Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins

Alex Sigal[1,6], Ron Milo[1,2,6], Ariel Cohen[1], Naama Geva-Zatorsky[1], Yael Klein[1], Inbal Alaluf[1], Naamah Swerdlin[1], Natalie Perzov[1], Tamar Danon[1], Yuvalal Liron[1], Tal Raveh[3], Anne E Carpenter[4], Galit Lahav[5] & Uri Alon[1,2]

We examined cell cycle–dependent changes in the proteome of human cells by systematically measuring protein dynamics in individual living cells. We used time-lapse microscopy to measure the dynamics of a random subset of 20 nuclear proteins, each tagged with yellow fluorescent protein (YFP) at its endogenous chromosomal location. We synchronized the cells *in silico* by aligning protein dynamics in each cell between consecutive divisions. We observed widespread (40%) cell-cycle dependence of nuclear protein levels and detected previously unknown cell cycle–dependent localization changes. This approach to dynamic proteomics can aid in discovery and accurate quantification of the extensive regulation of protein concentration and localization in individual living cells.

A long-term goal of biology is a quantitative understanding of gene and protein networks of human cells and their responses to stimuli. This requires the ability to accurately measure protein levels[1–4] and localizations[5–12]. An ideal proteomic measurement system would (i) work at the level of individual cells, as experiments that average over cell populations can overlook events happening in only a subset of cells, all-or-none effects and variability between cells; (ii) follow the same cell over extended periods of time to uncover phenomena such as oscillations and temporal programs; (iii) make minimal perturbations to the state of the cell and (iv) measure both protein levels and localization.

One cellular process that produces widespread changes in the proteome is the cell cycle. Systematic studies of cell cycle–regulated genes using DNA arrays showed that about 10% of genes were cell cycle dependent in yeast and *Arabidopsis thaliana*[13–16], compared with 2–3% of genes in human cells (whether primary fibroblast[17] or HeLa carcinoma[18]). These differences may reflect fundamental differences between plant and yeast cells on one hand and human cells on the other, or increased experimental noise arising from problems in synchronization of human cells relative to yeast and *A. thaliana*.

We examined the cell-cycle dependence of nuclear proteins (which constitute 90% of known cell cycle–dependent proteins[18]) as a proof of principle for a new dynamic proteomics approach in individual living cells. Our approach was based on automated time-lapse microscopy and image analysis of a library of cells, each with a different fluorescently tagged protein expressed from its endogenous chromosomal location. We were able to use a non-perturbing method of synchronization, monitoring tagged protein levels in non-synchronized individual cells from one cell division to the next and retroactively aligning protein dynamics of all cells between consecutive division events. We discovered that in the random subset of nuclear proteins we examined, 40% showed cell cycle–dependent dynamics.

RESULTS

Construction of a YFP CD-tagged reporter clone library

To measure protein level and localization in individual living cells, we developed a system for dynamic proteomics. We generated a library of cell clones in which each clone contained a different fluorescently tagged protein expressed from its endogenous chromosomal location. We fluorescently labeled endogenous proteins by the central dogma (CD) tagging approach[19,20]. For tagging, we used a retrovirus to insert the YFP coding region flanked by splice signals into the genome of H1299 lung carcinoma cells in a nondirected manner (**Fig. 1a**). When integrated in the proper frame and orientation into an intron of an expressed gene, the YFP was spliced into the mRNA as a new exon. This resulted in a full-length fluorescent protein fusion expressed from its endogenous chromosomal locus. The CD-tagging method allows construction of tagged mammalian cell libraries, resembling the library of endogenous GFP protein fusions constructed in yeast by homologous recombination[11]. Unlike gene trap methods[21–24], this tagging approach results in a full-length protein. CD-tagging preserves wild-type localization and function in a large fraction of the tagged proteins[19,25,26] and avoids overexpression concerns associated with fluorescent protein fusions expressed from exogenous promoters.

After tagging with the virus, we selected cells with tagged proteins by flow cytometric sorting of YFP-positive cells into multiwell plates; each tagged cell was expanded into a clone (**Fig. 1b,c**). In

[1]Department of Molecular Cell Biology and [2]Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel. [3]Department of Developmental Biology, Stanford University, Palo Alto, California 94305-5439, USA. [4]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA. [5]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. [6]These authors contributed equally to this work. Correspondence should be addressed to U.A. (uri.alon@weizmann.ac.il).

