www.sciencemag.org/cgi/content/full/1160165/DC1

Supporting Online Material for

Dynamic Proteomics of Individual Cancer Cells in Response to a Drug

A. A. Cohen,* N. Geva-Zatorsky, E. Eden, M. Frenkel-Morgenstern, I. Issaeva, A. Sigal, R. Milo, C. Cohen-Saidon, Y. Liron, Z. Kam, L. Cohen, T. Danon, N. Perzov, U. Alon

*To whom correspondence should be addressed. E-mail: ariel.cohen@weizmann.ac.il

Published 20 November 2008 on *Science* Express
DOI: 10.1126/science.1160165

This PDF file includes

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/1160165/DC1)

Movies S1 to S4

Materials and methods

1. Construction of tagged proteins library

A library of fluorescently tagged proteins was constructed in non-small cell lung carcinoma cell line (H1299) in a two stage process. In both stages a fluorescent reporter was integrated into the genome via Central Dogma tagging (CD-tagging) (*1-4*).

The first stage was carried out in order to produce a parental clone in which the nucleus is colored brighter than the cytoplasm and the cytoplasm is colored brighter than the background of the image. To achieve this, a red fluorescent protein, mCherry (*5*), was introduced in two rounds of CD-tagging. In the first round, clone H7a, expressing tagged protein XRCC5 localized to the nucleus, was selected. In the second round (carried out on the previously selected clone H7a), clone H7 expressing tagged DAP1 localized to the whole intracellular domain was selected. Following these two steps, a parental clone expressing two mCherry endogenously tagged proteins (XRCC5 and DAP1), stained in the cytoplasm and brighter in the nucleus was obtained. We find that the cell-cell variability of the red fluorescence in this clone is much smaller than the variability in clones generated in preliminary experiments by transfection with mCherry. This reduced variability was crucial for reliable image analysis of cell movies.

The second stage in the generation of the library was to use CD-tagging in order to tag different proteins with a second color eYFP or Venus (*6*) within the parental clone H7 (H1299-cherry). The CD-tagging protocol used in all rounds of library construction is described in detail by Sigal *et al.* (*4*). Briefly, a fluorescent protein (FP), flanked by splice acceptor and donor sequences was integrated into the genome as an artificial exon via retroviral vectors (U5000, U5001, U5002), each containing FP in one of 3 reading frames. Cells positive for relevant FP fluorescence were sorted using flow cytometry into 384 well plates and expanded into cell clones. Tagged protein identities were determined by 3'RACE, using a nested PCR reaction that amplified the section between the FP and the polyA tail of the mRNA of the host gene. The PCR product was sequenced directly and aligned to the genome. Our library of CD-tagged proteins is detailed at www.dynamicproteomics.net.

2. Long period time-lapse microscopy

Time-lapse movies were obtained (at 20x magnification) as described by Sigal *et al*. (*7*) with an automated, incubated (including humidity and CO2 control) Leica DMIRE2 inverted fluorescence microscope and an ORCA ER cooled CCD camera (Hamamatsu Photonics). The system was controlled by ImagePro5 Plus (Media Cybernetics) software which integrated time-lapse acquisition, stage movement, and software based auto-focus. During the experiment, cells were grown and visualized in 12-well coverslip bottom plates (MatTek) coated with 10μM fibronectin (Sigma). For each well, time lapse movies were obtained at four fields of view. Each movie was taken at a time resolution of 20 minutes and was filmed for at least three days (over 200 time points). Each time point included three images- phase contrast, red and yellow fluorescence.

3. Drug Materials

Camptothecin (C9911 Sigma), was dissolved in DMSO (Dimethyl sulphoxide hybri-max), (SIGMA, D2650) to achieve a stock solution of 10mM. In each experiment, the drug was diluted to 10μM in a transparent growth medium (RPMI, 1% PenStrep, 10%FCS, w/o riboflavin, w/o phenol red, Biological Industries, Kibbutz Beit Haemek, Israel). Growth medium (2ml) was replaced by the diluted drug (2ml) under the microscope. The same procedure was carried for the following drugs: Etoposide (E1383 Sigma), dissolved to a stock of 10mM and diluted to 40μM and for Cisplatinum (P4394 Sigma) dissolved to 100mM and diluted to 80μM. The stock solution for ActD (A1410 Sigma) was 1mg/ml and was diluted to 1μg/ml. The stock solution for Irinotecan was 10mM and was diluted to 250 μM in growth medium.

4. Image analysis of time lapse movies

We used a custom written image analysis tool developed using the Matlab image processing toolbox environment (Mathworks, Natick, MA). The main steps include; image correction, segmentation, tracking of the cells and automated identification of cell phenotypes (mitosis and cell death).

Image background correction (flat field correction and background subtraction) was carried out as previously described (*7*). No significant bleaching was observed (on average less than 3% over the duration of the experiment).

Cell and nuclei segmentation was based on the red fluorescent images – all clones in the library showed similar distribution of red fluorescence – bright in the cytoplasm and significantly brighter in the nucleus. The main steps of the segmentation process are: 1) Differentiation between cells and background by a global image threshold using Otsu's method (*1*); 2) Segmentation of neighboring cells by applying the seeded watershed segmentation algorithm. Seeds were obtained by smoothening the red intensity image and usage of bright nuclei as cell seeds (by identifying local maxima) – one seed per cell; 3) Nuclei segmentation following cell segmentation; each cell was independently stretched between zero and one and a fixed threshold was used to differentiate between the cytoplasm and the nucleus; 4) Tracking of cells was performed by analyzing the movie from end to start and linking each segmented cell to the cell in the previous image with the closest centroid.

5. Automated cell death and mitosis identification

The automated cell death identification algorithm utilized the morphological changes correlated with dying cells: rounding followed by blebbing and an explosion of the outer membrane or its collapse (for example see Supporting movie 1). We constructed an artificial neural network (ANN) algorithm that can identify each one of these morphological patters similar to the method previously described in (*8*). Briefly, we constructed two sets of images: The first contained 400 cell images in different stages of cell death and the second contained 400 live cell images. For each image we computed a collection of high-level image texture features. An example of such a feature is a measure of object roundness, which is relevant due to the rounding that typically occurs prior to cell death. This process transforms each image into a multi dimensional vector of features. Based on these features we train an ANN classifier in order to distinguish between live and dead cells resulting in a 96% sensitivity and specificity on a previously unseen test set.

We found that the dynamics of cell death measured by this algorithm follow closely the dynamics obtained by a standard neutral red assay (described below) on the cells (Fig. S3). The advantages of this algorithm is that it requires no external chemicals, and it allows direct count of the number of cells showing the correlates of death rather than a proxy (e.g. population average of ATP, membrane permeability or staining of organelles). Furthermore, the direct cell monitoring approach allows a high temporal resolution (every 20 minutes for over 48 hours), which is usually not feasible using the alternative approaches.

Cell division is identified by monitoring the temporal protein profile of cells and searching for sharp decreases in their protein content, which occur after divisions. These events are further filtered using morphology based considerations (using a similar framework as the one used for cell death) to improve detection accuracy.

6. Neutral red assay:

Cells were plated into 96 well plates at 1000 cells/well in normal growth conditions (37^oC and 8% CO_2). A day later drugs were added to the cells at the appropriate concentrations. Cell survival was determined by neutral red (NR) (Fluka Chemie, Buchs, cat #72210, Switzerland) accumulation as described (*9, 10*), and in the following way: growth medium was removed and neutral red reagent (4mg/ml, diluted to 1:100 in growth medium) was added for an incubation of 40 min at 37^oC. After 40 minutes, the neutral red reagent was removed from the cells and fixation buffer (1% $CaCl_2$, 0.5% Formaldehyde) was added at 150µl/well. Subsequently, the fixation buffer was removed and extraction buffer (1% Glacial Acetic Acid, 50% Ethanol) was added at 100µl/well. The plate was shaked vigorously and read in an ELISA plate reader (BIO-RAD Model 680 Microplate Reader). After subtraction of assay blanks, net optical density (570 nm) was computed as the average value of triplicate determinations. The NR levels in response to CPT are shown in Fig. S3B. Cell survival was calculated as the percentage of the dye accumulated in the treated vs. untreated controls.

7. Protein dynamics clustering

The five average population dynamics profiles depicted in Fig. 2 were generated in the following manner: The levels of each protein were smoothed using a median filter and

linearly scaled between -1 and 1. The distance between every pair of proteins was measured in terms of Pearson correlation and clustering was performed using a k-means algorithm (reproducibility of results using different seeds is > 99%). To choose the number of clusters we optimized over the average silhouette score (*11*), which measures the dissimilarity of a protein to its assigned cluster compared to other clusters.

8. Gene Ontology (GO) enrichment analysis

To systematically search for functions, processes and localizations common to proteins that show similar dynamics we performed a GO (*12*) enrichment analysis procedure. We devised a distance measure that uses both the protein amount and its localization changes through time. Formally, each protein *i* is represented by two vectors, c_i and n_i, describing the amount of protein in the nucleus and cytoplasm respectively in 141 sequential time points each.

The distance between each pair of proteins *i* and *j* was computed using the following formulas:

$$D_1(i,j) = \frac{1 - Corr(n_i + c_i, n_j + c_j)}{2}$$

$$D_2(i,j) = Euc\left(\frac{n_i}{n_i + c_i}, \frac{n_j}{n_j + c_j}\right)$$

$$D_{tot}(i,j) = w_1 \cdot D_1(i,j) + w_2 \cdot D_2(i,j)$$

D_1 is one minus the Pearson correlation between the total amounts of two proteins scaled between 0 and 1. D_2 is the normalized Euclidian distance between two vectors that depict the protein localization at each time point. Notice that at a given time *t*, $\frac{n(t)}{n(t) + c(t)}$ may range from 0 to 1 corresponding to a cytoplasmic and nuclear localization respectively.

D_{tot} is the weighted sum of the protein amount and protein localization distances where $w_1 + w_2 = 1$ (we used $w_1 = 0.5$ and $w_2 = 0.5$). The larger w2 is, the more emphasis is put on localization and consequentially the GO terms that were identified (see next paragraph) were more related to Cellular Compartments terms.

The GO enrichment procedure was performed as following: For each protein we generated a list containing all other proteins ranked according to their distance. Each protein can be thought of as a cluster center and all the other proteins are ranked according to their distance from that center. We wanted to find whether a subset of proteins that show similar dynamics, i.e. reside near the cluster center, also share a common GO term. To this end we used a flexible cutoff version of the Hyper Geometric score termed mHG (*13*). This analysis was done using GORILLA software [http://cbl-gorilla.cs.technion.ac.il/].

9. Quantification of nucleolar translocations

To detect translocation events between the nucleoli and the nucleoplasm, we followed a three step procedure. First we focused on a subgroup of clones that showed initial nuclear localization of the YFP tagged protein (i.e. pixels of the nucleus were the source of over

50 % of the total intensity). Then, for each of the selected clones we calculated the ratio of fluorescence intensity between the top and bottom ten percent pixels in individual nuclei and averaged over the population. Clones whose ratio changed during the experiment by over 20 percent were inspected manually to verify that the source of change in pixel intensity distribution was nucleolar translocation.
Finally, to quantify the extent and direction (nucleoli to nucleoplasm or vise versa) of the translocation, we calculated the ratio between mean fluorescence intensity of nucleoli vs. nucleoplasm ($R_{ncll/nuc}$) at the two time points were the max/min ratio was maximized and minimized.
For clarity in Fig. 3B and C, $R_{ncll/nuc}$ measurements (ranging from 0 to 3) were normalized to 0.5, 1 and 2 at the time of drug addition (t=0), based on the $R_{ncll/nuc}$ ratio at t=0; $R_{ncll/nuc}$ values <0.8 were normalized to 0.5, 0.8<$R_{ncll/nuc}$ <1.2 were normalized to 1 and $R_{ncll/nuc}$ values >1.2 were normalized to 2.

10. Determination of bimodal behavior
The coefficient of variance (CV defined as the ratio between the std between cells and the mean) was measured for 400 proteins for 47 hours following addition of CPT (at a 20 minute resolution) (see Supporting Fig. S15).All CVs were normalized to average 1 (CV(i,j)/mean(mean(CV)) where i is protein number (i = 1..400) and j is time point (j=1..141)). All proteins deviating 3 standard deviations from the average normalized CV were considered as bimodal candidates (N=53). Following manual inspection, 24 of these proteins listed in Supporting table 1 were classified as bimodal.

11. Immunoblots against 20 selected proteins
Total cell lysates were prepared with RIPA buffer (Pierce) according to manufacturer's instructions. The protein concentrations were determined by BCA protein assay kit (Thermo scientific). Equal amounts of proteins were resolved on SDS-PAGE and subjected to immunoblot analysis by using the antibodies listed below. The intensity of protein bands was quantified by using ImageJ software.
The following commercially available primary antibodies were used in the study:
Antibodies against AKAP8L (ab51342), Calmodulin/CALM1 (ab38590), Cyclophilin A/PPIA (ab3563), DDX5 (ab21696), Enolase/ENO1 (ab35075 and ab49256), eIF3K/EIF3S12R (ab50736), GAPDH (ab9285 and ab9484), HSP90 (ab13492 and ab34909), Nucleophosmin/NPM1 (ab15440), PBX3 (ab56239), CKS2 (ab54658), Topoisomerase1 (ab28432) and VPS26 (ab23892) were purchased from Abcam.
Antibodies against Calmodulin/CALM1 (FL-149), HDAC2 (H-54), Lamin A/C (H-110), RACK1/GNB2L (H-187 and B-3) and Fascin1/FAU (H-110) were obtained from Santa-Cruz.
Antibodies against RPL37 (A01), RPS7 (A01) and RPS3 (A01) proteins were obtained from Abnova.
Anti-Myosin IIA/MYH9 (M8064) and anti-GFP (11814460001) antibodies were obtained from Sigma and Roche, respectively.

12. siRNA

Dharmacon Accell SMARTpool E-003774-00-0050, Human DDX5, NM_004396, Dharmacon, was used. This pool contained 4 different siRNA sequences each of 19 base pairs.

Target DDX5 sequence	**Molecular Weight (g/mol)**	**Ext. Coeff (L/mol*cm)**
GCAUGUCGCUUGAAGUCUA	13,581.8	355,644
CUCUUUAUAUUGUGUGUUA	13,543.2	377,716
GCUGCACCUAUGAUUGGUU	13,575.3	349,948
GCUCUAAGUGGAUUGGAUA	13,524.5	359,204

Dharmacon, Accell siRNA was mixed with Accell delivery media supplemented with 3% FBS (fetal bovine serum) and 0.01% P/S (Penicillin, Streptomycin antibiotics). This mix was immediately added to the cells to a final concentration of 1µM of siRNA. For 12well format the mix included 1.562ml delivery media and 15.62µl siRNA. Cells were incubated for 72hours in normal condition of H1299 cells ($37^{o}C$, 8% CO_2). After 72hours medium was washed out and drug was added - 10µM of Camptothecin (CPT). siRNA efficiency was assessed by time-lapse fluorescence microscopy under incubated conditions (Figure S18) immediately after siRNA addition to the cells.

Dharmacon Accell Non-targeting Pool (A-D-001910-10-50) was used as control following the same procedures as above.

