High-throughput RNAi screening by time-lapse imaging of live human cells

Beate Neumann[1,4], Michael Held[1,4], Urban Liebel[1,4], Holger Erfle[1,4], Phill Rogers[1], Rainer Pepperkok[2] & Jan Ellenberg[2,3]

RNA interference (RNAi) is a powerful tool to study gene function in cultured cells. Transfected cell microarrays in principle allow high-throughput phenotypic analysis after gene knockdown by microscopy. But bottlenecks in imaging and data analysis have limited such high-content screens to endpoint assays in fixed cells and determination of global parameters such as viability. Here we have overcome these limitations and developed an automated platform for high-content RNAi screening by time-lapse fluorescence microscopy of live HeLa cells expressing histone-GFP to report on chromosome segregation and structure. We automated all steps, including printing transfection-ready small interfering RNA (siRNA) microarrays, fluorescence imaging and computational phenotyping of digital images, in a high-throughput workflow. We validated this method in a pilot screen assaying cell division and delivered a sensitive, time-resolved phenoprint for each of the 49 endogenous genes we suppressed. This modular platform is scalable and makes the power of time-lapse microscopy available for genome-wide RNAi screens.

After the completion of the human genome sequencing project[1], the task of functional genomics is to discover protein function genome-wide. Currently, RNAi is the method of choice to study loss-of-function phenotypes in human cells by specifically suppressing the expression of virtually any desired protein-coding gene. Indeed, several RNAi screens in human cells have already been reported[2–5]. These time-consuming and expensive primary screens have typically been based on endpoint assays of cells transfected in microtiter plates[2–5]. This allowed reasonable throughput but limited the information content of the phenotypic readout. This is problematic for two reasons. Endpoint assays capture only phenotypes for genes whose protein products are sufficiently reduced at the time the assay is performed, which depends on the stability of the targeted gene product. In a genome-wide screen this time varies between one and several days and endpoint assays are therefore typically scored late after RNAi transfection to detect as many phenotypes as possible. At late time points, however, most assays suffer from specificity problems, because for rapidly turned over proteins they can no longer differentiate between primary consequences of knockdown and secondary effects, and therefore lead to misinterpretation of phenotypes.

In principle, these limitations of RNAi screening can be overcome using the power of time-resolved live cell imaging, a strategy that has been used very successfully but with an enormous labor effort in the model system *Caenorhabditis elegans*[6]. Limitations in manual data acquisition of live movies as well as manual, and thus inherently biased and nonquantitative, data annotation have so far precluded the use of live-cell imaging for RNAi screening in vertebrate cell systems. Here we have overcome these bottlenecks by taking advantage of the miniaturized RNAi delivery offered by transfected cell microarrays, in which individually spotted siRNA transfection mixes are directly taken up from the solid phase by cells seeded on top of the array[7,8]. By spotting siRNA microarrays directly in live-cell imaging chambers, we were able to perform time-lapse microscopy of HeLa cells on the arrays. By massively increasing the throughput of fluorescence imaging and developing computerized analysis of the phenotypes by digital image processing, we established a fully automated high-throughput and high-content workflow of RNAi screening by time-lapse imaging (**Fig. 1**). We validated this method in a pilot experiment screening 49 endogenous human genes with a live-cell assay for chromosome segregation and structure in HeLa cells stably expressing the core histone 2B tagged with GFP. This platform is applicable to many assays that can be scored by *in vivo* (or fixed-cell) imaging, and is directly scalable to genome-wide RNAi screens.

RESULTS

siRNA validation

We chose to target 49 genes (**Supplementary Table 1** online) known or predicted to have a function in chromosome segregation or nuclear structure by one chemically synthesized siRNA each. For 43 of these siRNAs, mRNA knockdown efficiency could be measured in transfected HeLa cells by quantitative reverse-transcriptase PCR (qRT-PCR). We found that 74% (32) of the siRNAs suppressed the mRNA of the targeted gene by more than 70%, 88% (38) by at least 60% and all siRNAs showed partial knockdown (**Supplementary Fig. 1** online). In duplicate transfections of all 49 gene-specific siRNAs and one scrambled siRNA (randomized

[1]MitoCheck Project Group, [2]Cell Biology/Biophysics and [3]Gene Expression Programmes, European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [4]These authors contributed equally to this work. Correspondence should be addressed to J.E. (jan.ellenberg@embl.de).