RECEIVED 22 MARCH; ACCEPTED 23 MAY; PUBLISHED ONLINE 21 JUNE 2006; DOI:10.1038/NMETH892

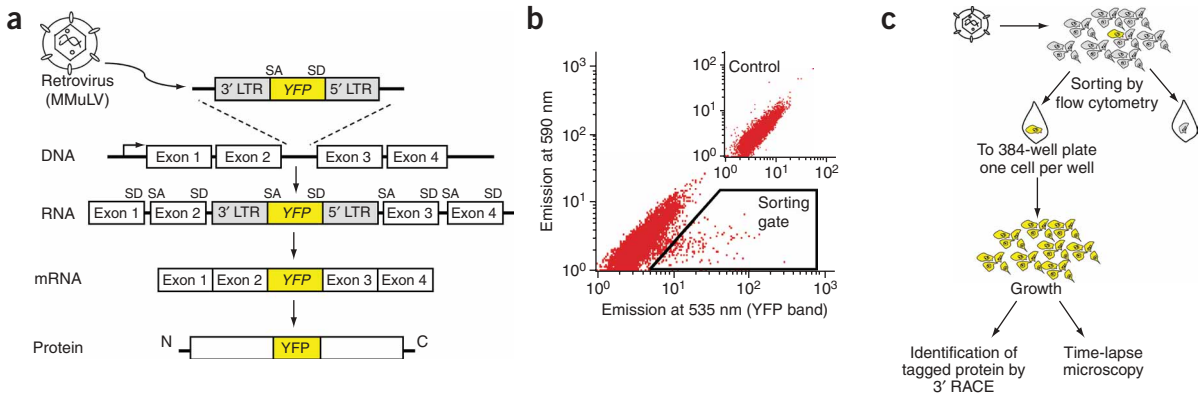

Figure 1 | Creation of a library of CD-tagged proteins. (**a**) CD-tagging with a retrovirally delivered YFP with no promoter, no start codon and no poly(A) signal, flanked by splice acceptor (SA) and splice donor (SD) sequences. YFP is retained as a new exon after splicing of the mRNA, leading to an internally labeled protein. MMuLV, Moloney murine leukemia virus. (**b**) Sorting of cells that contain CD-tagged proteins by flow cytometry. Tagged cells were collected by a sorting for high 535 nm (YFP) to 590 nm (autofluorescence) emission ratio. Inset shows cells infected with an empty retroviral vector as a control. The percentage of YFP-positive cells was about 0.5–1% for different experiments. Therefore, the likelihood that a cell contained two detectable fluorescently tagged proteins was small. (**c**) Cells that fell into the sorting gate shown in **b** were sorted at one cell per well into 384-well plates and expanded into clones. The YFP-tagged protein in each clone was identified by 3′ RACE and analyzed by time-lapse microscopy.

order to prevent multiple productive tagging events, multiplicity of infection was controlled so that $<1\%$ of cells were YFP positive. We identified the tagged protein in each clone by 3′ RACE[27], based on the amplification of the mRNA transcript of the tagged gene using nested PCR primers for YFP and for the poly(A) tail. We established a library of 200 cell clones, each with a different identified tagged protein (database of clones available online; http://www.dynamicproteomics.net. Additional data on localizations and integration sites are available in **Supplementary Fig. 1** online). We determined the localization of the protein in each labeled clone by fluorescence microscopy. The cell library contained proteins in diverse localizations. Of the proteins with previously characterized localizations in mammalian cells, 74% showed the same localization patterns with CD-tagging (**Supplementary Table 1** online). The library included 18 different ribosomal proteins. Of these proteins, two-thirds showed the correct cytoplasmic and nucleolar localization[28].

Quantitation and *in silico* synchronization of protein dynamics

We imaged 20 YFP CD-tagged nuclear proteins from the library using fluorescence microscopy under controlled CO_2, temperature and humidity. We concentrated on nuclear proteins in this study because image analysis of the fluorescence microscopy movies is readily feasible for nuclear localization patterns, owing to the regular shape of the nucleus and its distance from other nuclei. Of the clones in the library, we randomly chose 20 nuclear proteins.

We recorded movies of clones containing tagged proteins with a 20× objective at a temporal resolution of 10 min for 48–60 h. Cells showed normal proliferation over this period. Cell division events occurred throughout the movie, with a cell cycle time of 18 ± 2.5 h (for cell-cycle time distributions of individual clones, see **Supplementary Fig. 2** online). Each movie captured the dynamics of an initial population of 10–20 cells from a clone. This initial population divided to generate up to 150 cells. Examples of the dynamics of nine different tagged proteins with diverse localization patterns over one full cell cycle are shown in **Supplementary Video 1** online, and an example of a field of view of one clone with a labeled nuclear protein (regulatory histone H2AFV) is shown in **Supplementary Video 2** (transmitted light) and **Supplementary Video 3** (fluorescence) online. We developed automated image analysis tools for segmentation and tracking to quantify nuclear protein levels (**Fig. 2**, Methods). For a detailed description of the segmentation and tracking algorithms used, see **Supplementary Methods** online.

Nuclear fluorescence of all proteins increased over the cell cycle and showed an abrupt drop after each division to about half of the maximal level, as a result of protein segregation to the two daughter cells (**Fig. 2c**). We synchronized cells *in silico* by aligning the dynamics between two division events and using the fraction of the cell cycle that has elapsed as a time base instead of absolute time (**Fig. 2d**). Because cells differed in their cell-cycle timing and duration, *in silico* synchronization allowed a normalized and aligned time base with respect to the cell-cycle duration of each particular cell.

In silico synchronization allowed us to quantify the average rates of change of protein concentration in the nucleus of each cell over the course of the cell cycle. The rate of change, which we denote the nuclear accumulation rate, was defined as the temporal derivative of the nuclear protein fluorescence level divided by the nuclear area (see Methods). The nuclear accumulation rate is therefore the net result of changes in production, degradation and localization. We aligned nuclear accumulation rates between the cell division events of each cell and averaged them.

