Supporting information to:
Cell type-specific nuclear pores: A case in point for context-dependent stoichiometry of molecular machines

Alessandro Ori, Niccolò Banterle, Murat Iskar, Amparo Andrés-Pons, Claudia Escher, Huy Khanh Bui, Lenore Sparks, Victor Solis-Mezarino, Oliver Rinner, Peer Bork, Edward A Lemke and Martin Beck

Supporting information:

Figures S1-S6

Supplementary Materials and Methods

Supplementary References

The following supplementary material is provided separately:

Table S1, related to Figure 1 and 3, contains a summary of the stoichiometry of the human NPC (sheet 1), the full dataset of peptide measurements used for absolute quantification (sheet 2), and the MRM assays used for absolute (sheet 3) and relative (sheet 4) quantification of Nups.

Table S2, related to Figure 5, contains the manually curated nuclear protein complexes definitions and their references (sheet 1), a summary table for the quantified complexes (sheet 2) and the full statistics for comple-wise protein expression analysis (sheet 3).

Movie S1, shows a representative example of photo-conversion of mEos2-Nup107 in a purified nucleus. Photo converted single mEos2 appears as diffraction limited Gaussian spots lasting several frames before bleaching. The exposure time was 0.12 s/frame but is displayed at 0.04 s/frame for illustrative purposes. The movie is limited to 1,000 frames while the original data acquisition lasted 10,000 frames. Source data for this movie is available on the online supplementary information page (this file can be viewed using ImageJ software, freely available at http://rsbweb.nih.gov/ij/download.html).

Supplementary Figures:

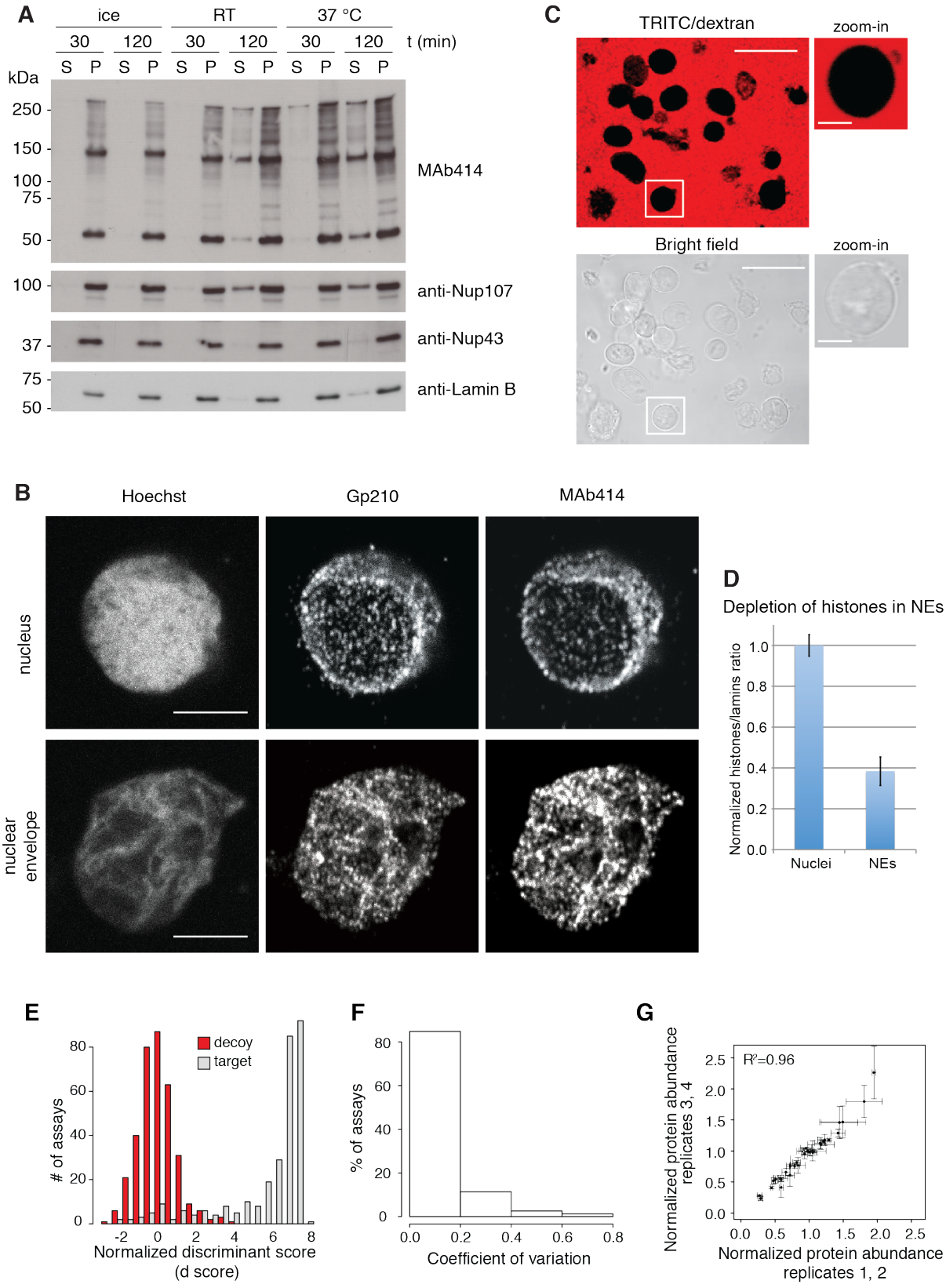

Figure S1, related to Figure 1. Quality assessment of nuclear envelope (NE) preparations and targeted proteomics measurements. (A) Western-blot analysis shows that major NPC components remain assembled under experimental conditions relevant for NEs isolation. No significant release of FG-containing (probed with MAb414) or scaffold Nups (Nup107 and Nup43) was observed for ~120 min on ice in the supernatant (S) after spinning nuclei (P). In contrast, leakage of Nups was observed when nuclei were incubated at higher temperatures. (B, C) NPCs remain intact in nuclear envelopes isolated from HeLa cells. The characteristic punctuate NPC immunofluorescence staining is observed in nuclei and NEs using anti-Gp210 and the MAb414 antibodies (B, scale bar: 5 μm). Although the transmembrane Nup Gp210 was shown to have a relatively short mean residence time of ~5 min at the NPC in live rat cells (Rabut et al, 2004), it remained stably integrated to nuclear pores during NE preparations from HeLa cells. (C) The permeability barrier of NPCs in isolated NEs was preserved as probed by their ability to exclude fluorescently labeled high molecular weight dextran (scale bars: 25 μm for overview; 5 μm for zoom-in). (D) Two thirds of the nuclear chromatin was removed in the NE preparations. The depletion of chromatin-associated proteins was quantified by comparing the ratio between the summed intensities of lamins and histones obtained from shotgun proteomics of nuclei and NEs. The displayed ratios correspond to mean ± standard error of the mean, obtained from three biological replicates. (E-G) Performance of targeted proteomic measurements. (E) The score distribution of the combined targeted proteomics measurements for absolute quantification obtained from four biological replicates were statistically validated using mProphet (Reiter et al), filtered with a normalized discrimination score > 4, corresponding to a false discovery rate (FDR) < 2.5×10^{-5}, and show an excellent discrimination of target from decoy transitions. (F) More than 80% of the selected multiple reactions monitoring (MRM) assays displayed a coefficient of variation < 0.2 across the four biological replicates, underlining the accuracy and reproducibility of the absolute quantification. (G) Absolute abundance across biological replicates is highly correlated. We assessed the reproducibility of the measurements at the protein level by arbitrarily splitting the four replicates in two groups and calculating the protein abundances independently for the two dataset. Nup abundances were calculated as the median of all the assays ± median absolute deviation and normalized to the Nup107-subcomplex as in Figure 1.

