Cover Page

Universiteit Leiden

Author: Di, Zi
Title: Development of automatic image analysis methods for high-throughput and high-content screening
Issue Date: 2013-12-10

Development of automatic image analysis methods for high-throughput and high-content screening

Zi Di

Development of automatic image analysis methods
for high-throughput and high-content screening

Zi Di

Development of automatic image analysis methods for high-throughput and high-content screening

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 10 December 2013
klokke 11.15 uur

door

Zi Di

邸 子

geboren te Chongqing, P. R. China
in 1984

Promotiecommissie

Promotor:

Prof. Dr. B. van de Water   LACDR, Leiden

Co-promotoren:

Dr. J. H. N. Meerman   LACDR, Leiden

Dr. Ir. F. J. Verbeek   LIACS, Leiden

Overige leden:

Prof. Dr. P. H. van der Graaf   LACDR, Leiden

Prof. Dr. J. N. Kok   LIACS, Leiden

Prof. Dr. H. J. Tanke   LUMC, Leiden

Prof. Dr. Ir. A. J. Koster   LUMC, Leiden

Prof. Dr. Ir. L. J. van Vliet   Delft University of Technology

Contents

Chapter 1

General introduction

1.1   Introduction of the thesis

This thesis focuses on the development of image analysis methods for ultra-high content analysis of high-throughput screens where cellular phenotype responses to various genetic or chemical perturbations that are under investigation. Our primary goal is to deliver efficient and robust image analysis platforms which can 1) automatically detect cellular structures of interest from florescence microscope images and 2) quantify dynamics and organization of multi-cellular systems with phenotypic features. To recover heterogeneity of cellular behavior, we aim to develop single-cell-based image analysis methods so that cell subpopulations can be distinguished and investigated. Furthermore, we intend to develop methods to extract an ultra-high level of phenotypic details from images. This would enable system-level studies of phenotype characterization.

To promote a further understanding of this thesis, this introductory chapter firstly provides the general background and essential knowledge related to the high-throughput and high-content screening. Next, the state-of-the-art techniques and image analysis methods that have been already applied to high-throughput and high-content screenings are given. Finally, the scope and structure of this thesis is presented at the end of this chapter.

1.2   High-throughput and high-content screening

High-throughput and high-content screening is a phenotypic screening technique that utilizes automated microscope systems to identify the functional role of substances such as small molecules, peptides or RNA interference (RNAi), in the context of a pivotal, pathology relevant, cellular process, ultimately enabling the identification of drug targets and/or novel drug molecules. To prepare the screening, cells are cultured in 96- or 384-well format micro plates and treatments with substances are applied. Changes in cell phenotype can be visualized by labeling structures and molecular components of the cells with fluorescent dyes, made visible by immunofluorescence

methodologies, or by expression of green fluorescent protein (GFP)-tagged proteins. Finally, micro plates are imaged by an automated microscope system to monitor cellular response to the different perturbations.

1.2.1 "High-throughput"

In recent years, florescence microscopy technology has been developed dramatically in respect to resolution, speed, complexity and scale. The emergence of automated microscope systems with robotic handling enables the investigation of a large volume of compounds or genetic players simultaneously. Combined with genome-wide RNAi approaches [1, 2], high-throughput small-molecule-based perturbations [3] or over-expression strategies [4], high-throughput and high-content screening has become a powerful technology to thoroughly study the regulation of biological pathways that underlie the function of intact cells.

1.2.2 "High-content"

Although other high-throughput techniques, such as mass spectrometry or DNA-microarrays, have been developed and successfully applied to study diverse cellular pathway and their possible involvement in disease, these techniques, despite their great usefulness, cannot provide adequate temporal and spatial information , in the context of the structural and functional integrity of cells. Most importantly, they do not directly show whether the identified molecules have a functional role in the cellular process that is under investigation. High-throughput and high-content screening technique fills this gap by probing the function of molecules in their nature environment with exquisite and ever increasing spatial and temporal resolution [5–8]. Moreover, integrated with other high-throughput techniques, interdisciplinary information is collected, which provides us a challenging opportunity to study whole biological systems more comprehensively.

1.3 Workflow of high-throughput and high-content screenings

High-throughput and high-content screenings consist of four major stages: sample preparation, image acquisition, image analysis and data analysis. In this chapter, a general introduction for each of these stages is provided and the state-of-the-art techniques that are used in each stage are summarized.

1.3.1 Sample preparation

According to the biological question, different types of assays can be used for screening. In general, they are categorized in two groups: fixed assays and live-cell assays. For the studies where only the final status of cells needs to be investigated, fixed assays are the primary technique to be considered. The most common one is a 2D

fixed assay. In this assay, cells are grown as a cell monolayer (therefore the term "2D" is used) and only a simple and specific read out is measured. Because this type of assay is relatively easy to be processed and imaged, it is often considered for higher-throughput experimentation and larger scale automation. Many projects have already achieved a great success using this type of assay [1, 3, 7, 9, 10]. One example is the work of Marino Zerial and his co-workers [1]: a genome-wide RNAi screen was performed in HeLa cells to explore the function of human kinases in two principal types of endocytosis: clathrin- and caveolae/raft-mediated endocytosis. They showed that a high number of kinases is involved in endocytosis, and that each endocytosis route is regulated by a specific kinase subset.

However, 2D fixed assays are not representative of the cellular environment found in organisms. In fact, tissue-specific architecture, mechanical and biochemical cues, and cell–cell communication are lost under such simplified and highly biased conditions. Efforts to address this problem led to the development of 3D cell cultures. They utilize an extracellular matrix (ECM) gel to re-establish physiological cell–cell and cell–ECM interactions, thus mimicking spatial organization of real tissues in their nature environment better than 2D cultures. Recently 3D cultures have shown many advantages in a broad range of cell biology studies [11, 12], including tumorigenesis [13, 14], cell adhesion [15–17], cell migration [18] and epithelial morphogenesis [19]. Many researchers also start to establish 3D assays for high-throughput and high-content screening, for example, in the study of tumor cell migration and invasion [20]. Two 3D screening assays based on mouse and human breast cancer cells are extensively described in **chapters 3 and 5** of this thesis.

Another type of assays are live-cell assays which are used to study the dynamics of cellular processes. It requires automated microscope to monitor these processes and often fluorescence tagged protein are used that are continuously imaged over a certain period of time. This allows the collection of much more detailed phenotypic information, especially temporal information is provided that otherwise cannot be obtained. In addition, live-cell assays can reveal primary defects and secondary consequences of the phenotype and thereby allow a more precise interpretation of the function of the molecules that are under investigation. This type of assays has been widely used in the study of embryogenesis [2], cell division [21], and intracellular translocation of molecules [22, 23], which so far are the most significant achievements among all high-throughput and high-content studies. These type of studies are based on the discrimination between the cell membrane, the nucleus and the cytoplasm, and the translocation of fluorescently labeled molecules between the distinct compartments. One example is the study of NF-κB nuclear translocation. Many studies on this subject have been published in recent years, notably the work of Covert and co-workers [24, 25]. They developed a screen platform with single cell resolution that can image and determine the NF-κB activation over time in a high-throughput manner. In **chapter 2** of this thesis, a novel NF-κB screening platform is described that also measures NF-κB activation dynamics at a single cell level but is able to apply for high cell density as well. Furthermore, it is able to quantify the characteristics of all single cell dynamics for further understanding of cell-to-cell heterogeneity.

1.3.2 Image acquisition in a high-throughput manner

One of the characteristics of high-throughput and high-content screening is that a large amount of images needs to be acquired. Therefore image acquisition systems that are fully automated and time efficient are required. Most researchers choose to use commercial systems that are already available on the market [7, 26, 27]. In this chapter, the techniques that are relevant in the context of this thesis are summarized.

1.3.2.1 Microscope models

Most high-throughput and high-content image acquisition systems are equipped with wide-field fluorescence microscopy. For this type of microscopy, cellular structures of interest are tagged with fluorescent proteins and exited by light of a specific wavelength. After absorbing energy from excitation light, the fluorescent proteins emit light of longer wavelength that is captured by a detector to generate images. Confocal microscopy is another frequently used and more advanced fluorescence microscopy method. It adds a pinhole aperture in front of the light source so that light excites only one optical plane (focal plane) at a time. In addition, another pinhole is added in front of the detector to filter out the emission light generated from the plane above and beneath the focal plane. Therefore, images from confocal microscopy contain only sharp in-focus information from the focal plane. This is especially useful when a specimen is relatively thick and a series of optical sections need to be acquired through the specimen. The disadvantage is the relatively slower imaging process and longer exposure time of the specimen. Especially when a large number of sections needs to be imaged, severe photobleaching can occur and that presents an as yet unresolved problem.

The new generation of high-throughput screening systems starts to explore two-photon excitation microscopy [28], spinning disc microscopy [3] and super resolution microscopy [29]. Unfortunately, due to their relatively higher demand for both hardware and software, only few researchers have so far developed and applied them successfully in high-throughput and high-content screening.

1.3.2.2 Magnification, resolution, sampling size and camera setting

Objective lenses are the most important component of an optical system because they are the predominant factor that defines image quality. In general, objective lenses can be classified based on their magnification and resolution. High-throughput and high-content screening systems are typically equipped with$< 10\times$, $20\times$ and $40\times$ magnification lenses. Different magnifications are applied according to the structure of analysis. Measurement of structures over large areas, such as cell networks or zebrafish embryos, requires lower magnification than measurement of sub-cellular structures, for example, nuclear repair foci in a DNA damage and repair assay.

Image resolution (which refers here to spatial resolution) is a term used to describe how closely two objects can be resolved in an image and is directly determined by numerical aperture (NA), which is a number describing the amount of light coming

from the focus that the objective can collect. According to the Rayleigh criterion, the relation between NA and resolution is formulated as:

$$r = 0.61 \times \frac{\lambda}{\text{NA}} \tag{1.1}$$

where λ is the wavelength of emitted light. For a fluorescence microscope of NA = 0.5 and λ = 500nm, this results a resolution limit of 610nm. Two distinct objects closer than this resolution will be imaged as a single object. Because in high-throughput and high-content screening image acquisition is often done (as well as in this thesis) with 4× or 10× lenses with NA equivalent to 0.13–0.3 (corresponding to resolution limits of 2346nm–1017nm at λ = 500nm), the objective lenses are the limiting factor in determining which sub-cellular structures can still be imaged.

Sampling size (also refers to sampling frequency) often refers to the distance between 2 signal-recording points. According to the Nyquist rate, image sampling size in xy-direction bigger than half of the objective lens resolution would result in a loss of information. The pixel size of the camera used often determines whether this can be achieved (the width of each pixel in the camera chip divided by the magnification gives the size of each pixel in the sample plane). Combining pixels in the camera (pixel binning) can increase the pixel size and signal intensity. The benefits of this setting are that exposure time and image file size can be reduced significantly, but the disadvantage is that spatial resolution will be decreased.

1.3.2.3 Balance between quality and quantity

Image quality is directly related to magnification, resolution, sampling size, pixel binning and other elements. Higher magnification, smaller sampling size or no pixel binning could increase the image quality and this could make complex and computational intensive image analysis methods redundant while much more detailed and accurate information can be extracted. However, better image quality sacrifices imaging time, processing time and requires more data storage. For example, reducing z sampling size from 10μm to 5μm not only doubles the size of the image file, but also doubles the time of imaging. Moreover, the limitations of hardware and software often make it impossible to deal with extremely large amount of data. For example, image files bigger than 1.5 gigabyte cannot be opened and analyzed by ImageJ on a 32-bit Windows operation system. Therefore finding the balance between image quality and quantity is very critical for high-throughput and high-content screening.

1.3.3 Image analysis

Once images are acquired, the next stage of screening is image analysis. Image analysis is a process to extract numerical information from images that are representative of cellular phenotypic responses. Those numerical parameters are then used for further data analysis, such as hit identification and compound characterization. Image analysis often comprises two general steps: segmentation and quantification.

1.3.3.1 Segmentation

Segmentation is a technique of defining regions of the images that contain signals from the molecules which are under investigation (also called region of interested or ROIs). Those molecules are often fused with fluorescence protein or antibodies, or stained with fluorescent dyes, and as a result, regions that contain these molecules present higher signal intensity in images than other regions. This difference of intensity between ROIs and background is used for most segmentation algorithms. The key of those algorithms is to automatically determine an intensity threshold to distinguish ROIs and background, or sometimes multiple-levels are selected. There are other advanced segmentation methods that are model based or using prior knowledge, for example active contour model [30, 31]. However, considering their complexity and specificity, they have not been widely applied in high-throughput and high-content screening.

Because quantification is often made based on the segmentation result, proper segmentation is the key to generate meaningful data and must be optimized for every screening. Several commercial systems already combine image acquisition and analysis [7], for example, the BD pathwayTM bioimaging system, KineticScan HCS Reader (Cellomics) and ImageXpress high-content imaging Systems (Molecular Devices). They provide many standard image analysis methods, however, it is not possible for researchers to substantially adapt these methods for more specific biological questions. Therefore many researchers choose to use independent image analysis packages or write their own macros. One of the most popular tools used in high-throughput and high-content screening is ImageJ (or Fiji, which is considered as a distribution of ImageJ). It contains various standard image analysis algorithms. More importantly, because of its free open-source feature, hundreds of state-of-the-art methods have been programmed in the form of ImageJ plugins and macros, and are provided online. Furthermore, it provides a user-friendly platform for developers to customize their own analysis. For example, image analysis algorithms developed in this thesis are programmed as ImageJ plugins.

1.3.3.2 Quantification

Quantification aims to extract numerical features from images. According to biological questions, features can be classified as morphological based, localization based and intensity based. Morphological based features usually refer to the morphological properties of structures that are under investigation, such as the area of the nucleus (2D) or surface of neuron cell (3D). Localization based features contain relative spatial information of interested structures, for example the position of focal adhesions relative to the cell border. Both features are measured based on the segmentation result, and can be measured on the single-cell level or cell population level, or sometimes even sub-cellular level [32]. Intensity based features can be measured from the segmentation result as well. For example, in the study of NF-κB translocation [33] (**chapter 2**), dynamic changes of fluorescence intensity in the nuclear and cytoplasmic area are quantified to analyze the kinetics of NF-κB translocation. Other intensity

based features are independent of segmentation result, such as moments and wavelet. Recently those features started to draw more and more attention, and some analysis packages already embed algorithms to measure them, such as CellProfiler [34] and PSLID [32, 35].

1.3.3.3 Tracking and analysis for image stacks

For some research projects, tracking techniques are required. Tracking refers to finding the location of objects (such as nuclei or cells) in each consecutive frame of time-lapse images. It works directly on the segmentation result and it connects objects in time on the basis of criteria such as the speed of motion, the shape of the trajectory and the possibility of the objects splitting (cell division) or merging (cell fusion). Once tracking is computed, the tracks and associated object properties (morphology and intensity changes) can be combined to yield a powerful description of the phenotypic evolution of cells.

For the study of 3D structures, series of optical sections are typically acquired and compose image stacks. Those stacks can be used to reconstruct the 3D geometric models of the sample, or can be collapsed (also called projected) into a single 2D image. For the latter, further analysis such as segmentation and quantification is performed on the 2D projected images, thus saving processing time and data storage significantly. However, information in z-direction is missing. 3D analysis provides much more detailed spatial information, but requires more complicated and computational intensive reconstruction methods and quantification methods. Although it has not been widely used for high-throughput screening, especially for large-scale experiments such as genome wide siRNA screens, due to the physiological and pathological relevance of 3D culture systems, image analysis for 3D high-throughput and high-content screening will become an important area for future innovative developments in image processing. This is the subject of **chapters 4 and 5** of this thesis.

1.3.4 Data analysis

In data analysis, numerical features derived from a single treatment, for example a single RNAi gene knockdown or exposure to a small molecule, are considered as data from one sample. The purpose of data analysis varies from experiment to experiment. One basic purpose is to identify "hits", meaning to identify genes or compounds which play a functional role in the cellular process that is under investigation (genes), or may - positively or negatively - affect the process (compounds). This is determined by comparing each sample with control treatments that are carried out under the same condition but induce no change in the cellular process. Before "hits" are identified, quality control and data normalization are performed to remove systematic errors and to allow comparison and combination of samples from different plates. There are many statistical methods for normalization and quality control for single readout screens which are summarized in references [36, 37]. For the high-throughput and high-content screen where multiple readouts are measured, multi-parametric tests are

applied, for example Mahalanobis distance. Because more aspects of phenotype are taken into account, multi-parametric tests often provide more precise hits for treatments that have high variability in a single readout metric. Several multi-parametric methods will be introduced and applied in this thesis.

Recently, machine learning started to show its powerful function in high-throughput and high-content screening [38]. Clustering is one of the typical techniques. It is a form of unsupervised machine learning and can be used to group the samples based on the phenotypic similarities. For example, clustering has been used to group drugs by their effects [39] and proteins by their sub-cellular patterns [40]. Another often used technique is classification [41], which is a form of supervised machine learning. The biologists define the classes of treatments according to the biological properties in advance. For example, genes which are involved in the same signaling pathway are assigned to the same class. Those samples with known classes are used as training data to train a "classifier", which can automatically determine the class of unknown sample and which of the features are informative for distinguishing the classes. Sometimes unsupervised machine learning and supervised machine learning are combined in one experiment, to for example investigate the relationship of known biological classes with morphological clusters.

1.4   Scope and structure of this thesis

Currently, image analysis is likely the major bottleneck in high-throughput and high-content screening studies. One of the problems is the lacking of robust single-cell based analysis. Few platforms perform single cell image analysis but rely on the cell density or image resolution. When the cells are too dense or image resolution and magnification is relatively low, for example in micro-tissue study, existing methods cannot work properly. Therefore, one goal of this thesis is to develop robust and efficient image analysis methods suitable for single-cell studies that should not be largely limited to cell type, cell density and image quantity

The reason that there is a great need for the methods of single cell analysis is that cells may respond to stimuli or perturbation differently, creating distinct subpopulations. Although many studies emphasize this effect, it has almost never been quantitatively identified and the information on subpopulation behavior has not been used for later screening data analysis. Thus how important the role is of subpopulation in many high-throughput and high-content screens is still unknown. This thesis aims to develop an analysis platform that takes into account the heterogeneity of cellular behavior so that subpopulation information is collected for the further data analysis.

In many high-throughput and high-content screens specific biological effects are expected, and in the analysis by conventional methods only representative features are measured from the images. Those methods largely depend on the expertise of biologists, but often even biologists are not certain of all possible effects. To solve this problem, many methods measure morphological parameters thoroughly, however,

they miss the information which cannot be revealed by morphological properties. This thesis aims to develop an ultra-high content analysis which collects maximum information of phenotype from images, which can be applied for various types of screening assays.

The structure of this thesis is as follows: **chapter 2** introduces an automated image analysis for the study of NF-κB nuclear translocation kinetics in high-throughput screening. **Chapter 3** moves away from the live-cell screening and presents another image analysis platform suitable for the study of 3D cultured micro-tissues. This platform firstly projects each image stack into a single in-focus 2D image, and then ultra-high content measurement is carried out on the 2D projected image. The work in **chapter 4** extends this 2D projection based analysis to real 3D analysis, and demonstrates that it is suited for screenings which use wide-field microscopy. In **chapter 5**, this 3D analysis platform is applied to analyze the effects of various anti-cancer drugs on 3D micro-issues of mouse breast cancer. The performance of 2D projection analysis and 3D analysis is compared in **chapter 5**. Finally, **chapter 6** provides a general discussion on the results obtained in our studies and on the implications for future research.

References

[1] Pelkmans, L. et al. Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. Nature 436, 78-86 (2005).

[2] Sonnichsen, B. et al. Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature 434, 462-469 (2005).

[3] Korn, K. & Krausz, E. Cell-based high-content screening of small-molecule libraries. Curr Opin Chem Biol 11, 503-510 (2007).

[4] Starkuviene, V. et al. High-content screening microscopy identifies novel proteins with a putative role in secretory membrane traffic. Genome Res 14, 1948-1956 (2004).

[5] Chudakov, D. M., Lukyanov, S. & Lukyanov, K. A. Fluorescent proteins as a toolkit for in vivo imaging. Trends Biotechnol 23, 605-613 (2005).

[6] Lippincott-Schwartz, J. & Patterson, G. H. Development and use of fluorescent protein markers in living cells. Science 300, 87-91 (2003).

[7] Pepperkok, R. & Ellenberg, J. High-throughput fluorescence microscopy for systems biology. Nat Rev Mol Cell Biol 7, 690-696 (2006).

[8] Tsien, R. Y. Building and breeding molecules to spy on cells and tumors. FEBS Lett 579, 927-932 (2005).

[9] Puigvert, J. C. et al. Systems Biology Approach Identifies the Kinase Csnk1a1 as a Regulator of the DNA Damage Response in Embryonic Stem Cells. Sci Signal 6 (2013).

[10] Zanella, F., Lorens, J. B. & Link, W. High content screening: seeing is believing. Trends Biotechnol 28, 237-245 (2010).

[11] Huh, D., Hamilton, G. A. & Ingber, D. E. From 3D cell culture to organs-on-chips. Trends Cell Biol 21, 745-754 (2011).

[12] Pampaloni, F., Reynaud, E. G. & Stelzer, E. H. K. The third dimension bridges the gap between cell culture and live tissue. Nat Rev Mol Cell Bio 8, 839-845 (2007).

[13] Petersen, O. W., Ronnov-Jessen, L., Howlett, A. R. & Bissell, M. J. Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. Proc Natl Acad Sci U S A 89, 9064-9068 (1992).

[14] Wang, F. et al. Reciprocal interactions between beta1-integrin and epidermal growth factor receptor in three-dimensional basement membrane breast cultures: a different perspective in epithelial biology. Proc Natl Acad Sci U S A 95, 14821-14826 (1998).

[15] Cukierman, E., Pankov, R., Stevens, D. R. & Yamada, K. M. Taking cell-matrix adhesions to the third dimension. Science 294, 1708-1712 (2001).

[16] Meshel, A. S., Wei, Q., Adelstein, R. S. & Sheetz, M. P. Basic mechanism of three-dimensional collagen fibre transport by fibroblasts. Nat Cell Biol 7, 157-164 (2005).

[17] Walpita, D. & Hay, E. Studying actin-dependent processes in tissue culture. Nat Rev Mol Cell Biol 3, 137-141 (2002).

[18] Zaman, M. H. et al. Migration of tumor cells in 3D matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis. Proc Natl Acad Sci U S A 103, 13897-13897 (2006).

[19] O'Brien, L. E., Zegers, M. M. & Mostov, K. E. Opinion: Building epithelial architecture: insights from three-dimensional culture models. Nat Rev Mol Cell Biol 3, 531-537 (2002).

[20] Echeverria, V. et al. An Automated High-Content Assay for Tumor Cell Migration through 3-Dimensional Matrices. J Biomol Screen 15, 1144-1151 (2010).

[21] Neumann, B. et al. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature 464, 721-727 (2010).

[22] Bartfeld, S. et al. High-throughput and single-cell imaging of NF-kappaB oscillations using monoclonal cell lines. BMC Cell Biol 11, 21 (2010).

[23] Azorsa, D. O. et al. High-content siRNA screening of the kinome identifies kinases involved in Alzheimer's disease-related tau hyperphosphorylation. BMC Genomics 11, 25 (2010).

[24] Lee, T. K. & Covert, M. W. High-throughput, single-cell NF-kappaB dynamics. Curr Opin Genet Dev 20, 677-683 (2010).

[25] Tay, S. et al. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. Nature 466, 267-271 (2010).

[26] Abraham, V. C., Taylor, D. L. & Haskins, J. R. High content screening applied to large-scale cell biology. Trends Biotechnol 22, 15-22 (2004).

[27] Starkuviene, V. & Pepperkok, R. The potential of high-content high-throughput microscopy in drug discovery. Br J Pharmacol 152, 62-71 (2007).

[28] Lakowicz, J. R., Gryczynski, I. I. & Gryczynski, Z. High Throughput Screening with Multiphoton Excitation. J Biomol Screen 4, 355-362 (1999).

[29] Ries, J., Kaplan, C., Platonova, E., Eghlidi, H. & Ewers, H. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat Methods 9, 582-584 (2012).

[30] Maksimovic, R., Stankovic, S. & Milovanovic, D. Computed tomography image analyzer: 3D reconstruction and segmentation applying active contour models–'snakes'. Int J Med Inform 58-59, 29-37 (2000).

[31] Maksimovic, R. et al. Computed tomography image analyzer: segmentation applying active contour models–"snakes". Stud Health Technol Inform 68, 395-399 (1999).

[32] Glory, E. & Murphy, R. F. Automated subcellular location determination and high-throughput microscopy. Dev Cell 12, 7-16 (2007).

[33] Di, Z. et al. Automated analysis of NF-κB nuclear translocation kinetics in high-throughput screening. PloS one 7, e52337 (2012).

[34] Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7 (2006).

[35] Huang, K. & Murphy, R. F. Boosting accuracy of automated classification of fluorescence microscope images for location proteomics. Bmc Bioinformatics 5 (2004).

[36] Birmingham, A. et al. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods 6, 569-575 (2009).

[37] Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. & Nadon, R. Statistical practice in high-throughput screening data analysis. Nat Biotechnol 24, 167-175 (2006).

[38] Murphy, R. F. An active role for machine learning in drug development. Nat Chem Biol 7, 327-330 (2011).

[39] Perlman, Z. E. et al. Multidimensional drug profiling by automated microscopy. Science 306, 1194-1198 (2004).

[40] Chen, X. & Murphy, R. F. Objective clustering of proteins based on subcellular location patterns. J Biomed Biotechnol 2005, 87-95 (2005).

[41] Boland, M. V. & Murphy, R. F. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. Bioinformatics 17, 1213-1223 (2001).

Chapter 2

Automated analysis of NF-κB nuclear translocation kinetics in high-throughput screening

Zi Di, Bram Herpers, Lisa Fredriksson, Kuan Yan, Bob van de Water, Fons J Verbeek & John H N Meerman

Abstract

Nuclear entry and exit of the NF-κB family of dimeric transcription factors play an essential role in regulating cellular responses to inflammatory stress. The dynamics of this nuclear translocation within a cell population may dramatically change e.g. upon drug exposure. Furthermore, there is significant heterogeneity in individual cell response upon stress signaling. In order to systematically determine factors that define NF-κB translocation dynamics, high-throughput screens that enable the analysis of dynamic NF-κB responses in individual cells in real time are essential. Thus far, only NF-κB downstream signaling responses of whole cell populations at the transcriptional level are in high-throughput mode. In this study, we developed a fully automated image analysis method to determine the time-course of NF-κB translocation in individual cells, suitable for high-throughput screenings in the context of compound screening and functional genomics. Two novel segmentation methods were used for defining the individual nuclear and cytoplasmic regions: watershed masked clustering (WMC) and best-fit ellipse of Voronoi cell (BEVC). The time profiles of NF-κB oscillatory response at the single cell and population level was coupled to automated extraction of 26 analogue parameters including number of translocation peaks, amplitude of each peak, and duration of each translocation. The automated image analysis method was validated through a series of statistical tests demonstrating computational efficient and accurate quantification of our algorithm. Both pharmacological inhibition of NF-κB and short interfering RNAs targeting the inhibitor of NF-κB, IκBα, demonstrated the ability of our method to identify compounds and genetic players that interfere with the nuclear translocation of NF-κB.

2.1 Introduction

NF-κB is a family of dimeric transcription factors consisting of homo- or heterodimers of different subunits (e.g. p65/RelA). It is involved in cellular stress responses to stimuli such as cytokines, free radicals, ultraviolet irradiation, oxidized LDL, and bacterial or viral antigens [1–5]. In resting cells, NF-κB dimers are located within the cytoplasm, bounding to the NF-κB inhibitor IκB. After being exposed to NF-κB activating stimuli such as TNFα or IL1β, the IKK (the inhibitor κB kinase) complex is activated, which in turn phosphorylates IκB [6] and NF-κB [7,8]. Phosphorylated IκB proteins are then ubiquitinated and degraded by the proteasome, thereby liberating NF-κB dimers that translocate into the nucleus and regulate the transcription of the target genes. However, NF-κB dimers do not stay in the nucleus permanently. IκBα, a member of IκB family, is a transcriptional target of NF-κB [9]. Transcription of IκBα creates a negative feedback loop: newly synthesized IκBα protein enters the nucleus and binds to NF-κB, leading to the export of complex back to the cytoplasm (Figure 2.1). This negative feedback loop creates an oscillation of NF-κB nuclear-to-cytoplasmic translocation. Such a response seems essential in modulating differential transcriptional responses under transient or sustained cytokine signaling [10]. Given the role of NF-κB in diverse (patho)physiological responses, understanding the cell population dynamics of this process is essential.
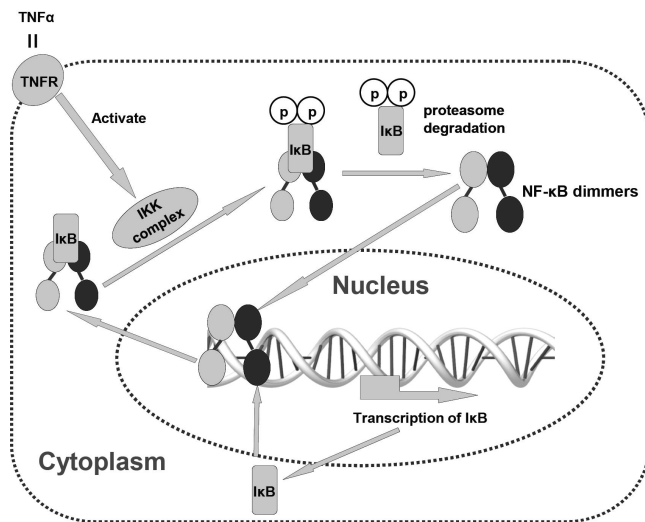
Figure 2.1: NF-κB oscillation is regulated by an auto-regulatory negative feedback loop. Simplified schematic overview of the TNFα-induced canonical NF-κB response. TNFα binding to the TNF receptor (TNFR) activates the inhibitor of κB kinase (IKK) complex, leading to phosphorylation of the inhibitor of NF-κB, IκB, upon which NF-κB is free to enter the nucleus to activate transcription of its target genes. One of the primary NF-κB target genes is IκB, which may retrieve NF-κB from the nucleus to maintain inactive IκB::NF-κB complex in the cytoplasm. Ongoing TNFR signaling can re-initiate the induction-inhibition cycle.

The most common approach taken in NF-κB translocation studies, which simply measures the NF-κB localization ratio between the total nuclear and total cytoplasmic region of a cell population, obscures the fact that not all cells respond to the stimulation synchronously [10, 11](Figure 2.2a and 2.2b). Similarly, recent studies of lipopolysaccharide-induced NF-κB activity showed that only half of the cells responded to the secondary TNFα autocrine signal, creating distinct subpopulations [12, 13]. Such cell-to-cell heterogeneity seems essential for the plasticity of tissue responses to inflammation [14, 15].

Furthermore, NF-κB responds to many different stimuli, each of which may lead to different activation dynamics. To understand NF-κB signaling under a wide variety of stimulation conditions, it is important to measure single-cell NF-κB dynamics in large cell populations. Obviously, studies of NF-κB translocation in just several individual cells are not sufficient for this purpose, although dedicated and sophisticated image analysis methods have been developed for this specific task [11,16,17]. In order to systematically determine factors that define NF-κB translocation dynamics, high-throughput screens need to be developed in relevant cell lines in the context of compound screening and functional genomics.