Supporting text

Limitations of library

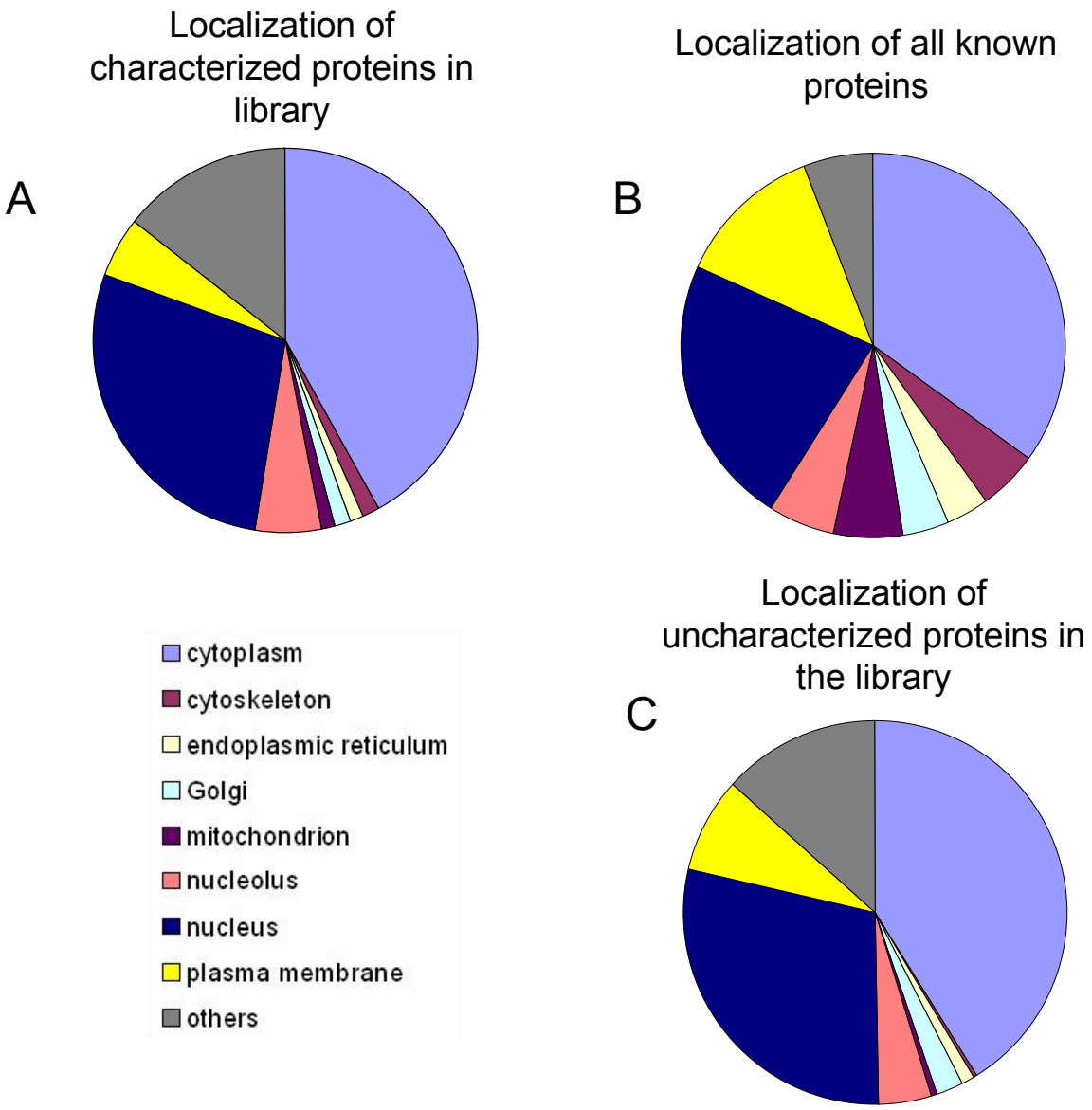
Details and limitations regarding library generation and expansion can be found at Sigal *et al.* Nature Protocols 2007 (*4*). In general, we used viral infection to introduce the YFP tag into the host cell. Viral integration sites are biased towards certain genes (hotspots) and to certain sites within genes, for example, the first intron. We observe an integration preference for specific gene functions and processes including RNA binding, ATPAse activity, splicing, translational elongation and ribosomal components (for details see S19 and S20). Our experience has been that using different conditions for viral infection and sorting, saturation is still far off and that one can expect to cover a significant fraction of the proteome. The library has been constructed at the rate of about 50 clones/person/month. Genes with only a single exon can not be captured. It is likely that once recombination in human cells becomes feasible on a large scale, N and C-terminal tagging libraries will be developed, similar to those developed in yeast (*14*). Directed targeting can be used to enrich the library with proteins of interest that are currently not included.

Comparison of response to Camptothecin vs. Irinotecan

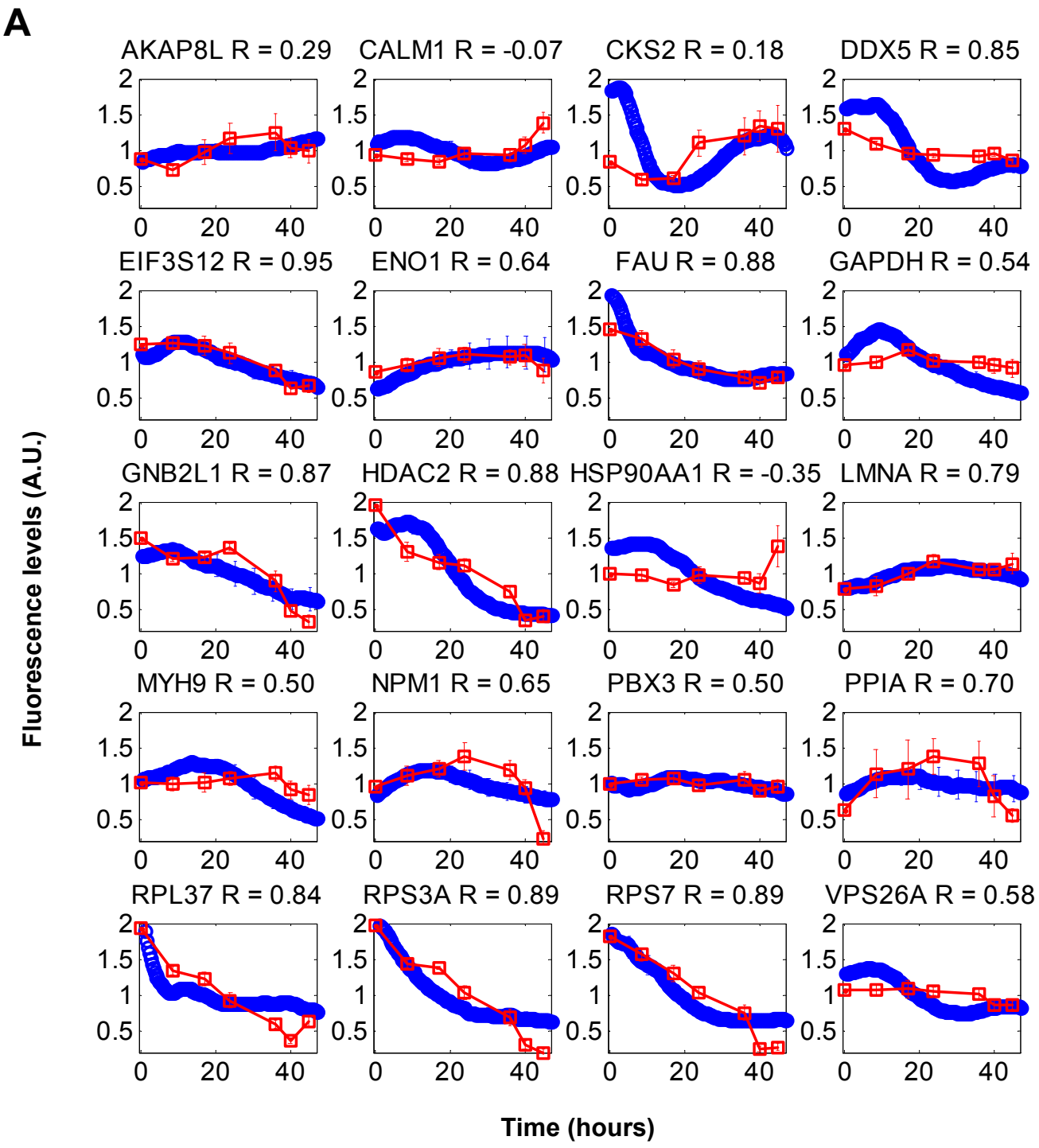
Campotothecin (CPT) has derivatives that are more commonly used in the clinic (i.e. topotecan, irinotecan and rubitecan). We asked how similar are dynamics observed with CPT to the dynamics with one of its derivatives – irinotecan (IT). To address this, we selected 10 representative proteins (chosen based on interesting dynamics following CPT) including: TOP1, TXNRD1, DDX5, RFC1, PFDN5, DNMT1, RPS3, CKS2, NCL and BAG2. We found that most of these proteins showed similarly shaped dynamic responses with CPT and IT (Fig. S12 and S13). In general, the amplitude was similar or milder for IT as compared to CPT at an equi-lethal dose. Out of four selected proteins that showed a bimodal response to CPT (BAG2, DDX5, RFC1 and RPS3), two showed bimodality in response to IT (DDX5 and RPS3). This suggests that the CPT derivative IT has similar qualitative protein dynamics to CPT, and also a few intriguing differences.

On the possible identification of drug targets using dynamic proteomics approach

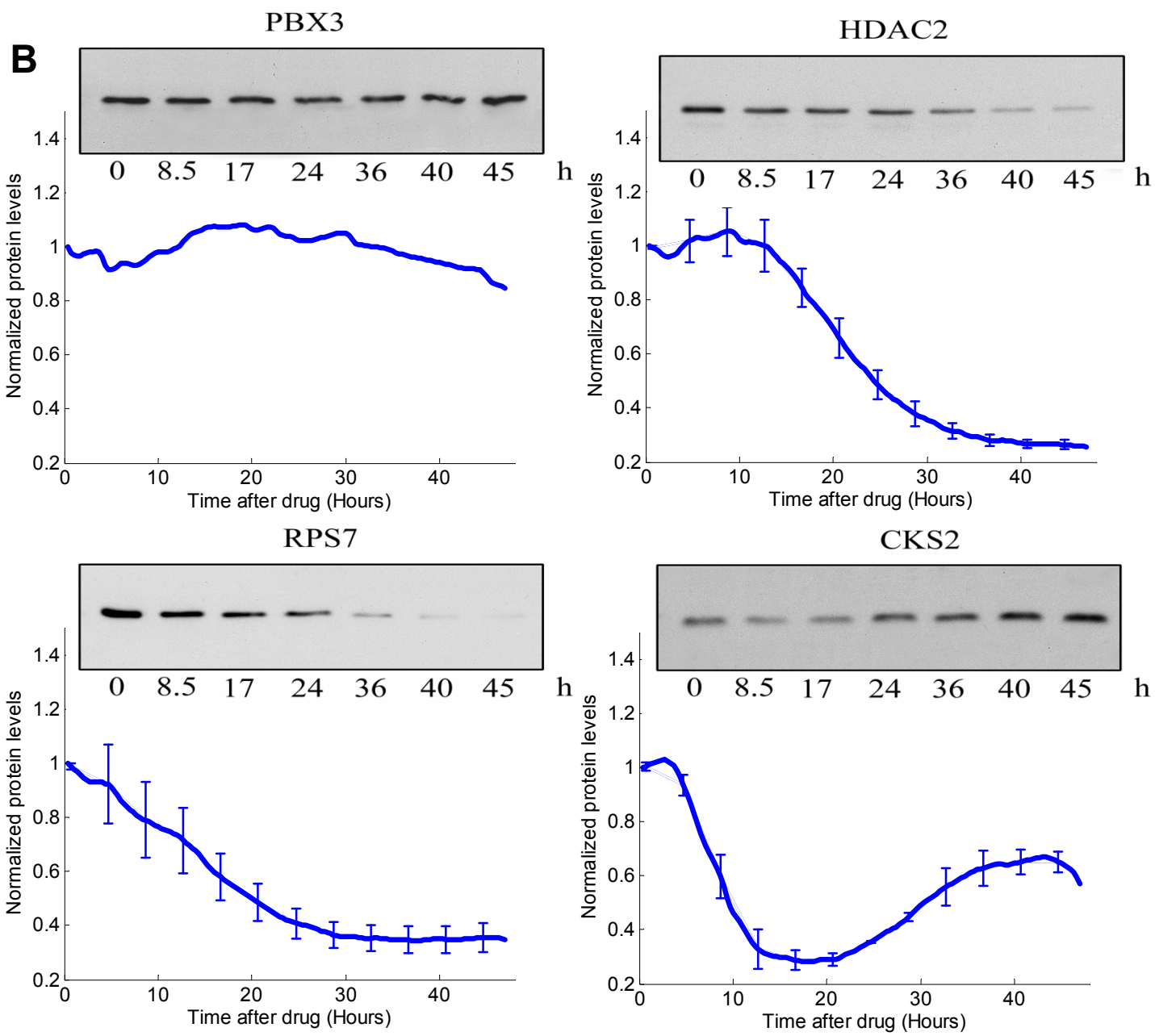
Topoisomerase 1, the camptothecin target, is among the first to respond in a number of ways that involve dynamics of protein levels (among the first to degrade) and localizations (exit from nucleoli and fragmental exit from nucleus). Based on this observation one might suggest that this approach can help to identify targets of drugs/compounds with unknown targets. However, one may expect that some drugs will not lead to changes in the level or localization of their direct target. For example, proteosome inhibitors do not necessarily lead to changes in expression level of localization of proteosome components. For such drugs, one may need to rely on secondary effects, for example, proteosome inhibitors may not change the level of the proteosome, but might affect the level of a range of proteosome targets.



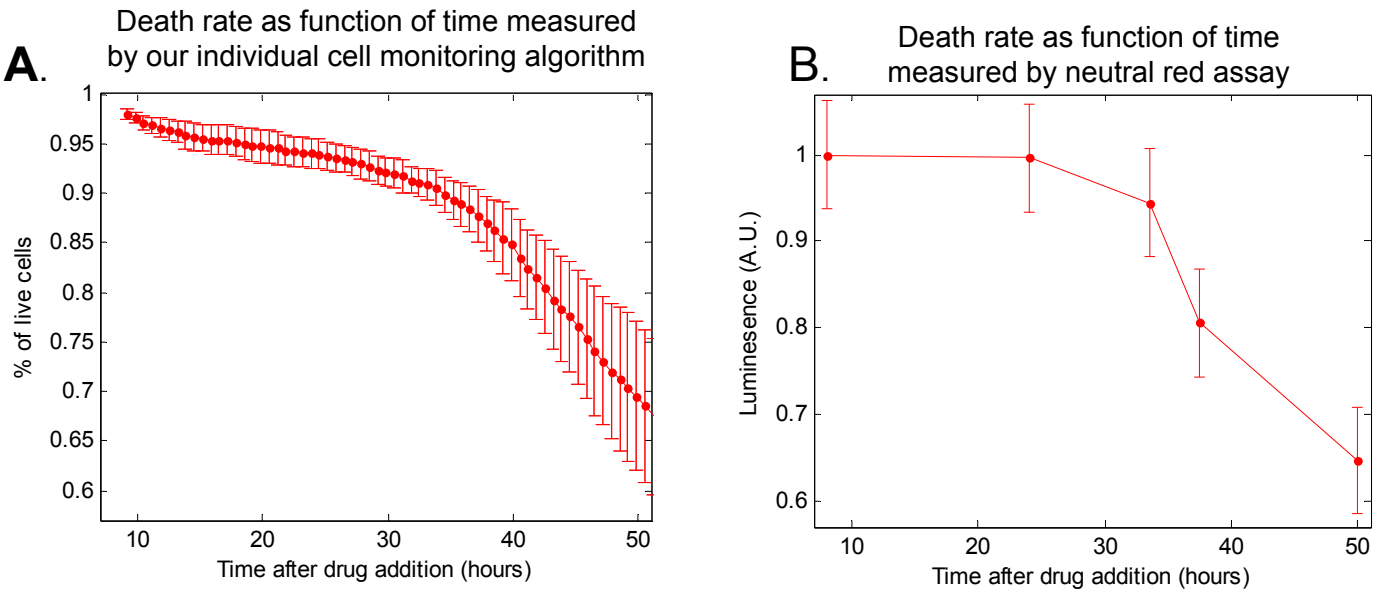
Supplementary figure 1. Tagged proteins in library are distributed across localizations similarly to the distribution of all known proteins. Distributions of protein localizations for A) All proteins in the library (LARC , Library of Annotated Reporter Clones) with published localization. B) Localization distribution of all proteins in the GO database. C) “Uncharacterized proteins in library based on manual inspection. (These proteins include hypothetical proteins and proteins encoded from regions in the genome denoted as ESTs and mRNA. These proteins have no published localization).



