-galactosidase expression from both hunt and prey vectors in yeast two-hybrid approach (Dr Iakowos Karakesisoglou, personal communication).

Taking into account all the above, an ideal method to co-precipitate NETs with their interacting partners would include, first of all, “freezing” of any interaction with the NET, followed by isolation of the “frozen” complexes under stringent conditions. Bifunctional covalent crosslinkers provide excellent means for “freezing” interactions, and a wide range of tags are now available which are capable of withstanding stringent conditions (reviewed in Lichty *et al.*, 2005).

IVCL approach might be useful not only for transmembrane proteins. Many interactions of soluble proteins are transient, and often lost during purification procedures. Thus, covalent cross-linking prior to purification theoretically provides a way to identify even such transient partners as kinases and phosphatases.

Crosslinker selection

Crosslinkers differ in the specificity of their functional groups, membrane permeability and cleavability.

Specificity of the crosslinker should be chosen based on the sequence of the protein of interest (Table 20). An excellent tool which helps to choose the most

suitable crosslinking compounds can be found online at the Pierce web site (http://www.piercenet.com/products/browse.cfm?fldID=26436A16-60A0-4A56-85F7-213A50830440).

The specificity of the crosslinker defines the probability of two proteins being crosslinked. The more amino acids possessing the targeted functional group in their side chain, the greater the probability. Homobifunctional crosslinkers targeting the C-terminal amino acid only have the least probability of covalently binding two given proteins, which decreases with increasing sizes of the proteins. Functional groups targeted by homobifunctional sulfhydryl crosslinkers are more frequent, as the frequency of cystein in an average protein is about 1.4% (RandSeq, computed from Swiss-Prot). Thus, the probability of two proteins being crosslinked by these compounds is no greater than 0.0196%[11]. The latter is based on the assumption that two interacting proteins are close to each other throughout their entire length. However this number ignores distances between the two cysteins, and since every crosslinker has different spacer arm length, this will affect the crosslinking probability. Occurrence of lysines is about 5.8% in average protein. Thus, a homobifunctional amine-targeting crosslinker has about 0.34% probability to detect an interaction, 17 times higher than with a homobifunctional sulfhydryl crosslinker. If a greater probability is required, non-selective and heterobifunctional crosslinkers are also available. However, excessive crosslinking might make it impossible to solubilize the protein extract, and thus the protein will not be effectively purified. It is advantageous to try a few crosslinkers of different specificity, and choose the optimal one by looking for either a band shift or absence of a corresponding band on Western blot[12].

Table 20: Crosslinker specificity

Target functional group	Reactive group	Instance crosslinkers
Carboxyl (C-terminus)	Carbodiimide	EDC (1-Ethyl-3-[3-dimethyl-

11 Due to rather complex nature of defining real probability of crosslinking of two proteins, the numbers here serve only to compare different types of crosslinkers, i.e. only relation of these numbers is meaningful.

12 High molecular weight often fail enter a gel or transfer effectively to a membrane

Target functional group	Reactive group	Instance crosslinkers
		aminopropyl]carbodiimide hydrochloride)
Amine (N-terminus, lysine and arginine side chains)	NHS ester Imidoester Aryl Halide PFP ester Hydroxymethyl phosphine	EGS (Ethylene glycol bis[succinimidyl-succinate]), DSP (Dithiobis[succinimidyl propionate]) DTBP (Dimethyl 3,3´-dithiobispropionimidate•2 HCl)
Sulfhydryl (cysteine side chain)	Maleimide Haloacetyl Br/I Pyridyldisulfide Vinyl Sulfone	BMB (1,4-Bismaleimidobutane), DTME (Dithiobismaleimidoethane)
Any	Diazirine Hydroxyphenyl Azide	NHS-ASA (N-Hydroxysuccinimidyl-4-azidosalicylic acid)

As the quintessence of the IVCL approach is to “freeze” interactions in the cells in a closest to native state, it is preferable to use membrane permeable compounds. This will allow a direct fixation of cells with crosslinker, similar to the widely used formaldehyde. However, if a non-permeable compound has to be used, it is should still be possible to use the smallest amount of mild detergent sufficient for the crosslinker to have access to the cell. Alternatively, cells can be pre-crosslinked with formaldehyde.

Cleavability of crosslinkers, on the other hand, is not as crucial as membrane permeability. Moreover, non-cleavable compounds add an opportunity to map exactly sites of extramolecular crosslinking, giving information about regions most likely to interact (Rappsilber *et al.*, 2000). However, for initial analysis having a cleavable crosslinker is beneficial as it allows specific elution step, which minimises the amount of non-specific contaminants in the eluate.

Determining optimal conditions of crosslinking

The reaction of crosslinking, as any chemical reaction, depends on many conditions such as concentration, temperature, pH and others.

The pH of the crosslinking reaction is the most important parameter, as even a

slight alteration might disrupt largely ionic binding of two proteins. Therefore, it is important to use a pH that is closest to the compartment of protein of interest and its possible binding partners (Table 21). Consequently, the pH of choice will limit the list of possible crosslinkers, as different reactive groups require different pH for optimal reaction. For example, the optimal pH for maleimides, functional groups of sulfhydryl crosslinkers, is 6.5-7.5, which is suitable for crosslinking proteins of the cytosol, but not those of the nucleus. On the other hand, imidoesters require more basic conditions of pH 8-10. This makes them more preferable in applications when nuclear or mitochondrial proteins are to be crosslinked.

Table 21: pH of cellular compartments

Compartment	pH	Reference
Cytosol	7.09-7.35	Bright *et al.*, 1987
Nucleus	7.74-7.95	Seksek and Bolard, 1996
Mitochondria	7.94-8.16	Abad *et al.*, 2004
Lysosome	4-5	Pillay *et al.*, 2002

The length and temperature of the crosslinking reaction are other factors to consider. If there is no particular requirement for the compound of choice, ambient temperature might be used for simplicity. In general, the higher the temperature, the faster the crosslinker reaction will proceed. However, excessive heat might denature proteins and strip their interactors. Consequently, 37°C should be an upper limit for IVCL when applied to most mammalian cells. Time, too, should be considered: a short time will introduce greater differences among samples, a long reaction time might lead to complexes dissociation or even proteolysis due to a slow rate of crosslinking. Furthermore, it will senselessly increase the total time of the procedure.

When the reaction's pH, temperature and time are set, the buffer of choice and crosslinker concentration are the next parameters to consider. The chosen buffer should support the set pH and must not contain competing functional groups. The crosslinker concentration should be tested empirically by band shifting on Western

blot: successful crosslinking will change the molecular mass of protein of interest by the mass of the crosslinked interactor.

Lysis and pull-down conditions

Immunoprecipitation with a specific antibody is the best method to purify a covalently crosslinked complex of proteins, both because the antibody is added after the crosslinking reaction, thus eliminating potential problems with tags. Additionally, it also allows crosslinking and purification of the endogenous protein, rather than an overexpressed exogenous one. However, it is not always available and, moreover, might fail to immunoprecipitate the complex due to an epitope being masked by a bound partner, which is especially relevant to monoclonal antibodies or polyclonal antibodies against a short peptide. Therefore, it is often necessary to use an affinity tag.

Conditions of lysis and purification depend largely upon the affinity tag used. Hence, the affinity tag should be chosen based on the crosslinker to be used. Ideally, the tag should allow for stringent binding and washing conditions and should lack the functional groups that are used for crosslinking. For example, the chitin binding tags affinity to chitin is lost after crosslinking with copper orthophenanthroline, a sulfhydryl crosslinker, most probably due to disruption of the tag's native structure (personal experience).

Since the crosslinking step introduces a covalent bond between the protein of interest and its partner protein, harsh conditions of lysis could be used to allow greater solubilization. Moreover, it is possible to use denaturating conditions for lysis and protein purification. However, these conditions should be compatible with the tag used. Only polyhistidines affinity to a metal ion column and biotins affinity to streptavidin are capable of withstanding denaturating conditions (Tagwerker *et al.*, 2006). On the other hand, non-ionic detergents such as Triton X-100, NP-40 and Tween 20 often may be used at high concentrations without altering tag affinities and column integrity (refer to the manufacturers manual for the solid matrix of choice).

Ultrasonic waves should be used for effective lysis, partially due to the genomic DNA shearing (McKee *et al.*, 1977). On the occasion when lysis conditions must be beyond the range of tag compatibility, a dialysis step might be introduced between lysis and purification to facilitate renaturation.

Overview of in vivo crosslinking approach

Chemical crosslinkers are widely used in biological applications, including in conjugation with fluorophores or enzymes to antibodies, haptens to cargos and more. Development of highly sensitive tandem mass spectrometry allowed use of these bifunctional compounds in structural biology by crosslinking purified complexes and mapping inter- and intramolecular interactions by analysing crosslinked products (Chen *et al.*, 2010).

The approach developed here depicted in Fig 5.3.1, differs from the latter considerably. First of all, purification and crosslinking steps are swapped in these two approaches. This allows the IVCL to catch transient interactions, but also might increase non-specificity. Second, the goal of IVCL is to identify partners of proteins that otherwise are difficult to co-immunoprecipitate, either because of harsh solubilising conditions or transiency of interactions. On the other hand, crosslinking of purified complexes aims to provide information about the structure of the complexes and their interacting domains. Although, once partners are identified, it is possible to use IVCL to gain information about domains that are most likely to interact by analysis of crosslinked products.

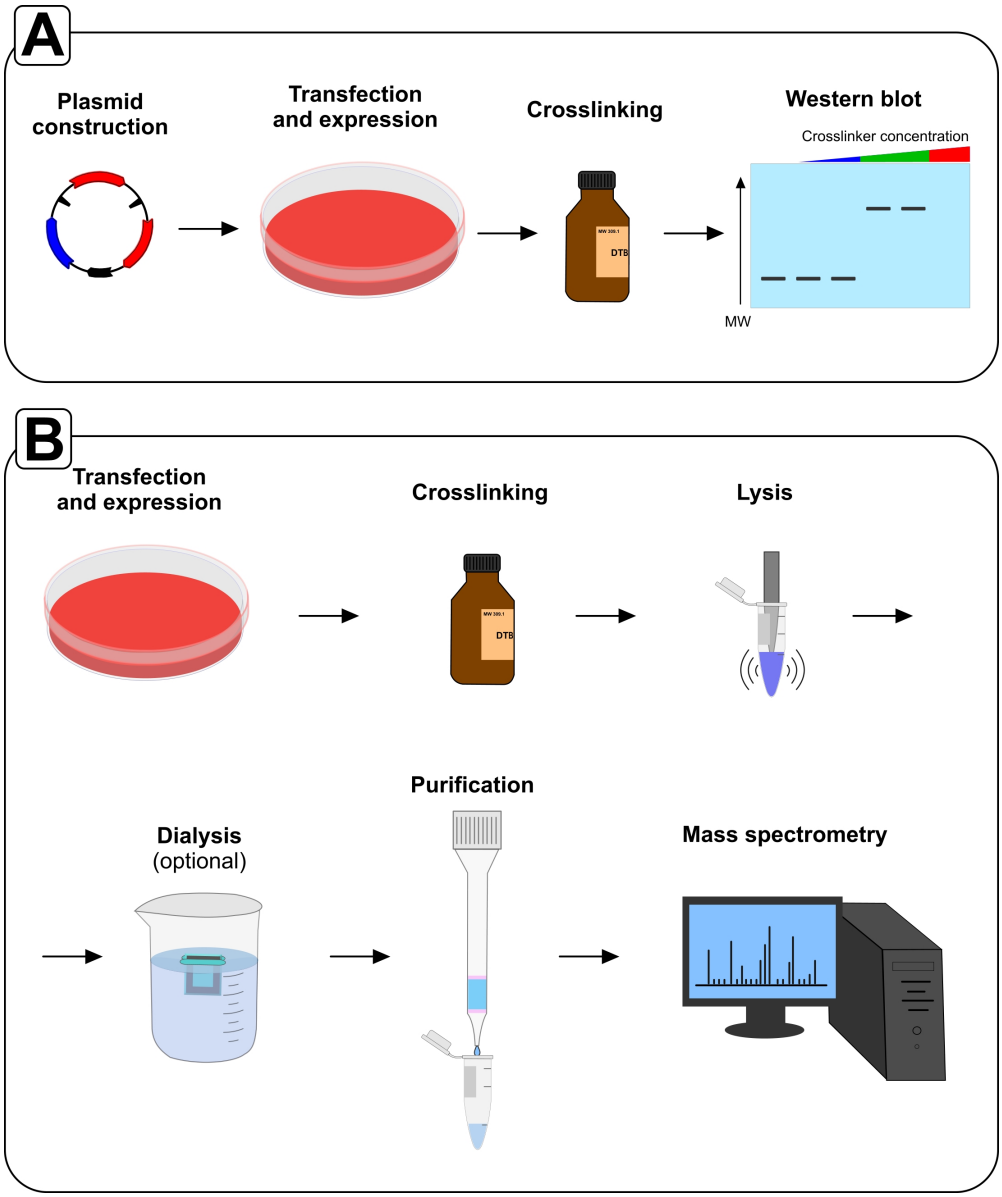


Fig 5.3.1: Schematic representation of the IVCL approach

***A – Initial testing**. DNA, encoding a protein of interest is subcloned into a plasmid, encoding an affinity tag and containing elements needed for expression of the fusion in target cells. The resulting plasmid is transfected into target cells and allowed to express. Different concentrations of crosslinker are applied and mobility of the fused protein is analysed with a specific antibody. The least concentration of the crosslinker that decreases mobility of the protein should be determined.*

***B – Final experiment**. Transfected cells are crosslinked using a determined concentration of crosslinker. After lysis, a dialysis step might be introduced to remove excess of detergent and salts. Cleavable crosslinkers allow gentle elution with washing buffer containing cleaving reagent. The eluate is then precipitated and analysed first by SDS PAGE, and then, if there are differences between the control and sample, by mass spectrometry.*

5.3.2 Application of the IVCL approach to NET29 and its mutant

Having established the outline of the method, IVCL was applied to SBP fusions of NET29 and its Y129F, Y133F mutant in order to determine its putative binding partner(s), responsible for the chromosome repositioning. The phospho-null mutant serves as the best control in this experiment, indeed, as it is different from wild type protein only by two amino acids. Hypothesising that phosphorylation of one or both of the tyrosines, mutated in the phospho-null mutant, is vital for an interaction with a protein directly or indirectly linked to chromatin, it is consequential, that the phospho-null mutant should lack or, rather, have much less of the putative interactor co-precipitated than wild type NET29.

Analysis of the region of NET29 that contains putative phosphorylation sites revealed a high number of lysines (**K**QA**K**FAY^{129}**K**DEY^{133}E**K**F**K**L), thus a homobifunctional amine-targeting membrane permeable cleavable crosslinker DTBP was chosen. In order to determine optimal crosslinking conditions, several concentrations of DTBP were used on 3T3-L1 cells at fixed pH and time. Resulting crosslinked material was lysed in non-reducing conditions and subjected to Western blotting using the anti-NET29 antibody (Fig 5.3.2 a). 0.01% DTBP shows insufficient crosslinking, 0.1% DTBP shows optimal crosslinking with NET29 monomer, dimer bands and a higher MW band that is specific to the wild type protein and not the phosphonull mutant. 0.2% DTBP shows excessive crosslinking. Based

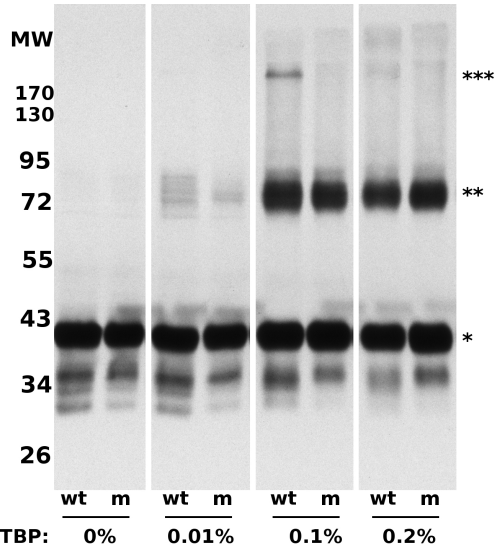


Fig 5.3.2 a: Determination of the optimal concentration of DTBP

*Different concentrations of DTBP were tested on 3T3-L1 cell expressing NET29 wild type protein (wt) or phosphonull mutant (m). * - NET29 monomer, ** - NET29 dimer, *** - NET29 dimer with a crosslinked binding partner.*

on this test, future experiments included crosslinking with 0.1% DTBP in triethanolamine pH 8.0 at room temperature for 15 min.

Next, the pulldown conditions were determined. NET29 was transfected into 3T3-L1 cells as a fusion with an SBP tag and different combinations of buffers and detergents were tried. SDS proved to abolish SBP binding to streptavidin at minimal concentrations. The optimal tested conditions for pull-down was 1% Triton X-100 in TEN_{500} buffer and excessive sonication (Fig 5.3.2 b, A).

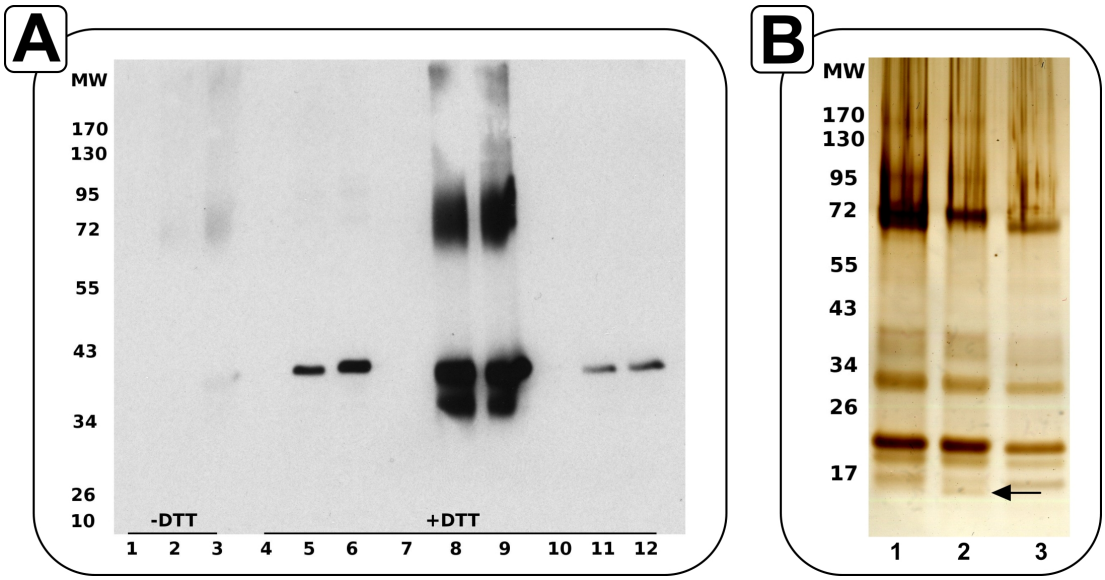


Fig 5.3.2 b: A potential 15 kDa mediator of chromosome repositioning by NET29

***A** - **IVCL followed by Western blot using anti-NET29 antibody.** 1,4 - SBP input; 2,5 - NET29-SBP input; 3,6 - NET29Y$^{129F/Y133F}$-SBP input; 7 - SBP pulled; 8 - NET29-SBP pulled; 9 - NET29Y$^{129F/Y133F}$-SBP pulled; 10 - SBP FT; 11 - NET29-SBP FT; 12 - NET29Y$^{129F/Y133F}$-SBP FT.*

***B** - **Silver stained gel of eluted material.** 1 - SBP eluate; 2 - NET29-SPB eluate; 3 - NET29$^{Y129F/Y133F}$ - SBP eluate. Arrow points at 15 kDa band that is co-precipitated with wild type NET29 fused to SBP.*

In the final experiment optimised conditions were used, but instead of harshly eluting everything from streptavidin beads, a gentle elution in washing buffer containing 100 mM of freshly added DTT was performed. This should have eluted only crosslinked material leaving bait (NET29-SBP) on the beads. Silver staining of the gel revealed a band of around 15 kDa that is present in NET29-SBP eluate but

not in phosphomimetic mutants eluate or the tag alone (Fig 5.3.2 b, B). Mass spectrometry analysis of a trypsin digest of this band is an ongoing collaboration with Dr Thierry Le Bihan.

5.4 Summary and conclusions

NET29 is evolutionary conserved and this suggests its important function in the cell. Combined analysis of its conservation and phosphorylation prediction allowed identification of its dominant-negative mutant. The mutant mimicking the state of the protein without phosphorylation does not reposition chromosome 6 to the nuclear periphery. Therefore it is likely that the mutations disrupt NET29 binding to a protein that mediates the chromosome repositioning effect. In order to understand the mechanism of the chromosome repositioning, it is important to identify this mediator.

Mutagenesis data suggests that NET29 exists mainly in its phosphorylated form, at least in 3T3-L1 cells. On the other hand, its dephosphorylation will lead to disengagement from the bound chromosome loci and, therefore, is likely to present a mechanism of the regulation of the chromatin distribution. The mechanism could function either in a time manner within the cell (i.e. in the cell cycle) or it could be distinct to a cell type (neurons, for example) and contribute to the tissue specific spatial genome organisation.

CHAPTER VI: Discussion

The NE is no longer considered as a simple barrier between the cytoplasm and components of the nucleus. Instead, it has been shown to actively participate in many vital processes. NETs have been linked to nuclear positioning, cell cycle, transcription, splicing, DNA replication and repair. Moreover, it is highly likely that INM proteins, functionally similar to lamins, are involved in regulation of chromatin organisation. This study aimed to uncover potential players of peripheral genome structuring and their possible post-translational regulation.

6.1 SUN2 and its potential links to chromatin

I first started out by looking at indication of differences in PTMs of abundant proteins in resting and activated lymphocytes. The cells have dramatic differences in the distribution of peripheral heterochromatin between the two states. Several proteins that seemed to have state-specific PTMs were re-analysed using the NE proteome datasets in order to confirm the predicted modifications. The analysis revealed many phosphorylated residues of SUN2, LAP2β, emerin and LBR. I have focused on SUN2 mainly because of its role in linking the cytoskeleton to the peripheral nucleoskeleton, and its involvement in telomere anchoring. However, no effect on heterochromatin or the cytoskeleton upon overexpression of the generated phosphomimetic or phospho-null mutants of SUN2 was found.

Immunoprecipitation followed by mass spectrometry identified histone H2AX as a potential SUN2 interactor and I hypothesised that the interaction brings persistent DNA double strand breaks, marked by phosphorylated H2AX, to the nuclear periphery. I have noticed many γ-H2AX-positive spots at the NE upon induced DNA damage. However, the soluble N-terminus of SUN2 did not show enrichment at these spots upon overexpression. The SUN2 part of the project was no longer pursued after this point, but there is a lot of potential remaining for SUN2 to be involved in chromatin rearrangement, cytoskeletal reorganisation or DNA double

strand break repair.

6.1.1 DNA repair and the NE

In yeast, DNA double strand breaks have been shown to relocate to the nuclear periphery and associate with a nuclear pore sub-complex (Nagai *et al.*, 2008). Oza *et al.* (2009) showed involvement of SUN domain-containing protein, Mps3p, in tethering of DNA DSB to the periphery. However, the molecular mechanism of this relocation remains unknown. One possible mechanism is suggested by my finding that the DNA damage indicator H2AX came down in immunoprecipitations with SUN2. Moreover, only the protocol that used DNAse I treatment before IP showed enrichment of H2AX in the co-precipitated material. Upon DNA damage, H2AX is rapidly phosphorylated and is recruited to the chromosome ends near DSBs. Therefore it would seem that the H2AX likely needs to be phosphorylated in response to DNA damage for it to interact with SUN2.

The same protocol also used a cysteine crosslinking reagent, therefore it is possible that intermolecular crosslinking rather than DNAse I treatment explains the histone enrichment. However the nucleoplasmic N-terminus of SUN2 does not contain cysteines and therefore, provided that the interaction takes place inside of the nucleus, crosslinking cannot explain this enrichment. Thus, I hypothesised that SUN2 interacts with γ-H2AX, the phosphorylated histone, and this interaction explains tethering of DSB to the NE.

In order to test the hypothesis, I first established conditions of DNA DSB generations. I used both a topoisomerase inhibitor etoposide and UV light to induce DNA damage, and noticed that a few hours after induction many γ-H2AX-positive loci were located at the NE. These most probably corresponded, as studies in yeast suggest, to slowly repaired breaks. Second, I generated a number of deletion mutants of SUN2, corresponding to its N-terminus. I showed that SUN2 is most likely to have 3 transmembrane domains and not one, as was described earlier (Hodzic *et al.*, 2004). This, however, preserves the previously suggested topology of the protein.

Contradictory to the initial hypothesis, the soluble N-terminus of SUN2 did not show enrichment at the γ-H2AX-positive loci.

There are a number of possible reasons why the soluble fragment of SUN2 did not co-localise with γ-H2AX and yet would not exclude the possibility of an interaction between full length SUN2 and the histone variant. First, folding of the soluble fragment could be different from the corresponding part of the full-length protein. However, this is something difficult to test as the structure of SUN2 is not yet resolved. Second, localisation at the NE could be mandatory for the interaction, either because of PTMs of SUN2 by a NE-anchored kinase or other enzyme, or because of forming heterodimers with another necessary component, for example SUN1. The latter raises possibility that it is SUN1 that interacts with γ-H2AX, and the histone was enriched in SUN2 IP due to SUN1/SUN2 heterodimerisation. This is supported by the observation in yeast where DSB were linked to a nuclear pore sub-complex (Nagai *et al.*, 2008). As it was mentioned in the introduction, SUN1 has also been shown to associate with the NPC (Liu *et al.*, 2007).

Thus, the interaction between SUN proteins and γ-H2AX remains to be confirmed. The most crucial experiment would be to effectively knock-down SUN2 (using siRNA) and quantify changes in peripheral γ-H2AX signal using the erosion macro that is used for chromosome repositioning quantification. If there would be a statistical difference between cells with and without SUN2, the project would prove viable. Then the possibility of interaction between γ-H2AX and C-terminus of SUN2 should also be investigated. Although the C-terminus of SUN2 is presumably located in the NE lumen (Hodzic *et al.*, 2004), it is still possible that a fraction of the protein adopts a different topology.

Mostly due to technical reasons (antibody availability and knockdown issues) and the time constrains, I had to switch the project. Several reasons led me to work on NET29. First of all, the fact that chromatin can be actively translocated to the nuclear periphery, as it was shown in the case of DNA double strand breaks. Secondly, a screen in our laboratory identified several NETs that are able to affect

positioning of the LacO array integrated into a chromosome. What's more, some of the NETs appeared to have tissue-specific expression raising the possibility of the NE to be involved in tissue-specific chromatin organisation.

6.2 NET29 contribution to positional gene regulation in adipogenesis

One of the hypotheses to explain mechanisms of the NE diseases implies involvement of NETs in chromatin organisation regulation. There are a number of chromatin-associated factors for which interaction with NETs and lamins have been shown. The factors include transcriptional repressors, DNA crosslinking factor BAF, histone modifying enzymes (histone deacetylase and histone acetyltransferase) and constitutive heterochromatin associated protein HP1 (refer to introduction). However, each of these proteins is widely expressed and cannot explain tissue-specific phenotypes of the nuclear envelopathies. Therefore, there must be tissue-specific NE components or chromatin-associated proteins, or combinations of both.

6.2.1 Tissue-specificity of NETs

Indeed, based on research I also contributed to, there seems to be considerable tissue-specificity in the NE proteome. Identification of a large number of putative NETs from three different tissues, liver, muscle and blood, allowed comparison of the NE proteome between the tissues. It is estimated that only about a third of all proteins identified by this approach were shared between the three tissues. Of course, it is only estimation and the number of identified putative NETs is expected to differ to some degree from the real “ultimate” NE proteome of a cell type. First of all, every putative NET needs to be confirmed to reside within the NE by staining with a specific antibody of the corresponding tissue section. We have shown that over 70% of the putative NETs show NE localisation upon overexpression of tagged fusion proteins (Malik *et al.*, 2010). The number might be higher as some of the proteins target to the NE only in specific cell lines, and not all possible combinations of protein/cell type were tested. Second of all, mass

spectrometry analysis of complex mixtures is inherently incomplete and multiple independent repeats should be done to identify the vast majority of the proteins in the mixture. We have also shown that the more repeats are done, the more comprehensive the list of identified proteins is, with 5 repeats being a reasonable number (N Korfali, GS Wilkie, SK Swanson, V Srsen, J de las Heras, DG Batrakou, ARW Kerr, L Florens, and EC Schirmer, paper in submission). Hence, for each tested tissue we have done at least 5 repeats of independent mass spectrometry runs. And third of all, the tissues analysed were from different species. Liver and muscle were isolated from rat, while peripheral blood was from human. However, great tissue variability of NETs is likely to reflect a general tendency, and the more tissues are analysed, the less the number of shared proteins would be.

Existing transcriptome data supports tissue-specificity of novel NETs. Moreover, we have done RT-PCR analysis of expression of some NETs in total RNA samples from multiple tissues and cell lines (Malik *et al.*, 2010; Wilkie *et al.*, 2010). This data also supports that certain NETs are more enriched in some tissues while being expressed at low levels in others. NET-specific antibody staining of commercial blots of normalised lysates from several human tissues also show significant differences in expression (N Zuleger, S Boyle, DA Kelly, J de Las Heras, DG Batrakou, V Lazou, GR Otti, N Korfali, KN Randles, GE Morris, DJ Harrison, WA Bickmore and EC Schirmer, in submission). Furthermore, we have shown that NETs may not localise to the NE in every cell line, suggesting that expression is not the only regulation mechanism of acquiring a tissue-specific NE proteome (Malik *et al.*, 2010).

6.2.2 NETs that are induced during adipogenesis

The original indication of NETs being upregulated in adipocytes came from transcriptome data available at biogps.org (Wu *et al.*, 2009). qRT-PCR confirmed upregulation of all three tested NETs; NET4, NET29 and NET33.

Potentially, NETs could play many different roles in adipogenesis. For

example, NET4 seems to withdraw cells from the cell cycle (Korfali *et al.*, 2011) and senescence is one of the distinct features of differentiated cells. NET33 belongs to family of scavenger receptors, molecules that recognise low-density lipoprotein and therefore might play a role in signalling or mechanically anchoring lipid droplets to the NE at the ONM. NET29 seems to play a role in chromatin reorganisation during adipogenesis.

The application of microarray analysis during the course of this study allowed the simultaneous screening of many affected genes. Analysis of only one time point of 3T3-L1 differentiation revealed over a thousand genes that are affected during adipogenesis, including at least 9 novel NETs including the three already tested by qRT-PCR. 5 of the other 6 NETs were found to be upregulated in adipogenesis. NET50 is an INM protein that is targeted almost exclusively to the NE upon overexpression. A half of the protein has similarity to dehydrogenase and therefore could be playing role in triglyceride biosynthesis. mNET20 that was identified in muscle NE proteome contains O-acyltransferase domain and also could be involved in triacylglycerol synthesis. NET62 contains similarity to both S-malonyl and acyl transferases. NET26 has an uncharacterised domain. Similarly, nothing is known about mNET22.

NET3 is downregulated in adipogenesis. It was shown to be a part of NPC (Mansfeld *et al.*, 2006). Interestingly, expression of many nucleoporins is diminished in adipogenesis. One possible explanation would be that differentiated cells are quiescent and no longer need to double the number of NPCs, so only nucleoporins with a short half-life are expressed at the same level.

6.2.3 Chromosome 6 relocation upon adipocyte differentiation

Historically, there were two models of chromatin organisation inside of the nucleus during interphase. The first model was proposed by Carl Rabl and enhanced by Theodor Boveri and postulated that interphase chromosomes have distinct territories. A second theory postulated that chromosomes are randomly tangled in a

"spaghetti" manner. The Cremer brothers proved the first theory correct (reviewed in Cremer and Cremer, 2001). However, later studies suggest that both theories were right. While it is clear that chromosomes do occupy distinct positions, studies visualising small chromosome loci clearly indicate that some of the loci are found far from boundaries of corresponding chromosome territories. Moreover, internally located chromosomes appear less compact and are, most probably, tangled to some degree. Thus, internally located chromosome 19 was retained in the nucleus upon high salt treatment, possibly due to interlinking with other central chromosomes (Croft *et al.*, 1999). Peripherally located chromosome 18, on the other hand, was released from the nuclei after the treatment.

A few studies have already shown chromosome relocation during differentiation (Stadler *et al.*, 2004; Chambeyron and Bickmore, 2004; Mayer *et* al., 2005), including adipogenesis (Kuroda *et al.*, 2004; Szczerbal *et al.*, 2009). Here we report that chromosome 6 is found more often at the nuclear periphery in 7 days differentiated 3T3-L1 adipocytes as compared to uninduced and 3 days differentiated cells.

It also has been reported that many chromosomes change their localisation upon the cells withdrawal from the cell cycle (Mehta *et al.*, 2010). This is clearly not the case with the chromosome 6 in 3T3-L1 cell differentiation. The 3T3-L1 cells were induced to differentiate after reaching confluency, and the mitotic expansion of the cells takes place during first two days of differentiation. However the chromosome 6 radial localisation is different between days 3 and 7 of differentiation.

The consequences of the chromosome movement are questionable. One expected outcome of this is a chromosome-wide effect on transcription. Thus, the ratio of up- and downregulated genes to the total number of genes on each chromosome was calculated and plotted[13] (Fig 6.2.3 a).

13 This analysis was done by Jose de las Jeras

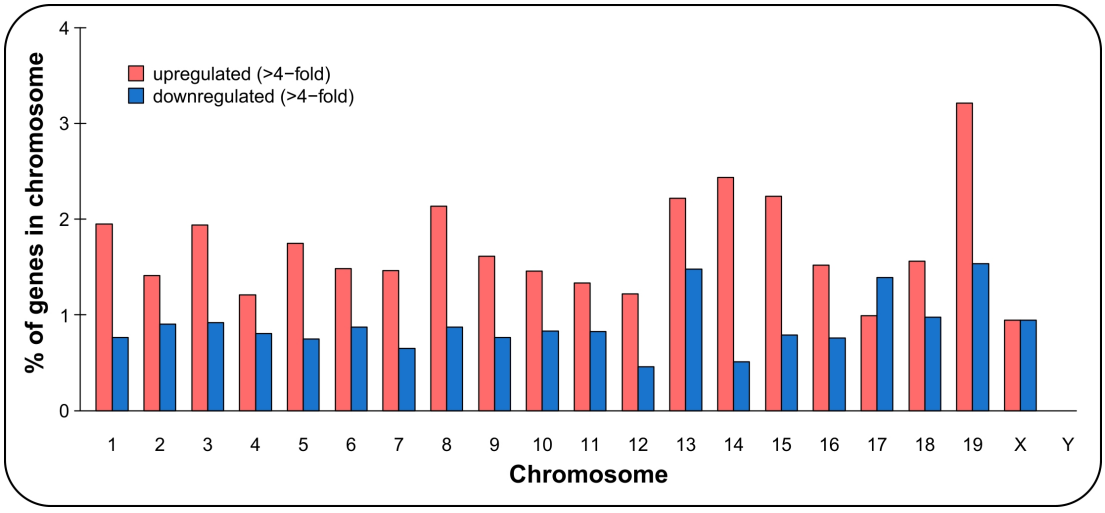


Fig 6.2.3 a: Relative changes in chromosomes expression during differentiation

Microarray data from undifferentiated and 9 days differentiated 3T3-L1 cells was analysed. For each chromosome the ratio of up- and downregulated genes to the total number of genes on the chromosome is plotted. The chromosome 6 expression profile seems similar to that of many other chromosomes.

However, changes in chromosome 6 did not stand out compared to changes in other chromosomes. This would seem to contradict the general idea of silencing at the nuclear periphery, as the number of activated genes is higher than the number of downregulated genes. However it is possible that the activation is associated with gene looping away from the periphery. FISH staining of specific genes loci should provide an answer to this possibility.

Of course, in order to draw any conclusions, repositioning analysis of all chromosomes should be done in this model system. Chromosomes 17 and 19 look particularly interesting as the first one seems to be comparatively most repressed and the latter - most activated in adipogenesis. Therefore analysis of the localisation of these two chromosomes in adipogenesis should indicate whether there is a correlation between chromosome positioning and whole chromosome transcription activation or repression.

Overall it seems that, at least in this model system, chromosome movement to

the nuclear periphery is not likely to directly correlate with general activation or repression of particular chromosomes. On the other hand, the effect could be more specific and affect particular loci on the chromosome. One possibility could be activation of specific genes by a membrane anchored transcription factor, or soluble transcription factor that is enriched at the NE due to binding to an NE-associated protein. In fact, SREBF1, a major transcription factor of adipogenesis, has been shown to interact with lamin A (Lloyd *et al.*, 2002). In this way, bringing a chromosome closer to the NE might facilitate activation of the genes that are under control of SREBF1. And, conversely, NE-associated transcription repressors could silence tethered loci. Mapping affected in adipogenesis genes to their chromosome loci shows that affected genes tend to cluster[14] (Fig 6.2.3 b). However, only extensive FISH analysis using BAC clones containing affected genes could provide evidence whether some of those loci are located at the NE. Alternatively, application of the DamID approach (Guelen *et al.*, 2008) to identify lamina-associated domains (LADs) of chromosomes in 3T3-L1 adipocytes, followed by comparison with the expression data, would also provide evidence regarding peripherally located chromosome loci.

14 This analysis was done by Jose de las Jeras

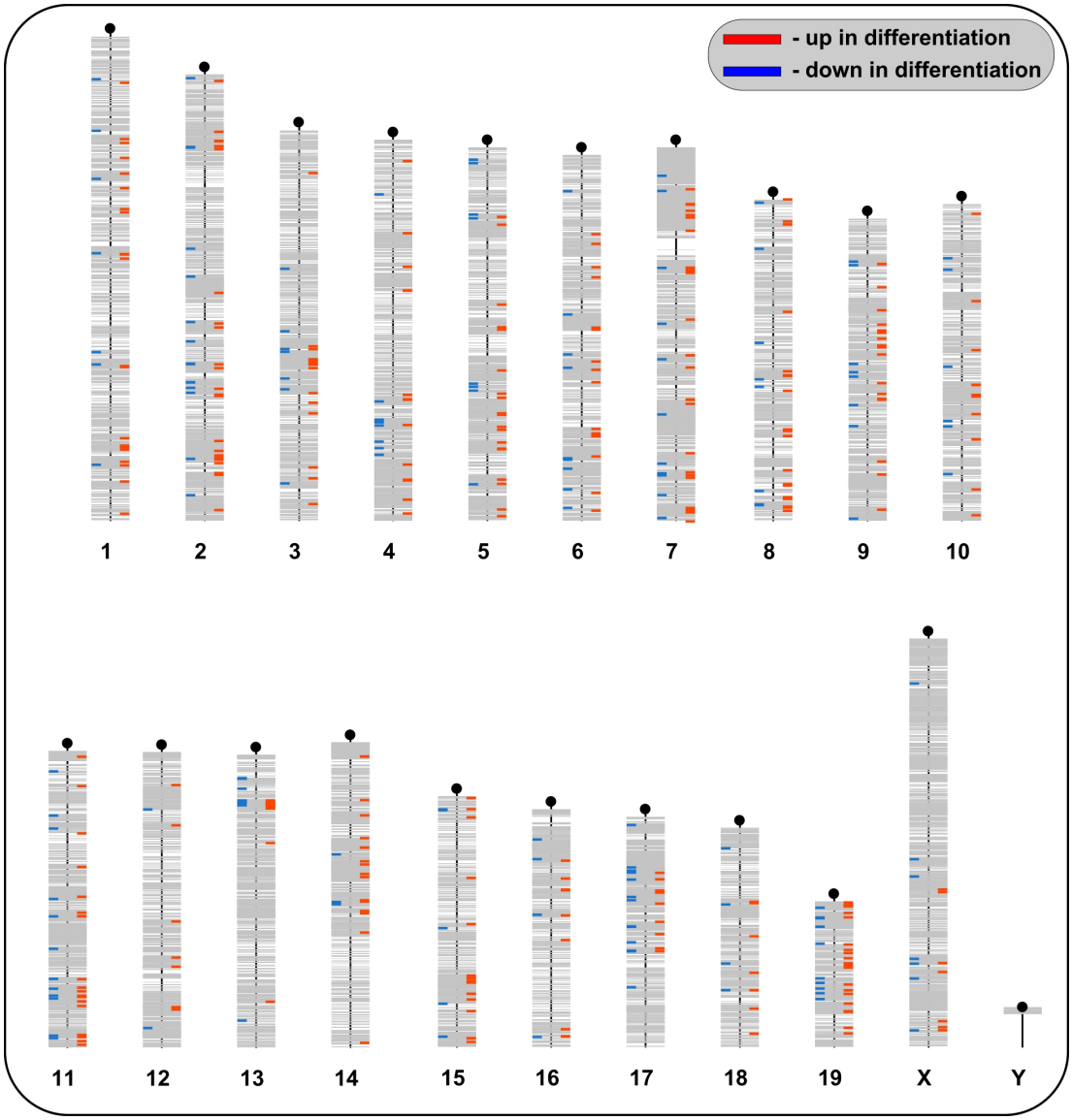


Fig 6.2.3 b: Genes affected in adipogenesis tend to cluster on chromosomes

Chromosome locations of genes upregulated (in red) and downregulated (in blue) during adipogenesis are mapped on the mouse chromosomes. The affected loci often coincide in the same chromosome region suggesting that the nuclear location of a chromosomal locus may affect several genes involved in adipogenesis.

Alternatively, chromosome repositioning might be neither a cause nor consequence of transcription activation/repression. One explanation could be that the chromosomes must be anchored to the NE in order to establish their chromosome

territories. Separation of individual chromosomes into the chromosome territories keeps them organised so that processes such as replication, mitotic condensation and radial repositioning would proceed flawlessly and without generating many DNA double strand breaks. However, in dividing cells they are anchored transiently to allow mitosis and, possibly, randomly to some degree. With cell quiescence, chromosomes should be attached permanently, which is achieved by expression of NE-associated proteins that possess affinity towards either specific chromosomes or a set of chromosomes. And since passive tethering that could occur during mitosis is not possible, cells have an active mechanism to translocate chromosomes towards the NE for anchoring.

Interestingly, a study utilising Giemsa-like FISH analysis of several human chromosomes argues that the radial chromatin distribution correlates with gene density (Küpper *et al.*, 2007). Thus more gene-rich domains are located in the centre of the nucleus, whereas gene-poor regions are found preferably at the periphery. Hence, chromosomes that contain more genes per chromosome size will have a more central location in the erosion macro analysis of the whole chromosome painting. It has also been shown that radial positioning of mouse chromosomes also correlates with gene density in several cell types, albeit to a lesser degree than in human cells (Mayer *et* al., 2005).

To address this hypothesis, ratios of known protein-coding genes to total DNA length have been plotted for each mouse chromosome (Fig 6.2.3 c). Consistent with the hypothesis, mouse chromosome 6 is not a gene-rich chromosome and therefore it is located peripherally in quiescent differentiated 3T3-L1 cells. Again, positional analysis of all chromosomes in differentiated adipocytes is needed to prove this correlation.

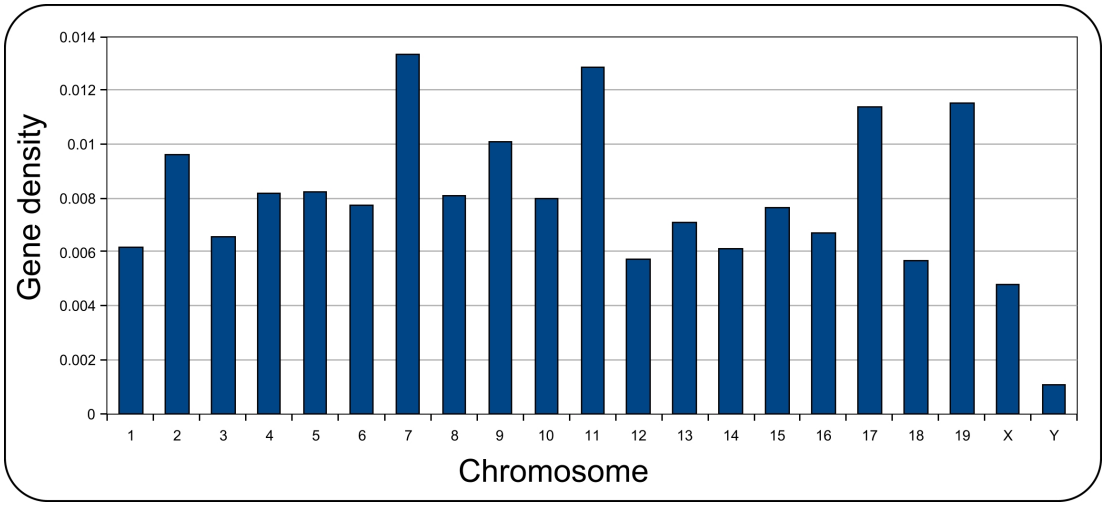
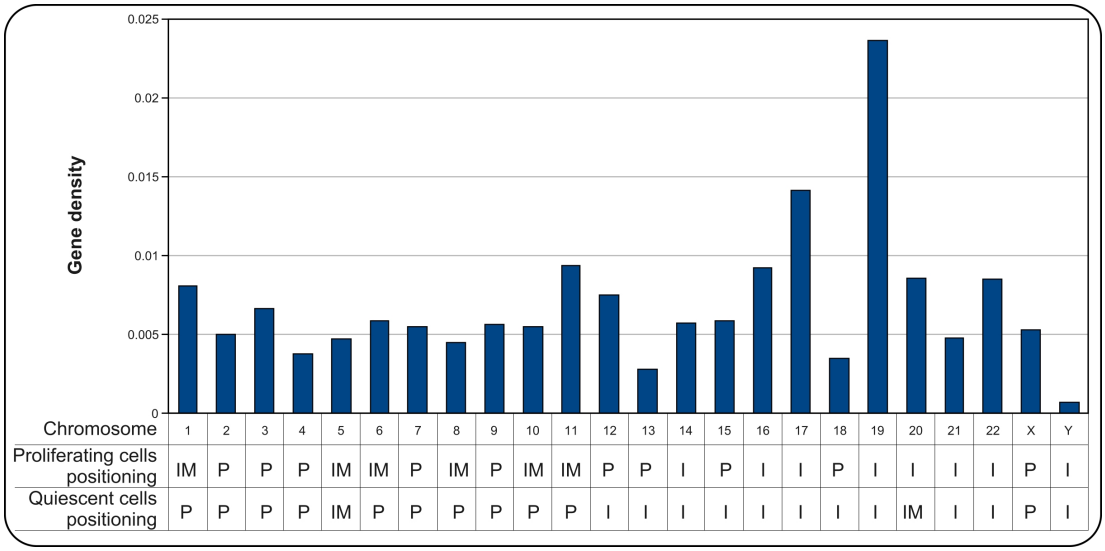


Fig 6.2.3 c: Relative gene densities of mouse chromosomes

Number of known protein-coding genes was plotted as a ratio to respective chromosome length in kb. Several chromosomes appear to be gene-rich: 2, 7, 9, 11, 17 and 19. Data was extracted from NCBIM37 database.

Data available for human cycling and quiescent fibroblasts chromosome positioning provides an opportunity to correlate a chromosomes gene density to its radial localisation in the nucleus (Mehta *et al.*, 2010). The ratio of a number of known genes to the corresponding human chromosome size was plotted (6.2.3 d). The radial positioning of every chromosome in proliferating and quiescent cells is marked as peripheral (P), internal (I) and intermediate (IM). However, no direct correlation is observed between gene density and chromosome radial positioning in any cells. In quiescent cells, however, chromosome positioning seems to correlate with their size (Mehta *et al.*, 2010). Thus, the large chromosomes are located peripherally and the small ones - internally.



Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y
Proliferating cells positioning	IM	P	P	P	IM	IM	P	IM	P	IM	IM	P	P	I	P	I	I	P	I	I	I	I	P	I
Quiescent cells positioning	P	P	P	P	IM	P	P	P	P	P	P	I	I	I	I	I	I	I	I	IM	I	I	P	I

Fig 6.2.3 d: Chromosome radial positioning does not correlate with gene density

Gene density is calculated as a ratio of a number of known coding genes to the corresponding chromosome size. ***P*** *- peripheral,* ***I*** *- internal,* ***IM*** *- intermediate. Chromosome positioning data are from Mehta et al., 2010; Bridger et al., 2000; Boyle et al., 2001; Meaburn et al., 2008 and Meaburn et al., 2007.*

6.2.4 Is NET29 likely to reposition during interphase or mitosis?

It is hypothesised that clonal expansion of 3T3-L1 cells, the two rounds of mitosis following addition of the differentiation cocktail, is needed to re-arrange chromatin within the nucleus and thus facilitate expression changes during differentiation (MacDougald and Lane, 1995). Although it has been shown that the majority of chromatin remodelling happens shortly after induction of differentiation, i.e. before mitoses (Siersbæk *et al.*, 2011), mitosis-dependent changes in whole chromosome localisations also seem to be important for cells. Thus, cells from patients with HGPS have abnormal chromosomes positioning (Mehta *et al.*, 2011). Consistent with this hypothesis, 3T3-L1 cells do not differentiate if clonal expansion is arrested at various cell cycle stages (Tang *et al.*, 2003).

On the other hand, several studies have shown that affinity mechanisms can drive changes in positioning of chromosome loci and even whole chromosomes

(Finlan *et al.*, 2008, Kumaran and Spector, 2008). Both studies used artificially created affinities and no endogenous NETs that are able to relocate chromosomes have been reported until now.

NET29, identified in the NE proteomics, had an effect in a chromosome repositioning screen carried out in our laboratory. The overexpression of NET29 in mouse pre-adipocytes leads to chromosome 6 being relocated to the nuclear periphery. Interestingly, chromosome 6 is also relocated to the nuclear periphery during normal adipocyte differentiation and that correlates with up-regulation of endogenous NET29. This suggests that NET29 might be one of the proteins that facilitate chromatin rearrangement during adipogenesis.

There seem to be at least two different mechanisms of chromosome repositioning: an ATP- and myosin-dependent and a mitosis-dependent. As NET29 is an inner nuclear membrane associated protein (Malik *et al.*, 2010), it might be involved in the cell cycle dependent mechanism. We have noticed that NET29 mediated chromosome repositioning was more pronounced after 3 days post transfection when compared to 2 days. Therefore we have assumed that cell division is required for the induced repositioning without directly testing this. However in order to support this, both behaviour of chromosome 6 and localisation of NET29 should be monitored in, and shortly after, mitosis.

Alternatively, NET29 could be a part of an active mechanism of bringing a chromosome to the periphery, similar to what was described for quiescent cells (Mehta *et al.*, 2010). In this case, NET29 would provide only anchor points at the NE. It is possible that the active mechanism would at least partially reuse the same machinery that brings persistent DNA double strand breaks to the periphery. In fact, there is a potential link between DNA repair and aberrant chromosome positioning. Thus, a particular family of microRNA was found to be upregulated in a HGPS mouse model and also upon DNA damage (Ugalde *et al.*, 2011). Proliferating HGPS fibroblasts, on the other hand, have a chromosome distribution similar to that of quiescent cells (Mehta *et al.*, 2011). Therefore chromosome repositioning might be

linked to DNA repair.

6.2.5 Effect of NET29 misregulation on transcription

A few studies aimed to find consequences of chromosome tethering to the NE. Thus, Kumaran and Spector (2008) have used LacO array integration into a human chromosome and lamin B fused to LacI provided tethering point for the integrated array. First, they showed that relocation of the array to the periphery required mitosis. Second, transcription of the integrated gene inserted into the same locus, under control of a minimal CMV promoter, was monitored. However its level was similar to that of control cells with the internally located locus. This lack of change could be due to the use of a viral promoter.

A similar mechanism was used by Reddy *et al.* (2008) with the exception that the GFP-LacI fusion was targeted to the NE by the C-terminal domain of emerin. This combination, however, showed repression of a gene included in the array marker. Finlan *et al.* (2008) also used an affinity between LacO and LacI, the last one being fused to the INM protein LAP2β. However they carried out microarray analysis of the cell line and found that, while chromosome-wide expression did not change significantly, genes located in 5 Mb regions flanking the array integration site did show repression.

Therefore, transcription of genes that are under control of endogenous promoters seems to be affected by tethering of chromatin to the nuclear periphery. The discrepancies between the effects this relocalisation had on the expression of genes within the vicinity of the tether maybe a consueqnce of differences in the choice of peripheral tethering protein, the interaction partners of the tethering protein or in the exent and distribution of heterochromatin in the various cell lines used. Nonetheless, both studies used artificially created affinities between chromatin and the NE. On the other hand, NET29-mediated tethering allowed us to investigate the *vivo* consequences of chromosome relocation to the NE following endogenous interactions.

Overexpression of NET29 in pre-adipocytes and, more so, down-regulation of NET29 in differentiating cells showed an effect on transcription. Affected genes were located on several different chromosomes, including chromosome 6. The other chromosomes could potentially be affected by NET29 in two ways. First, some of the chromosomes could also be tethered to the nuclear periphery by NET29 overexpression and, conversely, fail to relocate upon NET29 downregulation. Alternatively, repositioning of a chromosome to the nuclear periphery is likely to displace another peripherally located chromosome with weaker affinity to the NE due to spatial constraints. Hence, overexpression of NET29 could indirectly cause some of the chromosomes to move away from the periphery. Therefore, all affected chromosomes should be tested individually in order to further validate microarray data.

It was noticed that some of the affected genes are important for myogenesis. Myogenesis is one of the alternative differentiation pathways for mesenchymal cells, which give rise to adipocytes. Thus, the gene that encodes chemokine (C-X-C motif) ligand 12 (*Cxcl12*, also known as *Sdf1*), located on chromosome 6, was upregulated upon NET29 knockdown. *Cxcl12* has been shown to induce myogenic differentiation (Melchionna *et al,*. 2010). On the other hand prolactin family 2, subfamily c, member 2 (*PRL2C2*, also known as proliferin) was downregulated upon NET29 knockdown. It has been shown to inhibit myogenic differentiation (Wilder and Linzer, 1989). Therefore, the effect of NET29 upregulation in adipogenesis correlates with blocking myogenic differentiation. Given common progenitors of both fat and muscle tissue, it seems logical that activation of adipocyte differentiation should be accompanied by repression of differentiation into muscle or bone.

It is not entirely clear whether the observed changes in transcription are due to different localisation of chromosome 6 or any other untested chromosome, or if it is due to other unknown reasons. Moreover, it is not known whether chromosome 6 fails to relocate to the nuclear periphery in differentiating cells with downregulated NET29. I have made several attempts to find this out, but FISH on differentiated

cells has stopped giving a specific signal. This is most probably due to different reagents being used as some of the originally used reagents are no longer supplied. While experiments in pre-adipocytes show that NET29 is a sufficient factor of the chromosome movement, it is not known whether NET29 is the only factor responsible for chromosome 6 peripheral location in adipocytes. What's more, there is a NET29 paralog, TMEM120B, which is similar to NET29 and is also upregulated during adipocyte differentiation. Therefore a double knockdown might be needed in order to block chromosome 6 relocation in differentiation.

The possibility remains that the changes in gene expression are not completely due to direct tethering of the chromosome loci, but are a consequence of affected regulators of gene expression such as microRNA, transcription inhibitors and activators. Indeed, if tethering of a locus containing a transcription regulator to the periphery affects its transcription, expression of its target genes will also be affected despite the localisation of the latter.

6.2.6 The dominant negative mutant of NET29

Multiple alignment of the NET29 protein sequence from different species revealed a number of highly conserved regions in the protein, most notably a region within the N-terminus that is next to the NET29 first transmembrane sequence. As NET29 is an INM protein, it is likely that its N-terminus is involved in chromosome repositioning. Furthermore, this region contains highly conserved tyrosines that are predicted to be phosphorylated by available algorithms. NetPhosK (Blom *et al.*, 2004) weakly predicts the two tyrosins to be insulin receptor kinase targets. It has been shown that the insulin receptor, originally membrane anchored in the cytoplasm, is cleaved upon stimulation with insulin and translocated into the nucleus while retaining kinase activity (Podlecki *et al.*, 1987). Hence we hypothesised that adding differentiating media containing insulin to 3T3-L1 cells promotes insulin receptor cleavage and translocation into the nucleus, where it is able to phosphorylate NET29. This, in turn, might regulate the affinity between NET29 and chromatin. In order to test the hypothesis, the two tyrosines were mutated to either phenylalanines

or glutamic acids. Phenylalanine is very similar to tyrosine except that it is missing the OH-group and therefore cannot be phosphorylated. The glutamic acid residue, on the other hand, is negatively charged and, therefore, mimics the phosphorylated tyrosine state.

Overexpression of the double phosphomimetic mutant of NET29 had a similar to wild type effect on chromosome 6 relocation. However, the double phospho-null mutant abolished the effect and chromosome 6 was found in the centre of the nucleus even more often than in untransfected cells. This defines the phospho-null mutant as a dominant negative as it is able to diminish tethering effect of endogenously expressed wild type NET29 (Herskowitz, 1987). Comparison of the effects of phosphomimetic and phospho-null mutants to wild type protein suggests that NET29 is constitutively phosphorylated in pre-adipocytes. Hence chromosome 6 repositioning effects in 3T3-L1 cells depend largely upon NET29 upregulation rather than post-translationally regulation, in contrast to our original hypothesis. However, it also most probably requires phosphorylation and it is possible that PTMs play an important regulatory role in the effect in other cell types or, for example, in mitosis.

The dominant negative effect of the NET29 mutant suggests that the protein forms oligomers, and each monomer must be phosphorylated in order for complex to have affinity to chromatin. This way, if the complex includes a non-phosphorylated monomer, it is unable to interact with a chromatin-associated protein.

Identification of a dominant negative mutant of NET29 offers us a unique reagent for pulldown experiments aiming to find NET29 binding partner(s) that mediate the chromosome tethering effect.

6.2.7 What makes repositioned chromosomes special?

The fact that specific chromosomes are repositioned within the nucleus, either rapidly or in a mitosis-dependent manner, suggests that the chromosomes have distinct features that are part of the affinity mechanisms.

One possible way of recognition is by general features such as A,T-rich

regions corresponding to G bands, or a chromosome size. The last could be explained by low non-specific affinity between the NE and chromosomes, and therefore the bigger the chromosome, the more affinity towards the NE it has. This, however, does not account for differences in chromosome compaction where a larger chromosome could occupy less volume than a smaller one.

Another possibility is that each (or some) chromosome has a distinct feature that is potentially recognised by feature-specific factors. The feature is most likely to be sequence-specific, for example a short repetitive sequence within "junk" DNA, unique for each chromosome. This way, each chromosome potentially has a specific sequence that is recognised by a "reader" - a protein that binds that sequence. Hence, interaction of the "reader" with an NE-anchored protein might provide means of tethering the chromosome to the NE. And loss of the interaction, either due to downregulation of one of the interactors expression (or both) or through their PTMs, will release the chromosome into the nuclear interior.

In order to investigate the possibility of chromosome barcoding, I would like to set up a *de novo* pattern search. The proposed brute force attack algorithm scheme should generate and check all possible sequence variations within a defined window. If searched for a 10 nucleotide long pattern, 4^{10} (1048576) combinations are possible. Each combination should go through a pattern search in individual chromosome sequences and ratios between hit count and chromosome length should be compared between all chromosomes. A barcode pattern would be enriched in one chromosome compared to the other ones. However, if one combination takes about 1-2 min to compute (close to actual time), the whole search will take several years to complete. Therefore, either the task should be divided between many computers, or a faster search algorithm should be implemented.

According to this chromosome barcode hypothesis, NET29 might play a role of NE component that is able to tether a specific chromosome to the nuclear periphery by interacting with a "reader" protein.

6.3 Future directions

There are a number of questions that need to be answered before the role of NET29 in chromosome repositioning will be understood. First of all, whether the cell cycle is required for the repositioning effect to take place. This could be done by establishing a stable cell line transfected with a plasmid encoding NET29 under an inducible promoter. FISH analysis of chromosome 6 after simultaneously inducing quiescence and NET29 overexpression in the cells should point out a mitosis requirement, unless the chromosome undergoes repositioning upon cell quiescence in a control cell line. Second, investigation of chromosome 6 position in differentiated cells with downregulated NET29 must be carried out. This will show whether NET29 or a complex its associated with, is the only factor responsible for relocation of chromosome 6 in adipocytes. FISH staining using BAC clones containing genes affected by NET29 might provide evidence of correlation between specific loci positioning and transcription of its genes. Also, uncovering of a molecular mechanism of the chromosome re-localisation will be a priority task. Mass spectrometry analysis of purified NET29 might confirm involvement of PTMs in the mechanism. Despite the outcome of phosphorylation analysis, the identified dominant negative mutant of NET29 serves as an ideal control for pulldown experiments aiming to find mediators of the chromosome repositioning. To identify possible characteristic features of chromatin associated with NET29 at the nucleotide sequence level, ChIP followed by next generation sequencing should be applied. Alternatively, the DamID approach could be used.

6.4 Final remarks

To summarise the current model, NET29 is an INM phosphoprotein that is expressed at a low level in pre-adipocytes. Upon terminal differentiation it is upregulated and its increased concentration at the NE leads to chromosome repositioning (Fig 6.4, A), presumably by a mitosis-dependent affinity mechanism during the clonal expansion of 3T3-L1 cells (Fig 6.4, B). The chromosome

repositioning includes tethering of some chromosomes to the nuclear periphery, which, in turn, leads to releasing of other peripheral chromosomes into the nuclear interior. The rearrangement also affects transcription of several genes either directly (by bringing them to the silencing peripheral environment or releasing from it) or indirectly (by altering expression of transcription regulators such as transcription factors, repressors and microRNA). And, finally, changes in transcription facilitate differentiation by suppressing alternative differentiation pathways.