RECEIVED 6 JANUARY; ACCEPTED 22 MARCH; PUBLISHED ONLINE 20 APRIL 2006; DOI:10.1038/NMETH876

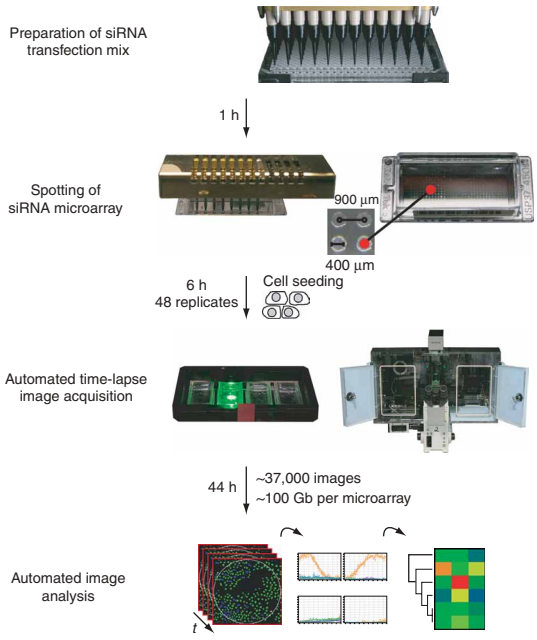

Figure 1 | Workflow of high-throughput RNAi screening by time-lapse imaging. Flowchart of the steps, including time lines of each process based on one live-cell microarray containing 384 siRNAs.

RNA sequence not targeting any human gene), we scored phenotypes manually by fluorescence microscopy of HeLa cells stably expressing *H2B-GFP* (see below). For 85% (42) of the targeted genes, the siRNA caused reproducible defects in nuclear structure or chromosome segregation, whereas cells transfected with the scrambled siRNA had no detectable phenotype (data not shown). We thus estimated the overall hit rate of this siRNA design as 74–85% and decided to use the 50 validated siRNAs (**Supplementary Table 1**) to develop the entire high-throughput workflow of live-cell RNAi screening by time-lapse imaging.

siRNA microarrays for live cell imaging

Application of live-cell imaging to genome-wide RNAi screens requires increasing the throughput of fluorescence microscopy by miniaturizing and parallelizing the siRNA transfection in a format compatible with imaging. We printed automatically prepared transfection cocktails for each of the 50 siRNAs directly in one-well live cell–imaging chambers with coverslip bottoms[8,9] (**Fig. 1**). A 400-µm spot size and 900-µm inter-spot distance allowed us to observe ~50 cells growing on each spot in a single field of view using a 10× microscope objective, and provided 384 spots per imaging chamber. The transfection efficiency of HeLa cells used in this study was 99% judged by uptake of fluorescently labeled scrambled siRNA (**Supplementary Protocol** online). In addition, A549, RPE and U2OS cell lines as well as primary skin fibroblasts could be transfected with more than 90% transfection efficiency (**Supplementary Protocol**).

siRNA uptake in HeLa cells on microarrays led to depletion of the targeted gene product as judged by immunofluorescence for two proteins, TPX2 and COPB. In both cases, specific protein staining was abolished in all (TPX2) or the majority (for the much more stable protein COPB) of HeLa cells growing on the siRNA spot (**Supplementary Protocol**). The siRNA microarrays could be stored for more than 7 months without loss of transfection efficiency as judged by ~70% of prometaphase-arrested HeLa cells after knockdown of *PLK1* (**Supplementary Protocol**). This has the advantage that the automated production of microarrays (**Supplementary Protocol**) can be uncoupled from the phenotypic analysis of the cells by imaging, allowing the use of replicates of identical microarrays over long time intervals to control the reproducibility of cell-based assays or the multiplexing of different assays on the same set of targeted genes.

Time-lapse imaging of chromosomes to assay cell division

The biological problem we investigated in the pilot screen was chromosome segregation during cell division. Fluorescent labeling of chromosomes allows dynamic processes such as mitosis to be phenotyped in great detail by fluorescence microscopy[10]. Mitosis is a rare event with typically less than 5% of a population of logarithmically growing HeLa cells undergoing division at one time. Furthermore, mitotic phenotypes such as chromosome segregation defects can lead to programmed cell death[11] and are often difficult to interpret in endpoint assays. We therefore decided to implement a time-lapse imaging assay for mitosis using a monoclonal HeLa cell line stably expressing the core histone 2B tagged with GFP[12], which highlights chromosomes and nuclei in all cell cycle stages. We imaged cells growing on scrambled siRNA every 30 min for 44 h starting 20 h after seeding, long enough to observe at least one division for each of the ~50 cells transfected on the spot. This provided the minimal temporal resolution to reliably score mitotic phases based on the changes in chromosome morphology (**Fig. 2a**). By starting the imaging 20 h after plating cells on siRNA microarrays, the total observation period allowed the scoring of early as well as late phenotypes (**Figs. 2b–e** and **3a**; and below).