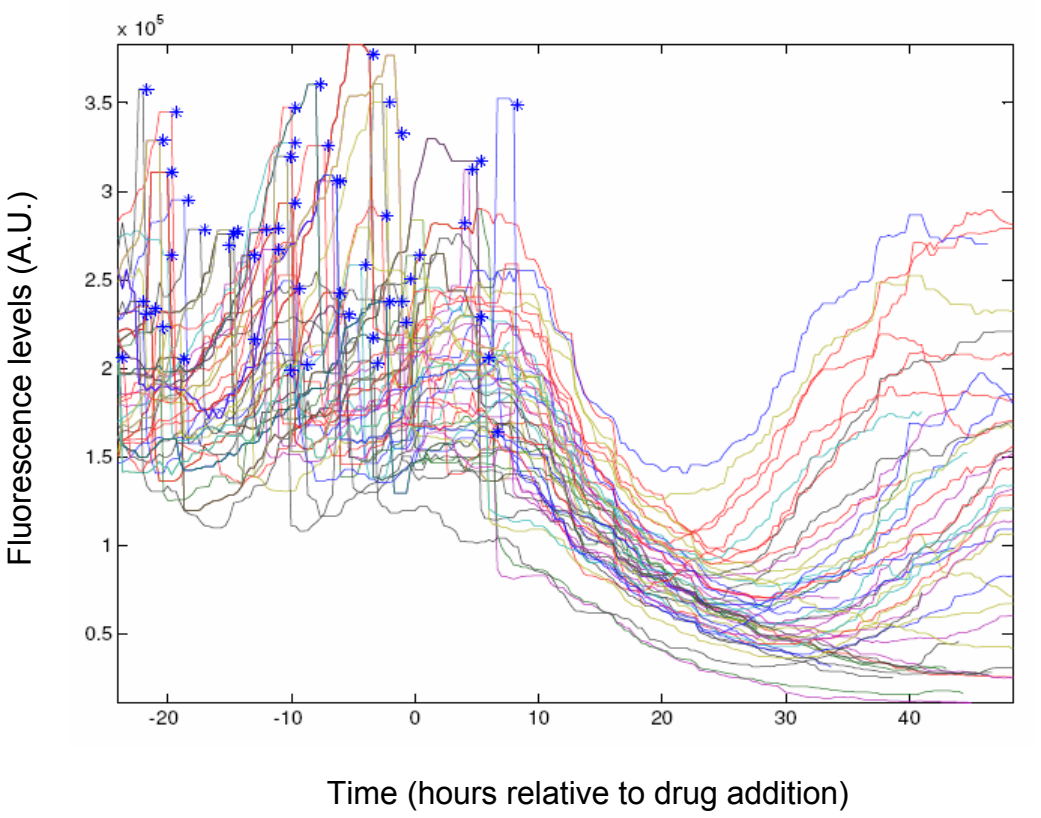
Supplementary figure 2. Immunoblots for twenty proteins show dynamics highly similar to those in the library movies. A) For each protein: red line denotes quantification of immunoblot analysis with antibodies for each protein (measurements taken at: 0, 8.5, 17, 24,36, 40 and 45 hours following CPT addition (3-8 repeats). Blue line denotes the average fluorescence intensity dynamics for each of the proteins in response to CPT (measurements taken every 20 minutes for 47 hours). Average correlation between the two measurements across all proteins is R= 0.6. Error bars denote standard errors.



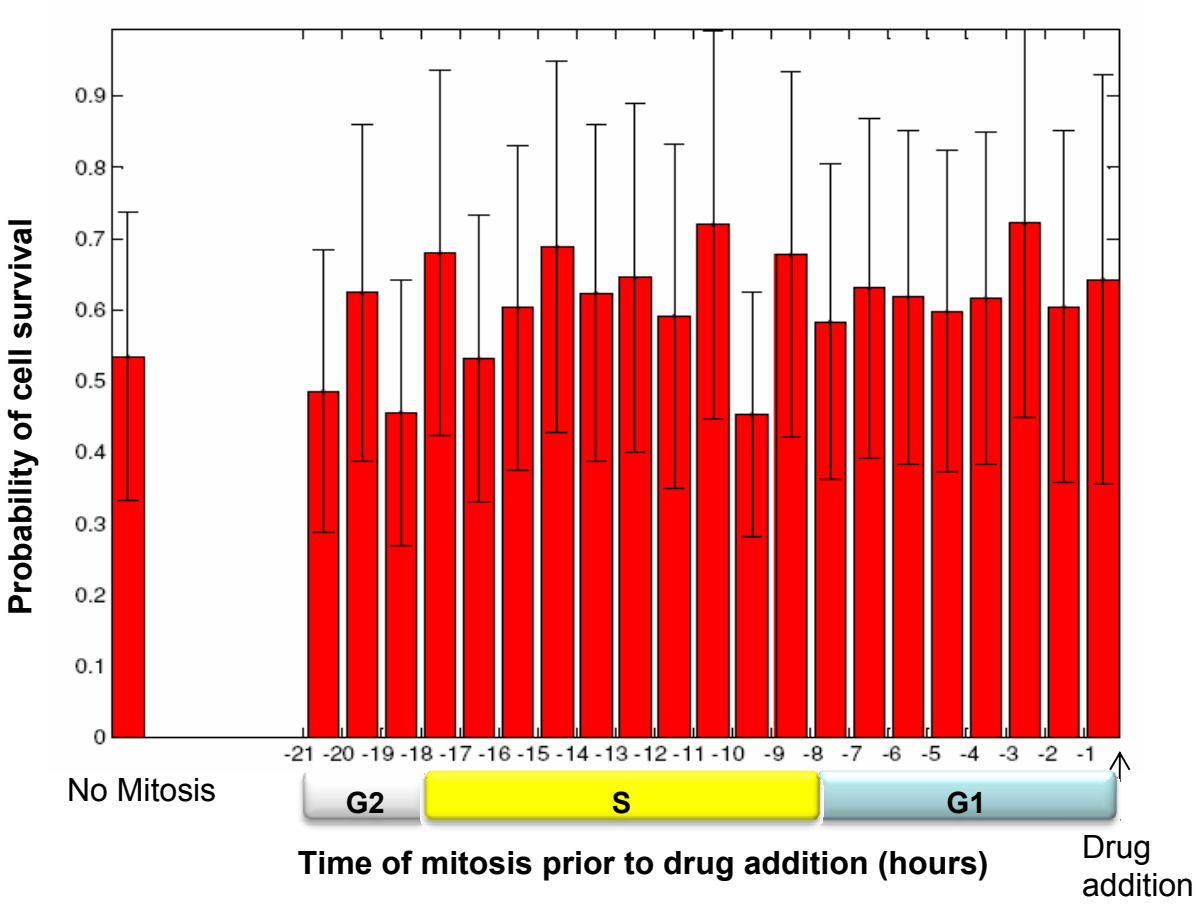
Supplementary figure 2 continued. B) Examples of immunoblots of four proteins which were quantified in (A). Graphs (blue lines) denote fluorescence measurements of tagged proteins from movies for ease of comparison.



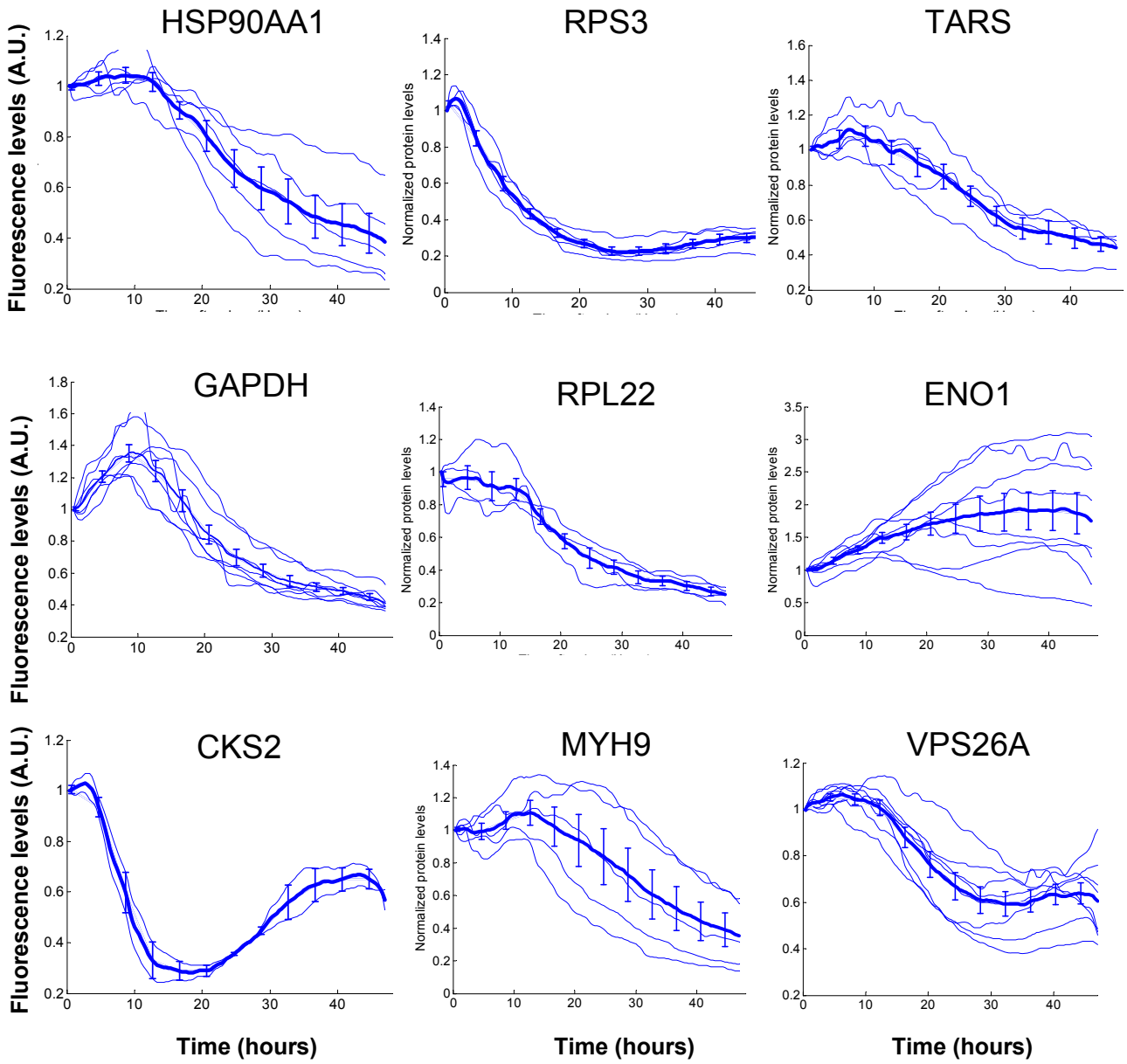
Supplementary figure 3. The cell death dynamics measured by our algorithm is similar to neutral red based measurements . (A) The percent of cell death in response to CPT addition was computed by monitoring the number of individual cells that showed cell rounding followed by blebbing (a morphology characteristic of apoptotic cells) divided by the total number of cells. Measurements were done on thousands of cells with a 20 minute time resolution. (B) Cell death in response to CPT was estimated using a neutral red (NR) assay which measures cell survival/viability (normalized to the first time point at t=8hrs). Both assays show similar dynamics with a slow death rate in the first 35 hours followed by an increase in rate from 35 to 50h.



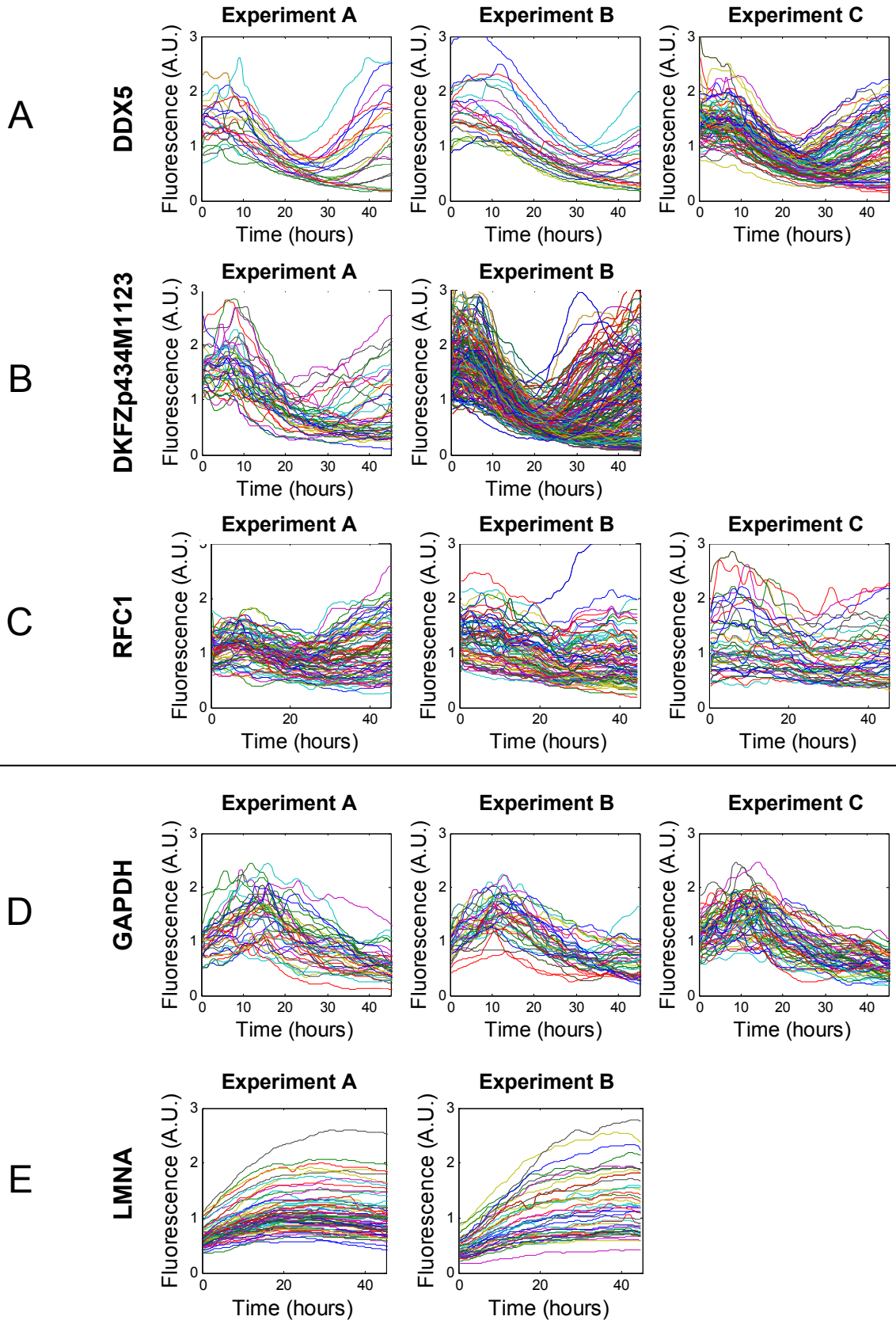
Supplementary figure 4. Automatic detection of cell division events shows that cells stop dividing at 5h after drug addition. Plotted are tracks of total tagged protein (DDX5) intensity in individual cells. Blue asterisks denote mitosis events. Time is relative to the time of drug (CPT) addition. Note that protein levels drop two-fold with each division.