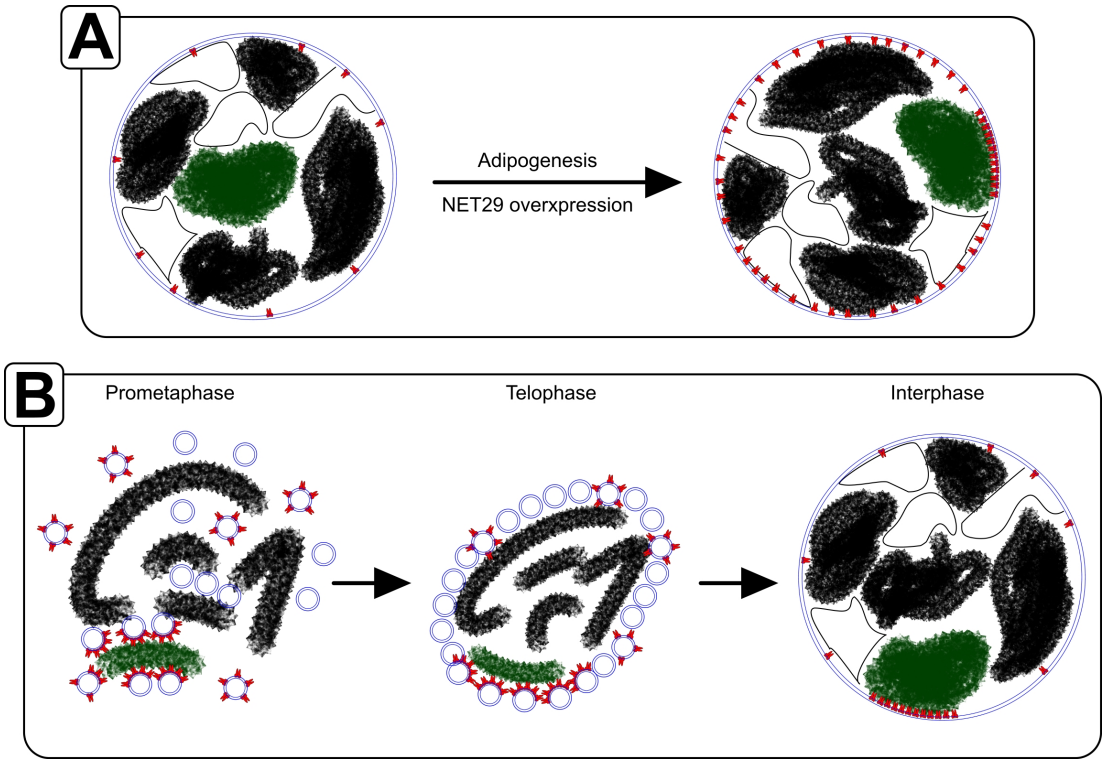


Fig 6.4: The current model of NET29-mediated chromosome repositioning effect

***A** - **Upregulation of NET29 leads to chromosome tethering.** In pre-adipocytes chromosome 6 (green) is located internally. Upregulation of either endogenous or exogenous NET29 leads to its increased concentration at the NE and a higher avidity to chromosome 6. Tethering of chromosome 6 to the periphery possibly displaces another chromosome into the nuclear interior.*

***B** - **Mitosis-dependent mechanism of chromosome tethering.** Upon NE breakdown in prometaphase membranes form micelles that comprise NETs. Micelles that include NET29 are concentrated at chromosome 6. In telophase the NE is assembled back from the micelles. Affinity between NET29 and chromosome 6 brings the chromosome to the nuclear periphery, where it stays during interphase.*

References

Abad MFC, Di Benedetto G, Magalhães PJ, Filippin L and Pozzan T. Mitochondrial pH monitored by a new engineered green fluorescent protein mutant. *J. Biol. Chem.* 2004

Akhtar A and Gasser SM. The nuclear envelope and transcriptional control. *Nat. Rev. Genet.* 2007

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997

Apel ED, Lewis RM, Grady RM and Sanes JR. Syne-1, a dystrophin- and Klarsicht-related protein associated with synaptic nuclei at the neuromuscular junction. *J. Biol. Chem.* 2000

Bae YH, Ding Z, Das T, Wells A, Gertler F and Roy P. Profilin1 regulates PI(3,4)P2 and lamellipodin accumulation at the leading edge thus influencing motility of MDA-MB-231 cells. *Proc. Natl. Acad. Sci. U.S.A.* 2010

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC and Kouzarides T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 2001

Bártová E and Kozubek S. Nuclear architecture in the light of gene expression and cell differentiation studies. *Biol. Cell* 2006

Batrakou DG, Kerr ARW and Schirmer EC. Comparative proteomic analyses of the nuclear envelope and pore complex suggests a wide range of heretofore unexpected functions. *J Proteomics* 2009

Bione S, Maestrini E, Rivella S, Mancini M, Regis S, Romeo G and Toniolo D. Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular

dystrophy. *Nat. Genet.* 1994

Blom N, Gammeltoft S and Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 1999

Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S and Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 2004

Boguslavsky RL, Stewart CL and Worman HJ. Nuclear lamin A inhibits adipocyte differentiation: implications for Dunnigan-type familial partial lipodystrophy. *Hum. Mol. Genet.* 2006

Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA and Bickmore WA. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum. Mol. Genet.* 2001

Bridger JM, Boyle S, Kill IR and Bickmore WA. Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. *Curr. Biol.* 2000

Bridger JM, Kill IR and Lichter P. Association of pKi-67 with satellite DNA of the human genome in early G1 cells. *Chromosome Res.* 1998

Bright GR, Fisher GW, Rogowska J and Taylor DL. Fluorescence ratio imaging microscopy: temporal and spatial measurements of cytoplasmic pH. *J. Cell Biol.* 1987

Burma S, Chen BP, Murphy M, Kurimasa A and Chen DJ. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J. Biol. Chem.* 2001

Callan HG and Tomlin SG. Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. *Proc. R. Soc. Lond., B, Biol. Sci.* 1950

Chambeyron S and Bickmore WA. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* 2004

Chen SS, Chen JF, Johnson PF, Muppala V and Lee YH. C/EBPbeta, when expressed from the C/ebpalpha gene locus, can functionally replace C/EBPalpha in liver but not in adipose tissue. *Mol. Cell. Biol.* 2000

Chen Z, Torrens JI, Anand A, Spiegelman BM and Friedman JM. Krox20 stimulates adipogenesis via C/EBPbeta-dependent and -independent mechanisms. *Cell Metab* 2005

Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills J, Nilges M, Cramer P and Rappsilber J. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* 2010

Chi Y, Haller K, Peloponese JJ and Jeang K. Histone acetyltransferase hALP and nuclear membrane protein hsSUN1 function in de-condensation of mitotic chromosomes. *J. Biol. Chem.* 2007

Chilson OP, Boylston AW and Crumpton MJ. Phaseolus vulgaris phytohaemagglutinin (PHA) binds to the human T lymphocyte antigen receptor. *EMBO J.* 1984

Conrad MN, Lee C, Chao G, Shinohara M, Kosaka H, Shinohara A, Conchello J and Dresser ME. Rapid telomere movement in meiotic prophase is promoted by NDJ1, MPS3, and CSM4 and is modulated by recombination. *Cell* 2008

Cousin W, Fontaine C, Dani C and Peraldi P. Hedgehog and adipogenesis: fat and fiction. *Biochimie* 2007

Cremer T and Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2001

Crisp M, Liu Q, Roux K, Rattner JB, Shanahan C, Burke B, Stahl PD and Hodzic D. Coupling of the nucleus and cytoplasm: role of the LINC complex. *J. Cell Biol.* 2006

Croft JA, Bridger JM, Boyle S, Perry P, Teague P and Bickmore WA. Differences in the localization and morphology of chromosomes in the human nucleus. *J. Cell Biol.* 1999

Deniaud E and Bickmore WA. Transcription and the nuclear periphery: edge of darkness?. *Curr. Opin. Genet. Dev.* 2009

Dölken L, Ruzsics Z, Rädle B, Friedel CC, Zimmer R, Mages J, Hoffmann R, Dickinson P, Forster T, Ghazal P and Koszinowski UH. High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA* 2008

Dreger M, Bengtsson L, Schöneberg T, Otto H and Hucho F. Nuclear envelope proteomics: novel integral membrane proteins of the inner nuclear membrane. *Proc. Natl. Acad. Sci. U.S.A.* 2001

Fajas L, Debril MB and Auwerx J. Peroxisome proliferator-activated receptor-gamma: from adipogenesis to carcinogenesis. *J. Mol. Endocrinol.* 2001

Fajas L, Fruchart JC and Auwerx J. Transcriptional control of adipogenesis. *Curr. Opin. Cell Biol.* 1998

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ and van Steensel B. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* 2010

Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR and Bickmore WA. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* 2008

Fischer JA, Acosta S, Kenny A, Cater C, Robinson C and Hook J. Drosophila klarsicht has distinct subcellular localization domains for nuclear envelope and microtubule localization in the eye. *Genetics* 2004

Flegal KM, Carroll MD, Ogden CL and Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. *JAMA* 2010

Foisner R and Gerace L. Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. *Cell* 1993

Francastel C, Magis W and Groudine M. Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc. Natl. Acad. Sci. U.S.A.* 2001

Fridolfsson HN, Ly N, Meyerzon M and Starr DA. UNC-83 coordinates kinesin-1 and dynein activities at the nuclear envelope during nuclear migration. *Dev. Biol.* 2010

Furukawa K. LAP2 binding protein 1 (L2BP1/BAF) is a candidate mediator of LAP2-chromatin interaction. *J. Cell. Sci.* 1999

Gant TM, Harris CA and Wilson KL. Roles of LAP2 proteins in nuclear assembly and DNA replication: truncated LAP2beta proteins alter lamina assembly, envelope formation, nuclear size, and DNA replication efficiency in Xenopus laevis extracts. *J. Cell Biol.* 1999

Gerace L, Comeau C and Benson M. Organization and modulation of nuclear lamina structure. *J. Cell Sci. Suppl.* 1984

Glass CK, Rose DW and Rosenfeld MG. Nuclear receptor coactivators. *Curr. Opin. Cell Biol.* 1997

Göb E, Schmitt J, Benavente R and Alsheimer M. Mammalian sperm head formation involves different polarization of two novel LINC complexes. *PLoS ONE* 2010

Greber UF, Senior A and Gerace L. 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.* 1990

Green H and Kehinde O. An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion. *Cell* 1975

Grønborg M, Kristiansen TZ, Stensballe A, Andersen JS, Ohara O, Mann M, Jensen ON and Pandey A. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol. Cell Proteomics* 2002

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W and van Steensel B. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 2008

Gyuris J, Golemis E, Chertkov H and Brent R. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 1993

Hallberg E, Wozniak RW and Blobel G. An integral membrane protein of the pore membrane domain of the nuclear envelope contains a nucleoporin-like region. *J. Cell Biol.* 1993

Haque F, Mazzeo D, Patel JT, Smallwood DT, Ellis JA, Shanahan CM and Shackleton S. Mammalian SUN protein interaction networks at the inner nuclear membrane and their role in laminopathy disease processes. *J. Biol. Chem.* 2010

Hedgecock EM and Thomson JN. A gene required for nuclear and mitochondrial attachment in the nematode Caenorhabditis elegans. *Cell* 1982

Hepperger C, Mannes A, Merz J, Peters J and Dietzel S. Three-dimensional positioning of genes in mouse cell nuclei. *Chromosoma* 2008

Herskowitz I. Functional inactivation of genes by dominant negative mutations. *Nature* 1987

Hodzic DM, Yeater DB, Bengtsson L, Otto H and Stahl PD. Sun2 is a novel

mammalian inner nuclear membrane protein. *J. Biol. Chem.* 2004

Hoffenberg S, Liu X, Nikolova L, Hall HS, Dai W, Baughn RE, Dickey BF, Barbieri MA, Aballay A, Stahl PD and Knoll BJ. A novel membrane-anchored Rab5 interacting protein required for homotypic endosome fusion. *J. Biol. Chem.* 2000

Hoffmann K, Dreger CK, Olins AL, Olins DE, Shultz LD, Lucke B, Karl H, Kaps R, Müller D, Vayá A, Aznar J, Ware RE, Sotelo Cruz N, Lindner TH, Herrmann H, Reis A and Sperling K. Mutations in the gene encoding the lamin B receptor produce an altered nuclear morphology in granulocytes (Pelger-Huët anomaly). *Nat. Genet.* 2002

Hofmann K and Stoffel W. TMBASE - A database of membrane spanning protein segments. , 1993.

Holaska JM, Lee KK, Kowalski AK and Wilson KL. Transcriptional repressor germ cell-less (GCL) and barrier to autointegration factor (BAF) compete for binding to emerin in vitro. *J. Biol. Chem.* 2003

Horton JD, Goldstein JL and Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin. Invest.* 2002

Inokuchi J. Insulin resistance as a membrane microdomain disorder. *Biol. Pharm. Bull.* 2006

Ishii K, Arib G, Lin C, Van Houwe G and Laemmli UK. Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* 2002

Jungwirth MT, Kumar D, Jeong DY and Goodchild RE. The nuclear envelope localization of DYT1 dystonia torsinA-deltaE requires the SUN1 LINC complex component. *BMC Cell Biol.* 2011

Kalocsay M, Hiller NJ and Jentsch S. Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. *Mol. Cell* 2009

Kampstra P. Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *Journal of Statistical Software, Code Snippets* 2008

Kennedy C, Sebire K, de Kretser DM and O'Bryan MK. Human sperm associated antigen 4 (SPAG4) is a potential cancer marker. *Cell Tissue Res.* 2004

Ketema M, Wilhelmsen K, Kuikman I, Janssen H, Hodzic D and Sonnenberg A. Requirements for the localization of nesprin-3 at the nuclear envelope and its interaction with plectin. *J. Cell. Sci.* 2007

Kim JB and Spiegelman BM. ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. *Genes Dev.* 1996

Kim JB, Sarraf P, Wright M, Yao KM, Mueller E, Solanes G, Lowell BB and Spiegelman BM. Nutritional and insulin regulation of fatty acid synthetase and leptin gene expression through ADD1/SREBP1. *J. Clin. Invest.* 1998

Kim JB, Spotts GD, Halvorsen YD, Shih HM, Ellenberger T, Towle HC and Spiegelman BM. Dual DNA binding specificity of ADD1/SREBP1 controlled by a single amino acid in the basic helix-loop-helix domain. *Mol. Cell. Biol.* 1995

King MC, Drivas TG and Blobel G. A network of nuclear envelope membrane proteins linking centromeres to microtubules. *Cell* 2008

Koppen A and Kalkhoven E. Brown vs white adipocytes: the PPARgamma coregulator story. *FEBS Lett.* 2010

Korfali N, Fairley EAL, Swanson SK, Florens L and Schirmer EC. Use of sequential chemical extractions to purify nuclear membrane proteins for proteomics identification. *Methods Mol. Biol.* 2009

Korfali N, Srsen V, Waterfall M, Batrakou DG, Pekovic V, Hutchison CJ and Schirmer EC. A Flow Cytometry-Based Screen of Nuclear Envelope Transmembrane Proteins Identifies NET4/Tmem53 as Involved in Stress-Dependent Cell Cycle Withdrawal. *PLoS ONE* 2011

Korfali N, Wilkie GS, Swanson SK, Srsen V, Batrakou DG, Fairley EAL, Malik P, Zuleger N, Goncharevich A, de Las Heras J, Kelly DA, Kerr ARW, Florens L and Schirmer EC. The leukocyte nuclear envelope proteome varies with cell activation and contains novel transmembrane proteins that affect genome architecture. *Mol. Cell Proteomics* 2010

Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG and Singh H. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 2002

Kracklauer MP, Wiora HM, Deery WJ, Chen X, Bolival BJ, Romanowicz D, Simonette RA, Fuller MT, Fischer JA and Beckingham KM. The Drosophila SUN protein Spag4 cooperates with the coiled-coil protein Yuri Gagarin to maintain association of the basal body and spermatid nucleus. *J. Cell. Sci.* 2010

Krogh A, Larsson B, von Heijne G and Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 2001

Kuhn K, Baker SC, Chudin E, Lieu M, Oeser S, Bennett H, Rigault P, Barker D, McDaniel TK and Chee MS. A novel, high-performance random array platform for quantitative gene expression profiling. *Genome Res.* 2004

Kumaran RI and Spector DL. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J. Cell Biol.* 2008

Küpper K, Kölbl A, Biener D, Dittrich S, von Hase J, Thormeyer T, Fiegler H, Carter NP, Speicher MR, Cremer T and Cremer M. Radial chromatin positioning is shaped by local gene density, not by gene expression. *Chromosoma* 2007

Kuroda M, Tanabe H, Yoshida K, Oikawa K, Saito A, Kiyuna T, Mizusawa H and Mukai K. Alteration of chromosome positioning during adipocyte differentiation. *J. Cell. Sci.* 2004

Lachner M, O'Carroll D, Rea S, Mechtler K and Jenuwein T. Methylation of

histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 2001

Lee KK, Haraguchi T, Lee RS, Koujin T, Hiraoka Y and Wilson KL. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J. Cell. Sci.* 2001

Lee YS, Sohn DH, Han D, Lee H, Seong RH and Kim JB. Chromatin remodeling complex interacts with ADD1/SREBP1c to mediate insulin-dependent regulation of gene expression. *Mol. Cell. Biol.* 2007

Li X, Kim JW, Grønborg M, Urlaub H, Lane MD and Tang Q. Role of cdk2 in the sequential phosphorylation/activation of C/EBPbeta during adipocyte differentiation. *Proc. Natl. Acad. Sci. U.S.A.* 2007

Liang Y, Chiu PH, Yip KY and Chan SY. Subcellular Localization of SUN2 Is Regulated by Lamin A and Rab5. *PLoS ONE* 2011

Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ and Tan S. Comparison of affinity tags for protein purification. *Protein Expr. Purif.* 2005

Lin F, Blake DL, Callebaut I, Skerjanc IS, Holmer L, McBurney MW, Paulin-Levasseur M and Worman HJ. MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. *J. Biol. Chem.* 2000

Liu J, Lee KK, Segura-Totten M, Neufeld E, Wilson KL and Gruenbaum Y. MAN1 and emerin have overlapping function(s) essential for chromosome segregation and cell division in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U.S.A.* 2003

Liu Q, Pante N, Misteli T, Elsagga M, Crisp M, Hodzic D, Burke B and Roux KJ. Functional association of Sun1 with nuclear pore complexes. *J. Cell Biol.* 2007

Lloyd DJ, Trembath RC and Shackleton S. A novel interaction between lamin A and SREBP1: implications for partial lipodystrophy and other laminopathies.

Hum. Mol. Genet. 2002

Lohka MJ and Masui Y. Formation in vitro of sperm pronuclei and mitotic chromosomes induced by amphibian ooplasmic components. *Science* 1983

Lu W, Gotzmann J, Sironi L, Jaeger V, Schneider M, Lüke Y, Uhlén M, Szigyarto CA, Brachner A, Ellenberg J, Foisner R, Noegel AA and Karakesisoglou I. Sun1 forms immobile macromolecular assemblies at the nuclear envelope. *Biochim. Biophys. Acta* 2008

MacDougald OA and Lane MD. Transcriptional regulation of gene expression during adipocyte differentiation. *Annu. Rev. Biochem.* 1995

Malik P, Korfali N, Srsen V, Lazou V, Batrakou DG, Zuleger N, Kavanagh DM, Wilkie GS, Goldberg MW and Schirmer EC. Cell-specific and lamin-dependent targeting of novel transmembrane proteins in the nuclear envelope. *Cell. Mol. Life Sci.* 2010

Malone CJ, Fixsen WD, Horvitz HR and Han M. UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development. *Development* 1999

Manilal S, Nguyen TM, Sewry CA and Morris GE. The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. *Hum. Mol. Genet.* 1996

Mansfeld J, Güttinger S, Hawryluk-Gara LA, Panté N, Mall M, Galy V, Haselmann U, Mühlhäusser P, Wozniak RW, Mattaj IW, Kutay U and Antonin W. The conserved transmembrane nucleoporin NDC1 is required for nuclear pore complex assembly in vertebrate cells. *Mol. Cell* 2006

Mansharamani M and Wilson KL. Direct binding of nuclear membrane protein MAN1 to emerin in vitro and two modes of binding to barrier-to-autointegration factor. *J. Biol. Chem.* 2005

Maraldi NM, Capanni C, Cenni V, Fini M and Lattanzi G. Laminopathies and lamin-associated signaling pathways. *J. Cell. Biochem.* 2011

Markiewicz E, Tilgner K, Barker N, van de Wetering M, Clevers H, Dorobek M, Hausmanowa-Petrusewicz I, Ramaekers FCS, Broers JLV, Blankesteijn WM, Salpingidou G, Wilson RG, Ellis JA and Hutchison CJ. The inner nuclear membrane protein emerin regulates beta-catenin activity by restricting its accumulation in the nucleus. *EMBO J.* 2006

Mattout-Drubezki A and Gruenbaum Y. Dynamic interactions of nuclear lamina proteins with chromatin and transcriptional machinery. *Cell. Mol. Life Sci.* 2003

Mayer R, Brero A, von Hase J, Schroeder T, Cremer T and Dietzel S. . Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. BMC Cell Biol. 2005

McGee MD, Rillo R, Anderson AS and Starr DA. UNC-83 IS a KASH protein required for nuclear migration and is recruited to the outer nuclear membrane by a physical interaction with the SUN protein UNC-84. *Mol. Biol. Cell* 2006

McKee JR, Christman CL, O'Brien WDJ and Wang SY. Effects of ultrasound on nucleic acid bases. *Biochemistry* 1977

Meaburn KJ, Cabuy E, Bonne G, Levy N, Morris GE, Novelli G, Kill IR and Bridger JM. Primary laminopathy fibroblasts display altered genome organization and apoptosis. *Aging Cell* 2007

Meaburn KJ, Newbold RF and Bridger JM. Positioning of human chromosomes in murine cell hybrids according to synteny. *Chromosoma* 2008

Medina-Gomez G, Virtue S, Lelliott C, Boiani R, Campbell M, Christodoulides C, Perrin C, Jimenez-Linan M, Blount M, Dixon J, Zahn D, Thresher RR, Aparicio S, Carlton M, Colledge WH, Kettunen MI, Seppänen-Laakso T, Sethi JK, O'Rahilly S, Brindle K, Cinti S, Oresic M, Burcelin R and Vidal-Puig A. The link between nutritional status and insulin sensitivity is dependent on the adipocyte-specific

peroxisome proliferator-activated receptor-gamma2 isoform. *Diabetes* 2005

Mehta IS, Amira M, Harvey AJ and Bridger JM. Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. *Genome Biol.* 2010

Mehta IS, Eskiw CH, Arican HD, Kill IR and Bridger JM. Farnesyltransferase inhibitor treatment restores chromosome territory positions and active chromosome dynamics in Hutchinson-Gilford Progeria syndrome cells. *Genome Biol.* 2011

Meindl U, Zhang D and Hepler PK. Actin microfilaments are associated with the migrating nucleus and the cell cortex in the green alga Micrasterias. Studies on living cells. *J. Cell. Sci.* 1994

Mekhail K and Moazed D. The nuclear envelope in genome organization, expression and stability. *Nat. Rev. Mol. Cell Biol.* 2010

Melchionna R, Di Carlo A, De Mori R, Cappuzzello C, Barberi L, Musarò A, Cencioni C, Fujii N, Tamamura H, Crescenzi M, Capogrossi MC, Napolitano M and Germani A. Induction of myogenic differentiation by SDF-1 via CXCR4 and CXCR7 receptors. *Muscle Nerve* 2010

Meyerzon M, Fridolfsson HN, Ly N, McNally FJ and Starr DA. UNC-83 is a nuclear-specific cargo adaptor for kinesin-1-mediated nuclear migration. *Development* 2009

Mislow JMK, Holaska JM, Kim MS, Lee KK, Segura-Totten M, Wilson KL and McNally EM. Nesprin-1alpha self-associates and binds directly to emerin and lamin A in vitro. *FEBS Lett.* 2002

Mori K, Giovannone B and Smith RJ. Distinct Grb10 domain requirements for effects on glucose uptake and insulin signaling. *Mol. Cell. Endocrinol.* 2005

Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM and Krogan NJ. Functional targeting of DNA

damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* 2008

Nakae J, Kitamura T, Kitamura Y, Biggs WH3, Arden KC and Accili D. The forkhead transcription factor Foxo1 regulates adipocyte differentiation. *Dev. Cell* 2003

Nery FC, Zeng J, Niland BP, Hewett J, Farley J, Irimia D, Li Y, Wiche G, Sonnenberg A and Breakefield XO. TorsinA binds the KASH domain of nesprins and participates in linkage between nuclear envelope and cytoskeleton. *J. Cell. Sci.* 2008

Newport JW and Forbes DJ. The nucleus: structure, function, and dynamics. *Annu. Rev. Biochem.* 1987

Nielsen H, Engelbrecht J, Brunak S and von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 1997

Nili E, Cojocaru GS, Kalma Y, Ginsberg D, Copeland NG, Gilbert DJ, Jenkins NA, Berger R, Shaklai S, Amariglio N, Brok-Simoni F, Simon AJ and Rechavi G. Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). *J. Cell. Sci.* 2001

Otto TC and Lane MD. Adipose development: from stem cell to adipocyte. *Crit. Rev. Biochem. Mol. Biol.* 2005

Oza P and Peterson CL. Opening the DNA repair toolbox: localization of DNA double strand breaks to the nuclear periphery. *Cell Cycle* 2010

Oza P, Jaspersen SL, Miele A, Dekker J and Peterson CL. Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. *Genes Dev.* 2009

Padmakumar VC, Libotte T, Lu W, Zaim H, Abraham S, Noegel AA, Gotzmann J, Foisner R and Karakesisoglou I. The inner nuclear membrane protein

Sun1 mediates the anchorage of Nesprin-2 to the nuclear envelope. *J. Cell. Sci.* 2005

Pan D, Estévez-Salmerón LD, Stroschein SL, Zhu X, He J, Zhou S and Luo K. The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-{beta} superfamily of cytokines. *J. Biol. Chem.* 2005

Park KW, Halperin DS and Tontonoz P. Before they were fat: adipocyte progenitors. *Cell Metab* 2008

Payne VA, Au W, Lowe CE, Rahman SM, Friedman JE, O'Rahilly S and Rochford JJ. C/EBP transcription factors regulate SREBP1c gene expression during adipogenesis. *Biochem. J.* 2010

Peters R. Use of Xenopus laevis oocyte nuclei and nuclear envelopes in nucleocytoplasmic transport studies. *Methods Mol. Biol.* 2006

Pillay CS, Elliott E and Dennison C. Endolysosomal proteolysis and its regulation. *Biochem. J.* 2002

Podlecki DA, Smith RM, Kao M, Tsai P, Huecksteadt T, Brandenburg D, Lasher RS, Jarett L and Olefsky JM. Nuclear translocation of the insulin receptor. A possible mediator of insulin's long term effects. *J. Biol. Chem.* 1987

Polioudaki H, Kourmouli N, Drosou V, Bakou A, Theodoropoulos PA, Singh PB, Giannakouros T and Georgatos SD. Histones H3/H4 form a tight complex with the inner nuclear membrane protein LBR and heterochromatin protein 1. *EMBO Rep.* 2001

Ragoczy T, Bender MA, Telling A, Byron R and Groudine M. The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev.* 2006

Rappsilber J, Siniossoglou S, Hurt EC and Mann M. A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and

mass spectrometry. *Anal. Chem.* 2000

Reddy KL, Zullo JM, Bertolino E and Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 2008

Rogakou EP, Pilch DR, Orr AH, Ivanova VS and Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* 1998

Rolls MM, Stein PA, Taylor SS, Ha E, McKeon F and Rapoport TA. A visual screen of a GFP-fusion library identifies a new type of nuclear envelope membrane protein. *J. Cell Biol.* 1999

Roos A and Boron WF. Intracellular pH. *Physiol. Rev.* 1981

Rosen ED and MacDougald OA. Adipocyte differentiation from the inside out. *Nat. Rev. Mol. Cell Biol.* 2006

Rosen ED, Hsu C, Wang X, Sakai S, Freeman MW, Gonzalez FJ and Spiegelman BM. C/EBPalpha induces adipogenesis through PPARgamma: a unified pathway. *Genes Dev.* 2002

Ross SE, Hemati N, Longo KA, Bennett CN, Lucas PC, Erickson RL and MacDougald OA. Inhibition of adipogenesis by Wnt signaling. *Science* 2000

Roux KJ, Crisp ML, Liu Q, Kim D, Kozlov S, Stewart CL and Burke B. Nesprin 4 is an outer nuclear membrane protein that can induce kinesin-mediated cell polarization. *Proc. Natl. Acad. Sci. U.S.A.* 2009

RTeam. R: A Language and Environment for Statistical Computing. 2010

Ruault M, Dubarry M and Taddei A. Re-positioning genes to the nuclear envelope in mammalian cells: impact on transcription. *Trends Genet.* 2008

Rubin CS, Hirsch A, Fung C and Rosen OM. Development of hormone receptors and hormonal responsiveness in vitro. Insulin receptors and insulin sensitivity in the preadipocyte and adipocyte forms of 3T3-L1 cells. *J. Biol. Chem.*

1978

Sadoni N, Targosz B, Englmann A, Fesser S, Koch J, Schindelhauer D and Zink D. Transcription-dependent spatial arrangements of CFTR and conserved adjacent loci are not conserved in human and murine nuclei. *Chromosoma* 2008

Salma N, Xiao H, Mueller E and Imbalzano AN. Temporal recruitment of transcription factors and SWI/SNF chromatin-remodeling enzymes during adipogenic induction of the peroxisome proliferator-activated receptor gamma nuclear hormone receptor. *Mol. Cell. Biol.* 2004

Salpingidou G, Rzepecki R, Kiseleva E, Lyon C, Lane B, Fusiek K, Golebiewska A, Drummond S, Allen T, Ellis JA, Smythe C, Goldberg MW and Hutchison CJ. NEP-A and NEP-B both contribute to nuclear pore formation in Xenopus eggs and oocytes. *J. Cell. Sci.* 2008

Samstag Y, Eibert SM, Klemke M and Wabnitz GH. Actin cytoskeletal dynamics in T lymphocyte activation and migration. *J. Leukoc. Biol.* 2003

Schirmer EC, Florens L, Guan T, Yates JR3 and Gerace L. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* 2003

Schmitt J, Benavente R, Hodzic D, Höög C, Stewart CL and Alsheimer M. Transmembrane protein Sun2 is involved in tethering mammalian meiotic telomeres to the nuclear envelope. *Proc. Natl. Acad. Sci. U.S.A.* 2007

Schöfer C and Weipoltshammer K. Gene dynamics and nuclear architecture during differentiation. *Differentiation* 2008

Seksek O and Bolard J. Nuclear pH gradient in mammalian cells revealed by laser microspectrofluorimetry. *J. Cell. Sci.* 1996

Sengupta K, Camps J, Mathews P, Barenboim-Stapleton L, Nguyen QT, Difilippantonio MJ and Ried T. Position of human chromosomes is conserved in mouse nuclei indicating a species-independent mechanism for maintaining genome

organization. *Chromosoma* 2008

Senior A and Gerace L. Integral membrane proteins specific to the inner nuclear membrane and associated with the nuclear lamina. *J. Cell Biol.* 1988

Shackleton S, Lloyd DJ, Jackson SN, Evans R, Niermeijer MF, Singh BM, Schmidt H, Brabant G, Kumar S, Durrington PN, Gregory S, O'Rahilly S and Trembath RC. LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. *Nat. Genet.* 2000

Shao X, Tarnasky HA, Lee JP, Oko R and van der Hoorn FA. Spag4, a novel sperm protein, binds outer dense-fiber protein Odf1 and localizes to microtubules of manchette and axoneme. *Dev. Biol.* 1999

Siersbæk R, Nielsen R, John S, Sung M, Baek S, Loft A, Hager GL and Mandrup S. Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. *EMBO J.* 2011

Smas CM and Sul HS. Control of adipocyte differentiation. *Biochem. J.* 1995

Smith PJ, Wise LS, Berkowitz R, Wan C and Rubin CS. Insulin-like growth factor-I is an essential regulator of the differentiation of 3T3-L1 adipocytes. *J. Biol. Chem.* 1988

Solovei I, Cavallo A, Schermelleh L, Jaunin F, Scasselati C, Cmarko D, Cremer C, Fakan S and Cremer T. Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). *Exp. Cell Res.* 2002

Somech R, Shaklai S, Geller O, Amariglio N, Simon AJ, Rechavi G and Gal-Yam EN. The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. *J. Cell. Sci.* 2005

Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T, Danuser

G and Misteli T. Positional stability of single double-strand breaks in mammalian cells. *Nat. Cell Biol.* 2007

Stadler S, Schnapp V, Mayer R, Stein S, Cremer C, Bonifer C, Cremer T and Dietzel S. The architecture of chicken chromosome territories changes during differentiation. *BMC Cell Biol.* 2004

Starr DA and Fridolfsson HN. Interactions between nuclei and the cytoskeleton are mediated by SUN-KASH nuclear-envelope bridges. *Annu. Rev. Cell Dev. Biol.* 2010

Starr DA and Han M. Role of ANC-1 in tethering nuclei to the actin cytoskeleton. *Science* 2002

Starr DA, Hermann GJ, Malone CJ, Fixsen W, Priess JR, Horvitz HR and Han M. unc-83 encodes a novel component of the nuclear envelope and is essential for proper nuclear migration. *Development* 2001

Stewart CL, Roux KJ and Burke B. Blurring the boundary: the nuclear envelope extends its reach. *Science* 2007

Student AK, Hsu RY and Lane MD. Induction of fatty acid synthetase synthesis in differentiating 3T3-L1 preadipocytes. *J. Biol. Chem.* 1980

Suh JM, Gao X, McKay J, McKay R, Salo Z and Graff JM. Hedgehog signaling plays a conserved role in inhibiting fat formation. *Cell Metab* 2006

Sullivan T, Escalante-Alcalde D, Bhatt H, Anver M, Bhat N, Nagashima K, Stewart CL and Burke B. Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J. Cell Biol.* 1999

Sulston JE and Horvitz HR. Abnormal cell lineages in mutants of the nematode Caenorhabditis elegans. *Dev. Biol.* 1981

Szczerbal I and Bridger JM. Association of adipogenic genes with SC-35 domains during porcine adipogenesis. *Chromosome Res.* 2010

Szczerbal I, Foster HA and Bridger JM. The spatial repositioning of adipogenesis genes is correlated with their expression status in a porcine mesenchymal stem cell adipogenesis model system. *Chromosoma* 2009

Tagwerker C, Flick K, Cui M, Guerrero C, Dou Y, Auer B, Baldi P, Huang L and Kaiser P. A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. *Mol. Cell Proteomics* 2006

Tang Q, Otto TC and Lane MD. Mitotic clonal expansion: a synchronous process required for adipogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 2003

Tapley EC, Ly N and Starr DA. Multiple mechanisms actively target the SUN protein UNC-84 to the inner nuclear membrane. *Mol. Biol. Cell* 2011

Temple KA, Basko X, Allison MB and Brady MJ. Uncoupling of 3T3-L1 gene expression from lipid accumulation during adipogenesis. *FEBS Lett.* 2007

Toczyski DP, Galgoczy DJ and Hartwell LH. CDC5 and CKII control adaptation to the yeast DNA damage checkpoint. *Cell* 1997

Tontonoz P, Hu E and Spiegelman BM. Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. *Cell* 1994

Tsuchiya Y. Till disassembly do us part: a happy marriage of nuclear envelope and chromatin. *J Biochem* 2008

Tsurutani Y, Fujimoto M, Takemoto M, Irisuna H, Koshizaka M, Onishi S, Ishikawa T, Mezawa M, He P, Honjo S, Maezawa Y, Saito Y and Yokote K. The roles of transforming growth factor-β and Smad3 signaling in adipocyte differentiation and obesity. *Biochem. Biophys. Res. Commun.* 2011

Ugalde AP, Ramsay AJ, de la Rosa J, Varela I, Mariño G, Cadiñanos J, Lu J, Freije JM and López-Otín C. Aging and chronic DNA damage response activate a regulatory pathway involving miR-29 and p53. *EMBO J.* 2011

Van Berlo JH, Voncken JW, Kubben N, Broers JLV, Duisters R, van Leeuwen REW, Crijns HJGM, Ramaekers FCS, Hutchison CJ and Pinto YM. A-type lamins are essential for TGF-beta1 induced PP2A to dephosphorylate transcription factors. *Hum. Mol. Genet.* 2005

Vázquez-Vela MEF, Torres N and Tovar AR. White adipose tissue as endocrine organ and its role in obesity. *Arch. Med. Res.* 2008

Vigers GP and Lohka MJ. A distinct vesicle population targets membranes and pore complexes to the nuclear envelope in Xenopus eggs. *J. Cell Biol.* 1991

Wagner N and Krohne G. LEM-Domain proteins: new insights into lamin-interacting proteins. *Int. Rev. Cytol.* 2007

Wang Q, Du X, Cai Z and Greene MI. Characterization of the structures involved in localization of the SUN proteins to the nuclear envelope and the centrosome. *DNA Cell Biol.* 2006

Wang W, Huang L, Huang Y, Yin J, Berk AJ, Friedman JM and Wang G. Mediator MED23 links insulin signaling to the adipogenesis transcription cascade. *Dev. Cell* 2009

Wang X, Sato R, Brown MS, Hua X and Goldstein JL. SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis. *Cell* 1994

Waterhouse AM, Procter JB, Martin DMA, Clamp M and Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009

Wettenhall JM and Smyth GK. limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics* 2004

Wilder EL and Linzer DI. Participation of multiple factors, including proliferin, in the inhibition of myogenic differentiation. *Mol. Cell. Biol.* 1989

Wilhelmsen K, Litjens SHM, Kuikman I, Tshimbalanga N, Janssen H, van den

Bout I, Raymond K and Sonnenberg A. Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. *J. Cell Biol.* 2005

Wilkie GS, Korfali N, Swanson SK, Malik P, Srsen V, Batrakou DG, de Las Heras J, Zuleger N, Kerr ARW, Florens L and Schirmer EC. Several novel nuclear envelope transmembrane proteins identified in skeletal muscle have cytoskeletal associations. *Mol. Cell Proteomics* 2010

Wilmut I, Schnieke AE, McWhir J, Kind AJ and Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature* 1997

Worman HJ and Bonne G. "Laminopathies": a wide spectrum of human diseases. *Exp. Cell Res.* 2007

Worman HJ, Yuan J, Blobel G and Georgatos SD. A lamin B receptor in the nuclear envelope. *Proc. Natl. Acad. Sci. U.S.A.* 1988

Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge CL, Haase J, Janes J, Huss JW3 and Su AI. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* 2009

Yadav H, Quijano C, Kamaraju AK, Gavrilova O, Malek R, Chen W, Zerfas P, Zhigang D, Wright EC, Stuelten C, Sun P, Lonning S, Skarulis M, Sumner AE, Finkel T and Rane SG. Protection from obesity and diabetes by blockade of TGF-β/Smad3 signaling. *Cell Metab* 2011

Ye Q and Worman HJ. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. *J Biol Chem* 1996

Zaremba-Czogalla M, Dubińska-Magiera M and Rzepecki R. Laminopathies: the molecular background of the disease and the prospects for its treatment. *Cell. Mol. Biol. Lett.* 2011

Zhang A, Gonzalez SM, Cantor EJ and Chong S. Construction of a mini-intein

fusion system to allow both direct monitoring of soluble protein expression and rapid purification of target proteins. *Gene* 2001

Zhang J, Fu M, Cui T, Xiong C, Xu K, Zhong W, Xiao Y, Floyd D, Liang J, Li E, Song Q and Chen YE. Selective disruption of PPARgamma 2 impairs the development of adipose tissue and insulin sensitivity. *Proc. Natl. Acad. Sci. U.S.A.* 2004

Zhang Q, Ragnauth CD, Skepper JN, Worth NF, Warren DT, Roberts RG, Weissberg PL, Ellis JA and Shanahan CM. Nesprin-2 is a multi-isomeric protein that binds lamin and emerin at the nuclear envelope and forms a subcellular network in skeletal muscle. *J. Cell. Sci.* 2005

Zhang Q, Skepper JN, Yang F, Davies JD, Hegyi L, Roberts RG, Weissberg PL, Ellis JA and Shanahan CM. Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. *J. Cell. Sci.* 2001

Zhen Y, Libotte T, Munck M, Noegel AA and Korenbaum E. NUANCE, a giant protein connecting the nucleus and actin cytoskeleton. *J. Cell. Sci.* 2002

Zink D, Amaral MD, Englmann A, Lang S, Clarke LA, Rudolph C, Alt F, Luther K, Braz C, Sadoni N, Rosenecker J and Schindelhauer D. Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *J. Cell Biol.* 2004

Appendix

JOURNAL OF PROTEOMICS 72 (2009) 56–70

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Review

Comparative proteomic analyses of the nuclear envelope and pore complex suggests a wide range of heretofore unexpected functions

*Dzmitry G. Batrakou, Alastair R.W. Kerr, Eric C. Schirmer**

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

ARTICLE INFO

Keywords:
Nuclear envelope
Nuclear pore complex (NPC)
Tissue variation
Phenylalanine–glycine (FG) repeats
Nuclear transport
Comparative proteomics

ABSTRACT

Since the discovery of several inherited diseases linked to the nuclear envelope the number of functions ascribed to this subcellular organelle has skyrocketed. However the molecular pathways underlying these functions are not clear in most cases, perhaps because of missing components. Several recent proteomic analyses of the nuclear envelope and nuclear pore complex proteomes have yielded not only enough missing components to potentially elucidate these pathways, but suggest an exponentially greater number of functions at the nuclear periphery than ever imagined. Many of these functions appear to derive from recapitulation of pathways utilized at the plasma membrane and from other membrane systems. Additionally, many proteins identified in the comparative nuclear envelope studies have sequence characteristics suggesting that they might also contribute to nuclear pore complex functions. In particular, the striking enrichment for proteins in the nuclear envelope fractions that carry phenylalanine–glycine (FG) repeats may be significant for the mechanism of nuclear transport. In retrospect, these findings are only surprising in context of the notion held for many years that the nuclear envelope was only a barrier protecting the genome. In fact, it is arguably the most complex membrane organelle in the cell.

Contents

* *Corresponding author*. Wellcome Trust Centre for Cell Biology, University of Edinburgh, Kings Building Campus, Swann 5.22, Mayfield Road, Edinburgh EH9 3JR, UK. Tel.: +44 131 650 7075; fax: +44 131 650 7360.
E-mail address: e.schirmer@ed.ac.uk (E.C. Schirmer).

1874-3919/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jprot.2008.09.004

1. Introduction

The structure of the nuclear envelope (NE) is complex (for reviews see [1–4]). It is a double membrane system continuous with the endoplasmic reticulum (ER) that consists of three connected but distinguishable membrane domains: the outer, inner and pore membranes (Fig. 1A). The outer nuclear membrane (ONM) is studded with ribosomes [5,6] and contains many ER proteins in addition to having a set of unique proteins, some of which appear to be involved in tethering the nucleus to cytoplasmic filament systems [7–12]. Some of these proteins in turn connect across the lumen of the nuclear envelope to the inner nuclear membrane (INM), which contains its own unique set of proteins (Fig. 1B; reviewed in [3,13]). Many of these have been shown to bind both to the intermediate filament lamin polymer [14–19] and chromatin (reviewed in [20]). Nuclear pore complexes (NPCs) are inserted at the pore membrane (PoM) that connects the ONM and INM (Fig. 1A; reviewed in [4]). The NPCs regulate directional transport of proteins and mRNA between the nucleus and cytoplasm that exceed the measured maximum free diffusion limit of ~40–60 kDa [21] and can accommodate very large molecules or complexes that have been experimentally tested up to 39 nm diameter [22] (to put this in perspective the longest dimension of an assembled ribosome is ~25 nm). To accommodate such large substrates, the NPCs are necessarily large, on the order of 44 to 125 MDa. These large complexes are tethered to the membrane by at least three unique transmembrane proteins [23–28] that also contribute to a ring component of the NPC in the NE lumen [29–31]. The lumen of the nuclear envelope is largely unexplored territory, but is likely to have its own unique functions.

NE proteins have now been shown to influence a wide range of functions, although it is unclear whether their effects are direct or indirect. These functions include nuclear morphology and stability [32–36], nuclear anchoring/migration within the cell [8–10], signaling cascades [37–40], and support of DNA replication [41–43], transcription [44–46], and RNA splicing [47]. Consistent with the notion that the nuclear lamina carries out or facilitates a diverse range of basic cellular functions, over a dozen inherited diseases and syndromes are linked to lamins and certain associated NE proteins. NPC proteins have also been linked to disease. These nuclear envelope diseases include muscular dystrophies, lipodystrophies, neuropathy, cardiomyopathies, dermopathy, bone disorders, and premature aging syndromes (reviewed in [48–51]). The proteins thus far linked to disease are lamins A/C, B1, and B2, emerin, LBR, LAP2, MAN1, Syne/Nesprin-1, FACE-1/ZMPSTE24, torsin A, and the NPC proteins Aladin, and Nup62. The favored hypotheses to explain how these proteins can yield so many different diseases are: 1) reduced resistance to mechanical stress, 2) disruption of gene regulation, 3) alterations in cell cycle and signaling pathways. However, none of these hypotheses can fully explain by themselves how mutations in the same widely expressed proteins can yield different diseases that each have distinctive tissue pathologies. Correspondingly, it is hard to imagine the many functions ascribed to lamins and associated proteins being due to diverse enzymatic activities encoded within the same proteins. The resolution in both cases likely involves additional partner proteins that provide these functions and have yet to be identified, hence the need for proteomic analysis.

The combination of its inclusion of cytoskeletal elements, the lamin polymer, integral membrane proteins, NPCs, a membrane and its lumenal content that is continuous with the ER together with the many associations with chromatin proteins, transport receptors and cargos, and indirectly attached peripheral components of the cytoskeleton give the NE a wide range of biochemical properties. This has the consequence that any biochemical fractionation methodology will necessarily remove some true NE components and bring along some contaminants. As the ONM is continuous with the ER and connected to cytoplasmic filament systems [9,52] these structures are difficult to separate. INM proteins bind chromatin (reviewed in [20]) and in yeast the NPCs are connected to nucleoplasmic filaments [53,54] that in turn connect to telomeres [55,56]. All of these connections on both cytoplasmic and nucleoplasmic faces of the NE and NPC further compound difficulties in their purification.

2. The nuclear pore proteome

2.1. *Pre-proteomics studies*

As one of the largest macromolecular complexes in biology it is not surprising that most of the original characterization of the NPC was through the electron microscope. This work determined that the Xenopus NPC had a diameter of ~120 nm with eight-fold radial symmetry perpendicular to the membrane and a predicted mass of roughly 125 MDa [29,31,57] while the yeast NPC was somewhat smaller at 55–72 MDa [58,59]. Thus the NPC could be made up of a very large number of distinct proteins. Determining its composition was therefore not trivial

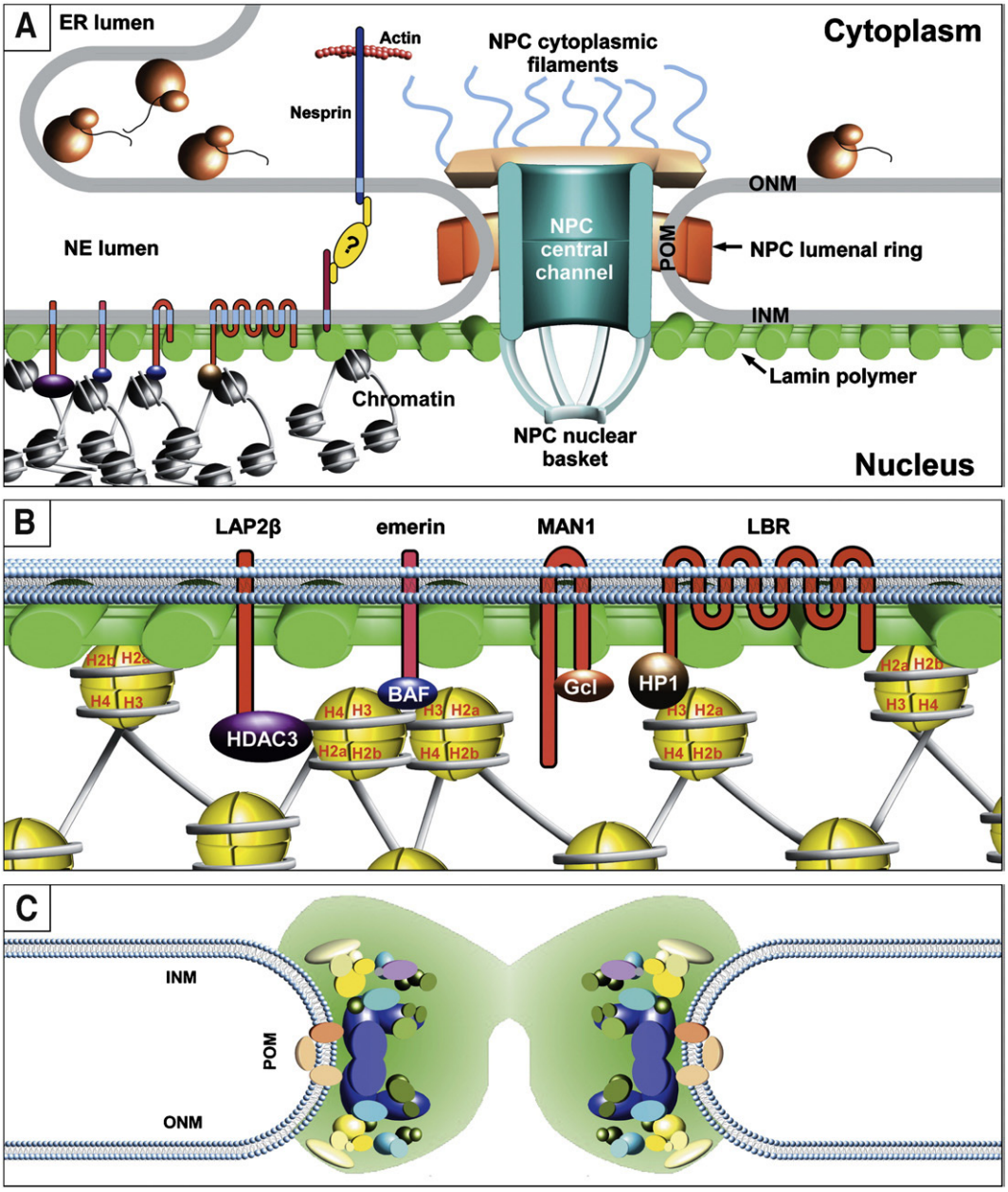


Fig. 1 – Schematic diagram of the NE with details of the NPC. (A) The nuclear envelope consists of outer and inner nuclear membranes connected at the "pore membrane". The outer nuclear membrane (ONM) is continuous with the rough and smooth endoplasmic reticulum (ER). The inner nuclear membrane (INM) contains many unique integral proteins, which commonly are associated with the intermediate filament lamin polymer. The pore membrane apposed to the nuclear pore complexes (NPCs) contains specific integral proteins involved in membrane tethering of NPCs. Depending on their topology and membrane subdomain, NE transmembrane proteins could have functions in the cytoplasm, nucleoplasm, or the perinuclear lumenal space. Some ONM-specific proteins have been identified such as the Syne/Nesprin protein families, but most proteins identified in this region such as ribosomal proteins share functions with the ER. Only a few characterized proteins have both been shown to contain most of their mass in the lumen, though more are likely based on proteomic results. (B) Many transmembrane proteins of the INM directly interact with the lamin polymer and/or chromatin proteins, though only a small number of the first identified proteins have been tested for such characteristics. Among these, LAP2ß interacts with the chromatin remodeling protein histone deacetylase 3 (HDAC3). LAP2ß, emerin, and MAN1 all interact with the chromatin crosslinking protein BAF and transcriptional repressors such as germ cell-less (Gcl). LBR binds to both the heterochromatin protein HP1 and histone H3, and even lamins can directly bind to the core histones H2A and H2B. (C) The NPC proteomic studies identified many proteins that recently have been mapped into a high resolution structure (compared to electron microscope images; [30]). A depiction of this structure shows the rough positions of many nucleoporins and a density gradient of unstructured phenylalanine–glycine (FG) repeat containing regions (represented as a haze around the NPC core components) that are thought to play a significant role in the transport process.

even in the eras of genomics and proteomics. As the average mammalian nucleus has 2–3000 NPCs [60] identification of NPC component proteins by biochemical fractionation was initially attempted and was successful in some cases. For example the first individual NPC component identified was gp210 that was isolated from a rat liver NE fraction and used to make antibodies that labeled the NPCs by immunogold electron microscopy ([61]; note this was originally called gp190 and renamed after the gene was cloned 7 years later [28]). Many other NPC proteins were soon identified using similar approaches and the proteins were called nucleoporins or NUPs [62–64]. Many of these antibodies cross-reacted among mammals, Xenopus and yeast [65–67] thus facilitating cloning of NUPs. Xenopus oocytes were the best system for biochemical purification while yeast was the most genetically tractable, so that many of the first NUP sequences were from yeast (e.g. NUP1 [66], NSP1 [68]). Once the first NUPs were identified it was discovered that they tend to form subcomplexes with 3–5 proteins and this rapidly facilitated further NUP identifications. For example p62 was among the first NUPs cloned [69] due to its abundance and strong antigenicity and it was subsequently found to be part of a complex with p54 and p58 in mammalian cells [70,71]. In yeast Nup170p was found to be in a complex with Nup53p and Nup59p [72] and Nic96p was identified through its interaction with Nsp1p [73]. Many individual studies over nearly two decades had identified 26 core NPC components in yeast (Table 1) by the time that the first major proteomic studies were done on the NPC. However, absent knowing how many copies of each protein were in each NPC it was impossible to gauge how many more components still needed to be identified and it was generally thought that between 50 and 100 proteins would be required to account for the estimated 55 to 125 MDa mass predicted by electron microscopy.

Table 1 – NPC core proteins [74,77] with number of phenyalanine-glycine (FG) repeats

Yeast	FG	Year (PubMed ID)	Mammalian	FG	Year (PubMed ID)
Nup1	17	**1990** (2190694)	Nup153	30	**1993** (8422679)
Nsp1	33	**1990** (2112428)	Nup62	6	**1990** (2295087)
Nic96	1	**1993** (7688296)	Nup93	3	**1997** (9348540)
Nup145N	14	**1994** (8044840)	Nup98	40	**1995** (7878057)
Nup145C	1	**1994** (8044840)	Nup96	1	**1999** (10087256)
Nup133	0	**1994** (7813444)	Nup133	2	**2001** (11564755)
Pom152	1	**1994** (8138573)	–	–	–
Nup42	29	**1995** (7634338)	NLP1/hCG1 (45)	14	**1999** (10358091)
Nup49	18	**1995** (1385442)	Nup58	11	**1991** (2050741)
Nup57	16	**1995** (7828598)	Nup54	8	**1991** (2050741)
Nup82	2	**1995** (7559750)	Nup88	2	**1997** (9049309)
Nup100	45	**1995** (1385442)	Nup98	40	**1995** (7878057)
Nup116	47	**1995** (1385442)	Nup98	40	**1995** (7878057)
Nup120	2	**1995** (8557736)	Nup160	3	**2001** (11684705)
Nup157	2	**1995** (8522578)	Nup155	0	**1993** (8458861)
Nup170	0	**1995** (8522578)	Nup155	0	**1993** (8458861)
Nup159	28	**1996** (8898365)	Nup214/CAN	45	**1994** (8108440)
Nup188	5	**1996** (8682854)	Nup188	3	**2000** (11029043)
Gle1(62)	1	**1996** (8848052)	hGle1(85)	0	**1998** (9618489)
Gle2(41)	0	**1996** (8970155)	Rae1/Gle2b (41)	3	**1997** (9370289)
Nup85	1	**1996** (8816998)	Nup75/Nup85	2	**1996** (8816998)
Nup84	0	**1997** (9166401)	Nup107	3	**1994** (8021268)
Ndc1 (74)	3	**1998** (9864355)	–	–	–
Nup53	4	**1998** (9864357)	Nup35	3	2002 (12196509)
Nup59	6	**1998** (9864357)	Nup35	3	2002 (12196509)
Nup192	4	**1999** (10428845)	Nup205	4	**2000** (11029043)
Seh1 (39)	0	2000 (10684247)	Sec13-like	0	2002 (12196509)
Nup60	0	2000 (10684247)	–	–	–
CDC31	0	2000 (10684247)	–	–	–
Pom34	1	2000 (10684247)	–	–	–
–	–	–	Pom121	24	**1993** (8335683)
–	–	–	Gp210	5	**1990** (2184032)
–	–	–	Nup358/ RanBP2	20	**1995** (7775481)
–	–	–	ALADIN (60)	1	2004 (15666842)
–	–	–	Nup37	0	2002 (12196509)
–	–	–	Nup43	1	2002 (12196509)
–	–	–	Tpr (266)	0	**1994** (7798308)
–	–	–	Nup50	5	**1997** (9073512)

2.2. *The yeast NPC proteome*

The first comprehensive determination of NPC composition was the product of collaboration between the Rout, Aitchison and Chait laboratories [74]. Critical to a successful proteomic study is the choice of experimental system and the purity of the fractions analyzed. This study used yeast (*Saccharomyces cerevisiae*) because it had the highest NPC:nuclear volume ratio of any organism tested [75]. A haploid yeast nucleus has 65–182 NPCs depending on the cell cycle stage [76]. Unlike mammalian cells yeast do not have a lamina, which has the advantage of less contamination from connections to such a promiscuous structure and the disadvantage of losing the added stability to NEs from the lamina during purification. Yeast cells were spheroplasted, lysed and nuclei were isolated on sucrose gradients. The chromatin was enzymatically digested and extracted with heparin to isolate NEs. NPCs were released from NEs with mild detergent followed by isolation through their partitioning on a continuous sucrose velocity gradient [58]. Three separate rounds of separation by HPLC and SDS-PAGE yeilded 465 protein bands that were digested and analyzed by matix-associated laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry and an additional 177 protein bands were analyzed by MALDI-ion trap tandem mass spectrometry. A total of 174 protein identifications were made [74]. Among these were all previously characterized NUPs, many proteins with known functions at the NPC such as transport factors and chaperones, many other proteins with known functions elsewhere that were considered to be contaminants, and 34 previously uncharacterized ORFs.

All previously uncharacterized ORFs and several known NUPs were genomically tagged with a protein A epitope and tested in yeast for their localization at the NE by immunofluorescence and specifically at the NPC by immuno-electron microscopy. Those that fulfilled these criteria and also were significantly enriched in a NE fraction by Western blot were classified as core NPC components, a total of 30 proteins (two of which are products of the same gene). Thus the yeast proteomic analysis identified only 4 additional core NPC components on top of those previously identified (Table 1). An additional ten NPC-associated proteins that did not fulfill their stringent criteria for inclusion as core components were

classified among those identified, leaving 134 of the 174 proteins identified classified as contaminants by the authors.

2.3. *The mammalian NPC proteome*

Subsequently the Matunis laboratory determined the composition of the mammalian NPC [77]. They used rat liver because previous work had developed procedures for isolating extremely clean fractions of nuclear envelopes from this material [61,78–80]. These procedures take advantage of the relative softness of liver as a tissue to lyse cells without breaking their nuclei and also benefit from the stability conferred to the NE by the lamin polymer to allow several steps of douncing, floating contaminating membranes on sucrose cushions, chromatin digestion and salt washes without fragmenting NEs. For fear that some NPC components might be also removed by high salt treatment, they however replaced the salt with heparin and spermidine so that extraction would favor removal of just histones. They next took advantage of the biochemical characteristics of the lamin polymer and its binding partners that are highly resistent to extraction by high salt and detergent treatments and tested a variety of mild detergents for those that would solubilize NPC components while leaving lamina components insoluble. They found that all well-characterized NPC components could be extracted while leaving most of the lamina components in the pellet using a hypotonic solution containing a low concentration (0.3%) of the detergent Empigen BB. Similarly to the yeast study, this material was separated by HPLC prior to SDS-PAGE and analyzed using a combination of single step and tandem mass spectrometry. Previously uncharacterized proteins were tagged with GFP and tested for NPC colocalization using an antibody (mAb414) that recognizes O-linked glycosylation on several nucleoporins [62]. This study identified 23 proteins that were classified as core NUPs and 18 proteins classified as NPC-associated proteins (Table 1).

2.4. *The revised NPC*

The big surprise from both the yeast and mammalian studies was the relatively small number of proteins identified. Original estimates of protein content based on electron microscopy studies had predicted between 50 and 100 distinct protein components for the NPC [29,31,57]. The total number of only ~30 can in part be explained by the observations that several components have well above average molecular weights and that many components are duplicated many times per NPC. Based on the eight-fold symmetry of the NPC core NUPs would be expected to be duplicated eight times, but the abundance of many components suggested that they are represented 16 and 32 times within a single NPC [74,77]. Furthermore, the yeast study excluded the Mlp (Tpr in mammals) proteins that make up much of a "nuclear basket" extending into the nucleus from the NPC that was observed in electron microscopy studies. Nonetheless, even taking into account these considerations, the total mass of the core NPC calculated from summing up the components identified in the proteomic analyses (44 MDa in yeast [74] and ~60 MDa in mammals [77]) was surprisingly lower than the mass predicted from electron microscopy studies (55–72 MDa in yeast [58,59] and 125 MDa in mammals [29,57].

A potential explanation for this discrepancy may lie in the additional proteins that both studies identified, but excluded because they did not fulfill their very conservative definition for core NPC components. There are a great many NPC proteins that have only transient associations during the transport process. For example the Ran GTPase is a very abundant protein that is involved in release of transport substrates from NPC proteins in the nucleus (Fig. 2A). Importins and exportins are transport receptors that bind to the transport cargo and interact with "core" NPC components to facilitate transport of the cargo through the central channel of the NPC (Fig. 2B and C). These interactions presumably occur through repeat motifs containing phenylalanine and glycine (FG repeats) that appear on both NUPs and many transport receptors. Because their associations are transient neither Ran proteins nor transport receptors were considered as core NPC components, yet due to their high abundance these proteins were also identified in these studies and could moreover account for a significant portion of the mass difference between the calculations from the "core" NPC components and those measured from electron microscopy studies. Even with these, however, the mass difference of almost 2-fold for the mammalian NPC would likely indicate the existence of additional proteins not yet identified (either core components or transient NPC proteins) that were extracted during the preparation of the core NPC fraction. Another possibility is that some of the proteins that were discounted because of other previously characterized cellular functions and localizations actually have separate functions at the NPC. For example the Sec13 proteins are known to function in vesicle formation in the ER [81], yet also were shown to function in nucleo-cytoplasmic transport [82]. Thus, additional NPC components might be identified if these studies were repeated using multiple and varied purification methodologies and if proteins identified with distinct known cellular functions were also tested for NPC function.

It is also worthy of note that the original 125 MDa mass estimate for the NPC of higher eukaryotes was actually principally determined from Xenopus, which was much more accessible to analysis than mammalian nuclei. Thus it is possible that the Xenopus NPC, not yet analyzed by proteomic approaches, is larger than either yeast or mammalian NPCs. Comparison of the yeast and mammalian NPCs would suggest that if this is the case the mass difference likely is accounted for mostly in more peripherally associated proteins because the differences between the yeast and mammalian proteins identified as core structural NPC components is smaller than it appears (Table 1). While several proteins did not have homologues, they nevertheless had analogues: although no clear sequence homology links yeast Pom152 to the mammalian Pom121, both are transmembrane proteins with similar functions. Likewise yeast Ndc1 had no clear mammalian homologue identifiable by genome mining and was not found in the mammalian NPC study, but the mammalian Ndc1 analogue was subsequently identified in a NE proteomic study ([26,83]; see below) giving direct support to the notion that differences in purification methodologies will also contribute to differences in identifications. The mammalian Nup50 appears to be an analogue of Nup2p in yeast, which was not considered in the yeast proteomic study to be part of the core

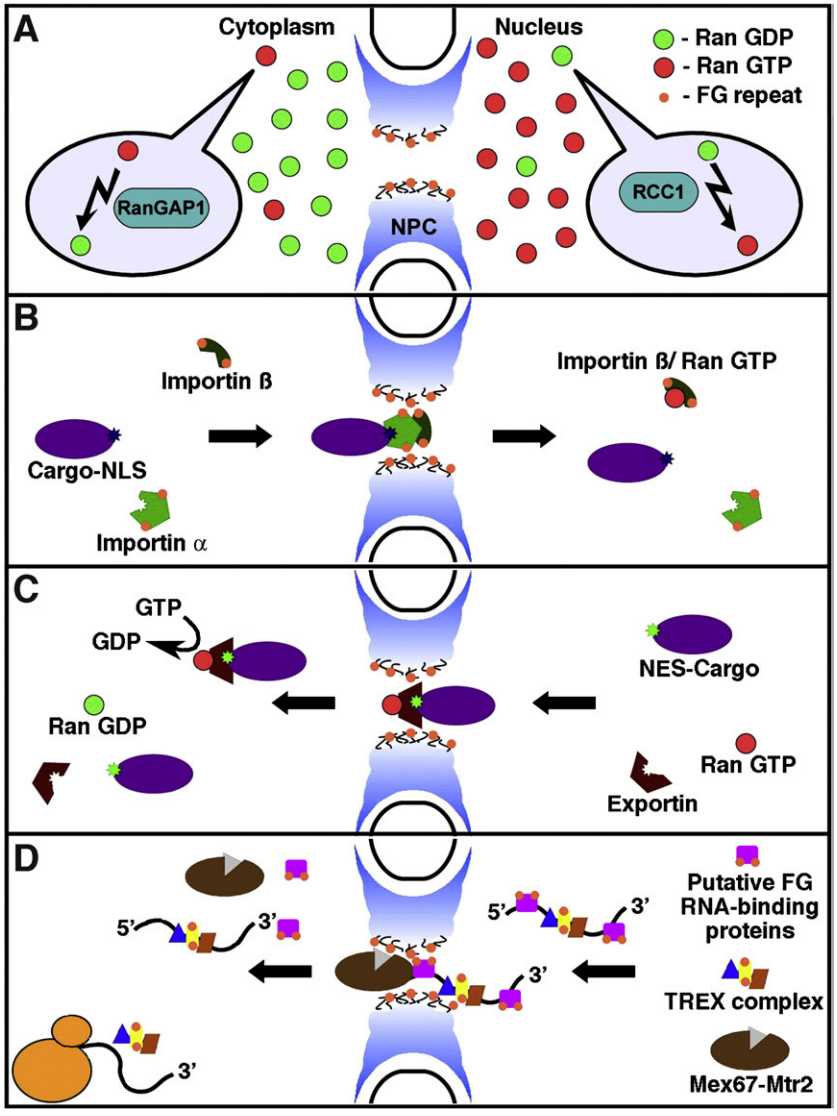


Fig. 2 – Transport through the NPC. (A) Transport through the central channel of the NPC is facilitated by the action of the small GTPase Ran. RanGTP is concentrated in the nucleus and RanGDP in the cytoplasm. RanGAPs (Ran GTPase Activating Proteins) facilitate conversion of the GTP to the GDP form in the cytoplasm while RanGEFs (Guanine nucleotide Exchange Factors) such as RCC1 mediate replenishment of RanGTP in the nucleus. Ran is thus a very abundant protein that will be associated with NPC proteins even though it is not part of the core NPC structure. (B) In nuclear import a transport cargo that has a nuclear localization signal (NLS) (depicted as "*") is recognized by the transport receptor importin α. Another transport receptor, importin ß, then recognizes the importin α-receptor cargo and facilitates its transport through the central channel of the NPC. This is thought to be mediated by interactions between the FG repeats on core NPC proteins and FG repeats on the transport receptors that now coat the cargo. Once the receptor-cargo complex has passed through the NPC RanGTP in the nucleus binds to importin ß to facilitate dissociation of the receptor-cargo complex and thus release of the free cargo into the nucleoplasm. (C) In nuclear export a different kind of transport receptor called an exportin recognizes a nuclear export signal (NES) (depicted as "*") on the cargo and binds together with RanGTP. The complex then transports through the central channel of the NPC and GTP hydrolysis initiated by the RanGAPs releases both the Ran and the exportin. (D) For RNA export several transport receptors have been identified such as the Mex67-Mtr2 heterodimer. Several additional proteins, some of which contain FGs, have also been found to be important that form a complex called the TREX complex in yeast. We propose that the many RNA-binding proteins that harbor FGs identified in NE proteomic datasets might be part of an even greater complex of RNA-binding proteins that coat the surface of the RNAs and expose FGs to interact with the FGs on the core NPC proteins and facilitate transport. The proteins shown in Table 2 are good candidates for this task.

because its association is too dynamic. Although we cannot tell if completely novel or as yet unidentified core components are part of the Xenopus NPC until it has been also analyzed by proteomics, nonetheless, blast searching yeast and mammalian nucleoporins against the *Xenopus laevis* and *Xenopus tropicalis* genomes indicates only Nup133, PoM152, Gle1,

Nup60, and PoM34 among the proteins listed in Table 1 do not have homologs in frogs.

