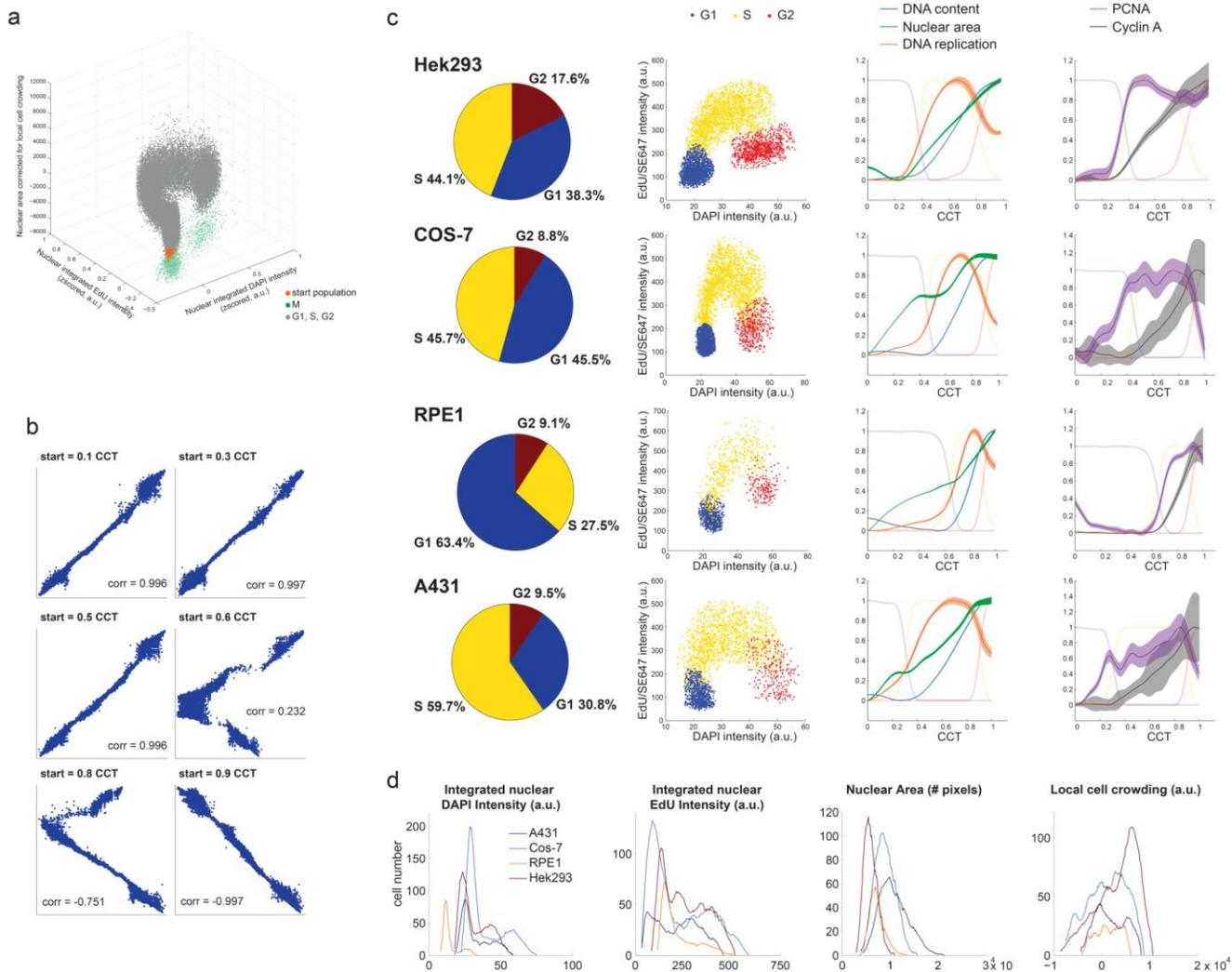


## Supplementary Figure 1

### Workflow of Cyclor and performance checks.

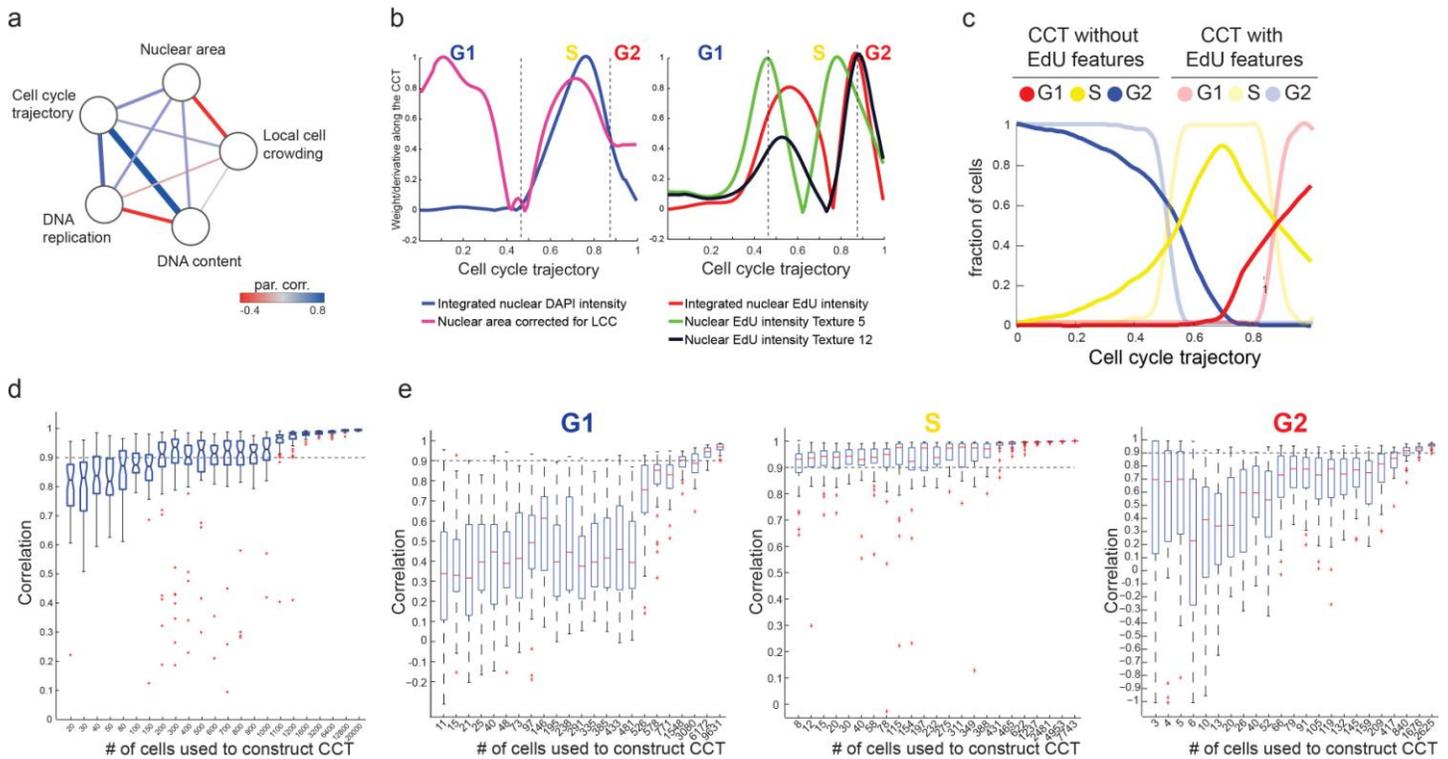
(a) Single-cell features are extracted from the images, and then, Cyclor is used to generate the cell-cycle trajectory (CCT). In parallel, Support Vector Machines (SVMs) are trained to detect S and M phase cells, G1 and G2 cells are subsequently classified by k means clustering and Gaussian fitting (Online Methods). (b) Distributions of the 5 single-cell features used by Cyclor to construct the CCT, black line represents the whole population, blue, yellow and red line represent G1, S and G2 population, respectively (c) All 5 features used by Cyclor to construct a CCT are plotted along the CCT (solid lines: weighted local mean). (d) Outlier cells classified by the SVM as G1, S or G2 cell (G1-, S- and G2-class) which are found in their non-corresponding phase along the CCT. Boundaries between G1-, S- and G2-phase are marked with a gray dashed line, arrow mark outlier cell population, which has been quantified. Percentage of miss-classified cells per cell cycle phase and class are displayed in the matrix. Misclassified cells never exceed 4% of a class. (e) Synchronized populations of cells released from mitotic block recapitulate dynamics of the Cyclor features along the CCT as a function of cell cycle progression (release time from the mitotic block). The violin plots represent the single cell distributions for the 5 Cyclor features, for every time point in a cell synchronization experiment (Online Methods) and are separated per cell cycle phase (G1 in blue (top row), S in yellow (middle row), G2 in red (bottom row)). The weighted mean of the individual features of a CCT generated from an independent and unsynchronized population of cells is plotted in gray over the violin plots. If a cell has been release for 4 hours from the mitotic block (early G1, arrow 1), it will have low DNA content (low DAPI intensity), low DNA replication activity (low EdU intensity) and a small nuclear area (which is corrected for local cell crowding (LCC)). As release time from the mitotic block increases (S, arrow 2) cells will start to replicate DNA (high EdU intensity), the DNA content will be intermediate (intermediate DNA intensity, because of only partially replicated genome) and the nuclear area corrected for LCC will have grown too. As release time is further increased (G2, arrow 3), DNA content (high DAPI intensity) and corrected nuclear area will have reached the maximum and DNA replication (EdU intensity) will be low again. Therefore cells transit through the feature space (with non-linear feature relations), used by Cyclor to construct the CCT, as a function of cell cycle progression (release time from mitotic block) and at rates reproduced by the CCT (gray solid line).