Cell-cycle dependence of nuclear proteins

To determine whether the nuclear accumulation profile obtained by *in silico* synchronization was cell cycle–dependent, we used criteria that took into account both experimental error and biological differences between clones (see Methods). We determined the average duration of the G1, S and G2 cell-cycle phases as a percentage of the cell cycle by flow cytometric analysis of DNA-stained cells (**Supplementary Fig. 3** online).

We scored proteins as cell cycle–dependent (**Fig. 3**) using two different tests. One was based on estimating the error from

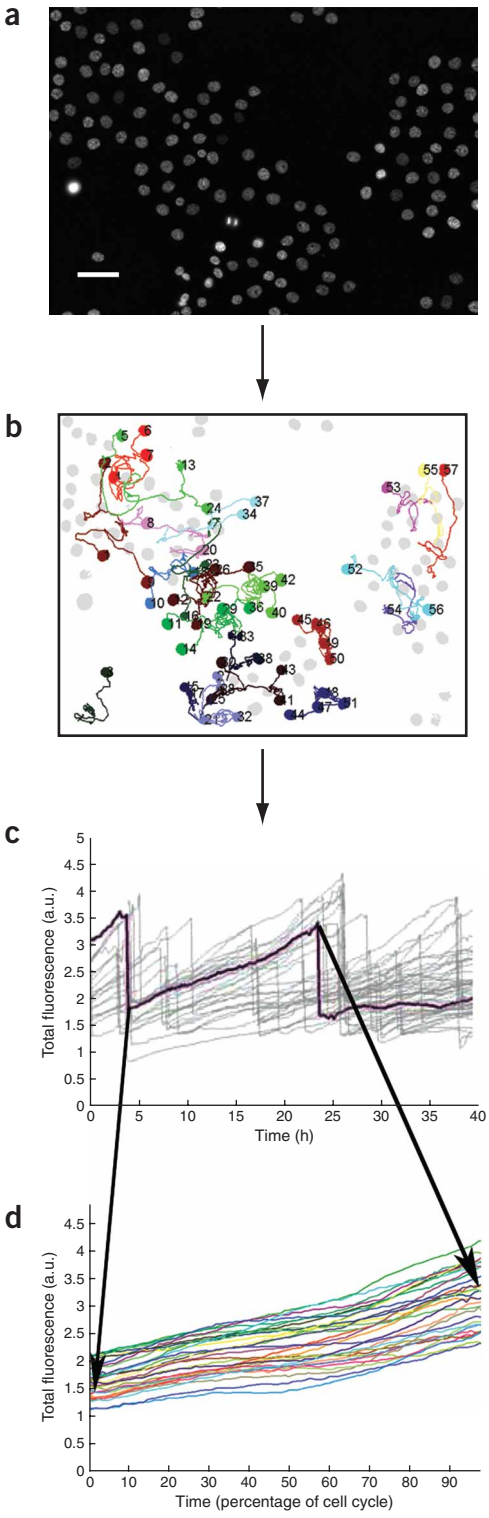

Figure 2 | *In silico* synchronization of individual cells. (**a**) Final frame from a time-lapse movie of tagged histone H2AFV. Movies were acquired with a resolution of 10 min and duration of 48–60 h. Nuclei were tracked and their total fluorescence determined in each frame. Bar, 50 μm. (**b**) Tracks of the cell nuclei in **a**, shown as lines. Cells with the same color and connected by a line are descended from the same parental cell. Cells in gray were not successfully analyzed. (**c**) Total fluorescence of each nucleus as a function of time. The dynamics of one typical nucleus is highlighted. Abrupt decreases are cell division events. (**d**) Cells were synchronized *in silico* by aligning the nuclear fluorescence between two cell divisions. The *x* axis shows the percentage of the cell cycle elapsed since the first division.

differences in cell-cycle pattern from independently derived clones with the same tagged protein. The other test used a *t*-test, based on within-clone variations (see Methods). We found significant ($P < 0.05$) cell cycle–dependent changes in 40% (8/20) of the proteins (**Fig. 3a–h**). The nuclear protein levels showed different types of dynamic profiles. We observed nearly sinusoidal variation, with increased accumulation at S phase, for the uncharacterized protein FLJ10154 and for the topoisomerase TOP1 (**Fig. 3a,d**). Other proteins showed high accumulation in G1 (**Fig. 3c,e,h**) or G2 (**Fig. 3b,f**).

We compared the results obtained here with mRNA profiles from a previous microarray study of the same genes in HeLa cells[18] (**Fig. 3**). The mRNA levels for three of the genes (**Fig. 3a,d,e**) were scored as cell cycle–dependent in the microarray study; the corresponding proteins were also scored as cell cycle–dependent in our study. An additional five genes that scored as non–cell cycle–dependent in the microarray study scored as cell cycle–dependent at the protein level in our study (**Fig. 3b,c,f–h**). Thus, we seem to detect two- to threefold more cell cycle–dependent genes with the present approach.