Phylogenetic analysis of the sequences of these core NPC proteins brought another striking observation. The yeast NPC proteins could be separated into pairs that clearly diverged after a gene duplication event [84]. This is consistent with predictions based on the yeast genome sequence of a general genome duplication that was followed by asynchronous differentiation of the duplicated genes [85]. The gene duplications expanded the range of phenylalanine–glycine (FG) repeat motif NUPs and of beta barrel/alpha solenoid fold NUPs. These latter are related to coatamer proteins and have been suggested to facilitate bending of the pore membrane during NE/NPC assembly [84]. Thus, though at the sequence level not all NPC proteins are conserved, the central evolutionary mechanism underlying their development argues that they will be essentially conserved at the core.

3. The nuclear envelope proteome

3.1. *Approaches prior to proteomics*

The abundance (~3 million copies per mammalian nucleus [60]) and biochemical properties of the intermediate filament lamins enabled their being the first NE proteins to be identified [78,86]. Following on this the first NE transmembrane proteins (NETs) identified were INM proteins that bound to lamins and so were resilient to extraction by procedures for isolating NEs that were based largely on the resistance of the supporting lamin polymer to high salt and detergents. The lamin B receptor (LBR) was so named because of its identification through its binding to lamin B1 [19] while LAP1 and LAP2 that also bind lamins were identified as major proteins in a NE/lamina fraction that was used to generate monoclonal antibodies [15,87]. The methods of discovering the next round of NETs were surprisingly varied ranging from autoimmune antibodies that stained the NE in the case of MAN1 [88] to a 2-hybrid screen for partners of a kinase of the postsynaptic membrane in muscle that identified the Syne/Nesprin-1 proteins [7].

As different groups were using different organisms to identify NETs and the timing correlated with the beginning of large-scale genome sequencing projects, once the first several proteins had been identified, attempts were made to identify related proteins through sequence similarity. This type of analysis found the Syne/Nesprin-2 NE protein family from its sequence similarity to Syne/Nesprin-1 [7], but it did not identify the Nesprin-3 family that was subsequently identified by proteomics [12,83]. Moreover, the two first identified *Drosophila* NETs do not have mammalian homologs (Otefin [89] and YA [90]). Thus the only way to determine the components of the NE would be directly through proteomic analyses.

3.2. *Strategies for NE proteome determination*

The continuity between the ONM and the peripheral ER on one side and many connections to chromatin on the other require creative approaches to the identification of NE proteins by proteomics, especially after the high number of proteins defined as contaminants in the NPC proteomes (provided they are truly contaminants). Two studies that were relatively similar in purifications yet different in strategies yielded strikingly different results [83,91]. Both studies used equivalent procedures for isolating mammalian NEs as were used in the mammalian NPC study; however, the NEs were either extracted to enrich for proteins associated with the intermediate filament lamin polymer or to enrich for proteins embedded in the membranes.

The first study from the Otto laboratory generated three separate NE fractions: a chaotrope-insoluble fraction, a non-ionic detergent-insoluble fraction and a salt-insoluble fraction [91]. Extraction with chaotropes (4 M Urea, 200 mM sodium carbonate) solubilizes the lamin polymer and most protein-protein interactions, but has no effect on membranes so that proteins embedded in the membrane are protected and maintained in the membrane fraction. The extraction with detergent (1% Triton X-100) should remove all membranes and so only proteins tightly bound to the lamin polymer should remain. Finally extraction with high salt (1 M NaCl) should also leave the lamina intact, but wash away soluble proteins that are weakly associated with it. Each fraction was separated on 2-D gels, and the protein spots were excised and analyzed by MALDI mass spectrometry. Proteins that were found in both the chaotrope and detergent-resistant fractions were considered as candidate NETs. In this way they could remove contamination from the ER and ONM proteins: as the chaotrope fraction contains both NE and ER transmembrane proteins, it alone is insufficient to distinguish INM proteins. This analysis identified most, but not all, previously characterized INM NETs, as well as mammalian Unc84A/SUN1 and a novel protein with no predicted functions that was named LUMA. Both novel mammalian NETs were shown to target to the NE by exogenous expression of the proteins fused to GFP [91].

Where the first study used a “comparative” approach to exclude peripheral ER proteins that also were present in the NE fraction, the second study from the Gerace and Yates laboratories used a “subtractive” approach [83]. In this case a microsomal membrane (MM) fraction was used to identify peripheral ER proteins. The MM fraction was analyzed separately from the NE fraction and all proteins appearing in both fractions were subtracted from the NE fraction. As there are no membranes in the nucleus besides the NE membrane and contaminating membranes of the NE fraction should in theory all also occur in the MM fraction, those transmembrane proteins in just the NE fractions were considered to be true NETs. Multiple NE fractions were also analyzed, but instead of using chaotropes an alkali extraction (0.1 N NaOH) was used to enrich for transmembrane proteins and salt and detergent (400 mM NaCl, 1% Triton X-100) were combined in one extraction to generate a cleaner lamina fraction. Combining both fractions resulted in a more comprehensive analysis of proteins with different biochemical characteristics as compared to the other study in which the different fractions were used to increase the stringency of inclusion. Fractions were analyzed using Multi-Dimensional Protein Identification Technology (MudPIT) [92,93], which couples tandem mass spectrometry with multiple liquid chromatography steps to analyze the complex mixture of peptides generated by direct digestion of isolated membranes. This avoids loss of

membrane proteins that are poorly resolved on 2-D gels [94]. Details of the method are given in [95]. The subtractive approach was validated by the identification of all expected previously characterized NETs in the NE fraction and their absence from the MM fraction. Moreover, the numbers of peptides recovered suggests that they are the most abundant NETs and this is why they were first identified by other means. In addition, 67 previously uncharacterized putative NETs were identified in the NE fraction, which were absent from the MM fraction [83]. All of the original eight tested targeted to the NE in the original study, suggesting that all would prove to be valid identifications; however of the 2/3 now tested only ~70% are valid NETs that are integral to the membrane and target to the NE (P. Malik, N. Korfali, N. Zuleger, V. Lazou, D. M. Kavanagh, G. S. Wilkie, D. G. Batrakou, and E. C. Schirmer, in preparation). Thus this method, like the NPC studies, still brings considerable contaminants. It appears, nonetheless, that some of this failure rate is not due to mis-identification or contamination, but rather to mis-prediction of transmembrane helices.

3.3. *Comparison of benefits and disadvantages between the two approaches*

The "subtractive" and "comparative" approaches used to identify NE-specific proteins both had limitations. The comparative approach disregarded NE-specific proteins that were not associated with the salt and nonionic detergent-insoluble lamina fraction, because no other basis was provided for distinguishing between the NE and ER transmembrane proteins that were present in the membrane-enriched fraction. The more conservative requirement of appearing in both membrane-enriched and lamina-enriched fractions also served to limit the number of identifications.

The subtractive approach had the disadvantage of disregarding proteins that have functions in both the ER and the NE. For example, a known ER protein, torsinA, appeared in both NE and MM fractions. However, torsinA is now known to move between the ER and INM where it interacts with multiple NETs [96].

While the subtraction limited the identifications, the combining all proteins in membrane-enriched and lamina-enriched datasets allowed for a more comprehensive analysis that covered a wider range of biochemical characteristics. This approach is not unreasonable considering the wide range of biochemical characteristics observed just for splice variants of one of the first identified NETs, LAP1 [87]. There were three variants of LAP1 recognized by a monoclonal antibody: the smallest was extracted with less than 200 mM salt, the intermediate sized variant was only about 50% extracted at this salt concentration while the largest remained fully associated with the lamin polymer [15]. Non-equivalence has also been observed for different transmembrane proteins of the NE with regard to their extraction to chaotropes and different detergents. Just as empigen BB favored extraction of NPC components over other integral membane proteins and lamins, partitioning of NETs in the first NE proteomic study indicated that emerin is more extractible by detergents than many other NETs while LBR is more extractible by chaotropes [91]. Thus the two studies highlight the persistent struggle between comprehensive identifications and contaminants.

Another reason for discrepancies between datasets is the randomness of getting single peptides into the mass spec and in obtaining good fragmentation to make identifications. In the present instance this was unlikely to be a primary contributor to variation as 92.7% of the proteins identified in the comparative study were also found in the subtractive study. Moreover 32.9% of the proteins identified in the NPC study were found in the comparative study (59.2% of core NPC proteins) and 89.3% in the subtractive study (if accepting variants or related proteins then this number jumps to 98.9%). This argues that the subtractive study was the most comprehensive, but the caveat as discussed above is that it likely also has the highest number of contaminants.

4. Recapitulation of other organellar functions in the nuclear envelope

4.1. *NPC and NE proteins with dual roles*

There are many previously characterized examples of proteins that have dual functions in the NE and other cellular structures or organelles. The first was Sec13p that functions in both ER vesicle formation [81] and NPC transport [82]. Additionally the DEAD-box helicases An3 and Dbp5 that bind RNA were also found to play an important role in RNA export [97–99]. Many core or peripheral NPC components have now been found to play separate roles on mitotic chromosomes when the NPCs are disassembled in mitosis. Ran, importinß and the Nup107–160 complex have been found to function in mitotic spindle assembly and on kinetochores [100–102].

NE proteins outside the NPC have also been found to have dual localizations and/ or functions. Lamins were originally thought to reside only under the inner nuclear membrane and provide structural support [60]; however in the past several years it has become apparent that lamins assemble and function also in the nucleoplasm [103] and on mitotic spindles [104]. Emerin is clearly an INM protein, binding to lamin A [14,18], the INM protein MAN1 [17], the chromatin binding protein BAF [105], the splicing factor YT521-B [106] and the transcriptional repressors germ cell-less, Btf, and Lmo7 [107–109], but emerin has recently been shown to function also in the outer nuclear membrane and ER [110]. Nesprins were named for their NE association, but are now recognized as a complex family containing many splice variants located throughout the cell [111]. Conversely Torsin A, like Sec13p, was originally characterized as an ER protein but is now known to normally sample the INM and accumulate there with certain point mutants associated with disease [112,113]. Torsin A has also been found to bind LAP1 in the INM and another protein that the authors renamed LULL1 and claimed resides in the ER [96]. However LULL1 was originally identified as NET9 [83] and has been confirmed in the INM [114]. Presumably NET9/LULL1, like Torsin A, can sample both compartments.

4.2. *Newly identified NE proteins with apparent dual roles*

Only a small percentage of the total proteins identified in the liver NE datasets were novel proteins with no predicted

functions. Even among the 67 novel putative NETs identified, nearly a third had predicted functions based on sequence homology. An analysis of both predicted and characterized functional regions indicates a wide range of functions represented in both the NE and MM datasets (Fig. 2A; [115,116]). These functions range from transport and signaling functions to specialized functions of differentiated cells. Some of the proteins associated with cell signaling might converge on or add additional pathways to those recently identified at the NE [37–40]. A clear enrichment in proteins associated with DNA and RNA functions was noticable in the NE compared to the MM. Lamins have been shown to have roles in DNA replication [41,42], transcription [45,46], and RNA splicing [47], but it is unclear whether lamin effects are direct or reflect their serving as recruitment sites for specific enzymes. It has already been shown that transcriptional regulators pRb and germ cell-less respectively interact with lamins [46] and several NETs [17,44,109]. These proteins are among those identified in the liver NE dataset and so other proteins in this dataset that function in DNA and RNA processes might also mediate such lamin effects. Looking in more detail at the types of DNA functions represented in the NE dataset shows a striking enrichment of proteins involved in chromatin organization and remodeling compared to the total proteins in the human genome that function on DNA (Fig. 3A). For proteins functioning on RNA there is both an abundance and enrichment for those involved in RNA splicing and pre-mRNA processing (Fig. 3B). Though less abundant, there is also enrichment for those involved in mRNA end processing and stability, RNA localization and rRNA metabolism. This indicates that there are many potential mechanisms that could direct disease pathology from the NE involving disruption of gene expression beyond those currently being investigated.

Within the transporter category there were many pathways at play in addition to the expected NPC transport-associated proteins. This is not surprising in retrospect as regulated transport of smaller molecules such as ions should be important in the nucleus and potentially for disease, yet there has

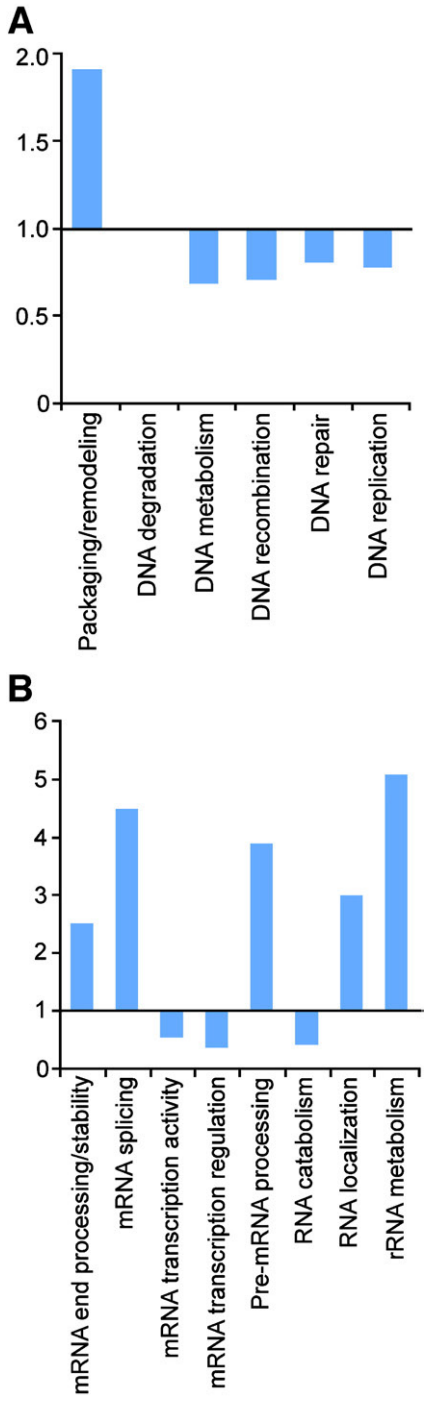


Fig. 3 – Functional distribution of proteins identified in the liver NE and MM proteomic study. (A) As subtraction of all proteins identified in the MM datasets would potentially miss interesting proteins that are nonetheless enriched in the NE, NE- and MM-enriched datasets were generated by comparing the number of spectra recovered per run for each dataset. If a NE protein had at least 5× more spectra than were recovered in the MM or vice-versa they were included in these datasets. The Panther (Protein ANalysis THrough Evolutionary Relationships) Classification System [115,116] was then used to organize proteins in each dataset according to functional annotations. A clear enrichment for proteins involved in DNA and RNA functions (DNA/RNA) was observed in the NE compared to the MM. As only 2 proteins (<1%) recovered in the liver MM had anything to do with DNA function and 4 (1.7%) with RNA function, the % distribution of all DNA or RNA functions in the human genome was compared with that for the proteins recovered in the NE-enriched dataset. The ratio of the percent of a function in the NE to its percent represented in the total human proteome is presented. (A) DNA functions represented 65/854 proteins in the NE-enriched dataset compared to 977/25431 proteins in the total human dataset (or 7.6% vs 3.8%). Within the DNA set ther was a strong enrichment at the NE for proteins involved in chromatin structure and remodeling. All other categories showed a relative decrease in functions at the NE. (B) RNA functions represented 330 of the 854 proteins in the NE-enriched dataset compared to 4295 of the 25431 proteins in the total human dataset (or 38.6% vs 16.8%). Within the RNA dataset functions in RNA splicing and pre-mRNA processing were both abundant and enriched for. Additionally functions in mRNA end processing and stability, RNA localization and rRNA metabolism were enriched for, though much less abundant.

been very little focus in this area within the now chromatin-oriented NE field. Inositol(1,4,5)P_3 receptors were long ago reported at the NE [117,118], and specific ones were also found in the liver NE proteomic study that have been since directly tested [119]. Other regulators of Ca^{2+} transport and signaling were also found such as the ryanodine receptor, which has since been shown to function in the NE from several studies [120–122]. Indeed Ca^{2+} oscillations have been shown to affect gene expression [123] and so regulation of ion transport could also affect disease pathology through gene expression. Zn^{2+} transporters, Na^{2+}/H^{+} exchangers and many other ion transporters were also found in the NE proteomic datasets, many of which have since been directly shown at the NE [124,125], and there was a striking enrichment for electron transporters.

Detailed analysis of the enrichment in signaling proteins indicates that lipid signaling mechanisms in particular could be important in transducing signals to the nucleus in a separate or backup mechanism to those that depend on transport of molecules through the NPC. In support of this NET39 that was identified in the liver NE proteomic study and confirmed at the NE [83] is a lipid phosphatase/lipid phosphotransferase of the candidate sphingomyelin synthase class and thus also now called CCS2ß [126]. This is the least studied subclass of the lipid phosphatase/transferases, but others have been shown to play roles in regulation of cell growth and survival. There appear to be many proteins in the NE involved in these processes as NET13, also identified in the liver NE study, has now been shown to be a sphingomyelin synthase [127].

The various biochemical functions observed for these proteins and our confirmation of their partial representation in the NE has significant implications for many aspects of NE biology. Some studies have argued that nuclear membrane growth during interphase requires vesicle fusion or that ER membranes perfuse around the NPCs, but the identification of several enzymes involved in lipid generation in both NE datasets argues for de novo generation of lipid during interphase when the NE grows 3- to 4-fold. The recently reported function of NET8 (LPGAT1) as a phosphatidylglycerol acyltransferase [128] could serve to direct lipid content in the nuclear membrane, which has been reported to be one of the subcellular membranes containing phosphatidylglycerol [129]. The membrane trafficking function of NET24 (ERGIC-32; [130]) could have relevance for NE disassembly as Rab5 (which also appeared in the NE proteomics datasets) functions in NE disassembly in addition to its normal vesicle transport role [131]. Similarly, the signaling roles of emerin, MAN1 and AKAP149 [37–40] could be facilitated by functions of NETs 45 and 55 that respectively have been reported to directly function in signaling and to be a homolog of a protein involved in signaling [132–134]. Thus these NETs could be involved in signal transduction pathways similar to those of the well-characterized NET MAN1 in Smad/BMP/TGFß signaling [135].

5. Future directions

5.1. *Tissue variation in the NE proteome*

There is precedent for tissue differences in NE proteins from the distribution of different lamin subtypes that have been shown to vary in relative concentrations during development and in different tissues [136–138]. There are also tissue-specific splice variants of lamins such as the lamin C2 that appears during spermatogenesis [139]. Although the first identified NETs were widely expressed, comparison of the novel NETs identified in the liver proteomic study to a transcriptome database [140] indicated that a significant percentage of the proteins recovered were preferentially expressed in liver and/or had restricted tissue expression [13,141]. Moderate tissue variation has been reported recently in proteomes from mitochondria and other organelles [142,143]. A current study analyzing the NE proteome from several different tissues confirms that there is not only considerable tissue variation in the NE proteome but further suggests that there is more variation in the NE than these other subcellular organelles (N. Korfali, G.S Wilkie, E.A.L. Fairley, S.K. Swanson, D.G. Batrakou, P. Malik, A.R.W. Kerr, L. Florens, and E.C. Schirmer, in preparation).

The finding of tissue variation in the NE proteome together with observations that different epitopes on NE proteins are occupied in different tissues [144] argue that tissue-preferential binding partners of the NE proteins mutated in disease might mediate the tissue-preferential phenotypes of the wide range of NE-related diseases [13,145]. Different diseases preferentially affect muscle, neurons, bone, skin, heart, fat, and immune cells. Some diseases have partial overlap between a subset of these tissues. Thus a particular point mutant might disrupt a functional complex in one tissue where a specific partner within the complex is expressed, but have only a minimal effect on a different functional complex in another tissue where a different partner that has slightly different binding characteristics occupies the same general binding site. If this is the case, then determining the protein composition of the NE in different tissues may be critical to understanding these diseases. As variation will likely mostly occur in terms of relative amounts of protein as opposed to all or none effects, it will also be important for such future analyses to be performed using recently developed quantitative approaches such as SILAC, iTRAQ, and protein correlation profiling.

This also applies for the NPC as a study on the integral NPC component gp210 revealed that it has restricted tissue expression in mammals [146]. It is nonetheless essential for viability in *C. elegans* by RNAi [147]. Gp210 was also reported to be essential in HeLa cells [147] but other studies found that it could be deleted and that, rather, another integral NPC component, NDC1/NET3, is essential in HeLa cells [26,148,149]. Interestingly, they also found that NDC1 was not essential in *C. elegans* [148]. It is possible thus that different combinations of NPC components are expressed in different tissues, thus rendering them essential depending on the tissue being sampled. Accordingly it will be important to directly compare knockdown of several different NUPs over a panel of tissues. This, however, also means that aspects of nuclear pore structure and function may differ between different tissues and indicates a need to compare the NPC proteome in different tissues using identical extraction methodologies and quantitative approaches so that differences in extractability due to different detergents or salt concentrations used are not misread as tissue variation.

5.2. *A further use of proteomics to gain structural insights into the NPC*

The two NPC proteomic studies were considered at the time to be the definitive conclusion to NPC proteomics, but a recent yeast study indicates that mass spectrometry can contribute much more to understanding the NPC even without counting potential tissue variation. This study performed rapid pull-downs with each component of the NPC followed by quantitative mass spectrometry to identify binding partners and their ratios [30,150]. Stringent conditions were used so that by comparing the ratios of NUPs in different fractions those most likely to be in direct contact with one another could be ascertained. These data were combined with position measurements from immuno-electron microscopy for epitopes on NUPs and other available data to generate what is by far the highest resolution structure of the NPC to date [30]. This structure gives a more detailed view of the internal structure of the core NPC and the organization of the many core NUPs that is diagrammed schematically in Fig. 1C.

5.3. *Additional NPC proteins in the NE proteome datasets?*

It is noteworthy that the integral NPC component NDC1/NET3 was not identified in the mammalian NPC proteomic study [77], apparently due to its resilience to the extraction with empigen BB that was used to isolate the NPC proteins as their method relied on extracting NPC components from the membrane and lamina. NDC1/NET3 was, however, identified in the less conservative study of the NE that enriched for transmembrane proteins [83] and was later shown to be the third integral component of the NPC in mammalian cells [26,27,148] (it had previously been identified in yeast, but was too divergent for mammalian homologs to be identified by genome searching [23]), and this suggests that other proteins may have been missed in the NPC studies that appear in the NE proteomic datasets. If this is the case, it might account in part for the large differences between the masses predicted by electron microscopy studies and the final masses calculated from the proteins identified in the proteomic studies.

Searching the liver NE datasets for characteristics of NPC proteins such as the appearance of FG repeats indicates that proteins containing this motif are significantly enriched in the NE compared to the whole genome. FG repeats were observed in some of the first NUPs identified and it soon became apparent that they represented a common motif among most of the NUPs when three NUPs were identified using an antibody that recognized regions containing a GLFG motif: NUP49, NUP100 and NUP116 were found to have respectively 13, 29 and 33 GLFG repeats [151]. It is noteworthy, however, that only 10 of the 30 core NUPs have six or more FG repeats and 13 have one to five FGs (Table 1). Thus very large numbers may not be required for some FG functions. The FG repeat regions of NUPs are thought to be disordered [152] and tend to have >30% of polar residues between them. It has been suggested that the FGs within unstructured regions interact within the central channel of the NPC to provide an entropic barrier to passive diffusion [153], though the exact nature of that barrier is the subject of active debate [154,155]. One possible mechanism for transport through the entropic barrier would be exchange of FG interactions in the barrier with FG motifs on the surface of proteins associated with transport cargos [156]. The transport receptor importinß has 3 FG repeats so that coating the surface of the substrate cargo with importinß could potentially provide these FG repeats on the outer surface to facilitate navigation through the central channel (Fig. 2B). If this hypothesis is correct, then other proteins that have FG repeats could similarly coat cargoes to facilitate their transport. A general search of all proteins in the human genome database reveals that only 1.86% have 5 or more FG repeats with 30% or more polar residues between at least 4 of them. In contrast, in the liver NE-enriched dataset of proteins 5.3% of proteins fulfill these criteria. It is particularly interesting that some DEAD box helicases have been shown to be important for nuclear export of mRNAs (including DEAD box protein 5 that has 5 FGs; [98,99]) and among the components of the TREX complex that is important for RNA export [157] is Sub2p, a DEAD box ATPase containing 2 FGs [158]. Strikingly, several of the liver NE proteins carrying multiple FGs are also RNA binding proteins including DEAD box proteins (Table 2). Thus if these stay bound to the mRNA during the process of transport they might facilitate the act of transport by coating the surface of the RNA with FGs (Fig. 2D).

Table 2 – RNA-binding proteins with FG repeats separated by polar residues found in the liver NE datasets

FG repeats	Protein name	Accession
9	hnRNP core protein A1	NP_058944.1
9	hnRNP A2/B1 isoform 1	NP_058086.2
8	RNA binding motif protein 19	XP_222200.4
6	hnRNP M	NP_446328.1
6	hnRNP A0	XP_001001311.1
6	U5 snRNP-specific protein, 116 kDa	XP_001081526.1
6	DEAD box polypeptide 42	XP_001081592.1

Criteria used to search for FG repeat containing proteins were 5 or more FG repeats, at least 4 of which have ≥30% polar residues between them (provided >3 residues separate the FGs).

These observations indicate that the NE has a great many more functions and diversity than ever considered. To gain clarity of these functions it will be necessary to sample further methods of enrichment to reduce the number of contaminants and also to sample from different tissues, in both cases using new quantitative approaches. Even in the absence of further sampling, mining of the existing NE proteome datasets should yield considerable treasure for many years to come.

Acknowledgements

D.G. Batrakou is supported by the Darwin Trust: A.R.W. Kerr and E.C. Schirmer are supported by the Wellcome Trust.

REFERENCES

[1] Hetzer MW, Walther TC, Mattaj IW. Pushing the envelope: structure, function, and dynamics of the nuclear periphery. Annu Rev Cell Dev Biol 2005;21:347–80.

[2] Prunuske AJ, Ullman KS. The nuclear envelope: form and reformation. Curr Opin Cell Biol 2006;18:108–16.
[3] Schirmer EC, Foisner R. Proteins that associate with lamins: many faces, many functions. Exp Cell Res 2007;313:2167–79.
[4] Suntharalingam M, Wente SR. Peering through the pore: nuclear pore complex structure, assembly, and function. Dev Cell 2003;4:775–89.
[5] Callan HG, Tomlin SG. Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. Proc R Soc of Lond [Biol] 1950;137:367–78.
[6] Fry DJ. The nuclear envelope in mammalian cells. In: Jameson GA, Robinson DM, editors. Mammalian Cell Membranes. Woburn, Mass: Butterworth; 1976. p. 197–265.
[7] Apel ED, Lewis RM, Grady RM, Sanes JR. Syne-1, a dystrophin- and Klarsicht-related protein associated with synaptic nuclei at the neuromuscular junction. J Biol Chem 2000;275:31986–95.
[8] Malone CJ, Fixsen WD, Horvitz HR, Han M. UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during *C. elegans* development. Development 1999;126:3171–81.
[9] Starr DA, Han M. Role of ANC-1 in tethering nuclei to the actin cytoskeleton. Science 2002;298:406–9.
[10] Starr DA, Han M. A genetic approach to study the role of nuclear envelope components in nuclear positioning. Novartis Found Symp 2005;264:208–19 discussion 219–230.
[11] Zhang Q, Skepper JN, Yang F, Davies JD, Hegyi L, Roberts RG, et al. Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. J Cell Sci 2001;114:4485–98.
[12] Wilhelmsen K, Litjens SH, Kuikman I, Tshimbalanga N, Janssen H, van den Bout I, et al. Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. J Cell Biol 2005;171:799–810.
[13] Schirmer EC, Gerace L. The nuclear membrane proteome: extending the envelope. Trends Biochem Sci 2005;30:551–8.
[14] Clements L, Manilal S, Love DR, Morris GE. Direct interaction between emerin and lamin A. Biochem Biophys Res Commun 2000;267:709–14.
[15] Foisner R, Gerace L. Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. Cell 1993;73:1267–79.
[16] Hodzic DM, Yeater DB, Bengtsson L, Otto H, Stahl PD. Sun2 is a novel mammalian inner nuclear membrane protein. J Biol Chem 2004;279:25805–12.
[17] Mansharamani M, Wilson KL. Direct binding of nuclear membrane protein MAN1 to emerin in vitro and two modes of binding to barrier-to-autointegration factor. J Biol Chem 2005;280:13863–70.
[18] Sakaki M, Koike H, Takahashi N, Sasagawa N, Tomioka S, Arahata K, et al. Interaction between emerin and nuclear lamins. J Biochem (Tokyo) 2001;129:321–7.
[19] Worman HJ, Yuan J, Blobel G, Georgatos SD. A lamin B receptor in the nuclear envelope. Proc Natl Acad Sci U S A 1988;85:8531–4.
[20] Mattout-Drubezki A, Gruenbaum Y. Dynamic interactions of nuclear lamina proteins with chromatin and transcriptional machinery. Cell Mol Life Sci 2003;60:2053–63.
[21] Paine PL, Moore LC, Horowitz SB. Nuclear envelope permeability. Nature 1975;254:109–14.
[22] Pante N, Kann M. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. Mol Biol Cell 2002;13:425–34.
[23] Chial HJ, Rout MP, Giddings TH, Winey M. Saccharomyces cerevisiae Ndc1p is a shared component of nuclear pore complexes and spindle pole bodies. J Cell Biol 1998;143:1789–800.
[24] Greber UF, Senior A, Gerace L. 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 1990;9:1495–502.
[25] Hallberg E, Wozniak RW, Blobel G. An integral membrane protein of the pore membrane domain of the nuclear envelope contains a nucleoporin-like region. J Cell Biol 1993;122:513–21.
[26] Mansfeld J, Guttinger S, Hawryluk-Gara LA, Pante N, Mall M, Galy V, et al. The conserved transmembrane nucleoporin NDC1 is required for nuclear pore complex assembly in vertebrate cells. Mol Cell 2006;22:93–103.
[27] Lau CK, Delmar VA, Forbes DJ. Topology of yeast Ndc1p: predictions for the human NDC1/NET3 homologue. Anat Rec A Discov Mol Cell Evol Biol 2006;288:681–94.
[28] Wozniak RW, Bartnik E, Blobel G. Primary structure analysis of an integral membrane glycoprotein of the nuclear pore. J Cell Biol 1989;108:2083–92.
[29] Akey CW, Radermacher M. Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. J Cell Biol 1993;122:1–19.
[30] Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, et al. The molecular architecture of the nuclear pore complex. Nature 2007;450:695–701.
[31] Hinshaw JE, Carragher BO, Milligan RA. Architecture and design of the nuclear pore complex. Cell 1992;69:1133–41.
[32] Lammerding J, Schulze P, Takahashi T, Kozlov S, Sullivan T, Kamm R, et al. Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. J Clin Invest 2004;113:370–8.
[33] Liu J, Ben-Shahar T, Riemer D, Treinin M, Spann P, Weber K, et al. Essential roles for *Caenorhabditis elegans* lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. Mol Biol Cell 2000;11:3937–47.
[34] Rowat AC, Lammerding J, Ipsen JH. Mechanical properties of the cell nucleus and the effect of emerin deficiency. Biophys J 2006;91:4649–64.
[35] Schirmer EC, Gerace L. The stability of the nuclear lamina polymer changes with the composition of lamin subtypes according to their individual binding strengths. J Biol Chem 2004;279:42811–7.
[36] Schirmer EC, Guan T, Gerace L. Involvement of the lamin rod domain in heterotypic lamin interactions important for nuclear organization. J Cell Biol 2001;153:479–89.
[37] Osada S, Ohmori SY, Taira M. XMAN1, an inner nuclear membrane protein, antagonizes BMP signaling by interacting with Smad1 in Xenopus embryos. Development 2003;130:1783–94.
[38] Pan D, Estevez-Salmeron LD, Stroschein SL, Zhu X, He J, Zhou S, et al. The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-{beta} superfamily of cytokines. J Biol Chem 2005;280:15992–6001.
[39] Steen RL, Martins SB, Tasken K, Collas P. 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 2000;150:1251–62.
[40] Markiewicz E, Tilgner K, Barker N, van de Wetering M, Clevers H, Dorobek M, et al. The inner nuclear membrane protein emerin regulates beta-catenin activity by restricting its accumulation in the nucleus. EMBO J 2006;25:3275–85.
[41] Ellis DJ, Jenkins H, Whitfield WG, Hutchison CJ. 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 1997;110:2507–18.
[42] Spann TP, Moir RD, Goldman AE, Stick R, Goldman RD. Disruption of nuclear lamin organization alters the

distribution of replication factors and inhibits DNA synthesis. J Cell Biol 1997;136:1201–12.
[43] Martins S, Eikvar S, Furukawa K, Collas P. HA95 and LAP2 beta mediate a novel chromatin-nuclear envelope interaction implicated in initiation of DNA replication. J Cell Biol 2003;160:177–88.
[44] Nili E, Cojocaru GS, Kalma Y, Ginsberg D, Copeland NG, Gilbert DJ, et al. Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). J Cell Sci 2001;114:3297–307.
[45] Johnson BR, Nitta RT, Frock RL, Mounkes L, Barbie DA, Stewart CL, et al. A-type lamins regulate retinoblastoma protein function by promoting subnuclear localization and preventing proteasomal degradation. Proc Natl Acad Sci U S A 2004;101:9677–82.
[46] Ozaki T, Saijo M, Murakami K, Enomoto H, Taya Y, Sakiyama S. Complex formation between lamin A and the retinoblastoma gene product: identification of the domain on lamin A required for its interaction. Oncogene 1994;9:2649–53.
[47] Kumaran RI, Muralikrishna B, Parnaik VK. Lamin A/C speckles mediate spatial organization of splicing factor compartments and RNA polymerase II transcription. J Cell Biol 2002;159:783–93.
[48] Capell BC, Collins FS. Human laminopathies: nuclei gone genetically awry. Nat Rev Genet 2006;7:940–52.
[49] Foisner R, Aebi U, Bonne G, Gruenbaum Y, Novelli G. 141st ENMC International Workshop inaugural meeting of the EURO-Laminopathies project "nuclear envelope-linked rare human diseases: from molecular pathophysiology towards clinical applications", 10-12 March 2006, Naarden, The Netherlands. Neuromuscul Disord 2007;17:655–60.
[50] Worman HJ, Bonne G. "Laminopathies": a wide spectrum of human diseases. Exp Cell Res 2007;313:2121–33.
[51] Shaklai S, Amariglio N, Rechavi G, Simon AJ. Gene silencing at the nuclear periphery. FEBS J 2007;274:1383–92.
[52] Crisp M, Liu Q, Roux K, Rattner JB, Shanahan C, Burke B, et al. Coupling of the nucleus and cytoplasm: role of the LINC complex. J Cell Biol 2006;172:41–53.
[53] Kosova B, Pante N, Rollenhagen C, Podtelejnikov A, Mann M, Aebi U, et al. Mlp2p, a component of nuclear pore attached intranuclear filaments, associates with nic96p. J Biol Chem 2000;275:343–50.
[54] Strambio-de-Castillia C, Blobel G, Rout MP. Proteins connecting the nuclear pore complex with the nuclear interior. J Cell Biol 1999;144:839–55.
[55] Galy V, Olivo-Marin JC, Scherthan H, Doye V, Rascalou N, Nehrbass U. Nuclear pore complexes in the organization of silent telomeric chromatin. Nature 2000;403:108–12.
[56] Scherthan H, Jerratsch M, Li B, Smith S, Hulten M, Lock T, et al. Mammalian meiotic telomeres: protein composition and redistribution in relation to nuclear pores. Mol Biol Cell 2000;11:4189–203.
[57] Reichelt R, Holzenburg A, Buhle Jr EL, Jarnik M, Engel A, Aebi U. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. J Cell Biol 1990;110:883–94.
[58] Rout MP, Blobel G. Isolation of the yeast nuclear pore complex. J Cell Biol 1993;123:771–83.
[59] Yang Q, Rout MP, Akey CW. Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. Mol Cell 1998;1:223–34.
[60] Gerace L, Burke B. Functional organization of the nuclear envelope. Annu Rev Cell Biol 1988;4:335–74.
[61] Gerace L, Ottaviano Y, Kondor-Koch C. Identification of a major polypeptide of the nuclear pore complex. J Cell Biol 1982;95:826–37.
[62] Davis LI, Blobel G. Identification and characterization of a nuclear pore complex protein. Cell 1986;45:699–709.
[63] Park MK, D'Onofrio M, Willingham MC, Hanover JA. A monoclonal antibody against a family of nuclear pore proteins (nucleoporins): O-linked *N*-acetylglucosamine is part of the immunodeterminant. Proc Natl Acad Sci U S A 1987;84:6462–6.
[64] Snow CM, Senior A, Gerace L. Monoclonal antibodies identify a group of nuclear pore complex glycoproteins. J Cell Biol 1987;104:1143–56.
[65] Aris JP, Blobel G. Yeast nuclear envelope proteins cross react with an antibody against mammalian pore complex proteins. J Cell Biol 1989;108:2059–67.
[66] Davis LI, Fink GR. The NUP1 gene encodes an essential component of the yeast nuclear pore complex. Cell 1990;61:965–78.
[67] Carmo-Fonseca M, Kern H, Hurt EC. Human nucleoporin p62 and the essential yeast nuclear pore protein NSP1 show sequence homology and a similar domain organization. Eur J Cell Biol 1991;55:17–30.
[68] Nehrbass U, Kern H, Mutvei A, Horstmann H, Marshallsay B, Hurt EC. NSP1: a yeast nuclear envelope protein localized at the nuclear pores exerts its essential function by its carboxy-terminal domain. Cell 1990;61:979–89.
[69] Starr CM, D'Onofrio M, Park MK, Hanover JA. Primary sequence and heterologous expression of nuclear pore glycoprotein p62. J Cell Biol 1990;110:1861–71.
[70] Hu T, Guan T, Gerace L. Molecular and functional characterization of the p62 complex, an assembly of nuclear pore complex glycoproteins. J Cell Biol 1996;134:589–601.
[71] Kita K, Omata S, Horigome T. Purification and characterization of a nuclear pore glycoprotein complex containing p62. J Biochem 1993;113:377–82.
[72] Marelli M, Aitchison JD, Wozniak RW. Specific binding of the karyopherin Kap121p to a subunit of the nuclear pore complex containing Nup53p, Nup59p, and Nup170p. J Cell Biol 1998;143:1813–30.
[73] Grandi P, Doye V, Hurt EC. Purification of NSP1 reveals complex formation with 'GLFG' nucleoporins and a novel nuclear pore protein NIC96. EMBO J 1993;12:3061–71.
[74] Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol 2000;148:635–51.
[75] Maul GG, Deaven L. Quantitative determination of nuclear pore complexes in cycling cells with differing DNA content. J Cell Biol 1977;73:748–60.
[76] Winey M, Yarar D, Giddings Jr TH, Mastronarde DN. Nuclear pore complex number and distribution throughout the *Saccharomyces cerevisiae* cell cycle by three-dimensional reconstruction from electron micrographs of nuclear envelopes. Mol Biol Cell 1997;8:2119–32.
[77] Cronshaw J, Krutchinsky A, Zhang W, Chait B, Matunis M. Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol 2002;158:915–27.
[78] Aaronson RP, Blobel G. Isolation of nuclear pore complexes in association with a lamina. Proc Natl Acad Sci U S A 1975;72:1007–11.
[79] Blobel G, Potter VR. Nuclei from rat liver: isolation method that combines purity with high yield. Science 1966;154:1662–5.
[80] Dwyer N, Blobel G. A modified procedure for the isolation of a pore complex-lamina fraction from rat liver nuclei. J Cell Biol 1976;70:581–91.
[81] Pryer NK, Salama NR, Schekman R, Kaiser CA. Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum in vitro. J Cell Biol 1993;120:865–75.
[82] Siniossoglou S, Wimmer C, Rieger M, Doye V, Tekotte H, Weise C, et al. A novel complex of nucleoporins, which

includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. Cell 1996;84:265–75.
[83] Schirmer EC, Florens L, Guan T, Yates, Gerace L. Nuclear membrane proteins with potential disease links found by subtractive proteomics. Science 2003;301:1380–2.
[84] Devos D, Dokudovskaya S, Williams R, Alber F, Eswar N, Chait BT, et al. Simple fold composition and modular architecture of the nuclear pore complex. Proc Natl Acad Sci U S A 2006;103:2172–7.
[85] Langkjaer RB, Cliften PF, Johnston M, Piskur J. Yeast genome duplication was followed by asynchronous differentiation of duplicated genes. Nature 2003;421:848–52.
[86] Gerace L, Blum A, Blobel G. Immunocytochemical localization of the major polypeptides of the nuclear pore complex-lamina fraction. Interphase and mitotic distribution. J Cell Biol 1978;79:546–66.
[87] Senior A, Gerace L. Integral membrane proteins specific to the inner nuclear membrane and associated with the nuclear lamina. J Cell Biol 1988;107:2029–36.
[88] Paulin-Levasseur M, Blake DL, Julien M, Rouleau L. The MAN antigens are non-lamin constituents of the nuclear lamina in vertebrate cells. Chromosoma 1996;104:367–79.
[89] Ashery-Padan R, Weiss AM, Feinstein N, Gruenbaum Y. Distinct regions specify the targeting of otefin to the nucleoplasmic side of the nuclear envelope. J Biol Chem 1997;272:2493–9.
[90] Goldberg M, Lu H, Stuurman N, Ashery-Padan R, Weiss AM, Yu J, et al. Interactions among Drosophila nuclear envelope proteins lamin, otefin, and YA. Mol Cell Biol 1998;18:4315–23.
[91] Dreger M, Bengtsson L, Schoneberg T, Otto H, Hucho F. Nuclear envelope proteomics: novel integral membrane proteins of the inner nuclear membrane. Proc Natl Acad Sci U S A 2001;98:11943–8.
[92] Washburn MP, Wolters D, Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 2001;19:242–7.
[93] Wolters DA, Washburn MP, Yates JR. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 2001;73:5683–90.
[94] Santoni V, Molloy M, Rabilloud T. Membrane proteins and proteomics: un amour impossible? Electrophoresis 2000;21:1054–70.
[95] Florens L, Korfali N, Schirmer EC. Subcellular fractionation and proteomics of nuclear envelopes. Methods Mol Biol 2008;432:117–37.
[96] Goodchild RE, Dauer WT. The AAA+ protein torsinA interacts with a conserved domain present in LAP1 and a novel ER protein. J Cell Biol 2005;168:855–62.
[97] Askjaer P, Bachi A, Wilm M, Bischoff FR, Weeks DL, Ogniewski V, et al. RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. Mol Cell Biol 1999;19:6276–85.
[98] Tseng SS, Weaver PL, Liu Y, Hitomi M, Tartakoff AM, Chang TH. Dbp5p, a cytosolic RNA helicase, is required for poly(A)+ RNA export. EMBO J 1998;17:2651–62.
[99] Zhao J, Jin SB, Bjorkroth B, Wieslander L, Daneholt B. The mRNA export factor Dbp5 is associated with Balbiani ring mRNP from gene to cytoplasm. EMBO J 2002;21:1177–87.
[100] Ohba T, Nakamura M, Nishitani H, Nishimoto T. Self-organization of microtubule asters induced in Xenopus egg extracts by GTP-bound Ran. Science 1999;284:1356–8.
[101] Wiese C, Wilde A, Moore MS, Adam SA, Merdes A, Zheng Y. Role of importin-beta in coupling Ran to downstream targets in microtubule assembly. Science 2001;291:653–6.
[102] Belgareh N, Rabut G, Bai SW, van Overbeek M, Beaudouin J, Daigle N, et al. An evolutionarily conserved NPC subcomplex, which redistributes in part to kinetochores in mammalian cells. J Cell Biol 2001;154:1147–60.
[103] Moir RD, Spann TP, Goldman RD. The dynamic properties and possible functions of nuclear lamins. Int Rev Cytol 1995:141–82.
[104] Tsai MY, Wang S, Heidinger JM, Shumaker DK, Adam SA, Goldman RD, et al. A mitotic lamin B matrix induced by RanGTP required for spindle assembly. Science 2006;311:1887–93.
[105] Lee KK, Haraguchi T, Lee RS, Koujin T, Hiraoka Y, Wilson KL. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. J Cell Sci 2001;114:4567–73.
[106] Wilkinson FL, Holaska JM, Zhang Z, Sharma A, Manilal S, Holt I, et al. Emerin interacts in vitro with the splicing-associated factor, YT521-B. Eur J Biochem 2003;270:2459–66.
[107] Holaska JM, Rais-Bahrami S, Wilson KL. Lmo7 is an emerin-binding protein that regulates the transcription of emerin and many other muscle-relevant genes. Hum Mol Gen 2006;15:3459–72.
[108] Bengtsson L, Wilson KL. Multiple and surprising new functions for emerin, a nuclear membrane protein. Curr Opin Cell Biol 2004;16:73–9.
[109] Holaska JM, Lee K, Kowalski AK, Wilson KL. Transcriptional repressor germ cell-less (GCL) and barrier to autointegration factor (BAF) compete for binding to emerin in vitro. J Biol Chem 2003;278:6969–75.
[110] Salpingidou G, Smertenko A, Hausmanowa-Petrucewicz I, Hussey PJ, Hutchison CJ. A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane. J Cell Biol 2007;178:897–904.
[111] Zhang Q, Ragnauth CD, Skepper JN, Worth NF, Warren DT, Roberts RG, et al. Nesprin-2 is a multi-isomeric protein that binds lamin and emerin at the nuclear envelope and forms a subcellular network in skeletal muscle. J Cell Sci 2005;118:673–87.
[112] Goodchild RE, Dauer WT. Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. Proc Natl Acad Sci U S A 2004;101:847–52.
[113] Naismith TV, Heuser JE, Breakefield XO, Hanson PI. TorsinA in the nuclear envelope. Proc Natl Acad Sci U S A 2004;101:7612–7.
[114] Chen IH, Huber M, Guan T, Bubeck A, Gerace L. Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis. BMC Cell Biology 2006;7:38.
[115] Mi H, Guo N, Kejariwal A, Thomas PD. PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. Nucleic Acids Res 2007;35:D247–252.
[116] Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003;13:2129–41.
[117] Humbert JP, Matter N, Artault JC, Koppler P, Malviya AN. Inositol 1,4,5-trisphosphate receptor is located to the inner nuclear membrane vindicating regulation of nuclear calcium signaling by inositol 1,4,5-trisphosphate. Discrete distribution of inositol phosphate receptors to inner and outer nuclear membranes. J Biol Chem 1996;271:478–85.
[118] Lanini L, Bachs O, Carafoli E. The calcium pump of the liver nuclear membrane is identical to that of endoplasmic reticulum. J Biol Chem 1992;267:11548–52.
[119] Cruttwell C, Bernard J, Hilly M, Nicolas V, Tunwell RE, Mauger JP. Dynamics of the inositol 1,4,5-trisphosphate receptor during polarisation of MDCK cells. Biol Cell 2005.
[120] Gerasimenko JV, Maruyama Y, Yano K, Dolman NJ, Tepikin AV, Petersen OH, et al. NAADP mobilizes Ca^{2+} from a thapsigargin-sensitive store in the nuclear envelope by activating ryanodine receptors. J Cell Biol 2003;163:271–82.
[121] Erickson ES, Mooren OL, Moore-Nichols D, Dunn RC. Activation of ryanodine receptors in the nuclear envelope

alters the conformation of the nuclear pore complex. Biophys Chem 2004;112:1–7.
[122] Pare GC, Easlick JL, Mislow JM, McNally EM, Kapiloff MS. Nesprin-1alpha contributes to the targeting of mAKAP to the cardiac myocyte nuclear envelope. Exp Cell Res 2005;303:388–99.
[123] Dolmetsch RE, Xu K, Lewis RS. Calcium oscillations increase the efficiency and specificity of gene expression. Nature 1998;392:933–6.
[124] Arteaga MF, Gutierrez R, Avila J, Mobasheri A, Diaz-Flores L, Martin-Vasallo P. Regeneration influences expression of the Na+, K+-atpase subunit isoforms in the rat peripheral nervous system. Neuroscience 2004;129:691–702.
[125] Ronaldson PT, Bendayan M, Gingras D, Piquette-Miller M, Bendayan R. Cellular localization and functional expression of P-glycoprotein in rat astrocyte cultures. J Neurochem 2004;89:788–800.
[126] Sigal YJ, McDermott MI, Morris AJ. Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions. Biochem J 2005;387:281–93.
[127] Krut O, Wiegmann K, Kashkar H, Yazdanpanah B, Kronke M. Novel tumor necrosis factor-responsive mammalian neutral sphingomyelinase-3 is a C-tail-anchored protein. J Biol Chem 2006;281:13784–93.
[128] Yang Y, Cao J, Shi Y. Identification and characterization of a gene encoding human LPGAT1, an endoplasmic reticulum-associated lysophosphatidylglycerol acyltransferase. J Biol Chem 2004;279:55866–74.
[129] Hallman M, Gluck L. Phosphatidylglycerol in lung surfactant. II. Subcellular distribution and mechanism of biosynthesis in vitro. Biochim Biophys Acta 1975;409:172–91.
[130] Breuza L, Halbeisen R, Jeno P, Otte S, Barlowe C, Hong W, et al. Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. J Biol Chem 2004;279:47242–53.
[131] Audhya A, Desai A, Oegema K. A role for Rab5 in structuring the endoplasmic reticulum. J Cell Biol 2007;178:43–56.
[132] Jacobsson JA, Haitina T, Lindblom J, Fredriksson R. Identification of six putative human transporters with structural similarity to the drug transporter SLC22 family. Genomics 2007;90:595–609.
[133] Coolen MW, Van Loo KM, Van Bakel NN, Pulford DJ, Serneels L, De Strooper B, et al. Gene dosage effect on gamma-secretase component Aph-1b in a rat model for neurodevelopmental disorders. Neuron 2005;45:497–503.
[134] Diao F, Li S, Tian Y, Zhang M, Xu LG, Zhang Y, et al. Negative regulation of MDA5- but not RIG-I-mediated innate antiviral signaling by the dihydroxyacetone kinase. Proc Natl Acad Sci U S A 2007;104:11706–11.
[135] Bengtsson L. What MAN1 does to the Smads TGFß/BMP signaling and the nuclear envelope. FEBS J 2007;274:1374–82.
[136] Broers JL, Machiels BM, Kuijpers HJ, Smedts F, van den Kieboom R, Raymond Y, et al. A- and B-type lamins are differentially expressed in normal human tissues. Histochem Cell Biol 1997;107:505–17.
[137] Jansen M, Machiels B, Hopman A, Broers J, Bot F, Arends J, et al. Comparison of A and B-type lamin expression in reactive lymph nodes and nodular sclerosing Hodgkin's disease. Histopathology 1997;31:304–12.
[138] Rober RA, Weber K, Osborn M. Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. Development 1989;105:365–78.
[139] Alsheimer M, Benavente R. Change of karyoskeleton during mammalian spermatogenesis: expression pattern of nuclear lamin C2 and its regulation. Exp Cell Res 1996;228:181–8.
[140] Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, et al. Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 2002;99:4465–70.
[141] Kavanagh DM, Powell WE, Malik P, Lazou V, Schirmer EC. Organelle proteome variation among different cell types: lessons from nuclear membrane proteins. Subcell Biochem 2007;43:51–76.
[142] Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, et al. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 2006;125:173–86.
[143] Mootha VK, Bunkenborg J, Olsen JV, Hjerrild M, Wisniewski JR, Stahl E, et al. Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. Cell 2003;115:629–40.
[144] Tunnah D, Sewry CA, Vaux D, Schirmer EC, Morris GE. The apparent absence of lamin B1 and emerin in many tissue nuclei is due to epitope masking. J Mol Histol 2005;36:337–44.
[145] Wilkie GS, Schirmer EC. Guilt by association: the nuclear envelope proteome and disease. Mol Cell Proteomics 2006;5:1865–75.
[146] Olsson M, Scheele S, Ekblom P. Limited expression of nuclear pore membrane glycoprotein 210 in cell lines and tissues suggests cell-type specific nuclear pores in metazoans. Exp Cell Res 2004;292:359–70.
[147] Cohen M, Feinstein N, Wilson KL, Gruenbaum Y. Nuclear pore protein gp210 is essential for viability in HeLa cells and Caenorhabditis elegans. Mol Biol Cell 2003;14:4230–7.
[148] Stavru F, Hulsmann BB, Spang A, Hartmann E, Cordes VC, Gorlich D. NDC1: a crucial membrane-integral nucleoporin of metazoan nuclear pore complexes. J Cell Biol 2006;173:509–19.
[149] Stavru F, Nautrup-Pedersen G, Cordes VC, Gorlich D. Nuclear pore complex assembly and maintenance in POM121- and gp210-deficient cells. J Cell Biol 2006;173:477–83.
[150] Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, et al. Determining the architectures of macromolecular assemblies. Nature 2007;450:683–94.
[151] Wente SR, Rout MP, Blobel G. A new family of yeast nuclear pore complex proteins. J Cell Biol 1992;119:705–23.
[152] Denning DP, Patel SS, Uversky V, Fink AL, Rexach M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. Proc Natl Acad Sci U S A 2003;100:2450–5.
[153] Lim RY, Huang NP, Koser J, Deng J, Lau KH, Schwarz-Herion K, et al. Flexible phenylalanine–glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. Proc Natl Acad Sci U S A 2006;103:9512–7.
[154] Lim RY, Fahrenkrog B, Koser J, Schwarz-Herion K, Deng J, Aebi U. Nanomechanical basis of selective gating by the nuclear pore complex. Science 2007;318:640–3.
[155] Frey S, Richter RP, Gorlich D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science 2006;314:815–7.
[156] Rexach M, Blobel G. Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. Cell 1995;83:683–92.
[157] Strasser K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, et al. TREX is a conserved complex coupling transcription with messenger RNA export. Nature 2002;417:304–8.
[158] Strasser K, Hurt E. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. Nature 2001;413:648–52.

Cell. Mol. Life Sci. (2010) 67:1353–1369
DOI 10.1007/s00018-010-0257-2

Cellular and Molecular Life Sciences

RESEARCH ARTICLE

Cell-specific and lamin-dependent targeting of novel transmembrane proteins in the nuclear envelope

Poonam Malik · Nadia Korfali · Vlastimil Srsen · Vassiliki Lazou · Dzmitry G. Batrakou · Nikolaj Zuleger · Deirdre M. Kavanagh · Gavin S. Wilkie · Martin W. Goldberg · Eric C. Schirmer

Received: 2 November 2009 / Revised: 21 December 2009 / Accepted: 5 January 2010 / Published online: 21 January 2010

Abstract Nuclear envelope complexity is expanding with respect to identification of protein components. Here we test the validity of proteomics results that identified 67 novel predicted nuclear envelope transmembrane proteins (NETs) from liver by directly comparing 30 as tagged fusions using targeting assays. This confirmed 21 as NETs, but 4 only targeted in certain cell types, underscoring the complexity of interactions that tether NETs to the nuclear envelope. Four NETs accumulated at the nuclear rim in normal fibroblasts but not in fibroblasts lacking lamin A, suggesting involvement of lamin A in tethering them in the nucleus. However, intriguingly, for the NETs tested alternative mechanisms for nuclear envelope retention could be found in Jurkat cells that normally lack lamin A. This study expands by a factor of three the number of liver NETs analyzed, bringing the total confirmed to 31, and shows that several have multiple mechanisms for nuclear envelope retention.

Keywords Inner nuclear membrane · Outer nuclear membrane · Nuclear lamina · Integral membrane protein · Cell-type specificity

P. Malik · N. Korfali · V. Srsen · V. Lazou · D. G. Batrakou · N. Zuleger · D. M. Kavanagh · G. S. Wilkie · E. C. Schirmer (✉)
Wellcome Trust Centre for Cell Biology, University of Edinburgh, Kings Buildings, Swann 5.22, Mayfield Road, Edinburgh EH9 3JR, UK
e-mail: e.schirmer@ed.ac.uk

M. W. Goldberg
School of Biological and Biomedical Sciences, Durham University, Durham, UK

Introduction

Several proteins of the nuclear envelope (NE) are linked to human diseases including muscular dystrophies, neuropathy, and progeroid aging syndromes [1, 2]. Nuclear envelope proteins associated with disease include the intermediate filament A/C lamins and several NE transmembrane proteins (NETs). Curiously, different variants of Emery-Dreifuss muscular dystrophy are caused by mutations in the *LMNA* gene, encoding lamins A and C [3] and NETs that interact with A/C lamins [4–7]. This may be a clue to explain how mutations in the widely expressed *LMNA* gene could lead to distinct diseases that each yield pathology in only a small subset of the tissues in which A/C lamins are expressed, e.g., specificity is conferred by different combinations of partner proteins in different cell types.

Already, a wide range of binding partners has been reported for both lamins and NETs [8, 9]. These proteins are likely to vary among cell types because different combinations of lamin and emerin antibodies stained different tissues, e.g., two of three lamin B1 antibodies stained human cardiomyocyte nuclei, whereas a different set of two stained hippocampal neurons [10]. Thus, different regions of the proteins might be occupied by binding partners in each tissue. It follows that as yet unidentified partners of lamins and NETs in tissue-specific complexes may mediate the phenotypes of the wide range of lamin-related diseases. Indeed, the three favored molecular mechanisms to explain NE disease pathology—mechanical instability from disruption of lamina-cytoskeletal interactions, altered expression of genes regulated from the nuclear periphery, and disabling of the cell cycle/stem cell maintenance [1, 11]—all likely involve additional associated proteins to produce pathology. Because both gene regulation and cytoskeletal connections have been

implicated, NE proteins involved could reside in either the inner nuclear membrane (INM) or outer nuclear membrane (ONM).

A proteomic study of liver NEs recently increased the number of putative NETs by fivefold [12]; however, some of these could be erroneous assignments or contaminants of the fractions, and so it is necessary to directly test them for NE localization. We sought to test the validity of the proteomic datasets by confirming the targeting of these putative NETs to the NE and also to gauge whether lamin interactions are likely to contribute to this targeting/retention. Work from several laboratories has so far confirmed only 13 of the 67 new putative NETs for NE targeting [12–15]. Here we investigate the targeting of 30 putative NETs, bringing the total characterized to 40.

Our results classify only 70% of those tested as NETs by their producing a distinctive 'rim' staining around the nucleus. However, the remaining 30% do not necessarily represent misidentifications in the proteomic analysis because some only targeted to the nuclear rim in certain cell types, likely reflecting the complexity of cell types found in liver and underscoring the potential for error in overuse of tissue culture systems in studying the NE. The majority of confirmed NETs targeted to the INM with only a few residing only in the ONM as determined using high-resolution structured illumination microscopy. Furthermore, most resisted a pre-fixation extraction with detergent—typically indicating association with the lamin polymer—yet only 4 out of 12 NETs tested targeted less to the NE in fibroblast cells deleted for lamin A, indicating that other lamins or lamina-associated NETs suffice for their NE retention. Surprisingly, among those that mistargeted in the absence of lamin A, those for which we had antibodies did not require lamin A for association with the NE in Jurkat cells that never had lamin A. This finding is important as it may explain in part how lamin A-interacting proteins could be involved in diseases where pathology is only observed in a subset of tissues: they have distinct mechanisms for NE retention in different cell types. This study gives a better view of NE composition and its potential functions, and indicates how its variability could contribute to the tissue specificity of NE diseases.

Methods and materials

Plasmid construction

IMAGE clones for human NETs were obtained from RZPD and Geneservice. NET numbers followed by IMAGE numbers or gene IDs in parentheses are listed: NETs 5 (199953-gene ID), 15 (5270233), 11 (4798194), 13 (6023304), 14 (3640219), 16 (5267120), 17 (4812681), 20 (3872837), 21 (84135-gene ID), 23 (5762441), 24 (4907240), 25 (5240212), 29 (6201334), 30 (4299899), 32 (4248728), 33 (4138639), 34 (4865469), 35 (3451350), 36 (4819093), 37 (30341915), 38 (4698763), 43 (5166101), 44 (4577143), 45 (3462452), 46 (5189722), 47 (4214662), 48 (3355282), 49 (3354945), 50 (3344010), 55 (4720647), 59 (3959506), and 62 (6052380). The only available IMAGE clone for NET5 was shorter than the original predicted hypothetical orf, but is likely to be a shorter splice variant. Coding sequences were amplified by PCR with added 5′ and 3′ restriction sites, sequenced from both ends in intermediate cloning vectors, and then inserted into mammalian expression vector pHHS10B that carries an amino-terminal HA epitope tag and/or pmRFP with a carboxyl-terminal tag (derived from Clontech pEGFP-N2 by replacing the GFP coding sequence with that of monomeric red fluorescent protein). Additionally, the coding sequences of previously characterized NETs 4, 31, 39, and 51 were moved from their original HA-tagged vector [12] to pmRFP and GFP vectors. Lamin A-GFP was obtained from Anne Straube (MCRI, Oxted, UK) and Calreticulin-GFP and RapM4-CFP from Tom Rapoport (Harvard, Boston).

Cell culture and transfections

Cell lines derived from human fibroblasts (HT1080), human embryonic kidney cells (293T), mouse myoblasts (C2C12), human myoblasts (RD), and human hepatocytes (HepG2) were maintained in high glucose DMEM (Lonza) supplemented with 10% fetal bovine serum (FBS), 100 μg/μl penicillin, and 100 μg/μl streptomycin sulfate. To this medium were added MEM non-essential amino acids and 1 mM sodium pyruvate for several mouse fibroblast cells: NIH3T3 and 3T3-L1 cell lines, and 216−/− lamin A knockout primary MEFs and their matched wild-type primary cells. Jurkat cells, a human lymphocyte cell line, were cultured in RPMI with 10% FBS.

Cells were plated on coverslips at ∼10% confluency to prevent their reaching confluency before fixation at 30 h post-transfection. DNA was transfected 12 h after plating using FuGENE 6 (Roche) according to the manufacturer's instructions.

Antibodies

Antibodies were: HA tag (mAb HA.11, Covance), lamin A and B1 (3262 and 3931; [16]), Nup153 (Covance), Nup358 (raised against recombinant human protein aa 2595–2881, kind gift of F. Melchior), Calreticulin (2891S, Cell Signaling), Calnexin (SPA-860, Stressgen), or NETs: rabbit anti-peptide polyclonal generated for this study by Millipore NET4 (11780), NET23 (11815), NET29 (11796), NET30 (11827), NET31 (11830), NET33 (11835), NET34

(11231), NET39 (11668), NET50 (11850), NET51 (11856), NET55 (11862), NET59 (11866), and SUN2 (11905). All fluorophore-conjugated secondary antibodies were minimal cross-reactivity from donkey (Jackson ImmunoResearch) or goat (MolecularProbes).

Quantitative Western blotting

Liver NE and microsomal membranes (MM) were prepared as previously described [17]. NE and MM were lysed in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 2 mM $MgCl_2$, 0.2% NP-40 in the presence of protease inhibitor cocktail (Roche 11 873 580 001) by heating at 65°C for 2 min and soncation in a sonibath at 4°C. Protein concentrations were determined using the Bradford Method (BioRad). An equal volume of protein sample buffer (100 mM Tris pH 6.8, 4 M Urea, 2% SDS, 50 mM DTT, and 15% sucrose) was added, and the samples were boiled at 100°C for 5 min. Equal amounts of protein were resolved by SDS-PAGE and transferred to nitrocellulose membrane (LI-COR Biosciences). Membranes were blocked in PBS, 5% milk, 0.2% tween. Primary antibodies were diluted in this buffer (1/200 Millipore NET peptide antibodies, 1/500 calreticulin, 1/200 Calnexin, 1/2000 lamin A) and allowed to incubate overnight at 4°C. Secondary antibodies IR800 conjugated goat anti-rabbit (LI-COR Biosciences) were added at a concentration 1/5,000 at RT for 2 h. Visualization and quantification were performed using a LI-COR Odyssey and software (Odyssey 3.0.16) using median background subtraction. A minimum of three independent blots were run for each NET and control. The averages from all three are presented in Fig. 3.

Microscopy

Cells were either directly fixed 7 min in 3.7% formaldehyde or washed with PBS, then extracted for 1 min with 1% Triton X-100, 25 mM Tris pH 8.0, 150 mM KOAc, 15 mM NaCl, and 5 mM $MgCl_2$, washed again with PBS, and then fixed with formaldehyde. In some cases cells were instead extracted with 2 × 1 min incubations with 0.1% Triton X-100 in PBS, which better maintained cell morphology but removed more cells from coverslips. For antibody staining, cells that were not pre-extracted were permeabilized 6 min in 0.2% Triton X-100 after fixation. Cells were then blocked with 10% FBS, 200 mM glycine in PBS, and reacted 40 min at RT with antibodies. All fluorophore-conjugated secondary antibodies were minimal cross-reactivity from donkey (Jackson ImmunoResearch) or goat (MolecularProbes). DNA was visualized with DAPI (4,6-diamidino-2 phenylindole, dihydrochloride) and coverslips mounted in fluoromount G (EM Sciences). Some cells were costained with ER membrane dyes DiOC6 and R_6 (MolecularProbes).

