

METHOD

Multidimensional profiling of drug-treated cells by Imaging Mass Cytometry

Alexandre Bouzekri, Amanda Esch and Olga Ornatsky

Proteomics R&D Department, Fluidigm Canada Inc., Markham, Canada

Keywords

anticancer drug treatment; cell-based screening; image-based computational analysis; Imaging Mass Cytometry; multiparametric similarities measurements

Correspondence

A. Bouzekri, Proteomics R&D Department, Fluidigm Canada Inc., 1380 Rodick Road, Markham, ON L3R 4G5, Canada
Tel: +1 905 943 0133
E-mail: alexandre.bouzekri@fluidigm.com

(Received 11 March 2019, revised 17 May 2019, accepted 26 June 2019)

doi:10.1002/2211-5463.12692

In pharmaceutical research, high-content screening is an integral part of lead candidate development. Measuring drug response *in vitro* by examining over 40 parameters, including biomarkers, signaling molecules, cell morphological changes, proliferation indices, and toxicity in a single sample, could significantly enhance discovery of new therapeutics. As a proof of concept, we present here a workflow for multidimensional Imaging Mass Cytometry™ (IMC™) and data processing with open source computational tools. CellProfiler was used to identify single cells through establishing cellular boundaries, followed by histoCAT™ (histology topography cytometry analysis toolbox) for extracting single-cell quantitative information visualized as t-SNE plots and heatmaps. Human breast cancer-derived cell lines SKBR3, HCC1143, and MCF-7 were screened for expression of cellular markers to generate digital images with a resolution comparable to conventional fluorescence microscopy. Predicted pharmacodynamic effects were measured in MCF-7 cells dosed with three target-specific compounds: growth stimulatory EGF, microtubule depolymerization agent nocodazole, and genotoxic chemotherapeutic drug etoposide. We show strong pairwise correlation between nuclear markers pHistone3^{S28}, Ki-67, and p4E-BP1^{T37/T46} in classified mitotic cells and anticorrelation with cell surface markers. Our study demonstrates that IMC data expand the number of measured parameters in single cells and brings higher-dimension analysis to the field of cell-based screening in early lead compound discovery.

Discovery of new treatments in oncology research relies extensively on the use of human-derived cell culture models [1]. High-content cell-based screens are widely applied in pharmaceutical drug development to prioritize lead molecules for animal testing [2]. These assays rely on the use of primary and cancer cell lines and mostly monitor cytotoxicity and proliferation. With the advent of sophisticated genomics and proteomics technologies for soluble proteins and mass

cytometry for single cells, it is becoming possible to increase the multidimensionality of *in vitro* screens. Databases of genotypic and phenotypic profiles [3–5] across cancer drug panels have provided a comparative analysis between *in vitro* studies and clinical therapeutic responses *in vivo* [6]. One of these datasets developed and shared by the National Cancer Institute 60 (NCI-60) in the late 1980s was the first *in vitro* discovery screening tool for pharmacologic compounds

Abbreviations

4E-BP1, eukaryotic initiation factor 4E-binding protein; cPARP, cleaved poly (ADP-ribose); CP, CellProfiler; EGF, epidermal growth factor; ER α , estrogen receptor alpha; FFPE, formalin-fixed, paraffin-embedded; HER2, human epidermal growth factor receptor 2; histoCAT, histology topography cytometry analysis toolbox; IMC, Imaging Mass Cytometry; LAMP-2, lysosome-associated membrane protein 2; MCF-7, Michigan Cancer Foundation-7; NCI-60, National Cancer Institute 60; ROI, regions of interest; RT, room temperature; SNE, stochastic neighbor embedding.

inducing growth inhibition across 60 cancer cell lines [7]. Over the years, several concerns were raised regarding the use of *in vitro* established human tumor-derived cell lines for drug testing. Concerns included genetic instability and dedifferentiation. However, the study of genetic mutations arising in immortalized cells remained a field of interest in drug discovery, providing knowledge on the dysregulation of cellular signaling pathways and the effects small-molecule inhibitors have on human tumor cell lines [8].

High-content imaging of cancer cell lines in response to drug treatment is a standard assay applied in pre-clinical studies for identification of different mechanisms of drug action [9–12]. The main criterion in every morphophenotypic screen for early drug discovery is the selection of biomarkers and detection modalities (antibodies, chemical and enzymatic probes, reporter detection tags). In the field of fluorescent cellular imaging microscopy, biomarker analysis in single cells is limited to 4–6 due to the overlapping spectra of fluorescent dyes [13]. Repetitive rounds of staining using the same biological sample are required to achieve multiplexing [14]. New imaging technologies are needed to significantly increase multiplicity in a single sample preparation experiment and perform replicate analyses [15]. High-content imaging assays applied to screening drug perturbations in heterogeneous cancer cell models could uncover additional modes of drug action and biomarkers for clinical trials [16].

Imaging Mass Cytometry is an emerging and transformative technique in the field of digital histopathology applied to complex tissue sections [17–19]. The Hyperion™ Imaging System (Fluidigm®, Toronto, Canada) can measure up to 40 parameters simultaneously in formalin-fixed, paraffin-embedded (FFPE), and frozen human tissue sections with subcellular resolution [20]. Sample preparation is very similar to standard immunohistochemical protocols [17], where the tissue section on a slide is first deparaffinized and treated with an antigen retrieval buffer for antigen epitope exposure in the case of FFPE, then stained once with a mixture of isotopically pure metal-labeled antibodies specific to structural and cell type biomarkers. After drying and inserting the slide into the ablation chamber, regions of interest (ROIs) are chosen and recorded by a camera integrated with the Hyperion Tissue Imager (Fluidigm®). These ROIs are then ablated in 1-μm steps as the slide moves under the laser. Each laser pulse vaporizes a 1-μm² area of the tissue and generates a plume of particles. The resulting particles are carried by a stream of helium/argon mixture into a CyTOF® instrument, an inductively coupled plasma

time-of-flight mass cytometer, described elsewhere [18]. Data are stored as raw binary files of ion counts for each mass channel (parameter) per pixel. These files can be converted into a stack of single-parameter matrices (grayscale images), which are used to create merged multiparametric pseudocolor images and perform quantitative pixel- or segmentation-based analysis. Bodenmiller *et al.* are developing multivariate computational tools to visualize and analyze multiplexed images of human tissue sections generated by IMC [21]. However, there is no single software package or analysis workflow that could currently be applied to answer specific biological questions.

In this proof-of-principle study, we set out to develop a comprehensive workflow for IMC data analysis based on recent advances in imaging algorithmic methods to visualize and measure multiple biomarkers in model cell lines cultured in chamber slides (Fig. 1). Several human breast cancer-derived cell lines, SKBR3, HCC1143, and MCF-7, were screened for expression of surface and intracellular markers. Predicted pharmacodynamic effects were studied in MCF-7 cells dosed with three target-specific compounds: growth stimulatory epidermal growth factor (EGF), microtubule depolymerization agent nocodazole, and genotoxic chemotherapeutic drug etoposide [22,23]. Our method analysis workflow demonstrates the high multiplexing capability of IMC for *in vitro* research and offers quantitation of multiple biomarkers at sub-cellular resolution. Future improvements of the technology toward a higher acquisition speed of multiple samples will expand IMC applications as an imaging platform for *in vitro* cell-based drug profiling.

Materials and methods

Cell culture

SKBR3 (HTB-30™), HCC1143 (CRL-2321™), and MCF-7 (HTB-22™) were purchased, tested for mycoplasma contamination, and authenticated with short tandem repeat DNA profiling by American Tissue Culture Collection (ATCC®, Manassas, VA, USA). Cells were cultured within 15 passages in their corresponding growth media supplemented with 10% bovine serum (HyClone™ Cosmic Calf™ Serum, Logan, UT, USA; Cat. No. SH30087.04) and penicillin/streptomycin (Thermo Fisher Scientific, Waltham, MA, USA; Cat. No. 15140-122); McCoy's 5a (ATCC 30-2007™) for SKBR3, RPMI-1640 (ATCC 30-2001™) for HCC1143, and DMEM (ATCC 30-2002™) for MCF-7 with 0.01 mg·mL⁻¹ insulin (Sigma-Aldrich®, St. Louis, MO, USA; Cat. No. I9278). Cells at 70–80% monolayer confluence were washed with Versene (Thermo Fisher Scientific; Cat. No. 15040-066), detached with trypsin/EDTA 0.25%

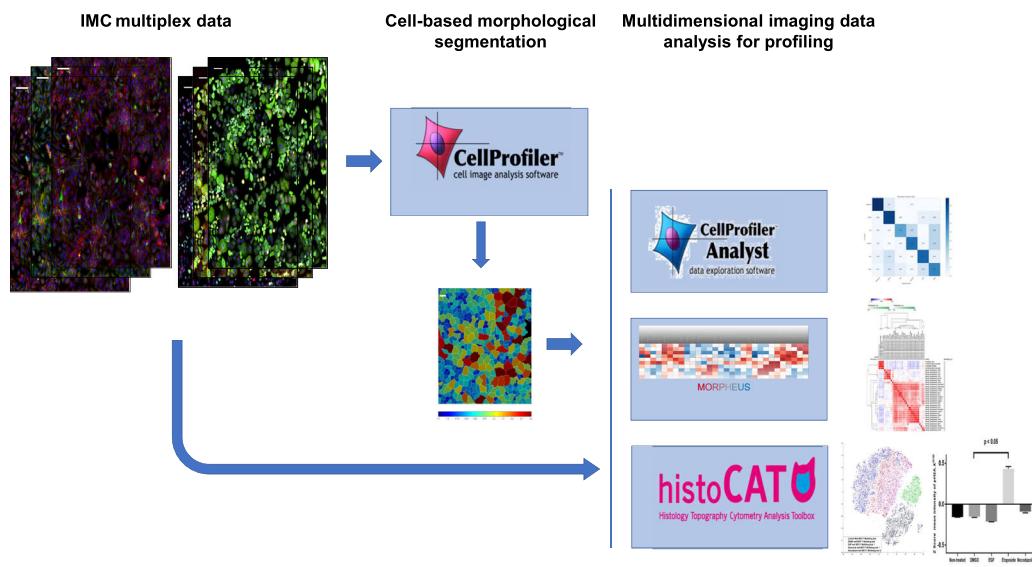


Fig. 1. Image-based IMC data analysis workflow with open source software.

(Thermo Fisher Scientific Cat. No. 25200-056), collected, and counted. For each cell line, chamber slides (BioCoatTM, Corning[®], New York, NY, USA Collagen I 8-Well Culture Slide, Corning[®] Cat. No. 354630) were used to seed cells at a concentration of $0.1 \times 10^6 \text{ mL}^{-1}$ in 0.5 mL per chamber and grown in 37 °C, 5% CO₂, 100% humidity for 48 h prior to drug compound addition. Each breast cancer cell line was seeded using individual chamber slides to avoid any cell cross-contamination, and each IMC experiment was carried out independently. In chronological order, seeding of HCC1143 was the first sample preparation, then we run the second experiment using SKBR3, and last we used MCF-7 for the drug-treated cells *in vitro* study.

In vitro drug treatment

All chemical compounds were dissolved and aliquoted at the supplier's recommended concentrations for long-term storage in 100% DMSO (Sigma-Aldrich Cat. No. 276855). Final concentrations for each compound were selected from previously published data [24–26] by premixing the required initial drug stock volume with full growth media and then transferring 0.5 mL volume to corresponding cell chambers (Table S1): 10 μM of etoposide (Cell Signaling Technology[®], Danvers, MA, USA; Cat. No. 2200S); 0.5 μM of nocodazole (Sigma-Aldrich Cat. No. M1404); 10 ng·mL⁻¹ of human EGF (Sigma-Aldrich Cat. No. E9644); 2% DMSO in growth media. One chamber was left with fresh growth media as nontreated control. Cells were exposed to compounds for 48 h at 37 °C, 5% CO₂, and 100% humidity. Preparation of additional eight chamber sample slides to replicate our experiment for IMC analysis was not

required due to the known *in vitro* predicted effect of the drugs we tested on this cell line.

Immunocytochemistry and antibody validation

Immunostaining of live cells was performed in chamber slides. Following drug treatment, fresh media in all chambers was supplemented with Cell-IDTM Intercalator-103Rh (Fluidigm Cat. No. 201103A) at 1 : 500 for 15 min at 37 °C, 5% CO₂ for dead cell identification. The chamber slides were washed with DPBS (Thermo Fisher Scientific Cat. No. 14190-144) at room temperature (RT) for 5 min. Surface metal-labeled antibody cocktail (Table S2) was made in 0.5% BSA/DPBS from antibody stocks at a dilution of 1 : 100 for each. Cell monolayers were stained with 250 μL antibody mix for 120 min, RT. Following a washing step, cells were fixed with 250 μL freshly made 1.6% formaldehyde (Thermo Fisher Scientific Cat. No. 28906) in DPBS for 15 min at RT. After fixative removal, cells were permeabilized with fresh 0.3% Saponin/DPBS (Sigma-Aldrich Cat. No. S7900-25G) for 30 min at RT and then blocked with 1% BSA/DPBS for 1 h at 37 °C. After blocking, cells were incubated overnight at 4 °C in 250 μL freshly prepared cocktail of metal-labeled antibodies targeting intracellular markers (Table S3) diluted 1 : 50 in 0.3% Saponin/DBPS. The following day cells were washed once with 0.5% BSA and stained with 250 μL of nuclear Cell-ID Intercalator-Ir for 30 min (1 : 1000), RT. After the last wash, each chamber was filled with 250 μL deionized water for 5 min for salt removal. Chambers were detached from slides and left to dry at RT until IMC analysis. The commercial antibodies of our panels have been previously

validated by immunocytochemistry applied to immunofluorescence analysis of adherent cells. A similar sample preparation method was applied to the validation of these antibodies labeled with metals regarding their antigen specificity detected by IMC. The selection of titers tested for surface markers (1 : 100) and intracellular markers (1 : 50) refers to rigorous quality control antibody validation functional assays using single-cell suspension Mass Cytometry, which has the same single isotopic detection modalities of antigens as IMC. The list of antibodies used for each breast cancer cell line experiment is cross-referenced in Tables S2 and S3. A total number of 14 antibodies were tested on HCC1143 cells, 20 markers for SKBR3 and 25 for MCF-7. The markers from each panel were selected based on their specificity to surface membrane, cytoplasmic and nuclear components of single cells, their previous applications in the context of breast cancer cell line studies, and their commercial availability. Expanding a panel to 37 markers requires additional custom labeling of target-specific monoclonal antibodies with nonused metal isotopes.

Data acquisition by IMC

Samples were analyzed with the Hyperion Imaging System (Fluidigm). The dried slide was loaded into the imaging module, where an optical previewing of the ROIs was recorded for laser ablation. Areas of dimension $1000 \times 1000 \mu\text{m}$ were acquired for HCC1143 and $1400 \times 1400 \mu\text{m}$ for SKBR3 cell lines experiments. Replicate ROIs of $1500 \times 1500 \mu\text{m}$ size were collected for MCF-7 exposed to chemical compounds, and a single ROI was ablated for nontreated control and DMSO-treated cells. Each ROI sample acquisition took 4.5 h at an ablation frequency of 200 Hz. The resulting data files were stored in MCD binary format. Multicolor images were generated with open source IMAGEJ 1.51 software [27]. Zoom-in regions for each multicolor image were made by selecting and cropping areas of dimension $400 \times 400 \mu\text{m}$ from the original picture without any change of the pixel size. The resolution of each IMC image shown is $1 \mu\text{m}$ for 1 pixel.

