Supplementary Methods

Genome-wide siRNA library provided by Ambion

Before production, the siRNA sequences were mapped (using BLAST) against the human genome using ENSEMBL genome database version 27. Only designs with 100% homology to the target gene, unique targeting in the genome and targeting all or most transcripts of the target gene were accepted for production. SiRNAs were redesigned if they failed these criteria. The silencer siRNAs were all chemically synthesized with a length of 21 bases and 2 bases overhang.

qRT-PCR

HeLa cells were reverse transfected with 30 nM gene-specific siRNA or 30 nM Silencer Negative Control #1 siRNA, using siPORT NeoFX Transfection Agent. 48 hours post-transfection, RNA was isolated, and then two-step qRT-PCR was performed using TaqMan Gene Expression Assays. 18S rRNA was used as an endogenous control. Validation results are expressed as average percent mRNA remaining, relative to negative control siRNA-treated samples, along with the standard deviation of multiple independent experiments.

Production of transfected cell microarrays for live cell imaging

A previously described protocol[21] was used with minor modifications. The siRNA-gelatine transfection solution was prepared in 384 well plates (NalgeNunc) as described using a Microlab STAR (Hamilton) liquid handling robot with the following modifications. 5 µl of siRNA solution (30µM), 3 µl Opti-MEM (Invitrogen) containing 0.4 M sucrose and 3.5 µl Lipofectamine 2000 (Invitrogen) were mixed and incubated for 20 minutes at room temperature. After incubation 7.25 µl of 0.2% gelatine and $3.5x10^{-4}$% fibronectin (both Sigma-Aldrich) were added. In some experiments 0.5 µl of a 40 µM marker solution of Cy3 labelled DNA oligonucleotide was added together with the siRNA to judge transfection efficiency. These final siRNA transfection cocktails were then arrayed onto single-well

chambered LabTek coverglass live cell imaging dishes (NalgeNunc) using a ChipWriter Compact Robot (Bio-Rad) with solid pins (Point Technologies) resulting in a spot volume of ~4 nl (containing ~5 ng siRNA) and a spot diameter of ~400 μm and a spot to spot distance of 1125 μm mirroring one 384 well plate per chamber. SiRNA microarrays were printed in 48 replicates then dried and stored in plastic boxes containing drying pearls orange (Fluka) at least over night. After drying, $1.1x10^5$ HeLa-H2B-*GFP* cells were plated on the microarrays in a total volume of 1.5 ml culture medium (DMEM containing 4.5 g/l glucose, 10 % heat-inactivated foetal calf serum, 2 mM glutamine, 100 U/ml penicillin and 100μg/ml streptomycin), and incubated for 16 h at 37°C and 5% CO_2. Cells were transfected simply by growing on siRNA spots without further manipulation. After 16 h incubation the 1.5ml culture medium was removed and replaced by 4 ml preheated CO_2-independent imaging medium (Invitrogen) (containing 10% heat-inactivated foetal calf serum, 2 mM glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin). In order to prevent any gas exchange the live cell microarrays were sealed using baysilone paste (Bayer) and kept for one additional hour in the pre-heated incubation chamber of an Olympus (IX-81; Olympus-Europe) microscope.

Production of coated 384 well plates for live cell imaging (validation screen)

The siRNA-gelatine transfection solution was prepared as described in the production of the transfected cell microarrays. 5 μl of the siRNA transfection solution was mixed with 130 μl of water using the manual 96 well liquid handling device Liquidator (Steinbrenner). 15 μl of this water diluted siRNA transfection solution was pipetted in 384 well imaging plates (BD) and immediately lyophilized using the concentrator miVAC (Genevac). In total 8 replicates were produced and stored in plastic boxes containing drying pearls orange (Fluka) until usage. 750 HeLa-H2B-*GFP* cells / per well were seeded into the coated 384 well imaging plates (BD) using the cell seeding device Multidrop Combi (Thermo Scientific) and incubated for 16 h at 37^0 C and 5% CO_2. After 16 h incubation the 50 μl culture medium was removed and replaced by 100 μl preheated CO_2-independent imaging medium (Invitrogen) (containing 10 % heat-inactivated foetal calf serum, 2 mM glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin). In order to prevent any gas exchange the 384 well plates were sealed using baysilone paste (Bayer) and kept for one additional hour in the pre-heated incubation chamber of an Olympus (IX-81; Olympus-Europe) microscope.

High-throughput time-lapse imaging

Images were acquired with an automated epifluorescence microscope (IX-81; Olympus-Europe) as previously described[1,22]. Stabilized light sources (MT20, Olympus-Biosystems, Munich), firewire cameras (DB-H1, 1300x1024, Olympus-Biosystems) and an in house modified version of the "ScanR" software (now commercially available with all new developments from OSIS), EMBL environmental microscope incubator (EMBL, GP 106) as well as a new objective (Plan10x, NA 0.4; Olympus-Europe) and filter sets for *GFP* (Chroma Inc.) were implemented. An image based autofocusing routine[22] was used to focus on the maximum number of interphase 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 48 hours without further autofocusing between the 30 min imaging intervals. As cells needed on average ~1 hour to go through mitosis, we obtained ~2 mitotic images per cell enabling us to detect delays ≥ 30 minutes, or a mitotic index change from 5 % to 7.5 %. Higher time resolution would increase the sensitivity to even more subtle delays but increase overall data volume. Illumination of the specimen was reduced to the minimum necessary for sufficient signal to noise ratio for automated phenotyping (see below) with exposure times of 18-50 ms (*GFP*). Data was compressed without loss (ZIP) on the fly, buffered locally on the microscope computers and then saved on a 5 TB network-attached-storage. Microscopes were connected via a 2 GBit/s network to the NAS to sustain uninterrupted data flow.

Quality Control

Imaging data from the siRNA microarrays underwent a two-step quality control procedure based on automatic and manual inspection of the data, first on the level of the whole microarray and second on the level of individual siRNA spots to ensure stringent quality control criteria for technical quality and performance of positive and negative controls. Imaging data from each siRNA microarray was first checked manually in order to assure that the images were in focus, that the positive (siRNAs targeting INCENP, KIF11, COPB1) and negative (scrambled siRNA) controls behaved as expected, the cells looked healthy and that the image covered the spot region. Second, automatic quality control rules were applied: a slide was accepted for further evaluation if at least in 5 out of 7 (6 out of 8) negative controls no significant phenotype was detected automatically and the cell density was within reasonable limits (initial cell count between 15 and 200 and end cell count under 500) and if

in addition 7 out of 11 positive controls showed the expected phenotype. After application of these quality control steps (both manual and automatic), data from 512 out of 674 imaged siRNA microarrays (~76 %) were kept for further evaluation.

A second level of quality control concerned single siRNA spots in order to deal with local quality issues like floating dead cells or contamination by cells from neighbouring spots. These situations were detected automatically and checked manually, which led to the removal of 9382 (4.8 %) individual spots.

Image Processing

Each time lapse experiment was analyzed fully automatically via an in house developed object recognition pipeline[2]. In order to identify the nuclei/mitotic chromosome sets (both referred to as nuclei for brevity in the following) in each frame, the image was prefiltered (Median Filter, Toggle Mappings[24]) and thresholded after background subtraction. For each of the identified nuclei, 190 features were calculated describing both, shape and texture (basic grey level and shape features, Haralick features[25], morphological granulometries[26,27], morphological dynamics[27], standard moments[28] and moment invariants[29], statistical geometric features, convex hull features). The representation of each nucleus as a vector in a 190-dimensional feature space allowed its classification by support vector machines (software: libSVM[30]) into one out of 16 predefined morphological classes (Fig. 1b and Supplementary Fig. 1).