For cryosections, rat liver was cut into 2–3-mm cubes, embedded in OCT (Tissue-Tek), snap-frozen in liquid nitrogen, and maintained at −80°C. Sections were cut on a Leica CM 1900 Cryostat at 6–8-μm thickness and immediately fixed in −20°C methanol. After rehydration with PBS, sections were incubated with anti-NET antibodies overnight at 4°C followed by secondary antibodies as above.

Most images were obtained using a Nikon TE-2000 microscope equipped with a 1.45 NA 100× objective, Sedat quad filter set, PIFOC Z-axis focus drive (Physik Instruments), and CoolSnapHQ High Speed Monochrome CCD camera (Photometrics) run by IPLab image acquisition software. Image stacks (0.2 μm steps) were deconvolved using AutoquantX. Structured illumination images (Fig. 5b) were taken on the OMX system at the University of Dundee microscopy facility (details described at http://microscopy.lifesci.dundee.ac.uk/omx/). High-resolution images to distinguish ONM from INM were also taken using a Deltavision (Applied Precision) microscope system with 100× 1.4 NA objective, and 0.2-μm stacks were deconvolved using DeconQ; images were processed using SoftWorks (Fig. 5c, d). Cryosection images were recorded using an SP5 confocal system (Leica) with 63× oil 1.4 NA objective using argon and UV lasers. Micrographs were saved from source programs as TIF files and prepared for figures using Photoshop 8.0.

Immuno-EM

Immuno-electron microscopy was performed on HeLa cells transiently transfected with NET51, NET55, emerin, or SUN2 all fused to GFP. Cells were fixed in 4% paraformaldehyde, pelletted and infiltrated with 2.3 M sucrose, then frozen by plunging into liquid nitrogen. Frozen pellets were sectioned on a cryo-ultramicrotome (Leica, UC6 with FC6 cryo-attachment). Cryosections were thawed, rinsed in PBS with 1% glycine, incubated in PBS with 1% BSA, incubated with rabbit anti-GFP antibody (Abcam, Cambridge, MA) at 1:400 dilution, rinsed in PBS then incubated with the secondary anti-rabbit IgG antibody conjugated to 5-nm colloidal gold (Agar Scientific, UK). Grids were then rinsed in PBS, transferred to 1% glutaraldehyde (Agar Scientific, UK) in PBS, washed in water. and embedded in 2% methyl cellulose containing 0.4% uranyl acetate (Agar Scientific, UK). Images were taken on a Hitachi H7600 electron microscope at 100 kV and 80,000×–100,000× magnification.

Transmembrane prediction

Sequences were analyzed for transmembrane helices using TMPred 'http://www.ch.embnet.org/software/TMPRED_form.html' [18] or TMHMM 2.0 'http://www.cbs.dtu.dk/services/TMHMM-2.0/' [19]. Transmembrane

helices predicted by TMPred were sometimes only weakly predicted by TMHMM; in these cases the percent probability is listed.

RT-PCR

Cells were lysed with Tri-Reagent (Sigma) and total RNA extracted according to the manufacturer's instructions. Reactions were carried out with 100 ng of total RNA using the Titan one-tube RT-PCR system (Roche) according to the manufacturer's instructions, except that the dNTP concentration was increased to 500 μM and $MgCl_2$ increased to 3 mM. Typical reaction conditions were 30 min reverse transcription at 50°C, 2 min denaturation at 94°C, then 24 cycles of 94°C for 30 s, 60°C for 30 s, and 68°C for 45 s. Human primer sets used were: NET11 5′-CTGAAGCTGGGAAGACCAAC-3′ and 5′-AATGCTC AACCCCTCATGTC-3′; NET13 5′-CTCTCATGGCTGG GCTTTAG-3′ and 5′-GAGGTGGTAGCGACAGAAGC-3′; NET32 5′-ATTCAAGCTGTGCGGGTAAC-3′ and 5′-TC TTGCTGTTGGAAGCAATG-3′; NET45 5′-TGCTGGTT TCATAGGGAAGG-3′ and 5′-TGGTCGAGCATGAGTT TCAC-3′; NET59 5′-ACCTGGACCACACAGACTCC-3′ and 5′-ATCCTCGTGTTACGGGTCAG-3′; emerin 5′-CTT TCGGATACCGAGCTGAC-3′ and 5′-CGTTCCCTATCC TTGCACTC-3′; Ppia 5′-CACCGTGTTCTTCGACATTG-3′ and 5′-TCGAGTTGTCCACAGTCAGC-3′.

Peptidylprolyl isomerase A (Ppia) was used as a loading control, and reactions were repeated at least three times.

In vitro extraction assay

Jurkat clone E6-1 (freshly obtained from ATCC) was nucleofected (Kit V, C-016) with lamin A-GFP in pCDNA3.1 linearized with *Bgl* II. Clones expressing GFP-lamin A stably integrated in the genome were selected with 1 mg/ml G418 and then further enriched by fluorescence activated cell sorting (FACS). Lamin A-GFP Jurkats, wild-type Jukat cells that lack lamin A, normal wild-type mouse fibroblasts (NIH 3T3), and lamin A-null mouse fibroblasts (216−/−) were recovered from dishes and divided equally into two tubes. One tube (non-extracted) was directly lysed by first adding 1% Triton X-100 and 8 M urea followed by SDS-sample buffer and sonication in a sonibath. The second tube (extracted) was extracted by adding 1% Triton X-100, 50 mM NaCl, 300 mM KOAc, 20 mM HEPES pH 8, 2 mM $MgCl_2$, and 8% sucrose with rapid inversion for 2 min followed by centrifugation for 15 min at 13,000 rpm. The pellet was resuspended in SDS-sample buffer containing 8 M urea. Lysates were loaded based on cell number onto Western blots, probed with NET and lamin antibodies, and bands quantified on a LI-COR Odyssey.

Results

Thirteen novel proteins were confirmed as NETs by their resistance to a pre-fixation detergent extraction in cells

A proteomic study of rodent liver NEs identified 67 new putative NETs, 8 of which were confirmed by targeting to the NE when exogenously expressed [12]. To further test the validity of this dataset and gain a more comprehensive view of the human NE proteome, cDNAs were obtained for the 32 putative NETs that were available in the human IMAGE clone collection, in addition to the eight first characterized. The coding sequences were placed into a vector for fusing their C-termini to mRFP in most cases and/or a vector for fusing their N-termini to a Heamagglutinin (HA) Tag. These were transiently transfected into HT1080 human fibrosarcoma cells stably expressing lamin A-GFP, and after 30 h cells were either directly fixed or first extracted with detergent and then fixed. For NET15 and NET48, no transfected cells were detected in the cell lines tested, although sequencing suggested the constructs were correct. This could reflect instability of the fusion proteins or indicate that these NETs require very specific backgrounds for stable expression. Binding to the intermediate filament lamin polymer is thought to drive INM accumulation of NETs [20]. Thus, as the nuclear lamina is resistant to pre-extraction with detergent (up to 1% Triton X-100) and high salt (up to 1 M NaCl), similar biochemical resistance to extraction of NETs typically indicates INM association. Some proteins, however, may be in the ONM and also resist extraction because of interactions with cytoskeletal components (e.g., nesprins; [21]).

After detergent extraction, the emerin control was clearly retained at the NE, whereas both ER membrane (stained with DiOC6) and overexpressed ER proteins calreticulin and RapM4 were completely removed by the same treatment (Fig. 1 and data not shown). Thirteen of the putative NETs tested were retained at the NE similarly to emerin as affirmed by colocalization with Lamin A, thus defining them as bona fide NETs (Fig. 1a). Untransfected adjacent cells in some of the images show that there was no bleedthrough of lamin signal (green) into the NET channel (red). Nuclei were sometimes misshapen as a result of the detergent treatment, but the colocalization with lamins (yellow merge) confirms that some of the protein was retained at the NE. All NETs shown in Fig. 1 resisted detergent extraction in each of multiple experiments, except for NETs 34 and 50 that resisted extraction in three of five experiments using the same HT1080 cells. It is possible that these NETs are tethered more weakly, similar to some LAP1 isoforms

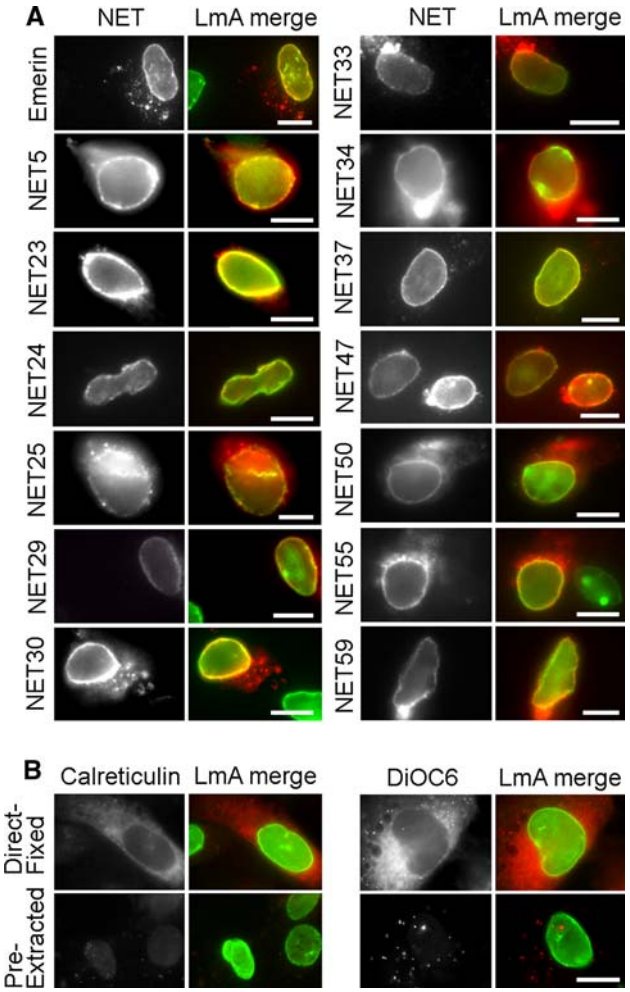


Fig. 1 NE localization and detergent resistance of novel NETs. (**a**) NETs fused to mRFP at their carboxy-termini (except NET25 fused at its amino-terminus to an HA epitope tag) were transiently expressed in HT1080 cells also expressing lamin A-GFP. Cells were pre-extracted with Triton X-100 prior to fixation to remove membranes and soluble proteins, which typically also distorts morphology. NETs alone (*left*) and the merge between the NET (*red*) and lamin A (*green*) are shown. In some panels adjacent untransfected cells are shown, confirming that NE signal is not due to bleedthrough from the lamin channel. Like the emerin control (*top*), all NETs shown resisted detergent pre-extraction. Such resistance typically indicates association with the lamin polymer. *Scale bars* 7.5 μm. (**b**) Controls showing that the ER was fully removed by the detergent pre-extraction. *Left* calreticulin fused to GFP (but colored *red*) was overexpressed in cells either directly fixed or pre-extracted with detergent. No colocalization with lamins (*green*) was observed, and no calreticulin remained after extraction, though lamins did remain. Both direct fixed and pre-extracted images were exposed for 2 s. *Right* untransfected cells were similarly treated, then stained with the ER lipid dye DiOC6. Endogenous ER staining did not exhibit notable accumulation at the NE and was completely removed by the pre-extraction. Both images were exposed for 500 ms. *Scale bars* 20 μm

that were previously found to extract at lower salt concentrations than others [22]. This may indicate a hierarchy in NET binding strength, but could also be a simple consequence of a lower relative abundance of binding partners to tether these NETs in the NE.

Some putative NETs accumulated at the NE but did not resist pre-extraction with detergent

As the INM is only half of the NE, the NE proteomic datasets likely include many ONM proteins that would be less likely to resist detergent pre-extraction. Emerin has in fact been shown to have both INM and ONM populations [23]. In cells that were not pre-extracted, both emerin and the NETs confirmed in Fig. 1 also accumulated in the ER, yet distinct nuclear rim accumulation and colocalization with lamins could still be observed. This is in contrast to cells overexpressing calreticulin and RapM4 where a distinct rim was not observed (Fig. 2, top and data not shown). Therefore, putative NETs that did not resist pre-extraction with detergent were compared in directly fixed cells to emerin and calreticulin controls, revealing four (20, 38, 46, and 62) that yielded a strong and distinctive rim against the background of the overexpressed protein in the ER (Fig. 2). The remaining 13 NETs out of the 30 analyzed either yielded no distinctive rim against the ER background or different cellular targeting in the HT1080 cells.

Antibody staining of endogenous NETs confirms that their primary residence is in the NE

To test if the accumulation in both the NE and ER with overexpressed proteins reflects the targeting of the endogenous NETs, peptide antibodies were generated to several NETs, of which nine recognized bands of the expected size for various splice forms on Western blots (data not shown). Equivalent amounts of total protein from NE and microsome preparations (as determined by Bradford protein assay) were compared for NET levels by quantitative Western blotting, and the percentage of the total signal in NE or microsomal fractions was plotted (Fig. 3). Most NETs had nearly all signal in the NE fraction, and the ER controls calnexin and calreticulin were mostly in the ER fraction. Another study compared NE and microsome factions based on the assumption that calnexin is similarly in the ONM as the ER instead of by equal protein loading [14]: we also see calnexin in the NE fractions when we overload NEs, but this means that our NETs would be even more enriched in the NE using their method. NET23 and NET33 had roughly a third of the signal in the microsome fraction, and NET34 had nearly all signal in the microsome fraction. Nonetheless, NET34 also resisted the pre-fixation detergent extraction, likely indicating that some small pool of NET34 resides in the INM.

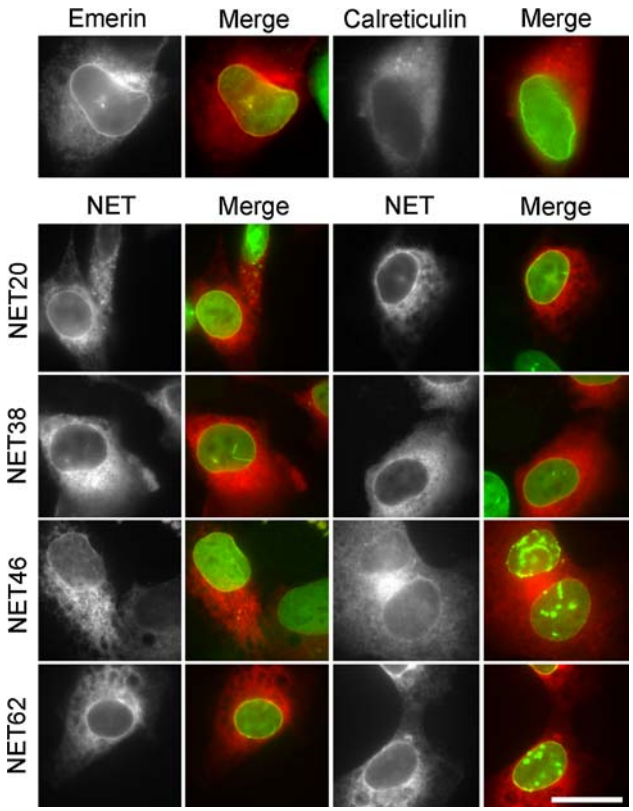


Fig. 2 NETs that did not resist detergent pre-extraction re-tested for NE accumulation in directly fixed cells. HA-tagged NET62 and mRFP-tagged NETs 20, 38, and 46 were transiently transfected into HT1080 cells expressing lamin A-GFP. Tagged emerin and calreticulin were separately overexpressed as controls. The NET alone (*black* and *white*) and the merge between the NET (*red*) and lamin A (*green*) are shown. The new NETs yielded clear nuclear rims against the high cytoplasmic accumulation, colocalizing with lamins (*yellow* merge) similarly to emerin (*top left*). This staining pattern clearly differs from the ER localization of calreticulin (*top right*). *Scale bars* 20 μm

Rat liver cryosections were also stained with several of the NET antibodies. These antibodies yielded nuclear rim staining in multiple cells within any given field (Fig. 4). In some cases very little staining was observed outside the nuclear rim in the tissue sections, whereas in others some other distinctive areas of cells were stained, particularly the ER, consistent with the partial localization in microsomes in Fig. 3. This is similar to the case of the well-characterized NET emerin, which has been shown to also stain in the cytoplasm of myotubes and interstitial discs in heart tissue [24, 25].

Most of the NETs target to the INM

The wide range of NET characteristics observed here highlighted the need to better distinguish their localization between INM and ONM. It was also considered that NETs that failed to resist the pre-fixation extraction with Triton

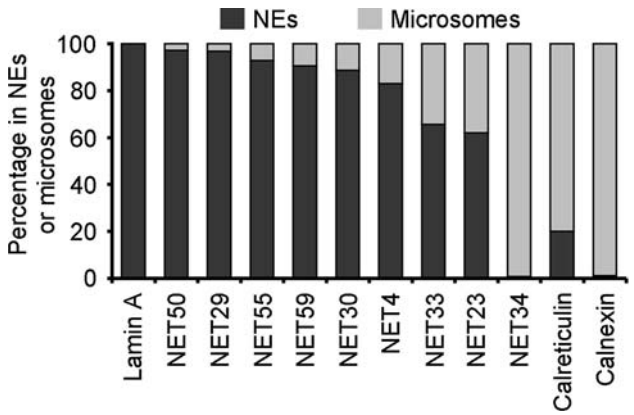


Fig. 3 Relative amounts of NETs in the NE vs ER. NEs and microsomal membranes representing ER were prepared, and an equal amount of total protein for each fraction loaded on SDS-PAGE based upon measurements with Bradford assay. The proteins were transferred to nitrocellulose membranes, reacted with NET antibodies and fluorescent secondary antibodies, and fluorescent signals quantified using a LI-COR Odyssey. The averaged results of three separate Western blots are plotted showing the percentage of the combined signal coming from NE and microsome lanes. Most NETs were principally in the NE fraction, but NET34 was principally in the microsome fraction

X-100 might preferentially localize to the ONM. A recent study was able to distinguish between INM and ONM localization by comparing staining for lamin B1 with nucleoporins from the nuclear basket (Nup153) and cyoplasmic filaments (Nup358) of the NPC using 3D structured illumination microscopy (OMX; [26]). If a protein localizes to the INM, it should co-localize in the same plane as Nup153, but yield a separable more internal staining compared to Nup358 (Fig. 5a). This same system was directly applied to the well-characterized NET LAP2ß fused to a fluorescent protein tag as a control and to several similarly fused NETs (Fig. 5b). For all but NET23 and the ER controls of those tested, the NET (red) and Nup153 (green) spots were observed in the plane of the INM, and an internal NET (red) ring was observed compared to Nup358 (green). The appearance of alternating spots as opposed to co-localization further supports the quality of the imaging as NPCs are positioned at perforations in the nuclear membrane, whereas the NETs would be in the plane of the inner membrane and so should not actually co-localize. Although differing levels of the exogenously expressed NETs had accumulated in the ER and so reduced the clarity of the planar separation, the inner ring was still clearly distinguishible.

Nup153 and Nup358 were also clearly distinguishable from one another when stained in the same cell with deconvolved images generated using a Deltavision microscope system (Fig. 5c). In this case Nup153 is colored red and Nup358 green to compare the two Nups in the same cell. Remaining NETs were then tested using this

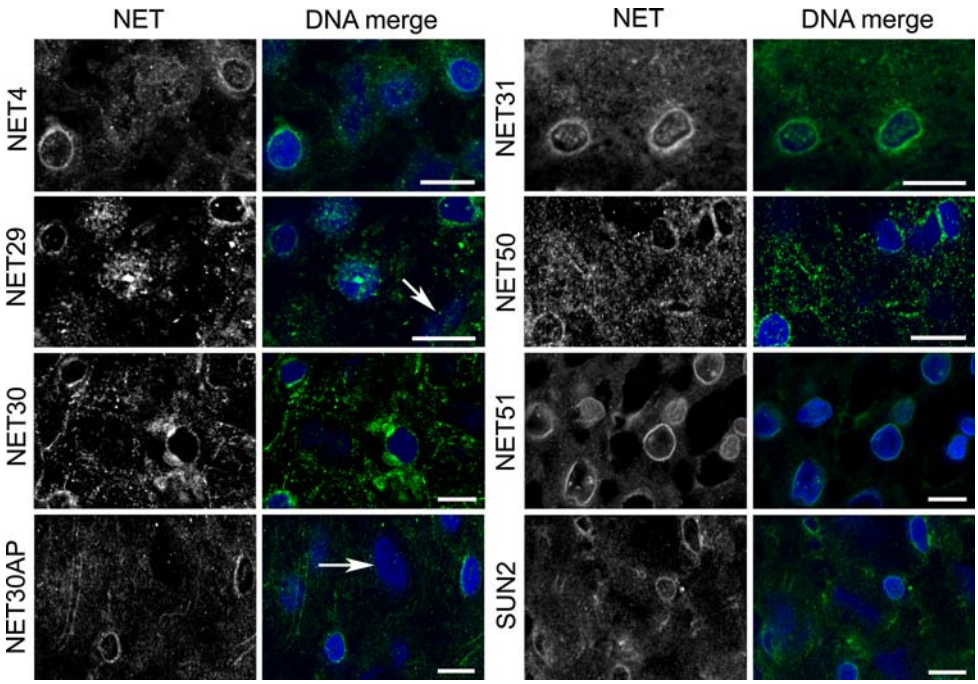


Fig. 4 Antibody staining in liver tissue sections. Cryosections of rat liver were stained with various NET antibodies and imaged on a confocal microscope. Nuclear rim staining could be observed in multiple cells in all fields, though for NET29 and NET30 not all cells in a given field yielded nuclear rims (cells indicated by *arrows*, as determined by comparing DAPI staining for mid-sectioned nuclei). Some background staining is observed with all NETs in the cytoplasm: this was slightly diminished with use of an affinity-purified antibody for NET30 (NET30AP), but may also indicate multiple cellular localizations for NETs. *Scale bars* 10 μm

system. LAP2ß, calreticulin, and NET55 were tested again (in red), revealing that INM and ONM could be similarly discriminated, though with slightly less clarity in resolution. Most of the remaining NETs tested similarly yielded alternating red and green spots with Nup153 in the plane of the inner ring and an inner NET ring compared to Nup358 (Fig. 5d). However, NET4, NET24, and NET31 joined NET23 in appearing only in the ONM.

To confirm the validity of INM localization determinations using this system, two NETs and controls were tested by immunogold-EM with anti-GFP antibodies for localization to the INM. The secondary antibody-conjugated gold particles were observed at the INM for both controls emerin and SUN2, and additionally gold particles were observed in both the ER and ONM, consistent with expected accumulation in the ER when saturating binding sites in the INM when overexpressed (Fig. 5e). The INM could be readily distinguished from the ONM and ER because of the denser appearance of chromatin and absence of additional membrane systems on one side of the double membrane and the lighter staining and appearance of ER and mitochondrial membranes on the other. NET51 and NET55 yielded similar patterns of gold particle distribution with NET55 yielding an even higher proportion of gold particles in the INM than the well-characterized INM proteins SUN2 and emerin.

Some NETs only target in specific cell types

Intriguingly, in the liver tissue sections nuclear rim accumulation was only observed in a subset of the cells within any given field for NETs 29 and 30, as can be distinguished by DAPI-stained nuclei sectioned in mid-plane that lack nuclear rim staining with the antibody (Fig. 4). This suggested that the failure of some NETs to target to the NE in the HT1080 cells used in tagged NET expression experiments might reflect the specific cell type used; i.e., they might target in other cell types. Moreover, antibodies to two NETs (NET29 and NET39) did not stain some cell lines tested, but did others (data not shown). These two NETs are expressed in a limited set of tissues according to a large-scale transcriptome study [27]. Therefore, several NETs that had not targeted in HT1080 cells were tested in other cell lines.

NET45 in particular might be expected to have restricted targeting because, according to transcriptome data [27], it is expressed higher in liver than any other tissue. In the HT1080 fibrosarcoma cells, NET45-mRFP was extractible and accumulated too strongly in the cytoplasm for a firm conclusion on NE targeting to be made, whereas in HepG2 cells that are derived from liver a distinct rim staining was obtained (Fig. 6a). Thus, the liver cells might have partners that tether NET45 in the NE that are absent in the HT1080 cells.

Thus, the NETs that failed to target in HT1080 cells were retested in a variety of cell types: HepG2 (hepatocyte), 293T (kidney), C2C12 (muscle), and 3T3-L1 (pre-adipocyte) cell lines (Fig. 6 and Table 1). Controls emerin and NET59 targeted to the NE in all five lines. NET32, previously reported to target to the NE and resist pre-extraction in C2C12 cells [14], also targeted in 293T cells, but not in HT1080, HepG2, or 3T3-L1 cells.

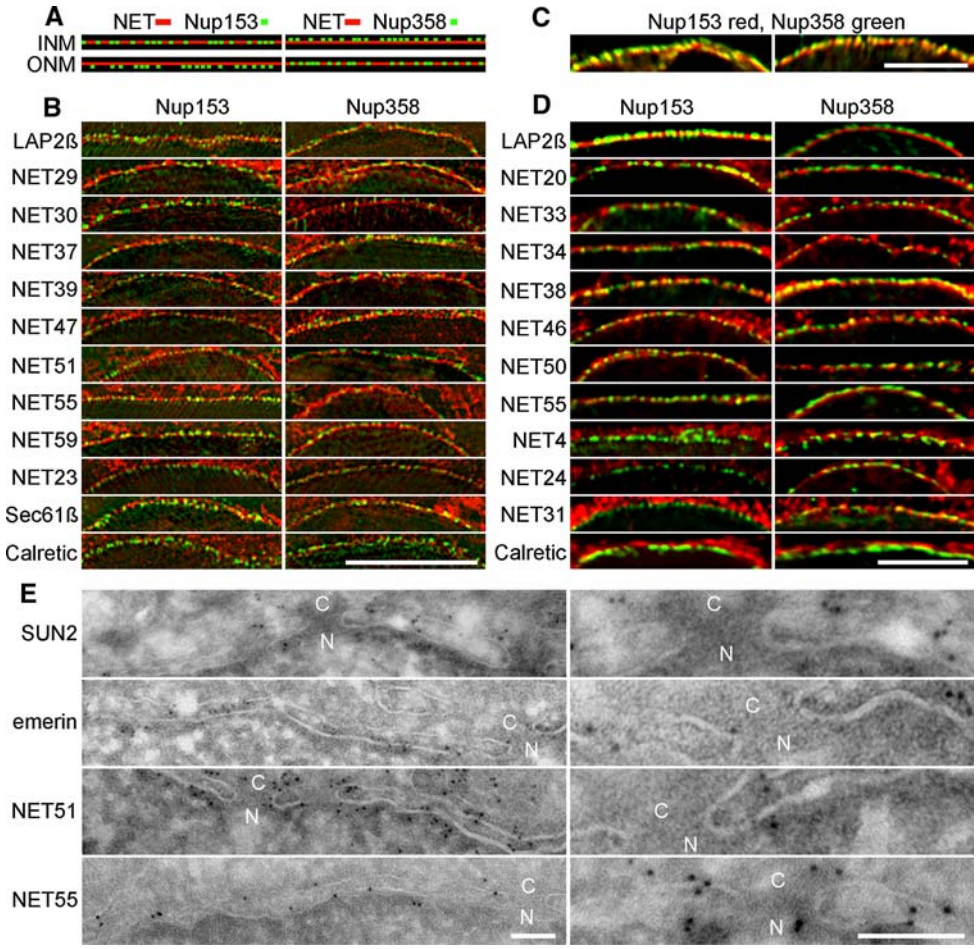
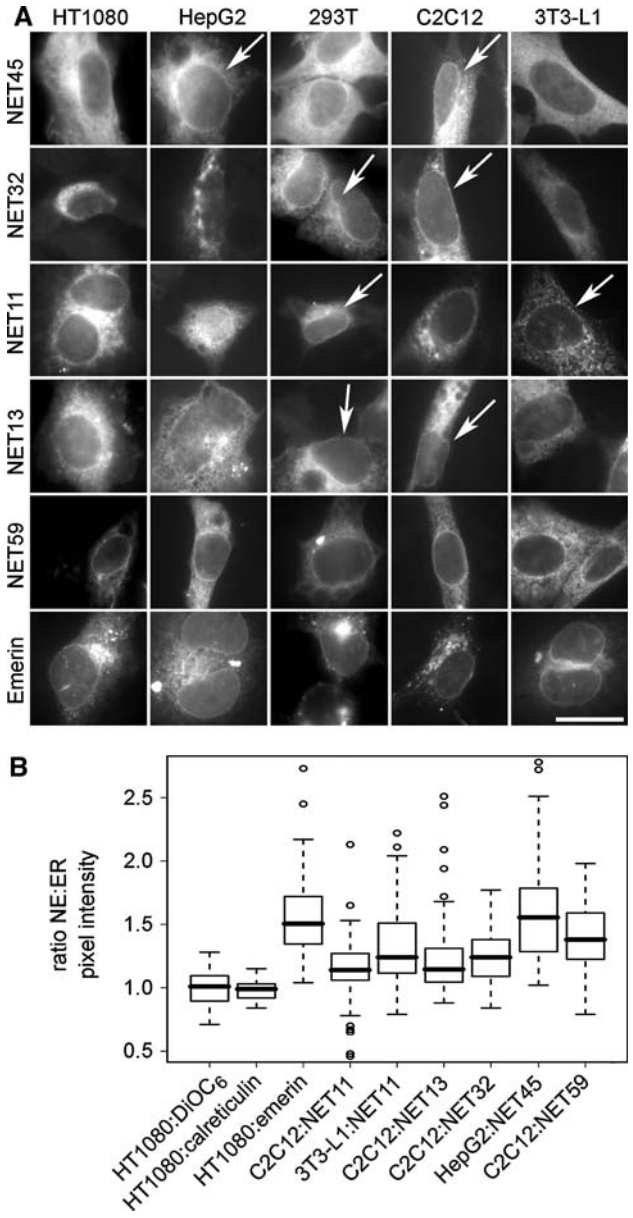


Fig. 5 INM versus ONM targeting. NETs were imaged using high-resolution systems for localization in relation to the two sides of the NPC using Nup153 from the nuclear basket and Nup358 from the cytoplasmic filaments. (**a**) Schematic of expected patterns indicating INM or ONM localization. If a protein is in the INM, the NET and Nup153 signals should occur in the same plane, and the NET signal should appear internal to the Nup358 signal. If the protein is in the ONM, the Nup153 signal should be internal to the NET signal and the Nup358 signal should occur in the same plane as the NET signal. (**b**) Images using structured illumination show both characterized (LAP2ß) and many novel NETs in the same plane of the inner nuclear membrane with Nup153 and internal to Nup358. Only NET23 and controls Sec61ß and calreticulin yielded the pattern expected for ONM residence. (**c**) High-resolution deconvolved Deltavision images also can distinguish inner from outer nuclear membranes with Nup153 shown in *red* and Nup358 shown in *green*. (**d**) Many additional novel NETs appeared in the INM using again LAP2ß as a control and NET55 that had been separately tested with the OMX system. In contrast NETs 4, 24, and 31 together with the calreticulin control yielded ONM targeting. *Scale bars* for **b**-**d** 5 µm. (**e**) Immunogold-EM confirms the validity of OMX and Deltavision results as 5-nm gold particles recognizing GFP antibodies for expressed NET51 and NET55 proteins appeared in the INM, similarly to controls. C and N denote cytoplasmic and nucleoplasmic sides where NPCs are inserted in the membrane. *Bars* 100 nm

NET11 and NET13 also targeted to the NE in a subset of the cell lines (Fig. 6). To support this visual readout, image pixel intensities were measured and a ratio of NE over ER values generated (Fig. 6b). For an ER dye and overexpressed calreticulin protein, the ratio was very close to 1, indicating no enrichment at the NE. However, the control emerin and novel NETs all had values between $\sim$1.3 and 1.5 in the cells where they targeted, and the p values indicated extremely high confidence that these numbers were statistically significant compared to the membrane and calreticulin controls (Table 2). It was not possible to test if they resisted pre-extraction in the cells where they targeted because transfection efficiencies were low, and many cells are washed away in the extraction procedure.

The cell-type specific targeting of exogenously expressed NETs suggested that endogenous expression levels of NETs might be prognostic for their NE targeting. Indeed, slightly lower expression levels were observed by RT-PCR in the HT1080 cells for NET11, NET13, and NET45, and higher levels of NET45 were observed in the HepG2 cells where it targeted best (Fig. 7). However, over the range of cell lines tested there was no correlation between expression levels and targeting.



◀ **Fig. 6** Some NETs only accumulate at the NE in certain cell types. (**a**) NETs that failed to target to the NE in HT1080 (human fibrosarcoma) cells were re-tested in other cell lines derived from different tissues: HepG2 (human liver tumor), 293T (human embryonic kidney), C2C12 (mouse skeletal muscle), and 3T3-L1 (mouse pre-adipocyte). HA-tagged NET32 and mRFP-tagged NETs 11, 13, 45, 59, and emerin were transiently transfected into the different cells. To ensure that rim accumulation was not due to bleedthrough or cross-reactivity with NE markers, cells were not co-transfected with or stained for other NE proteins. *Arrows* mark cells in which different NETs yielded discernible rim staining by having a strong distinct rim as opposed to one that could be accounted by ER condensed against the nucleus. *Scale bars* 20 μm. (**b**) To further validate NE targeting, the relative pixel intensities in ER and NE were quantified compared to ER controls. Pixel intensity was measured at a point in the nuclear rim (based on DAPI staining) and at a point approximately 2 μm distant into the ER, and the NE/ER ratio was calculated. Eight such measurements were taken from each NET from 5 different cells, and Tukey's boxplots [45] for the 40 ratios for each NET in each cell line are shown with the median (*central line*), two quartiles above and below (*box*) and third quartile (*error bars*) shown. We compared each sample to each control (DiOC6 or calreticulin expressed in HT1080 cells) with the null hypothesis that 'control to sample differences are by random chance.' After analysis we reject the null hypothesis for each sample at $P < 0.001$ by Mann-Whitney (Wilcoxon) U test (Table 2). The ER dye DiOC6 and calreticulin-GFP both were very similar in intensity between the ER and NE, yielding a ratio of ~ 1. Though NET ratios tended to be in the 1.3–1.5 range, they were highly statistically significant, even for NET11 in C2C12 cells where a strong rim was not visually evident

Some putative NETs appear to not be integral to the membrane

Most remaining putative NETs that failed to target were tested in the same range of cell lines, some with both N-terminal HA and C-terminal RFP epitope tags in case one or the other tag or tag location interfered with proper folding or blocked binding sites required for NE retention (Table 1). In several cases (NETs 14, 17, 21, 35, 43, and 49) the tagged protein accumulated in the nucleoplasm/nucleolus instead of the NE (data not shown). As this suggested that the original membrane helix predictions using TMPred [18] were erroneous, all putative NETs were re-evalutated using the more stringent TMHMM v2.0 [19]. TMHMM predicted transmembrane helices in only 33 of the 67 putative NETs. Of those tested here that TMHMM failed to predict as transmembrane, 36% nonetheless gave clear nuclear rim and ER distribution consistent with membrane association, whereas 90% of those with TMHMM membrane predictions yielded NE targeting (Table 3). As TMHMM failed to predict membrane spans for 22 of the 27 NETs that have not been tested for targeting, few of them are likely to be NETs.

Several NETs exhibit significantly reduced NE accumulation in lamin A/C knockout cells

The NE retention of NETs is thought to be driven by their binding to lamins or chromatin [28], e.g., emerin accumulates in the ER in cells lacking A/C lamins [29] and has been shown to directly bind to A/C lamins [5]. To gauge the percentage of NETs in this large dataset likely to depend on lamin A for targeting, we compared the distribution of the tagged proteins in matched mouse embryonic fibroblasts that either expressed the endogenous lamin A protein (Lmna +/+) or carried a disruption in the *LMNA* gene (Lmna −/−; [29]). Most had no difference between the Lmna +/+ and the Lmna −/− cells, but several were altered in distribution. The following scenarios were observed: (1) distinctive NE accumulation in both cell lines (NETs 8, 20, 29, 37, 46, 51, 55, and 59); (2) a significant and reproducible lack of NE accumulation in the Lmna

Table 1 Failed putative NETs tested in different cell lines

NETs tested in	14-mRFP	HA-14	16-GFP	17-mRFP & HA-17	21-mRFP	HA-35	35-mRFP	36-mRFP	43-mRFP & HA-43	44-mRFP	49-mRFP
HT1080	n/d	n/d	–	–	–	–	n/d	–	–	–	–
293T	n/d	n/d	–	–	n/d	–	n/d	–	–	–	–
C2C12	–	–	–	–	n/d	–	–	–	–	–	–
NIH 3T3-L1	–	n/d	–	–	n/d	–	–	–	–	–	–
HepG2	n/d	n/d	–	–	n/d	–	n/d	–	–	–	–
U2OS	–	–	n/d	n/d	n/d	–	–	–	n/d	n/d	n/d

–, NETs did not target in this cell line; n/d, NETs not tested in this cell line

Table 2 Mann-Whitney U-test (Wilcoxon) *P*-values for comparison of each protein to the DiOC6 control or the calreticulin control for Tukey's boxplot shown in Fig. 6b (all are significant)

NET in cell line	With control DiOC6		With control calreticulin	
	W	*P*-value	W	*P*-value
Emerin in HT1080	30	1.303e-13	8	2.571e-14
NET11 in C2C12	459.5	0.001065	336	8.13e-06
NET11 in NIH-3T3	303	1.765e-06	239.5	7.032e-08
NET13 in C2C12	343.5	1.135e-05	220	2.416e-08
NET32 in C2C12	279	5.441e-07	189.5	4.309e-09
NET45 in HT1080	346	1.270e-05	236	5.791e-08
NET45 in HepG2	58	9.585e-13	17	5.014e-14
NET59 in C2C12	168.5	1.255e-09	140	2.184e-10

−/− cells compared to the Lmna +/+ cells (NETs 23, 26, 34, 50); (3) weak or no NE accumulation in the Lmna +/+ cells so that a comparison could not be made (NETs 4, 5, 30, 33, 47 and 56) (Fig. 8). As both lines tested are mouse, it is unlikely that differences between human and mouse proteins would account for the mistargeting in the Lmna −/− cells, the more so as they are highly conserved with NET26, NET34, and NET50 all having over 77% amino acid identity between human and mouse and NET23 having 69% identity.

Alternate mechanisms for NE retention of NETs in cells lacking lamin A

To further assess potential interactions of the NETs with lamins, we investigated whether their resistance to detergent extraction was altered by the presence or absence of lamin A. This could only be assayed for the NETs with working antibodies because the overexpressed NET accumulation in the ER would interfere with quantification. Wild-type Jurkat cells that normally lack lamin A and Jurkat cells stably transfected with lamin A-GFP were extracted with 1% Triton X-100/50 mM NaCl. The insoluble material (e.g., lamins and associated NETs) that resisted extraction was measured by immunoblotting and quantified as a percentage of the relative levels in non-extracted lysates, which were set to 100 (Fig. 9a). To improve solubility of the lamina in the non-extracted lysates, they were first treated with detergent and urea prior to heating in SDS-sample buffer; however, lamin levels detected on Western blots were about 30% higher in the extracted lysates than in the non-extracted lysates. This could be due to greater accessibility for subsequent solubilization in the sample buffer after removing chromatin proteins such as histones.

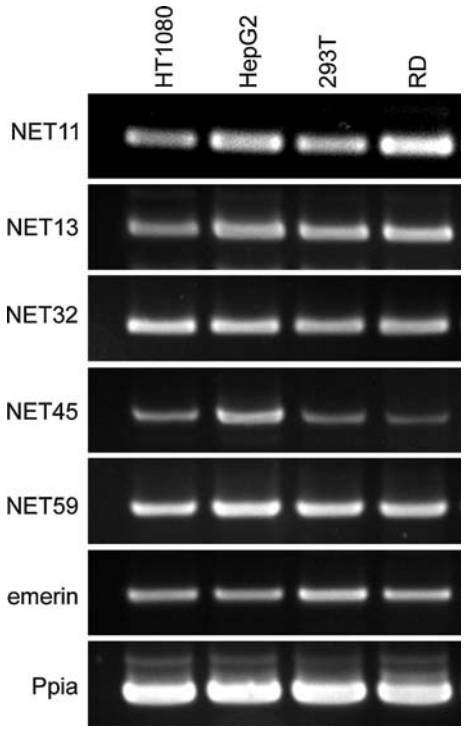


Fig. 7 Expression profiles of NETs only sometimes correlate with their targeting. mRNAs were prepared from the human cell lines used in Fig. 6. RD myoblast cells (*human*) were used to represent muscle because C2C12 myoblast cells are derived from mouse. RT-PCR reactions were performed to determine the relative NET levels in the cell lines, using peptidylprolyl isomerase A (Ppia) as a loading control. Each was repeated at least three times, and representative gels are shown

Table 3 Summary of NET targeting and transmembrane predictions

NET-GENE	Common protein name	Mr (kDa)	NE Localization	Resist Triton	Confirmed w/Ab	Predicted TM spans	
						TMPred	TMHMM
3-TMEM48	Transmembrane protein 48; NDC1	76.3	Yes [12]	Yes [12]	nd	5	5
5-TMEM201	Transmembrane protein 201	43.3	Yes	Yes	nd	6	6
8-LPGAT1	Lysophosphatidylglycerol acyltransferase 1	43.0	Yes [12]	Yes [12]	nd	3	4
9-TOR1AIP2	Torsin A interacting protein 2	51.2	Yes [14]	Yes [14]	nd	1	1
23-TMEM173	Transmembrane protein 173; MITA; STING; MYPS	42.1	Yes	Yes	Yes	3	3
24-ERGIC1	Endoplasmic reticulum-golgi intermediate compartment 32-kDa protein isoform 1 (*Homo sapiens*); (ERGIC-32) (*Rattus norvegicus*)	32.5	Yes	Yes	nd	2	2
25-LEMD2	LEM domain containing 2	56.9	Yes [13, 14]	Yes [13, 14]	nd	1	2
29-TMEM120	Transmembrane protein 120; transmembrane protein induced by tumor necrosis factor alpha	42.8	Yes, cell specific	Yes	Yes	5	5
30-MOSPD3	Motile sperm domain containing 3	25.5	Yes	Yes	Yes	2	2
32-ERLIN2	ER lipid raft associated 2 isoform 1 (*Homo sapiens*); similar to SPFH domain family, member 2 (*Rattus norvegicus*)	37.8	Yes [14], cell specific	Yes [14]	nd	1	0 (60%)
33-SCARA5	Scavenger receptor class A, member 5 (*Homo sapiens*); metallothionein-like protein 4C; similar to protease, serine, 12 (*Rattus norvegicus*)	43.2	Yes	Yes	Yes	1	1
34-SLC39A14	Solute carrier family 39	52.8	Yes	Yes	nd	7	7
37-KIAA1161	Similar to CG11909-PA (*Rattus norvegicus*); domain from glycosyl hydrolase family 31	81.0	Yes [12, 14]	Yes [12, 14]	nd	2	1
39-PPAPDC3	Phosphatidic acid phosphatase type 2 domain containing 3	21.9	Yes [12], cell specific	Yes [12]	Yes	3	3
47-TM7SF2	Transmembrane 7 superfamily member 2	46.4	Yes	Yes	nd	7	6
50-DHRS7	Dehydrogenase/reductase (SDR family) member 7	38.2	Yes	Yes	Yes	2	1

Table 3 continued

NET-GENE	Common protein name	Mr (kDa)	NE Localization	Resist Triton	Confirmed w/Ab	Predicted TM spans	
						TMPred	TMHMM
51-C14orf1	Chromosome 14 open reading frame 1 (*Homo sapiens*); similar to Probable ergosterol biosynthetic protein 28 (*Rattus norvegicus*)	15.8	Yes [12]	Yes [12]	Yes	3	4
53-C14orf49	Nesprin-3	112.2	Yes [15]	Yes [15]	nd	1	0 (2%)
55-APH1B	Anterior pharynx defective 1B	28.4	Yes	Yes	nd	6	6
56-DULLARD	Dullard homolog	28.3	Yes [12]	Yes [12]	nd	1	1
59-NCLN	Nicalin	62.9	Yes	Yes	nd	3	2
11-SCCPDH	Saccharopine dehydrogenase (putative)	47.2	Yes, cell specific	nd	nd	1	1
13-SMPD4	Neutral sphingomyelinase 3 isoform 1 (*Homo sapiens*); nSMase4, sphingomyelin phosphodiesterase 4 (*Mus musculus*)	83.0	Yes, cell specific	nd	nd	1	2
4-TMEM53	Transmembrane protein 53	31.5	Yes [12]	No [12]	Yes	1	1
20-FAM105A	Family with sequence similarity 105, member A	42.1	Yes	No	nd	1	1
26-TMEM14c	Transmembrane protein 14C	11.5	Yes [12]	No [12]	nd	4	4
31-TMEM209	Transmembrane protein 209	62.9	Yes [12]	No [12]	Yes	2	2
38-ALG2	Asparagine-linked glycosylation 2 homolog (yeast, alpha-1,3-mannosyl transferase), transcript variant 1	47.0	Yes	No	nd	1	0 (20%)
45-DAK	Dihydroxyacetone kinase 2	58.9	Yes, cell specific	No	nd	1	0 (60%)
46-SLC22A24	Solute carrier family 22, member 24	35.9	Yes	No	nd	6	5
62-MCAT	Similar to Malonyl CoA-acyl carrier protein transacylase	42.9	Yes	No	nd	1	0 (20%)
14-WDR33	WD repeat domain 33, transcript variant 2	38.2	No, nucleoplasm	nd	nd	2	1
16-WDR75	WD repeat domain 75	94.4	No	nd	nd	1	0 (10%)
17-C3orf17	Chromosome 3 open reading frame 17, transcript variant 1	64.4	No, nucleoplasm	nd	nd	1	0 (50%)

Table 3 continued

NET-GENE	Common protein name	Mr (kDa)	NE Localization	Resist Triton	Confirmed w/Ab	Predicted TM spans	
						TMPred	TMHMM
21-UTP15	U3 small nucleolar ribonucleoprotein, homolog (*S. cervisiae*); similar to Src-associated protein SAW (*Rattus norvegicus*)	58.4	No, nucleolar	nd	nd	1	0 (2%)
35-KIAA1967	p30 DBC protein	40.8	No	nd	nd	1	0 (5%)
36-TMEM74	Transmembrane protein 74	33.3	No	nd	nd	2	2
43-NAT10	N-acetyltransferase 10; hALP	115.7	No, nucleolar	nd	nd	2	0 (70%)
44-SLC25A22	Solute carrier family 25 (mitochondrial carrier: glutamate), member 22; mitochondrial glutamate carrier 1	34.4	No, mitochondria	nd	nd	3	1
49-NOC4L	Nucleolar complex associated 4 homolog	58.4	No, nucleoplasm	nd	nd	1	0 (25%)
15-NOC2L	Nucleolar complex associated 2 homolog (*S. cerevisiae*)	84.2	No expression	nd	nd	4	1
48-C12orf11	Chromosome 12 open reading frame 11 (sarcoma antigen NY-SAR-95) isoform 1	80.2	No expression	nd	nd	1	0 (0%)

Gene names are given next to NET numbers. The protein molecular weights given are for the NET coding sequence obtained from IMAGE clones used in this study. For most NETs in their endogenous forms, several different splice variants are predicted from database information. Numbers of transmembrane helices predicted by each algorithm are given in the columns marked TMPred and TMHMM. Several possible transmembrane helices predicted by TMPred and not by TMHMM nonetheless had partial probability scores from TMHMM. The percent probability that TMHMM assigned for these possible transmembrane helices is given in parentheses

nd Not done

As expected, the negative control LAP2ß that has been shown to bind lamin B [22] yielded no difference to extraction between the Jurkat cells containing or lacking lamin A; neither, however, did NETs 23 and 50 that were strongly mislocalized in the lamin A/C null fibroblasts nor the positive control emerin. Only NET33 showed a minimal difference in resistance to extraction between the Jurkat cells containing or lacking lamin A. NET23 and the lamins alone fully resisted the detergent extraction from Jurkat cells, even though these novel NETs resisted extraction in multiple adherent cell types. LAP2ß resisted extraction only to 80% of non-extracted levels and emerin only to 50%. Novel NETs 33 and 50 weakly resisted the extraction, retaining ~30% of non-extracted levels.

As this result was unexpected, the same assay was applied to two mouse fibroblast lines, one of which was disrupted for endogenous lamin A as confirmed by immunoblotting. In this case emerin, which was expected to depend on lamin A for NE retention, was ~2× more resistant to detergent extraction in the wild-type fibroblasts compared to those lacking A/C lamins (Fig. 9b). Yet nearly 40% resisted extraction in the lamin A null cells, indicating that lamin A-independent mechanisms also exist for retention of emerin at the NE in these fibroblasts. NET33, which in Jurkat cells had shown only a slight benefit from the presence of lamin A for its resistance to detergent extraction, became entirely dependent on the presence of lamin A for its resistance to detergent in the fibroblasts. NET50 remained indifferent to lamin A. Surprisingly, only about 50% of the LAP2ß resisted extraction in the lamin A knockout fibroblasts, whereas most had resisted extraction in both Jurkat lines. Lamin B1 levels measured were higher in the Lmna −/− cells. This could be due to upregulation by the cell to compensate for the loss of lamin A or might reflect greater solubility of the B-type lamins in the absence of lamin A that has previously been reported [30]. Thus, multiple mechanisms must exist for the targeting and association of the same NETs in different cell types.

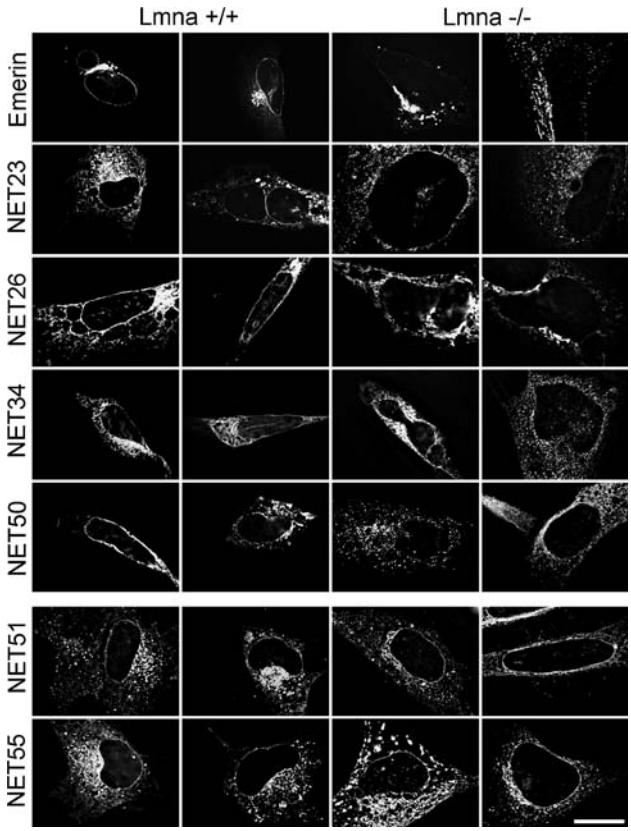


Fig. 8 NETs that mistarget in cells lacking lamin A/C. Mouse embryonic fibroblasts extracted from a wild-type mouse (*Lmna* +/+) or from a matched *LMNA* knockout mouse (*Lmna* −/−) [29] were transfected with NET fusion constructs. At 30 h post-transfection, cells were directly fixed with formaldehyde and processed for immunofluorescence microscopy. Upper panels above the break show emerin and NETs that produced a distinctive rim in the Lmna +/+ cells, but did not in the Lmna −/− cells. Other NETs tested yielded no striking or reproducible differences in presence or absence of lamin A/C (only NET51 and NET55 are shown in *bottom panels*). Deconvolved images are shown. *Scale bars* 20 µm

Discussion

This study contributes three important findings: first it tests the validity of proteomic results that previously greatly expanded the number of putative NETs [12], confirming many, but finding that roughly a third were erroneously named 'NET.' Second, it shows that some valid NETs only target to the NE in certain cell types, a finding that indicates the need to carefully match studies of NETs with appropriate cell types. Third, the differences among cell types and between in vivo and in vitro results for both extraction and targeting studies indicate that NETs likely have different mechanisms for targeting to the NE in different cell types.

While this study has confirmed many novel NETs, it has also shown that transmembrane predictions for others were erroneous. Extrapolating from the NETs now tested, we estimate that 30% of the 67 putative NETs cannot properly be called NETs because of the absence of transmembrane spans. Nonetheless, the failure of newer algorithms to predict transmembrane helices for several NETs that we and others have confirmed [14, 15] leaves open the possibility that other NETs may be validated. Some of those apparently lacking transmembrane segments may have additional splice variants that encode membrane spans, whereas others might function at the NE without a membrane span: NET43/hALP is recruited to the NE through binding SUN1 at the end of mitosis where it contributes to chromosome decondensation [31]. Thus, in addition to

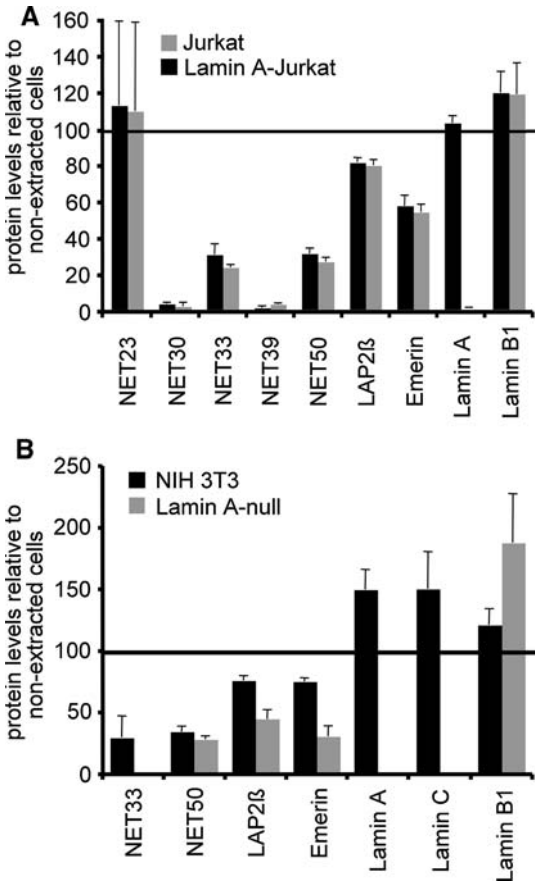


Fig. 9 Indirect assay for lamin A interactions. If a NET that resists detergent extraction depends on lamin A for its NE retention either through direct or indirect binding, then it would be expected to be more resistant to detergent extraction in cells that express lamin A. Relative NET and lamin protein levels were quantified between lamin A expressing and not expressing cells from lysates run on Western blots using anti-NET and anti-lamin antibodies. The graphs show the levels of the protein left in cells extracted with 1% Triton X-100 as a percentage of the levels measured in unextracted cells. The numbers used for NETs were normalized according to the amount of lamin B1 remaining after extraction, which is shown here as absolute values, and three separate experiments were averaged to generate standard deviations. (**a**) Jurkat cells are suspension cells derived from a T-cell lymphoma that do not normally express lamin A. To test if adding lamin A to these cells would affect NET resistance to detergent extraction, Jurkats stably expressing lamin A-GFP were generated. Several NETs that resisted detergent extraction in the HT1080 fibroblasts grown on coverslips did not resist extraction in the Jurkat cells whether or not lamin A was present. Those that did resist showed no difference between the lamin A expressing and non-expressing cells. (**b**) Lmna −/− mouse fibroblasts ($216^{-/-}$; [29]) and control mouse fibroblasts (NIH3T3) were also compared. In the fibroblast system emerin, LAP2ß, and NET33 differed in their extractibility between the lamin A null cells and the lamin A-expressing cells

being at the NE only in certain cell types, some proteins may have multiple subcellular locations and only reside transiently at the NE during certain cell-cycle stages.

Transient nuclear envelope accumulation was also reported previously for NET59/nicalin, which tends to reside in the ER but accumulates in the NE when certain co-factors are present [32]. Moreover, we have shown here that its NE accumulation is in the INM and it resists detergent extraction. Many NETs likely have multiple cellular localizations as emerin, one of the most established INM NETs [33], also functions in the ONM, ER, and interstitial discs [23–25]. Moreover, a recent elegant proteomic study suggests that at least a third of all proteins have multiple cellular localizations [34]. Nonetheless, most of the NETs for which we have antibodies yeilded their principal accumulation in the NE compared to the ER.

The finding that some NETs only target to the NE in particular cell types both by overexpression analysis and antibody staining in tissue sections is consistent with the hypothesis that specificity of NE functions in different tissues and/or cell types might contribute to the tissue preference of NE disease pathologies [35]. Although the primary tissue-specific cell type of liver is hepatocytes (estimated at ~80%), there are also Kuppfer cells that are liver specific, lipocytes (similar to adipocytes), as well as epithelial cells and an extensive vasculature providing endothelial cells. As functions have yet to be found for most new NETs, it is early to speculate detailed molecular mechanisms that would yield pathology. However, NETs 13, 39, 45, 55, and 59 have either been reported to directly function in signaling or to be related to signaling proteins [32, 36–40]. Thus, these NETs could be involved in signal transduction pathways similar to those of the well-characterized NET MAN1 in Smad/BMP/TGFß signaling [41]. Intriguingly, two of these NETs accumulate in the inner membrane, whereas the other three accumulate in the outer membrane. Moreover, NETs functioning in signaling pathways might vary in localization between ONM and INM depending on timing in the cell cycle. The restriction in targeting to certain cell types for three of these signaling-related NETs (NETs 13, 39, and 45) might reflect cell-specific use of these signaling pathways, a possible mechanism towards disease that further underscores the need to test putative NETs in multiple cell types.

Historically, NE proteins have been thought to reside exclusively in the inner membrane despite the fact that this represents only half of the NE. Of the 18 NETs tested by high-resolution microscopy, 4 appeared to localize to the ONM only and not to the INM, indicating the need to redefine the ONM as a separate and distinct cellular compartment and not just a subcompartment of the ER. INM localization only generally correlated with detergent resistance, with 79% of INM tested proteins resisting detergent pre-extraction and 50% of ONM proteins tested not resisting the pre-extraction.

Few NETs have been directly tested for binding to lamins in general or to particular subtypes. Some bind multiple lamin subtypes (e.g., LAP1 [22]) while others

bind specific subtypes (e.g., LBR with lamin B [22, 42] and emerin with lamin A [5]). The mistargeting of some NETs in cells lacking lamin A suggests that they are tethered at the NE, either directly or indirectly, through binding to lamin A. However, differences were observed in NETs for lamin A-dependent resistance to detergent extraction between the lymphoblasts and fibroblasts tested. Thus, multiple mechanisms must exist for the targeting and association of NETs with lamins in different cell types. This can explain in part the tissue-preference of pathology in NE diseases as it suggests that the NET's NE retention only depends on lamin A in certain cell types.

The confirmation of such a large set of NETs is an important step in understanding the functioning and shared characteristics of proteins in this organelle. For example, the 31 of the liver NETs identified by proteomics now confirmed may be enough for analyses by bioinformatics experts to identify NE targeting sequences. Moreover, some NETs that have more tissue-specific expression likely interact with lamins and other NETs, and thus might contribute to complexes involved in NE diseases. Of particular note is NET39, which resisted pre-extraction with detergent and was in the INM. It is expressed preferentially in skeletal muscle in mouse and has been shown to be strongly induced during myogenesis and to play a direct role in signaling mechanisms for myogenesis [14, 43]. In humans NET39 is preferentially expressed in heart [27]: as cardiomyopathy has been linked to the lamina in humans [44], this protein might be relevant for the tissue-specificity of this disease. The differences we observed for several NETs in cellular localizations in different cell types and targeting in the Lmna −/− cells thus provide an important starting point from which to address lamin-NET complexes.

Acknowledgments We would like to thank David Kelly for assistance with image quantification, Alastair Kerr for statistical analysis, Emma King for OMX assistance, Tom Rapoport and Anne Straube for plasmids, Sue Shackleton for cell lines, Billy Powell and Hanae Gourier for technical assistance, and Maria Vogelauer (UCLA) for critical reading of the manuscript. Use of the OMX microscope was supported by the Scottish University Life Sciences Alliance. P.M. was supported by a Royal Society Fellowship, D.G.B. by a Darwin Trust scholarship, and this work by a Wellcome Trust Senior Research Fellowship to E.C.S.