Image analysis software tools

For each recorded ROI, stacks of 16-bit and 32-bit single-channel TIFF files were exported from MCD binary files using MCD™ Viewer 1.0 (Fluidigm). Cell-based morphological segmentation was carried out with two image processing pipelines for 16-bit or 32-bit TIFF files and CELLPROFILER 2.2.0 (CP, Broad Institute, Cambridge, MA, USA), a widely adopted software in the open source image analysis community which has been continually improved since its availability in 2005 to read and analyze cell images using advanced algorithms [28]. The 16-bit TIFF unstacked image format was used as input data in CELLPROFILER to

run a segmentation pipeline with a set of sequential modules to generate and save an unsigned 16-bit integer single-cell mask TIFF image. Inputs of 16-bit TIFF images with their corresponding segmentation mask were uploaded in histoCAT to open a session data analysis. Due to the high number of 29 single isotopic channel TIFF images for each replicate and drug treatment condition, histoCAT analysis was generated faster with a 16-bit format instead of a 32-bit image format which is bigger in size. The 32-bit TIFF unstacked images were loaded in CELLPROFILER to run a segmentation and mean intensity multiparametric measurement analysis of individual cells. The resulting outputs were exported and saved as SQLite single-cell object measurements database files directly compatible with CELLPROFILER ANALYST for downstream computational analysis [29]. Spatial distribution maps, dimensionality reduction, and unsupervised clustering for 16-bit single images were performed using the histoCAT 1.73 open source toolbox [21]. Supervised analysis by support vector machine-learning classification on 32-bit datasets was processed with CELLPROFILER ANALYST 2.2.1 [30]. Hierarchical clustering and correlation heat maps of classified cell populations were prepared using the web-based tool Morpheus (GenePattern, Broad Institute) [31]. Clustering and visualization of different types of markers in classified populations of cells across all controls and drug treatments were performed with the free software environment r 3.5.3, the packages pheatmap 1.0.12, igraph 1.2.4 and edgebundleR 0.1.4 [32].

Statistical analysis

Data from nontreated control, DMSO-only, and drug-treated cells were analyzed using a nonparametric Mann–Whitney test with a two-tailed *P* value for every channel independently. Statistical significance was defined as *P* < 0.05. All statistical analysis was performed with GRAPH-PAD™ PRISM® 7.04 software [33].

Results

Multiparametric characterization of breast cancer cell lines by IMC

Imaging Mass Cytometry analysis revealed heterogeneity in the expression of surface, intracellular, and nuclear markers in human breast cancer cell lines cultured in chamber slides. As an example, we selected the SKBR3 cell line, which is characterized by an invasive phenotype and increased proliferation *in vitro* and is used as a model to delineate mechanisms of resistance to ErbB2-targeted clinical therapies [34,35]. As seen in Fig. 2, the culture consists of mostly small- and medium-size rounded epithelial cells and occasional large multinucleated cells. Most cells show high

levels of the cytoskeletal marker pan-keratin. Human epidermal growth factor receptor 2 (HER2) and EGFR are seen at the surface membrane, while tumor suppressor and transcription factor p53 are localized in the nuclei. Presence of lysosomal organelles is outlined in the cytoplasm with an antibody targeting the lysosome-associated membrane protein 2 (LAMP-2), known as CD107b. Image resolution is similar to fluorescence microscopy. Ductal breast carcinoma-derived triple-negative HCC1143 cell line was used to demonstrate mesenchymal morphology (Fig. 3) with high vimentin detection in the cytoplasm, nuclear localization of p53, and cytoplasmic distribution of the basal-like breast cancer marker cytokeratin 5 [36,37]. The proliferation marker Ki-67 identifies cells in the active state of the cell cycle. Combination of multiple markers in an *in vitro* cell profiling assay allows more precise analysis of cellular states of differentiation and invasiveness characteristic for breast cancer tumors.

In vitro drug effect profiling by IMC

To evaluate how IMC may be translated into future applications in the field of cell-based drug profiling, we initiated a limited drug treatment study using one model breast cancer cell line. The phenotypic and functional responses of wild-type p53-expressing MCF-7 cells to etoposide, nocodazole, or EGF compounds were investigated. In preclinical *in vitro* drug discovery, this model is considered to be reproducible, fast, and inexpensive despite the lack of clinical correlation to drug response *in vivo* due to breast tumor heterogeneity [38]. To identify how these chemical compounds induce cytostatic and cytotoxic effects on this noninvasive luminal breast cancer cell line, we generated multiplexed pseudocolor images of affected cellular compartments (Fig. 4). For example, nuclear size and proliferative activity of each cell were analyzed using the nuclear intercalator-Ir stain (Cell-ID Intercalator-Ir), cell cycle-S phase marker (Cell-ID

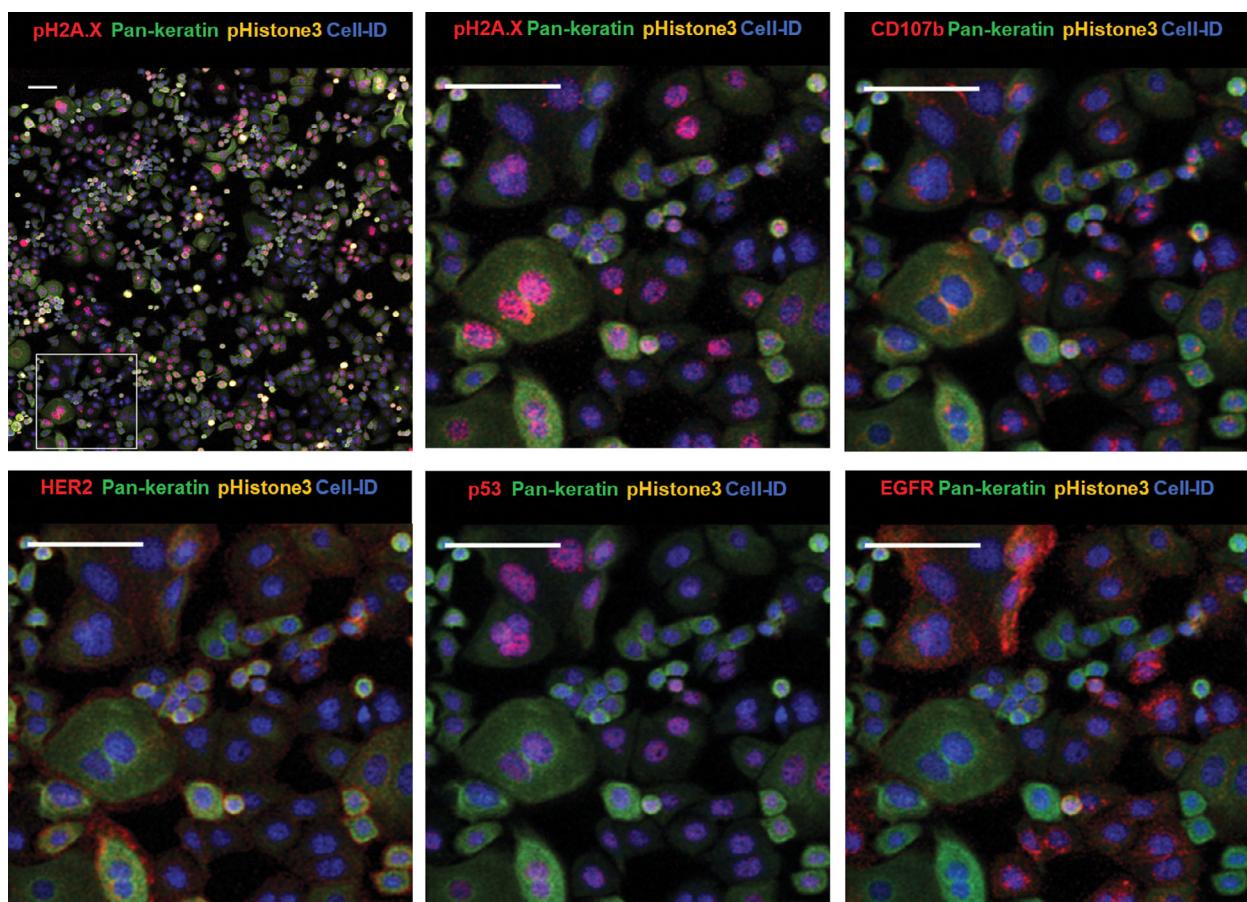


Fig. 2. Imaging Mass Cytometry images of cellular and nuclear markers expressed by SKBR3 with zoom-in area. Cell-ID: nuclei labeled with Ir-intercalator. Zoom-in images are rendered with multiple combinations of protein markers detected in different cellular compartments. Scale bar = 100 μ m.

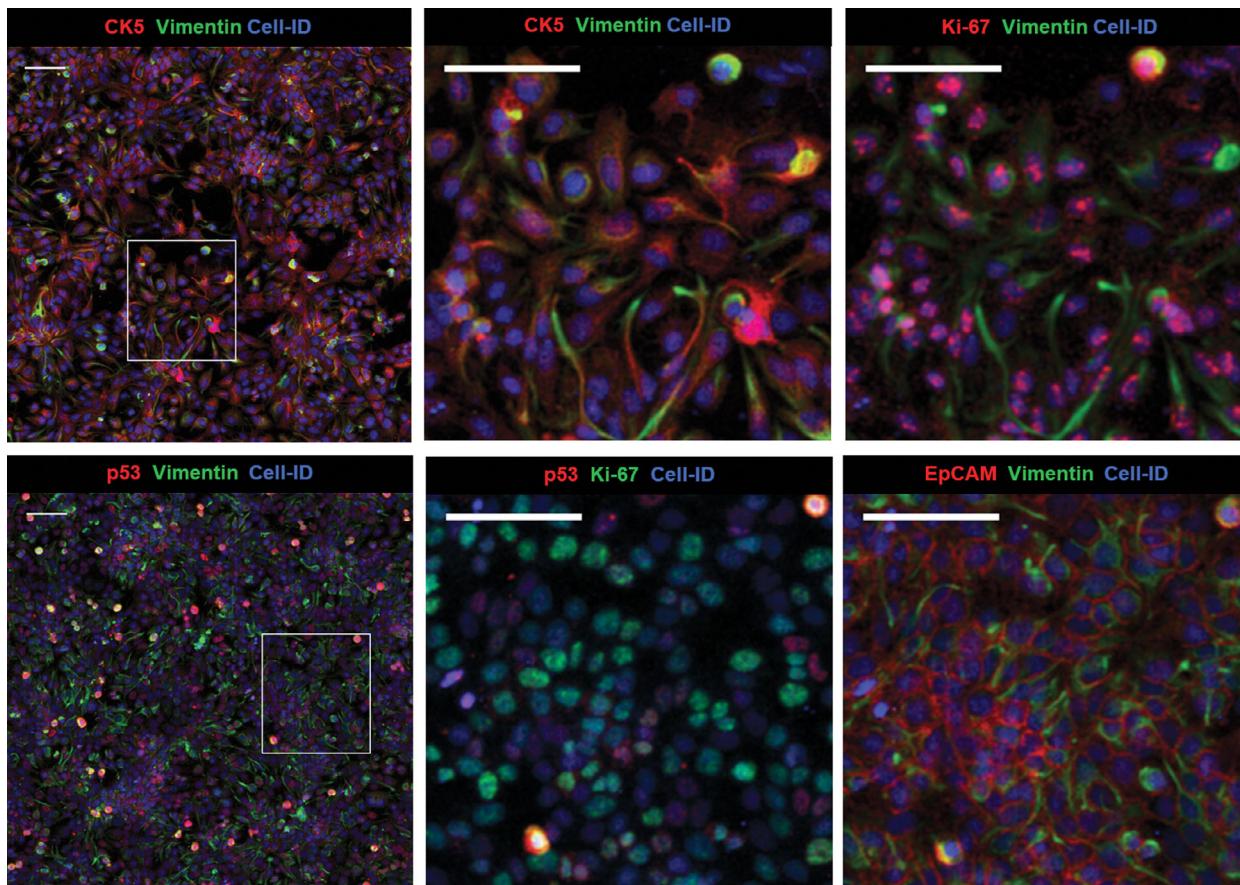


Fig. 3. Imaging Mass Cytometry images of cellular and nuclear markers expressed by HCC1143 with zoom-in areas (two ROI). Cell-ID: nuclei labeled with Ir-intercalator. Each zoom-in composite image is rendered with a selection of different markers. Scale bar = 100 μ m.

IdU), proliferation marker Ki-67, cell cycle regulatory proteins cyclin B1 and D3, and estrogen receptor alpha (ER α). Antibody against pH2A.X S139 was used as a nuclear marker for DNA double-strand breaks and pHistone3 S28 for mitotic cells. Cytoskeletal markers pan-keratin and cytokeratin 19 were used to follow changes in morphology and cell size. Drug response of MCF-7 morphology, adhesion, and cell-to-cell interaction were visualized with surface membrane proteins, such as the membrane-bound mucin marker MUC-1 overexpressed in breast carcinoma [39], cell adhesion molecule EpCAM, integrin CD29, tetraspanin CD81, the low-expressed epidermal growth factor receptor EGFR, regulators of integrin-mediated cell adhesion CD98 and CD47 [40], and E-cadherin as epithelial interaction protein (Table S2).

Long-term exposure of MCF-7 cells to EGF induced significant morphological changes and an increase in cell size. Lineage shift from a classic epithelial cobblestone monolayer to a more invasive mesenchymal type was observed. The presence of large spindle-shaped

cells that seemed to lose contact with other cells and migrate into empty spaces may be related to an increase in cell motility (Fig. 4A). Activation of the EGFR signal transduction pathway by EGF is known to play a role in cellular motility and size increase in the tumor aggressiveness [41].

Screening of MCF-7 cells treated with etoposide, nocodazole, or EGF shows multivariate phenotypic response profiles compared to nontreated and DMSO-treated controls (Fig. S1).