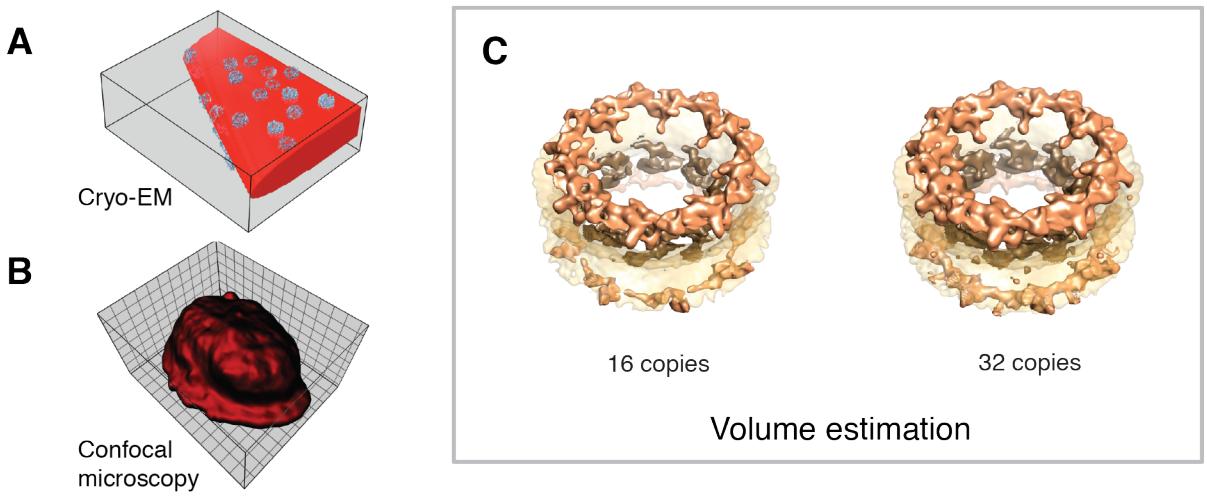

Figure S2. Estimation of the number of NPC per nucleus by cryo electron- and fluorescence microscopy. (A) Representative cryo electron tomogram of a HeLa NE. NPCs are displayed in light blue and the fitted surface used to estimate NPC density is displayed in red. (B) Representative image of HeLa nucleus used to estimate nuclear surface area. The surface (dark red) was obtained by fitting the confocal z-stacks of nuclei stained with a fluorescent membrane dye. (C) Volumetric matching of the mass of all Nups into the cryo-EM map of the human NPC different absolute stoichiometries underlines the plausibility of a total molecular weight of 110 MDa of the human NPC based on the assumption that 32 copies of the Nup107 subcomplex occur per NPC. Isosurface representations of the cryo-EM map of the human NPC (Maimon et al, 2012) are shown at different molecular weights that were calculated based on 16 copies of the majority of all scaffold Nups per NPC and alternatively for 32 copies. Transmembrane and FG-containing Nups were omitted in the calculation because they are likely averaged out in the cryo-EM map. The thresholds for the isosurface representations were set in a way that the enclosed volume matches the calculated molecular weight. While the electron optical density enclosed by the isosurface largely skeletonizes in case of 16 copies, it looks more realistic in case of 32 copies.

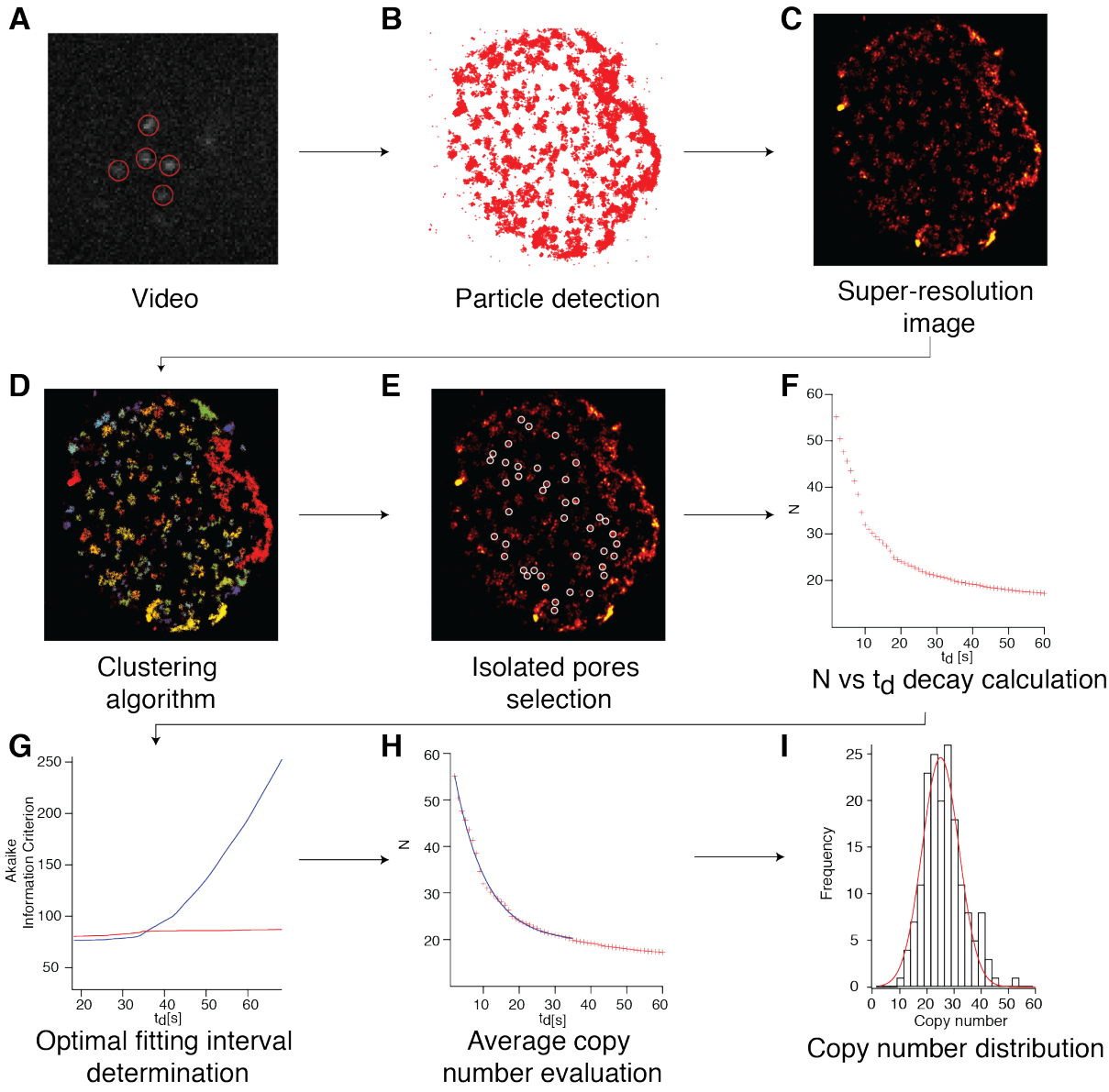

Figure S3, related to Figure 2. The workflow for the measurement of Nup107 copies per NPC by super-resolution microscopy is shown in (A-I, referred to by Supplementary Materials and Methods).