Our goal was to develop a methodology for quantification of NF-κB translocation dynamics in single cells, suitable for high throughput screening (HTS). For this, we used HepG2/GFP-p65 cells which show a dynamic nuclear-to-cytoplasmic translocation response upon TNFα stimulation (Figure 2.2a). To derive quantitative information of this shuttling in the entire cell population, we set out a strategy for the image analysis (Figure 2.2c). We describe two novel segmentation methods that are required for this purpose: one for the segmentation of individual nuclei, and one for the segmentation of individual cells. Next, cell tracking was done based on the segmentation results of nuclei. Finally, methods for the quantification of NF-κB translocation dynamics, and the extraction of informative parameters from the NF-κB translocation time profiles are described. In addition, procedures and results for the validation of each step in the quantification methodology are presented.

2.2   Results

2.2.1   Image collection and pre-processing

First, dual channel confocal images were collected (the two channels are Hoechst stained nuclear channel and GFP-p65 channel) in a 6 hour time-lapse series with a recording interval of 6 minutes (see Methods for details). Next, image pre-processing was applied separately for each of the two channels (Figure 2.3a and 2.3e). For the nuclear channel, images were sharpened first in order to enhance the contrast between background and intensity signals (by ImageJ http://rsbweb.nih.gov/ij/). This was implemented by an unsharp filter. It equals to subtracting a Gaussian blurred copy of the image and rescales the image to obtain the same contrast of large (low-frequency) structures as in the input image. We empirically defined the optimal radius of the
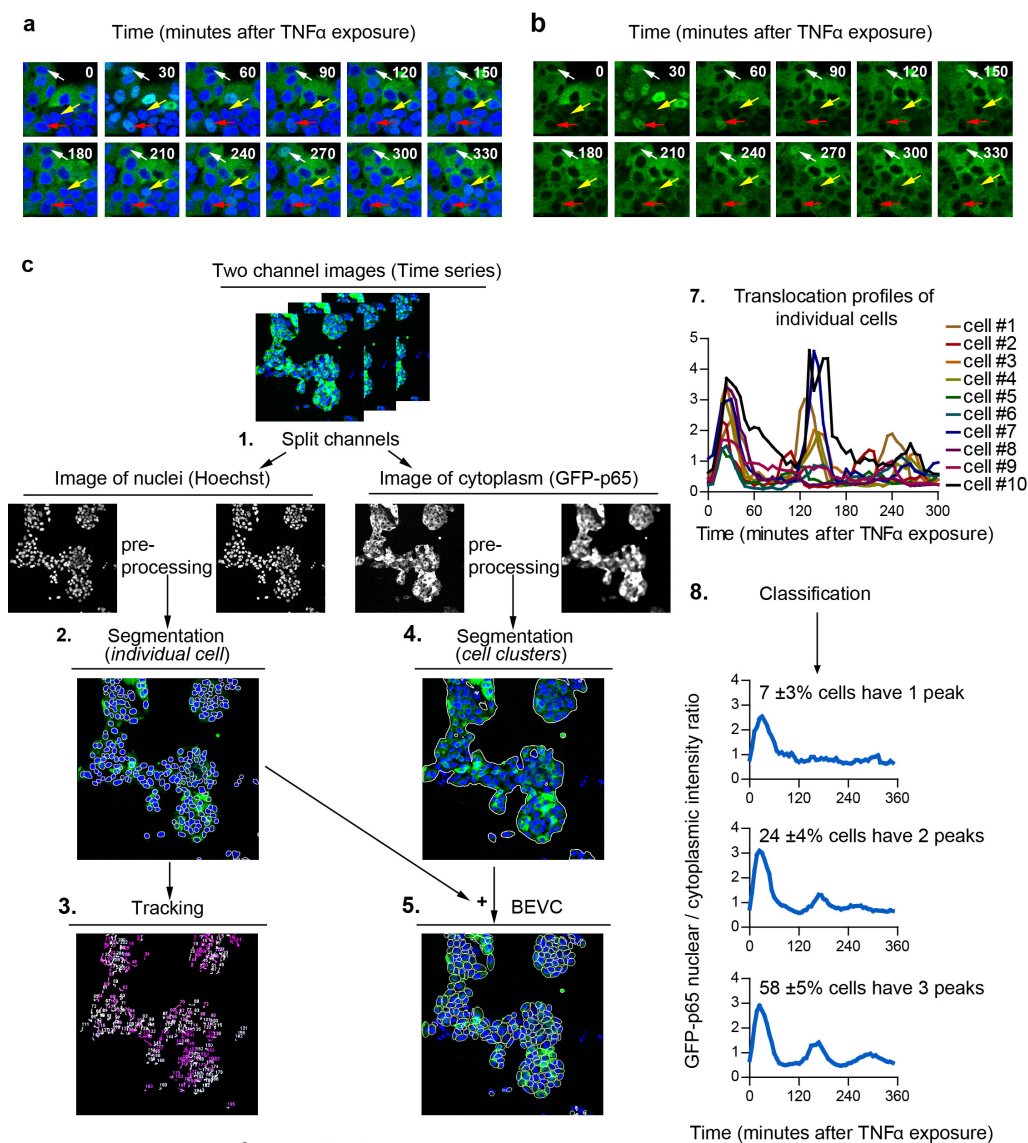
Figure 2.2: Image analysis of NF-κB nuclear translocation. (a, b) An example time-lapse image series of GFP-p65 expressing HepG2 cells after stimulated with 10 ng/mL TNFα. (a) Both Hoechst stained nuclear channel and GFP-p65 channel. (b) Only GFP-p65 channel to provide clear view of NF-κB translocation. Heterogeneous cell behavior is shown here. The white arrow indicates a cell with multiple nuclear translocations at 30, 150 and 270 minutes, while for another cell (yellow arrow) the translocation occurs at 30, 120, 210 and 330 minutes. The red arrow indicates a cell with only one, long, nuclear translocation event. (c) Flowchart of the single-cell based image analysis for the quantification of NF-κB translocation: 1. Splitting two fluorescence channels of time series images. 2. Segmentation of individual nuclei for the Hoechst stained nuclear channel. 3. Tracking of nuclear masks throughout the time series. 4. Segmentation of cell clusters for the GFP-p65 channel. 5. Identification of individual cells by best-fit ellipse of Voronoi cell (BEVC). 6. Quantification of GFP intensity ratio between the nuclear and cytoplasmic region per time-point, per cell. 7. Nuclear translocation profiles of individual cells. 8. One example of subpopulations analysis.

Gaussian filter [18] to be 3 pixels, and the scaling of the filter to be 0.6. Next, the so-called Rolling Ball method [19] was used to remove the unevenly illuminated background by subtracting an averaged image intensity within a circular kernel around each pixel (by ImageJ). The size of the kernel was chosen to be slightly larger than the radius of the largest nucleus. A pre-processed image of the nuclear channel is shown in Figure 2.3b. To define the overall cell cluster regions in images, the GFP-p65 channel was processed with a Median filter [20], resulting in smooth cellular regions (Figure 2.3f).

2.2.2 Segmentation for individual nuclei: Watershed Masked Clustering (WMC)

The segmentation of individual nuclei was accomplished by watershed masked clustering [21, 22]. This method uses a watershed segmentation to divide images into separated regions, each of which contains only one nucleus. Subsequently, within each region K-means clustering [23] was applied to define the nuclear region (Figure 2.3c). This method is based on the assumption that each nucleus is evenly illuminated and the contrast between nuclei and background is sufficiently high. Over-segmentation is a well-known issue of watershed segmentation. In order to address this issue, pre-processed images (Figure 2.3b) were convolved with a Gaussian filter to smooth discrete intensity signals, using an optimized kernel size. Once watersheds were obtained from this image, the pre-processed images prior to Gaussian convolving (Figure 2.3b) were used to apply K-means clustering.

2.2.3 Cell tracking

The nuclear masks obtained from the segmentation were used for the cell-tracking. In our NF-κB translocation experiments, we observed that most of the cells moved over short distances between two consecutive image frames, and also a negligible number of cell divisions occurred during the image acquisition period. Given these conditions, the maximum overlap ratio OLR (Equation 2.1) is a feasible and applicable criterion. For every labeled nuclear region in the current frame n_i^f, where i represents corresponding label and f represents the frame index, we connect its track with the labeled nucleus in the next frame n_j^{f+1} which maximizes the OLR with n_i^f:

$$OLR_{ij} = \frac{n_i^f \cap n_j^{f+1}}{max(Area(n_i^f),\ Area(n_j^{f+1}))} \tag{2.1}$$

Given the short imaging interval, cells should not disappear from one frame to another except when moving out of the frame borders. Disappearing cells are thus likely to be caused by under-segmentation. In order to avoid fragmented cell traces, as may occur in the more condensed cell clusters, we applied an extra tracking image buffer [24] to store disappearing cell regions until they are recovered again in one of the subsequent frames that maximizes the $OLRs$. As a result, nuclei that are not consistently detected in every frame can still be tracked correctly.

2.2.4 Segmentation for individual cells: Best-fit Ellipse of Voronoi Cell (BEVC)

The objects that need to be extracted in this particular live-cell NF-κB imaging application are cells that grow in clusters. These cells touch and may slightly overlap with each other thus making it sometimes difficult to uniquely identify the cellular edges. Therefore the classic edge detection methods which locate the maximum intensity gradient are not applicable to this particular case. Instead, we propose a single cell simulation algorithm called best-fit ellipse of Voronoi cell (BEVC). The algorithm produces an estimate of the single cellular areas based on the topology of the cells, which is derived from the distribution of nuclei. In principle, it consists of three steps:

Step 1, generate Voronoi diagram according to the topology of nuclei: A Voronoi diagram [25] was generalized based on a set of disjointed nuclear masks $b_i^{nucleus}$ for $i =1,2,3,\ldots, D_f$, (Figure 2.3c) with $b_i^{nucleus} \bigcap b_j^{nucleus} = 0$ when $i \neq j$, where D_f is number of nuclear masks in the frame f. Each Voronoi cell V_i containing $b_i^{nucleus}$ (Figure 2.3d) is defined as a region which includes all pixels r closer to the boundary of $b_i^{nucleus}$ than to the other nuclear masks. The formula is presented as following:

$$V_i(b_i^{nucleus}) = \left\{ r \in \min_{s \in b_i^{nucleus}} \mid r - s \mid < \min_{s' \in \bigcup_{j \neq i}^{D} b_j^{nucleus}} \mid r - s' \mid \right\} \qquad (2.2)$$

Step 2, obtain the Voronoi diagram within the cell cluster regions: A global threshold was applied to the pre-processed images of the GFP-p65 channel (Figure 2.3f) to obtain the binary masks of cell clusters (Figure 2.3g). Subsequently the masks were multiplied (AND operation) with the Voronoi diagram obtained from step 1, so that only the Voronoi cells within the binary masks of cell clusters were preserved (Figure 2.3h).

Step 3, estimate the cell shape per Voronoi cell: The underlying model for BEVC is that cells are ellipsoid shaped objects. Based on this assumption, we simulated the region, or better shape, of an individual cell as the best-fit ellipse in each Voronoi cell V_i by calculating the major and minor axis from the centralized moments (Figure 2.3i) [26, 27].

2.2.5 Quantification of NF-κB translocation dynamics

In order to exclude improper segmentation, both cellular masks and nuclear masks were validated by supervised two-class classifiers, based on their morphological features (Table S2.1, Figure S2.1-S2.3 and see Supplementary note for details). Next, the NF-κB translocation dynamics D_i^t of a single cell i at a time point t is defined as the ratio of average fluorescence intensities between the nuclear area $b_i^{nucleus}$ and cytoplasmic area, where the cytoplasmic area is defined as total cellular area minus nuclear area $b_i^{cell} \cap \overline{b_i^{nucleus}}$:

$$D_i^{t,f} = \frac{\frac{1}{N} \sum_{p \in b_i^{nucleus}} Intensity(p)}{\frac{1}{M} \sum_{p \in b_i^{nucleus} \cap \overline{b_i^{nucleus}}} Intensity(p)} \tag{2.3}$$

where p represents a pixel and f represents the frame index. N is the number of pixels in the nuclear mask of cell i, and M is the number of pixels in the cytoplasmic mask of cell i.
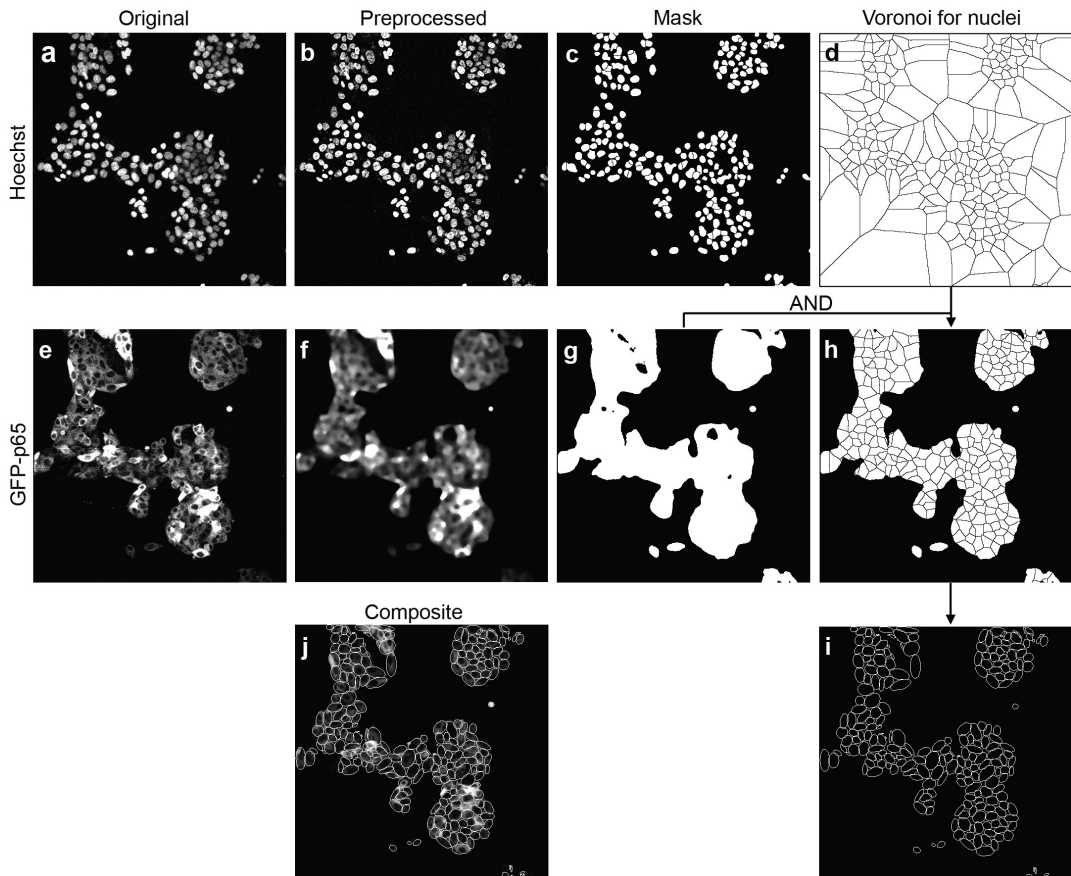
Figure 2.3: Stepwise demonstration of the image analysis method. An original image from the Hoechst stained nuclear channel (a) is pre-processed by image sharpening and background subtraction (b), followed by WMC for the nuclear mask identification (c). Subsequently, the Voronoi diagram (d) is generated based on the disjointed nuclear masks. For the GFP-p65 channel, the original image (e) is pre-processed by a smoothing filter (f) for the identification of cell cluster regions (g). By multiplication of the cell cluster masks (g) with the Voronoi diagram (d), the Voronoi mask is defined for each cell (h). The cytoplasmic areas are refined as the best-fit ellipse of Voronoi cells (i). Figure(j) shows the composite view of the original image from GFP-p65 channel and the corresponding BEVC segmentation result.

For cells with tracks that disappeared in three or less than three consecutive

frames, linear interpolation was applied to generate missing data. For cells with tracks that disappeared in more than three consecutive frames, interpolation becomes too inaccurate and the corresponding translocation profiles were removed from the final data. Generally, <30% of cells were removed by this procedure.

2.2.6 Quantification of NF-κB translocation analogue parameters

One advantage of the proposed analysis method is its ability to automatically quantify analogue parameters from each individual translocation profile (Supplementary note, Figure S2.5, Table S2.2). We first defined translocation events which start at a local minimum of a profile (NF-κB stays in the cytoplasmic region), include the next local maximum (NF-κB translocates to the nuclear region), and end at the next local minimum (NF-κB translocates back to the cytoplasmic region). We calculated the number of translocation events, various properties for each translocation event, nuclear entry and exit rates and time between consecutive peaks; in total 26 analogue parameters. More detailed information and pseudo code are presented in the Supplementary note.

2.2.7 Statistical validation of the NF-κB quantification method

We validated our quantification method in a three step process using five randomly selected time-lapse image series. First, we compared our BEVC method for cell segmentation with other segmentation methods that are used to segment touching or overlapping cells. One approach that is often used to define the cytoplasmic topological region is to dilate the corresponding binary nuclear mask by a few iterations. However, the extent of the dilation requires fine-tuning for different cell sizes to avoid overlap between individual cells. Another approach is to define the cell region by only applying the Voronoi diagram. Our method (BEVC) extends the topology information from the Voronoi diagrams with a best-fitting ellipse, which leads to a more stringent definition of the cellular area.

To compare these three methods (Dilation, Voronoi and BEVC), we first generated the binary images using different methods. For Dilation, we used a circular kernel with a radius of three pixels to define cytoplasmic regions, based on the general cell size in our images. Next, we assessed each segmentation result by comparing with human perception. For this, five test frames from different image series were used with a total of 1116 nuclei. For each frame f, a score named "error rate" ∂_f was calculated to measure the segmentation accuracy:

$$\partial_f = \frac{\sum_{i=1}^{D_f} B(b_i^{cell})}{D_f} \times 100\% \tag{2.4}$$

where $B(b_i^{cell})$ is a binary indicator that $B(b_i^{cell}) = \begin{cases} 1, & if\ b_i^{cell} \in b_i^{cell,Original} \\ 0, & if\ b_i^{cell} \notin b_i^{cell,Original} \end{cases} \cdot b_i^{cell}$ is the cellular mask of cell i obtained from one of three methods. $b_i^{cell,original}$ is the

cellular mask of $i'th$ cell obtained by human perception. D_f is the total number of cells in the image frame f, detected by one of three methods. The use of the Voronoi combined with best fit ellipse (BEVC) yielded the smallest error rate for cytoplasmic area definition ($10.3\% \pm 2.2\%$), compared to a Dilation or Voronoi method ($14.5\% \pm 3.2\%$ and $11.8\% \pm 1.4\%$ respectively) (Figure 2.4a).

Next, we validated our NF-κB translocation quantification method by comparing automatically generated translocation profiles with benchmark profiles that were produced from cells with validated segmentation and tracking profiles by human perception. Five randomly selected time-lapse image series each with 47 frames were used in this test. From each test image series, three benchmarks were generated separately by three independent individuals (Figure S2.4a-e), in order to compensate for possible human bias. Subsequently, a split-plot ANOVA [28] was applied (by Statistical Computing Seminars Repeated Measures Analysis with R) to test the difference between the benchmark profiles generated by three test persons and the computational result, in total four groups. The metric is the NF-κB Nuclear/Cytoplasmic intensity ratio, and two independent factors are time and group. The statistical tests indicate that the variation between the three benchmarks is not significant; moreover there are no significant differences between the benchmarks and the computational result (Figure S2.4f). This indicates that the designed algorithm provides an accurate estimation of NF-κB translocation profiles.

2.2.8 Computational efficiency of the algorithm

We tested computational efficiency of the algorithm on the dataset obtained from HepG2/GFP-p65 cells (see Methods for details). The computational complexity of this algorithm is $O(nlogn)$. We analyzed six sets of 60 time-lapse image series (six times 3.51 GB). Each series contains two channels, and each channel consists of 60 frames. On average, 250 cells were analyzed per movie. The analysis of this dataset was completed in 83 ± 2 minutes on a desktop PC (Intel Core i7-3770, 3.40 GHz with 8 GB of RAM and Microsoft Windows 7 Professional, SP1). The most computationally intensive part is the background subtraction on the nuclear channel followed by the segmentation of the nuclei by WMC. This takes ˜64 seconds per series. Tracking of the nuclei is done in 6 to 7 seconds.

2.2.9 Identification of heterogenous cell populations

One of the main purposes to quantify single-cell NF-κB nuclear translocation dynamics, especially in the context of high-throughput screens, is to study the heterogeneity between cell subpopulations. Therefore it is necessary to validate whether our quantification method correctly identifies specific subpopulations of cells and does not create a bias towards any particular subpopulation. To establish this, we benchmarked five time series images (with 1116 cells) by manually counting the cell subpopulations. We performed three separate tests, comparing the computational
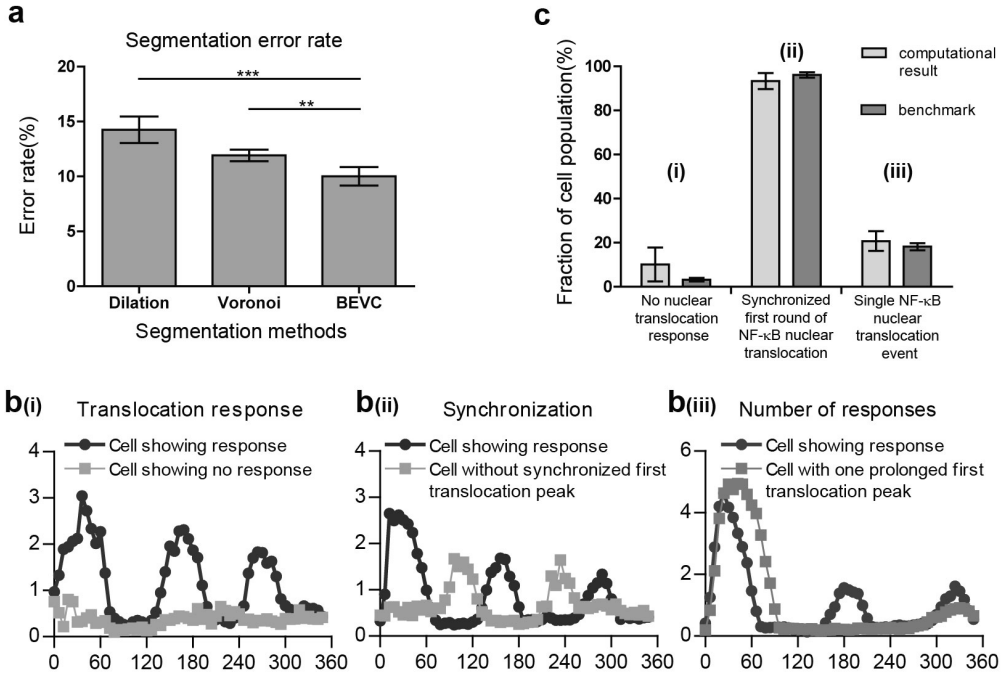
Figure 2.4: Statistical validation of the automated image segmentation and NF-κB translocation quantification. (a) Comparison of 3 cytoplasmic segmentation methods based on the criterion of error rate (Equation 2.4). The error rate of the Dilation method is $14.5\% \pm 3.2\%$; of Voronoi it is $11.8\% \pm 1.4\%$; and of BEVC it is $10.3\% \pm 2.2\%$.* $p-value < 0.05$; ** $p-value < 0.005$; Paired t-test (b) Example translocation profiles of (i) cells without translocation and cells with translocation, (ii) cells with and without a synchronized first round of NF-κB translocation, (iii) cells with NF-κB translocation occurring only once and cells with more than one NF-κB translocation event. (c) Bias assessment of our quantification method by comparison of the computational results with the benchmark for different subpopulations. No significant differences ($p - value > 0.1$) were found between the computational results and the benchmark for different cell subpopulations within a 6 hour imaging time frame.

results with the benchmark. In each test, cells were clustered into two complementary categories. In the first test, cells were clustered in cells without translocation response versus cells with translocation response (Figure 2.4b(i)). In the second test, we distinguished cells with a synchronized first peak of NF-κB translocation, from non-synchronized responders (Figure 2.4b(ii)). In this category, synchronization was defined as the first NF-κB translocation peak occurring within three frames from the first peak of population averaged profile. The third test clustered cells into (a) cells with no or only one (prolonged) NF-κB translocation event, and (b) cells with more than one NF-κB translocation event (Figure 2.4b(iii)). The reason for defining these three tests is their simplicity for human counting. For all three tests, we obtained $p - values$ greater than 0.1, indicating that there is no significant difference between our computational result and the benchmarks. Therefore, we conclude that our algorithm can efficiently be used to perform population studies on NF-κB nuclear

translocation profiles.

2.2.10 Biological validation of the NF-κB quantification method

In order to validate the sensitivity of our algorithm for perturbation of the biological system, a pilot experiment was performed by pre-exposing the HepG2/GFP-p65 cells for 2 hours with increasing concentrations of an IKK-inhibitor, BMS-345541 (0.5, 2.0 and 4.0μM) before TNFα stimulation. Inhibition of IKK prevents NF-κB nuclear translocation (Figure 2.1). The experiment was performed in 96-well plates on two different days, with two replicates per plate. In the first analysis step, the average GFP-p65 nuclear/cytoplasmic ratio profile over whole cell population of each image series were generated from our quantification method. Already at very low inhibitor concentration (0.5μM), the second and third NF-κB nuclear translocation peaks were delayed and the amplitude of the first peak was decreased. Increasing the concentration of BMS-345541 to 2.0μM and 4.0μM prolonged the first nuclear translocation event. Without TNFα stimulation, no NF-κB oscillation was observed (Figure 2.5a).

Next, the individual GFP-p65 nuclear translocation profiles were analyzed for the number of translocation events within the 6 hours imaging period after TNFα stimulation. In non-stimulated cells, 5% of the cells show spontaneous nuclear translocation, which is non-synchronous (Figure 2.5b). After TNFα stimulation, there is nuclear translocation with either one, two or three peaks, in 90% of the cells (Figure 2.5b). The average nuclear translocation profiles for cell subpopulations with either one, two or three peaks, clearly show that the percentage of cells with only one translocation peak increases with the concentration of BMS-345541, and that the percentage of cells with 3 translocation peaks decreases. In addition, we compared the time distribution of each translocation peak between the control (TNFα stimulated cells) and BMS-345541 pre-treatment. The result indicated that already at a low concentration (0.5μM) a significant delay occurred for the second and third translocation peaks (Figure 2.5c). In conclusion, all these data indicate that the quantification method can be used to perform cell-population studies, to identify rare events, and to study drug-dependent effects, even at low concentrations.

2.2.11 Application of the NF-κB quantification method in high-throughput screening assays

Having validated our NF-κB translocation quantification approach for segmentation accuracy, for correct subpopulation identification, and for sensitivity to biological perturbation of the system, we validated whether our quantification method can successfully be applied in the context of high-throughput functional genomics screening. For this screening, the approach of gene silencing by transient transfection of short interfering RNAs (siRNAs) was applied. We used three different siRNAs as control: the positive control is siNFKBIA that targets IκBα, upon which knockdown the NF-κB response will be affected [11]; the negative control is siCASP8 that targets caspase
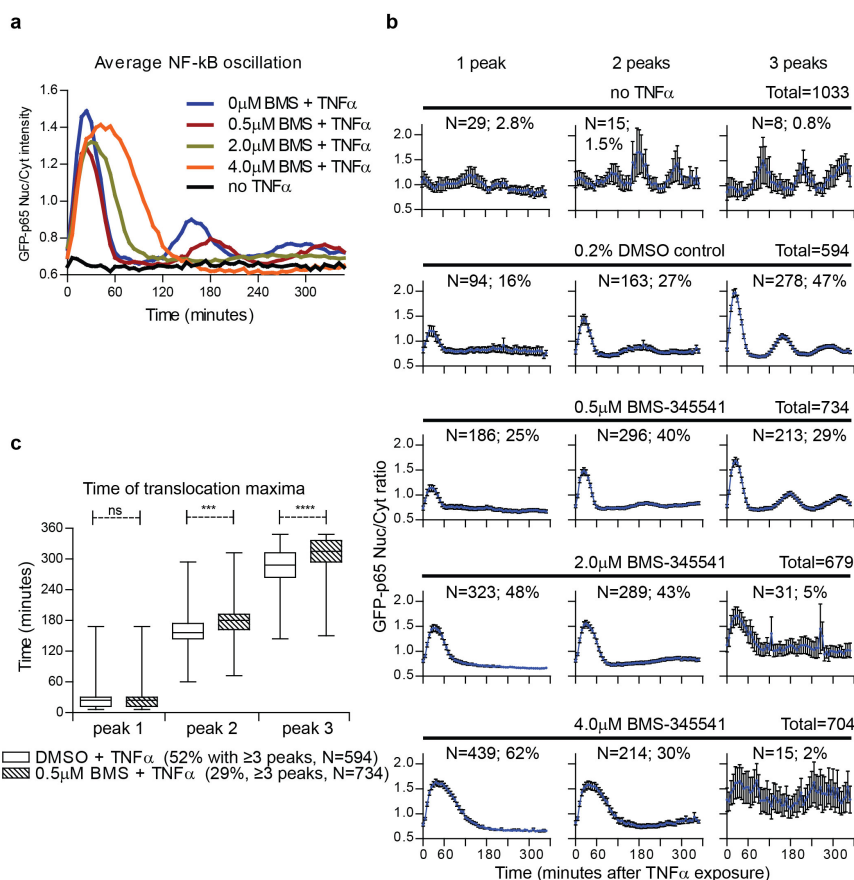
Figure 2.5: Population analysis of NF-κB nuclear translocation perturbation by the IKKβ inhibitor BMS-345541. Cells were pre-treated for 2 hours with increasing concentrations of BMS-345541 before TNFα stimulation. (a) Average nuclear translocation profiles over whole cell population of image series. (b) Average nuclear translocation profiles with standard error bars for cells with one, two or three translocation peaks. The total number of cells, the number (N) and percentage of cells which show responding number of peaks are presented. (c) Analysis of the time distribution of the 1st, 2nd and 3rd nuclear translocation peak in TNFα stimulated and TNFα stimulated plus 0.5μM BMS pre-treated cells. ns: No significant difference; *** $p - value < 0.001$; **** $p - value < 0.0001$.

8, which is a downstream effector of the TNFR, but does not affect the NF-κB activation; and siRNA control #1 (targeting luciferase) which also should not affect the NF-κB activation. These siRNAs were tested in 12 different 96-well plates (2 replicates per plate) on 4 different days, allowing an accurate analysis of the robustness of the assay.

First, we calculated the population average of GFP-p65 nuclear/cytoplasmic ratio profiles for each control, as well as for the cells that were not transfected with siRNAs, but exposed to a transfection reagent (mock) (Figure 2.6a). We did not detect an effect of caspase 8 knockdown on NF-κB oscillation compared to mock treatment;

yet surprisingly, siCntrl#1 slightly decreased the peak amplitude. IκBα knockdown however, strongly impaired NF-κB oscillation as expected.

Next, for each well, we calculated the average for each of the 26 analogue parameters (Table S2.2) over the cell population and derived further subpopulation information including the percentage of cells (%) showing 0 to 4 translocation peaks: in total 32 parameters (Figure 2.6b). In order to remove the cross-plate variance, a robust $z - score$ [29] was calculated to normalize each of the 32 parameters for each well.

