CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging

Michael Held[1,2], Michael H A Schmitz[1,2], Bernd Fischer[3], Thomas Walter[4], Beate Neumann[5], Michael H Olma[1], Matthias Peter[1], Jan Ellenberg[4] & Daniel W Gerlich[1,2]

Fluorescence time-lapse imaging has become a powerful tool to investigate complex dynamic processes such as cell division or intracellular trafficking. Automated microscopes generate time-resolved imaging data at high throughput, yet tools for quantification of large-scale movie data are largely missing. Here we present CellCognition, a computational framework to annotate complex cellular dynamics. We developed a machine-learning method that combines state-of-the-art classification with hidden Markov modeling for annotation of the progression through morphologically distinct biological states. Incorporation of time information into the annotation scheme was essential to suppress classification noise at state transitions and confusion between different functional states with similar morphology. We demonstrate generic applicability in different assays and perturbation conditions, including a candidate-based RNA interference screen for regulators of mitotic exit in human cells. CellCognition is published as open source software, enabling live-cell imaging–based screening with assays that directly score cellular dynamics.

The availability of RNA interference (RNAi) technology for high-throughput gene inactivation experiments, fluorescent protein labeling and automated microscopy has opened a new era of screening possibilities in higher eukaryotes[1]. Indeed, imaging-based RNAi and chemical-compound screening has become an important discovery tool for the identification of new gene function, for example, in the regulation of DNA damage and repair[2], endocytosis[3] and mitosis[4–6].

Imaging-based screens typically assay altered incidence of cells with specific features in a population of fixed, fluorescently labeled cells. The development of computational methods, such as machine learning for supervised classification of cellular morphologies, were key for the automated annotation of high-throughput imaging data and the establishment of microscopy-based screening as a routine technology in a wide research community[7–12].

Many biological processes depend on stochastic events and occur in an unsynchronized and transient manner, which limits the applicability of single–time-point assays. Complex dynamic processes such as cell division or intracellular trafficking demand time-resolved, live-cell imaging[13]. Automated microscopes now enable high-throughput live-cell imaging with excellent spatiotemporal resolution[1,7,14]. Computational analysis of such data is challenging and existing machine learning and classification approaches do not provide sufficient accuracy to correctly annotate cellular trajectories with multiple time points. Published live-cell imaging–based RNAi screens have scored phenotypes either exclusively in cell populations[6,7] or relied on visual evaluation of single-cell dynamics[4]. However, cell population analysis cannot be used to detect stochastic and transient phenotypes, and visual interpretation of morphological dynamics is very time-consuming and often unreliable.

To improve the classification accuracy of machine learning methods, the temporal context can be taken into account. For example, if the biological process underlying an assay is well known, a biological model can be explicitly defined in an error correction scheme that suppresses illegitimate stage transitions. This has been applied to study the pattern of mitotic chromatin morphology changes[11,12]. However, temporal error correction based on biological a priori models limits the detection of unexpected phenotypic variations, and the adaptation to different biological questions requires re-implementation of the underlying models by the user for each new assay.

Here we present CellCognition, an integrated computational strategy that combines machine learning methods for supervised classification and hidden Markov modeling to measure morphological dynamics in live-cell microscopic movies. Our error-correction method does not require a priori definition of the temporal progression, which enables its application to a wide range of assays and phenotypic variations. We demonstrate efficiency and sensitivity of the methodology in various assays and perturbation conditions.
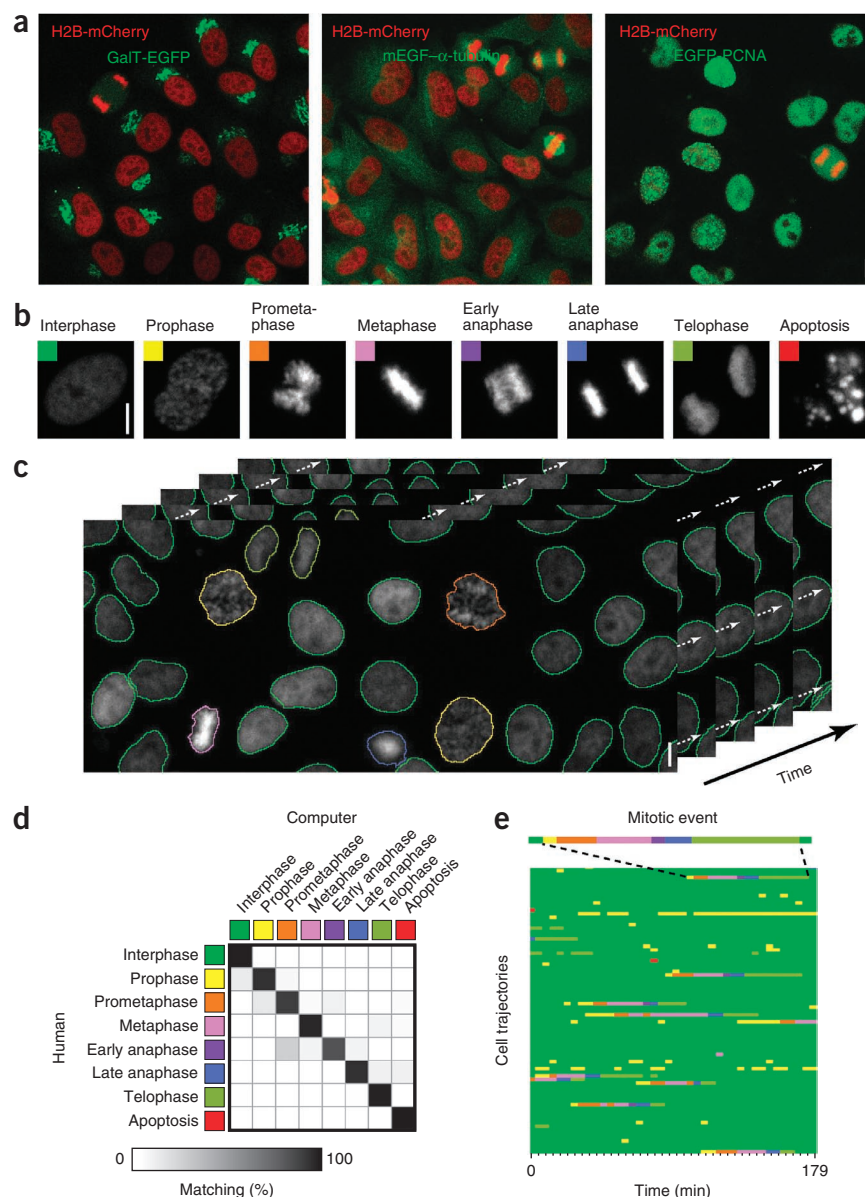
RESULTS
High-throughput imaging of cellular dynamics
To visualize morphological dynamics of various cellular structures, we generated a collection of human HeLa reporter cell

