

## RESEARCH ARTICLE SUMMARY

## CELL BIOLOGY

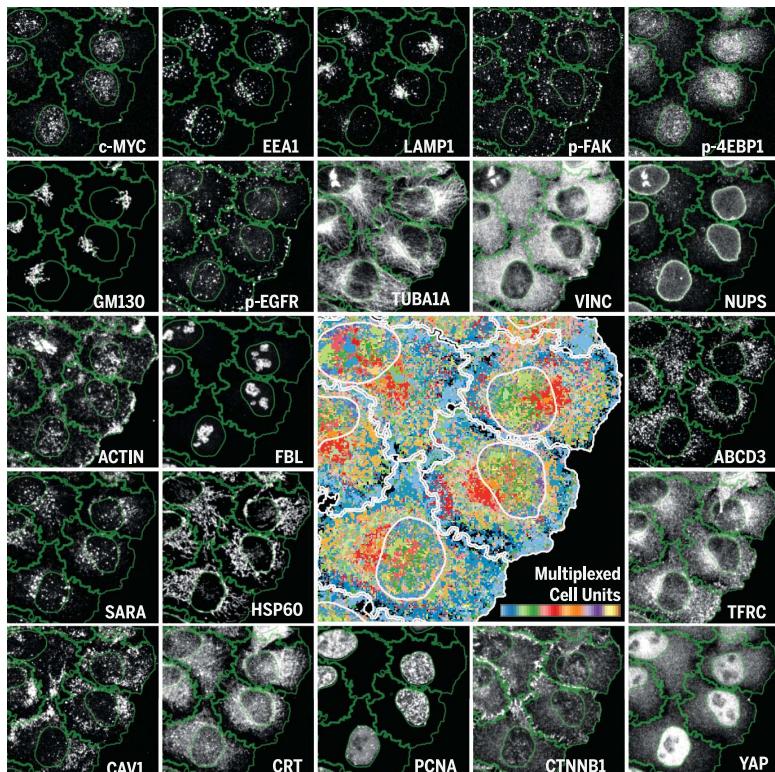
# Multiplexed protein maps link subcellular organization to cellular states

Gabriele Gut<sup>\*†</sup>, Markus D. Herrmann<sup>\*</sup>, Lucas Pelkmans<sup>†</sup>

**INTRODUCTION:** Obtaining multiplexed molecular readouts from large numbers of single cells *in situ* is an important technological goal to facilitate scientific discoveries in basic and translational research. Various methods have been developed in recent years to achieve spatially resolved multiplexed measurements of the abundance of large sets of mRNAs or proteins in biological samples. These technologies have brought the promise that, through large-scale efforts, all functionally relevant cell types of an organism will emerge from such multiplexed data in an unbiased manner. Furthermore, these cell types will be able to be

mapped within their physical context within a tissue.

**RATIONALE:** The involvement of an mRNA or protein (state) in cellular function depends on its specific intracellular location and interaction with other molecules and cellular structures. Moreover, the phenotype of an individual cell is determined by the functional state, abundance, morphology, and turnover of its organelles and cytoskeletal structures. To functionally interpret molecular multiplexed information, such measurements will thus need to resolve the intracellular length scale.



**Iterative indirect immunofluorescence imaging (4i).** 4i obtains 40-plexed protein readouts at high spatial detail in thousands of cells with the use of off-the-shelf antibodies. Multiplexed protein maps derived from such images provide a comprehensive quantitative description of compartmentalized intracellular protein composition. These maps can identify new cellular states and allow quantitative comparisons of intracellular organization between single cells in different cell cycle states, microenvironments, and drug treatments.

**RESULTS:** We report a simple, robust, and nondegrading protocol that achieves 40-plex protein staining in the same biological sample using off-the-shelf antibodies called iterative indirect immunofluorescence imaging (4i). In conjunction with high-throughput automated microscopy and computer vision, 4i allows highly reproducible multiplexed measurements from surface areas of several mm<sup>2</sup> subsampled by pixels of 165 nm by 165 nm. This approach simultaneously captures functionally relevant properties that emerge at the cell population, cellular, and intracellular level. 4i can thus quantify the influence of local cell crowding on

protein abundance, the effect of cell cycle position on protein phosphorylation, patterns of protein subcompartmentalization, and organelle morphology all in the same single cell and across thousands of cells.

We developed a data-driven computer vision approach that generates multiplexed protein maps (MPMs). MPMs comprehensively quantify intracellular protein composition with high spatial detail in large numbers of single cells. They are not confounded by the specific relative geometry and orientation of an individual cell. Thus, MPMs allow systematic comparisons of subcellular spatial protein distribution between single cells that experience different cell cycle states, microenvironments, growth conditions, or exposure to drugs. Using the example of subcellular relocalization of epidermal growth factor (EGF) receptor upon exposure to EGF, we demonstrate that MPMs allow the systematic definition of cellular states undetectable by multiplexed whole-cell measurements. The findings are functionally relevant and can be connected to multiple molecular and phenotypic properties apparent at the cellular and cell population scale.

**CONCLUSION:** By preventing photocrosslinking during imaging, 4i enables multiplexed immunofluorescence with off-the-shelf antibodies, both in small-scale and high-throughput experiments. 4i datasets cover multiple length scales, eliminating the need for extrapolation or inference in the interpretation of results. Integration of these length scales in one dataset reveals a richness of scale-crossing connections that current models of biological processes do not yet consider. These connections determine how gene expression is adapted to the cellular state, how a cell type is determined, how a pathological cellular phenotype emerges, or how a tumor cell responds to a drug. ■

\*These authors contributed equally to this work.

†Corresponding author. Email: gabriele.gut@uzh.ch (G.G.); lucas.pelkmans@imls.uzh.ch (L.P.)

Cite this article as G. Gut et al., *Science* **361**, eaar7042 (2018). DOI: 10.1126/science.aar7042

## RESEARCH ARTICLE

## CELL BIOLOGY

# Multiplexed protein maps link subcellular organization to cellular states

Gabriele Gut<sup>1,2\*</sup>, Markus D. Herrmann<sup>1,3\*†</sup>, Lucas Pelkmans<sup>1†</sup>

Obtaining highly multiplexed protein measurements across multiple length scales has enormous potential for biomedicine. Here, we measured, by iterative indirect immunofluorescence imaging (4*i*), 40-plex protein readouts from biological samples at high-throughput from the millimeter to the nanometer scale. This approach simultaneously captures properties apparent at the population, cellular, and subcellular levels, including microenvironment, cell shape, and cell cycle state. It also captures the detailed morphology of organelles, cytoskeletal structures, nuclear subcompartments, and the fate of signaling receptors in thousands of single cells *in situ*. We used computer vision and systems biology approaches to achieve unsupervised comprehensive quantification of protein subcompartmentalization within various multicellular, cellular, and pharmacological contexts. Thus, highly multiplexed subcellular protein maps can be used to identify functionally relevant single-cell states.

**V**arious methods have revolutionized the ability to obtain multiplexed measurements of the abundance of hundreds or thousands of different molecular species from single cells (5–7). These have brought the promise that through large-scale efforts, all functionally relevant cell types of a human body will, in an unbiased manner, emerge from such data (8, 9). Because some of these methods can be applied *in situ*, the identified cell types can then be placed within the context of a cell population or tissue. However, the abundance of a protein or protein state, or of an RNA transcript, may not be directly informative about its involvement in cellular function. This depends on the specific intracellular location and interaction with other molecules and intracellular structures, which may only involve a small fraction of the total cellular pool (10, 11). Moreover, the phenotype of an individual cell is determined by the functional state, abundance, morphology, and turnover of its intracellular organelles and cytoskeletal structures. Therefore, to obtain functionally relevant information, these unbiased large-scale methods need to extend the length scale of molecular multiplexing into the intracellular domain and ultimately acquire temporal information (12). Recently, a tour-de-force study achieved intracellular im-

munofluorescence imaging of 12,000 proteins, from which an average subcellular map of the human proteome was created (8). However, to understand how the subcellular distribution of the proteome is functionally linked to the phenotypic state of a cell and its microenvironment and how it responds to varying conditions, such maps must be directly measured in the same single cell and across many cells *in situ*. Various powerful methods exist that can achieve spatially resolved antibody multiplexing on tissues or single cells with subcellular resolution (4, 5, 13–21). None, however, combines the flexibility of indirect immunofluorescence with a high-throughput approach that can simultaneously obtain multiplexed protein readouts from the multicellular, single-cell, and intracellular length scales with suborganelle resolution for multiple conditions and extracts the rich amount of biological information that such scale-crossing data contains.

## Multiplexed immunofluorescence with off-the-shelf antibodies

We combined our high-throughput multivariate imaging and computer vision approach (22–24) with an automated liquid handling platform that applies multiple iterations of staining, signal removal, and re-staining, a principle used in other fluorescence-based multiplexing approaches (13, 15, 18). Because photobleaching for signal removal was not practical at these scales, we chose chemical antibody elution. Moreover, to be unrestricted in our choice of antibodies and not require primary antibody conjugations, we turned to conventional indirect immunofluorescence (Fig. 1A and fig. S1, A and B). We noticed that when we imaged one site of the sample with an automated spinning-disk confocal microscope

after a standard immunofluorescence protocol (first imaging), eluted the bound antibodies from the sample, re-stained with only secondary antibody, and checked the efficiency of antibody elution (second imaging), we observed that the elution efficiency was strongly compromised in the region that was exposed to light, but not in regions that were not exposed to light during the first round of imaging (Fig. 1A). We reasoned that this may be caused by the formation of singlet oxygen radicals during the excitation of fluorophores (25, 26), which can introduce covalent bonds between reactive amino acids (methionine and cysteine) of proteins near the fluorophore, resulting in the cross-linking of antibodies to the sample during exposure to high-energy light used in microscopy (27–29). Because this has gone unnoticed in previous antibody elution approaches (15, 18, 21), it may have prevented complete elution and led to the use of particularly harsh conditions that degrade the sample.

We next screened combinations of reagents that prevent such photoinduced cross-linking during imaging without reducing the efficiency of photon emission. These should allow complete antibody elution under very mild conditions that do not visibly remove or degrade antigens (fig. S1C). This approach identified an imaging buffer that contains a radical scavenger as well as an acceptor for free radical-induced photocrosslinking, a very mild elution buffer, which relies on a reducing agent, low pH and chaotropic salts, and a blocking buffer that blocks free sulphydryl groups in the sample. For more than 40 different antibodies, this combination achieved complete elution of primary and secondary antibodies over as many as 21 iterations of staining and elution, the maximum number of cycles tested, while preventing any detectable loss or morphological changes in staining (Fig. 1, B and C, and fig. S1D). Composite images obtained with the same antibodies in a 1st, 11th, and 21st iteration resulted in nearly identical gray-scale images for various types of intracellular structures (Fig. 1D and fig. S1E) and in very high single-cell and single-pixel intensity correlations between staining iterations (Fig. 1, E and F, and fig. S1, F to H). A 2 × 2 pixel smoothing further showed that small fluctuations occasionally seen for some antigens between different rounds of staining mainly involve single-pixel shifts. Thus, our high-throughput automated iterative indirect immunofluorescence imaging approach (which we refer to as 4*i*) can obtain multiplexed signals from areas as large as several mm<sup>2</sup> that are quantitatively reproducible over at least 20 iterations from subsamples as small as 165 nm by 165 nm, which corresponds to the surface area of a single pixel when acquired with a 40× objective and a scientific complementary metal oxide semiconductor (sCMOS) camera.

## 40-plex 4*i* across multiple length scales

We then applied 4*i* on populations of human tissue culture cells (Fig. 2A). Image acquisition covered the full cell height with 18 zplanes, which were computationally integrated by maximum

<sup>1</sup>Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland. <sup>2</sup>Molecular Life Sciences PhD Program, Life Science Zurich Graduate School, University of Zurich, Zurich, Switzerland. <sup>3</sup>MD-PhD and Systems Biology PhD Program, Life Science Zurich Graduate School, University of Zurich, Zurich, Switzerland.

\*These authors contributed equally to this work.