In addition, the nuclei were tracked over time by a constrained nearest-neighbour method using Euclidean distance of nuclei centres. Nuclei with no predecessor at time t-1 and all their followers were not considered any further (i.e. only nuclei in the spot region at the first time point and their daughter nuclei were used for the hit detection).

Hence, each time-lapse experiment was represented by a set of time series each of which corresponded to the temporal evolution of the percentage of nuclei in one of the predefined morphological classes. Each time series was then smoothed by local regression[31]. For each time-lapse experiment and for each morphological class, a score was derived as the maximal difference over time between the experiment curve and the corresponding average negative control curve (the average was calculated over all negative control series on the same siRNA microarray). The siRNA score for each morphological class was defined as the "upper median" of the corresponding technical replicate scores, where "upper median" means the median in case of an odd number of replicates and the higher of the two centre values in case of an even number of replicates. Hence, if the upper median is higher than a certain threshold,

at least 50% of the technical replicates are also higher than this threshold. A siRNA was a hit if its score was higher than a significance threshold (see "Detection of hits" for the thresholds in the different phenotypic categories).

Several morphological classes were merged into one class for hit detection: the mitotic delay class was defined as the sum of "prometaphase" and "MAP" (metaphase alignment problems), as these morphologies were a strong indication of mitotic delays or arrests (with a time-lapse of 30 minutes, prometaphases are only rarely detected in normal cells). We did not find significant deviations in normal "metaphase" or "anaphase" classes and therefore did not include them in the class relevant for the detection of mitotic hits. The dynamic class was defined as the sum of "hole", "small irregular" and "folded" morphologies. These joint morphological classes indicated similar biological phenotypes; the individual classes were only kept separately for subsequent analysis, where an increased phenotypic resolution might be of interest. We joined these classes for hit detection because we required for reproducibility the technical replicates to score in the same phenotypic class. This is certainly the most stringent way of obtaining siRNA hits. However, if the phenotypic resolution is too high, i.e. if phenotypic classes describe subtle differences of overall similar biological processes (like "MAP" and "prometaphase"), this can lead to a decrease in sensitivity, as siRNAs scoring reproducibly in strongly related phenotypic classes, but not reproducibly in any of them separately, would then not be considered as hits. The merged classes are calculated as the sum of the single classes before calculation of the replicate score (i.e. the time-series were added, rather than the final scores).

To summarize the siRNA scoring: for each replicate, we calculated one score in each morphological class (and joint classes) as the maximal difference over time between the time series in that class and the average negative control time series. The siRNA score was defined as the upper median of the corresponding replicate scores, and a gene was considered as a "potential hit" if at least one targeting siRNA had a score above the threshold. A gene was considered as a "validated hit" if two or more siRNAs resulted in a consistent phenotype.

Detection of "potential hits" based on nuclear morphology: Mitosis, survival, nuclear size and shape.

The list of potential mitotic hits (one or more siRNAs) was the union of four hit lists, each of which was obtained by applying the following thresholds to the scores obtained as described above: Mitotic Delay (sum of prometaphase and MAP) 4.0 %, Binuclear 9.2 %, Polylobed

11.0 % and Grape 3.0 % (all thresholds refer to the maximal difference to the average negative control curve in the corresponding phenotypic category).
For the non mitotic morphological classes, the following thresholds were applied: Cell Death 7.2 %, Dynamic (sum of "Hole", "Folded" and "Small Irregular") 6.197 % and Large 6.76 % (all thresholds refer to the maximal difference to the average negative control curve in the corresponding phenotypic category).
All thresholds were manually determined and more than three standard deviations above the average score of the entire data set (with control experiments removed), except the cell death threshold (2.5 standard deviations above the average) (Fig. 1c and Supplementary Fig. 9). In addition, manually identified false positives were removed from the list of potential mitotic hits (union of "Mitotic Delay class", "Binuclear", "Polylobed" and "Grape").

Proliferation hit detection

The proliferation of a replicate was defined as the ratio of end cell count to initial cell count. In order to deal with slide-to-slide-variations, the average proliferation of the negative controls was calculated for each slide and subtracted from the experiment proliferation. The siRNA proliferation score was defined as the upper median of the replicate proliferation scores. A manually defined threshold (1.22, more than 3 standard deviations above the mean) was applied in order to obtain a list of siRNAs resulting in an accelerated proliferation; the supposedly targeted genes were the potential hits in this class.

Identification of migration hits

After automatic registration of stage repositioning inaccuracies, the tracking information allowed the quantification of nuclear movement in the experiments.
Two parameters were considered for movement quantification: speed and displacement. For the first, the frame-to-frame distance was determined and averaged over all interphase nuclei in each frame. This resulted in a time series of average frame to frame distance values, which was fitted by local regression. The maximal difference to the corresponding average negative control curve formed the score of the speed category for each experiment. As before, the upper median value of the technical replicate scores was defined as the siRNA score, and a siRNA was considered a hit if its score exceeded a manually defined threshold (1.841, more than 3 standard deviations above the mean). For the displacement category, simple subtracks without divisions and fusions were determined and the average Euclidean distance between start and end points of these subtracks was calculated. The difference to the corresponding

negative control average distance formed the replicate score. The siRNA score was defined as the upper median score of its replicates, and a siRNA was considered a hit if its score exceeded a manually fixed threshold (11.3, more than 3 standard deviations above the mean). The list of candidate genes whose depletion potentially resulted in increased nuclear motility was the set of genes supposedly targeted by these siRNAs.

Event Order Maps (EOM)

Event Order Maps visualize the typical temporal succession of phenotypic events for a group of genes[2]. Starting from the representation of each time lapse experiment by a set of time curves in the classes "Mitotic Delay", "Binuclear", "Polylobed", "Grape", "Large", "Dynamic" and "Cell Death", such an order can be defined by comparing characteristic time points for each of the corresponding curves. For this purpose, a characteristic time point was defined for each class as the first time point for which the difference between the time curve and the corresponding control curve reached at least 70% of its maximal value and exceeded the class threshold. The distribution of these time points for different phenotypic categories is shown in Supplementary Fig. 10.

The phenotypic events of each technical replicate were then ordered in accordance with these characteristic time points. In order to derive an order of events for a siRNA from several replicate event orders, the best matching permutation of reproducible events was chosen by assigning to each possible permutation the number of event order pairs from the replicates that were respected and then choosing the permutation with the highest score[2].

For each of the phenotypic classes and the corresponding list of siRNAs (one per gene), the event orders were visualized by centring them on the class of interest and dividing them into three groups: one where the centre phenotypic class was the first event and had consequences, one where it was not the first event and one where it was the only phenotype observed. Each of these groups was rearranged separately with maximal compliance between the event orders from left to right in such a way that the more frequent event order sub-patterns were on top. Within groups of identical event orders, the siRNAs were ordered according to the penetrance in the centred phenotype.

Rescue experiments

HeLa-Kyoto cells and BAC cell line pools stably expressing the mouse specific gene used for the RNAi knock down were seeded on individual wells of a "ready to transfect" 8 well chambered cover glass. Before seeding the cells, 150 µl of D-MEM was added to each well

followed by 100 µl of 5 x 10^4 cells / ml of the respective cell line. Both cell lines were fixed with 4% PFA and stained with Hoechst 33342 (1µg/ml) to visualize the chromosomes. The time of fixation after seeding was chosen as the maximal penetrance time of the individual phenotypes extracted from the genome-wide data set. Images were taken with CellR-Software on a scan^R fluorescence microscope from Olympus with a 20 × / 0.7 UplanApo air objective.

Production of 8 well chambered cover glasses "ready to transfect"

The detailed description of the protocol has been previously published[23].