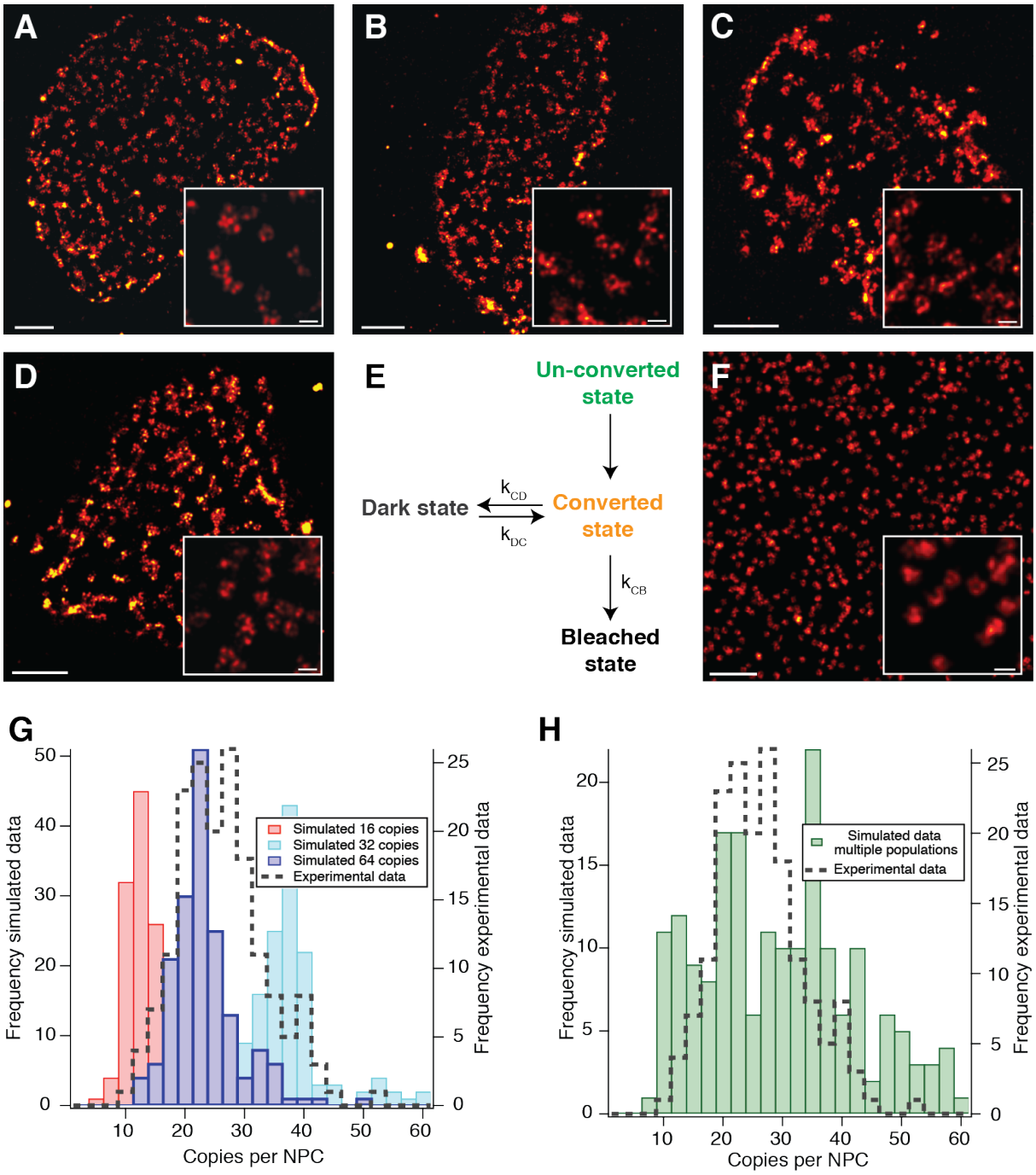

Figure S4, related to Figure 2. Experimental super-resolution images of mEos2-Nup107 and synthetic dataset generation and analysis. (A-D) Super-resolution images of nuclei isolated from transfected HEK293 cells used for the measurement of Nup107 copies per NPC. (E) Schematic representation of the theoretical states of the fluorophore and constants used for Monte-Carlo simulations. (F) Simulated super-resolution image of NPCs obtained assuming 32 fluorophores per pore and a labeling efficiency of 80%. (G and H) Distribution of the estimated copies of mEos2-Nup107 per NPC obtained from super-resolution based counting for simulated data assuming 16 (red), 32 (blue) or 64 (cyan) copies of the fluorophore (G), or a mixed dataset generated assuming an equal distribution of the three values (H). For comparison, both distributions are overlaid with the experimental data (dashed dark line) obtained for mEos2-Nup107, as shown in Figure 2C. In (G) and (H) copies

per NPC are shown after correction for the labeling efficiency but without correction for the maturation efficiency, which was set to 80% in the simulations but is unknown for the experimental data. The very bright spots seen in images (A-D) are fiducial markers used for drift correction. Scale bars: 1 μm and 150 nm for zoomed images.

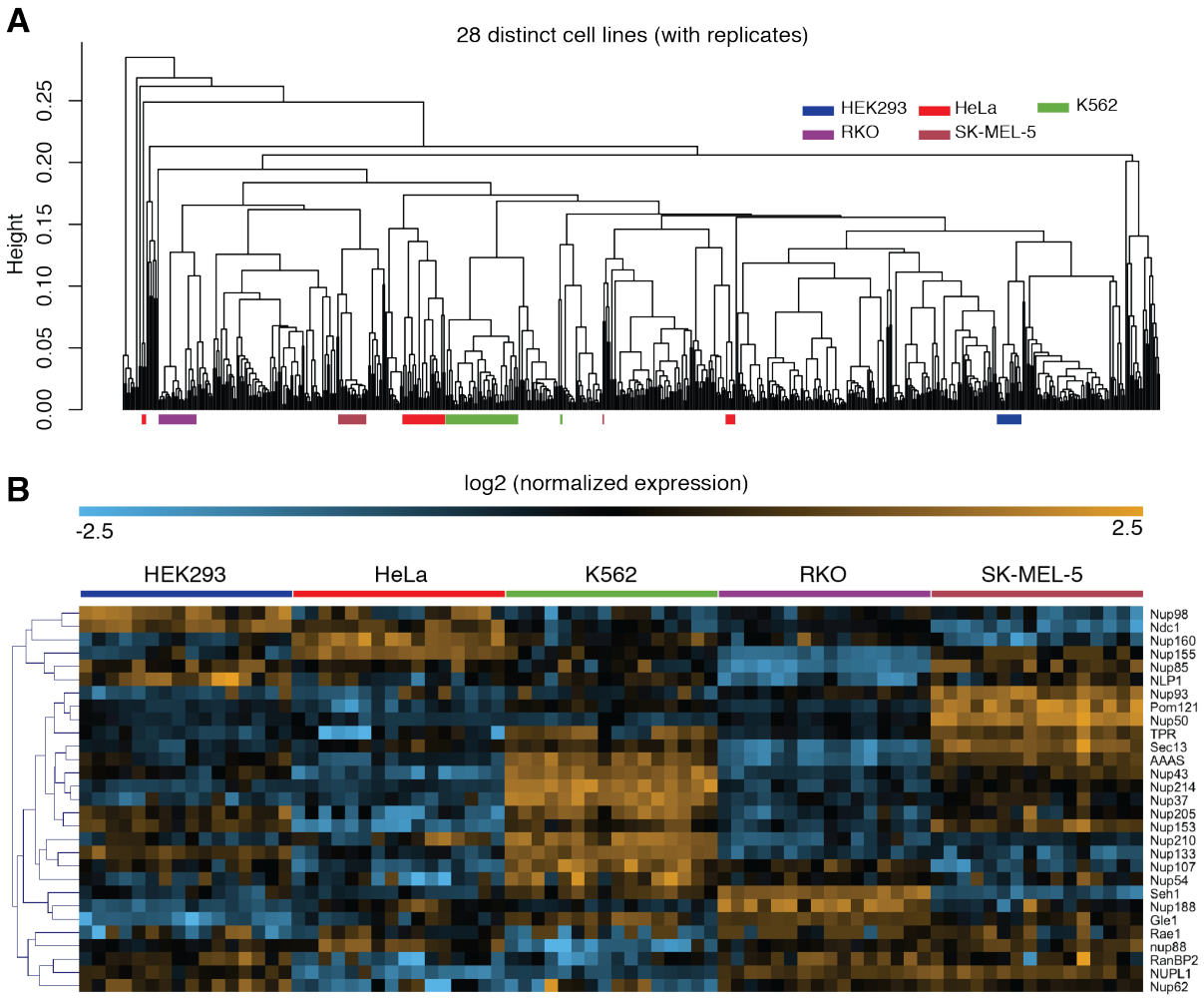

Figure S5, related to Figure 3. Selection of human cell lines for the analysis of NPC composition from gene expression data. (A) Average-linkage hierarchical clustering of Nups expression across 28 cell lines. The main clusters of the five selected cell lines are highlighted. (B) Normalized gene expression shows coherent, robust and characteristic expression profiles of Nup genes across the five cell lines as apparent from the consistent clustering of the 16 replicates per cell line shown. For details on data source and analysis see Supplementary Materials and Methods.

