Dynamics of re-constitution of the human nuclear proteome after cell division is regulated by NLS-adjacent phosphorylation

Gergely Róna[1,*], Máté Borsos[1,2], Jonathan J Ellis[3], Ahmed M. Mehdi[3], Mary Christie[3], Zsuzsanna Környei[5], Máté Neubrandt[5], Judit Tóth[1], Zoltán Bozóky[1], László Buday[1], Emília Madarász[5], Mikael Bodén[3,6,7], Bostjan Kobe[3,4,6], Beáta G. Vértessy[1,8,*]

[1] Institute of Enzymology, RCNS, Hungarian Academy of Sciences, H-1117 Budapest, Hungary, [2] Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Dr. Bohr-Gasse 3, 1030 Vienna, Austria, [3] School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane Queensland 4072, Australia, [4] Australian Infectious Diseases Research Centre, The University of Queensland, Brisbane Queensland 4072, Australia, [5] Institute of Experimental Medicine of Hungarian Academy of Sciences, H-1083 Budapest, Hungary, [6] Institute for Molecular Bioscience, The University of Queensland, Brisbane Queensland 4072, Australia, [7] School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane Queensland 4072, Australia, [8] Department of Applied Biotechnology, Budapest University of Technology and Economics, H-1111 Budapest, Hungary

[*]**Corresponding authors:**
Gergely Róna (rona.gergely@ttk.mta.hu) and Beáta G. Vértessy (vertessy.beata@ttk.mta.hu)

Author to communicate with the Editorial and Production offices:
Beáta G. Vértessy, phone: +3613826707, fax: +3614665465, e-mail: vertessy.beata@ttk.mta.hu
Address: Institute of Enzymology, RCNS, HAS, Magyar Tudósok körútja 2, H-1117 Budapest, Hungary

Keywords: trafficking, dUTPase, importin, cell cycle, phosphorylation

Abstract

Phosphorylation by the cyclin-dependent kinase 1 (Cdk1) adjacent to nuclear localization segments (NLSs) is an important mechanism of regulation of nucleocytoplasmic transport. However, no systematic survey has yet been performed in human cells to analyze this regulatory process, and the corresponding cell-cycle dynamics have not yet been investigated. Here, we focused on the human proteome and found that numerous proteins, previously not identified in this context, are associated with Cdk1-dependent phosphorylation sites adjacent to their NLSs. Interestingly, these proteins are involved in key regulatory events of DNA repair, epigenetics, or RNA editing and splicing. This finding indicates that cell-cycle dependent events of genome editing and gene expression profiling may be controlled by nucleocytoplasmic trafficking. For in-depth investigations, we selected a number of these proteins and analyzed how point mutations, expected to modify the phosphorylation ability of the NLS segments, perturb nucleocytoplasmic localization. In each case, we found that mutations mimicking hyper-phosphorylation abolish nuclear import processes. To understand the mechanism underlying these phenomena, we performed a video microscopy-based kinetic analysis to obtain information on cell-cycle dynamics on a model protein, dUTPase. We show that the NLS-adjacent phosphorylation by Cdk1 of human dUTPase, an enzyme essential for genomic integrity, results in dynamic cell cycle-dependent distribution of the protein. Non-phosphorylatable mutants have drastically altered protein re-import characteristics into the nucleus during the G1 phase. Our results suggest a dynamic Cdk1-driven mechanism of regulation of the nuclear proteome composition during the cell cycle.

Introduction

Eukaryotic cells distribute proteins into various cellular compartments and these intracellular trafficking processes are under multiple levels of control. The transport of large macromolecules into and out of the nucleus depends on karyopherins, a group of proteins specifically recognizing short peptide segments in cargo proteins [1]. Peptide segments involved in these processes are termed nuclear localization signals and nuclear export signals (NLSs and NESs, respectively) [2]. The recognition of NLS and NES sequences by karyopherins can be effectively modulated by introducing negatively charged phosphate groups adjacent to the localization signals via protein phosphorylation. Such modifications have been shown to have a drastic effect on changing the cellular distribution of several proteins[3-5]. The most extensive studies in this respect have been performed in yeast, while the human proteome has not yet been systematically investigated in this context [6].

The best-characterized nuclear import pathway employs the transport factors importin-α (Impα; also known as karyopherin-α), and importin-β (Impβ). Impα recognizes the cargo in the cytoplasm through binding to the classical nuclear localization signals (cNLSs), and the cargo enters the nucleus through nuclear pore complexes as a trimeric Impα:Impβ:cargo complex. The directionality of transport through the nuclear pore complexes is determined by the small GTPase Ran, which has an asymmetric distribution of its nucleotide-bound states between the cytoplasm and the nucleus; it binds to Impβ inside the nucleus in its GTP-bound state, dissociating the import complex and releasing the cargo. Most cNLSs contain either one (monopartite) or two (separated by a linker sequence of usually 10-12 residues; bipartite) clusters of positively charged amino acids. Combining structural studies and interaction data for various cNLSs and their mutants has enabled the molecular understanding of the cNLS:Impα recognition and the definition of cNLS consensus sequences. Impα contains two cNLS-binding regions, the major and minor sites. Bipartite cNLSs span both binding sites, while monopartite cNLSs usually bind preferentially to the major site. Individual amino acids in the cNLS bind to specific pockets in the cNLS-binding sites; in the bipartite cNLS consensus $KRX_{10-12}KRRK$ (X corresponds to any amino acid) [7], the N-terminal basic cluster corresponding to positions P1'–P2' binds to the minor site, while the C-terminal cluster corresponding to positions P2–P5 binds to the major site.

Phosphorylation of proteins is mediated by protein kinases. The specificity of phosphorylation by a particular kinase depends on the composition of residues flanking the phosphorylation site (so-called peptide specificity) [8], although it is further influenced by the context that the kinase finds itself in, including various forms of substrate recruitment [9]. For example, the core consensus sequence for cyclin-dependent kinase 1 (Cdk1)-dependent phosphorylation sites has been described as [S/T*]-P-X-[K/R], where S/T* is the phosphorylated serine or threonine [10-12]. More extensive analyses of known substrates and other available experimental data have uncovered further more subtle determinants of specificity for Cdk1 and other protein kinases [8, 13, 14]

Both NLSs and phosphorylation motifs can be described as linear motifs. Linear motifs are short sequences found most frequently in the disordered regions of proteins, and usually function in cellular signaling and regulation, by binding to protein interaction domains or by being the target of post-translational modifications [15]. Although they pose difficulty for computational analysis because of their small size, significant progress has been made in the recent years in the computational identification of a number of different types of linear motifs and the integration of diverse types of experimental data into these computational approaches [15]. In particular, we have developed some of the most reliable approaches for the identification of NLSs [16] and protein phosphorylation sites [13, 17]. While computational predictions are often hampered by less than desired accuracies, combined prediction of two associated motifs can in fact lead to increased accuracy [18].

The composition of the nuclear proteome defines the availability of the different proteins within the cell nucleus for dedicated functions. The role of cell-cycle dependent nucleocytoplasmic trafficking in regulatory processes of gene expression regulation, DNA damage and repair, and other genome editing pathways have been partially investigated in yeast [6] but have not yet been assessed systematically in mammalian cells. A fundamental difference between the two systems is the closed mitosis of yeast. During closed mitosis the nuclear membrane remains intact and the microtubule-based spindle extends within the nucleus [19]. In case of open mitosis, cells temporally lack their nuclear envelope in M phase. Thus, after mitosis the nuclear proteome has to be reconstituted from proteins that had passively diffused into the cytoplasm (expect for the ones chromatin associated) and become excluded from the nucleus as the newly forming nuclear envelope is initially tightly attached to chromatin [20]. The modulation of the nuclear re-

import of cargoes could give an additional layer of regulation of nuclear proteome composition throughout the cell cycle.

In the present study, we first carried out a computational analysis of the human proteome to identify putative regulatory Cdk1-dependent phosphorylation sites in the vicinity of NLSs in nuclear proteins. We then validated the computational predictions experimentally for a number of proteins, comparing the nucleocytoplasmic localization of hyper-phosphorylation mimicking (hyper-P; with a glutamic acid substitution) and non-phosphorylatable hypo-phosphorylation (hypo-P; with a glutamine substitution) mutants to the phosphorylatable wild-type (WT) protein in a new cellular assay. Finally, we selected one particular protein as an in-depth case study to analyze the dynamics of phosphorylation-regulated nuclear transport during the cell-cycle. We selected human dUTPase, a protein involved in genomic integrity [21], for this case study; where we have previously characterized the molecular and structural basis of NLS-adjacent phosphorylation on nuclear import [22]. Namely, the introduction of negative charge into the P-1 position rearranges the accommodation pattern of the dUTPase NLS in the importin-α NLS binding site. This results in the loss of critical hydrogen bonds between the importin-α surface and the NLS peptide, impairing nuclear import [22]. Here, we show that Cdk1-dependent phosphorylation of dUTPase results in a scheduled dynamic pattern of nuclear availability in the newly-formed daughter cells. Jointly, our results uncover a ubiquitous mechanism for the regulation of nuclear trafficking of human proteins by Cdk1 during the cell cycle and provide a molecular explanation for the negative regulation of nuclear import by NLS-adjacent Cdk1-dependent phosphorylation.

Results/Discussion

The effect of NLS-adjacent phosphorylation on nucleocytoplasmic protein distribution during the cell cycle

Earlier publications identified several yeast [6] and human proteins [3, 4, 23], where NLS-adjacent phosphorylation was shown to inhibit nuclear import. In these cases, phosphorylation took place at either the P0 or the P-1 positions of the NLS, immediately N-terminal to the large basic cluster of the NLS [24]. Yeast Cdc28 or its human orthologue Cdk1 were proposed to be responsible for most of these phosphorylation events, thus giving these proteins a cell cycle-specific localization pattern [6]. Cdc28 and Cdk1 phosphorylate a number of proteins that control critical cell cycle events, including DNA replication and segregation, transcriptional programs and cell morphogenesis [25]. The available results clearly argue for the importance of Cdk1-kinase-regulated nuclear transport for several yeast proteins involved in the regulation of cell cycle progression, DNA replication, DNA damage recognition and repair. As we presently show, similar regulation may occur for human proteins of similar function (Table 1).

We set out to perform a human proteome-wide bioinformatics screen with the aim of identifying human proteins possessing a Cdk1-dependent phosphorylation site at either P0 or P-1 positions of their NLS. We combined two state-of-the-art bioinformatics tools for prediction of NLSs (NucImport [16]) and phosphorylation sites (Predikin [13]) that we ourselves developed previously. NucImport uses a probabilistic (Bayesian network) approach to recognize a variety of NLSs by integrating amino acid sequence and interaction data and predicts the sequence position of the NLS, out-performing other available methods [16]. Predikin uses the concept of specificity-determining residues to predict peptide specificity of protein kinases and identify substrates for protein kinases [14, 17, 26]; the tool outperformed other competing tools in the protein kinase section of the Peptide Recognition Domain specificity prediction category of the 2009 DREAM4 challenge (an independent test using unpublished data) [13]. We first used Predikin [13] to determine how often a Ser or Thr residue was predicted to be phosphorylated, regardless of the import status or presence of NLS. This background frequency of phosphorylation was determined to be 0.136 (c.f. Experimental Procedures). We then used NucImport [16] to predict the location of classical NLSs. For each NLS location, we determined the frequency of (predicted)

phosphorylation of the P0 position to be 0.393 (p = 2.348e-30) and of the P-1 position to be 0.234 (p = 9.530e-05). Hence, both positions are significantly enriched for phosphorylation.

Overall, with a conservative setting of both predictors (see Experimental Procedures for details), we found 92 proteins with a phosphorylation site in the P0 position and 44 proteins with a phosphorylation site in the P-1 position, considering all protein isoforms (if considering only parent proteins, this corresponds to 50 and 22, respectively; Table S1). Among these proteins, there are numerous examples for which, to our knowledge, no previous experimental data have been reported as being relevant to phosphorylation-dependent nuclear translocation. Using Gene Ontology (GO) term annotations, we found proteins involved in DNA damage recognition and repair, gene expression, epigenetics, RNA-editing and several transcription factors (Table 1 and S2). For any of these functions, strict and regulated scheduling of nuclear availability has clear and imminent significance, arguing for the need for further direct experimental study. We therefore selected several identified proteins for experimental validation.

Cellular screen to evaluate NLS function for selected proteins

To efficiently test the effect of phosphorylation on nuclear import for a number of proteins, we designed a sensitive model system. We chose DsRed-monomer labeled β-galactosidase, a well-described bacterial protein, as an inert fluorescent cargo, upon which different NLSs can be attached. The construct is strictly cytoplasmic, unless fused to a functional NLS, such as the well-established SV40 large T-antigen (TAg) NLS (Figure 1A). In order to evaluate phenotypic characteristics of any further NLSs, we set a measure scale for five distinct cellular distribution patterns (Figure 1B). We tested this NLS reporter system using the WT and mutant NLSs of the Swi6 protein (Figure 2A), which has been previously described to be phosphorylated at the P-1 position of its NLS by Cdc28, resulting in the inhibition of its nuclear transport [6, 27-29]. The mutations were introduced in such a way that the negative charge mimicking the phosphorylated residue was introduced either at the P-2, P-1 or P0 positions [24], while the structurally important proline residue was not perturbed (Figure 2B).

The results argue that the exact position of the phosphorylated residue is a crucial determining factor in the localization of NLS-containing cargo. Our results are in good agreement with previous work proposing that phosphorylation at P0 or P-1 positions impedes nuclear import,

while on the other hand, phosphorylation at upstream positions, for example in the P-2 position, have the opposite effect by enhancing nuclear import [30].

Using the validated screening system, we tested a selection of proteins involved in a variety of cellular functions (Table 1 and references therein). Our selections included six proteins with predicted Cdk1 phosphorylation sites at the P0 position, and seven proteins with predicted Cdk1 phosphorylation sites at the P-1 position (Table S3). The resulting localization data in Figure 3 show that in each case, substitution of the appropriate Ser or Thr, predicted to be phosphorylated by Cdk1, by Glu (hyper-P mimicking) consistently leads to significantly weaker nuclear accumulation or even to complete nuclear exclusion, compared to the WT protein. The efficiency of nuclear targeting differed among the different NLSs. Our experimental data further argue for the inhibitory effect of the phosphorylation at the P-1 and P0 positions on nuclear import, and confirms the validity of the *in silico* analysis. Among the hits of the *in silico* screening dUTPase was not present due to the strict settings NucImport. However its NLS was tested in our reporter system, which showed nuclear exclusion upon Glu substitution in the P-1 position in agreement with our previous data on full length protein (Figure S1A) [22]. Our analysis thus identified numerous human proteins potentially sharing a similar Cdk1-driven regulatory pattern (Table 1 and S1). These proteins are involved in crucial cellular functions such as DNA damage recognition and repair, transcriptional regulation, cell cycle control, epigenetics and RNA editing.

For several proteins where NLS-adjacent Cdk1-driven phosphorylation has been reported previously, its actual effects could not be properly deciphered when using only hypo-P mimicking mutants in static or kinetic experiments. Therefore, we checked the effect of hyper-P mimicking mutations at the previously established Cdk1 sites of UNG2 (S14 phosphorylation [31, 32]), UBA1 (S4 phosphorylation [33]) and p53 (S315 and S312 phosphorylation in human and in mouse, respectively [34]) (Figure 4). The NLS segments were cloned into our pGal-DsRed NLS reporter construct and the full length ORFs were fused with DsRed-monomer for localization studies. These phosphorylation sites are predicted to be located in the P-2 position adjacent to their NLSs. As expected, a negative charge introduced at this position did not abolish the nuclear localization of either of these constructs (both with the NLS reporter and the full length proteins). However, if we used mutagenesis to move this negative charge to the P-1 position, the impeding effect on nuclear import is clearly observable in case of UBA1 and UNG2 (Figure 4) with the NLS reporter constructs. In case of the UNG2 it is clearly visible that the nuclear localization is

enhanced in the S14E mutant (in P-2 position), and the nuclear targeting capability of the WT NLS is evident when compared to the NLS impaired mutant, K18N (Figure S1B and C) [35]. Phosphorylation of the P-2 positions might enhance nuclear accumulation of these proteins after mitosis. Possible reason why the full length UNG2 does not show the same localization pattern as the NLS reporter construct is that it has a complex NLS which not exclusively relies on the S^{14}PARKRHA sequence. Perturbation of this segment does not lead to complete nuclear exclusion, other sequences also have a role in proper localization of UNG2 [35]. p53, which harbors a bipartite NLS sequence, might have the flexibility to compensate these negative effects by the additional binding to the minor NLS-binding site.

Video microscopy-based kinetic analyses

To have a better understanding of the effect of the described phosphorylation on the dynamic distribution of proteins throughout the cell cycle, we used dUTPase as a model. This enzyme catalyzes the hydrolysis of dUTP into dUMP and inorganic pyrophosphate [36-38] preventing dUMP incorporation into DNA [39, 40]. dUTPase is an important contributor to genomic integrity from bacteria to human [21, 41-46] and possesses a nuclear isoform in different eukaryotes [41, 47, 48]. We have previously shown that the cell cycle-dependent phosphorylation of dUTPase by Cdk1 at the S11 position (which is located in the P-1 position of its NLS [47, 49]) abolishes its nuclear import and is linked to M phase [23].