Figure 1 | Supervised machine learning and classification of morphologies. (**a**) Confocal images of live HeLa cells stably expressing a chromatin marker (H2B-mCherry), together with GalT-EGFP to visualize the Golgi apparatus, with mEGFP–α-tubulin or with the replication factory marker EGFP-PCNA. The images show maximum intensity projections of five z dimension sections. (**b**) Live imaging of HeLa cells expressing H2B-mCherry at different cell cycle stages or apoptosis (two-dimensional time series imaged with widefield epifluorescence 20× dry objective). Colors indicate H2B-mCherry morphology classifications used in subsequent figures. (**c**) Object detection (contours) and classification (colors) of cellular morphologies over time corresponding to mitotic stages defined in **b**. Scale bars, 10 μm (**a**–**c**). (**d**) Confusion matrix displays the matching of human annotations versus annotation of support vector machines with radial basis functions. (**e**) Automated annotation of cell trajectories over time as illustrated in **c**. Displayed are 80 randomly selected trajectories (rows) over 40 time frames (columns); time lapse, 4.6 min. Colors refer to morphology classes as defined in **b**. Tick marks indicate sampled time points. Mitotic events were rare, and the trajectories contain many single frames of mitotic annotations, likely owing to classification errors.

lines stably expressing different combinations of fluorescent markers. All cell lines expressed a red chromatin marker (core histone 2B fused to monomeric (m)Cherry; H2B-mCherry). In this background, we expressed markers for microtubules (monomeric enhanced GFP (mEGFP)–α-tubulin), the Golgi apparatus (galactosyl transferase (GalT)–EGFP), or DNA replication factories (proliferating cell nuclear antigen (PCNA)–mEGFP). These diverse secondary markers (**Fig. 1a**) are a well-suited test case for the implementation of a generic annotation method. With these cells, we performed multilocation time-lapse imaging on an automated widefield epifluorescence microscope[14]. We typically recorded 96 movies in parallel, with a temporal resolution less than 5 min over a total duration of 24 h, generating datasets of about 100,000 images, or 200 gigabytes, per day and microscope. The analysis of such a single experiment requires annotation of up to 25 million cellular morphologies derived from about 260,000 objects per movie with a 10× microscope objective.
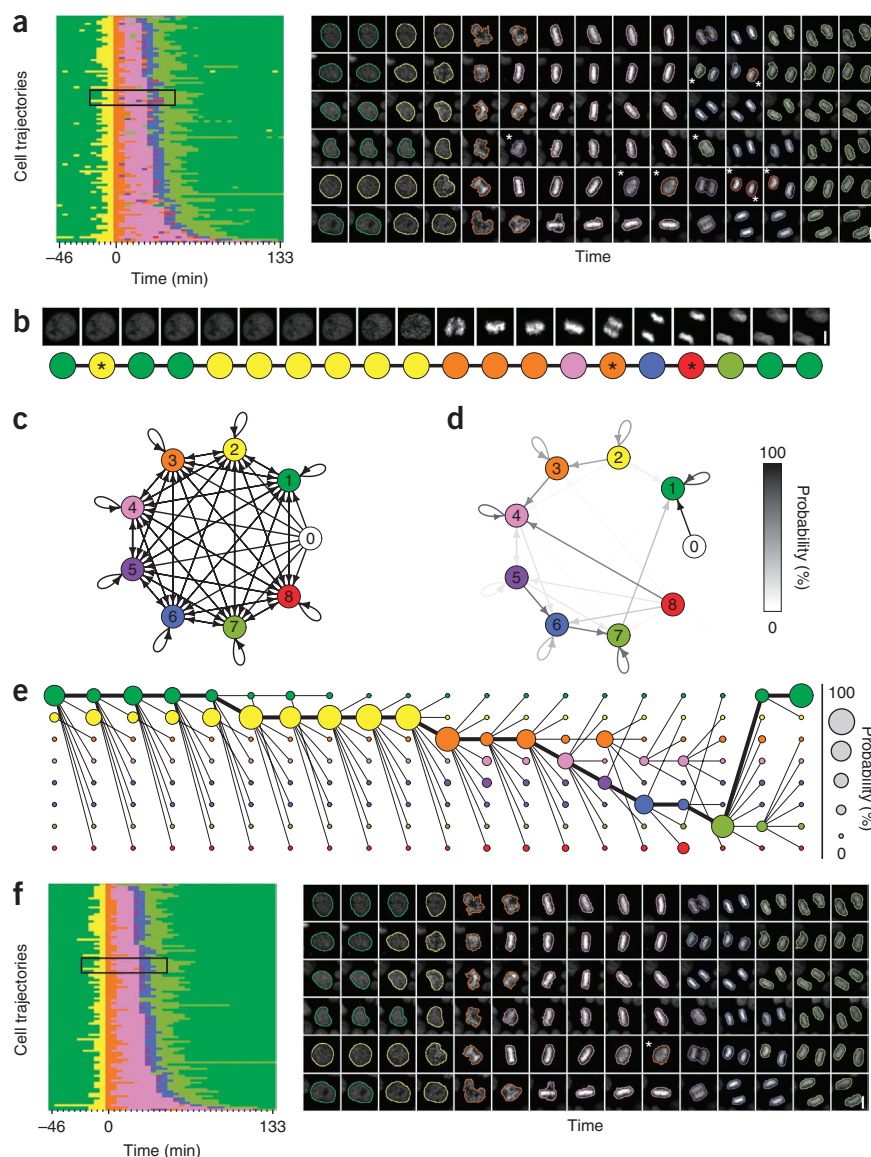
Machine learning and classification of morphologies

Timing measurements in live-cell imaging data are often based on the progression through distinct morphologies that relate to specific biological states. An excellent example for this is mitosis, for which the chromatin morphology can be used to annotate the canonical mitotic stages (**Fig. 1b** and **Supplementary Movie 1**). We used this classic assay as a test case to measure timing events at the single cell level.

We first implemented a canonical strategy for automated annotation of morphological classes[7–9,15], based on object detection, multivariate feature extraction and supervised machine learning (**Fig. 1c**). We used local adaptive thresholding[7], followed by a watershed split-and-merge segmentation error correction[16] to detect individual cells at an accuracy of 95.7% (1,876 objects; 2.6% oversegmented (falsely cut objects) and 1.7% undersegmented (falsely merged objects)). For each object, we then calculated 186 quantitative features[17,18] (**Supplementary Table 1** and **Supplementary Fig. 1**) describing texture and shape. Next, we trained a support vector machine classifier[19] for the discrimination of eight different object morphologies (**Fig. 1b**; interphase, six different mitotic stages and apoptosis). We defined these classes by manual annotation of 28–195 example objects. The agreement between human and computer annotation was 94.6% (mean of all classes; fivefold cross-validation), ranging between 75.0% for the early anaphase class and 99.0% for interphase class (**Fig. 1d**). This performance was similar to that of several previously reported supervised

Figure 2 | Hidden Markov modeling of progression through morphology stages. (**a**) Automated extraction of mitotic events. Cells were synchronized *in silico* to the prophase to prometaphase transition. The plot displays a random selection of 100 mitotic events (from a total of 172 mitotic events from eight movies; time lapse, 4.6 min). For plotted data in the boxed region, contour overlays on image data are shown on the right. Predicted morphology classes were color-coded as in **Figure 1b**. Asterisks indicate classification errors. (**b**) Images of a single cell and corresponding trajectory of class labels. Asterisks denote classification errors. (**c**) Graph for all possible transitions between classes. Node 0 is the start node; all other nodes are color-coded as defined in **Figure 1b**. (**d**) Learned class transition probabilities based on the trajectories shown in **a**. Normalization of probabilities was per node. (**e**) Trellis diagram showing all class prediction estimates for the cell shown in **b**. Vertical columns correspond to single time points, aligned to the images in **b**. Rows correspond to morphology classes, labeled as defined in **Figure 1b**. Probability estimates derived from the support vector machine are coded by size. The Viterbi algorithm was used to decode the overall most likely sequence (thick black line). Thin black lines indicate the most likely preceding state of a label at each given time point. (**f**) Error correction as in **e** was performed for all trajectories shown in **a**. Scale bars, 10 μm.

machine learning applications[7,9,11,20]. Next, a nearest-neighbor algorithm that supports trajectory splitting (for example, cell division) and merging (for example, cell-to-cell fusion) tracked individual cells over time. The automated tracking matched 99.8% of the human annotated object-over-frame connections (1,942 connections), a value comparable to the performance of previous studies on cell tracking[11,21].