Cell cycle–dependent changes in localization

Changes in protein levels associated with the cell cycle may involve not only changes in protein accumulation rates but also changes in localization. We examined cell cycle–dependent localization changes using a display termed a 'synchrogram'. The synchrogram allows comparison of localization dynamics across the cell cycle, spanning the time between two consecutive division events (**Fig. 4a**). This display is an effective way to observe localization dynamics but not changes in protein amounts (which are hard to distinguish by eye and are best quantified). In **Figure 4a**, we show a subset of labeled proteins in the library that includes several different cellular localizations. These include the Na-K channel component ATP1A1, localized to the plasma membrane; moesin (MSN), a cytoplasmic protein linking the cytoskeleton to receptors on the plasma membrane; RTN4, a protein found in the endoplasmic reticulum; RPL4, ribosomal protein found in the cytoplasm and nucleoli; fibrillarin (FBL), involved in nucleolar organization; LaminB1 (LMNB1), one of the components of the nuclear envelope; and the nuclear proteins LMNA and USP7 and the regulatory histone H2AFV. This display is generated by custom software when cell locations and division times have been either automatically or manually determined (see Methods).

Proteins show diverse cell cycle–dependent localization patterns as the cells go through defined events in the cell cycle, such as the localization of LMNB1 to the nuclear envelope after 1–2% of the cell cycle has elapsed, decondensation and condensation of DNA associated H2AFV (after 3–5% and 97% of the cell cycle has elapsed, respectively) and the nucleolar localization of FBL (3–5% elapsed) and RPL4 (5–10% elapsed).

To illustrate one example of cell cycle–dependent translocation in detail, we focus on USP7, a ubiquitin-specific nuclear protease. This protein is known to assume either a diffuse localization in the nucleus or localize to nuclear bodies[29]. We find that tagged USP7 showed a previously unknown cell-cycle dependence in