To validate the reproducibility of controls and the quality of assay, we first calculated the standard $Z' - factor$ for each individual parameter (Figure 2.6b) to quantify the stability of both positive and negative controls, as well as the difference between positive and negative controls [30]. However, this conventional method of quality control is developed for assays with only a single readout and our quantification method provides readouts with multiple parameters. To enable the comparison of our assay with assays using only a single readout, we integrated multiple parameters into one value by Fisher's linear discriminant [31, 32] as suggested recently for integration of multiple readouts for quality control in high-content screening [33]. A direction ω was first identified to maximize the separation between positive control and negative control:

$$\omega = (S_{positive} + S_{negative})^{-1}(\mu_{positive} - \mu_{negative}) \qquad (2.5)$$

where $\mu_{positive}$, $\mu_{negative}$, $S_{positive}$ and $S_{negative}$ is the mean vector of positive control, mean vector of negative control, covariance matrix of positive control and negative control, respectively. They were derived from the $z - score$ vector $[z - score1, z - score2, ... , z - scoreN]$ of each well, where N is the number of parameters used for the calculation of multi-parametric $Z' - factor$ and $N \leqslant 32$. The $z - scores$ were then linearly projected onto this dimension ω according to Equation 2.6 and a multi-parametric $Z' - factor$ (Equation 2.7) can be calculated from the projected values.

$$P_i = \omega \cdot z - score_i \qquad (2.6)$$

$$Z' - factor = 1 - \frac{3(Std(P_{positive}) + Std(P_{negative}))}{\mid Mean(P_{positive}) - Mean(P_{negative}) \mid} \qquad (2.7)$$

where $z - score_i$ is the $z - score$ vector of well i. We measured both standard $Z' - factor$ (Equation 2.7) and robust $Z' - factor$ (Equation 2.8) by calculating the mean and standard deviation (Std), and the median and median absolute deviation (MAD), respectively:

$$\text{robust } Z' - factor = 1 - \frac{3(MAD(P_{positive}) + MAD(P_{negative}))}{\mid Median(P_{positive}) - Median(P_{negative}) \mid} \qquad (2.8)$$

Figure 2.6b shows univariate standard $Z' - factors$ for all 32 parameters. The highest value was 0.32 and the highest robust $Z' - factor$ was 0.54, both for the parameter "% 3 or more peaks" in concordance with the strong reduction of the number of oscillations upon IκBα knockdown. According to an established criteria [30], a $Z' - factor > 0.5$ indicates an assay suitable for HTS. The high value of the robust $Z' - factor$ therefore may validate our method for HTS. When we integrated the parameters with univariate $Z' - factor > 0$, in this case 6 parameters (Figure 2.6b), to calculate multi-parametric $Z' - factor$, a higher value of 0.61 for standard $Z' - factor$ and a very good robust $Z' - factor$ of 0.78 were obtained which further validates our assay for HTS (Figure 2.6c).

2.3 Discussion and conclusions

Controlling cellular fate in response to external stimuli is an important event in many physiological and pathological processes and in the action pharmacologically active compounds. Signaling routes that are involved herein frequently modulate gene-transcription by activation of nuclear transcription factors, such as NF-κB. In order to obtain a better insight in underlying processes that lead to the activation of these transcription factors, their subsequent translocation to the nucleus, and in the downstream events that follow their activation, methods need to be developed that enable the study of such events at the individual cell level and in high throughput fashion. In this study, we successfully developed such a methodology based on a novel method for cytoplasm definition (BEVC) and nuclei segmentation (WMC). Our method can easily be adapted to study the activation and nuclear cycling of other nuclear transcription factors as well.

The cell line used in this study (HepG2) is an epithelial-like hepatoma cell line, showing clustered and stacked cell growth. This influences the readout for GFP-p65 translocation by epifluorescence microscopy: superimposed, yet out of focus nuclei decrease the accuracy of single cell tracking and measurements. By adopting confocal microscopy in this study, the resolution and accuracy of single cell measurements are increased. Furthermore, we introduced the BEVC algorithm for accurate cytoplasm definition based on cell topology. Combined with WMC segmentation for the nuclear mask, and a series of quantification processes such as linear interpolation, the NF-κB translocation profile of each individual cell can be constructed. In order to validate our method, three sets of tests were applied on five time-lapse image series. These tests evaluated the proposed quantification method from three different perspectives, i.e. (1) accuracy of BEVC algorithm, (2) accuracy of calculated NF-κB translocation profiles, and (3) correct identification of cell subpopulations. In our test, only 10% of cells were segmented incorrectly by BEVC algorithm. Compared with a 14% error rate obtained by the Dilation method and a 12% error rate by the Voronoi method, we can state that the BEVC algorithm provides sufficiently accurate cell segmentation. The BEVC algorithm is also highly efficient, which is a key consideration for HTS analysis.
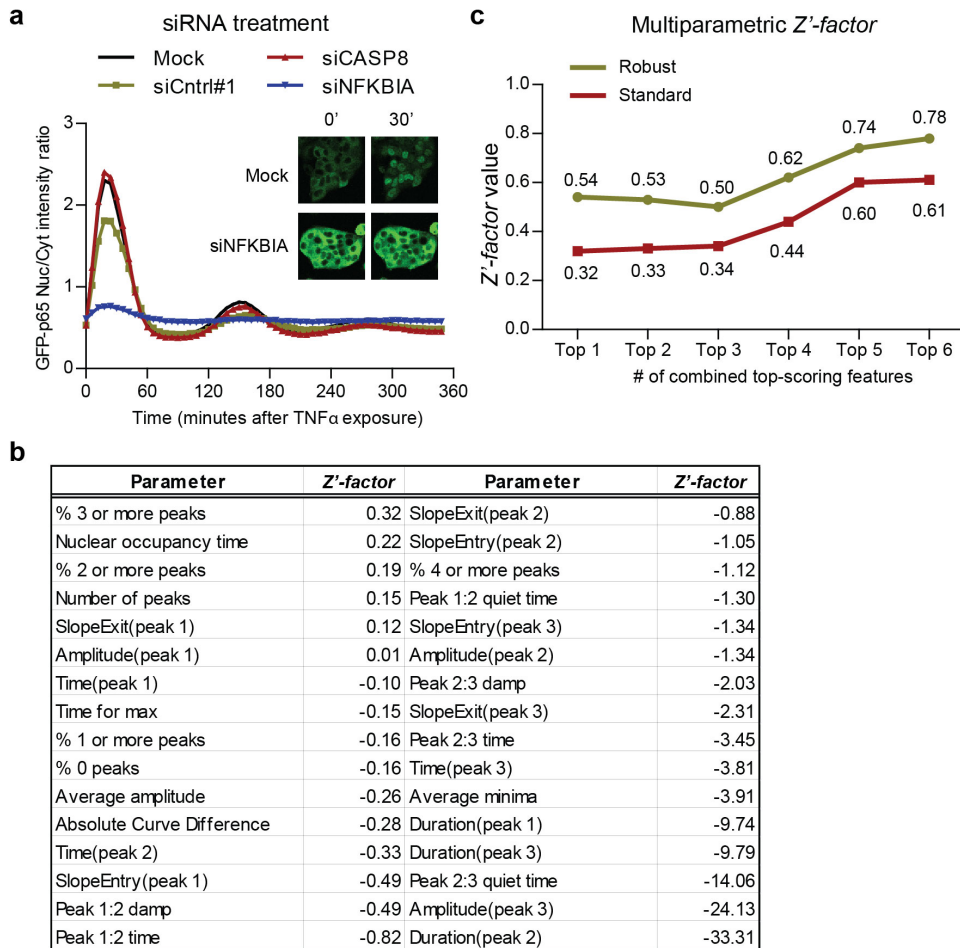
Figure 2.6: Application of the NF-κB nuclear translocation analysis method in siRNA screening assays. (a) The population average of GFP-p65 nuclear/cytoplasmic ratio profiles for negative control siCASP8, siCntrl#1, transfection reagent without siRNA (mock), and positive control siNFKBIA. Inset: representative images of mock and siNFKBIA treated GFP-p65 cells, at 0 and 30 minutes after TNFα stimulation (b) Table showing the ordered univariate $Z' - factors$ for all 32 individual parameters. The definitions of 26 analogue parameters are given in Table S2.2. Absolute Curve Difference indicates the absolute point-by-point difference between average profiles of control and treatment. (c) Multi-parametric $Z' - factors$ calculated based on the top-scoring univariate $Z' - factors$. Both the standard as well as robust multi-parametric $Z' - factors$ exceed the confidence threshold of 0.5 by combining more than 5 top-scoring univariate $Z' - factors$ with linear projection.

Other algorithms, such as contours derived from an active shape model [34], would possibly define a more precise cell edge, yet at the cost of analysis speed. Moreover, due to the uniform distribution of GFP-p65 in the cytoplasm, exact detection of cell boundaries is considered less relevant.

In the second and third validation test, we evaluated the accuracy of the calculated

NF-κB translocation profiles and the accuracy of cell subpopulation identification respectively. Our results indicate no significant differences between human generated benchmarks and the results obtained from the automated computational procedures, thereby validating our methods for studying NF-κB translocation, not only in the context of overall effects on the translocation response, but more importantly, also at the individual cell level.

In order to establish the sensitivity of our algorithm for perturbation of the biological system, an experiment was performed by pre-exposing the HepG2/GFP-p65 cells for 2 hours with an IKK-inhibitor, BMS-345541 before TNFα stimulation. The results show that already at the lowest concentration of BMS-345541, perturbation of NF-κB translocation was observed, thus validating our method for studying factors that affect this translocation. We also validated that our quantification method can successfully be applied in the context of high-throughput functional genomics screening. For this screening, the approach of gene silencing by transient transfection of siRNAs was applied. Based on calculation of multi-parametric $Z' - factors$, we demonstrated that our NF-κB quantification method can be used in HTS assays to identify genetic players that interfere with the nuclear translocation of NF-κB.

In this study we demonstrated the effect of IκBα silencing by siNFKBIA treatment on NF-κB oscillation. Theoretically the expected effect of IκBα loss would be persistent nuclear presence of NF-κB, however, this is not observed. The 3-day siRNA treatment instead led to an increased expression of the GFP-p65 construct, which was strongly retained in the cytoplasm, even upon TNFα stimulation (Figure 2.6a). Western-blot analyses showed that the loss of IκBα had resulted in a basic up-regulation of NF-κB target genes, including A20 and IκBα itself (data not shown), indicating that upon siNFKBIA treatment, the reporter cells had undergone multiple rounds of NF-κB translocation that most likely prevented further activation at the time of imaging.

NF-κB signaling is a complex process, and the balance of cytokine production and intracellular signaling transduction controls cellular fate in innate immunity and inflammation responses [14, 35]. It has been established by several groups that the individual cell response to cytokines may be very heterogeneous and is characterized by a full response of a few cells at low TNFα concentrations, and a similar response, but now for almost all cells at high concentrations, thereby creating distinct subpopulations of cells [10, 14, 15]. We show that for the HepG2 cells 5% of the cells oscillate spontaneously when no stimulation was applied. Spontaneous nuclear translocation has also been reported in neuroblastoma cells, although at a slightly higher level (18%) [15]. It is thought that this cellular variation serves biological important goals such as stability in acute tissue responses that are made up from highly heterogeneous individual oscillatory cell responses [14]. Therefore, it is an important goal and a major challenge to quantify cell subpopulations within the NF-κB response pathway. Several methods have been described that partially meet this demand [10, 14, 15, 36]. However, none of these is suitable for HTS because they either lack fully automated image analysis and require human intervention at some point, or require special equipment that prohibits massive parallel screening. The development

of a methodology suitable for HTS in the context of NF-κB signaling as presented in this study, whereby time courses of NF-κB translocation can be recorded in hundreds of individual cells over a period of many hours, presents a major breakthrough in this field. It now becomes possible to identify factors that govern NF-κB signaling at a genome wide scale. We are currently performing siRNA screening using this model to identify novel kinases and ubiquitinases that affect TNF-induced NF-κB nuclear shuttling.

Finally, the analogue parameters that we acquire from all the individual translocation profiles can be used as variables to model the sinusoidal oscillation of NF-κB translocation by systems biology approaches [14, 15, 36].

2.4 Methods

2.4.1 Cell line and cell culture

HepG2 cells stably expressing N-terminally GFP-tagged p65 (GFP-p65) [37] were maintained in Dulbecco's modified Eagle's medium (DMEM) with high glucose, 10% (v/v) FBS and 25μg/mL penicillin/25μg/mL streptomycin. HepG2/GFP-p65 cells were seeded on Greiner micro-clear 96well black plates (20,000 cells/well) and grown at 37°C, 5% CO_2 for 2-3 days.

2.4.2 Treatment of cells

The human cytokine TNFα (R&D Systems) was used in all experiments at 10ng/mL. The IKK-inhibitor BMS-345541 was from Sigma-Aldrich. Transient knockdown of NFKBIA was achieved using siGENOME NFKBIA SMARTpool siRNA (50nM; Dharmacon Thermo Fisher Scientific, Landsmeer, the Netherlands) and transfected into the HepG2 cells 3 days before imaging with INTERFERin (Polyplus transfection, Leusden, the Netherlands). Transfections with siGENOME SMARTpool CASP8 siRNA were used as negative controls in these experiments. Prior to imaging, nuclei were labelled with 100ng/ml Hoechst 33342 in culture medium for 45 minutes. For confocal fluorescence microscopy, upon recording the first frame of the time-series, TNFα was added as 10μL to each well containing 190μL medium.

2.4.3 Fluorescence microscope

The NF-κB nuclear translocation in the HepG2/GFP-p65 cells was imaged using a Nikon TiE2000 microscope equipped with a Perfect Focus System at 37°C with 5% CO_2 delivery to the sample plate location. Both the Hoechst-nuclear channel (excitation 405 nm, emission: 450 nm) and the GFP-p65 channel (excitation 488 nm, emission 515 nm) were recorded with the laser excitation confocal system. Images were acquired with a 20$\times$ (NA 0.75) dry Plan Apochromat objective and the image acquisition was controlled by EZ-C1 software (Nikon). In each well, an image from

the same position was acquired every 6 minutes for a period of 6 hours. The time-lapse series were exported in TIFF files as 16-bit digital images with 512×512 pixels frame.

2.4.4   Image analysis and statistical analysis

Image analysis was implemented using ImageJ (http://rsbweb.nih.gov/ij/). In-house plugins were written for quantification of both translocation profile and analogue parameters (Supplementary note). R (http://www.r-project.org/) was used to calculate the ANOVA test, t-test and multi-parametric $Z' - factor$.

References

[1] Gilmore, T. D. Introduction to NF-kappaB: players, pathways, perspectives. Oncogene 25, 6680-6684 (2006).

[2] Brasier, A. R. The NF-kappaB regulatory network. Cardiovasc Toxicol 6, 111-130 (2006).

[3] Perkins, N. D. Integrating cell-signalling pathways with NF-kappaB and IKK function. Nat Rev Mol Cell Biol 8, 49-62 (2007).

[4] Gilmore, T. D. The Rel/NF-kappaB signal transduction pathway: introduction. Oncogene 18, 6842-6844 (1999).

[5] Tian, B. & Brasier, A. R. Identification of a nuclear factor kappaB-dependent gene network. Recent Prog Horm Res 58, 95-130 (2003).

[6] DiDonato, J. A., Hayakawa, M., Rothwarf, D. M., Zandi, E. & Karin, M. A cytokine-responsive IκB kinase that activates the transcription factor NF-κB. Nature 388, 548-554 (1997).

[7] Jiang, X., Takahashi, N., Matsui, N., Tetsuka, T. & Okamoto, T. The NF-κB activation in lymphotoxin beta receptor signaling depends on the phosphorylation of p65 at serine 536. J Biol Chem 278, 919-926 (2003).

[8] Sakurai, H., Chiba, H., Miyoshi, H., Sugita, T. & Toriumi, W. IκB kinases phosphorylate NF-κB p65 subunit on serine 536 in the transactivation domain. J Biol Chem 274, 30353-30356 (1999).

[9] Sun, S. C., Ganchi, P. A., Ballard, D. W. & Greene, W. C. NF-κB controls expression of inhibitor IκBα: evidence for an inducible autoregulatory pathway. Science 259, 1912-1915 (1993).

[10] Tay, S. et al. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. Nature 466, 267-271 (2010).

[11] Ashall, L. et al. Pulsatile Stimulation Determines Timing and Specificity of NF-kappaB-Dependent Transcription. Science 324, 242-246 (2009).

[12] Covert, M. W., Leung, T. H., Gaston, J. E. & Baltimore, D. Achieving stability of lipopolysaccharide-induced NF-kappaB activation. Science 309, 1854-1857 (2005).

[13] Lee, T. K. et al. A noisy paracrine signal determines the cellular NF-kappaB response to lipopolysaccharide. Sci Signal 2, ra65 (2009).

[14] Paszek, P. et al. Population robustness arising from cellular heterogeneity. P Natl Acad Sci USA 107, 11644-11649 (2010).

[15] Turner, D. A. et al. Physiological levels of TNFα stimulation induce stochastic dynamics of NF-kappaB responses in single living cells. J Cell Sci 123, 2834-2843 (2010).

[16] Nelson, D. E. et al. Oscillations in NF-kappaB signaling control the dynamics of gene expression. Science 306, 704-708 (2004).

[17] Lee, T. K. & Covert, M. W. High-throughput, single-cell NF-kappaB dynamics. Curr Opin Genet Dev 20, 677-683 (2010).

[18] Haddad, R. A. & Akansu, A. N. A Class of Fast Gaussian Binomial Filters for Speech and Image-Processing. IEEE Trans Image Process 39, 723-727 (1991).

[19] Sternberg, S. R. Biomedical Image-Processing. Computer 16, 22-34 (1983).

[20] Mount, E. Nonlinear signal processing; a statistical approach. Sci-Tech News 59, 53-54 (2005).

[21] Kuan Yan & Verbeek, F. J. Segmentation for high-throughput image analysis: watershed masked clustering. In: SoLA'12 Proceedings of the 5th international conference on Leveraging Applications of Formal Methods, Verification and Validation: applications and case studies. (eds T. Margaria & B. Steffen) 25-41 (Springer-Verlag).

[22] Le Devedec, S. E. et al. Systems microscopy approaches to understand cancer cell migration and metastasis. Cell Mol Life Sci 67, 3219-3240 (2010).

[23] MacQueen, J. B. Some Methods for classification and Analysis of Multivariate Observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. (eds L. M. Le Cam & J. Neyman) 281-297 (University of California Press).

[24] Tang, C. & Bengtsson, E. Automatic Tracking of Neural Stem Cells. In: Proceedings of WDIC 2005. (eds B. C. Lovell & A. J. Maeder) 61-66 (The University of Queensland).

[25] Voronoi, G. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. J. Reine Angew. Math. 133, 97–178 (1907).

[26] Verbeek, F. J. Three dimensional reconstruction from serial sections including deformation correction. PhD Thesis, Delft, The Netherlands, 1-85 (1995).

[27] Verbeek, F. J. Theory and practice of 3D-reconstructions from serial sections. In: Image Processing, A Practical Approach. Oxford University Press, Oxford, 153-195 (2000).

[28] Stevens, J. Applied multivariate statistics for the social sciences. 3rd edn, 151-188 (Lawrence Erlbaum Associates, 1996).

[29] Birmingham, A. et al. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods 6, 569-575 (2009).

[30] Zhang, J. H., Chung, T. D. & Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen 4, 67-73 (1999).

[31] Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics 7 179-188 (1936).

[32] McLachlan, G. J. Discriminant analysis and statistical pattern recognition. Vol. 235 52-54 (John Wiley & Son, 1992).

[33] Kummel, A. et al. Integration of Multiple Readouts into the Z ' Factor for Assay Quality Assessment. Journal of Biomolecular Screening 15, 95-101 (2010).

[34] Verbeek, F. J., Rodrigues, D. D., Spaink, H. & Siebes, A. Data submission of 3D image sets to a bio-molecular database using active shape models and a 3D reference model for projection. Proceedings of the SPIE 5304, 13-23 (2003).

[35] Hayden, M. S. & Ghosh, S. Shared principles in NF-kappaB signaling. Cell 132, 344-362 (2008).

[36] Kalita, M. K. et al. Sources of Cell-to-cell Variability in Canonical Nuclear Factor-kappaB (NF-kappaB) Signaling Pathway Inferred from Single Cell Dynamic Images. Journal of Biological Chemistry 286, 37741-37757 (2011).

[37] Fredriksson, L. et al. Diclofenac inhibits tumor necrosis factor alpha-induced nuclear factor-kappaB activation causing synergistic hepatocyte apoptosis. Hepatology (2011).

[38] Raudys, S. & Duin, R. P. W. Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix. Pattern Recogn Lett 19, 385-392 (1998).

[39] Liu, C. J. & Wechsler, H. Robust coding schemes for indexing and retrieval from large face databases. IEEE Trans Image Process 9, 132-137 (2000).

[40] Duda, R. O., Hart, P. E. & Stork, D. G. Pattern Classification (2nd Edition) 215-268 (John Wiley and Sons, 2001).

[41] Webb, A. R. Statistical Pattern Recognition (2nd Edition) 123-163 (John Wiley & Sons, 2002).

[42] Cortes, C. & Vapnik, V. Support-Vector Networks. Mach Learn 20, 273-297 (1995).

[43] Theodoridis, S. & Koutroumbas, K. Pattern recognition. 3rd edn, 93-118 (Academic Press, 2006).

Supplementary note

1. Nuclear mask validation by classification

a. Overview

After defining the nuclear mask by WMC, we observed that some masks were not segmented accurately. Especially masks that extended over nuclear boundaries would generate inaccurate average GFP intensities of nuclei. To solve this problem, we decided to train a classifier which can automatically recognize incorrect nuclear masks and discard them. This classifier can be used for experiments carried out on different dates or with different treatments, as long as the microscope settings and cell line remain the same.

b. Training data

5 frames were randomly selected from 5 different time-lapse image series. We manually identified incorrect nuclear masks from the segmentation result of WMC . For all test images, 1179 nuclear masks were validated as accurate segmentation results (Figure S2.1a), and 127 nuclear masks were considered as incorrect (Figure S2.1b). Next morphological parameters (Table S2.1) were calculated using ImageJ, on both correct masks and incorrect masks, and then used to train the classifier.

c. Classification

Feature selection was performed to avoid the curse of dimensionality, using "forward" search algorithm. To define the optimal number of features, multiple classification methods were applied and a 10-fold cross-validation was used to evaluate the number of features, as well as the accuracy of each classification algorithm. The classification methods we tested were k-nearest neighbor classification with k = 1 or 2, linear Bayes normal classification, quadratic Bayes normal classification, nearest mean classification, Fisher linear classification and support vector machine with linear kernel. Figure S2.2 showed the cross-validation error rate for each classification method calculated on certain number of features. The result showed that when 2 features were selected, quadratic Bayes normal classification gave a quite low error rate of 5.58%. Those 2 features are circularity and area. In the end, a quadratic classifier was obtained which can be used to automatically validate the nuclear mask for the whole experiment. All the functions were implemented using PRtools on MATLAB.

2. Cellular mask validation by classification

a. Overview

We have also validated the cellular masks. Since we used the best-fit ellipse of Voronoi cell to simulate the cellular region, the circularity for each mask would all be 1, so that can not be used for the classification. Therefore we decided to only use an area threshold to identify the incorrect cellular masks. For regions where cells grow on top of each other, cell layers are not perfectly aligned with the focal plane where images are acquired and therefore these regions are not in focus (Figure S2.3a). As a

result, nuclei can not be detected and no nuclear masks were obtained in those regions (Figure S2.3b). Consequently, very big Voronoi cells (Figure S2.3c) and ellipses are generated (Figure S2.3d). Overlap of the GFP channel with ellipses clearly showed that those big ellipses contained multiple cells. To discard those incorrect cellular masks, we manually distinguished those ellipses from the rest and calculated an area threshold to automatically identify cellular masks with incorrect size.

b. Training data and classification
5 images which contain out of focus regions were chosen. Next, incorrect cellular masks which covered multiple cellular areas where the corresponding nuclear masks were missing were manually identified. Afterwards, the areas of these identified masks were measured, as well as those of correctly identified masks. In the end, an optimal area threshold was set up to minimize false classification.

3 Quantification of analogue parameters

a. Overview
One advance of our fully automated method is its ability to provide analogue parameters automatically for each time course profile. These analogue parameters translate the profiles into numerical parameters, such as number of peaks and amplitude of each peak (Table S2.2). This is very useful for categorizing different cell subpopulations according to their analogue parameters so that we can study not only the influence of various conditions on the whole population but also on subpopulations.

The outline for the quantification is shown in the Figure S2.5. For each time course profile, we first located the maximum value. Then each profile was smoothened to remove the small spikes which may be caused by intensity noise. Next, we started from the maximum point and scanned in both directions along the translocation profile, to search for the neighboring local minimum and local maximum. After defining all local maxima and local minima which represent peaks and valleys of each nuclear translocation events, parameters were measured to characterize each translocation event.

b. Smoothening of each time course profile
The main idea of smoothening is to remove noisy spikes on each profile so that we can locate the local maximum and local minimum precisely. We used a 1×5 mean filter window | 1 | 1 | 1 | 1 | 1 | to slide, frame by frame, convolving over the entire profile.

c. Local maximum and local minimum
Logically, local maximum and local minimum should appear alternatively. Combining this rule with following 4 rules, we experimentally define a point as local minimum:

1) The GFP-p65 nuclear/cytoplasmic ratio of this point is smaller than the ratio of neighbor points;

2) The adjacent local maximum should be at least 2 frames away from this point;

3) The ratio of this point should be below the half of maximum value;

4) The ratio difference between adjacent local maximum and this point should be at least 0.1.

For the local maximum, the rules are:

1) The nuclear/cytoplasmic ratio of this point is bigger than the ratio of neighbor points;

2) The adjacent local minimum should be at least 2 frames away from this point;

3) The ratio difference between adjacent local minimum and this point should be at least 0.1.

According to above rules, we are giving pseudo code as following:

```
record =new int[frame];          //This array contains the flag for each time frame. -1 is local minimum and +1 is local maximum
record [mT] =1;                  //Firstly tag the flag +1 for global maximum. mT is the time frame of global maximum

/***Now start from the mT, we scan in both left and right direction***/

//Flag for right direction which indicates whether we are searching for local minimum (-1) or local maximum(+1)
int flagr=-1;

//Flag for left direction which indicates whether we are searching for local minimum (-1) or local maximum(+1)
int flagl=-1;

/***The direction: left***/
for(time frame i=mT; i>0;i--)

    if((flagl==-1)&&(islocalminimum(i) ==true)&&(abs(index(neighbor local maximum)-i)>2))
    {//Find a local minimum

        if(((ratio(neighbor local maximum)-ratio(i))>=0.10)&&(ratio(i)<0.5*ratio(mT)))
        {
                record[i]=-1;
                flagl=1;         //Next, we are going to search for local maximum
                ratio(neighbor local minimum)=ratio(i);
                index(neighbor local minimum)=i;
        }
    }
    if((flagl==1)&&(islocalmiaximum(i) ==true)&&(abs(index(neighbor local minimum)-i)>2))
    {//Find a local maximum

        if((ratio(i)-ratio(neighbor local minimum))>=0.10)
        {
                record[i]=1;
                flagl=-1;        //Next, we are going to search for local minimum
                ratio(neighbor local maximum)=ratio(i);
                index(neighbor local maximum)=i;
        }
    }

/***The direction: right***/
Similar as left direction

return record;
```

d. Nuclei entry and exit time points

Some of parameters, such as SlopEntry and SlopExit of each translocation event, require us to first define where nuclei entry and exit time points are (Figure S2.5c), and then to calculate the gradient on that point. Those parameters can provide the information about how fast NF-κB translocates into or exits nuclei. The method

to calculate them is similar to the 2D edge detection by a Sobel operator. The filter $\boxed{-1 \quad 0 \quad 1}$ is convolved over the smoothened profile to calculate the gradient approximation for each frame, and then the frames with local maximum magnitude were assigned as SlopeExit or SlopeEntry according to the direction of the gradient.

Supplementary figures
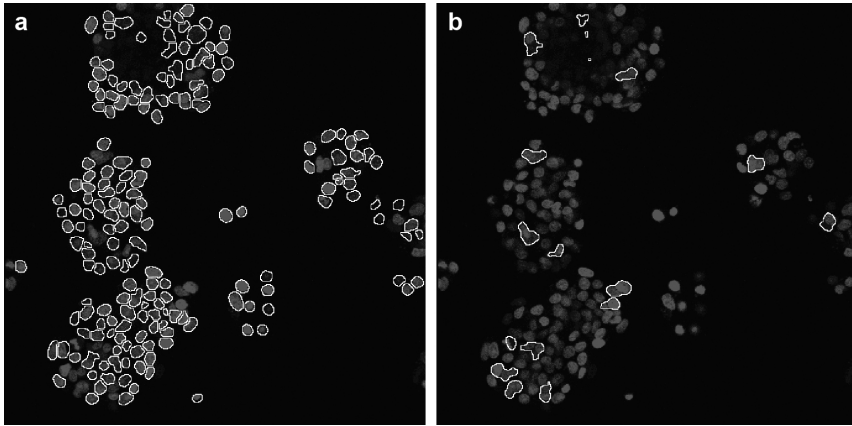
Figure S2.1: Validation of nuclear masks. (a) Accurately identified nuclear masks overlapping with Hoechst stained nuclei and (b) incorrect masks which were manually identified.
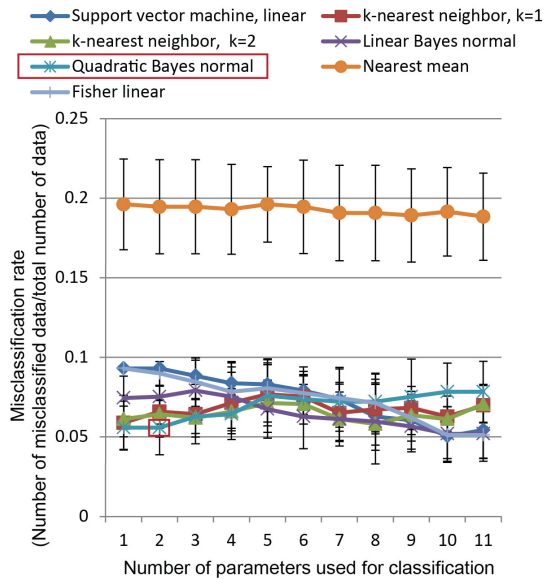
Figure S2.2: 10-fold cross-validation result from 7 different classification methods. The best result was obtained from quadratic Bayes normal classification when 2 features were selected (marked by red box).