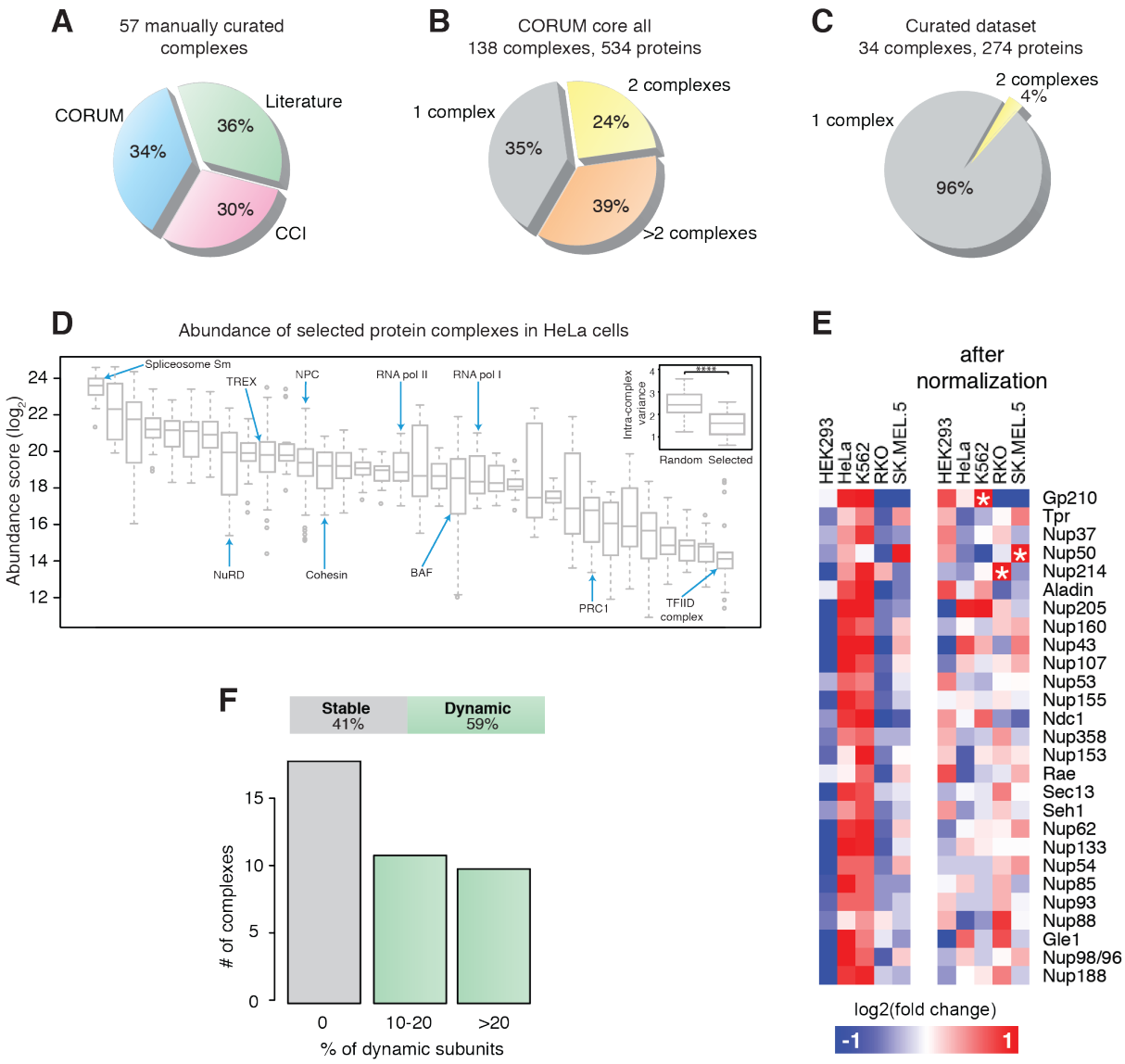

Figure S6, related to Figure 5. (A-D) Selection of nuclear protein complexes. (A) Pie chart displaying the sources of the protein complex definitions used. Complexes composed of at least five subunits were manually selected from the CORUM database (Ruepp et al, 2010), complex-complex interaction resource (CCI) (Malovannaya et al, 2011), and the available literature. (B-C) Pie charts comparing subunit redundancy in complexes definitions in the entire CORUM core dataset (Release February 2012) (B) and the manually curated dataset (C) by counting the number of proteins assigned to one or more entries. Complexes were filtered in order to reduce redundant and overlapping complex definitions and to ensure a minimum number of components quantified in our MS dataset (see Materials and Methods). (D) The absolute abundance of the 34 selected complexes spreads over several orders of magnitude in HeLa cells. Selected examples are indicated with light blue arrows. The variance of the absolute abundance of their subunits is significantly lower as compared to a set of randomly defined complexes of identical size (significance was assessed using Welch's t–test, **** = p–value < $1x10^{-8}$) (inset). (E) Heat map showing that a two-step normalization of protein abundance data is necessary to unravel compositional rearrangements of protein complexes, using the NPC as a test case. Protein abundance across the 5 cell lines was analyzed using label-free shotgun proteomics. When data are analyzed

without normalization for nuclear pore abundance (to take into account the fact that different cell lines have a different number of NPCs), the variation in Nup abundance is dominated by consistent up- or down-regulation of all components across cell lines (left). However, if variations in the total amount of each protein complex across cell types are taken in account, compositional rearrangements (changes in relative stoichiometry between subunits) are revealed (right). Starred squares indicate significant cases (adjusted p value < 0.01 and absolute $\log_2$-fold change > 0.9). Nups are ordered as in Figure 3A for comparison (targeted proteomics). Although not all seven dynamic Nups identified by targeted proteomics passed the significance threshold when applied to the less accurate shotgun data, we observed the same trend in almost all cases. (F) Complexes remodeling was analyzed using an alternative source of complex definitions (Havugimana et al, 2012). As in Figure 5, complexes were identified as dynamic if the expression of at least one subunit was significantly changed across the cell lines. Despite a slight inflation in the proportion of dynamic complexes, the results are in agreement with our analysis (Figure 5), indicating that a significant fraction of nuclear protein complexes undergo cell-type specific compositional rearrangements.

Supplementary Materials and Methods:

Cell lines

HeLa Kyoto cell line was maintained in Dulbecco's modified Eagle medium (DMEM) medium containing 1 g/L glucose supplemented with 2 mM L-glutamine; human embryonic kidney cells 293 (HEK293-Flp-In-T-REx) (Life Technologies, Darmstadt, Germany) were grown in DMEM medium containing 5 g/L glucose; K562 (ATCC number: CCL-243) in Iscove's Modified Dulbecco's Medium supplemented with 4 mM L-glutamine; RKO (ATCC number: CRL-2577) and SK-MEL-5 (ATCC number: HTB-70) in Eagle's Minimum Essential Medium (EMEM) supplemented with 2 mM L-glutamine, 1% non-essential aminoacids, and 1 mM sodium pyruvate. All media were supplemented with 10% heat inactivated fetal calf serum (FCS). All cell lines were propagated using standard procedures and grown at 37°C in 5% CO_2.

Isolation of nuclei and NEs from human cell lines

All following procedures were performed on ice or at 4°C. Cells growing in exponential phase (60-70% confluent) were harvested by scraping, collected by centrifugation and washed with cold PBS-Dulbecco (2.7 mM KCl, 1.5 mM KH_2PO_4, 137 mM NaCl, 8.1 mM Na_2HPO_4). Cells were allowed to swell in hypotonic buffer (0.2 M sucrose, 10 mM Tris-HCl pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$) for HEK293; 50 mM Tris-HCl pH 7.5 for the other cell lines, both supplemented with protease inhibitors (aprotinin 10 μg/mL, leupeptin 5 μg/mL) for 40 min, then lysed with 10-23 gentle strokes in a Dounce homogenizer using a tight pestle. Nuclei were collected by centrifugation at 1,000x *g* for 13 min, washed once with buffer NB (0.25 M sucrose supplemented with 50 mM Tris-HCl pH 7.5, 25 mM KCl, 5 mM $MgCl_2$, 2 mM DTT and protease inhibitors). For RKO cells, final concentrations of 10 mM KCl, 5 mM $MgCl_2$ and 0.25 M sucrose were added before centrifugation. Typically between 1 and 3e7 nuclei were then resuspended in 3 mL of a 2:1 mixture of buffer NC (2.3 M sucrose supplemented with 50 mM Tris pH 7.5, 25 mM KCl, 5 mM $MgCl_2$, 2 mM DTT and protease inhibitors) and buffer NB, underlayed with 1.5 mL buffer NC and spun for 30 min at 15,000x *g* (45 min at 17,800x *g* for HEK293). Pellet was resuspended with 200 μL buffer NB, diluted to 5 mL with buffer NB and spun for 13 min at 1,000x *g*. Recovered nuclei were resuspended in buffer NB counted with hemocytometer and either further processed or directly lysed for MS analysis (see below).

For NE isolation, purified nuclei (typically between 0.5 and 1e7) were resuspended in 1 mL buffer NE-A (0.1 mM $MgCl_2$, 1 mM DTT, 30 units DNaseI (Sigma-Aldrich GmbH, Munich, Germany) and 250 units RNase A (Sigma-Aldrich GmbH) and rapidly diluted with 4 mL buffer NE-B (10% (w/v) sucrose, 20 mM Tris-HCl pH

8.5, 0.1 mM $MgCl_2$, 1 mM DTT). Chromatin digestion was allowed to proceed for 15 min at room temperature and inspected by light microscopy. The digestion solution was then underlayed with 2 mL buffer NE-C (30% (w/v) sucrose, 20 mM Tris-HCl pH 7.5, 0.1 mM $MgCl_2$, 1 mM DTT) and spun at 1,500x *g* for 15 min. NEs were finally resuspended in 200 µL buffer NE-D (10% (w/v) sucrose, 20 mM Tris-HCl pH 7.5, 0.1 mM $MgCl_2$, 1 mM DTT), counted with hemocytometer and directly lysed for MS analysis.

FACS sorting

MoFlo cell sorter (Beckman Coulter GmbH, Krefeld, Germany) equipped with a 70 µm nozzle was used in the flow cytometric analysis and cell sorting of HeLa nuclei. Coherent Innova 90C Argon ion laser (Coherent Inc., Santa Clara CA, USA), tuned to 488 nm TEM00 mode (200 mW), was used as primary laser. Laser illumination, Moflo's L-configuration optical layout, and sorting were optimized using Flow-CheckTM Fluorospheres (Beckman Coulter GmbH). BD FACSFlowTM sheath (Becton Dickinson GmbH, Heidelberg, Germany), filtered in-line through a PALL Fluorodyne II filter 0.2 μm (Pall GmbH, Dreieich, Germany), was used in the acquisition and sorting of samples. The temperature on both sample and collection tubes was kept at 4°C during sorting. Data was analyzed using Moflo Summit software. Nuclei were sorted in Purify 1-drop mode and the sort decision gate was based on a combination of light scatter and pulse width parameters. Sorted nuclei were processed directly after sorting by adding solid urea, 2% (w/v) Rapigest (Waters GmbH, Eschborn, Germany) and 1 M NH_4HCO_3 to a final concentration of 4 M, 0.1% (w/v) and 100 mM, respectively.