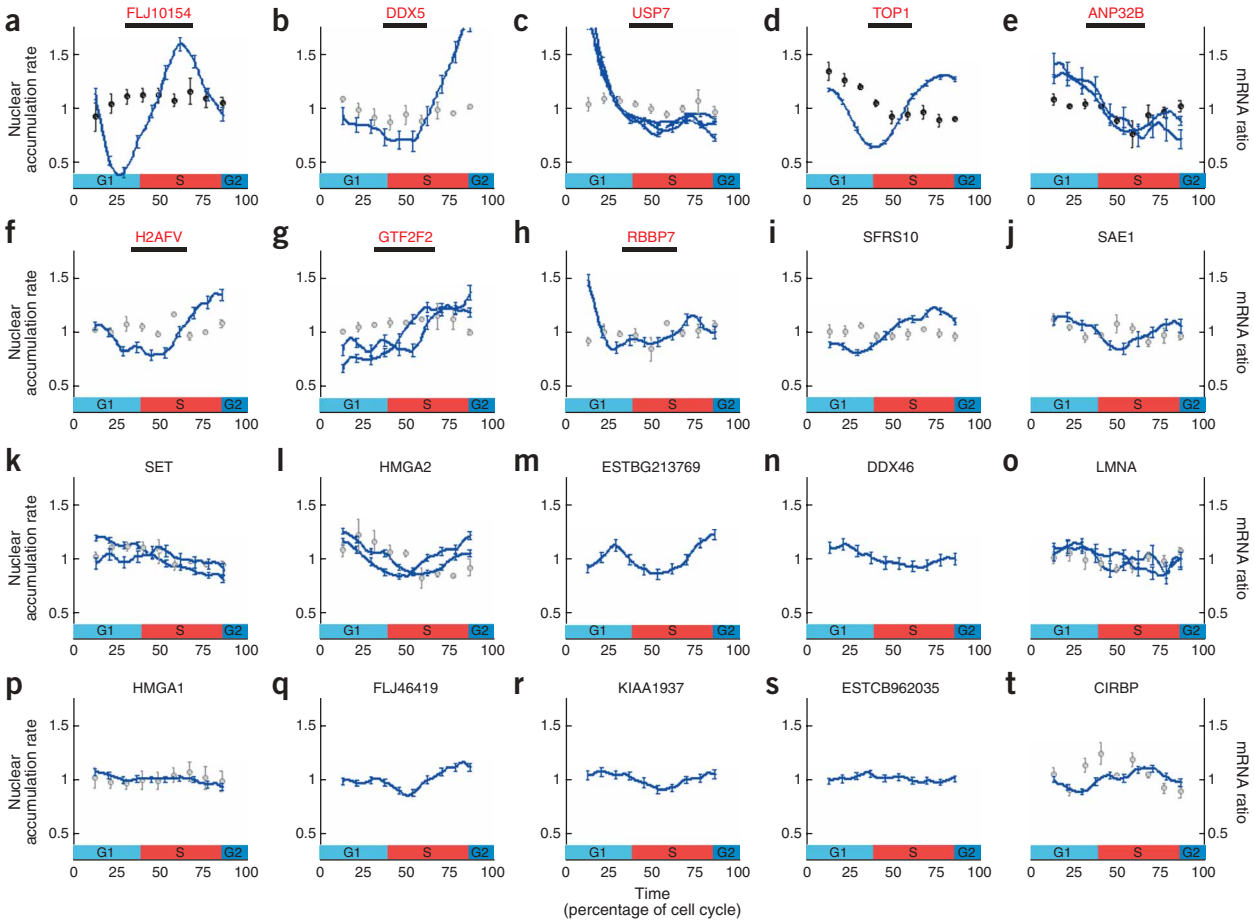

Figure 3 | Nuclear accumulation as a function of cell cycle for 20 nuclear proteins. Blue lines represent average nuclear accumulation rate ± s.e.m. for each clone. Proteins in **a**–**h** (red, underlined names) have significant ($P < 0.05$) cell cycle–dependent nuclear accumulation rates (see Methods). Two or more lines for the same protein denote independently derived clones in which the same protein was labeled. In the clones expressing HMGA2, integrations occurred in different introns. Proteins are shown in order of decreasing statistical significance with respect to cell-cycle dependence: (**a**) uncharacterized protein FLJ10154, (**b**) RNA helicase DDX5, (**c**) ubiquitin-specific protease USP7, (**d**) topoisomerase TOP1, (**e**) uncharacterized protein ANP32B, (**f**) regulatory histone H2AFV, (**g**) general transcription factor GTF2F2, (**h**) retinoblastoma binding protein RBBP7, (**i**) splicing factor SFRS10, (**j**) SUMO-1 activating enzyme subunit SAE1, (**k**) histone binding protein SET, (**l**) the transcription regulating factor HMGA2, (**m**) an unknown protein corresponding to ESTBG213769, (**n**) RNA helicase DDX46, (**o**) nuclear intermediate filament LMNA, (**p**) transcription regulating factor HMGA1, (**q**) uncharacterized protein FLJ46419, (**r**) uncharacterized protein KIAA1937, (**s**) unknown protein corresponding to ESTCB962035 and (**t**) RNA binding protein CIRBP. Relative fraction of time spent by cells in G1, S and G2 cell cycle phases was determined by flow cytometric analysis of DNA content. Gray and black circles denote average mRNA levels ± s.e.m. as reported by microarrays of chemically synchronized HeLa cells[18]; gray circles represent mRNA levels scored as cell cycle–independent by microarrays, and black circles mRNA levels scored as cell cycle–dependent.

translocation to nuclear bodies (**Fig. 4b**; see also **Supplementary Video 4** online). These bodies appeared at about the halfway point of the cell cycle. **Figure 4c** shows the appearance of nuclear body localization in 79 different cells. The majority of cells showed nuclear body localization only at the 40–95% cell cycle fraction. Notably, there was some cell-to-cell variability in localization: some cells showed nuclear body localization through most of the cell cycle, whereas others showed it for only brief periods or not at all.

DISCUSSION

Observation of dynamic processes at the population level, such as cell-cycle dependence of proteins, is challenging when these processes are not synchronized between cells. Looking at protein dynamics in individual cells is perhaps the most direct approach to investigate nonsynchronized processes. Observing protein behavior in single cells, combined with accurate quantification and *in silico* synchronization, allowed us to detect widespread changes in the amount and localization of proteins as a function of the cell-cycle stage.

The present study uses a new method for retroactive synchronization of cells based on detection of division events in individual cells in time-lapse movies. This approach avoids mechanical or chemical synchronization of cells, which have the potential to alter their physiological state[30]. The present approach also avoids a common problem of synchronization in which cells begin to lose their initial synchronization to each other even in the first cell cycle. With *in silico* synchronization, cells can be synchronized for as long as they grow exponentially in the time-lapse movie. Such retroactive *in silico* synchronization is not limited to the cell cycle and can be extended to align protein behavior with respect to any cellular event with automatically detectable start and end points. In