Confocal microscopy

HeLa cells stably expressing GFP-tubulin and Histone 2B-mCherry were seeded on "ready to transfect" 8 well chambered cover glasses. 5 x 10^3 HeLa cells were seeded per well. 19 h after seeding the culture medium (DMEM containing 4.5 g/l glucose, 10% heat-inactivated foetal calf serum, 2mM glutamine, 100U/ml penicillin and 100µg/ml streptomycin) was replaced by CO_2-independent imaging medium containing 10% heat-inactivated foetal calf serum, 2mM glutamine, 100 U/ml penicillin and 100µg/ml streptomycin (Invitrogen). The chambered cover glasses were sealed with baysilone paste and kept for one additional hour in the pre-heated incubation chamber of a Leica SP5 confocal micrcroscope. 20 hours after cell seeding Leica Application Suite Advanced Fluorescence (LASAF) and Leica Matrix Screening Application (MSA) was used to image cells for 50 hours in 3D with a time resolution of 12 min. A 63 x / 1.40 -0.60 oil λ blue objective, a 561 nm and a 488 nm laser were used to acquire 12 optical sections of 1.25 µm thickness (FWHM) over a 15 µm range. Maximum intensity projections of the 12 sections are shown in the images and supplemental movies.

siRNAs to genes mapping

Ensembl 53 was used as a reference genome. The guide strand of each siRNA was mapped to the human transcriptome using blastn with the following parameters:

-E=10 -B=100 -W=7 -M=4 -N=-3 -Q=3 -R=3 -nogaps

All hits with less than 3 mismatches were recorded. For hit gene identification, only siRNAs with a unique target gene were considered i.e. a siRNA has a unique target gene if it maps to exactly one gene with no mismatch. However, at www.mitocheck.org, all targets with less than 3 mismatches are also shown.

Gene Ontology annotations

GO term annotations of genes were retrieved from the Ensembl database. For Fig. 3a and Supplementary Fig. 3, a flat slim ontology was created by selecting pertinent terms from the full ontology and gene annotations were then mapped to this slim ontology. To make the categories non-redundant with respect to the genes, the list of GO terms in the slim ontology was arbitrarily ordered by relevance to mitosis and each gene was only assigned to the first term from the list it had been annotated with.

Hierarchical Clustering of hits

For each gene, a representative siRNA was chosen according to the maximal ratio between phenotypic score and corresponding threshold over all mitotic classes. In order to cluster genes (each represented by one siRNA) in accordance with temporal phenotypic similarity, the phenotypic time-series were averaged over all replicates for this siRNA. Hence, each gene was represented by a set of 6 time series describing the trajectory in a six dimensional phenotype space, where the six dimensions corresponded to the six phenotypic classes "Mitotic Delay", "Binuclear", "Polylobed", "Grape", "Large", "Dynamic" and "Cell Death". Each trajectory was then represented by two vectors (LSE approximation) in order to catch the global tendencies rather than small variations. In order to define a pairwise distance between genes, the distance between the two representing vector pairs was calculated as a combination of Euclidean and angular distances[2]. Ward's method was used for clustering.

Manual annotation

Development of the automatic hit detection and controlling the quality of the data was only possible by performing manual annotation of a part of the data set in order to define the relevant morphological classes, assess and judge the quality of the data and verify the outcome of the various image processing steps. We have in total annotated 18423 movies (9333 negative and positive control movies and 9090 experiment movies; movies that did not pass quality control are not counted) documenting the presence or absence of phenotypes of 2271 siRNAs targeting 1545 genes. Out of the 1545 genes annotated, 710 (46%) scored as manual hits. This annotation occurred during several stages of the project: before availability of the complete automatic algorithm, only proliferation arrests were automatically scored and – as proliferation arrests are highly unspecific - manually annotated. During training of the classification algorithm, a set of genes was annotated including also randomly picked genes. The largest set of movies was annotated during quality control and evaluation of the

thresholds for hit detection. Most of the manually annotated genes were also found by our automatic method (71%). The remaining 29% were either just below the threshold or showed a phenotype that was not scored by our method (e.g. metaphase delays or chromosome bridges).

The following phenotype classes have been used: Prometaphase delay, metaphase delay, metaphase alignment problems, segregation problems, failure in condensation, strange nuclear shape, nuclei stay close together and condensation followed by de-condensation.

A siRNA was considered as a hit if the majority of the corresponding movies was annotated with the same phenotype category. The results of the manual annotation can also be found at: www.mitocheck.org.

References

23. **Erfle et al. Work flow for multiplexing siRNA assays by solid-phase reverse transfection in multiwell plates. J Biomol Screen 13(7), 575-80 (2008).**

24. **Meyer, F. & Serra, J. Contrasts and Activity Lattice. *Signal Processing* 16, 303-317 (1989).**

25. **Haralick, R. M. & Dinstein, I. Spatial Clustering Procedure for Multi-Image Data. *IEEE Transactions on Circuits and Systems* CA22, 440-450 (1975).**

26. **Serra, J. *Image Analysis and Mathematical Morphology* (Academic Press, Inc., Orlando, FL, USA, 1983).**

27. **Soille, P. *Morphological Image Analysis: Principles and Applications* (Springer-Verlag Berlin, Heidelberg, New York, 1999).**

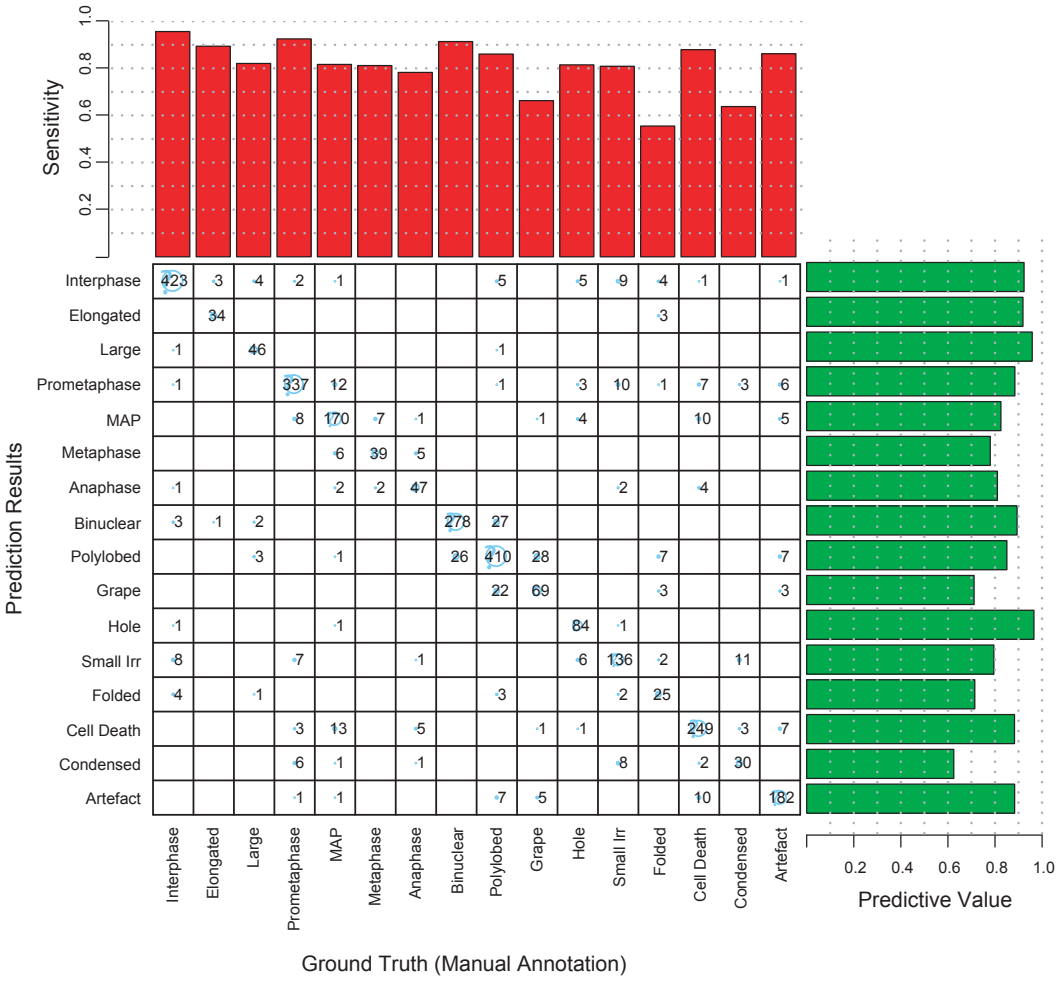
28. **Prokop, R. J. & Reeves, A. P. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing* 54, 438--460 (1992).**