High-throughput time-lapse microscopy of cell microarrays

After establishing the assay conditions for a single spot, the next step was to image as many microarray spots as possible within a time lapse of 30 min for large-scale screening applications. By further developing previously established hardware and software methods[13], we implemented a fully automated and 37 °C incubated wide-field epifluorescence screening microscope. After an initial autofocus of the entire microarray, we acquired single fluorescence images of each spot sequentially throughout the array with minimal illumination. Using real-time control of the fully motorized microscope hardware we could revisit up to 1,536 positions (four siRNA microarrays) every 30 min. To our knowledge, this is the first time that such a throughput in live-cell imaging has been described and makes live-cell imaging feasible for genome-wide RNAi screening. We used this system to produce two-day time-lapse movies of HeLa *H2B-GFP* cells on triplicates of a custom microarray that contained one siRNA spot for each of the 49 endogenous genes and three spots of a scrambled siRNA as negative controls (data not shown). Notably, cells transfected by the scrambled siRNA proliferated normally throughout the observation period (**Supplementary Protocol**), showing no toxicity by the repeated illumination and prolonged incubation on the screening microscope.

Automatic and quantitative phenotype annotation

After acquisition, the frames comprising each movie (89 per siRNA) had to be scored for mitotic phenotypes. We developed an automatic method also for phenotyping for two reasons. First,