Treatment of MCF-7 cells with the antineoplastic agent nocodazole resulted in an increase in the population of pHistone3 S28 -positive cells with a mitotic index threefold higher compared to DMSO-treated or nontreated control (Table S4). Another cytostatic phenotype was observed after prolonged drug exposure: subpopulations of cyclin D3-positive and tetraploid cells exiting mitosis without cytokinesis as a result of microtubule disruption by nocodazole (Fig. 4B). Presence of cyclin D3-positive and tetraploid cells in the G1/S

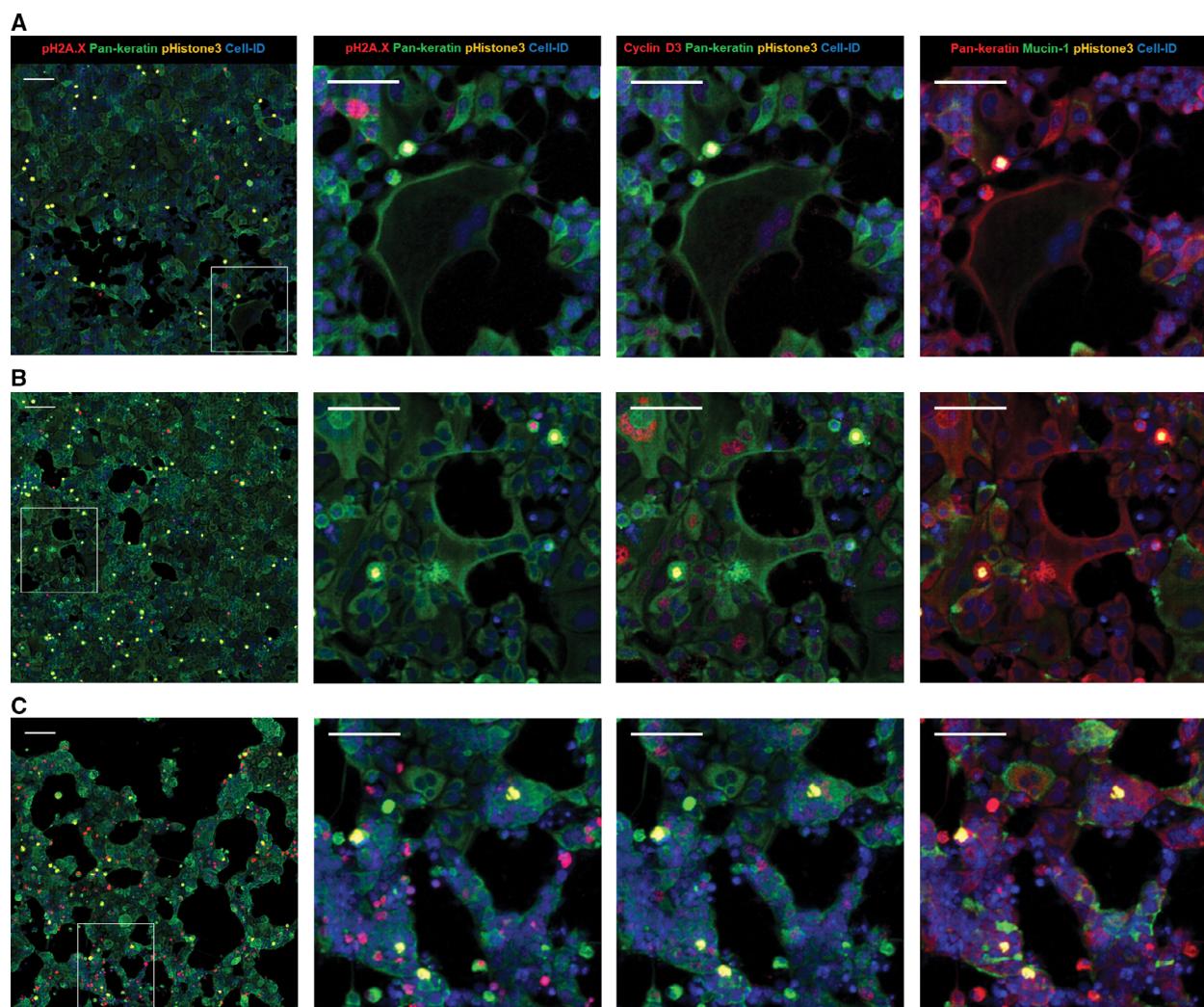


Fig. 4. Multiplexed IMC images of MCF-7 cells treated with different compounds. (A) EGF, (B) nocodazole, (C) etoposide. Composite images of 1500 × 1500 μm size (scale bar = 160 μm) with 400 × 400 μm zoom-in areas (scale bar = 100 μm). Columns images correspond to the same combination of proteins detected per drug treatment (row).

interphase confirms nocodazole effect on MCF-7 cell cycle [24].

Incubation of MCF-7 with the anticancer agent etoposide shows inhibition of cell proliferation, elevated numbers of cells positive for nuclear pH2A.X^{S139} protein, and compaction of the monolayer compared to EGF- or nocodazole-treated samples (Fig. 4C). Exposure of tumor-derived cell lines to this inhibitor induces the formation of a stable covalent topoiso-merase II-cleaved DNA complex, which causes multiple breaks in double-stranded DNA. Etoposide treatment of MCF-7 activates phosphorylation of H2A.X, which is required for checkpoint-mediated cell cycle arrest, promotes DNA repair, and maintains genomic stability. Thus, pH2A.X^{S139} is a sensitive

biomarker for measuring levels of drug-induced double-stranded DNA breaks in cancer cells [42].

High-dimensional phenotype analysis of *in vitro* drug response

Multivariate phenotypic responses of MCF-7 cells following chemical perturbations were quantified with a set of available software tools. Unfortunately, there is no single software package for this type of analysis. Therefore, we employed a workflow that includes a wide variety of tools designed for cellular imaging. Hopefully, this work will underscore the need for software specifically developed for analysis of multiparametric data acquired by IMC.

First, we used CELLPROFILER 2.2.0 (CP) to identify single cells through establishing nuclear and cellular boundaries [43]. Accuracy and robustness of cell segmentation were assessed by using Cell-ID Intercalator-Ir to identify cell nuclei and pan-keratin for cytoplasm (Fig. 5A, Figs S2A and S3). The segmentation masks were generated for each ROI and then exported for further downstream analysis of changes in MCF-7 cellular phenotypes in response to drug treatment.

Second, to visualize and analyze data we used the open source platform histoCAT, which enables highly multiplexed, quantitative, and detailed analysis of cell phenotypes, microenvironment interactions, and tissue architecture [21]. The workflow of histoCAT consists of overlaying the segmentation masks previously generated by CP to extract single-cell level information.

To evaluate the diversity of phenotypic profiles collected from drug-treated MCF-7 cells, we then selected and generated Z-score normalized mean intensities of pHistone 3^{S28} and pH2AX S^{139} as drug-sensitive protein markers. Spatial heat maps of both channels were displayed for various drug-treated MCF-7 ROIs using a 99th percentile cutoff (Fig. 5B,C, Fig. S2B). This

quantitative and visual approach allows us to compare at the same scale the spatial distribution of these parameters between ROIs. For example, pH2A.X S^{139} nuclear expression level in etoposide-treated MCF-7 cells is higher than in nocodazole or EGF-treated cells on average. Mitotic cell marker pHistone 3^{S28} expression level in nocodazole-treated MCF-7 is twofold higher than that identified for other drugs. Related single-cell Z-score mean intensity values for nuclear markers (pHistone 3^{S28} , pH2A.X S^{139} , Ki-67, p4EBP1 T^{37}/T^{46} , p53, cyclin D3), surface markers (CD98, CD81, CD29, CD49e, CD47), and cell size parameters (area, perimeter, major and minor axis lengths) were exported from each drug treatment dataset to assess statistically significant changes of protein expression levels and cellular morphology (Fig. 5D and Fig. S4). We can see differences between both controls and drug-treated samples in the expression of nuclear markers, with a coincrease of pH2A.X S^{139} and Ki-67 levels, and phosphorylation of the translation repression protein 4E-BP1 under etoposide treatment. This suggests the presence of a nonquiescent population of MCF-7 undergoing DNA repair and maintaining high proliferative activity. For

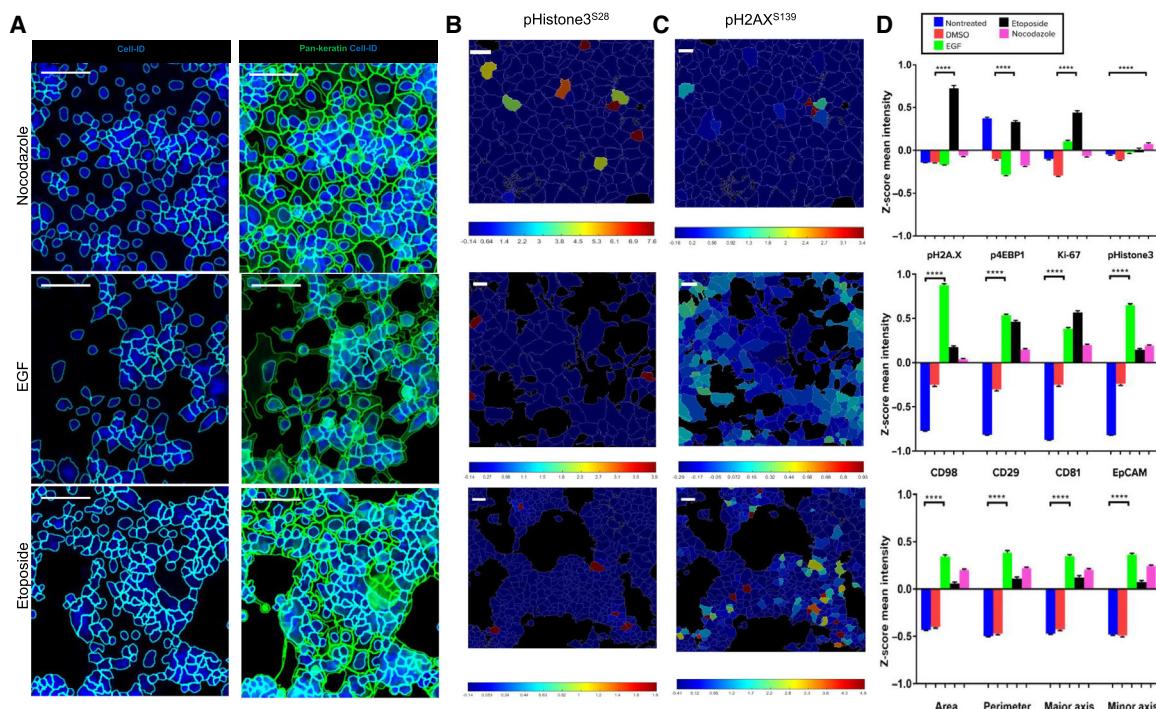


Fig. 5. Image processing by cell-based segmentation and spatial parametric distribution across ROIs of drug-treated MCF-7 cells. (A) Cell and nuclear masks overlaid on color IMC images. Scale bar = 100 μ m. Zoom-in areas of selected parameters visualized by Z-score normalized color heat maps for (B) pHistone 3^{S28} (C) pH2AX S^{139} . Scale bar = 10 μ m. (D) Z-score mean intensity levels of nuclear markers (top), surface markers (middle), and cellular size parameters (bottom) for control and drug treatment conditions. Data are presented as mean \pm SEM. ****P < 0.0001 (unpaired t-test).

nocodazole and EGF treatments, low Ki-67 levels may be related to the disruption of cell cycle and lineage shift induced in response to respective compound [44]. EGF treatment shows increased cell size and expression of adhesion surface markers compared to other conditions. These quantitative results are in agreement with multicolor images of treated cells.

Finally, the heterogeneity of MCF-7 cell phenotypes was analyzed using the t-distributed stochastic neighbor embedding (t-SNE) algorithm, a data dimensionality reduction method based on the similarity of selected markers or cell types [45]. We processed replicates of ROI for each compound using Z-score-normalized protein markers and cell size parameters to generate two-dimensional t-SNE plots. This visualization tool shows signal distribution of phosphoproteins, size parameters, and nuclear and cell surface membrane markers over the different ROIs (Fig. 6A and Fig. S5). For example, the population of mitotic cells was grouped in a single region on the map with high expression levels of p4EBP1^{T37/T46}, pHistone3^{S28}, and Ki-67. The high level of pH2A.X^{S139} was detected in the region of etoposide-treated cells and correlates with the strong Ir-intercalator signal due to compaction of cellular DNA in response to the drug. Area plot reveals the presence of larger size cells in EGF and nocodazole treatments, in line with EGF-induced lineage shift and cytokinesis inhibition by nocodazole (Fig. 6B).

Classification of drug-treated MCF-7 phenotypic changes by support vector machine learning

Here, we applied the support vector machine-learning classification ruler termed Fast Gentle Boosting for accurate IMC dataset multiparametric analysis of nuclear states of compound-treated MCF-7 cells [29]. This classifier was integrated into CELLPROFILER ANALYST (CPA) and trained [30] to identify five different nuclear classes of MCF-7 cells: pHistone3^{S28}-positive mitotic cells, pH2A.X^{S139} for cells undergoing DNA repair, p53 as tumor suppressor activation, cyclin D3 for G1/S transition in interphase, and Ki-67 for proliferating cells. A composite image of all nuclear markers is shown in Fig. 7A. These single-cellular objects were then scored and classified for a supervised analysis of every protein marker and cell size parameter per specified nuclear subpopulation. Accuracy of image classification on each dataset of MCF-7 drug compound was evaluated by cross-validation between the true labeled cells as classification training set and the image full scoring as predicted label set for each nuclear class. Classification reports were displayed as confusion

matrices where each value corresponds to average cross-validation metrics between the trained (true label) and scored (predicted label) single-cell objects for each nuclear class on a performance scale from 0 to 1 (Fig. 7B and Fig. S6). As shown, classification of each nuclear class per drug compound dataset shows relatively good performance between 0.5 and 1.

Raw mean intensity values of protein expression for each of the five nuclear markers were exported per classified MCF-7 population and compared to evaluate statistical significance (Fig. 7C and Fig. S7). Bar charts show various levels of protein markers among the five predicted classes of nuclear states, with higher protein detection in the corresponding class, except for the Ki-67 proliferative marker, which shows stronger detection in pHistone3^{S28} mitotic cells.

Multiparametric MCF-7 drug compound data prelabeled by nuclear classification were reprocessed using hierarchical clustering similarity heat maps. First, we computed pairwise similarities for each class of cells per drug exposure with Pearson correlation coefficient between each measured parameter. This method is applicable to non-normalized data and can be used when data measurements vary between samples. Similarity Pearson correlation heat maps were computed using the browser-based tool Morpheus [31]. The core interface in Morpheus is a graphical color heat map representing protein marker and size parameter measurements applied to each single-cell object. Determination of the Pearson correlation coefficient between each measured parameter generates a reorganized heat map of multiparametric similarity relationships for each classified MCF-7 drug dataset [46]. Using these hierarchical similarity clustering heat maps, we compared the grouping pattern of mitotic pHistone3^{S28}-positive cells across the multiple ROIs of compound-treated MCF-7 (Fig. 8 and Fig. S8). Naturally occurring isotopes of iridium, ¹⁹¹Ir and ¹⁹³Ir, are present in the nuclear staining reagent Cell-ID Intercalator-Ir and are grouped consistently across all heat maps. The two cytoskeleton biomarkers pan-keratin and cytokeratin 19 strongly correlate as well, since anti-pan-keratin and anti-CK19 antibodies identify cytokeratin isoforms highly expressed in MCF-7. Size parameters including perimeter, area, and axis length of each mitotic cell are grouped together after hierarchical clustering across all drug treatments.