29. **Hu, M.-K. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* IT, 179-187 (1962).**

30. **Chang, C.-C. & Lin, C.-J. LIBSVM: a library for support vector machines (2001).**

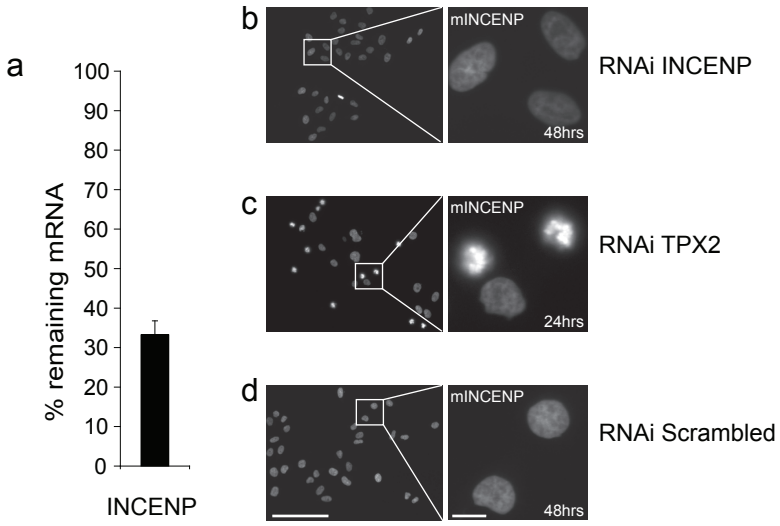
31. **Loader, C. *Local Regression and Likelihood* (Springer, 1999).**

Confusion matrix : Results for the classification of chromosome configurations

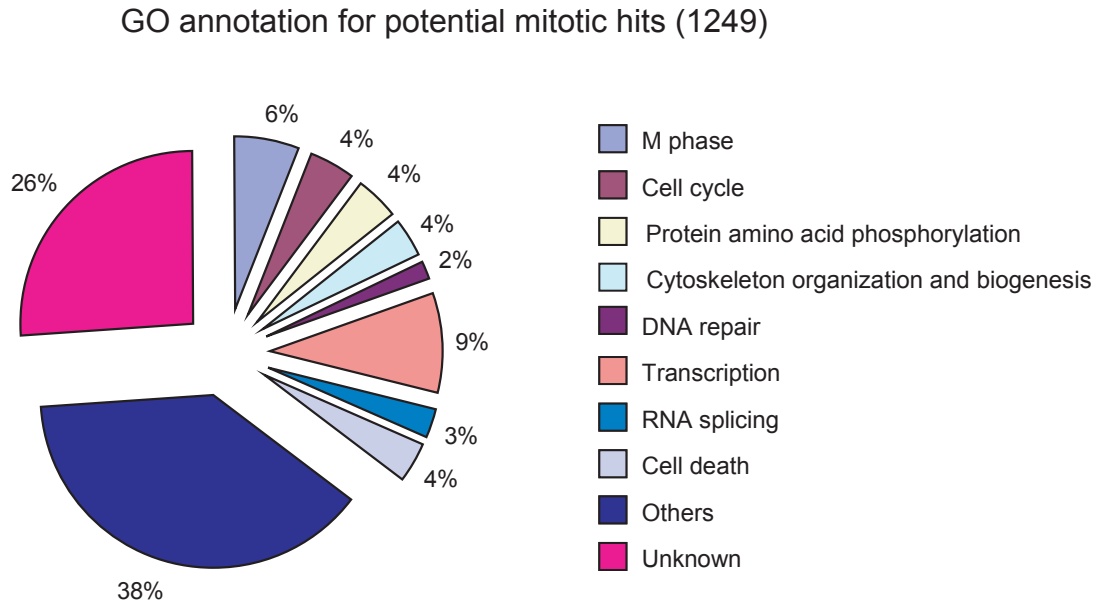


Supplementary Figure 1 | Confusion matrix. The confusion matrix (bottom left) shows the comparison between manual and automatic annotation for the nuclei of the training set. The rows correspond to the prediction result, the columns to the manually assigned labels (ground truth). The confusion matrix has been determined by 20-fold cross validation. The class specific sensitivity is shown on the top, the class specific prediction value on the right.

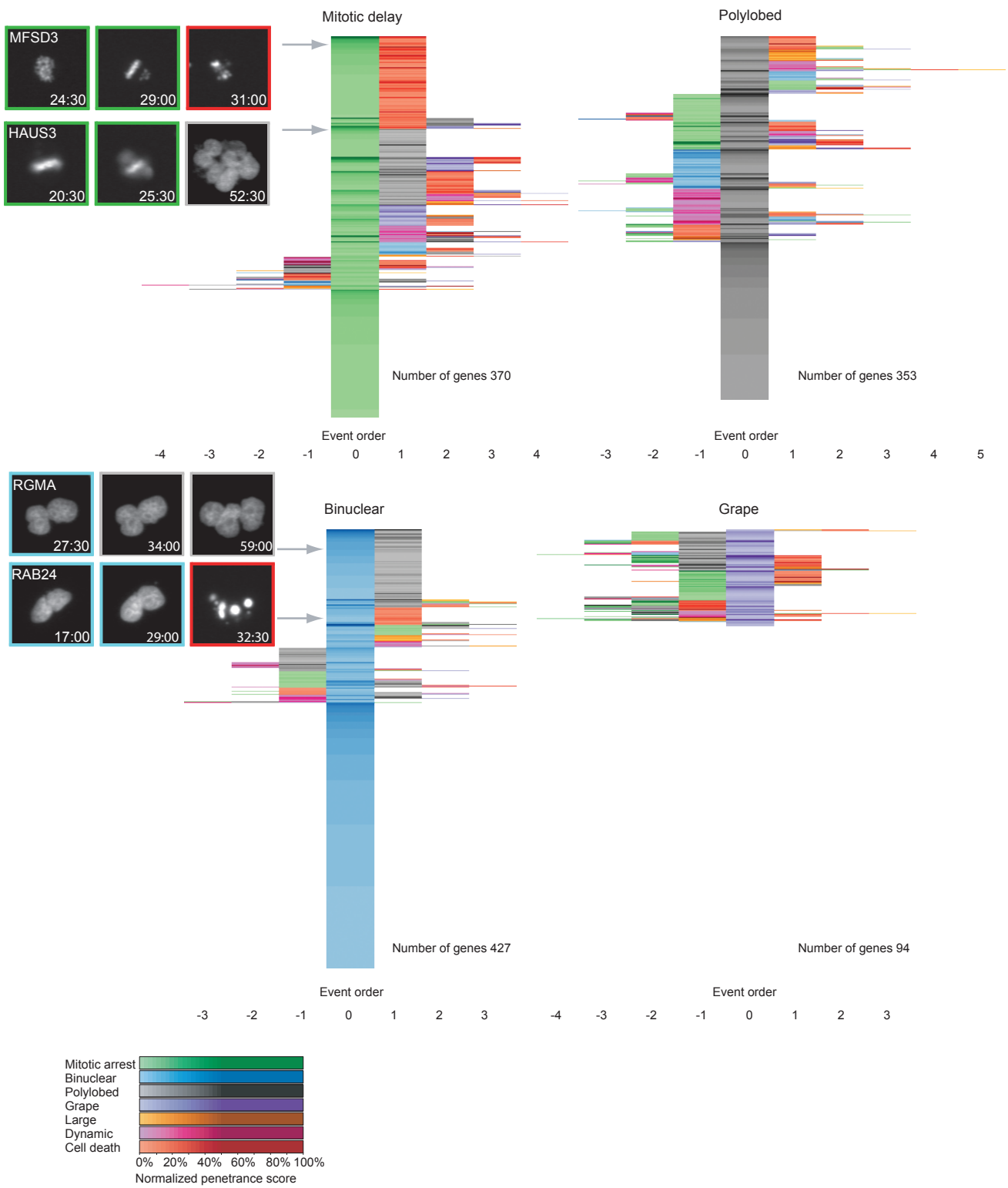
Rescue experiments with a BAC cell line expressing mouse INCENP protein