References

1. Worman HJ, Bonne G (2007) "Laminopathies": a wide spectrum of human diseases. Exp Cell Res 313:2121–2133
2. Mattout A, Dechat T, Adam SA, Goldman RD, Gruenbaum Y (2006) Nuclear lamins, diseases and aging. Curr Opin Cell Biol 18:335–341
3. Bonne G, Di Barletta MR, Varnous S, Becane HM, Hammouda EH, Merlini L, Muntoni F, Greenberg CR, Gary F, Urtizberea JA, Duboc D, Fardeau M, Toniolo D, Schwartz K (1999) Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. Nat Genet 21:285–288
4. Bione S, Maestrini E, Rivella S, Mancini M, Regis S, Romeo G, Toniolo D (1994) Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. Nat Genet 8:323–327
5. Clements L, Manilal S, Love DR, Morris GE (2000) Direct interaction between emerin and lamin A. Biochem Biophys Res Commun 267:709–714
6. Zhang Q, Bethmann C, Worth NF, Davies JD, Wasner C, Feuer A, Ragnauth CD, Yi Q, Mellad JA, Warren DT, Wheeler MA, Ellis JA, Skepper JN, Vorgerd M, Schlotter-Weigel B, Weissberg PL, Roberts RG, Wehnert M, Shanahan CM (2007) Nesprin-1 and -2 are involved in the pathogenesis of Emery-Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. Hum Mol Genet 16:2816–2833
7. Zhang Q, Ragnauth CD, Skepper JN, Worth NF, Warren DT, Roberts RG, Weissberg PL, Ellis JA, Shanahan CM (2005) Nesprin-2 is a multi-isomeric protein that binds lamin and emerin at the nuclear envelope and forms a subcellular network in skeletal muscle. J Cell Sci 118:673–687
8. Bengtsson L, Wilson KL (2004) Multiple and surprising new functions for emerin, a nuclear membrane protein. Curr Opin Cell Biol 16:73–79
9. Zastrow MS, Vlcek S, Wilson KL (2004) Proteins that bind A-type lamins: integrating isolated clues. J Cell Sci 117:979–987
10. Tunnah D, Sewry CA, Vaux D, Schirmer EC, Morris GE (2005) The apparent absence of lamin B1 and emerin in many tissue nuclei is due to epitope masking. J Mol Histol 36:337–344
11. Dorner D, Gotzmann J, Foisner R (2007) Nucleoplasmic lamins and their interaction partners, LAP2alpha, Rb, and BAF, in transcriptional regulation. FEBS J 274:1362–1373
12. Schirmer EC, Florens L, Guan T, Yates JRr, Gerace L (2003) Nuclear membrane proteins with potential disease links found by subtractive proteomics. Science 301:1380–1382
13. Brachner A, Reipert S, Foisner R, Gotzmann J (2005) LEM2 is a novel MAN1-related inner nuclear membrane protein associated with A-type lamins. J Cell Sci 118:5797–5810
14. Chen IH, Huber M, Guan T, Bubeck A, Gerace L (2006) Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis. BMC Cell Biol 7:38
15. Wilhelmsen K, Litjens SH, Kuikman I, Tshimbalanga N, Janssen H, van den Bout I, Raymond K, Sonnenberg A (2005) Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. J Cell Biol 171:799–810
16. Schirmer EC, Guan T, Gerace L (2001) Involvement of the lamin rod domain in heterotypic lamin interactions important for nuclear organization. J Cell Biol 153:479–489
17. Florens L, Korfali N, Schirmer EC (2008) Subcellular fractionation and proteomics of nuclear envelopes. Methods Mol Biol 432:117–137
18. Hofmann K, Stoffel W (1993) TMbase: a database of membrane spanning proteins segments. Biol Chem Hoppe-Seyler 374:166
19. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580
20. Soullam B, Worman HJ (1995) Signals and structural features involved in integral membrane protein targeting to the inner nuclear membrane. J Cell Biol 130:15–27

21. Starr DA, Fischer JA (2005) KASH 'n Karry: the KASH domain family of cargo-specific cytoskeletal adaptor proteins. Bioessays 27:1136–1146
22. Foisner R, 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
23. Salpingidou G, Smertenko A, Hausmanowa-Petrucewicz I, Hussey PJ, Hutchison CJ (2007) A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane. J Cell Biol 178:897–904
24. Cartegni L, di Barletta MR, Barresi R, Squarzoni S, Sabatelli P, Maraldi N, Mora M, Di Blasi C, Cornelio F, Merlini L, Villa A, Cobianchi F, Toniolo D (1997) Heart-specific localization of emerin: new insights into Emery-Dreifuss muscular dystrophy. Hum Mol Genet 6:2257–2264
25. Lattanzi G, Ognibene A, Sabatelli P, Capanni C, Toniolo D, Columbaro M, Santi S, Riccio M, Merlini L, Maraldi NM, Squarzoni S (2000) Emerin expression at the early stages of myogenic differentiation. Differentiation 66:208–217
26. Schermelleh L, Carlton PM, Haase S, Shao L, Winoto L, Kner P, Burke B, Cardoso MC, Agard DA, Gustafsson MG, Leonhardt H, Sedat JW (2008) Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. Science 320:1332–1336
27. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, Patapoutian A, Hampton GM, Schultz PG, Hogenesch JB (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA 99:4465–4470
28. Gruenbaum Y, Margalit A, Goldman RD, Shumaker DK, Wilson KL (2005) The nuclear lamina comes of age. Nat Rev Mol Cell Biol 6:21–31
29. Sullivan T, Escalante-Alcalde D, Bhatt H, Anver M, Bhat N, Nagashima K, Stewart CL, Burke B (1999) Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. J Cell Biol 147:913–920
30. Schirmer EC, 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 279:42811–42817
31. Chi YH, Haller K, Peloponese JM Jr, Jeang KT (2007) Histone acetyltransferase hALP and nuclear membrane protein hsSUN1 function in de-condensation of mitotic chromosomes. J Biol Chem 282:27447–27458
32. Haffner C, Dettmer U, Weiler T, Haass C (2007) The Nicastrin-like protein Nicalin regulates assembly and stability of the Nicalin-nodal modulator (NOMO) membrane protein complex. J Biol Chem 282:10632–10638
33. Manilal S, Nguyen TM, Sewry CA, Morris GE (1996) The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. Hum Mol Genet 5:801–808
34. 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 125:187–199
35. Wilkie GS, Schirmer EC (2006) Guilt by association: the nuclear envelope proteome and disease. Mol Cell Proteomics 5:1675–1865
36. Jacobsson JA, Haitina T, Lindblom J, Fredriksson R (2007) Identification of six putative human transporters with structural similarity to the drug transporter SLC22 family. Genomics 90:595–609
37. Coolen MW, Van Loo KM, Van Bakel NN, Pulford DJ, Serneels L, De Strooper B, Ellenbroek BA, Cools AR, Martens GJ (2005) Gene dosage effect on gamma-secretase component Aph-1b in a rat model for neurodevelopmental disorders. Neuron 45:497–503
38. Diao F, Li S, Tian Y, Zhang M, Xu LG, Zhang Y, Wang RP, Chen D, Zhai Z, Zhong B, Tien P, Shu HB (2007) Negative regulation of MDA5- but not RIG-I-mediated innate antiviral signaling by the dihydroxyacetone kinase. Proc Natl Acad Sci USA 104:11706–11711
39. Sigal YJ, McDermott MI, Morris AJ (2005) Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions. Biochem J 387:281–293
40. Krut O, Wiegmann K, Kashkar H, Yazdanpanah B, Kronke M (2006) Novel tumor necrosis factor-responsive mammalian neutral sphingomyelinase-3 is a C-tail-anchored protein. J Biol Chem 281:13784–13793
41. Worman HJ (2006) Inner nuclear membrane and regulation of Smad-mediated signaling. Biochim Biophys Acta 1761:626–631
42. Worman HJ, Yuan J, Blobel G, Georgatos SD (1988) A lamin B receptor in the nuclear envelope. Proc Natl Acad Sci USA 85:8531–8534
43. Liu GH, Guan T, Datta K, Coppinger J, Yates J 3rd, Gerace L (2009) Regulation of myoblast differentiation by the nuclear envelope protein NET39. Mol Cell Biol 29:5800–5812
44. Fatkin D, MacRae C, Sasaki T, Wolff MR, Porcu M, Frenneaux M, Atherton J, Vidaillet HJ Jr, Spudich S, De Girolami U, Seidman JG, Seidman C, Muntoni F, Muehle G, Johnson W, McDonough B (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
45. Chambers JM, Cleveland WS, Kleiner B, Tukey PA (1983) Graphical Methods for Data Analysis. Duxbury Press, Boston

⌘ *Author's Choice*

The Leukocyte Nuclear Envelope Proteome Varies with Cell Activation and Contains Novel Transmembrane Proteins That Affect Genome Architecture*[S]

Nadia Korfali‡, Gavin S. Wilkie‡, Selene K. Swanson§, Vlastimil Srsen‡, Dzmitry G. Batrakou‡¶, Elizabeth A. L. Fairley‡, Poonam Malik‡‖, Nikolaj Zuleger‡, Alexander Goncharevich‡, Jose de las Heras‡, David A. Kelly‡, Alastair R. W. Kerr‡, Laurence Florens§, and Eric C. Schirmer‡

A favored hypothesis to explain the pathology underlying nuclear envelopathies is that mutations in nuclear envelope proteins alter genome/chromatin organization and thus gene expression. To identify nuclear envelope proteins that play roles in genome organization, we analyzed nuclear envelopes from resting and phytohemagglutinin-activated leukocytes because leukocytes have a particularly high density of peripheral chromatin that undergoes significant reorganization upon such activation. Thus, nuclear envelopes were isolated from leukocytes in the two states and analyzed by multidimensional protein identification technology using an approach that used expected contaminating membranes as subtractive fractions. A total of 3351 proteins were identified between both nuclear envelope data sets among which were 87 putative nuclear envelope transmembrane proteins (NETs) that were not identified in a previous proteomics analysis of liver nuclear envelopes. Nuclear envelope localization was confirmed for 11 new NETs using tagged fusion proteins and antibodies on spleen cryosections. 27% of the new proteins identified were unique to one or the other of the two leukocyte states. Differences in expression between activated and resting leukocytes were confirmed for some NETs by RT-PCR, and most of these proteins appear to only be expressed in certain types of blood cells. Several known proteins identified in both data sets have functions in chromatin organization and gene regulation. To test whether the novel NETs identified might include those that also regulate chromatin, nine were run through two screens for different chromatin effects. One screen found two NETs that can recruit a specific gene locus to the nuclear periphery, and the second found a different NET that promotes chromatin condensation. The variation in the protein milieu with pharmacological activation of the same cell population and consequences for gene regulation suggest that the nuclear envelope is a complex regulatory system with significant influences on genome organization. *Molecular & Cellular Proteomics 9: 2571–2585, 2010.*

The nuclear envelope (NE)[1] is a double membrane system consisting of the intermediate filament nuclear lamin polymer and associated proteins attached to the inner nuclear membrane (INM) (1), nuclear pore complexes (NPCs) that direct transport of soluble macromolecules in and out of the nucleus (2), and the outer nuclear membrane (ONM) and associated proteins. Structurally, the ONM is continuous with the endoplasmic reticulum (ER) and is studded with ribosomes (3), yet it also contains unique proteins, many of which connect the cytoskeleton to the NE (4). On the other side, lamins and many INM proteins directly connect chromatin to the NE. Lamins and an increasing number of nuclear envelope transmembrane proteins (NETs) have been linked to a similarly increasing number of diseases ranging from muscular dystrophy to neuropathy, dermopathy, lipodystrophy, bone disorders, and progeroid aging syndromes (5, 6).

A favored hypothesis to explain how different NE proteins can produce such a wide range of disease pathologies is that chromatin-NE connections are disrupted with NE protein mutations, yielding changes in gene regulation. This hypothesis is supported by observations that the distribution of dense peripheral chromatin is affected in fibroblasts from patients

From the ‡Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom and the §Stowers Institute for Medical Research, Kansas City, Missouri 64110

Received, July 6, 2010

⌘ *Author's Choice*—Final version full access.

Published, MCP Papers in Press, August 6, 2010, DOI 10.1074/mcp.M110.002915

[1] The abbreviations used are: NE, nuclear envelope; NET, nuclear envelope transmembrane protein; INM, inner nuclear membrane; ONM, outer nuclear membrane; PBMC, peripheral blood mononuclear cell; PHA, phytohemagglutinin; MudPIT, multidimensional protein identification technology; NPC, nuclear pore complex; ER, endoplasmic reticulum; TCEP, tris(2-carboxylethyl)phosphine hydrochloride; IAM, iodoacetamide; FDR, false discovery rate; dNSAF, distributed normalized spectral abundance factor; ID, identity; GO, gene ontology; mRFP, monomeric red fluorescent protein; lacO, lac operator; NuRD, nucleosome remodeling and deacetylase.

with NE-linked muscular dystrophy, cardiomyopathy, mandibuloacral dysplasia, and progeria (7–10). Furthermore, many binding partners have been identified for NETs that are either chromatin proteins, enzymes that modify chromatin proteins, or regulators of gene expression (1, 11). These include markers of silent chromatin such as heterochromatin protein 1 (12) and proteins that modify chromatin to a silent conformation such as histone deacetylase 3 (13). The importance of the NE to global genome organization has been underscored by several recent studies that showed that affinity-based recruitment of a specific chromosome locus by the NE both pulled entire chromosomes to the periphery and affected gene regulation in complex ways (14–16).

To identify NE proteins likely to be involved in genome organization, we turned to lymphocytes as a model system. Lymphocytes in the resting state tend to have massive amounts of dense peripheral chromatin as determined by electron microscopy studies. Upon activation with phytohemagglutinin, this dense chromatin largely dissipates as the cells actively express genes (17–20). Thus, to identify proteins that might be involved in tethering heterochromatin to the NE or in changing its organization, we analyzed the NE proteomes of leukocyte populations (~70% lymphocytes) in both the resting and phytohemagglutinin (PHA)-activated states. The previously validated subtractive approach was applied (21) using microsomes and mitochondria, the principal membrane contaminants expected, as subtractive fractions.

Many new NE proteins were identified that had not been identified in previous NE proteomics investigations using liver and neuroblastoma cells (21, 22). NE residence was confirmed for 12 novel NETs by expression of epitope-tagged versions and using antibodies on tissue cryosections.

Roughly one-quarter of the proteins identified varied between the resting and activated states. Some NET differences between the two data sets were confirmed by RT-PCR. Among the known proteins identified were several that suggest that changes in NE composition associated with PHA activation contribute to gene regulation. Novel NETs identified also appear to play significant roles in genome organization/regulation as we found that several can either recruit a specific locus to the nuclear periphery or promote chromatin condensation. As several studies have implicated misregulation of chromatin organization in NE diseases (7, 8), these newly identified NETs may contribute to the diverse pathologies associated with NE diseases.

EXPERIMENTAL PROCEDURES

Preparation and Activation of Leukocytes—Human buffy coats from healthy donors were obtained anonymously from the Scottish National Blood Transfusion Service according to approved ethics protocols. Buffy coats were diluted with phosphate-buffered saline (PBS), and a peripheral blood mononuclear cell (PBMC) fraction was isolated by centrifugation on Ficoll-Histopaque-1077 (Sigma) at 250 × *g* for 30 min. Cells were diluted in PBS, pelleted, and resuspended in RPMI 1640 medium containing 10% fetal bovine serum (FBS), 100 μg/μl penicillin, and 100 μg/μl streptomycin sulfate (Invitrogen). PBMCs were activated by addition of 2.5 μg/ml PHA (Sigma) for 7 days. To determine the composition of PBMC preparations, cells were analyzed by FACS using antibodies from Caltag Medsystems to CD19 for resting B-cells and CD19/CD4 for activated B-cells, CD3 for resting T-cells and CD3/CD45 for activated T-cells, CD3/CD4 for resting helper T-cells and CD4/CD45 for activated helper T-cells, CD3/CD8 for resting cytotoxic T-cells and CD8/CD45 for activated cytotoxic T-cells, and CD11b/CD14 for myeloid cells.

Preparation of Fractions—NEs were isolated using procedures specifically modified for blood leukocytes and detailed previously (23). In brief, nuclei were first isolated from PBMCs by hypotonic lysis in 10 mM HEPES, pH 7.4, 1.5 mM $MgCl_2$, 10 mM KCl, 2 mM DTT using a "loose" Dounce homogenizer (Wheaton). Nuclei were pelleted at 1000 × *g* for 10 min to separate them from small vesicles and mitochondria that require higher speeds to pellet. To float/remove contaminating membranes, nuclei were resuspended in SHKM (50 mM HEPES, pH 7.4, 25 mM KCl, 5 mM $MgCl_2$, 1 mM DTT and 1.8 M sucrose) and pelleted through a 5 ml 2.1 M sucrose cushion in an SW28 swinging bucket rotor (Beckman) at 4 °C for 2 h at 82,000 × *g*. NEs were then prepared from isolated nuclei by two rounds of digestion with DNase and RNase in 0.3 M sucrose, 10 mM HEPES, pH 7.4, 2 mM $MgCl_2$, 0.5 mM $CaCl_2$, 2 mM DTT for 20 min followed by layering onto the same buffer with 0.9 M sucrose and centrifugation at 6000 × *g* for 10 min at 4 °C.

The PBMC cell populations had very little ER, so microsomes representing contaminant membrane fractions were instead isolated from rat liver using well established procedures (24, 25). In brief, liver was diced and homogenized in SHKM using a motorized Dounce homogenizer. After removing nuclei as for NE preparations, 0.5 mM EDTA was added to inhibit metalloproteinases, and mitochondria and other debris from postnuclear supernatants were also removed by pelleting at 10,000 × *g* for 15 min. The supernatant was made to 2 M sucrose with SHKM and then overlaid with 1.86 and 0.25 M sucrose layers. This was then subjected to centrifugation in an SW28 swinging bucket rotor (Beckman) at 4 °C for 4 h at 57,000 × *g* to float microsomes. The material between the 1.86 and 0.25 M layers was then diluted 4-fold with 0.25 M SHKM and pelleted at 152,000 × *g* in a type 45 Ti rotor (Beckman) for 1 h.

NE preparations were divided, and one portion was further extracted on ice with 400 mM NaCl, 25 mM HEPES, pH 8.0, 5 mM $MgCl_2$, 10 mM DTT, 1% β-D-octyl glucoside, and material used for mass spectrometry analysis was pelleted at 15,000 × *g* for 30 min and washed in sterile H_2O. Another portion of NEs and separate microsomes were further extracted with 0.1 N NaOH, 10 mM DTT, and the material used for mass spectrometry was pelleted by centrifugation at 150,000 × *g* for 30 min and washed three times in sterile H_2O.

MudPIT—Two types of digestion were applied to NE and microsome pellets (supplemental Table S1). For the first, dried membrane pellets were incubated overnight in 90% formic acid and 500 mg/ml cyanogen bromide (CNBr). Ammonium hydroxide (NH_4OH) was then added to pH 8.5, solid urea was added to 8 M, and tris(2-carboxylethyl)phosphine hydrochloride (TCEP) was added to 5 M and incubated for 30 min. Next free cysteines were alkylated with 20 mM iodoacetamide (IAM) for 30 min prior to addition of endoproteinase Lys-C (Roche Applied Science) at a 1:100 (w/w) enzyme to protein ratio for 6 h at 37 °C. Urea was diluted to 2 M with 0.1 M Tris-HCl, pH 8.5, and $CaCl_2$ (0.5 mM) and modified trypsin (1:100, w/w) were added for 12 h at 37 °C. Digestions were quenched with 5% formic acid (25). For the second digestion (Ti), pellets were solubilized in 0.1 M Tris-HCl, pH 8.5, 8 M urea, 5 mM TCEP. IAM was added to 10 mM for 30 min, and endoproteinase Lys-C and trypsin digestion was performed as above. Samples were centrifuged for 30 min at 17,500 × *g*. Supernatants were analyzed by MudPIT, whereas pellets were resus-

pended in 0.1 M sodium carbonate (Na_2CO_3), pH 11.5, 8 M urea, 5 mM TCEP for 30 min and then in 10 mM IAM for 30 min and then further digested with proteinase K for 4 h at 37 °C and also analyzed by MudPIT.

Each digested sample was analyzed independently by MudPIT as described previously (26–28) using a flow rate of 200–300 nl/min. 12–15 cycles (supplemental Table S1) of 120 min each of increasing salt concentrations followed by organic gradients (5–80% acetonitrile with 0.1% formic acid) slowly released peptides directly into the mass spectrometer (26). The last two or last five (of 12 or 15) chromatography steps consisted of a high salt wash with 5% acetonitrile, 0.1% formic acid, 500 mM ammonium acetate followed by the acetonitrile gradient. The distal application of a 2.5-kV voltage electrosprayed the eluting peptides directly into ion trap mass spectrometers equipped with a nano-LC electrospray ionization source (ThermoFinnigan). Each full MS scan (from 400 to 1600 *m/z*) was followed by three (XP-Deca) or five (LTQ) MS/MS events (supplemental Table S1) using data-dependent acquisition where the first most intense ion was isolated and fragmented by collision-induced dissociation (at 35% collision energy) followed by the second to third or fifth most intense ions. The raw mass spectrometric data may be downloaded from ProteomeCommons.org by going to the URL indicated in Table I and then inserting the hash identifiers from Table I.

RAW files were extracted into ms2 file format (29) using RAW_Xtract v.1.0 (30). For the human PBMC data set (supplemental Table S2), MS/MS spectra were queried for peptide sequence information using SEQUEST v.27 (revision 9) (31) against 30,552 human proteins (non-redundant NCBI sequences on March 4, 2008) plus 162 sequences from usual contaminants (*e.g.* human keratins etc.) and 313 previously identified NETs (21). To estimate false discovery rates, each non-redundant protein entry was randomized and added to the database bringing the total search space to 61,864 sequences (supplemental Table S1). For the mouse (supplemental Table S3) (21) and rat (supplemental Table S4) liver microsomal membrane data sets, the species-specific databases consisted of 29,998 mouse or 28,400 rat proteins (non-redundant NCBI sequences on June 22, 2007 and July 10, 2006, respectively) complemented as described above with sequences for contaminants and previously identified NETs as well as their corresponding randomized sequences (supplemental Table S1). MS/MS spectra were searched without specifying differential modifications, but +57 Da were added statically to cysteine residues to account for carboxamidomethylation. No enzyme specificity was imposed during searches, setting a mass tolerance of 3 amu for precursor ions and of ±0.5 amu for fragment ions.

Results from different runs were compared using DTASelect and Contrast (32). Spectrum/peptide matches were retained if at least 7 amino acids long with ends complying with proteolytic specificity, *e.g.* Met, Lys, or Arg before the amino terminus and at the carboxyl terminus of peptides for those cleaved with CNBr prior to trypsin digestion. Other criterion were $\Delta Cn \geq 0.08$, $XCorr \geq 1.8$ for singly, 2.0 for doubly, and 3.0 for triply charged spectra, and a maximum Sp rank of 10. For proteinase K-digested samples, no specific peptide ends were imposed, but the ΔCn cutoff was increased to 0.15 (33), and XCorr minima were increased to 2.5 for doubly and 3.5 for triply charged spectra (SEQUEST parameters for the spectrum to peptide matches leading to the identification of proteins from human PBMC NEs, mouse liver microsomes, and rat liver microsomes are provided in supplemental Tables S2A, S3A, and S4A, respectively). Peptide hits from all analyses were merged to establish a master list of proteins identified by at least two peptides or one peptide with two independent spectra (see supplemental Tables S2B, S3B, and S4B for the detailed peptide and spectral counts of proteins detected by MudPIT analysis of human PBMC NEs, mouse liver microsomes, and rat liver microsomes, respectively). Based on the merged detected peptides, proteins could fall into three categories following the parsimony principle. (i) Proteins detected by the exact same set of peptides were grouped together because they could not be distinguished based on the available peptide data (see column named "Proteins in Group" in supplemental Tables S2–S4), and only one arbitrarily selected representative protein entry is reported for such groups of proteins (see column labeled "Locus" in supplemental Tables S2–S4). (ii) Proteins with at least one peptide uniquely mapping to them were considered unique entries. (iii) Subset proteins for which no unique peptides were detected were removed from the final list of identified proteins.

Identifications mapping to shuffled peptides were used to estimate false discovery rates. Spectral FDR was calculated as follows.

$$\text{Spectral FDR} = \frac{2 \times \text{SHUFFLED\_SpectralCounts}}{\text{Total\_SpectralCounts}} \times 100 \quad \text{(Eq. 1)}$$

Protein level FDR was calculated as follows.

$$\text{Protein FDR} = \frac{\text{SHUFFLED\_Proteins}}{\text{Total\_Proteins}} \times 100 \quad \text{(Eq. 2)}$$

Under these criteria, the final FDRs at the protein and peptide levels were on average 1.5 and 0.5%, respectively (supplemental Tables S2B–S4B, bottom).

To estimate relative protein levels, normalized spectral counts were calculated for each non-redundant protein as described (34–36).

$$(\text{NSAF})_i = \frac{(\text{SpectralCount/Length})_i}{\sum_{k=1}^{N} (\text{SpectralCount/Length})_k} \quad \text{(Eq. 3)}$$

We implemented a new algorithm on this data set to refine spectral counts to deal with peptides shared between multiple proteins (37). For each run, dNSAFs were calculated based on distributed spectral counts in which shared spectral counts were distributed based on spectral counts unique to each isoform (supplemental Tables S2B–S4B).

Bioinformatics Analysis—To be able to directly compare the lists of proteins detected in the human PBMC NEs and rodent liver microsome data sets, proteins were mapped to an Ensembl gene to remove redundancy from differences in protein annotations. Human gene orthologous groups were identified with Ensembl release 48 (38) to remove redundancy and false variation that might have resulted from differences in human and rodent gene assignments. Orthologous group IDs were sorted according to frequency of detection in the NE and microsome runs, relative levels (determined by averaged dNSAF values), membrane helix prediction (determined using TMHMM 2.0; http://www.cbs.dtu.dk/services/TMHMM-2.0/ (39)), and dNSAF ratios between values measured in NEs *versus* microsomes (supplemental Table S5).

Biologically interesting GO terms and their corresponding child terms were retrieved from the mySQL database AmiGO (http://amigo.geneontology.org (40)). To ensure a fair comparison for term enrichment, only human mapped genes in our data set were considered. These were compared with the genomics data set of human Ensembl genes as well as those GO defined as having nuclear localization. For a given GO term, the fraction of genes containing that term or any of the child terms was calculated for all data sets. The -fold difference was calculated by dividing this fractional value from our data set of interest by the value from the reference group.

TABLE I

Hash identifiers to access raw mass spectrometry files for each MudPIT run

Individual runs are organized by cell fraction, extraction, and protease to generate peptides for analysis. MM, microsomal membrane.

Cell fraction	Extraction/digestion	Sample name used in supplemental Tables S1–S5	At https://www.proteomecommons.org/tranche/downloads.jsp, the raw mass spectrometry files can be downloaded using the following hash identifiers. In all cases, there are two files ranging from 100 MB to 1.5 GB that are archived and compressed tar.gz files: use tar −xzf Linux command to uncompress to *.RAW files (1 per MudPIT step)
PBMC NEs	NaOH/CNBr-trypsin	HsNaOH-NE_NT_lymphocytes_CNBrTi_1	4mWjs+LvJQb4FUwQD8O3X5MkkLit6/3bHJfJZK3owfOX4jwMtCUQFhRv95ORUTT8gT5HKx6RdVxGlCjy/nOHw/pjLeMAAAAAAAAGUg==
PBMC NEs	NaOH/trypsin	HsNaOH-NE_NT_lymphocytes_Ti_1	5aTYg17TJJuGrxJ/wra09yMtKGoVwtlgYWi+EK7Np92HD73ed5417f2/x0ndBFzSjLxwFldZheB74gYe829Ql8CDDiwAAAAAAAAFrQ==
PBMC NEs	NaOH/proteinase K	HsNaOH-NE_NT_lymphocytes_PK_1	MBcaJl2EOh07R/EEOLUVt85OXiVbebs7/PiRkN3HvoazDKMmEZxw53RLG96Vx9pgGhlxM0Dxk2O05V0QiEiQi0+oUqoAAAAAAAAGlQ==
PBMC NEs	Salt-detergent/trypsin	HsSD-NE_NT_lymphocytes_Ti_1	rgMMQvG4dw8tJ9WeX0DHi0k6vvEclU34F8mzsnExggk9AbENEtfaBPxvTuBlhJ/iAnDU7XQLLNr8fJV2W72mziU6KrMAAAAAAAAFsQ==
PBMC NEs	Salt-detergent/proteinase K	HsSD-NE_NT_lymphocytes_PK_1	bQnUrYCBdxRh8wM32qka9Jg80U4ow7Z0HK+SmdbPbxPVo3TzLgeTqCWgzKPDYAkUltDGXybehz2ZFVq2EJ+G/et3AiwAAAAAAAAGmQ==
PHA-treated PBMC NEs	NaOH/CNBr-trypsin	HsNaOH-NE_PHA_lymphocytes_CNBrTi_1	/0iTLp3cH9L60CR441Dxg/REq/GKoYRz9Kd/QloEd+3h265EsD1S213sb8ic267IWS5gGDdQHV3xR+T7qZ36K4+O/ZUAAAAAAAAAGXA==
PHA-treated PBMC NEs	NaOH/trypsin	HsNaOH-NE_PHA_lymphocytes_Ti_1	gKqv8i+5tPSdMhZM9IZW93nSIXTeQUpYhgyc6SEXUgYJAj+AqeF49+VneqJ5RTQQrEnOM9t2ANPZQJWRZAJaKY0uMBAAAAAAAAAFtQ==
PHA-treated PBMC NEs	NaOH/proteinase K	HsNaOH-NE_PHA_lymphocytes_PK_1	pi3VbyTD3uo8WvUrRgmJn0phN7KpMp5yNInoWpvXSInLu/N0cNKjTY79b3okjGy6+e5zSkQATmN2dQzQ4giyJMdVmvMAAAAAAAAGnw==
PHA-treated PBMC NEs	Salt-detergent/trypsin	HsSD-NE_PHA_lymphocytes_Ti_1	5nw33XkNPBMdAuQNg6HRD2MPyTa7osxDObK2cxSdAWtVHsRV+sC3kW6omUK3PrvBNa4WCarVz3Hnskjrbn4e7JuU6+sAAAAAAAAFuw==
PHA-treated PBMC NEs	Salt-detergent/proteinase K	HsSD-NE_PHA_lymphocytes_PK_1	QPGdJRqWff6OE9isf9e9zSnB9D0xmV0rQzyR7+Jy5Z6hZZNQDzuGYKNO5+vZJFLxasQ7/Tx4lWHp/Jd5yztAJ095NuUAAAAAAAAGpA==
Liver MMs	NaOH/CNBr-trypsin	MmNaOH-MM_liver_CNBrTi_1	Y+ITtCT8NMID0geT1IJiVcUIdaOHJmM08umEd9yhyGOq/rqjAkuft0Txw1yeznfC8iXNbkfR/0v35Y0qLBhVyvrgaSwAAAAAAAGBg==
Liver MMs	NaOH/CNBr-trypsin	MmNaOH-MM_liver_CNBrTi_2	2KBwq4C/GrCKjyrUMRoRdZ55hzMbJr3wc+O3kZqwnFLV/90uaYuK64FxNneIdmnQ7HO1dRNBaNG51EGUnz3EV18OrusAAAAAAAAGAA==
Liver MMs	NaOH/trypsin	RnNaOH-MM_liver_Ti_1	bpp6s9dcwf3IzcIX0cPDJCJ01f97QesmXl6anF4PVaUNuNmfdIRrtqkELKcBM6l2X08B/d69Hys6o4vUq66TYHBc4CQAAAAAAAAFaQ==
Liver MMs	NaOH/trypsin	RnNaOH-MM_liver_Ti_2	7wO8Wz/6iv2rZU5WACjIm1sY5sIkiMWTC5X2+Tjyt5SN+fPoCR/LxuWiD3DetTh+oTWNkU7wd6gp2c3XKx8OybxI6HMAAAAAAAAFaQ==
Liver MMs	NaOH/trypsin	RnNaOH-MM_liver_Ti_3	ElAxLpoQLOHTnIp0FpP1FeRrEckCbdz0VHUt4h9ISX8v8sMEoutFOarYhBK3+d+jUv1Ihe5rRfXhhR7UcXqx0Bicw4sAAAAAAAAFaQ==
Liver MMs	NaOH/trypsin	RnNaOH-MM_liver_Ti_4	cnJ7dLVqF73eGbFqgZyTn3Egb/XL2RzMWrP7gXG41VfvbRri1eHa1uaHWjvMjbXGdvpq4XCQoz1u+gU//rR/MO7/hckAAAAAAAAFaQ==
Liver MMs	NaOH/proteinase K	RnNaOH-MM_liver_PK_1	ENorMHs0SFhkyD9Y5w4gbEh2bVCdI7gT5SF6dnp0zebaxoFG7kYpPTwYFQXGmhPBjKR/ZG8ewsnnSZZcEzmJZvYq8ecAAAAAAAAGVA==
Liver MMs	NaOH/proteinase K	RnNaOH-MM_liver_PK_2	H36mWYfElnzq6ISRTgc6v6VJVs4WxrUkCNa0LTmufeP+OoMPQaxkkBxn6nHlEv8iCgcw8ghryUbgIY9i+Udjk4k8XnsAAAAAAAAGUg==

Comparison of expression levels in different blood cell lineages was done by downloading numbers for each NET and blood cell type from BioGPS (http://biogps.gnf.org/#goto=welcome (41)) and calculating the -fold expression over the median value from all cell types tested in this transcriptome database. Thus, the numbers not only reflect differences among the blood cell lineages but also differences with other tissues.

Antibodies and Western Blotting—Antibodies used were GAPDH (Enogene, E1C604), Calreticulin (Cell Signaling Technology, 2891S), Calnexin (Stressgen, SPA-860), lamins A and B1 (3262 and 3931; Ref. 42), Nup153 (Covance), and Nup358 (raised against recombinant human protein amino acids 2595–2881; a kind gift from F. Melchior). NET antibodies were rabbit polyclonal antibodies generated to peptides from human sequences by Millipore, Tmem126 (06-1037), C17orf62 (06-1033), C17orf32 (06-1035), and MARCH5 (06-1036), and the previously characterized NETs NET31/Tmem209 (06-1020), SUN2 (06-1038). All fluorophore-conjugated secondary antibodies were minimally cross-reactivity from donkey (Jackson ImmunoResearch Laboratories) or goat (Molecular Probes).

To increase lamina solubility, PBMC or liver NEs were incubated on ice in 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 0.2% Nonidet P-40 in the presence of protease inhibitor mixture (Roche Applied Science, 11 873 580 001) for 15 min and then sonicated in a 4 °C sonic bath. Microsomes were treated similarly. Protein concentrations of liver NEs and microsomes (see Fig. 5) were determined by Bradford assay before adding sample buffer (100 mM Tris, pH 6.8, 4 M urea, 2% SDS, 50 mM DTT, 15% sucrose), heating at 65 °C, and then sonicating again in a 4 °C sonic bath. PBMC NEs and microsomes were adjusted for equal loading based on Coomassie Blue staining on polyacrylamide gels (see Fig. 3*A*). Proteins were resolved by SDS-PAGE and transferred to nitrocellulose membrane (LI-COR Biosciences). Membranes were blocked in PBS, 5% milk, 0.2% Tween. Primary antibodies were diluted in this buffer (1:200 for Millipore NET peptide antibodies, 1:500 for calreticulin, 1:200 for calnexin, and 1:2000 for lamin A) and allowed to incubate overnight at 4 °C. Secondary IR800-conjugated goat anti-rabbit antibodies (LI-COR Biosciences) were added at 1:5000 dilution at RT for 2 h. Visualization and quantification were performed using a LI-COR Odyssey system and software (Odyssey 3.0.16) using median background subtraction. Three independent blots were run for each NET and control, and averages from all three are presented in Fig. 5*B*.

Plasmid Construction—IMAGE clones for human NETs were obtained from RZPD and Geneservice. NET gene names are followed by IMAGE numbers or gene IDs in parentheses: STT3A (3891543), TAPBPL (3916213), METTL7A (2900478), C17orf62 (2822930), IAG2 (3834858), Tmem109 (3453616), C20orf3 (3530962), C17orf32 (5429060), Sec11C (3895819), Tmem126A (1667589), AADACL1 (4815834), and Tmem41A (3838280). Coding sequences were amplified by PCR with added 5′ and 3′ restriction sites, sequenced from both ends in intermediate cloning vectors, and then inserted into the mammalian expression vector pmRFP with a carboxyl-terminal tag (derived from Clontech pEGFP-N2 by replacing the GFP coding sequence with that of monomeric red fluorescent protein). Lamin A-GFP was obtained from Anne Straube (Marie Curie Research Institute, Oxted, UK), and calreticulin-GFP was obtained from Tom Rapoport (Harvard, Boston, MA).

Cell Culture and Transient Transfection—HT1080, HT1080 5.1 (carrying a lac operator (lacO) array and lac repressor-GFP), HeLa cells stably transfected with H2B-GFP, COS-7, and C2C12 cells were maintained in high glucose DMEM with 4.5 g/liter glucose (Invitrogen) supplemented with 10% fetal bovine serum (FBS), 100 μg/μl penicillin, and 100 μg/μl streptomycin sulfate (Invitrogen). PBMCs were maintained in RPMI 1640 medium with 10% FBS and antibiotics. PBMCs were activated with 2.5 μg/ml PHA (Sigma) for 7 days. Adherent cells destined for microscopy were plated on coverslips, and DNA was transfected the next day using FuGENE 6 (Roche Applied Science) according to the manufacturer's instructions.

Immunofluorescence Microscopy—After 30 h, cells were either directly fixed for 7 min in 3.7% formaldehyde or washed with PBS; then extracted for 1 min with 1% Triton X-100, 25 mM Tris, pH 8.0, 150 mM KOAc, 15 mM NaCl, 5 mM $MgCl_2$; washed again with PBS; and then fixed with formaldehyde. For antibody staining, cells that were not pre-extracted were permeabilized for 6 min in 0.2% Triton X-100 after fixation. Cells were then blocked with 10% FBS, 200 mM glycine in PBS and incubated for 40 min at RT with relevant antibodies. DNA was visualized with Hoechst 33342 or 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI) and coverslips mounted in Fluoromount-G (EM Sciences). For structured illumination (OMX) microscopy, Alexa Fluor secondary antibodies (Molecular Probes) were used. Most images were obtained using a Nikon TE-2000 microscope equipped with a 1.45 numerical aperture 100× objective, Sedat quad filter set, and CoolSnapHQ High Speed Monochrome charge-coupled device camera (Photometrics). Structured illumination images (Fig. 4*C*) were taken on the OMX system at the University of Dundee microscopy facility (details described at http://microscopy.lifesci.dundee.ac.uk/omx/).

For cryosections, fresh rat spleen cut into 2–3-mm cubes were embedded in optimal cutting temperature compound (Tissue-Tek), snap frozen in liquid nitrogen, and maintained at −80 °C. Sections were cut on a Leica CM 1900 cryostat at 6–8-μm thickness and fixed in −20 °C methanol. After rehydration, sections were incubated with NET antibodies overnight at 4 °C followed by secondary antibodies as above. Cryosection images were recorded using an SP5 confocal system (Leica) with 63× oil 1.4 numerical aperture objective using argon and UV lasers. Micrographs were saved from source programs as .tif files and prepared for figures using Photoshop 8.0.

The positional distribution of the lacO array was determined using a macro (available on request) written in Visual Basic within Image Pro Plus. In brief, the total nuclear area was automatically measured on DAPI images, and the area was then divided into five equal regions of interest through eroding 20% of the total area from the outer limits of the DAPI-defined nucleus. The nuclear sector containing the lacO spot was determined and exported to Microsoft Excel where the results from each cell were summed.

Images of chromatin condensation in Fig. 8 were generated from 60-h post-transfection cells fixed with formaldehyde using a wide field DeltaVision microscope (Applied Precision) using a 60× Plan Apo oil objective. Image stacks (0.2-μm steps) were deconvolved using DeconQ and processed using SoftWorks.

Electron Microscopy—For transmission electron microscopy, whole PBMCs or isolated NEs were fixed in 3% glutaraldehyde in 0.1 M sodium cacodylate, pH 7.3 for 2 h; washed in 0.1 M sodium cacodylate; postfixed in 1% osmium tetroxide for 45 min; washed again; and dehydrated in 50, 70, 90, and 100% normal grade acetones for 10 min each and then for a further two 10-min changes in analar acetone. Samples were then embedded in Araldite resin. 60-nm-thick sections were cut on a Reichert OMU4 ultramicrotome (Leica), stained in uranyl acetate and lead citrate, and then viewed in a Phillips CM120 transmission electron microscope. Images were taken at 100 kV at 11,000× magnification using a Gatan Orius charge-coupled device camera.

RESULTS

Generation of Lymphocyte-enriched Populations and NE Fractions—NEs were prepared from PBMCs using methods specially optimized for blood cells (23) (Fig. 1*A*). Central to these methods is isolating clean nuclei from the rest of the cell (Fig. 1*B*) that are subsequently digested with DNase and

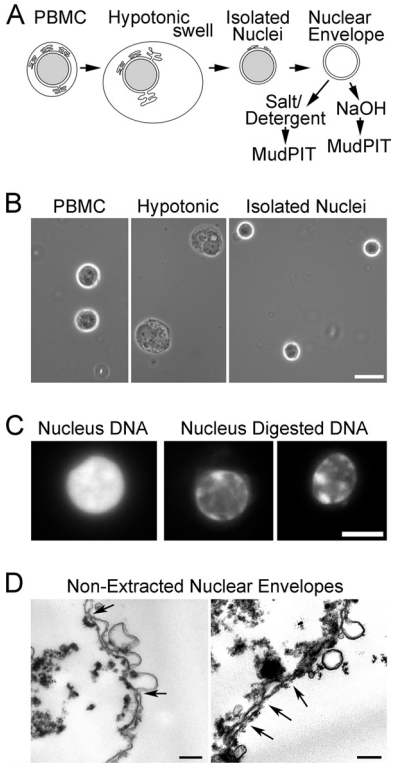


FIG. 1. **Cellular fractionation of PBMCs.** *A*, method schematic. Crude nuclei prepared from hypotonic lysis of PBMCs were cleaned of contaminating cellular structures by gradient centrifugation to float contaminating membranes while pelleting the denser nuclei. Crude NEs were prepared by digesting/extracting nuclear contents from isolated nuclei. Before MudPIT analysis, these were further extracted with 1% β-octyl glucoside, 400 mM NaCl or 0.1 N NaOH to enrich for proteins associated with the insoluble lamin polymer or integral membrane proteins, respectively. *B*, buffy coats from human blood were separated on Ficoll-Hypaque density gradients to enrich for mononuclear leukocytes (*left panel*). The cells were swollen hypotonically (*middle panel*) and Dounce homogenized to release nuclei (*right panel*), which were then further purified on sucrose gradients. Phase-contrast light microscope images are shown. *Scale bar*, 10 μm. *C*, enrichment for NEs by chromatin digestion. DAPI staining for DNA visualizes significant nuclear chromatin content in an isolated PBMC nucleus (*left panel*) and the loss of most of this material after two rounds of digestion with DNase and RNase, each followed by salt washes (*two right panels*). A fluorescence microscope image is shown. *Scale bar*, 5 μm. *D*, ultrastructure of isolated NEs. Electron micrographs of PBMC NEs show that most membranes in the population are the characteristic double membrane with little contamination from single membrane vesicles. *Arrows* point to positions of NPCs. In some places, the hypotonic treatment used to swell the NEs while digesting/extracting most of the chromatin resulted in membrane blebbing. These NEs were further extracted with salt and detergent to enrich for proteins associated with the intermediate filament lamin polymer or with an alkaline treatment to enrich for transmembrane proteins prior to analysis by MudPIT. After such treatment, no structure remains that can be readily discerned by EM with the characteristics of NEs. *Scale bar*, 0.2 μm.

RNase and salt-washed to remove most of the nuclear contents (Fig. 1*C*). Even after such isolation, NEs maintain strong connections to partners such as peripheral chromatin (Fig. 1*D*). Prior to mass spectrometry, NEs were further extracted with 1% β-octyl glucoside, 400 mM NaCl or 0.1 N NaOH that enriched for proteins associated with the insoluble intermediate filament lamin polymer or integral membrane proteins, respectively. Some well characterized NETs distribute to one or the other fraction (21, 22); hence, both detergent/salt- and NaOH-extracted fractions were analyzed to be comprehensive. Both extractions remove most contaminants so that only strongly bound and relevant partners should remain.

PBMCs were ~70% lymphocytes as determined by FACS and antibody staining for CD antigens (Fig. 2*A*). PBMC NEs were isolated with or without PHA stimulation, which mimics immune responses (43) and increased activated cells >3-fold (Fig. 2, *B* and *C*). PHA activation also results in dissolution of dense chromatin at the nuclear periphery (17–20). Staining of cells with the DNA-binding dye acridine orange confirmed the reduction of denser chromatin at the periphery. Acridine orange has been reported to bind more strongly to single-stranded DNA and thus gives a brighter signal in actively transcribed chromatin regions with condensed chromatin not allowing efficient intercalation (44). Thus, a weaker signal concentrated largely at the periphery was observed in unstimulated lymphocytes, whereas a brighter signal that was more diffuse through the nucleoplasm was observed for PHA-activated cells (Fig. 2*D*). This was further confirmed at the ultrastructural level as electron microscope images of the cells revealed significant differences in the chromatin distribution. Although patches could still be observed at the periphery in PHA-activated cells, they were less dense and broken (Fig. 2*E*). Proteins identified in the PBMC data sets would thus likely include proteins involved in chromatin remodeling, and some of these proteins might be highlighted by differences between proteins identified in the unstimulated and PHA-activated PBMC NE fractions.

Microsomes were prepared separately as a comparative/subtractive fraction as they are rich in the main expected transmembrane protein contaminants from the ER, which is continuous with the NE. Thus, they provide a mechanism to distinguish INM proteins and ONM-enriched proteins from those such as ribosomes that inhabit both the ONM and ER. Because isolated lymphocytes have very little ER, we used rodent liver for production of microsomes. Actual contamination of NE fractions with ER was minimal because, although some single membrane vesicles can be observed by electron microscopy (see Fig. 1*D*), ER markers such as calreticulin and calnexin were undetectable in the crude PBMC NE fractions (Fig. 3). Conversely, NE proteins such as lamin A and the NET LAP2β were not detectable in microsome fractions (Fig. 3), indicating that the hypotonic lysis step only disrupted the plasma membrane and did not contribute NE membranes to the microsome fraction.

To remove donor variation as a factor in potential differences measured between unstimulated and PHA-activated PBMCs, each PBMC preparation was divided in two: half was

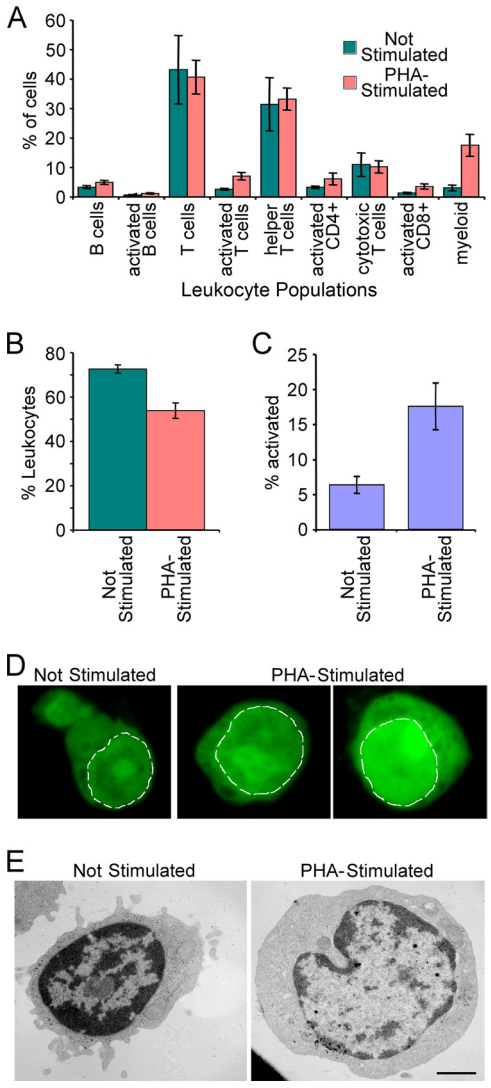


FIG. 2. **PBMC composition.** *A*, cells treated as for NE purification were incubated with CD markers for different blood cell types and analyzed by flow cytometry. The percentage of each cell type for PBMC composition in leukocytes is graphed. *B*, percentage of leukocytes in the total population. *C*, percentage of lymphocyte activation in the different PBMC populations measured by appearance of surface markers. *Error bars* in *A*, *B*, and *C* indicate standard deviation. *D*, PHA activation gauged by changes in nucleotide distribution using acridine orange dye. Acridine orange intercalates more strongly with single-stranded DNA than double-stranded DNA, so intensity to some degree measures "open" DNA that is being actively transcribed. A *dashed white line* delineates the periphery of the nucleus. The unstimulated PBMCs have weak staining at the periphery of the nucleus and no staining in some central areas. PHA-activated lymphocytes have more uniform distributions and brighter staining, indicating more open DNA. *E*, PHA activation effects on chromatin organization assessed by electron microscopy. The unstimulated cell is smaller and has more compacted chromatin concentrated at the nuclear periphery. In the PHA-activated cell, much of this dense peripheral chromatin has become less compact, and both the cell and nucleus have increased in volume as the activated cells are now rapidly transcribing DNA and making protein. Images were taken at 11,000×. *Scale bar*, 1 μm.

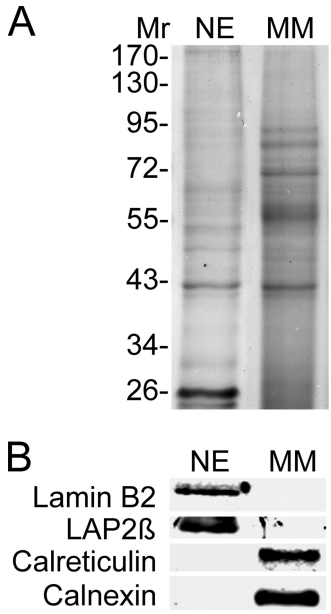


FIG. 3. **Fraction purity.** *A*, Coomassie-stained gel of NE and microsome fractions analyzed. *B*, Western blot of the above fractions stained with organelle markers. ER markers calreticulin and calnexin were absent from NEs, whereas NE markers lamin B2 and the NET LAP2β were absent from microsomes. Similar amounts of total protein were loaded. *MM*, microsomal membranes.

treated with PHA, and the other half was used as unstimulated. Some degree of variation between blood donors was expected because of the wide variation in human immune responses, so donor variation was averaged out within each condition by combining NEs isolated from at least 15 separate preparations for analysis. Combining this large number of samples was necessary for the independent reason that the vast majority of material was lost in isolating clean NEs from PBMCs. In this study, it was not possible to further control for variation due to factors such as age or racial differences as all blood donor information was anonymous, but all buffy coats were obtained within 2 days of donation. The combined NE preparations were divided and extracted with either 0.1 M NaOH or 1% β-octyl glucoside, 400 mM NaCl, and insoluble material was pelleted. The alkali extraction should have enriched for transmembrane proteins, and so vesicles were pelleted at 150,000 × *g* for 30 min, whereas the detergent/salt extraction should have removed membranes and enriched for both soluble and transmembrane proteins strongly associated with the higher order intermediate filament lamin polymer; these proteins were pelleted at 15,000 × *g* for 30 min.

MudPIT Analysis—After extraction, pellets of isolated NE fractions were digested with either a combined cyanogen bromide-endoproteinase Lys-C-trypsin treatment or with a combined endoproteinase Lys-C-trypsin treatment. For the latter, treatment material that could be pelleted at 17,500 × *g* for 30 min was subsequently digested with proteinase K. The multiple and processive digestions were favored over direct

replicates on the limiting material to better access material tightly associated with the intermediate filament lamin polymer that is resistant to 2% Triton X-100, 2 M NaCl, and even 4 M urea. Indeed, additional identifications were made from the proteinase K-digested material (supplemental Table S2B). Five separate MudPIT (27, 28) runs were performed for each of the unstimulated and PHA-activated PBMC NE fractions, three for the NaOH-extracted and two for the salt/detergent-extracted NEs. Thus, using different extraction conditions and different digestions and performing multiple runs were expected to provide comprehensive coverage to identify all proteins in the PBMC NE. After removing redundancy by converting protein IDs to orthologous gene groups, 3351 proteins were identified between the unstimulated and PHA-activated PBMC NE data sets, 2923 from the unstimulated and 2839 from the PHA-activated cells, 2411 of which were in both (supplemental Table S2B).

To complement previously published microsome data sets isolated from mouse liver (21), microsome data sets were separately generated from rat liver as a subtractive fraction (supplemental Table S4). These newly acquired microsome factions were treated similarly to the NE fractions, *i.e.* digested with trypsin and proteinase K and analyzed on LTQ ion trap mass spectrometers (supplemental Table S1). Another potential contaminant is mitochondria, but as the mitochondrial proteome has been extensively studied, lists of mitochondrial proteins were generated from a previous mass spectrometry study of that organelle (45) for subtraction instead of reanalyzing mitochondria (supplemental Table S5). After subtracting 177 proteins that were previously identified in mitochondria, 2756 proteins remained in unstimulated PBMC NEs, 2760 remained in PHA-activated PBMC NEs, and 3174 remained in total. Using microsomes strictly as an absolute subtractive fraction would have resulted in loss of several well characterized NETs that have been demonstrated to inhabit both the NE and ER. For example, emerin, one of the most established INM NETs (46), was detected in both NE and microsome data sets (supplemental Table S5) in agreement with the fact that emerin also functions in the ONM, ER, and interstitial discs (47–49). Moreover, as a recent elegant proteomics study suggests that at least a third of all proteins have multiple cellular localizations (50), we reasoned that absolute subtraction would also remove relevant identifications. Thus, proteins were ordered according to an estimate of abundance based on normalized spectral counts (dNSAF) (35), and NE data sets were restricted to those proteins either not detected in the microsomes or at least 5× more abundant than in the microsome fractions based on their NSAF values, leaving 2183 proteins in unstimulated PBMC NEs, 2141 in PHA-activated, and 2542 in total (supplemental Table S5).

This large number of proteins associated with the NE includes chromatin proteins and other factors that are tethered directly or indirectly to both INM and ONM proteins and so, although they can associate with the NE, are not necessarily primarily resident at the NE. The subset of NETs is more likely to be enriched at the NE and constitutes the surest NE core components in PBMCs. These proteins include all 13 of the previously well characterized NETs, 39 NETs previously detected in the liver NE data sets (21), and 423 proteins predicted to contain at least one transmembrane domain by TMHMM or membrane anchor by SignalP. Based on available GO terms and annotations, we selected 87 of these proteins as putative blood NETs (supplemental Table S5). It is of note that five of 12 proteins we further characterized for NE targeting (see below) did not pass the criterion of 5-fold enrichment over microsomes once the new rat liver microsome data set was acquired. However, NE residence was confirmed for all five, indicating that our selection criteria were very stringent and that additional NETs are likely to be found in this data set.

Confirmation of NE Residence—The putative NETs identified could be in the INM and/or in the ONM that is continuous with the ER or be contaminants. To directly test the validity of novel protein identifications, 12 putative NETs from PBMC data sets were cloned as tagged fusions and expressed in cells to determine whether they target to the NE. Adherent tissue culture cells were used instead of freshly isolated PBMCs because lymphocytes are very difficult to transfect, and their small cytoplasm made it difficult to distinguish the NE from the ER. To be able to clearly distinguish the NE from the ER and other cell membranes, adherent flat HT1080 cells were used for this analysis because lymphocytes are round suspension cells with little ER that is packed closely around the nucleus. Another issue is that, when overexpressed, NETs tend to saturate binding sites at the NE and accumulate in the ER, making clear confirmation of NE targeting difficult (Fig. 4*A*, *left panels*). To get around this problem, most of this extraneous material was removed with a prefixation detergent extraction. Most previously characterized NETs resist a prefixation extraction with 1% Triton X-100 and 400 mM KCl presumably because of tight associations with the intermediate filament lamin polymer (*e.g.* emerin; Fig. 4*A*, *top panels*), whereas most ER proteins (*e.g.* calreticulin; Fig. 4*A*, *middle panels*) are extracted with this treatment. Staining for lamins is also shown as co-localization (*yellow*) with this well defined NE protein, indicating NE localization. One of the novel NETs identified here, C20orf3, behaved similarly to the well characterized NET emerin in this assay (Fig. 4*A*, *bottom panels*). In all, ten of the 12 putative NETs tested were retained at the NE after a prefixation detergent extraction (Fig. 4; all except Tmem126A and METTL7A). The putative NETs Tmem41A and Tmem126A both failed to be expressed in the HT1080 cells first tested, but Tmem41A was expressed and targeted to the NE in COS-7 cells and resisted the detergent pre-extraction (Fig. 4*B*, *bottom left*). Tmem126A also failed to be expressed in COS-7 cells but was expressed and targeted in lymphocytes, which due to their small size and being suspension cells could not be properly tested for resistance to detergent pre-extraction (Fig. 4*B*, *bottom right*). METTL7A exhibited a

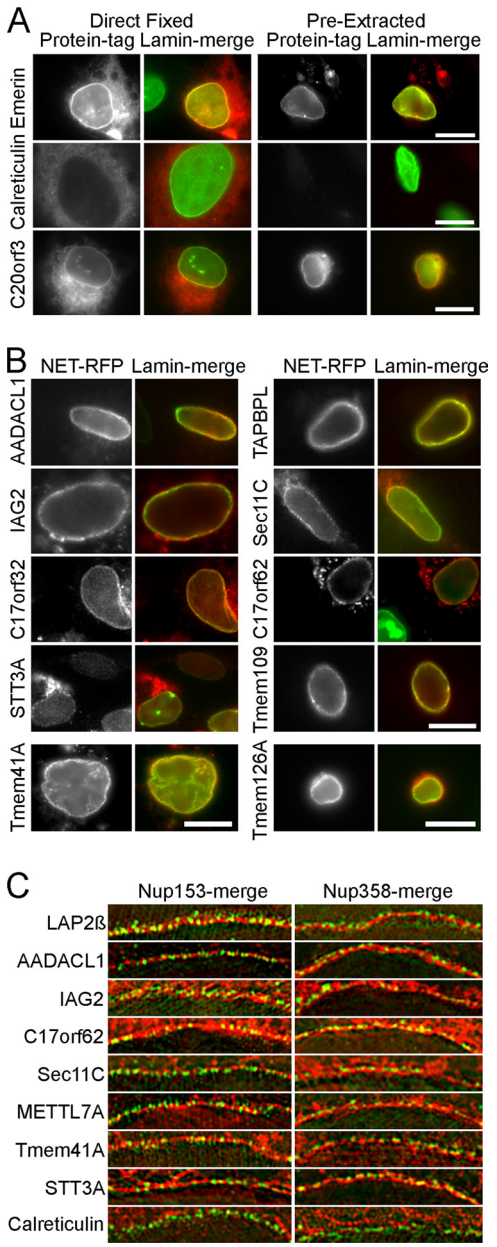


FIG. 4. **mRFP fusions confirm NE targeting for several novel PBMC NETs.** *A*, HT1080 cells expressing NETs fused to mRFP were directly fixed (*left*) or extracted with Triton X-100 prior to fixation (*right*). The NE is marked by lamin A in *green* so that *yellow* indicates co-localization of the NET at the NE. The directly fixed emerin image has part of an untransfected cell, confirming that none of the NET staining at the nuclear rim is due to bleed-through from the lamin A channel. Note that prefixation extraction affects morphology and sometimes leaves aggregated protein in the cytoplasm. The emerin control and new NET C20orf3 are retained after extraction, whereas the ER protein calreticulin is not. *Scale bar*, 10 μm. *B*, other NET-mRFP fusions were similarly pre-extracted and retained at the NE. Tmem41A is shown in COS-7 cells because it was not expressed in HT1080 cells, and Tmem126A is shown in Jurkat cells because it failed to be expressed in either HT1080 or COS-7 cells. *Scale bar*, 10 μm. *C*, inner *versus* outer nuclear membrane targeting. If a NET (*red*) is in the INM it should appear in the same plane as the nuclear basket protein Nup153 (*green*, *left*) and internal to the cytoplasmic filament protein Nup358 (*green*, *right*) using structured illumination microscopy. Characterized NET LAP2β and most new NETs tested appeared in the INM. One INM NET, METTL7A, did not resist Triton pre-extraction. *Scale bars*, 5 μm.

concentration at the nuclear periphery that co-localized with lamins but was never observed after a pre-fixation detergent extraction (data not shown).

As the NE is comprised equally of INM and ONM, some NETs may preferentially accumulate in one or the other subcompartment. A recent study used three-dimensional structured illumination microscopy (OMX) to distinguish INM from ONM localization of lamins in cells when co-stained for proteins from the nuclear pore nucleoplasmic (Nup153) or cytoplasmic (Nup358) face (51). We have previously used this system to test several NETs identified in the earlier liver proteomics data sets (52). Control NET LAP2β and all the new NETs tested here accumulated in the INM as both the NET (*red*) and Nup153 (*green*) localize to the same ring, whereas an inner NET ring was observed compared with Nup358 (Fig. 4*C*). In contrast to the NETs, a separable inner Nup153 ring was observed against the ER protein calreticulin. *Yellow* co-localization was not generally observed in this system because the resolution of OMX microscopy is sufficient to distinguish the NPC proteins that are at ~120-nm gaps in the membrane from NETs that are inserted in the membrane itself. Even METTL7A, the putative NET that was not resistant to the detergent pre-extraction, appeared in the INM.

To test localization of endogenous proteins, polyclonal antibodies were generated to three novel NETs (C17orf32, Tmem126A, and C17orf62). Each recognized a protein of the correct size in immunoblots (Fig. 5*A*). First, to test whether NETs were principally in the NE or inhabited multiple cellular localizations as has been indicated for roughly 40% of the human proteome (50), antibodies were used on Western blots comparing NE and microsome fractions. Tmem126A and C17orf62 were exclusively in the NE fraction, but about a third of C17orf32 occurred in the microsome fraction (Fig. 5*B*). This is similar to the previously characterized NET31/Tmem209 that was detected in both the PBMC NE data sets and the earlier liver NE data set (52).

The Tmem126A and C17orf62 antibodies yielded strong and distinctive staining at the nuclear periphery when used on cryosections from rat spleen (Fig. 5*C*). The nuclei are clearly visualized with the DAPI DNA stain, and the antibody staining is concentrated in a crisp rim just surrounding the nucleus that is very similar to that observed with antibodies to the well characterized NET SUN2. C17orf62 and C17orf32 both resisted pre-extraction with Triton X-100 when expressed as fusions in cell lines, and similarly, the antibodies to these NETs yielded a rim stain in cultured cells pre-extracted with detergent. Adherent tissue culture cells were used again because lymphocytes have little cytoplasm and do not with-

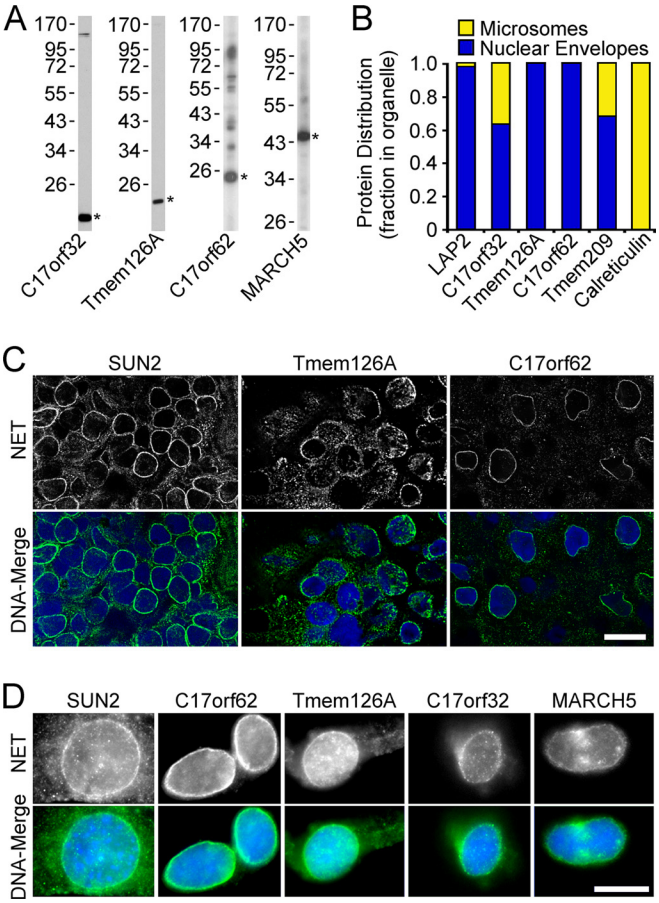


FIG. 5. **Antibody staining confirms novel NET identifications.** *A*, NET antibody validation by Western blotting. C17orf32 and Tmem126A antibodies were tested on human PBMC lysates, whereas C17orf62 antibodies were tested on a C2C12 lysate to check background in the cell line used to test pre-extraction in *D*. MARCH5 antibodies were tested using the Jurkat human blood cell line. *Asterisks* indicate expected molecular weight. *B*, comparison of relative NET amounts in ER (microsome) and NE fractions by quantitative Western blotting. LAP2 and Tmem209 antibodies are NET controls, whereas calreticulin antibodies are an ER control. Blots were quantified using a LI-COR Odyssey system (three repeats averaged). *C*, cryosections of rat spleen stained with antibodies to new NETs and a SUN2 control. Nuclear rim staining was clearly observed for all NETs tested. *Scale bars*, 10 μm. *D*, antibody staining on Triton-pre-extracted C2C12 cells. A nuclear rim staining was observed for the control NET SUN2 and all novel NETs tested.

stand the prefixation extraction well. However, C2C12 cells were used instead of the HT1080 cells used earlier for targeting of fusion proteins. According to available transcriptome data, all the NETs for which we had antibodies were at least weakly expressed in C2C12 cells, whereas some were not expressed at all in HT1080 cells. This allowed confirmation that Tmem126A also resists detergent pre-extraction (Fig. 5*D*). Several proteins identified in these data sets were previously characterized proteins that had other reported cellular localizations. Among these was MARCH5, previously reported to be a mitochondrial protein that possibly contributes to perinuclear localization of mitochondria (53). Although widely staining throughout the cytoplasm in non-extracted cells (data not shown), antibodies to MARCH5 (Fig. 5*A*) also yielded nuclear rim staining after the detergent pre-extraction (Fig. 5*D*, *right panels*), indicating that the endogenous protein has a separate function at the NE. Thus MARCH5 is among the estimated 40% of cellular proteins that have multiple cellular localizations (50).

Changes in PBMC NE Composition following PHA Activation—Of the 87 core putative new PBMC NETs (supplemental Table S5), 14% uniquely appeared in the unstimulated PBMC NE fraction, and 14% uniquely appeared in the PHA-activated PBMC NE fractions with 72% shared between both conditions. These numbers were similar when considering all transmembrane proteins identified (73% shared) or both transmembrane and soluble proteins identified (72% shared). Differences between the unstimulated and PHA-stimulated fractions were also observed in RNA expression levels for NETs that appeared in both fractions that correlated with peptide recovery. For example, four peptides were recovered for AADACL1 in the PHA-activated *versus* seven in the unstimulated lymphocytes (Fig. 6*A* and supplemental Table S2B), and correspondingly, AADACL1 transcript levels were reduced in PHA-activated lymphocytes to 60% of the level in unstimulated lymphocytes (Fig. 6*B*). The trends for expression changes also agreed with estimated protein levels for C17orf62, TAPBPL, and FAM3C.