The overall accuracy of the individual computational steps appeared to be high. However, considering >500 frames per cell trajectory for our time-resolved datasets, we obtained almost no error-free trajectories by this approach (**Fig. 1e** and **Supplementary Movie 2**).

Detecting scarce events in long-term movies

Mitotic events are scarce in comparison to the much longer duration of interphase (**Fig. 1e**). To improve the sensitivity for mitotic-stage annotation, we automatically selected mitotic events based on a morphology class sequence motif of prophase-prometaphase. This yielded a subgraph highly enriched for mitotic events (**Fig. 2a** and **Supplementary Movie 3**; 81.5% of all mitotic events were automatically extracted; 294 mitotic events in three movies). This set of trajectories contained 2.1% misclassifications per object (a posteriori compared with human annotation).

Untrained biological users may annotate the classifier training set less reliably. To test the sensitivity of the support vector

machine toward annotation errors, we randomized the labels on fractions of training objects and measured the overall classification accuracy (**Supplementary Fig. 2**). Randomization of the labels on 50% of the training objects reduced the overall annotation accuracy only slightly below 90%. This demonstrates that classification by support vector machine is relatively insensitive to annotation errors.

Hidden Markov model for time-lapse imaging

Single object–based machine learning and classification does not take the temporal context into account, but objects with ambiguous morphologies invariably occur in a typical context of preceding and following morphologies. This context could help derive the correct annotation for the ambiguous object. This could be particularly relevant for gradual morphology changes at stage transitions, where single object–based classification is relatively inaccurate (for example, interphase, prophase, interphase, interphase, prophase, prometaphase; **Fig. 2b** and **Supplementary Movie 3**).

We reasoned that taking the history of a cell into account might provide a means to correct for such noise at stage transitions, as

well as confusion between closely related morphology classes. We assumed that the true state of a cell at a given time point (the mitotic stage in this assay) is not known, but that it correlates with an observed state (the morphology class prediction probabilities). We also assumed that the progression to the next state entirely depends on a given present state. This fulfils the criteria for a hidden Markov model, which can be used for error correction in time-resolved data[22].

We built a model with five components: (i) hidden states, representing the true morphology classes (for example, mitotic stages), (ii) observed states (the class prediction probability vectors of the support vector machine), (iii) probabilities of hidden state transitions, (iv) observation probabilities and (iv) initial probabilities of hidden states. All elements of this model were computationally derived from the data without additional user interaction. We defined the hidden states by the initial class annotation, as described above (**Fig. 1b**). The support vector machine yielded observed states as a vector of class prediction probabilities for each time point. The hidden state probabilities were initialized at the first time point by the support vector machine predictions. Transition probabilities between hidden states were calculated based on the support vector machine prediction probabilities of all cellular trajectories per experimental condition (**Fig. 2c,d**), and the observation probabilities between hidden and observed states were estimated based on the confusion matrix of the support vector machine. We derived the overall maximum likelihood path for the progression through mitosis by the Viterbi algorithm[23] (**Fig. 2e**). This increased the overall per-object accuracy to 99.0%. Iterative learning of transition probabilities by the expectation-maximization algorithm[24,25] did not improve prediction accuracy (98.1% after five iterations). We suspected that the confusion matrix overestimated observation probabilities, as classes that are difficult to discriminate (prophase and early anaphase) were overrepresented in the annotation data. We therefore tested the performance of temporal error correction with lower error rates in the observation probabilities (0.1% for all transitions) and found that this eliminated noise at state transitions and more efficiently corrected single frames of misclassified objects, yielding overall accuracy of 99.4% per object and 91% completely error-free trajectories (100 trajectories; 4,000 objects; **Fig. 2f**, **Supplementary Fig. 3** and **Supplementary Movie 4**).

We next tested whether incorporation of a priori biological knowledge on state transitions increased the annotation accuracy. Specifically, we constrained the state transition graph of three consecutive classes to the forward direction and defined apoptosis as a terminal state (**Supplementary Fig. 4a,b**). The probability matrix for constrained state transitions improved the error correction performance of the hidden Markov model to 99.7% per object, yielding 94% completely error-free trajectory annotations (100 trajectories; 4,000 objects; **Supplementary Fig. 4c**).

We expected temporal error correction by the hidden Markov model to depend on good estimates of the predicted morphology classes. We therefore investigated the robustness of temporal error correction toward simulated classification noise. We randomized the class prediction probability vectors of fractions of objects, then trained the hidden Markov model on the noisy trajectories and applied it to correct classification errors (**Supplementary Fig. 5**). Comparison with manually annotated data demonstrated that the hidden Markov model–based error correction improved the overall accuracy at all noise levels.

We also tested whether the temporal error correction was sensitive to changes in the time-lapse interval by generating trajectories sampled to every second time point up to every sixth time point (**Supplementary Fig. 6**). Comparison with the manually annotated labels showed that the hidden Markov model increased the overall annotation accuracy at all sampling intervals.

Hidden Markov modeling provided a robust and efficient means to eliminate misclassifications and noise at morphology state transitions. The combination of mitotic event selection and hidden Markov model error correction reduced the per-object error rate about tenfold compared to single time point–based classification.

Generic strategy for annotation of cellular dynamics

We next used our tools for simultaneous analysis of multiple markers in the same cell, for example, to address temporal coordination of mitotic processes. We defined cytoplasmic areas based on their relative position to the chromatin marker, using non-overlapping 'region growing' of the contours derived from the chromatin channel (**Supplementary Fig. 7a,b**). Although this may be less precise than segmenting in the secondary channel, it proved to be robust over many different assays and was insensitive to temporal dynamics (**Figs. 3** and **4**). We applied tracking results of the primary channel to the secondary channel and performed all subsequent analysis of temporal dynamics independently for primary and secondary channels, as outlined above (**Supplementary Fig. 7c**).