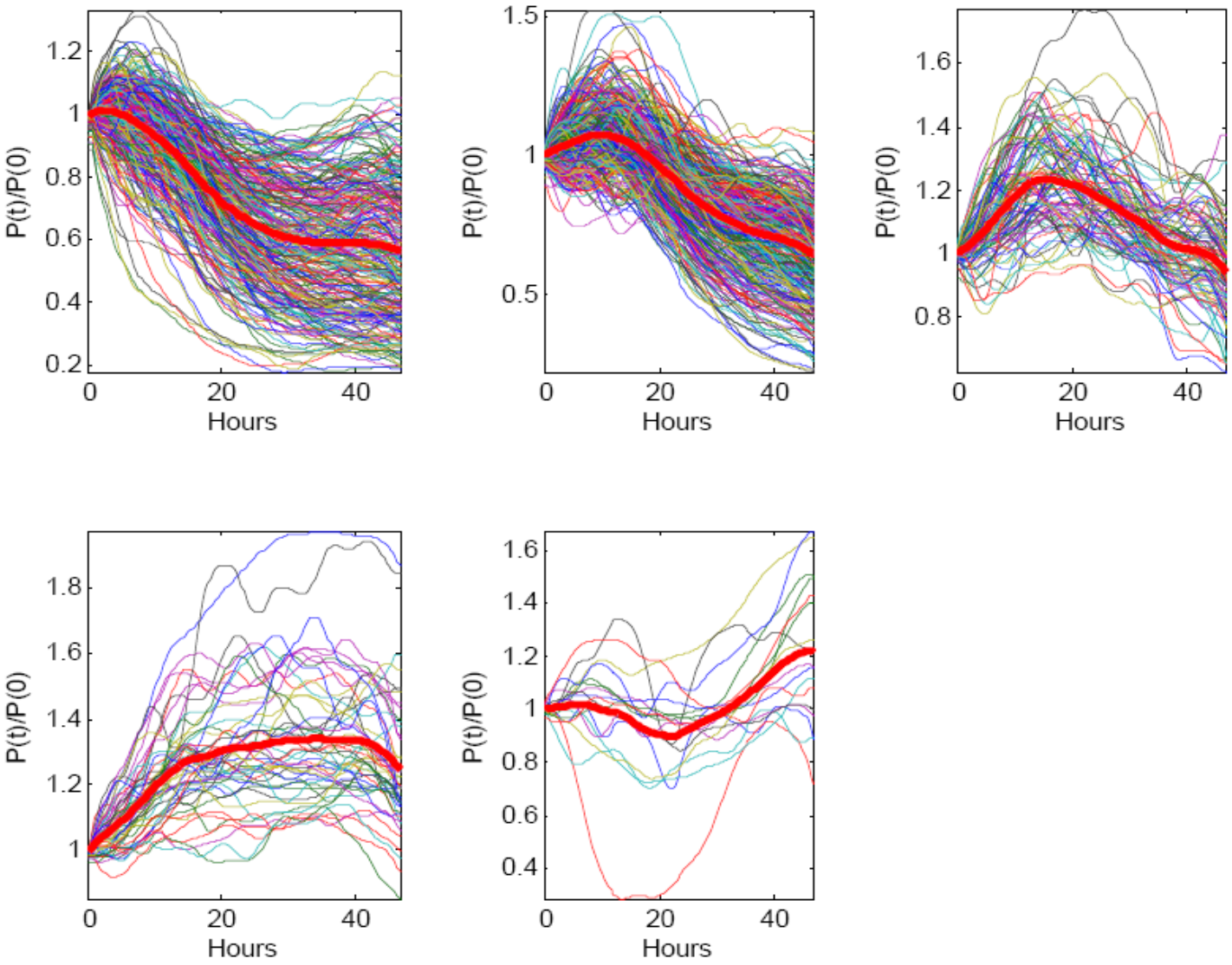
Supplementary figure 5. Cell cycle stage at time of drug addition does not correlate with cell fate. Cells were tracked -20 hours prior drug addition until 48 hours after drug addition. Mitosis and cell death timing were automatically identified in over 300 movies taken in 7 different dates covering a wide range of clones. It was previously shown that the cell cycle of H1299 cells used in this study is 18 +- 2.5 hours, and the relative fraction of time spent in G1, S and G2 cell cycle stages was determined by flow cytometry [Sigal *et al.* Nature Methods 2006]. We thus used the time of last mitosis prior to drug addition as an estimate to the cell cycle stage at the time of drug addition. The y axis is the probability of cell survival. Standard error bars correspond to variability between movies. We do not find a statistically significant correlation between the cell cycle phase of an individual cell at drug addition and that cell's probability to survive .



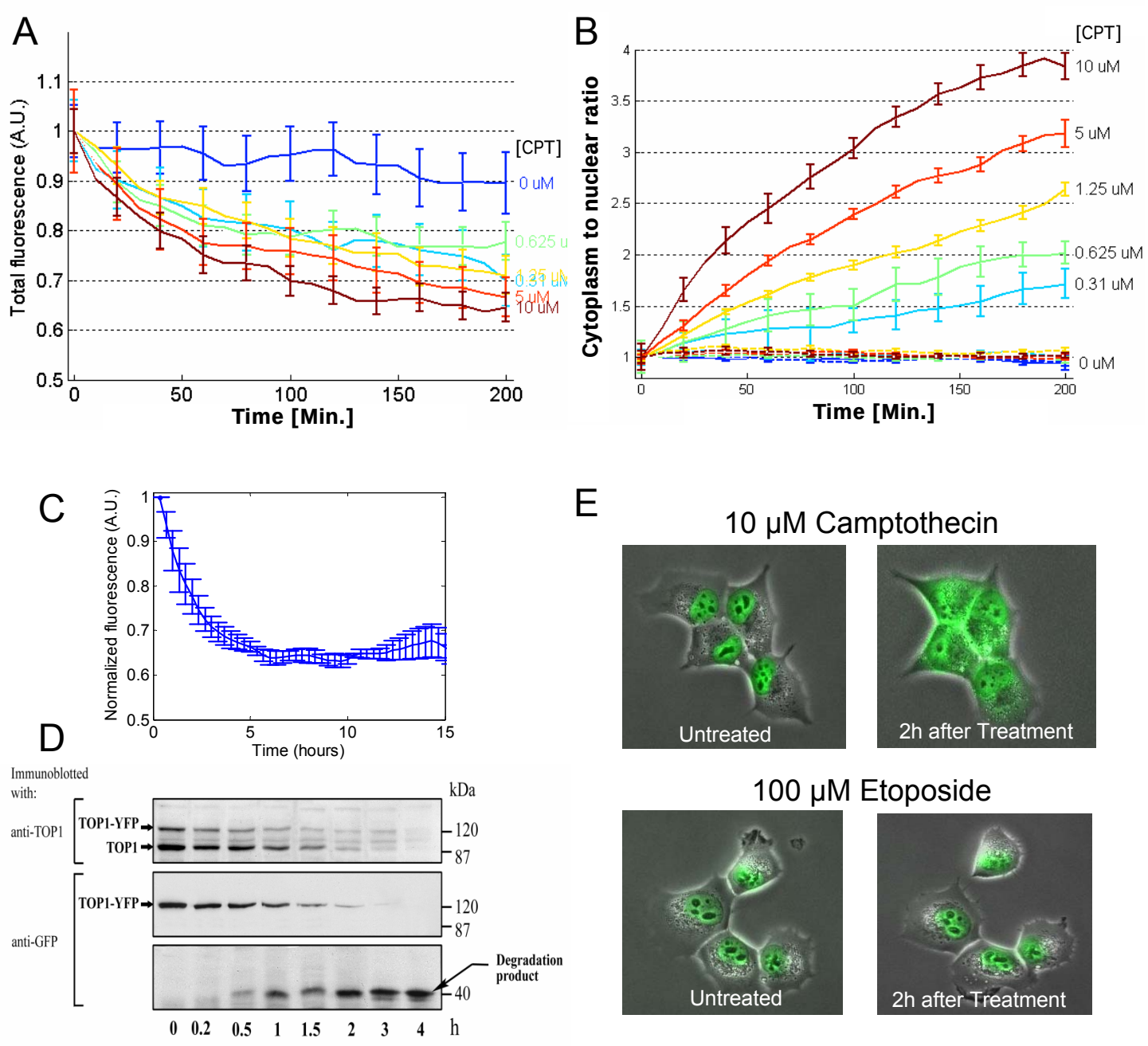
Supplementary figure 6. Experiments on different days show highly repeatable dynamics. Each experiment was repeated between 2 to 8 times. Thin blue lines denote normalized total fluorescence averaged over many cells in one experiment, bold line denotes average over all days, error bars denote standard error. Mean error (std/mean) over all clones and all time points of all proteins is 0.13 (mean correlation between experiments at different dates is R=0.8).



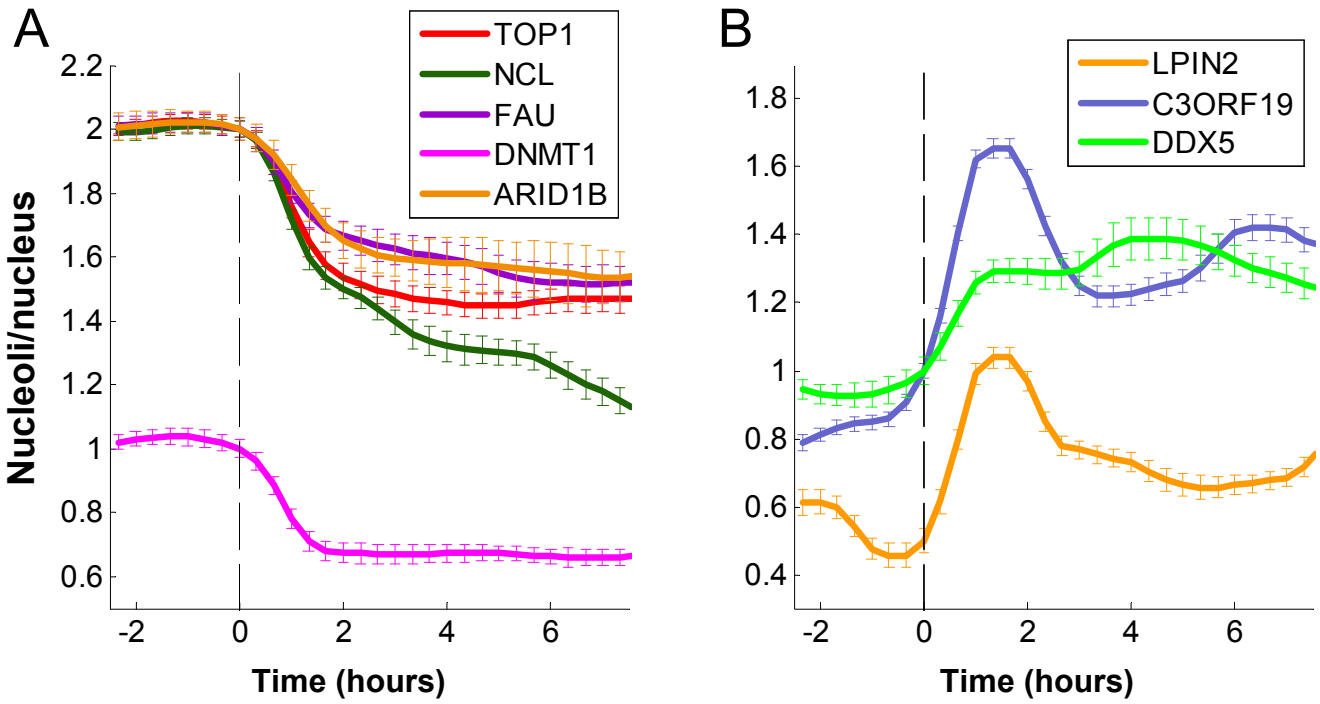
Supplementary figure 6 continued. Individual cell data is reproducible between experiments performed on different days. A-C) Examples of three bimodal proteins DDX5, DKFZ434M1123 and RFC1. D-E) Examples of two non-bimodal proteins, GAPDH and LMNA.