In order to follow the dynamic alterations of dUTPase localization pattern during the cell cycle, we followed individual cells after transfection with the appropriate fluorescent constructs by video microscopy. The dUTPase pool exhibits marked cell cycle-dependent dynamic behavior (Figure 5A). When the new nuclear envelope appears, dUTPase is excluded from the nucleus. Following cytokinesis, it takes a considerable time before the nuclear space is again re-populated with WT dUTPase (Video S1 left panel). Interestingly, for the S11Q mutant, the nuclear repopulation dynamics is markedly different (Video S1 right panel). The S11E mutant, by contrast, remains cytoplasmic during the entire cell cycle (Video S2).

Using this approach, we could measure apparent rate constants for nuclear re-accumulation of the WT and the S11Q mutant dUTPases (Figure 5B). The same approach cannot be applied to the S11E mutant, as it never enters the nucleus in our experiments. We found that the major

difference between the WT protein and the S11Q mutant is that the WT protein re-enters the nucleus with a considerable lag. Following the lag phase, the apparent rate constants of nuclear accumulation are identical for the WT protein (k_{obs} = 0.0044 min^{-1} ± 7%) and for the S11Q mutant (k_{obs} = 0.0043 min^{-1} ± 8%). Importantly, the mean total fluorescence of the cells (F_{n+c}) did not change during our observations (Figure 5B inset), indicating the steady-state of the investigated fluorescent protein pool. The apparent single exponential kinetics we observe likely represents the result of multiple undistinguishable transport events. The fact that the wild type protein and the non-phosphorylatable mutant S11Q show different kinetic behaviour is clearly due to the change in their phosphorylatable properties.

In order to directly address the pattern of nucleocytoplasmic trafficking of a given protein pool, we repeated the video-microscopy experiments by transfecting the fluorescent proteins themselves, instead of plasmids that lead to continuous expression. Cells transfected with recombinantly expressed WT and S11Q DsR-DUT proteins show the same dynamic events of dUTPase pool distribution as those in the plasmid transfection experiments (Figure S2 and Videos S3-S4).

To investigate whether the exogenous DsR-DUT constructs (originating either from transfected plasmids or the recombinant protein itself) used in the video-microscopy experiments can be phosphorylated similarly to the endogenous protein, we performed western blot experiments (Figure 5C and Figure S3B). We used a dUTPase-specific antibody generated against the full-length protein (anti-hDUT [44]), in combination with the dUTPase S11-phosphoserine specific antibody (anti-S11P-hDUT [23]). Figure 5C shows that 293T cells transfected with the appropriate fusion protein-encoding plasmids produce a WT DsR-DUT protein pool that can be phosphorylated. The recognition of dUTPase by the anti-S11P-hDUT antibody is observed only if Ser11 can be phosphorylated, providing evidence for the specificity of this antibody. The endogenous dUTPase pool is also visible at a lower molecular mass position. Neither forms of recombinant proteins produced in *E. coli* are recognized by the anti-S11P-hDUT antibody, indicating that they are not phosphorylated at Ser11 (Figure S3B). Within the cells transfected with recombinantly produced DsR-DUT protein itself, however, the cognate phosphorylation event targeting Ser11 can take place. The anti-hDUT antibody recognizes all dUTPase construct forms, as well as the endogenous dUTPase pool, independently of the point mutation or phosphorylation state (Figure S3B).

The P-mimicking mutation leads to the exclusion from the nucleus, resulting in an interesting observation that the nuclei of the daughter cells become populated with dUTPase protein only after a significant delay. These results indicate that the potential of the WT protein to be phosphorylated within the nucleus may have physiological implications manifested in its retarded re-import in the daughter cells. The non-phosphorylatable S11Q construct does not exhibit this behavior. We suggest that this mechanism may also operate for the proteins listed in Table S1, because they contain similar phosphorylatable NLSs. Assuming similar nuclear re-import characteristics as that of dUTPase, Cdk1 kinase-induced phosphorylation at these NLS positions would significantly alter the nuclear proteome re-establishment in the daughter cells after the M phase (Figure 6).

Biological significance of cell-cycle dependent re-shaping of the nuclear proteome

Our studies suggested that Cdk1 kinase-induced phosphorylation of many human proteins potentially alter their localization patterns during the cell cycle. Specifically, the most detailed kinetic analysis performed in our case study using dUTPase indicated that the re-import into the nucleus is delayed significantly if the relevant site close to the NLS segment is phosphorylated. Although the mechanism by which this cell-cycle-dependent localization pattern is governed seems to be general (cf Figure 6), the exact physiological consequences of these effects depend on the actual protein and its role in cellular pathways. Below, we discuss these protein-specific characteristics.

In the case of dUTPase, Cdk1-induced phosphorylation of the protein within the nucleus at the G2/M phase will have a prominent effect on the dUTPase pool localization in daughter cells. Namely, dUTPase nuclear import is hampered until the phosphate moiety is removed, thus nuclear re-population in the daughter cells takes place only after a considerable time delay. dUTPase nuclear accumulation reaches its maximal extent around the S phase and the protein remains strictly nuclear until the onset of mitosis (Figure 5A). Recently, it has been shown that nuclear localization of the *de novo* thymidylate biosynthesis pathway is required for the maintenance of genomic integrity [50]. This is achieved by sumoylation-mediated nuclear transport of the enzymes of the pathway, composed of thymidylate synthase (TYMS), dihydrofolate reductase (DHFR), and serine hydroxymethyltransferase (SHMT1 and SHMT2α) [51, 52]. For all

these enzymes, this partial nuclear translocation takes place at the onset of S phase, and they remain in the nucleus until the G2/M phases, while they are cytoplasmic during G1 phase, enabling *de novo* thymidylate synthesis during DNA replication and repair [53]. dUTPase catalyzes the hydrolysis of dUTP into pyrophosphate and dUMP; ensuring the substrate for TYMS, and also low cellular dUTP/dTTP ratios, thus inhibiting uracil accumulation in the DNA [21]. Here we show that dUTPase nuclear accumulation also reaches its maximal extent during the S phase, similarly to the *de novo* thymidylate biosynthesis enzymes. Because it was suggested that *de novo* thymidylate biosynthesis does not occur in the cytoplasm at rates sufficient to prevent uracil misincorporation into DNA [50], it is reasonable to propose that dUTPase might also be necessary to accompany this enzyme complex into the nucleus for proper genomic DNA maintenance. Partially due to S phase activation of ribonucleotide reductase subunits, regulated by transcriptional and post-transcriptional processes, the dNTP pool in mammalian cells increases 20-fold at this cell cycle stage compared to G1 [54]. Thus scheduled nuclear availability of *de novo* thymidylate biosynthesis enzymes, along with dUTPase and ribonucleotide reductase, may ensure strictly regulated dNTP pool composition for DNA polymerases.

Table 1 provides a list for other human proteins where we found potential Cdk1 sites and suggest that phosphorylation regulates their nuclear import and thus their availability in the nucleus. These proteins are involved in crucial cellular functions such as DNA damage recognition and repair, transcriptional regulation, cell cycle control, epigenetics and RNA editing. Clearly, for proteins involved in such functions, the fine-tuned regulation of nuclear availability is of high significance. For example, cullin-4B plays a role in cell cycle regulation together with cyclin-L2, which is also involved in pre-mRNA splicing, alongside with apoptosis induction and cell-cycle arrest in cancer cells. The ataxia telangiectasia and Rad3-related protein, BRCA1-A complex subunit RAP80 and histone acetyltransferase p300 are key components of DNA damage repair. We also found that many of these proteins may act in an interconnected manner; for example, during DNA damage, the protein kinase ATR phosphorylates the bromodomain adjacent to Zn-finger domain protein 2a, the BRCA1-A complex subunit RAP80, the Ras-responsive element binding protein 1, and the Ser/Arg repetitive matrix protein 2 [55]. For proteins involved in DNA repair (ATR, cullin 4B, BRCA1-A subunit RAP80, transcription factor AP-4), the tight connection between cell-cycle checkpoints and DNA damage recognition and response pathways may be the underlying reason for their scheduled absence or presence within the nucleus [56]. Such regulation of protein subcellular distribution may assist in maintaining the

correct logistics of scheduling and executing different tasks during the cell cycle along with the regulation of mRNA nucleocytoplasmic trafficking [57]. In such a way, the use of crucial metabolites involved in energy and signal transduction can also be correctly distributed among cell division, replication and repair tasks. Interestingly, examination of SNP databases revealed two instances where a given SNP may overwrite the Cdk1 driven regulation, however, no data for either frequency or physiological relevance of these SNPs has been reported (Table S1).

Conclusions

Dynamic exchange of macromolecules between the cytoplasm and the nucleus is regulated by several mechanisms. Here we suggest that nuclear import is significantly delayed for those cellular proteins where a Cdk1 kinase-dependent phosphorylation event occurs during the M phase at a relevant site in the vicinity of the NLS (Figure 6).

Two bioinformatics tools, namely NucImport and Predikin, were used to identify the scope of the hypothesized mechanism and to isolate candidate targets. We established the statistical basis for identifying relevant hits and their functional (gene ontology) associations, in ways not supported by the tools individually. Albeit not implemented as a tool in its own right, our integrated approach, may help to develop further studies that aim to understand how (other) post-translational modifications can dynamically modulate functions of sequence motifs (including localization signals). Our analysis showed that (i) positions P-1 and P0 relative to predicted human NLSs are both significantly enriched for predicted Cdk1 phosphorylation; and (ii) 44 and 92 protein isoforms (with phosphorylation of NLS P-1 and P0 sites, respectively) are associated with a range of functions that require strict and regulated scheduling of nuclear availability.

Our cellular reporter assay confirmed the computational predictions of proteins regulated by NLS-adjacent phosphorylation and showed that Cdk1 phosphorylation at P-1 and P0 positions of human NLSs impedes nuclear import. Although the observed effects may be modulated in the full-length proteins, our results clearly provide proof-of-concept. Namely, we propose that the cell cycle-dependent changes in the nuclear proteome may have an important role in selecting the correct set of proteins to be present in the nucleus during the different stages of the cell cycle. Cdk1 phosphorylation events at the M phase will result in proteins that cannot be re-imported

into the nucleus after cytokinesis, because phosphorylation of the P-1 and P0 positions impedes their binding to importin-α. Therefore, several proteins will only appear in the nucleus of the daughter cells after a significant delay following dephosphorylation or *de novo* protein synthesis. Dephosphorylation in the cytoplasm may require some time and might be under yet another level of regulation, giving cells further plasticity in fine-tuning their nuclear proteome. The regulatory pattern we described may prevent accumulation of proteins within the nucleus that could perturb cellular functions, for example by initiating expression of genes with an incorrect schedule. This regulation might also be further fine-tuned by cytoplasmic anchoring processes, facilitated by phosphorylation events [58]. However, the exact purpose of this regulation might differ for each protein, and should be checked individually in detail.

Cell cycle-dependent changes in the nuclear proteome are of utmost importance in the prompt regulation of cellular events, and protein kinases such as Cdk1 cooperate to control the cell cycle dynamics. After cell mitosis, daughter cells form their own nuclear envelope and begin with a limited set of proteins that remain strictly adherent to the chromosomes during cytokinesis [20]. Organism, like yeast, with closed mitosis rely on active protein transport in every phase of their cell-cycle while cells with open mitosis thus have the unique opportunity of re-setting the protein composition within the nucleus of daughter cells after every division. We conclude that Cdk1-driven phosphorylation at P-1 or P0 positions of the NLSs makes a significant contribution to this re-shaping process of the nuclear proteome after the M phase.

Materials and methods

Computational analyses of NLSs and phosphorylation motifs

The computational tools Predikin [13] and NucImport [16] were used to analyze the human proteome. Protein sequences were obtained from UniProt [59] as the complete human proteome including all known isoforms, as defined by UniProt complete proteome sets (representing a total of 71,809 sequences).

To predict the location of nuclear localization signals, we used NucImport [16]. NucImport predicts the probability of nuclear import, type of classical NLS (as categorized by [60]) and its exact location in any query protein sequence. Apart from sequence properties, the prediction is based on known (human) protein interactions that are retrieved from BioGRID [61]. We refer to the predicted proteome set as those proteins that were assigned a type-1 classical NLS with a probability of 0.95 or greater.

Cdk1 phosphorylation sites were predicted for all potential sites, i.e., all Ser and Thr residues, using Predikin [13]. As we used the Cdk1 matrix to score all potential phosphorylation sites in the human proteome, we obtained the complete distribution of scores associated with Cdk1. Converting these to a cumulative density allowed us to (empirically) determine *p*-values associated with each Predikin score (the *p*-value is the probability of achieving a score at least as high as the one observed).

We looked for enrichment of phosphorylation at the P0 and P-1 sites relative to the (predicted) NLS in each protein in the nuclear proteome by counting, for each NLS site, all potential phosphorylation sites (i.e., all Ser and Thr sites) that do not occur at the NLS site of interest and recorded whether they are above or below a threshold (Predikin *p*-value = 0.1). From these counts, the ratio of phosphorylation sites/potential sites can be calculated for the background, P0 and P-1 positions. We assessed whether observations at P0 and P-1 differed from the background by performing a χ^2 analysis.

Gene Ontology (GO) term enrichment analysis was performed for identified proteins using Fisher's exact test. Specifically, we used all proteins predicted to have a type-1 classical NLS and a predicted phosphorylation site at either P0 or P-1 as a foreground, and all “reviewed” human proteins in UniProtKB as background. We used the Gene Ontology official release of human annotations (as of February 2012). For each biological process GO term, we counted the number of proteins in the foreground set and the background set with this term. The one-tailed Fisher’s exact test establishes the *p*-value of the term: the probability of finding this protein count

or more extreme (greater proportion in the foreground). The *p*-value was corrected for multiple testing (shown as *E*-value). A term is thus assigned a small *E*-value only if proteins annotated with that term occur in the foreground set with a higher prevalence than can be statistically explained by chance (i.e. proteins picked randomly from the background set).

Cell culture and constructs

293T cells were kindly provided by Prof. Yvonne Jones (Cancer Research UK, Oxford). Cells were cultured in DMEM/F12 HAM (Sigma) supplemented with Penicillin–Streptomycin solution (50 µg/ml; Gibco) and 10% FBS (Gibco). dUTPase nuclear isoform (DUT) fused to DsRed-Monomer (DsR-DUT) was described in [23]. DsR-DUT was further cloned into the NdeI/XhoI sites of the vector pET-20b (Novagen) for recombinant protein expression (with oligonucleotides dutpETF and dutpETR). Human tumor protein p53 cDNA was purchased from OriGene (NM_000546.2). p53 was fused to DsRed-Monomer, by cloning it into the XhoI/BamHI sites of a modified pEGFP-C1 vector (Clontech) (with primers p53_F and p53_R), where EGFP was replaced by DsRed-Monomer (within the NheI/XhoI sites of the vector). Ubiquitin-activating enzyme E1 (UBA1) cDNA was purchased from OriGene (NM_003334.2) and the fusion construct was made cloning it into the KpnI/BamH1 sites of the pDsRed-Monomer-N1 vector (Clontech) (with primers UBA1_F and UBA1_R). Human Uracil-DNA glycosylase 2 (UNG2) cDNA was a generous gift of Professor Salvatore Caradonna and was cloned into the XhoI/KpnI sites of the pDsRed-Monomer-N1 vector (with primers UNG2_F and UNG2_R). Site-directed mutagenesis was performed by the QuickChange method (Stratagene). The NLS reporter construct was created by fusing β-galactosidase with DsRed-Monomer (termed pGal-DsRed). β-galactosidase was amplified lacking its start codon from the vector pCAUG (with oligos galN1F and galN1R), and was cloned into the KpnI/EcoRI sites of the vector pDsRed-M-N1, thus generating the vector termed pGal-DsRed. Single-stranded oligonucleotide pairs, listed in Table S4, encoding different NLS peptides were cloned into the NheI/EcoRI sites of the pGal-DsRed vector after annealing. In addition, the vector pHM830 (Addgene plasmid 20702) (AflII/XbaI sites) was also used to generate constructs for the NLSs that showed a strong tendency for aggregation when used in context of the previously described pGal-DsRed construct [62]. Primers used for cloning and mutagenesis were synthesized by Eurofins MWG GmbH and are summarized in TableS4. All constructs were verified by sequencing at Eurofins MWG GmbH.

Fluorescence imaging and analysis of DsRed-tagged constructs

For DNA transfections, Lipofectamine™ LTX (Invitrogen) was used according to the manufacturer's instruction. Briefly, subconfluent cultures of 293T cells grown in 35 mm Petri dishes were incubated with 1-2 µg DNA along with 10 µl LTX reagent in serum-free medium, for 16 hours. Protein transfection was performed according to the manufacturer's protocol using Pro-DeliverIN™ reagent (OZ Biosciences). In brief, 8-10 µg protein and 15 µl transfection reagent was used to deliver DsRed-tagged proteins into the cells for 14-18 hours. Image analysis to quantify relative subcellular localization was performed from single-cell measurements using ImageJ 1.46j (NIH, Bethesda), where the mean nuclear (F_n) and cytoplasmic (F_c) fluorescence ratio ($F_{n/c}$) was measured within each cell. Statistical analysis of the relative subcellular localization changes was carried out by the InStat 3.05 software (GraphPad Software, San Diego California, USA) using the non-parametric Mann-Whitney test. Differences were considered statistically significant at $p<0.05$. Images were either acquired with a Leica DM IL LED Fluo microscope equipped with a Leica DFC345 FX monochrome camera.