We first applied our methods to analyze movies from cells expressing mEGFP–α-tubulin to annotate mitotic spindle assembly and disassembly (**Fig. 3a** and **Supplementary Movie 5**) and movies from cells expressing GalT-EGFP to study mitotic breakdown and reassembly of the Golgi apparatus (**Fig. 3b** and **Supplementary Movie 6**). We trained classifiers for six (α-tubulin) or five (GalT) distinct morphology classes. The mean accuracy of object class predictions was 96.5% for mEGFP–α-tubulin and 97.3% for GalT-EGFP (fivefold cross-validation, computational versus visual scoring). This yielded 55% (α-tubulin) or 38% (GalT) completely error-free trajectories. By using hidden Markov model error correction, the accuracy increased to 89% completely error-free trajectories for α-tubulin (**Fig. 3d** and **Supplementary Movie 7**) and 90% for GalT (**Fig. 3e** and **Supplementary Movie 8**; $n = 100$ for both assays; **Fig. 3g,h**).

To apply our methods to nonmitotic cellular dynamics, we annotated the timing of S-phase progression. We imaged a HeLa cell line stably expressing H2B-mCherry and EGFP-PCNA, a marker for DNA replication foci, which visualizes a characteristic pattern of morphology changes during S-phase progression (**Fig. 3c** and **Supplementary Movie 9**). We trained classifiers for six distinct PCNA morphology classes and established a hidden Markov model for error correction. This yielded 98.2% correctly annotated objects and 90% completely error-free trajectories (100 trajectories containing 15,000 objects; **Fig. 3f,i** and **Supplementary Movie 10**). The high performance in this diverse set of assays demonstrated generic applicability of our computational methods.
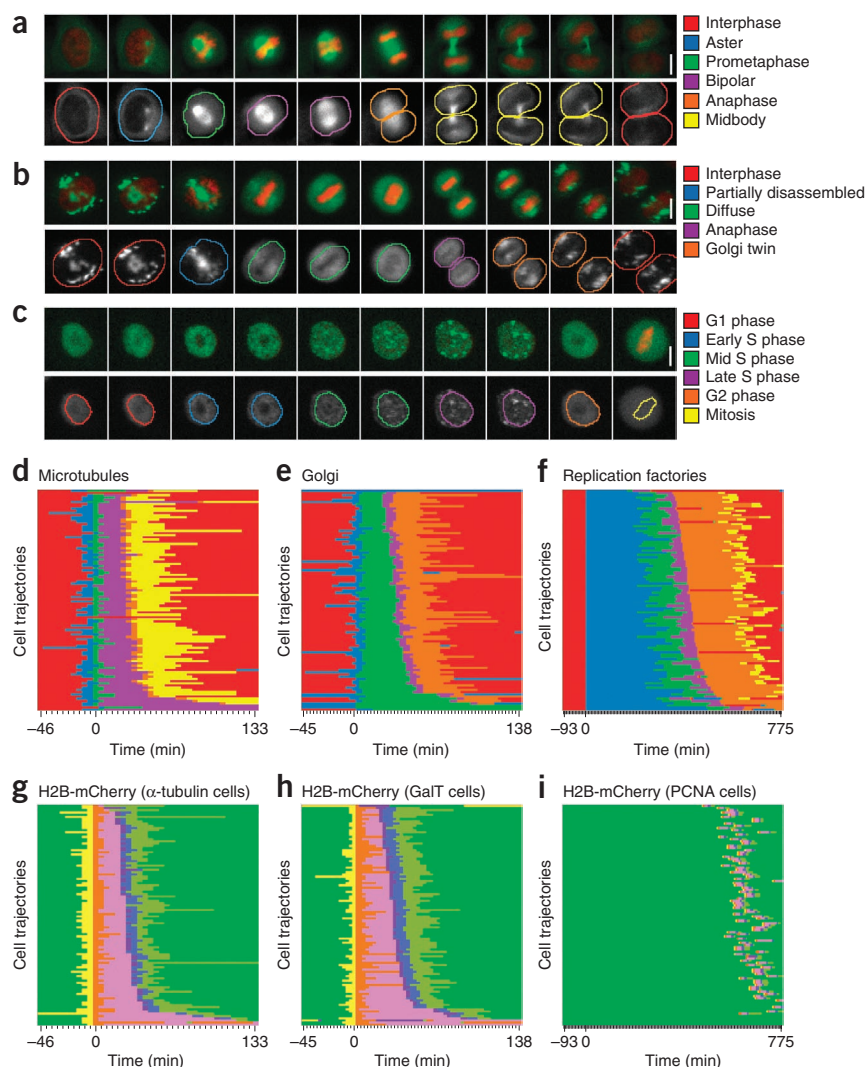
Figure 3 | Automated annotation of mitotic spindle and Golgi dynamics, and replication factory patterns during S-phase progression. (**a**) Live imaging of mitotic spindle dynamics of a cell expressing H2B-mCherry (red) and mEGFP–α-tubulin (green) (20× objective; 4.6 min time lapse). Automated hidden Markov model–corrected classification of spindle morphology, color labeled as indicated. (**b**) Live imaging of mitotic Golgi dynamics in a cell line expressing H2B-mCherry (red) and GalT-EGFP (green) (10× objective; 2.8 min time lapse). Colors indicate automated hidden Markov model–corrected annotation of Golgi morphologies. (**c**) Live imaging of DNA replication factory dynamics in a cell line expressing H2B-mCherry (red) and PCNA-EGFP (green) (10× objective; 5.9 min time-lapse). Colors indicate automated hidden Markov model–corrected annotation of S-phase progression based on PCNA morphology. Scale bars, 10 μm (**a**–**c**). (**d**) Automated annotation of a high-throughput imaging dataset. One hundred randomly selected mitotic events were derived and *in silico* synchronized to the prophase-prometaphase transition based on the H2B-mCherry annotation (**Fig. 2**). The secondary channel annotation was calculated independently from the H2B-mCherry channel, as indicated in **a**. (**e**) Automated annotation of Golgi dynamics, processed as in **d**. (**f**) Automated annotation of S-phase progression. Cells were *in silico* synchronized to the G1–early S phase transition based on the EGFP-PCNA classification. (**g**–**i**) Hidden Markov model-corrected annotations of H2B-mCherry morphologies for the cells shown in **d**–**f**. Colors label classes as defined in **Figure 1b**.

Quantitative phenotyping and kinetic measurements

We designed our methods for the detection of timing phenotypes. We therefore established perturbation conditions that are known to delay or shorten particular stages of mitosis. First, we used the microtubule-depolymerizing drug nocodazole, which arrests cells in prometaphase by permanent activation of the spindle check-point (**Fig. 4a** and **Supplementary Movie 11**). Our computational

tools reliably detected this (96.2% completely error-free annotated trajectories, 154 trajectories; **Fig. 4b**).

Next, we depleted the essential spindle checkpoint component Mad2 by RNAi, which is known to accelerate the timing from mitotic entry until anaphase onset in HeLa cells by about twofold[26] (**Fig. 4a** and **Supplementary Movie 12**). We evaluated the accuracy of automated timing measurements, scoring the time from prometaphase until anaphase onset based on the chromatin marker (cells that did not segregate chromosomes were omitted). Automated measurements of 47.2 ± 20.0 min (mean ± s.d.; $n = 195$) in control cells did not significantly differ from manual annotation of the same dataset (48.5 ± 18.0 min; two-sided Mann-Whitney-Wilcoxon test, $P = 0.12$). Automated timing measurements in Mad2 RNAi cells demonstrated mitotic acceleration (13.0 ± 3.6 min), agreeing with manual annotation (12.4 ± 3.4 min; two-sided Mann-Whitney-Wilcoxon test, $P = 0.23$). As expected from the known biological function of Mad2, the mitotic acceleration in Mad2 RNAi cells was mainly due to a shortened metaphase stage (1.6 ± 1.1 min in Mad2 RNAi cells; 36.5 ± 16.6 min in control; **Fig. 4b**).