†Corresponding author. Email: gabriele.gut@uzh.ch (G.G.); lucas.pelkmans@imls.uzh.ch (L.P.). ‡Present address: MGH & BWH Center for Clinical Data Science, Boston, MA 02114, USA.

intensity projections, on a total surface area of  $6.25 \text{ mm}^2$ , capturing ~20,000 single cells, for each of 11 different experimental conditions (Fig. 2, B and C, and fig. S2A). Collectively, these images contain multivariate molecular and phenotypic information across several orders of magnitude of spatial scales, providing the possibility of directly studying connections between multiple levels

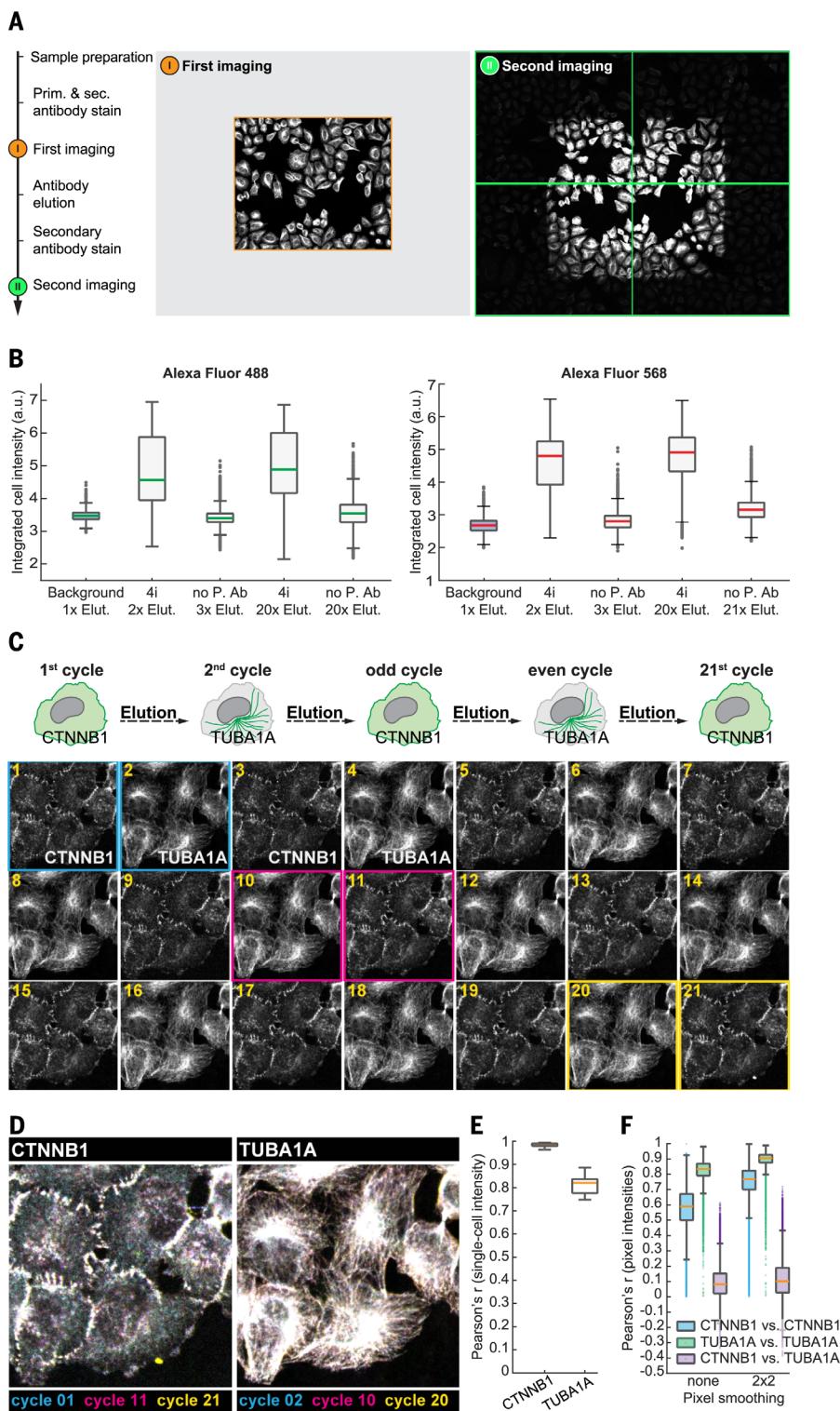
of biological organization. For instance, they allowed visualization of the influence of local cell crowding on the abundance of cytoskeleton-associated proteins, an emergent property at the cell-population level (Fig. 2D, left panel), or of how position in the cell cycle affected the phosphorylated state of a protein, a property at the cellular level (Fig. 2D, right panel). At the same

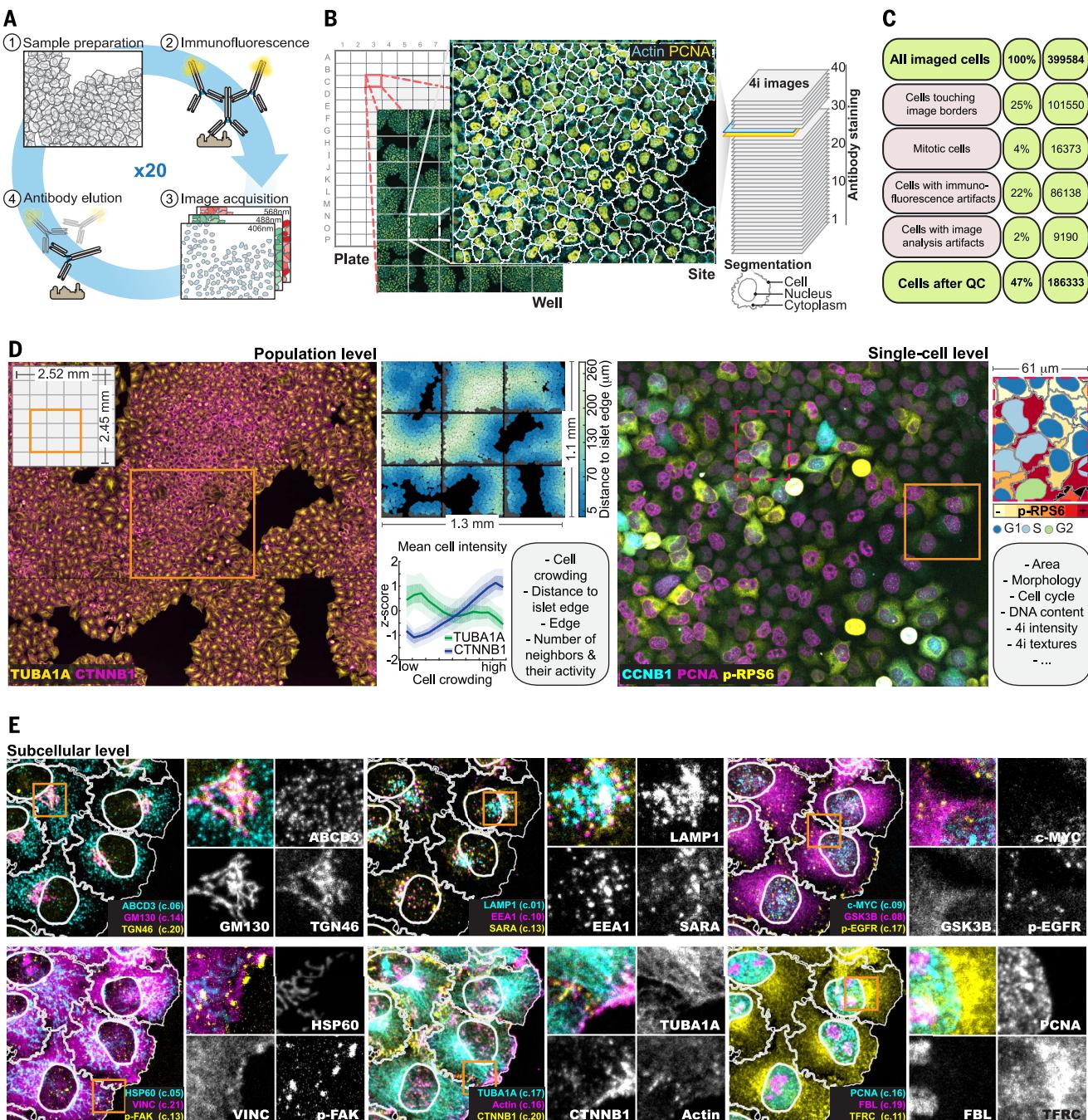
time, it allowed assessment of protein subcompartmentalization, a property at the subcellular level, with enough resolution to capture the detailed morphology of cytoskeletal structures and organelles, such as microtubules, actin ruffles and focal adhesions, multiple types of endosomes, individual tubules of mitochondria, and ribbons of the Golgi complex (Fig. 2E), all in the same single cell and across hundreds of thousands of cells.

### Multiplexed single-cell analysis

By applying a series of image processing and cellular computer vision methods (22–24, 30) (fig. S2, C to E), we generated a 4i image dataset with  $\sim 10^9$  single-cell measurements (fig. S2B). As expected, these measurements were well suited to the types of analysis usually performed with other multiplexed single-cell methods (31–33). Through the ability to combine various types of single-cell measurements apparent at different length scales, 4i could provide insights that are generally not considered in single-cell approaches. For instance, the variance in phenotypic properties emergent at the cell population and cellular scales, such as in cell size or DNA content, is strongly underestimated by clustering single cells based on the mean intensities of their multiplexed molecular stains only (fig. S2, F to J). Analyzing

**Fig. 1. 4i achieves 40-plexed spatially resolved molecular readouts. (A)** IIF against TUBA1A without Imaging Buffer (First imaging, orange area), after which antibodies were eluted, sample was re-stained with secondary antibodies (sec. abs.) only, and a larger area was imaged (Second imaging, green areas). Images from first and second imaging are rescaled independently. **(B)** 4i elutes primary and secondary antibodies repeatedly from the sample and preserves dynamic range. Box plots of integrated cell intensity for multiple primary antibodies detected by AF488- (left) or AF568-labeled (right) sec. abs. Background: Cells incubated solely with secondary antibodies. **(C)** Alternating application of 4i against CTNNB1 and TUBA1A in the same cells for 21 cycles. The cyan-, magenta-, and yellow-boxed images were used to generate composite images in (D). Images from different cycles were rescaled differently. **(D)** Composite images from the boxed images in (C). Images from different cycles were rescaled differently prior to composite image creation. **(E)** Box plots of pairwise single-cell correlations of integrated cell intensity measured for either CTNNB1 or TUBA1A between 21 cycles. Number of cells: 16,000. **(F)** Box plots of pairwise pixel intensity correlations between 21 cycles for CTNNB1 (cyan), TUBA1A (green), and between CTNNB1 and TUBA1A (purple) in images either unsmoothed or smoothed by a  $2 \times 2$  mean filter. Number of cells: 16,000. Box plots indicate population median (central mark), interquartile range (box), 99.3% of population range (whiskers), and outliers (dots).





**Fig. 2.** 4i in high-throughput generates information across multiple spatial scales. **(A)** Schematic representation of the 4i protocol. (1) The sample is blocked to prevent nonspecific antibody binding and blocking of free sulfohydryls by thioether reaction with maleimide. (2) Indirect immunofluorescence (IF) against two antigens is performed. (3) Sample is imaged in Imaging Buffer. (4) Elution Buffer elutes antibodies from the sample, which can subsequently undergo another round of 4i. **(B)** 4i protocol performed in 384-well plates in a high-throughput workflow combined with automated spinning-disk microscopy. Wells were imaged at 40 $\times$  magnification in a 7  $\times$  6 tiled fashion for 21 4i cycles. Image analysis was performed to identify cells and generate segmentation masks of cell, cytoplasm, and nucleus. **(C)** Diagram of cell numbers excluded from analysis through quality control. **(D)** Tiled overview (3  $\times$  3 imaging sites) of a HeLa cell population stained

for TUBA1A (yellow) and CTNNB1 (magenta). Upper right: Segmented cells color-coded for distance to cell islet edge. Lower right: Quantification of differential protein mean concentration as a function of cell crowding. Solid lines, median; shading, interquartile range. Orange square: region used for the single-cell-level panel on the right, depicting stains for CCNB1 (cyan), PCNA (magenta), and p-RPS6 (yellow). Red-dashed square, region visualized on the right, where segmented cytoplasm is color-coded for p-RPS6 mean cell intensity and segmented nuclei for position in the cell cycle. Orange square: region used for (E). **(E)** Visualization of the subcellular distribution of 18 4i stains. Each subpanel consists of two three-channel composites (whole region and higher magnification of orange squares) and three grayscale images of the individual 4i stains at higher magnification. The number next to each staining label indicates their corresponding 4i acquisition cycle.

the covariance in phenotypic properties and molecular readouts could, among others, reveal a differential regulation of cytoskeletal components as a function of local cell crowding, and a discrete stepwise increase in phosphorylated S6 (p-RPS6) as a function of cell cycle phase in the same single cells (fig. S3, A to H), the mechanisms of which are not well understood. Although these are important observations for future studies, we here decided to specifically focus on a scale in the dataset that has been even less studied in conjunction with the other scales—namely, that of multiplexed single-pixel measurements.

### Multiplexed single-pixel analysis framework

To capture the full amount of information that the pixel scale contains within its 16-bit single-pixel intensity measurements, we represented each pixel as a vector of 34 multiplexed 4*i* intensities and applied a two-step clustering approach that identifies, in an unsupervised manner, different pixel types in the dataset, ensuring that rare, but relevant types were preserved (fig. S4, A and B, and table S2). A pixel type represents a subset of single pixels that share a distinct multiplexed intensity profile, which is different from the other identified pixel types. Because the clustering is derived from a large number of different, independently obtained multiplexed signals, they are robust to technical noise in an individual stain. This also reduced the influence that superimposition artifacts introduced by maximum intensity projection of Z stacks have on the clustering of single pixels, particularly from the nuclear region. Although some fluorescent signal of a protein stain, which is localized in the cytoplasm, may be measured in pixels assigned to the nuclear region (as some of the fluorescent signal could be localized under or on top of the nucleus), multiplexed pixel clustering will remain robust, because it is dominated by multiple true nuclear signals whose fractions of total signal in those pixels is much higher. After pixel types were determined, we then assigned each two-dimensional (2D) pixel position in each analyzed cell to its corresponding type, which we term multiplexed cell units (MCUs) (Fig. 3A). The classification into MCUs was unsupervised and based only on information contained within the 34-plex intensity profiles of a very large number of pixels ( $\sim 10^7$ ) sampled from a large number of single cells (ranging from 300 to 2000), which can be phenotypically very different. Nevertheless, MCUs demonstrated highly specific multivariate profiles and extensive intracellular organization (fig. S4, C to H), indicating that they contain generalizable information about protein subcompartmentalization at high spatial detail (Fig. 3A and fig. S4H).