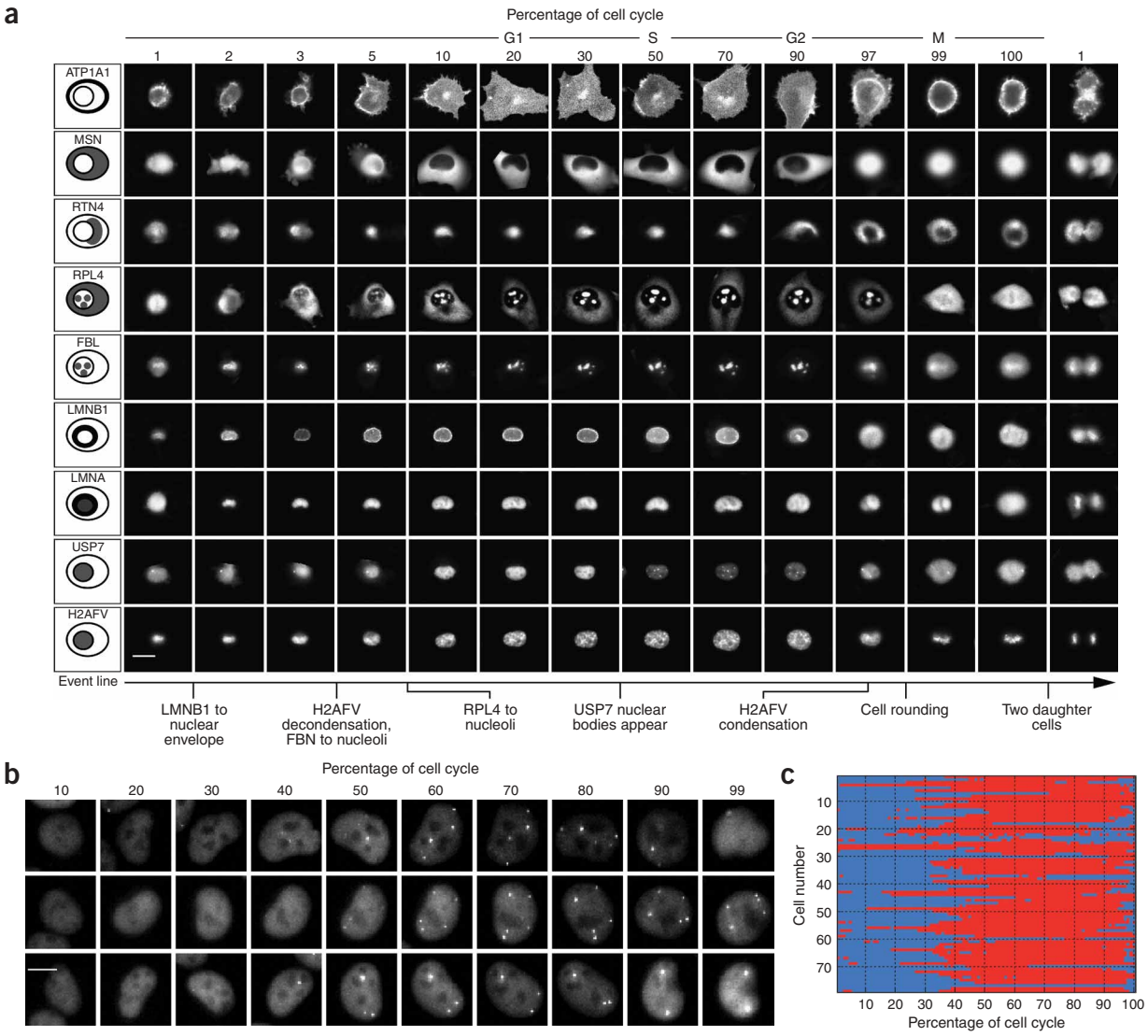

Figure 4 | Cell cycle–dependent translocation of proteins. (**a**) The synchrogram shows one cell for each tagged protein. Images are ordered according to the fraction of the time elapsed between two division events. The cells are automatically centered and aligned so that their long axes are parallel. Four nuclear proteins are included, as well as five other proteins: Na-K channel component ATP1A1 (plasma membrane), MSN (moesin, cytoplasmic), RTN4 (endoplasmic reticulum), RPL4 (ribosomal protein, cytoplasm and nucleoli) and FBL (fibrillarin, nucleolar). The percentage of elapsed cell cycle and cell cycle stage are indicated at the top, and an event line at the bottom of the synchrogram marks several visible changes in protein localization. Bar, 20 μm. (**b**) Cell cycle–dependent appearance of nuclear bodies in USP7. Shown are three typical cells. Bar, 10 μm. (**c**) Presence of nuclear bodies in 79 cells as a function of the cell cycle. The vertical axis represents individual cells. Each row represents the state of one cell. Red denotes the existence of nuclear bodies, and blue denotes their absence.

addition, our approach detects regulation at the protein level, including changes in localization.

We are presently extending this dynamic proteomics approach to cover a large fraction of the proteins expressed in this cell line. The present CD-tagging protocol can generate ~1,000 tagged and identified clones in a few months. High-speed and high-throughput microscopy systems can generate time-lapse movies on hundreds of clones per week per microscopy station[31,32]. We therefore believe that with the present approach it will become feasible to study proteome dynamics of human cells under diverse conditions and stimuli.

In conclusion, this study presents a system for dynamic proteomics in individual living cells, based on a library of CD-tagged clones and automated microscopy, combined with image analysis. We found that many of the tested nuclear proteins show cell cycle–dependent changes in their nuclear accumulation rate. This study indicates that single-cell dynamic proteomics can complement present microarray and proteomic methods for a high-resolution, quantitative view of protein networks in individual living human cells.

METHODS

Long period time-lapse microscopy. We obtained time-lapse movies with a Leica DMIRE2 inverted fluorescence microscope (Leica Microsystems) equipped with a 20×, 0.7 numerical aperture (NA) phase objective and a filter cube containing a filter set

for YFP (Chroma), a motorized stage (Maertzhaeuser), automated Uniblitz shutters (Vincent Associates) and an ORCA ER cooled charge-coupled device (CCD) camera (Hamamatsu Photonics). The microscope was placed within a temperature-controlled incubator, with an internal chamber for regulated CO_2 and humidity (Leica Microsystems). The system was controlled by ImagePro5 Plus (Media Cybernetics) macros, which integrated time-lapse acquisition, stage movement and software-based autofocus. During the experiment, cells were grown and visualized in 12-well coverslip-bottom plates (MatTek) coated with 0.001% fibronectin (Sigma-Aldrich). The normal RPMI medium was replaced with RPMI without phenol red (Biological Industries) to decrease autofluorescence.

Image analysis of time-lapse movies. Full details of the image analysis algorithm are given in **Supplementary Methods**. We used a custom-written image analysis tool developed using the Matlab image processing toolbox environment (Mathworks). The main steps were the segmentation of the cells, tracking and detection of cell division events. Segmentation code used a version of an algorithm now available and documented in the open-source software package CellProfiler (http://www.cellprofiler.org; A.E.C., unpublished data). The algorithm is based on a combination of thresholding and watershed segmentation; we applied it to images after performing flat field correction and background subtraction. An example of nucleus segmentation is shown in **Supplementary Figure 4** online.