Simultaneous measurements of morphological dynamics and the state of regulatory factors are a powerful approach for mechanistic dissection of perturbation phenotypes. Here we combined the annotation of mitotic stages with kinetic measurements of Securin degradation, which is required for anaphase initiation[27] (**Fig. 4a** and **Supplementary Movies 11–13**). In the normalized degradation kinetic profiles (**Fig. 4c**), we found that in control cells the Securin-mEGFP degradation initiated briefly before anaphase (**Fig. 4b,c**), consistent with spindle checkpoint inactivation at this stage. In nocodazole-arrested cells, almost all Securin-mEGFP remained stable during the measurement period of 138 min, consistent with an efficient and permanent activation of the spindle checkpoint. Degradation of Securin-mEGFP in

Mad2 RNAi cells initiated directly after mitotic entry, at a stage where chromosomes were still in prometaphase configuration. This indicates that the anaphase-promoting complex was activated before complete chromosome congression, as expected for a compromised spindle checkpoint function. These experiments demonstrate accurate timing phenotype annotation in RNAi-and drug-perturbed cells.

Figure 4 | Timing phenotypes and kinetic measurements. (**a**) Mitotic progression assayed by H2B-mCherry morphology and degradation of Securin-mEGFP. Examples are shown for an untreated control cell, a cell with Mad2 RNAi–inactivated spindle checkpoint, and a cell arrested in prometaphase by nocodazole (Noc.). Time lapse, 2.7 min; every third frame is shown. Scale bar, 10 μm. (**b**) Automated classification of mitotic stage progression as in **Figure 2f** for the three experimental conditions shown in **a**. (**c**) Securin-mEGFP degradation kinetics for the same cells shown in **b**. Normalization was per trajectory to the first prometaphase frame.

RNAi screen for mitotic exit regulators

To test the sensitivity and performance of our computational methods in a high-throughput application, we performed a screen for regulators of mitotic exit. Specifically, we aimed to identify regulators of post-anaphase stages of mitosis, for which RNAi phenotypes have not been reported so far. Mitotic exit control is well understood in budding yeast, yet it is unclear whether homologs of the yeast factors also control mitotic exit in higher eukaryotes[28]. We therefore designed a library of 283 small interfering RNAs (siRNAs) targeting 93 candidate mitotic exit regulator genes, including all known human genes with homology to budding yeast mitotic exit regulators and some additional genes known to be involved in mitotic regulation (**Supplementary Table 2**). As

an assay for mitotic exit timing, we scored the timing from anaphase onset, based on the chromatin marker H2B-mRFP, until postmitotic nuclear envelope reassembly, based on the nuclear import substrate IBB-EGFP (**Fig. 5a** and **Supplementary Movie 14**).
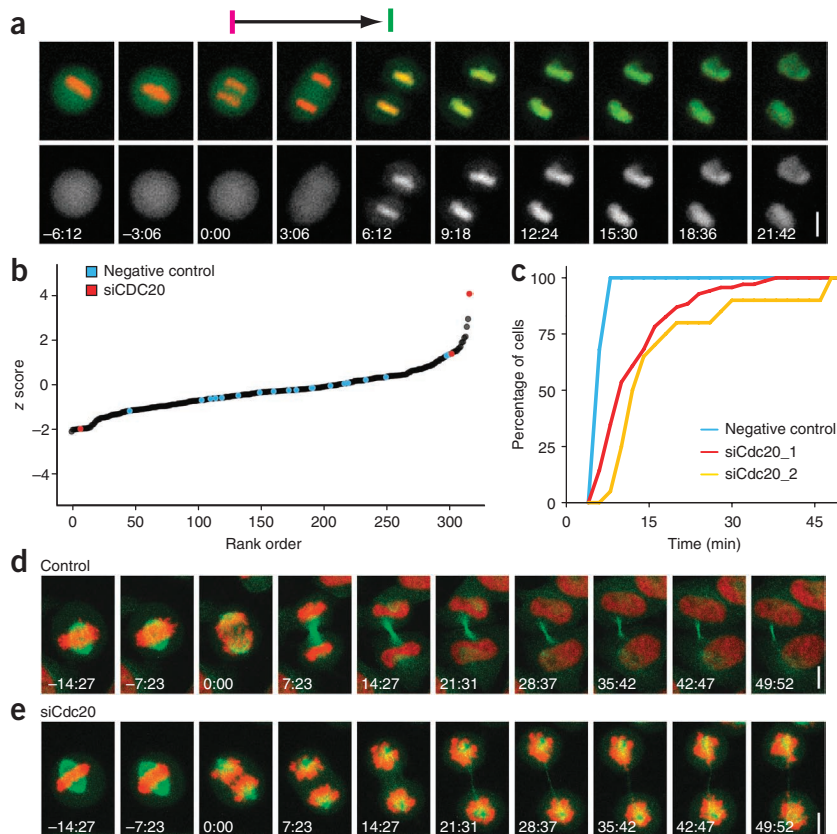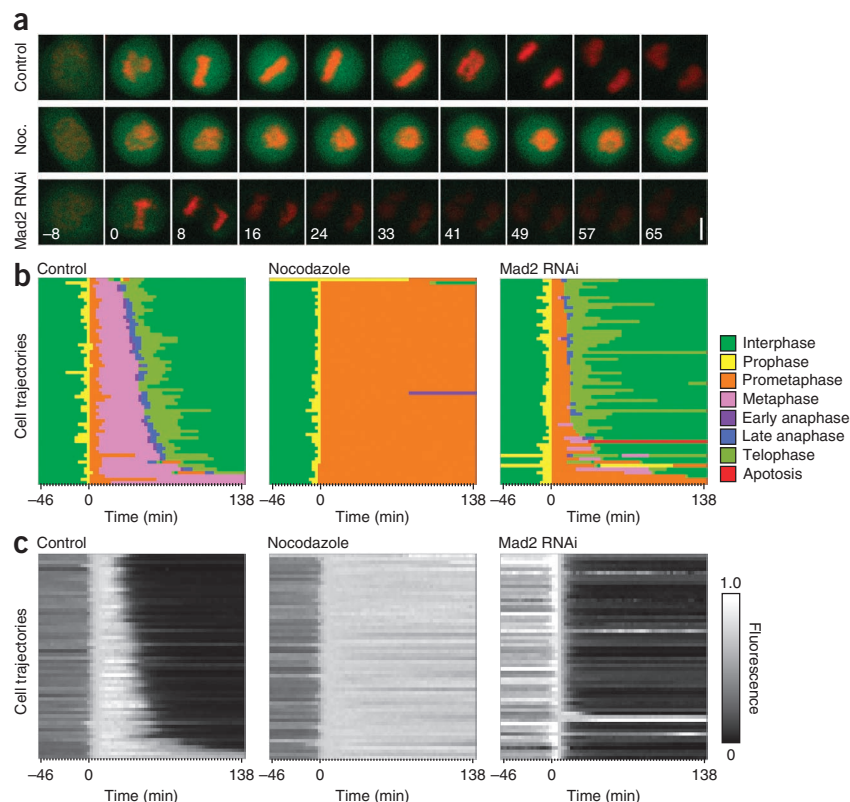