**Supplementary Figure 2**

**Robustness and general applicability of Cycler.**

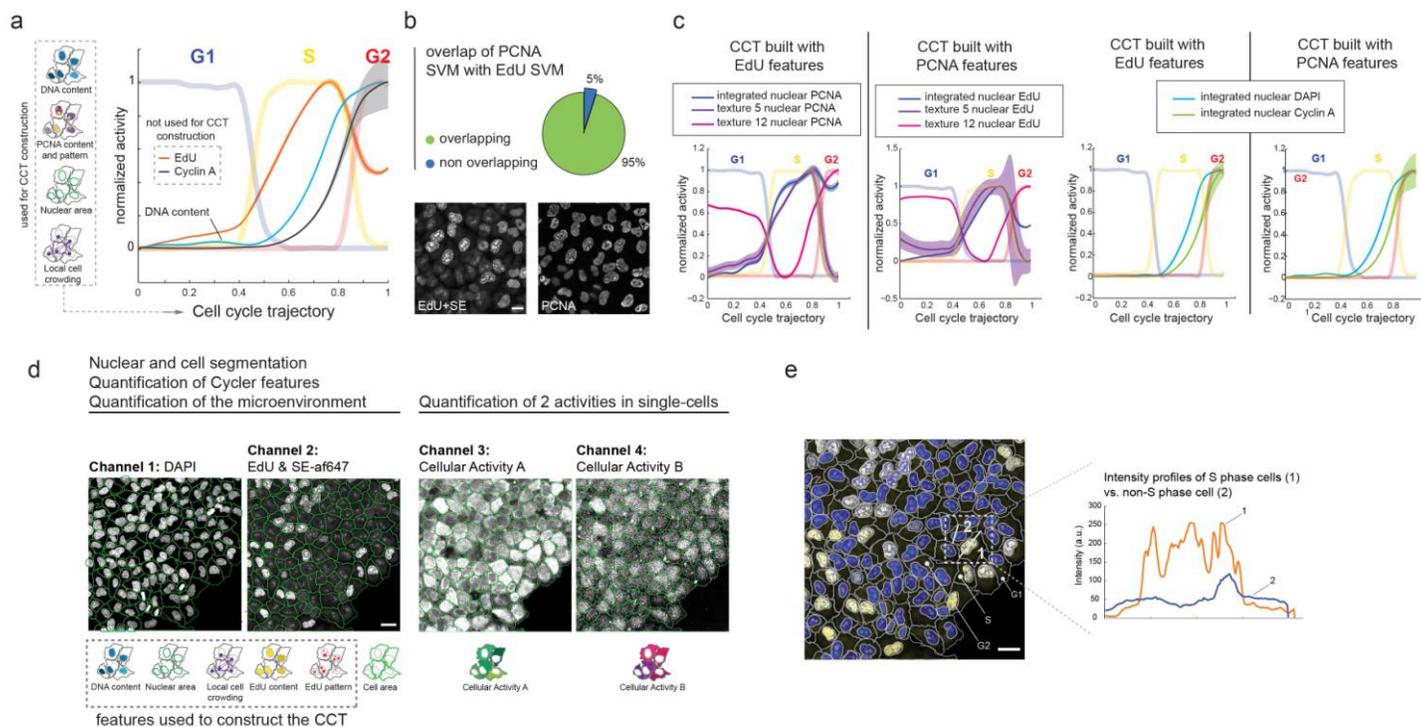
(a) Scatter plot of a cell population. Interphase cells (G1, S and G2 cells) are colored in gray, mitotic cells in green, the population of cells from which Cycler is started is colored in orange and is the subset of G1 cells, which is most similar to mitotic cells with completed meta phase (DNA content is halved) with respect to nuclear DNA content, nuclear area corrected local cell crowding and nuclear EdU content. (b) Robustness analysis for the selection of the start population to construct a reproducible cell-cycle trajectory (CCT). A CCT has been constructed starting from a start population of G1 cells with the smallest nuclear area and lowest nuclear DAPI content (x-axis, baseline trajectory). Subsequently Cycler has been re-run ten times with the start cell parameter advancing across the baseline trajectory a new start population has been selected Cycler has been re-run six times with the start cell parameter advancing across the baseline trajectory. X-axes are baseline trajectory, Y-axes are Cycler's output for given start population. Each dot is a cell along the trajectory. Correlations displayed are measured using Pearson's correlation. The CCT is well correlated with the baseline for start $\leq$ 0.5 and is inversely correlated (the algorithm detects the reverse trajectory) for start=0.9 and 1.0. For other values of start, Cycler broke the trajectory in half and ordered each half either correctly or inversely. (c) CCTs were constructed using Cycler in 4 cell lines, with highly variable cell cycle phase propensities and cellular characteristics, generating consistent dynamics of nuclear DNA content, nuclear EdU content, nuclear area, nuclear PCNA and nuclear cyclin A content along the CCT. Beginning from the left: ratio of G1 S and G2 phase observed in a population with heterogeneous microenvironments; scatterplot of integrated nuclear DAPI intensity and integrated nuclear EdU intensity; plots of DNA content, nuclear area and DNA Replication (EdU) along the CCT (solid lines: weighted local mean, semi-transparent area: weighted local standard deviation); plots of integrated nuclear PCNA and cyclin A intensity along the CCT (solid lines: weighted local mean, semi-transparent area: weighted local standard deviation). (d) Histograms of 4 single-cell features used by Cycler to construct CCTs in the 4 cell lines.