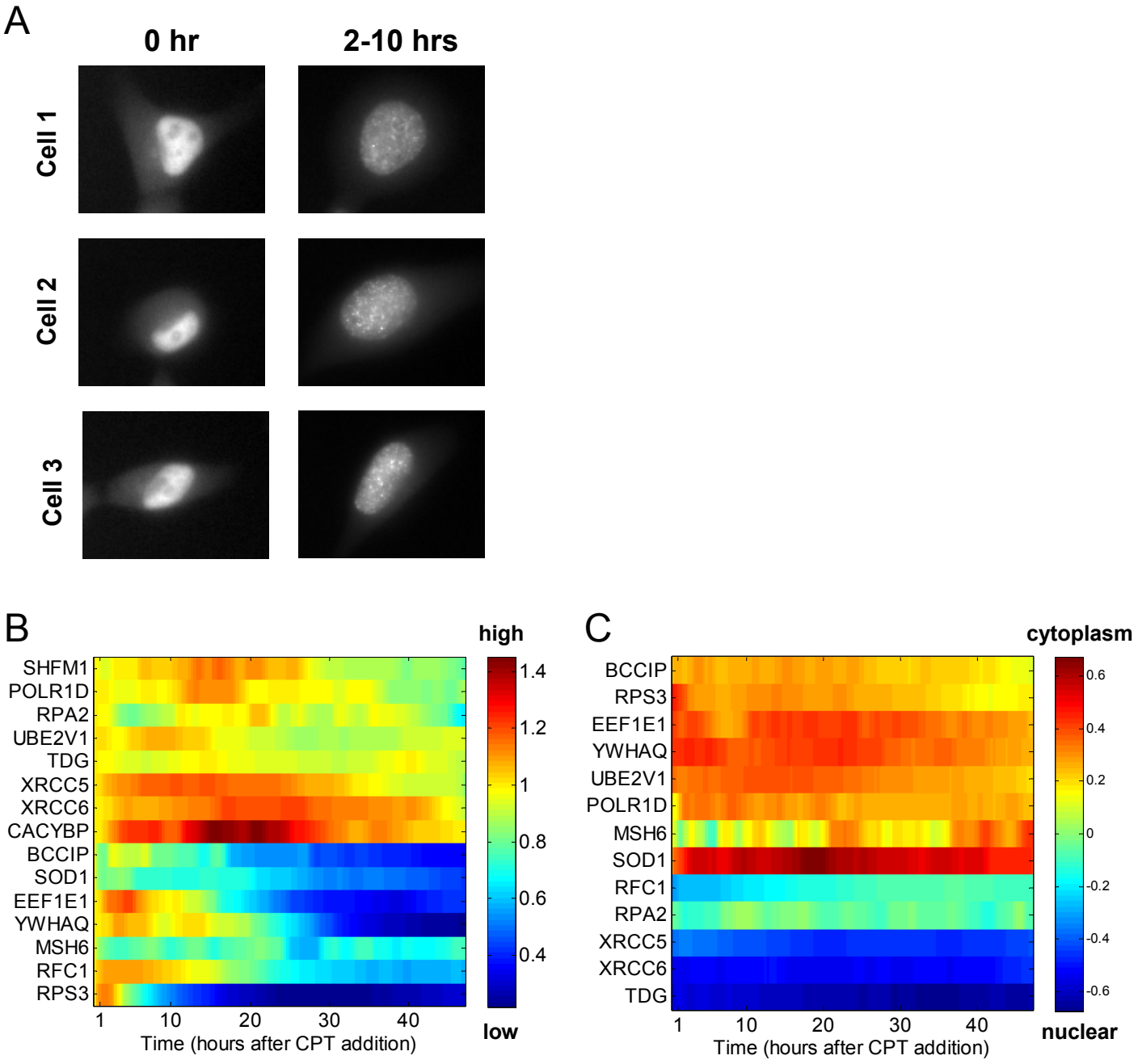
Supplementary figure 7. Protein dynamics following CPT shows several types of profiles following CPT addition. Proteins are divided into five profiles of behavior as in Figure 2 in the main text. Each line is an average of at least 30 cells. Protein levels are normalized to initial protein levels as measured at the time the drug was added. Bold red line denotes mean profile of each cluster.



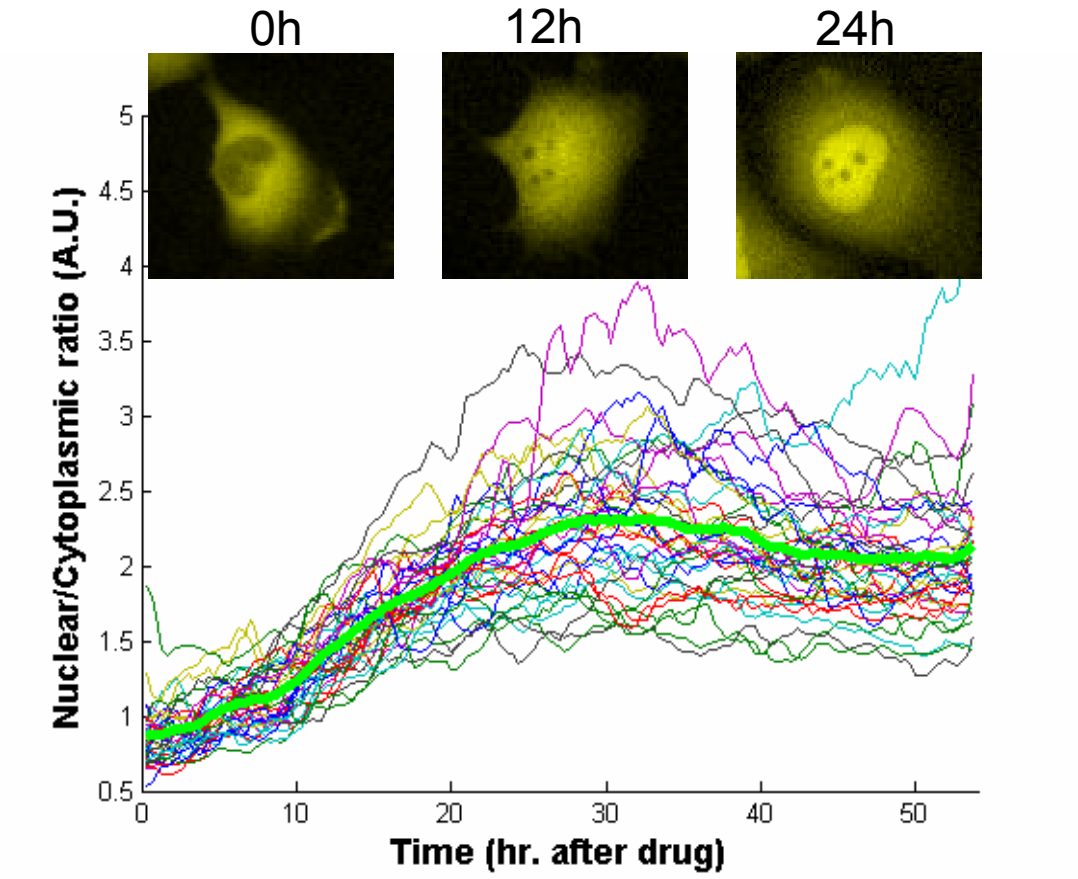
Supplementary figure 8. TOP1, the drug target shows degradation and dose dependent translocation, specific to CPT. A) Tagged TOP1 intensity decreases in a CPT dose dependant manner. Each line denotes different concentration of CPT added at time 0 onto the cells. B) Tagged TOP1 exits from the nucleus to the cytoplasm in a CPT dose dependent manner (full lines). A control nuclear protein expressed in the same cells (XRCC5-mCherry) does not exit the nucleus at any of the CPT doses (dashed lines). Each line is the mean of all cells at each time-point. C) Fluorescence of tagged TOP1 decreases to 65% of the initial fluorescence levels with half life of 1.5 hours. D) immunoblotting with anti-TOP1 and Anti-GFP shows fast decrease in endogenous and tagged TOP1 levels within 4 hours. Immunoataining with Anti-GFP Ab also reveals a degradation product of 40 kDa rising in anticorrelation with TOP1 degradation. E) Nuclear exit of tagged TOP1 does not occur with an equivalently lethal dose of etoposide, a topoisomerase-2 inhibitor drug.



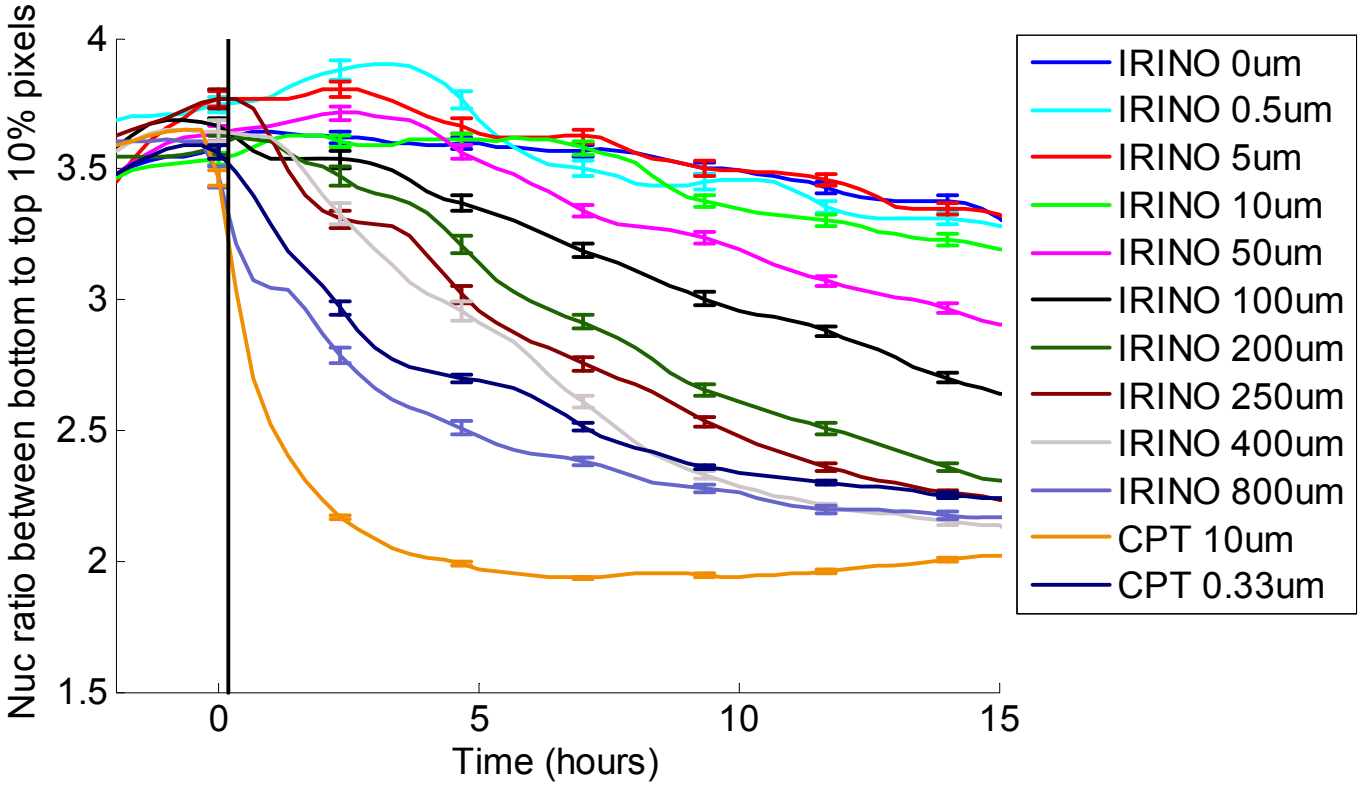
Supplementary figure 9. Actinomycin-D causes nucleolar localization changes very similar to those caused by CPT (FIG 3 in main text). Localization changes of proteins in response to addition of 1 μ g/ml of Actinomycin-D (a transcription inhibitor). Time units in hours is relative to Actinomycin-D addition. Data was normalized to protein levels at time of addition (t=0). For clarity, normalized protein levels at time t=0 were shifted from 1, to values of 0.5, 1 and 2 according to the localization of the protein (nucleoli vs. the nucleoplasm). Proteins excluded from the nucleoli were denoted at t=0 as 0.5, proteins equally dispersed were denoted as 1 and Proteins localized preferentially to the nucleoli were denoted 2. A) Tagged proteins that show a rapid decrease in nucleolar intensity. B) Tagged proteins that show a rapid increase in nucleolar/nucleoplasm ratio followed in some cases by a decrease back to basal levels.



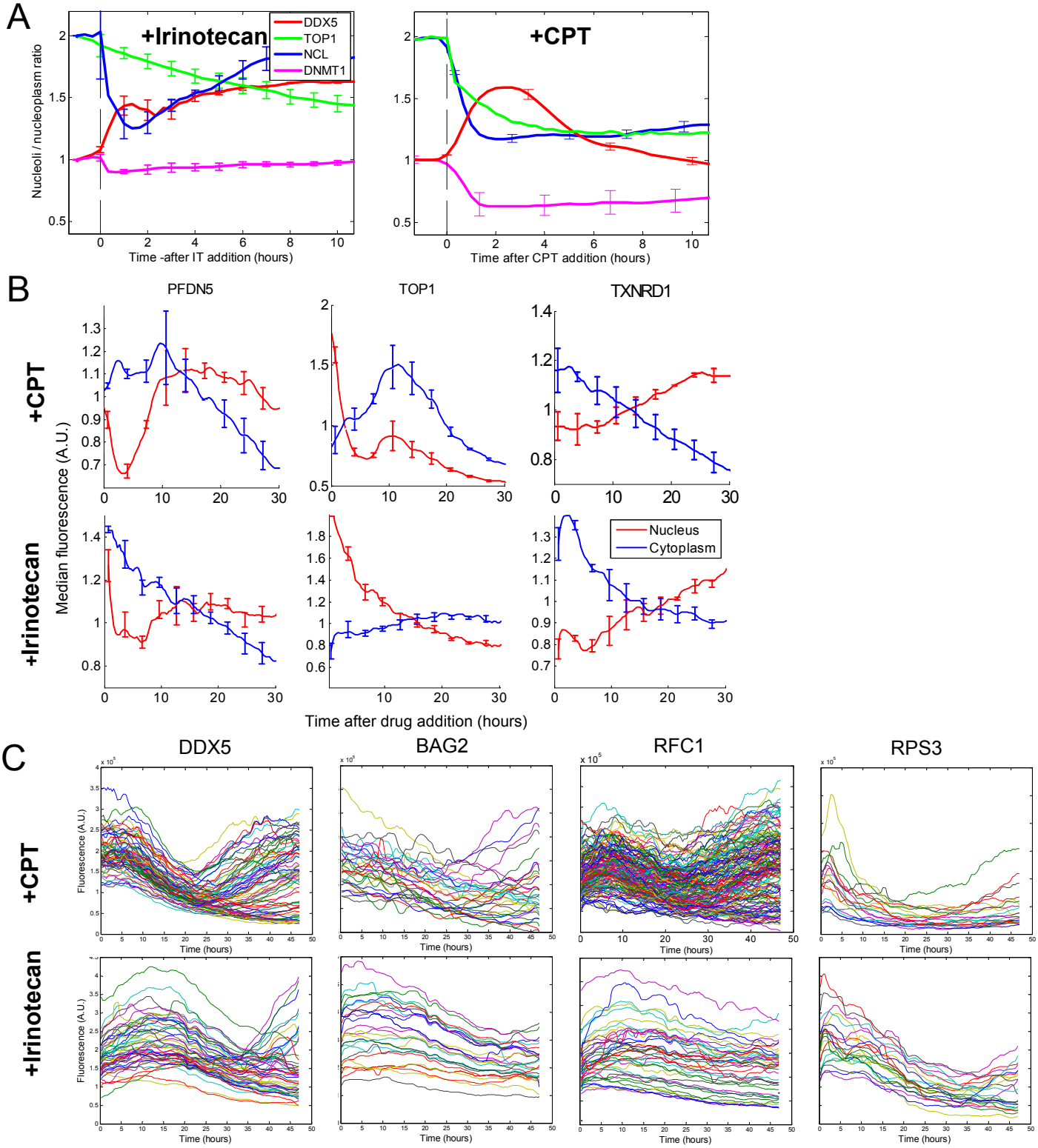
Supplementary figure 10. DNA damage protein RPA2 shows translocation to nuclear foci; other DNA damage related proteins show temporal changes in abundance. (A) Replication protein A2 **(**RPA2), localizes to nuclear foci upon CPT treatment, a well characterized response to DNA breaks. B) Average fluorescence over time following drug addition of 15 DNA damage related proteins, normalized to the initial protein levels at time zero. Blue denotes low and red denotes high. C) Localization of fluorescence for 13 DNA damage related proteins proteins (nuclear intensity divided by total intensity, blue – nuclear, red - cytoplasmic) over time following drug addition. .