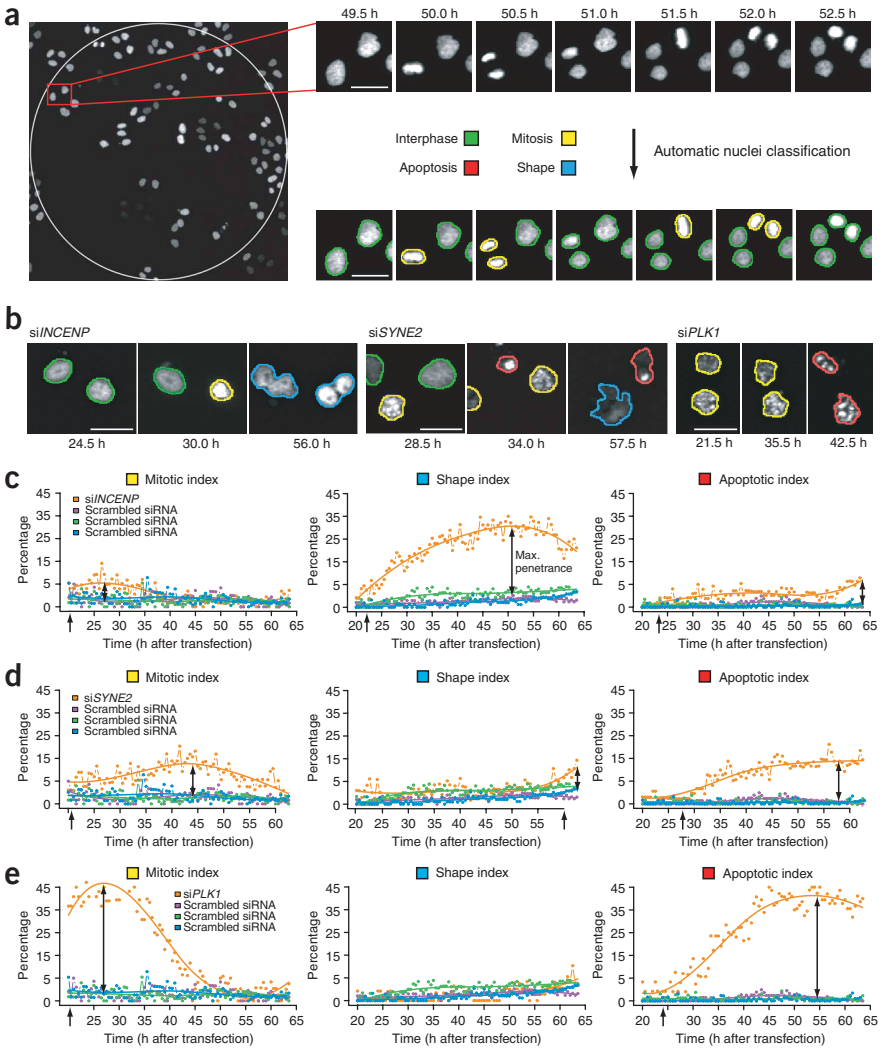

Figure 2 | Imaging of cell microarrays and automatic phenotype analysis. (**a**) HeLa H2B-EGFP cells on a scrambled siRNA spot (white circle represents spot diameter; left). Division of two cells (right, top row). Interphase and mitotic chromosomes recognized by the automatic classification are marked as indicated (right, bottom row; **Supplementary Video 1**). (**b**) Mitotic RNAi phenotypes followed by either binucleated arrested cells (*INCENP* siRNA), or by apoptosis-like *SYNE2* siRNA and PLK1 siRNA. Interphase, mitotic, apoptotic and binucleated ("shape") cells are automatically recognized and labeled in green, yellow, red and blue, respectively (**Supplementary Videos 2–4**). (**c–e**) Indices of "mitosis," "shape" and "apoptosis" classes for three scrambled siRNAs and one *INCENP* siRNA movie are plotted over time. Data was smoothed and fitted with a third-order polygon by local regression (solid line). Divergence between the average scrambled siRNA and the *INCENP* siRNA was computed for all 89 time points. The earliest statistically significant divergence is marked as the "first phenotype appearance" (arrows) and the maximal divergence as the "phenotype penetrance" (double-headed arrows) for each index. (**d**,**e**) The same analysis is shown for exemplary *SYNE2* (**d**) and *PLK1* (**e**) siRNA movies.

only automatic annotation by image processing allows the quantitative, unbiased and reproducible analysis of phenotypes that can later be easily parameterized and statistically evaluated to compare phenotypic signatures of genes. Second, because genome-wide RNAi screens would produce tens of thousands of movies, it was clear that annotation would become the rate-limiting step in the workflow for large-scale screening experiments.

The software we developed for automatic phenotyping proceeds in three major steps: (i) location of chromosome sets within single cells for each frame of the movie, (ii) classification of the chromosomes according to their morphology and (iii) detection of cell-cycle phenotypes based on the classification results for the entire movie.

For the first step, we defined the boundary of fluorescent chromosomes by segmentation using optimized local adaptive thresholding[14], which identified both interphase nuclei and mitotic chromosome sets with more than 99% accuracy compared to manual image analysis (data not shown).

In the second step we computed the texture and shape properties of each segmented set of chromosomes based on our past experience in identifying subcellular localization patterns[15] and used them to assign the cell to a biological class. The software distinguishes between "interphase," "mitosis," "apoptosis" and "shape" cells by machine-learning using multiclass support vector machines (SVM; see Methods for details). The classification was 97% accurate compared with manual annotation. We carried out segmentation and classification analysis for two normal cell divisions on a scrambled siRNA spot (**Fig. 2a**). The software correctly assigned "interphase" and "mitotic" chromosomes, the latter including all stages of M phase from late prophase to telophase. The correct identification of "apoptotic" cells was illustrated for the two distinct cell death morphologies resulting from suppression of the nuclear envelope proteins SYNE2 or PLK1 (**Fig. 2b**). By contrast, knockdown of the inner centromere protein INCENP resulted in bi- and multinucleated cells as a result of segregation defects that were correctly assigned to the "shape" class (**Fig. 2b**). Examples of the automatic classification for three additional gene knockdown experiments are shown (**Supplementary Fig. 2** online).

Finally we analyzed the classification results for an entire movie to decide whether the suppression of the targeted gene caused a chromosome segregation or nuclear structure phenotype. The software computes the number of cells in each biological class relative to the total number of cells within the spot area of a movie. We generated "mitotic," "shape" and "apoptotic" index plots for the *PLK1* knockdown experiment (**Fig. 2e**). We then compared the percentage of cells in a biological class to the three scrambled siRNAs on the same cell microarray. If any of the three classes showed a statistically significant difference to the controls, the earliest time at which the difference became significant is defined as the phenotype appearance and the maximum difference as the phenotype penetrance. For example, suppression of the essential