To capture how these MCUs localize relative to each other, we next calculated the extent to which the 2D pixel positions of one MCU borders the 2D pixel positions of another MCU (Fig. 3A) and used this quantification to create a map of MCU proximities. We call such maps multiplexed protein maps (MPMs) (Fig. 3B), where each MCU

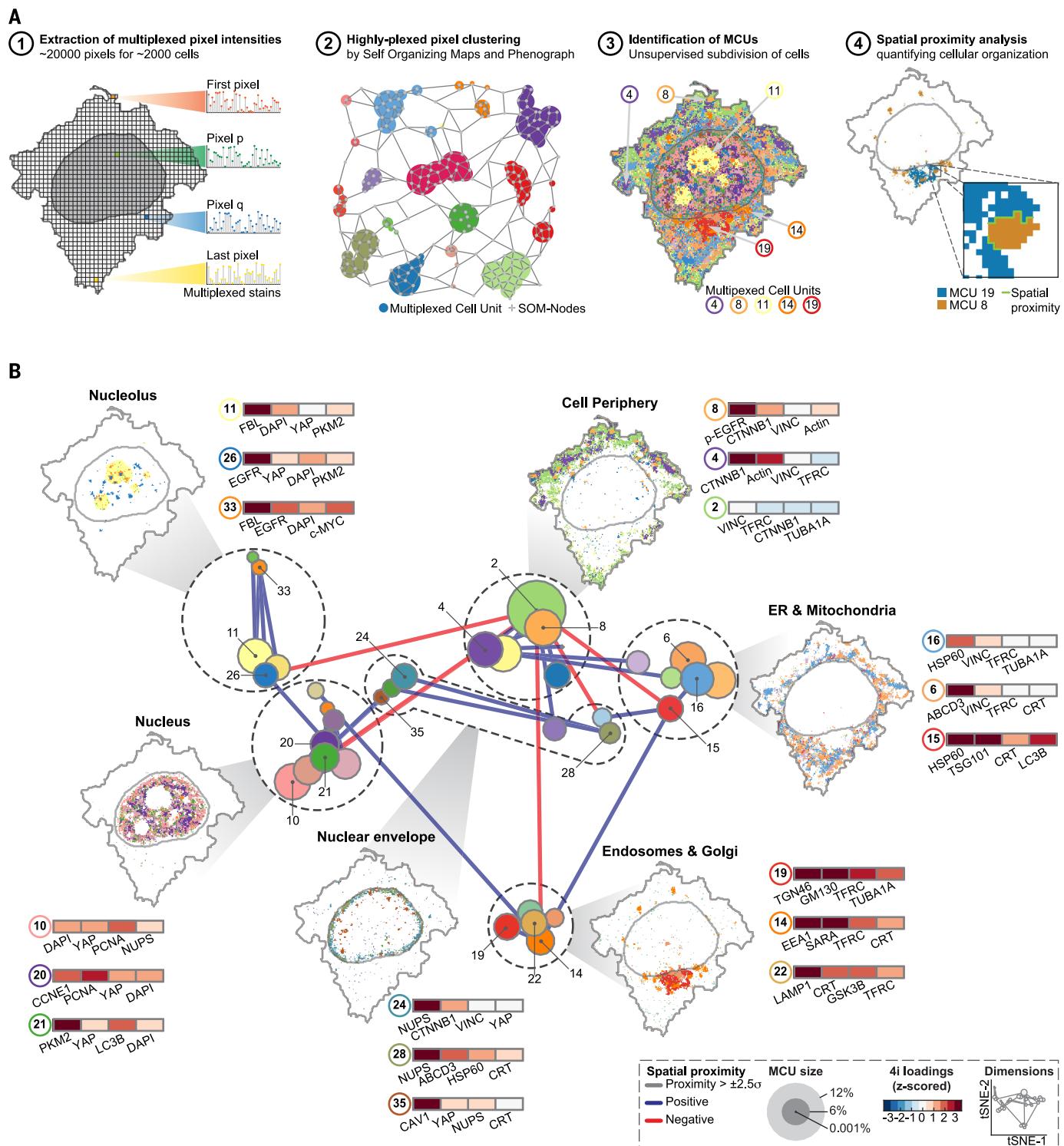
is represented as a node. The size of a node reflects the fraction of pixels that were assigned to that MCU; the distance between nodes reflects their similarity in bordering pixels from other MCUs (fig. S4F); and the node edges reflect direct spatial proximity or avoidance between two MCUs. As illustrated in one particular single cell (fig. S4, H and I), the resulting MPM identified multiple MCUs within the nucleus, based on specific combinations of DNA stain, nuclear cytoskeleton, DNA replication factors, proteins involved in the cell cycle, transcription factors, and ribosomal RNA processing factors (Fig. 3B). It also identified separate MCUs within the inner and outer nuclear envelope (fig. S4G) and multiple MCUs around the microtubule-organizing center (MTOC), including those enriched in specific combinations of Golgi complex and various endosome markers. In addition, it identified multiple MCUs in the extended cytoplasm enriched in various combinations of endoplasmic reticulum (ER), mitochondria, peroxisome, and late endosome markers including microtubules, as well as multiple MCUs in the cell periphery, reflecting specific combinations of markers of focal or cell adhesions, phosphorylated growth factor receptors, and the actin cytoskeleton (Fig. 3B). Thus, we developed a data-driven, unsupervised computer vision approach to comprehensively quantify protein organization with submicrometer spatial detail in hundreds of single cells without being confounded by the specific relative geometry and orientation of an individual cell. This allowed systematic comparisons of subcellular spatial protein distribution between single cells that experience different states, microenvironments, growth conditions, or exposure to drugs.

### Proof-of-principle findings with multiplexed single-pixel analysis

To explore, as a proof of principle, what such analyses could reveal, we compared the abundance of MCUs and their spatial interactions between cells in different phases of the cell cycle (Fig. 4, A and B). As expected, MCUs containing cell cycle markers showed strong differences in abundance between different phases. A cytoplasmic MCU enriched for cyclin B1 (MCU 17) was more abundant in cells in G<sub>2</sub>, whereas a nuclear MCU enriched for cyclin E (MCU 20) was much more prevalent in G<sub>1</sub> cells. The MPM also identified other nuclear MCUs (MCU 9, 10, and 13), which contained specific combinations of c-Myc, phospho-4EBp1, and YAP and which were more abundant in cells in G<sub>2</sub>. Similarly, it identified cytoplasmic MCUs enriched in phospho-S6 (MCU 3 and 29) as being more abundant in cells in G<sub>2</sub>. This reflects the adaptation of signaling activities and their downstream effects on transcription factors to the cell cycle (34–37). The approach also correctly identified the duplication of the centrosome (PCNT) in G<sub>2</sub>, as well as the concomitant increase in abundance of Golgi complex markers (38) (TGN46, GM130) (MCU 27), which are key functional differences between single cells in different positions of the cell cycle that are only represented by MCUs covering a

few pixels. Highlighting all these MCUs in particular single cells (fig. S5, A to D) underscores their accuracy and sensitivity (fig. S5, A to D). When we compared single cells that experience either low or high local cell crowding (Fig. 4, C and D, and fig. S5, E to H), we observed that MCUs containing markers of multiple endocytic organelles in various subcellular locations (MCU 12, 14, 15, 19, 22, and 37) were more abundant, relative to cell size, in cells experiencing high crowding. In contrast, MCUs containing markers of mitochondria (HSP60) and peroxisomes (ABCD3) (MCU 6, 16, 25, 28, and 29) were more abundant, relative to cell size, in cells experiencing low crowding. This is in line with previous observations for lysosomes (23) and mitochondria (39) and suggests the existence of a mechanism that inversely adapts the abundance of organelles involved in catabolism versus biosynthesis to the available space that a single cell has for growth. Thus, our unsupervised data-driven approach accurately and sensitively quantifies changes in cellular subcompartmentalization at high spatial detail and enables the meaningful interpretation of intracellular complexity by integrating multiple small differences in each of the multiplexed measurements.

As a second proof of principle, we created a new MPM from cell populations that were exposed to nine frequently used small-compound inhibitors or different growth conditions. To this end, we randomly selected 200 cells from each of the nine perturbations and 200 control cells to construct an MPM from a total of 2000 single cells (fig. S6, A to C). This MPM revealed extensive and widespread changes in the spatial organization of proteins, both expected and non-expected. For instance, a 3-hour treatment of cells with brefeldin A, an inhibitor of the guanosine triphosphatase Arf1 used to study ER-to-Golgi traffic, resulted in the expected disappearance of the MCU containing markers of the Golgi complex (MCU 19) (Fig. 5A, B and fig. S6D). Unexpectedly, we also observed relocalization of mitochondria away from the MTOC (MCU 11) and toward the periphery (MCU 5), an increase in enlarged peroxisomes in the vicinity of the ER (MCU 26), and an increase in actin-positive structures in the cell periphery (MCU 29), as visualized on single-cell examples (Fig. 5, C to E, and fig. S6E). Treatment of cells with nocodazole (Fig. 5, F and G), which depolymerizes microtubules and thus depletes  $\alpha$ -tubulin staining across many cytoplasmic MCUs (fig. S6, B and F), also resulted in an increase in nuclear MCUs enriched in YAP (MCU 13, 18) (Fig. 5, G and H, and fig. S6H), in MCUs of intermediate filaments (vimentin) (MCU 6, 30) (Fig. 5, G and H, and fig. S6G), and in MCUs of cell adhesion structures containing signaling kinases in the cell periphery (MCU 3, 27, 28, 38) (Fig. 5, H and I, and fig. S6I). Treatment of cells with baflomycin (Fig. 5, J and K), an inhibitor of the vacuolar adenosine triphosphatase that mediates endolysosome acidification, resulted in an expected increase in MCUs containing markers of lysosomes (LAMP1) and autophagosomes (LC3B) (MCU 17, 31) (fig. S6,



**Fig. 3. Multiplexed protein maps of multiplexed cell units.** (A) Schematic of the statistical analysis for the identification of MCUs by two-step clustering of multiplexed pixels and the generation of MPMs (fig. S4A). (1) All multiplexed pixels of cells are extracted and (2) clustered using self-organizing maps (SOMs) and Phenograph. (3) Identified clusters are called MCUs. MCUs represent areas in cells comprising pixels with common profiles of 16-bit multiplexed 4i marker intensities. (4) Pairwise spatial proximity scores (SPSs) between all MCUs are calculated. (B) Graphical representation of an MPM for a population of 300 HeLa cells. MCUs, represented as nodes, are placed

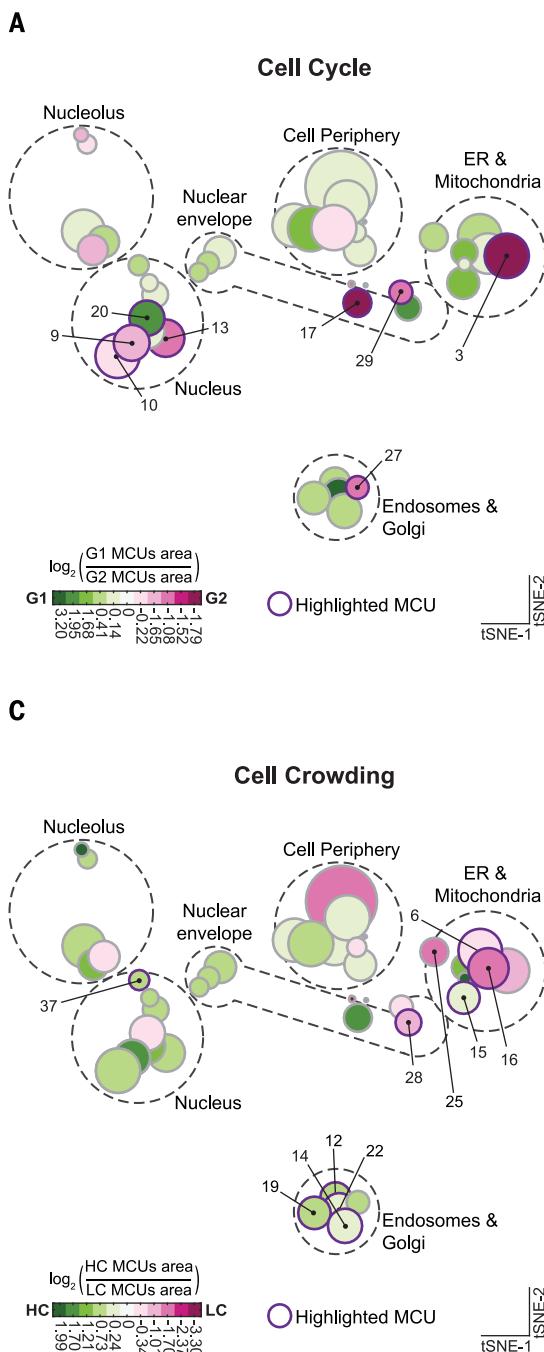
within a 2D plane using t-SNE. Node diameter represents the fraction of pixels assigned to that MCU. Nodes are connected by their pairwise SPS. SPS values  $> 2.5$  standard deviations away from the mean are depicted as edges and colored red when negative and blue when positive. Dashed black circles, spatial proximity domains (SPDs), drawn manually. MCUs within an SPD are projected onto the cell segmentation of a representative cell; the color coding of MCU projections and nodes in the MPM are the same. Heatmaps next to the cell segmentations are z-scored intensity values for the four stains with the highest intensities measured in each of the projected MCUs.