37
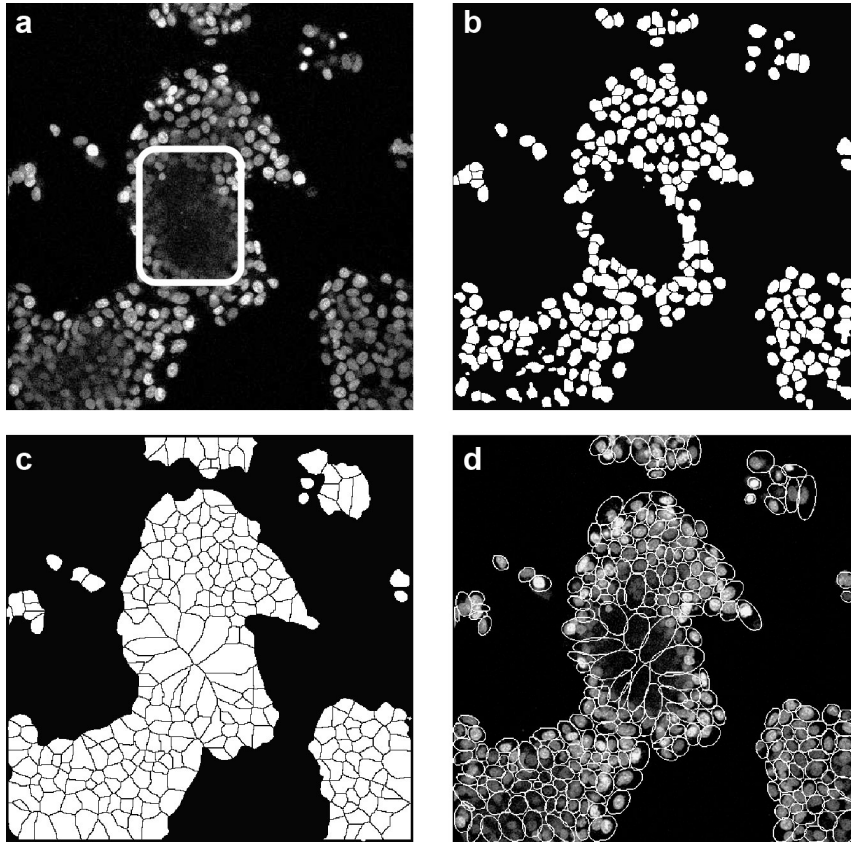
Figure S2.3: Cellular mask validation. (a) Out-of-focus region are formed when cells are clustered on top of each others (in the white box). (b) No nuclear masks were identified in those out of focus regions. (c) Very big Voronoi cells were generated due to the missing nuclear masks, and consequently big ellipses were generated (d). Overlap of the GFP channel with ellipses (d) clearly showed that those big ellipses contained multiple cells. An area threshold was trained to discard incorrect cellular masks.
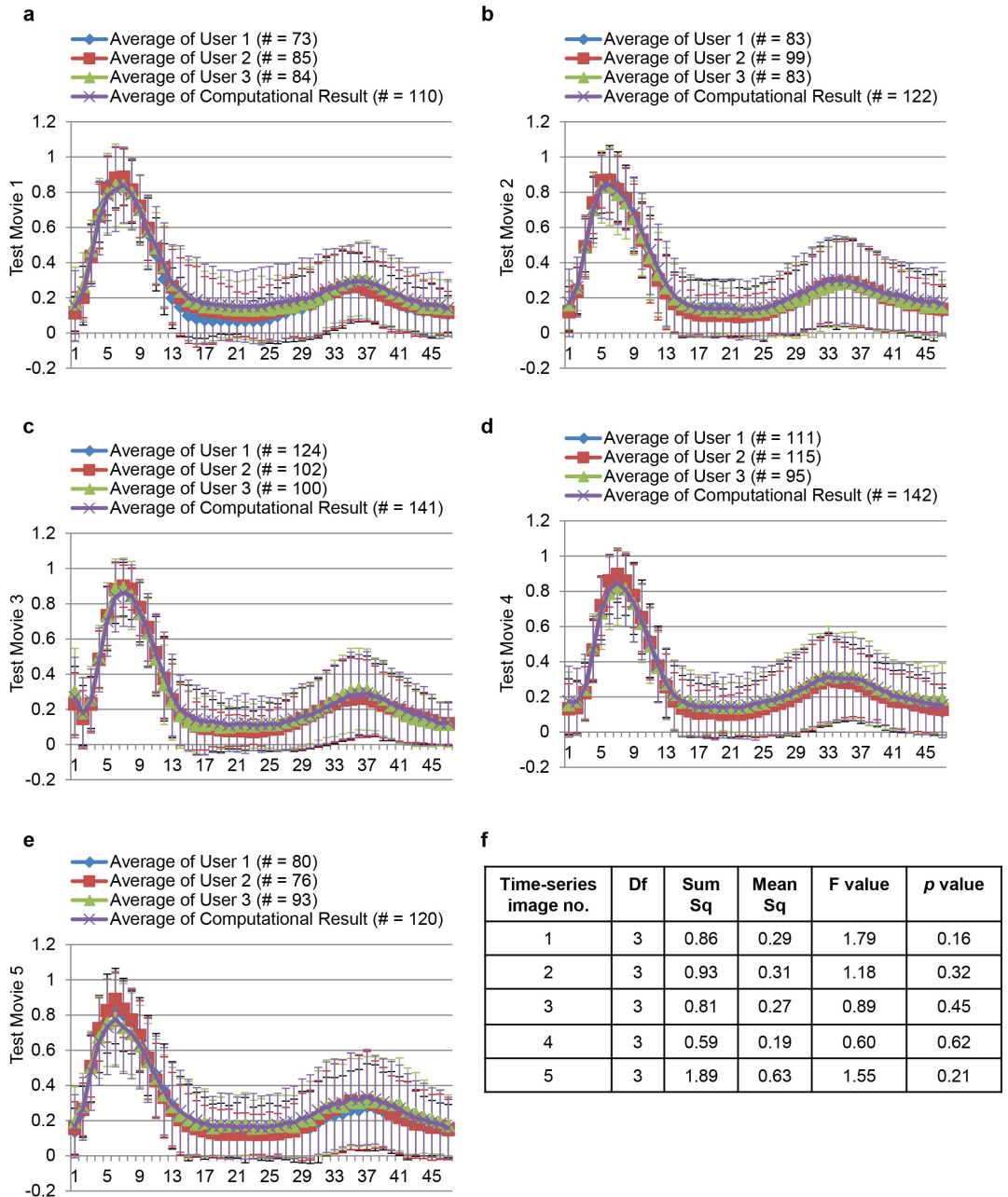
Figure S2.4: Validation of the automated NF-κB translocation quantification method. (a-e) For five test time-lapse series, we show the numbers of cells (#) and average GFP-p65 nuclear/cytoplasmic ratio profiles obtained from different individuals and computational result. (f) The accuracy validation results from the split-plot ANOVA analysis. Df: degrees of freedom. Sum Sq: sum of squares. Mean Sq: mean of squares.
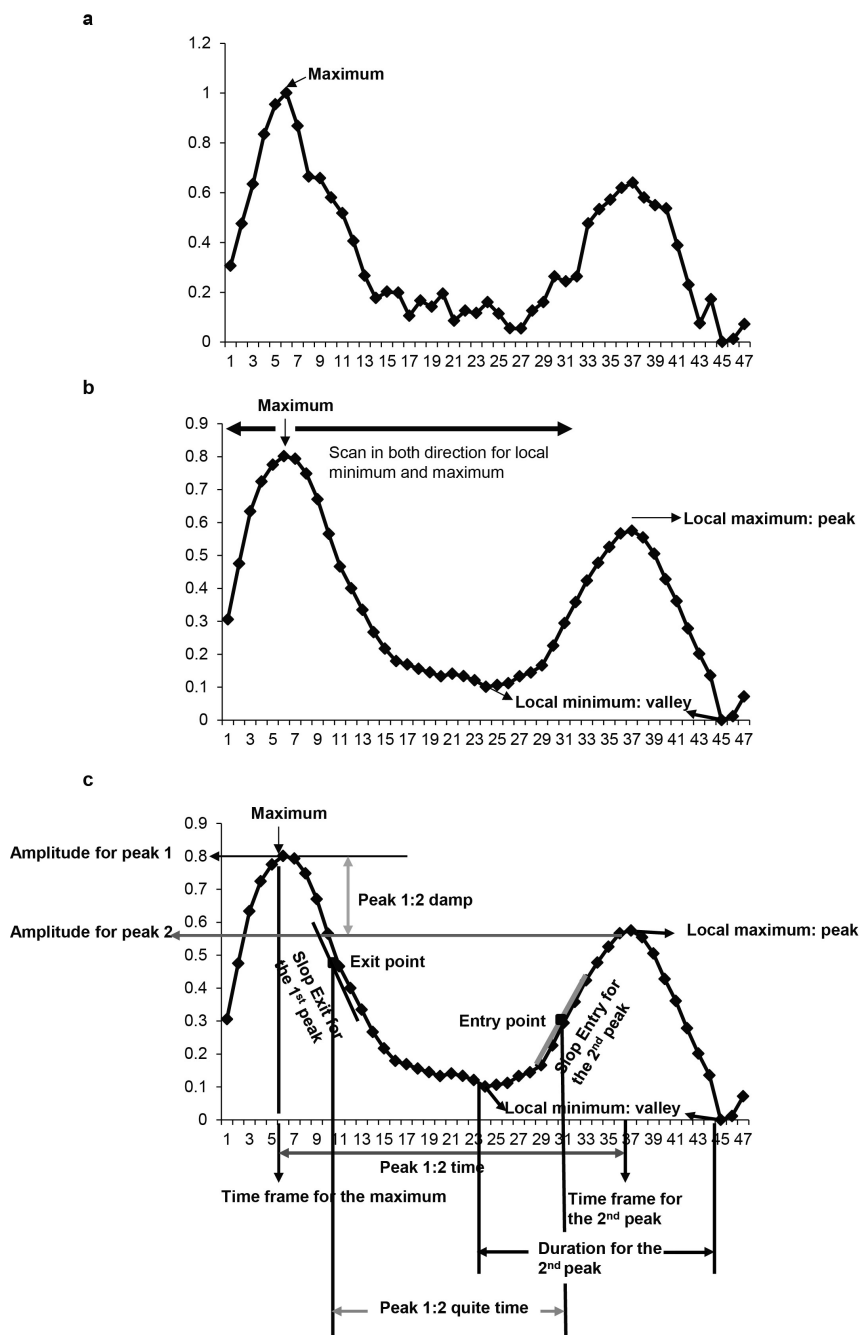
Figure S2.5: The outline to quantify the analogue parameters from individual time course profiles. (a) One example time course profile of one cell. (b) Smoothed profile on which we defined the local maximum (peak of translocation) and local minimum (valley of translocation). (c) Afterwards, parameters were measured to characterize the profile dynamic.

Supplementary tables

Index	Measurement	Description
1:	Area	The number of pixel in the current mask
2:	Perimeter	The length of the outside boundary of the current mask
3:	MajorAxis	The primary axis of the best fitting ellipse of the current mask
4:	MinorAxis	The secondary axis of the best fitting ellipse of the current mask
5:	Angle	The angle between the primary axis and a line parallel to the X-axis of the image
6:	Circularity	$4\pi \times \frac{[\text{Area}]}{[\text{Perimeter}]^2}$
7:	Maximum caliper	The longest distance between any two points along the boundary of the current mask
8:	Minimum caliper	The shortest distance between any two points along the boundary of the current mask
9:	Axis ratio	$\frac{[\text{MajorAxis}]}{[\text{MinorAxis}]}$
10:	Roundness	$4 \times \frac{[\text{Area}]}{\pi \times [\text{MajorAxis}]^2}$
11:	Solidity	$\frac{[\text{Area}]}{[\text{ConvexArea}]}$

Table S2.1: Morphological parameters for training the classifier to validate nuclear masks.

Analogue parameters	Parameter description
Number of peaks	The number of nuclear translocation events shown in the translocation profile
Time for max	The time when N/C ratio reaches the maximum along the profile
Time(peak 1)	The time when N/C ratio reaches the first translocation maximum of the profile
Amplitude (peak 1)	The N/C ratio of the first translocation maximum
Duration(peak 1)	The duration of first nuclear translocation event
SlopeEntry(peak 1)	The slope of nuclear entry of the first nuclear translocation event
SlopeExit(peak 1)	The slope of nuclear exit of the first nuclear translocation event
Time(peak 2)	The time when N/C ratio reaches the second translocation maximum of the profile
Amplitude(peak 2)	The N/C ratio of the second translocation maximum
Duration(peak 2)	The duration of second nuclear translocation event
SlopeEntry(peak 2)	The slope of nuclear entry of the second nuclear translocation event
SlopeExit(peak 2)	The slope of nuclear exit of the second nuclear translocation event
Time(peak 3)	The time when N/C ratio reaches the third translocation maximum of the profile
Amplitude(peak 3)	The N/C ratio of the third translocation maximum
Duration(peak 3)	The duration of third nuclear translocation event
SlopeEntry(peak 3)	The slope of nuclear entry of the third nuclear translocation event
SlopeExit(peak 3)	The slope of nuclear exit of the third nuclear translocation event
Peak 1:2 time	The duration time between the first translocation maximum and the second maximum
Peak 1:2 damp	The difference of N/C ratio between the first translocation maximum and the second maximum
Peak 1:2 quiet time	The time between end of the first nuclear exit and start of next nuclear entry event
Peak 2:3 time	The duration time between the second translocation maximum and the third translocation maximum
Peak 2:3 damp	The difference of N/C ratio between the second translocation maximum and the third translocation maximum
Peak 2:3 quiet time	The time between end of the second nuclear exit and start of next nuclear entry event
Average amplitude	The average N/C ratio of all the translocation maxima
Average minima	The average N/C ratio of all the translocation minima
Nuclear occupancy time	The total amount of time when the N/C localization ratio>half of the maximum amplitude

Table S2.2: Definition of analogue parameters measured for each individual cell translocation profile.

Chapter 3

Ultra high-content analysis and phenotype profiling of 3D cultured micro-tissues

Zi Di, Maarten J D Klop, Vasiliki-Maria Rogkoti, Sylvia E Le Dévédec, Bob van de Water, Fons J Verbeek, John H N Meerman & Leo S Price

Abstract

In many situations, 3D cell cultures mimic the natural organization of tissues more closely than 2D cultures, enabling cells to develop complex phenotypes. Conventional methods for phenotyping such 3D cultures use either single or multiple simple parameters which limits system-level study of phenotype characteristics. Here, we have developed a new image analysis platform to automatically profile 3D cultured micro-tissues with 598 phenotypic parameters including morphological parameters, topological parameters, texture parameters such as wavelet and image moments, as well as subpopulation related parameters. As proof of concept, we analyzed mouse breast cancer cells in a 384-well plate format following exposure to a diverse set of compounds at different concentrations. The result showed concentration dependent phenotypic trajectories for different biologically active compounds that could be used to classify compounds based on their biological target. To demonstrate the wider applicability of our method, we analyzed the phenotypes of 44 human breast cancer cell lines cultured in 3D and showed that our method correctly distinguished basal-A, basal-B, luminal and ERBB2+ cell lines in a supervised nearest neighbor classification method.

3.1 Introduction

Over the past decade, in vivo models and 2D cell cultures represented the two principle approaches used to study cellular process. The extreme low throughputs of in vivo models and poor (patho-) physiological relevance of over-simplified monolayer cell cultures motivated the development of 3D cell cultures. In many situations, 3D cell cultures mimic the natural organization of tissues more closely than 2D cultures, enabling cells to develop complex micro-tissue phenotypes. Especially for the study of tissue development where the spatial organization, architecture and interaction with the extracellular matrix are critical, 3D cell culture models may bridge the gap between in vivo studies and simple 2D cell mono-layer cultures [1]. One example is tubulogenesis of epithelial cells in 3D extracellular matrix (ECM) hydrogels, which result in in vivo-like structures that cannot be recapitulated in 2D cell cultures [2]. 3D cell cultures are also used frequently in tumor studies, allowing the effects of ECM, stromal cells and individual genes on tumor growth and invasion to be studied [3–7].

One potential application of 3D cultures is for the high-throughput screening (HTS) and high-content analysis (HCA) of pharmacologically active compounds [8]. For such purposes, 3D cultures are treated with compound libraries in 96- or 384-well micro plates. After incubation for an appropriate time, cells are labeled with fluorescent dyes, made visible by immunofluorescence methodologies, or by expression of green fluorescent protein (GFP)-tagged proteins. Next, fluorescent images are captured by automated microscopy. Here, 3D cultures have presented a challenge in collecting image data with sufficient resolution that is not prohibitively time consuming for a significant throughput of screening. Many methodologies have been established for the image analysis of HTS [9], although the quantification of phenotypes is mostly performed with single or multiple simple parameters. Therefore methods are needed to extract the maximum amount of useful information from the phenotypic complexity of the cultured tissues to allow the system-level study of the full range of effects of test compounds.

Our goal was to develop an automated multi-parametric profiling platform which is suited for HTS and is able to quantify cellular phenotypes exhaustively. Such a platform should apply rapid image preprocessing and segmentation methods that are suited for the images with limited resolution. For each well, collected attributes should provide a full spectrum of phenotypic information including properties of morphology, fluorescence staining intensity and topology, spatial attributes of the nuclei, as well as texture information such as image moments and wavelets. More importantly, this platform should be able to recover heterogeneous cell behavior. For example, it should be able to identify epithelial cells that develop branched structures from those that do not in response to a specific treatment. Such cell-to-cell heterogeneity seems essential for the plasticity of tissue responses e.g. in response to inflammation or associated with tumor invasion [10, 11].

As proof of concept, we investigated mouse breast cancer cells (4T1) after they formed micro-tissues in 3D by monitoring their cellular phenotypic response to a diverse set of compounds, using a novel 3D screening and ultra high-content analysis

(uHCA) technique. After image acquisition, we extensively mined images for feature data which we used for multi-parametric phenotypic profiling. To investigate phenotypic patterns, principle component analysis (PCA) was first used to reduce the dimensionality of the dataset. Subsequently, we compared various multi-parametric tests to identify biologically active compounds. Next, polynomial regression modeling was applied to characterize concentration dependent trajectories for each biologically active compound, and the distance between the trajectories was used for hierarchical clustering of compounds. Finally, multiple classification models were used to identify distinct phenotypic patterns.

3.2 Results

3.2.1 4T1 breast cancer cells acquire a complex phenotype in 3D culture, which is perturbed by biologically active compounds

To generate 3D micro-tissues, mouse triple negative breast cancer cells (4T1) were cultured in an extracellular matrix protein-rich hydrogel in 384-well high-content imaging micro plates. After seeding, cells were exposed to 29 compounds with different biological activities including tyrosine kinase inhibitors, cytostatic drugs, and Wnt-signalling activators (Table S3.1) at 6 different concentrations in quadruplicate (see Methods). To visualize nuclei and F-actin, the cultured micro-tissues were fixed and stained with Hoechst 33258 and rhodamine-phalloidin, respectively. Z-stacks of 17 xy epifluorescence image slices were collected using a BD Pathway 855 automated microscope with a $4\times$ objective and 0.16 Numerical Aperture (NA). In control wells, cells spontaneously formed a heterogeneous array of multi-cellular structures comprising spheroids and branched micro-tissues that were interconnected to form a complex network (Figure 3.1a). Exposure to many of the 29 test compounds resulted in a change in various aspects of network formation, such as branch length and thickness, number of branches, and the proportion and shape of spheroids (Figure 3.1b-3.1f)). Some compounds, such as the protease inhibitor bortezomib, and several compounds at higher concentrations, showed apparent toxicity, characterized by complete inhibition of network formation and pronounced inhibition of cell growth (Figure 3.1f). Exposure to other compounds, such as dasatinib, resulted in complete inhibition of branch formation and only multi-cellular spheroids were visible (Figure 3.1c).

3.2.2 Multi-parametric image analysis for phenotypic profiling

To capture the complexity and variation in phenotype in response to different treatments, we developed an automated image analysis pipeline for the multi-parametric phenotype profiling of the 3D cell cultures (Figure 3.2, Figure S3.1). First, a projection of in-focus information was generated from each fluorescence channel of the image stack (Figure 3.2a-3.2d), using an ImageJ plugin "Stack_Focuser" (see Methods). Next, the projected nuclei and F-actin images were segmented automatically to define the region of nuclei and cell clusters, respectively (Figure 3.2e, 3.2f). For
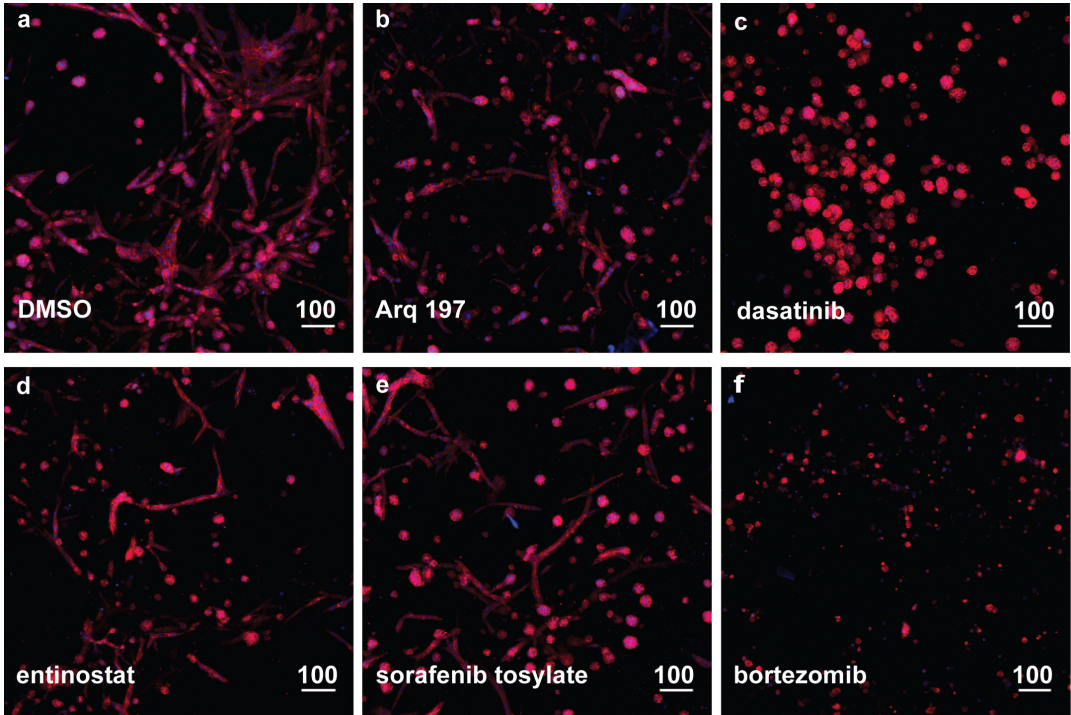
Figure 3.1: Mouse breast cancer cell (4T1) exposed to different compounds in 3D cell culture. For a clear representation of cellular phenotypic responses to different compounds, these images were acquired by a Nikon Eclipse Ti microscope in confocal mode. We used a dry air lens with 4× magnification and 0.2 NA. Two channel (Hoechst stained nuclei channel and rhodamine stained F-actin channel) z-stacks of 32 xy epifluorescence image slices were collected from each well, with acquisition step size in z-direction 50μm. Maximum intensity projection was applied to compress 3D image stacks to 2D image representation. Concentration of all compounds shown here was 0.316μM. Scale bar represents 100μm. (a) Untreated cells cultured in 0.2% DMSO, (b) cells exposed to compound Arq 197, (c) cells exposed to dasatinib, (d) cells exposed to entinostat, (e) cells exposed to sorafenib tosylate, (f) cells exposed to bortezomib.

the Hoechst stained nuclei channel, watershed masked clustering (WMC) [12,13] was applied to retrieve the binary masks of individual nuclear regions. For the rhodamine stained F-actin channel, the local Niblack algorithm [14] was used to define the region of cell clusters (see Methods). Quantification algorithms assembled from the literature [15,16] were incorporated into an ImageJ plugin to extract morphological and fluorescence intensity parameters from the images (Figure 3.2g, Table S3.2-S3.4 and Methods). We observed that the 3D micro-tissues formed different subpopulations: 1) spherical cell clusters and 2) branched and interconnected complex networks. Upon exposure to different compounds, the proportion and the shape of these two subpopulations often changed in a specific pattern and these patterns may contain compound specific information. To be able to quantitatively study these changes, an automated classifier for these two subpopulations was developed and embedded in the image analysis pipeline (see Methods and Figure S3.2). Relevant information was collected for each subpopulation based on fluorescence intensity and morphology

(Table S3.5). In total, 598 parameters were measured from the image of each well for the whole population and two subpopulations.

3.2.3 Identification of biologically active compounds

Our first goal was to identify biologically active compounds which significantly affect the cellular phenotype compared to controls (cells growing in 0.2% DMSO). For this purpose, we firstly calculated the robust $z-score$ [17] for each of the 598 parameters for the purpose of cross-plate normalization (see Methods). Next, the $z-scores$ of all 598 parameters were used for PCA and we retrieved 9 principle components which preserved 90% data variation. A 3D plot of data points on the first 3 principle components is given in Figure 3.3a. The parameters that contribute most to the first 3 principle components are shown in Table S3.6. These include subpopulation parameters, intensity parameters and morphological parameters. To identify biologically active compounds, we compared three multi-parametric tests based on the 9 principle components: Mahalanobis distance, Chi-square and Wilks' lambda test (see Methods). Each test was validated manually (by eye) by comparing control and compound-treated images. Mahalanobis distance at α=0.05 came the closest to visual scoring a treatment as having an effect on phenotype. Figure 3.3b shows the false positives and false negatives for the different statistical tests. From a total 29 (test compounds) × 6 (concentrations) = 174 conditions, 60 conditions were identified as active when we set the significance level α=0.05 for Mahalanobis distance. The Mahalanobis distance to control (DMSO) of all active compounds is shown in Figure 3.3c. Table S3.7 lists the active compounds and the concentrations at which a statistically significant effect on the phenotype was detected. Of all the 29 compounds that were tested, 21 induced a significant phenotypic change at least at one tested concentration.

3.2.4 Concentration dependent phenotypic trajectories of biologically active compounds

We found for many biologically active compounds, that data points seem to move away from negative controls in trajectories with increasing concentration (Figure 3.3a). This is shown clearly in a PCA plot of the first 2 principle components (Figure 3.4a). Interestingly, the trajectories of different active compounds separate the most at medium concentrations but converge at higher concentrations. This may be explained by the fact that at high concentrations severe toxicity is induced which inhibits not only cell invasion and branching but also proliferation and growth, leading to a similar phenotype. This is shown for the compounds Arq 197, dasatinib, entinostat, and sorafenib tosylate at the concentration of 10μM (Figure 3.4b). At the lowest concentration (0.03μM), except for dasatinib that already induced apparent inhibition of network formation, all other compounds had only a marginal effect on cellular network. Distinct phenotypes of different compounds were observed at the medium concentration of 0.316μM. Dasatinib inhibited branching but not prolifera-
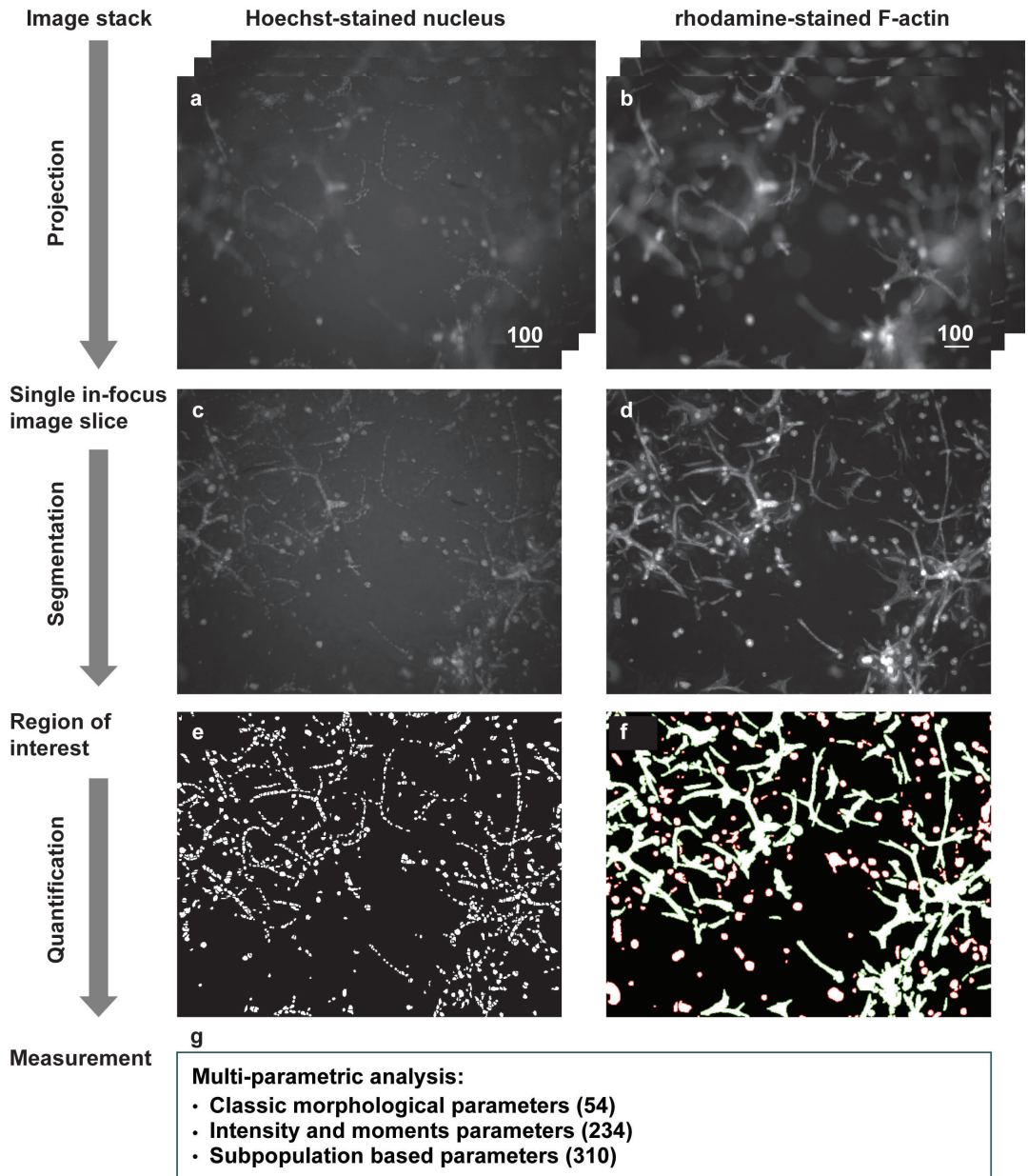
Figure 3.2: Stepwise demonstration of the image analysis method. Images were produced by a BD pathway 855 microscope in the wide-field mode with 4× air objective (NA=0.16). The scale bar represents 100μm. (a-b) Image stack obtained from the Hoechst stained nuclei channel and rhodamine stained F-actin channel. (c-d) In-focus 2D images projected from the stacks of Hoechst stained nuclei channel and stacks of rhodamine stained F-actin channel. (e) Binary nuclear mask after segmentation by Watershed Masked Clustering. (f) Binary cellular mask after segmentation by local Niblack algorithm. The green contour represents branched and interconnected complex networks. The red contour represents spherical multi-cellular structures. (g) Phenotypic parameters measured for each well of the 384-well plates are categorized in three classes: 54 classic morphological parameters, 234 intensity and moments parameters, and 310 subpopulation based parameters. See Table S3.2-S3.5 and Methods for further description.
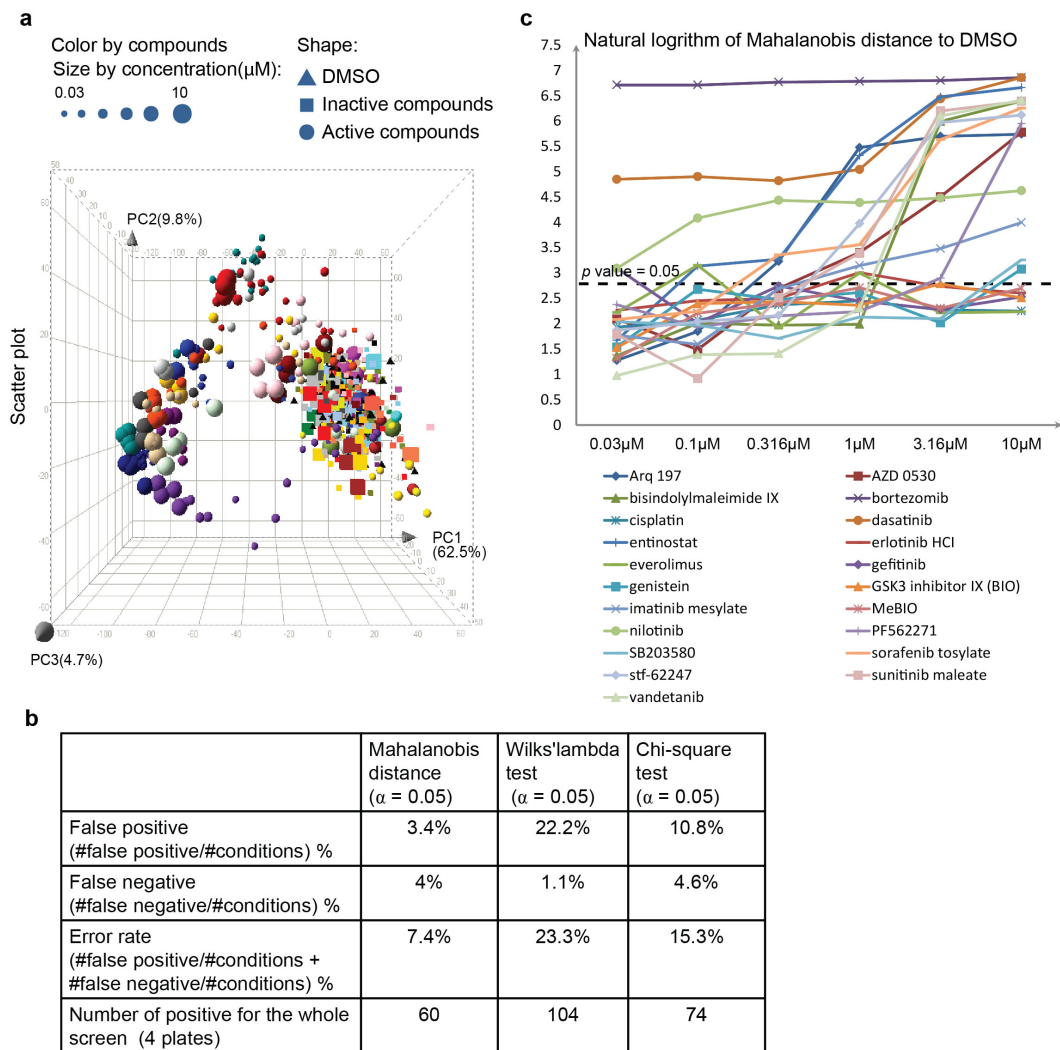
Figure 3.3: Identification of biologically active compounds. (a) 3D PCA plots of all compounds and concentrations. Compounds are marked with different colors and the concentration is represented by the size of data points. Control, active compounds and inactive compounds identified by Mahalanobis distance were represented by different shape of data points. Percentages of data variation preserved in each principle component are shown with each axis. (b) Comparison of three multi-parametric tests for the identification of biologically active compounds. "positive" indicates correctly identified active concentration of a test compound ($p - value \leqslant \alpha$). "False positive" indicates the concentration which is identified as active but no obvious difference was observed compared to control images. "negative" indicates correctly identified inactive concentration of a test compound ($p - value > \alpha$). "False negative" indicates the concentration which is identified as inactive but obvious differences were observed compared to control images. "#" means "number of". (c) Natural logarithm of Mahalanobis distance to DMSO control of all active compounds. Compounds are marked with different colors and shapes. Black dashed line corresponds to $p - value = 0.05$.

tion so that bigger cell clusters were formed, while entinostat induced much thinner branches compared to the control (Figure 3.4b). Sorafenib tosylate and Arq 197 caused formation of much shorter branches, indicative of inhibition of invasion. Most strikingly, we found that trajectories of compounds that inhibit the same biological target were more similar to each other than to trajectories of compounds with different targets (Figure 3.4a), indicating that phenotypic development of 4T1 cells is effected by different biological targets in characteristic ways and that we can identify this with our uHCA methodology.