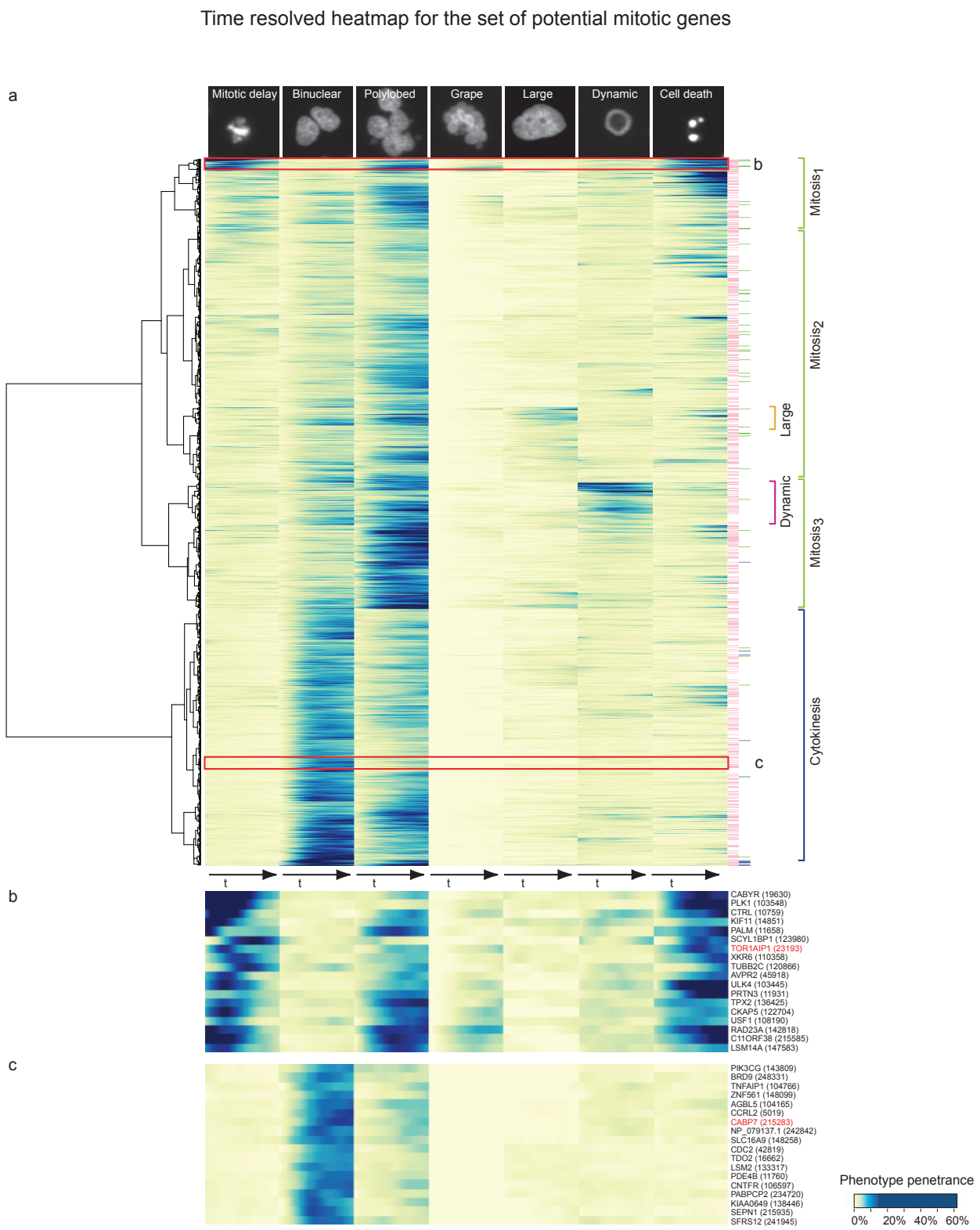
Supplementary Figure 2 | **Rescue control experiments.** a, After knock down of the human INCENP mRNA, b, the mouse protein is able to rescue the INCENP phenotype. c, TPX2 RNAi knock down experiments clearly show the expected Prometaphase arrest phenotype, demonstrating that the expression of the mouse INCENP protein is not interferring with the RNAi efficiency in general. d, Negative control siRNA does not have any phenotypic effect on the mouse BAC cell line.



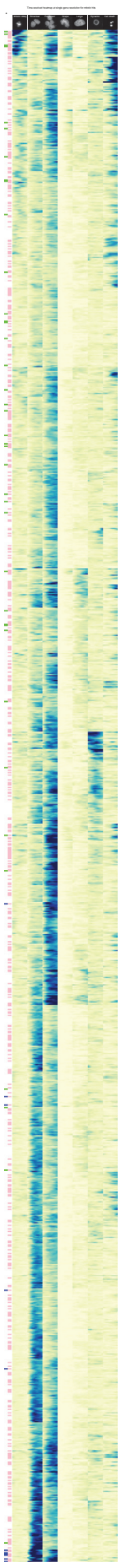
Supplementary Figure 3 | GO analysis. Biological process annotations of the 1249 potential mitotic hits (identified by at least one siRNA). To deal with multiple annotations for each gene, categories were arbitrarily ordered by relevance to mitosis and each gene was assigned to the first term from this list it had been annotated with.

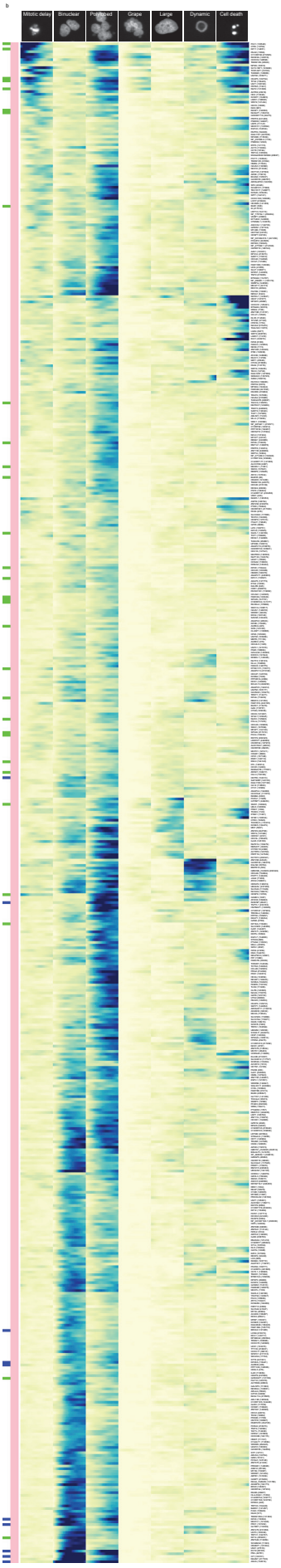


Supplementary Figure 4 | Event Order Maps (EOM) for potential mitotic hits. To each gene of the whole set of potential mitotic hits (reproducibly scoring for at least one siRNA), a representative order of phenotypic events and a normalized penetrance score in each of the morphological classes is assigned. For each gene, the event order can be visualized by a sequence of coloured fields where different colours correspond to different phenotypic classes, the colour intensity to the corresponding penetrance and the colour order to the phenotypic event order. The event orders are then centred on the phenotypic classes “mitotic delay”, “binuclear”, “polylobed” and “grape”. Genes with similar event orders are grouped together resulting in a centred event order map, where the rows correspond to genes and the columns to the event order relative to the main phenotypic event.