Dextran staining

After isolation, NEs were diluted in wash buffer and transferred to a cold Lab-Tek 8 well imaging dish (Thermo Fisher Scientific GmbH, Bremen, Germany). The isolated organelles were spun gently at 4°C, 600x *g* for 6 min. After centrifugation, the Lab-Tek dish was kept on ice until imaging. TRITC-Dextran (molecular weight > 155 kDa) (Sigma-Aldrich GmbH), a kind gift of the Mattaj Lab, was added 10 min before imaging. Images were obtained on a Leica SP5 confocal microscope using a 63x NA 1.4 oil λ objective (Leica Microsystems GmbH, Wetzlar, Germany) and the 561 nm laser line for fluorophore excitation. Images were acquired with Leica LAS AF imaging software.

Immunofluorescence

Immunofluorescence was performed following standard protocols using 0.2% (v/v) Triton X-100 in PBS as permeabilizing buffer, 2% (v/v) FCS in PBS supplemented with 2% (w/v) bovine serum albumin (BSA) as blocking buffer and 2% (w/v) paraformaldehyde in PBS as fixation buffer. Samples were incubated for 1 h at room temperature with primary antibodies diluted in blocking buffer (1:2,500 for MAb414, Covance, Princeton NJ, USA, and 1:100 for anti-Gp210 antibody, kind gift of Dr. Joseph Glavy) and 1h at room temperature with 1:1,000 dilution of secondary antibodies (anti-mouse IgG-Cy3, Sigma-Aldrich GmbH, and anti-rabbit IgG-Alexa Fluor 488, Life Technologies) in blocking buffer. Samples were counterstained with Hoechst-33342 (Sigma-Aldrich GmbH) at 1 µg/mL for 5 min at room temperature. Samples were mounted with Mowiol (Carl Roth GmbH, Karlsruhe, Germany) on glass coverslips for imaging purpose. Imaging was performed on a LSM780 (Zeiss GmbH, Oberkochen, Germany) confocal microscope with a 63x objective (Plan-Apochromat 63x/1.4 Oil Zeiss) using the 514, 488 and 405 nm laser lines for fluorophore excitation.

Nup-release Western Blot

Nuclei were purified as described above, and resupended to a final concentration of 1e7 nuclei/mL. 100-200 µL were used per incubation condition. Each sample was incubated at the corresponding temperature and time, then spun at 10,000x *g* for 10 min at 4°C, and supernatant and pellet were collected. 4x Laemmli sample buffer was added to the supernantant fraction, and the pellet was resuspended in 1x-Laemmli buffer to the same final volume. Samples were incubated at 95°C for 10 min, and equal volumes were loaded for SDS-PAGE and WB. Standard protocols were used for semi-dry transfer of the proteins onto a nitrocellulose membrane followed by WB. The primary antibodies used were MAb414 (Covance), anti-laminB1/B2 IQ176 (clone X223, Immuquest, Seamer, UK), anti-Nup107 and anti-Nup43 (kind gifts of Dr. Joseph Glavy). Secondary antibodies conjugated to HRP were from Jackson ImmunoResearch Inc. (West Grove PA, USA).

Sample preparation for mass spectrometry

Nuclei and NE were resuspended in the appropriate wash buffer without protease inhibitors at a concentration between 0.5 and 1.5e7/mL. Proteins were solubilized by adding urea buffer (10 M urea in 250 mM NH_4HCO_3) and 2% (w/v) Rapigest (Waters) to a final concentration of 4 M and 0.2% (w/v), respectively. Samples were then sonicated in a vial tweeter for 3x 30 s, each time followed by one minute rest on ice, and stored at -80°C until further processing.

Samples were rapidly thawed in a bath sonicator and further sonicated in a vial tweeter for 3x 30 s, each time followed by one minute rest on ice. Samples were reduced using 10 mM DTT (30 min at 37°C) and cysteines were alkylated with 15 mM iodoacetamide (30 min in the dark). Protein digestion was performed in two steps, first using 1:100 (w/w) LysC (Wako Chemicals GmbH, Neuss, Germany) for 4 h at 37°C then finalized with 1:50 (w/w) trypsin (Promega GmbH, Mannheim, Germany) overnight at 37°C, after the urea concentration was diluted to 1.5 M. Samples were then acidified with 10% (v/v) TFA and the cleavage of Rapigest was allowed to proceed for 30 min at 37°C. Samples were clarified by centrifugation at 17,000x *g* for 5 min at room temperature and desalted using C18 spin columns (Harvard Apparatus, Holliston MA, USA) following standard procedures. For absolute quantification, AQUA peptides mix (Thermo Fisher Scientific GmbH) was spiked into the samples prior the addition of LysC at a concentration of 1 pmol per peptide / 1e6 nuclei or NE. For relative quantification, crude synthetic peptides mix was spiked prior acidification. In both cases, the samples were spiked with 1:20 (v/v) retention time kit (RT-kit) (Biognosys AG, Schlieren, Switzerland) prior the injection in the mass spectrometer. Between 2,000 and 5,000 nuclei or NE equivalents (corresponding approximately to 300-700 ng protein) were analyzed by LC-MS/MS or targeted MS.

LC-MS/MS

Peptides were analyzed using a nanoAcquity UPLC system (Waters GmbH) connected online to a LTQ-Orbitrap Velos or LTQ-Orbitrap Velos Pro instrument (Thermo Fisher Scientific GmbH). Peptides were separated on a BEH300 C18 (75 μm x 250 mm, 1.7 μm) nanoAcquity UPLC column (Waters GmbH) using a stepwise 145 min gradient between 3 and 85% (v/v) ACN in 0.1% (v/v) FA. Data acquisition was performed using a TOP-20 strategy where survey MS scans (m/z range 375-1,600) were acquired in the orbitrap (R = 30,000 FWHM) and up to 20 of the most abundant ions per full scan were fragmented by collision-induced dissociation (normalized collision energy = 35, activation Q = 0.250) and analyzed in the LTQ. Ion target values were 1,000,000 (or 500 ms maximum fill time) for full scans and 10,000 (or 50 ms maximum fill time) for MS/MS scans. For the development of multiple reaction monitoring (MRM) assays, data acquisition was performed using a TOP-10 strategy where up to 10 of the most abundant ions per full scan were fragmented by higher-energy collisionally activated dissociation (HCD, normalized collision energy = 35, activation Q = 0.100) and analyzed in the Orbitrap (R = 7,500 FWHM). In this case, ion target value for MS/MS scans was 50,000 (or 100 ms maximum fill time) for MS/MS scans. Charge states 1 and unknown were rejected. Dynamic exclusion was enabled with repeat count = 1, exclusion duration = 60 s, list size = 500 and mass window ± 15 ppm.

LC-MRM

A key step for the establishment of a successful targeted proteomic assay is the accurate selection of proteotypic peptides (PTPs) for the targets of interest. Therefore, we first performed a discovery phase aimed at the selection of PTPs, which was based on a deep proteomic characterization of NE samples. NE tryptic digests from four biological replicates were analyzed by LC-MS/MS (see LC-MS/MS section). At least two technical replicates per sample were analyzed to ensure the highest proteome coverage. In total, more than 2,500 NE-associated proteins were identified including all the known Nups. In particular, more than 1,000 unique Nup-derived peptides were identified. From this dataset, we selected PTPs using a semi-automated pipeline that assigned suitability scores based upon:

1. Absence of tryptic missed-cleavage sites;
2. Peptide intensity based on spectral counting (this also account for reproducibility of detection between different runs);
3. Absence of methionine residues;
4. Absence of known post-translation modifications (this was assessed by querying protein databases, including PhosphoELM (Dinkel et al, 2011), PHOSIDA (Gnad et al, 2011) and UniProt (UniProt-Consortium, 2009), and large scale proteomic datasets (Choudhary et al, 2009; Zielinska et al, 2010);
5. Hydrophobicity scores, in order to discard extremely hydrophobic peptides that could present solubility issues during chemical synthesis and storage, and extremely hydrophilic peptides that could present reproducibility issues in their chromatographic behavior.

Additionally, suppliers inspected the suitable peptides in order to avoid difficult sequence compositions for chemical synthesis. Using this pipeline, we selected 2-4 PTPs per target protein that were synthesized isotopically labeled (carboxy-terminal Arg(13C6; 15N4) or Lys(13C6; 15N2)) using spot-synthesis, and supplied in a crude form (~70% purity) by JPT Peptide Technologies GmbH (Berlin, Germany). PTPs were resuspended with 20% (v/v) acetonitrile in 0.1% (v/v) formic acid, pooled and desalted using standard procedures.