3.2.5 Trajectory modeling and phenotypic pattern recognition

To further characterize the different phenotypes, we used a 2nd order polynomial regression modeling to build the trajectory for each identified active compound. First, we investigated data variation for each of the 9 principle components. We compared the data between control and active compounds using a two sample Kolmogorov-Smirnov (KS) test (Figure S3.3, Supplementary note). The principle components with no significant difference between negative control and active compounds, or equal data variation or bigger variation in negative controls than in active compounds, were excluded to avoid overtraining, resulting in retaining only the first 2 components. Next, a 2nd order polynomial regression model of the trajectory for each compound was trained and the difference between the trajectories of two compounds i and j was computed based on the coefficient of determination R^2:

$$Distance(i,j) = \sqrt{1 - R^2_{mean}} \tag{3.1}$$

$$R^2_{mean} = \frac{R^2_{(i,j)} + R^2_{(j,i)}}{2} \tag{3.2}$$

where $R^2_{(i,j)}$ is the coefficient of determination to indicate how well data points of compound j fit a trajectory which is modeled for compound i (see also Supplementary note). Subsequently, a hierarchical clustering with complete linkage [18] was applied on the distance matrix defined above (Figure 3.5a). Consistent with our finding in the PCA analysis, compounds with the same biological target cluster together but compounds with different targets are separated. To further validate our hypothesis that phenotypic responses are specific to the biological target that is inhibited, we applied classification. Five classes of compounds were defined based on their biological targets (Figure 3.5b). Only active concentrations were taken into account, and compound classes with less than 15 data points were not included in order to avoid the curse of dimensionality. We tested several classification algorithms on the 598 $z-scores$ space, including k-nearest neighbor classification, linear Bayes normal classification [19–21], quadratic Bayes normal classification [19,21], nearest mean classification, support vector machine classification (SVC) [22,23] with different kernels, and Fisher linear classification [19,21,24]. For each classification method, a forward feature selection with the criterion Mahalanobis distance was firstly performed and
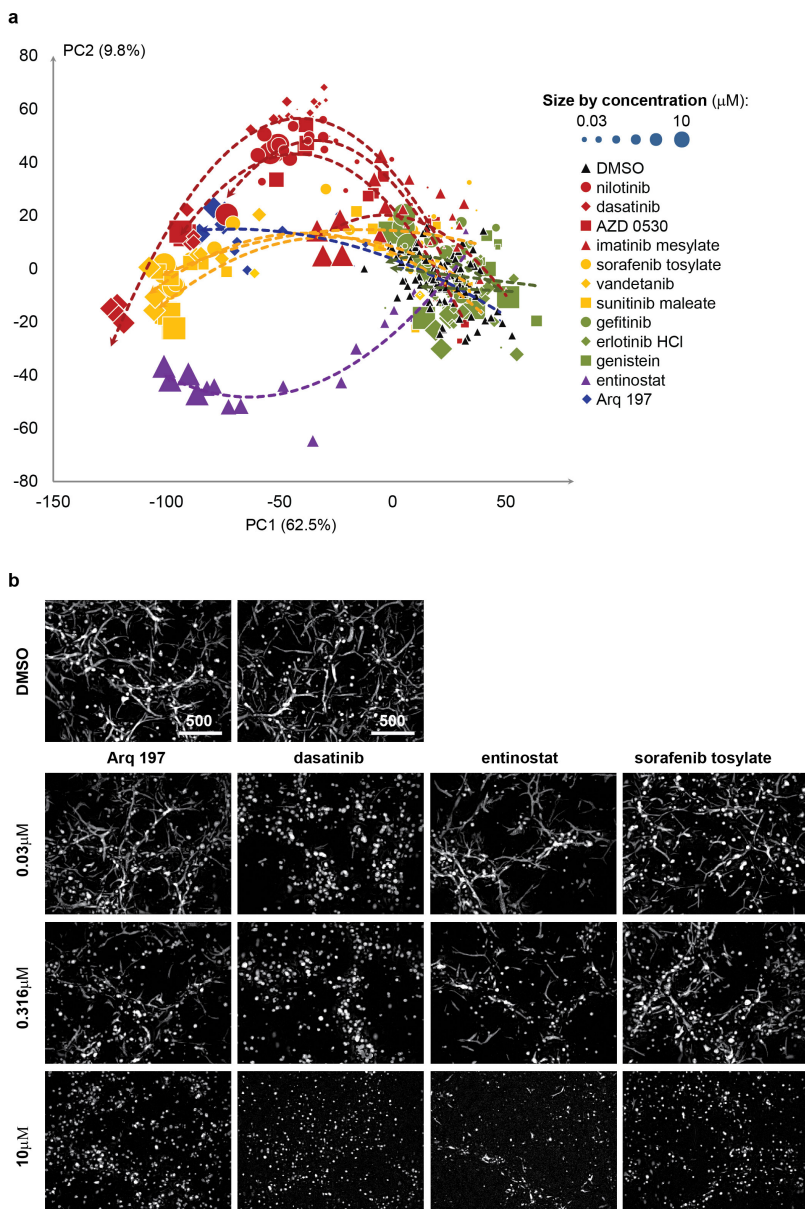
Figure 3.4: 2D PCA plot of phenotype profiles for various active compounds and their concentration dependent phenotypic trajectories. (a) 2D PCA plot of phenotype profiles for negative control (DMSO) and 12 active compounds at different concentrations. Percentages of data variation preserved in each principle component are shown with each axis. Compounds with the same biological target are colored identically. Red: BCR-ABL target inhibitor; Yellow: VEFGR inhibitor; Green: EGFR inhibitor; Purple: HDAC inhibitor; Blue: c-MET inhibitor. Concentration is represented by the size of data points. The trend lines were added for each effective compound using 2nd polynomial regression models. (b) Comparison of microscope images of four example compounds with two DMSO control images. Each compound has a different biological target. 2D projected images from the rhodamine stained F-actin channel are shown here. The scale bar represents $500 \mu m$.

only the selected features were used for classification. To evaluate the performance of feature selection and classification, 10-fold cross-validation was used (Figure 3.5c). For each of 10 tests, a classification error rate (%) was calculated as follows:

$$\sum \frac{number\ of\ erroneously\ classified\ objects\ per\ class}{size\ of\ class} \times prior\ probability\ of\ class$$

(3.3)

where the prior probability is equal for each class (20%). According to the feature selection result, we found that the most frequently selected discriminative parameters included morphology- and intensity-based parameters and parameters from both the whole population and two subpopulations. The classification result is shown in Figure 3.5c. The lowest mean classification error rate of 12.8% was obtained from a linear kernel based SVC with 7 features selected. 1-nearest neighbor classification also showed a relatively high classification accuracy with an error rate of 13.1% when 6 features were selected. With this high classification accuracy of different classification methods, further evidence is provided for the potential of our phenotypic pattern recognition method to identify the different biological targets of compounds.

3.2.6 Contribution of the different parameter classes to the classification

To recover the heterogeneity of responses of 4T1 cells to the different compounds within the cell population, we have analyzed different multi-cellular subpopulations and quantified phenotype parameters for each subpopulation separately. In order to establish the value of analyzing these subpopulations separately, we repeated the above analysis excluding subpopulation-related parameters, collecting only 288 parameters from each well. After normalization, we applied PCA and retrieved 7 principle components which preserved 90% data variation. After computing the trajectories for the active compounds and hierarchical clustering as described above, the result showed that excluding subpopulation parameters caused a failure to co-cluster all BCR-ABL inhibitors together, and EGFR inhibitors together (Figure S3.4a). We also repeated classification without subpopulation parameters (Figure S3.4b) and the classification accuracy was decreased compared to Figure 3.5c. SVC with linear kernel still performed relatively better than other classification methods when 8 features were selected, but the error rate increased to 16.5% compared to the error rate of 12.8% when subpopulation parameters were used. Using 1-nearest neighbor classification, the lowest error rate was obtained when 11 features were selected resulting an error rate of 16.6%.

We also repeated the analysis without moments and intensity parameters. Only morphological parameters were measured from the whole object population and the two subpopulations, resulting in 152 parameters extracted from each well for classification (Figure S3.4c). The classification results of SVC with linear kernel showed that omission of moments and intensity parameters increased the misclassification error to 15.4%. These results show that simplification of the analysis by omitting either

a

Compound	Biological Target	Cluster
AZD 0530	SRC, BCR-ABL	1
dasatinib	BCR-ABL, SRC	1
nilotinib	BCR-ABL	1
imatinib mesylate	PDGFR, KIT, BCR-ABL	1
Arq 197	c-MET	2
SB203580	p38MAPK, PKB	2
stf-62247	autophagy inducer	2
sunitininb maleate	PDGFR, KIT, VEGFR	2
vandetanib	EGFR, VEGFR	2
sorafenib tosylate	PDGFR, VEGFR	2
PF562271	FAK	2
bisindolylmaleimide IX	PKC, RSK, GSK3	2
bortezomib	proteasome inhibitor	3
entinostat	HDAC	3
cisplatin	genotoxic, DNA-crosslinker	4
GSK3 inhibitor IX (BIO)	GSK3	4
MeBIO	negative control for BIO, AhR-ligand	4
erlotinib HCl	EGFR	4
gefitinib	EGFR	4
genistein	EGFR, Caspases, PPAR gamma, Topoisomerase	4
everolimus	mTORC1, immunosuppressant	5

Natural logrithm of $\sqrt{(1 - R^2_{mean})}$

b

Class index	Target	Compounds	Number of data points
1	BCR-ABL	AZD 0530, dasatinib, nilotinib, imatinib mesylate	12 24 24 12
2	VEGFR	sunitinib maleate, vandetanib, sorafenib tosylate	12 8 16
3	EGFR	erlotinib HCl, gefitinb, genistein	8 8 4
4	HDAC	entinostat	20
5	c-MET	Arq 197	16

c

Error rate vs. Number of selected features for classification

- Linear Bayes normal
- Quadratic Bayes normal
- Nearest mean
- K-nearest neighbor (K=1)
- K-nearest neighbor (K=2)
- K-nearest neighbor (K=3)
- Fisher linear
- SVC(kernel "linear")
- SVC(kernel "polynomial", order=2)
- SVC(kernel "polynomial", order=3)
- SVC(kernel "radial_basis", order=1)
- SVC(kernel "radial_basis", order=2)
- SVC(kernel "radial_basis", order=3)

Figure 3.5: Characterization of cellular phenotypes by clustering and classification. (a) Hierarchical clustering result using a distance matrix based on Equation 3.1. The scale of dendrogram is the natural logarithm of $\sqrt{(1 - R^2_{mean})}$. (b) Five defined classes of test compounds and corresponding compounds and number of data points. (c) Classification result using multiple classification methods. Feature selection with search algorithm "forward" and criterion "Mahalanobis distance" was applied to detect the optimal number of features. For each classification method and each number of selected features, 10-fold cross-validation was applied, resulting in 10 error rates. The average error rates are shown in the chart with standard deviation as error bar. SVC means support vector machine classification.

subpopulation parameters or intensity and moments parameters e.g. to decrease time for computation, compromises the quality of the analysis.

3.2.7   Computational efficiency

Our complete method takes approximately 260 minutes of computational time for 384 wells on an Intel i7 3770 model with 4 GB of RAM with Windows XP professional 2002. As the costs of these computational systems are relatively modest, we do not regard the computation time of our method as a major drawback because the load of the analysis can easily be spread across multiple computers.

3.2.8   Comparison to other analysis methods

In order to compare the performance of our newly developed analysis method with other published methods, we have analyzed our images with PhenoRipper [25] and CellProfiler [26]. PhenoRipper is a platform using a segmentation-free approach: it breaks down images into small blocks and clusters the blocks to different types according to the pixel intensity distribution. Then it quantifies images by proportions of different types of block. As the quantification is not executed based on segmentation, this approach is highly computational efficient, using ~5 minutes to analyze four complete 384 well plates of our data set. As correct block size is essential, we tested different block widths ranging from 20 to 80 pixels. However, after plotting profiles of all data points on a 2D multidimensional scaling (MDS) plot using this platform, we found that the distances between compounds in the plot did not reflect the similarities or dissimilarities in the images observed by eye (Figure S3.5a-S3.5c). For example, the negative control (DMSO) data points are closely located to data points of the positive control Arq 197 with concentration $3.16\mu M$ in the MDS plot even though these two conditions showed clearly discrete phenotypes.

CellProfiler is segmentation dependent software. It is able to calculate morphological parameters including Zernike moments, object intensity parameters, topological parameters, texture parameters and image intensity parameters. In total we measured 395 parameters (including the per-image mean and standard deviation for object measurement) for each well using CellProfiler. After robust $z-score$ normalization, we applied PCA and plotted the concentration trajectories for the active compounds. However, these trajectories were not biological activity specific (Figure S3.6a). This is also reflected in the hierarchical clustering result which did not show co-clustering of compounds with the same biological target (Figure S3.6b). Finally we applied Mahalanobis distance (α=0.05) to identify the active concentrations of the biologically active compounds, and those concentrations were used for classification (still five classes). The lowest error rate of 26.8% was obtained when SVC with linear kernel was applied with 14 parameters selected. Compared to the classification error rate of 12.8% obtained using our method, this higher error rate indicates that subpopulation information which is lacking in CellProfiler plays a very important role in our method and should be taken into account for compound characterization.

3.2.9   Reproducibility of the methodology

To validate the reproducibility of our multi-parametric profiling platform, the 4T1 cell screen was repeated independently on a different occasion. Twelve biologically active compounds (Table S3.8) were included in this screen. We used the same data analysis pipeline as described above for the multi-parametric profiling of those compounds. After robust $z-score$ normalization and PCA, a 2D PCA plot with control was generated (Figure S3.7a). A large extent of similarity was observed with Figure 3.4a, which validates the high reproducibility of our platform. Subsequently, we applied Mahalanobis distance (α=0.05) to identify the biologically active compounds and corresponding active concentration (Table S3.8). Compounds at concentrations which were detected as active in this screen but not active in the first screen are highlighted with bold style. 7 out of 49 concentrations were identified differently between the two independent experiments. Finally we defined the classes according to the biological targets and applied classifications. (Figure S3.7b). The error rate was 17.8% when we used 1-nearest neighbor classification and 9 features were selected. The best classification accuracy was obtained using quadratic Bayes normal classification. The corresponding error rate was 15.1% when 8 features were used for classification. Excluding subpopulation parameters increased the classification error rate of quadratic Bayes normal classification to 19.1%, while the lowest classification error of 1-nearest neighbor classification increased to 19.8%. We also repeated the classification without moments, wavelet and intensity parameters. The lowest classification error rate of quadratic Bayes normal classification and 1-nearest neighbor classification were 17.6% and 19.0%, respectively, which is higher than the 15.1% and 17.8% obtained when these parameters were included in the analysis, confirming their importance for the quality of the analysis.

3.2.10   Other applications of the methodology: classification of human breast cancer cell lines

To investigate the wider applicability of our multi-parametric image analysis platform, we used it to classify 44 known human breast cancer cell lines (Table S3.9) that have been categorized as basal-A, basal-B, luminal or luminal/ERBB2+ based on their gene expression profiles [27–30]. The cell lines were cultured in ECM-rich hydrogel in 384-well high-content imaging micro plates, with each cell line having 3 or 6 replicates (Table S3.9 and Methods). Image stacks were obtained from the Hoechst stained nuclei channel and the rhodamine stained F-actin channel as described for the experiments with the 4T1 cells. We firstly applied our platform for multi-parametric profiling (including automatic identification of spherical cell clusters and branched network subpopulations) on the in-focus projections of z-stack images. After robust $z-score$ normalization, we applied supervised forward feature selection with criterion Mahalanobis distance and different classification methods for categorizing basal-A, basal-B, luminal and ERBB2+ cell lines. 10-fold cross-validation was used to assess classification performance (Figure 3.6a). The mean classification error was lowest

when 8 features were selected and 1-nearest neighbor classification method was used, resulting in an error rate of 5.9%. The selected features include intensity parameters from the whole population and subpopulations, morphological parameters and topological parameters from the whole population and subpopulations. Based on these 8 selected features, a PCA was applied (Figure 3.6b), which clearly shows the separation between the various human breast cancer cell classes. If we used only classic morphological parameters, moments, and wavelet parameters for the whole population only, thus excluding parameters from the subpopulations, we obtained a significantly increased classification error rate (24.9%) when a 1-nearest neighbor classification method was used. Similarly, omitting wavelets, moment and intensity features from our feature set increased the error rate to 35%.
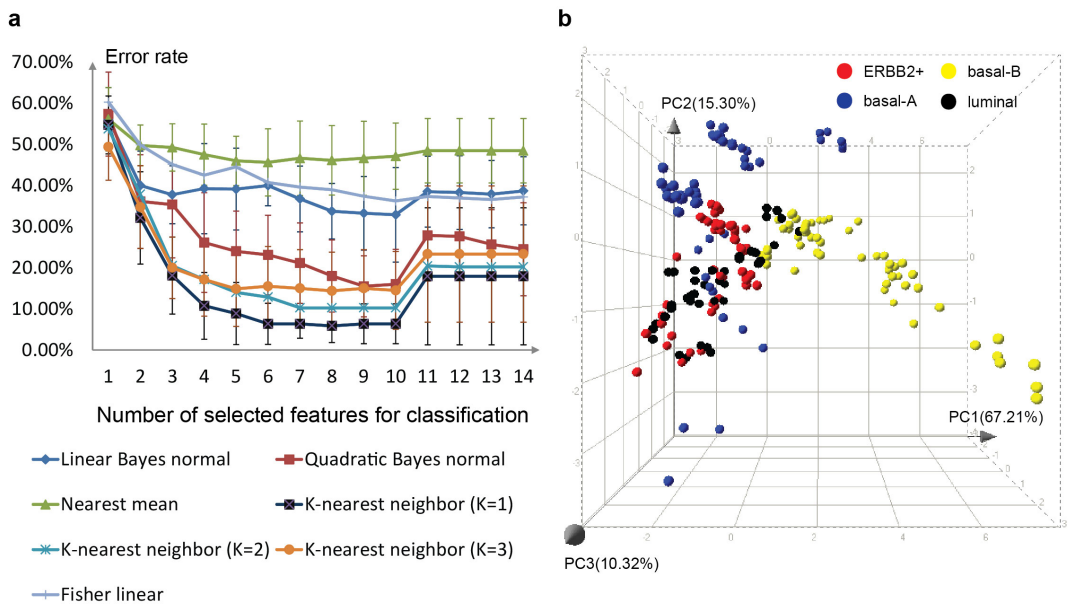
Figure 3.6: Classification of human breast cancer cell lines. (a) Classification result using 7 classification methods. Feature selection with search algorithm "forward" and criterion "Mahalanobis distance" was applied to detect the optimal number of features. For each classification method and each number of selected features, 10-fold cross-validation was applied, resulting in 10 error rates. The average error rates are shown in the chart with standard deviation as error bar. (b) According to the cross-validation result, the smallest error rate was achieved when 8 features were selected. A 3 dimensional PCA plots was generated based these 8 selected features. Percentages of data variation preserved in each principle component were shown with each axis. Different categories of breast cancer cells are colored differently to show the separation between the various human breast cancer cell classes.

3.3   Discussion and conclusions

In this study, we developed a new methodology for 3D cell culturing in conjunction with high-throughput imaging for the characterization of compounds according to their effect on the phenotype of cultured micro-tissues. First, we developed a novel method for 3D cell culturing in 384-well plates and next an image analysis platform, which automatically measures 598 parameters, including morphological parameters, fluorescent staining intensity parameters, topological parameters, spatial attributes of the cellular nuclei, subpopulation information, as well as image moments and wavelets using both segmentation dependent and independent approaches. This platform provides a very detailed level of phenotypic description which opens the door to system-level study of the full activity of test compounds.

Our method is the first for high-content analysis of 3D micro-tissues that is capable of classifying complex phenotypes of 3D cell cultures and can do this in an automated, high-throughput fashion. This represents an important tool e.g. for the classification of new and existing pharmacologically active compounds, but also for the characterization of different tumor sub-types. It is also the first quantification method which takes into account heterogeneous cell behavior. Our method requires only an initial human-based step (defining branched and spheroidal structures) to train an automatic classifier. This simple procedure is then followed up by a largely automated method to classify multi-cellular structures and retrieve the subpopulation related parameters. The classifier can be applied for a broad class of phenotypes. For example, we applied it to classify 44 different breast cancer cell lines successfully (our manuscript) and also to invasive prostate cancer cells and invasive lung cancer cells (not yet published).

This platform can be used for many different applications including identification of biologically active compounds, concentration dependent trajectory construction, clustering and classification. We revealed that trajectories of different compounds varied and are likely to be specific for biological activities. We also demonstrated that our platform can be used to correctly classify a collection of human breast cancer cell lines into known subclasses. Although the group of Bissell et al. has pioneered automated phenotyping of human breast cancer cell lines [31], a complete classification of these cell lines could only be achieved so far with more elaborate and expensive techniques such as gene expression analysis. Our platform present an opportunity to identify new treatments/drugs aimed at reverting cancer cells to phenotypes with better prognosis and/or drug responsiveness, rather than direct tumor cell killing per se.

For each application, we analyzed the contribution of different parameter classes to the phenotype characterization of the 4T1 cell cultures. With PCA, we found that the parameters which contributed most to the variation in the first 3 principle components included subpopulation parameters, intensity parameters, morphological parameters and topological parameters. For classification of compounds with different biological activities, the feature selection method that we used selected morphological and intensity parameters, and parameters from both total populations and subpop-

ulations. Altogether, these findings indicate that to systematically study phenotypes associated with modulation of different cellular pathways, it is necessary to quantify images with a full spectrum of phenotypic information. It will not only reveal details that are otherwise not resolved, but also enable adaption to different cell lines or new biological questions. Although the use of 598 parameters may seem redundant initially, it enables us to analyze 3D-cellular phenotypes under a wide variety of conditions, while our feature selection methodologies automatically identify those features that contribute most to the separation and characterization of the particular phenotypes under study.

Compared to other methods, we demonstrated that our methodology offers significant advantages in terms of recognition of specific phenotypes within still reasonable computational demands. While some other methods such as PhenoRipper [25] are suitable for analyzing sub-cellular data in 2D imaging-based assays, they are not appropriate for the type of images used in this study. Another advantage of our methodology is that it can be incorporated in the user friendly, freely available ImageJ environment written in Java, and therefore can be run with various operating systems (Linux, Windows, Mac OS X).

3.4 Methods

3.4.1 Cell culturing, fluorescence staining, and image acquisition

To generate 3D micro-tissues, mouse triple negative breast cancer cells (4T1) were cultured in a hydrogel containing extracellular matrix proteins and inert hydrogel (O4-Gel, OcellO, Leiden, The Netherlands) for 4 days in 384-well high-content imaging micro plates. 24 hours after seeding, cells were exposed to 29 different compounds with different biological activities (Table S3.1) at 6 different concentrations (0.03μM, 0.1μM, 0.316μM, 1μM, 3.16μM, 10μM) in quadruplicate. After 72 hours of exposure, the cultured micro-tissues were fixed and stained with Hoechst 33258 (final concentration 0.4 μg/mL) and rhodamine-phalloidin (final concentration 0.1μM) to visualize nuclei and F-actin, respectively.

For the classification of the 44 breast cancer cell lines (Table S3.7), these were cultured similarly as described above for the 4T1 cells with each cell line having 3 or 6 replicates. 96 hours after seeding, they were stained with Hoechst and rhodamine-phalloidin as described above. The human breast cancer cell lines were from ATCC (Manassas, VA, USA) or as described [32] and provided to us by Prof. Dr. John A. Foekens and Dr. John W. Martens from Erasmus University Medical Center-Daniel den Hoed Cancer Center, Rotterdam, The Netherlands.

3.4.2 Image analysis for multi-parametric phenotype profiling

a. Overview

For each well of a 384-well plate, 2 channels (corresponding to the rhodamine stained F-actin and the Hoechst stained nuclei) of 16-bit image stacks were collected using

an automated microscope system: BD pathway 855, equipped with a 4× magnification/0.16 NA UPlanSApo objective. For each image slice, pixel size was 1.6μm and step size in z-direction was 50μm. For quantitative phenotype profiling, image analysis was performed on both channels of image stacks in four steps. First, each stack was projected to a 2D plane. Next, two segmentation algorithms were applied on nuclei and F-actin stained channels separately to retrieve binary masks of the nuclear and cell cluster regions. After binary multi-cellular objects were obtained, a user predefined classifier was used to distinguish two cellular subpopulations: spherical objects and branched objects. In the last step, morphological parameters and intensity based parameters were measured on both the whole population and subpopulations.

b. 2D projection

In order to achieve high-throughput in our methodology, we used wide-field microscopy for screening as its imaging process is much faster than confocal laser scanning microscopy. However, due to its limited depth of field, each image slice includes both in-focus regions and out-of-focus regions of specimen that are bigger than the depth of field. To extract only the in-focus information, a free open source plugin of ImageJ "Stack Focuser" was used to compose 2D image slices by projecting only in-focus regions from each slice of image stacks. First, a median filter was used to remove the noise. We empirically defined the filter kernel size as 3×3 pixels; using a bigger size would affect object size. Next, a Sobel filter [33] was used for edge detection and subsequently a maximum filter was applied to map the local focal strength, by taking the maximum value in a kernel of the specified size. In this project, we defined the kernel size as 10×10 pixels, which is slightly bigger than the average size of the nuclei (8×8 pixels). In the last step, for each coordination (x,y) the pixel value from the slice with the highest focal strength was copied to the composed 2D image (Figure S3.1a, S3.1d).

c. Segmentation

Two different segmentation algorithms were used for the two image channels. For the Hoechst stained nuclear channel, watershed masked clustering (WMC) [12, 13] was applied to retrieve the binary masks for individual nuclear regions. This algorithm first generates watersheds on the Gaussian filter ($\sigma = 2.0$) convolved images to separate the adjacent nuclei into individual compartments (Figure S3.1b). Convolving with a Gaussian filter prevented the influence of noise from causing artificial local maxima. Next, K-means clustering was then applied on the images prior to convolving to refine the region of the nucleus in each compartment (Figure S3.1c).

Before segmentation on the rhodamine stained F-actin channel, median filter (kernel size 3×3 pixels) and rolling ball (radius = 30 pixels) were applied to remove the background and reduce the noise level (Figure S3.1e). The radius of rolling ball was chosen to be slightly bigger than the biggest cell cluster. Next, the Local Niblack algorithm [14] was used to define regions of cell clusters (Figure S3.1f). This algo-

rithm calculates a threshold T for each pixel (x, y) according to the intensity mean $mean(x, y)$ and standard deviation $\sigma(x,y)$ within a specified kernel size (we defined this size as 30×30 pixels which is slightly bigger than the biggest cell cluster) centered at pixel (x, y):

$$T(x, y) = mean(x, y) + k * \sigma(x, y) \tag{3.4}$$

where k indicates how much the standard deviation influences the threshold and its value was determined empirically.

d. Subpopulation classification

To automatically identify the spherical cell clusters and branched networks in the 3D micro-tissues, a classifier was trained. Five test image stacks were randomly selected from DMSO controls which contain both spherical structures and tubular structures. 2D projection and segmentation were applied for the rhodamine stained F-actin channel to define the binary masks of multi-cellular regions (Figure S3.2a). We defined each connected binary region (in 8-connectivity) as a multi-cellular object and in total 1318 binary multi-cellular objects were obtained. Next, we manually identified spherical objects (Figure S3.2b) and branched objects (Figure S3.2c). For each of binary multi-cellular masks, a skeleton (Figure S3.2d) was retrieved using a "Skeletonize" plugin of ImageJ, which iteratively removes pixels from the edge of object in the binary image, reducing them to the single pixel wide skeleton. We labeled each pixel of the skeleton with one of the following categories depending on their neighbors in 8-connectivity:

- End point: the pixel with only 1 skeleton pixel in neighbor.
- Junction point: the pixel with more than 2 skeleton pixels in neighbor.
- Single junction point: the connected junction points in 8-connectivity.
- Slab point: the pixel with 2 skeleton pixels in neighbor
- Triple point: the pixel with 3 skeleton pixels in neighbor.
- Quadruple point: the pixel with 4 skeleton pixels in neighbor.