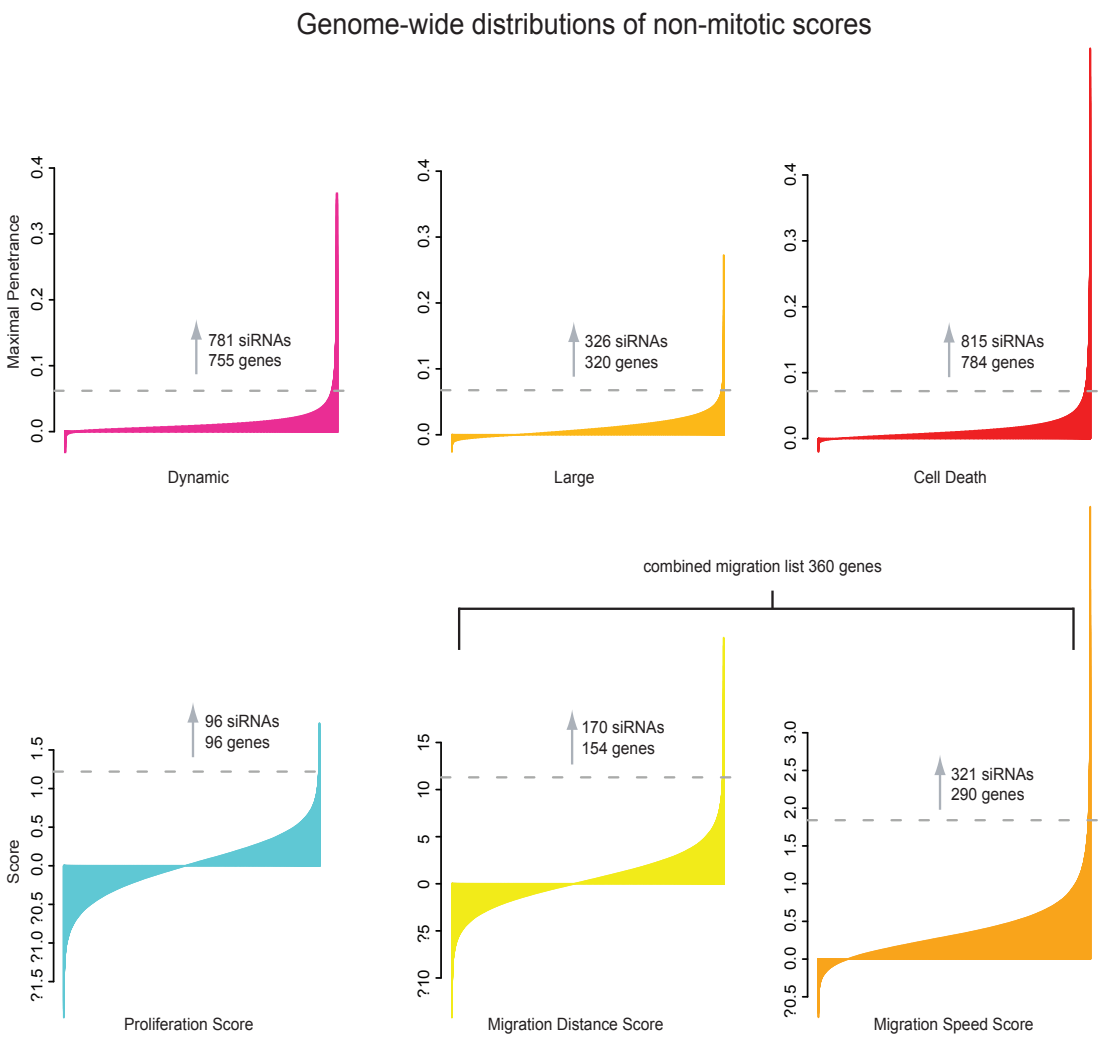


Supplementary Figure 6 | Time resolved Heatmap for the whole set of potential mitotic genes. a, The quantitative time-resolved phenoprints for all potential mitotic hit genes (1249) can be used for phenotypic clustering, taking into account both the penetrance values and the joint temporal evolution in several phenotypic classes illustrated at the top of each column. Colour code at the right: validated genes are marked in pink, reference genes (Supplementary Table 5) with known function are marked in blue (cytokinesis) or green (early mitotic phenotype). On the right, interesting clusters are highlighted. b, Zoom in: early mitotic phenotypes (corresponding to the top red square in Supplementary Fig. 6a); the rescued gene TOR1AIP1 is highlighted in red. c, Zoom in: binuclear phenotypes (corresponding to the bottom red square in Supplementary Fig. 6a); the rescued gene CABP7 is highlighted in red.