Cell surface protein markers (CD29, CD98, CD81, CD47, EpCAM, and CD49e) generated a cluster with high correlation and redundancy in the pHistone3^{S28} mitotic cell subclass. However, EGF treatment influenced expression of these markers, and some level of dissimilarity occurred between the surface membrane

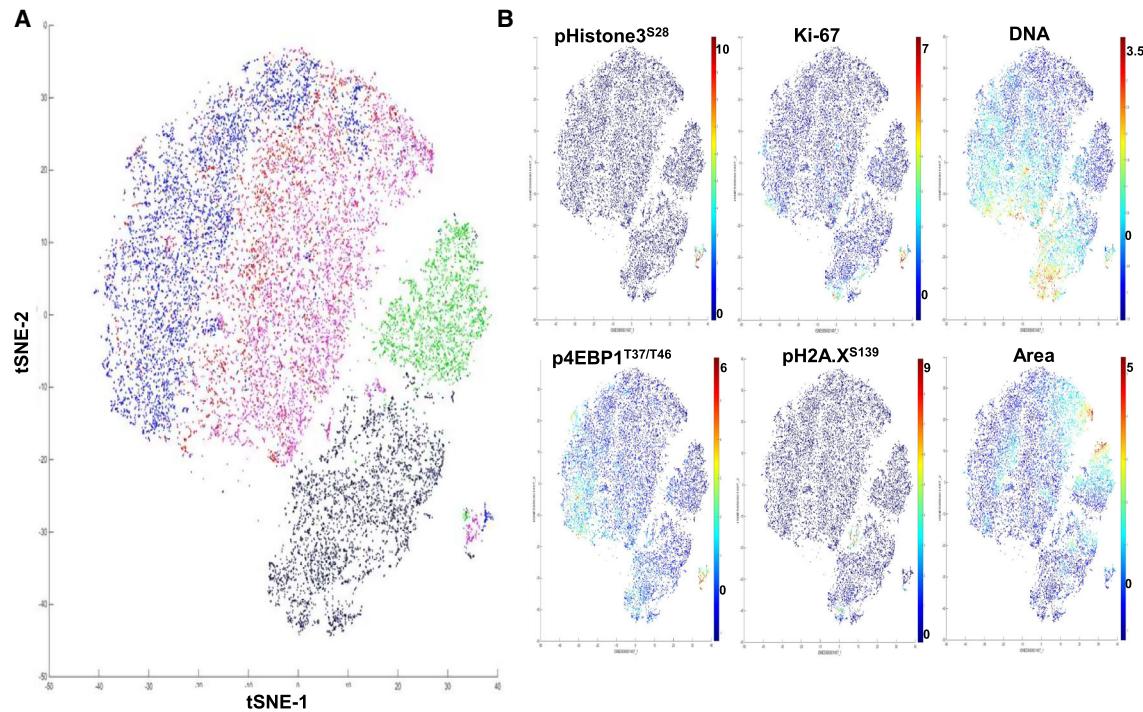


Fig. 6. Unsupervised cluster identification by t-SNE of control and compound-treated MCF-7. (A) Combined colored t-SNE maps of individual cells (EGF, green; etoposide, black; nocodazole, purple), DMSO (red), and nontreated control (blue). (B) t-SNE maps colored according to the expression level of nuclear markers and cell size (area).

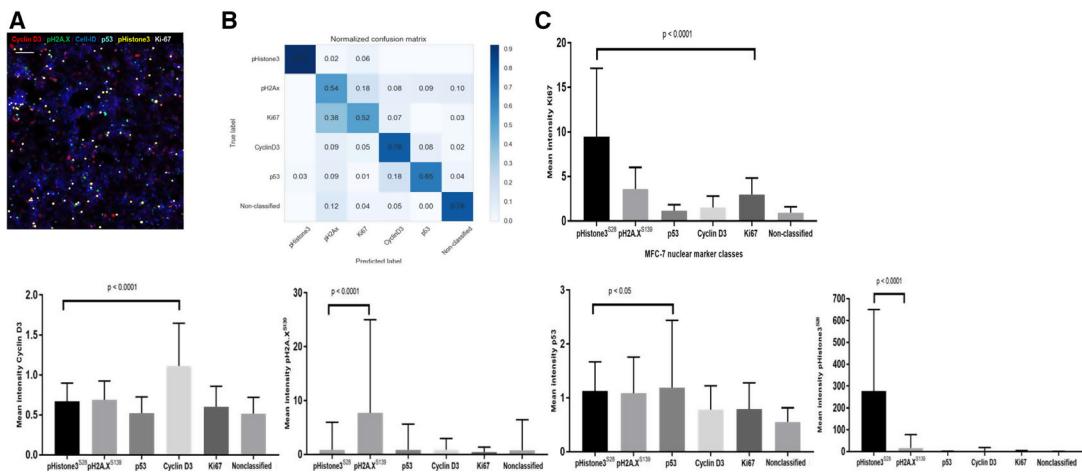


Fig. 7. Support vector machine-learning classification and multiparametric hierarchical similarity clustering of drug-treated MCF-7. (A) 6-pixel color image of nocodazole-treated cells generated by CPA Image Viewer. Scale bar = 160 μ m. (B) Normalized confusion matrix of six different cell nuclear phenotypes. (C) Protein expression levels of nuclear markers used for cell classification of MCF-7 treated with nocodazole. Graphs represent the non-normalized mean intensity (\pm SD) of nuclear proteins within each assigned category. $P < 0.05$ and $P < 0.0001$ (unpaired t -test) are included.

cluster and two integrin-mediated regulators, CD98 and CD47. This is consistent with the fact that MCF-7 cells switch to a more mesenchymal phenotype after EGF treatment and higher motility, with a

concomitant increase in the expression level of some adhesion proteins.

The intracellular and nuclear biomarkers pHistone3^{S28}, p4EBP1^{T37/T46}, and Ki-67 for mitotic cells

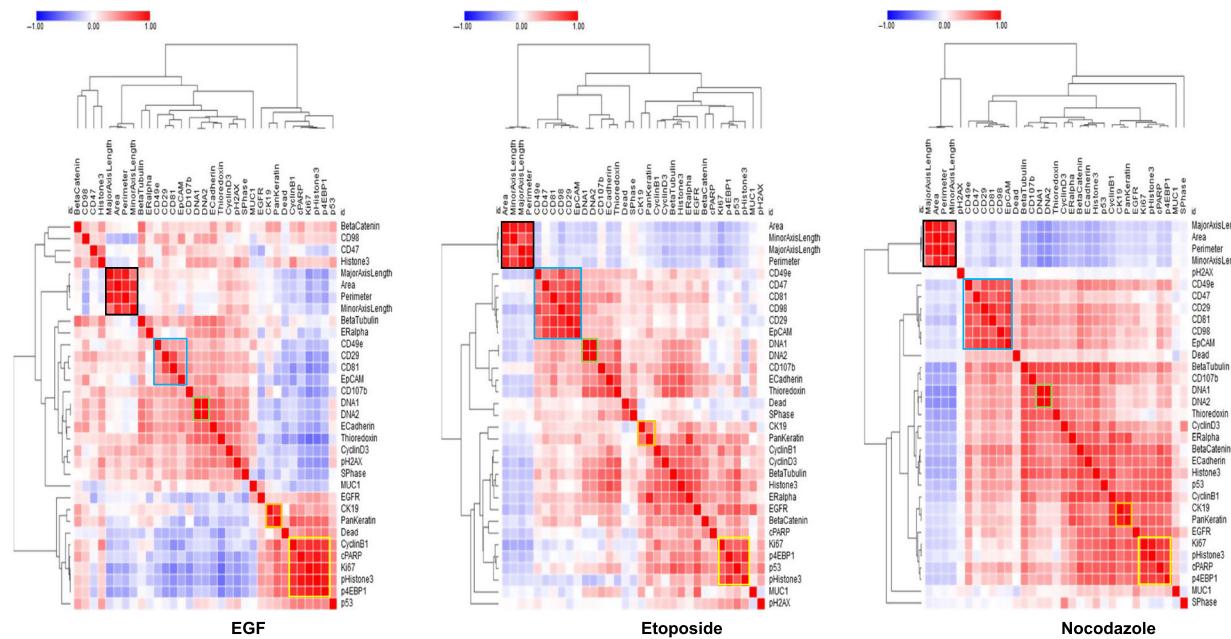


Fig. 8. Hierarchical clustering heat maps of Pearson coefficients of multiple parameters for drug-treated pHistone3^{S28}-positive cells. Clusters of cells are highlighted according to cells size parameters (black), CD surface markers (blue), Cell-ID Intercalator-Ir isotopes (green), pan-keratin and CK19 markers (orange), and nuclear markers (yellow).

cluster together in all conditions, showing a specific phenotypic profile. By looking at the correlation between the mitotic marker pHistone3^{S28} and other markers, we see that p53, cyclin B1, and cleaved poly (ADP-ribose) (cPARP) show variations in their correlation similarities depending on the drug treatment. p53 and cPARP exhibit close relationships with other nuclear markers, since both bind to nuclear DNA to be functional. Cyclin B1 shows strong correlation with pHistone3^{S28} in cells from nontreated control and EGF-treated ROIs. This correlation is less pronounced in the case of etoposide and nocodazole.

The clustering of different types of markers in the class of pHistone3^{S28} mitotic cells was measured by comparing the similarity level of each target-specific cluster across all controls and drug treatments. First, we reprocessed cell surface membrane CD markers for each treatment as a new hierarchical similarity heatmap by using Pearson correlation as distance metric and complete linkage method (Fig. 9A). As expected, each individual cluster is aligned and identified around the matrix diagonal component. Both nontreated and DMSO controls group of protein markers show close similarities to each other, and low or negative correlation with drug-treated markers. We used a hierarchical visualization edge bundle circle plot diagram to highlight the positive intra- and intersimilarities across

surface membrane protein parameters (Fig. 9B). Visualization of the high pairwise correlations within each cluster is shown as edge connecting nodes, and bundled lines correspond to close inter-relationships between both control conditions. Drug-treated clusters of proteins show poor or no interdependency. A similar visual workflow was applied to the multivariate analysis by pairwise Pearson correlation of specific nuclear markers and cytoplasmic markers types (Fig. S9).

Discussion

The goal of this study was to demonstrate a multidimensional IMC workflow and quantitative data processing using open source computational tools for *in vitro* drug response research. The multiplexed measurements acquired from a single image may provide a more informative and reliable screening of drug leads for further clinical development [47]. In this work, phenotypic features of tumor-derived cell lines such as SKBR3, HCC1143, and MCF-7, which recapitulate different subtypes found in breast tumors, were characterized. The wild-type p53-expressing MCF-7 epithelial cell line, a model of luminal breast adenocarcinoma, was selected for *in vitro* drug treatment and exposed for 48 h to three classes of bioactive compounds: EGF,

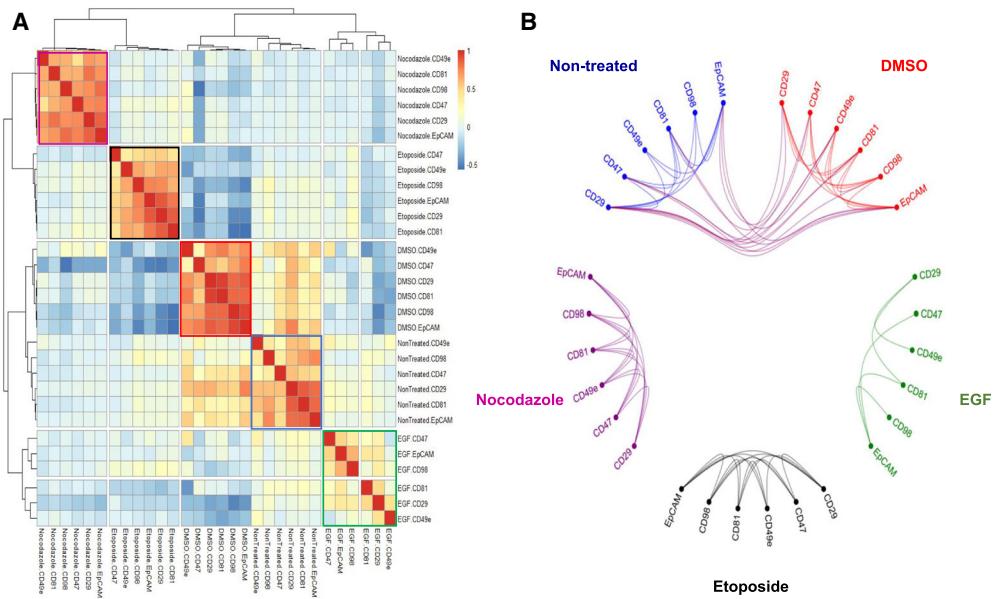


Fig. 9. Comparative analysis of multiparametric pairwise correlations between surface membrane markers expressed by MCF-7 pHistone3^{S28}-positive mitotic cells across controls and drug treatments. (A) Hierarchical similarity heatmap of CD98, CD29, EpCAM, CD81, CD47, and CD49e protein markers. Individual clusters are highlighted around the diagonal component for each condition (nontreated, blue; DMSO, red; EGF, green; etoposide, black; nocodazole, purple). (B) Hierarchical edge bundle visual graph showing connecting edges between parametric nodes. Bundled lines between nodes correspond to positive pairwise correlations with a cutoff value > 0.3. The bundled interconnections show similarities of surface membrane phenotyping profiles between nontreated and DMSO controls.

nocodazole, and etoposide. The cells were stained with a mixture of metal-tagged antibodies against cell membrane, cytoplasmic, and nuclei markers and subjected to IMC. The panel of metal tag antibodies used in our study was designed to minimize any potential spill over cross-interference between neighbor isotopic mass channels [48,49]. When analyzing our IMC data, we did not encounter strong and significant channel contamination from one metal-labeled antibody to another. The collected data allowed us to develop a workflow for cellular imaging analysis, which relies on established computational methods [50]. Multiplexed images of cell compartments give a detailed and broad visualization of biomarker localization and cell size parameter changes. Cell-based morphological segmentation of generated digital images is a prerequisite to identify and export multiple features of interest at single-cell resolution. Exploratory downstream analysis was performed with the histoCAT platform to generate, normalize, measure, and compare multiple images according to their multidimensional content in the form of spatial distribution heat maps, dimensionality reduction, and unsupervised clustering. In this report, we also describe cell classification by machine learning to prelabel cells before measuring their multivariate similarities. Prior knowledge of different cellular states is

important for the calculation of distance and similarity metrics between each marker in the evaluation of the heterogeneity of cells and the effects of different compounds on cancer cells [22]. In the example of pHistone3^{S28}-positive mitotic cells, we generated a multivariate high-dimensional heat map that allows a comparative study of mechanisms of action of multiple drugs. Additional nuclear classes such as pH2A.X^{S139}, p53, cyclin D3, and Ki-67 were analyzed using similar multiparametric clustering approaches (data not shown). The strategy could enable new ways to perform deep proteomic drug screening on different models for translational medicine and preclinical trials [51,52].