19 parameters (Figure S3.2e) were finally calculated from each cellular cluster, including basic morphological parameters and topological parameters from the skeletons. The topological parameters were calculated based on the category of each skeleton pixel defined above. Single branch is defined as the part of skeleton between single junction points, end points, or single junction point and end point. Branch length of a single branch is defined as the sum of Euclidian distance between two adjacent skeleton pixels over all pixels on the single branch. The definition and equation of the 11 morphological parameters are defined as follows:

M1 Area: the number of pixels in the binary object.

M2 Solidity: $\frac{\text{Area}}{\text{Area of convex hull}}$

M3 Major axis [33]: the primary axis of the best fitting ellipse derived from 2nd order moment η

M4 Minor axis [33]: the secondary axis of the best fitting ellipse derived from 2nd order moment η

M5 Axis ratio: $\frac{\text{Major axis}}{\text{Minor axis}}$

M6 Perimeter: the length of the contour of the binary object.

M7 Roundness: $4 \times \frac{[\text{Area}]}{\pi \times [\text{Major Axis}]^2}$

M8 Equivdiameter: the length of the diameter of the perfect circle that had the same area as the binary object.

M9 Elongation [33]: measures how much the shape must be compressed along its major axis in order to minimize the extension.

M10 Extension [33]: measures how much the shape differs from the circle.

M11 Dispersion [33]: the minimum extension that can be attained by uniform compression of the shape

where $\eta = \frac{\mu_{pq}}{\mu_{pq}^\gamma}$, $\gamma = \frac{p+q}{2} + 1$, $\mu_{pq} = \sum_0^{width-1} \sum_0^{height-1} (x - \bar{x})^p (y - \bar{y})^q B(x, y)$. *width* and *height* indicate the width and height of the object in pixels. $(\bar{x}, \bar{y})$ is the center coordinate of the binary object $B(x, y)$.

For classification, we used a Matlab toolbox, PRTools. Firstly, feature selection was performed to avoid the curse of dimensionality, using search algorithm 'forward' with criterion "Mahalanobis distance" [34]. For this two subpopulations situation, Mahalanobis distance criterion optimizes:

$$J_{maha} = (\mu_2 - \mu_1)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_2 - \mu_1) \tag{3.5}$$

where μ_1, μ_2, Σ_1, Σ_2 indicate the mean vector of subpopulation 1 data, mean vector of subpopulation 2 data, the covariance matrices of subpopulation 1 and 2. To define the optimal number of features, multiple classification methods were applied and the 10-fold cross-validation was applied to evaluate the number of features, as well as the accuracy of each classification method. The classification methods tested were k-nearest neighbor classification (k=1), linear Bayes normal classification [19–21], quadratic Bayes normal classification [19, 21], nearest mean classification and Fisher linear classification [19, 21, 24]. Figure S3.2f showed the average error rate obtained from 10-fold cross-validation for each classification method with a certain number of features. The classification error rate was calculated according to Equation 3.3. In this study, prior probability was defined equally for each of two subpopulation classes. The result revealed that when two features were selected, linear Bayes normal classification gave the least average error rate of 1.67%. The two features which were selected most often are perimeter and major axis. Finally, a linear Bayes normal classifier based on these two parameters was trained and embedded in the platform to automatically classify all cellular binary masks of the whole experiment. To increase the robustness of this platform, the user can also define the classification method and

features for classification empirically.

e. Quantification

Quantification algorithms were assembled from the literature [15, 16] and incorporated into an ImageJ plugin to extract different phenotypic parameters from the segmentation results or projected images. According to the function, parameters can be categorized into three groups: 1) classic morphological parameters (M), 2) moment parameters of the binary objects as shape descriptor (S) and 3) intensity-based parameters (I).

1. (***M1-M29***)**Classic morphological parameters** include a series of shape properties that were calculated for each binary object. A binary object was defined as a connected binary region (8-connectivity) on the binary image.

   - Basic morphological parameters measured for each binary multi-cellular object: These parameters describe the basic dimensional and geometry information of the cell clusters. As noted in session d above, we already explained the Area, Solidity, Major axis, Minor axis, Axis Ratio, Perimeter, Roundness, Equivdiameter, Elongation, Extension and Dispersion. The remaining parameters were defined as following:
     - ***M12*** Feret's Diameter: the longest distance between any two points along the selection boundary.
     - ***M13*** MinFeret: the minimum distance between any two points along the selection boundary.
   - Skeleton parameters measured for each binary multi-cellular object: These parameters measure the properties of skeletons obtained after skeletonization of binary objects (Figure S3.2d). As noted in session d above, we already explained the basic properties: end point, junction point, single junction point, slab point, triple point, quadruple point, single branch and branch length. Parameters shown as following were measured based on those properties:
     - Number of end point (***M14***), number of junction point (***M15***), number of single junction point (***M16***), number of single branch (***M17***), number of triple point (***M18***) and number of quadruple point (***M19***).
     - Branch length parameters: maximum single branch length (***M20***), average branch length (***M21***), average length of the single branch between two single junction points (***M22***), accumulated branch length (***M23***).
   - The number of nuclei (***M24***) and the number of multi-cellular structures (***M25***).
   - Total area of nuclei (***M26***) and total area of cells (***M27***): This is defined as the sum of all binary objects' area in the Hoechst stained nuclei channel and the rhodamine stained F-actin channel, respectively.

- Density of nuclei (**M28**) and density of cell clusters (**M29**): For each binary object, we defined its local density as the mean of its peer-to-peer distance to the other binary objects. The distance was calculated between the gravity center coordinate of two objects. The density for whole image was defined as the average of all the objects' local density.

2. (**S1-S49**)**Moment parameters** measured for each binary multi-cellular object: All parameters in this list were calculated on the binary masks as shape descriptors. We only considered the moments which are invariant under translation and rotation.

   - Zernike moment parameters (**S1**-**S49**) [16, 35]: Calculating the Zernike moments was composed of three stages. In the first stage, the center coordinate was calculated for each binary object. Next, we calculated the corresponding radius and mapped the pixel (x, y) of the object to a unit circle. Finally, Zernike moments up to degree 12 were calculated ($n \leq 12$), providing 49 numbers for describing each binary object.

3. (**I1-I140**)**Intensity-based parameters**: In this list, all parameters were calculated on the projected grey value images.

   - Total intensity of the projected images for the Hoechst stained nuclei channel (**I1**) and rhodamine stained F-actin channel (**I2**).

   - The average intensity calculated on the projected images of the Hoechst stained nuclei channel (**I3**) and rhodamine stained F-actin channel (**I4**).

   - The standard deviation of the intensity calculated on the projected images of the Hoechst stained nuclei channel (**I5**) and rhodamine stained F-actin channel (**I6**).

   - The maximum intensity value calculated on the projected images of the Hoechst stained nuclei channel (**I7**) and rhodamine stained F-actin channel (**I8**).

   - The minimum intensity value calculated on the projected images for the Hoechst stained nuclei channel (**I9**) and rhodamine stained F-actin channel (**I10**).

   - Zernike moment parameters of the projected images for the rhodamine stained F-actin channel (**I11-I59**). Different from the Zernike moments calculated on binary objects, the intensity value was also taken into account. We firstly normalized the projected images to a unit circle with radius of 1. Next, the Zernike moments polynomials up to degree 12 were calculated.

   - Hu moment [36] calculated on the projected images for the rhodamine stained F-actin channel (**I60-I66**): 7 Hu moments were calculated.

- The Gabor wavelet parameters [35, 37] calculated on the projected images for the rhodamine stained F-actin channel (**I67-I136**): The Gabor wavelet parameters are formed by a set of multi-scale and multi-orientation coefficients to describe texture variations in an image. The method to calculate 70 numbers (the mean μ and the standard deviation σ of the magnitude of coefficients from the Gabor wavelet transformation when we used four scale $S = 4$, and six orientation $K = 6$) is describe in the literature [35].

- Intensity-based parameter measured for each multi-cellular structure: We calculated the average intensity (**I137**), standard deviation of the intensity (**I138**), maximum intensity (**I139**) and minimum intensity (**I140**) for each binary object masked region on the projected grey value image.

f. Phenotypic profiling

For each of the classic morphological parameters and moment parameters which were measured on the binary objects, the mean and standard deviation were calculated over the whole population, as well as the subpopulations (Table S3.2-S3.5), providing 598 parameters to profile phenotype of each well.

3.4.3   Identification of the biologically active compounds

a. Overview

After quantification, a workflow of multi-parametric analysis to identify the biologically active compounds was set up in KNIME. It is composed of three stages. Firstly, normalization was performed to remove artifact errors from the data and to allow comparison and combination of data from different plates in the screen. Next, we performed a PCA to extract the dimensions which contained most of the data variation. Then we compared different multivariate statistics to identify the biologically active compounds and their active concentrations (hit identification).

b. Normalization

A simple statistics robust $z - score$ [17] was calculated for each of 598 parameters and used for plate normalization and further data analysis:

$$robust\ z - score\ for\ each\ parameter = \frac{x_i - median_{negative}}{MAD_{negative}} \qquad (3.6)$$

where x_i is the original measurement on the ith well, $median_{negative}$ and $MAD_{negative}$ are the median and median absolute deviation of this parameter calculated for the negative control. Compared with $z - score$, robust $z - score$ is less sensitive to outliers.

c. Principle component analysis

PCA was performed to project the normalized 598 $z-scores$ to independent principle component space. We obtained 9 principle components which retained 90% of data variation. We plotted the data points (all compounds with different concentrations) on the first three dimensional PC space and 6 outliers were observed. These outliers were confirmed by visual inspection of the original image stacks and were due to errors occurring at the image acquisition stage. Since the outliers would result in bias in the PCA, they were removed and PCA was performed again on the outlier removed dataset. As the result, another 9 principle components were obtained which maintained 90% of dataset variation. The top 5 parameters which have the highest coefficient representing the influence of the respective input dimension to the first 3 principal components are shown in Table S3.6.

d. Multi-parametric test

For hits identification, we compared three commonly used multivariate tests: Mahalanobis distance [34, 38], Chi-square test [39] and Wilks' lambda [40]. They were performed on the 9 principle components obtained from the PCA. The Mahalanobis distance D calculates the similarity of a data vector to the negative control based on the correlation between variables:

$$D_i^2 = (x_i - \mu_{negative})^T \Sigma_{negative}^{-1} (x_i - \mu_{negative}) \tag{3.7}$$

where x_i is a 9 principle components vector of the ith well and $\mu_{negative}$ indicates mean vector of the negative control. $\Sigma_{negative}$ indicates the covariance matrix of the negative control. Chi-square calculates

$$\chi^2 = (x_i - \mu_{negative})^2 \tag{3.8}$$

When the variables are multivariate normal distributed, the squared Mahalanobis distance D^2 and χ^2 follow a Chi-square distribution with n degrees of freedom (n is number of variables). Wilks' lambda is a statistics used as a measure of the class center separation when classes are multinomial with identical covariance matrix:

$$\Lambda = \frac{det(S_{wg})}{det(S_{wg} + S_{bg})} \tag{3.9}$$

where det indicates determinant of a matrix, S_{wg} and S_{bg} represent within-group and between-group sum-of-square and cross-products matrix (SSCP) correspondingly, when two groups in comparison are defined as negative control and tested compound with certain concentration. For all three methods, a $p-value$ was calculated for each compound and each concentration and appropriate threshold of significance level α was calibrated by manually checking a selection of control and compound-treated images.

3.4.4   Software

For image analysis, ImageJ plugins were developed in-house, and written in Java. KN-IME was used to develop the hits identification pipeline including robust $z - score$ normalization, PCA and multivariate tests. Mahalanobis distance, Chi-square and Wilks' lambda test were coded in R snippets and integrated in KNIME. For phenotypic pattern recognition including clustering, classification and feature selection, we used the Matlab toolbox, PRTool.

References

[1] Haycock, J. W. 3D cell culture : methods and protocols. Vol. 695 (Humana Press/Springer, 2011).

[2] Zegers, M. M., O'Brien, L. E., Yu, W., Datta, A. & Mostov, K. E. Epithelial polarity and tubulogenesis in vitro. Trends Cell Biol 13, 169-176 (2003).

[3] Weigelt, B. & Bissell, M. J. Unraveling the microenvironmental influences on the normal mammary gland and breast cancer. Semin Cancer Biol 18, 311-321 (2008).

[4] Debnath, J. & Brugge, J. S. Modelling glandular epithelial cancers in three-dimensional cultures. Nat Rev Cancer 5, 675-688 (2005).

[5] Wolf, K. et al. Collagen-based cell migration models in vitro and in vivo. Semin Cell Dev Biol 20, 931-941 (2009).

[6] Gaggioli, C. & Sahai, E. Melanoma invasion - current knowledge and future directions. Pigment Cell Res 20, 161-172 (2007).

[7] Muschler, J. & Streuli, C. H. Cell-matrix interactions in mammary gland development and breast cancer. Cold Spring Harb Perspect Biol 2, a003202 (2010).

[8] Justice, B. A., Badr, N. A. & Felder, R. A. 3D cell culture opens new dimensions in cell-based assays. Drug Discov Today 14, 102-107 (2009).

[9] Korn, K. & Krausz, E. Cell-based high-content screening of small-molecule libraries. Curr Opin Chem Biol 11, 503-510 (2007).

[10] Paszek, P. et al. Population robustness arising from cellular heterogeneity. P Natl Acad Sci USA 107, 11644-11649 (2010).

[11] Turner, D. A. et al. Physiological levels of TNFalpha stimulation induce stochastic dynamics of NF-kappaB responses in single living cells. J Cell Sci 123, 2834-2843 (2010).

[12] Kuan Yan & Verbeek, F. J. Segmentation for high-throughput image analysis: watershed masked clustering. In: SoLA'12 Proceedings of the 5th international conference on Leveraging Applications of Formal Methods, Verification and Validation: applications and case studies. (eds T. Margaria & B. Steffen) 25-41 (Springer-Verlag).

[13] Cao, L., Yan, K., Winkel, L., Graauw, M. D. & Verbeek, F. J. Pattern recognition in high-content cytomics screens for target discovery: case studies in endocytosis. In: PRIB'11 Proceedings of the 6th IAPR international conference on Pattern recognition in bioinformatics (eds Marco Loog, Marcel J. T. Reinders, Dick De Ridder, & Lodewyk Wessels) 330-342 (Springer-Verlag).

[14] Niblack, W. An introduction to digital image processing. 1-215 (Prentice-Hall International, 1986).

[15] Bakal, C., Aach, J., Church, G. & Perrimon, N. Quantitative morphological signatures define local signaling networks regulating cell morphology. Science 316, 1753-1756 (2007).

[16] Wang, J. et al. Cellular phenotype recognition for high-content RNA interference genome-wide screening. J Biomol Screen 13, 29-39 (2008).

[17] Birmingham, A. et al. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods 6, 569-575. (2009).

[18] Defays, D. Efficient Algorithm for a Complete Link Method. Comput J 20, 364-366 (1977).

[19] Duda, R. O., Hart, P. E. & Stork, D. G. Pattern Classification (2nd Edition) 215-268 (John Wiley and Sons, 2001).

[20] Liu, C. & Wechsler, H. Robust coding schemes for indexing and retrieval from large face databases. IEEE Trans Image Process 9, 132-137 (2000).

[21] Webb, A. R. Statistical Pattern Recognition (2nd Edition) 123-163 (John Wiley & Sons, 2002).

[22] Cortes, C. & Vapnik, V. Support-Vector Networks. Mach Learn 20, 273-297 (1995).

[23] Theodoridis, S. & Koutroumbas, K. Pattern recognition. 3rd edn, 93-118 (Academic Press, 2006).

[24] Raudys, S. & Duin, R. P. W. Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix. Pattern Recogn Lett 19, 385-392 (1998).

[25] Rajaram, S., Pavie, B., Wu, L. F. & Altschuler, S. J. PhenoRipper: software for rapidly profiling microscopy images. Nat Methods 9, 635-637 (2012).

[26] Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7 (2006).

[27] Hollestelle, A. et al. Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines. Breast cancer research and treatment 121, 53-64 (2010).

[28] Kao, J. et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PloS one 4, e6146 (2009).

[29] Lehmann, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. The Journal of clinical investigation 121, 2750-2767 (2011).

[30] Neve, R. M. et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer cell 10, 515-527 (2006).

[31] Han, J. et al. Molecular predictors of 3D morphogenesis by breast cancer cell lines in 3D culture. PLoS computational biology 6 (2010).

[32] Hollestelle, A. et al. Four human breast cancer cell lines with biallelic inactivating alpha-catenin gene mutations. Breast cancer research and treatment 122, 125-133 (2010).

[33] Gonzalez, R. C. & Woods, R. E. Digital image processing. 2nd edn, 136-137 (Prentice Hall, 2002).

[34] Mahalanobis, P. C. On the generalised distance in statistics. Proceedings of the National Institute of Sciences of India 2, 49-55 (1936).

[35] Hasan Abdel Qader, A. R. R., Syed Al-Haddad. Fingerprint recognition using zernike moments. The International Arab Journal of Information Technology 4, 372-376 (2007).

[36] Hu, M. Visual-Pattern Recognition by Moment Invariants. Ire T Inform Theor 8, 179-187 (1962).

[37] Manjunath, B. S. & Ma, W. Y. Texture features for browsing and retrieval of image data. IEEE Trans Pattern Anal Mach Intell 18, 837-842 (1996).

[38] McLachlan, G. J. Discriminant analysis and statistical pattern recognition. Vol. 235 52-54 (John Wiley & Son, 1992).

[39] Greenwood, P. E. & Nikulin, M. S. A guide to chi-squared testing. Vol. 280 1-280 (John Wiley & Son, 1996).

[40] Mardia, K. V., Kent, J. T. & Bibby, J. M. Multivariate analysis. 1-521 (Academic Press, 1979).

Supplementary note

Phenotypic trajectory modeling using 2nd order polynomial regression model

To further characterize different phenotypes induced by the biologically active compounds, we used a 2nd order polynomial regression model to build the concentration dependent phenotypic trajectory for each identified active compound. We firstly calculated the data variation for each of the 9 principle components. The mean and standard deviation were calculated for pool of all identified active compounds and control, respectively. Two sample Kolmogorov-Smirnov test was used to compare control data and active compounds data. The principle components with no significant difference between control and compounds, equal data variation, or bigger variation in negative controls were excluded to avoid over training of the regression model. As the result, the first 2 components were retained.

For each compound we used all concentrations including inactive concentrations, active concentrations and the average of DMSO controls for the regression model. To quantify the distance between two trajectories, for example the trajectories of compound i and compound j, we firstly calculated the $R^2_{(i,j)}$ using the trajectory model built for compound i to fit the data points of compound j , and also for the other way around $R^2_{(j,i)}$.

$$R^2_{(i,j)} = 1 - \frac{SS_{err(i,j)}}{SS_{(i,j)}}$$

$$SS_{err(i,j)} = \sum_{n_j}(x_2 - x'_2)^2$$

$$SS_{tot(i,j)} = \sum_{n_j}(x_2 - \overline{x_2})^2$$

where x_2 is the observed value of a data point of compound j on the PC2. x'_2 is the predicted value of the same data point of compound j on the PC2 using the trajectory model of compound i. $\overline{x_2}$ is the mean of x_2 over all data points of compound j. n_j represents number of data points for compound j. Finally the distance between two trajectories was calculated according to the Equation 3.1 and Equation 3.2 in the Results section.

This modeling method can be easily extended to more dimensions. For example, to model a 3D trajectory, our task becomes to fit a model

$$x'_2 = \beta_a x_1^2 + \beta_b x_1 + \varepsilon_1$$

$$x'_3 = \beta_c x_1^2 + \beta_d x_1 x'_2 + \beta_e x_1 + \beta_f x'_2 x'_2 + \beta_g x'_2 + \varepsilon_2$$

for each compound, where β represents regression coefficient and ε is an error term. x_3 and x'_3 is the observed value and predicted value of a data point of compound j on

the 3$^{\text{rd}}$ dimensional principle component, respectively. Then $SS_{err(i,j)}$ and $SS_{tot(i,j)}$ would be defined as

$$SS_{err(i,j)} = \sum_{n_j}(x_3 - x_3')^2$$

$$SS_{tot(i,j)} = \sum_{n_j}(x_3 - \overline{x_3})^2$$

Supplementary figures
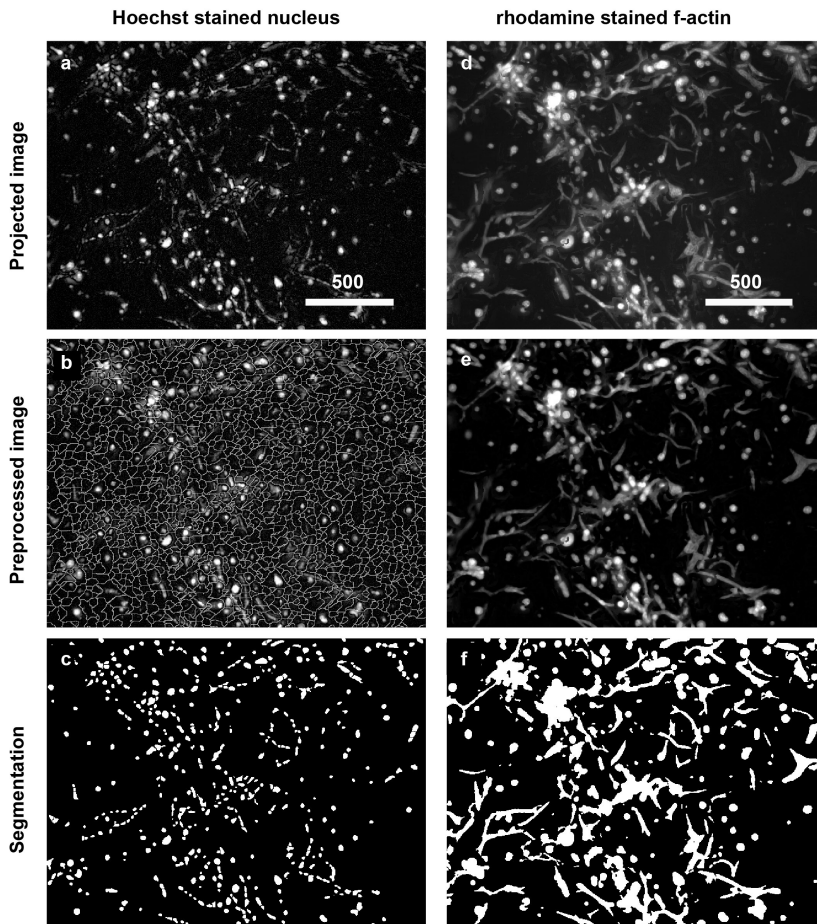
Figure S3.1: Segmentation of the projected images from the Hoechst stained nuclei channel and rhodamine stained F-actin channel. (a-c) An example of the stepwise result of "Projection-Preprocessing-Segmentation pipeline" of an image from the Hoechst stained nuclei channel (d-f) Stepwise result of "Projection-Preprocessing-Segmentation" pipeline of the corresponding image from the rhodamine stained F-actin channel.
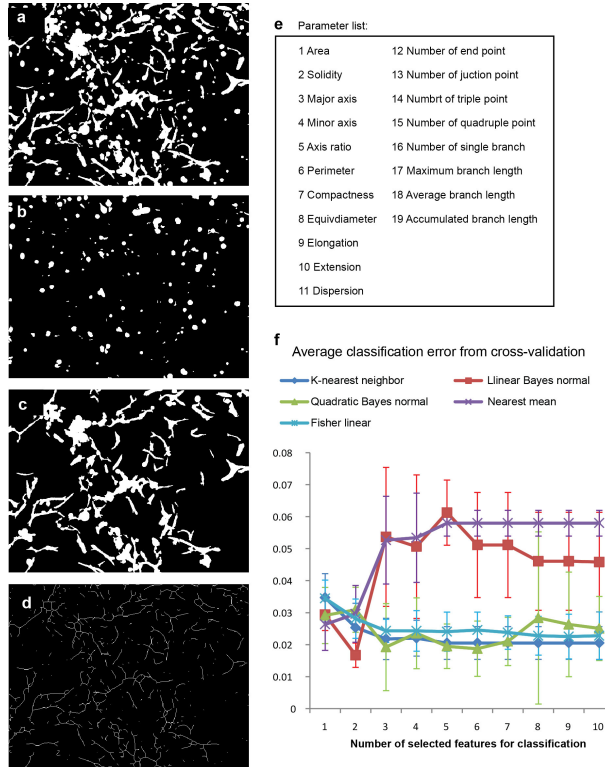
Figure S3.2: Subpopulation classification. (a) Segmentation result of a projected rhodamine stained F-actin image. (b) Manually selected spherical objects and (c) branched objects. (d) Skeleton of each binary object. (e) Features calculated from each binary object for subpopulation classification. (f) Cross-validation result for comparing different classification methods and identifying optimal number of features for classification. Average error rate of a 10-fold cross-validation is shown in the chart with standard deviation as error bar.
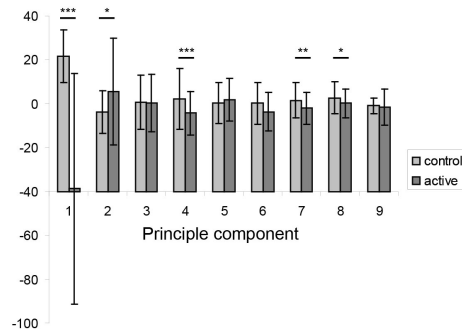
Figure S3.3: Selection of principle components for phenotypic trajectory modeling. Comparison of data variation of active compounds and negative control for each one of the 9 principle components. In the chart the mean and standard deviation are shown for all active compounds and control. Two sample Kolmogorov-Smirnov test was used to compare the data variation between active compounds and negative control. *: $p-value < 0.05$, **: $p-value < 0.005$, ***: $p-value < 0.0005$.

73

Figure S3.4: Clustering and classification of 4T1 compounds screen when subpopulation or intensity parameters are not used. (a) Hierarchical clustering with complete linkage was applied on the distance matrix defined according to Equation 3.1. All the subpopulation parameters were not taken into account when we computed the PCA. The scale of the dendrogram is the natural logarithm of $\sqrt{(1 - R_{mean}^2)}$. (b) Classification result using multiple classification methods when the subpopulation parameters were not taken into account. (c) Classification result when all the wavelet, moments and intensity parameters were not taken into account. For (b) and (c), the classes were defined according to the Figure 3.5b. Feature selection with search algorithm "forward" and criterion "Mahalanobis distance" was applied to detect the optimal number of features. For each classification method and each number of selected features, 10-fold cross-validation was applied, resulting in 10 error rates. The average error rates are shown in the chart with standard deviation as error bar. * SVC stands for support vector machine classification.

Figure S3.5: Comparison to PhenoRipper. (a) A two dimensional MDS plot after using PhenoRipper for analysis of the 4T1 compounds screen. The block size used was 50 pixels. Compounds are identified by different color. The highlighted orange point corresponds to a well treated with Arq 197 at the concentration of 3.16μM, and the highlighted purple point corresponds to a control well. (b-c) Phenotype images corresponding to the highlighted points in (a).
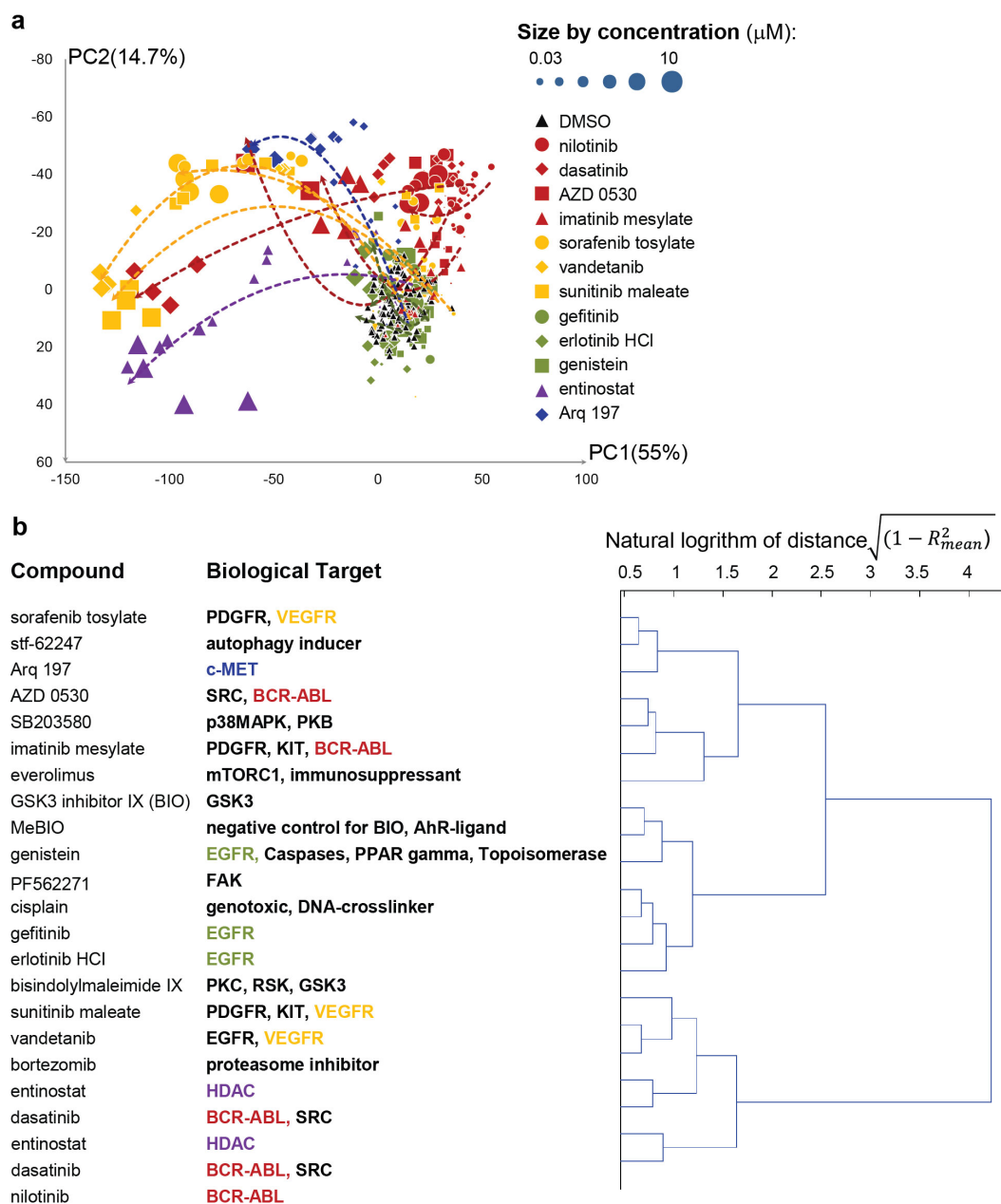
Figure S3.6: Comparison to CellProfiler. (a) A 2D PCA plot of phenotype profiles for negative control (DMSO) and 12 active compounds at different concentrations using CellProfiler to profile compounds in the 4T1 screen (not all compounds are shown). Percentages of data variation preserved in each principle component are shown with each axis. Compounds are marked with different shapes and colors. Compounds with the same biological target are colored identically. Red: BCR-ABL target inhibitor; Yellow: VEFGR inhibitor; Green: EGFR inhibitor; Purple: HDAC inhibitor; Blue: c-MET inhibitor. Concentration is represented by the size of data points. The trend lines were added for each effective compound using 2nd order polynomial regression model. (b) Hierarchical clustering result for all active compounds using a distance matrix defined based on Equation 3.1. The scale of dendrogram is the natural logarithm of $\sqrt{(1 - R_{mean}^2)}$.

a

b
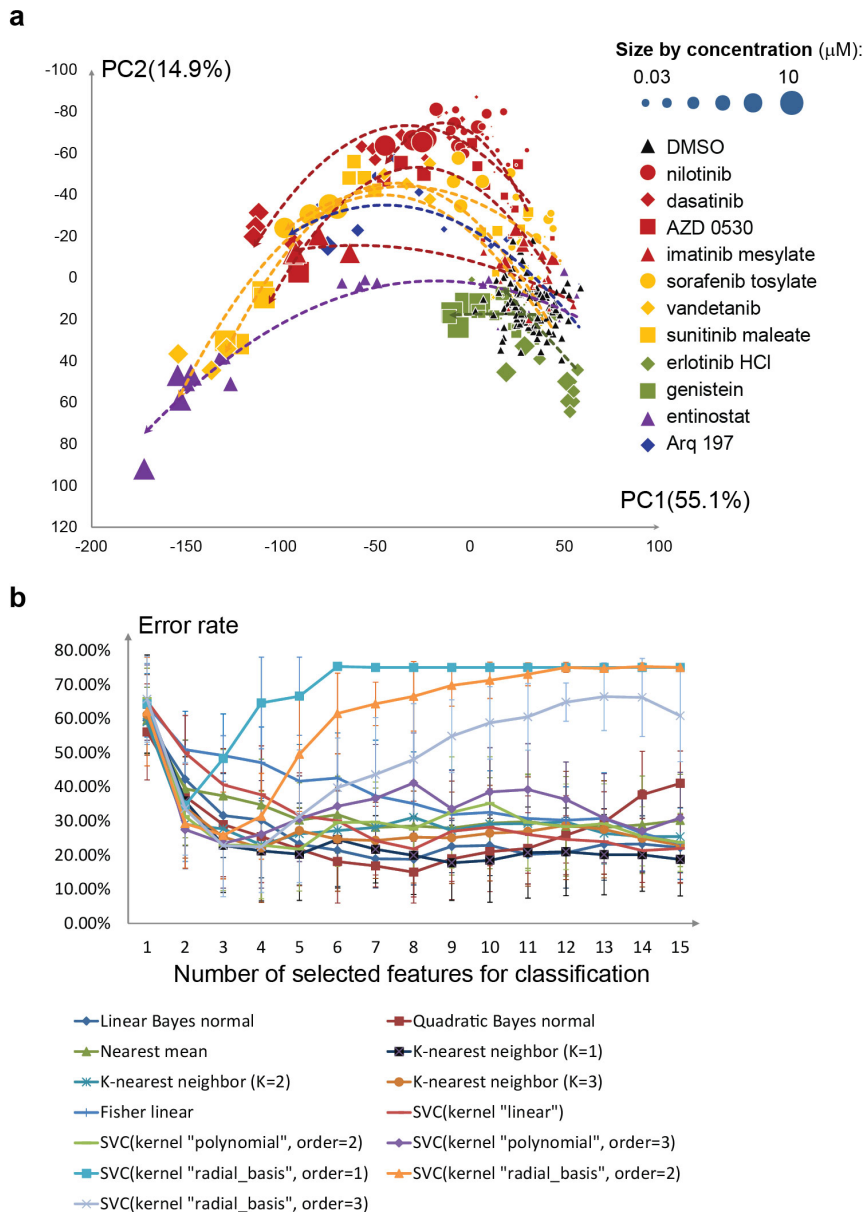
Figure S3.7: Reproducibility of our methodology. (a) A two dimensional PCA plot of phenotype profiles for negative control (DMSO) and 11 active compounds from a repeated experiment (not all compounds are shown). Percentages of data variation preserved in each principle component are shown with each axis. Compounds are marked with different shapes and colors. Compounds with the same biological target are colored identically. Red: BCR-ABL target inhibitor; Yellow: VEFGR inhibitor; Green: EGFR inhibitor; Purple: HDAC inhibitor; Blue: c-MET inhibitor. Concentration is represented by the size of data points. The trend lines were added for each effective compound using 2nd order polynomial regression model. (b) Classification result using multiple classification methods. Feature selection and classification algorithms are the same as in the first experiment.