Live-cell microscopy and evaluation

Time-lapse recordings were performed on a Zeiss 200M inverted microscope equipped with an AxioCam Mnr camera and controlled by the AxioVision 4.8 software. Cells were cultured in Ibidi dishes and kept at 37°C in a humidified 5% CO_2 atmosphere within custom-made microscope stage incubator (CellMovie). Images were acquired every 5 minutes for at least 24 hours using a 10X magnification objective. After transfection, the cells were washed three times with a serum-containing medium. Time-lapse imaging started one hour after changing the medium. Addition of serum resulted in the flattening of the cells and mitogenic serum factors boosted cell proliferation.

Plasmid transfection experiments. The kinetic treatment of the imaging data addresses the gross kinetics of nuclear dUTPase accumulation and does not aim to carry out a detailed analysis of the underlying processes. The quantification of fluorescence in single cells from each frame was performed using ImageJ 1.46j (NIH, Bethesda), where the mean nuclear (F_n) and cytoplasmic (F_c) fluorescence were measured. Data points represent mean values extracted from 16 cells in triplicates. The time axis was defined relative to the visual observation of cytokinesis i.e. t = 0 at cytokinesis termination. The observed fluorescence intensity increase in the nucleus could be analysed, as the total fluorescence of the cytoplasmic and nuclear compartments (F_{n+c}) was constant during the time period of the analysis. Single exponential kinetics fit well to the rising phase of the nuclear accumulation curves in both the WT and the S11Q mutant cell lines. The

considerable lag in nuclear fluorescence accumulation in the WT cells was not included in the kinetic analysis due to the lack of information on building a comprehensive kinetic model for the whole trafficking process.

Protein transfection experiments. These image sequences were not subjected to densitometric analyses due to lower intensity of the intracellular fluorescent signal as well as to the higher background (Videos S3 and S4). The time elapsed between the onset of cytokinesis and the appearance of fluorescent signal within the nucleus (Figure S2) was determined by careful visual observation. Considerable nuclear accumulation of fluorescent proteins was declared when the fluorescent intensity within the nucleus exceeded that within the cytoplasm. Parallel phase contrast images were used to determine the onset of cell cleavage.

Both DNA and protein transfection-based experiments yielded the same conclusions regarding the dynamic distribution pattern of the WT and S11Q mutant DsR-DUT. This is potentially due to the fact that the DsRed-labeled proteins can only be detected after a considerable time delay following protein translation, partially because of the time required for maturation of DsRed fluorophore and because of the time required for detectable fluorophor accumulation. Furthermore, newly maturing DsRed molecules (which also went through phosphorylation in M phase) might be in steady state with a degradation process. Because of these effects, the DsR-DUT pool translated during the recording time of video-microscopy used for analysis (~12 hours) does not contribute to the fluorescent signal. The observable fluorescent signal of the mature folded protein molecules thus necessarily originates from the protein pool translated during the cell cycle(s) completed prior to start of the video recording.

Immunoblot analysis

Phosphorylation of the constructs after cellular delivery was investigated using immunblot analysis. Cells were collected, washed twice with PBS, and resuspended in the lysis buffer (50 mM TRIS·HCl pH=7.4; 140 mM NaCl; 0,4% NP-40; 2 mM dithiothreitol (DTT); 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride; 5 mM benzamidin, Complete^{TM} EDTA free protease inhibitor cocktail tablet (Roche), PhosSTOP^{TM} phophatase inhibitor cocktail tablet (Roche)). Cell lysis was achieved by sonication. Insoluble fraction was removed by centrifugation (20,000 x *g* x 15 min at 4°C). Protein concentration was measured with Bio-Rad Protein Assay to ensure equivalent total protein load per lane. Products were resolved under denaturing and reducing conditions on a 15% polyacrylamide gel and transferred to PDVF membrane (Immobilon-P, Millipore). Membranes were blocked with 5% nonfat dried milk, incubated with primary antibodies for 2

hours at room temperature. After washing the membranes secondary antibodies coupled with horseradish peroxidase were applied (Amersham Pharmacia Biotech and Sigma). Immunoreactive bands were visualized by enhanced chemiluminescence reagent (Amersham) and recorded on X-ray film (Kodak). Antibodies to detect the following proteins were used in western blotting: anti-hDUT (1:5000) [44], anti-S11P-hDUT (1:200, GenScript).

Recombinant protein production

DsRed-tagged dUTPase constructs were expressed in Rosetta BL21 (DE3) pLysS bacteria strain and purified using Ni-NTA affinity resin (Qiagen). Transformed cells growing in Luria broth medium were induced at A_{600nm}=0.6 with 0.6 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 16 hours at 20ºC. Cells were harvested and lysed in lysis buffer (50 mM TRIS·HCl, pH=8.0, 300 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100, 10 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride; 5 mM benzamidin, Complete™ EDTA free protease inhibitor cocktial tablet (Roche)) with sonication and cell debris was pelleted by centrifugation at 20.000 x *g* for 30 minutes. Supernatant was applied onto a Ni-NTA column and washed with lysis buffer containing 50 mM imidazole. dUTPase was finally eluted with elution buffer (50 mM HEPES, pH=7.5, 30 mM KCl, 500 mM imidazole, 10 mM 2-merchaptoethanol). dUTPase constructs were dialyzed against buffer containing: 20 mM HEPES, pH=7.4, 140 mM NaCl, 1 mM $MgCl_2$ and 2 mM dithiothreitol (DTT). The proteins were >95 % pure as assessed by SDS-PAGE.

Acknowledgements

We are grateful to Professor Salvatore Caradonna for providing the UNG2 cDNA. This work This work was supported by grants from the by the Hungarian Scientific Research Fund (OTKA NK 84008, OTKA K109486), the Baross program of the New Hungary Development Plan (3DSTRUCT, OMFB-00266/2010 REG-KM-09-1- 2009-0050), the Hungarian Academy of Sciences (TTK IF-28/2012), the MedinProt program of the Hungarian Academy of Sciences, and the European Commission FP7 BioStruct-X project (contract No. 283570), to BGV. GR is the recipient of Young Researcher Fellowships from the Hungarian Academy of Sciences. JT is the recipient of the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

1. Gorlich D, Vogel F, Mills AD, Hartmann E, Laskey RA. Distinct functions for the two importin subunits in nuclear protein import. Nature 1995; 377:246-8.
2. Sorokin AV, Kim ER, Ovchinnikov LP. Nucleocytoplasmic transport of proteins. Biochemistry Biokhimiia 2007; 72:1439-57.
3. Jans DA. The regulation of protein transport to the nucleus by phosphorylation. The Biochemical journal 1995; 311 (Pt 3):705-16.
4. Jans DA, Hubner S. Regulation of protein transport to the nucleus: central role of phosphorylation. Physiol Rev 1996; 76:651-85.
5. Nardozzi JD, Lott K, Cingolani G. Phosphorylation meets nuclear import: a review. Cell communication and signaling : CCS 2010; 8:32.
6. Kosugi S, Hasebe M, Tomita M, Yanagawa H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proceedings of the National Academy of Sciences of the United States of America 2009; 106:10171-6.
7. Fontes MR, Teh T, Jans D, Brinkworth RI, Kobe B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. J Biol Chem 2003; 278:27981-7.
8. Kobe B, Kampmann T, Forwood JK, Listwan P, Brinkworth RI. Substrate specificity of protein kinases and computational prediction of substrates. Biochim Biophys Acta 2005; 1754:200-9.
9. Zhu G, Liu Y, Shaw S. Protein kinase specificity. A strategic collaboration between kinase peptide specificity and substrate recruitment. Cell Cycle 2005; 4:52-6.
10. Holmes JK, Solomon MJ. A predictive scale for evaluating cyclin-dependent kinase substrates. A comparison of p34cdc2 and p33cdk2. The Journal of biological chemistry 1996; 271:25240-6.
11. Nigg EA. Cellular substrates of p34(cdc2) and its companion cyclin-dependent kinases. Trends in cell biology 1993; 3:296-301.
12. Songyang Z, Blechner S, Hoagland N, Hoekstra MF, Piwnica-Worms H, Cantley LC. Use of an oriented peptide library to determine the optimal substrates of protein kinases. Current biology : CB 1994; 4:973-82.
13. Ellis JJ, Kobe B. Predicting protein kinase specificity: Predikin update and performance in the DREAM4 challenge. PLoS One 2011; 6:e21169.
14. Brinkworth RI, Breinl RA, Kobe B. Structural basis and prediction of substrate specificity in protein serine/threonine kinases. Proc Natl Acad Sci U S A 2003; 100:74-9.
15. Kobe B, Boden M. Computational modelling of linear motif-mediated protein interactions. Current topics in medicinal chemistry 2012; 12:1553-61.
16. Mehdi AM, Sehgal MS, Kobe B, Bailey TL, Boden M. A probabilistic model of nuclear import of proteins. Bioinformatics 2011; 27:1239-46.
17. Saunders NF, Brinkworth RI, Huber T, Kemp BE, Kobe B. Predikin and PredikinDB: a computational framework for the prediction of protein kinase peptide specificity and an associated database of phosphorylation sites. BMC Bioinformatics 2008; 9:245.
18. Yaffe MB, Leparc GG, Lai J, Obata T, Volinia S, Cantley LC. A motif-based profile scanning approach for genome-wide prediction of signaling pathways. Nature Biotechnol 2001; 19:348-53.
19. Ding R, West RR, Morphew DM, Oakley BR, McIntosh JR. The spindle pole body of Schizosaccharomyces pombe enters and leaves the nuclear envelope as the cell cycle proceeds. Mol Biol Cell 1997; 8:1461-79.
20. Swanson JA, McNeil PL. Nuclear reassembly excludes large macromolecules. Science 1987; 238:548-50.

21. Vertessy BG, Toth J. Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. Acc Chem Res 2009; 42:97-106.
22. Rona G, Marfori M, Borsos M, Scheer I, Takacs E, Toth J, Babos F, Magyar A, Erdei A, Bozoky Z, et al. Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. Acta crystallographica Section D, Biological crystallography 2013; 69:2495-505.
23. Rona G, Marfori M, Borsos M, Scheer I, Takacs E, Toth J, Babos F, Magyar A, Erdei A, Bozoky Z, et al. Phosphorylation in the vicinity of the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. Acta Crystallogr D Biol Crystallogr 2013.
24. Marfori M, Mynott A, Ellis JJ, Mehdi AM, Saunders NF, Curmi PM, Forwood JK, Boden M, Kobe B. Molecular basis for specificity of nuclear import and prediction of nuclear localization. Biochim Biophys Acta 2011; 1813:1562-77.
25. Enserink JM, Kolodner RD. An overview of Cdk1-controlled targets and processes. Cell division 2010; 5:11.
26. Saunders NF, Kobe B. The Predikin webserver: improved prediction of protein kinase peptide specificity using structural information. Nucleic Acids Res 2008; 36:W286-90.
27. Geymonat M, Spanos A, Wells GP, Smerdon SJ, Sedgwick SG. Clb6/Cdc28 and Cdc14 regulate phosphorylation status and cellular localization of Swi6. Mol Cell Biol 2004; 24:2277-85.
28. Harreman MT, Kline TM, Milford HG, Harben MB, Hodel AE, Corbett AH. Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. The Journal of biological chemistry 2004; 279:20613-21.
29. Sidorova JM, Mikesell GE, Breeden LL. Cell cycle-regulated phosphorylation of Swi6 controls its nuclear localization. Mol Biol Cell 1995; 6:1641-58.
30. Kosugi S, Hasebe M, Entani T, Takayama S, Tomita M, Yanagawa H. Design of peptide inhibitors for the importin alpha/beta nuclear import pathway by activity-based profiling. Chem Biol 2008; 15:940-9.
31. Dephoure N, Zhou C, Villen J, Beausoleil SA, Bakalarski CE, Elledge SJ, Gygi SP. A quantitative atlas of mitotic phosphorylation. Proceedings of the National Academy of Sciences of the United States of America 2008; 105:10762-7.
32. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA, et al. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Science signaling 2010; 3:ra3.
33. Stephen AG, Trausch-Azar JS, Handley-Gearhart PM, Ciechanover A, Schwartz AL. Identification of a region within the ubiquitin-activating enzyme required for nuclear targeting and phosphorylation. The Journal of biological chemistry 1997; 272:10895-903.
34. Lee MK, Tong WM, Wang ZQ, Sabapathy K. Serine 312 phosphorylation is dispensable for wild-type p53 functions in vivo. Cell Death Differ 2010; 18:214-21.
35. Otterlei M, Haug T, Nagelhus TA, Slupphaug G, Lindmo T, Krokan HE. Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively. Nucleic acids research 1998; 26:4611-7.
36. Fiser A, Vertessy BG. Altered subunit communication in subfamilies of trimeric dUTPases. Biochemical and biophysical research communications 2000; 279:534-42.
37. Mustafi D, Bekesi A, Vertessy BG, Makinen MW. Catalytic and structural role of the metal ion in dUTP pyrophosphatase. Proceedings of the National Academy of Sciences of the United States of America 2003; 100:5670-5.

38. Vertessy BG, Persson R, Rosengren AM, Zeppezauer M, Nyman PO. Specific derivatization of the active site tyrosine in dUTPase perturbs ligand binding to the active site. Biochemical and biophysical research communications 1996; 219:294-300.
39. Kovari J, Barabas O, Varga B, Bekesi A, Tolgyesi F, Fidy J, Nagy J, Vertessy BG. Methylene substitution at the alpha-beta bridging position within the phosphate chain of dUDP profoundly perturbs ligand accommodation into the dUTPase active site. Proteins 2008; 71:308-19.
40. Nemeth-Pongracz V, Barabas O, Fuxreiter M, Simon I, Pichova I, Rumlova M, Zabranska H, Svergun D, Petoukhov M, Harmat V, et al. Flexible segments modulate co-folding of dUTPase and nucleocapsid proteins. Nucleic acids research 2007; 35:495-505.
41. Bekesi A, Zagyva I, Hunyadi-Gulyas E, Pongracz V, Kovari J, Nagy AO, Erdei A, Medzihradszky KF, Vertessy BG. Developmental regulation of dUTPase in Drosophila melanogaster. The Journal of biological chemistry 2004; 279:22362-70.
42. Kovari J, Barabas O, Takacs E, Bekesi A, Dubrovay Z, Pongracz V, Zagyva I, Imre T, Szabo P, Vertessy BG. Altered active site flexibility and a structural metal-binding site in eukaryotic dUTPase: kinetic characterization, folding, and crystallographic studies of the homotrimeric Drosophila enzyme. The Journal of biological chemistry 2004; 279:17932-44.
43. Lari SU, Chen CY, Vertessy BG, Morre J, Bennett SE. Quantitative determination of uracil residues in Escherichia coli DNA: Contribution of ung, dug, and dut genes to uracil avoidance. DNA repair 2006; 5:1407-20.
44. Merenyi G, Kovari J, Toth J, Takacs E, Zagyva I, Erdei A, Vertessy BG. Cellular response to efficient dUTPase RNAi silencing in stable HeLa cell lines perturbs expression levels of genes involved in thymidylate metabolism. Nucleosides Nucleotides Nucleic Acids 2011; 30:369-90.
45. Muha V, Horvath A, Bekesi A, Pukancsik M, Hodoscsek B, Merenyi G, Rona G, Batki J, Kiss I, Jankovics F, et al. Uracil-containing DNA in Drosophila: stability, stage-specific accumulation, and developmental involvement. PLoS genetics 2012; 8:e1002738.
46. Pecsi I, Hirmondo R, Brown AC, Lopata A, Parish T, Vertessy BG, Toth J. The dUTPase enzyme is essential in Mycobacterium smegmatis. PLoS One 2012; 7:e37461.
47. Ladner RD, Carr SA, Huddleston MJ, McNulty DE, Caradonna SJ. Identification of a consensus cyclin-dependent kinase phosphorylation site unique to the nuclear form of human deoxyuridine triphosphate nucleotidohydrolase. The Journal of biological chemistry 1996; 271:7752-7.
48. Toth J, Varga B, Kovacs M, Malnasi-Csizmadia A, Vertessy BG. Kinetic mechanism of human dUTPase, an essential nucleotide pyrophosphatase enzyme. The Journal of biological chemistry 2007; 282:33572-82.
49. Tinkelenberg BA, Fazzone W, Lynch FJ, Ladner RD. Identification of sequence determinants of human nuclear dUTPase isoform localization. Experimental cell research 2003; 287:39-46.
50. MacFarlane AJ, Anderson DD, Flodby P, Perry CA, Allen RH, Stabler SP, Stover PJ. Nuclear localization of de novo thymidylate biosynthesis pathway is required to prevent uracil accumulation in DNA. The Journal of biological chemistry 2012; 286:44015-22.
51. Anderson DD, Eom JY, Stover PJ. Competition between sumoylation and ubiquitination of serine hydroxymethyltransferase 1 determines its nuclear localization and its accumulation in the nucleus. The Journal of biological chemistry 2012; 287:4790-9.
52. Woeller CF, Anderson DD, Szebenyi DM, Stover PJ. Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. The Journal of biological chemistry 2007; 282:17623-31.
53. Anderson DD, Woeller CF, Chiang EP, Shane B, Stover PJ. Serine hydroxymethyltransferase anchors de novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis. The Journal of biological chemistry 2012; 287:7051-62.