Figure 5 | RNAi screen for mitotic exit regulators. (**a**) Live-cell imaging of a cell line expressing H2B-mCherry (red) and IBB-EGFP (green; also shown separately in the lower images) to assay mitotic exit timing. The timing from anaphase onset (magenta line) until onset of nuclear accumulation of IBB-EGFP (green line) was used to define mitotic exit timing (arrow). Time is in min:s. (**b**) Mitotic exit timing in an RNAi screen for 300 different RNAi conditions. We recorded 108 movies of different siRNA transfections in parallel over 46 h, to collect the entire dataset in four experiments. Time lapse, 3.7 min. Each point in the graph indicates the z score for one siRNA. Each gene was targeted by three different siRNA oligonucleotides. (**c**) Cumulative percentage of cells exiting mitosis after onset of chromosome segregation ($t = 0$ min). The curves represent all mitotic events from two experimental replica. Cells were transfected in liquid phase with two different siRNA targeting Cdc20 (siCdc20_1 and siCdc20_2), or a non-targeting oligo for control, as indicated in the legend. (**d**) Confocal time-lapse imaging of a cell stably expressing H2B-mCherry and mEGFP–α-tubulin. Time is in min:s; maximum intensity projection of five z dimension slices. (**e**) Confocal imaging as in **d** for a Cdc20 RNAi cell. Scale bars, 10 μm.

For solid-state transfection of siRNAs into HeLa cells, we used a high-density transfection array with 300 spots of different siRNA transfection solutions printed to the glass surface of a chambered coverslip[7]. We seeded the cells onto this array and 20 h later started parallel imaging of 108 movies per experiment, for a total duration of 46 h and with 3.7 min time resolution. We automatically annotated the mean mitotic exit timing per experimental condition within the 1.6 terabyte data containing 646,754 images and 16,314 mitotic events. Only one siRNA delayed mitotic exit above a z-score threshold of 3.0 (**Fig. 5b** and **Supplementary Fig. 8a**; 6.8 ± 2.0 min mean $\pm$ s.d.; $n = 50$ mitotic events). This oligo depleted the anaphase-promoting complex co-activator Cdc20, as validated by western blotting (**Supplementary Fig. 8b**). We confirmed the specificity of the phenotype in two additional replicas with standard liquid-phase transfection and with an additional siRNA (**Fig. 5c**).

To test whether Cdc20 was required for other cellular reorganization processes during mitotic exit, we assayed chromosome decondensation and mitotic spindle disassembly. High resolution confocal time-lapse imaging of cells expressing both H2B-mCherry and mEGFP–α-tubulin (**Fig. 5d,e** and **Supplementary Movies 15** and **16**) showed that 100% of control cells (30 cells tested) started chromosome decondensation within 14 min after chromosome segregation, whereas only 54% (36 cells tested) did so after Cdc20 depletion. Thirty-one percent (36 cells tested) of Cdc20-depleted cells started kinetochore fiber spindle disassembly 7 min after anaphase onset, in contrast to 87% (30 cells tested) in control cells. These data suggest a requirement of Cdc20 for various cellular processes leading to postmitotic reassembly of interphase cells. This is unexpected given that Cdc20 has so far been thought to act mainly at pre-anaphase stages of mitosis and it has not been noticed in previous phenotypic analysis of Cdc20 RNAi cells[29].

DISCUSSION

Building on existing machine learning methodologies, the design of a generic workflow for annotation of morphological dynamics faced two main challenges. First, the classification noise at continuous morphology stage transitions impairs coherent trajectory annotation. Second, some biologically distinct classes appear morphologically similar, which leads to high classification confusion. By hidden Markov modeling, our methods efficiently correct both types of errors based on the temporal context. The hidden Markov models are learned individually for each experimental condition, without any human supervision. This allows the software to automatically adapt the error correction scheme to phenotypic deviations.

Biological a priori knowledge to suppress state transitions that are assumed to be impossible can also be used to improve annotation accuracy[11,12] but such explicit error correction schemes cannot be applied to new markers or assay systems without adaptation, and they may not apply to phenotypes with potentially altered stage progression. We found that the gain in accuracy by biological a priori constraints on the temporal progression was only minor. Our hidden Markov implementation modeled time series analysis in a high dimensional feature space with an intrinsic class-discriminant dimensionality reduction. Unlike principle component analysis[12,30], our method preserves context-specific structures. This may explain the large gain in accuracy compared to the previous implementations (**Supplementary Tables 3** and **4**).

Compared to previously described models[11,12], our model is the only one that can handle arbitrary relationships between phenotypic cell classes, providing a powerful and generic solution for time-resolved cellular phenotyping.

We demonstrated that our analysis methods can be used for a broad range of biological assays. We are not aware of any constraints that would preclude the use of our methods in other biological contexts, for example, apoptosis or cellular differentiation. However, the texture and shape features implemented in our software do not enable assays relying on absolute object counts, for example, in centrosome duplication assays. Also, assays scoring rapid intracellular dynamics would require integration of motion feature extraction methods.

Supervised machine learning, as in this study, requires user-defined morphology classes. It is therefore not possible to detect aberrant phenotypic morphologies that do not occur in the control conditions used for annotating the classifier training set. This limitation may be overcome in future studies by implementing unsupervised machine learning methods for the analysis of image time series.

We integrated our methods into the platform-independent software package CellCognition, with graphical user interface and supporting high-throughput batch processing on computer clusters. CellCognition is published as open source software (current version 1.0.7; **Supplementary Software**), along with high-quality reference image data at http://www.cellcognition.org/. With the increased availability of live-cell screening microscopes, we anticipate that time-resolved imaging assays will soon dominate a considerable fraction of high content screening and systems biology applications.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

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

and of the image analysis workflow. B.N. and J.E. generated the siRNA cell transfection array. M.H.O. and M.P. established live imaging of EGFP-PCNA. D.W.G. designed assays and the general strategy for image processing and wrote the paper.