Supplementary tables

Compound Name	Biological activity	Active in screen
DMSO	Negative control	no
AG 538	tyrosine kinase inhibitor (EGFR)	no
Arq 197	tyrosine kinase inhibitor (c-MET), Positive control	yes
AZD 0530	tyrosine kinase inhibitor (SRC family, BCR-ABL)	yes
bisindolylmaleimide IX	serine/threonine kinase inhibitor (PKC, GSK3)	yes
bortezomib	proteasome inhibitor	yes
cisplatin	genotoxic, DNA-crosslinker	yes
dasatinib	tyrosine kinase inhibitor (Src family, BCR-ABL)	yes
entinostat	HDAC inhibitor	yes
erlotinib HCl	tyrosine kinase inhibitor (EGFR)	yes
everolimus	threonine kinase inhibitor (mTORC1), immunosuppressant	yes
gefitinib	tyrosine kinase inhibitor (EGFR)	yes
genistein	tyrosine kinase inhibitor (EGFR), Caspase, PPAR gamma and Topoisomerase II inhibitor, Estrogenic activity	yes
GSK3 inhibitor IX (BIO)	serine/threonine kinase inhibitor (GSK3)	yes
imatinib mesylate	tyrosine kinase inhibitor (PDGFR, KIT, BCR-ABL)	yes
lapatinib dihydrate	tyrosine kinase inhibitor (EGFR, HER2)	no
LY364947	serine/threonine kinase inhibitor (TGF-β RI, TGF-β RII, p38 MAPK, MLK-7)	no
MeBIO	negative control for BIO, AhR-ligand	yes
nilotinib	tyrosine kinase inhibitor (BCR-ABL, KIT, LCK, EPHA3, EPHA8, DDR1, DDR2, PDGFR, MAPK11 and ZAK)	yes
olaparib	PARP inhibitor	no
PF562271	tyrosine kinase inhibitor (FAK, PTK2B)	yes
SB203580	serine/threonine kinase inhibitor (p38MAPK, PKB)	yes
sorafenib tosylate	tyrosine kinase inhibitor (VEGFR, PDGFR, Raf kinases)	yes
SP600125	serine/threonine kinase inhibitor (JNK1, JNK2, JNK3, Aurora A, Flt3, TRKA)	no
stf-62247	autophagy inducer	yes
SU6668	tyrosine kinase inhibitor (PDGFR,FGFR1, Flk-1/KDR)	no
sunitinib maleate	tyrosine kinase inhibitor (VEGFR, EGFR, RET, KIT)	yes
vandetanib	tyrosine kinase inhibitor (VEGFR, EGFR, RET, KIT)	yes
wortmannin	phosphoinositide (PI) 3-kinase inhibitor	no
Y27632	serine/threonine kinase inhibitor (ROCK)	no

Table S3.1: List of compounds used in the 4T1 cell screen.

Parameter index	Description
From 1-23	The mean of **M1-M23** of all cell clusters in one image
From 24-46	The standard deviation of **M1-M23** of all cell clusters in one image
47	The number of nuclei in one image **(M24)**
48	The number of cell clusters in one image **(M25)**
49	The total area of nuclei in one image **(M26)**
50	The total area of cell clusters in one image **(M27)**
51	The mean density of nuclei in one image **(M28)**
52	The mean density of cell clusters in one image **(M29)**
53	The maximum branch length of all cell clusters in one image
54	The accumulated branch length of all cell clusters in one image

Table S3.2: List of classic morphological parameters calculated for the whole object population. The definition of parameters **M1-M29** is in the Methods section.

Parameter index	Description
From 1-49	The mean of **S1-S49** of all cell clusters in one image
From 50-98	The standard deviation of **S1-S49** of all cell clusters in one image

Table S3.3: List of moment parameters calculated for the whole object population. The definition of parameters **S1-S49** is in the Methods section.

Parameter index	Description
From 1-136	The parameters **I1-I136** measured on the projected images of the Hoechst stained nuclei channel and the rhodamine stained F-actin channel

Table S3.4: List of intensity parameters calculated for the raw intensity projected images. The definition of parameters **I1-I136** is in the Methods section.

Parameter index	Description
1	The number of cell clusters which are classified as spherical object
2	The number of cell clusters which are classified as branched object
3	The accumulated area of cell clusters which are classified as spherical object
4	The accumulated area of cell clusters which are classified as branched object
5	The accumulated branch length of cell clusters which are classified as spherical object
6	The accumulated branch length of cell clusters which are classified as branched object
From 7-29	The mean of **M1-M23** of the cell clusters which are classified as spherical object
From 30-52	The standard deviation of **M1-M23** of the cell clusters which are classified as spherical object
From 53-56	The mean of **I137-I140** of cell clusters which are classified as spherical object
From 57-60	The standard deviation of **I137-I140** of cell clusters which are classified as spherical object
From 61-83	The mean of **M1-M23** of the cell clusters which are classified as branched object
From 84-106	The standard deviation of **M1-M23** of the cell clusters which are classified as branched object
From 107-110	The mean of **I137-I140** of cell clusters which are classified as branched object
From 111-114	The standard deviation of **I137-I140** of cell clusters which are classified as branched object
From 115-212	The mean and the standard deviation of **S1-S49** of cell clusters which are classified as spherical object
From 213-310	The mean and standard deviation of **S1-S49** of cell clusters which are classified as branched object

Table S3.5: List of parameters calculated for the subpopulations. The definition of parameters **M1-M23**, **I137-I140**, **S1-S49** is in the Methods section.

Principle component 1 (62.45%)	Principle component 2 (9.81%)	Principle component 3 (4.69%)
The average dispersion of cell clusters which are classified as branched object	The average solidity of cell clusters which are classified as branched object	The average solidity of cell clusters which are classified as branched object
The average extension of cell clusters which are classified as branched object	The mean intensity of cell clusters which are classified as branched object	The average extension of cell clusters which are classified as branched object
The average solidity of cell clusters which are classified as branched object	The mean(minimum intensity of each cell cluster which is classified as branched object)	The average elongation of cell clusters which are classified as branched object
The average Feret's diameter of cell clusters	The mean(maximum intensity of each cell cluster which is classified as branched object)	The average Zernike moments (order 4) measured on the binary mask of the cellular clusters
The standard deviation of cellular solidity	The intensity standard deviation of rhodamine channel	The average Zernike moments (order 14) measured on the binary mask of cellular clusters which are classified as branched object

Table S3.6: Top 5 parameters which contribute most to each of the 3 first principle components in the 4T1 cell screen. * The percentage (%) after each principle component indicates what percentage of data variance it accounts for.

Active compounds	Active concentrations					
Arq 197	10 μM	3.16 μM	1 μM	0.316 μM		
AZD 0530	10 μM	3.16 μM	1 μM			
bisindolylmaleimide IX	10 μM	3.16 μM				
bortezomib	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	0.03 μM
cisplatin			1 μM			
dasatinib	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	0.03 μM
entinostat	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	
erlotinib HCl		3.16 μM	1 μM			
everolimus			1 μM		0.1 μM	
gefitinib				0.316 μM		0.03 μM
genistein	10 μM					
GSK3 inhibitor IX		3.16 μM				
imatinib mesylate	10 μM	3.16 μM	1 μM			
MeBIO	10 μM					
nilotinib	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	0.03 μM
PF	10 μM	3.16 μM				
SB203580	10 μM					
sorafenib tosylate	10 μM	3.16 μM	1 μM	0.316 μM		
stf-62247	10 μM	3.16 μM	1 μM			
sunitinib maleate	10 μM	3.16 μM	1 μM			
vandetanib	10 μM	3.16 μM				

Table S3.7: Biologically active compounds and corresponding active concentration identified using Mahalanobis distance (α=0.05).

Active compounds	Active concentrations					
Arq 197	10 μM	3.16 μM	1 μM	0.316 μM		
AZD 0530	10 μM	3.16 μM	1 μM	**0.316** μM	**0.1** μM	
dasatinib	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	0.03 μM
entinostat	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	
erlotinib HCl	**10** μM	3.16 μM	1 μM			
genistein	10 μM	**3.16** μM				
imatinib mesylate	10 μM	3.16 μM	1 μM			
nilotinib	10 μM	3.16 μM	1 μM	0.316 μM	0.1 μM	0.03 μM
sorafenib tosylate	10 μM	3.16 μM	1 μM	0.316 μM	**0.1** μM	**0.03** μM
stf-62247	10 μM	3.16 μM	1 μM			
sunitinib maleate	10 μM	3.16 μM	1 μM			
vandetanib	10 μM	3.16 μM	**1** μM			

Table S3.8: Compounds tested in the repeated 4T1 cell screen and corresponding active concentration identified using Mahalanobis distance (α=0.05). Concentrations which were detected as active in this screen but not active in the first screen are highlighted in bold.

Cell line index	Cell line	Categories	Replicates
1	BT20	basal-A	6
2	BT474	ERBB2+	3
3	BT483	luminal	3
4	BT549	basal-B	6
5	CAMA-1	luminal	6
6	EVSA-T	ERBB2+	6
7	HCC1143	basal-A	6
8	HCC1395	basal-B	6
9	HCC1569	basal-A	3
10	HCC1806	basal-A	3
11	HCC1937	basal-A	3
12	HCC1954	basal-A	3
13	HCC202	ERBB2+	3
14	HCC70	basal-A	6
15	Hs578T	basal-B	6
16	MCF7	luminal	6
17	MDA-MB-134VI	luminal	3
18	MDA-MB-175VII	luminal	3
19	MDA-MB-231	basal-B	6
20	MDA-MB-361	ERBB2+	3
21	MDA-MB-415	luminal	3
22	MDA-MB-435s	basal-B	6
23	MDA-MB-436	basal-B	6
24	MDA-MB-453	ERBB2+	6
25	MDA-MB-468	basal-A	6
26	MPE600	luminal	3
27	OCUB-F	ERBB2+	6
28	OCUB-M	ERBB2+	6
29	SK-BR-3	ERBB2+	6
30	SK-BR-7	basal-B	6
31	SUM102PT	basal-B	6
32	SUM1315MO2	basal-B	6
33	SUM149PT	basal-A	6
34	SUM159PT	basal-B	6
35	SUM185PE	luminal	6
36	SUM190PT	ERBB2+	3
37	SUM225CWN	ERBB2+	3
38	SUM229PE	basal-A	6
39	SUM44PE	luminal	3
40	SUM52PE	luminal	3
41	T47D	luminal	3
42	UACC812	ERBB2+	3
43	UACC893	ERBB2+	3
44	ZR-75-1	luminal	3

Table S3.9: Breast cancer cell lines (basal-A, basal-B, luminal or ERBB2+) used for classification.

Chapter 4

Automated analysis pipeline for 3D surface reconstruction and phenotype profiling of 3D cultured micro-tissues suitable for high-content and high-throughput screening

Zi Di, Maarten J D Klop, Bob van de Water, Leo S Price,
John H N Meerman & Fons J Verbeek

Abstract

3D cell cultures have been widely applied for high-content screening to investigate cellular phenotypic responses to different genetic or chemical perturbations. In order to study complex micro-tissue architectures that cells develop in 3D cultures, confocal laser scanning microscopy is often used to visualize specimens by shifting a focal plane through their entire form. In this manner high-resolution images are generated using point-by-point laser excitation and application of a filtering pinhole to eliminate out-of-focus information from adjacent focal planes. However, the slow scanning process is a major drawback so that image acquisition takes a large amount of time, which limits its application for high-throughput screening. To overcome this, we developed a high-content analysis pipeline that is able to perform phenotypic profiling of 3D cultured micro-tissues based on automated wide-field microscopy. Image stacks of two fluorescent channels were acquired for each well of a standard multi-well micro plate by shifting a focal plane in z-direction. We first applied a deconvolution method to restore the image signals, which were degraded by light scattering. Next, two novel segmentation methods were developed to define single nucleus and multi-cellular regions, respectively. For each nuclear structure, we calibrate its dimension in z-direction using images of fluorescent beads with a known size. After surface reconstruction, 3D morphological, topological features, moments and localization properties were measured from the reconstructed structures. To validate our method, we generated multiple image stacks using a confocal laser scanning microscope with a high resolution objective lens. The quantification results from the two imaging techniques were compared statistically and no significant differences were obtained. Therefore, we conclude that our analysis pipeline can retrieve 3D properties of micro-tissue structures from wide-field microscope images that are comparable to the information extracted from confocal microscope images, but at much less cost of imaging and computational time allowing higher throughput.

4.1 Introduction

Recent advances in 3D cell cultures have provided novel insights into various aspects of cell behavior. 3D cultures mimic spatial organization of real tissues by using extracellular matrix (ECM) gel to re-establish physiological cell-cell and cell-ECM interactions, and therefore enable cells to develop more in vivo like tissue architectures [1, 2]. To investigate those complex micro-tissue structures, molecular components of the cells are labeled with fluorescent dyes and 3D fluorescence microscopes are used to scan specimens over their entire depth range by shifting a focal plane in z-direction [3, 4], yielding stacks of sequential image slices and each of them contains information of the focal plane. One of the most common 3D fluorescence microscopy is confocal laser scanning microscopy. Although the generated images have much higher resolution compared to conventional fluorescence microscopy, a major drawback is that it needs to scan whole specimen point-by-point. Especially when the specimen is thick or at micro-tissue scale, this slow scanning process not only limits the application of confocal microscopy for high throughput experiments, but also causes a severe bleaching problem when the screening time becomes too long. Therefore, conventional wide-field microscopy needs to be considered as an alternative solution for high-throughput screening of 3D cultured micro-tissues. Wide-field microscopes use an excitation light-source to illuminate entire specimen so that the imaging speed is dramatically increased. However, each image slice is degraded by out-of-focus signals because the emission light that composes the image comes from the focal plane as well as the planes above and below the focal plane. A major challenge is to recover 3D structures correctly from those low resolution images. This requires an advanced, accurate and efficient image analysis method.

Another challenge is the phenotypic profiling of 3D cultured micro-tissues. Although increased popularity of high-content screening has fueled the development of image analysis techniques, until recently, quantification of cellular phenotypic features is still limited to single or multiple 2D parameters [5]. Since cells are not flat and together develop much higher levels of tissue architecture in 3D cell cultures, it is necessary to develop an image analysis method to measure more relevant and sophisticated 3D parameters.

Here we aimed to establish a wide-field microscopy-based high-content analysis pipeline for the high-throughput screening of 3D cultured micro-tissues that involves the challenges mentioned above (Figure 4.1). After fluorescent staining, images stacks of two channels (for nucleus and microfilament signal) were collected from an automated wide-field microscope system. We first used a deconvolution technique to enhance the quality of the image stacks by removing the out-of-focus signal. Next, we developed two segmentation methods to define single nucleus and multi-cellular micro-tissue regions, respectively. As defined objects appeared to be much elongated in the vertical direction, we introduced a simple calibration method to normalize the dimension of each individual nucleus in z-direction, using images of fluorescent beads with known size. After 3D surface reconstructions, three categories of 3D parameters were measured on the reconstructed nucleus and micro-tissue structures, respectively.
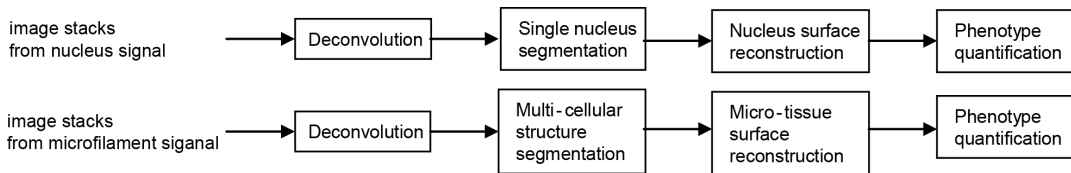
Figure 4.1: Overview of the proposed image analysis pipeline for high-content screening of 3D cultured micro-tissues based on wide-field florescence microscopy.

To validate our method, we investigated human prostate cancer cells (PC3) in a 384-well plate format after they were exposed to hepatocyte growth factor and formed invasive micro-tissue structures in the 3D cell culture. The micro-tissues were fixed and stained with Hoechst 33258 to visualize nuclei and with rhodamine-phalloidin to visualize F-actin. We used a wide-field microscope system as well as a confocal laser scanning microscope to collect image stacks from the same wells. Subsequently, our analysis pipeline described above was applied to extract phenotypic parameters from the image stacks acquired from the wide-field microscope. For the image stacks that were collected by the confocal microscope, we applied automated segmentation and the results were validated by human evaluation. Based on the segmentation results, 3D surface reconstruction was performed with a normalization of the nuclei in z-direction similar to the normalization of the nuclei performed for the wide-field microscope. Next, the same phenotypic parameters were measured from the reconstructed nuclei and micro-tissue structures, respectively. Finally, we statistically compared the quantification results extracted by the two different microscopes. A substantial gain in time efficiency was shown when we used our analysis pipeline in combination with the wide-field microscope and the quantification results were comparable to the results from the confocal microscope.

4.2 Results

4.2.1 Method development for 3D surface reconstruction

4.2.1.1 Deconvolution

A major disadvantage of wide-field microscopy is that each generated image slice contains out-of-focus signals. According to the optical principles, this blurring effect is mainly caused by light scattering and can be formulated by a point spread function (PSF). One way to eliminate this blurring effect is deconvolution. It computes the PSF based on the optical principles and then deconvolves microscope images with that PSF so that the process of image degradation is inversed and the image quality is improved. There are many methods available for calculating the PSF. In this project, we used Huygens Software (http://www.svi.nl/HuygensProducts) to compute a theoretical PSF that is based on the microscope model and microscope

parameters (Table S4.1). First, the signal-to-noise ratio was estimated on randomly selected sample images. This ratio was relatively high (> 80), and considering the time efficiency we used a fast maximum likelihood estimation algorithm [6] for the image restoration (see examples before and after deconvolution in Figure 4.2a-4.2h). A large extent of out-of-focus signals was removed and the the signal-to-noise ratio was enhanced for both Hoechst stained nuclei channel and rhodamine stained F-actin channel.
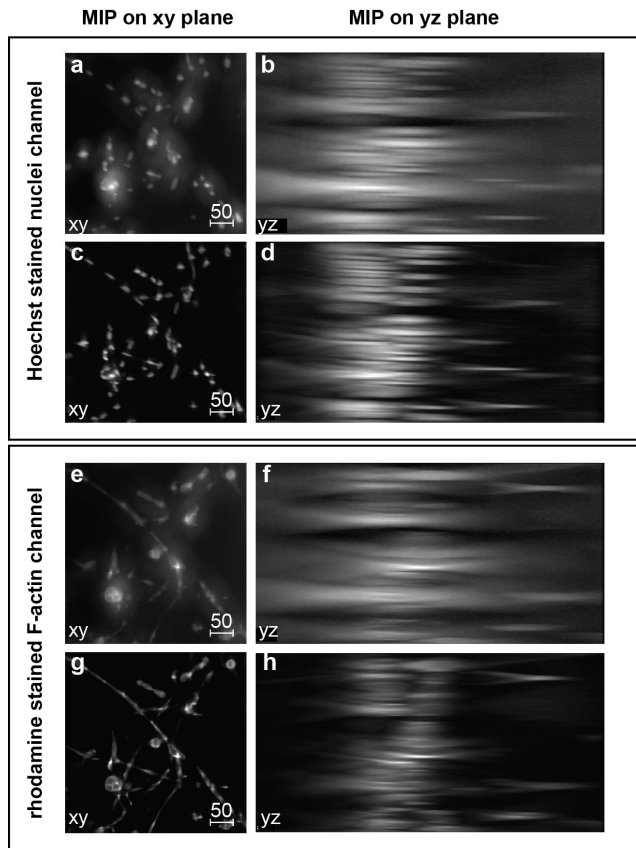
Figure 4.2: An example of deconvolution results for both Hoechst stained nuclei channel and rhodamine stained F-actin channel. For the Hoechst stained nuclei channel, we cropped a part of an original image stack and showed its maximum intensity projection (MIP) on (a) xy-plane and on (b) yz-plane. After deconvolution of this image stack, we showed the MIP of the same part of image stack on (c) xy-plane and on (d) yz-plane. For the rhodamine stained F-actin channel, we showed the MIP of corresponding part of image stack before deconvolution (e-f) and after deconvolution (g-h). The scale bar represents 50μm.

4.2.1.2 Segmentation for each individual nucleus

Before segmentation, a 2D median filter with kernel size 3×3 pixels was applied on each slice to reduce systemic noise. Next, a Rolling Ball algorithm [7] was applied to correct for uneven illumination in the background. As a rule of thumb, the rolling ball radius should be at least as large as the radius of the largest object in the image. Here, the average nuclear radius equals about 8 pixels, thus the radius of the rolling ball was defined at 15 pixels. To segment individual nuclei, a 2D watershed masked segmentation method [8, 9] was extended to its equivalent 3D version. First, a 3D watershed algorithm [10, 11] was applied to divide image stacks into 3D compartments each of which contains one nucleus. Subsequently, K-means clustering [12, 13] was applied within each compartment to refine nuclear regions (Figure 4.3a). As the watershed segmentation is sensitive to signals of discrete intensity that may cause artificial local maximum and an over-segmentation problem, images were first convolved with a Gaussian filter (kernel size 3×3 pixels) to remove noises before using watershed segmentation. Once 3D watersheds were obtained, the preprocessed images prior to Gaussian convolution were used to apply K-means clustering [8, 9].

4.2.1.3 Segmentation for multi-cellular micro-tissue structures

A novel segmentation method was developed to define multi-cellular structures on the image stacks with relatively low resolution. Due to the low NA, deconvolution cannot remove all the out-of-focus signal from each image slice. Moreover, we observed that the level of sharpness varied over different slices. It is not feasible to calculate a global threshold for the whole image stack. Instead, a segmentation method which dynamically calculates a threshold intensity value for each slice according to the estimation of its sharpness level was developed.

To estimate the sharpness level for each image slice, we firstly calculated the magnitude of the gradient (GM) for each pixel using a 2D Sobel filter [14]. A simple sharpness metric SL of a certain slice was then defined as the average gradient magnitude (Equation 4.1).

$$SL(s) = \frac{\sum_{p \in s} GM_p}{n_{s,GM>0}} \tag{4.1}$$

where s indicates the sth slice of an image stack and p indicates each pixel of slice s. n_s means number of pixels of which $GM > 0$ in slice s. This is based on the principle that sharper images should have much more intensity variation, and thus the SL would be relatively higher. In contrast, blurred images contain more out-of-focus regions where intensity varies more smoothly which results in an SL that would be decreased.

In blurred images where there are much more out-of-focus regions than in-focus regions (though the intensity of our-of-focus regions are still lower than that of in-focus regions) the conventional K-means clustering cannot work properly as it is based on the assumption that the intensity variations in the foreground and background are

equal. In order to take into account the difference in variation between foreground and background, we modified the conventional K-means clustering method [12] to adjust the threshold for each image slice with the SL. It iteratively updated a new threshold T_i as:

$$T_i = \frac{k}{SL} \times m_f(T_{i-1}) + (1 - \frac{k}{SL}) \times m_b(T_{i-1}) \qquad (4.2)$$

where i represents i^{th} iteration. $m_f(T)$, $m_b(T)$, represents the average foreground intensity and background intensity when the threshold is T. k is a rational number and can be defined empirically according to the signal-to-noise ratio of image stacks and is often consistent through the whole screening process. The iterations terminate when the changes $|T_i - T_{i-1}|$ become sufficiently small. Actually, we can rewrite Equation 4.2 to

$$\frac{k}{SL} = \frac{T_i - m_b(T_{i-1})}{m_f(T_{i-1}) - m_b(T_{i-1})} \qquad (4.3)$$

which indicates that $\frac{k}{SL}$ can determine the proportion of background intensity variation in the total intensity variation of an image slice. For the conventional K-means clustering method $\frac{k}{SL} = \frac{1}{2}$. After being incorporated with SL, the blurred image slice with smaller SL would increase the proportion of background intensity variation, and thus the corresponding threshold value would be higher. Figures 4.3b-4.3d show segmentation results of two image slices obtained from the same image stack but with different sharpness levels. Although in Figure 4.3c there are out-of-focus regions with high intensity value, due to high $\frac{k}{SL}$, these regions are defined as background, whereas the in-focus regions with lower intensity in Figure 4.3d are recognized as foreground.

4.2.1.4 3D surface reconstruction and normalization of nuclei in z-direction

Our 3D surface reconstruction consists of two important steps: 1) 3D labeling which assigns a label to each foreground pixel so that pixels with the same label define one single object; 2) 3D reconstruction which generates a geometrical 3D model for each labeled object. For the Hoechst stained nuclei channel, 3D watershed already created compartments for each single nucleus. Foreground pixels identified in the same compartment were assigned with the same object label. For 3D labeling of the multi-cellular micro-tissue structures, we applied a sequential labeling algorithm based on 18-connected connectivity as follow: suppose one foreground pixel with coordinates (x, y, z) has already been labeled as l, 18 neighboring pixels with coordinates $(x, y, z \pm 1)$, $(x \pm 1, y, z \pm 1)$, $(x, y \pm 1, z \pm 1)$, $(x \pm 1, y, z)$, $(x, y \pm 1, z)$, $(x \pm 1, y \pm 1, z)$ are pushed into a first-in-last-out stack to be assessed later. Every time one pixel is removed from the beginning of the stack. If this pixel is a foreground pixel, label l is assigned to the pixel and its 18 neighbors are pushed into the stack. The labeling of
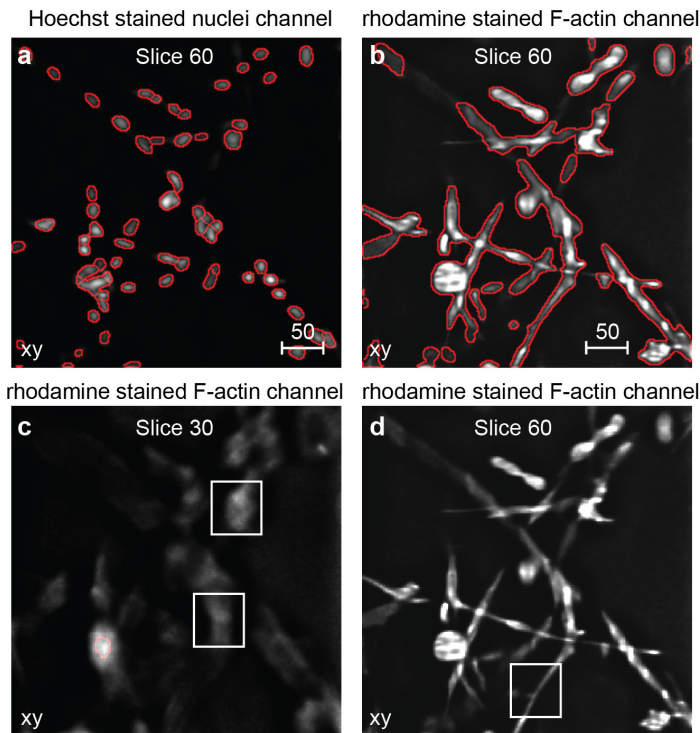
Figure 4.3: Segmentation results for both Hoechst stained nuclei channel and rhodamine stained F-actin channel. (a) Segmentation result of one cropped image slice (deconvolved) for the Hoechst stained nuclei channel. Red marks the contour of the segmentation result. (b) Segmentation result for the corresponding image (deconvolved) from the rhodamine stained F-actin channel. (c) A deconvolved image slice with low sharpness level, from the same stack. White boxes mark the regions which are out-of-focus but with high intensity. Red marks the contour of the segmentation result. (d) A deconvolved image slice with higher sharpness level than (c). White box marks the region which is in-focus but has lower intensity than the regions marked in (c). The segmentation result is shown in (b). The scale bar represents 50μm.

object l is finished when the stack is empty and a new labeling process $l + 1$ starts when an unsigned foreground pixel is found.