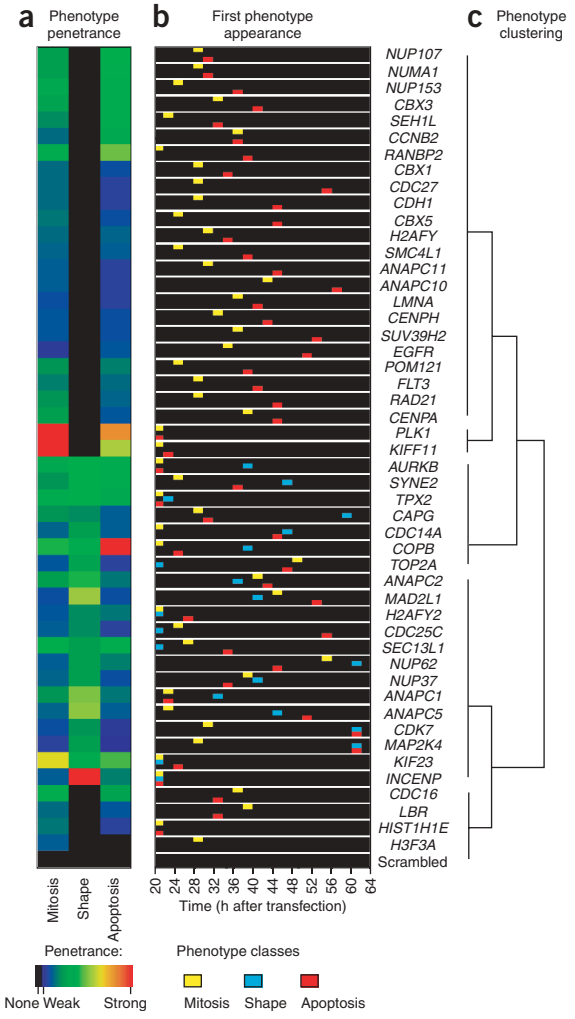

Figure 3 | Clustering of genes by time-resolved phenoprints. (**a**) Heat map of the phenotype penetrances for the three phenotypic classes "mitosis," "shape" and "apoptosis" automatically computed for all 49 targeted genes of the pilot screen and the scrambled siRNA. Shown are mean penetrances of three movies for each siRNA. The dynamic range of phenotype penetrance displayed in pseudocolor was scaled for every phenotypic class individually according to the maximal penetrance of each class, "mitosis" 30.3% (*KIF11*), "shape" 24.2% (*INCENP*) and "apoptosis" 37.5% (*COPB*). (**b**) First appearance of a phenotype in each of the three classes "mitosis" (yellow), "shape" (blue) or "apoptosis" (red) plotted over time after cell seeding is shown for all 49 targeted genes and the scrambled siRNA. (**c**) Genes were hierarchically clustered by penetrance and relative phenotype sequence shown in **b**. The dendrogram was reduced to four branches, yielding groups of genes related by their time-resolved phenotypic fingerprint.

mitotic kinase PLK1 led to prometaphase arrest[11] and consequently the mitotic index was elevated to 52% (compared to <5% in the controls) at the beginning of the movie, and the first appearance of the mitosis phenotype was thus calculated as 20 h after seeding (**Fig. 2e**). The maximum mitotic index was reached at 27 h yielding a phenotype penetrance of 48% (**Fig. 2e**). The apoptotic index plot of the same movie reveals the long term consequence of the prometaphase arrest. Corresponding to the decrease in mitotic index after 27 h, the apoptotic index increased to a maximum penetrance of 44% at 54 h (**Fig. 2e**). As expected, the shape index indicative of binucleated cells was not affected, as cells lacking *PLK1* do not initiate chromosome segregation but enter apoptosis directly from prometaphase[11] (**Fig. 2e**). We carried out an analogous analysis for *INCENP* and *SYNE2* knockdown (**Fig. 2c**,**d**). In summary, time-resolved automatic phenotyping provides information on when the suppressed gene product becomes limiting and allowed us to distinguish between primary (for example, prometaphase arrest) and secondary (for example, apoptosis) phenotypes.

Time-resolved phenoprints can cluster genes by phenotype

We next used the quantitative image processing software to detect phenotypes in the entire pilot screen dataset. We benchmarked the automatic phenotyping against manual annotation of both mitotic as well as apoptotic phenotypes in the same dataset, which we assumed to be 100% correct. For mitotic phenotypes, automatic and manual annotation agreed for 86% (42) of the targeted genes, with a false-positive rate of 14% (7) and no false negatives. Similarly for apoptosis, the automatic method had 88% (43) agreement with 10% (5) false positives and 2% (1) false negatives. Thus the automatic phenotyping was highly accurate and delivered unbiased and quantitative parameters about phenotype appearance and penetrance that allowed us to cluster this set of candidate genes according to phenotypic similarities. Genes were clustered by the kind of phenotype they showed (increased mitotic, shape or apoptotic index; **Fig. 3a**), by the mean maximal penetrance of each phenotype (ranging from 2–38%; **Fig. 3a**) and by the relative order in which multiple phenotypes, if present, appeared (**Fig. 3b**). This multiparameter clustering (**Fig. 3c**) revealed groups of genes related by their time-resolved phenotypic signature, or phenoprint. For example, a large group of genes that initially caused a mitotic phenotype followed later by apoptosis is represented by *PLK1* (**Fig. 2b**,**e**) and *CDC16* (**Supplementary Fig. 2**). Other groups of genes produced only apoptotic phenotypes (for example, *NUMA1*, **Supplementary Fig. 2**) or complex phenotypes in all classes, for example the gene *SYNE2*, encoding a nuclear envelope protein (**Fig. 2d**). It should be noted that the purpose of this study was to develop and apply a fully automated method for high-throughput RNAi screening by time-lapse imaging of live human cells and not to fully define the biological phenotype of each single gene in our pilot set. Because we only used one siRNA per gene, we cannot exclude that some phenotypes are weak hypomorphs or may result at least in part from off-target effects. But the high specificity of our mitosis assay and the very good correspondence of our data with previously published phenotypes (see below) make it unlikely that off-target effects had a major contribution to the observed phenotypes.