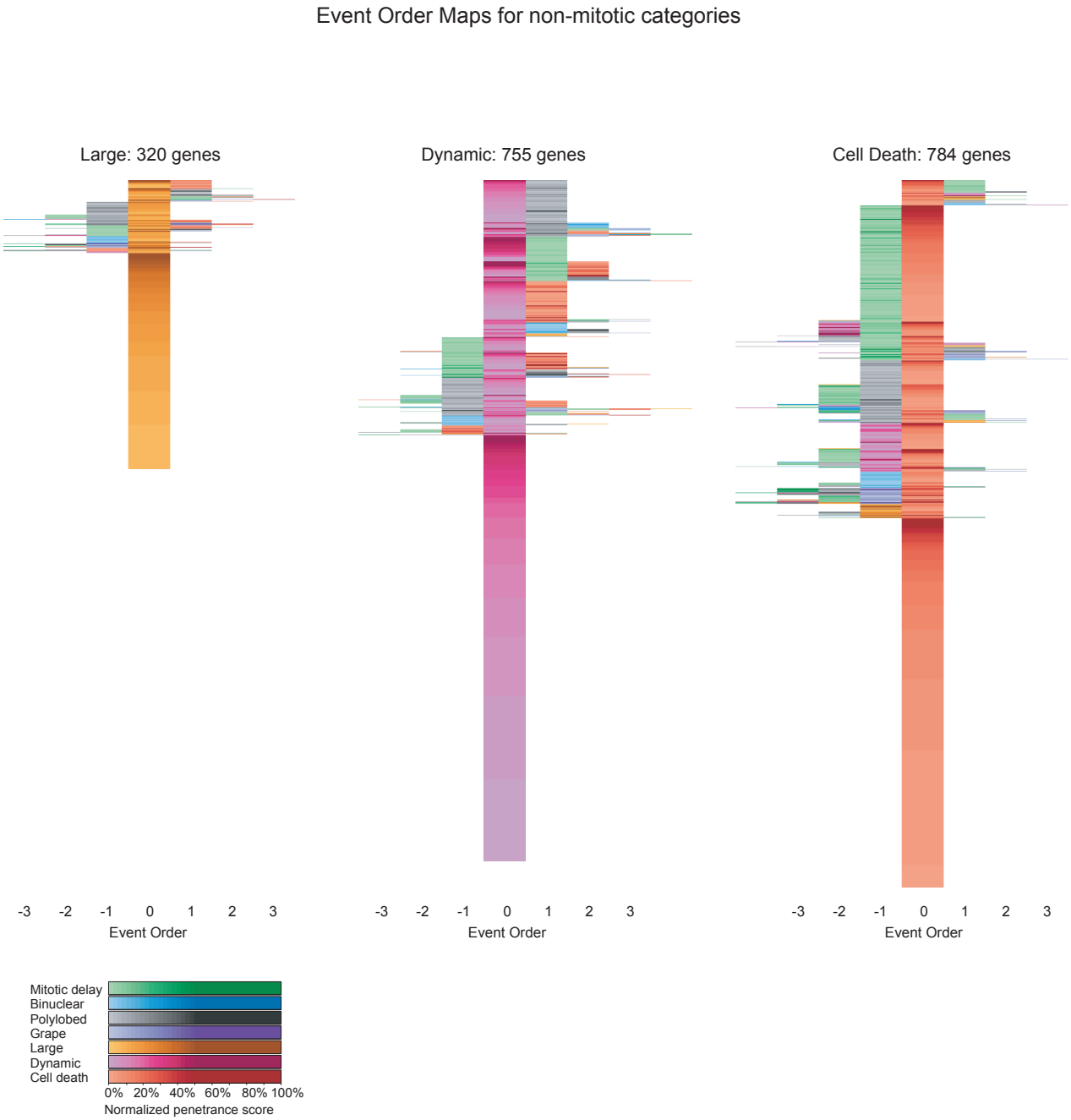

Time-resolved heatmap of single gene resolution for mitotic hits



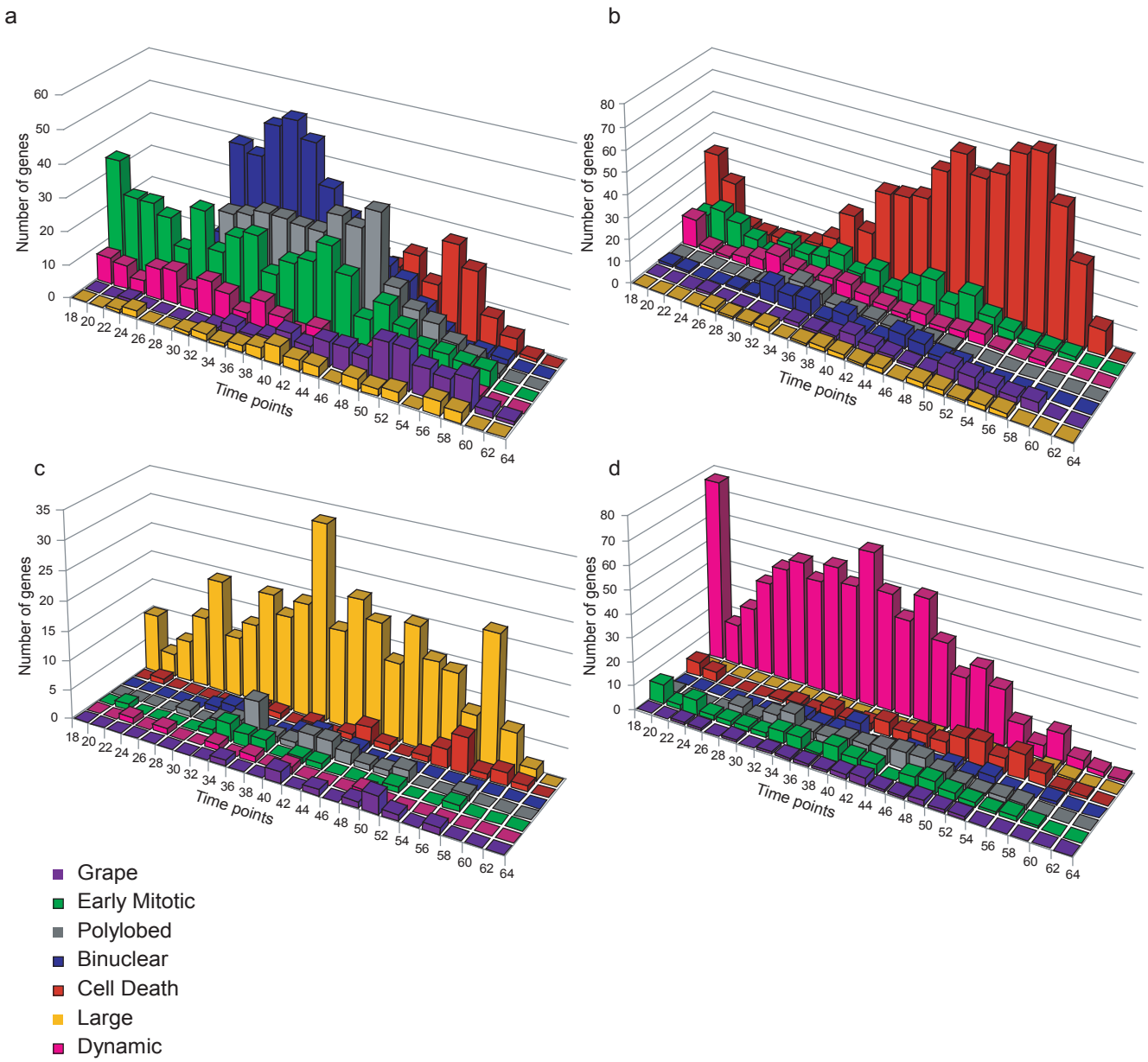
Supplementary Figure 7 | Time-resolved heatmap at single gene resolution for mitotic hits. Each row corresponds to one siRNA targeting one gene (gene symbols and siRNA identifier are shown at the right side of the plot; the corresponding raw data and plots can be seen at www.mitocheck.org). At the left side of the plot, genes with known function are highlighted by the colors blue (cytokinesis) and green (early mitotic phenotype), validated genes (i.e. genes with a consistent phenotype for two siRNAs or more) are marked in pink. a. All potential mitotic hits including non-validated genes (1249). b. All validated mitotic hits (972).



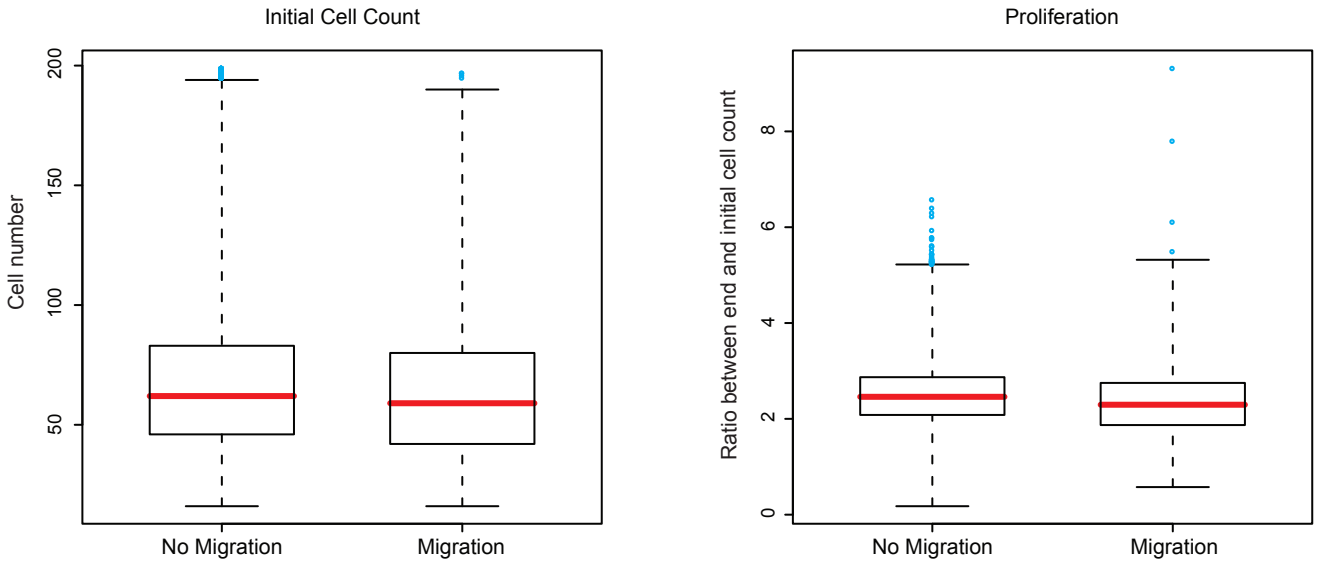
Supplementary Figure 8 | Genome-wide distributions of non-mitotic scores. To each siRNA, a score is assigned in each of the non-mitotic categories. Each siRNA is considered as a hit, if the median score of its replicates exceeds a manually defined threshold (dotted lines); a gene is defined as a “potential hit” if is targeted by at least one siRNA from the hit list. These “potential hits” were not validated by independent siRNAs. Top row: score distribution for the maximal penetrances in the categories "Dynamic", "Large" and "Cell death". Bottom row: score distributions for proliferation, distance (subtrack distance) and speed (frame to frame distance)