Many of the novel NETs identified that were directly tested here varied not only between the unstimulated and PHA-activated leukocytes but appeared to be uniquely expressed in different blood cell lineages. Extracting relative transcript levels from a large scale transcriptome data set that compared RNA transcript levels across multiple tissues (41), most of the novel NETs had completely distinct patterns of expression among 10 different blood cell types ranging from erythroid cells to B-cells (Fig. 6*C*). Only IAG2 and Sec11C of those shown had similar general patterns of expression, and only Tmem41A and C17orf32 (not shown) were expressed similarly in all lineages. The data are expressed relative to a median value that was generated from a much wider range of tissues, and these NETs ranged from ~10× the median value to nearly 60× (Fig. 6*C*). Thus, many of the novel NETs identified from the PBMC proteomics are reasonably restricted in the tissues and cell types in which they are expressed. The interpretation of these results is tempered by the observation that some changes in the population composition occur upon PHA activation (Fig. 2*A*). Myeloid cells were found to increase, and this is consistent with some of the NETs being more specific to myeloid dendritic cells (Fig. 6*C*); however, the percentage of myeloid cell-associated proteins at the NE did not increase. This could be attributed to having already made a full identification of myeloid cell-associated proteins in the unstimulated cells. Interestingly, when considered as a percentage of the total myeloid cell-associated proteins, roughly 10% occur

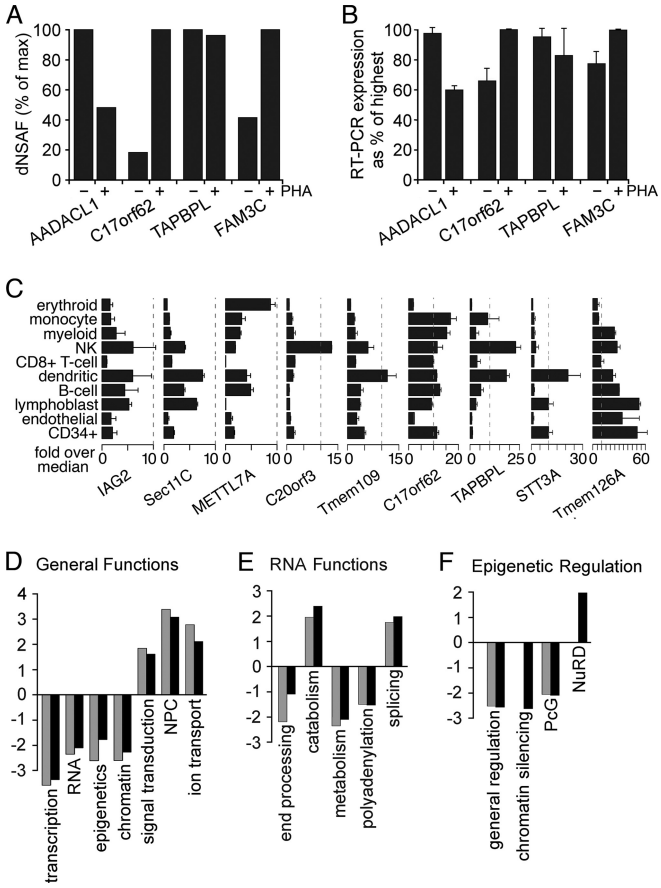


FIG. 6. **Differences in NE composition with PHA activation and effects on genome organization.** *A*, differences in abundance estimated from dNSAF values for some NETs with or without PHA activation. dNSAF values were taken from the two most equivalent runs based on total protein coverage and identifications. *B*, RT-PCR of the NETs in *A* revealed reproducible differences in expression that are consistent with their abundance estimates based on peptide recoveries. *C*, transcript levels for several of the NETs identified differ among blood cell lineages. Data from the BioGPS transcriptome study comparing 84 different tissues/cell types are plotted relative to the median value over all 84 tissues sampled. *Ticks* along the *bottom* are for increments of 5-fold above the median. Each NET had its own unique expression pattern. Surface markers defining cell populations were CD71 for early erythroid cells, CD14 for monocytes, CD33 for myeloid cells, CD56 for natural killer (*NK*) cells, CD8 for T-cells, BDCA4 for myeloid dendritic cells, CD19 for B-cells, 721 for B-lymphoblasts, CD105 for endothelial cells, and CD34 for polyploidy progenitor cells. *Error bars* in *B* and *C* indicate standard deviation. *D*, within the subset of proteins in NE data sets with GO annotations, the fraction with a particular functional annotation was calculated. Similar fractions were calculated against all "nuclear"-annotated proteins in the GO database. The ratio of NE/nuclear fractions was then calculated, setting a 1:1 ratio to 0 so that positive values are -fold enrichment and negative values are -fold deficiency at the NE compared with the whole nucleus. Resting and activated PBMC data sets are represented by *gray* and *black bars*, respectively. *E*, the same analysis applied to the specific subset associated with RNA functions. *F*, the same analysis applied to the specific subset associated with functions in epigenetic regulation. *PcG*, Polycomb group.

at the NE, indicating this organelle as a dynamic player in cell type specification.

PHA-induced changes include physiologically relevant functions such as re-entering the active cell cycle (43) and altering peripheral chromatin organization (17). Individual NETs could contribute to chromatin organization and transcriptional regulation to facilitate one of the individual lineages observed in Fig. 6*C*. To gauge the likelihood of NE contributions to genome functions, the known proteins in the NE data sets (both NETs and soluble proteins) that had functional annotations in the GO database were extracted. A separate data set was extracted for all proteins in the GO database associated with the nucleus. GO functional annotations were then applied to these data sets, and the fraction of proteins within each data set with a particular GO functional annotation was calculated. These fractions were compared with one another to give the -fold enrichment of a particular function at the NE compared with that function as a fraction of activity within the nucleus in general (Fig. 6, *D–F*). For example, transcription takes place throughout the nucleus, and the nuclear periphery is considered to be a generally transcriptionally repressive environment (54): the fraction of proteins identified at the NE involved in transcription, as thus expected, is much lower than that for the nucleus, yielding a negative -fold value of −3.5 (Fig. 6*D*). In contrast, proteins involved in nucleocytoplasmic transport would be expected to be enriched at the NE compared with the nucleus as a whole, and the value obtained for NPC proteins showed enrichment of 3.5-fold at the NE (Fig. 6*D*).

The general functional categories varied little between unstimulated and PHA-activated PBMC NEs; however, analysis of more specific subcategories revealed interesting differences. RNA functions, as might be expected, varied little upon PHA activation except that a few more end-processing proteins were found at the NE (Fig. 6*E*), consistent with an increased burden of RNAs to be exported from the nucleus with the significant gene activation resulting from PHA treatment. Functions in epigenetic regulation yielded more dynamic changes upon leukocyte activation (Fig. 6*F*). Chromatin silencing functions were less enriched at the NE with PHA activation, clearly indicating that changes at the NE play a role in the major chromatin reorganization that takes place. NuRD complex proteins in contrast were specifically enriched at the NE upon PHA activation.

Novel NET Functions in Chromatin/Genome Organization—The finding of so many known proteins with chromatin regulatory functions in these data sets that change between the data sets in ways that match the physiological changes observed in chromatin upon PHA activation argues that some novel proteins identified might also have chromatin functions. To screen for such NETs, we tested nine PBMC NETs in assays for effects on genome organization.

The first assay utilized an HT1080-derived cell line carrying a lacO repeat insertion in chromosome 5, the position of

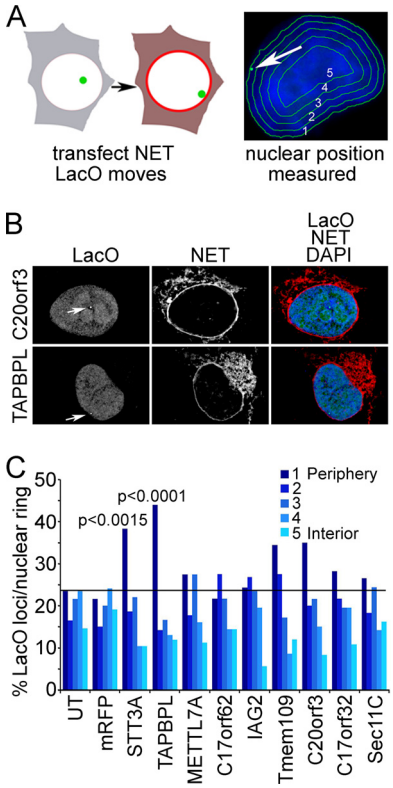


FIG. 7. **Cell-based screen for PBMC NETs that promote recruitment of chromosome loci to nuclear periphery.** *A*, schematic representation of a screen to determine whether overexpression of a particular NET can recruit a specific chromatin locus to the NE. NETs were transfected into cells containing a lacO repeat integration that is typically in the interior. The lacO position was visualized with GFP-lacI (*green*) and measured using an algorithm that partitions the nucleus based on DAPI staining (*blue*) into five concentric circles of roughly equal area. *B*, example of NET-transfected cells. The position of the lacO locus is highlighted by the *white arrows*. The lacO locus position is unaffected by C20orf3 expression but moves to the periphery with TAPBPL expression. DAPI staining added to the merged image confirms that the movement of the locus is not an artifact of generalized chromatin condensation at the periphery. Deconvolved images are shown. *C*, the ring containing the locus was recorded in roughly 100 transfected cells. *p* values were calculated for NETs that increased the locus at the periphery in comparison with untransfected (*UT*) or mRFP-transfected control cells using a χ^2 test.

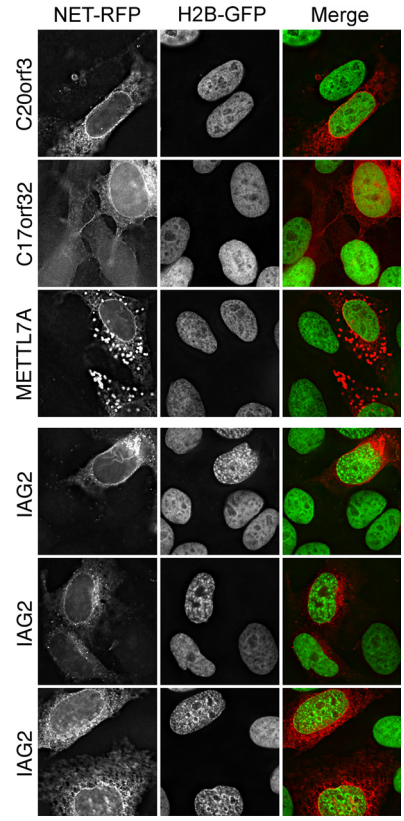


FIG. 8. **Screen for PBMC NETs that promote chromatin compaction.** HeLa cells stably expressing H2B-GFP were transfected with the same set of nine NETs used in Fig. 7. Only IAG2 displayed strong chromatin condensation, whereas all other NETs (METTL7A, C17orf32, and C20orf3 are shown) exhibited no differences in chromatin compaction or distribution compared with adjacent untransfected cells. It is noteworthy that IAG2 had no effect in the lacO screen and that NETs STT3A and TAPBPL, which altered lacO positioning, had no effect on chromatin compaction.

which is visualized with a lac repressor-GFP fusion that binds to the array. This lacO array tends to be in the nuclear interior. These cells were transfected with expression constructs for NET-mRFP fusions and assayed for a change in the distribution of the lacO array position with respect to the nuclear periphery. If a NET altered the position of the lacO array by pulling it to the nuclear periphery it could be considered to function in genome organization (Fig. 7*A*, *left*). To determine lacO positioning, a previously published algorithm was utilized that erodes the nucleus based on DAPI staining for DNA into five concentric rings of roughly equal area (55). The ring containing the lacO array was marked in each cell and recorded (Fig. 7*A*, *right*). The NET C20orf3 had no effect on the position of the lacO array; however, TAPBPL expression increased the occurrence of the array at the periphery (Fig. 7*B*). The ring containing the lacO array was marked in ~100 cells, and the distribution was plotted (Fig. 7*C*). The lacO array tended to be in the nuclear interior, occurring in the peripheral ring only ~20% of the time. Most blood NETs tested yielded no significant differences, but in cells expressing STT3A and TAPBPL, the occurrence at the periphery roughly doubled in multiple separate experiments with *p* values <0.0015 and 0.0001, respectively (Fig. 7*C*).

For the second assay, the same set of blood NETs was transfected into a HeLa cell line stably expressing H2B-GFP, and cells were observed for general changes in overall chromatin distribution. The two NETs that had strong effects in the lacO screen had no effect in this screen, further indicating that the interactions pulling the locus to the periphery are specific. A different NET, IAG2, had a strong effect on chromatin condensation (Fig. 8; compare with cells transfected with C20orf3, C17orf32, and METTL7A in *upper panels*), and this effect was repeated in all of three independent experiments. A

ring of condensed chromatin is noticeable at the periphery, although the effect has clearly propagated throughout the nucleus. Distribution of the NET alone in the *left panels* clearly indicates that, whereas some of the overexpressed protein has accumulated in the ER, none is in the nucleoplasm: thus, the effect must be propagated from the NE.

DISCUSSION

This analysis of the leukocyte NE proteome identified many more NE-associated proteins and NETs than an earlier study of liver NEs (21). Moreover, some proteome differences were observed between unstimulated and PHA-activated leukocytes, many of which correspond to proteins that are expressed in only specific blood cell lineages. Many of these new NETs appear to be tissue-specific, and our results further show that NE protein composition changes in response to chemical stimulation. These findings indicate the potential of other organelle proteomes to vary with chemical treatments or between tissues.

Large scale changes have been reported in the organization of leukocyte peripheral chromatin upon PHA stimulation (17–20). These data sets should contain both NETs and tightly bound chromatin proteins involved in making these changes. As several NETs have known associations with chromatin-silencing proteins (12, 13), it is not surprising that several of the proteins identified here in association with the NE had functions in gene regulation and RNA processing. The PHA-induced general decrease in epigenetic silencing factors is consistent with the observed reduction in dense peripheral chromatin. Contrasting with this, the specific relative increase in NuRD complex proteins at the NE is interesting as this chromatin remodeling complex plays roles in development and was recently shown to function in progeria defects caused by NE mutations (56).

NET developmental roles are consistent with the blood cell type specificity of the two NETs that affected positioning of the lacO array (~25-fold higher in blood cells over other tissues). Several tissue-specific genes move from the NE to the nuclear interior during differentiation (*e.g. IgH* locus, *Mash1*, and *CFTR*) (57–59); however, the NE proteins that regulate NE tethering and release have yet to be identified. Recent studies showed that tethering lac repressor to an NE protein could recruit the lacO array to the periphery by an affinity-based mechanism (14–16). As lac repressor was *not* fused to the NETs used in our screen, the specific effects of TAPBPL and STT3A likely operate through affinity for other chromosome regions on the same chromosome as the lacO array. This is consistent with the earlier studies showing that the whole chromosome accompanied the array to the periphery (14, 16). The interaction cannot be due to a more general affinity for chromatin because if so it would increase chromatin density at the periphery, and these two NETs had no effect in the second screen for chromatin compaction. Thus, our identification here of endogenous NETs that can alter the positioning of genes/chromosomes in the nucleus appears to reflect the unique chromatin characteristics of lymphocytes. As several studies have implicated misregulation of chromatin organization in NE diseases (7, 8), these new blood NETs may contribute to the diverse pathologies associated with NE diseases.

Acknowledgments—We thank D. Tollervey and W. C. Earnshaw for comments; W. Bickmore, F. Melchior, T. Rapoport, J. Stewart, M. Waterfall, E. King, D. Kelly, T. Guan, and S. Mitchell for reagents/technical assistance; and the Scottish National Blood Transfusion Service.

* This work was supported in part by the Wellcome Trust (a senior research fellowship to E. C. S.) and the Stowers Institute for Medical Research (to L. F.). Use of the OMX microscope was supported by the Scottish University Life Sciences Alliance.

S This article contains supplemental Tables S1–S5.

¶ A Darwin Trust student.

‖ A Royal Society Dorothy Hodgkin fellow.

** To whom correspondence should be addressed: Wellcome Trust Centre for Cell Biology, University of Edinburgh, Kings Bldgs., Swann 5.22, Mayfield Rd., Edinburgh EH9 3JR, UK. Tel.: 441316507075; Fax: 441316507360; E-mail: e.schirmer@ed.ac.uk.

REFERENCES

1. Schirmer, E. C., and Foisner, R. (2007) Proteins that associate with lamins: many faces, many functions. *Exp. Cell Res.* **313,** 2167–2179
2. Suntharalingam, M., and Wente, S. R. (2003) Peering through the pore: nuclear pore complex structure, assembly, and function. *Dev. Cell* **4,** 775–789
3. Callan, H. G., and Tomlin, S. G. (1950) Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. *Proc. R. Soc. Lond. B Biol. Sci.* **137,** 367–378
4. Starr, D. A., and Fischer, J. A. (2005) KASH 'n Karry: the KASH domain family of cargo-specific cytoskeletal adaptor proteins. *BioEssays* **27,** 1136–1146
5. Stewart, C. L., Roux, K. J., and Burke, B. (2007) Blurring the boundary: the nuclear envelope extends its reach. *Science* **318,** 1408–1412
6. Worman, H. J., and Bonne, G. (2007) "Laminopathies": a wide spectrum of human diseases. *Exp. Cell Res.* **313,** 2121–2133
7. Fidziańska, 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
8. Goldman, R. D., Shumaker, D. K., Erdos, M. R., Eriksson, M., Goldman, A. E., Gordon, L. B., Gruenbaum, Y., Khuon, S., Mendez, M., Varga, R., and Collins, F. S. (2004) Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 8963–8968
9. Maraldi, N. M., Squarzoni, S., Sabatelli, P., Lattanzi, G., Ognibene, A., and Manzoli, F. A. (2002) Emery-Dreifuss muscular dystrophy, nuclear cell signaling and chromatin remodeling. *Adv. Enzyme Regul.* **42,** 1–18
10. Ognibene, A., Sabatelli, P., Petrini, S., Squarzoni, S., Riccio, M., Santi, S., Villanova, M., Palmeri, S., Merlini, L., and Maraldi, N. M. (1999) Nuclear changes in a case of X-linked Emery-Dreifuss muscular dystrophy. *Muscle Nerve* **22,** 864–869
11. 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
12. Ye, Q., and Worman, H. J. (1996) Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. *J. Biol. Chem.* **271,** 14653–14656
13. Somech, R., Shaklai, S., Geller, O., Amariglio, N., Simon, A. J., Rechavi, G., and Gal-Yam, E. N. (2005) The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. *J. Cell Sci.* **118,** 4017–4025

14. Finlan, L. E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J. R., and Bickmore, W. A. (2008) Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* **4,** e1000039
15. Kumaran, R. I., and Spector, D. L. (2008) A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J. Cell Biol.* **180,** 51–65
16. Reddy, K. L., Zullo, J. M., Bertolino, E., and Singh, H. (2008) Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452,** 243–247
17. Drings, P., and Sonnemann, E. (1974) Phytohemagglutinin-induced increase of euchromatin contents in human lymphocytes. *Res. Exp. Med.* **164,** 63–76
18. Hirschhorn, R., Decsy, M. I., and Troll, W. (1971) The effect of PHA stimulation of human peripheral blood lymphocytes upon cellular content of euchromatin and heterochromatin. *Cell. Immunol.* **2,** 696–701
19. Manteĭfel', V. M., Andreĭchuk, T. N., and Karu, T. I. (1992) A comparative study of chromatin from lymphocyte nuclei upon activation of transcription by irradiation from an He-Ne-laser or phytohemagglutinin. *Mol. Biol.* **26,** 1054–1062
20. Pompidou, A., Rousset, S., Macé, B., Michel, P., Esnous, D., and Renard, N. (1984) Chromatin structure and nucleic acid synthesis in human lymphocyte activation by phytohemagglutinin. *Exp. Cell Res.* **150,** 213–225
21. Schirmer, E. C., Florens, L., Guan, T., Yates, J. R., 3rd, and Gerace, L. (2003) Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* **301,** 1380–1382
22. Dreger, M., Bengtsson, L., Schöneberg, 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
23. Korfali, N., Fairley, E. A., Swanson, S. K., Florens, L., and Schirmer, E. C. (2009) Use of sequential chemical extractions to purify nuclear membrane proteins for proteomics identification. *Methods Mol. Biol.* **528,** 201–225
24. Walter, P., and Blobel, G. (1983) Preparation of microsomal membranes for cotranslational protein translocation. *Methods Enzymol.* **96,** 84–93
25. Florens, L., Korfali, N., and Schirmer, E. C. (2008) Subcellular fractionation and proteomics of nuclear envelopes. *Methods Mol. Biol.* **432,** 117–137
26. Florens, L., and Washburn, M. P. (2006) Proteomic analysis by multidimensional protein identification technology. *Methods Mol. Biol.* **328,** 159–175
27. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19,** 242–247
28. Wolters, D. A., Washburn, M. P., and Yates, J. R., 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **73,** 5683–5690
29. McDonald, W. H., Tabb, D. L., Sadygov, R. G., MacCoss, M. J., Venable, J., Graumann, J., Johnson, J. R., Cociorva, D., and Yates, J. R., 3rd (2004) MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Commun. Mass Spectrom.* **18,** 2162–2168
30. Venable, J. D., Dong, M. Q., Wohlschlegel, J., Dillin, A., and Yates, J. R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* **1,** 39–45
31. Eng, J., McCormack, A., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989
32. Tabb, D. L., McDonald, W. H., and Yates, J. R., 3rd (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1,** 21–26
33. Zybailov, B. L., Florens, L., and Washburn, M. P. (2007) Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. *Mol. Biosyst.* **3,** 354–360
34. Florens, L., Carozza, M. J., Swanson, S. K., Fournier, M., Coleman, M. K., Workman, J. L., and Washburn, M. P. (2006) Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **40,** 303–311
35. Paoletti, A. C., Parmely, T. J., Tomomori-Sato, C., Sato, S., Zhu, D., Conaway, R. C., Conaway, J. W., Florens, L., and Washburn, M. P. (2006) Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 18928–18933
36. Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L., and Washburn, M. P. (2006) Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **5,** 2339–2347
37. Zhang, Y., Wen, Z., Washburn, M. P., and Florens, L. (2010) Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal. Chem.* **82,** 2272–2281
38. Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Gräf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kähäri, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A., and Searle, S. (2008) Ensembl 2008. *Nucleic Acids Res.* **36,** D707–D714
39. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305,** 567–580
40. Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., and Lewis, S. (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics* **25,** 288–289
41. 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., Patapoutian, A., Hampton, G. M., Schultz, P. G., and Hogenesch, J. B. (2002) Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 4465–4470
42. 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
43. Hirschhorn, K., Bach, F., Kolodny, R. L., Firschein, I. L., and Hashem, N. (1963) Immune Response and Mitosis of Human Peripheral Blood Lymphocytes in Vitro. *Science* **142,** 1185–1187
44. Ichimura, S., Zama, M., and Fujita, H. (1971) Quantitative determination of single-stranded sections in DNA using the fluorescent probe acridine orange. *Biochim. Biophys. Acta* **240,** 485–495
45. Mootha, V. K., Bunkenborg, J., Olsen, J. V., Hjerrild, M., Wisniewski, J. R., Stahl, E., Bolouri, M. S., Ray, H. N., Sihag, S., Kamal, M., Patterson, N., Lander, E. S., and Mann, M. (2003) Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115,** 629–640
46. 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
47. Cartegni, L., di Barletta, M. R., Barresi, R., Squarzoni, S., Sabatelli, P., Maraldi, N., Mora, M., Di Blasi, C., Cornelio, F., Merlini, L., Villa, A., Cobianchi, F., and Toniolo, D. (1997) Heart-specific localization of emerin: new insights into Emery-Dreifuss muscular dystrophy. *Hum. Mol. Genet.* **6,** 2257–2264
48. Lattanzi, G., Ognibene, A., Sabatelli, P., Capanni, C., Toniolo, D., Columbaro, M., Santi, S., Riccio, M., Merlini, L., Maraldi, N. M., and Squarzoni, S. (2000) Emerin expression at the early stages of myogenic differentiation. *Differentiation* **66,** 208–217
49. Salpingidou, G., Smertenko, A., Hausmanowa-Petrucewicz, I., Hussey, P. J., and Hutchison, C. J. (2007) A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane. *J. Cell Biol.* **178,** 897–904
50. Foster, L. J., de Hoog, C. L., Zhang, Y., Zhang, Y., Xie, X., Mootha, V. K., and Mann, M. (2006) A mammalian organelle map by protein correlation profiling. *Cell* **125,** 187–199
51. Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M. C., Agard, D. A., Gustafsson, M. G., Leonhardt, H., and Sedat, J. W. (2008) Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320,** 1332–1336
52. Malik, P., Korfali, N., Srsen, V., Lazou, V., Batrakou, D. G., Zuleger, N.,

Kavanagh, D. M., Wilkie, G. S., Goldberg, M. W., and Schirmer, E. C. (2010) Cell-specific and lamin-dependent targeting of novel transmembrane proteins in the nuclear envelope. *Cell Mol. Life Sci.* **67,** 1353–1369
53. Karbowski, M., Neutzner, A., and Youle, R. J. (2007) The mitochondrial E3 ubiquitin ligase MARCH5 is required for Drp1 dependent mitochondrial division. *J. Cell Biol.* **178,** 71–84
54. Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M., and van Steensel, B. (2006) Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nat. Genet.* **38,** 1005–1014
55. Croft, J. A., Bridger, J. M., Boyle, S., Perry, P., Teague, P., and Bickmore, W. A. (1999) Differences in the localization and morphology of chromosomes in the human nucleus. *J. Cell Biol.* **145,** 1119–1131
56. Pegoraro, G., Kubben, N., Wickert, U., Göhler, H., Hoffmann, K., and Misteli, T. (2009) Ageing-related chromatin defects through loss of the NURD complex. *Nat. Cell Biol.* **11,** 1261–1267
57. Kosak, S. T., Skok, J. A., Medina, K. L., Riblet, R., Le Beau, M. M., Fisher, A. G., and Singh, H. (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296,** 158–162
58. Williams, R. R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jørgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M., and Fisher, A. G. (2006) Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *J. Cell Sci.* **119,** 132–140
59. Zink, D., Amaral, M. D., Englmann, A., Lang, S., Clarke, L. A., Rudolph, C., Alt, F., Luther, K., Braz, C., Sadoni, N., Rosenecker, J., and Schindelhauer, D. (2004) Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *J. Cell Biol.* **166,** 815–825

Several Novel Nuclear Envelope Transmembrane Proteins Identified in Skeletal Muscle Have Cytoskeletal Associations*[S]

Gavin S. Wilkie‡§, Nadia Korfali‡§, Selene K. Swanson¶, Poonam Malik‡‖, Vlastimil Srsen‡, Dzmitry G. Batrakou‡, Jose de las Heras‡, Nikolaj Zuleger‡, Alastair R. W. Kerr‡, Laurence Florens¶, and Eric C. Schirmer‡ ‡‡**

Nuclear envelopes from liver and a neuroblastoma cell line have previously been analyzed by proteomics; however, most diseases associated with the nuclear envelope affect muscle. To determine whether muscle has unique nuclear envelope proteins, rat skeletal muscle nuclear envelopes were prepared and analyzed by multidimensional protein identification technology. Many novel muscle-specific proteins were identified that did not appear in previous nuclear envelope data sets. Nuclear envelope residence was confirmed for 11 of these by expression of fusion proteins and by antibody staining of muscle tissue cryosections. Moreover, transcript levels for several of the newly identified nuclear envelope transmembrane proteins increased during muscle differentiation using mouse and human *in vitro* model systems. Some of these proteins tracked with microtubules at the nuclear surface in interphase cells and accumulated at the base of the microtubule spindle in mitotic cells, suggesting they may associate with complexes that connect the nucleus to the cytoskeleton. The finding of tissue-specific proteins in the skeletal muscle nuclear envelope proteome argues the importance of analyzing nuclear envelopes from all tissues linked to disease and suggests that general investigation of tissue differences in organellar proteomes might yield critical insights. ***Molecular & Cellular Proteomics 10: 10.1074/mcp. M110.003129, 1–16, 2011.***

The nuclear envelope (NE)[1] is an impenetrable membrane barrier between the nucleus and the cytoplasm perforated by nuclear pore complexes (NPCs) that regulate transport of soluble macromolecules in and out of the nucleus (1, 2). Structurally, the NE consists of the outer nuclear membrane (ONM) that is continuous with the endoplasmic reticulum (ER) (3), a lumen, the inner nuclear membrane (INM), and associated proteins including the NPCs and the intermediate filament nuclear Lamin polymer (4). Both the ONM and the INM have unique sets of transmembrane proteins, sometimes called NETs for nuclear envelope transmembrane proteins. Lamins and several NETs have been linked to an increasing number of relatively rare diseases that range from forms of muscular dystrophy to neuropathy, dermopathy, lipodystrophy, bone disorders, and accelerated aging syndromes (5, 6).

The three favored molecular mechanisms to explain NE disease pathology are mechanical instability from disruption of Lamin-cytoskeleton interactions, altered expression of genes regulated from the nuclear periphery, and disabling of the cell cycle/stem cell maintenance (6, 7). All of these may involve additional associated proteins to produce pathology. Indeed, it would seem that some proteins must be missing from the system as those so far mutated in disease are widely expressed, yet each disease exhibits pathology in a particular subset of the tissues in which the protein is expressed. Because both gene regulation and cytoskeletal connections have been implicated, NE proteins involved could reside in either the INM or the ONM.

The mechanical stability of the NE is derived largely from the nuclear Lamin polymer. Loss of Lamins (8–10) or their mutation (11) greatly disturbs nuclear morphology and stability. This stability itself varies among different cell types as, for example, neutrophils have highly lobulated nuclei and have higher relative concentrations of the Lamin B2 subtype (12), which is the least stable of the different Lamin subtypes (13, 14). In contrast, muscle cells must withstand high shear forces and have high concentrations of Lamin A, the most stable subtype. Many NETs bind Lamins, some of which also make connections across the lumen of the NE to

From the ‡Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom and ¶Stowers Institute for Medical Research, Kansas City, Missouri 64110

Received, July 13, 2010, and in revised form, September 14, 2010
⌘ *Author's Choice*—Final version full access.
Published, MCP Papers in Press, September 27, 2010, DOI 10.1074/mcp.M110.003129

[1] The abbreviations used are: NE, nuclear envelope; NET, nuclear envelope transmembrane protein; INM, inner nuclear membrane; SR, sarcoplasmic reticulum; ONM, outer nuclear membrane; MudPIT, multidimensional protein identification technology; NPC, nuclear pore complex; ER, endoplasmic reticulum; EDMD, Emery-Dreifuss muscular dystrophy; qRT-PCR, quantitative RT-PCR; FDR, false discovery rate; NSAF, normalized spectral abundance factor; dNSAF, distributed normalized spectral abundance factor; CKM, creatine muscle kinase; MYOG, myogenin; MYH1, myosin heavy chain 1; Hs, homo sapiens; Mm, mus musculus.

ONM proteins (15), which in turn connect the NE to the cytoskeleton (16, 17).

Mutations in the *LMNA* gene encoding Lamins A and C cause forms of Emery-Dreifuss muscular dystrophy (EDMD) (18, 19), limb-girdle muscular dystrophy (LGMD-1B) (20), and dilated cardiomyopathy with conduction defect (CMD1A) (21), which each affect different muscle groups, although all are often also associated with cardiac conduction defects. Mutations in the transmembrane proteins Emerin and Nesprin 1 cause other forms of EDMD (22, 23). Nesprins have been shown to connect to cytoskeletal proteins (16, 17, 24, 25). Thus, both Lamins and NETs involved in connecting Lamins to the cytoskeleton can cause muscle disease in humans.

Postulating that other, more muscle-specific proteins might also contribute to NE-cytoskeleton interactions, we sought to determine whether additional NETs could be found in the NE proteome of skeletal muscle. The previously validated subtractive approach was applied (26) using microsomes/sarcoplasmic reticulum (SR) and mitochondria, the principal membrane contaminants expected, as subtractive fractions. Many new NE proteins were identified that had not been identified in previous NE proteomics investigations using liver, blood, and neuroblastoma cells (26–28). NE residence was confirmed for 11 novel NETs by expression of epitope-tagged versions and using antibodies on tissue cryosections.

Muscle-specific expression was determined for several NETs both according to transcriptome databases (29) and by direct testing with qRT-PCR. Moreover, some of the novel skeletal muscle NETs were up-regulated in myogenesis. Expression of NETs as tagged fusions revealed that NETs TMEM214 and KLHL31 concentrate at the nuclear surface with partial overlap with microtubules, suggesting some connectivity with the cytoskeleton. Moreover, TMEM214 and the NET WFS1 have a unique distribution at the base of the spindle poles in mitosis similar to that recently reported for NET5/TMEM201/Samp1 (30). Thus, some of these newly identified muscle NETs could potentially contribute to cytoskeletal defects in NE muscle diseases.

EXPERIMENTAL PROCEDURES

Preparation of Skeletal Muscle NEs and Microsomes/SR—Six to 10 week-old Sprague-Dawley rats were obtained from the University of Edinburgh animal facility according to university ethics procedures. For time and gradient size constraints, each individual preparation was limited to six rats. The procedure is given with full details in Ref. 31, but briefly, hind leg muscle between the femur and hip was isolated after discarding the sciatic nerve, and the muscle was crudely chopped into roughly 0.5-cm pieces, rinsed, and weighed. This was minced three times at 4 °C in 50 ml of ice-cold homogenization buffer (10 mM HEPES, pH 7.4, 60 mM KCl, 0.5 mM spermidine, 0.15 mM spermine, 2 mM EDTA, 0.5 mM EGTA, and 300 mM sucrose with freshly added protease inhibitors and 2 mM DTT) and then brought to a volume of 5 ml/g of starting muscle mass with homogenization buffer. A Potter-Elvehjem motorized Dounce homogenizer with a 0.1–0.15-mm-clearance Teflon pestle was used at 900–1000 rpm to produce a crude homogenate that was then poured through four layers of sterile cheesecloth on ice in a 4 °C room. Crude nuclei were pelleted from the filtrate at 1000 × *g* in a swinging bucket rotor for 10 min at 4 °C. The supernatant was carefully removed and processed for mitochondria and microsome/SR membranes (see below). The crude nuclear pellet was washed with Percoll gradient buffer (10 mM HEPES, pH 7.4, 60 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, and 300 mM sucrose), repelleted, and resuspended in 26.7 ml of ice-cold Percoll gradient buffer. Percoll (Sigma 77237) was added to a final concentration of 27% (v/v), and nuclei were pelleted at 27,000 × *g* for 30 min at 4 °C in a swinging bucket rotor. Some fragmented myofibrils have a mass similar to that of nuclei but a different density, so the isopycnic gradient accumulates them at the top (32). Nuclei were collected from where they banded near the bottom of the tube. Fractions were checked for nuclear enrichment by microscopy. The nuclear fraction was diluted 10-fold with ice-cold SHKM buffer (50 mM HEPES, pH 7.4, 25 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, and 0.25 M sucrose) and pelleted by centrifugation at 4000 × *g* for 20 min at 4 °C in a swinging bucket rotor. This pellet still contains other membrane and cytoplasmic contaminants, so the pellets were resuspended in 11 ml of ice-cold SHKM buffer using a hand-held Dounce homogenizer. 39 ml of SHKM buffer with 2.3 M sucrose was added to bring the sucrose concentration to 1.85 M. 25 ml was removed to each SW28 ultracentrifuge tube and underlaid with 5 ml of ice-cold SHKM buffer with 2.15 M sucrose and then a further 5 ml of SHKM buffer with 2.8 M sucrose using a 14-gauge needle. After centrifugation in an SW28 swinging bucket rotor for 60 min at 82,000 × *g* at 4 °C, the nuclei occur at the interface of the 2.8 and 2.15 M sucrose layers. Nuclei were collected, diluted 10-fold in SHKM buffer, centrifuged at 4000 × *g* for 20 min at 4 °C, and then resuspended in SHKM buffer, and nuclei were counted using a hemocytometer. At this stage, no contaminants could be observed under the microscope.

NEs were then prepared from isolated nuclei by two rounds of digestion with DNase I (Sigma D4527) and RNase A (Sigma R4875) in 0.3 M sucrose, 10 mM HEPES, pH 7.4, 2 mM $MgCl_2$, 0.5 mM $CaCl_2$, and 2 mM DTT (using a Dounce homogenizer to resuspend nuclei) for 20 min followed by layering onto the same buffer with 0.9 M sucrose and centrifugation at 4000 × *g* for 10 min at 4 °C to wash out and float digested chromatin. The first round used nuclei at 1–2 million/ml with 10 units/ml DNase and 1.4 mg/ml RNase, and the second round used nuclei at 2–4 million/ml with 50 units/ml DNase and 5 mg/ml RNase. Digestions were followed by rapid staining of nuclei with 4,6-diamidino-2-phenylindole dihydrochloride (DAPI) using a fluorescence microscope and extended if most of the DAPI signal had not been washed out. NEs were snap frozen in liquid nitrogen and maintained at −80 °C.

To prepare mitochondria and microsomes enriched in SR membranes, 0.5 mM EDTA was added to the postnuclear supernatant to inhibit metalloproteases. This homogenate was first centrifuged at 10,000 × *g* at 4 °C for 20 min to pellet mitochondria. The postmitochondrial supernatant was then subjected to centrifugation at 100,000 × *g* for 45 min at 4 °C to generate a crude microsomal/SR pellet. This was resuspended in 5 ml of ice-cold SHKM buffer, and the sucrose concentration was increased to 2 M by adding 2.7 volumes of ice-cold 2.8 M SHKM buffer. 28 ml of these membranes was added to SW28 ultracentrifuge tubes and overlaid with 7 ml of ice-cold SHKM buffer with 1.85 M sucrose and again with 3 ml of regular SHKM buffer. Microsomes were floated at 57,000 × *g* for 4 h at 4 °C in an SW28 swinging bucket rotor. Microsomes were recovered by aspiration of material at the interphase between the 1.85 M sucrose layer and the uppermost 0.25 M sucrose layer, diluted with 4 volumes of SHKM buffer, and pelleted at 152,000 × *g* in a 45 Ti rotor for 75 min.

NE preparations were accumulated and divided into aliquots. One aliquot was extracted on ice with 400 mM NaCl, 25 mM HEPES, pH 8.0, 5 mM $MgCl_2$, 10 mM DTT, and 1% β-D-octyl glucoside. A separate NE aliquot and a microsomal membrane aliquot were extracted with

0.1 N NaOH and 10 mM DTT. The material used for mass spectrometry was pelleted from the extracted material by centrifugation at 150,000 × *g* for 30 min and washed three times in filtered H_2O prior to digestion.

MudPIT—Two biological replicate preparations of NaOH-extracted NEs and SR were analyzed in duplicate, whereas two technical replicates were obtained for one preparation of salt- and detergent-extracted NEs (Table I and supplemental Table S1). Pellets were solubilized in 0.1 M Tris-HCl, pH 8.5, 8 M urea, and 5 mM tris(2-carboxylethyl)phosphine hydrochloride. Next, free cysteines were alkylated with 10 mM iodoacetamide for 30 min, and Endoproteinase Lys-C (Roche Applied Science) was added at a 1:100 (w/w) enzyme to protein ratio for 6 h at 37 °C. Urea was diluted to 2 M with 0.1 M Tris-HCl, pH 8.5, and $CaCl_2$ (0.5 mM) and modified Trypsin (1:100, w/w) were added for 12 h at 37 °C. Digestions were quenched with 5% formic acid (33). Samples were centrifuged for 30 min at 17,500 × *g*. Supernatants (Ti Table 1) were directly analyzed by MudPIT, whereas pellets were further resuspended in 0.1 M sodium carbonate (Na_2CO_3), pH 11.5, 8 M urea, and 5 mM tris(2-carboxylethyl)phosphine hydrochloride for 30 min; incubated with 10 mM iodoacetamide for 30 min; and then further digested with Proteinase K for 4 h at 37 °C and separately analyzed by MudPIT (34).

MudPIT was performed as described previously (35) using a 200–300 nl/min flow rate over 12–15 cycles (Table I and supplemental Table S1) of 120 min each of increasing salt concentrations followed by organic gradients (5–80% acetonitrile with 0.1% formic acid) (35). The last two to five chromatography steps consisted of a high salt wash with 5% acetonitrile, 0.1% formic acid, and 500 mM ammonium acetate followed by the acetonitrile gradient. Distal application of a 2.5-kV voltage electrosprayed eluting peptides directly into a linear ion trap mass spectrometer (ThermoFinnigan). Each full MS scan (from 400 to 1600 *m/z*) was followed by five MS/MS events using data-dependent acquisition (at 35% collision energy). The raw mass spectrometric data may be downloaded from proteomecommons.org Tranche using the hash identifiers provided in Table I and supplemental Table S1.

RAW files extracted into ms2 file format (36) using RAW_Xtract v.1.0 (37) were queried for peptide sequence information using SEQUEST v.27 (revision 9) (38) against 28,400 rat proteins (non-redundant NCBI sequences on July 10, 2006) plus 197 human and mouse homologs of previously identified NETs (26) and 172 sequences from usual contaminants (*e.g.* human keratins). To estimate false discovery rates, each non-redundant protein entry was randomized and added to the database, bringing the total search space to 57,538 sequences (supplemental Table S1). MS/MS spectra were searched without specifying differential modifications, but +57 Da were added statically to cysteine residues to account for carboxamidomethylation. No enzyme specificity was imposed during searches, setting a mass tolerance of 3 amu for precursor ions and of ±0.5 amu for fragment ions.

Different runs were compared using DTASelect and Contrast (39). Spectrum/peptide matches were retained if at least 7 amino acids long with ends complying with proteolytic specificity. Other criterion were ΔCn ≥ 0.08; XCorr ≥ 1.8 for singly, 2.0 for doubly, and 3.0 for triply charged spectra; and a maximum Sp rank of 10. For Proteinase K-digested samples, no specific peptide ends were imposed, but the ΔCn cutoff was increased to 0.15 (40), whereas XCorr minima were increased to 2.5 for doubly and 3.5 for triply charged spectra (SEQUEST parameters for the spectrum to peptide matches for all detected proteins are provided in supplemental Table S2). Peptide hits were merged from all analyses to establish a master list of proteins identified by at least two peptides or one peptide with two independent spectra. In addition, six previously described liver and blood NETs as well as three newly characterized skeletal

TABLE I

Hash identifiers to access raw mass spectrometry files for each MudPIT run

Individual runs are organized by cell fraction (NE or SR), extraction (NaOH or salt/detergent (SD)), and protease (Trypsin (Ti) or Proteinase K (PK)) to generate peptides for analysis.

At https://www.proteomecommons.org/tranche/downloads.jsp the raw mass spectrometry files can be downloaded using the following hash identifiers. In all cases, there are two files ranging from 100 MB to 1.5 GB that are archived and compressed tar.gz files: use tar —xzf linux command to uncompress to *.RAW files (1 per MudPIT step)	Description of Tranche (run name used in supplemental tables)
U4oMnK/mHWCdGJ8KpNw5iv2ZPYyxYeVOvNzuYOQ9A7qrBmtMNy1dCO7Kbpmjj0WN0fvJSyhOBK0rbDCYlPvyq33KHbYAAAAAAAAFkg==	RnNaOH-NE_muscles_Ti_1
eHaU1zhkwqfSFMUO4UEkzUFbJPw6vyS45tO318icJJGuYWHjUMK/11UF8+COjwqVrdOf9JrS6ma3WSBSgTBJQAqykMMAAAAAAAAFIA==	RnNaOH-NE_muscles_Ti_2
QVcyypxCuUz96X63QHoyirM3VMeo7Z7CmAUOCxR+kt1Rk/N0HF95t9n1Ik/nlla4pFMAawYyaKRQYxYCHEwqLFyVcWoAAAAAAAAGZQ==	RnNaOH-NE_muscles_PK_1
JFoLKqlkxRP2oX97exWAW80opnPuqRPvaDv72NMWJ8hSeMs8bIERnXOFC/nKhbB8INlKMcXlo09rESd1 × 5bskRSVepAAAAAAAAAFlQ==	RnNaOH-NE_muscles_Ti_3
LxSuw9vAbjDC/bWD2ZO3wf+Ypt5gp+SKVJMncZZxpCvFgbLSRGlzbL0vovsfmWLUo2DuLKyCudi4v9anLRg0k6JoqBUAAAAAAAAAFjQ==	RnNaOH-NE_muscles_Ti_4
39VSpfd6gB1OtpV5LTY6Wu7MxEBUwvjGF8ixiUE0nZR2l8QCbl21e0PZpOwnhKH3HwNv+ToF8AaWm18iUVegCFPrVKQAAAAAAAAAFeg==	RnSD-NE_muscles_Ti_1
4p8SXlJPyzBw5yVMEQVPAHONj7H8neFm4BD9bC+dL5XNsYTo4sllmyBpwpPMrlHy++HM6So6odAATGlzmP5G3TDqgwwAAAAAAAAAFeg==	RnSD-NE_muscles_Ti_2
anQNY4xCeTOsKvMrdqopEgAh9rja7xmYNC/mGwN6wumhDqUMCf5AbFGgDY4I6AGPpQziUJE9NiG5Kpmj3TkpzQUhlWgAAAAAAAAFkA==	RnNaOH-SR_muscles_Ti_1
aDpnQPjyNEgCqiN6Onra9dqzBDKWkQD9btbY8ye3gZAxVdFSGM9ddqr+Od7/a+Gx8/qdENJij2OlPbtsUfD2lBjTTS0AAAAAAAAAFkA==	RnNaOH-SR_muscles_Ti_2
nOl0FKF+b+56faEzU+sMQwEd46id4tTQlk1N/hhl1kWOezoF9HtAyXiQGcbY5bQgxZhrv/xGGslPQTga7G2gJYPlczcAAAAAAAAGZA==	RnNaOH-SR_muscles_PK_1
JZMpidblCAlMpTaUDeiiZdSpbAmm4L8SH7w2xB83m2LtAjPtz0cTrbe1XcCBm1smSjUH5T6PWj4fZZFV5kGcyC3ZELoAAAAAAAAAFkQ==	RnNaOH-SR_muscles_Ti_3
xTESWmC9cv1UKXyctOVK35MKMTH7Iw5XZ0rv9n5BefVgOM32ubzLabhUHDQC1TBhxtlCZsSyqw9N7lWXR3/UqmHXq0MAAAAAAAAAFkQ==	RnNaOH-SR_muscles_Ti_4
CT5hCf/Y3/wzsYaPfsoNZ75InelvcN+arCYacOxlOb0+OgLXno5a3nRhlGbzK/ckuCtAwQwXpOk6lcPZFk5WMAJRf4UAAAAAAAAGYQ==	RnNaOH-SR_muscles_PK_2

muscle NETs were manually selected although detected by single peptides (annotated spectra for these proteins are provided in supplemental Fig. S1). Based on the merged detected peptides, proteins could fall into three categories following the parsimony principle. (i) Proteins detected by the exact same set of peptides were grouped together because they could not be distinguished based on the available peptide data (see column named "Proteins in Group" in supplemental Table S3), and only one arbitrarily selected representative protein entry is reported for such groups of proteins (see column labeled "Locus" in supplemental Table S3). (ii) Proteins with at least one peptide uniquely mapping to them were considered unique entries. (iii) Subset proteins for which no unique peptides were detected were removed from the final list of identified proteins.

Identifications mapping to shuffled peptides were used to estimate false discovery rates. Spectral FDR was calculated as shown in Equation 1.

$$\text{Spectral FDR} = \frac{2 \times \text{SHUFFLED\_SpectralCounts}}{\text{Total\_SpectralCounts}} \times 100 \quad \text{(Eq. 1)}$$

Protein level FDR was calculated as shown in Equation 2.

$$\text{Protein FDR} = \frac{\text{SHUFFLED\_Proteins}}{\text{Total\_Proteins}} \times 100 \quad \text{(Eq. 2)}$$

Under these criteria, the final FDRs at the protein and peptide levels were on average 2.2 ± 0.9 and 0.3 ± 0.09%, respectively (supplemental Table S3).

Normalized spectral counts (NSAF) were calculated for each nonredundant protein to estimate relative protein levels as described (41–43).

$$(\text{NSAF})_i = \frac{(\text{SpectralCount/Length})_i}{\sum_{k=1}^{N} (\text{SpectralCount/Length})_k} \quad \text{(Eq. 3)}$$

To refine spectral counts to deal with peptides shared between multiple proteins (44), dNSAFs were calculated based on distributed spectral counts for each run in which shared spectral counts were distributed based on spectral counts unique to each isoform (supplemental Table S3). For each protein, averaged dNSAF and standard deviations (supplemental Table S3) were calculated across several replicate analyses of NE and SR samples (supplemental Table S1).

Bioinformatics Analysis—Proteins were mapped to an Ensembl gene to remove redundancy from differences in protein annotations and cross-referenced to the human genes with Ensembl release 48 (45). Orthologous group identities were sorted according to frequency of detection in the NE and microsome runs, relative levels (determined by averaged dNSAF values), membrane helix prediction (determined using TMHMM 2.0; http://www.cbs.dtu.dk/services/TMHMM-2.0/; Ref. 46), and dNSAF ratios between values measured in NEs *versus* microsomes/SR (supplemental Table S3). Biologically interesting gene ontology terms were retrieved from the mySQL database (http://amigo.geneontology.org; Ref. 47) and added to tables. Comparison of expression levels in different tissues was done by downloading relative values from BioGPS (http://biogps.gnf.org/#goto=welcome; Refs. 29 and 48) and calculating the -fold expression over the median value from a wide variety of mouse tissues tested in this transcriptome database.

Plasmid Construction—IMAGE clones for human NETs *TMEM38A*, *TMEM194*, *POPDC2*, *KLHL31*, *WFS1*, *TMEM70*, *MBOAT5*, *CKAP4*, *TMEM214*, *RHBDD1*, and *TMEM201* were obtained from RZPD and Geneservice. Coding sequences were amplified by PCR with added 5′ and 3′ restriction sites, sequenced from both ends in intermediate cloning vectors, and then inserted into mammalian expression vector pEGFP-N2 (Clontech) or pmRFP-N2 (derived from pEGFP-N2 by replacing the GFP coding sequence with that of monomeric red fluorescent protein). *Sec61β-GFP* was obtained from Tom Rapoport (Harvard, Boston, MA).

Antibodies and Western Blotting—Antibodies used were glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Enogene E1C604), Calreticulin (Cell Signaling Technology 2891S), Calnexin (Stressgen SPA-860), Lamin A (3262; Ref. 11), Nup153 (Covance MMS-102P), Nup358 (raised against recombinant human protein amino acids 2595–2881; a kind gift from F. Melchior), SUN2 (Millipore 06-1038), Tubulin (Sigma T6074), TMEM38A (Millipore 06-1005), POPDC2 (Millipore 06-1007), TMEM201 (Millipore 06-1013), and LOC203547 (Millipore 06-1008). All fluorophore-conjugated secondary antibodies were minimally cross reactive from donkey (Jackson ImmunoResearch Laboratories) or goat (Molecular Probes).

Mitochondria prepared as described above were directly lysed in sample buffer. Muscle NE and microsomes/SR were incubated on ice in 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM MgCl2, and 0.2% Nonidet P-40 with protease inhibitors; then heated at 65 °C for 2 min; and sonicated in a 4 °C sonic bath. Protein concentrations were determined by Bradford assay before adding sample buffer. Mitochondrial loading was based on Coomassie staining compared with NE and SR lanes.

All blots shown were run according to standard procedures, and protein bands were visualized and quantified with IR800-conjugated secondary antibodies using a LI-COR Odyssey imaging system and software (Odyssey 3.0.16) using median background subtraction. For Fig. 2, example blots are shown, whereas for Fig. 5, a graph from averages of three independent blots for each NET and controls is presented.

Cell Culture, Differentiation, and Transient Transfection—HT1080, HeLa, U2OS, RD, and C2C12 cells were maintained in high glucose DMEM (4.5 g/liter glucose; Invitrogen) supplemented with 10% fetal bovine serum (FBS), 100 μg/μl penicillin, and 100 μg/μl streptomycin sulfate (Invitrogen). RD and C2C12 cells were differentiated after 2 days of confluence by replacing the medium with 0.2% FBS medium containing 10 μg/ml 12-*O*-tetradecanoylphorbol-13-acetate and 0.1% FBS medium containing 5 μg/ml insulin and 5 μg/ml transferrin, respectively. Successful differentiation into myotubes was confirmed by light microscopy, counting that more than half of cell bodies exhibited the lengthening characteristic of myotubes and were multinucleate. Cells were extracted for RNA typically 2–3 days after myotubes became prevalent. Adherent cells destined for microscopy were plated on coverslips, and when indicated, DNA was transfected the next day using FuGENE 6 (Roche Applied Science) according to the manufacturer's instructions.

Immunofluorescence Microscopy—After 30 h post-transfection, cells were directly fixed for 7 min in 3.7% formaldehyde and processed for microscopy after permeabilization for 6 min in 0.2% Triton X-100. For the prefixation extractions in Fig. 5, cells were first washed with PBS; then extracted for 1 min with 1% Triton X-100, 25 mM Tris, pH 8.0, 150 mM KOAc, 15 mM NaCl, and 5 mM $MgCl_2$; washed again with PBS; and then fixed with formaldehyde. In both cases, cells were then blocked with 10% FBS and 200 mM glycine in PBS and incubated for 40 min at RT with relevant antibodies. DNA was visualized with Hoechst 33342 or DAPI, and coverslips were mounted in Fluoromount-G (EM Sciences). Most images were obtained using a Nikon TE-2000 microscope equipped with a 1.45 numerical aperture 100× objective, PIFOC Z-axis focus drive (Physik Instruments), Sedat quad filter set, and CoolSnapHQ High Speed Monochrome charge-

coupled device camera (Photometrics). Those in Fig. 4*A* were deconvolved from 0.2-μm sections using AutoquantX. Structured illumination images in Fig. 4*B* were taken on the OMX system at the University of Dundee microscopy facility, and Alexa Fluor secondary antibodies (Molecular Probes) were used (other details are available upon request).

For cryosections, fresh rat heart or leg muscle was cut into 2–3-mm cubes, embedded in optimal cutting temperature compound (Tissue-Tek), snap frozen in liquid nitrogen, and maintained at −80 °C. Sections were cut on a Leica CM 1900 cryostat at 6–8-μm thickness and fixed in −20 °C methanol. After rehydration, sections were incubated with NET antibodies overnight at 4 °C followed by secondary antibodies as above. Cryosection images were recorded using an SP5 confocal system (Leica) with 63× oil 1.4 numerical aperture objective using argon and UV lasers. Micrographs were saved from source programs as .tif files and prepared for figures using Photoshop 8.0.

Electron Microscopy—Isolated skeletal muscle NEs were fixed in 3% glutaraldehyde in 0.1 M sodium cacodylate, pH 7.3, for 2 h; washed in 0.1 M sodium cacodylate; postfixed in 1% osmium tetroxide for 45 min; washed again; and dehydrated in 50, 70, 90, and 100% normal grade acetones for 10 min each and then for a further two 10-min changes in analar acetone. Samples were then embedded in Araldite resin. 60-nm-thick sections were cut on a Reichert OMU4 ultramicrotome (Leica), stained in uranyl acetate and lead citrate, and then viewed in a Phillips CM120 transmission electron microscope. Images were taken at 100 kV at 11,000× magnification using a Gatan Orius charge-coupled device camera.

Semiquantitative and Quantitative RT-PCR—For differentiation experiments, cells were lysed with TRI Reagent (Sigma), and total RNA was extracted according to the manufacturer's instructions from C2C12 or RD myoblasts or 4-day differentiated myotubes. RT-PCRs were carried out with 100 ng of total RNA using the Titan one-tube RT-PCR system (Roche Applied Science) in accordance with the manufacturer's instructions except that the dNTP concentration was increased to 500 μM and $MgCl_2$ was increased to 3 mM. Typical reaction conditions were 30 min of reverse transcription at 50 °C; 2 min of denaturation at 94 °C; and then 24 cycles of 94 °C for 30 s, 60 °C for 30 s, and 68 °C for 45 s. Peptidylprolyl isomerase A (*PPIA*) was used as a loading control, and creatine muscle kinase (*CKM*), myogenin (*MYOG*), and myosin heavy chain 1 (*MYH1*) were used as positive controls. Human (Hs) and mouse (Mm) primer sets used are given in supplemental Table S4.

For tissue expression experiments, human tissue RNAs were purchased from Stratagene except for blood RNA that was prepared from peripheral blood mononuclear cells isolated from buffy coats obtained from the Scottish National Blood Transfusion Service. 8 μg of RNA from each tissue was first converted into cDNA using a First Strand cDNA Synthesis kit (Thermoscript, Invitrogen) using a mixture of reverse primers for the cDNA to be amplified at a final concentration of 3 μM. The quantitative real time RT-PCR control was *CKM* in heart muscle. Standard curves were performed using a Roche Lightcycler for each primer pair by serial dilutions of cDNA products of heart muscle (1:5, 1:25, 1:125, and 1:625) to estimate PCR efficiency. *GAPDH* was used as the internal control to normalize the cDNA samples across the different tissues, and the relative abundance of each mRNA within the tissues was calculated using heart as the standard reference dilution. To avoid amplification of contaminating genomic DNA, primers from each set were specific to different exons. Primers used are given in supplemental Table S4.

RESULTS

Generation of Muscle NE Fractions—Rat leg muscle nuclei were prepared using a method that first depleted collagen from crude lysates (confirmed by Western blot; data not shown), thus presumably removing most connective tissue and enriching for myofiber nuclei. Myofibrillar material and contaminating membranes were then removed on Percoll and sucrose gradients (31) (Fig. 1, *A* and *B*). Isolated nuclei were subsequently digested with DNase and RNase and salt-washed to remove most of the nuclear contents (Fig. 1*C*). Even after such isolation, NEs maintain strong connections to various protein partners including a fraction of peripheral chromatin as shown by electron microscopy (Fig. 1*D*). Therefore, to further purify NEs prior to mass spectrometry analysis, NE fractions were extracted with 1% β-octyl glucoside and 400 mM NaCl or with 0.1 N NaOH (Fig. 1*A*), which enrich for proteins associated with the insoluble intermediate filament Lamin polymer and integral membrane proteins, respectively. Both extracted fractions were separately analyzed (supplemental Table S1) because previous NE proteome analyses demonstrated that some well characterized NETs are distributed to one or the other fraction (26, 27).

Microsomes enriched in SR were separately prepared as a comparative/subtractive fraction. Some SR would be expected to contaminate NEs because the ONM is continuous with the SR, and also during nuclear isolation, some SR vesicles may stick to the surface of nuclei. In contrast, as nuclei are stable and pellet fully at low speeds, there should be no NE contamination in the microsome/SR fraction. Indeed, neither Lamins A/C nor the NET LAP2β, used as NE markers, were observed in the SR fraction by Western blot analysis (Fig. 2, *A* and *B*). Therefore, as the only membrane in the nucleus is the NE, all membrane proteins enriched in the NE fractions compared with the SR fractions should be true NETs. As expected, there were similar levels of the SR protein Calnexin in NE fractions because half of the NE, the ONM, is continuous with the SR (Fig. 2*B*). Mitochondria are also expected to stick to NEs during isolation, although most should in theory be removed by Dounce homogenization and sucrose gradients that float many contaminating membranes (Fig. 1*A*). Indeed, just as the NE Lamin A/C proteins were undetectable in a separately prepared mitochondrial fraction, the mitochondrial protein Porin was not detectable in the crude NE fraction (Fig. 2*C*).

As the vast majority of material was lost in isolating clean NEs from the skeletal muscle, NEs isolated from at least 15 separate preparations were combined prior to alkali or salt/detergent extraction. Thus, variation between preparations should have been averaged out.

MudPIT Analysis—After extraction, pellets of isolated NE fractions were digested with Endoproteinase Lys-C and Trypsin. Because some membrane proteins that associate with the intermediate filament Lamin polymer resist extraction with 2% Triton X-100 and 2 M NaCl, and the polymer itself is not completely solubilized even in 8 M urea, a second digestion was performed on material that remained insoluble after the Endoproteinase Lys-C and Trypsin digestion. Thus, material that could be pelleted at 17,500 × *g* for 30 min was subse-

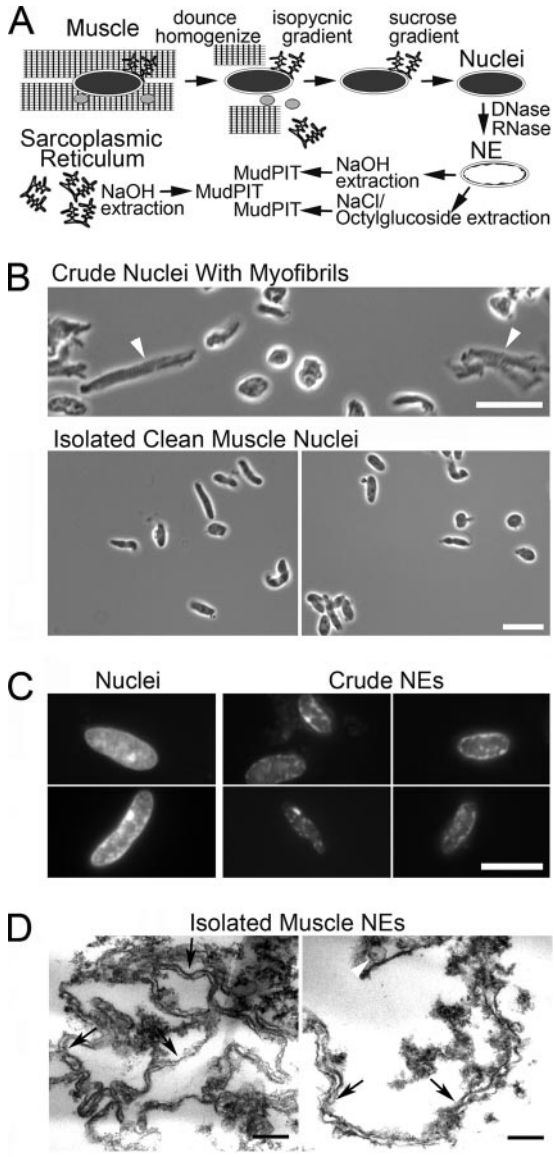


FIG. 1. **Cellular fractionation of rat muscle.** *A*, method schematic. Nuclei were first prepared from diced and homogenized rat leg muscle and cleaned of contaminating cellular structures by centrifugation first on isopycnic gradients and then on sucrose to float contaminating membranes while pelleting the denser nuclei. Crude NEs were prepared by digesting/extracting nuclear contents from isolated nuclei. Before MudPIT analysis, these were further extracted with 1% β-octyl glucoside and 400 mM NaCl or with 0.1 N NaOH to enrich for proteins associated with the insoluble Lamin polymer and integral membrane proteins, respectively. Microsomes/sarcoplasmic reticulum preparations were generated separately and analyzed for comparison/subtraction as they contain most expected membrane contaminants of the NE. *B*, *top panel*, before running both gradients, the crude nuclear fractions contained many chunks of myofibrillar material (*e.g.* Z-bands; highlighted with *white arrowheads*) released during homogenization. *Bottom panels*, after both gradients, isolated nuclei were clean of these contaminants. Phase-contrast light microscope images are shown. *Scale bar*, 10 μm. *C*, enrichment for NEs by chromatin digestion. DAPI staining for DNA visualizes significant nuclear chromatin content in an isolated muscle nucleus (*left panels*) and the loss of most of this material after two rounds of digestion with DNase and RNase each followed by salt washes (*two right panels*). Fluorescence microscope images are shown. *Scale bar*, 5 μm. *D*, ultrastructure of isolated NEs. Electron micrographs of crude muscle NEs show that most material in the population has the characteristic double membrane structure of the NE with little contamination from single membrane vesicles (a likely SR single membrane vesicle contaminant is highlighted by the *white arrowhead*). *Arrows* point to positions of NPCs. These NEs were further salt/detergent- or alkali-extracted prior to analysis by MudPIT. After such treatment, no structure remains that can be readily discerned by EM with the characteristics of NEs. *Scale bar*, 0.2 μm.