DISCUSSION

RNAi screening by time-lapse imaging of living cells has key advantages over traditional endpoint assays. Most importantly, it allows one to discriminate the primary phenotype caused by gene suppression from secondary cellular responses. In addition it can detect early as well as late phenotypes resulting from different stabilities of the targeted gene product. Time-lapse imaging is also very sensitive and can detect hypomorphic phenotypes such as mitotic delays rather than arrest, which can occur with partial depletion by RNAi. But phenotypes caused by partial depletion can be very different from those caused by complete removal of a protein[16], and therefore a hit in a primary screening method as the

one presented here should be followed up by secondary experiments in which the depletion level of individual proteins can be quantitated. The three phenotypes shown (**Fig. 2b**) would all be misinterpreted or missed in a typical endpoint determination 48 h after siRNA transfection. *SYNE2* would already show more apoptotic than mitotic cells; PLK1 would be wrongly scored as an apoptotic phenotype, which is the secondary consequence of the earlier prometaphase arrest. It is worth noting that the HeLa cells we used in this study are known to have reduced p53 levels, which has recently been reported to enhance the apoptosis response caused by depletion of mitotic genes such as *PLK1* (ref. 17). It is therefore likely that the high percentage of apoptotic phenotypes is at least in part due to this cell type. *INCENP* would be scored correctly as a binucleated cell phenotype, but the segregation defects causing this phenotype would not be detected because they already occur 20 h after seeding. The wide range of first phenotype appearances in our pilot screen 20–44 h after seeding (**Fig. 3b**) demonstrates the need to sample phenotypes at different times. In addition to classification and penetrance of the primary phenotype itself, the temporal sequence of phenotypes caused by suppression of a gene gives important clues about its function and can therefore be used to cluster genes into groups (**Fig. 3b**,**c**). We used only the relative timing of phenotypes for hierarchical clustering because the first appearance depends not only on stability of the targeted protein but also on the efficiency of the siRNA knockdown. Nevertheless we expect that in genome-wide screens the first appearance will provide valuable, albeit indirect, clues about average protein stability, as we find that genes encoding short-lived cyclically regulated proteins, such as *PLK1*, *AURKB* and *TPX2*, cause phenotypes early, whereas those encoding stable proteins, such as *CENPA* and *LMNA*, cause late phenotypes (**Fig. 3b**).

The automatic phenotyping software we developed matched manual annotation for 86–88% of the genes and correctly classified many phenotypes known from the literature (for example, mitotic phenotypes *PLK1* (ref. 11), *TPX2* (ref. 18), *KIF11* (ref. 5) and *KIF23* (ref. 19)). Machine learning, however, also has limitations, because it will only correctly recognize phenotypes that were used to train the classification software. This explains the up to 14% false-positive rate of the automatic annotation we observed in this pilot screen. An example of a false-positive mitotic recognition is the gene *NUMA1* (**Fig. 3a**,**b**).

In conclusion, we report for the first time an integrated high-throughput workflow that combines two powerful methods of functional genomics and cell biology: RNAi screening and live-cell time-lapse microscopy. This method delivers detailed time-resolved and quantitative phenoprints of human genes and is directly scalable to genome-wide screens. This method is not limited to HeLa cells, and can be applied to adherent cell lines derived from different tissues (**Supplementary Protocol**). It is worth noting, however, that a different cell line will require adaptation of the imaging conditions (slower cell cycle, for example, in primary skin fibroblasts, would require longer time-lapse sequences) and image processing (a different nuclear morphology, for example, in RPE cells, would have to be retrained for accurate classification). In general, this method can be adapted to any assay that can be detected by fluorescence microscopy and scored by computerized image processing, most prominently the host of *GFP* reporters available in the cell biology community for diverse cellular functions.