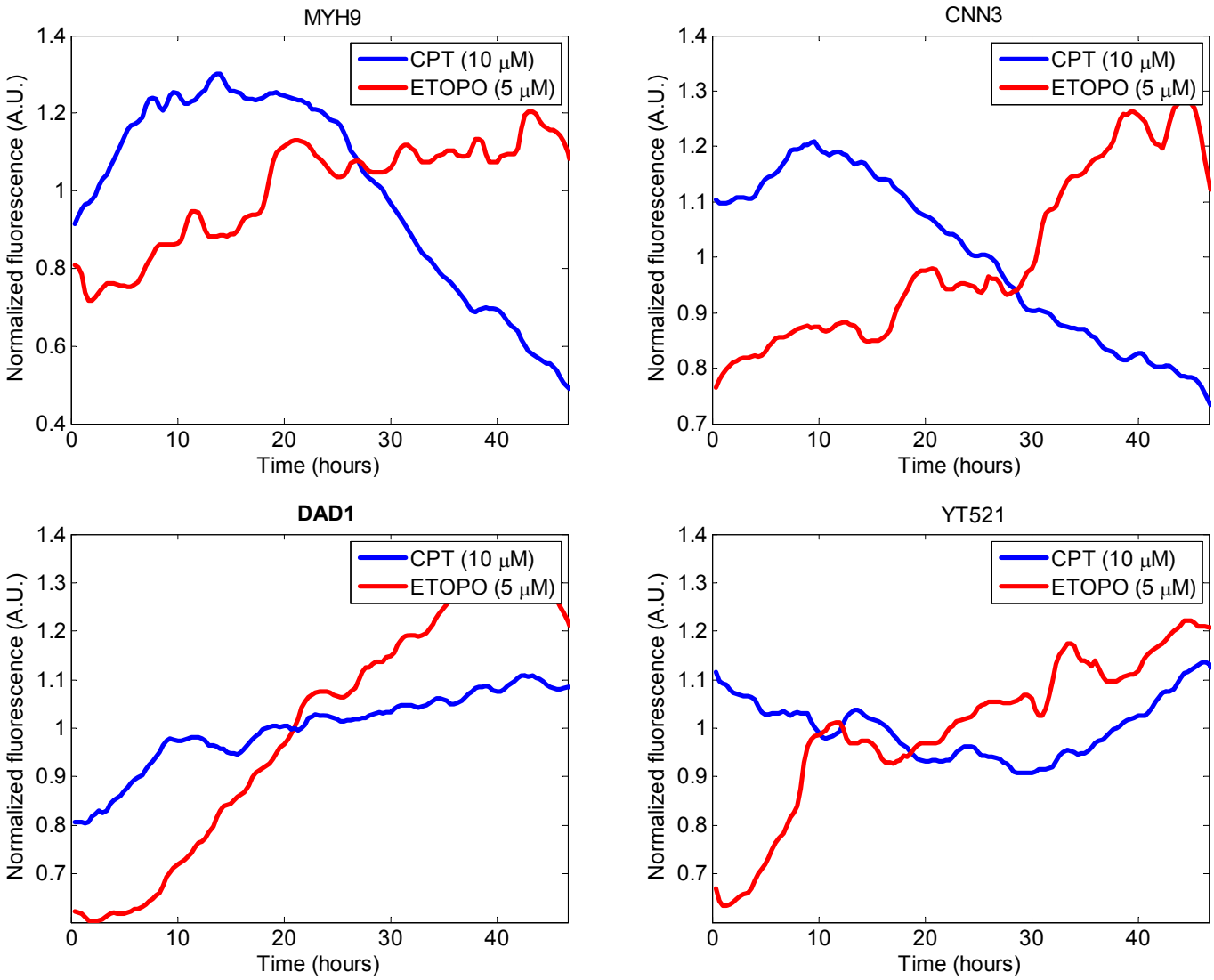
Supplementary figure 11. Nuclear to cytoplasmic ratio of TXNRD1 increases following CPT addition. Top panel: a fluorescent image of a cell with tagged TXNRD1. Time after drug (10 μM CPT) addition is indicated above each image (first image was taken before drug addition). Each line denotes the nuclear to cytoplasmic ratio measured for an individual cell tracked over 50 hours. Bold green line denotes the average.



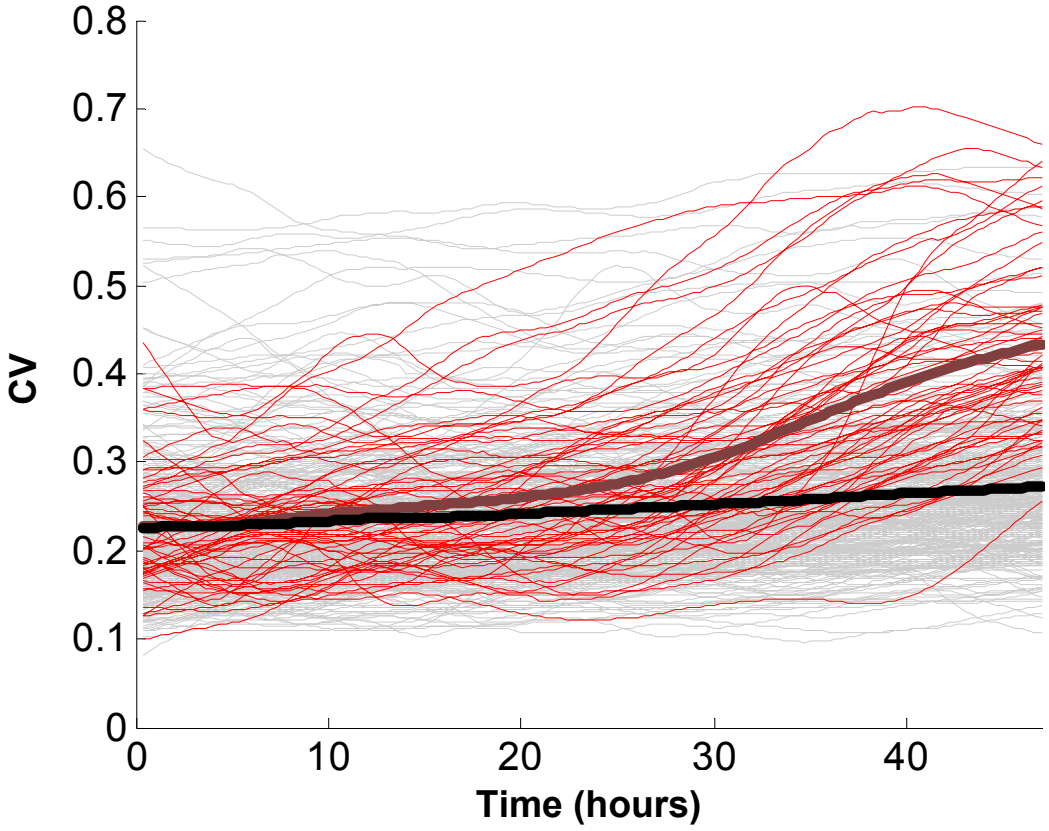
Supplementary figure 12. TOP1 exists the nucleoli upon irinotecan treatment, with dynamics similar to CPT but lower amplitude. Nucleoli exit was measured by dividing the top 10% pixels by the bottom 10% pixels within the nucleus. TOP1 exits the nucleoli at dose dependent rates, with larger effects of CPT as compared to irinotecan.



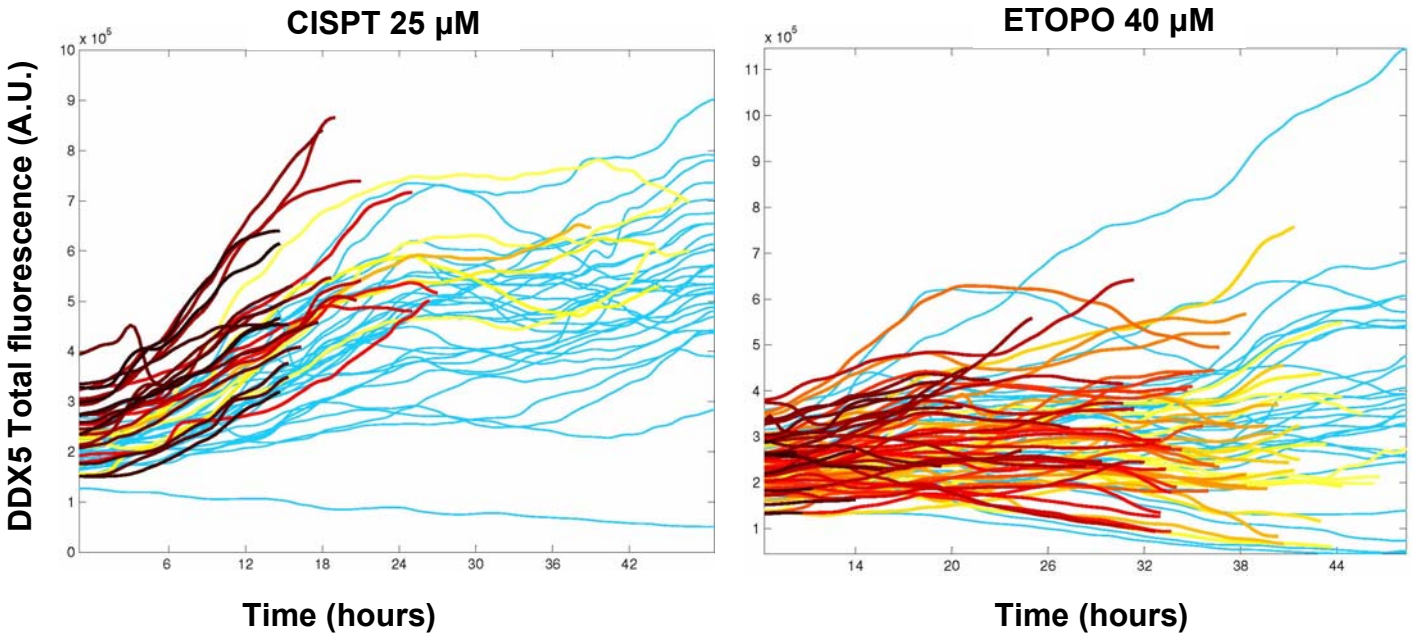
Supplementary figure 13. Protein dynamics and localization changes upon treatment with irinotecan are similar to those observed with CPT. (A) Comparison of translocation dynamics between nucleoli and nucleoplasm in response to Irinotecan (250 μM) and CPT (10 μM). Shown are 4 proteins, DDX5 (DEAD box protein 5), TOP1 (Topoisomerase 1), NCL (Nucleolin) and DNMT1 (DNA methylase 1). (B) Comparison of translocation dynamics between nucleus and cytoplasm for three proteins PFDN5 (Prefoldin 5), TOP1 (Topoisomerase 1)and TXNRD1 (thioredoxin reductase 1). (C) Comaprison of bimodality for four proteins, DDX5 (DEAD box protein 5), BAG2 (BCL2-associated athanogene 2), RFC1 (replication factor C large subunit) and RPS3 (ribosomal protein small subunit 3).



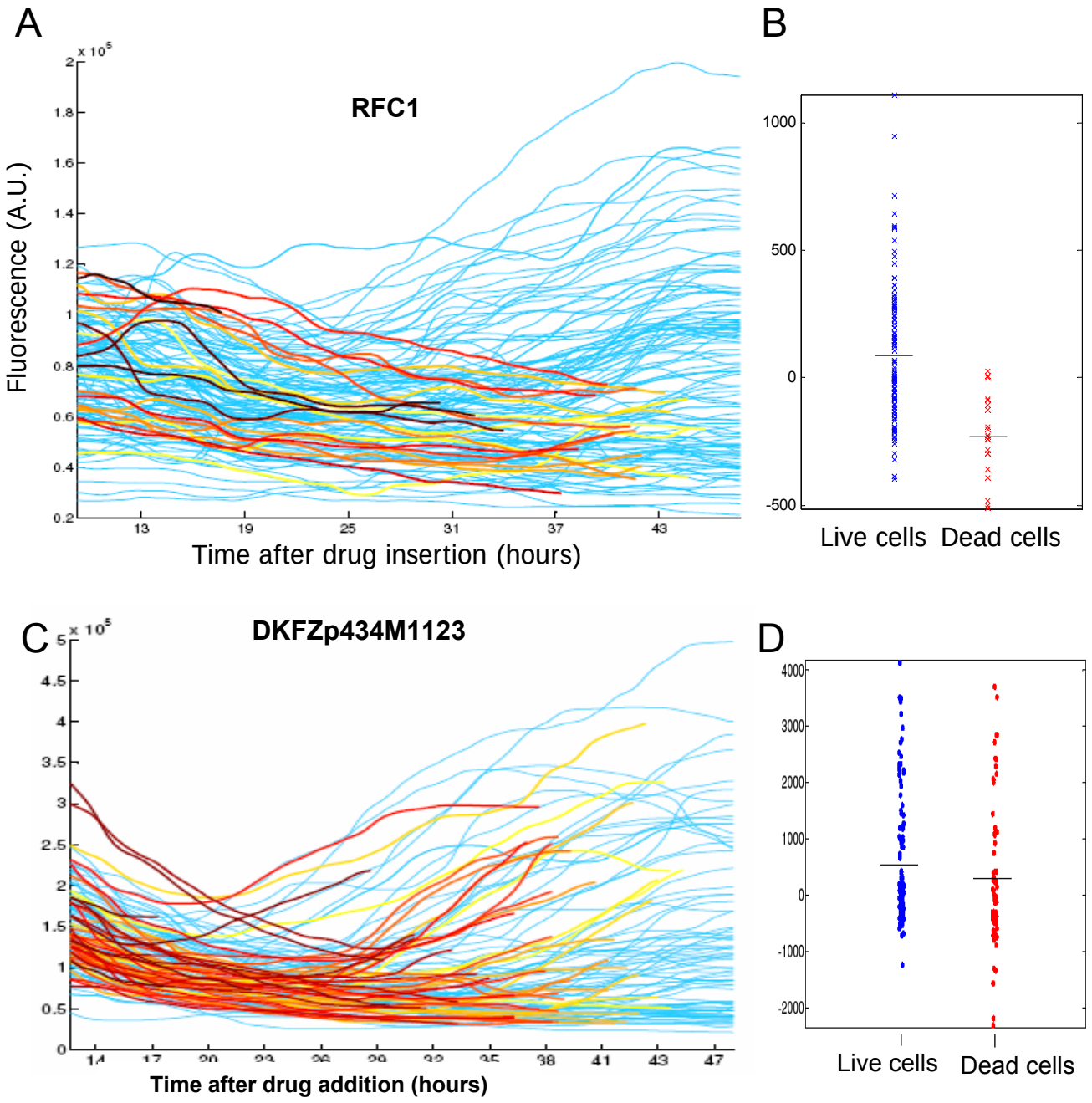
Supplementary Figure 14. Protein dynamics upon treatment with etoposide (5μM) differ from those observed with CPT (10 μM). Shown are four proteins: Myosin heavy chain 9 (MYH9), Calponin 3 (CNN3), defender against the death 1 (DAD1) and the splicing associated factor YT521. CPT average profile in blue and etoposide average profile in red.