Other *in vitro* assays using stem cells, or cocultures of different cell types, could be studied by IMC as an additional approach to designing high-content analysis to investigate cell-to-cell interactions and signaling. A recent targeted therapy against cancer oncogenes described a cytotoxic small-molecule compound that activated the steroid receptor signaling pathways and led to cell stress and death [53]. IMC may assist researchers to better understand perturbations in this and other pathways through high-dimensional analysis. Three-dimensional tumor spheroids grown *in vitro* are more representative of the complexity of cancer

Table 1. Key similarities and differences of biological samples processing used for IMC

Sample type	Tissue sections (Frozen and FFPE)	Cell smear on microscopic slide (Liquid biopsy)	Adherent cell lines grown in chamber slides
Sample quality and integrity criteria	Thin thickness of the cut tissue slice (5 µm) Effectiveness of tissue fixation for FFPE (temperature, fixing chemicals, pH, time) Storage conditions for tissue preservation (temperature, humidity, light exposure)	Cell monolayer smear on slide by surface tension Fixation method Low-temperature storage conditions Preidentification of cells by immunofluorescence staining and scanning	High cell viability and proliferation rate Adherence efficiency of growing cells to the surface of the slide Optimal seeding density on slide and growth conditions per cell line tested
Sample staining critical steps	Optimization of dewaxing and hydration procedures for FFPE tissue Removal of tissue freezing matrix for frozen tissue Antigen retrieval and blocking conditions Metal isotope labeling of validated antibody clones for tissue staining Titration range testing of metal tag antibodies and DNA metallo-intercalators	Use of hydrophobic circle barrier around cells in specific regions for liquid staining on slide Blocking conditions Preparation of metal tag antibodies liquid cocktail Washing and drying of the sample slide before acquisition	Metal tag antibodies panel specific to surface membrane, cytoplasm, and nucleus of adherent cells Chamber well liquid volume measurements for cell staining procedure Fixation and permeabilization conditions Removal of chamber wells from the slide
Sample acquisition by IMC	Simultaneous acquisition of all markers Whole ablation of tissue regions No autofluorescence	Plotting of two-dimensional coordinates of cells preidentified by immunofluorescence Laser ablation of areas with good cellular density and morphology	Regions of interest with intact cells Ablation of areas with even cell monolayers rather than densely packed overlapping cells
Data analysis challenges	Pathology expertise required for IMC imaging data visualization Annotation of structural and cellular regions Segmentation of individual and distinct cells within tissue samples requires nuclei and membrane markers	Single-cell segmentation of small-sized cells Identification of nuclei with DNA metal intercalator Use of membrane stain for detection of cellular boundaries Classification of cellular subpopulations Identification of circulating tumor metastatic cells	Single-cell morphological segmentation analysis Nuclei segmentation with DNA metal intercalator Whole-cell detection using cytoplasmic marker Classification of different cell states
Multiplexing biological applications	Immuno-oncology Solid tumor oncology Quantitative pathology	Hematology Immunology	Drug discovery Cell lines authentication and characterization Single-cell systems biology

tissue than two-dimensional cell cultures and can be analyzed by IMC through the preparation of sections of FFPE pelleted spheroids [54]. We summarized in Table 1 several advantages and challenges during preparation on slide, staining procedure, acquisition by IMC, and software data analysis for three different models of sample. IMC was useful in the biodistribution study of the platinum-containing anticancer drug cisplatin in pancreatic adenocarcinoma xenograft model and could be applied to the emerging metal-containing chemotherapies and photodynamic therapies, which are considered as alternative approaches to classic compounds [55,56]. Therefore, the goal for IMC in preclinical profiling of metallodrugs will be to

efficiently and effectively identify cellular responses of *in vitro* tumor-derived cell lines for lead optimization in target-based drug discovery [57]. A comparison of IMC performance metrics to other high-content imaging system technologies in the field of single-cell analysis is shown in Table 2.

In conclusion, we demonstrate a comprehensive image analysis workflow for cell-based IMC data analysis of surface and intracellular markers in drug-treated cells. The correlation coefficient for pairs of multiple parameters and the similar distances among a collection of treatment profiles facilitate downstream analysis and allow for direct data visualization. Image-based cell profiling studies compute statistical estimates of the

Table 2. Performance comparison of IMC to other high-content single-cell imaging technologies

High-content imaging technology	Imaging Mass Cytometry	Imaging Flow Cytometry	Imaging Mass Spectrometry (SIMS MIBI, MALDI)	Single-cell fluorescence microscopy	Super-resolution microscopy
Reagents used	Metal-tagged antibodies DNA/metal intercalators	Fluorescent labeled antibodies DNA dyes	Metal-tagged antibodies (SIMS MIBI) Label free, matrix solutions (MALDI)	Fluorescent labeled antibodies DNA dyes	Fluorescent probes
Sample format	Chamber slide (2-, 4-, 8-, 16-well format) Tissue on slide Cell smear on slide	Multi-well plates Tubes	Tissue on slide Cell smear on slide	Multi-well plates (96-, 384-, 1536-well format)	Single coverslip Multi-well plates
Sample preparation time	6 h	6 h	6 h	6 h per staining cycle	6 h
Spatial resolution	1 μm	1 μm	< 1 μm (SIMS MIBI) 5–200 μm (MALDI)	1–100 μm	< 1 μm
Throughput time	200 pixels per second	1000–5000 cells per second	10 000 pixels per second (SIMS MIBI)	100–300 cells per second	10 cells per hour
Number of cells analyzed	10^3 – 10^4	10^4 – 10^5	100 pixels per second (MALDI) 10^3 – 10^4 (SIMS MIBI)	10^3 – 10^5	10^{-10^2}
Sample types	Fixed adherent and suspension cells FFPE and frozen tissue sections > 37 markers	Cells in suspension	Fixed adherent and suspension cells FFPE and frozen tissue sections 12 markers	Live and fixed cells	Live and fixed cells
Number of targets			7 markers per scan (SIMS MIBI)	6 markers per scan	4 markers per scan
References	[17,20]	[51]	0–100 000 Dalton <i>m/z</i> spectra range of molecules (MALDI) [58,59]	[14,59]	[59,60]

likelihood of equivalence between two drug-induced profiles. Similarity measurements quantify proximity between profiles, since they detect deviations from one sample to another regardless of the absolute magnitude. This procedure is useful in finding relations and groups of samples that share common properties in high-content screening. Predicted pharmacodynamic effects were visualized and quantified in MCF-7 cells dosed with three target-specific compounds. Strong pairwise correlation between nuclear markers pHistone $^3S^{28}$, Ki-67, and p4E-BP1 T^{37}/T^{46} in mitotic cells and anticorrelation with cell surface markers CD29, CD98, CD81, CD47, and EpCAM was demonstrated.

Acknowledgements

The authors thank the employees of Fluidigm Canada Inc. (Markham, Ontario) for their support of this research. This research received no specific grant from any funding agency.

Conflict of interest

AB, AE, and OO are employees of, and receive remuneration from, Fluidigm Corporation. Fluidigm, Cell-ID, CyTOF, Imaging Mass Cytometry, and IMC are trademarks and/or registered trademarks of Fluidigm Corporation in the United States and/or other countries. All other trademarks are the sole property of their respective owners. For research use only. Not for use in diagnostic procedures.

Author contributions

AB designed the study and drafted the manuscript. AB, AE, and OO took part to the manuscript writing. All authors have read and approved the manuscript.

References

- Masters JR (2000) Human cancer cell lines: fact and fantasy. *Nat Rev Mol Cell Biol* **1**, 233–236.
- An WF and Tolliday N (2010) Cell-based assays for high-throughput screening. *Mol Biotechnol* **45**, 180–186.
- Wilding JL and Bodmer WF (2014) Cancer cell lines for drug discovery and development. *Cancer Res* **74**, 2377–2384.
- Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves E, Barthorpe S, Lightfoot H *et al.* (2016) A landscape of pharmacogenomic interactions in cancer. *Cell* **166**, 740–754.
- Haverty PM, Lin E, Tan J, Yu Y, Lam B, Lianoglou S, Neve RM, Martin S, Settleman J, Yauch RL *et al.* (2016) Reproducible pharmacogenomic profiling of cancer cell line panels. *Nature* **533**, 333–337.
- Goodspeed A, Heiser LM, Gray JW and Costello JC (2016) Tumor-derived cell lines as molecular models of cancer pharmacogenomics. *Mol Cancer Res* **14**, 3–13.
- Shoemaker RH (2006) The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* **6**, 813–823.
- Sharma SV, Haber DA and Settleman J (2010) Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. *Nat Rev Cancer* **10**, 241–253.
- Lang P, Yeow K, Nichols A and Scheer A (2006) Cellular imaging in drug discovery. *Nat Rev Drug Discov* **5**, 343–356.
- Kang J, Hsu CH, Wu Q, Liu S, Coster AD, Posner BA, Altschuler SJ and Wu LF (2016) Improving drug discovery with high-content phenotypic screens by systematic selection of reporter cell lines. *Nat Biotechnol* **34**, 70–77.
- Gustafsdottir SM, Ljosa V, Sokolnicki KL, Anthony Wilson J, Walpita D, Kemp MM, Petri Seiler K, Carrel HA, Golub TR, Schreiber SL *et al.* (2013) Multiplex cytological profiling assay to measure diverse cellular states. *PLoS One* **8**, e80999.
- Wu PH, Phillip JM, Khatau SB, Chen WC, Stirman J, Rosseel S, Tschudi K, Van Patten J, Wong M, Gupta S *et al.* (2015) Evolution of cellular morpho-phenotypes in cancer metastasis. *Sci Rep* **5**, 18437.
- Boutros M, Heigwer F and Laufer C (2015) Microscopy-based high-content screening. *Cell* **163**, 1314–1325.
- Lin JR, Izar B, Wang S, Yapp C, Mei S, Shah PM, Santagata S and Sorger PK (2018) Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife* **7**, e31657.
- Wawer MJ, Li K, Gustafsdottir SM, Ljosa V, Bodycombe NE, Marton MA, Sokolnicki KL, Bray MA, Kemp MM, Winchester E *et al.* (2014) Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. *Proc Natl Acad Sci USA* **111**, 10911–10916.
- Wagner BK and Schreiber SL (2016) The power of sophisticated phenotypic screening and modern mechanism-of-action methods. *Cell Chem Biol* **23**, 3–9.
- Chang Q, Ornatsky OI, Siddiqui I, Loboda A, Baranov VI and Hedley DW (2017) Imaging mass cytometry. *Cytometry A* **91**, 160–169.
- Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick JE and Tanner SD (2009) Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* **81**, 6813–6822.

- 19 Lou X, Zhang G, Herrera I, Kinach R, Ornatsky O, Baranov V, Nitz M and Winnik MA (2007) Polymer-based elemental tags for sensitive bioassays. *Angew Chem Int Ed Engl* **46**, 6111–6114.
- 20 Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, Schüffler PJ, Grolimund D, Buhmann JM, Brandt S *et al.* (2014) Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* **11**, 417–422.
- 21 Schapiro D, Jackson HW, Raghuraman S, Fischer JR, Zanotelli VRT, Schulz D, Giesen C, Catena R, Varga Z and Bodenmiller B (2017) histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat Methods* **14**, 873–876.
- 22 Ljosa V, Caie PD, Ter Horst R, Sokolnicki KL, Jenkins EL, Daya S, Roberts ME, Jones TR, Singh S, Genovesio A *et al.* (2013) Comparison of methods for image-based profiling of cellular morphological responses to small-molecule treatment. *J Biomol Screen* **18**, 1321–1329.
- 23 Bougen-Zhukov N, Loh SY, Lee HK and Loo LH (2017) Large-scale image-based screening and profiling of cellular phenotypes. *Cytometry A* **91**, 115–125.
- 24 Blajeski AL, Phan VA, Kottke TJ and Kaufmann SH (2002) G(1) and G(2) cell-cycle arrest following microtubule depolymerization in human breast cancer cells. *J Clin Invest* **110**, 91–99.
- 25 Osborne CK, Hamilton B, Titus G and Livingston RB (1980) Epidermal growth factor stimulation of human breast cancer cells in culture. *Cancer Res* **40**, 2361–2366.
- 26 Hau A, Ceppi P and Peter ME (2012) CD95 is part of a let-7/p53/miR-34 regulatory network. *PLoS One* **7**, e49636.
- 27 Schulz D, Zanotelli VRT, Fischer JR, Schapiro D, Engler S, Lun XK, Jackson HW and Bodenmiller B (2018) Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry. *Cell Syst* **6**, 25–36.
- 28 Bray MA, Vokes MS and Carpenter AE (2015) Using cell profiler for automatic identification and measurement of biological objects in images. *Curr Protoc Mol Biol* **109**, 1–13.
- 29 Jones TR, Carpenter AE, Lamprecht MR, Moffat J, Silver SJ, Grenier JK, Castoreno AB, Eggert US, Root DE, Golland P *et al.* (2009) Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc Natl Acad Sci USA* **6**, 1826–1831.
- 30 Dao D, Fraser AN, Hung J, Ljosa V, Singh S and Carpenter AE (2016) Cell profiler analyst: interactive data exploration, analysis and classification of large biological image sets. *Bioinformatics* **32**, 3210–3212.
- 31 Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK *et al.* (2017) A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437–1452.
- 32 Holten D (2006) Hierarchical edge bundles: visualization of adjacency relations in hierarchical data. *IEEE Trans Vis Comput Graph* **12**, 741–748.
- 33 Swift ML (1997) GraphPad Prism, data analysis, and scientific graphing. *J Chem Inf Comput Sci* **37**, 411–412.
- 34 Henjes F, Bender C, von der Heyde S, Braun L, Mannsperger HA, Schmidt C, Wiemann S, Hasmann M, Aulmann S, Beissbarth T *et al.* (2012) Strong EGFR signaling in cell line models of ERBB2-amplified breast cancer attenuates response towards ERBB2-targeting drugs. *Oncogenesis* **1**, e16.
- 35 Holliday DL and Speirs V (2011) Choosing the right cell line for breast cancer research. *Breast Cancer Res* **13**, 215–224.
- 36 Prat A, Karginova O, Parker JS, Fan C, He X, Bixby L, Harrell JC, Roman E, Adamo B, Troester M *et al.* (2013) Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. *Breast Cancer Res Treat* **142**, 237–255.
- 37 Lacroix M and Leclercq G (2004) Relevance of breast cancer cell lines as models for breast tumors: an update. *Breast Cancer Res Treat* **83**, 249–289.
- 38 Caie PD, Walls RE, Ingleston-Orme A, Daya S, Houslay T, Eagle R, Roberts ME and Carragher NO (2010) High-content phenotypic profiling of drug response signatures across distinct cancer cells. *Mol Cancer Ther* **9**, 1913–1926.
- 39 Bafna S, Kaur S and Batra SK (2010) Membrane-bound mucins: the mechanistic basis for alterations in the growth and survival of cancer cells. *Oncogene* **29**, 2893–2904.
- 40 Kim SM and Hahn JH (2008) CD98 activation increases surface expression and clustering of beta1 integrins in MCF-7 cells through FAK/Src- and cytoskeleton-independent mechanisms. *Exp Mol Med* **40**, 261–270.
- 41 Garcia R, Franklin RA and McCubrey JA (2006) EGF induces cell motility and multi-drug resistance gene expression in breast cancer cells. *Cell Cycle* **23**, 2820–2826.
- 42 Ji J, Zhang Y, Redon CE, Reinhold WC, Chen AP, Fogli LK, Holbeck SL, Parchment RE, Hollingshead M, Tomaszewski JE *et al.* (2017) Phosphorylated fraction of H2AX as a measurement for DNA damage in cancer cells and potential applications of a novel assay. *PLoS One* **12**, e0171582.
- 43 Molnar C, Jermyn IH, Kato Z, Rahkama V, Östling P, Mikkonen P, Pietäinen V and Horvath P (2016) Accurate morphology preserving segmentation of overlapping cells based on active contours. *Sci Rep* **6**, 32412.