METHODS

High-throughput time-lapse imaging. We acquired images with an automated epifluorescence microscope (IX-81; Olympus Europe) as described[13] with the following new developments for time-lapse acquisition and increased throughput. We implemented stabilized light sources (MT20; Olympus Biosystems), firewire cameras (DB-H1, 1,300 × 1,024, Olympus Biosystems) and an in house–modified version of the "ScanR" software (now commercially available with all new developments from Olympus Biosystems), EMBL environmental microscope incubator (European Molecular Biology Laboratory, GP 106) as well as a new objective (Plan, 10×; numerical aperture (NA), 0.4; Olympus Europe) and filter sets for GFP (Chroma Inc.). We used an image-based autofocussing procedure[13] to focus on the maximum number of healthy cells (scoring size, intensity, contrast) in a field of view. The focus z coordinates of the transfected cell microarray were saved during the first round of imaging. The temperature stability of the incubated system allowed us to reuse the saved focus map for 44 h without further autofocussing between the 30-min imaging intervals. As cells needed on average ~1 h to go through mitosis, we obtained ~2 mitotic images per cell that enabled us to detect delays ≥30 min, or a mitotic index change from 5% to 7.5%. Higher time resolution would increase the sensitivity to subtle delays but reduce throughput and increase overall data volume. We reduced illumination of the specimen to the minimum necessary for sufficient signal-to-noise ratio for automated phenotyping (see below) with exposure times of 18–50 ms (for GFP). Stage movements thus limited throughput. Data was losslessly compressed on the fly and continuously saved on a 5 Tb network-attached storage. We connected microscopes via a 2 Gbit/s network to the network-attached storage to sustain uninterrupted data flow.

Automated phenotyping by image processing. We segmented chromosome sets by an optimized method of local adaptive thresholding, reducing computational costs by 90% compared to the original implementation[20]. We extracted numerical features of shape and texture from each segmented set of chromosomes resulting in a vector of 214 features each. Automatic classification into biologically defined classes was by machine-learning using multiclass SVMs[14,15,21,22]. Classes were "interphase," "mitotic," "apoptotic" (several cell death phenotypes with supercompacted or fragmented nuclei including, but not limited to, apoptosis), "shape" (nuclei of abnormal shape including, but not limited to, a high percentage of binucleated cells), and "artifact" (containing other objects). Feature selection was on ~200 manually annotated examples per class. We trained the SVM classifier on the 45 best-ranked features of all ~1,000 manually labeled cells. The classifier achieved an overall accuracy of 97% on all samples measured by fivefold cross-validation. In each single frame of time-lapse movies for each spotting position, the classifier then automatically assigned cells to the classes (**Fig. 2a**) resulting in kinetic plots of the fraction of a given class relative to the total number of cells for the entire movie (class index; **Fig. 2c–e**). Three classes "mitosis," "shape" and "apoptosis" were mined individually for the presence of a cell-cycle phenotype by statistical analysis. To increase the robustness and to cope with occasional false classifications, the data of the knockdown experiment of interest was first smoothed by local regression, which provides a fitted value (**Fig. 2e**) and a confidence interval (data not shown) for the fit at each time point. We then compared

the fitted value to the mean value of the three regression fitted control experiments on the same microarray, taking into consideration the standard deviation (s.d.) of the controls. If the divergence between the class index of the RNAi experiment and the mean of the negative controls on the same cell array was above a certain level of significance ("mitosis" 2× s.d., "apoptosis" 3× s.d., "shape" 2× s.d.), the earliest difference was calculated as the first appearance of the phenotype and the maximum difference as its maximal penetrance in the cell population (**Fig. 2c–e**). Besides statistical significance, we also assessed biological significance by requiring each phenotypic class index to be above a certain absolute threshold ("mitosis" and "apoptosis" at least 1%, "shape" at least 7%), which we determined empirically to remove false positives resulting from noise in the data (experimental and classification errors) that could not be removed by local regression smoothing. We excluded mitotic hits with an onset of more than 6 h after apoptosis based on *a priori* biological knowledge. We analyzed triplicates separately and averaged the results. We applied hierarchical clustering on relative phenotype sequence and maximum penetrance for these three phenotypic classes to the entire pilot screen dataset (**Fig. 3c**). We performed hierarchical cluster analysis on a dissimilarity structure, which was generated by the Euclidean distance between hextuples containing the maximum penetrance and the relative phenotype appearance for the three classes "mitosis," "shape" and "apoptosis." As there was no *a priori* reason to emphasize the categories penetrance and appearance or the classes "mitosis," "shape" and "apoptosis," all values of the hextuple were equally weighted for clustering. The automatic phenotyping software, which is coded in C++ and Python, is available upon request for academic researchers.

Additional methods. See **Supplementary Table 1** for siRNA sequences and **Supplementary Table 2** for reverse-transcriptase PCR primers used in this study; the **Supplementary Protocol** for description of the siRNA transfection methodology and its performance in different cell lines; **Supplementary Methods** for details of qRT-PCR; and **Supplementary Videos 1–7** for sample images of selected phenotypes.