54. Niida H, Shimada M, Murakami H, Nakanishi M. Mechanisms of dNTP supply that play an essential role in maintaining genome integrity in eukaryotic cells. Cancer Sci 2010; 101:2505-9.
55. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, 3rd, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 2007; 316:1160-6.
56. Langerak P, Russell P. Regulatory networks integrating cell cycle control with DNA damage checkpoints and double-strand break repair. Philos Trans R Soc Lond B Biol Sci 2012; 366:3562-71.
57. Chakraborty P, Wang Y, Wei JH, van Deursen J, Yu H, Malureanu L, Dasso M, Forbes DJ, Levy DE, Seemann J, et al. Nucleoporin levels regulate cell cycle progression and phase-specific gene expression. Dev Cell 2008; 15:657-67.
58. Fulcher AJ, Roth DM, Fatima S, Alvisi G, Jans DA. The BRCA-1 binding protein BRAP2 is a novel, negative regulator of nuclear import of viral proteins, dependent on phosphorylation flanking the nuclear localization signal. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 2010; 24:1454-66.
59. Uniprot C. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic acids research 2012; 40:D71-5.
60. Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, Yanagawa H. Six classes of nuclear localization signals specific to different binding grooves of importin alpha. The Journal of biological chemistry 2009; 284:478-85.
61. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. Nucleic acids research 2006; 34:D535-9.
62. Sorg G, Stamminger T. Mapping of nuclear localization signals by simultaneous fusion to green fluorescent protein and to beta-galactosidase. Biotechniques 1999; 26:858-62.
63. Gnad F, Gunawardena J, Mann M. PHOSIDA 2011: the posttranslational modification database. Nucleic acids research 2011; 39:D253-60.
64. Gnad F, Ren S, Cox J, Olsen JV, Macek B, Oroshi M, Mann M. PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. Genome Biol 2007; 8:R250.
65. Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol 2008; 9:616-27.
66. Kim H, Chen J, Yu X. Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. Science 2007; 316:1202-5.
67. Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, Xia B, Livingston DM, Greenberg RA. RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. Science 2007; 316:1198-202.
68. Guerrero-Santoro J, Kapetanaki MG, Hsieh CL, Gorbachinsky I, Levine AS, Rapic-Otrin V. The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. Cancer Res 2008; 68:5014-22.
69. Ku WC, Chiu SK, Chen YJ, Huang HH, Wu WG. Complementary quantitative proteomics reveals that transcription factor AP-4 mediates E-box-dependent complex formation for transcriptional repression of HDM2. Mol Cell Proteomics 2009; 8:2034-50.
70. Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, Wang C, Brindle PK, Dent SY, Ge K. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. EMBO J 2011; 30:249-62.
71. Liu H, Hew HC, Lu ZG, Yamaguchi T, Miki Y, Yoshida K. DNA damage signalling recruits RREB-1 to the p53 tumour suppressor promoter. The Biochemical journal 2009; 422:543-51.

72. Thiagalingam A, De Bustros A, Borges M, Jasti R, Compton D, Diamond L, Mabry M, Ball DW, Baylin SB, Nelkin BD. RREB-1, a novel zinc finger protein, is involved in the differentiation response to Ras in human medullary thyroid carcinomas. Mol Cell Biol 1996; 16:5335-45.
73. Iyer NG, Ozdag H, Caldas C. p300/CBP and cancer. Oncogene 2004; 23:4225-31.
74. Egawa T, Littman DR. Transcription factor AP4 modulates reversible and epigenetic silencing of the Cd4 gene. Proceedings of the National Academy of Sciences of the United States of America 2011; 108:14873-8.
75. Imai K, Okamoto T. Transcriptional repression of human immunodeficiency virus type 1 by AP-4. The Journal of biological chemistry 2006; 281:12495-505.
76. Guetg C, Lienemann P, Sirri V, Grummt I, Hernandez-Verdun D, Hottiger MO, Fussenegger M, Santoro R. The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. EMBO J 2010; 29:2135-46.
77. Strohner R, Nemeth A, Jansa P, Hofmann-Rohrer U, Santoro R, Langst G, Grummt I. NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. EMBO J 2001; 20:4892-900.
78. Jia S, Kobayashi R, Grewal SI. Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. Nat Cell Biol 2005; 7:1007-13.
79. Blencowe BJ, Bauren G, Eldridge AG, Issner R, Nickerson JA, Rosonina E, Sharp PA. The SRm160/300 splicing coactivator subunits. RNA 2000; 6:111-20.
80. Yang L, Li N, Wang C, Yu Y, Yuan L, Zhang M, Cao X. Cyclin L2, a novel RNA polymerase II-associated cyclin, is involved in pre-mRNA splicing and induces apoptosis of human hepatocellular carcinoma cells. The Journal of biological chemistry 2004; 279:11639-48.
81. Zou Y, Mi J, Cui J, Lu D, Zhang X, Guo C, Gao G, Liu Q, Chen B, Shao C, et al. Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression. The Journal of biological chemistry 2009; 284:33320-32.
82. Miranda-Carboni GA, Krum SA, Yee K, Nava M, Deng QE, Pervin S, Collado-Hidalgo A, Galic Z, Zack JA, Nakayama K, et al. A functional link between Wnt signaling and SKP2-independent p27 turnover in mammary tumors. Genes Dev 2008; 22:3121-34.
83. Higa LA, Yang X, Zheng J, Banks D, Wu M, Ghosh P, Sun H, Zhang H. Involvement of CUL4 ubiquitin E3 ligases in regulating CDK inhibitors Dacapo/p27Kip1 and cyclin E degradation. Cell Cycle 2006; 5:71-7.
84. Li HL, Wang TS, Li XY, Li N, Huang DZ, Chen Q, Ba Y. Overexpression of cyclin L2 induces apoptosis and cell-cycle arrest in human lung cancer cells. Chin Med J (Engl) 2007; 120:905-9.
85. Kondo T, Sheets PL, Zopf DA, Aloor HL, Cummins TR, Chan RJ, Hashino E. Tlx3 exerts context-dependent transcriptional regulation and promotes neuronal differentiation from embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America 2008; 105:5780-5.
86. Worman HJ. Nuclear lamins and laminopathies. J Pathol 2011; 226:316-25.
87. Dittmer TA, Misteli T. The lamin protein family. Genome Biol 2011; 12:222.

Figure legends:

Figure 1. Representation of relative NLS activity.

(A) The β-galactosidase-DsRed (pGal-DsRed) reporter construct is strictly cytoplasmic, unless fused to a functional NLS, such as the SV40 large T-antigen (TAg) NLS. Scale bar represents 20 μm.

(B) The reporter construct was fused with various NLSs to generate constructs with different extents of nuclear localization. Localization was categorized into five types: N (completely nuclear), Nc (mainly nuclear), NC (homogenous distribution between the nucleus and cytoplasm), nC (mainly cytoplasmic), C (completely cytoplasmic). Scale bar represents 20 μm.

Figure 2. Position-specific effect of phosphorylation on NLSs.

(A) Performance of the reporter system (pGal-DsRed) after fusing Swi6 WT and mutant NLSs to the construct. $F_{n/c}$ ratios (± standard error of the mean) were determined as described in the Experimental Procedures section. Localization pattern was categorized according to Figure 1B. Scale bar represents 20 μm.

(B) Glutamic acid at P-2 or P0 was introduced by insertion of Ala in P0 or deletion of Leu in P1. Sequences were aligned as predicted to bind to the NLS-binding pockets of importin-α [24].

Figure 3. Evaluation of the proteins identified by computational analysis: cellular screens for NLS activity.

Localization patterns of proteins selected based on proteome-wide analysis. DNA corresponding to NLSs was cloned into the pGal-DsRed reporter system, and localization was tested in 293T cells. P-mimicking mutations at the appropriate Ser/Thr position in most cases significantly reduced nuclear accumulation. Scale bar represents 20 μm.

Figure 4. Effect of phosphorylation on localization of Cdk1 substrate proteins.

UNG2 (residue S14), UBA1 (residue S4) and p53 (residue S315) phosphorylation at the P-2 positions of their NLSs were mimicked by Glu substitution in the pGal-DsRed reporter system or were mutated in full length proteins fused to DsRed-monomer. Localization was tested in 293T cells.. Scale bar represents 20 μm.

Figure 5. Phosphorylation-dependent cellular localization patterns of dUTPase.

(A) Live-cell microscopy of daughter cells. Transfected 293T cells were observed during at least one full cell cycle. Still images were taken from Video S1. The once-nuclear pool gets slowly re-imported into the nucleus.

(B) Kinetic analysis of protein re-import dynamics of the daughter cells indicate similar import kinetics but different lag phases for the WT protein and the S11Q mutant (k_{obs} = 0.0044 min^{-1} ± 7% and 0.0043 min^{-1} ± 8%, respectively).

(C) Western blot shows cognate phosphorylation of exogenous dUTPases.

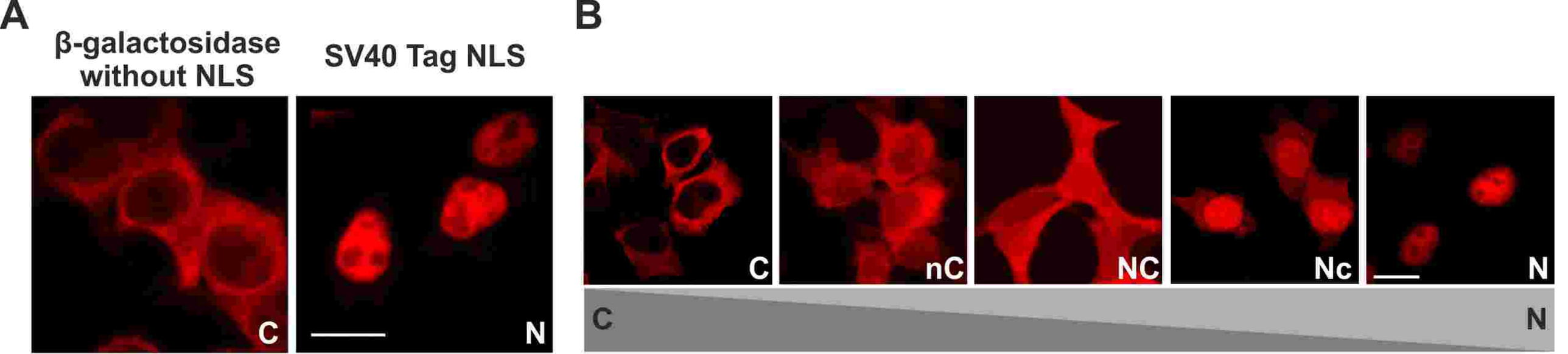
Figure 6. Reconstitution of the nuclear proteome after cell division.

Schematic diagram of the model of nuclear proteome re-setting through regulation of nuclear import by Cdk1 phosphorylation during the cell cycle in human cells.

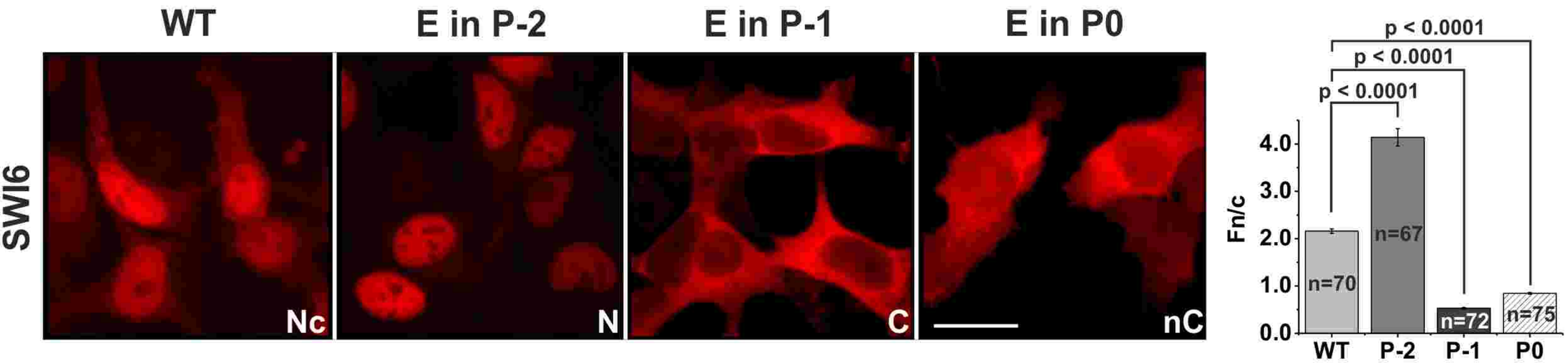
Table 1. Proteins covering different functions within the cell selected for further analysis from the proteome-wide screen

Proteins for which the phosphorylation of the particular NLS adjacent residues were experimentally confirmed, according to derived from the Phosida database (http://www.phosida.com/) [63, 64] are indicated in italics.

Function	Protein name	Abbreviation	NLS sequence	Ref.
DNA damage recognition and repair	*Ataxia telangiectasia and Rad3-related protein*	ATR	**S**PKRRRLS	65
	BRCA1-A complex subunit RAP80	UIMC1	**S**VKRKRRL	66, 67
	Cullin-4B	CUL4B	**TS**AKKRKL	68
	Transcription factor AP-4	TFAP4	**S**PKRRRAE	69
	Histone acetyltransferase p300	EP300	**S**AKRPKLS	70
	Ras-responsive element-binding protein 1	RREB1	**S**PLKRRRL	71
Regulation of gene expression	Ras-responsive element-binding protein 1	RREB1	**S**PLKRRRL	72
	Histone acetyltransferase p300	EP300	**S**AKRPKLS	73
	Transcription factor AP-4	TFAP4	**S**PKRRRAE	69, 74, 75
Epigenetics	Histone acetyltransferase p300	EP300	**S**AKRPKLS	70
	Bromodomain adjacent to zinc finger domain protein 2A	BAZ2A	**S**PSKRRRL	76, 77
	Cullin-4B	CUL4B	**TS**AKKRKL	78
RNA editing/splicing	Ser-Arg repetitive matrix protein 2	SRRM2	**T**PAKRKRR	79
	Cyclin-L2	CCNL2	**S**PKRRKSD	80
Cell cycle regulation	Cullin-4B	CUL4B	**TS**AKKRKL	81-83
	Cyclin-L2	CCNL2	**S**PKRRKSD	80, 84
Development	T-cell leukemia homeobox protein 3	TLX3	**T**PPKRKKP	85
	Transcription factor AP-4	TFAP4	**S**PKRRRAE	69
Nuclear skeleton	*Lamin A*	LMNA	**S**VTKKRKL	86, 87