- 44 Sobecki M, Mrouj K, Colinge J, Gerbe F, Jay P, Krasinska L, Dulic V and Fisher D (2017) Cell-cycle regulation accounts for variability in Ki-67 expression levels. *Cancer Res* **10**, 2722–2734.
- 45 Van der Maaten LJP and Hinton GE (2008) Visualizing high-dimensional data using t-SNE. *J Mach Learn Res* **9**, 2579–2605.
- 46 Zhou JX, Isik Z, Xiao C, Rubin I, Kauffman SA, Schroeder M and Huang S (2016) Systematic drug perturbations on cancer cells reveal diverse exit paths from proliferative state. *Oncotarget* **7**, 7415–7425.
- 47 Moffat JG, Vincent F, Lee JA, Eder J and Prunotto M (2017) Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat Rev Drug Discov* **16**, 531–543.
- 48 Takahashi C, Au-Yeung A, Fuh F, Ramirez-Montagut T, Bolen C, Mathews W and O'Gorman WE (2017) Mass cytometry panel optimization through the designed distribution of signal interference. *Cytometry A* **91**, 39–47.
- 49 Chevrier S, Crowell HL, Zanotelli VRT, Engler S, Robinson MD and Bodenmiller B (2018) Compensation of signal spillover in suspension and imaging mass cytometry. *Cell Syst* **6**, 612–620.
- 50 Caicedo JC, Cooper S, Heigwer F, Warchal S, Qiu P, Molnar C, Vasilevich AS, Barry JD, Bansal HS, Kraus O et al. (2017) Data-analysis strategies for image-based cell profiling. *Nat Methods* **14**, 849–863.
- 51 Doan M, Vorobjev I, Rees P, Filby A, Wolkenhauer O, Goldfeld AE, Lieberman J, Barteneva N, Carpenter AE and Hennig H (2018) Diagnostic potential of imaging flow cytometry. *Trends Biotechnol* **36**, 649–652.
- 52 Rahmoune H and Guest PC (2017) Application of multiplex biomarker approaches to accelerate drug discovery and development. *Methods Mol Biol* **1546**, 3–17.
- 53 Wang L, Yu Y, Chow DC, Yan F, Hsu CC, Stossi F, Mancini MA, Palzkill T, Liao L, Zhou S et al. (2015) Characterization of a steroid receptor coactivator small molecule stimulator that overstimulates cancer cells and leads to cell stress and death cancer. *Cell* **28**, 240–252.
- 54 Stock K, Estrada MF, Vidic S, Gjerde K, Rudisch A, Santo VE, Barbier M, Blom S, Arundkar SC, Selvam I et al. (2016) Capturing tumor complexity in vitro: comparative analysis of 2D and 3D tumor models for drug discovery. *Sci Rep* **6**, 28951.
- 55 Chang Q, Ornatsky OI, Siddiqui I, Straus R, Baranov VI and Hedley DW (2016) Biodistribution of cisplatin revealed by imaging mass cytometry identifies extensive collagen binding in tumor and normal tissues. *Sci Rep* **6**, 36641.
- 56 Mjos KD and Orvig C (2014) Metallodrugs in medicinal inorganic chemistry. *Chem Rev* **114**, 4540–4563.
- 57 Gillet JP, Varma S and Gottesman MM (2013) The clinical relevance of cancer cell lines. *J Natl Cancer Inst* **105**, 452–458.
- 58 Angelo M, Bendall SC, Finck R, Hale MB, Hitzman C, Borowsky AD, Levenson RM, Lowe JB, Liu SD, Zhao S et al. (2014) Multiplexed ion beam imaging of human breast tumors. *Nat Med* **20**, 436–442.
- 59 Pomerantz AK, Sari-Sarraf F, Grove KJ, Pedro L, Rudewicz PJ, Fathman JW and Krucker T (2019) Enabling drug discovery and development through single-cell imaging. *Expert Opin Drug Discov* **14**, 115–125.
- 60 Beghin A, Kechkar A, Butler C, Levet F, Cabillic M, Rossier O, Giannone G, Galland R, Choquet D and Sibarita JB (2017) Localization-based super-resolution imaging meets high-content screening. *Nat Methods* **14**, 1184–1190.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. IMC multiplex image of MCF-7 non-treated and DMSO-treated control cells. Zoom-in colored area of 4-plex non-treated (A) and DMSO-treated control (B) with different combinations of biomarkers, 400 × 400 µm cropped size. Scale bar = 100 µm.

Fig. S2. Cell segmentation and single-marker distribution of MCF-7 non-treated and DMSO-treated controls. (A) CellProfiler overlays of nucleus and cell segmentation masks on color composite images (scale bar = 100 µm). (B) Spatial distribution heat map zoom-in areas of pHistone3^{S28}, Ki-67, pH2A.X^{S139} and cell area (scale bar = 10 µm).

Fig. S3. Segmentation accuracy of nuclei and cytoplasmic contours of identified MCF-7 cells drug-treated using binary image overlap. The identified nucleus and cytoplasm from each cell as object is converted to a binary format called test ‘image’ and overlaid to the binary converted tiff image used for segmentation (DNA-Ir for nuclei identification and pan-keratin for cell identification) as ‘ground truth’. Both binary images are overlapped to calculate statistics of the closeness from the test image to its ground truth. Performance segmentation statistics parameters is ranging on a scale from 0 to 1, where 1 means perfect overlap and 0 no overlap. Nuclei segmentation shows high precision (number of true positive pixels/(number of true positive pixels + number of false positive pixels) close to 1 over the different images. For cell segmentation the precision is slightly lower but still higher than 0.5 value.

Fig. S4. Comparison of average Z-score expression level of p53, cyclin D3, CD47, CD49e and Cell-ID Intercalator-Ir between each control and drug treatment condition. Data are presented as mean ± SEM. ***P < 0.0001 (unpaired t-test).

Fig. S5. Unsupervised cluster identification by t-SNE of control and compound-treated MCF-7. (a) Combined colored t-SNE maps of individual cells (EGF, green; etoposide, black; nocodazole, purple), DMSO (red) and non-treated control (blue). (b) t-SNE maps colored according to the expression level of nuclear markers and cell size (Area).

Fig. S6. Fast Gentle Boosting classification matrices of stitched replicates of MCF-7 ROIs per control and drug compound.

Fig. S7. Comparative expression levels of protein markers Ki-67, cyclin D3, pH2A.X^{S139}, p53 and pHistone3^{S28} (rows) per nuclear class and drug treatment condition (column) of MCF-7 cells. Bar charts represent non-normalized mean intensity (+SD) of each nuclear marker to its related class. $P < 0.05$ and $P < 0.0001$ (unpaired t-test) are included.

Fig. S8. Heat maps of Pearson correlation coefficients of multiple parameters for non-treated and DMSO-treated controls ROIs on mitotic pHistone3^{S28} cells. Cluster highlighting of cell size parameters (black), surface markers (blue), Cell-ID iridium isotopes (green), pan-keratin and CK19 (orange) and nucleus markers (yellow).

Fig. S9. Graphical representations of Pearson correlation coefficients for nuclei markers (A) and

cytoplasmic markers (B) across all controls and drug treatments in the classified population of mitotic pHistone3^{S28} MCF-7 cells. Hierarchical similarities heatmaps show all pairwise correlations values between each protein, with diagonal components highlighted as multiparametric clusters identified for each condition (Non-treated, blue; DMSO, red; EGF, green; etoposide, black; nocodazole, purple). Intra and inter-relationships between these clusters are shown as Hierarchical edgebundle visual graphs. Bundled lines connecting parametric protein nodes correspond to positive correlation values higher than 0.3.

Table S1. List of chemical compounds used for MCF-7 cell treatment, 48 h exposure.

Table S2. Metal-labeled antibody panel for cell surface markers.

Table S3. Metal-labeled antibody panel for intracellular markers.

Table S4. Mitotic index of compound-treated MCF-7 generated by support vector machine classification with Fast Gentle Boosting ruler. Population of pHistone3^{S28} of all ROIs replicates per drug treatment and control were counted and divided by the total number of cells detected to determine the mitotic index of each condition.

SUPPORTING INFORMATION

Multidimensional Profiling of Drug-Treated Cells by Imaging Mass Cytometry

Alexandre Bouzekri[†], Amanda Esch[†], Olga Ornatsky[†]

[†] Fluidigm Canada Inc., 1380 Rodick Road, Markham, Ontario, L3R 4G5, Canada

Supporting Figures

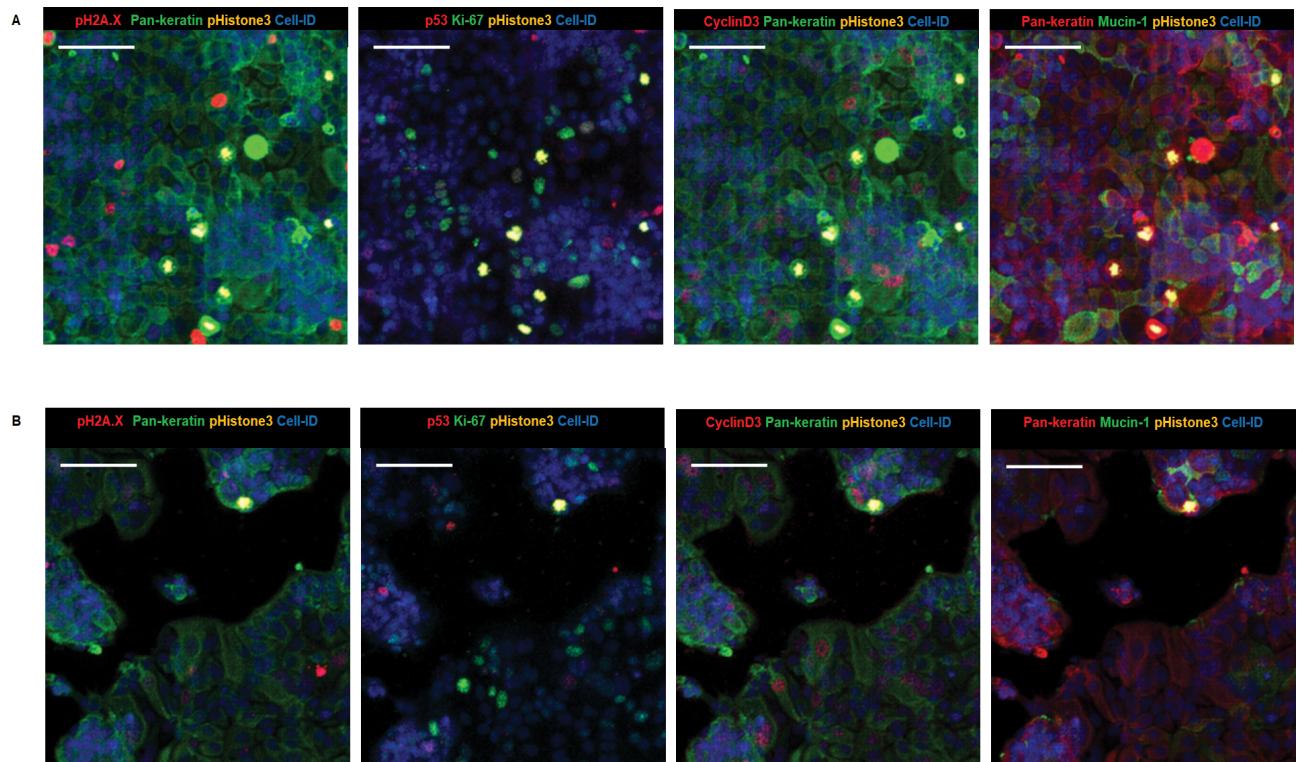


Figure S1. IMC multiplex image of MCF-7 non-treated and DMSO-treated control cells.

Zoom-in colored area of 4-plex non-treated (A) and DMSO-treated control (B) with different combinations of biomarkers, 400 x 400 μm cropped size. Scale bar = 100 μm .

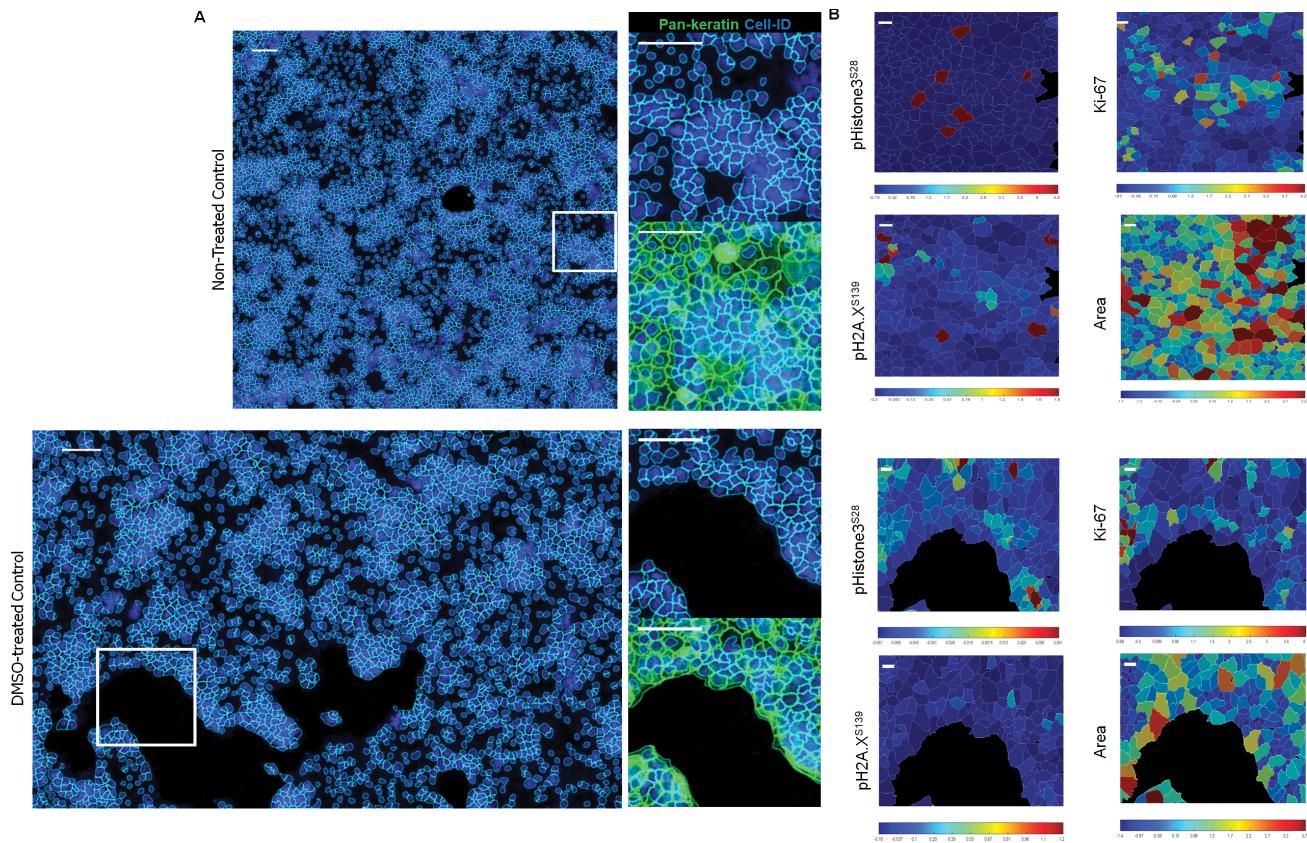


Figure S2. Cell segmentation and single-marker distribution of MCF-7 non-treated and DMSO-treated controls. (A) CellProfiler overlays of nucleus and cell segmentation masks on color composite images (scale bar = 100 μm). (B) Spatial distribution heat map zoom-in areas of pHistone3^{S28}, Ki-67, pH2AX^{S139} and cell area (scale bar = 10 μm).