Cell division was detected by a sharp twofold drop in total fluorescence level between consecutive images. The main parameters measured for each cell nucleus were the x and y centroid locations, cell area, total fluorescence, and mean and variance of the fluorescence.

Calculation of nuclear accumulation rates. The production rate at time t was calculated by the difference equation $(f(t + \tau/2) - f(t - \tau/2)) / \tau$, where $f(t)$ is the total fluorescence at time t. Both t and τ are in units of 'percentage of elapsed cell cycle'. We used a value of $\tau = 10\%$; we obtained similar results for values of $\tau = 6\%$, $\tau = 20\%$ and $\tau = 30\%$. We did not calculate production rates at times $t < 15\%$ or $t > 85\%$ of the cell cycle because of fluctuations in the segmentation near the division event. For each protein, the nuclear accumulation rate calculation is based on tracks of 50–400 cells. Each track comprises $\sim$100 frames covering the complete cell cycle. We obtained $>$5,000 full cell cycles altogether, using $>$500,000 cell segments. The nuclear accumulation rate was defined as the temporal derivative of the nuclear protein fluorescence level divided by the nuclear area. To correct for systematic measurement errors on the order of 20%, we normalized all curves by a standard curve: the nuclear accumulation rate was normalized by the median accumulation profile of the 25% least-variable proteins. We obtained similar scoring of cell cycle profiles when we omitted normalization (**Supplementary Fig. 5** online).

Analysis of cell-cycle dependence. To determine whether a nuclear accumulation profile was cell cycle dependent, we used criteria that took into account both experimental error and biological differences between clones. This was possible because some proteins were labeled more than once in the library. For 6 out of the 20 proteins examined (**Fig. 3c**,**e**,**g**,**k**,**l**,**o**), we measured the nuclear accumulation dynamics of 2–3 clones in which the same protein was tagged. We then estimated the variability in the accumulation rate between clones of the same protein (interclone variability). The nuclear accumulation of proteins was scored as significant if its root mean squared deviation from a constant accumulation rate was two or more interclone standard deviations above the mean interclone variability (that is, $P < 0.05$). Note that this criterion is stringent in the sense that brief dynamic changes are not picked up. Therefore, we also included one protein (RBBP7) that obviously had a cell cycle–dependent profile with variation mainly at one point.

We used a second criterion to assess cell-cycle dependence, similar to that used in microarray studies. Cell-cycle dependence was defined based on a criterion of at least twofold difference in accumulation rate across the cell cycle and a difference of at least eight standard errors between highest and lowest nuclear accumulation rates. Such a definition gave very similar results to the first method described above: all proteins that were scored as cell cycle–dependent according to the first criterion also passed the second criterion. The second criterion also scored as cell cycle–dependent an additional three proteins: HMGA2, RBBP7 and SFRS10. We found similar results when we considered nuclear accumulation rate without normalization to the cell area or to a standard curve (**Supplementary Fig. 5** online).

Synchrogram visualization. Cells were tracked either automatically using the custom-written Matlab image analysis software described above (for nuclear proteins) or using MetaMorph (Universal Imaging) software (for other localizations). We used custom Matlab code to create the synchrogram. The beginning and end of a cell cycle were marked by a twofold decrease in total fluorescence. Linear interpolation between these time markers served to find the frames that represented any given percentage of the cell cycle. We automatically or manually separated each cell image from its neighboring cells and deleted the neighboring cell images. We then centered the cell image and rotated it to place the cell's long axis in the horizontal direction. **Supplementary Figure 6** online shows the synchrogram without this manipulation. Each frame shows an 80 $\times$ 80 pixel (40 $\times$ 40 μm) neighborhood centered at the cell's centroid. We normalized the pixel intensity scale for each image to keep a small fixed percentage of the pixels in saturation (usually 0.1%, to avoid image darkening by singular bright pixels). This scaling allowed spatial details to be seen for clones with different levels of signal. Some proteins showed an apparent increase in total intensity (for example, as RPL4 did, between 97% to 99% of the cell cycle) owing to redistribution fluorescence leading to changes in the maximal pixel intensity and a corresponding change in the scaling.

Quantitation of nuclear bodies. To detect nuclear bodies, we calculated for each cell the ratio of fluorescence intensity between the top five pixels in the nucleus and the average fluorescence within the nucleus. Cells with a ratio $>$1.5 were scored as containing nuclear bodies. Cells without nuclear bodies had a ratio of around 1.2.