LC-MRM assays were generated from high-resolution LC-MS/MS runs. Peptides were diluted, spiked with the iRT-kit (Biognosys AG) according to the manufacturer's instructions by LC-MS/MS in an Orbitrap Velos using HCD for peptide fragmentation. The best spectrum of each peptide served as template to derive the 12 most intense fragment ions (Picotti et al, 2009) and approximate retention times relative to the iRT-peptides (Escher et al, 2012). The resulting 12 transitions were measured in scheduled MRM mode after retention time calibration

according to the iRT-kit peptides' retention times on a TSQ Vantage triple quadrupole mass spectrometer (Thermo Fisher Scientific) equipped with a standard nano-electrospray source connected to a Proxeon Easy nLC II system (Thermo Fisher Scientific). Samples were loaded onto a nanoLC column prepared in house by packing PicoFrit emitters (360 μm OD, 75 μm ID, 10 μm tip, New Objective Inc., Wolburn MA, USA) with 11 cm of Magic C18 AQ resin (3 μm particle size, 200 Å pore size, Michrom Bioresources Inc., Auburn CA, USA). Gradient elution was performed using solvents A = 3% (v/v) acetonitrile in 0.1% (v/v) formic acid in water and B = 3% (v/v) water in 0.1% (v/v) formic acid in acetonitrile according to the following gradient program: 0-30 min: 5-35% B, 30-32 min: 35-100% B, 32-40 min: 100% B. The flow rate was set to 300 nL/min. The LC eluent was electrosprayed using an ionization voltage of 1.9 kV. MRM analysis was performed using Q1 and Q3 resolutions of 0.7 Da with a cycle time of 2.5 s. Collision energies were calculated for the TSQ Vantage by a linear regression according to the vendor's specifications.

Data were processed using mProphet (Reiter et al, 2011). Relative fragment intensities for the triple quadrupole mass spectrometer and exact iRTs were extracted. The five most intense transitions were selected for the final LC-MRM assay library that was split in three or four runs (corresponding to an average of 38 assays per run) with optimal spreading of iRTs over the runs in order to maximize the dwell time per peptide. Measurement in scheduled MRM mode using 3 min acquisition windows was carried out after calibration of retention times according to the iRT-kit (LC-MRM parameters as above). Assays were manually validated by the detection of co-eluting light and heavy traces in spiked NE samples.

A further selection was applied in order to establish the assays to be used for absolute quantification using AQUA peptides. Therefore, technical and biological replicates of spiked NE samples were measured in MRM-mode and analyzed using mProphet. Only assays that could be reproducibly detected in all the replicates measured were considered as candidates for absolute quantification. If more than two candidates were available for a target protein, the two assays with the highest average mProphet score across replicates were selected. Using this procedure, we selected 76 high-confidence PTPs that were purchased as isotopically labeled AQUA-Ultimate peptides (Thermo Fisher) and supplied as accurately quantified (± 5%) stock solutions. Stock solutions were pooled, acetonitrile was added to a final concentration of 20%, and aliquots were stored at -80°C. Each aliquot was thawed only once and then discarded. Peptides were spiked at 1 pmol per $1x10^6$ nuclei or NEs extract.

Data processing: LC-MS/MS

Acquired raw files were converted to centroid mzXML format using readW and then to mgf format using the MzXML2Search script part of the Trans Proteomic Pipeline (TPP) (Deutsch et al, 2010). After conversion, MS/MS spectra were searched using Mascot v2.2.03 (Matrix Science, London, UK) against a concatened target-decoy database derived from the human Swiss-Prot entries of the Uniprot release 2010_09. The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues, unless followed by proline); 2 missed cleavages were allowed; carbamidomethylation (C) was set as fixed modification; oxidation (M), acetylation (protein N-term), 13C6-15N2 (K) and 13C6-15N4 (R) were applied as variable modifications, if applicable; mass tolerance of 10 ppm (precursor) and 0.6 Da (fragments). The database search results were further processed using the PeptideProphet (Keller et al, 2002) and ProteinProphet (Keller et al, 2005) program. Protein hits were filtered at a false discovery rate of 1% using a target-decoy strategy (Elias & Gygi, 2007).

For quantitative label-free analysis, raw files were analyzed using MaxQuant v1.2.2.5 (Cox & Mann, 2008) and protein abundance scores were calculated from the summed intensity of proteotypic peptides, similarly to what already described in (Eberl et al, 2012; Geiger et al, 2012; Schwanhausser et al, 2011). MS/MS spectra were searched against the human Swiss-Prot entries of the Uniprot release 2011_12 using the Andromeda search engine (Cox et al, 2011). The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues, unless followed by proline); 2 missed cleavages were allowed; carbamidomethylation (C) was set as fixed modification; oxidation (M) and acetylation (protein N-term) were applied as variable modifications, if applicable; mass tolerance of 20 ppm (precursor) and 0.5 Da (fragments). The reversed sequences of the target database were used as decoy database. Peptide hits were filtered at a false discovery rate of 1% using a target-decoy strategy (Elias & Gygi, 2007). After the analysis with MaxQuant, the output data was processed as follows: i) peptides matching to at least one reverse or contaminant hit were filtered out; ii) peptides matching to a unique protein and having at least a second peptide partner matching to the same protein were kept (i.e., two unique peptides per protein); iii) protein intensity was calculated by summing up the intensity of all the unique peptides matching that protein; iv) finally, the intensity was normalized by dividing it over the molecular weight of the protein, and log2-transformed to obtain the final abundance score.

Data processing: LC-MRM

Raw files were converted to mzXML format using readW and processed using mProphet (Reiter et al, 2011) using a "reference peptide spike-in" workflow. The resulting peak groups were filtered using a normalized discrimination score (d_score) > 4, generally resulting in a FDR << 0.001, and manually inspected. L/H ratios were calculated from the apex sum of the selected transitions (at least 2, generally 5) using the "apex sum without outlier" option of mProphet. mProphet output files were processed using routines written in R statistical programming language.

For absolute quantification, four biological replicates were combined and protein ratios were calculated from the median value of all the measurements ± median absolute deviation (Table S1). All protein ratios were normalized to the average ratio of the iso-stoichiometric components of the Nup107 subcomplex (Nup107, Nup133, Nup160, Nup43, Nup37, Nup85 and Nup96), which was set as 1 (Figure 1). Since we detected a significant enrichment for Nup50 (1.5 fold, p = 1.08e-06, t-test) and Tpr (1.4 fold, p = 2.07e-06, t-test) in the whole nuclear extracts when compared to NE, we estimated the proportion of Nup50 and Tpr that co-purified with remaining chromatin-associated proteins in the NE extracts and corrected their abundance for this factor using the following formula:

$$\text{ratio(NEpool)} = \text{ratio(NEmeasured)} - (\text{NCF}*(\text{ratio(Nmeasured)}) - \text{ratio(NEpool)}))$$

where NCF (Nuclear Contamination Factor) is a general estimate of the fraction of chromatin-associated proteins that remains in the NE fraction after RNase and DNase digestion. NCF was derived from LC-MS/MS analysis of 3 biological replicates each of nuclei and NEs from HeLa cells as the ratio between the summed intensities peptides deriving from histones and lamins. The ratio histones/lamins should therefore provide an estimate of the chromatin-associated pool of proteins that has been depleted during the procedure. Peptides' intensities were extracted using MaxQuant; only proteotypic peptides from histone H4 (UniProt accession: P62805) and histone H1.5 (UniProt accession: P16401)), Prelamin A/C (UniProt accession: P02545), Lamin B1 (UniProt accession: P20700) and Lamin B2 (UniProt accession: Q03252)) were used. NCF was thus estimated at 0.384 ± 0.070, indicating that more 60% of the chromatin-associated pool of proteins was depleted during our procedure (Figure S1D).

For relative quantification, samples were analyzed in biological triplicates and selected ratios were normalized across samples using quantile normalization (Bolstad et al, 2003) as implemented in the package 'affy' of Bioconductor (Gentleman et al, 2004). Differential protein levels of human Nups in each cell line versus all others were evaluated using linear mixed-effect models (as implemented in R package 'nlme' (version 3.1-103))

with cell line as the fixed factor and multiple peptides as the random factor. Significant differences in protein levels of Nups were noted in individual cell lines with p-value cut-off 0.01 after false discovery rate (FDR) correction for multiple hypothesis testing.

Generation of HEK293_mEos2-Nup107_ miRNAs-αNup107 cell line

The HEK293-Flp-In-T-REx cell line was stably transfected with an inducible plasmid expressing mEos2-Nup107 and two microRNAs (miRNAs) against the endogeneous Nup107 mRNA (derived from the BLOCK-iT Inducible Pol II miRNA RNAi Expression Vector w/EmGFP from Life Technologies). Cells were treated for 96 h with 1 μg/mL of tetracycline to induce the expression of mEos2-Nup107 and the miRNAs.