A
β-galactosidase without NLS
SV40 Tag NLS
C
N
B
C
nC
NC
Nc
N
C
N

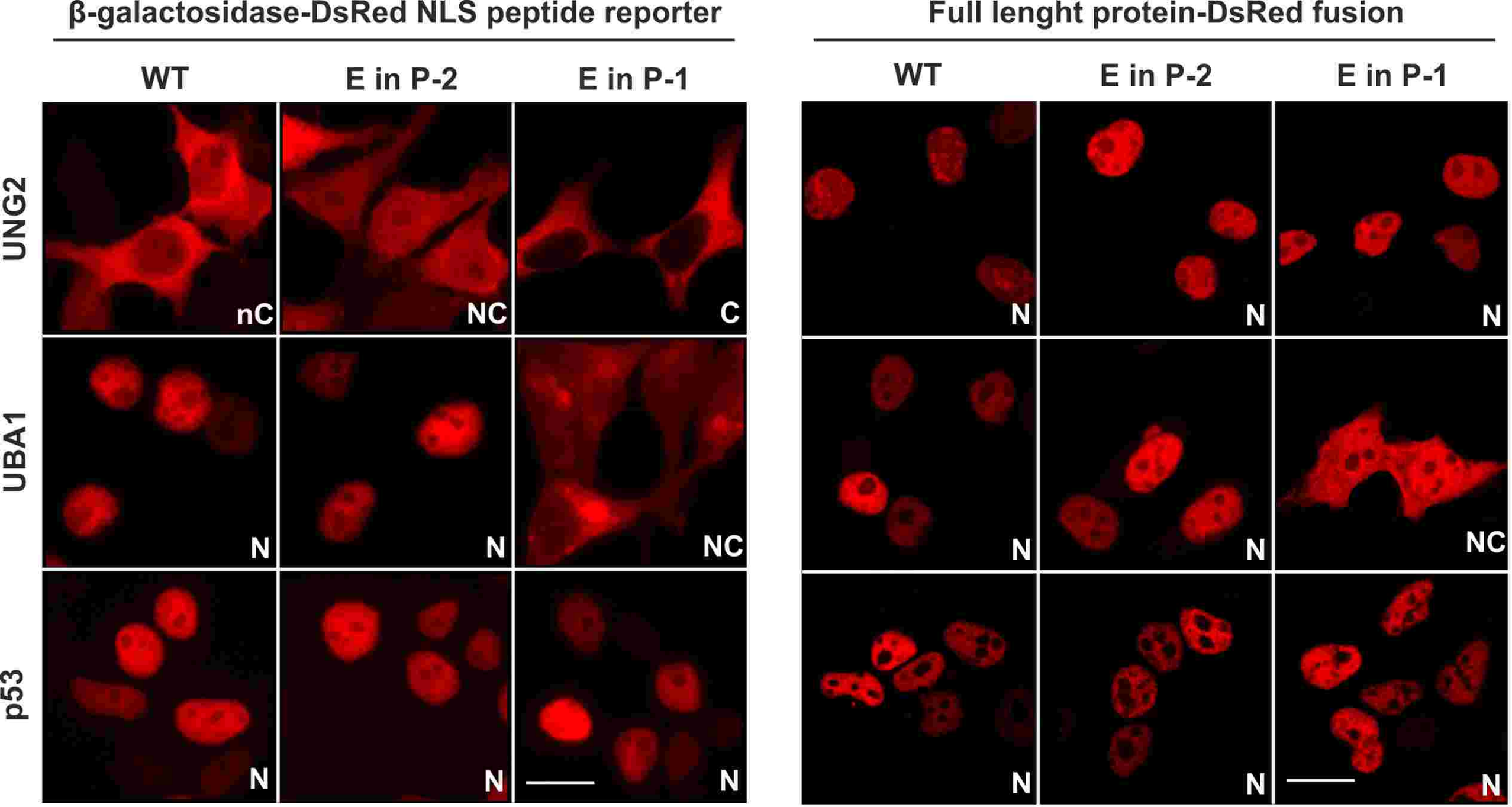
A



B

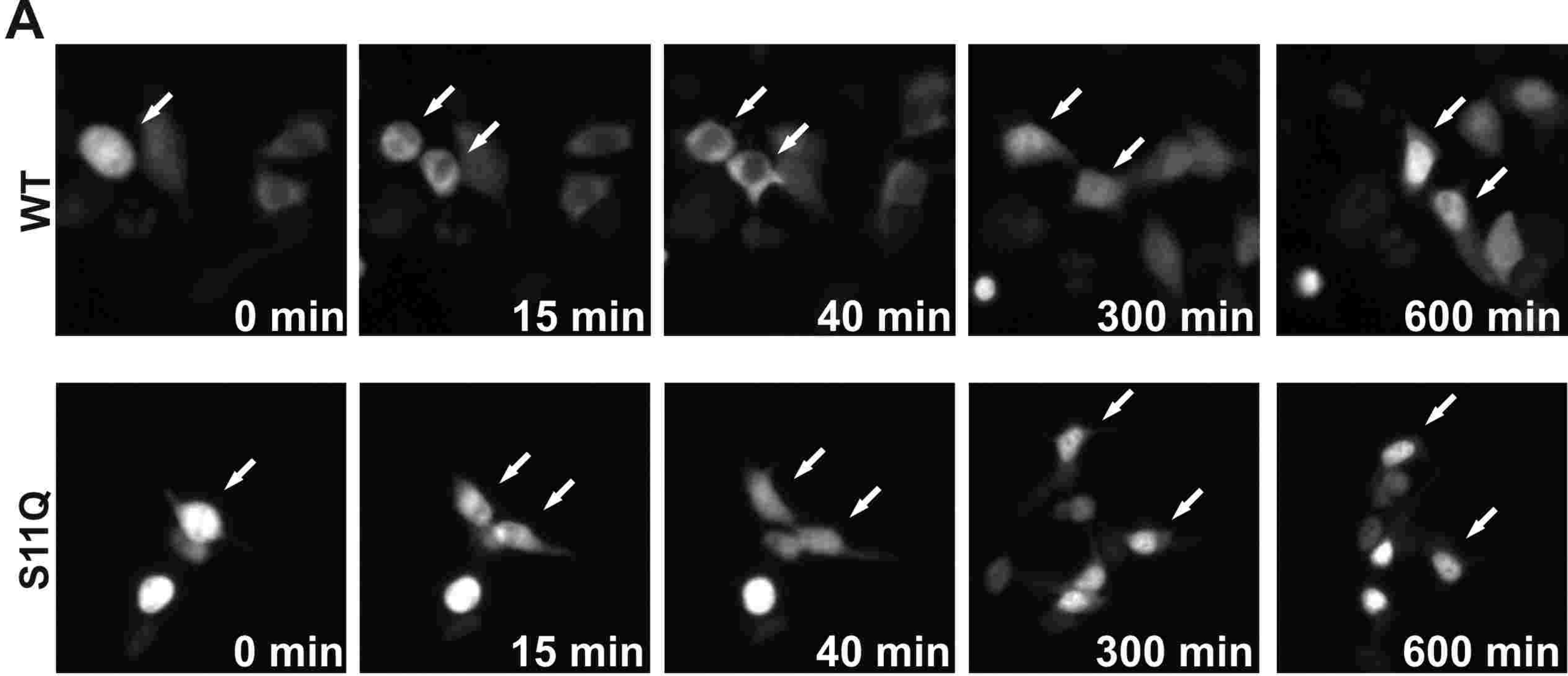
	NLS in major binding pocket of importin-α									Localization
	P-2	**P-1**	**P0**	P1	P2	P3	P4	P5	P6	
Swi6 wt	A	**S**	P	L	K	K	L	K	I	Nc
Swi6 E in P-1	A	**E**	P	L	K	K	L	K	I	C
Swi6 E in P-2	**E**	P	A	L	K	K	L	K	I	N
Swi6 E in P0	G	A	**E**	P	K	K	L	K	I	nC

WT
E in P-1
RREB1
N
NC
p < 0.0001
Fn/c
n=63
n=64
WT
hyper-P
LMNA
Nc
NC
n=81
n=52
BAZ2A
N
NC
n=79
n=53
TLX3
Nc
nC
n=66
n=66
SRRM2
N
NC
n=65
n=65
WT
E in P0
ATR
N
Nc
p < 0.0001
n=67
n=65
UIMC1
NC
C
n=61
n=65
p = 0.0005
CCNL2
Cn
C
n=147
n=110
EP300
N
Nc
n=66
n=178
WT
E in P0
E in P-1
CUL4B
N
N
Nc
p < 0.0001
p = 0.0915
n=62
n=107
n=100
WT
hyper-P0
hyper-P-1
TFAP4
NC
NC
C
p < 0.0001
p = 0.1255
n=65
n=73
n=68

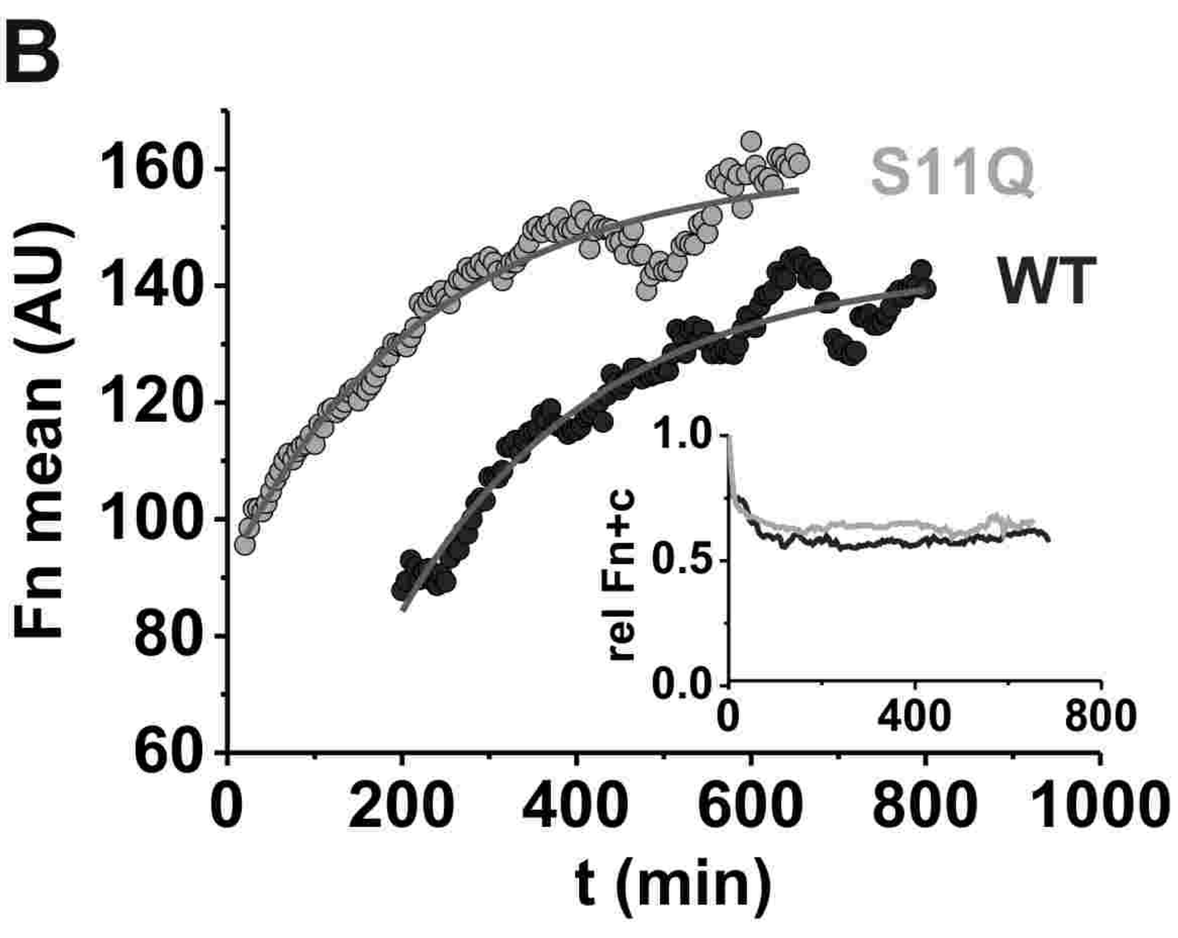


	minor pocket					NLS in major binding pocket of importin-α							Localization
	P1'	P2'	linker	**P-2**	**P-1**	P0	P1	P2	P3	P4	P5	P6	NLS peptide
UNG2 wt				**S**	P	A	R	K	R	H	A	P	nC
UNG2 E in P-2				**E**	P	A	R	K	R	H	A	P	NC
UNG2 E in P-1					**E**	P	R	K	R	H	A	P	C
UBA1 wt				**S**	P	L	S	K	K	R	R	V	N
UBA1 E in P-2				**E**	P	L	S	K	K	R	R	V	N
UBA1 E in P-1					**E**	P	S	K	K	R	R	V	NC
p53 wt	K	R	…	**S**	P	Q	P	K	K	K	P	L	N
p53 E in P-2	K	R	…	**E**	P	Q	P	K	K	K	P	L	N
p53 E in P-1	K	R	…		**E**	P	P	K	K	K	P	L	N

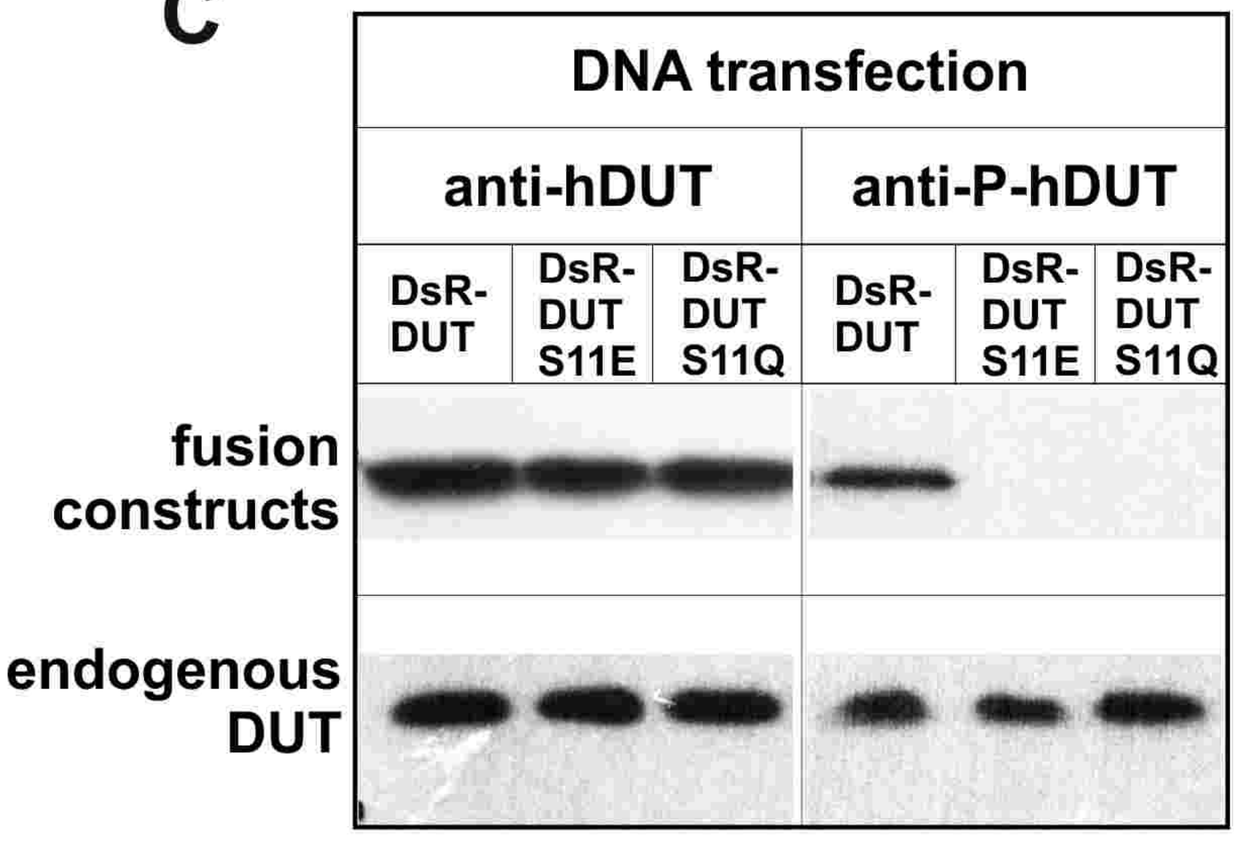
A

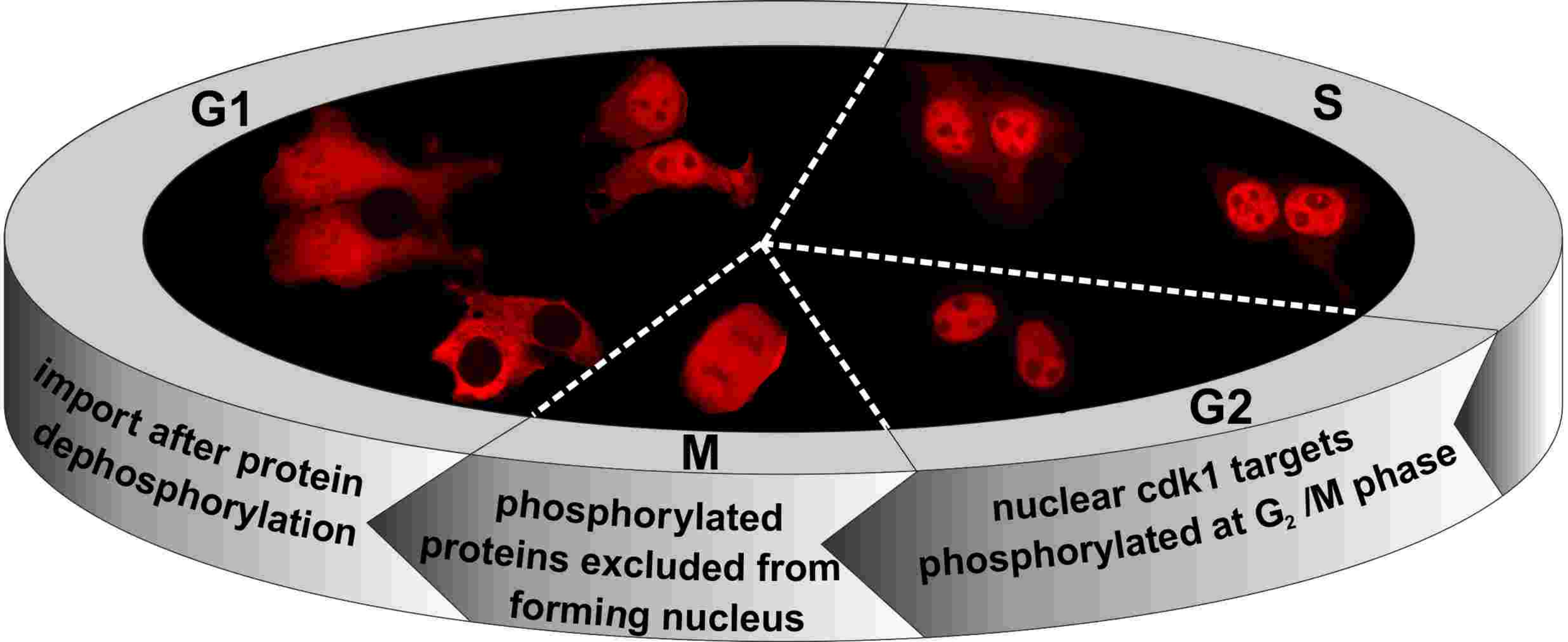

WT
S11Q
0 min
15 min
40 min
300 min
600 min
0 min
15 min
40 min
300 min
600 min

B


Fn mean (AU)
160
140
120
100
80
60
0
200
400
600
800
1000
t (min)
S11Q
WT
rel Fn+c
1.0
0.5
0.0
0
400
800