Additional information. Descriptions of retroviral constructs, cells and infection; sorting of YFP fluorescent cells by flow cytometry; identification of tagged protein by 3′ RACE; analysis of DNA content by flow cytometry to estimate relative duration of

cell cycle phases and image analysis and fluorescence quantification are available in **Supplementary Methods**.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank Quantomix for plasmid construction; Z. Kam for microscopy guidance; B. Zimmerman for cytoskeleton staining; E. Ariel and A. Sharp from the Weizmann flow cytometry unit for assistance, and M. Springer, P. Bordalo, O. Zuk and members of the Alon lab for discussions of the manuscript. We thank the Kahn Family Foundation and the Israel Science Foundation for support. R.M. thanks the Horowitz Complexity Science Foundation for support. A.E.C. thanks the Novartis Life Science Foundation for support.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Andersen, J.S. *et al.* Nucleolar proteome dynamics. *Nature* **433**, 77–83 (2005).
2. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
3. Ideker, T. *et al.* Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929–934 (2001).
4. Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. & Elowitz, M.B. Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005).
5. Perlman, Z.E. *et al.* Multidimensional drug profiling by automated microscopy. *Science* **306**, 1194–1198 (2004).
6. Bannasch, D. *et al.* LIFEdb: a database for functional genomics experiments integrating information from external sources, and serving as a sample tracking system. *Nucleic Acids Res.* **32**, D505–D508 (2004).
7. Gerlich, D., Beaudouin, J., Gebhard, M., Ellenberg, J. & Eils, R. Four-dimensional imaging and quantitative reconstruction to analyse complex spatiotemporal processes in live cells. *Nat. Cell Biol.* **3**, 852–855 (2001).
8. Tvarusko, W. *et al.* Time-resolved analysis and visualization of dynamic processes in living cells. *Proc. Natl. Acad. Sci. USA* **96**, 7950–7955 (1999).
9. Conrad, C. *et al.* Automatic identification of subcellular phenotypes on human cell arrays. *Genome Res.* **14**, 1130–1136 (2004).
10. Kumar, A. *et al.* Subcellular localization of the yeast proteome. *Genes Dev.* **16**, 707–719 (2002).
11. Huh, W.K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
12. Chen, X. & Murphy, R.F. Objective clustering of proteins based on subcellular location patterns. *J. Biomed. Biotechnol.* **2005**, 87–95 (2005).
13. Menges, M., de Jager, S.M., Gruissem, W. & Murray, J.A. Global analysis of the core cell-cycle regulators of *Arabidopsis* identifies novel genes, reveals multiple and highly specific profiles of expression and provides a coherent model for plant cell-cycle control. *Plant J.* **41**, 546–566 (2005).
14. Spellman, P.T. *et al.* Comprehensive identification of cell-cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–3297 (1998).
15. Cho, R.J. *et al.* A genome-wide transcriptional analysis of the mitotic cell-cycle. *Mol. Cell* **2**, 65–73 (1998).
16. Rustici, G. *et al.* Periodic gene expression program of the fission yeast cell-cycle. *Nat. Genet.* **36**, 809–817 (2004).
17. Cho, R.J. *et al.* Transcriptional regulation and function during the human cell-cycle. *Nat. Genet.* **27**, 48–54 (2001).
18. Whitfield, M.L. *et al.* Identification of genes periodically expressed in the human cell-cycle and their expression in tumors. *Mol. Biol. Cell* **13**, 1977–2000 (2002).
19. Jarvik, J.W. *et al. In vivo* functional proteomics: mammalian genome annotation using CD-tagging. *Biotechniques* **33**, 852–860 (2002).
20. Jarvik, J.W., Adler, S.A., Telmer, C.A., Subramaniam, V. & Lopez, A.J. CD-tagging: a new approach to gene and protein discovery and analysis. *Biotechniques* **20**, 896–904 (1996).
21. Scheel, J.R., Ray, J., Gage, F.H. & Barlow, C. Quantitative analysis of gene expression in living adult neural stem cells by gene trapping. *Nat. Methods* **2**, 363–370 (2005).
22. Whitney, M. *et al.* A genome-wide functional assay of signal transduction in living mammalian cells. *Nat. Biotechnol.* **16**, 1329–1333 (1998).
23. Friedrich, G. & Soriano, P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5**, 1513–1523 (1991).
24. Skarnes, W.C., Auerbach, B.A. & Joyner, A.L. A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev.* **6**, 903–918 (1992).
25. Clyne, P.J., Brotman, J.S., Sweeney, S.T. & Davis, G. Green fluorescent protein tagging *Drosophila* proteins at their native genomic loci with small P elements. *Genetics* **165**, 1433–1441 (2003).
26. Morin, X., Daneman, R., Zavortink, M. & Chia, W. A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **98**, 15050–15055 (2001).
27. Zambrowicz, B.P. *et al.* Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* **392**, 608–611 (1998).
28. Nomura, M. *Ribosomal RNA Genes, RNA Polymerases, Nucleolar Structures, and Synthesis of rRNA in the Yeast Saccharomyces cerevisiae* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001).
29. Muratani, M. *et al.* Metabolic-energy-dependent movement of PML bodies within the mammalian cell nucleus. *Nat. Cell Biol.* **4**, 106–110 (2002).
30. Shedden, K. & Cooper, S. Analysis of cell-cycle-specific gene expression in human cells as determined by microarrays and double-thymidine block synchronization. *Proc. Natl. Acad. Sci. USA* **99**, 4379–4384 (2002).
31. Liron, Y., Paran, Y., Zatorsky, N.G., Geiger, B. & Kam, Z. Laser autofocusing system for high-resolution cell biological imaging. *J. Microsc.* **221**, 145–151 (2006).
32. Ghosh, R.N. *et al.* Quantitative cell-based high-content screening for vasopressin receptor agonists using transfluor technology. *J. Biomol. Screen.* **10**, 476–484 (2005).