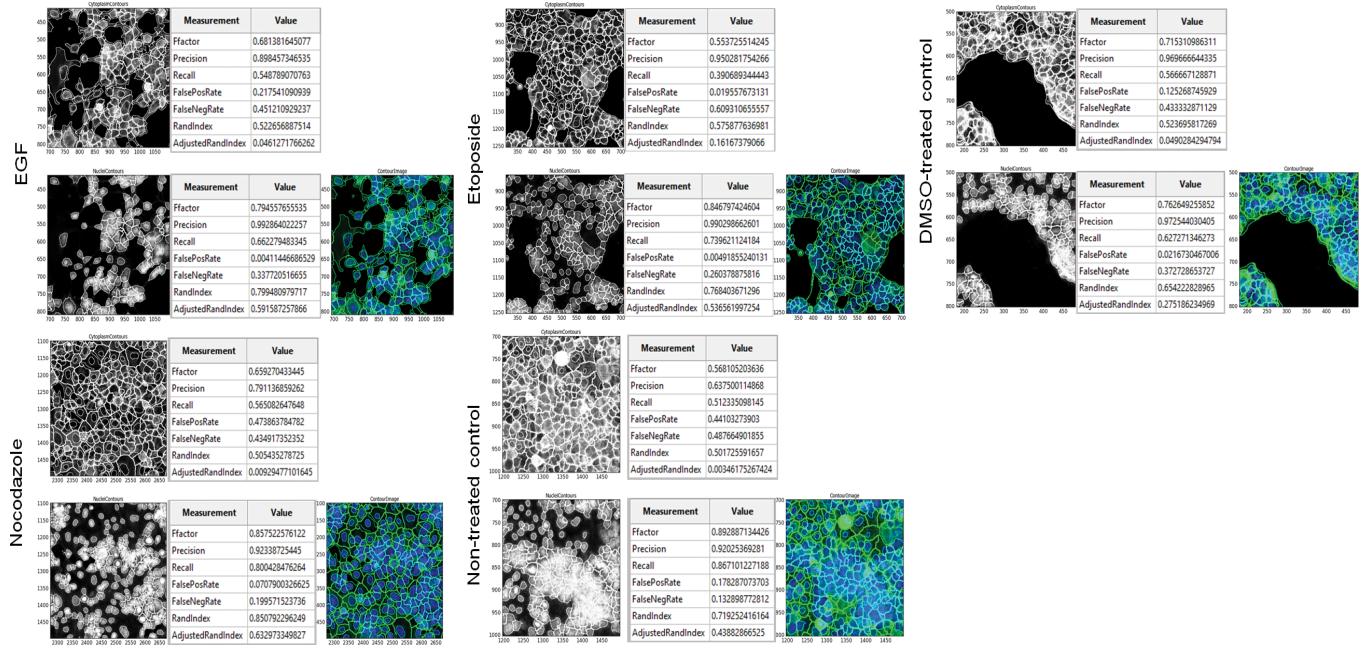


Figure S3. Segmentation accuracy of nuclei and cytoplasmic contours of identified MCF-7 cells drug-treated using binary image overlap. The identified nucleus and cytoplasm from each cell as object is converted to a binary format called test “image” and overlaid to the binary converted tiff image used for segmentation (DNA-Ir for nuclei identification and pan-keratin for cell identification) as “ground truth”. Both binary images are overlapped to calculate statistics of the closeness from the test image to its ground truth. Performance segmentation statistics parameters is ranging on a scale from 0 to 1, where 1 means perfect overlap and 0 no overlap. Nuclei segmentation shows high precision (number of true positive pixels / (number of true positive pixels + number of false positive pixels) close to 1 over the different images. For cell segmentation the precision is slightly lower but still higher than 0.5 value.

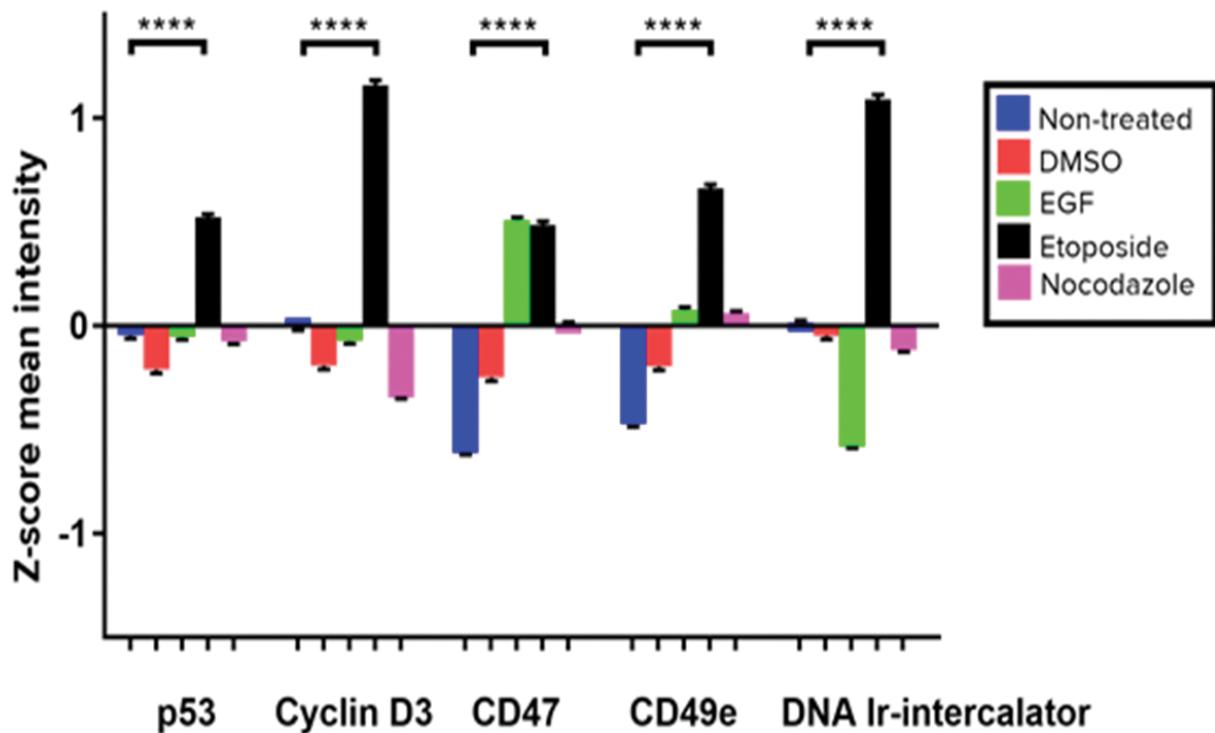


Figure S4. Comparison of average Z-score expression level of p53, cyclin D3, CD47, CD49e and Cell-ID Intercalator-Ir between each control and drug treatment condition. Data are presented as mean \pm SEM. **P <0.0001 (unpaired t-test).**

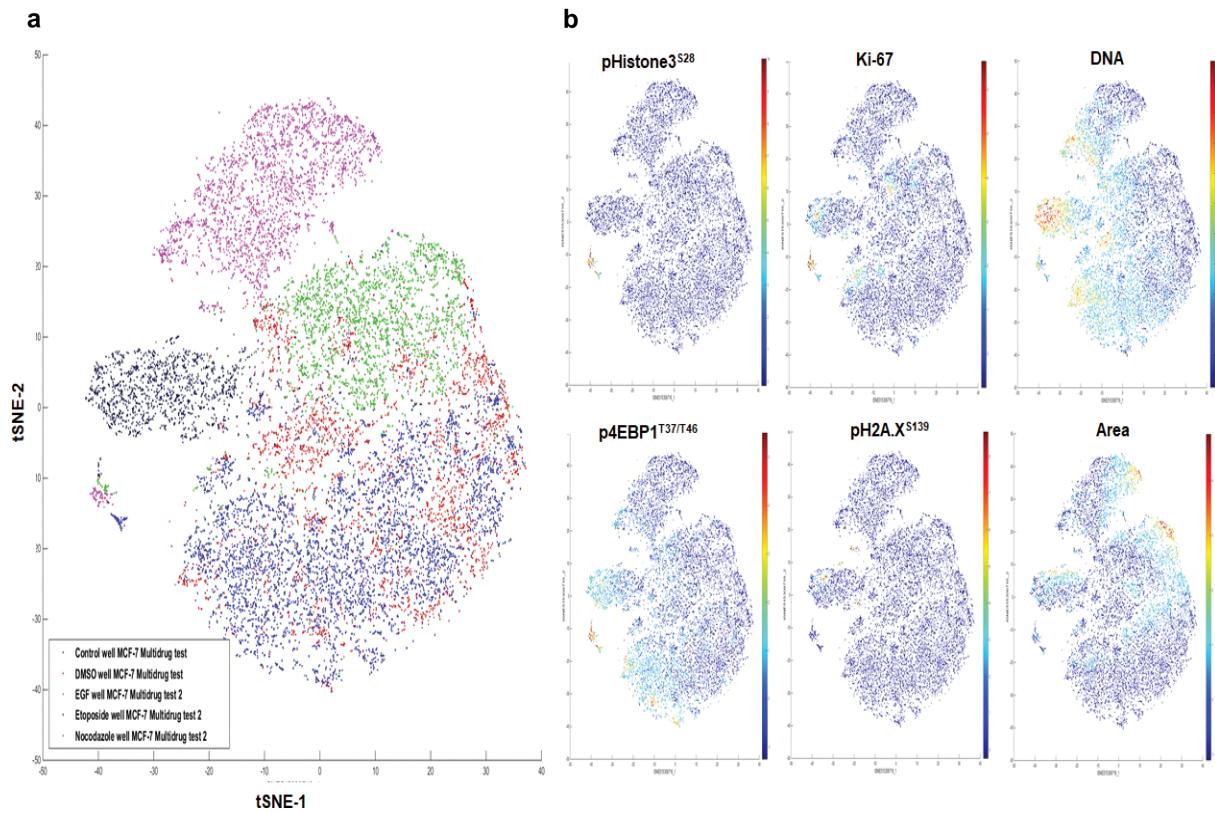


Figure S5. Unsupervised cluster identification by t-SNE of control and compound-treated MCF-7. **(a)** Combined colored t-SNE maps of individual cells (EGF, green; etoposide, black; nocodazole, purple), DMSO (red) and non-treated control (blue). **(b)** t-SNE maps colored according to the expression level of nuclear markers and cell size (Area).

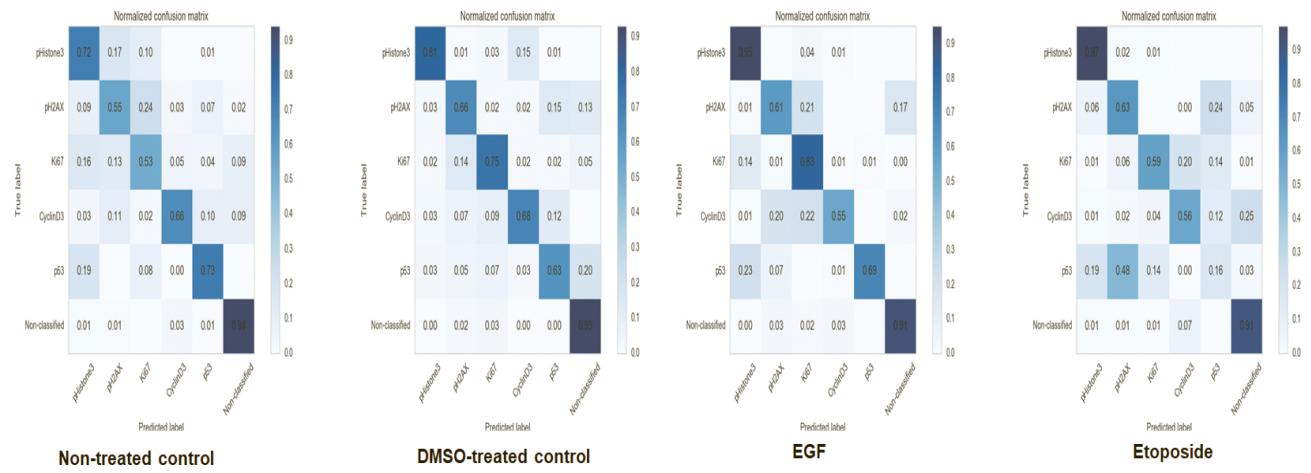


Figure S6. Fast Gentle Boosting classification matrices of stitched replicates of MCF-7 ROIs per control and drug compound.

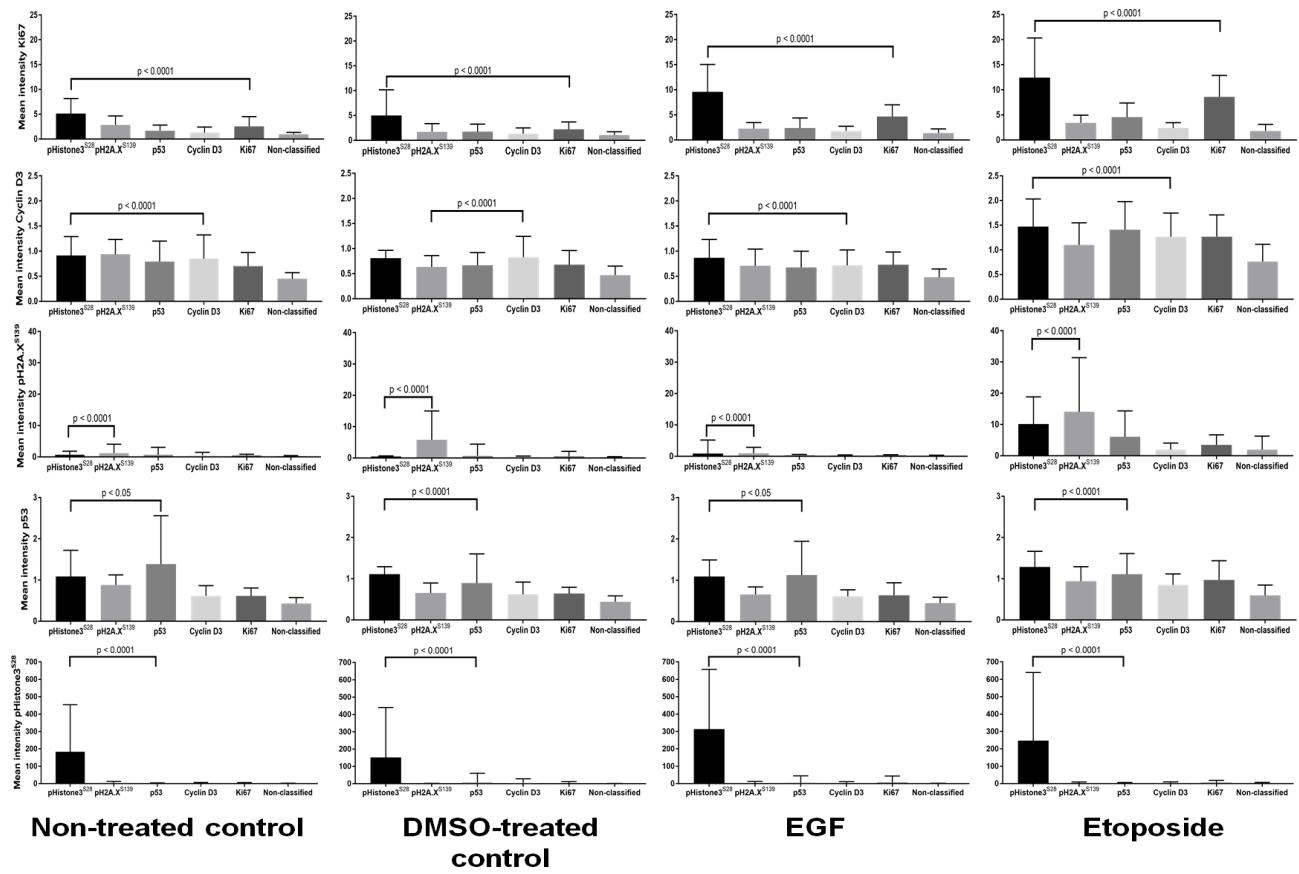


Figure S7. Comparative expression levels of protein markers Ki-67, cyclin D3, pH2A.X^{S139}, p53 and pHistone3^{S28} (rows) per nuclear class and drug treatment condition (column) of MCF-7 cells. Bar charts represent non-normalized mean intensity (+SD) of each nuclear marker to its related class. P < 0.05 and P < 0.0001 (unpaired t-test) are included.