C


DNA transfection
anti-hDUT
anti-P-hDUT
DsR-DUT
DsR-DUT S11E
DsR-DUT S11Q
DsR-DUT
DsR-DUT S11E
DsR-DUT S11Q
fusion constructs
endogenous DUT


G1
S
G2
M
import after protein dephosphorylation
phosphorylated proteins excluded from forming nucleus
nuclear cdk1 targets phosphorylated at G_2 /M phase

Dynamics of re-constitution of the human nuclear proteome after cell division is regulated by NLS-adjacent phosphorylation

Gergely Róna[1,*], Máté Borsos[1,2], Jonathan J Ellis[3], Ahmed M. Mehdi[3], Mary Christie[3], Zsuzsanna Környei[5], Máté Neubrandt[5], Judit Tóth[1], Zoltán Bozóky[1], László Buday[1], Emília Madarász[5], Mikael Bodén[3,6,7], Bostjan Kobe[3,4,6], Beáta G. Vértessy[1,8,*]

[1] Institute of Enzymology, RCNS, Hungarian Academy of Sciences, H-1117 Budapest, Hungary, [2] Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Dr. Bohr-Gasse 3, 1030 Vienna, Austria, [3] School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane Queensland 4072, Australia, [4] Australian Infectious Diseases Research Centre, The University of Queensland, Brisbane Queensland 4072, Australia, [5] Institute of Experimental Medicine of Hungarian Academy of Sciences, H-1083 Budapest, Hungary, [6] Institute for Molecular Bioscience, The University of Queensland, Brisbane Queensland 4072, Australia, [7] School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane Queensland 4072, Australia, [8] Department of Applied Biotechnology, Budapest University of Technology and Economics, H-1111 Budapest, Hungary

[*]**Corresponding authors:**
Gergely Róna (rona.gergely@ttk.mta.hu) and Beáta G. Vértessy (vertessy.beata@ttk.mta.hu)

Author to communicate with the Editorial and Production offices:
Beáta G. Vértessy, phone: +3613826707, fax: +3614665465, e-mail: vertessy.beata@ttk.mta.hu
Address: Institute of Enzymology, RCNS, HAS, Magyar Tudósok körútja 2, H-1117 Budapest, Hungary

Keywords: trafficking, dUTPase, importin, cell cycle, phosphorylation

Supplemental Data

Figure S1. dUTPase and UNG2 NLSs in the pGal-DsRed reporter assay

(A) dUTPase NLS was cloned into the pGal-DsRed reporter assay. Localization was tested in 293T cells. P-mimicking mutations at the appropriate Ser/Thr position in most cases significantly reduced nuclear accumulation. Scale bar represents 20 μm.

(B) UNG2 NLS was cloned into the pGal-DsRed reporter assay. Localization was tested in 293T cells. Comparision of the WT NLS peptide with the K18N mutated NLS argues for the NLS function of the segment. Scale bar represents 20 μm.

(C) Positioning of the NLS residues within the major NLS-binding site of importin-α. Localization pattern is indicated for each NLS.

Figure S2. Live cell microscopy after protein transfection.

Still images taken from Movies S4 and S5 showing 293T cells expressing WT and S11Q mutant dUTPase during cell division and G1 phase, after protein transfection. The differences observed between in the dynamics of shuttling between the two forms correlate well with those described with the plasmid-transfected cells.

The inserted table shows the time (minutes) elapsed between the onset of cytokinesis and the appearance of considerable fluorescent signal within the nucleus. Data were collected from fluorescent time-lapse image sequences taken in 5 min intervals. Parallel phase contrast images were used to determine the onset of daughter cell separation (cleavage).

Figure S3. S11 phosphorylation status of recombinant dUTPase constructs

(A) After protein transfection both recombinant DsR-DUT and DsR-DUT S11Q localize to the nucleus, as in the case of plasmid transfection. Scale bar represents 20 μm.

(B) Neither forms of recombinant proteins produced in *E. coli* are recognized by the anti-S11P-hDUT antibody, indicating that they are not phosphorylated at Ser11. DsR-DUT can be phosphorylated similarly to the endogenous protein.

Video S1

Localization dynamics of WT DsR-DUT as compared to the non-phosphorylatable mutant S11Q DsR-DUT during the cell cycle in 293T-HEK cells 16 hours after plasmid transfection.

Video S2

Localization dynamics of the hyper-phosphorylation mimick S11E DsR-DUT construct during the cell cycle in 293T-HEK cells 16 hours after plasmid transfection.

Video S3

Localization dynamics of the WT DsR-DUT recombinantly produced construct during the cell cycle in 293T-HEK cells 14-18 hours after protein transfection.

Video S4

Localization dynamics of the S11Q DsR-DUT recombinantly produced construct during the cell cycle in 293T-HEK cells 14-18 hours after protein transfection.

Table S1. Proteins with a predicted type-1 classical NLS and a predicted Cdk1 phosphorylation site at the P0 position or P-1 position.

The table excludes isoforms of the parent protein. Proteins for which the phosphorylation of the particular NLS adjacent residues were experimentally confirmed, according to data derived from the Phosida database (http://www.phosida.com/) are indicated in italics.

Supplemented as an individual .xls file

Table S2. Gene Ontology term enrichment report for proteins with predicted class-1 NLS and Cdk1 phosphorylation site at the P0 or P-1 position.

The table shows the output of a GO enrichment analysis of proteins with predicted NLSs and phosphorylation sites (72 proteins (isoforms excluded) from UniProtKB with 1310 unique biological process terms) relative to the complete human proteome (20246 proteins retrieved from UniProtKB with 10726 unique biological process terms) using Fisher's exact test. The *E*-value is a Bonferroni corrected *p*-value. We only list GO terms with *E*-value < 0.01. For each significant term, we also specify the number of assigned "hit" proteins vs. the "total" number found in the complete background set.

Supplemented as an individual .xls file

Table S3. Predicted binding to the major NLS-binding site of importin-α and nuclear localization of proteins selected by our proteome-wide analysis and their mutants

	NLS binding to major binding site of importin-α									Localization
	P-2	**P-1**	**P0**	P1	P2	P3	P4	P5	P6	
ATR wt[a]	A	I	**S**	P	K	R	R	R	L	N
ATR mut[b]	A	I	**E**	P	K	R	R	R	L	Nc
CCNL2 wt[a]	A	A	**S**	P	K	R	R	K	S	nC
CCNL2 mut[b]	A	A	**E**	P	K	R	R	K	S	C
TFAP4 wt[a]	A	S	**S**	P	K	R	R	R	A	NC
TFAP4 mut[b]	A	S	**E**	P	K	R	R	R	A	NC
CUL4B wt[a]	A	T	**S**	A	K	K	R	K	L	N
CUL4B mut[b]	A	T	**E**	A	K	K	R	K	L	N
EP300 wt[a]	A	P	**S**	A	K	R	P	K	L	N
EP300 mut[b]	A	P	**E**	A	K	R	P	K	L	Nc
UIMC1 wt[a]	A	V	**S**	V	K	R	K	R	R	NC
UIMC1 mut[b]	A	V	**E**	V	K	R	K	R	R	C
RREB1 wt[a]	A	**S**	P	L	K	R	R	R	L	N
RREB1 mut[b]	A	**E**	P	L	K	R	R	R	L	NC
LMNA wt[a]	L	**S**	V	T	K	K	R	K	L	Nc
LMNA mut[b]	L	**E**	V	T	K	K	R	K	L	NC
BAZ2A wt[a]	L	**S**	P	S	K	R	R	R	L	N
BAZ2A mut[b]	L	**E**	P	S	K	R	R	R	L	NC
TFAP4 wt[a]	A	**S**	S	P	K	R	R	R	A	NC
TFAP4 mut[b]	A	**E**	S	P	K	R	R	R	A	C
CUL4B wt[a]	A	**T**	S	A	K	K	R	K	L	N
CUL4B mut[b]	A	**E**	S	A	K	K	R	K	L	Nc
TLX3 wt[a]	A	**T**	P	P	K	R	K	K	P	Nc
TLX3 mut[b]	A	**E**	P	P	K	R	K	K	P	nC
SRRM2 wt[a]	A	**T**	P	A	K	R	K	R	R	Nc
SRRM2 mut[b]	A	**E**	P	A	K	R	K	R	R	NC

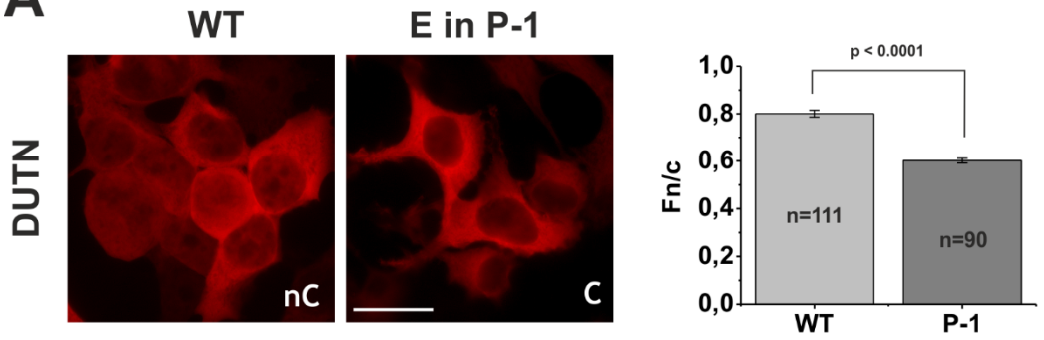
[a] wild-type

[b] hyper-P mimic mutant

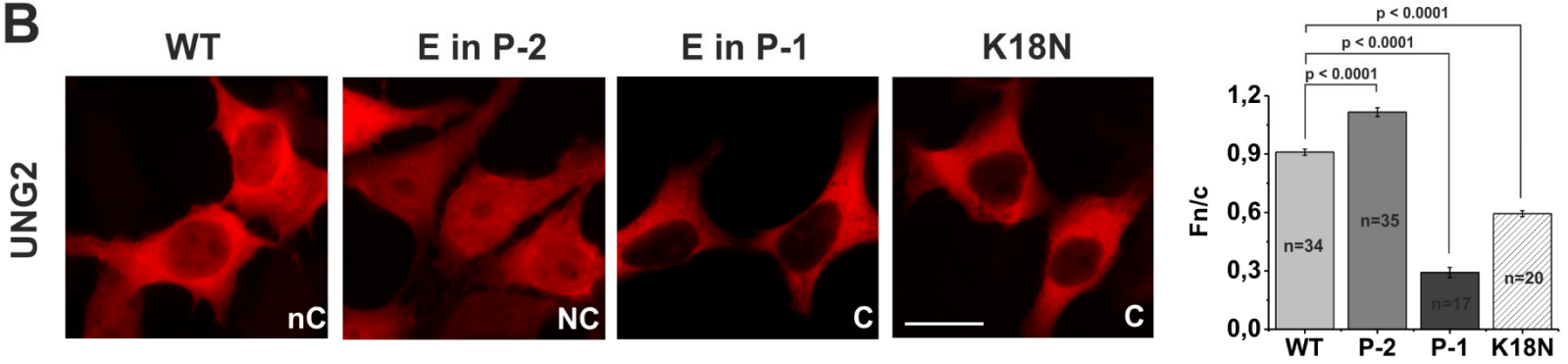
Table S4. Oligonucleotides and peptides used in the study

Supplemented as an individual .xls file

A



B



C

	NLS in major binding pocket of importin-α									Localization
	P-2	**P-1**	P0	P1	P2	P3	P4	P5	P6	
DUT wt		**S**	P	S	K	R	A	R	P	nC
DUT E in P-1		**E**	P	S	K	R	A	R	P	C
UNG2 wt	**S**	P	A	R	K	R	H	A	P	nC
UNG2 E in P-2	**E**	P	A	R	K	R	H	A	P	NC
UNG2 E in P-1		**E**	P	R	K	R	H	A	P	C
UNG K18N	S	P	A	R	**N**	R	H	A	P	C

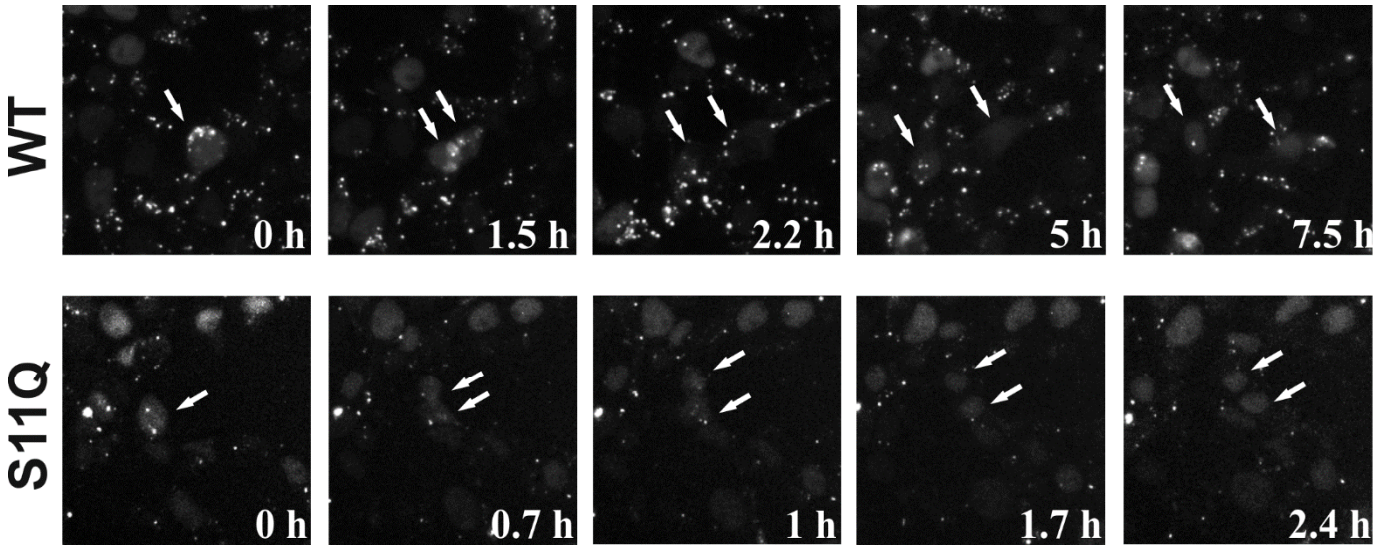
S1. dUTPase and UNG2 NLSs in the pGal-DsRed reporter assay

Pase NLS was cloned into the pGal-DsRed reporter assay. Localization was tested in 293T
g mutations at the appropriate Ser/Thr position in most cases significantly reduced nuclear accu
r represents 20 μm.

2 NLS was cloned into the pGal-DsRed reporter assay. Localization was tested in 293T cells. Co
WT NLS peptide with the K18N mutated NLS argues for the NLS function of the segment. S
ts 20 μm.

tioning of the NLS residues within the major NLS-binding site of importin-α. Localization pattern is
NLS.