### Supplementary Figure 3

#### The role of cell crowding and other single-cell features and the number of single cells in Cycler's performance.

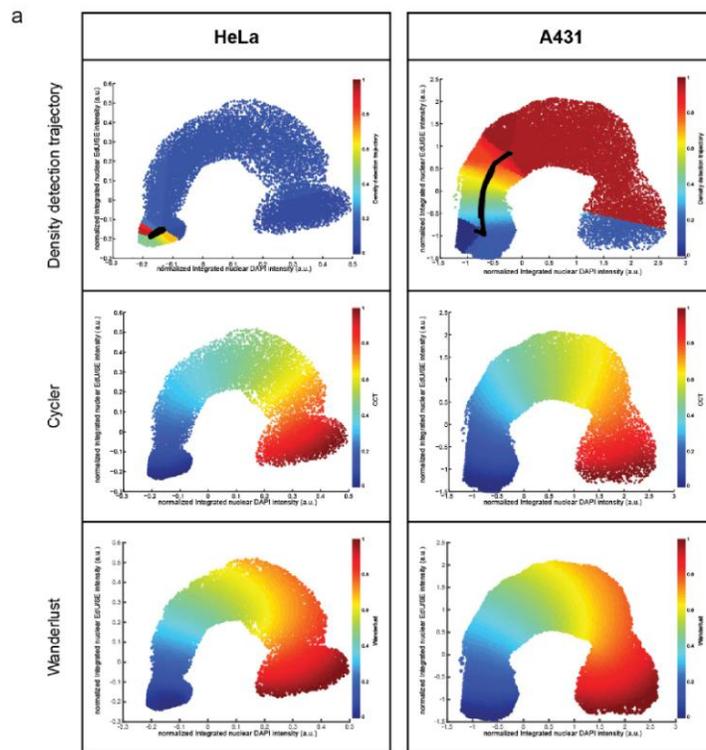
(a) Network where each edge denotes the partial correlation between cellular characteristics represented on the nodes. Size of the edge represents the strength of the correlation. Blue edges are positive correlations, while red edges are negative correlations. (b) Weights (derivatives along the CCT) of the features used by Cycler to build a CCT, reveal that nuclear area corrected for local cell crowding is the strongest determinant of CCT progression in G1, the EdU Texture features 5 and 12 drive the CCT through G1/S and S/G2 transition, whereas DAPI intensity and EdU intensity are mainly responsible for CCT progression in S phase. (c) Distribution of the discrete cell cycle phases G1, S and G2 along a CCT built with and without nuclear EdU intensity and texture features, respectively. A CCT constructed only with nuclear area corrected for local cell crowding and nuclear integrated DAPI intensity yields already a good estimate of a cell cycle progression trajectory but the individual cell cycle phases are poorly separated. Including features of DNA replication (EdU intensity and EdU texture), allows to separate G1, S and G2 thoroughly (d) Analysis of reproducibility of Cycler (Online Methods). The correlation was measured between a CCT constructed with 24,000 cells and lower cell numbers (x-axis). Boxplot represent the correlation calculated between 50 CCTs with a given cell number (x-axis) and the CCT generated with 24,000 cells, CCTs constructed with 1,200 cells reach 0.98 correlation and a CCT with 12,800 cells achieves maximal correlation at 0.95. (e) Reproducibility of Cycler per cell cycle phase (Online Methods). Correlations were measured between a CCT constructed with 24000 cells and lower cell numbers (x-axis) for every cell cycle phase independently. Boxplot represent the correlations calculated between 50 CCTs with a given cell number in a certain cell cycle phase and the CCT generated with 24,000 cells in a certain cell cycle phase, X-axis label represent the mean cell number of cells in a certain cell cycle phase of 50 randomly subsampled populations at a given cell number (see x-axis of Supplementary Fig. 3d).