J and K). Unexpectedly, we also observed an increase in a nuclear MCU containing YAP (MCU 18) (Fig. 5K and fig. S6L) and the heterogeneous appearance of a small MCU near the MTOC positive for caveolin-1 and vimentin (MCU 44) (Fig. 5L). Similarly, we found that other well-known inhibitors—such as latrunculin A, which prevents actin polymerization, wortmannin, which inhibits phosphatidylinositol 3-kinase, and rapamycin, which inhibits mechanistic target of rapamycin (mTOR)—as well as differential exposure to growth factors, all had multiple effects on the intracellular organization of cells (fig. S7). Thus, a single MPM analysis reveals,

in an unsupervised fashion, systems-level changes of intracellular protein organization in hundreds of single cells from 10 different treatment conditions enabling a detailed profiling of the phenotypic responses of cells to drugs or different growth conditions.

As a third proof of principle, we asked whether MPMs can provide insights into the heterogeneity that underlies the responses of single cells to perturbations. For each perturbed condition, we clustered the 200 randomly selected single cells based on their profile of MCU sizes (the fraction of pixels of that single cell belonging to an MCU) and compared this to a clustering of single cells

based on a profile of mean single-cell intensities of the multiplexed markers. This comparison revealed that each method identifies different subpopulations and that pairwise comparisons of single cells based on MCUs contain more information for seven out of nine conditions (Fig. 6A and fig. S8A). This result indicates that MPMs provide additional information for the identification of cellular states. To explore the functional relevance of states identified by MPMs, we focused on cells exposed for 3 hours to epidermal growth factor (EGF). Among these cells, the MCU-based clustering identified seven different subpopulations (fig. S8, B and C), which showed



**Fig. 4. MPMs reveal subcellular reorganization by cellular state and microenvironment.**

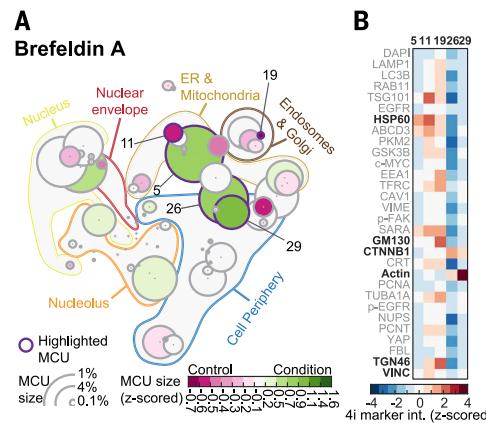
(A) MPM in which MCUs are color-coded for being more abundant in cells in G<sub>1</sub> versus cells in G<sub>2</sub> phase of the cell cycle. Green: MCU more abundant in G<sub>1</sub> cells. Magenta: MCU more abundant in G<sub>2</sub> cells. Nodes of special interest are numbered as in Fig. 3B and outlined in purple. (B) Heatmaps of z-scored intensity loadings of 4i stains of highlighted MCUs in (A). First heatmap, MCUs associated with cell cycle markers (17 and 20), loaded with CCNB1 and CCNE1, respectively (yellow boxes). Second heatmap, nuclear MCUs (9, 10, and 13) loaded with growth signaling-related transcription factors c-MYC, p-4EBP1, and YAP (light green boxes). Third heat map, cytoplasmic MCUs associated with growth signaling (3 and 29), loaded by p-RPS6 (light green boxes). Fourth heatmap, MCU 27, which demarcates centrosomes stained by PCNT and contains Golgi markers (GM130, TGN46) (magenta boxes). (C) MPMs in which MCUs are color-coded based on their abundance in cells experiencing high (HC, green) versus low (LC, magenta) crowding. Nodes of special interest are numbered as in Fig. 3B and outlined in purple. (D) Heat maps of z-scored 4i intensity loadings of 4i stains of highlighted MCUs in (C). First heatmap, MCUs associated with endomembrane system (12, 14, 15, 19, 22, 37, dark green boxes), which are more abundant in cells at HC. Second heatmap, MCUs associated with mitochondria and peroxisomes (6, 16, 25, and 28, pink boxes), which are more abundant in cells at LC.

strong differences in the abundance of various MCUs containing markers for early endosomes (EEA1) (MCU 19, 20, 34), cell adhesion structures such as phosphorylated focal adhesion kinase (pFAK) (MCU 3, 28, 38) and  $\beta$ -catenin (CTNNB1) (MCU 27), autophagosomes (LC3B) (MCU 31), and late endosomes and lysosomes (LAMP1) (MCU 32, 34, 39, 40) (Fig. 6, B and C). To il-

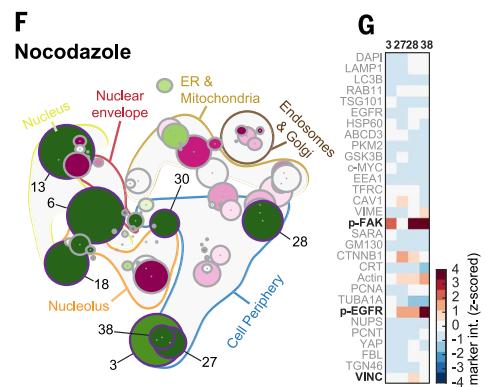
lustrate these differences and appreciate their subcellular organization, we depicted these MCUs in selected single cells from each of the seven subpopulations (Fig. 6B). When we quantified the MCUs in which epidermal growth factor receptor (EGFR) concentrated upon addition of EGF, we observed markedly different responses in the seven subpopulations (Fig. 6D). In one of

these (subpopulation 1), EGFR remained in a non-organellar MCU (MCU 2), similar to where EGFR was located in untreated cells (Fig. 6D), and accumulated in serum-starved cells (fig. S8E). However, in other subpopulations (2 and 3), EGFR concentrated largely in early endosomes (MCU 20); in both early and late endosomes (MCU 20, 32) (subpopulations 4 to 6); or primarily

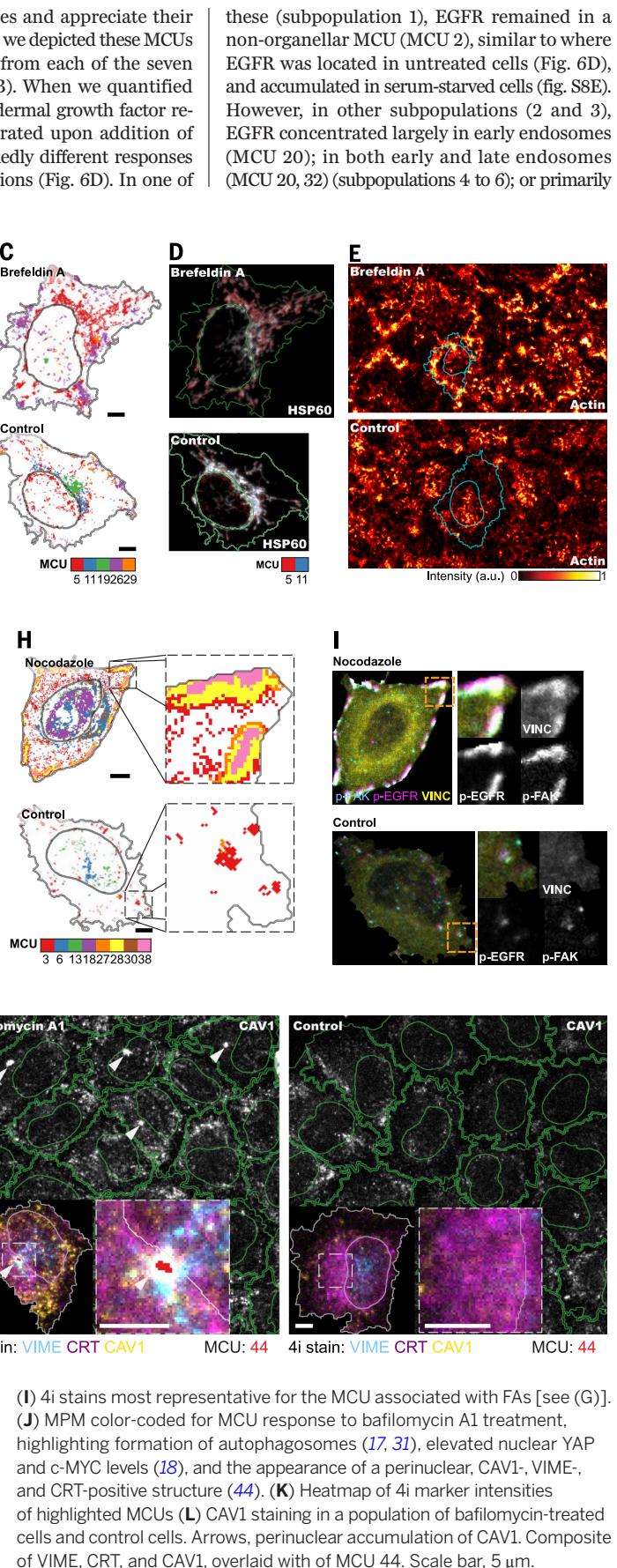
**Fig. 5. MPMs reveal subcellular reorganization upon pharmacological perturbations.** (A) MPM generated as in Fig. 3, but from a total of 2000 cells, by pooling 200 cells each from 10 different conditions. Node color: z-score—normalized MCU abundance in cells from a specific condition (here brefeldin A drug treatment) relative to control cells. Green: MCU more abundant in treatment condition. Magenta: MCU more abundant in control. MCUs of special interest are numbered and outlined in purple.



(B) Heatmap of z-scored 4i marker intensities for MCUs associated with mitochondria (5, 11), Golgi apparatus (19), and the actin cytoskeleton (26, 29). (C) Spatial distribution of highlighted MCUs in BFA-treated cells compared to control cells. Projections of highlighted MCUs from (A) onto representative cell segmentation of a BFA-treated cell and a control cell. Scale bar, 5  $\mu$ m.



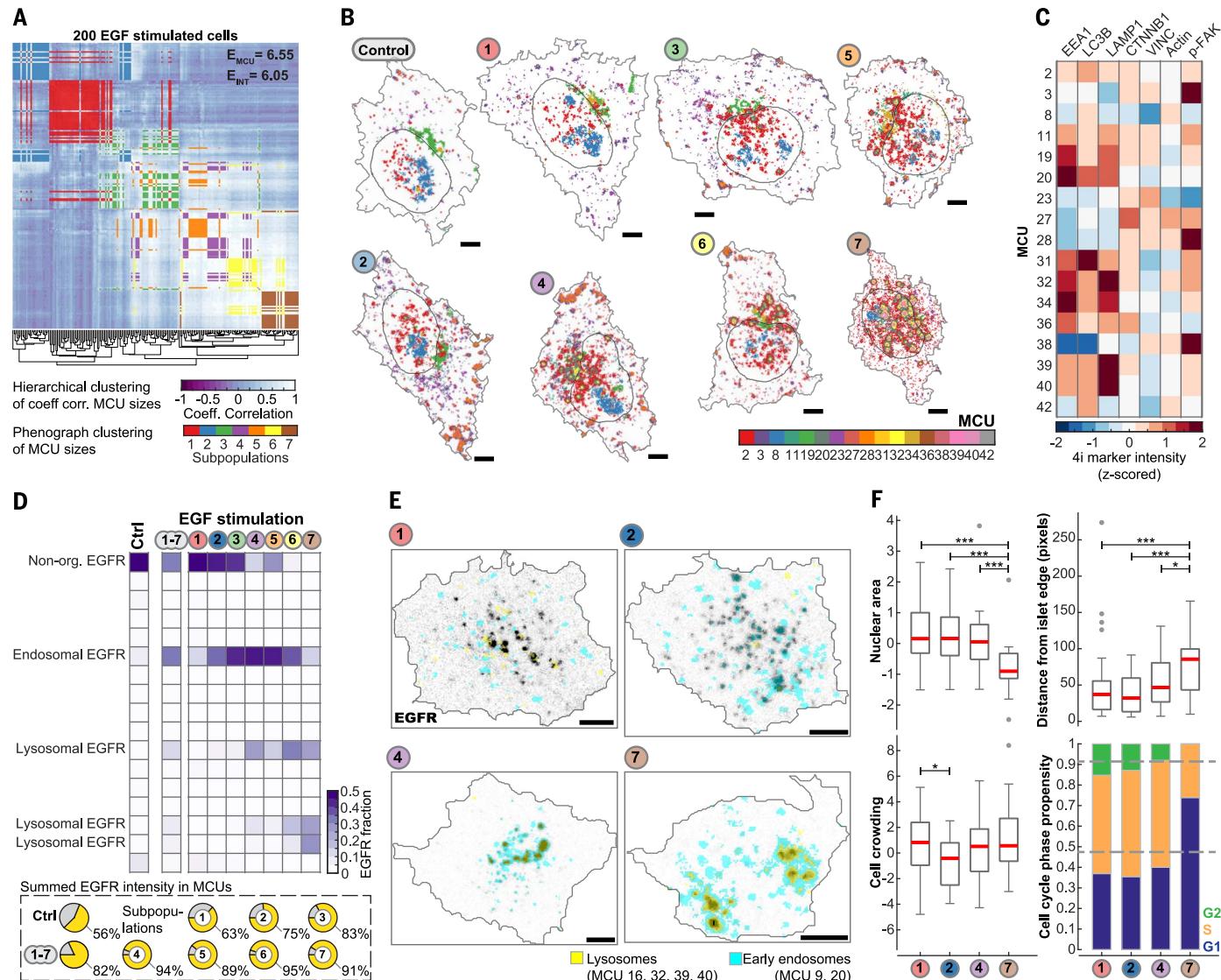
(D) Projection of MCU 5 and 11 onto HSP60 staining. MCU 5, densely packed perinuclear mitochondria; MCU 11, reticulated mitochondria. (E) Image of actin in BFA-treated and control cells. (F) MPM color-coded for MCU response to nocodazole treatment, highlighting enlarged focal adhesions (FA) with increased p-EGFR signal, increased nuclear YAP and c-MYC levels, and fortified intermediary filaments. (G) Heatmap of 4i marker intensities for MCUs marking FAs. (H) Spatial distribution of highlighted MCUs in nocodazole-treated and control cells. Scale bar, 5  $\mu$ m.