Quantification of mEos2-Nup107 expression

Nuclei from HEK293_mEos2-Nup107_ miRNAs-αNup107 were purified as described above. Purified nuclei were boiled with Laemmli sample buffer, and loaded for SDS-PAGE and WB with anti-Nup107 antibody. The bands were visualized using a chemiluminiscence solution (Amersham ECL Western Blotting Detection Reagents) and the membrane was exposed onto X-ray films (Amersham Hyperfilm ECL). At least 3 different film exposures were quantified. The films were scanned in a Canon flatbed scanner in grayscale at a resolution of 1200 dpi. The analysis was done with ImageJ 1.46h (Abramoff et al, 2004) as follows: background was subtracted with the default parameters (rolling ball method, with radius=50), lanes were selected and the profile was plotted, a straight horizontal line was drawn to define the x-axis background, each peak was divided in half by a vertical line, and the area under the curve (of half peak) was measured. The percentage corresponding to each band (mEos2-Nup107 and endogenous Nup107) was calculated.

Measurement of Nup107 copies per NPC by super-resolution microscopy

The following procedure describes an imaging and analysis scheme that robustly permits counting the minimal number of mEos2 proteins in the nuclear envelope. In our conservative analysis, we do not take any maturation efficiency of mEos2 into account, which can reduce the number of photoactive and emitting fluorophores (Ulbrich & Isacoff, 2007). The actual copy number will thus even be higher than the minimal value determined.

Imaging was performed on a custom built microscope centered around a 100x high numerical aperture total internal reflection fluorescence (TIRF) objective (UAPON 100XOTIRF, N.A. 1.49, Olympus GmbH, Hamburg, Germany) equipped with 405 nm (150 mW OBIS, Coherent Inc.), 488 nm (100 mW OBIS, Coherent Inc.) and

568 nm (100 mW Sapphire, Coherent Inc.) laser lines for excitation. A linear stage (M505, PI, Karlsruhe, Germany) was used to select the beam position at the back port to switch between epi-fluorescence and TIRF mode. Fluorescence was separated from excitation light with a quadband dichroic (Chroma Technology Corp., Bellows Falls VT, USA) and a double notch (Stopline 568/647, Semrock Inc., Rochester NY, USA). The collected fluorescence was spectrally split through an Optoslit II (Cairn Research, Faversham, UK) equipped with a 555 dichroic (Brightline, Semrock) and filtered with a 525/50 (Semrock) and a 609/57 (Brightline, Semrock) bandpass filter. The image was recorded by a 512x512 pixel iXon X3 EMCCD Camera (Andor Technology, Belfast, Northern Ireland).

For imaging, purified nuclei were spun at 100x *g* for 9 min on an 8 well Lab-Tek chamber slide (Thermo Fisher Scientific GmbH) together with gold nanospheres (100 nm diameter, Corpuscolar Inc, Cold Spring NY, USA), used as drift fiducial during image acquisition. Imaging using the 488 nm laser line at low power (~1 mW at the back aperture) was used to focus on the nuclear envelope resting on the coverslip. Photo-conversion with 405 nm light (15 ms exposure time) converted sparse mEos2 from the green emitting state to the orange emitting state. This was followed by 17 images of the orange fluorescent state with a frame exposure time of 0.12 s using the 568 nm laser (100 mW). Cycles of photo-conversion and photo-bleaching were performed until complete bleaching of the sample (~1200 s). Complete photo-conversion was then verified by absence of green fluorescence using 488 nm excitation at the end of the experiment. In order to compensate for the decrease in the number of fluorophores during the experiment a hyperbolic power increase excitation strategy for the 405 laser was adopted. Power of the 405 laser at time *t* after the start of the image series was set according to:

$$I(t) = I_b + \frac{I_0}{1 - \frac{I_0 t}{\tau}}$$

Where I(t) is the excitation intensity $I_b = 4.7$ mW, $I_0 = 100\ \mu$W and $\tau = 121000$ s μW.

In order to count the number of mEos2-Nup107 molecules in the sample, single fluorophores were activated in different NPCs in the TIRF field. To calculate the number of fluorophores in each NPC the different fluorescent spots need to be detected and assigned to the corresponding NPC. To achieve this goal the different NPCs were identified in the following way; a super-resolution image was reconstructed by summing all the single particles detected in the entire image series (Shroff et al, 2008) and isolated pores were located with a density based

clustering algorithm. The analysis steps to perform such a task are shown in detail in Figure S3 and described hereof:

1. Event detection: converted mEos2-Nup107 show up as individual spots in the orange channel. These spots were detected in each frame (red circles in Figure S3A) using the "Localizer package" (Dr. Peter Dedecker, Katholieke University, Louvain, Belgium) for IgorPro (WaveMetrics, Portland OR, USA) by applying a threshold on the Generalized Likelihood Ratio Test. The center of the spot was localized through Gaussian fitting with the known Point Spread Function with sub-pixel resolution (Figure S3B). The resolution, as given by the localizer package (i.e. the precision of the localization), was at full width half maximum (FWHM)= 37 nm.
2. Super-resolution image reconstruction: a super-resolution image (15 nm*15 nm pixel size) was then reconstructed from binning all the detected events and convolving the resulting image with a Gaussian width according to the determined resolution of 37 nm (Figure S3C).
3. Isolated pores selection: in order to detect isolated pores we analyzed the detected events with a density based clustering algorithm, DBSCAN (Ester et al, 1996). We used a threshold parameter for the clustering algorithm to ensure the retrieval of at least 80% of the data. The result of the clustering algorithm overlapped with the super-resolution image is shown in Figure S3D. To select isolated pores we filtered the resulting clusters on the base of the cluster size (calculated as average distance from the center of mass) and roundness (calculated as variance of distance from the center of mass). To remove noise particles outside of the nuclear rim (such as fluorescent particles sticking to the coverslip) from further analysis we fitted a 2D Gaussian function on the entire image and selected the detected particles lying within 1.5 sigma from the center of the fitted Gaussian. The result of the pore selection is shown in Figure S3E.
4. Average pore number estimation: mEos2 is known to transiently undergo transitions into a non-fluorescent state (Annibale et al, 2010). Such a transition can in theory lead to over-counting, e.g. if a particle recovered from the dark state is misinterpreted as a newly activated fluorophore. To eliminate this concern we followed the method presented by Annibale et al. to reliable extract the correct number of mEos2 particles from such an iterative photo-conversion / photo-bleaching experiment (Annibale et al, 2011). Every particle appearing in the same pixel within a time t_d was counted as one. t_d was varied from 0 to 63 s and the average number of particles N in one NPC was calculated as a function of t_d, which generated the N vs t_d histogram in Figure S3G. A multi exponential decay is visible, that originates from blinking and photo-conversion kinetics. According to Annibale et al., the first part of the curve is purely mono-exponential as it contains

only blinking kinetics. The exact mono-exponential fitting range was further validated by comparing the Akaike (Akaike, 1974) information criterion (AIC) for both mono-exponential (blue line in Figure S3G) and bi-exponential decay (red line in Figure S3G) for increasing fitting range. The N vs t_d histogram was fit (blue line in Figure S3H) until the maximal td that still allowed a significantly better mono-exponential fit (AIC for the mono-exponential decay lower than for the bi-exponential one) with the semi-empirical function

$$N(t_d) = N(1 + n_{blink} e^{\frac{1-t_d}{t_{off}}})$$

Being N the number of fluorophores (which is the copy number of mEos2 per single NPC), n_{blink} is the fraction of fluorophores blinking and t_{off} is the average permanence time into the dark state.

5. mEos2-Nup107 distribution: the final copy number (N) frequency distribution was obtained from 172 clusters imaged from 5 nuclei and corrected for labeling efficiency (Figure S3I).

Monte-Carlo simulation of super-resolution images of NPCs

Simulated super-resolution images of NPCs were generated using the following steps:

1. Number of pores was set according to the estimated experimental pore density. NPC positions were generated extracting a random number from a uniform distribution along each axis.
2. For each pore, N fluorophores were distributed around the central axis at a fixed radius (45nm) following an 8-fold symmetry. For each fluorophore, a uniformly distributed random number between 0 and 1 was extracted and compared to the expected overall labeling efficiency, i.e. 0.64 assuming the measured 0.8 gene replacement efficiency and a realistic value for fluorophore maturation of 0.8 (Ulbrich & Isacoff, 2007). If the extracted number was higher than the efficiency value the fluorophore was not considered for further analysis.
3. Taking the known photo physics properties of mEos2 into account (Lee et al, 2012) for each fluorophore, a 4-level system photodynamic approximation was assumed (Figure S4E). The photo-conversion was randomly extracted from a uniform distribution between the pulse periods along the total duration of the experiment (pulse frequency, total duration of the experiment and frame rate were set according to the actual experimental conditions). The rate constant value for k_{CD} and k_{DC} and k_B were taken from literature (Lee et al, 2012). For each frame in which the fluorophore was found in the converted emitting state the fluorophore

position was recorded and noise was introduced by extracting a value from a gaussian distribution with a standard deviation of 20 nm. An image resulting from the simulation is shown in Figure S4F.