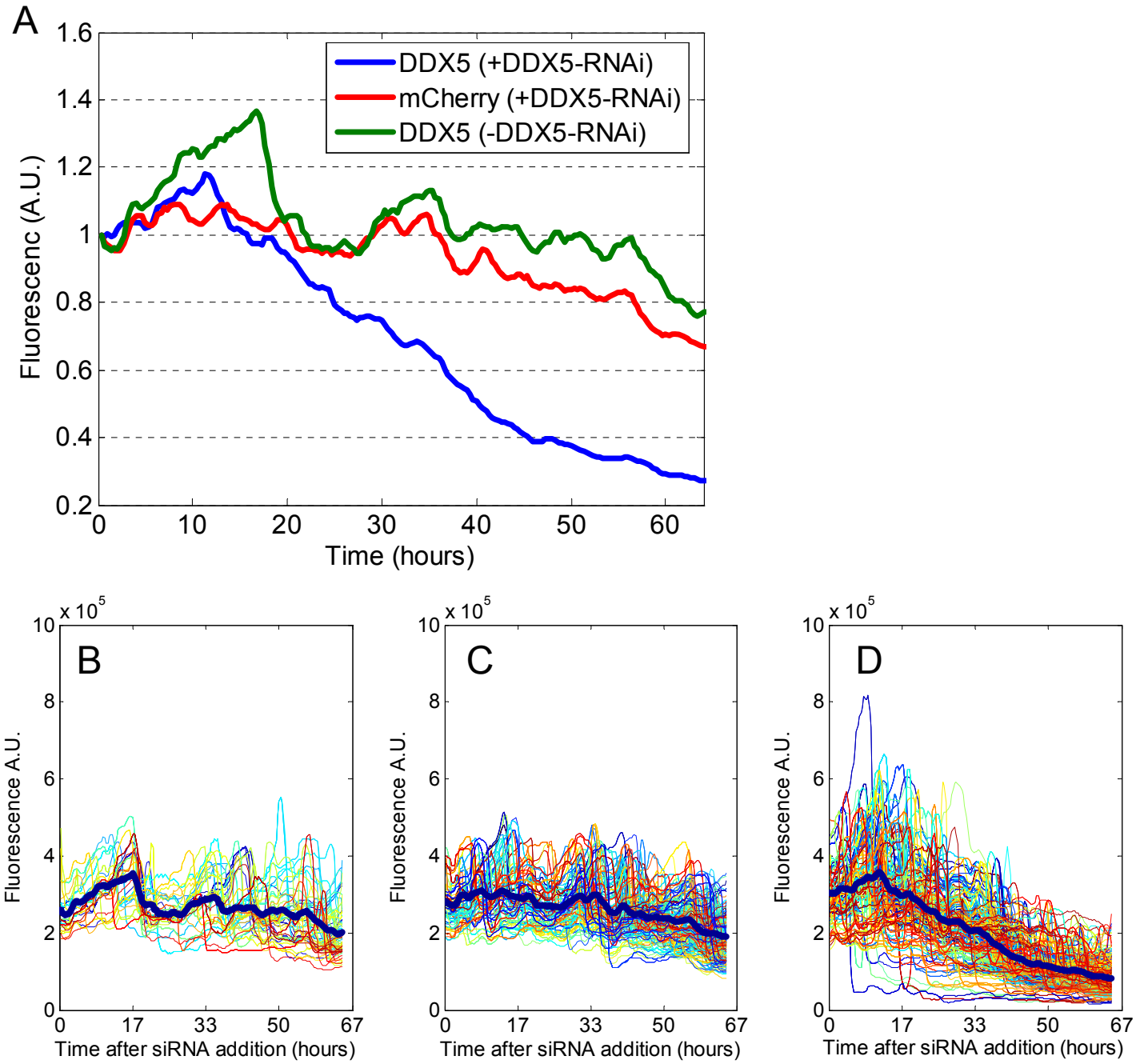
Supplementary figure 15. Cell-cell variability of most proteins rises slightly after drug, and rises sharply at 25h for a subset of proteins. Each line denotes CV of a different protein (total 400 proteins shown), where CV is the Coefficient of variance (std/mean of cells in the clone at each time point). In red: proteins that show CV of over 3 standard deviations from the average normalized CV of all proteins. Average CV of all 400 proteins is bold black and that of the 53 'bimodal' proteins is bold brown.



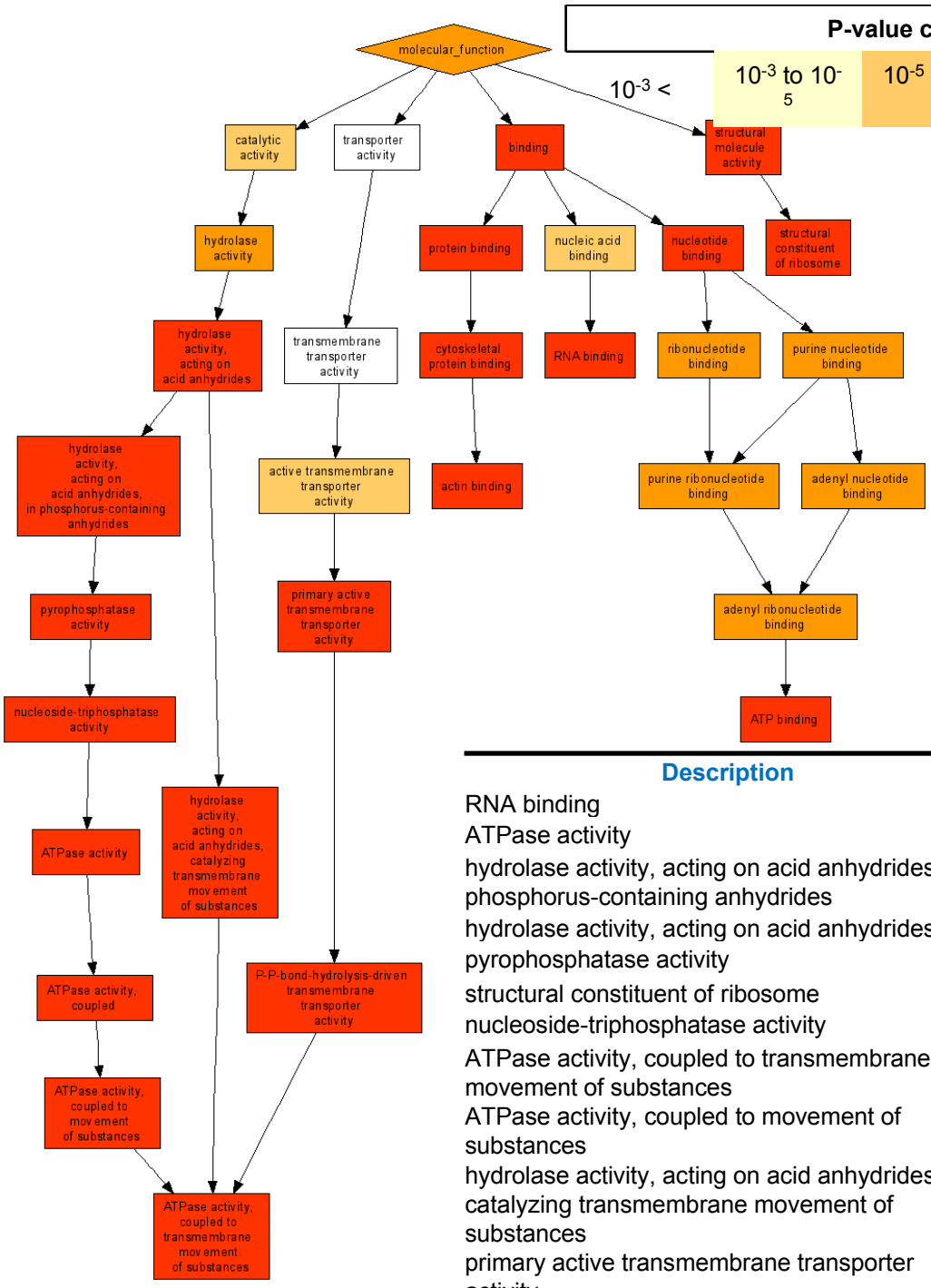
Supplementary figure 16. DDX5 shows different dynamics in response to different drugs. Response of DDX5 to Cis-Platinum 25 μM and Etoposide 40 μM. Each line denotes total fluorescence measured for a single cell.



Supplementary figure 17. RFC1 - a bimodal protein that correlates with cell fate, and an example of a bimodal protein that does not. A) Replication factor RFC1 shows an increase in intensity in cells that survive the drug after 48h, and a decrease in cells that show the morphological changes associated with cell death. Heavy colored lines are cells that die, with darker colors corresponding to earlier cell death. Blue lines are cells that do not die in the movie. B) Cells that show the morphological correlates of cell death have significantly higher slopes of RFC1 fluorescence accumulation than cells that do not (T-test P<10-6). Slopes are defined as in Fig. 4 of the main text. C-D) The hypothetical protein DKFZp434M1123 shows similar dynamics to DDX5, (coloring as described in A. However, no significant correlation is found between the dynamics and cell fate (T-test P=0.18).

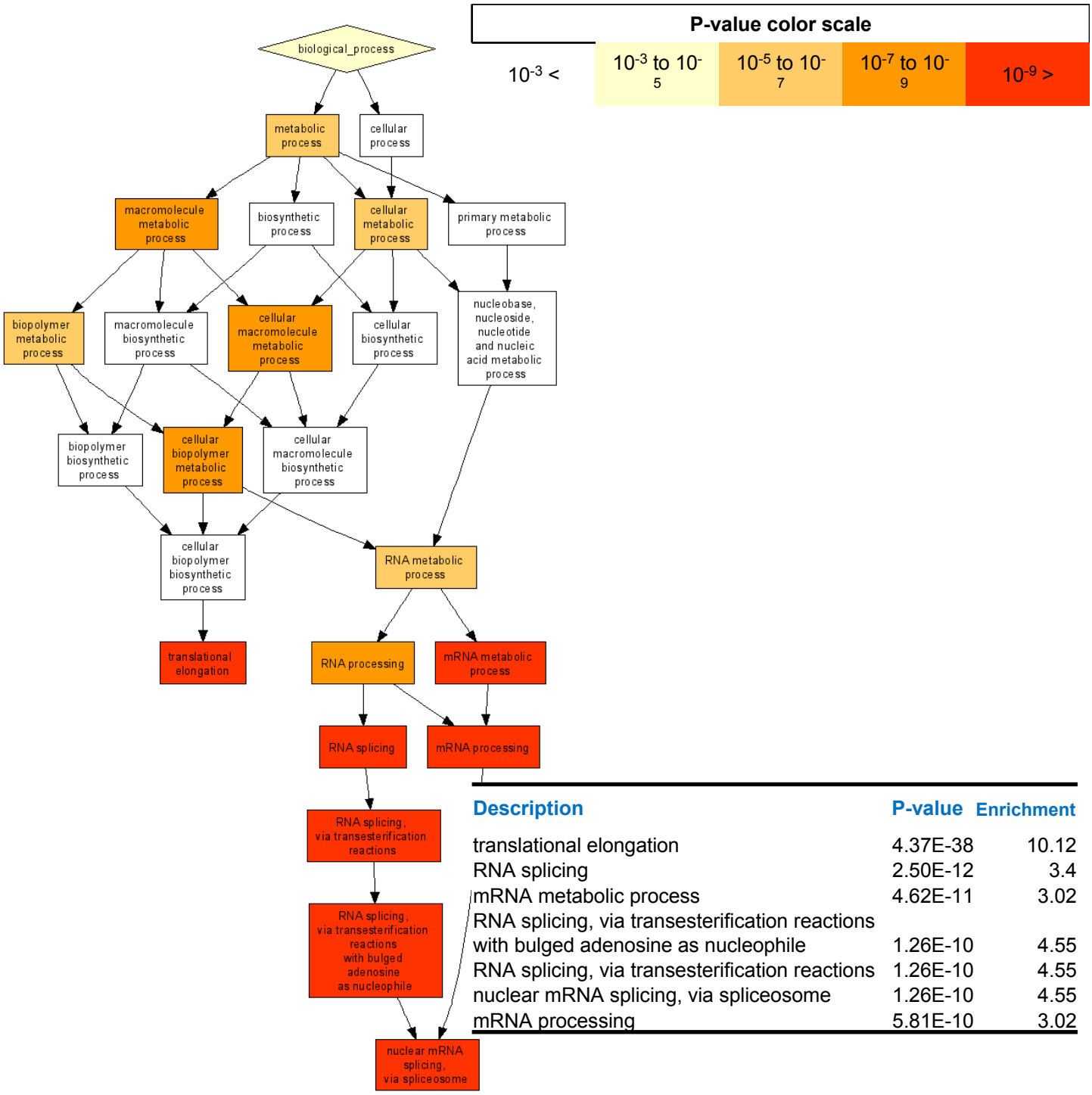


Supplementary figure 18. Reduction in DDX5 levels following siRNA addition. A) Average fluorescence levels of DDX5 with DDX5-RNAi (blue line), without DDX5-RNAi (green line) and of the cherry tag following DDX5-RNAi (red line). Data was normalized to first time point. B-D) Single cell data, from which averages were calculated on. B) data of tagged DDX5 with OUT DDX5-RNAi, C) data of cherry tag following DDX5-RNAi, D) data of tagged DDX5 following DDX5-RNAi.



Description	P-value	Enrichment
RNA binding	5.73E-35	3.61
ATPase activity	1.55E-27	4.68
hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	2.57E-26	3.43
hydrolase activity, acting on acid anhydrides	3.42E-26	3.42
pyrophosphatase activity	7.26E-26	3.41
structural constituent of ribosome	1.01E-25	5.73
nucleoside-triphosphatase activity	2.22E-25	3.45
ATPase activity, coupled to transmembrane movement of substances	1.63E-19	6.43
ATPase activity, coupled to movement of substances	2.29E-19	6.37
hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances	3.20E-19	6.31
primary active transmembrane transporter activity	6.46E-19	6
P-P-bond-hydrolysis-driven transmembrane transporter activity	6.46E-19	6
protein binding	1.19E-18	1.37
ATPase activity, coupled	3.34E-18	4.17
structural molecule activity	1.85E-16	2.62
nucleotide binding	3.58E-15	1.8
actin binding	1.59E-12	3.22
binding	5.33E-12	1.15
cytoskeletal protein binding	2.83E-11	2.66
ATP binding	8.86E-10	1.78

Figure S19: Viral integration is bias to genes that are involved in specific functions and processes. See next figure for description.



Description	P-value	Enrichment
translational elongation	4.37E-38	10.12
RNA splicing	2.50E-12	3.4
mRNA metabolic process	4.62E-11	3.02
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	1.26E-10	4.55
RNA splicing, via transesterification reactions	1.26E-10	4.55
nuclear mRNA splicing, via spliceosome	1.26E-10	4.55
mRNA processing	5.81E-10	3.02

Figure S20: Viral integration is bias to genes that are involved in specific functions and processes. Viral integration preference was measured by computing the hyper geometric GO enrichment of genes in our library compared to all human genes. Several function and processes are highly enriched ($P < 10^{-9}$) including RNA binding, ATPAse activity, splicing, translational elongation and ribosomal components. Images generated by the Gorilla software (http://cbl-gorilla.cs.technion.ac.il/).

Supporting tables

Supporting Table 1: Proteins tagged with YFP are full length fusions. Comparison between observed size of YFP-fused protein and expected size for 23 proteins that showed a detectable band when immunoblotted with anti-GFP.