in lysosomes (MCU 32, 39, 40) (subpopulation 7) (Fig. 6, D and E). Although the endocytic trafficking of EGFR to lysosomes is well understood (40–42), this result shows that the subcellular fate of EGFR upon stimulation by EGF can be highly heterogeneous in a population of genetically identical cells and that this can be quantified in an

unsupervised high-throughput manner using a 4i. Having captured this observation within a 4i dataset allowed us to relate these different subcellular fates of EGFR to molecular, phenotypic, or microenvironmental properties apparent at higher scales. Analyzing such relationships revealed that cells from subpopulation 2, which

do not relocalize their EGFR to lysosomes, displayed sustained phospho-EGFR specifically in MCUs corresponding to actively signaling focal adhesion structures (MCU 3, 27, 28, 38) (Fig. 6C and fig. S8D). These cells also had a larger surface area and experienced particularly low local cell crowding (Fig. 6F), possibly contributing to



**Fig. 6. MCU-based quantification of heterogeneous subcellular relocalization of EGFR.** (A) Two hundred EGF-stimulated cells hierarchically clustered by their pairwise coefficients of correlation of MCU sizes (white-to-purple-scale image, dendrogram), overlaid by Phenograph clustering of MCU sizes (colors). Top right corner, entropy measured in MCU size- and mean cell intensity-based dissimilarity matrix. (B) MCUs most characteristic of each of the seven subpopulations of cells were projected onto the cell and nucleus outlines of a representative cell from a nontreated control and from each of the seven EGF-treated subpopulations (1 to 7). Scale bar, 5  $\mu$ m. (C) Heatmap of 4i marker intensities (z-scored) of MCUs characteristic for the seven subpopulations. MCU 2: a non-organelle-specific region in cells. MCU 3, 27, 28, 38: focal adhesions with strong loadings of p-FAK, CTNNB1, VINC, and actin. MCU 19, 20: early endosomes with strong loadings of EEA1. MCU 32, 34: endosomes with both EEA1 and LAMP1. MCU 39,

40: lysosomes with strong loadings of LAMP1. (D) Heatmaps of the relative fraction of EGFR intensity within selected MCUs [MCU order and IDs as in (C)] in control cells (Ctrl), averaged over all subpopulations (1 to 7), and in each subpopulation (1, 2, 3...7). Pie charts represent fraction of EGFR intensity in selected MCUs compared to all MCUs. (E) Cells from subpopulation 1, 4, and 7 stained for EGFR (black) and overlaid with spatial projections of early endosome- (cyan) and lysosome- (yellow) associated MCUs. Scale bar, 5  $\mu$ m. (F) Box plots of nuclear area, distance from cell islet edge, cell crowding, and cell cycle phase distribution of cells belonging to subpopulations 1, 2, 4, and 7. Dashed gray lines: cell cycle phase distribution in control cells. Significant differences measured by pairwise two-sided Kolmogorov-Smirnov tests. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . Box plots: population median (central mark), interquartile range (box), 99.3% of population range (whiskers), and outliers (dots).

the emergence of cell crowding-dependent social cellular behavior (43). In contrast, cells that relocalized their EGFR to lysosomes (subpopulation 7) showed less phosphorylated EGFR in the cell periphery but more on endosomes (MCU 20) (fig. S8D). These cells were smaller, further away from cell colony edges, and predominantly in G<sub>1</sub> phase of the cell cycle (Fig. 6F). Using mean single-cell intensities would not have identified these functional differences (fig. S8F). Thus, cellular states specifically defined by the subcellular organization of their proteins are functionally relevant and can be connected to multiple molecular and phenotypic properties apparent at the cellular and cell population scales (fig. S8G).

## Discussion

We have developed 4i, a simple and robust protocol for multiplexed indirect immunofluorescence, which can detect at least 40 protein states in the same cell. Its uncomplicated design makes it compatible with high-throughput setups for the simultaneous study of multiple conditions. Because of its high reproducibility at the subcellular level, we could develop an analysis workflow, which quantifies subcellular organization at high spatial detail in thousands of single cells based on multiplexed single-pixel profiles. 4i builds on a well-established high-throughput multivariate imaging platform (22–24) combined with automated liquid handling that applies the proven principle (13, 15, 18) of iterative staining and signal removal. It uses off-the-shelf antibodies without the need for special conjugations, resulting in high signal yield due to the use of bright fluorophores and signal amplification by a secondary antibody. By preventing photocrosslinking during imaging, the approach enabled complete removal of both primary and secondary antibodies with a mild elution buffer, while preserving the sample even at the smallest spatial scale across a large number of cycles. This comes with the added advantage that epitope masking upon the detection of multiple antigens in close proximity is precluded. Clearly, because 4i is based on conventional indirect immunofluorescence, it shares the same limitations, such as possible artifacts arising from cell fixation and permeabilization, or compromised antibody specificity. In addition, although already greatly minimized, some remaining artifacts due to iterative sample handling and staining cannot be avoided. In the future, new fixation and sample embedding methods, combined with complete automation, may even further reduce these limitations and speed up duration of the 4i protocol.

To make full use of 4i data, it was important to develop a computational strategy that extracts, in an unsupervised manner, quantitative information from various length scales. In particular, the unsupervised subdivision of cells into meaningful, subcellular regions using the full 16-bit information from each of the 40 multiplexed channels of millions of single pixels provides a handle on subcellular complexity that the human eye and brain cannot process. It allows for a

comprehensive quantitative analysis and comparison of cellular subcompartmentalization without being confounded by the specific relative geometry and orientation of each cell. This computational analysis can now be combined with computer vision approaches that specifically quantify geometric properties of cells and subcellular objects, for which there are readily available tools (44, 45), to study the interplay between subcellular protein composition and cell and organelle orientation and shape. This might be especially powerful in the study of cellular processes with a well-described list of molecular markers, such as vesicle trafficking. Although here we restricted our analysis to 2D projections of Z stacks of single tissue culture cells, our approach could also be applied to voxels, which should further reduce possible superimposition artifacts and will allow for the multiplexed characterization of more advanced 3D tissue culture systems such as organoids and organotypic slice cultures. In addition, 4i and multiplexed pixel profile analysis could be used in translational applications, which could lead to the development of multivariate intracellular biomarkers to characterize patient-derived samples. Moreover, the ability to obtain quantitatively reproducible single-pixel measurements promises that 4i will be applicable in even higher-resolution light microscopy, including super-resolution approaches, as previously suggested (46). The ability to bridge biological length scales is one of the major challenges in the life sciences. Usually, extrapolation or inference is applied. However, to predict how properties at a higher scale emerge from multiple interactions occurring at a lower scale, and how these feed back on each other, it is necessary to cover multiple length scales within one measured dataset. Such datasets contain a richness of connections between scales that current models of biological processes do not yet consider. However, it is exactly these connections, by which gene expression is adapted to the cellular state, that determine how a cell type is specified, how a pathological cellular phenotype emerges, or how a tumor cell responds to a drug.

## Materials and methods

### Materials

#### Cell line

HeLa Kyoto (human cervical epithelial cell line, Prof. J. Ellenberg laboratory, EMBL, Germany). Cells were tested for identity by karyotyping (47) and tested for absence of mycoplasma before use.

#### Complete medium(CM)

CM consists of 10% Fetal Bovine Serum (FBS), and 5% Glutamine in DMEM. DMEM (Life-technologies), Fetal Bovine Serum (Sigma Aldrich), Glutamine (Life-technologies).

#### Pharmacological perturbations

Epidermal Growth Factor (EGF) (Milipore), Nocodazole (Sigma Aldrich), Latrunculin A (Sigma Aldrich), Bafilomycin A1 (Sigma Aldrich), Brefeldin A (Sigma Aldrich), Wortmannin (Sigma Aldrich), Rapamycin (Sigma Aldrich).

### 4i blocking solution (sBS)

sBS consists of 1% Bovine Serum Albumine (BSA), and 150mM Maleimide in phosphate buffered saline (PBS). Maleimide is added to aqueous solution just before Blocking step in 4i protocol. BSA (Sigma Aldrich), Maleimide (Sigma Aldrich)

### Conventional blocking solution (cBS)

cBS consists of 1% Bovine Serum Albumine (BSA) in phosphate buffered saline (PBS). BSA (Sigma Aldrich)

### Imaging buffer (IB)

IB consists of 700mM N-Acetyl-Cysteine (NAC) in ddH<sub>2</sub>O. Adjust to pH 7.4. NAC (Sigma Aldrich)

### Elution buffer (EB)

EB consists of 0.5M L-Glycine, 3M Urea, 3M Guanidinium chloride (GC), and 70mM TCEP-HCl (TCEP) in ddH<sub>2</sub>O. Adjust to pH 2.5. L-Glycine (Sigma-Aldrich), Urea (Sigma-Aldrich), GC (Sigma-Aldrich), TCEP (Sigma-Aldrich).

### Primary antibodies

Antibodies were selected based on the following criteria: 1. Successful use of antibody in immunofluorescence has been published in the past in scientific literature. 2. Antibody is raised against epitopes on bona fide markers of organelles. 3. To ensure same number of antibodies raised in both mouse and rabbit. While testing antibodies for this study, two (antibody against epitope on TOM20 (Abcam ab56783) and CAT (Abcam ab110292)) out of more than 50 antibodies were identified to not work with the 4i protocol. Information on antibodies used in this study is in found in Supplementary Table S1.

### Secondary antibodies

Anti-mouse AlexaFluor-488 was diluted 1:600 and anti-rabbit AlexaFluor-568 was diluted 1:300 in cBS respectively. Anti-mouse AlexaFluor-488 (Life-technologies), anti-rabbit AlexaFluor-568 (Life-technologies)

### DNA stain solution (DSS)

4', 6-Diamidino-2-phenylindole (DAPI) diluted 1:250 to 1:50 in PBS. DAPI concentration was increased with increasing numbers of elutions to compensate for signal lost due to depurination of DNA, and the resulting reduced binding affinity of DAPI to DNA.DAPI (Life-technologies)

### Computational infrastructure

Image analysis steps were performed on the high-performance cluster computer Brutus at ETH Zürich. Extraction of multiplexed pixel profiles, as well as their clustering using Self-Organizing Map algorithms were performed on Science Cloud UZH. All other described computational methods were executed on a desktop computer.

### Methods

#### Cell Culture

Cells were cultured in Complete Medium at 37°C, 95% Humidity and 5% CO<sub>2</sub>, 750 cells per

well were seeded in a 384-well plate (Greiner) and were grown for 3 days in the above-mentioned conditions.

#### Pharmacological and metabolic perturbations

All compounds were diluted in to their respective final concentration using Complete Medium, except for EGF, which was diluted in DMEM only. Cells were incubated for 3 hours with one of the following compounds: Nocodazole (NOC) 500 ng/ml, Latrunculin A (LATA) 0.2 mM, Bafilomycin A1 (BAF) 100 nM, Brefeldin A (BRF) 2.5 mg/ml, Wormannin (WRT) 1mM, Rapamycin (RPA) 0.5mM. Metabolic perturbations were performed as following. Overnight growth factor starvation in Opti-Mem (GFS), GFS followed by 3 hours of EGF stimulation with 100 ng/ml of EGF (S + EGF), and 3 hours EGF stimulation with 100 ng/ml of EGF (EGF).

#### Microscopy

An automated spinning disk microscope from Yokogawa (CellVoyager 7000) with an enhanced CSU-W1 spinning disk (Microlens-enhanced dual Nipkow disk confocal scanner, wide view type) was used in combination with a 40× Olympus objective of 0.95 NA, and Neo sCMOS cameras (Andor, 2,560 × 2,160 pixels) to acquire microscopy images. 18 z-planes with a 500 nm z-spacing were acquired per site and a maximum intensity projection was computed and used for subsequent image analysis. UV (406 nm), green (488 nm) and red (568 nm) signals were acquired sequentially.