4. Four different datasets were generated assuming different copy number distribution. The first three simulations assumed respectively a copy number of 16, 32 and 64 protein per NPC (Figure S4G). The last simulation was generated assuming an equal distribution of the three values (Figure S4H). The synthetic datasets were analyzed with the same routines as used for the experimental data (see above).

Gene expression of Nups: microarray data source, pre-processing and data analysis

Gene expression data were retrieved from the Gene Expression Atlas (available through ArrayExpress), which contains 5,372 human samples integrated from 206 public experiments (HG-U133A platform) including 369 different cell- and tissue-types, disease states and cell lines (Kapushesky et al, 2012; Lukk et al, 2010). Raw data were acquired and pre-processed separately for cell lines and tissue types using GCRMA within R Bioconductor framework (Gentleman et al, 2004; Wu et al, 2004). In total, 28 cell lines having at least 10 replicates and 44 tissue types (excluding universal-reference) having at least 6 replicates were retained for this study. Additionally, a large-scale dataset of disease-associated microarray experiments were collected from the publicly available repository, NCBI GEO (Barrett et al, 2011) using a previously published method (Butte & Chen, 2006; Dudley et al, 2009). After extensive filtering with manual annotation (at least 8 disease and control states) and pre-processing, expression profiles of 30 unique diseases derived from disease-associated primary tissues were generated for this study. The analysis was performed using 29 single Nup probe sets selected on the basis of their specificity, sensitivity (GeneAnnot (Chalifa-Caspi et al, 2004)) and variance (based on the cell lines dataset). In addition to pre-processing, the expression of Nups across samples was median-centered to be able to focus on condition specific changes of individual subunits rather than the whole complex. In order to avoid analysis bias towards conditions represented by higher number of replicates, we retained equal number of replicates for selected five cell lines (16 replicates for HeLa, HEK293, K562, RKO, SK-MEL-5), tissues (5 replicates) and diseases (8 disease against 8 controls). To achieve these, we discarded additional replicates by recursively filtering the sample with the lowest Pearson correlation (using the 29 Nups) across replicates. Differentially expressed genes were detected by comparing expression values of gene in each cell line/tissue against the expression in all the rest of cell lines/tissues. In case of disease-related gene expression profiles, disease samples were compared against control samples in order to reveal condition-specific changes. In both

analyses, the significance of condition-specific differential expression was assessed using an unpaired, two-sided t-test (FDR corrected, adjusted p–value < $1x10^{-2}$ with an additional constraint of absolute log_2-fold change > 1).

Supplementary References:

Abramoff MD, Magalhães PJ, Ram SJ (2004) Image processing with ImageJ. *Biophotonics international* **11:** 36-42

Akaike H (1974) A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **19:** 716-723

Annibale P, Scarselli M, Kodiyan A, Radenovic A (2010) Photoactivatable Fluorescent Protein mEos2 Displays Repeated Photoactivation after a Long-Lived Dark State in the Red Photoconverted Form. *J Phys Chem Lett* **1:** 1506-1510

Annibale P, Vanni S, Scarselli M, Rothlisberger U, Radenovic A (2011) Quantitative Photo Activated Localization Microscopy: Unraveling the Effects of Photoblinking. *PLoS One* **6:** e22678

Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Holko M, Ayanbule O, Yefanov A, Soboleva A (2011) NCBI GEO: archive for functional genomics data sets-10 years on. *Nucleic Acids Res* **39:** D1005-D1010

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19:** 185-193

Butte AJ, Chen R (2006) Finding disease-related genomic experiments within an international repository: first steps in translational bioinformatics. *AMIA Annu Symp Proc***:** 106-110

Chalifa-Caspi V, Yanai I, Ophir R, Rosen N, Shmoish M, Benjamin-Rodrig H, Shklar M, Stein TI, Shmueli O, Safran M, Lancet D (2004) GeneAnnot: comprehensive two-way linking between oligonucleotide array probesets and GeneCards genes. *Bioinformatics* **20:** 1457-1458

Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M (2009) Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* **325:** 834-840

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26:** 1367-1372

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J Proteome Res* **10:** 1794-1805

Deutsch EW, Mendoza L, Shteynberg D, Farrah T, Lam H, Tasman N, Sun Z, Nilsson E, Pratt B, Prazen B, Eng JK, Martin DB, Nesvizhskii AI, Aebersold R (2010) A guided tour of the Trans-Proteomic Pipeline. *Proteomics* **10:** 1150-1159

Dinkel H, Chica C, Via A, Gould CM, Jensen LJ, Gibson TJ, Diella F (2011) Phospho.ELM: a database of phosphorylation sites-update 2011. *Nucleic Acids Res* **39:** D261-D267

Dudley JT, Tibshirani R, Deshpande T, Butte AJ (2009) Disease signatures are robust across tissues and experiments. *Mol Syst Biol* **5:** 307

Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M, Mann M (2012) A Map of General and Specialized Chromatin Readers in Mouse Tissues Generated by Label-free Interaction Proteomics. *Mol Cell* **49:** 368-378

Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4:** 207-214

Escher C, Reiter L, Maclean B, Ossola R, Herzog F, Chilton J, Maccoss MJ, Rinner O (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12:** 1111-1121

Ester M, Kriegel H-p, Jörg S, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)***:** 226-231

Geiger T, Wehner A, Schaab C, Cox J, Mann M (2012) Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins. *Mol Cell Proteomics* **11:** M111.014050

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge YC, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G et al (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5:** R80

Gnad F, Gunawardena J, Mann M (2011) PHOSIDA 2011: the posttranslational modification database. *Nucleic Acids Res* **39:** D253-D260

Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL, Li Z, Wang PI, Boutz DR, Fong V, Phanse S, Babu M, Craig SA, Hu P, Wan C, Vlasblom J, Dar VU, Bezginov A, Clark GW, Wu GC, Wodak SJ et al (2012) A census of human soluble protein complexes. *Cell* **150:** 1068-1081

Kapushesky M, Adamusiak T, Burdett T, Culhane A, Farne A, Filippov A, Holloway E, Klebanov A, Kryvych N, Kurbatova N, Kurnosov P, Malone J, Melnichuk O, Petryszak R, Pultsin N, Rustici G, Tikhonov A, Travillian RS, Williams E, Zorin A et al (2012) Gene Expression Atlas update-a value-added database of microarray and sequencing-based functional genomics experiments. *Nucleic Acids Res* **40:** D1077-D1081

Keller A, Eng J, Zhang N, Li XJ, Aebersold R (2005) A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol Syst Biol* **1:** 2005.0017

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

Lee SH, Shin JY, Lee A, Bustamante C (2012) Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc Natl Acad Sci U S A* **109:** 17436-17441

Lukk M, Kapushesky M, Nikkila J, Parkinson H, Goncalves A, Huber W, Ukkonen E, Brazma A (2010) A global map of human gene expression. *Nat Biotechnol* **28:** 322-324

Maimon T, Elad N, Dahan I, Medalia O (2012) The Human Nuclear Pore Complex as Revealed by Cryo-Electron Tomography. *Structure* **20:** 998-1006

Malovannaya A, Lanz RB, Jung SY, Bulynko Y, Le NT, Chan DW, Ding C, Shi Y, Yucer N, Krenciute G, Kim BJ, Li CS, Chen R, Li W, Wang Y, O'Malley BW, Qin J (2011) Analysis of the Human Endogenous Coregulator Complexome. *Cell* **145:** 787-799

Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R (2009) Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell* **138:** 795-806

Rabut G, Doye V, Ellenberg J (2004) Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nat Cell Biol* **6:** 1114-1121

Reiter L, Rinner O, Picotti P, Huttenhain R, Beck M, Brusniak MY, Hengartner MO, Aebersold R (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat Methods* **8:** 430-435

Ruepp A, Waegele B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, Frishman G, Montrone C, Mewes HW (2010) CORUM: the comprehensive resource of mammalian protein complexes-2009. *Nucleic Acids Res* **38:** D497-D501

Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* **473:** 337-342

Shroff H, White H, Betzig E (2008) Photoactivated localization microscopy (PALM) of adhesion complexes. *Curr Protoc Cell Biol* **Chapter 4:** 21.21-21.27

Ulbrich MH, Isacoff EY (2007) Subunit counting in membrane-bound proteins. *Nat Methods* **4:** 319-321

UniProt-Consortium (2009) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res* **38:** D142-148

Wu ZJ, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F (2004) A model-based background adjustment for oligonucleotide expression arrays. *J Am Statist Assoc* **99:** 909-917

Zielinska DF, Gnad F, Wisniewski JR, Mann M (2010) Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. *Cell* **141:** 897-907