## Supplementary Figure 4

### Enabling Cyclus to use stains for fixed cells combined in two channels.

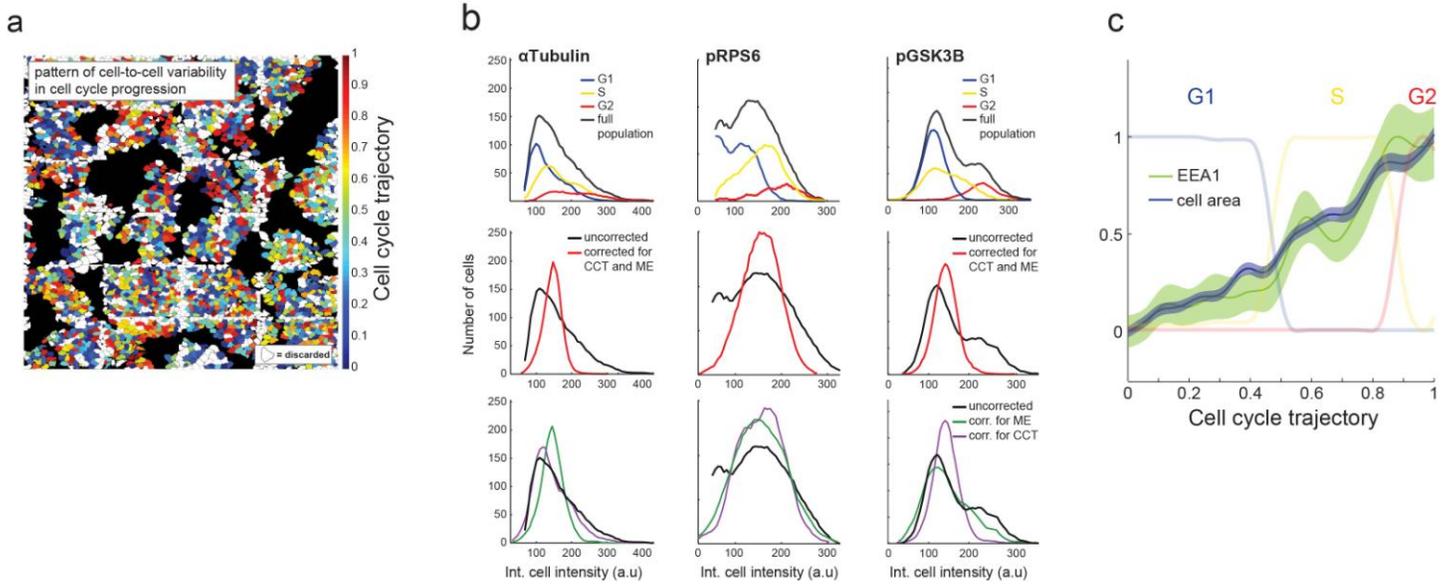
(a) Cyclus can use multiple marker for S-phase to construct bona fide cell-cycle trajectories (CCTs). CCT built using nuclear PCNA content, nuclear PCNA texture feature 5 and 12 instead of EdU features. Both EdU and cyclin A abundances and DNA content show expected dynamics over the CCT almost identical to dynamics observed along a CCT constructed with EdU features. (b) To validate that PCNA and EdU can be interchanged both for the construction of a CCT and the training of the S-phase SVM two independent SVMs were trained to identify S phase cells in the same dataset. The first SVM was trained with intensity and texture features of the nuclear PCNA signal, the second using intensity and texture features of the nuclear EdU signal. The overlap between the two classifications was quantified at the single-cell level. 95% of all cells classified as S phase cells by the EdU SVM were also classified by the PCNA SVM as S phase. This reveals that EdU (costained with SE-af647) is a bona fide marker for S phase cells. Scale bar 15  $\mu$ m. (c) Further comparison of cellular features and staining along CCTs constructed either with EdU or PCNA, reveals striking robustness of Cyclus to construct equivalent CCT with both S-phase markers. First column: integrated nuclear PCNA intensity and texture features plotted along the CCT built with EdU (Intensity and texture) as S-phase marker. Second column: integrated nuclear EdU intensity and texture features plotted along the CCT built with PCNA as S-phase marker. Third column: integrated nuclear DAPI intensity and integrated cellular cyclin A intensity plotted along the CCT built with EdU as S-phase marker. Fourth column: integrated nuclear DAPI intensity and integrated cellular cyclin A intensity plotted along the CCT built with PCNA as S-phase marker. (d) Overview of the enhanced staining protocol allows to integrate Cyclus in standard immunofluorescence protocols without the need to dedicate individual channels specifically to Cyclus. Channel 1 (DAPI, 405nm excitation) is used to segment nuclei, measure nuclear area and DNA content (integrated nuclear DAPI intensity), and quantify the microenvironment of single cells. Channel 2 (647nm excitation): EdU and SE-af647 are costained in the same acquisition channel, DNA Replication (integrated nuclear EdU intensity) and patterns of DNA replication (nuclear EdU texture #5 and #12) are quantified and the cell outline segmented. Channel 3 and 4 (488nm and 546nm excitation) can therefore be used to measure 2 single-cell activities of interest. Scale bar 30  $\mu$ m. (g) Finely balancing the cell outline staining (SE-af647) with the EdU staining allows to segment cell outline and measure DNA replication activity in the same channel. Selected image of cells stained simultaneously with EdU and SE-af647 in the far-red channel (647nm excitation). Intensity profiles of cell 1 (in S-phase and positive to EdU) and cell 2 (cell negative to EdU) show the differences in cell intensities that allows an informative costaining of the two signals. Scale bar 30  $\mu$ m.