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

ACKNOWLEDGMENTS

We thank S. Narumiya (Kyoto University, Kyoto) and T. Hirota (Institute of Molecular Pathology; IMP; Vienna) for HeLa 'Kyoto' cells; W. Huber (European Bioinformatics Institute; EBI; Hinxton) for advice on statistical analysis of kinetic data; O. Gruss (Zentrum für Molekulare Biologie Heidelberg; ZNBH; Heidelberg) for TPX2 antibody; J.-M. Peters (IMP; Vienna) for RPE cells; I. Hoffmann (Deutsches Krebsforschungszentrum; DKFZ; Heidelberg) for U2OS cells; H. Runz (Univ. Heidelberg) for primary human fibroblasts; Chroma Inc. for providing customized emission filter sets free of charge; EMBL's IT Services group (B. Kindler, M. Hemberger, R. Lück) for support; Olympus Biosystems, Hamilton and Bio-Rad for continuous support; Cenix BioScience GmbH for siRNA design and for providing the A549 cells; and Ambion Europe, Ltd. for providing siRNAs for validation. This project was funded by grants to J.E. within the MitoCheck consortium by the European Commission (FP6-503464) as well as in part by the Federal Ministry of Education and Research (BMBF) in the framework of the National Genome Research Network (NGFN) (NGFN-2 SMP-RNAi, FKZ01GR0403 to J.E. and NGFN-2 SMP-Cell FKZ01GR0423, NGFN-1 FKZ01GR0101, FKZ01KW0013 to R.P.).

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. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
2. Kittler, R. *et al.* An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* **432**, 1036–1040 (2004).
3. Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).
4. Pelkmans, L. *et al.* Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. *Nature* **436**, 78–86 (2005).
5. Zhu, C. *et al.* Functional analysis of human microtubule-based motor proteins, the kinesins and dyneins, in mitosis/cytokinesis using RNA interference. *Mol. Biol. Cell* **16**, 3187–3199 (2005).
6. Sonnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469 (2005).
7. Ziauddin, J. & Sabatini, D.M. Microarrays of cells expressing defined cDNAs. *Nature* **411**, 107–110 (2001).
8. Erfle, H. *et al.* siRNA cell arrays for high-content screening microscopy. *Biotechniques* **37**, 454–458, 460, 462 (2004).
9. Wheeler, D.B., Carpenter, A.E. & Sabatini, D.M. Cell microarrays and RNA interference chip away at gene function. *Nat. Genet.* **37** (Suppl.), S25–S30 (2005).
10. Gerlich, D. & Ellenberg, J. 4D imaging to assay complex dynamics in live specimens. *Nat. Cell Biol.* **4** (Suppl.), S14–S19 (2003).
11. Sumara, I. *et al.* Roles of polo-like kinase 1 in the assembly of functional mitotic spindles. *Curr. Biol.* **14**, 1712–1722 (2004).
12. Kanda, T. & Wahl, G.M. The dynamics of acentric chromosomes in cancer cells revealed by GFP-based chromosome labeling strategies. *J. Cell. Biochem.* (Suppl.) **35**, 107–114 (2000).
13. Liebel, U. *et al.* A microscope-based screening platform for large-scale functional protein analysis in intact cells. *FEBS Lett.* **554**, 394–398 (2003).
14. Huang, K. & Murphy, R.F. From quantitative microscopy to automated image understanding. *J. Biomed. Opt.* **9**, 893–912 (2004).
15. Conrad, C. *et al.* Automatic identification of subcellular phenotypes on human cell arrays. *Genome Res.* **14**, 1130–1136 (2004).
16. Meraldi, P. & Sorger, P.K. A dual role for Bub1 in the spindle checkpoint and chromosome congression. *EMBO J.* **24**, 1621–1633 (2005).
17. Liu, X., Lei, M. & Erikson, L. Normal cells, but not cancer cells, survive severe plk1 depletion. *Mol. Cell. Biol.* **26**, 2093–2108 (2006).
18. Gruss, O.J. *et al.* Chromosome-induced microtubule assembly mediated by Tpx2 is required for spindle formation in HeLA cells. *Nat. Cell Biol.* **4**, 871–879 (2002).
19. Zhu, C., Bossy-Wetzel, E. & Jiang, W. Recruitment of MKLP1 to the spindle midzone/midbody by INCENP is essential for midbody formation and completion of cytokinesis in human cells. *Biochem. J.* **389**, 373–381 (2005).
20. Hirota, T. *et al.* Distinct functions of condensin I and II in mitotic chromosome assembly. *J. Cell Sci.* **117**, 6435–6445 (2004).
21. Seul, M., Lawrence, O. & Sammon, M. *Practical Algorithms for Image Analysis* (Cambridge Univ. Press, Cambridge, UK, 2000).
22. Huang, K. & Murphy, R.F. Boosting accuracy of automated classification of fluorescence microscope images for location proteomics. *BMC Bioinformatics* **5**, 78 (2004).