Supplementary Figure 9 | EOMs for non-mitotic categories. For the potential hit lists (identified by at least one siRNA) in the categories "Large", "Dynamic" and "Cell Death", the characteristic order of events has been determined for each potential hit gene. For each gene, the event order can be visualized by a sequence of coloured fields where different colours correspond to different phenotypic classes, the colour intensity to the corresponding penetrance and the colour order to the phenotypic event order. The genes are rearranged according to their event order pattern, centred on the respective hit category (”Large”, “Dynamic” and “Cell Death”). The centred hit category gets the order index 0; all categories before have a negative, all categories after a positive order index.



Supplementary Figure 10 | Distributions of characteristic timepoints for different phenotypic categories. For each potential hit gene (reproducibly scoring for at least one siRNA in one of the relevant morphological classes), the characteristic timepoints of all reproducibly and significantly deviating morphologies were determined as described in the text. a. Mitotic (1249 genes). b. Cell Death (784 genes). c. Large (320 genes). d. Dynamic (755 genes).



Supplementary Figure 11 | Initial cell count and proliferation distribution for migration data. a. The initial cell count was determined for experiments with and without increased migration score. b. The proliferation (ratio between end cell count and initial cell count) is shown for experiments with and without increased migration score.

Information for Supplementary Movies 1-40

Group1:

Supplementary Movie 1: Interphase and anaphase nuclei

Supplementary Movie 2: Elongated interphase nuclei due to VIPR RNAi

Supplementary Movie 3: Large interphase nuclei due to NFKB RNAi

Supplementary Movie 4: Prometaphase arrested nuclei due to KIF11 RNAi

Supplementary Movie 5: Metaphase alignment problems (MAP) due to CKAP5 RNAi

Supplementary Movie 6: Metaphase delay due to FAM33A RNAi

Supplementary Movie 7: Binuclear cells due to CIT RNAi

Supplementary Movie 8: Polylobed nuclei due to INCENP RNAi

Supplementary Movie 9: Grape shaped nuclei due to RAD23 RNAi

Supplementary Movie 10: Interphase nuclei showing hole structures due to ZADH1 RNAi

Supplementary Movie 11: Interphase nuclei showing small irregular shapes due to SARS2 RNAi

Supplementary Movie 12: Interphase nuclei showing folded shapes due to PIGR RNAi

Supplementary Movie 13: Fragmented nuclei (cell death) due to COPB2 RNAi

Supplementary Movie 14: Cells showing condensed nuclei due to PABPC1 RNAi

Supplementary Movie 15: Two cell nuclei (interphase and cell death) on top of each other representing the artefact class

Group2:

Supplementary Movie 16: Cutout of the highlighted cell in Supplementary Movie 1 shows normal interphase and anaphase morphology.

Supplementary Movie 17: Cutout of the highlighted cell in Supplementary Movie 2 shows an elongated interphase nucleus (RNAi VIPR).

Supplementary Movie 18: Cutout of the highlighted cell in Supplementary Movie 3 shows a large interphase nucleus (RNAi NFKB).

Supplementary Movie 19: Cutout of the highlighted cell in Supplementary Movie 4 shows a nucleus lasting for hours in Prometaphase resulting in cell death (RNAi KIF11).

Supplementary Movie 20: Cutout of the highlighted cell in Supplementary Movie 5 is showing metaphase alignment problems for hours resulting in grape shaped nuclei (RNAi CKAP5).

Supplementary Movie 21: Cutout of the highlighted cell in Supplementary Movie 6 is showing a metaphase delay (RNAi FAM33A).

Supplementary Movie 22: Cutout of the highlighted cell in Supplementary Movie 7 is showing a binuclear cell (RNAi CIT).

Supplementary Movie 23: Cutout of the highlighted cell in Supplementary Movie 8 is showing a cell with no segregation in the first mitotic event and with segregation problems when re-entering mitosis resulting in a polylobed nucleus (RNAi INCENP).

Supplementary Movie 24: Cutout of the highlighted cell in Supplementary Movie 9 is showing metaphase alignment problem resulting in a grape shaped nucleus (RNAi RAD23A).

Supplementary Movie 25: Cutout of the highlighted cell in Supplementary Movie 10 is showing dynamic nuclear morphologies e.g. hole like structures (RNAi ZADH1).

Supplementary Movie 26: Cutout of the highlighted cell in Supplementary Movie 11 is showing dynamic nuclear morphologies e.g. small irregular shaped nucleus (RNAi SARS2).

Supplementary Movie 27: Cutout of the highlighted cell in Supplementary Movie 12 is showing dynamic nuclear morphologies e.g. folded nucleus (RNAi PIGR).

Supplementary Movie 28: Cutout of the highlighted cell in Supplementary Movie 13 is showing a fragmented nucleus indicating cell death due to COPB2 RNAi.

Supplementary Movie 29: Cutout of the highlighted cell in Supplementary Movie 14 is showing a condensed nucleus without going into mitosis (RNAi PABPC1).

Supplementary Movie 30: Cutout of the highlighted cell in Supplementary Movie 15. Artefact class – Two cell nuclei on top of each other. One dead cell is floating on top of an interphase cell representing one example of the artefact class.

Group3:

Supplementary Movie 31: RNAi Negative control

Supplementary Movie 32: RNAi AURKB – No segregation occurs resulting in malformed nuclei followed by cell death.

Supplementary Movie 33: RNAi INCENP – During the first cell cycle no segregation occurs resulting in large nuclei, which show a segregation problem during the second cell division leading to polylobed cell nuclei.

Supplementary Movie 34: RNAi TOR1AIP1 - Metaphase alignment problem resulting in cell death.

Supplementary Movie 35: RNAi CENPE - Metaphase alignment problems resulting in binucleated cells followed by segregation problems during the second cell cycle leading to polylobed nuclei.

Supplementary Movie 36: RNAi OGG1 – Segregation problems in both cell cycles resulting in polylobed cell nuclei. Formation of a multipolar spindle in the second cell cycle is visible.

Supplementary Movie 37: RNAi PTGER2 – Cytokinesis failure during the first cell cycle resulting in arrest of binucleated cells without entering mitosis again.

Supplementary Movie 38: RNAi ECT2 – Cytokinesis failure followed by the formation of a multipolar spindle during the second cell cycle resulting in multinucleated cells.

Supplementary Movie 39: RNAi CABP7 – Cytokinesis failure followed by the formation of a multipolar spindle during the second cell cycle resulting in multinucleated cells.

Supplementary Movie 40: RNAi C13orf23 – Cytokinesis failure followed by the formation of a multipolar spindle during the second cell cycle resulting in multinucleated cells followed by cell death.