#### Iterative Indirect Immunofluorescence Imaging (4i)

Sample preparation was performed as following. Cells were fixed in 4% Paraformaldehyde (Electron Microscopy Sciences) for 30min. Cells were then permeabilized with 0.5% Triton X-100 for 15 min. Cells were washed 6 times with PBS both before and after permeabilization. Fixation and permeabilization were performed at room temperature. Each of the subsequent steps was performed in sequence of their mentioning and in every cycle of 4i. If not stated differently, all steps were performed at room temperature. (1) Antibody Elution. Sample was washed 6 times with ddH<sub>2</sub>O. Residual ddH<sub>2</sub>O was aspirated to minimal volume. Subsequent actions are repeated 3 times: EB was added to sample and shaken at 100 rpm for 10 min. Then EB was aspirated to minimal volume possible. (2) Blocking. SBS was added to sample and shaken at 100 revolutions per minutes (rpm) for 1 hour. After 1h sample was washed 6 times with PBS. (3) Indirect immunofluorescence, primary antibody stain. Primary antibody solution was added to sample and shaken at 100 rpm for 2 hours. After 2 hours, the sample was washed 6 times with PBS. (4) Indirect immunofluorescence, secondary antibody stain. Secondary antibody solution was added to sample and shaken at 100 rpm for 2 hours. After 2 hours, the sample was washed 6 times with PBS. (5) Nuclear staining. DSS was added to

sample and shaken at 100 rpm for 10 min. After 10min sample was washed 6 times with ddH<sub>2</sub>O. Residual ddH<sub>2</sub>O was aspirated to minimal volume. (6) Imaging. IB was added to sample and sample was imaged. Perform step 1 to 6 until required plexity is achieved. All liquid dispensing and washing steps of the 4i protocol were performed using a Washer Dispenser ELA406 (BioTek). Primary and secondary antibodies were dispensed using a Bravo liquid handling platform from Agilent Technologies.

#### Computation of single-cell pixel correlations

Pixel correlations were calculated between two 4i signals. If the signals were not recorded during the same acquisition, image alignment was performed prior to the correlation measurement. First, the same background value was subtracted from both images. Next, single pixel intensities of the two different 4i signals originating from the same cell were correlated in the segmented areas (Cell, Cytoplasm, Nucleus). This was done for every cell individually. Pixel correlations were calculated either with unsmoothed images or on images smoothed by either a 2x2, 3x3, 5x5, 7x7, or 10x10 pixel mean filter. Pixel correlations in Fig. 1F and fig. S1F were calculated as following. Pixel intensities of the first CTNNB1 staining (cycle 1) were correlated with each other CTNNB1 stain (odd cycles) up to the 21<sup>st</sup> cycle. Pixel intensities of the first TUBA1A staining (cycle 2) were correlated with each other TUBA1A stain (even cycles) up to the 20th cycle (green box plots). Pixel intensity correlations between CTNNB1 and TUBA1A were calculated between the first CTNNB1 stain (cycle 1) and each TUBA1A staining (cycle 2, 4, 6, 8, 10, 12, 14, 16, 18, 20). The calculated correlations over all cycles for each stain were aggregated in one box plot each and calculated after the corresponding images were smoothed by a mean filter of increasing size (none, 2x2, 3x3, 5x5, 7x7, and 10x10 pixel) pixels correlations of all cycles. Custom CP modules used for the calculation of single-cell pixel correlations between 4i signals (MeasureCorrelationInTrans, LoadImagesInTrans, LoadSegmentationInTrans, AlignTransImages).

#### Image alignment of acquisition from different 4i cycles

Microscopy images of different cycles from the same site require image alignment, as slight shifts in X and Y occur in between acquisitions due to imperfect stage repositioning. Image registration based on Fast Fourier Transform (48) was performed on DAPI images of two cycles. 488nm and 568nm acquisition, and segmentations masks were shifted by the calculated offset, resulting in aligned microscopy sites. Image alignment, and image and segmentation mask shifting were performed using custom CellProfiler.

#### Image analysis and feature extraction

CellProfiler 1 (49) (CP) was used to perform image analysis and feature extraction using both custom and standard CP modules on the image analysis platform IBRAIN (50). Every image was

corrected for illumination biases prior to image analysis. Nuclei were segmented with the CP Module IdentifyPrimaryIterative using images of DAPI signal. Cellular segmentations were achieved by combining signals of GSK3B (cycle 08), CTNNB1 (cycle 15) and YAP (cycle 19) in a ratio of 2:1:2 for the same site. Subsequently the resulting images were analyzed by a watershed algorithm with varying thresholds to detect cell outlines (IdentifySecondaryIterative). Single-cell features of signal intensity and texture, and area shape of objects were then extracted for nuclei, cytoplasm and cells using standard CP modules (MeasureObjectIntensity, MeasureTexture, and MeasureObjectAreaShape).

#### Data normalization

Cell and pixel data was correct for inhomogeneities in signal intensity due to plate position and experimental handling as following: Mean and standard deviation of pixel intensities for each 4i marker of 2000 randomly selected unperturbed cells originating from four plate rows (see fig. S2A) were calculated per individual row. Subsequently, every 4i pixel value was z-scored using the mean and standard deviation corresponding to the pixel's origin from the plate. Cell measurements were normalized as follows: In a first step, measured intensities of unperturbed and DMSO treated cells were combined. Next, mean and standard deviation of each 4i marker stain intensity was calculated per row (see fig. S2A, each row contains 2 wells of unperturbed cells and 2 wells of DMSO treated cells). Intensity values for all wells in a row were then z-scored using the corresponding mean and standard deviation. In a second step, mean and standard deviation for each condition were calculated. Next, every well was z-scored independently and data from each well was z-scored independently, then z-scored values of each wells were multiplied with the condition specific standard deviation and the condition specific mean was added (inverse operation of z-scoring). In a third step, the newly normalized intensity values were z-scored with the mean and standard deviation measured in cells from the control condition (DMSO treated cells).

#### Cell cycle phase classification: G<sub>1</sub>, S, G<sub>2</sub>, and M phase

Cells were classified into their respective cell cycle phase using Support Vector Machine (SVM) Classification, k-means clustering. First, a classifier was trained to identify M phase cells based on intensity and texture features of nuclear DAPI, as well as features of nuclear area. Then a classifier was trained to identify S phase cells based on intensity and texture features of nuclear PCNA staining (cycle 16). CellClassifier (51) was used to train both SVM classifiers. Cells classified as M and S phase cells were then excluded from the nuclear DAPI distribution, which was derived from calculating the mean of nuclear DAPI integrated intensity of single cells from cycle 2 to 8. The resulting histogram consisted of two separated Gaussians representing G<sub>1</sub> and G<sub>2</sub> cells.

K-means clustering with 2 groups was used to assign single cells into either one of the groups (G1 or G2 cells), where G1 was defined as the group with its median nuclear DAPI total intensity half the value of G2.

#### 4i readout SOM clustering

Mean cell intensities measured for 40 4i markers in cells originating from 10 different conditions (see fig. S2A, DMSO control, GFS, S + EGF, EGF, NOC, LATA, BAF, BRF, WRT, RPA) were used for the analysis. Clustering of the high-dimensional dataset (~165000 cells \* 40 4i intensities) was performed by FlowSOM (52), an R implementation of Self-Organizing Maps, using 400 nodes and Euclidean distance as distance metric.

#### Cellular state SOM

Cellular state SOM was generated by FlowSOM using 37 cellular state features of 154714 cells from 9 perturbation condition and control condition. The used features describe DNA content, cell cycle position (cell cycle markers such as PCNA, CCNE, CCNBI), Cellular, cytoplasmic and nuclear size and shape, as well as a cells distance from the population edge, its cell crowding and whether it is located at the edge of a population. Features were z-scored using the mean and the standard deviation measured in control cells prior to clustering using FlowSOM. Cell data was clustered using 400 nodes, Euclidean as distance metric and 20 runs.

#### Partial correlation analysis

Partial correlations between cell mean intensities of 38 4i stainings were calculated as follows. 40% of cells assigned to each node of the cellular state SOM were selected randomly. Subsequently, the mean of the 38 4i stainings was calculated for each SOM node. Partial correlation analysis (Matlab function partialcorr) was then performed with the resulting 400 means (one per node) and stored. Random subsampling of 400 cells and subsequent partial correlation analysis was performed 1600 times. The mean partial correlation between each node was calculated by calculating the mean of the 1600 bootstrapped partial correlations.

#### Generation of multiplexed cell units

The aim of this analysis is to generate Multiplexed Cell Units (MCUs) from multiplexed pixel profiles (MPP). An MPP corresponds to a one-dimensional vector which contains the intensity value for every 4i staining measured at a 2D pixel position. In step 1, images of the same cells but from different 4i cycles were computationally aligned (see above, table S2, 1. Image acquisition and analysis). In step 2, MPP are generated for each 2D pixel position within a cell's segmentation. All MPPs are then concatenated into a large matrix of m \* n, where m represents MPPs and n represents the 4i intensities measured. Step 2 was repeated for each cell included in the analysis, MPP profiles of each cell are concatenated along the m dimension to create a MPP Matrix. Thus, MPP Matrix 2D matrix contains all MPPs of all

2D positions within the cell segmentations of all cells included in the analysis. Each row represents a pixel, each column a 4i stain (table S2, 2. Generation of Multiplexed Pixel Profile Matrix). Step 3, background subtraction was performed on the intensity values in the MPP Matrix (camera chip background values is 110), followed by an intensity. All pixel of each 4i staining were rescaled between 0 and 1, where 0 is background value and 1 is 98th intensity quantile measured for each 4i staining respectively. When MPP of different conditions were compared to control, rescaling was based on the 98<sup>th</sup> pixel intensity quantile measured from control cells. MPP exclusively containing rescaled intensity values below 0.33 were excluded from further analysis (table S2, 3. Process MPP Matrix). In step 4, the rescaled MPPs were clustered by a clustering algorithm called Self-Organizing Maps (SOM, R implementation FlowSOM), using Euclidean distance as a distance metric, and 20 runs (table S2, 4. First clustering step, SOM nodes). SOMs are already widely used in the analysis of multiplexed data derived from Flow cytometry and Mass cytometry. The 2 variables are the product of SOM clustering. One is a matrix which contains SOM node positions. Each row of the matrix represents a SOM node and each column its positions in multivariate 4i intensity space. SOM nodes represent groups of MPP with similar 4i intensity profiles. The other is a vector which contains SOM node assignments of every MPP. Each row of the vector represents one MPP and its assignment to the SOM node closest to it in multidimensional space. Step 5, The median multiplexed intensities of cells assigned to each SOM nodes were further clustered by Phenograph. Neighborhood value for clustering was selected as the inflection point of the function of cluster number at a given neighborhood value (e.g., fig. S4B) and 'jaccard' as graph type. The resulting clusters from the Phenograph analysis were called Multiplexed Cell Units (MCUs) (table S2, 5. Second clustering, MCU identification). This two-step clustering approach obtains the assignment of every pixel in a cell to an MCU. They represent areas in cells comprising pixels with common profiles of multiplexed 4i marker intensities. For Fig. 3, 300 unperturbed cells were selected randomly and analyzed using the aforementioned clustering approach and their MPP were clustered by FlowSOM with 2500 nodes. For Figs. 5 and 6, 200 cells from each of the ten different conditions (DMSO, GFS + S, GFS, Stimulation, Nocodazole, Latrunculin, Bafilomycin, Brefeldin, Wortmannin, Rapamycin) were randomly selected. Their MPP were combined and clustered with FlowSOM using 3600 nodes.

#### Calculation of MCU size and spatial proximity score

MCU sizes per cell were calculated by dividing the number pixels assigned to each MCU per cell by the total number of pixels assigned to any MCU in the cell. Spatial Proximity Score (SPS) was calculated for each cell individually and for each MCU sequentially. (1) MCU of in-

terest (see Section Generation of Multiplexed Cell Units) was projected back on to the cell segmentation (e.g., MCU $\alpha$ , consisting of 2 adjacent pixels). (2) Then, 8-connectivity morphological expansion was performed on MCU (e.g., MCU $\alpha$  is expanded by one pixel in all directions, MCU $\alpha$  consists now of 10 pixels, 2 original and 8 new pixels). (3) Pixels, newly occupied by the expanded MCU were analyzed for their original MCU assignment and their occurrence was divided by the number of newly occupied pixels (e.g., 2 out of newly occupied pixels belong to MCU $\beta$  (0.25), 4 to MCU $\gamma$  (0.5), 2 to MCU $\delta$  (0.25), 0 to MCU $\epsilon$  (0)). (4) In order to control for MCU size the positions of all MCUs were randomized, and steps 1-6 were repeated 300 times. The mean of all frequencies measured in the 300 randomizations was calculated and subtracted from the observed frequencies (e.g., randomization mean for MCU $\beta$  0.1, thus SPS is 0.15, etc.) (Supplementary Table S2, 6. MCU proximity and size measurements). The resulting values represent the SPS for a given MCU with all others. SPS reach from -1 to +1, where positive SPSs reflect a spatial proximity between two MCUs and negative SPSs reflect a spatial avoidance between two MCUs. SPS of zero show no trend in proximity compared to randomization control. Pixels which did not neighbor any other pixel, assigned to the same MCU, were excluded from MCU size and SPS analysis.