## Supplementary Figure 5

Comparison among three algorithms used for the construction of a trajectory of cell-cycle progression.

(a) Trajectories of cell cycle progression generated by density detection trajectory (top), Cyler (middle) and Wanderlust (bottom) are color-coded on a single-cell scatterplot with nuclear DAPI content as x and nuclear EdU content as y-axis in population with heterogeneous microenvironments for HeLa (left) and A431 (right) cells (Online Methods). Each cell is colored-coded based on their position along the CCT, from blue (early in the cell cycle) to red (late in the cell cycle). For density detection, the black line represents the ridge of peak density, as identified by the respective algorithm.



## Supplementary Figure 6

### Heterogeneous patterns and bimodality in cell populations are explained by Cyclcr.

(a) Population of cells in heterogeneous microenvironments is color-coded based on their position along the cell-cycle trajectory (CCT) revealing non-trivial patterns of cell to-cell variability. (b) Distributions of  $\alpha$ -tubulin, pRPS6, and pGSK3B in unsynchronized populations of cells growing in heterogeneous microenvironments. Accounting for variability caused by local cell crowding and cell cycle in populations of cells explains complex multimodal distributions. Top row shows the distributions of the full population (in black) and the distributions of cells in individual cell cycle phases: G1 (blue), S (yellow) and G2 (red). Middle row shows the distributions of the full populations of cells (black) and the same distributions of abundances corrected for both the microenvironment and the cell cycle progression (red). Bottom row shows the uncorrected distributions of the full populations of cells (black) and corrected distributions for microenvironment (ME, green) and cell cycle progression (CCT, purple). (c) Cell area and integrated cellular EEA1 intensity plotted along the CCT (solid lines: weighted local mean, semi-transparent area: weighted local standard deviation).