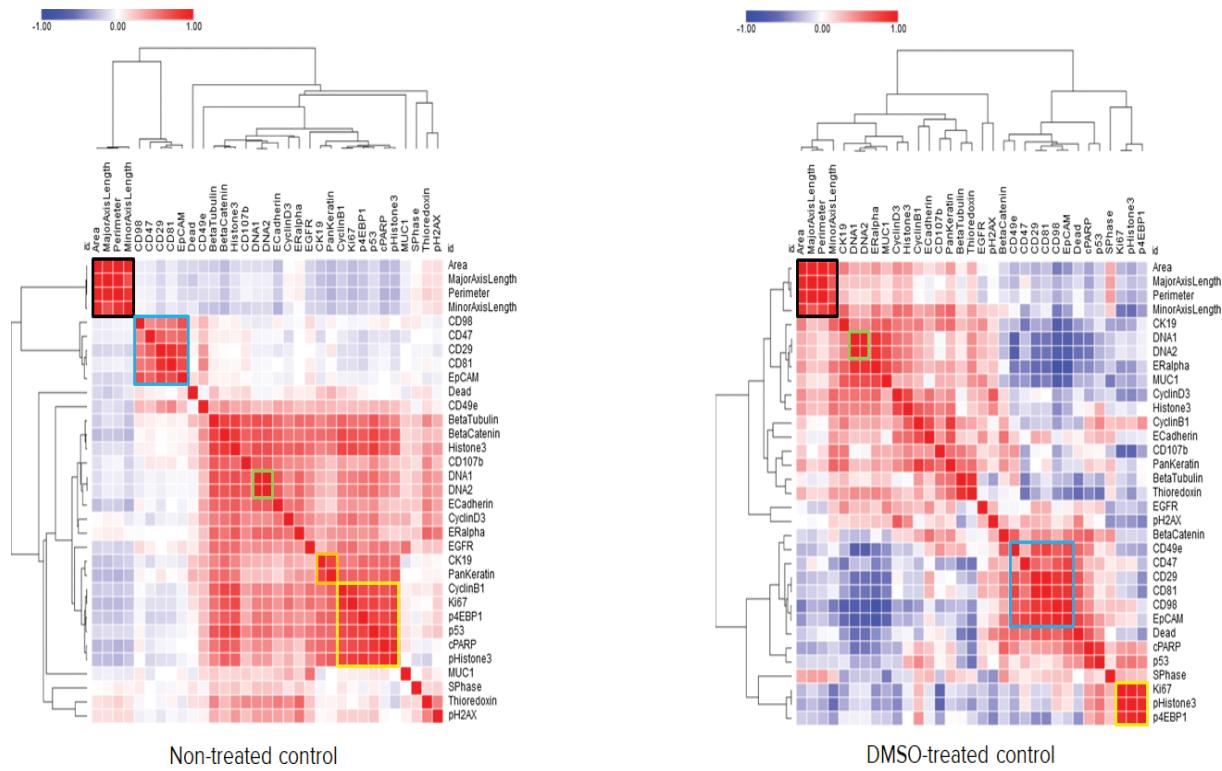


Figure S8. Heat maps of Pearson correlation coefficients of multiple parameters for non-treated and DMSO-treated controls ROIs on mitotic pHistone3^{S28} cells. Cluster highlighting of cell size parameters (black), surface markers (blue), Cell-ID iridium isotopes (green), pan-keratin and CK19 (orange) and nucleus markers (yellow).

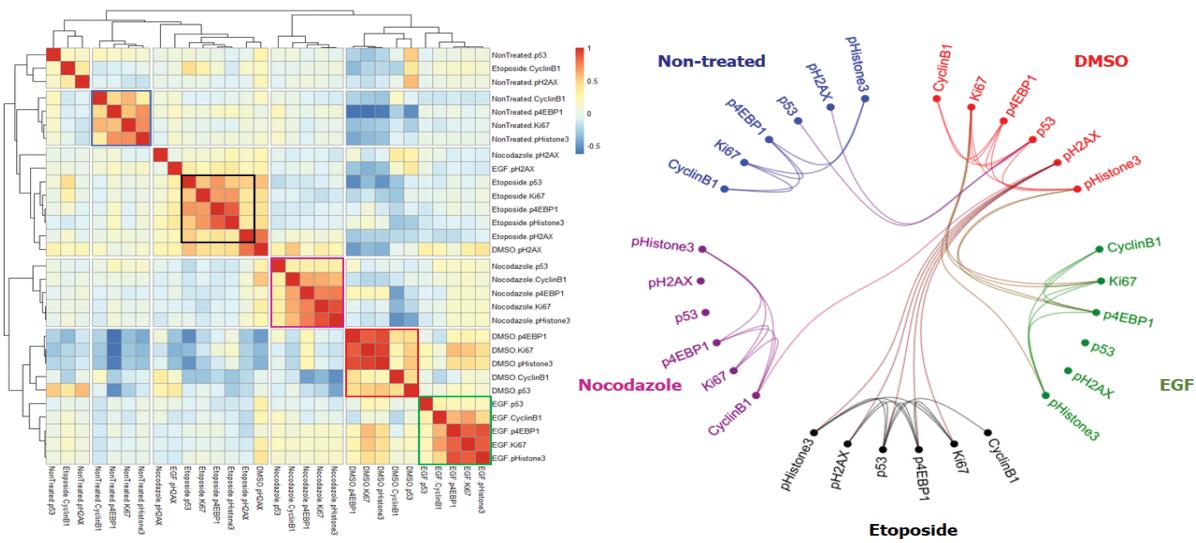
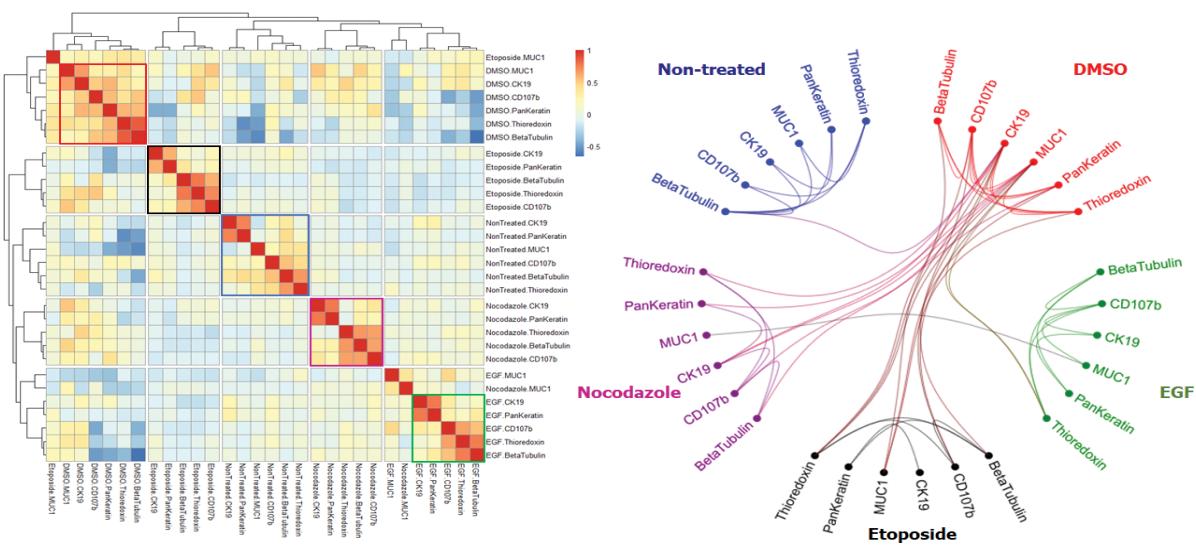
A**B**

Figure S9. Graphical representations of Pearson correlation coefficients for nuclei markers (A) and cytoplasmic markers (B) across all controls and drug treatments in the classified population of mitotic pHistone3^{S28} MCF-7 cells. Hierarchical similarities heatmaps show all pairwise correlations values between each protein, with diagonal components highlighted as multiparametric clusters identified for each condition (Non-treated, blue; DMSO, red; EGF, green; etoposide, black; nocodazole, purple). Intra and inter-relationships between these clusters are shown as Hierarchical edgebundle visual graphs. Bundled lines connecting parametric protein nodes correspond to positive correlation values higher than 0.3.

Supporting Tables

Compound	Stock Concentration	Working Concentration	Chamber Number
Nocodazole	3.3 mM	0.5 µM	6, 7
Etoposide	50 mM	10 µM	2, 3
hEGF	0.2 mg/mL	10 ng/mL	4, 8

Table S1. List of chemical compounds used for MCF-7 cell treatment, 48 h exposure.

Surface Marker Antibody	Metal isotope	Company/Catalogue Number	Clone	Marker	Breast Cancer Cell Line Tested
CD326	¹⁴¹ Pr	Fluidigm/3141006B	9C4	Epithelial cell membrane	SKBR3, HCC1143, MCF-7
CD81	¹⁴⁵ Nd	Fluidigm/3145007B	5A6	Tetraspanin	SKBR3, MCF-7
HER2	¹⁴⁸ Nd	Fluidigm/3148011A	29D8	Human epidermal growth factor receptor 2	SKBR3, HCC1143
CD29	¹⁵⁶ Gd	Fluidigm/3156007B	TS2/16	Integrin β1 adhesion	SKBR3, MCF-7
CD98	¹⁵⁹ Tb	Fluidigm/3159022B	UM7F8	Integrin β1-mediated adhesion	SKBR3, HCC1143, MCF-7
CD49e	¹⁶⁰ Gd	Fluidigm/3160015B	NKI-SAM-1	Integrin α5 adhesion	HCC1143, MCF-7
EGFR	¹⁷⁰ Er	Fluidigm/3170009B	AY13	Epidermal Growth Factor Receptor	SKBR3, HCC1143, MCF-7
MUC-1(CD227)	¹⁷³ Yb	BioLegend/355604	16A	Transmembrane-bound mucin	SKBR3, MCF-7
CD47	²⁰⁹ Bi	Fluidigm/3209004B	CC2C6	Integrin-associated protein	SKBR3, HCC1143, MCF-7

Table S2. Metal-labeled antibody panel for cell surface markers.

Intracellular Markers Antibody	Metal isotope	Company/Catalogue Number	Clone	Marker	Breast Cancer Cell Line Tested
CK5	¹⁴² Nd	BioLegend/905504	Poly19055	Keratin cytoskeleton	HCC1143
cPARP	¹⁴³ Nd	Fluidigm/3143011A	F21-852	Apoptosis	SKBR3, HCC1143, MCF-7
Thioredoxin	¹⁴⁶ Nd	Fluidigm/3146016B	2G11/TRX	Oxidative damage	HCC1143, MCF-7
pH2AX ^{S139}	¹⁴⁷ Sm	Fluidigm/3147016A	JBW301	Nucleus	SKBR3, HCC1143, MCF-7
p4E-BP1	¹⁴⁹ Sm	Fluidigm/3149005A	236B4	Nucleus	SKBR3, MCF-7
p53	¹⁵⁰ Nd	Fluidigm/3150024A	DO7	Nucleus	SKBR3, HCC1143, MCF-7
CK19	¹⁵⁵ Gd	BioLegend/628502	A53-B/A2	Keratin cytoskeleton	SKBR3, MCF-7
Vimentin	¹⁵⁶ Gd	Fluidigm/3156023A	RV202	Vimentin cytoskeleton	HCC1143
E-cadherin	¹⁵⁸ Gd	Fluidigm/3158021A	2.4E+11	Cell-cell interaction	HCC1143, MCF-7
CD107b	¹⁶¹ Dy	Fluidigm/3161017B	H4B4	Lysosomes	SKBR3, MCF-7
Pan keratin	¹⁶² Dy	Fluidigm/3162027A	C11	Cytoskeleton	SKBR3, HCC1143, MCF-7
ER alpha	¹⁶³ Dy	Fluidigm/3163024A	D8H8	Estrogen receptor	MCF-7
Cyclin B1	¹⁶⁴ Dy	Fluidigm/3164010A	GNS-1	Cell cycle G2/M interphase	SKBR3, MCF-7
Beta-catenin	¹⁶⁵ Ho	Fluidigm/3165027A	D13A1	Cell adhesion	SKBR3, MCF-7
Cyclin D3	¹⁶⁶ Er	CST/2936	DCS22	Cell cycle G1/S interphase	MCF-7
Ki-67	¹⁶⁸ Er	Fluidigm/3168007B	B56	Proliferation	SKBR3, HCC1143, MCF-7
Beta-tubulin	¹⁷¹ Yb	Abcam/ab6046	Rabbit polyclonal	Microtubules	MCF-7
pHistone3 ^{S28}	¹⁷⁵ Lu	Fluidigm/3175012A	HTA28	Mitotic cell	SKBR3, MCF-7
Histone3	¹⁷⁶ Yb	Fluidigm/3176016A	D1H2	Nucleus	SKBR3, HCC1143, MCF-7

Table S3. Metal-labeled antibody panel for intracellular markers.

Compound	Number of pHistone3^{S28} positive cells	Total number of cells	Mitotic Index
Nocodazole	316	6248	5.1 %
Etoposide	130	3264	3.9 %
hEGF	93	3816	2.4 %
Non-treated	127	4052	3.1 %
DMSO	14	2029	0.7 %

Table S4. Mitotic index of compound-treated MCF-7 generated by support vector machine classification with Fast Gentle Boosting ruler. Population of pHistone3^{S28} of all ROIs replicates per drug treatment and control were counted and divided by the total number of cells detected to determine the mitotic index of each condition.