Protein name	Clone ID	Size of YPF-fused protein (kDa)	
		Expected	Observed
CALM1	150506pl1E2	~47 (20+27)	~47
CKS2	010806pl4A1	~47 (10+27)	~48
DDX5	090505pl3D6	~95 (68+27)	~95
	010806pl2F1		
EIF3S12	041206pl1C1	~55 (28+27)	~55
	041206pl5H5		~57
ENO1	150506pl2F1	~77(50+27)	~77
FAU	170407pl2A5	~41 (14+27)	~45
FSCN1	010806pl1E12	~82 (55+27)	~85
GAPDH	310806pl2C2	67 (40+27)	~66
GNB2L1	310806pl1H12	~64 (37+27)	~66
HDAC2	041206pl1E4	~87 (60+27)	~87
HSP90AA1	310506pl1B9	~120 (90+27)	~120
LMNA/C	310806pl1H11	Lamin A: ~96 (69+27)	~96
		Lamin C: ~89 (62+27)	~89
MYH9	010506pl1C12	~227 (200+27)	~227
NPM1	010806pl2H1	~60 (33+27)	~67
PBX3	041206pl3C8	~67 (40+27)	~70
PEPP-2	010806pl2B4	~59 (32+27)	~58
	010806pl2D11		
PPIA	310506pl4C1	~47 (20+27)	~49
	031206pl3B6		~47
RPL18	150506pl1C8	~47 (20+27)	~ 47
RPS3A	150506pl1B7	~63 (36+27)	~66
RPS7	310506pl1G12	~54 (27+27)	~54
TJP1	050707pl1C9	~227 (200+27)	~227
TOP1	200906pl1C12	~117(90+27)	~120
	200306pl1H1		
VPS26A	010506pl1B1	~67 (40+27)	~70
	050707pl1B11		
	211007pl2A8		

Supporting Table 2: List of proteins that changed localization in response to drug

Protein name	Clone ID	Description	Response to CPT
TOP1	200906pl1C12	topoisomerase 1	Rapid decrease in nucleoli followed by decrease in nucleus
NCL	160507pl3D6	nucleolin	Rapid decrease in nucleoli
DNMT1	310806pl2C3	DNA methylase 1	Rapid decrease in nucleoli
ARID1B	041206pl7B1	AT rich interactive domain 1B	Rapid decrease in nucleoli
RPL39	200906pl2H12	Ribosomal protein large subunit 39	Rapid decrease in nucleoli
FAU	170407pl2A5	fusion protein consisting of the ubiquitin-like protein fubi at the N terminus and ribosomal protein S30 at the C terminus	Rapid decrease in nucleoli
UPF3	041206pl6A3	UPF3 regulator of nonsense transcripts homolog A	Rapid increase in nucleoli followed by nucleolar decrease
DDX5	010806pl2F1	RNA helicase p68	Rapid increase in nucleoli followed by nucleolar decrease
C3ORF19	010506pl1E8	hypothetical protein	Rapid increase in nucleoli followed by nucleolar decrease
LPIN2	170407pl1B6	lipin 2	Rapid increase in nucleoli followed by nucleolar decrease
PFDN5	160507pl1E3	prefoldin subunit 5 (c-myc binding protein)	Rapid nuclear decrease followed by nuclear increase
TXNRD1	010506pl1E8	thioredoxin reductase 1	Slow nuclear increase
TXN	160507pl1B6	thioredoxin	Slow nuclear increase

Supporting Table 3: List of proteins showing bimodal behavior

Protein name	Clone ID	description	Reference to association of protein with cell death
ARL3	050707pl3G1	ADP-ribosylation factor-like 3	
BAG2	010806pl1C7	BCL2-associated athanogene 2	
BAG3	170407pl3D4	BCL2-associated athanogene 3	(*15*)
C9ORF40	130207pl1E1	hypothetical protein LOC55071	
CALM1	150506pl1E2	calmodulin 1	(*16, 17*)
CAV1	170407pl1C2	caveolin 1	(*18*)
CCDC23	310506pl2C3	coiled-coil domain containing 23	
DDX5	010806pl2F1	p68 RNA helicase	(*19*)
DKFZP434M1123	160507pl1B11	hypothetical protein	
EIF1AX	010806pl2B11	eukaryotic translation initiation factor 1A, X-linked	
FABP5	200906pl1B6	fatty acid binding protein 5	
FSCN1	010806pl1E12	fascin homolog 1, actin-bundling protein	
PCMTD2	010506pl2D2	protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing	
PDCD5	170407pl1B5	programmed cell death 5	
PFN1	050707pl2E5	profilin 1	
NPM1	010806pl2H1	Nucleophosmin (B23)	(*20*)
PPP1R2	010806pl1G5	protein phosphatase 1, regulatory (inhibitor) subunit 2	
PTTG1	310506pl2C2	pituitary tumor-transforming 1	
RFC1	050707pl1B12	replication factor C (activator 1)	
RPS3	150506pl2B7	ribosomal protein S3	(*21*)
SLBP	010506pl2E6	stem-loop binding protein	(*22*)
SPCS1	050707pl2F4	signal peptidase complex subunit 1 homolog	
TOMM70A	170407pl3H11	translocase of outer mitochondrial membrane 70 homolog A	
YT521	010806pl1F2	YTH domain containing 1	

Supporting Table 4: List of Expressed Sequence Tags (ESTs) and hypothetical proteins and their measured localizations. For each entry (EST or hypothetical protein) we measured the ratio of total fluorescence in the cytoplasm vs. total fluorescence in the whole cell, above or equal to 0.5 is denoted as nuclear localization and below 0.5 as cytoplasmic localization.

	names	Cytoplasm / whole cell	Nucleus	Cytoplasm
1	AA282714	0.7866	0	1
2	AA479512	0.779	0	1
3	AA843465	0.3618	1	0
4	AA928516	0.4001	1	0
5	AF086125	0.8349	0	1
6	AF087973	0.7233	0	1
7	AI027434	0.2965	1	0
8	AI208228	0.7128	0	1
9	AI434862	0.7284	0	1
10	AI671392	0.3552	1	0
11	AI733141	0.5479	0	1
12	AI801879	0.2595	1	0
13	AI870477	0.7639	0	1
14	AK022356	0.6871	0	1
15	AK023312	0.7707	0	1
16	AK023856	0.2276	1	0
17	AK024998	0.6494	0	1
18	AK057505	0.8767	0	1
19	AK091021	0.7426	0	1
20	AK091830	0.6938	0	1
21	AK092541	0.691	0	1
22	AK092875	0.3468	1	0
23	AK095109	0.7859	0	1
24	AK097658	0.3469	1	0
25	AK098306	0.6876	0	1
26	AK124927	0.1741	1	0
27	AK127572	0.5898	0	1
28	AK127877	0.7119	0	1
29	AK130903	0.7623	0	1
30	AK131516	0.8201	0	1
31	AV741821	0.7017	0	1
32	AW070221	0.6662	0	1
33	AW592040	0.8192	0	1
34	AW662723	0.623	0	1
35	AY054401	0.7634	0	1
36	AY176665	0.7225	0	1
37	BC033363	0.8908	0	1
38	BC034424	0.6379	0	1
39	BC035195	0.6273	0	1
40	BC035377	0.4531	1	0

41	BC038752	0.7525	0	1
42	BC039104	0.8318	0	1
43	BC040610	0.7936	0	1
44	BC042060	0.7563	0	1
45	BC042816	0.7201	0	1
46	BC042855	0.8326	0	1
47	BC043574	0.685	0	1
48	BC044257	0.6643	0	1
49	BC044741	0.3626	1	0
50	BC053955	0.6361	0	1
51	BC054862	0.8227	0	1
52	BC078172	0.8116	0	1
53	BC108263	0.8339	0	1
54	BC127846	0.8948	0	1
55	BE745782	0.2625	1	0
56	BE785612	0.7293	0	1
57	BE044435	0.7093	0	1
58	BF062994	0.714	0	1
59	BF245041	0.7327	0	1
60	BF594738	0.2631	1	0
61	BF688062	0.2489	1	0
62	BG189068	0.6341	0	1
63	BG201613	0.194	1	0
64	BG203790	0.2773	1	0
65	BI462136	0.3108	1	0
66	BI559775	0.727	0	1
67	BI825982	0.7214	0	1
68	BM461531	0.4477	1	0
69	BM690995	0.7291	0	1
70	BQ184944	0.7141	0	1
71	BQ233546	0.6304	0	1
72	BU533525	0.6682	0	1
73	BU534173	0.303	1	0
74	BU619815	0.3354	1	0
75	BX089034	0.8095	0	1
76	BX090666	0.7584	0	1
77	BX100329	0.7407	0	1
78	BX100818	0.7962	0	1
79	BX103408	0.3196	1	0
80	BX103636	0.8348	0	1
81	BX104605	0.7985	0	1
82	BX537644	0.7389	0	1
83	BX537772	0.8385	0	1
84	BX648555	0.6607	0	1
85	BX648926	0.3742	1	0
86	C12ORF43	0.3436	1	0
87	C12ORF45	0.3186	1	0

88	C14ORF112	0.4427	1	0
89	C14ORF2	0.7418	0	1
90	C16ORF14	0.4108	1	0
91	C19ORF33	0.622	0	1
92	C19ORF43	0.4308	1	0
93	C19ORF53	0.7672	0	1
94	C19ORF61	0.7063	0	1
95	C20ORF24	0.7255	0	1
96	C21ORF59	0.7483	0	1
97	C2ORF25	0.7598	0	1
98	C3ORF19	0.3994	1	0
99	C3ORF6	0.7952	0	1
100	C7ORF48	0.787	0	1
101	C8ORF44	0.4354	1	0
102	C9ORF40	0.7684	0	1
103	CB045860	0.724	0	1
104	CD692919	0.6126	0	1
105	CN267986	0.6675	0	1
106	CN280387	0.7509	0	1
107	CN398253	0.7986	0	1
108	CR593740	0.7132	0	1
109	CR604408	0.8164	0	1
110	CR623475	0.6816	0	1
111	CR626360	0.7563	0	1
112	CR627148	0.7868	0	1
113	CR737784	0.8232	0	1
114	CR994463	0.659	0	1
115	DB049861	0.8422	0	1
116	DB054822	0.7785	0	1
117	DB186251	0.2773	1	0
118	DB331110	0.2272	1	0
119	DB514539	0.7233	0	1
120	DB522524	0.7956	0	1
121	DC347972	0.6791	0	1
122	DKFZP434M1123	0.8034	0	1
123	EF565105	0.5012	0	1
124	ESTDB089792	0.7495	0	1
125	FLJ10154	0.3163	1	0
126	FLJ10292	0.7617	0	1
127	FLJ20105	0.8155	0	1
128	FLJ31951	0.8475	0	1
129	FLJ32065	0.4847	1	0
130	FLJ35776	0.752	0	1
131	KIAA1064	0.1932	1	0
132	KIAA1430	0.3263	1	0
133	LOC130074	0.6583	0	1
134	LOC284184	0.6882	0	1

135	LOC286016	0.7608	0	1
136	LOC441161	0.6789	0	1
137	LOC541471	0.6685	0	1
138	MGC16824	0.7113	0	1
139	MGC5509	0.7338	0	1
140	MGC71993	0.5969	0	1
141	MRNA-AK098520	0.2283	1	0
142	NM_001093732	0.6534	0	1
143	NM_015681	0.6197	0	1
144	T85821	0.7951	0	1
145	T85822	0.7259	0	1
146	T85823	0.815	0	1
147	T85824	0.8146	0	1
148	AI342698	0.6337	0	1
149	AK094352	0.6052	0	1
150	AK094903	0.3903	1	0
151	AK128457	0.3942	1	0
152	AW418496	0.4929	1	0
153	AX748230	0.7376	0	1
154	BC005233	0.5561	0	1
155	BC036259	0.6996	0	1
156	BG221753	0.6439	0	1
157	BX648475	0.795	0	1
158	C12ORF48	0.3315	1	0
159	C14ORF166	0.6383	0	1
160	C6ORF106	0.5592	0	1
161	C7ORF11	0.4211	1	0
162	LOC729416	0.7143	0	1
163	N68399	0.6699	0	1
164	NT_022171	0.6871	0	1

Supporting references

1. N. Otsu, *IEEE Transactions on Systems, Man, and Cybernetics* **9**, 62-66 (1979).
2. J. W. Jarvik, S. A. Adler, C. A. Telmer, V. Subramaniam, A. J. Lopez, *Biotechniques* **20**, 896-904 (May, 1996).
3. J. W. Jarvik *et al.*, *Biotechniques* **33**, 852-4, 856, 858-60 passim (Oct, 2002).
4. A. Sigal *et al.*, *Nat Protoc* **2**, 1515-27 (2007).
5. N. C. Shaner *et al.*, *Nat Biotechnol* **22**, 1567-72 (Dec, 2004).
6. T. Nagai *et al.*, *Nat Biotechnol* **20**, 87-90 (Jan, 2002).
7. A. Sigal *et al.*, *Nat Methods* **3**, 525-31 (Jul, 2006).
8. E. Eden, Rudzsky, M., Brod, V., Waisman, D., Bitterman, H., Sabo, E. and Rivlin E, *IEEE, Transactions on Medical Imaging* **24**, 1011-1024 (2005, 2005).
9. E. Borenfreund, J. A. Puerner, *Toxicol Lett* **24**, 119-24 (Feb-Mar, 1985).
10. S. Z. Zhang, M. M. Lipsky, B. F. Trump, I. C. Hsu, *Cell Biol Toxicol* **6**, 219-34 (Apr, 1990).
11. R. K. Blashfield, *Journal of Classification* **8**, 277-279 (1991).
12. M. Ashburner *et al.*, *Nat Genet* **25**, 25-9 (May, 2000).
13. E. Eden, D. Lipson, S. Yogev, Z. Yakhini, *PLoS Comput Biol* **3**, e39 (Mar 23, 2007).
14. W. K. Huh *et al.*, *Nature* **425**, 686-91 (Oct 16, 2003).
15. P. Bonelli *et al.*, *Leukemia* **18**, 358-60 (Feb, 2004).
16. O. Cohen, E. Feinstein, A. Kimchi, *Embo J* **16**, 998-1008 (Mar 3, 1997).
17. Y. Shirasaki, Y. Kanazawa, Y. Morishima, M. Makino, *Brain Res* **1083**, 189-95 (Apr 14, 2006).
18. C. C. Ho *et al.*, *Lung Cancer* **59**, 105-10 (Jan, 2008).
19. L. Yang, C. Lin, S. Y. Sun, S. Zhao, Z. R. Liu, *Oncogene* **26**, 6082-92 (Sep 6, 2007).
20. Y. Qing, G. Yingmao, B. Lujun, L. Shaoling, *J Neurol Sci* **266**, 131-7 (Mar 15, 2008).
21. C. Y. Jang, J. Y. Lee, J. Kim, *FEBS Lett* **560**, 81-5 (Feb 27, 2004).
22. Y. Kodama, J. H. Rothman, A. Sugimoto, M. Yamamoto, *Development* **129**, 187-96 (Jan, 2002).

Supporting movies

Supporting Movie 1:
Time-lapse movie of transmitted light images of the clone with YFP CD-tagged DDX5. Movie duration is 72 hours, 24 hours before drug addition followed by 48 hours after drug addition (time-lapse: 1 frame per 20 minutes).

Supporting Movie 2:
Time-lapse movie of yellow fluorescence images of the clone with YFP CD-tagged DDX5. Movie duration is 72 hours, 24 hours before drug addition followed by 48 hours after drug addition (time-lapse: 1 frame per 20 minutes).

Supporting Movie 3:
Time-lapse movie of red fluorescence images of the clone with YFP CD-tagged DDX5. Movie duration is 72 hours, 24 hours before drug addition followed by 48 hours after drug addition (time-lapse: 1 frame per 20 minutes).

Supporting Movie 4:
Time-lapse movie of segmentation and tracking of the clone with YFP CD-tagged DDX5. Movie duration is 72 hours, 24 hours before drug addition followed by 48 hours after drug addition (time-lapse: 1 frame per 20 minutes).