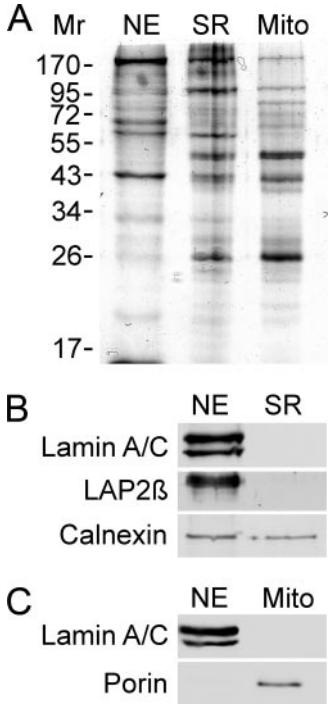


FIG. 2. **Fraction purity.** *A*, Coomassie-stained gel of NE and microsome/SR muscle fractions analyzed and also of a separately isolated muscle mitochondrial fraction (*Mito*). *B*, the SR marker Calnexin was present in both the SR and NE fractions because the ONM is continuous with the SR and shares many proteins. In contrast, the SR fraction was completely free of NE-specific markers, the NET LAP2β and Lamins A/C. Loading was the same as for *A* with similar amounts of total protein loaded. *C*, the mitochondrial marker Porin was absent from NEs, whereas Lamins A/C were absent from mitochondria. Loading was the same as in *A* and *B*.

quently digested with Proteinase K at high pH (34). The results of each digestion and replicates can be compared in supplemental Table S3. Five separate MudPIT runs were performed for NaOH-extracted NEs (one with Proteinase K digestion), and two were performed for the salt/detergent-extracted NEs. Six MudPIT runs were performed on the SR material extracted with NaOH (two with Proteinase K digestion). Thus, using different extraction conditions and different digestions and performing multiple runs were expected to provide comprehensive coverage to identify all proteins in the

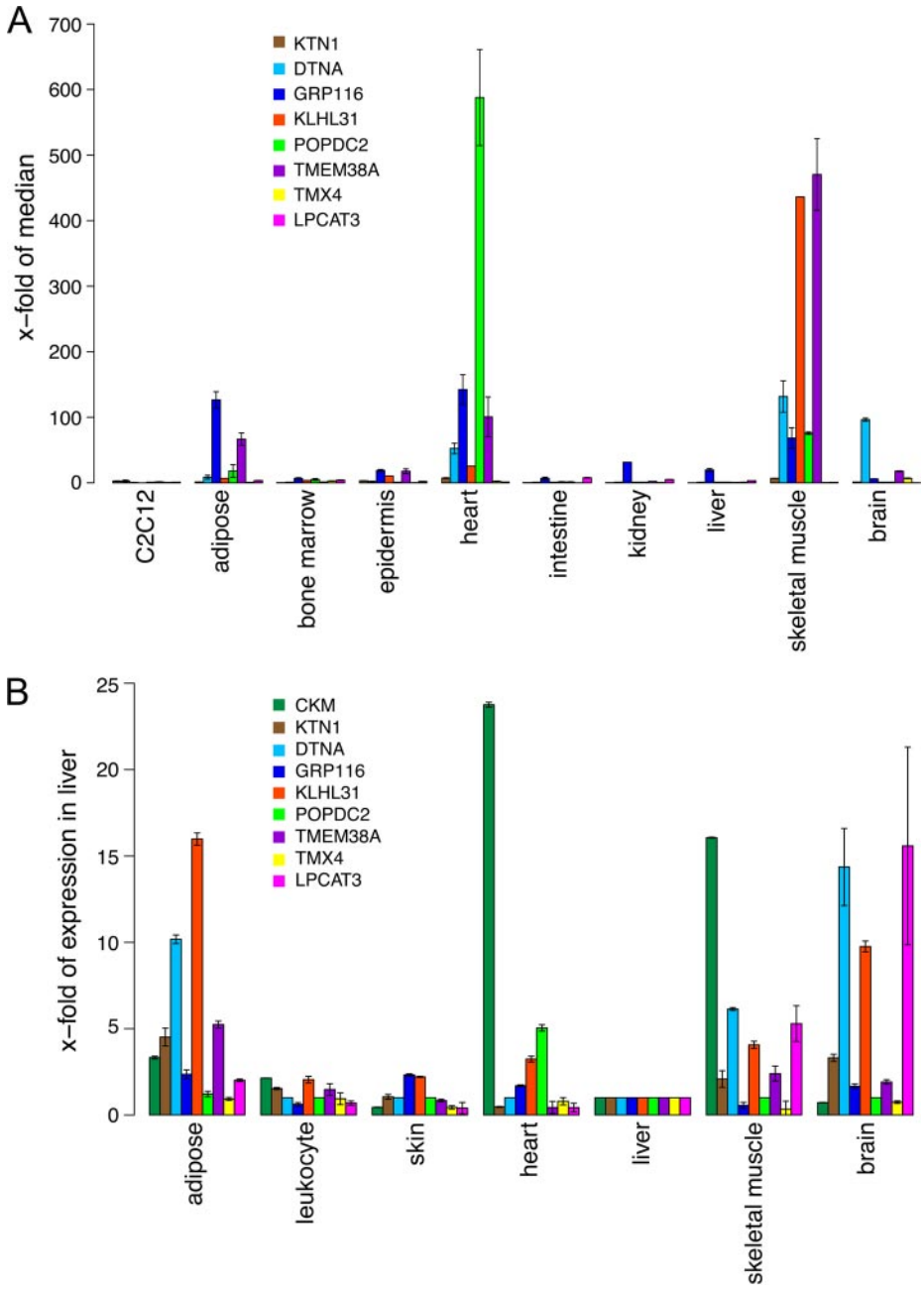


FIG. 3. **NET tissue specificity of expression.** *A*, expression data were downloaded from the BioGPS transcriptome database that compared over 60 mouse tissues on microarrays. The median expression value was obtained over the complete set of tissues, and data for a subset of tissues are graphed relative to this median value. Of the novel muscle putative NETs represented on the arrays, only one was not expressed notably above the median value in either heart or skeletal muscle (*LPCAT3*) with six expressed more than 100-fold over the median. In many cases, the median value occurred at background levels of expression (see Table II), so NETs that tended to be expressed around the median in most tissues besides muscle are more uniquely expressed in muscle. Error bars are standard deviations from fluorescence values from multiple microarrays. *B*, quantitative real time PCR of the same NETs and the control *CKM* over a smaller set of tissues. The values were generated in relation to *GAPDH* expression in the same reactions, and each NET is graphed as the -fold expression over its levels in liver. Tissue-preferential expression was also observed for several NETs directly tested by qRT-PCR; however, expression was sometimes observed in tissues that were at background levels according to the transcriptome data. Of particular note, *LPCAT3*, the only NET that was not up-regulated in skeletal muscle or heart according to transcriptome data, was up-regulated in skeletal muscle more than 5-fold by qRT-PCR. The expression of these new muscle NETs in blood and liver was at low/background levels by both methods, consistent with their having been uniquely identified in the skeletal muscle NE. Error bars are standard deviations from 3 Q-RT PCR replicates.

skeletal muscle NE. A total of seven MudPIT runs were acquired on the NE fractions because we postulated that some tissue-specific NETs might not be highly abundant. To test this, the frequency of NET detection of the seven NE runs was plotted against an estimate of abundance based on normalized spectral counts (dNSAF) (42, 44). A clear relationship was observed with the more abundant NETs tending to be detected not only in multiple analyses of the same tissue but also in independent analyses of different tissues (supplemental Fig. S2).

After removing redundancy by converting protein identities to orthologous gene groups, 1267 proteins were identified in total for the skeletal muscle NE data sets (supplemental Table S3). Separately, 1408 proteins were identified in the SR data sets (supplemental Table S3). Although fractions appeared to be clean of mitochondria by Western blot analysis, to more fully test mitochondrial contamination, data sets were compared with lists of mitochondrial proteins from a previous mass spectrometry study of that organelle (49) (supplemental Table S3). After subtracting 74 proteins that were previously identified in mitochondria, 1193 proteins remained in the skeletal muscle NE fraction. Using microsomes/SR strictly as an absolute subtractive fraction would have resulted in the loss of several well characterized

NETs including LBR, Emerin, SUN2, and Nesprin 1 (50–54). To some degree this is expected; for example, Emerin is not unique to the INM, and separate functions have been reported in the ONM, ER, and interstitial discs (55–57). Furthermore, a recent proteomics study suggests that at least a third of all proteins have multiple cellular localizations (58). Nonetheless, the number of spectra recovered from NEs was much greater than for ER for all the previously characterized NETs. Adding spectra from all runs, LBR had only two spectra recovered from the SR fraction compared with 94 from the NE fraction, Emerin had one in the SR compared with 33 in the NE fraction, SUN2 had 24 in the SR compared with 1490 in the NE, and Nesprin 1 had two in the SR compared with 281 in the NE (supplemental Table S3). Thus, instead of absolute subtraction, proteins were ordered according to their dNSAF estimate of abundance (42, 44), and NE data sets were restricted to those proteins either not detected in the microsomes or at least 5× more abundant than in the microsome fractions based on their NSAF values, which, again after subtracting all mitochondrial proteins, left 710 proteins in the skeletal muscle NE fraction (supplemental Table S3).

This protein set includes not only membrane proteins but also many soluble proteins. Some of these soluble proteins may have binding sites throughout the nucleus but nonetheless have relevant interactions at the NE. In contrast, the subset of proteins harboring transmembrane spans (NETs) is the most likely to be enriched at the NE and constitutes the surest core NE components. This set of proteins include all 13 of the original well characterized NETs, which were ubiquitously detected in our large scale analyses of nuclear envelopes isolated from liver (26), blood (28), and now muscle; these ubiquitous NETs were also recovered at high levels (dNSAF values) and reproducibly in multiple analyses of muscle NE (supplemental Fig. S2A and Table S5). An additional 28 liver NETs had been previously detected in liver NE data sets (26), 14 of which have confirmed NE residence (supplemental Fig. S2B and Table S5). Another 109 proteins were predicted to contain at least one transmembrane domain by TMHMM v2.0 or membrane anchor by SignalP. Some of the transmembrane proteins identified in the data sets were known proteins that had not been detected previously at the NE, many being ion transporters that likely function in muscle SR. Others had no annotations at the time of data set generation, and we selected 53 of these as putative novel NETs, 24 of which were also detected in a concurrent proteomics analysis of NEs from peripheral blood leukocytes (Ref. 28; supplemental Fig. S2C and Table S5). Within the "muscle-only" muscle NETs, one protein of high abundance was detected in six of seven runs (TMEM38A), another group of 11 proteins of lower abundance was detected in at least two runs, whereas 21 proteins were detected only once and at the lowest levels (supplemental Fig. S2D and Table S5). The nuclear envelope targeting of over half of the proteins (48 of 94) plotted in supplemental Fig. S2 has been confirmed. In

TABLE II

NETs with high expression levels in muscle tissues expressed as normalized microarray intensity units and relative to median (in parentheses) for all tissues sampled (see "Experimental Procedures")

Bold values, >10× median.

NET	Tissue of highest expression (Hs)	Hs heart	Hs skeletal muscle	Hs smooth muscle	Tissue of highest expression (Mm)	Mm heart	Mm skeletal muscle	Mm liver	C2C12
WFS1	Lung (8.0)	93.45 **(11.8)**	4.9 (0.6)	14.2 (1.8)	Nucleus accumbens (210.5)	1815.1 (8.6)	923.4 (4.4)	211.2 (1.0)	513.5 (2.4)
POPDC2	Heart (5.3)	194.6 **(36.7)**	16.3 (3.1)	5.85 (1.1)	Heart (4.6)	2703.8 **(583.2)**	350.3 **(75.6)**	4.6 (1.0)	4.6 (1.0)
KLHL31	Below threshold	—[a]	—	—	Skeletal muscle (70.4)	1792.4 **(25.5)**	30702.1 **(436.1)**	71.6 (1.0)	[b]
GPR116	Fetal lung (6.7)	8.7 (1.3)	9.6 (1.4)	6.8 (1.0)	Lung (14.3)	2006.7 **(140.5)**	950.4 **(66.5)**	278.5 **(19.5)**	8.0 (0.6)
ATG9A	Testis (29.3)	101.7 (3.5)	31.1 (1.1)	33.65 (1.1)	Testis (107.4)	652.8 (6.1)	1474.7 **(13.7)**	573.2 (5.3)	74.6 (0.7)
CKAP4	Smooth muscle (38.0)	63.6 (1.7)	32.1 (0.8)	1333.8 **(35.1)**	Osteoblast day 14 (345.7)	470.7 (1.4)	158.5 (0.5)	7.5 (0.0)	2656.9 (7.7)
DTNA	Below threshold	—	—	—	Lens (4.8)	221.1 **(45.7)**	541.3 **(111.9)**	4.8 (1.0)	39.7 (8.2)
TMEM214	Smooth muscle (150.1)	161.5 (1.1)	193.05 (1.3)	398.05 (2.7)	Lacrimal gland (297.3)	565.5 (1.9)	343.9 (1.2)	605.2 (0.2)	407.6 (1.4)
TMEM38A	[c]	—	—	—	Skeletal muscle (20.4)	1960.8 **(95.9)**	9550.4 **(467.2)**	15.2 (0.7)	29.4 (1.4)

[a] —, not in the BioGPS transcriptome database.
[b] Cell line not studied in the experiment.
[c] Gene not represented in the microarray.

particular, 15 of the low abundance NETs have been directly tested by us and by others (see below and supplemental Table S5) and confirmed at the NE.

Muscle-specific NETs—The identification of new NETs in skeletal muscle that were not in previous liver, blood, and neuroblastoma cell data sets (26–28) could reflect either their preferential expression in muscle or differences in the sensitivity of the mass spectrometers between the time of the different studies. To distinguish between these possibilities, a subset of novel putative rat skeletal muscle NETs was searched for expression levels in the BioGPS transcriptome database that compared relative expression for RNAs from over 60 mouse tissues on microarrays (mouse was used because very little rat transcriptome data were available) (29, 48). Although many NETs were widely expressed according to the transcriptome data, others were preferentially expressed in heart and/or skeletal muscle over other tissues (Fig. 3*A*). A few of these were just 3–10-fold higher in the heart and/or skeletal muscle than the median value for all other tissues, but several reached levels greater than 100-fold over the median value (Fig. 3*A* and Table II). Most of the eight putative NETs shown were expressed at roughly background levels in liver and bone marrow, explaining why they were not found in previous liver (26) or leukocyte (28) NE data sets and further reflecting tissue specificity in the NE proteome.

To confirm these results, qRT-PCR was performed on the same eight muscle NETs over seven human tissues including cardiac and skeletal muscle. Using *GAPDH* as a standard and *CKM* as a positive control, all but *TMX4* were more highly expressed than in liver in at least one of the two muscle tissues (Fig. 3*B*). The results did not precisely match the BioGPS data for which there are many possible explanations, but most of the NETs were preferentially expressed in muscle over several other tissues in both systems.

The tissue specificity of NET expression based on the transcriptome data was not always shared between species (Table II). The BioGPS transcriptome database also compared relative RNA expression levels for over 70 human tissues (29, 48). Some NETs like *WFS1* and *POPDC2* were highly expressed in heart muscle in both human and mouse systems. However, *GPR116* was very preferentially expressed in mouse heart and skeletal muscle yet was expressed at the median level in all human muscle groups (Table II). This may indicate species differences as well as tissue differences in the NE proteome.

Confirmation of NE Residence—To directly test NE residency for novel NETs as a measure of the validity of novel protein identifications, 10 putative NETs from the skeletal muscle data set were cloned as tagged fusions (GFP and/or monomeric red fluorescent protein) and expressed in tissue culture cells. Some were not expressed in the HT1080 cells first used, but when a variety of different cell lines were utilized, all 10 putative NETs were expressed and yielded a discernible rim at the nuclear periphery (Fig. 4*A*). However,

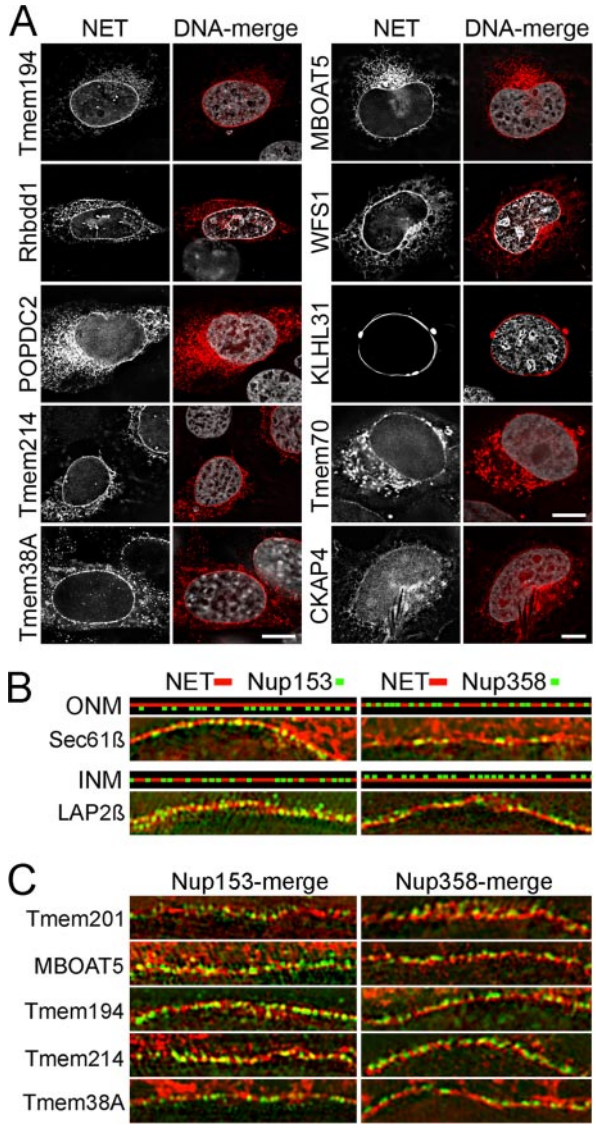


FIG. 4. **Confirmation of NE residence for NETs by targeting tagged fusion proteins.** *A*, NETs were tested as tagged fusions for NE targeting, which is determined by their enrichment in a rim at the limits of DNA staining. NET fusions are visualized in *white* in the *left panels* and in *red* in the *right panels*, whereas DNA is visualized in *white* in the *right panels*. NETs were sometimes expressed or targeted better in certain cell types: all panels shown are U2OS osteosarcoma cells except for TMEM38A in mouse C2C12 cells and TMEM70 and CKAP4 in HeLa cells. All micrographs are on the same scale except for CKAP4 with all *scale bars* at 10 μm. *B* and *C*, inner *versus* outer nuclear membrane targeting. Structured illumination microscopy can distinguish proteins in the INM from those in the ONM when co-stained with nuclear basket protein Nup153 and cytoplasmic filament protein Nup358. *B*, an ONM NET or ER/ONM protein (*red*) should be in the same plane with Nup358 (*green*) but should be external to Nup153 (*green*), and this is observed for the control ER protein Sec61β (*upper schematic* and *images*). Correspondingly, an INM NET should be in the same plane as Nup153 and internal to Nup358 as is observed for the control NET LAP2β (*lower schematic* and *images*). *C*, all new NETs tested appeared in the same plane of the INM with Nup153 and formed an internal ring to Nup358, indicating INM residence. *Bars*, 5 μm.

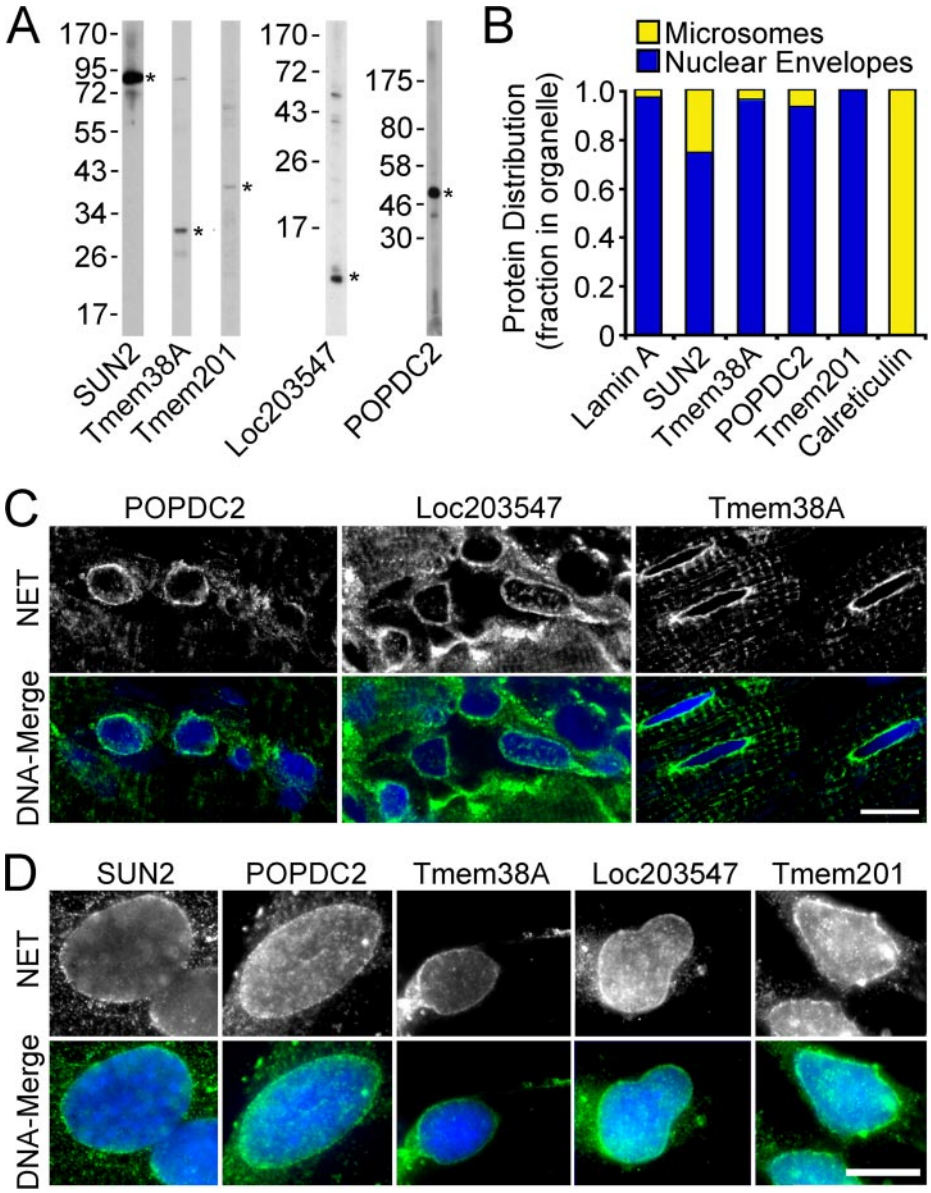


FIG. 5. **Antibody staining confirms novel NET identifications.** *A*, Western blots testing antibodies generated to novel muscle NETs TMEM38A, LOC203547, and POPDC2, control NET SUN2, and a NET previously identified in liver that appears to be much more abundant in muscle, TMEM201/NET5. The correct size for each NET is indicated by an *asterisk*. *B*, comparison of the relative amounts of proteins in ER (microsome) and NE fractions. Lamin A was used as a control for a NE-specific protein, whereas Calreticulin was used as a control for an ER protein. Western blots of the two fractions similarly loaded for total protein were quantified on a LI-COR Odyssey imaging system, and the percentage of the total signal between both microsome and NE lanes was plotted with NE signal in *blue* and microsome signal in *yellow* (three repeats were averaged). The new muscle NETs were more enriched in the NE fraction than was the control NET SUN2. *C*, cryosections of rat muscle stained with NET antibodies. NE staining was clearly observed for all NETs tested as determined by a rim around the DAPI-stained DNA. Leg muscle is shown for TMEM38A, and heart is shown for the other two NETs. *Bars*, 10 μm. *D*, antibody staining on pre-extracted cells. All cells shown are C2C12 mouse myoblasts except in the case of TMEM201 where HeLa cells are shown. Cells were pre-extracted with detergent (1% Triton X-100) to remove membranes and most soluble cytoplasmic material and then fixed and incubated with NET antibodies. A nuclear rim staining was observed for the control NET SUN2 and all novel NETs tested. The resistance to detergent pre-extraction is consistent with the INM localization indicated by OMX results in Fig. 4.

they were quite variable in the percentage of the protein accumulating at the NE. For example, TMEM194 tended to have a very crisp nuclear rim, whereas POPDC2 had higher levels in the cytoplasm, although a rim could still be discerned. Those that had both NE and ER accumulation can be explained by three possibilities. First, binding sites for the NET at the NE might have become saturated due to overexpression, a phenomenon common among NETs. A corollary of this possibility is that partner proteins involved in tethering the NET in the NE are of low abundance in the cell lines used. This

is the possibility we consider most likely because of the high degree of tissue specificity in expression of these NETs. Second, the NET might normally have multiple cellular localizations as a recent proteomics study has suggested for nearly 40% of the mammalian proteome (58). Third, it is possible that the GFP/red fluorescent protein fusion altered the NET protein structure so that it is targeting aberrantly. In this latter case, it is actually more likely that a reduction in affinity at an NE-tethering site results in artifactual accumulation in the ER than that the NE targeting is artifactual, although both are possible. KLHL31 staining was highly filamentous and only sometimes completely encircled the NE. High resolution deconvolved images confirm its proximal association with the NE, but it appeared in some cases of overexpression to concentrate in membrane invaginations protruding into chromatin and in others to form filaments that extended into the cytoplasm beyond the NE.

The NE equally comprises INM and ONM, so NETs could occur in one or the other subcompartment. Structured illumination microscopy (OMX) can distinguish INM from ONM localization by co-staining for proteins from the NPC nuclear basket (Nup153) or cytoplasmic filaments (Nup358) (59). The NPC nuclear basket and cytoplasmic filaments both distend more than 50 nm away from the NE, which itself is roughly 50 nm from the ONM to the INM. Therefore, the combined separation between NETs and these distal NPC components is within the ~100-nm-resolution limit of the OMX system. If a NET is in the ONM, one would observe localization to the same plane as Nup358 and some planar separation with Nup153 (Fig. 4*B*, *upper schematic*), and this is what was observed for the ONM control protein Sec61β (Fig. 4*B*, *upper image*). The reverse should apply for an INM protein with localization to the same plane as Nup153 and some planar separation with Nup358 (Fig. 4*B*, *lower schematic*), and this is what we observed for the INM control protein LAP2β (Fig. 4*B*, *image*). Yellow co-localization is generally not observed in this system because the OMX also resolves the ~120-nm gaps in the nuclear membrane where the NPCs are inserted from the NETs in the membrane itself. Similarly to the control NET LAP2β, all the new skeletal muscle NETs tested here were found to accumulate in the INM as both the NET (*red*) and Nup153 (*green*) localized to the inner ring, whereas an inner NET ring could be distinguished from an outer Nup358 ring (Fig. 4*C*).

To test localization of endogenous protein, polyclonal antibodies were generated to four novel NETs (TMEM38A, POPDC2, LOC203547, and TMEM201). Antibodies were also obtained for the control NET SUN2. Each antibody recognized a protein of the correct size in immunoblots (Fig. 5*A*). Antibodies against each NET were incubated with similar amounts of NE and microsome preparations. Signals from the Western blots were detected with fluorescent antibodies and quantified. The total signal intensity from both NE and microsome bands were measured, and the percentage of total signal for each was plotted. Surprisingly, the control SUN2 had the highest percentage appearing in the microsome fraction (Fig. 5*B*). Thus, when not overexpressed, these novel NETs are more exclusive to the NE than one of the best characterized NETs.

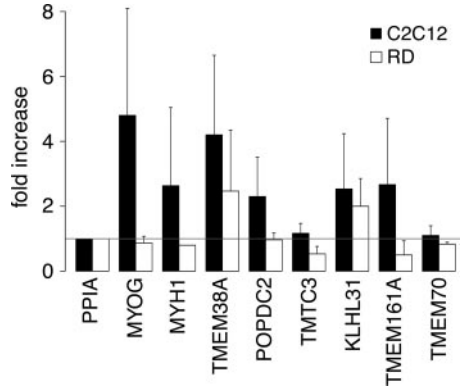


FIG. 6. **NE composition changes during muscle differentiation.** Expression levels of several muscle NETs were tested in both C2C12 (mouse) and RD (human) myoblast to myotube differentiation systems. RNA extracted from both untreated myoblast and chemically differentiated myotube populations were subjected to RT-PCR for each NET, and expression levels were quantified. The relative change between myoblast and myotube populations was determined after first normalizing values to the control peptidylprolyl isomerase A (*PPIA*). Relative transcript levels for both human and mouse systems are shown, and *error bars* indicate standard deviations between three replicates for each differentiation system. Four NETs were induced similarly to positive differentiation controls, *MYOG* and *MYH1*.

POPDC2 and LOC203547 antibodies produced distinctive nuclear rim staining on rat heart muscle cryosections (Fig. 5*C*). The nuclei were clearly visualized with DAPI, and the antibody staining was concentrated in a crisp rim just surrounding the nucleus. The TMEM38A antibodies produced strong nuclear rim staining in rat leg skeletal muscle cryosections in addition to fainter staining in the SR (Fig. 5*C*). This is consistent with supplemental data from a previous report on this protein that introduced it as an SR protein (60). The appearance of TMEM38A in the INM as well as the SR resembles the well characterized NET Emerin that has been reported in the cytoplasm of myotubes and interstitial discs of heart tissue (55, 56, 61). Consistent with the immunofluorescence results, some peptides were obtained in the mass spectrometry analysis of the SR fraction for TMEM38A, Emerin, and SUN2 (supplemental Table S3).

Many previously characterized NETs resist a prefixation extraction with 1% Triton X-100 and 400 mM KCl presumably due to tight associations with the intermediate filament Lamin polymer. The functioning antibodies enabled testing of this on mouse C2C12 myoblast cells. C2C12 cells grown on coverslips were extracted with 1% Triton X-100 prior to fixation. The antibodies concentrated in a rim around the DNA, consistent with retention of these NETs after the detergent preextraction and consistent with observations of INM targeting by OMX. One of the NETs tested, LOC203547, had not been

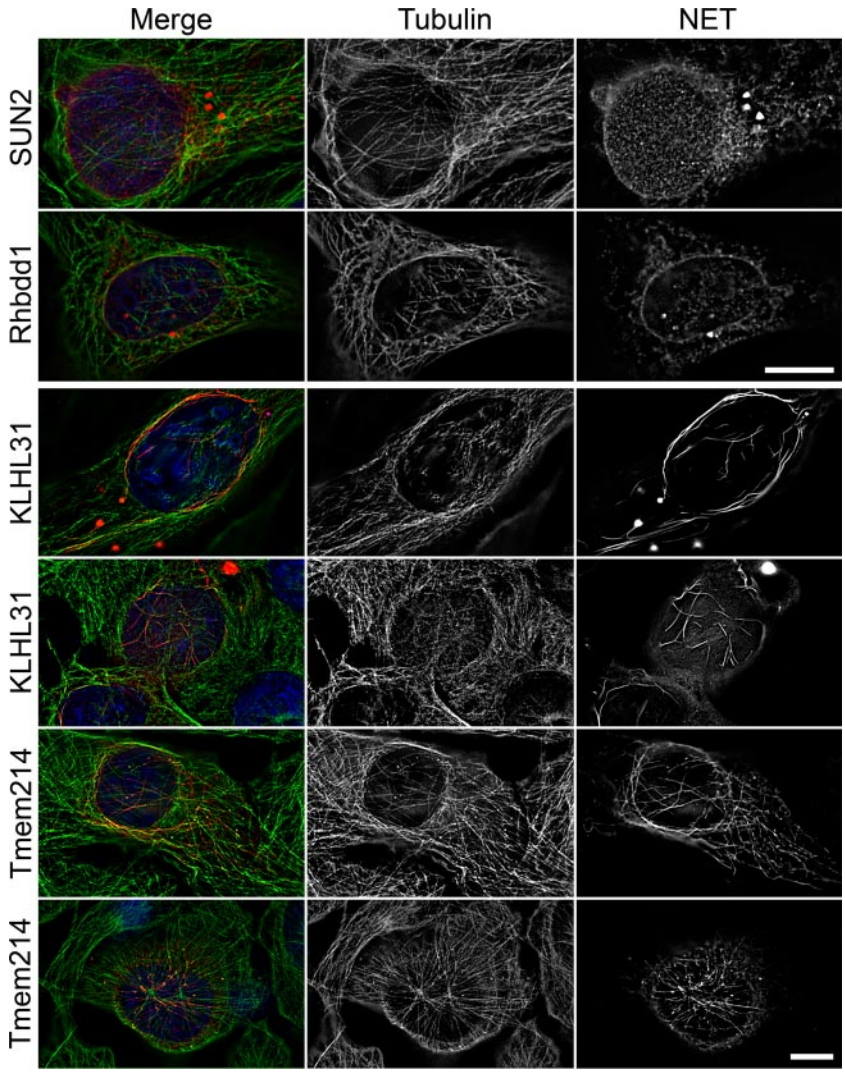


FIG. 7. **Muscle NET alignment with microtubules.** U2OS cells expressing various NET-GFP fusions were fixed and stained for microtubules. Sections from deconvolved images are shown for a focal plane at the upper nuclear surface. DAPI staining for DNA is in *blue*, microtubules are in *green*, and the NET is in *red*. In the *right columns*, the microtubules or NET are shown in *grayscale*. Microtubule filaments are observed crossing the nuclear surface in the SUN2 control-transfected cell, but the SUN2 staining on the surface is spotty, consistent with previous reports. The same was observed for RHBDD1 and most muscle NETs tested; however, KLHL31, which had a more filamentous appearance in Fig. 4, exhibited filaments also on the nuclear surface that partly tracked with the microtubule filaments (note both *grayscale* images for each and some *yellow* co-localization in the *merge panel*). TMEM214, although consistently yielding a much crisper and continuous rim at the nuclear periphery, also exhibited filamentous accumulation at the nuclear surface that exhibited even more overlap with microtubules. *Scale bars*, 10 μm.

cloned or tested by exogenous expression, bringing the number of new NETs with confirmed NE residence to 11. Another of the NETs shown, TMEM201, was previously identified in the liver data set as NET5 (26) but was tested here because it had many more peptides in the skeletal muscle data set (supplemental Table S3).

Several Muscle NETs Are Up-regulated during Muscle Differentiation—Muscle-specific NETs could have functions related to myogenic differentiation. Thus, expression levels of several skeletal muscle NETs were analyzed in myogenic differentiation systems where myoblast cell lines are grown to confluency and then moved to reduced serum medium containing pharmacological agents to induce differentiation. The myoblasts then fuse to form multinucleate myotubes with several characteristics of differentiated muscle. In both human (RD; Ref. 62) and mouse (C2C12; Ref. 63, 64) systems, negative controls were unchanged, whereas *TMEM38A* and *KLHL31* were induced similarly to the positive controls, myogenin (*MYOG*) and myosin heavy chain (*MYH1*) (Fig. 6). *POPDC2* and *TMEM161A* were both up-regulated only in the mouse system. The up-regulation of certain NETs in myogenic differentiation is consistent with their tissue-restricted expression patterns and the fact that the antibodies yielded more distinctive nuclear rim staining in skeletal and heart muscle cryosections than in cell lines.

Three Novel Skeletal Muscle NETs Appear to Associate with the Cytoskeleton—Nesprins 1 and 2 in the ONM bind to the actin cytoskeleton and contribute to nuclear positioning (17), whereas Nesprin 3, originally identified by proteomics of liver NEs (NET53; Ref. 26), binds plectin and thus presumably connects the NE to the intermediate filament cytoskeleton (24). These ONM NETs are also connected to the INM SUN proteins that in turn bind Lamins, hence connecting the cytoskeleton to the nucleoskeleton (15). Thus, NETs in both the ONM and INM are involved in connecting cytoskeletal filaments to the nucleoskeleton. The importance of such connections is illustrated by the recent discovery that mutations in Nesprins 1 and 2 (23) also cause EDMD.

NETs connecting microtubules to the nucleus have not previously been identified, but it is clear that microtubules have some interactions with the NE because they associate with the NE prior to NE disassembly in mitosis (65), and NET5/TMEM201/Samp1 associates with both the centrosome and the mitotic spindle (30, 66). As muscle cells have many unique cytoskeletal features, the novel NETs from skeletal muscle were screened for effects indicating any associations with microtubules.

The NETs were overexpressed in U2OS cells and co-visualized with the microtubule cytoskeleton. Overexpression of most NETs yielded no visible alterations in microtubule organization. However, cell sections at the upper surface of the nucleus revealed a filamentous distribution for KLHL31 and TMEM214 (Fig. 7) in nearly all transfected cells. This strongly contrasted with the typical punctate distribution of NETs at the nuclear surface (SUN2 and RHBDD1 shown). Overlaying the NET and microtubule staining patterns in these sections for KLHL31 and TMEM214 revealed significant co-localization; however, the percentage of microtubules visible at the nuclear surface that tracked with the NETs was highly variable from cell to cell. The *upper* cell for TMEM214 shows also how some of the overexpressed NET distributed in the ER also tracked with the microtubules in the cytoplasm and the lack of a clear microtubule-organizing center. The *lower* cell shows a different defect with two microtubule-organizing centers at the center of the nuclear surface.

Observations of NPC-associated proteins and the NE Lamin proteins on the mitotic spindle (*e.g.* Lamin B, Importin β, and Ran; Refs. 67–69) provided the first evidence that NE proteins have separate roles in mitosis. A more recent study showed that NET5/TMEM201/Samp1 associates with the base of the mitotic spindle in mitosis (30). During mitosis, many NETs are redistributed into small vesicles that spread throughout the cell except from the spindle area. This pattern was observed for most NETs with SUN2 and RHBDD1 shown as examples (Fig. 8). In contrast, TMEM214 exhibited an increased concentration by the spindle base in most cells, although it was excluded from the spindle poles themselves (Fig. 8, *upper right TMEM214 panel*). Some cells exhibited an even more striking microtubule association with clear co-localization throughout the length of many spindle microtubules (Fig. 8, *lower TMEM214 panels*). KLHL31 was distributed differently in mitosis, associating more with the mitotic chromosomes than with microtubules. WFS1, although not tracking with microtubules in interphase, exhibited a distribution similar to that of TMEM214 with a concentration around the base of the spindle in mitosis (Fig. 8, *bottom panels*).

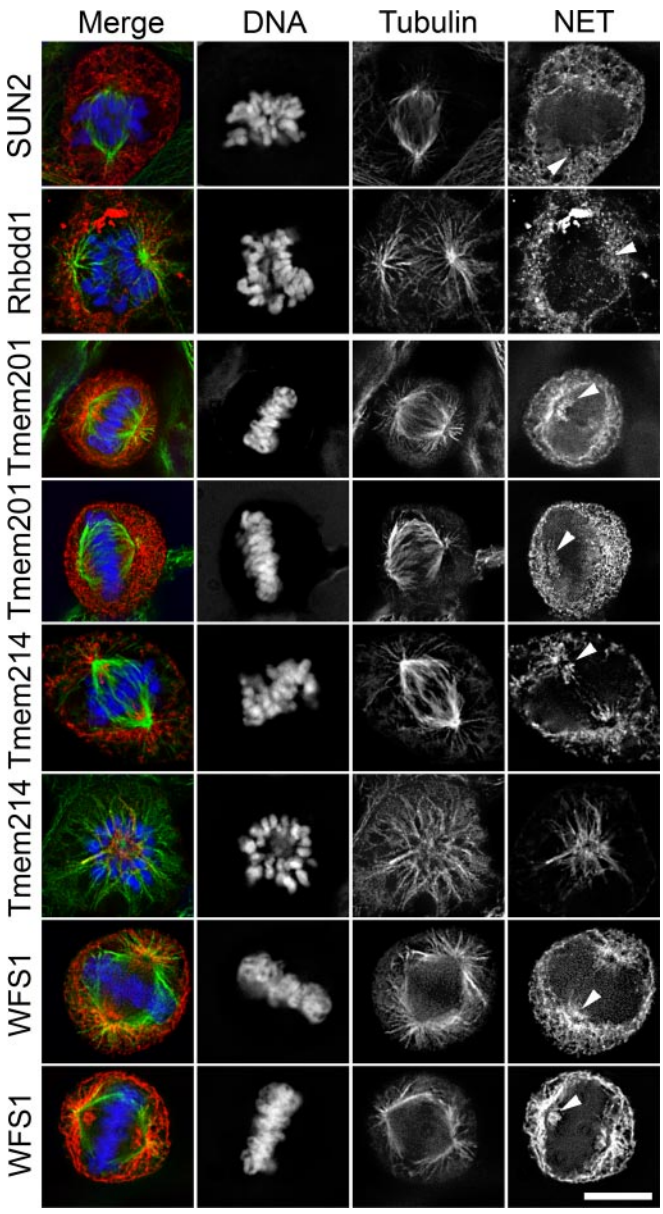


FIG. 8. **Partial NET-microtubule associations are also observed in mitosis.** In mitosis, the NE breaks down, and several NPC proteins and Lamins have previously been reported to partly assemble on the mitotic spindle. The negative control SUN2 and the new muscle-specific NET RHBDD1 were generally distributed throughout the mitotic cell but excluded from the spindle (*white arrowheads* point to one of the spindles in each cell). No particular accumulation or change in the NET distribution was observed with respect to the mitotic spindle. As a positive control, TMEM201 (NET5/Samp1) is shown as it has previously been shown to associate with the spindle. TMEM214 exhibited in most cells a strong accumulation at the base of the spindles close to the poles but was excluded from the poles themselves. In other cells, the distribution was even more focused on the spindle itself with significant co-localization with the microtubules. A third muscle NET, WFS1, consistently yielded the first phenotype of TMEM214; however, in many cases, an even more pronounced accumulation surrounding the spindle poles was observed. One spindle pole in each cell (except the prophase TMEM214 cell) is marked by an *arrowhead*. *Scale bar*, 10 μm.

DISCUSSION

This analysis of the skeletal muscle NE proteome identified many new NE-associated proteins and NETs not identified in earlier studies of liver and neuroblastoma NEs (26, 27) or a more recent study of blood NEs (28). Moreover, some of these

proteins had very restricted tissue expression and were specifically up-regulated in myogenic differentiation systems. Thus, each tissue likely has a different combination of proteins in its NE proteome. That some of these more tissue-restricted NETs track with the microtubule cytoskeleton has the potential to explain some of the tissue specificity of NE disease phenotypes. The three favored hypotheses to explain how NE protein mutation can yield pathology are altered expression of genes regulated from the nuclear periphery, mechanical instability from disruption of lamina-cytoskeleton interactions, and disabling of the cell cycle/stem cell maintenance (6, 7). That NET-microtubule associations indicated here appear to occur in both interphase and mitosis provides potential mechanisms toward both mechanical instability and cell cycle regulation. That some of these proteins are tissue-restricted provides a potential explanation for the tissue preferences in NE disease pathologies.

Much additional work will be required to determine such disease links and whether associations are direct or indirect. Nonetheless, the tracking of TMEM214 and KLHL31 with microtubules on the nuclear surface is both striking and a novel finding among all NETs identified thus far. The association of TMEM214 and WFS1 at the base of the spindles in mitosis resembles that observed for NET5/TMEM201/Samp1 (30). This finding followed observations that Ima1, the yeast homolog of this NET, is involved in connecting the NE to the centrosome (66). Thus, the potential link to cell cycle regulation is even stronger.

The idea of more tissue-specific NETs participating in protein complexes at the NE whose disruption could lead to more muscle-specific pathologies makes sense (70). It is hard otherwise to explain how mutations in ubiquitous proteins like Lamin A, Emerin, or the Nesprin proteins could result in tissue-specific pathologies (18–23). Interestingly, WFS1 has already been linked to three disorders, although the function of the protein remains obscure (71–73). These initial successes in identifying tissue-specific NETs that have links to disease indicate the importance of investigating the NE proteome in all tissues linked to disease.

Acknowledgments—Use of the OMX microscope is supported by the Scottish University Life Sciences Alliance. We thank D. Tollervey and W. C. Earnshaw for comments and F. Melchior, T. Rapoport, E. King, D. Kelly, T. Guan, and S. Mitchell for reagents/technical assistance.

* This study was supported by the Wellcome Trust (senior research fellowship to E. C. S.) and The Stowers Institute for Medical Research (to L. F.).

S This article contains supplemental Figs. S1 and S2 and Tables S1–S5.

§ Both authors contributed equally to this work.

‖ A Royal Society Dorothy Hodgkin fellow.

** A Darwin Trust student.

‡‡ To whom correspondence should be addressed: Wellcome Trust Centre for Cell Biology, University of Edinburgh, Kings Bldgs., Swann 5.22, Mayfield Rd., Edinburgh EH9 3JR, UK. Tel.: 441316507075; Fax: 441316507360; E-mail: e.schirmer@ed.ac.uk.

REFERENCES

1. Prunuske, A. J., and Ullman, K. S. (2006) The nuclear envelope: form and reformation. *Curr. Opin. Cell Biol.* **18,** 108–116
2. Suntharalingam, M., and Wente, S. R. (2003) Peering through the pore: nuclear pore complex structure, assembly, and function. *Dev. Cell* **4,** 775–789
3. Callan, H. G., and Tomlin, S. G. (1950) Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. *Proc. R. Soc. Lond. B Biol. Sci.* **137,** 367–378
4. Schirmer, E. C., and Foisner, R. (2007) Proteins that associate with lamins: many faces, many functions. *Exp. Cell Res.* **313,** 2167–2179
5. Stewart, C. L., Roux, K. J., and Burke, B. (2007) Blurring the boundary: the nuclear envelope extends its reach. *Science* **318,** 1408–1412
6. Worman, H. J., and Bonne, G. (2007) "Laminopathies": a wide spectrum of human diseases. *Exp. Cell Res.* **313,** 2121–2133
7. Dorner, D., Gotzmann, J., and Foisner, R. (2007) Nucleoplasmic lamins and their interaction partners, LAP2alpha, Rb, and BAF, in transcriptional regulation. *FEBS J.* **274,** 1362–1373
8. Lenz-Böhme, B., Wismar, J., Fuchs, S., Reifegerste, R., Buchner, E., Betz, H., and Schmitt, B. (1997) Insertional mutation of the Drosophila nuclear lamin Dm0 gene results in defective nuclear envelopes, clustering of nuclear pore complexes, and accumulation of annulate lamellae. *J. Cell Biol.* **137,** 1001–1016
9. Liu, J., Rolef 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
10. Lammerding, J., Schulze, P. C., Takahashi, T., Kozlov, S., Sullivan, T., Kamm, R. D., Stewart, C. L., and Lee, R. T. (2004) Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. *J. Clin. Investig.* **113,** 370–378
11. 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
12. Olins, A. L., Herrmann, H., Lichter, P., Kratzmeier, M., Doenecke, D., and Olins, D. E. (2001) Nuclear envelope and chromatin compositional differences comparing undifferentiated and retinoic acid- and phorbol ester-treated HL-60 cells. *Exp. Cell Res.* **268,** 115–127
13. 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.* **279,** 42811–42817
14. Lammerding, J., Fong, L. G., Ji, J. Y., Reue, K., Stewart, C. L., Young, S. G., and Lee, R. T. (2006) Lamins A and C but not lamin B1 regulate nuclear mechanics. *J. Biol. Chem.* **281,** 25768–25780
15. 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
16. Starr, D. A., and Fischer, J. A. (2005) KASH 'n Karry: the KASH domain family of cargo-specific cytoskeletal adaptor proteins. *BioEssays* **27,** 1136–1146
17. Starr, D. A., and Han, M. (2002) Role of ANC-1 in tethering nuclei to the actin cytoskeleton. *Science* **298,** 406–409
18. Bonne, G., Di Barletta, M. R., Varnous, S., Bécane, H. M., Hammouda, E. H., Merlini, L., Muntoni, F., Greenberg, C. R., Gary, F., Urtizberea, J. A., Duboc, D., Fardeau, M., Toniolo, D., and Schwartz, K. (1999) Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **21,** 285–288
19. Raffaele Di Barletta, M., Ricci, E., Galluzzi, G., Tonali, P., Mora, M., Morandi, L., Romorini, A., Voit, T., Orstavik, K. H., Merlini, L., Trevisan, C., Biancalana, V., Housmanowa-Petrusewicz, I., Bione, S., Ricotti, R., Schwartz, K., Bonne, G., and Toniolo, D. (2000) Different mutations in the LMNA gene cause autosomal dominant and autosomal recessive Emery-Dreifuss muscular dystrophy. *Am. J. Hum. Genet.* **66,** 1407–1412
20. 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
21. Fatkin, D., MacRae, C., Sasaki, T., Wolff, M. R., Porcu, M., Frenneaux, M.,

Atherton, J., Vidaillet, H. J., Jr., Spudich, S., De Girolami, U., Seidman, J. G., Seidman, C., Muntoni, F., Müehle, G., Johnson, W., and McDonough, B. (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

22. 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
23. Zhang, Q., Bethmann, C., Worth, N. F., Davies, J. D., Wasner, C., Feuer, A., Ragnauth, C. D., Yi, Q., Mellad, J. A., Warren, D. T., Wheeler, M. A., Ellis, J. A., Skepper, J. N., Vorgerd, M., Schlotter-Weigel, B., Weissberg, P. L., Roberts, R. G., Wehnert, M., and Shanahan, C. M. (2007) Nesprin-1 and -2 are involved in the pathogenesis of Emery-Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. *Hum. Mol. Genet* **16,** 2816–2833
24. 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
25. Zhang, Q., Ragnauth, C. D., Skepper, J. N., Worth, N. F., Warren, D. T., Roberts, R. G., Weissberg, P. L., Ellis, J. A., and Shanahan, C. M. (2005) Nesprin-2 is a multi-isomeric protein that binds lamin and emerin at the nuclear envelope and forms a subcellular network in skeletal muscle. *J. Cell Sci.* **118,** 673–687
26. Schirmer, E. C., Florens, L., Guan, T., Yates, J. R., 3rd, and Gerace, L. (2003) Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* **301,** 1380–1382
27. Dreger, M., Bengtsson, L., Schöneberg, 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
28. Korfali, N., Wilkie, G. S., Swanson, S. K., Srsen, V., Batrakou, D. G., Fairley, E. A., Malik, P., Zuleger, N., Goncharevich, A., de Las Heras, J., Kelly, D. A., Kerr, A. R., Florens, L., and Schirmer, E. C. (September 13, 2010) The leukocyte nuclear envelope proteome varies with cell activation and contains novel transmembrane proteins that affect genome architecture. *Mol. Cell. Proteomics* 10.1074/mcp.M110.002915
29. 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., Patapoutian, A., Hampton, G. M., Schultz, P. G., and Hogenesch, J. B. (2002) Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 4465–4470
30. Buch, C., Lindberg, R., Figueroa, R., Gudise, S., Onischenko, E., and Hallberg, E. (2009) An integral protein of the inner nuclear membrane localizes to the mitotic spindle in mammalian cells. *J. Cell Sci.* **122,** 2100–2107
31. Wilkie, G. S., and Schirmer, E. C. (2008) Purification of nuclei and preparation of nuclear envelopes from skeletal muscle. *Methods Mol. Biol.* **463,** 23–41
32. Hahn, C. G., and Covault, J. (1990) Isolation of transcriptionally active nuclei from striated muscle using Percoll density gradients. *Anal. Biochem.* **190,** 193–197
33. Florens, L., Korfali, N., and Schirmer, E. C. (2008) Subcellular fractionation and proteomics of nuclear envelopes. *Methods Mol. Biol.* **432,** 117–137
34. Wu, C. C., MacCoss, M. J., Howell, K. E., and Yates, J. R., 3rd (2003) A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol.* **21,** 532–538
35. Florens, L., and Washburn, M. P. (2006) Proteomic analysis by multidimensional protein identification technology. *Methods Mol. Biol.* **328,** 159–175
36. McDonald, W. H., Tabb, D. L., Sadygov, R. G., MacCoss, M. J., Venable, J., Graumann, J., Johnson, J. R., Cociorva, D., and Yates, J. R., 3rd (2004) MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Commun. Mass. Spectrom.* **18,** 2162–2168
37. Venable, J. D., Dong, M. Q., Wohlschlegel, J., Dillin, A., and Yates, J. R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* **1,** 39–45
38. Eng, J. K., McCormack, A. L., and Yates, J. R., 3rd (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989
39. Tabb, D. L., McDonald, W. H., and Yates, J. R., 3rd (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1,** 21–26
40. Zybailov, B. L., Florens, L., and Washburn, M. P. (2007) Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. *Mol. Biosyst.* **3,** 354–360
41. Florens, L., Carozza, M. J., Swanson, S. K., Fournier, M., Coleman, M. K., Workman, J. L., and Washburn, M. P. (2006) Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **40,** 303–311
42. Paoletti, A. C., Parmely, T. J., Tomomori-Sato, C., Sato, S., Zhu, D., Conaway, R. C., Conaway, J. W., Florens, L., and Washburn, M. P. (2006) Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 18928–18933
43. Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L., and Washburn, M. P. (2006) Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **5,** 2339–2347
44. Zhang, Y., Wen, Z., Washburn, M. P., and Florens, L. (2010) Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal. Chem.* **82,** 2272–2281
45. Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Gräf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kähäri, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A., and Searle, S. (2008) Ensembl 2008. *Nucleic Acids Res.* **36,** D707–D714
46. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305,** 567–580
47. Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., and Lewis, S. (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics* **25,** 288–289
48. Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C. L., Haase, J., Janes, J., Huss, J. W., 3rd, and Su, A. I. (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* **10,** R130
49. Mootha, V. K., Bunkenborg, J., Olsen, J. V., Hjerrild, M., Wisniewski, J. R., Stahl, E., Bolouri, M. S., Ray, H. N., Sihag, S., Kamal, M., Patterson, N., Lander, E. S., and Mann, M. (2003) Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115,** 629–640
50. 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
51. Hodzic, D. M., Yeater, D. B., Bengtsson, L., Otto, H., and Stahl, P. D. (2004) Sun2 is a novel mammalian inner nuclear membrane protein. *J. Biol. Chem.* **279,** 25805–25812
52. 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
53. 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
54. 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
55. Cartegni, L., di Barletta, M. R., Barresi, R., Squarzoni, S., Sabatelli, P., Maraldi, N., Mora, M., Di Blasi, C., Cornelio, F., Merlini, L., Villa, A., Cobianchi, F., and Toniolo, D. (1997) Heart-specific localization of emerin: new insights into Emery-Dreifuss muscular dystrophy. *Hum. Mol. Genet* **6,** 2257–2264
56. Lattanzi, G., Ognibene, A., Sabatelli, P., Capanni, C., Toniolo, D., Columbaro, M., Santi, S., Riccio, M., Merlini, L., Maraldi, N. M., and Squarzoni,

S. (2000) Emerin expression at the early stages of myogenic differentiation. *Differentiation* **66,** 208–217
57. Salpingidou, G., Smertenko, A., Hausmanowa-Petrucewicz, I., Hussey, P. J., and Hutchison, C. J. (2007) A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane. *J. Cell Biol.* **178,** 897–904
58. Foster, L. J., de Hoog, C. L., Zhang, Y., Zhang, Y., Xie, X., Mootha, V. K., and Mann, M. (2006) A mammalian organelle map by protein correlation profiling. *Cell* **125,** 187–199
59. Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M. C., Agard, D. A., Gustafsson, M. G., Leonhardt, H., and Sedat, J. W. (2008) Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320,** 1332–1336
60. Yazawa, M., Ferrante, C., Feng, J., Mio, K., Ogura, T., Zhang, M., Lin, P. H., Pan, Z., Komazaki, S., Kato, K., Nishi, M., Zhao, X., Weisleder, N., Sato, C., Ma, J., and Takeshima, H. (2007) TRIC channels are essential for Ca^{2+} handling in intracellular stores. *Nature* **448,** 78–82
61. Wheeler, M. A., Warley, A., Roberts, R. G., Ehler, E., and Ellis, J. A. (2010) Identification of an emerin-beta-catenin complex in the heart important for intercalated disc architecture and beta-catenin localisation. *Cell. Mol. Life Sci.* **67,** 781–796
62. McAllister, R. M., Melnyk, J., Finkelstein, J. Z., Adams, E. C., Jr., and Gardner, M. B. (1969) Cultivation in vitro of cells derived from a human rhabdomyosarcoma. *Cancer* **24,** 520–526
63. Blau, H. M., Pavlath, G. K., Hardeman, E. C., Chiu, C. P., Silberstein, L., Webster, S. G., Miller, S. C., and Webster, C. (1985) Plasticity of the differentiated state. *Science* **230,** 758–766
64. Yaffe, D., and Saxel, O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* **270,** 725–727
65. Salina, D., Bodoor, K., Eckley, D. M., Schroer, T. A., Rattner, J. B., and Burke, B. (2002) Cytoplasmic dynein as a facilitator of nuclear envelope breakdown. *Cell* **108,** 97–107
66. King, M. C., Drivas, T. G., and Blobel, G. (2008) A network of nuclear envelope membrane proteins linking centromeres to microtubules. *Cell* **134,** 427–438
67. Ohba, T., Nakamura, M., Nishitani, H., and Nishimoto, T. (1999) Self-organization of microtubule asters induced in Xenopus egg extracts by GTP-bound Ran. *Science* **284,** 1356–1358
68. Tsai, M. Y., Wang, S., Heidinger, J. M., Shumaker, D. K., Adam, S. A., Goldman, R. D., and Zheng, Y. (2006) A mitotic lamin B matrix induced by RanGTP required for spindle assembly. *Science* **311,** 1887–1893
69. Wiese, C., Wilde, A., Moore, M. S., Adam, S. A., Merdes, A., and Zheng, Y. (2001) Role of importin-beta in coupling Ran to downstream targets in microtubule assembly. *Science* **291,** 653–656
70. Wilkie, G. S., and Schirmer, E. C. (2006) Guilt by association: the nuclear envelope proteome and disease. *Mol. Cell. Proteomics* **5,** 1865–1875
71. Bespalova, I. N., Van Camp, G., Bom, S. J., Brown, D. J., Cryns, K., DeWan, A. T., Erson, A. E., Flothmann, K., Kunst, H. P., Kurnool, P., Sivakumaran, T. A., Cremers, C. W., Leal, S. M., Burmeister, M., and Lesperance, M. M. (2001) Mutations in the Wolfram syndrome 1 gene (WFS1) are a common cause of low frequency sensorineural hearing loss. *Hum. Mol. Genet.* **10,** 2501–2508
72. Domènech, E., Gómez-Zaera, M., and Nunes, V. (2002) WFS1 mutations in Spanish patients with diabetes mellitus and deafness. *Eur. J. Hum. Genet.* **10,** 421–426
73. Strom, T. M., Hörtnagel, K., Hofmann, S., Gekeler, F., Scharfe, C., Rabl, W., Gerbitz, K. D., and Meitinger, T. (1998) Diabetes insipidus, diabetes mellitus, optic atrophy and deafness (DIDMOAD) caused by mutations in a novel gene (wolframin) coding for a predicted transmembrane protein. *Hum. Mol. Genet.* **7,** 2021–2028

OPEN ACCESS Freely available online

PLoS one

A Flow Cytometry-Based Screen of Nuclear Envelope Transmembrane Proteins Identifies NET4/Tmem53 as Involved in Stress-Dependent Cell Cycle Withdrawal

Nadia Korfali[1☯], Vlastimil Srsen[1☯], Martin Waterfall[2], Dzmitry G. Batrakou[1], Vanja Pekovic[3], Christopher J. Hutchison[3], Eric C. Schirmer[1]*

1 The Wellcome Trust Centre for Cell Biology and Institute of Cell Biology, University of Edinburgh, Edinburgh, United Kingdom, **2** Institute of Immunology and Infection Research, University of Edinburgh, Edinburgh, United Kingdom, **3** School of Biological and Biomedical Sciences, Durham University, Durham, United Kingdom

Abstract

Disruption of cell cycle regulation is one mechanism proposed for how nuclear envelope protein mutation can cause disease. Thus far only a few nuclear envelope proteins have been tested/found to affect cell cycle progression: to identify others, 39 novel nuclear envelope transmembrane proteins were screened for their ability to alter flow cytometry cell cycle/DNA content profiles when exogenously expressed. Eight had notable effects with seven increasing and one decreasing the 4N:2N ratio. We subsequently focused on NET4/Tmem53 that lost its effects in $p53^{-/-}$ cells and retinoblastoma protein-deficient cells. NET4/TMEM53 knockdown by siRNA altered flow cytometry cell cycle/DNA content profiles in a similar way as overexpression. NET4/TMEM53 knockdown did not affect total retinoblastoma protein levels, unlike nuclear envelope-associated proteins Lamin A and LAP2α. However, a decrease in phosphorylated retinoblastoma protein was observed along with a doubling of p53 levels and a 7-fold increase in p21. Consequently cells withdrew from the cell cycle, which was confirmed in MRC5 cells by a drop in the percentage of cells expressing Ki-67 antigen and an increase in the number of cells stained for ß-galactosidase. The ß-galactosidase upregulation suggests that cells become prematurely senescent. Finally, the changes in retinoblastoma protein, p53, and p21 resulting from loss of NET4/Tmem53 were dependent upon active p38 MAP kinase. The finding that roughly a fifth of nuclear envelope transmembrane proteins screened yielded alterations in flow cytometry cell cycle/DNA content profiles suggests a much greater influence of the nuclear envelope on the cell cycle than is widely held.

Citation: Korfali N, Srsen V, Waterfall M, Batrakou DG, Pekovic V, et al. (2011) A Flow Cytometry-Based Screen of Nuclear Envelope Transmembrane Proteins Identifies NET4/Tmem53 as Involved in Stress-Dependent Cell Cycle Withdrawal. PLoS ONE 6(4): e18762. doi:10.1371/journal.pone.0018762

Editor: Joanna Mary Bridger, Brunel University, United Kingdom

Received January 7, 2011; **Accepted** March 17, 2011; **Published** April 14, 2011

Funding: This work was funded by a Wellcome Trust Senior Research Fellowship to ECS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: e.schirmer@ed.ac.uk

☯ These authors contributed equally to this work.

Introduction

Several proteins of the nuclear envelope are linked to human diseases ranging from muscular dystrophies to neuropathy, bone diseases, and progeroid aging syndromes [1,2]. These proteins include the intermediate filament A/C Lamins and several proteins integral to the nuclear membrane. Favored molecular mechanisms to explain how mutations in nuclear envelope proteins produce pathology include loss of nuclear mechanical stability, alterations in gene expression, and cell cycle/stem cell maintenance defects (reviewed in [2,3,4]). However, the known functions of the proteins mutated in disease are insufficient to fully explain the pathologies observed without assistance from partner proteins that thus far have not been identified.

The first indication of a link between nuclear envelope diseases and the cell cycle came from studies with specific mutations in the nuclear envelope transmembrane protein (NET) Emerin linked to Emery-Dreifuss muscular dystrophy. It was reported that two disease-linked mutations prolonged S-phase from 12 h to 22 h when overexpressed in COS-7 cells [5]; however similar effects were not observed in all disease mutants and so this was not investigated in further detail. In *C. elegans* disruption of Emerin alone did not have a strong effect on the cell cycle, but when combined with disruption of a second NET, MAN1, it did [6]. Loss of Emerin has also been reported to interfere with retinoblastoma protein (pRb)-regulated genes in mouse and consequently with myogenic differentiation [7], and the same pRb-dependent cell cycle exit is disrupted in nuclear envelope-linked muscular dystrophy [8]. pRb is a tumor suppressor that regulates the cell cycle at the G1/S transition by regulating the E2F family of transcription factors (reviewed in [9]). pRb also interacts with Lamin A [10], but this is thought to principally involve the nucleoplasmic and not the nuclear envelope pool of Lamin A because it operates in a complex with LAP2α, a soluble splice variant of the nuclear envelope protein LAP2 that is principally found in the nucleoplasm [11,12,13].

To determine if any of several newly identified nuclear envelope proteins play a role in the cell cycle, 39 novel confirmed NETs were screened for their ability to alter flow cytometry cell cycle/DNA content profiles when exogenously expressed. These NETs

were identified in two recent proteomic analyses of liver and blood cells [14,15]. Seven of the NETs tested showed an increase in the 4N:2N ratio while one showed a decrease. To determine if pathways affected by these NETs involved the p53 master cell cycle regulator, these eight NETs were retested in $p53^{-/-}$ cells. The change in 4N:2N ratios still occurred in the absence of p53 for most NETs, but the effect of NET4/Tmem53 and NET59/Ncln was lost. NET4/Tmem53 was selected for a more detailed analysis of how it interacts with the p53 pathway. Knockdown of NET4/TMEM53 resulted in cell cycle withdrawal, apparently through activation of the p38 kinase with consequent upregulation of p53 and p21 and downregulation of phosphorylated pRb.

Results

A screen for NETs that alter flow cytometry profiles

To identify nuclear envelope proteins that might contribute to cell cycle progression, a collection of 39 NETs were screened for their ability to affect flow cytometry cell cycle/DNA content profiles. All NETs were fused to a monomeric red fluorescent protein (mRFP) tag at their carboxyl-termini and were previously confirmed to target to the nuclear envelope [14,15,16]. HEK293T human embryonic kidney cells were used for the screen because this cell line is efficiently transfected, easily recovered from plates for the flow cytometry experiments, and has a relatively stable karyotype compared to other commonly used lines such as HeLa, U2OS or HT1080 cells. Tagged NETs were transiently transfected into the HEK293T cells and after 40–48 h of expression the frequency of live cells with 2N or 4N DNA content was measured by flow cytometry.

DNA profiles were acquired for both the transfected cells (mRFP positive) and the untransfected population for each transfection. Thus the use of transient transfections provided an internal control for each experiment that removed any cell cycle variation between plates and/or due to the transfection reagent. For each NET at least three independent flow cytometry experiments were performed, each on different days and with a minimum of 1,000 singlet transfected cells (and in most cases >5,000 cells) analyzed. For those NETs where a strong effect was observed, additional repeats were done with 20,000 transfected cells analyzed to increase confidence.

Examples of flow cytometry cell cycle/DNA content profiles are shown in Figure 1 with the untransfected cell traces (blue) overlaid with those of the mRFP-expressing population (red). Cell fragments and apoptosing cells were excluded based on propidium iodide (PI) staining and FSC/SSC (from light scattering). The flow cytometry profiles for the mRFP control and many other NETs tested were indistinguishable from those of untransfected cells in the same population or only exhibited minor differences. By contrast, NET11/Sccpdh, NET31/Tmem209, NET59/Ncln, Tmub1, Fam3c, Magt1 and Tmem126a all yielded striking accumulations of cells with a 4N DNA content suggesting an increased G2/M population (Figure 1). NET4/Tmem53 yielded a different effect, exhibiting a reduction in cells with 4N DNA content suggestive of more cells in the G1 phase of the cell cycle. The percentages of cells in G1, S, and G2/M phases based on DNA content are listed in Table 1. While effects of these eight NETs were the most striking and reproducible, one cannot discount that some NETs that caused minor changes might also be relevant to the cell cycle.

It is possible that some NETs positive in the screen could have altered flow cytometry DNA content profiles because of aberrant nuclear morphologies as opposed to effects on cell cycle progression (though this could in turn reflect problems with cytokinesis). To test if this was a likely explanation, cells from the transfected populations were imaged by fluorescence microscopy. Representative images revealed no gross aberration in nuclear morphology within the transfected population, indicating this is unlikely to have affected the flow cytometry results (Figure 2). Nuclear envelope targeting is not always extremely clear in the HEK293T cells either because NETs have multiple localizations or because the high expression saturates binding sites at the nuclear envelope in these cells; however, all NETs tested here were previously confirmed to target to the nuclear envelope [14,15,16].