#### MCU size in different microenvironments and cell cycle phase

Cells were binned into 3 cell crowding bins—dense, intermediate and sparse—based on their local cell crowding value (LCC). Cells with crowding values between 85<sup>th</sup> and the 99.9<sup>th</sup> were classed as high LCC, cells with values between 35<sup>th</sup> and 65<sup>th</sup> percentile were classed as intermediate LCC, and cells with crowding values between 0.1 and 15<sup>th</sup> percentile were classed as low LCC. Cell cycle phase classification was performed as described above. This results in 9 categories. Dense G<sub>1</sub>, S dense, G<sub>2</sub> dense, intermediate G<sub>1</sub>, etc. MCU sizes per individual cell cycle phase were then computed by calculating the mean over the three corresponding density categories (e.g., MCU size for G<sub>1</sub> is the mean of dense G<sub>1</sub>, intermediate G<sub>1</sub>, and dense G<sub>1</sub>). Similarly, MCU sizes for individual crowding bins were computed by calculating the mean over the three corresponding cell cycle categories (e.g., MCU size of Dense is the mean of dense G<sub>1</sub>, dense S, and dense G<sub>2</sub>).

#### MCU analysis and MPM construction for pharmacological perturbations

MCUs and SIS were computed as described above for 200 cells randomly drawn from each perturbation and 200 randomly drawn from the DMSO control. Resulting MCU size distributions for all 2000 cells (200 cells × 10 conditions) were transformed by inverse hyperbolic sine operation and z-scored with the mean and the standard deviation measured for each MCU size distribution in control cells (see fig. S7B, middle

panel). SPS from all 2000 cells were combined to calculate one shared Multiplexed Protein Map (MPM, see Fig. 5 and figs. S7 and S8). Spatial distribution of the 69 MCU nodes in 2D space was calculated by t-SNE based on SPS of MCUs of all 2000 cells.

#### *Identification of heterogeneous cell responses to EGF stimulation and selection of their most specific MCUs*

First, each of the 69 MCU size distributions of the EGF stimulated cells (200 cells) was z-scored using the mean and standard deviation of the respective MCU size distribution observed in control cells (200 cells). Next, MCU size distributions of EGF stimulated cells were tested for significant differences to MCU size distributions derived from control cells by Two-sample Kolmogorov-Smirnov test (kstest2 function in MatLab) at a 1% significance level. The z-scored values of MCU size distributions which were identified to be significantly different to those observed in control cells (sdMCUs) were further used for identification of heterogenous cell responses by clustering. Next correlation coefficients based on the sdMCUs were calculated for the EGF stimulated cells. The resulting matrix was then clustered using hierarchical clustering (clustergram function in MatLab, default input values), revealing distinct subpopulations of cells. Next, the EGF stimulated cells were clustered based on their sdMCU sizes by Phenograph using 10 as number of neighbors input (inflection point of function cluster number at any given neighborhood value) and ‘jaccard’ as graph type. The resulting clusters were called Subpopulations (SP) and represent phenotypes of MCU size configurations observed in HeLa cells as response to EGF stimulation. The mean of each sdMCU was calculated per SP and was called rpMCU. To identify MCUs which were most specific to individual SB cells were randomly assigned to clusters of the same size as those identified by Phonograph when SB were identified. Subsequently mean size for each sdMCU was calculated for the randomly generated clusters. Random clustering and calculation of mean sdMCU sizes was performed 5000 times, resulting in distributions of sizes for each sdMCU derived from random cell clustering (rsdMCU). Both, mean and standard deviation were calculated for each rsdMCUs and used to z-score rpMCU values, resulting in a value called normalized response phenotype MCU (nrpMCU). nrpMCU with a size greater than 2.5 were classed as specific for a SP.

#### *Single-cell clustering by MCU sizes and cell mean intensities*

MCU size based single-cell clustering was compared to single-cell clustering based on mean cell intensities by their capability to identify subpopulations in 200 cells randomly drawn from 200 cells from 9 conditions. First, MCU selected for single cell clustering were selected as stated in “Identification of heterogeneous cell responses to EGF stimulation and selection of their most

specific MCUs” (see above). Next, 4i stainings characteristic for MCUs used for clustering were selected for cell-intensity-based clustering as follows. 4i stainings with z-scored intensity in the MCU loadings (fig. S6 B) bigger or smaller than  $\pm 2.375$  STD were selected. Next, single-cell clustering based on MCU sizes and cell intensities was performed as follows. Coefficients of correlation between cells were measured based on their profiles of mean cell intensities or MCU sizes. Next, Pair-wise dissimilarity matrix was calculated for the resulting coefficients of correlation using Euclidean distance and Hierarchically clustering was performed using “average” linkage (MatLab function clustergram) (fig. S8D). Information content of MCU size and cell intensity-based dissimilarity matrix was quantified by calculating entropy of the matrices.

#### *EGFR and p-EGFR signal flux analysis*

Signal intensity (pixel values) of both EGFR and p-EGFR within each MCU of a cell was summed up and normalized by the total signal intensity measured in all MCUs present in a cell. This was repeated for every cell (200 cells). Then, the relative fraction of EGFR and p-EGFR signal present in every MCU was computed by calculating the mean of normalized signal intensities measured in each cell. To control for unspecific signal accumulation in MCUs due to their different sizes, assignment of pixels to MCU was randomized and the fraction of EGFR and p-EGFR to randomized MCUs was measured. This step was repeated 200 times per cell. Randomization results were aggregated by calculating the mean over all randomizations. Mean values were then subtracted from originally observed EGFR and p-EGFR values per cell, prior to calculating the mean over all cells.

#### *Antibody elutability and dynamic range preservation over 20 4i cycles*

To test sample stability, elution of primary and secondary antibody from the sample, and potential back ground signal increase from non-specific binding of secondary antibody over 20 4i cycles the following experiment was performed. (1) The sample was first treated with EB (1x Elut.), (2) then stained only with secondary antibody to record the fluorescence background level (SecAb only). (3) Subsequently the sample was treated with EB (2x Elut.) and then (4) incubated with both primary and secondary antibodies in test wells and only with secondary antibodies in control wells (IF). (5) Primary and secondary antibodies were eluted from the sample (3x Elut.) and (6) the sample was incubated with secondary antibodies only (SecAb only) in both test and control wells. Next, (7) 5 cycles of 4i were performed with the sample without antibody staining and image acquisition, (8) followed by another round of IF (8x Elut.). (9) Primary and secondary antibodies were eluted from the sample (9x Elute) and (10) the sample was incubated with secondary antibodies only (SecAb only). Steps 7 to 10 were repeated twice (15x Elut., 21x Elut.).

#### *Quality control of multiplexed data set*

Data clean-up was performed for each multiplexing cycle independently and consisted of a set of SVM classifiers using CellClassifier (51), as well as automated discretion of any cell touching the image border. SVM Classifiers were trained to identify cells with immunofluorescence artifacts, and missegmented cells and/or nuclei. Such cells were earmarked. Only cells, which were not earmarked by SVM classifiers over the whole set of multiplexing cycles were used for subsequent cell and pixel analysis.

#### **REFERENCES AND NOTES**

1. A. P. Frei et al., Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat. Methods* **13**, 269–275 (2016). doi: [10.1038/nmeth.3742](https://doi.org/10.1038/nmeth.3742); pmid: [26808670](https://pubmed.ncbi.nlm.nih.gov/26808670/)
2. T. Hashimshony, F. Wagner, N. Sher, I. Yanai, CEL-Seq: Single-cell RNA-Seq by multiplexed linear amplification. *Cell Reports* **2**, 666–673 (2012). doi: [10.1016/j.celrep.2012.08.003](https://doi.org/10.1016/j.celrep.2012.08.003); pmid: [23099981](https://pubmed.ncbi.nlm.nih.gov/23099981/)
3. S. Islam et al., Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* **11**, 163–166 (2014). doi: [10.1038/nmeth.2772](https://doi.org/10.1038/nmeth.2772); pmid: [2436023](https://pubmed.ncbi.nlm.nih.gov/2436023/)
4. C. Giesen et al., Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014). doi: [10.1038/nmeth.2869](https://doi.org/10.1038/nmeth.2869); pmid: [24584193](https://pubmed.ncbi.nlm.nih.gov/24584193/)
5. M. Angelo et al., Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442 (2014). doi: [10.1038/nm.3488](https://doi.org/10.1038/nm.3488); pmid: [24584119](https://pubmed.ncbi.nlm.nih.gov/24584119/)
6. J. H. Lee et al., Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi: [10.1126/science.1250217](https://doi.org/10.1126/science.1250217); pmid: [24758530](https://pubmed.ncbi.nlm.nih.gov/24758530/)
7. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015). doi: [10.1126/science.aaa6090](https://doi.org/10.1126/science.aaa6090); pmid: [25858977](https://pubmed.ncbi.nlm.nih.gov/25858977/)
8. P. J. Thul et al., A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017). doi: [10.1126/science.aal3321](https://doi.org/10.1126/science.aal3321); pmid: [28495876](https://pubmed.ncbi.nlm.nih.gov/28495876/)
9. O. Rozenblatt-Rosen, M. J. T. Stubbington, A. Regev, S. A. Teichmann, The Human Cell Atlas: From vision to reality. *Nature* **550**, 451–453 (2017). doi: [10.1038/550451a](https://doi.org/10.1038/550451a); pmid: [29072289](https://pubmed.ncbi.nlm.nih.gov/29072289/)
10. A. R. Buxbaum, G. Haimovich, R. H. Singer, In the right place at the right time: Visualizing and understanding mRNA localization. *Nat. Rev. Mol. Cell Biol.* **16**, 95–109 (2015). doi: [10.1038/nrm3918](https://doi.org/10.1038/nrm3918); pmid: [25549890](https://pubmed.ncbi.nlm.nih.gov/25549890/)
11. V. I. Korolchuk et al., Lysosomal positioning coordinates cellular nutrient responses. *Nat. Cell Biol.* **13**, 453–460 (2011). doi: [10.1038/ncb2204](https://doi.org/10.1038/ncb2204); pmid: [21394080](https://pubmed.ncbi.nlm.nih.gov/21394080/)
12. A. M. Valin et al., Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017). doi: [10.1038/nature22369](https://doi.org/10.1038/nature22369); pmid: [28538724](https://pubmed.ncbi.nlm.nih.gov/28538724/)
13. W. Schubert et al., Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat. Biotechnol.* **24**, 1270–1278 (2006). doi: [10.1038/nbt1250](https://doi.org/10.1038/nbt1250); pmid: [17013374](https://pubmed.ncbi.nlm.nih.gov/17013374/)
14. M. J. Gerdes et al., Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11982–11987 (2013). doi: [10.1073/pnas.1300136110](https://doi.org/10.1073/pnas.1300136110); pmid: [23818604](https://pubmed.ncbi.nlm.nih.gov/23818604/)
15. C. Wahlby, F. Erlandsson, E. Bengtsson, A. Zetterberg, Sequential immunofluorescence staining and image analysis for detection of large numbers of antigens in individual cell nuclei. *Cytometry* **47**, 32–41 (2002). doi: [10.1002/cyto.10026](https://doi.org/10.1002/cyto.10026); pmid: [11774347](https://pubmed.ncbi.nlm.nih.gov/11774347/)
16. H. Y. Lan, W. Mu, D. J. Nikolic-Paterson, R. C. Atkins, A novel, simple, reliable, and sensitive method for multiple immunoenzyme staining: Use of microwave oven heating to block antibody crossreactivity and retrieve antigens. *J. Histochem. Cytochem.* **43**, 97–102 (1995). doi: [10.1177/43.1.722770](https://doi.org/10.1177/43.1.722770); pmid: [7822770](https://pubmed.ncbi.nlm.nih.gov/7822770/)
17. E. C. Stack, C. Wang, K. A. Roman, C. C. Hoyt, Multiplexed immunohistochemistry, imaging, and quantitation: A review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods* **70**, 46–58 (2014). doi: [10.1016/j.ymeth.2014.08.016](https://doi.org/10.1016/j.ymeth.2014.08.016); pmid: [25242720](https://pubmed.ncbi.nlm.nih.gov/25242720/)