1. Conrad, C. & Gerlich, D.W. Automated microscopy for high-content RNAi screening. *J. Cell Biol.* **188**, 453–461 (2010).
2. Doil, C. *et al.* RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell* **136**, 435–446 (2009).
3. Collinet, C. *et al.* Systems survey of endocytosis by multiparametric image analysis. *Nature* **464**, 243–249 (2010).
4. Sonnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469 (2005).
5. Goshima, G. *et al.* Genes required for mitotic spindle assembly in *Drosophila* S2 cells. *Science* **316**, 417–421 (2007).
6. Neumann, B. *et al.* Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* **464**, 721–727 (2010).
7. Neumann, B. *et al.* High-throughput RNAi screening by time-lapse imaging of live human cells. *Nat. Methods* **3**, 385–390 (2006).
8. Loo, L.H., Wu, L.F. & Altschuler, S.J. Image-based multivariate profiling of drug responses from single cells. *Nat. Methods* **4**, 445–453 (2007).
9. Conrad, C. *et al.* Automatic identification of subcellular phenotypes on human cell arrays. *Genome Res.* **14**, 1130–1136 (2004).
10. Glory, E. & Murphy, R.F. Automated subcellular location determination and high-throughput microscopy. *Dev. Cell* **12**, 7–16 (2007).
11. Harder, N. *et al.* Automatic analysis of dividing cells in live cell movies to detect mitotic delays and correlate phenotypes in time. *Genome Res.* **19**, 2113–2124 (2009).
12. Zhou, X., Li, F., Yan, J. & Wong, S.T. A novel cell segmentation method and cell phase identification using Markov model. *IEEE Trans. Inf. Technol. Biomed.* **13**, 152–157 (2009).
13. Gerlich, D. & Ellenberg, J. 4D imaging to assay complex dynamics in live specimens. *Nat. Cell Biol.* **5**, S14–S19 (2003).
14. Schmitz, M.H. & Gerlich, D.W. Automated live microscopy to study mitotic gene function in fluorescent reporter cell lines. *Methods Mol. Biol.* **545**, 113–134 (2009).
15. Boland, M.V. & Murphy, R.F. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics* **17**, 1213–1223 (2001).
16. Wahlby, C., Sintorn, I.M., Erlandsson, F. & Borgefors, G. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *J. Microscopy* **215**, 67–76 (2004).
17. Walker, R. & Jackway, P. Statistical geometric features-extensions for cytological texture analysis. *Proc. 13th Int. Conf. Pattern Recognition* **2**, 790–794 (1996).
18. Haralick, R., Dinstein & Shanmugam Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* **3**, 610–621 (1973).
19. Boser, B.E., Guyon, I. & Vapnik, V. A training algorithm for optimal margin classifiers. *COLT '92: Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (1992).
20. Wang, M. *et al.* Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. *Bioinformatics* **24**, 94–101 (2008).
21. Chen, X., Zhou, X. & Wong, S.T. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE Trans. Biomed. Eng.* **53**, 762–766 (2006).
22. Durbin, R.R., Eddy, S., Krogh, A. & Mitchison, G. *Biological sequence analysis: probabilistic models of proteins and nucleic acids.* (Cambridge University Press, 1998).
23. Viterbi, A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inf. Theory* **13**, 260–269 (1967).
24. Baum, L.E., Petrie, T., Soules, G. & Weiss, N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164–171 (1970).
25. Dempster, A.P., Laird, N.M. & Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **39**, 1–3 (1977).
26. Meraldi, P., Draviam, V.M. & Sorger, P.K. Timing and checkpoints in the regulation of mitotic progression. *Dev. Cell* **7**, 45–60 (2004).
27. Hagting, A. *et al.* Human securin proteolysis is controlled by the spindle checkpoint and reveals when the APC/C switches from activation by Cdc20 to Cdh1. *J. Cell Biol.* **157**, 1125–1137 (2002).
28. Bollen, M., Gerlich, D.W. & Lesage, B. Mitotic phosphatases: from entry guards to exit guides. *Trends Cell Biol.* **19**, 606–616 (2009).
29. Wolthuis, R. *et al.* Cdc20 and Cks direct the spindle checkpoint-independent destruction of cyclin A. *Mol. Cell* **30**, 290–302 (2008).
30. Wang, M., Zhou, X., King, R.W. & Wong, S.T. Context based mixture model for cell phase identification in automated fluorescence microscopy. *BMC Bioinformatics* **8**, 32 (2007).

ONLINE METHODS

Cell culture, RNAi and cell transfection arrays, and western blotting. HeLa 'Kyoto' cells were cultured in DMEM (Gibco) supplemented with 10% FCS (PAA Laboratories) and 1% penicillin-streptomycin (Invitrogen) and grown on LabTek chambered coverslips (Nunc) for live microscopy. All experiments were performed with monoclonal cell lines stably expressing combinations of the fluorescent markers as indicated throughout the manuscript. Live imaging was in DMEM containing 10% FCS and 1% penicillin-streptomycin, but without phenol red and riboflavin to reduce autofluorescence of the medium. Cell transfection arrays for live-cell RNAi screening were produced and used as described previously[7,31]. All other RNAi interference experiments were performed using single RNAi duplexes (Qiagen) that were liquid phase–transfected with either Oligofectamine (Invitrogen) or HiPerfect (Qiagen) as the transfection reagent according to the manufacturers' protocols. Final siRNA concentrations were 50 nM for Oligofectamine or 10 nM for HiPerfect. Cdc20 siRNA validation oligos were obtained from Qiagen with the following target sequences: 5′-AACCTTGTGGATTGGAGTTCT-3′ (Cdc20_1), 5′-CACCACCATGATGTTCGGGTA-3′ (Cdc20_2). Total HeLa cell lysates for SDS-PAGE analysis were prepared according to standard procedures. Rabbit anti-human Cdc20 (diluted 1:5,000) was from Bethyl laboratories.

Fluorescent reporter plasmid constructs. For efficient generation of cell lines stably expressing fluorescently tagged marker proteins, genes derived from plasmids in refs. 32–36 were subcloned into pIRES-puro2 and pIRES-neo3 vectors (Clontech) that allow expression of resistance genes and tagged proteins from a single transcript. For details on the plasmids, see **Supplementary Table 5**.

Stably expressing cell lines. For generation of stably expressing cell lines, HeLa Kyoto cells were first transiently transfected using FuGENE6 (Roche) following the manufacturer's instructions. Cells were then seeded to clonal density and grown in culture medium supplemented with 500 µg ml^{-1} geneticin (Invitrogen) and/or 0.5 µg ml^{-1} puromycin (Merck/Calbiochem) for 3 weeks. Individual colonies of resistant cells were picked, expanded and validated for homogeneous expression levels and correct subcellular localization of fluorescent proteins. All cell lines used in this study had a normal morphology and cell-cycle progression as compared to the maternal line. For details on the stable cell lines, see **Supplementary Table 6**.

Live microscopy. Automated microscopy with reflection-based laser auto focus was performed on a Molecular Devices ImageXpressMicro screening microscope equipped with a 10× 0.5 numerical aperture (NA) and 20× 0.8 NA S Fluor dry objectives (Nikon) and recorded as two-dimensional time series. The microscope was controlled by in-house–developed Metamorph macros (PlateScan software package, available at http://www. bc.biol.ethz.ch/people/groups/gerlichd). Cells were maintained in a microscope stage incubator at 37 °C in humidified atmosphere of 5% CO_2 throughout the entire experiment. We adjusted illumination conditions such that cell death rate was below 5% in untreated control cells[14]. Confocal microscopy was performed on a customized Zeiss LSM 510 Axiovert microscope using a 63×, 1.4 NA oil Plan-Apochromat objective (Zeiss). The microscope

was equipped with piezo focus drives (piezosystemjena), custom-designed filters (Chroma) and EMBL incubation chamber, providing a humidified atmosphere at 37 °C with 5% CO_2.