Some NETs depend on p53 and/or pRb for their effects

The cell cycle protein p53 is often referred to as a master regulator because it has a role in a large number of cell cycle pathways. The ability of the eight NETs identified in the initial screen to alter the flow cytometry DNA content profiles was re-examined in the HCT116 $p53^{-/-}$ cell line. The controls mRFP and NET51/C14orf1 had not yielded changes in the flow cytometry DNA content profile in the HEK293T cells and similarly yielded no significant changes in the $p53^{-/-}$ cells (Figure 3A). Correspondingly, NET11/Sccpdh, NET31/Tmem209, Tmub1, Fam3c, Magt1, and Tmem126a that had exhibited increases in the 4N population in the HEK293T cells yielded similar changes in the $p53^{-/-}$ cells; so p53 does not appear to be involved in the potential effects of these NETs on the cell cycle. By contrast, in the $p53^{-/-}$ cells NET59/Ncln no longer exhibited a 4N increase and NET4/Tmem53 no longer exhibited a 4N decrease as had been observed in the HEK293T cells. Thus changes in the flow cytometry profiles from expression of NET59/Ncln and NET4/Tmem53 appear to be p53 dependent. To better compare the results in the p53 positive and negative cells, the 4N:2N ratios from both cell lines were plotted for this set of NETs (Figure 3B). For NETs other than the two that lost their effects a similar pattern was observed between the two cell lines, although in some cases the 4N:2N ratio increase was slightly higher in the HCT116 $p53^{-/-}$ cell line. NET4/Tmem53, a previously uncharacterized protein with no known functional domains, was subsequently followed in more detail.

Cell cycle effects dependent on p53 often involve changes in pRb [17] and links have previously been identified between pRb and nuclear envelope-associated proteins [10,11,12,13]. To determine if pRb also plays a role in the NET4/Tmem53-directed effects, flow cytometry cell cycle/DNA content profiles were determined for cells expressing NET4/Tmem53 or controls in HEK293T cells with normal or reduced levels of pRb, using siRNA oligos to knock down pRb. The reduction in the 4N:2N ratio caused by NET4/Tmem53 was lost in the pRb-depleted cells (Figure 4A). As pRb phosphorylation is crucial for its role in cell cycle progression [18,19,20], antibodies to a form of pRb phosphorylated at serine 780 were utilized to test if phosphorylated pRb levels were affected in the NET4/Tmem53 transfected cells. Because of low and variable transfection efficiencies this could not be assayed at a population level; thus transfected cells were stained with the antibodies and the levels of the phosphorylated pRb in the nucleoplasm were quantified by measuring the average pixel intensity (Figure 4B). Plotting these values revealed a significant loss of phosphorylated pRb in cells expressing NET4/Tmem53.

RNAi knockdown of NET4/Tmem53

Further study of the pathways through which NET4/Tmem53 affects cell cycle regulation would not be practical using exogenous expression because transfection efficiencies were too low (5–10%) to be able to quantify changes in pathway components by Western

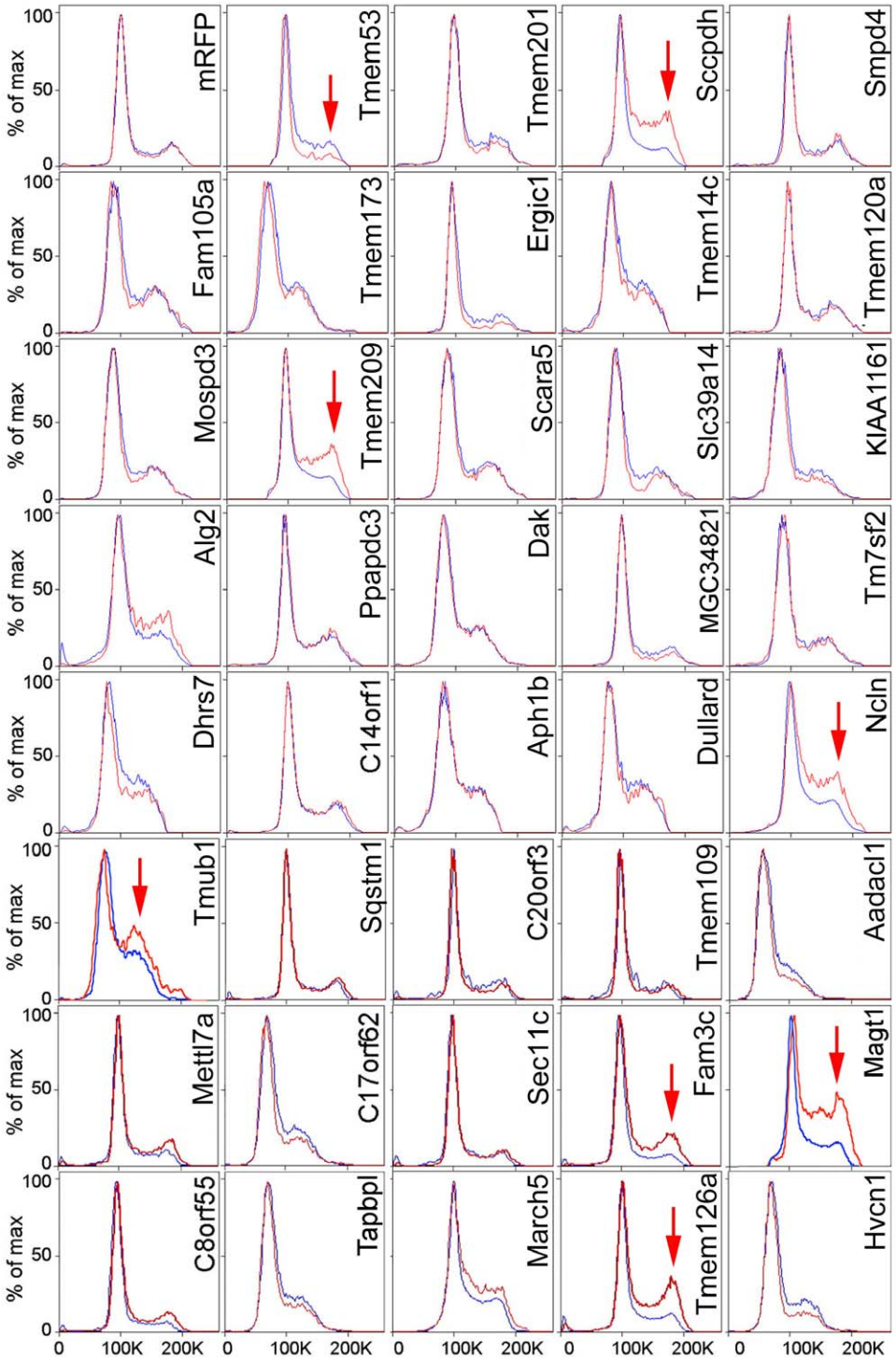


Figure 1. Changes in flow cytometry cell cycle profiles for cells overexpressing NETs. HEK293T cells expressing mRFP-NET fusions were recovered by trypsinization and analyzed by flow cytometry at 48 h post-transfection. Data were analyzed using FlowJo software and histogram overlays are displayed as %Max, scaling each curve to mode = 100%. The red line is the mRFP expressing cells in the population while the blue line is the untransfected cells in the population (the majority of cells were not transfected). The transfected and untransfected populations were both set on the scale to 100 for the 2N population so that increases or decreases in the 4N peak reveal changes in the cell distribution. Arrows indicate significant changes in cell cycle profile between transfected and non-transfected cells.
doi:10.1371/journal.pone.0018762.g001

blot. The p53, p38 and p21 antibodies used in subsequent assays were tested by immunocytochemistry, but proved inadequate for quantification due both to a diffuse distribution throughout the cell body and cell-to-cell variation in intensities that appears to result from induction of stress pathways in some cells during transfection (data not shown). Therefore, to further elucidate the pathways through which NET4/Tmem53 affects the cell cycle, its knockdown was attempted.

Table 1. Percentage of cells in each cell cycle phase by flow cytometry upon exogenous expression of NET-mRFP fusions.

NET	mRFP negative			mRFP positive		
	% G1	% S	% G2/M	% G1	% S	% G2/M
untransfected	60	14	21	-	-	-
mRFP	61	14	20	62	13	20
NET4/Tmem53	62	14	20	74	8	14
NET11/Sccpdh	60	14	20	37	18	38
NET31/Tmem209	58	15	21	39	15	38
NET51/C14orf1	62	16	18	60	17	19
NET59/Ncln	59	15	22	37	19	35
Fam3c	62	10	18	44	10	30
Magt1	65	12	19	53	14	26
Tmub1	63	13	19	50	14	31
Tmem126a	65	13	14	53	17	26

doi:10.1371/journal.pone.0018762.t001

There are both long and short splice variants of NET4/TMEM53 (Figure 5A). To determine whether it would be better to design siRNA oligos to knock down one or both (the original screen used the shorter variant), long and short splice variants of NET4/TMEM53 with the GFP moiety at either terminus were cloned and then tested to determine if both could produce the flow cytometry cell cycle/DNA content ratio effect of the original mRFP construct. All constructs yielded the same effect (Figure 5B). Multiple siRNA oligos were then generated that should knock down both long and short splice variants. Although the HEK293T cells were ideal for the flow cytometry-based screen due to their generally high transfection efficiencies and comparatively stable DNA content, for subsequent more directed cell cycle experiments MRC5 primary fibroblasts [21] and the U2OS osteosarcoma cell line were used because they have comparatively more operational checkpoint machinery. Moreover, the primary fibroblasts can enter a state of senescence that is not possible for HEK293T cells. Before proceeding, the ability of NET4/Tmem53-mRFP expression to alter flow cytometry DNA content profiles was evaluated in both MRC5 primary fibroblasts and the U2OS cells, confirming that a comparable effect to that observed in HEK293T cells occurred in the MRC5 and U2OS cells (Figure 5C).

Two siRNA oligos (si1 and si2) that should each in theory knock down both long and short splice variants (Figure 6A) effectively knocked down NET4/TMEM53 transcripts in MRC5 cells (Figure 6B). The si2 was slightly more effective than the si1 and therefore initially used in preference, though all relevant findings were subsequently verified with both oligos. As *NET4/TMEM53* is an uncharacterized gene with no direct evidence for its full range of splicing possibilities, it is possible that additional splice variants exist that the two siRNA oligos do not knock down. Therefore, an esiRNA that should knock down any and all possible splice variants of NET4/Tmem53 was also tested and found to be effective in reducing NET4/TMEM53 transcripts in MRC5 cells (Figure 6B). Knockdown of NET4/TMEM53 transcripts was also successful in U2OS cells (Figure 6C).

Available NET4/Tmem53 antibodies could not be used to determine if the NET4/Tmem53 protein was also being knocked down because they recognized several closely migrating bands where NET4/Tmem53 protein is expected to migrate on Western blot. Nonetheless, the endogenous protein is very likely to be reduced because cells expressing NET4/Tmem53-GFP prior to addition of siRNA oligos for NET4/TMEM53 knockdown exhibited a loss of the fusion protein with both GFP and NET4/Tmem53 antibodies (Figure 6D and data not shown). This indirect approach typically indicates knockdown of the endogenous protein [22].

The effect of NET4/TMEM53 knockdown on flow cytometry cell cycle/DNA content profiles was next determined. An increase was observed in the population of cells with a 2N amount of DNA that was similar to the increase resulting from exogenous NET4/Tmem53 expression (Tables 1 and 2). A corresponding reduction was observed in the population of cells with 4N DNA content both with the knockdown and with overexpression. Both MRC5 and U2OS cell lines were subsequently used in parallel for most experiments to ascertain if any observed effects required the more reliable cell cycle checkpoints of primary cells (MRC5) and thus were not recapitulated in transformed cell lines such as the U2OS

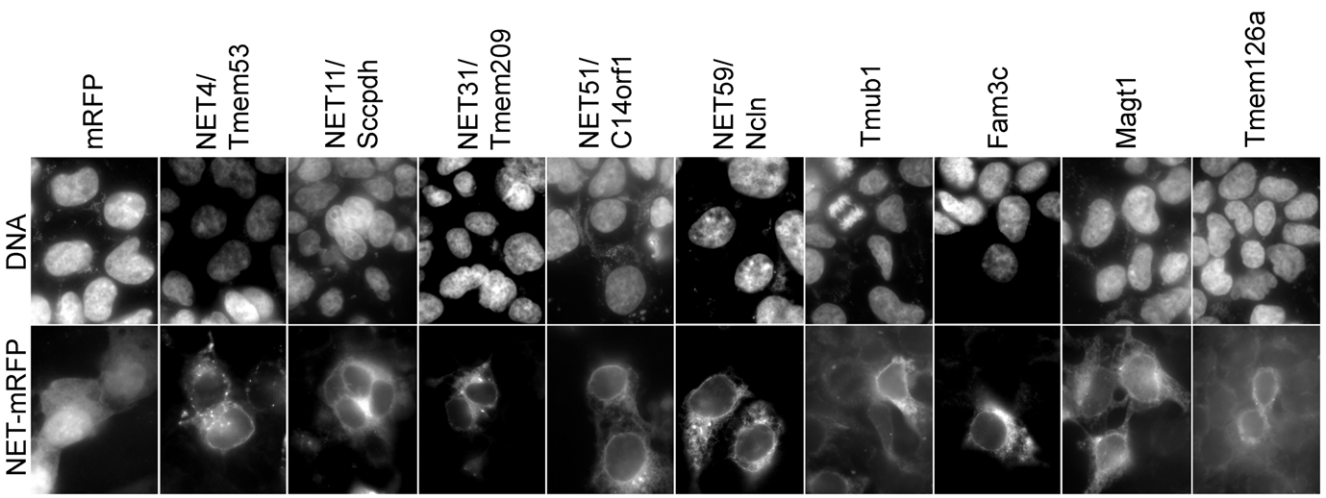


Figure 2. The changes in flow cytometry profile of cells expressing NETs are not caused by aberrant nuclear morphology of cells. Images of HEK293T cells expressing NET-mRFP shows nuclear morphology depicted by DAPI staining (upper panel) compared to mRFP signal (lower panel) indicating cells expressing NETs.
doi:10.1371/journal.pone.0018762.g002

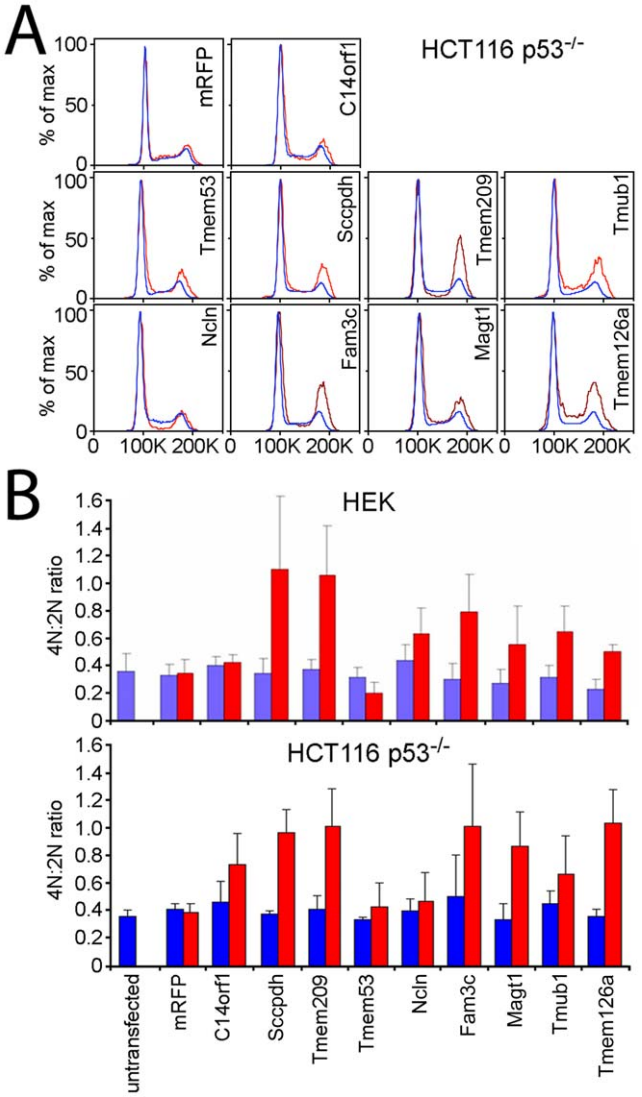


Figure 3. Cell cycle effects of NET4/Tmem53 and NET59/Ncln depend on p53. (A) Flow cytometry profiles of cells expressing NET-mRFP fusions in HCT116 p53−/− cells. Most NETs that had produced increases in the 4N population in Figure 1 yielded similar increases in the 4N population in HCT116 p53−/− cells; however, NET4/Tmem53 and NET59/Ncln lost their effects. (B) The percentage of cells in the 4N and 2N populations were calculated and 4N:2N ratios were plotted from at least three separate experiments with standard errors shown. The results for the HEK293T cells are shown above those for the HCT116 p53−/− cells.
doi:10.1371/journal.pone.0018762.g003

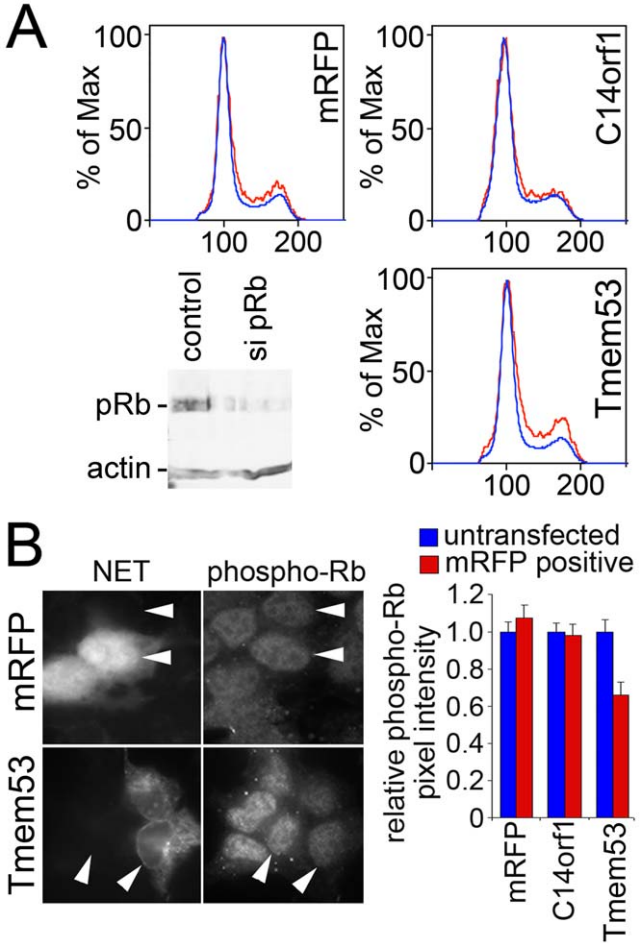


Figure 4. pRb is also involved in NET4/Tmem53 effects. (A) pRb was knocked down by siRNA in HEK293T cells subsequently transfected with NETs. Cell cycle profiles are shown for these cells expressing NET4/Tmem53 or controls of mRFP alone or NET51/C14orf1 that had no effect in the original flow cytometry screen. The 2N accumulation effect of NET4/Tmem53 on the cell cycle profile observed in HEK293T cells was lost when pRb levels were reduced. The knockdown of pRb is confirmed in the lower left corner. (B) Photomicrographs of cells overexpressing the RFP alone control or NET4/Tmem53. Cells were stained with an antibody that recognizes pRb phosphorylated on serine 780. NET/mRFP signal is shown in the left panels to identify transfected cells and phospho-pRb staining in the right panels. In each panel two adjacent cells are marked by arrowheads, one transfected and the other not transfected. All micrographs were taken at the same exposure time. The graph on the right shows quantification of the pixel intensity for the phospho-pRb staining. The pixel intensity for untransfected controls internal to each micrograph was set to 1 and average relative values from 40 transfected cells and 40 untransfected cells for each transfection are shown with standard errors.
doi:10.1371/journal.pone.0018762.g004

line that has stable p53- and pRb-dependent checkpoints but is defective for p16INK [23,24].

NET4/TMEM53 knockdown causes premature senescence in primary fibroblasts but only a cell cycle delay in the transformed U2OS cells

When MRC-5 cells were transfected with siRNA oligos or esiRNA for NET4/TMEM53 the number of cells per dish appeared to be lower than in cultures transfected with the scrambled control siRNA oligo, yet there did not seem to be an increase in apoptotic cells. In order to establish the cause of this difference, cell proliferation was assessed using antibodies to the nuclear protein Ki-67 that is present only in proliferating cells [25,26]. In cells knocked down for NET4/TMEM53 (Figure 7A, upper graph) there was a notable decrease in the frequency of Ki-67 positive cells from 57.6% in the control scrambled oligo transfected cells to 26.9% for the NET4/TMEM53 si2 transfected cells and 10.9% for the NET4/TMEM53 si1 transfected cells. The esiRNA also confirmed the phenotype (39.6% Ki-67 positive cells), further indicating its specificity to the NET4/Tmem53 knockdown.

The marked reduction in cell proliferation upon NET4/TMEM53 knockdown could indicate a permanent arrest or a temporary arrest from which cells could subsequently recover. ß-

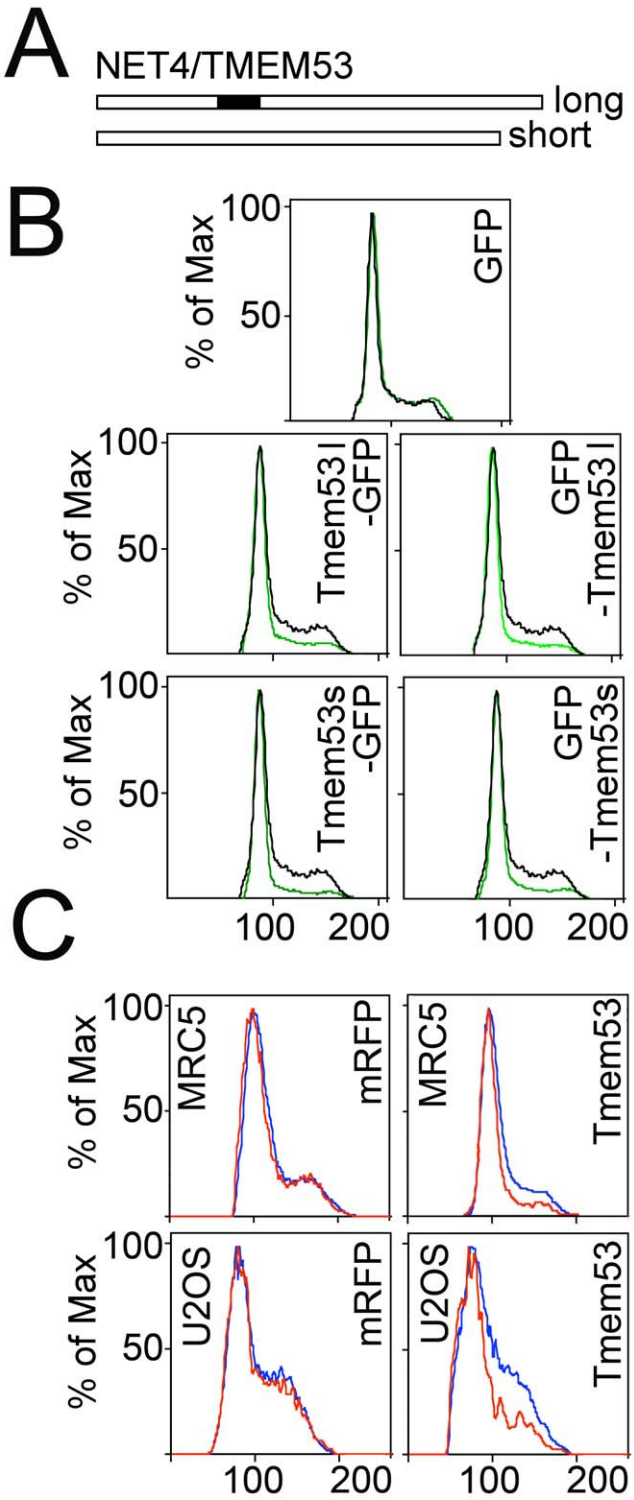


Figure 5. Splice variants of NET4/TMEM53 and the effect of knockdown in U2OS and MRC5 cells. (A) Two splice variants of NET4/TMEM53 were cloned that differed only by a 33 amino acid insertion roughly a third of the way into the protein. (B) Both long (l) and short (s) splice variants had the same effect on cell cycle/DNA content profiles and effects were independent of whether the GFP tag was on the amino terminus or carboxyl terminus. (C) Because siRNAs failed to knock down NET4/TMEM53 in HEK293T cells, the cell cycle/DNA content profile effect of the original NET4/TMEM53-mRFP construct was tested on U2OS and MRC5 cells as alternatives for knockdowns.
doi:10.1371/journal.pone.0018762.g005

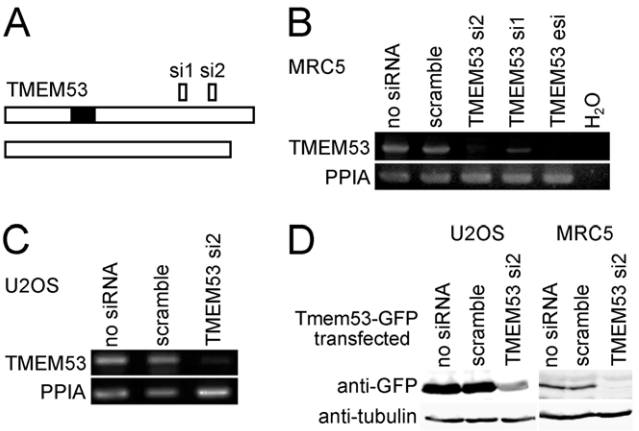


Figure 6. Knockdown of NET4/TMEM53 in U2OS and MRC5 cells. (A) Splice variants and position of knockdown oligos. (B) NET4/TMEM53 transcript levels were effectively knocked down in MRC5 cells using both siRNA oligos and an esiRNA. Peptidylprolyl isomerase A (PPIA) was used as a control for normalization. (C) NET4/TMEM53 transcript levels were also knocked down in the U2OS cell line using the siRNA oligo TMEM53 si2. The scramble siRNA oligo had no effect. PPIA was used as a control for normalization. (D) Western blot demonstrating the knockdown of NET4/Tmem53. Because antibodies to NET4/Tmem53 recognized multiple bands on Western blot at the expected molecular weight, knockdown of the protein was tested using a NET4/Tmem53-GFP fusion and GFP antibodies. The siRNAs were transfected 24 h after transfection of NET4/Tmem53-GFP in order to enable protein to be generated from the plasmid prior to the beginning of knockdown.
doi:10.1371/journal.pone.0018762.g006

galactosidase activity, a characteristic feature of senescent cells [27], was therefore assessed in MRC5 primary fibroblasts knocked down for NET4/TMEM53. At 72 h post transfection, the percentage of ß-galactosidase positive cells was ~3 times higher in cell cultures transfected with NET4/TMEM53 siRNA oligos or esiRNA as compared to untransfected cultures or those transfected with the control scramble oligo (Figure 7A, lower graph). This suggests a tendency of NET4/Tmem53 deficient MRC-5 cells to withdraw from the cell cycle due to premature senescence, which is a programmed cell response to many extra- and intracellular stresses with features similar to proliferative senescence and leading to permanent cell cycle exit.

By contrast, when NET4/TMEM53 siRNA oligo si2 was transfected into the transformed osteosarcoma U2OS cell line no decrease of Ki-67 expression or increase of ß-galactosidase was detected (data not shown), indicating that these cells do not fully exit the cell cycle. This difference is likely due to the absence in U2OS cells of p16INK, as this kinase has been reported to be important for establishment of senescence [23,24]. Confirming these reports, transcript levels of p16INK were undetectable in

Table 2. Percentage of cells in each cell cycle phase by flow cytometry for knockdown of NET4/TMEM53 and controls in U2OS cells.

	%G1	%S	%G2/M
untransfected	58	10	32
scramble	59	11	30
oligo si 2	65	12	22

doi:10.1371/journal.pone.0018762.t002

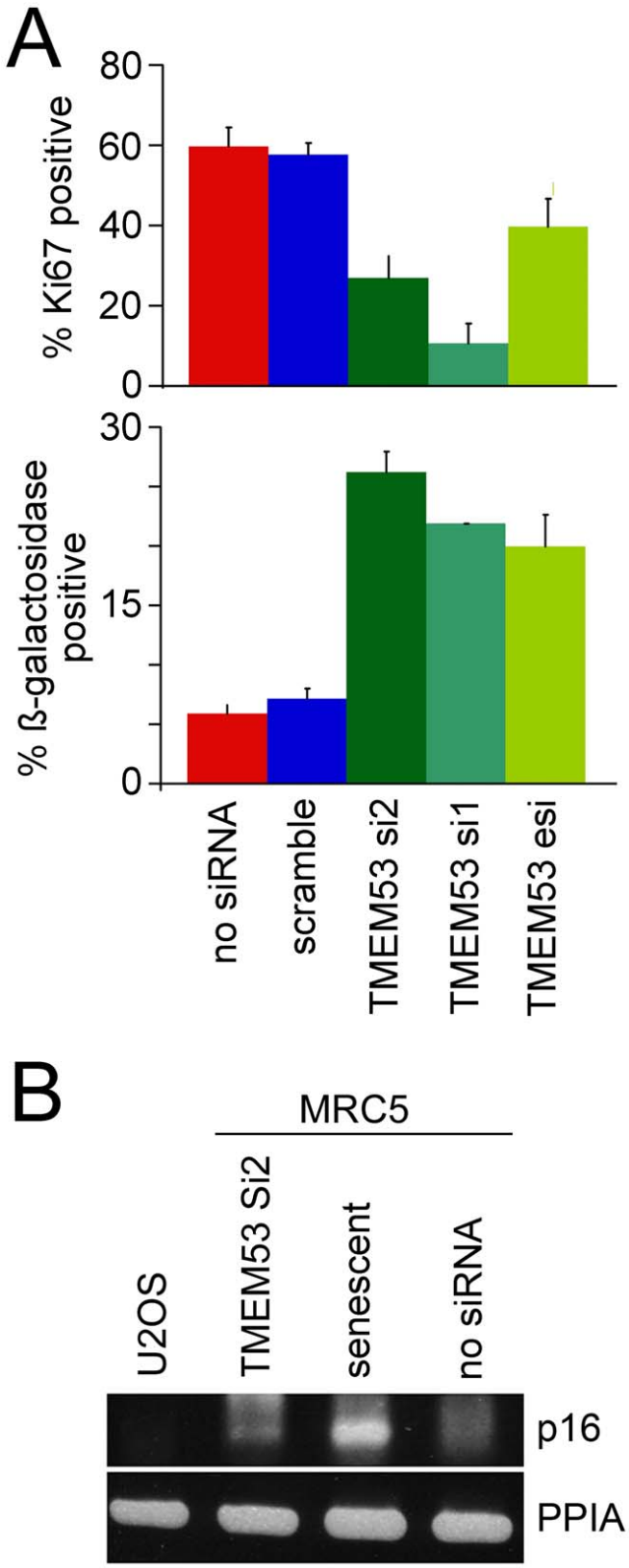


Figure 7. NET4/TMEM53 knockdown results in cell cycle withdrawal. (A) NET4/TMEM53 knockdown with both siRNAs and the esiRNA all resulted in a reduced number of actively proliferating cells as measured by Ki-67 staining (upper graph) and an increase in senescent cells as measured by ß-galactosidase staining (lower graph). The results from three separate experiments with standard errors are shown. (B) Differences between MRC5 and U2OS cells in NET4/Tmem53-induced senescence are likely due to the absence of p16INK in U2OS cells. RT-PCR was used to detect transcript levels of p16INK. In U2OS cells no transcripts were observed while in senescent MRC5 cells transcript levels were greatly increased compared to cycling untransfected MRC5 cells. Knockdown of NET4/Tmem53 also induced p16INK expression in MRC5 cells, albeit less so than by full senescence.
doi:10.1371/journal.pone.0018762.g007

U2OS cells and strongly upregulated in senescent MRC5 cells (Figure 7B). Knockdown of NET4/TMEM53 resulted in a small increase in p16INK transcript levels (Figure 7B), consistent with the hypothesis that its absence in U2OS cells underlies the failure to become senescent. Despite the inability to senesce, some effects on the cell cycle still occurred in U2OS cells knocked down for NET4/TMEM53 as BrdU incorporation was reduced by nearly 50% (data not shown), indicating that the frequency of cells in S phase was greatly reduced. This could indicate that these cells spend an extended period in G1, consistent with the flow cytometry data.

NET4/TMEM53 knockdown results in alteration of levels of several cell cycle regulators

To further understand the pathways mediating the effect of NET4/Tmem53 on the cell cycle, levels of various cell cycle regulators before and after NET4/Tmem53 depletion were quantified by Western blot. In both MRC5 and U2OS cells the p53 protein level was increased ~2-fold upon loss of NET4/Tmem53 while p21 levels increased ~7-fold (Figure 8A,B). Total levels of pRb remained relatively unchanged, but pRb became hypophosphorylated when NET4/Tmem53 was knocked down and phosphorylated p38MAP kinase levels increased (Figure 8A,B). It should be noted that although total pRb levels were unchanged on average over several experiments with NET4/TMEM53 knockdown, there was reasonable variability in the levels of total pRb among experiments. Thus the ratio of phosphorylated pRb to total pRb was likely also somewhat variable, but levels of the phosphorylated pRb were notably reduced in all experiments. To confirm that the effect observed was specific for NET4/TMEM53 knockdown the experiment was repeated using oligos si1, si2, and the esiRNA. The same result was observed with p53 and p21 increasing upon loss of NET4/Tmem53 (Figure 8C).

By contrast no changes were observed in levels for other NE proteins previously shown to contribute to cell cycle regulation Emerin, LAP2 (both soluble and transmembrane splice variants alpha and beta), or Lamin A (Figure 9A). Neither were differences observed in cyclins A, C, D, or E (Figure 9A,B). Cyclin B levels may have been slightly reduced; however, considerable variability was observed between all experiments and so no clear conclusion could be drawn (data not shown).

NET4/Tmem53 effects on cell cycle protein levels are mediated by the p38MAP kinase

The p38MAP kinase that was upregulated by NET4/TMEM53 knockdown mediates cell response to range of stresses, such as UV irradiation, osmotic shock, heat shock, starvation and cytokines [28]. p38 phosphorylates and stabilizes p53, which activates transcription of the cdk inhibitor p21 that subsequently blocks the cell cycle [17]. As this study has shown that p21 is also upregulated by NET4/TMEM53 knockdown, it seemed likely that the effects on cell cycle proteins p53, pRb and p21 are mediated by the p38MAP kinase.

To further test if the increases in p53 and p21 were dependent on p38MAP kinase, a specific p38 inhibitor SB203580 (Calbio-

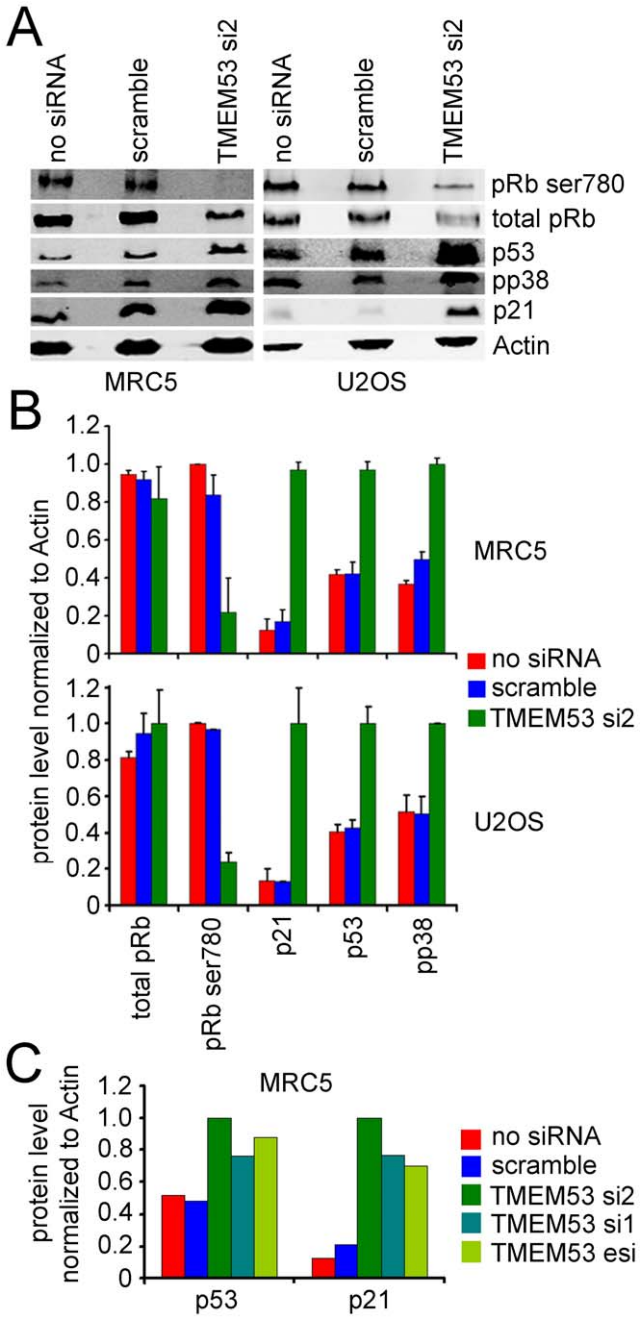


Figure 8. NET4/TMEM53 knockdown alters levels of several cell cycle proteins. Protein lysates were recovered from both MRC5 and U2OS cells that were either untransfected, transfected with a scrambled siRNA, or transfected with the NET4/TMEM53 si2 oligo. (A) Western blots reacted with antibodies to pRb phosphorylated on serine 780, total pRb, p53, phosphorylated p38MAP kinase, p21, and actin as a loading control. Note that some variation in total pRb levels was observed between experiments, causing the ratio of phosphorylated pRb to total pRb to be also somewhat variable. Because the phosphorylated form migrates slightly differently on SDS-PAGE the change in the spread of the band gives a greater appearance of such variability than is actually measured upon quantification. (B) Quantification of levels of each protein normalized to actin using a LI-COR Odyssey and fluorescent secondary antibodies. Averages from three experiments are shown with standard errors. (C) To ensure that results reflected effects of NET4/TMEM53 knockdown and not off-target effects both siRNA oligos and the esiRNA were tested for effects on p53 and p21 levels. Similar results were observed with all NET4/TMEM53 knockdowns.
doi:10.1371/journal.pone.0018762.g008

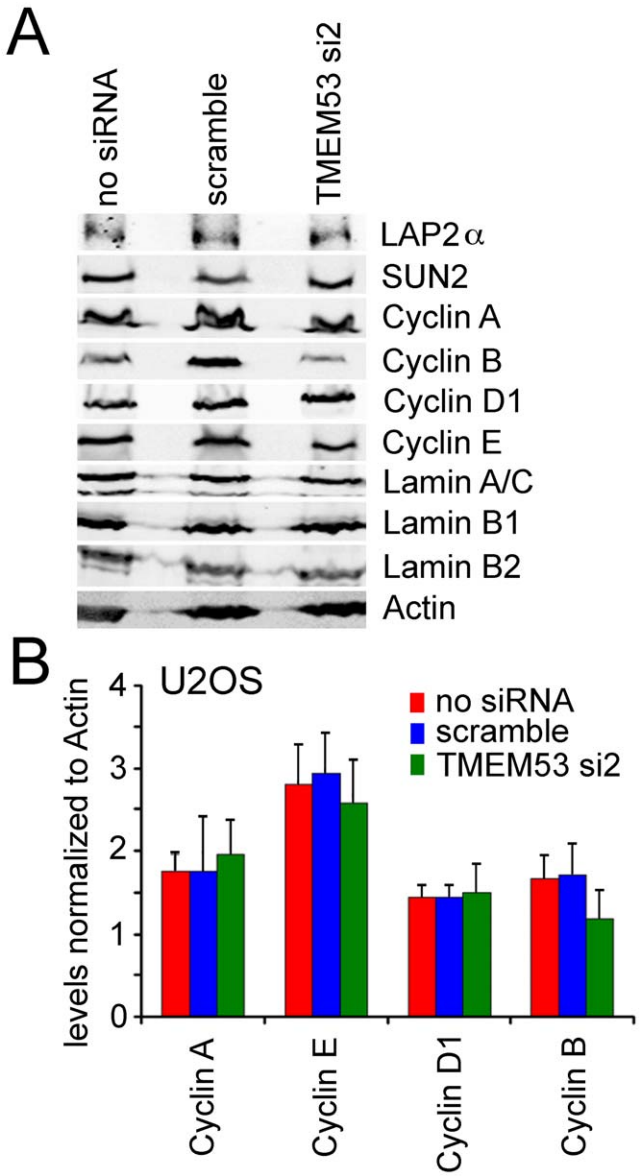


Figure 9. NET4/TMEM53 knockdown did not affect levels of other nuclear envelope proteins linked to the cell cycle or cyclins. (A) Protein lysates were recovered from U2OS cells treated as in Figure 8 and reacted on Western blots with antibodies to the various proteins. This experiment was repeated 3 times and a representative Western blot is shown. (B) Cyclin levels were quantified from three separate experiments analyzed by LI-COR using fluorescent secondary antibodies and are plotted normalized to the actin control. Standard errors are shown.
doi:10.1371/journal.pone.0018762.g009

chem) was added to the cultures 24 h after transfection with the siRNA oligos. The cells were harvested 48 h later and levels of p53, p21, and phosphorylated p38 measured by Western blot. In all cases the increases resulting from knockdown of NET4/TMEM53 were lost when p38MAP kinase was inhibited (Figure 10A,B). Moreover, the NET4/TMEM53 knockdown-mediated decrease in Ki-67 positive cells was partially abrogated and the increase in ß-galactosidase staining cells was completely abrogated (Figure 10C,D). Thus, p38MAP kinase appears to be activated by NET4/TMEM53 knockdown and mediates most of the NET4/Tmem53 effects on the cell cycle.

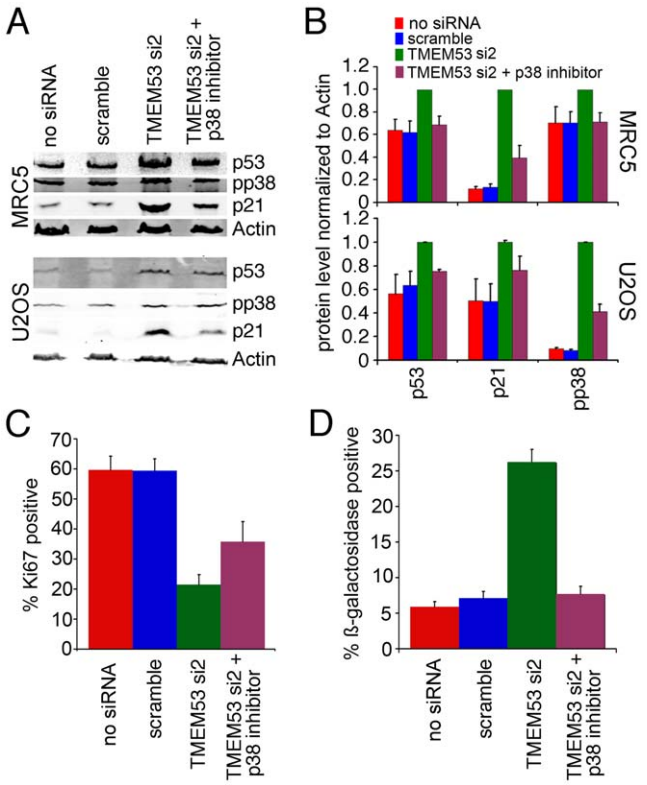


Figure 10. p38MAP kinase mediates the NET4/Tmem53 effects on cell cycle proteins. MRC5 and U2OS cells transfected with either siRNA oligos or controls were treated 24 h after transfection with a specific inhibitor of the p38MAP kinase. (A) Protein lysates were generated after a further 48 h and analyzed by Western blot for levels of p53, phosphorylated p38, and p21. This experiment was repeated 3 times and a representative Western blot is shown. (B) Levels of p53, p21 and phosphorylated p38 were quantified from three separate experiments analyzed by LI-COR using fluorescent secondary antibodies and are plotted normalized to the actin control with standard errors. In all cases the p38MAP kinase inhibitor blocked the effects of NET4/TMEM53 knockdown on levels of these cell cycle proteins. (C) The change in the percentage of Ki-67 positive cells induced by NET4/TMEM53 knockdown is partially abrogated by blocking the function of p38 kinase. (D) Effects of NET4/TMEM53 knockdown on senescence, as determined by ß-galactosidase staining, are abrogated by blocking p38MAP kinase function.
doi:10.1371/journal.pone.0018762.g010

Discussion

This flow cytometry-based screen identified several NETs with likely cell cycle functions. These NETs are generally uncharacterized proteins, many of which have conserved domains of unknown function. However, NET59/Ncln has separately been linked to a pathway that could intersect with cell cycle regulation: this NET interacts with Nomo1 that binds Smad proteins and thus intersects with TGFß signaling pathways [29]. The NET MAN1 has also been found to interact with Smad proteins [30,31] while the NET Emerin interacts with ß-catenin [32] and disruption of these two NETs yields cell cycle defects [6]. Thus, combined signaling/cell cycle effects appear to be a common trait among NETs and it is likely that further analysis of the other NETs identified in this screen will yield some similar intersections of nuclear envelope proteins with cell cycle pathways. That these 8 NETs represent roughly 20% of the proteins tested in the flow cytometry cell cycle/DNA content screen and that several recent studies have found NETs associated with centrosomes or the mitotic spindle [33,34,35] suggest that a great many nuclear envelope proteins will be found to contribute to cell cycle regulation.

Here, additionally, NET4/Tmem53 has been demonstrated to influence cell cycle regulation, affecting maintenance of the proliferative state of cells. Although the mechanism by which NET4/Tmem53 expression levels activate p38MAP kinase has thus far been elusive, it is clear that the effect of NET4/Tmem53 on proliferation is through p38MAP kinase-dependent upregulation of p53 and p21 with subsequent reduction in levels of phosphorylated pRb. Thus unlike Lamin A and LAP2α that exert effects directly on pRb to influence cell proliferation [11,12,13], NET4/Tmem53 acts upstream through stress signaling pathways.

NET4/Tmem53 was originally identified in a proteomic analysis of liver nuclear envelopes [14] and high-resolution microscopy on tagged exogenously expressed protein suggests that it resides in the outer nuclear membrane [16]. Antibodies generated to NET4/Tmem53 were previously shown to principally stain the nuclear periphery in liver tissue cryosections and nuclear envelope versus microsomal fractions isolated from rodent liver [16]; however, despite reactivity on tissue sections, these antibodies recognize several additional bands at the expected molecular weight of NET4/Tmem53 in cultured cells and so could not be exploited in the current study. The indication of NET4/Tmem53 as an outer nuclear membrane protein is not inconsistent with NE functions on the cell cycle. For example, Emerin, originally thought to reside only in the inner nuclear membrane [36], affects signaling pathways and has recently been shown to function also in the outer nuclear membrane and ER [37].

A potential functional relevance for NET4/Tmem53-induced cell cycle regulation is suggested by observations that NET4/TMEM53 is highly expressed in liver cells and adipocytes according to the BioGPS transcriptome database [38,39]. Recent studies have shown that the differentiation of mesenchymal stem cells (MSCs) is influenced by mechanical strain. In response to the application of physical strain MSCs are prevented from differentiating along an adipogenic lineage and instead are redirected towards an osteoblast lineage [40], This gives rise to the possibility the physical stress provides a homeostatic mechanism for mesenchymal lineage selection. Since it has already been shown the NE plays a key role in propagating stress signaling and sensing [41], it is important to understand which NE proteins interact with stress signaling pathways and how they might influence stem cell self renewal and differentiation. We believe that our data suggests that NET4/Tmem53 has such a role. Future studies will focus on putative regulation of adipogenic stem cells by NET4 and its possible involvement in Laminopathies such as Dunnigan-type familial partial lipodystrophy (FPLD; [42,43]) or mandibuloacral dysplasia (MADA; [44]).

Methods

Plasmid construction

NET expressing plasmids were generated from IMAGE clones as previously described [14,15,16]. All were under regulation of the CMV promoter with their carboxyl-termini fused to monomeric red fluorescent protein (mRFP).

Cell culture and transient transfection

HEK293T (from ATCC; [45]), U2OS (from ATCC; [24]), MRC5 (from the EUACC; [21]) and HCT116 p53$^{-/-}$ (kind gift of B. Vogelstein, Johns Hopkins University; [46]) cells were maintained in high glucose DMEM (Invitrogen) supplemented

with 10% fetal calf serum, 100 units/ml penicillin and 100 µg/ml streptomycin sulfate (Invitrogen). For U2OS siRNA assays cells were seeded at 30% confluency and the next day transfected using Dharmafect transfection reagent 1 according to the manufacturer recommendations (Dharmacon RNA Technologies). MRC5 cells were trypsinized at ~80% confluency and 3×10^6 cells were aliquoted for nucleofection using Kit R, Program V-020 (Lonza). The cells were transfected with either 8 µg of siRNA oligo or 12 µg of esiRNA (esiRNA human TMEM53, Sigma) and plated to 10 cm dishes with several coverslips. At 48 or 72 h coverslips were removed and fixed for immunocytochemistry and the remaining cells on the plate were harvested by trypsinization and lysed in sample buffer for SDS-PAGE and Western blotting. Alternatively they were lysed in Trizol reagent (Sigma) for isolation of RNA. HEK293T cells were either nucleofected using Kit V Program Q-001 (Lonza) or were transfected using Fugene 6 or HD (Roche). For microscopy, cells were generally plated on coverslips in 24 well dishes after nucleofection or were first plated and then transfected using Fugene. In this case, plating was at ~10% confluency so that cells did not reach confluency before 30 h post-transfection when they were fixed. Then, roughly 12 h after plating, DNA was transfected using Fugene. For p38MAP kinase inhibitor experiments the specific inhibitor SB203580 (Calbiochem) was added to cultures at 10 µM concentration 12 h after transfection with siRNA oligos.

Cell cycle assay

Plasmids encoding different NETs fused to mRFP were transfected into HEK293T cells using Fugene 6 transfection reagent (Roche). At 48 h post-transfection, the DNA stain Hoechst 33342 (Molecular Probes) was added to the cells at a final concentration of 5 µg/ml and left to incubate at 37°C for a period of 30 min to 60 min. Cells were harvested by trypsinization, trypsin was inactivated with serum and cell pellets were collected by centrifugation at 250×g for 5 min at RT, washed once in PBS and resuspended in 1 ml of PBS. Cells were immediately analyzed on an LSR II flow cytometer (BD Bioscience, UK) equipped with 488 nm and 350 nm lasers and appropriate filters. Cells with fragmented DNA that might be undergoing necrosis or apoptosis and cell aggregates were excluded from analysis by application of electronic gates. Cell cycle analysis was carried out on the live singlets gate using *FlowJo* software (TreeStar, Inc). At least 10,000 cells were scored for the total live singlets and 1,000 cells for the mRFP positive live singlets. Each experiment was repeated at least 2 times for negative results and 3 times for those with a positive effect on the cell cycle. For further confirmation NETs that showed an effect on the cell cycle in 3 independent experiments (NET4/Tmem53, NET11/Sccpdh, NET31/Tmem209, NET59/Ncln, Tmub1, Fam3c, Magt1 and Tmem126a) were repeated a 4th time where at least 20,000 mRFP positive intact singlets were counted. Data are displayed in the form of histogram overlays using %Max option, which scales each population curve to mode = 100%.

Antibodies

Antibodies to the following proteins were used: Ki-67 (610968, BD Transduction Lab), total Rb (4H1 9309, Cell signaling), phospho-Rb (9307, Cell Signaling), p21 (556430, BD Transduction lab), p53 rabbit (9282, Cell signaling), p53 mouse (NCL-p53-DO1, Leica), p38 total (9212, Cell Signaling), active p38 (V3281 Anti-active MAPK Family Sampler, Promega), cyclin E mAb clone HE12 (32-1600, Invitrogen), cyclin A mAb clone Cy-A1 (4710, Sigma), cyclin D (2922, Cell Signaling), cyclin B (SC245, Santa Cruz), LAP2ß (06-1002, Millipore), Lap2α previously described in [47], Lamin A and B1 (3262 and 3931) previously described in [48]. All fluorophore-conjugated secondary antibodies used for immunofluorescence were minimally cross-reactive from donkey (Jackson ImmunoResearch) or goat (Molecular Probes). For Western blotting IR800 conjugated goat anti-rabbit antibodies (LI-COR Biosciences) were used.

Immunofluorescence microscopy

Cells transfected with NETs were fixed for 7 min in 3.7% formaldehyde, permeabilized for 6 min in 0.1% Triton X-100, blocked with 3% BSA in PBS, and reacted for 40 min at RT with antibodies to Ki-67 or phospho-Rb. After washing, appropriate fluorophore-conjugated secondary antibodies were added for 30 min at RT and washed. Cells were also stained with Hoechst 33342 (Molecular Probes) to visualize nuclei and mounted in fluoromount G (EM Sciences). For ß-galactosidase assays, histochemical staining at pH 6.0 was performed as described in [27].

Images were obtained using a Nikon TE-2000 microscope equipped with a 1.45 NA 100× objective, Sedat quad filter set, PIFOC Z-axis focus drive (Physik Instruments), and CoolSnapHQ High Speed Monochrome CCD camera (Photometrics) run by IPLab image acquisition software. Micrographs were saved from source programs as .tif files and prepared for figures using Photoshop 8.0.

Quantification of grey levels for NETs was performed using .tif images from IPLab imported into Image Pro Plus. Cells fixed and stained with phosho-Rb antibodies were imaged using identical settings. Average grey scale values (pixel) intensity were measured for at least 40 mRFP positive cells and as many untransfected cells on the same coverslips.

Quantitative Western blotting

Cells were scraped and lysed in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 2 mM $MgCl_2$, 0.2% NP-40 in the presence of protease inhibitor cocktail (Roche 11 873 580 001) by heating at 65°C for 2 min and sonication in a sonibath at 4°C. Protein concentrations were determined using the Bradford Method (BioRad). An equal volume of protein sample buffer (100 mM Tris pH 6.8, 4 M Urea, 2% SDS, 50 mM DTT and 15% sucrose) was added and the samples were boiled at 100°C for 5 min then sonicated in a sonibath with high frequency for 10 min with 30 sec interval on/off. Equal amounts of protein were resolved by SDS-PAGE and transferred to Nitrocellulose membrane (LI-COR Biosciences). Membranes were blocked in PBS, 5% milk, 0.2% tween-20. Primary antibodies were diluted in this buffer and allowed to incubate overnight at 4°C. Secondary antibodies IR800 conjugated goat anti-rabbit (LI-COR Biosciences) were added at concentration 1/5000 at RT for 2 h. Visualization and quantification were performed using a LI-COR Odyssey and software (Odyssey 3.0.16) using median background subtraction. A minimum of three independent blots was run for each NET and control. The averages from all three are presented in figures with standard error shown.

siRNA oligos

NET4/TMEM53 oligo si1: 5′-AGAAGUGGGUGUGGAAGAGGG-3′. NET4/TMEM53 oligo si2: 5′- UAGUAAGUAGGGUAGUCACGG-3′. Scramble control oligo 5′-UCGAAGUAUUCCGCGUACG-3′. Sigma NET4/TMEM53 esiRNA EHU027971.

pRb siRNA oligo: GCCCUUACAAGUUUCCUAG [49].

RT-PCR analysis

Cells were lysed on tissue culture plates with Tri-Reagent (Sigma), and total RNA was extracted according to the manufacturers instructions. RT-PCR reactions were carried out with 100 ng of total RNA using the Titan one tube RT-PCR system (Roche) in accordance with the manufacturer's instructions, except that the dNTP concentration was increased to 500 mM and $MgCl_2$ increased to 3 mM. Typical reaction conditions were 30 min reverse transcription at 50°C, 2 min denaturation at 94°C, then 24 cycles of 94°C for 30 s, 60°C for 30 s and 68°C for 45 s. Peptidylprolyl isomerase A (PPIA) was used as a loading control and reactions were repeated at least three times.

The human primer sets used were: NET4/TMEM53 5′-AAGCTGCTCGAGCTGCTC3-3′ and 5′-CAGAGGCTTGTGTAGTAA-3′; PPIA 5′-CACCGTGTTCTTCGACATTG-3′ and 5′-TCGAGTTGTCCACAGTCAGC-3′.

Acknowledgments

We would like to thank Ewa Markiewicz for advice, David Kelly for assistance with image quantification, Alastair Kerr for statistical analysis, and Bert Vogelstein for HCT116 $p53^{-/-}$ cells.

Author Contributions

Conceived and designed the experiments: NK VS CJH VP ECS. Performed the experiments: NK VS MW DGB. Analyzed the data: NK VS CJH VP MW. Contributed reagents/materials/analysis tools: CJH VP MW. Wrote the paper: ECS.

References

1. Mounkes L, Stewart CL (2004) Structural organization and functions of the nucleus in development, aging, and disease. Curr Top Dev Biol 61: 191–228.
2. Worman HJ, Bonne G (2007) "Laminopathies": a wide spectrum of human diseases. Exp Cell Res 313: 2121–2133.
3. Bridger JM, Foeger N, Kill IR, Herrmann H (2007) The nuclear lamina. Both a structural framework and a platform for genome organization. FEBS J 274: 1354–1361.
4. Dorner D, Gotzmann J, Foisner R (2007) Nucleoplasmic lamins and their interaction partners, LAP2alpha, Rb, and BAF, in transcriptional regulation. FEBS J 274: 1362–1373.
5. Fairley E, Riddell A, Ellis J, Kendrick-Jones J (2002) The cell cycle dependent mislocalisation of emerin may contribute to the Emery-Dreifuss muscular dystrophy phenotype. J Cell Sci 115: 341–354.
6. Liu J, Lee K, Segura-Totten M, Neufeld E, Wilson K, et al. (2003) MAN1 and emerin have overlapping function(s) essential for chromosome segregation and cell division in Caenorhabditis elegans. Proc Natl Acad Sci U S A 100: 4598–4603.
7. Melcon G, Kozlov S, Cutler DA, Sullivan T, Hernandez L, et al. (2006) Loss of emerin at the nuclear envelope disrupts the Rb1/E2F and MyoD pathways during muscle regeneration. Hum Mol Genet 15: 637–651.
8. Bakay M, Wang Z, Melcon G, Schiltz L, Xuan J, et al. (2006) Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. Brain 129: 996–1013.
9. Genovese C, Trani D, Caputi M, Claudio PP (2006) Cell cycle control and beyond: emerging roles for the retinoblastoma gene family. Oncogene 25: 5201–5209.
10. Ozaki T, Saijo M, Murakami K, Enomoto H, Taya Y, et al. (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.
11. Markiewicz E, Dechat T, Foisner R, Quinlan R, Hutchison C (2002) Lamin A/C binding protein LAP2alpha is required for nuclear anchorage of retinoblastoma protein. Mol Biol Cell 13: 4401–4413.
12. Johnson BR, Nitta RT, Frock RL, Mounkes L, Barbie DA, et al. (2004) A-type lamins regulate retinoblastoma protein function by promoting subnuclear localization and preventing proteasomal degradation. Proc Natl Acad Sci U S A 101: 9677–9682.
13. Pekovic V, Harborth J, Broers JL, Ramaekers FC, van Engelen B, et al. (2007) Nucleoplasmic LAP2alpha-lamin A complexes are required to maintain a proliferative state in human fibroblasts. J Cell Biol 176: 163–172.
14. 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.
15. Korfali N, Wilkie GS, Swanson SK, Srsen V, Batrakou DG, et al. (2010) The leukocyte nuclear envelope proteome varies with cell activation and contains novel transmembrane proteins that affect genome architecture. Mol Cell Proteomics 9: 2571–2585.
16. Malik P, Korfali N, Srsen V, Lazou V, Batrakou DG, et al. (2010) Cell-specific and lamin-dependent targeting of novel transmembrane proteins in the nuclear envelope. Cell Mol Life Sci 67: 1353–1369.
17. Agarwal ML, Taylor WR, Chernov MV, Chernova OB, Stark GR (1998) The p53 network. J Biol Chem 273: 1–4.
18. Chen PL, Scully P, Shew JY, Wang JY, Lee WH (1989) Phosphorylation of the retinoblastoma gene product is modulated during the cell cycle and cellular differentiation. Cell 58: 1193–1198.
19. DeCaprio JA, Furukawa Y, Ajchenbaum F, Griffin JD, Livingston DM (1992) The retinoblastoma-susceptibility gene product becomes phosphorylated in multiple stages during cell cycle entry and progression. Proc Natl Acad Sci U S A 89: 1795–1798.
20. Mihara K, Cao XR, Yen A, Chandler S, Driscoll B, et al. (1989) Cell cycle-dependent regulation of phosphorylation of the human retinoblastoma gene product. Science 246: 1300–1303.
21. Jacobs JP, Jones CM, Baille JP (1970) Characteristics of a human diploid cell designated MRC-5. Nature 227: 168–170.
22. Wu W, Hodges E, Redelius J, Hoog C (2004) A novel approach for evaluating the efficiency of siRNAs on protein levels in cultured cells. Nucleic Acids Res 32: e17.
23. Grossel MJ, Baker GL, Hinds PW (1999) cdk6 can shorten G(1) phase dependent upon the N-terminal INK4 interaction domain. J Biol Chem 274: 29960–29967.
24. Ponten J, Saksela E (1967) Two established in vitro cell lines from human mesenchymal tumours. Int J Cancer 2: 434–447.
25. Gerdes J, Lemke H, Baisch H, Wacker HH, Schwab U, et al. (1984) Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. J Immunol 133: 1710–1715.
26. Scholzen T, Gerdes J (2000) The Ki-67 protein: from the known and the unknown. J Cell Physiol 182: 311–322.
27. Dimri GP, Lee X, Basile G, Acosta M, Scott G, et al. (1995) A biomarker that identifies senescent human cells in culture and in aging skin in vivo. Proc Natl Acad Sci U S A 92: 9363–9367.
28. Zarubin T, Han J (2005) Activation and signaling of the p38 MAP kinase pathway. Cell Res 15: 11–18.
29. Haffner C, Dettmer U, Weiler T, Haass C (2007) The Nicastrin-like protein Nicalin regulates assembly and stability of the Nicalin-nodal modulator (NOMO) membrane protein complex. J Biol Chem 282: 10632–10638.
30. Osada S, Ohmori SY, Taira M (2003) XMAN1, an inner nuclear membrane protein, antagonizes BMP signaling by interacting with Smad1 in Xenopus embryos. Development 130: 1783–1794.
31. Pan D, Estevez-Salmeron LD, Stroschein SL, Zhu X, He J, et al. (2005) The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-{beta} superfamily of cytokines. J Biol Chem 280: 15992–16001.
32. Markiewicz E, Tilgner K, Barker N, van de Wetering M, Clevers H, et al. (2006) The inner nuclear membrane protein emerin regulates beta-catenin activity by restricting its accumulation in the nucleus. EMBO J 25: 3275–3285.
33. Buch C, Lindberg R, Figueroa R, Gudise S, Onischenko E, et al. (2009) An integral protein of the inner nuclear membrane localizes to the mitotic spindle in mammalian cells. J Cell Sci 122: 2100–2107.
34. King MC, Drivas TG, Blobel G (2008) A network of nuclear envelope membrane proteins linking centromeres to microtubules. Cell 134: 427–438.
35. Wilkie GS, Korfali N, Swanson SK, Malik P, Srsen V, et al. (2011) Several novel nuclear envelope transmembrane proteins identified in muscle have cytoskeletal associations. Mol Cell Proteomics 10: M110.003129.
36. Manilal S, Nguyen TM, Sewry CA, Morris GE (1996) The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. Hum Mol Genet 5: 801–808.
37. Salpingidou G, Smertenko A, Hausmanowa-Petrucewicz I, Hussey PJ, Hutchison CJ (2007) A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane. J Cell Biol 178: 897–904.
38. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, et al. (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 99: 4465–4470.
39. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, et al. (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol 10: R130.
40. Sen B, Xie Z, Case N, Ma M, Rubin C, et al. (2008) Mechanical strain inhibits adipogenesis in mesenchymal stem cells by stimulating a durable beta-catenin signal. Endocrinology 149: 6065–6075.
41. Lammerding J, Schulze P, Takahashi T, Kozlov S, Sullivan T, et al. (2004) Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. J Clin Invest 113: 370–378.

42. Cao H, Hegele RA (2000) Nuclear lamin A/C R482Q mutation in Canadian kindreds with Dunnigan- type familial partial lipodystrophy. Hum Mol Genet 9: 109–112.
43. Shackleton S, Lloyd DJ, Jackson SN, Evans R, Niermeijer MF, et al. (2000) LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. Nat Genet 24: 153–156.
44. Novelli G, Muchir A, Sangiuolo F, Helbling-Leclerc A, D'Apice M, et al. (2002) Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding lamin A/C. Am J Hum Genet 71: 426–431.
45. DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, et al. (1987) Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol Cell Biol 7: 379–387.
46. Bunz F, Hwang PM, Torrance C, Waldman T, Zhang Y, et al. (1999) Disruption of p53 in human cancer cells alters the responses to therapeutic agents. J Clin Invest 104: 263–269.
47. Dechat T, Gotzmann J, Stockinger A, Harris CA, Talle MA, et al. (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.
48. Schirmer EC, Guan T, Gerace L (2001) Involvement of the lamin rod domain in heterotypic lamin interactions important for nuclear organization. J Cell Biol 153: 479–489.
49. Srsen V, Gnadt N, Dammermann A, Merdes A (2006) Inhibition of centrosome protein assembly leads to p53-dependent exit from the cell cycle. J Cell Biol 174: 625–630.