18. J. R. Lin, M. Fallahi-Sichani, P. K. Sorger, Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat. Commun.* **6**, 8390 (2015). doi: [10.1038/ncomms9390](https://doi.org/10.1038/ncomms9390); pmid: [26399630](https://pubmed.ncbi.nlm.nih.gov/26399630/)
19. P. Zraževskiy, X. Gao, Quantum dot imaging platform for single-cell molecular profiling. *Nat. Commun.* **4**, 1619 (2013). doi: [10.1038/ncomms2635](https://doi.org/10.1038/ncomms2635); pmid: [23511483](https://pubmed.ncbi.nlm.nih.gov/23511483/)
20. R. M. Schweller *et al.*, Multiplexed *in situ* immunofluorescence using dynamic DNA complexes. *Angew. Chem. Int. Ed. Engl.* **51**, 9292–9296 (2012). doi: [10.1002/anie.201204304](https://doi.org/10.1002/anie.201204304); pmid: [22893271](https://pubmed.ncbi.nlm.nih.gov/22893271/)
21. R. Gendusa, C. R. Scalia, S. Buscone, G. Cattoretti, Elution of high-affinity ( $>10^9$  K<sub>D</sub>) antibodies from tissue sections: Clues to the molecular mechanism and use in sequential immunostaining. *J. Histochem. Cytochem.* **62**, 519–531 (2014). doi: [10.1369/0022155414536732](https://doi.org/10.1369/0022155414536732); pmid: [24794148](https://pubmed.ncbi.nlm.nih.gov/24794148/)
22. B. Snijder *et al.*, Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* **461**, 520–523 (2009). doi: [10.1038/nature08282](https://doi.org/10.1038/nature08282); pmid: [19710653](https://pubmed.ncbi.nlm.nih.gov/19710653/)
23. P. Liberali, B. Snijder, L. Pelkmans, A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell* **157**, 1473–1487 (2014). doi: [10.1016/j.cell.2014.04.029](https://doi.org/10.1016/j.cell.2014.04.029); pmid: [24906158](https://pubmed.ncbi.nlm.nih.gov/24906158/)
24. N. Battich, T. Stoeger, L. Pelkmans, Control of Transcript Variability in Single Mammalian Cells. *Cell* **163**, 1596–1610 (2015). doi: [10.1016/j.cell.2015.11.018](https://doi.org/10.1016/j.cell.2015.11.018); pmid: [26687353](https://pubmed.ncbi.nlm.nih.gov/26687353/)
25. S. Devanathan, T. A. Dahl, W. R. Midden, D. C. Necker, Readily available fluorescein isothiocyanate-conjugated antibodies can be easily converted into targeted phototoxic agents for antibacterial, antiviral, and anticancer therapy. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2980–2984 (1990). doi: [10.1073/pnas.87.8.2980](https://doi.org/10.1073/pnas.87.8.2980); pmid: [2109321](https://pubmed.ncbi.nlm.nih.gov/2109321/)
26. J. C. Liao, J. Roider, D. G. Jay, Chromophore-assisted laser inactivation of proteins is mediated by the photogeneration of free radicals. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2659–2663 (1994). doi: [10.1073/pnas.91.7.2659](https://doi.org/10.1073/pnas.91.7.2659); pmid: [8146171](https://pubmed.ncbi.nlm.nih.gov/8146171/)
27. P. Yan *et al.*, Fluorophore-assisted light inactivation of calmodulin involves singlet-oxygen mediated cross-linking and methionine oxidation. *Biochemistry* **45**, 4736–4748 (2006). doi: [10.1021/bi052395a](https://doi.org/10.1021/bi052395a); pmid: [16605242](https://pubmed.ncbi.nlm.nih.gov/16605242/)
28. A. Michaeli, J. Feitelson, Reactivity of singlet oxygen toward amino acids and peptides. *Photochem. Photobiol.* **59**, 284–289 (2017). doi: [10.1111/j.1751-1097.1994.tb05035.x](https://doi.org/10.1111/j.1751-1097.1994.tb05035.x); pmid: [8016206](https://pubmed.ncbi.nlm.nih.gov/8016206/)
29. K. W. Marek, G. W. Davis, Transgenically encoded protein photoactivation (FIash-FALI): Acute inactivation of synaptotagmin I. *Neuron* **36**, 805–813 (2002). doi: [10.1016/S0896-6273\(02\)01068-1](https://doi.org/10.1016/S0896-6273(02)01068-1); pmid: [12467585](https://pubmed.ncbi.nlm.nih.gov/12467585/)
30. G. Gut, M. D. Tadmor, D. Pe'er, L. Pelkmans, P. Liberali, Trajectories of cell-cycle progression from fixed cell populations. *Nat. Methods* **12**, 951–954 (2015). doi: [10.1038/nmeth.3545](https://doi.org/10.1038/nmeth.3545); pmid: [26301842](https://pubmed.ncbi.nlm.nih.gov/26301842/)
31. S. Van Gassen *et al.*, FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636–645 (2015). doi: [10.1002/cyto.a.22625](https://doi.org/10.1002/cyto.a.22625); pmid: [25573116](https://pubmed.ncbi.nlm.nih.gov/25573116/)
32. L. M. Weber, M. D. Robinson, Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry A* **89**, 1084–1096 (2016). doi: [10.1002/cyto.a.23030](https://doi.org/10.1002/cyto.a.23030); pmid: [2799211](https://pubmed.ncbi.nlm.nih.gov/2799211/)
33. J. H. Levine *et al.*, Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184–197 (2015). doi: [10.1016/j.cell.2015.05.047](https://doi.org/10.1016/j.cell.2015.05.047); pmid: [26095251](https://pubmed.ncbi.nlm.nih.gov/26095251/)
34. J. Zanconato *et al.*, Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat. Cell Biol.* **17**, 1218–1227 (2015). doi: [10.1038/ncb3216](https://doi.org/10.1038/ncb3216); pmid: [26258633](https://pubmed.ncbi.nlm.nih.gov/26258633/)
35. D. Boyer, R. Quintanilla, K. K. Lee-Fruman, Regulation of catalytic activity of S6 kinase 2 during cell cycle. *Mol. Cell. Biochem.* **307**, 59–64 (2008). doi: [10.1007/s11010-007-9584-5](https://doi.org/10.1007/s11010-007-9584-5); pmid: [17786541](https://pubmed.ncbi.nlm.nih.gov/17786541/)
36. M. K. Mateyak, A. J. Obaya, S. Adachi, J. M. Sedivy, Phenotypes of c-Myc-deficient rat fibroblasts isolated by targeted homologous recombination. *Cell Growth Differ.* **8**, 1039–1048 (1997). pmid: [9342182](https://pubmed.ncbi.nlm.nih.gov/9342182/)
37. A. Seth, S. Gupta, R. J. Davis, Cell cycle regulation of the c-Myc transcriptional activation domain. *Mol. Cell. Biol.* **13**, 4125–4136 (1993). doi: [10.1128/MCB.13.7.4125](https://doi.org/10.1128/MCB.13.7.4125); pmid: [8321217](https://pubmed.ncbi.nlm.nih.gov/8321217/)
38. J. Shorter, G. Warren, Golgi architecture and inheritance. *Annu. Rev. Cell Dev. Biol.* **18**, 379–420 (2002). doi: [10.1146/annurev.cellbio.18.030602.133733](https://doi.org/10.1146/annurev.cellbio.18.030602.133733); pmid: [12142281](https://pubmed.ncbi.nlm.nih.gov/12142281/)
39. C. A. Wurm *et al.*, Nanoscale distribution of mitochondrial import receptor Tom20 is adjusted to cellular conditions and exhibits an inner-cellular gradient. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13546–13551 (2011). doi: [10.1073/pnas.1107553108](https://doi.org/10.1073/pnas.1107553108); pmid: [21799113](https://pubmed.ncbi.nlm.nih.gov/21799113/)
40. S. Sigismund *et al.*, Threshold-controlled ubiquitination of the EGFR directs receptor fate. *EMBO J.* **32**, 2140–2157 (2013). doi: [10.1038/embj.2013.149](https://doi.org/10.1038/embj.2013.149); pmid: [23799367](https://pubmed.ncbi.nlm.nih.gov/23799367/)
41. F. Capuani *et al.*, Quantitative analysis reveals how EGFR activation and downregulation are coupled in normal but not in cancer cells. *Nat. Commun.* **6**, 7999 (2015). doi: [10.1038/ncomms8999](https://doi.org/10.1038/ncomms8999); pmid: [26264748](https://pubmed.ncbi.nlm.nih.gov/26264748/)
42. L. K. Opresko *et al.*, Endocytosis and lysosomal targeting of epidermal growth factor receptors are mediated by distinct sequences independent of the tyrosine kinase domain. *J. Biol. Chem.* **270**, 4325–4333 (1995). doi: [10.1074/jbc.270.9.4325](https://doi.org/10.1074/jbc.270.9.4325); pmid: [7876194](https://pubmed.ncbi.nlm.nih.gov/7876194/)
43. M. Frechin *et al.*, Cell-intrinsic adaptation of lipid composition to local crowding drives social behaviour. *Nature* **523**, 88–91 (2015). doi: [10.1038/nature14429](https://doi.org/10.1038/nature14429); pmid: [26009010](https://pubmed.ncbi.nlm.nih.gov/26009010/)
44. H. Mahboubi, M. Kodiba, U. Stochaj, Automated detection and quantification of granular cell compartments. *Microsc. Microanal.* **19**, 617–628 (2013). doi: [10.1017/S1431927613000159](https://doi.org/10.1017/S1431927613000159); pmid: [23552615](https://pubmed.ncbi.nlm.nih.gov/23552615/)
45. C. Sommer, C. Straehle, U. Köthe, F. A. Hamprecht, in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. (2011), pp. 230–233.
46. W. Schubert, A. Gieselear, A. Krusche, P. Serocka, R. Hillert, Next-generation biomarkers based on 100-parameter functional super-resolution microscopy TIS. *N. Biotechnol.* **29**, 599–610 (2012). doi: [10.1016/j.nbt.2011.12.004](https://doi.org/10.1016/j.nbt.2011.12.004); pmid: [22209707](https://pubmed.ncbi.nlm.nih.gov/22209707/)
47. N. Battich, T. Stoeger, L. Pelkmans, Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat. Methods* **10**, 1127–1133 (2013). doi: [10.1038/nmeth.2657](https://doi.org/10.1038/nmeth.2657); pmid: [24097269](https://pubmed.ncbi.nlm.nih.gov/24097269/)
48. M. Guizar-Sicairos, S. T. Thurman, J. R. Fienup, Efficient subpixel image registration algorithms. *Opt. Lett.* **33**, 156–158 (2008). doi: [10.1364/OL.33.000156](https://doi.org/10.1364/OL.33.000156); pmid: [18197224](https://pubmed.ncbi.nlm.nih.gov/18197224/)
49. A. E. Carpenter *et al.*, CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006). doi: [10.1186/gb-2006-7-10-r100](https://doi.org/10.1186/gb-2006-7-10-r100); pmid: [17076895](https://pubmed.ncbi.nlm.nih.gov/17076895/)
50. B. Snijder *et al.*, Single-cell analysis of population context advances RNAi screening at multiple levels. *Mol. Syst. Biol.* **8**, 579 (2012). doi: [10.1038/msb.2012.9](https://doi.org/10.1038/msb.2012.9); pmid: [22531119](https://pubmed.ncbi.nlm.nih.gov/22531119/)
51. P. Rämö, R. Sacher, B. Snijder, B. Begemann, L. Pelkmans, CellClassifier: Supervised learning of cellular phenotypes. *Bioinformatics* **25**, 3028–3030 (2009). doi: [10.1093/bioinformatics/btp524](https://doi.org/10.1093/bioinformatics/btp524); pmid: [19729371](https://pubmed.ncbi.nlm.nih.gov/19729371/)
52. S. Van Gassen *et al.*, FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636–645 (2015). doi: [10.1002/cyto.a.22625](https://doi.org/10.1002/cyto.a.22625); pmid: [25573116](https://pubmed.ncbi.nlm.nih.gov/25573116/)

**ACKNOWLEDGMENTS**

We thank all members of the Pelkmans lab and Klemm lab for constructive discussions and comments during the preparation of the manuscript. In particular, we thank R. Holtackers for experimental support at the beginning of this study and R. Klemm and V. Eastham (University of Zurich) for expert feedback throughout the study. **Author contributions:** L.P. initiated the study. G.G., M.H., L.P. designed experiments. G.G. performed and analyzed the experiments. G.G. developed analysis algorithms. L.P. and G.G. wrote the manuscript. **Funding:** L.P. acknowledges financial support from the University of Zurich, the Swiss initiative in Systems Biology SystemsX.ch (MorphogenetiX, and PhosphoNetPPM), and the Swiss National Science Foundation. **Data and materials availability:** Custom Cell Profiler 1 modules and scripts to generate MCUs and MPMs can be downloaded from [www.github.com/Pelkmanslab](https://www.github.com/Pelkmanslab) and [www.cellatlas.org](https://www.cellatlas.org). Single-cell data supporting the findings of this study can be found in the supplementary materials (tables S3 and S4). **Competing interests:** L.P., M.H., and G.G. have filed a patent on the 4i technology, and L.P. and G.G. have filed a patent on the approach to generate and use MCUs and MPMs.

**SUPPLEMENTARY MATERIALS**

[www.science.org/content/361/6401/eaar7042/suppl/DC1](https://www.science.org/content/361/6401/eaar7042/suppl/DC1)  
Figs. S1 to S8  
Tables S1 to S4

11 December 2017; resubmitted 23 March 2018

Accepted 21 June 2018

10.1126/science.eaar7042

## Multiplexed protein maps link subcellular organization to cellular states

Gabriele Gut, Markus D. Herrmann and Lucas Pelkmans

Science 361 (6401), eaar7042.  
DOI: 10.1126/science.aar7042

### Making multiplexed subcellular protein maps

Being able to visualize protein localizations within cells and tissues by means of immuno-fluorescence microscopy has been key to developments in cell biology and beyond. Gut *et al.* present a high-throughput method that achieves the simultaneous quantification of their expression levels in thousands of single cells; captures their detailed subcellular distribution to various compartments, organelles, and cellular structures within each of these single cells; and places all this information within a multicellular context. Such a scale-crossing dataset empowers artificial intelligence-based computer vision algorithms to achieve a comprehensive profiling of intracellular protein maps to measure their responses to different multicellular, cellular, and pharmacological contexts, and to reveal new cellular states.

Science, this issue p. eaar7042

#### ARTICLE TOOLS

<http://science.scienmag.org/content/361/6401/eaar7042>

#### SUPPLEMENTARY MATERIALS

<http://science.scienmag.org/content/suppl/2018/08/01/361.6401.eaar7042.DC1>

#### REFERENCES

This article cites 51 articles, 11 of which you can access for free  
<http://science.scienmag.org/content/361/6401/eaar7042#BIBL>

#### PERMISSIONS

<http://www.scienmag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2018, American Association for the Advancement of Science