2

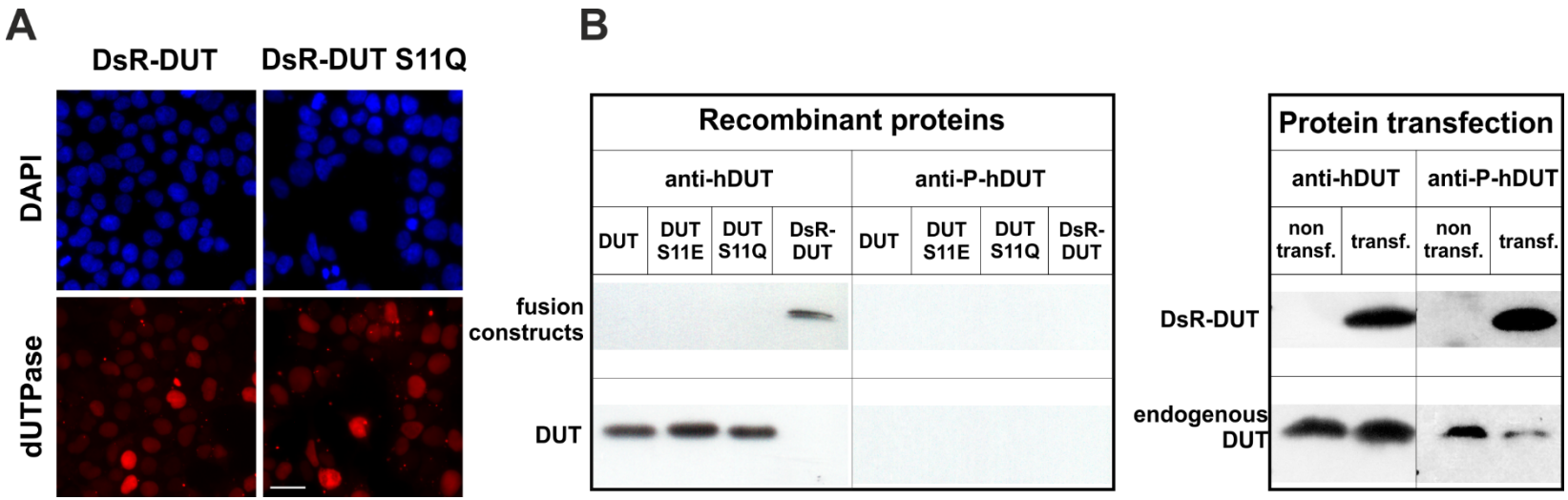


n=15	WT	S11Q
1	370	130
2	360	125
3	240	235
4	305	60
5	235	145
6	230	170
7	285	70
8	335	60
9	255	95
10	240	105
11	245	140
12	370	175
13	230	260
14	305	170
15	225	125
MEAN	282	138
SD	55	58

S2. Live cell microscopy after protein transfection.

ges taken from Movies S4 and S5 showing 293T cells expressing WT and S11Q mutant dUTPas
sion and G1 phase, after protein transfection. The differences observed between in the dynamics of
n the two forms correlate well with those described with the plasmid-transfected cells.

erted table shows the time (minutes) elapsed between the onset of cytokinesis and the appea
rable fluorescent signal within the nucleus. Data were collected from fluorescent time-laps
ces taken in 5 min intervals. Parallel phase contrast images were used to determine the onset of
aration (cleavage).


A
DsR-DUT
DsR-DUT S11Q
DAPI
dUTPase
B
Recombinant proteins
anti-hDUT
anti-P-hDUT
DUT
DUT S11E
DUT S11Q
DsR-DUT
fusion constructs
DUT
Protein transfection
anti-hDUT
anti-P-hDUT
non transf.
transf.
DsR-DUT
endogenous DUT

S3. S11 phosphorylation status of recombinant dUTPase constructs

r protein transfection both recombinant DsR-DUT and DsR-DUT S11Q localize to the nucleus, as in
nid transfection. Scale bar represents 20 μm.

ther forms of recombinant proteins produced in *E. coli* are recognized by the anti-S11P-hDUT a
g that they are not phosphorylated at Ser11. DsR-DUT can be phosphorylated similarly to the end

			P0 position
Uniprot	**Phosphorylation position**	**NLS sequence**	**Protein name**
Q9H9L7	22	**S**PKRRRCA	Akirin-1
Q53H80	21	**S**PKRRRCA	*Akirin-2*
Q15699	129	**S**SKKRRHR	ALX homeobox protein 1
Q13535	428	**S**PKRRRLS	Serine/threonine-protein kinase ATR
Q9NYF8	114	**S**PKRRSVS	Bcl-2-associated transcription factor 1
Q9H0E9	77	**T**PKRKRGE	Bromodomain-containing protein 8
Q96S94	400	**S**PKRRKSD	Cyclin-L2
O14646	1328	**S**SKRRKAR	*Chromodomain-helicase-DNA-binding protein 1*
P49759	27	**S**HKRRKRS	Dual specificity protein kinase CLK1
Q8N684	417	**S**RKRHRSR	*Cleavage and polyadenylation specificity factor*
Q13620	53	**S**AKKRKLN	Cullin-4B
Q09472	12	**S**AKRPKLS	Histone acetyltransferase p300
P17096	53	**T**PKRPRGR	*High mobility group protein HMG-I/HMG-Y*
P52926	44	**S**PKRPRGR	High mobility group protein HMGI-C
Q9NV88	564	**S**GKKRKRV	Integrator complex subunit 9
O43679	254	**T**TKRRKRK	LIM domain-binding protein 2
P43364	16	**S**IKRKKKR	Melanoma-associated antigen 11
Q02078	496	**S**VKRMRMD	Myocyte-specific enhancer factor 2A
Q06413	461	**S**VKRMRLS	Myocyte-specific enhancer factor 2C
Q8N5Y2	334	**T**PKRRKAE	Male-specific lethal 3 homolog
Q659A1	757	**S**KKRKKIR	NMDA receptor-regulated protein 2
O00712	268	**S**SKRPKTI	Nuclear factor 1 B-type
O00567	556	**S**KKKRKFS	Nucleolar protein 56
Q14207	1366	**T**TKKRKIE	Protein NPAT
Q7Z417	220	**T**PKKRKAR	*Nuclear fragile X mental retardation-interacting protein 2*
Q96QT6	240	**S**SKRRRKE	PHD finger protein 12
Q8TF01	551	**S**PKRKKRH	Arginine/serine-rich protein PNISR
Q8NAV1	245	**S**PKRRSPS	Pre-mRNA-splicing factor 38A
Q13523	349	**S**PKRRSLS	*Serine/threonine-protein kinase PRP4 homolog*
Q8NDT2	19	**S**AKRPRER	Putative RNA-binding protein 15B
Q96LT9	115	**S**EKKKRSD	RNA-binding protein 40

Q13127	541	**T**KKKKKVE	RE1-silencing transcription factor
O15446	459	**S**TKKRKKQ	DNA-directed RNA polymerase I subunit RPA34
Q14690	41	**S**TKRKKSQ	*Protein RRP5 homolog*
Q9Y6X0	620	**T**KKRKRRR	SET-binding protein
O95104	442	**S**PKRRRSR	Arginine/serine-rich splicing factor 15
O15042	991	**T**PKRSRRS	U2 snRNP-associated SURP motif-containing protein
Q08170	391	**S**KKKKKED	Serine/arginine-rich splicing factor 4
Q16629	217	**S**PKRSRSP	Serine/arginine-rich splicing factor 7
Q8IWZ8	378	**T**VKRKRKS	SURP and G-patch domain-containing protein 1
P54274	347	**S**TKKKKES	Telomeric repeat-binding factor 1
Q12789	1214	**S**QKRKRLK	General transcription factor 3C polypeptide 1
Q01664	124	**S**PKRRRAE	Transcription factor AP-4
Q15583	162	**S**GKRRRRG	Homeobox protein TGIF1
Q13769	5	**S**SKKRKPK	THO complex subunit 5 homolog
O15405	243	**T**PKKKKKK	TOX high mobility group box family member 3
Q9NPG3	187	**S**PKKRKLK	Ubinuclein-1
Q5T4S7	3366	**S**TKKSKKE	E3 ubiquitin-protein ligase UBR4
Q96RL1	29	**S**VKRKRRL	BRCA1-A complex subunit RAP80
Q96JG9	3786	**S**TKRKKGQ	Zinc finger protein 469

			P-1 positions
Uniprot	**Phosphorylation position**	**NLS sequence**	**Protein name**
Q9UIF9	1783	**S**PSKRRRL	Bromodomain adjacent to zinc finger domain protein 2A
Q9NSI6	905	**S**PPKRRRK	Bromodomain and WD repeat-containing protein
Q8N684	416	**S**SRKRHRS	*Cleavage and polyadenylation specificity factor*
Q13620	52	**T**SAKKRKL	Cullin-4B
O75618	175	**S**QRKRRKS	Death effector domain-containing protein
Q9NPF5	459	**S**SVKKAKK	DNA methyltransferase 1-associated protein 1
Q03001	1382	**S**PVKRRRM	Dystonin
Q92522	31	**S**PSKKRKN	*Histone H1x*
P02545	414	**S**VTKKRKL	*Prelamin-A/C*
Q15788	30	**S**TEKRRRE	Nuclear receptor coactivator 1
Q14207	1365	**S**TTKKRKI	Protein NPAT
O94913	475	**S**TRKRSRS	Pre-mRNA cleavage complex 2 protein Pcf11

Q8TF01	550	**S**SPKRKKR	Arginine/serine-rich protein PNISR
Q8NDT2	18	**S**SAKRPRE	Putative RNA-binding protein 15B
Q13127	540	**S**TKKKKKV	RE1-silencing transcription factor
Q92766	161	**S**PLKRRRL	Ras-responsive element-binding protein 1
Q9UQ35	2599	**T**PAKRKRR	Serine/arginine repetitive matrix protein 2
Q08170	390	**S**SKKKKKE	Serine/arginine-rich splicing factor 4
Q01664	123	**S**SPKRRRA	Transcription factor AP-4
O43763	153	**T**PPKRKKP	T-cell leukemia homeobox protein 2
O43711	162	**T**PPKRKKP	T-cell leukemia homeobox protein 3
Q5T4S7	3365	**S**STKKSKK	E3 ubiquitin-protein ligase UBR4

Gene name	SNPs at P0 position
AKIRIN1 C1orf108	
AKIRIN2 C6orf166	
ALX1 CART1	
ATR FRP1	
BCLAF1 BTF KIAA0164	
BRD8 SMAP SMAP2	
CCNL2 SB138	
CHD1	
CLK1 CLK	
CPSF7	
CUL4B KIAA0695	
EP300 P300	
HMGA1 HMGIY	
HMGA2 HMGIC	
INTS9 RC74	
LDB2 CLIM1	
MAGEA11 MAGE11	
MEF2A MEF2	
MEF2C	
MSL3 MSL3L1	
NARG2 BRCC1 UNQ3101/PRO10100	
NFIB	rs146765479
NOP56 NOL5A	
NPAT CAND3 E14	
NUFIP2 KIAA1321 PIG1	
PHF12 KIAA1523	
PNISR C6orf111 SFRS18 SRRP130 HSPC261 HSPC306	
PRPF38A	
PRPF4B KIAA0536 PRP4 PRP4H PRP4K	
RBM15B OTT3	
RNPC3 KIAA1839 RBM40 RNP	

REST NRSF XBR	
CD3EAP ASE1 CAST PAF49	
PDCD11 KIAA0185	
SETBP1 KIAA0437	
SCAF4 KIAA1172 SFRS15	
U2SURP KIAA0332 SR140	
SRSF4 SFRS4 SRP75	
SRSF7 SFRS7	
SUGP1 SF4	
TERF1 PIN2 TRBF1 TRF TRF1	
GTF3C1	
TFAP4 BHLHC41	
TGIF1 TGIF	
THOC5 C22orf19 KIAA0983	
TOX3 CAGF9 TNRC9	
UBN1	
UBR4 KIAA0462 KIAA1307 RBAF600 ZUBR1	
UIMC1 RAP80 RXRIP110	
ZNF469 KIAA1858	

Gene names	SNPs at P-1 position
BAZ2A KIAA0314 TIP5	
BRWD1 C21orf107 WDR9	
CPSF7	
CUL4B KIAA0695	
DEDD DEDPRO1 DEFT KE05	
DMAP1 KIAA1425	
DST BP230 BP240 BPAG1 DMH DT KIAA0728	
H1FX	
LMNA LMN1	
NCOA1 BHLHE74 SRC1	
NPAT CAND3 E14	
PCF11 KIAA0824	

PNISR C6orf111 SFRS18 SRRP130 HSPC261 HSPC306	
RBM15B OTT3	
REST NRSF XBR	
RREB1 FINB	
SRRM2 KIAA0324 SRL300 SRM300 HSPC075	
SRSF4 SFRS4 SRP75	
TFAP4 BHLHC41	
TLX2 HOX11L1 NCX	
TLX3 HOX11L2	rs139496015
UBR4 KIAA0462 KIAA1307 RBAF600 ZUBR1	

E-value	*p*-value	GO term	Hits/Total (with GO term)	GO term description
1.50E-17	1.40E-21	GO:0090304	39/1924	Nucleic acid metabolic process
8.40E-14	7.83E-18	GO:0006139	39/2453	Nucleobase-containing compound metabolic process
1.50E-13	1.40E-17	GO:0016070	30/1298	RNA metabolic process
7.50E-12	6.99E-16	GO:0045934	24/847	Negative regulation of nucleobase-containing compound metabolic process
9.30E-12	8.67E-16	GO:0051172	24/855	Negative regulation of nitrogen compound metabolic process
1.50E-11	1.40E-15	GO:0010467	30/1538	Gene expression
1.90E-11	1.77E-15	GO:0034641	39/2871	Cellular nitrogen compound metabolic process
3.20E-11	2.98E-15	GO:0051252	39/2916	Regulation of RNA metabolic process
6.20E-11	5.78E-15	GO:0051253	22/741	Negative regulation of RNA metabolic process
7.20E-11	6.71E-15	GO:0006807	39/2986	Nitrogen compound metabolic process
2.30E-10	2.14E-14	GO:0044260	44/4028	Cellular macromolecule metabolic process
3.90E-10	3.64E-14	GO:0006396	20/630	RNA processing
4.80E-10	4.48E-14	GO:0019219	40/3341	Regulation of nucleobase-containing compound metabolic process
9.10E-10	8.48E-14	GO:0010468	39/3221	Regulation of gene expression
1.10E-09	1.03E-13	GO:0051171	40/3422	Regulation of nitrogen compound metabolic process
2.30E-09	2.14E-13	GO:0010605	24/1098	Negative regulation of macromolecule metabolic process
3.40E-09	3.17E-13	GO:2000113	21/802	Negative regulation of cellular macromolecule biosynthetic process
3.70E-09	3.45E-13	GO:0031324	24/1123	Negative regulation of cellular metabolic process
3.70E-09	3.45E-13	GO:0006397	16/386	mRNA processing
6.50E-09	6.06E-13	GO:0010558	21/830	Negative regulation of macromolecule biosynthetic process
2.10E-08	1.96E-12	GO:0009892	24/1217	Negative regulation of metabolic process
2.70E-08	2.52E-12	GO:0008380	14/300	RNA splicing
3.60E-08	3.36E-12	GO:0031327	21/908	Negative regulation of cellular biosynthetic process
4.20E-08	3.92E-12	GO:0060255	41/4013	Regulation of macromolecule metabolic process
4.40E-08	4.10E-12	GO:0045892	19/715	Negative regulation of transcription, DNA-dependent
4.90E-08	4.57E-12	GO:0009890	21/923	Negative regulation of biosynthetic process
7.30E-08	6.81E-12	GO:0006351	16/470	Transcription, DNA-dependent
9.80E-08	9.14E-12	GO:0043170	44/4749	Macromolecule metabolic process
2.10E-07	1.96E-11	GO:0080090	41/4214	Regulation of primary metabolic process
2.50E-07	2.33E-11	GO:0010629	19/790	Negative regulation of gene expression

4.30E-07	4.01E-11	GO:0031323	41/4302	Regulation of cellular metabolic process
7.10E-07	6.62E-11	GO:0006355	33/2829	Regulation of transcription, DNA-dependent
8.40E-07	7.83E-11	GO:2001141	33/2846	Regulation of RNA biosynthetic process
9.10E-07	8.48E-11	GO:0052472	6/21	Modulation by host of symbiont transcription
9.10E-07	8.48E-11	GO:0043921	6/21	Modulation by host of viral transcription
9.20E-07	8.58E-11	GO:0032774	16/557	RNA biosynthetic process
1.10E-06	1.03E-10	GO:2000112	34/3051	Regulation of cellular macromolecule biosynthetic process
1.20E-06	1.12E-10	GO:0052312	6/22	Modulation of transcription in other organism involved in symbiotic interaction
1.80E-06	1.68E-10	GO:0010556	34/3111	Regulation of macromolecule biosynthetic process
2.10E-06	1.96E-10	GO:0045893	19/896	Positive regulation of transcription, DNA-dependent
3.20E-06	2.98E-10	GO:0016071	16/606	mRNA metabolic process
3.50E-06	3.26E-10	GO:0048523	30/2478	Negative regulation of cellular process
5.40E-06	5.03E-10	GO:0051254	19/947	Positive regulation of RNA metabolic process
6.20E-06	5.78E-10	GO:0051851	6/28	Modification by host of symbiont morphology or physiology
6.70E-06	6.25E-10	GO:0010628	19/960	Positive regulation of gene expression
7.20E-06	6.71E-10	GO:0045935	20/1083	Positive regulation of nucleobase-containing compound metabolic process
8.60E-06	8.02E-10	GO:0031326	34/3291	Regulation of cellular biosynthetic process
9.50E-06	8.86E-10	GO:0019222	41/4731	Regulation of metabolic process
1.00E-05	9.32E-10	GO:0043922	5/14	Negative regulation by host of viral transcription
1.00E-05	9.32E-10	GO:0051173	20/1106	Positive regulation of nitrogen compound metabolic process
1.10E-05	1.03E-09	GO:0009889	34/3319	Regulation of biosynthetic process
1.20E-05	1.12E-09	GO:0016568	13/392	Chromatin modification
1.20E-05	1.12E-09	GO:0051702	6/31	Interaction with symbiont
1.70E-05	1.58E-09	GO:0006325	14/489	Chromatin organization
1.90E-05	1.77E-09	GO:0010557	19/1023	Positive regulation of macromolecule biosynthetic process
2.00E-05	1.86E-09	GO:0006366	12/330	Transcription from RNA polymerase II promoter
2.20E-05	2.05E-09	GO:0032897	5/16	Negative regulation of viral transcription
2.30E-05	2.14E-09	GO:0006357	19/1034	Regulation of transcription from RNA polymerase II promoter
3.20E-05	2.98E-09	GO:0048519	30/2715	Negative regulation of biological process
4.30E-05	4.01E-09	GO:0044419	13/436	Interspecies interaction between organisms
1.50E-04	1.40E-08	GO:0006369	6/46	Termination of RNA polymerase II transcription
1.50E-04	1.40E-08	GO:0031328	19/1160	Positive regulation of cellular biosynthetic process
1.70E-04	1.58E-08	GO:0044237	44/5889	Cellular metabolic process