Image analysis. Cell nuclei were detected by local adaptive thresholding[7], which is robust toward variable expression levels of the fluorescent chromatin marker in individual cells and inhomogeneous illumination typical for widefield microscopy. To improve segmentation accuracy, we implemented a split-and-merge approach. First, we split objects containing directly adjacent nuclei, using watershed transformation based on object contours. In some cases, this incorrectly split single objects. Thus we implemented object merging based on a priori definition of size and circularity criteria[16]. Regions of interest for the secondary marker were derived by region growing of the chromatin segmentation to a fixed size but constrained by regions of neighboring cells. Depending on the marker, we defined nuclear, cytoplasmic or total cellular areas. This segmentation strategy turned out to be more precise than direct segmentation in the secondary channel, as many secondary markers dramatically changed in intensity levels or pattern throughout the time course of the experiment. Texture and shape features[17,18] (**Supplementary Table 1**) were extracted from the two channels and all regions individually. For secondary region classification, only texture features were used because the shape information only depended on the chromatin segmentation.

Samples for morphology classes were manually annotated on the original images overlaid with the segmentation contours to establish a training set for supervised classification. Support vector classification with radial-based kernel and probability estimates[37] was then computed with libSVM. Classification performance was calculated with fivefold cross-validation. Samples and feature plots for all classifiers used in this study can be accessed online through a web browser interface (see http://www.cellcognition.org/).

Tracking cells over time was achieved by a constrained nearest-neighbor approach based on the Euclidian distance between objects[21]. As tracks might be lost because of segmentation errors or migration of cells into the field of view, the tracking must be able to create new tracks for all objects without incoming edges. To detect cell division events, or potential cell-to-cell fusion events, the tracking algorithm needed to support both splitting and merging. This yielded a hierarchical directed graph of isolated tracks for each cell over time. Tracking errors resulted mostly from segmentation errors and lead to wrong edges between the cell tracks. Secondary objects are tracked indirectly by the primary objects associated with them. Mitotic motifs were detected in this graph structure by the transition from prophase to prometaphase. Subgraphs (mitotic trajectories) were extracted by considering a predefined number of frames preceding and following this mitotic motif, resulting in synchronized mitotic trajectories of equal length, as displayed in the figures.

Hidden Markov model and statistical analysis. A hidden Markov model λ is defined as $\lambda = (\mathbf{X}, \mathbf{A}, \mathbf{Y}, \mathbf{B}, \boldsymbol{\pi})$, in which $\mathbf{X}$ is the set of hidden states, $\mathbf{A}$ is a matrix of transition probabilities from one state to another, $\mathbf{Y}$ is the set of observable variables per state, $\mathbf{B}$ is a matrix of observation probabilities storing the probability of observation k being produced from state j (also termed emission or observation probability), and $\boldsymbol{\pi}$ is a vector of probabilities of the initial state (first time point) in the trajectory.

The hidden states **X** are the true cellular stages expressed by the class labels (eight classes for fluorescent histone H2B; **Fig. 1b**). The hidden Markov model is learned by maximum likelihood estimates from the aligned trajectories of estimated prediction probabilities of the support vector machine, which is a three-dimensional array over trajectories, time points and classes. Transition probabilities **A** are learned from the prediction probabilities along the trajectories on the underlying graph structure. In a free model all transitions between morphology classes were allowed (**Fig. 2c**). In a constrained model some transitions were suppressed based on biological a priori knowledge (transition probabilities were set to 0 for edges missing in the graph; **Supplementary Fig. 6a**). For the initial probabilities **π** the prediction probabilities of all trajectories at the first time point are considered. The observables **Y** are the class labels. The observation probabilities were either set to an error rate of 0.1%, or derived from the confusion matrix of support vector machine training.

Using the Viterbi algorithm, each trajectory was corrected based on its sequence of support vector machine probability estimates and the trained hidden Markov model for a given experimental condition (decode problem). This correction scheme was calculated individually for each marker and experimental perturbation condition.

To detect the onset of nuclear envelope breakdown and nuclear envelope reformation the time series of IBB-EGFP, intensity ratios of individual cells were analyzed. We computed the ratio by a shrunken area of the chromatin object and a ring around. The onset was defined as the time point where the ratio was 1.5-fold increased above the ratio at the time point of chromosome segregation.

For data normalization in **Figure 5b** we computed the z scores of mitotic exit timing for all siRNA conditions (mean over all values of one condition). The z score was computed by the mean of negative controls and the s.d. of the entire dataset.

Implementation and performance. The basic image processing was implemented in C++ using VIGRA (vision with generic algorithm) (http://hci.iwr.uni-heidelberg.de/vigra/) and in house-developed extensions. The C++ code was then wrapped for Python, which is a programming language particularly well-suited for handling complex data structures and integration of external modules. Statistical analysis and plots were performed with the R project (http://www.r-project.org/). The entire software package is platform-independent, and was compiled for Mac OS X and Windows environments.

Computation of each movie required 4–20 s per image and processor node, consuming 500–1,500 megabytes of RAM, depending of the number of frames and objects per frame. As an example, a single movie of **Figure 2** with 206 frames and ~37,000 objects required a total processing time of 34 min on a single processor node. For high-throughput analysis, we implemented distributed computing on a farm of desktop computers (four MacPro 2.2GHz, 28 cores total).

Software and data resources. CecogAnalyzer is a platform-independent graphical user interface, which covers the entire workflow presented in this paper. The software is publicly available in source and binary versions and was tested on MacOS X Leopard/SnowLeopard and Windows XP/7. We use a subversion repository for concurrent software development by remote contributors and tracking of software changes. Our website is based on the project management tool TRAC (http://trac.edgewall.org/), which allows coordination of this open-source project by milestones, tickets, wiki pages and browsing of code changes.

The software, a subset of raw images presented here, the classifiers and parameters used for generating the figures are available online at http://www.cellcognition.org/. The classifiers datasets consisting of annotated samples and extracted features are interactively visualized by Adobe Flex and can be browsed online at http://flex.cellcognition.org/.

The MetaMorph journals developed for fast and robust acquisition of the time-lapse experiments presented here are available on our group website: http://www.bc.biol.ethz.ch/people/groups/gerlichd.

31. Erfle, H. *et al.* Reverse transfection on cell arrays for high content screening microscopy. *Nat. Protoc.* **2**, 392–399 (2007).
32. Snapp, E.L. *et al.* Formation of stacked ER cisternae by low affinity protein interactions. *J. Cell Biol.* **163**, 257–269 (2003).
33. Dultz, E. *et al.* Systematic kinetic analysis of mitotic dis- and reassembly of the nuclear pore in living cells. *J. Cell Biol.* **180**, 857–865 (2008).
34. Schaub, B.E., Berger, B., Berger, E.G. & Rohrer, J. Transition of galactosyltransferase 1 from trans-Golgi cisterna to the trans-Golgi network is signal mediated. *Mol. Biol. Cell* **17**, 5153–5162 (2006).
35. Leonhardt, H. *et al.* Dynamics of DNA replication factories in living cells. *J. Cell Biol.* **149**, 271–280 (2000).
36. Steigemann, P. *et al.* Aurora B-mediated abscission checkpoint protects against tetraploidization. *Cell* **136**, 473–484 (2009).
37. Wu, T.F., Lin, C.J. & Weng, R.C. Probability estimates for multi-class classification by pairwise coupling. *J. Machine Learning Res.* **5**, 975–1005 (2004).