1.90E-04	1.77E-08	GO:0051817	6/48	Modification of morphology or physiology of other organism involved in symbiotic interaction
1.90E-04	1.77E-08	GO:0035821	6/48	Modification of morphology or physiology of other organism
2.00E-04	1.86E-08	GO:0009891	19/1179	Positive regulation of biosynthetic process
3.10E-04	2.89E-08	GO:0000377	9/194	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile
3.10E-04	2.89E-08	GO:0000398	9/194	Nuclear mRNA splicing, via spliceosome
3.10E-04	2.89E-08	GO:0044403	7/89	Symbiosis, encompassing mutualism through parasitism
3.80E-04	3.54E-08	GO:0000375	9/199	RNA splicing, via transesterification reactions
4.90E-04	4.57E-08	GO:0043484	6/56	Regulation of RNA splicing
5.10E-04	4.75E-08	GO:0051276	14/640	Chromosome organization
7.00E-04	6.53E-08	GO:0048024	5/30	Regulation of nuclear mRNA splicing, via spliceosome
7.30E-04	6.81E-08	GO:0034645	19/1281	Cellular macromolecule biosynthetic process
8.00E-04	7.46E-08	GO:0044238	44/6177	Primary metabolic process
8.70E-04	8.11E-08	GO:0016570	9/219	Histone modification
9.70E-04	9.04E-08	GO:0016569	9/222	Covalent chromatin modification
9.80E-04	9.14E-08	GO:0009059	19/1305	Macromolecule biosynthetic process
1.00E-03	9.32E-08	GO:0010604	20/1454	Positive regulation of macromolecule metabolic process
1.80E-03	1.68E-07	GO:0046782	6/69	Regulation of viral transcription
2.30E-03	2.14E-07	GO:0031124	6/72	mRNA 3'-end processing
2.60E-03	2.42E-07	GO:0031325	20/1536	Positive regulation of cellular metabolic process
2.80E-03	2.61E-07	GO:0050684	5/39	Regulation of mRNA processing
3.40E-03	3.17E-07	GO:0048524	6/77	Positive regulation of viral reproduction
5.30E-03	4.94E-07	GO:0006353	6/83	Transcription termination, DNA-dependent
5.60E-03	5.22E-07	GO:0009893	20/1613	Positive regulation of metabolic process
6.60E-03	6.15E-07	GO:0031123	6/86	RNA 3'-end processing
8.90E-03	8.30E-07	GO:2000242	5/49	Negative regulation of reproductive process
9.20E-03	8.58E-07	GO:0090343	3/6	Positive regulation of cell aging

	Analyzed Ser position	Oligo name	Oligonucleotides 5'-3'
	testing NLS function on pGal-DsRed	SV40_WT_F	CTAGCATGGGAGCTTCACCCAAGAAGAAGAGAAAGGTGGG
		SV40_WT_R	AATTCCCACCTTTCTCTTCTTCTTGGGTGAAGCTCCCATG
	S160	Swi6_WT_F	CTAGCATGGGAGCTTCACCCCTGAAGAAGCTGAAGATCGACGG
		Swi6_WT_R	AATTCCGTCGATCTTCAGCTTCTTCAGGGGTGAAGCTCCCATG
		Swi6_P-1E_F	CTAGCATGGGAGCTGAACCCCTGAAGAAGCTGAAGATCGACGG
		Swi6_P-1E_R	AATTCCGTCGATCTTCAGCTTCTTCAGGGGTTCAGCTCCCATG
		Swi6_P-2E_F	CTAGCATGGGAGCTGAACCCGCCCTGAAGAAGCTGAAGATCGACGG
		Swi6_P-2E_R	AATTCCGTCGATCTTCAGCTTCTTCAGGGCGGGTTCAGCTCCCATG
		Swi6_P0_F	CTAGCATGGGAGCTGAACCCAAGAAGCTGAAGATCGACGG
		Swi6_P0_R	AATTCCGTCGATCTTCAGCTTCTTGGGTTCAGCTCCCATG
	S428	ATR_WT_F	CTAGCATGGGAGCTATCAGCCCCAAGAGAAGAAGACTGGG
		ATR_WT_R	AATTCCCAGTCTTCTTCTCTTGGGGCTGATAGCTCCCATG
		ATR_E_F	CTAGCATGGGAGCTATCGAACCCAAGAGAAGAAGACTGGG
		ATR_E_R	AATTCCCAGTCTTCTTCTCTTGGGTTCGATAGCTCCCATG
	S400	CCLN2_WT_F	CTAGCATGGGAGCTGCCAGCCCCAAGAGAAGAAAGAGCGG
		CCLN2_WT_R	AATTCCGCTCTTTCTTCTCTTGGGGCTGGCAGCTCCCATG
		CCLN2_E_F	CTAGCATGGGAGCTGCCGAACCCAAGAGAAGAAAGAGCGG
		CCLN2_E_R	AATTCCGCTCTTTCTTCTCTTGGGTTCGGCAGCTCCCATG
	S124 (P0) S123 (P-1)	TFAP4_WT_F	CTAGCATGGGAGCTAGCAGCCCCAAGAGAAGAAGAGCCGG
		TFAP4_WT_R	AATTCCGGCTCTTCTTCTCTTGGGGCTGCTAGCTCCCATG
		TFAP4_EP0_F	CTAGCATGGGAGCTAGCGAGCCCAAGAGAAGAAGAGCCGG
		TFAP4_EP0_R	AATTCCGGCTCTTCTTCTCTTGGGCTCGCTAGCTCCCATG
		TFAP4_EP-1_F	CTAGCATGGGAGCTGAGAGCCCCAAGAGAAGAAGAGCCGG
		TFAP4_EP-1_R	AATTCCGGCTCTTCTTCTCTTGGGGCTCTCAGCTCCCATG
	S53 (P0) T52 (P-1)	CUL4B_WT_F	CTAGCATGGGAGCTACCAGCGCCAAGAAGAGAAAGCTGGG
		CUL4B_WT_R	AATTCCCAGCTTTCTCTTCTTGGCGCTGGTAGCTCCCATG
		CUL4B_EP0_F	CTAGCATGGGAGCTACCGAGGCCAAGAAGAGAAAGCTGGG
		CUL4B_EP0_R	AATTCCCAGCTTTCTCTTCTTGGCCTCGGTAGCTCCCATG
		CUL4B_EP-1_F	CTAGCATGGGAGCTGAGAGCGCCAAGAAGAGAAAGCTGGG
		CUL4B_EP-1_R	AATTCCCAGCTTTCTCTTCTTGGCGCTCTCAGCTCCCATG
er		EP300_WT_F	CTAGCATGGGAGCTCCCAGCGCCAAGAGACCCAAGCTGGG

Cloning into pGal-DsRed reporte

S12	EP300_WT_R	AATTCCCAGCTTGGGTCTCTTGGCGCTGGGAGCTCCCATG
	EP300_E_F	CTAGCATGGGAGCTCCCGAAGCCAAGAGACCCAAGCTGGG
	EP300_E_R	AATTCCCAGCTTGGGTCTCTTGGCTTCGGGAGCTCCCATG
S29	UIMC1_WT_F	CTAGCATGGGAGCTGTGAGCGTGAAGAGAAAGAGAAGAGG
	UIMC1_WT_R	AATTCCTCTTCTCTTTCTCTTCACGCTCACAGCTCCCATG
	UIMC1_E_F	CTAGCATGGGAGCTGTGGAAGTGAAGAGAAAGAGAAGAGG
	UIMC1_E_R	AATTCCTCTTCTCTTTCTCTTCACTTCCACAGCTCCCATG
S161	RREB1_WT_F	CTAGCATGGGAGCTAGCCCCCTGAAGAGAAGAAGACTGGG
	RREB1_WT_R	AATTCCCAGTCTTCTTCTCTTCAGGGGGCTAGCTCCCATG
	RREB1_E_F	CTAGCATGGGAGCTGAACCCCTGAAGAGAAGAAGACTGGG
	RREB1_E_R	AATTCCCAGTCTTCTTCTCTTCAGGGGTTCAGCTCCCATG
S162	TLX3_WT_F	CTAGCATGGGAGCTACCCCCCCCAAGAGAAAGAAGCCCGG
	TLX3_WT_R	AATTCCGGGCTTCTTTCTCTTGGGGGGGGTAGCTCCCATG
	TLX3_E_F	CTAGCATGGGAGCTGAGCCCCCCAAGAGAAAGAAGCCCGG
	TLX3_E_R	AATTCCGGGCTTCTTTCTCTTGGGGGGCTCAGCTCCCATG
S2599	SRRM2_WT_F	CTAGCATGGGAGCTACCCCCGCCAAGAGAAAGAGAAGAGG
	SRRM2_WT_R	AATTCCTCTTCTCTTTCTCTTGGCGGGGGTAGCTCCCATG
	SRRM2_E_F	CTAGCATGGGAGCTGAGCCCGCCAAGAGAAAGAGAAGAGG
	SRRM2_E_R	AATTCCTCTTCTCTTTCTCTTGGCGGGCTCAGCTCCCATG
S4	UBA E1_WT_F	CTAGCATGGGAGCTTCGCCGCTGTCCAAGAAACGTCGCGTGGG
	UBA E1_WT_R	AATTCCCACGCGACGTTTCTTGGACAGCGGCGAAGCTCCCATG
	UBA E1_EP-2_F	CTAGCATGGGAGCTGAACCGCTGTCCAAGAAACGTCGCGTGGG
	UBA E1_EP-2_R	AATTCCCACGCGACGTTTCTTGGACAGCGGTTCAGCTCCCATG
	UBA E1_EP-1_F	CTAGCATGGGAGCTGAACCGTCCAAGAAACGTCGCGTGGG
	UBA E1_EP-1_R	AATTCCCACGCGACGTTTCTTGGACGGTTCAGCTCCCATG
S14	UNG_WT_F	CTAGCATGGGAGCTTCACCCGCCAGGAAGCGACACGCCCCCGG
	UNG_WT_R	AATTCCGGGGGCGTGTCGCTTCCTGGCGGGTGAAGCTCCCATG
	UNG_EP-2_F	CTAGCATGGGAGCTGAACCCGCCAGGAAGCGACACGCCCCCGG
	UNG_EP-2_R	AATTCCGGGGGCGTGTCGCTTCCTGGCGGGTTCAGCTCCCATG
	UNG_EP-1_F	CTAGCATGGGAGCTGAACCCAGGAAGCGACACGCCCCCGG
	UNG_EP-1_R	AATTCCGGGGGCGTGTCGCTTCCTGGGTTCAGCTCCCATG
K18N	K18N_F	CTAGCATGGGAGCTAGCCCCGCCAGGAACCGACACGCCCCCGG
	K18N_R	AATTCCGGGGGCGTGTCGGTTCCTGGCGGGGCTAGCTCCCATG

	S315	p53_WT_F	CTAGCATGGGAGCTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGGG
		p53_WT_E	AATTCCCAGTGGTTTCTTCTTTGGCTGGGGAGAGGAGCTGGTGTTGTTGGGCAGTGCTCGCTTAGCTCCCATG
		p53_EP-2_F	CTAGCATGGGAGCTAAGCGAGCACTGCCCAACAACACCAGCTCCGAACCCCAGCCAAAGAAGAAACCACTGGG
		p53_EP-2_R	AATTCCCAGTGGTTTCTTCTTTGGCTGGGGTTCGGAGCTGGTGTTGTTGGGCAGTGCTCGCTTAGCTCCCATG
		P53_EP-1_F	CTAGCATGGGAGCTAAGCGAGCACTGCCCAACAACACCAGCTCCGAACCCCCAAAGAAGAAACCACTGGG
		P53_EP-1_R	AATTCCCAGTGGTTTCTTCTTTGGGGGTTCGGAGCTGGTGTTGTTGGGCAGTGCTCGCTTAGCTCCCATG
Cloning into pHM830 reporter	S11	DUT_WT_F	TTAAGGGCCTTGCCTCACCCAGTAAGCGGGCCCGGCCTGCGT
		DUT_WT_R	CTAGACGCAGGCCGGGCCCGCTTACTGGGTGAGGCAAGGCCC
		DUT_E_F	TTAAGGGCCTTGCCGAACCCAGTAAGCGGGCCCGGCCTGCGT
		DUT_E_F	CTAGACGCAGGCCGGGCCCGCTTACTGGGTTCGGCAAGGCCC
	S414	LMNA_WT_F	TTAAGGGCCTTGCCAGCGTGACCAAGAAGAGAAAGCTGT
		LMNA_WT_R	CTAGACAGCTTTCTCTTCTTGGTCACGCTGGCAAGGCCC
		LMNA_E_F	TTAAGGGCCTTGCCGAAGTGACCAAGAAGAGAAAGCTGT
		LMNA_E_R	CTAGACAGCTTTCTCTTCTTGGTCACTTCGGCAAGGCCC
	S1783	BAZ2A_WT_F	TTAAGGGCCTTGCCAGCCCCAGCAAGAGAAGAAGACTGT
		BAZ2A_WT_R	CTAGACAGTCTTCTTCTCTTGCTGGGGCTGGCAAGGCCC
		BAZ2A_E_F	TTAAGGGCCTTGCCGAACCCAGCAAGAGAAGAAGACTGT
		BAZ2A_WT_R	CTAGACAGTCTTCTTCTCTTGCTGGGTTCGGCAAGGCCC
Cloning β-gal into pDsRed-Monomer-N1		galN1F	GCAAGAATTCCAGCATCGTTTACTTTGACCAACAAGAACG
		galN1R	AATTGGTACCGCTTTTTGACACCAGACCAACTGGTAATGG
Cloning dUTPase into pET20b		dutpETF	GGAATTCCATATGCCCTGCTCTGAAGAGACAC
		dutpETR	CCGCTCGAGCTGGGAGCCGGAGTGG
Cloning p53 into pDsRed-Monomer-C1		p53_F	AGCTCTCGAGATGGAGGAGCCGCAGTCAGATC
		p53_R	CACTGGATCCTCAGTCTGAGTCAGGCCCTTCTG
Cloning UBA1 into pDsRed-Monomer-N1		UBA1_F	CAGCTGGTACCATGTCCAGCTCGCCGCTGTCC
		UBA1_R	AGATGGATCCGCTCCGCGGATGGTGTATCGGACATAGG
Cloning UNG2 into pDsRed-Monomer-N1		UNG2_F	CGATCTCGAGATGATCGGCCAGAAGACGC
		UNG2_R	ACGGTACCGCGATGTACCTGTAGGTGTCCAGC
	dUTPase (in pET20b)	S11E_F	GAGACACCCGCCATTGAACCCAGTAAGCGGGC
		S11E_R	GCCCGCTTACTGGGTTCAATGGCGGGTGTCTC
		S11Q_F	GAGACACCCGCCATTCAACCCAGTAAGCGGGC
		S11Q_R	GCCCGCTTACTGGGTTGAATGGCGGGTGTCTC
		S315E_F (P-2)	CCCAACAACACCAGCTCCGAACCCCAGCCAAAGAAGAAACC

Mutagenesis oligos	**p53**	S315E_R (P-2)	GGTTTCTTCTTTGGCTGGGGTTCGGAGCTGGTGTTGTTGGG
		Q317Δ_F (P-1)	ACACCAGCTCCGAACCCCCAAAGAAGAAACCACTGGATG
		Q317Δ_F (P-1)	CATCCAGTGGTTTCTTCTTTGGGGGTTCGGAGCTGGTGT
	UBA1	S4E_F (P-2)	CGACGGTACCATGTCCAGCGAACCGCTGTCCAAGAAACGTC
		S4E_R (P-2)	GACGTTTCTTGGACAGCGGTTCGCTGGACATGGTACCGTCG
		L6Δ_F (P-1)	GTACCATGTCCAGCGAACCGTCCAAGAAACGTCGCGTGTC
		L6Δ_R (P-1)	GACACGCGACGTTTCTTGGACGGTTCGCTGGACATGGTAC
	UNG2	S14E_F (P-2)	CTACTCCTTTTTCTCCCCCGAACCCGCCAGGAAGCGAC
		S14E_R (P-2)	GTCGCTTCCTGGCGGGTTCGGGGGAGAAAAAGGAGTAG
		A16Δ_F (P-1)	CCTTTTTCTCCCCCGAACCCAGGAAGCGACACGCCCC
		A16Δ_R (P-1)	GGGGCGTGTCGCTTCCTGGGTTCGGGGGAGAAAAAGG