3D reconstruction is a process of constructing 3D geometrical models by triangularization of 3D surface area and connecting mesh of surface triangles based on foreground voxels (pixels transform to voxels by calibrating z-sampling size). Here, we applied the marching cube algorithm [15] for surface reconstruction of both nuclei and micro-tissue structures. Figure 4.4 shows an example of reconstructed nuclei and micro-tissue structure. We observed that both nuclei and micro-tissues were elongated in z-direction (Figure 4.4b, 4.4d). This can be caused by two reasons: one is due to the low NA which significantly degrades the imaging vertical resolution and the other one is the spherical aberration caused by refractive index mismatch between the objective immersion medium (Air/1.0) and cell culture medium (Collagen-
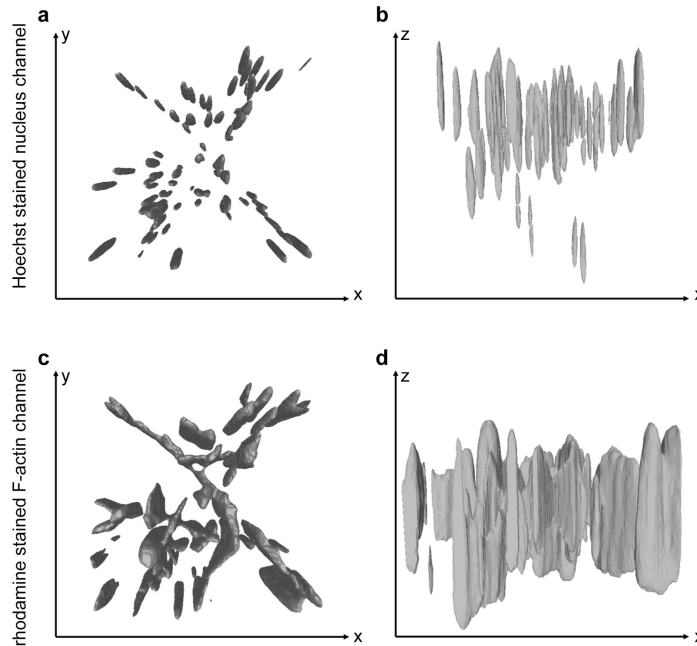
Figure 4.4: 3D reconstruction result for nuclei and multi-cellular micro-tissue structures. (a-b) 3D reconstructed nuclei in xy- and xz-direction, based on the same image stack as shown in Figure 4.3. (c-d) 3D reconstructed multi-cellular structures in xy- and xz-direction.

Matrigel/1.36). This elongation artifact was partially corrected by the deconvolution process; however, it still affects the reconstruction results so that the obtained 3D geometrical models do not resemble the correct object size in z-direction. To solve this problem, we developed a normalization method to calibrate the dimension of nuclei in z-direction according to the images of fluorescent beads with a known size. Blue fluorescent-labeled microspheres (Molecular Probes) with a 10μm diameter were used, as this size was close to the diameter of a nucleus. We firstly embedded the microspheres into the medium which we used to culture the PC3 cells. Subsequently, images were collected using the same microscope according to the parameters shown in Table S4.1. After deconvolution, we investigated the intensity profile of the beads in x-, y-, z-direction. Figures 4.5a and 4.5b show the same image slice through the middle plane of a bead and Figure 4.5c shows the intensity profile along the lines indicated in Figure 4.5a and 4.5b. These figures clearly show that in both x- and y-direction the diameter of the beads is ˜ 10μm when we set the intensity threshold as $1.4E + 04$ (Figure 4.5c). With the same intensity threshold, the diameter of the bead in z-direction is 150.17μm (Figure 4.5d, 4.5e). We measured beads located at different positions in several image stacks and obtained an average diameter of 150μm in z-direction, indicating that the spherical beads appeared to have a diameter in z-direction equal to 15 times their diameter in x-, y-direction. According to this result, we corrected the elongation effect of a nucleus by normalizing the coordinates

of nuclear foreground voxels as follows:

Step 1: Calculate the centroid (C_x, C_y, C_z) for each foreground voxel (x_p, y_p, z_p) with label l:

$$C_x = \frac{\sum_{p \in l} x_p}{n_l}, \; C_y = \frac{\sum_{p \in l} y_p}{n_l}, \; C_z = \frac{\sum_{p \in l} z_p}{n_l} \tag{4.4}$$

where p indicates a foreground voxel assigned with label l . n_l represents number of voxels assigned to l.

Step 2: Normalize z coordinate for each voxel with a factor of $\frac{1}{15}$

$$z_p^{calibrated} = \frac{z_p - C_z}{15} + C_z \tag{4.5}$$

Figure 4.5f shows the reconstructed nuclei after correction of the elongation artifact. Compared to Figure 4.4b which was generated from the same part of image stacks before correction, nuclei appear much more spherical in shape and this is more consistent with our expectation, while the distance between nuclear centroids is not affected. However, this normalization method is not suitable for the micro-tissue structures obtained from the rhodamine stained F-actin images because of their irregular shapes and sizes. We found that beads with different orders of magnitude (size) were elongated by different factors using the same microscope and medium. Normalizing all micro-tissue structures according to one factor would cause incorrect reduction of the elongation effect.

4.2.2 Phenotype measurement

4.2.2.1 Phenotype measurement for individual nuclei

Parameters to profile the phenotype of each individual nucleus can be categorized in three classes: morphological parameters, localization parameters and image moment parameters (Table 4.1). Morphological parameters include a series of shape properties and are measured from the reconstructed 3D geometrical models. In addition to the basic shape properties such as volume and surface, we computed a convex hull [16] and the best-fit ellipsoid [17] for each nucleus, and relative geometrical parameters were measured, for example the volume of the convex hull and semi-axis of the best-fit ellipsoid. Localization parameters estimate the nuclear density by calculating distance between pairs of nuclear centroids. Moment parameters include centroid coordinates and Eigenvalues calculated by Eigen decomposition of the covariance matrix of voxel coordinates. To calculate Eigen decomposition, all coordinates were normalized by moving centroid to origin according to Equation 4.6. Furthermore, we measured the inertia tensor matrix $\begin{matrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{matrix}$ for each nucleus l [17], according to the parallel axis theorem (Equation 4.7-4.12) [18] :
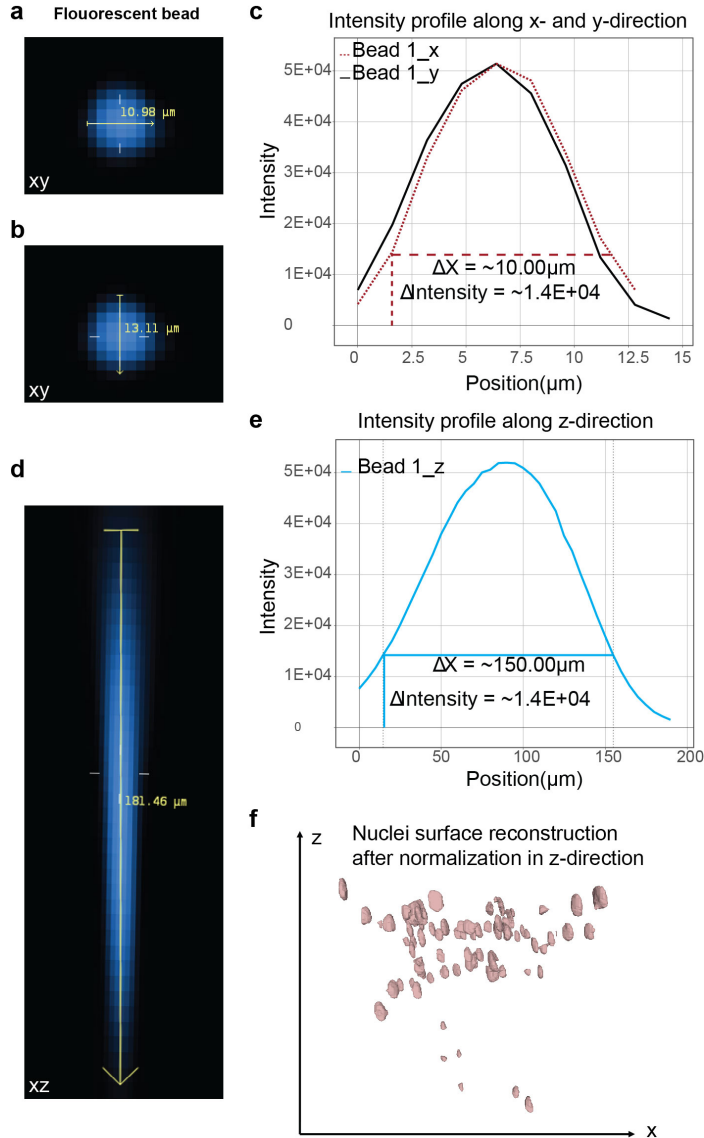
Figure 4.5: Normalization of nuclei in z-direction. (a-b) Image slices through the middle of one fluorescent bead in x- and y-direction. (c) The intensity profiles along the lines indicated in (a) and (b). Red dash line represents the intensity profile in x-direction and black solid line represents the intensity profile in y-direction. (d-e) A xz-plane through the middle of the same bead as in (a) and (b) and corresponding intensity profile. (f) 3D reconstruction for nuclei in xz-direction after normalization. The reconstruction result of the same image stacks before normalization is shown in Figure 4.4b.

$$x'_p = x_p - C_x, \; y'_p = y_p - C_y, \; z'_p = z_p^{calibrated} - C_z \qquad (4.6)$$

$$I_{xx} = \sum_{p \in l} (y'^2_p + z'^2_p + \frac{yScale^2 + zScale^2}{12}) \qquad (4.7)$$

$$I_{yy} = \sum_{p \in l} (x'^2_p + z'^2_p + \frac{xScale^2 + zScale^2}{12}) \qquad (4.8)$$

$$I_{zz} = \sum_{p \in l} (x'^2_p + z'^2_p + \frac{xScale^2 + yScale^2}{12}) \qquad (4.9)$$

$$I_{xy} = \sum_{p \in l} - (x'_p \times y'_p) \qquad (4.10)$$

$$I_{xz} = \sum_{p \in l} - (x'_p \times z'_p) \qquad (4.11)$$

$$I_{yz} = \sum_{p \in l} - (y'_p \times z'_p) \qquad (4.12)$$

where *xScale*, *yScale* and *zScale* equal the sampling size in x-,y-,z-direction. Subsequently, principle axes I_1, I_2, I_3 were computed by the Eigen decomposition of the inertia tensor matrix.

4.2.2.2 Phenotype measurement for multi-cellular micro-tissue network

In the 3D cell cultures, the invasive cancer cells spontaneously develop elongated and branched micro-tissue structures that are interconnected to form a complex network. To investigate the organization of those networks, we quantified the phenotypic properties based on whole image stacks. For the morphological profiling, we calculated the total size, volume and surface (Table 4.2) of all reconstructed multi-cellular structures found in one image stack (One example of reconstructed multi-cellular structure is shown in figure 4.6a). In addition, morphological parameters used to describe geometrical properties of whole micro-tissue networks (Table 4.2) were calculated. The convex hull of a micro-tissue network (Figure 4.6c) was constructed using the Quick-Hull algorithm after assigning all foreground voxels of the stack with one object label. For image moment parameters, Eigenvalues and principle axes were calculated based on the coordinates of all foreground voxels. In addition to morphological parameters and image moment parameters, we also quantified topological features for each multi-cellular structure, based on a topological skeleton (Figure 4.6b) computed by a 3-D thinning algorithm [20]. Every voxel that was part of the skeleton was labeled with different categories according to their 18-connected neighbors; voxels with one, two or more than two skeleton voxels in neighbor were respectively labeled as "End-point", "Slab-point" or "Junction-point". Next, the properties of the skeleton were calculated

based on the category of skeleton voxels, for example, the number of End-points and the branch length which is defined as the Euclidian distance between two adjacent skeleton voxels over all voxels on the skeleton. Finally, sum of those properties (Table 4.2) were calculated over all reconstructed multi-cellular structures in the image stack.

Morphological Parameters for each nucleus
Size: Number of foreground voxels assigned to a nucleus
Volume: Volume of a nucleus calculated by summarizing the volume of triangular pyramids that compose the 3D geometrical model.
Surface: Surface of a nucleus calculated by summarizing the area of triangles that compose the surface of the 3D geometrical model.
Width: Width of the 3D bounding rectangular box for a nucleus
Height: Height of the 3D bounding rectangular box for a nucleus
Thickness: Thickness of the 3D bounding rectangular box for a nucleus
Sphericity [19]**:** Compactness measure. Sphericity $= \frac{\pi^{\frac{1}{3}} \times (6 \times \text{Volume})^{\frac{2}{3}}}{\text{Surface}}$
SAV: Surface to volume ratio. SAV $= \frac{\text{Surface}}{\text{Volume}}$
Volume of convex hull: Convex hull was calculated using QuickHull algorithm [16]
Surface of convex hull
Solidity: Solidity $= \frac{\text{Volume}}{\text{Volume of convex hull}}$
Major axis: Length of the longest radius of the best-fit ellipsoid. The best-fit ellipsoid was calculated based on a least-square optimization algorithm according to a ImageJ plugin BoneJ [17]
Median axis: Length of the middle radius of the best-fit ellipsoid
Minor axis: Length of the shortest radius of the best-fit ellipsoid
Mass of the best-fit ellipsoid: Mass $= \frac{4}{3} \times \pi \times$ Major axis $\times$ Median axis $\times$ Minor axis
Moments of inertia of the best-fit ellipsoid: $I_{xx}^{ellipsoid} = \frac{1}{5} \times Mass \times (Median\,axis^2 + Minor\,axis^2)$ $I_{yy}^{ellipsoid} = \frac{1}{5} \times Mass \times (Major\,axis^2 + Minor\,axis^2)$ $I_{zz}^{ellipsoid} = \frac{1}{5} \times Mass \times (Major\,axis^2 + Median\,axis^2)$
Localization Parameters
Average distance to other nuclear centroids
Moments Parameters
Centroid(C_x, C_y, C_z), this is used to calculate Eigen decomposition.
Eigenvalue from Eigen decomposition: $\lambda_1, \lambda_2, \lambda_3$
Principle axes: I_1, I_2, I_3

Table 4.1: Phenotypic parameters measured for each nucleus.

Morphological Parameters
Total size: Number of foreground voxels in an image stack
Total volume: Sum of the volume of all micro-tissue structures within an image stack
Total surface: Sum of the surface of all micro-tissue structures of an image stack
Thickness: Thickness of a 3D bounding rectangular box for micro-tissue network
Sphericity: Sphericity(binary image stack) $= \frac{\pi^{\frac{1}{3}} \times (6 \times \text{Total volume})^{\frac{2}{3}}}{\text{Total surface}}$
SAV: SAV(binary image stack) $= \frac{\text{Total surface}}{\text{Total volume}}$
Volume of convex hull of micro-tissue network
Surface of convex hull of micro-tissue network
Solidity: Solidity(binary image stack) $= \frac{\text{Total volume}}{\text{Volume\_of\_convex\_hull}}$
Topological Parameters
Total number of End-point, Junction-point, Triple-point and Quadruple-point: Triple point is one kind of junction point which has 3 skeleton voxels in neighbor . Quadruple point is one kind of junction point which has 4 skeleton voxels in neighbor.
Total number of branches: Branch is defined as the part of skeleton between junction points, end points, or junction point and end point
Total length of all branches
Moments Parameters
Centroid of a micro-tissue network, which is used to calculate Eigen decomposition
Eigenvalue from Eigen decomposition: $\lambda_1, \lambda_2, \lambda_3$
Principle axes of a micro-tissue network: I_1, I_2, I_3

Table 4.2: Phenotype parameters measured for multi-cellular micro-tissue network.

4.2.3 Validation by comparison to the confocal laser scanning microscope results

To validate our method, we imaged the same fields of multiple wells using both the wide-field microscope and a confocal laser scanning microscope with a higher resolution and magnification objective (Table S4.2). The quantification results obtained from the confocal microscope were used to evaluate the quantification results obtained from our analysis pipeline in combination with the wide-field microscope. As image files generated by the confocal microscope were too large (4.05GB for each channel per well) to perform image processing on whole image stacks, we randomly cropped segments from each channel of confocal image stacks and the same fields of image stacks were cropped from the wide-field microscope image stacks (Figure 4.7). For the Hoechst stained nuclei channel, five image segments were generated including 125 nuclei, while ten image segments were generated for the rhodamine stained F-actin channel.

Image analysis on the wide-field microscope images was performed using our pipeline (Figure 4.7b, 4.7d). To define individual nuclei and multi-cellular structures correctly in the confocal microscope image stacks, we performed automated segmentation methods and then the segmentation results were validated by visual inspection.
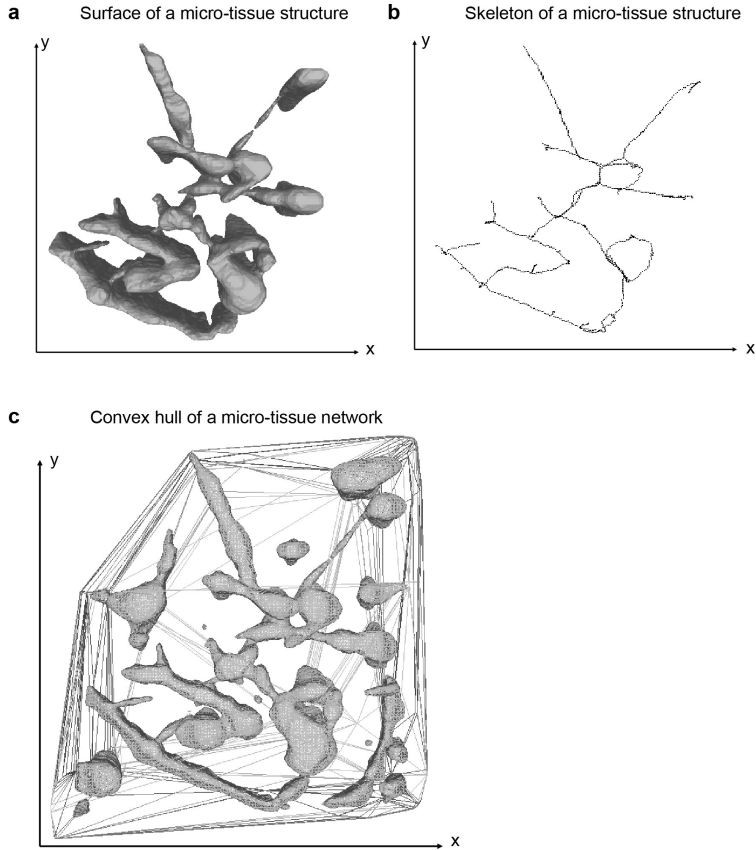
Figure 4.6: Morphological properties for micro-tissue structure and network. (a) Surface of a micro-tissue structure, and (b) corresponding skeleton. (c) Convex hull of a micro-tissue network which contains the structure shown in (a).

Before segmentation, we applied the rolling ball algorithm (radius = 50) and a median filter (kernel size 3×3 pixels) to the image stacks, to remove uneven background illumination and image noise. To segment individual nucleus, we performed Otsu segmentation method [22] on each image slice and then used a 2D watershed to divide connected nuclei (Figure 4.7c). Over-segmentation was corrected manually. For the segmentation of the rhodamine channel images, we applied a log K-means clustering algorithm, which firstly transformed images by taking the natural logarithm and then performed K-means clustering to define multi-cellular micro-tissue structures (Figure 4.7e). A sequential labeling algorithm based on 18-connected neighbor pixels was used to label each nucleus and micro-tissue structure after segmentation. Similar to the correction of elongated nuclei in the wide-field microscope images, we imaged blue fluorescent-labeled microspheres with a 10μm diameter for the calculation of a normalization factor using the same confocal microscope settings as for imaging the micro-tissues (Table S4.2). A normalization factor $\frac{1}{5}$ was obtained for the confocal microscope and then used to calibrate dimensions of nuclei in z-direction. Finally, we
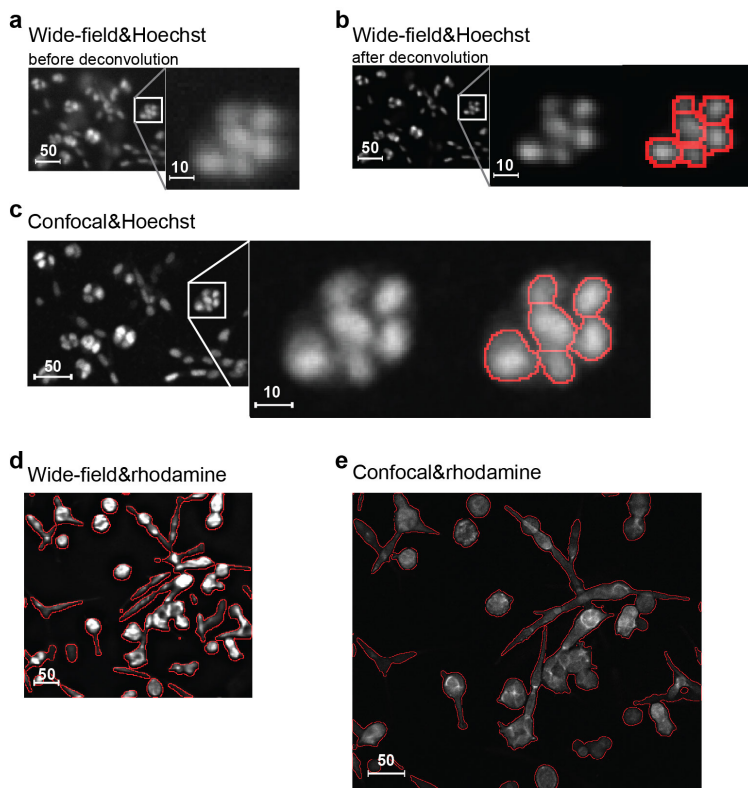
Figure 4.7: Comparison of wide-field microscope images with confocal microscope images. (a) A cropped segment of an original image slice obtained from the wide-field microscopy and Hoechst stained nuclei channel. A part of segment (white box) is magnified 5× to provide a clear perception of resolution. (b) The same image part after deconvolution. Red marks the contour of segmentation result. (c) The same field was cropped from an image slice obtained by the confocal microscope, and the contour of segmentation result is presented in red. (d) A cropped segment of a deconvolved image slice with segmentation result for rhodamine stained F-actin channel. (e) Corresponding region and segmentation result from the confocal microscope image. The unit of the scale bar is μm.

reconstructed nuclei and micro-tissue structures (Figure S4.1) based on the marching cube algorithm, and the phenotypic parameters (Table 4.1-4.2) were measured for those reconstructed structures from the confocal microscope.

4.2.3.1 Comparison of the quantification results of nuclei obtained from the wide-field microscope images and confocal microscope images.

We firstly compared the number of nuclei obtained by applying our image analysis pipeline to the wide-field microscope images with the manual counting of nuclei in the confocal microscope images (Table 4.3). Although the low resolution of the wide-field microscopy affected segmentation so that a slightly higher number of nuclei was detected using our automated method, the difference was not significant. Next we calculated a two-sample KS test for each of the parameters presented in Table 4.1

Cropped segments	Number of nuclei (Wide-field)	Number of nuclei (Confocal)
Segment 1	8	8
Segment 2	14	14
Segment 3	21	20
Segment 4	28	26
Segment 5	63	57
$p - value$ of two-sample KS test	> 0.9	

Table 4.3: Comparison of number of nuclei between confocal microscope images and wide-field microscope images.

(Figure 4.8a), except for the size and principle axes which are highly dependent on the image resolution and number of foreground voxels. The result shows that for most parameters no significant difference ($\alpha = 0.01$) was obtained comparing the quantification results of wide-field microscope images with confocal microscope images. In the end, we investigated the Pearson product-moment correlation coefficient for each parameter (Figure 4.8b) and obtained significant correlation ($p - values < 0.01$ for the hypothesis of no correlation) for most of parameters, further validating that by using our image analysis pipeline for the nuclear channel we are able to obtain comparable quantification results from the wide-field microscope images as from the confocal microscope images with relatively higher resolution. Figure 4.8c shows the scatter plot of volume for each nucleus measured from the two microscope techniques.

4.2.3.2 Comparison of the quantification result of micro-tissue networks obtained from the wide-field microscope images and confocal microscope images

Due to the different resolutions of the two microscopes and the fact that we did not calibrate the dimension of micro-tissue structures in z-direction, morphological parameters obtained from those two different image modalities cannot be compared directly to each other using KS tests. Nevertheless, the topological parameters that are independent of calibration result, such as the number of branches, should be comparable. Therefore, we firstly investigated the topological parameters using both Pearson's correlation and the two-sample KS test, and the result is presented in Table 4.4, which clearly shows a high correlation and no significant difference between the results obtained from two images modalities. We found that for most test segments, slightly more branches were detected from the confocal microscope images than from the wide-field images although the differences were not significant. We presume that this is due to the higher resolution of confocal microscope images so that more subtle details were preserved (Figure 4.9a).

Next, we calculated Pearson's correlation for each of the parameters that are presented in the Table 4.2. Figure 4.9b shows the coefficients for all parameters of which each has a high value, indicating the quantification results of multi-cellular micro-
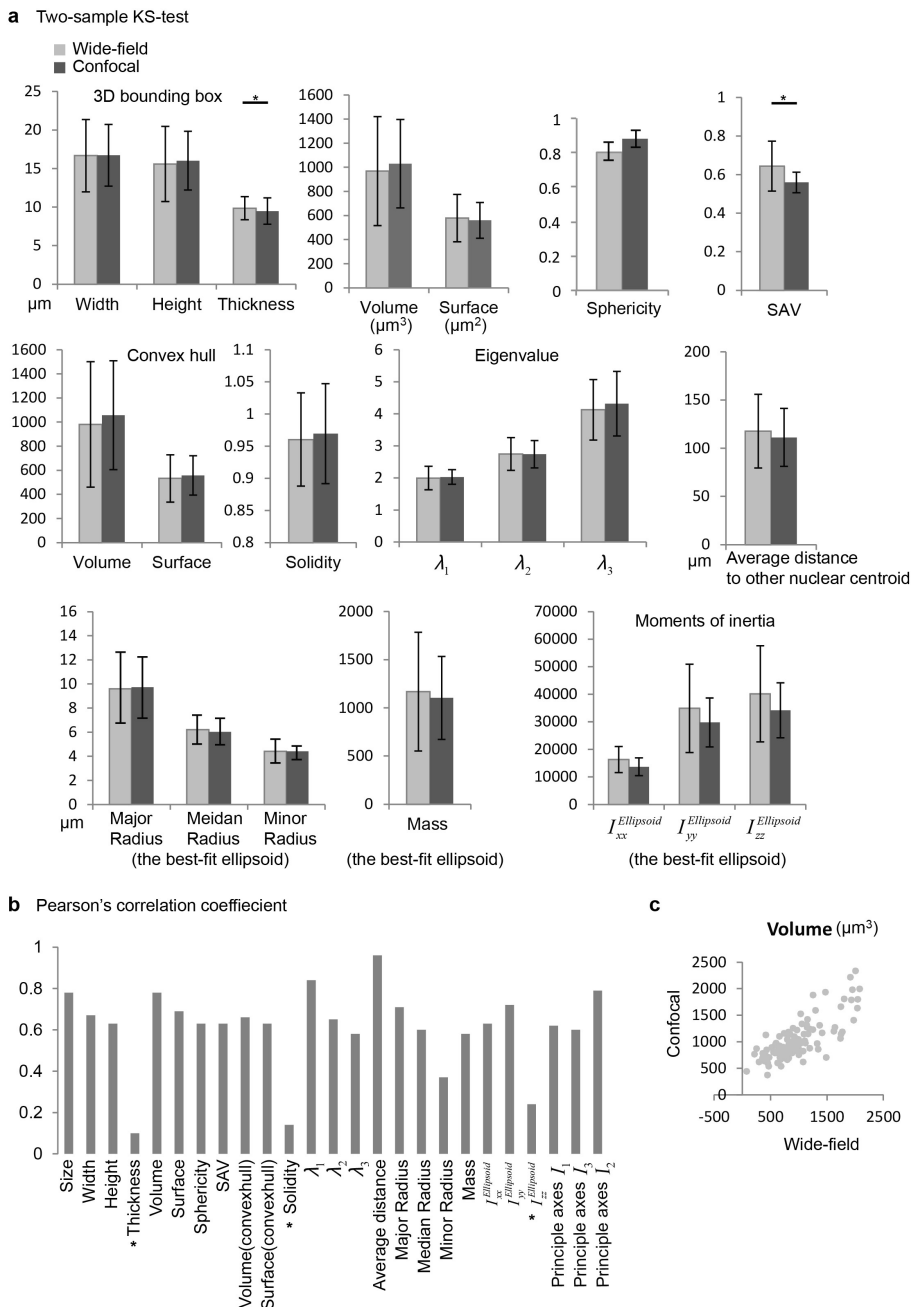
Figure 4.8: Comparison of the quantification results of the nuclei obtained from the wide-field microscope images and the confocal microscope images. (a) The result of the two-sample KS tests, comparing the quantification result between wide-field microscope and confocal microscope for the Hoechst stained nuclei channel. (b) Pearson's correlation coefficient between the quantification result from the wide-field microscope images and the results from the confocal microscope images. * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated. (c) The scatter plot of volume for each nucleus. x-axis represents the value from wide-field microscope images and y-axis represents the corresponding value from confocal microscope images.

cropped segments	#End		#Junction		#Triple		#Quad.		#Branches	
	WF	CF	WF	CF	WF	CF	WF	CF	WF	CF
Segment 1	88	88	75	68	65	64	1	1	157	146
Segment 2	147	157	82	95	66	82	9	8	203	223
Segment 3	45	51	39	36	35	33	2	1	80	82
Segment 4	57	74	21	35	14	33	2	1	58	91
Segment 5	106	132	70	76	55	67	9	4	163	181
Segment 6	65	80	44	56	41	51	3	1	107	115
Segment 7	56	57	31	37	27	33	3	3	77	85
Segment 8	111	128	77	79	66	68	5	7	175	186
Segment 9	38	53	34	34	37	31	2	0	71	76
Segment 10	83	94	72	80	62	72	3	4	151	168
Pearson's correlation	0.98		0.95		0.91		0.77		0.97	
Two-sample KS-test $p-value$	0.98		0.98		0.31		0.31		0.68	

Table 4.4: Comparison of topological parameters between the wide-field microscope and the confocal microscope for micro-tissue structures. "#" means "number of"; "End", "Junction", "Triple" and "Quad." mean "end point", "junction point", "triple point" and "quadruple point", respectively.

tissue structures obtained from the wide-field microscope images are comparable to the quantification results obtained from the confocal microscope images. Figure 4.9c and 4.9d show the scatter plots of the total volume and total surface of the micro-tissue network for each image segment, calculated from the two techniques.

4.2.3.3 Comparison of the time efficiency between the wide-field microscope and confocal microscope

Due to the slow scanning process, the confocal microscope requires more imaging time compared to the wide-field microscope, and therefore limits its application in high-throughput screening where time efficiency is an important consideration. To quantitatively illustrate this point, we compared the image acquisition time between the confocal microscope and the wide-field microscope that were used in this study. For the BD Pathway wide-field microscope system, approximately 150 seconds were required to collect a two channel image stack from one well, while our Nikon confocal microscope took 6 hours to capture the same area of one well. Due to this large amount of imaging time, severe bleaching was observed in the last image slices as well as in neighboring wells. Besides, larger image files were generated when we used higher magnification of objective for the confocal microscope, and this increased the computational time and required more computer memory to perform image analysis.

We also recorded the image acquisition time for the BD Pathway microscope system with the same settings as in Table S4.1, but now with laser scanning confocal mode: 105 minutes were required to collect a two channels of image stack from one well which compares very unfavorable to the image acquisition time in the wide-field
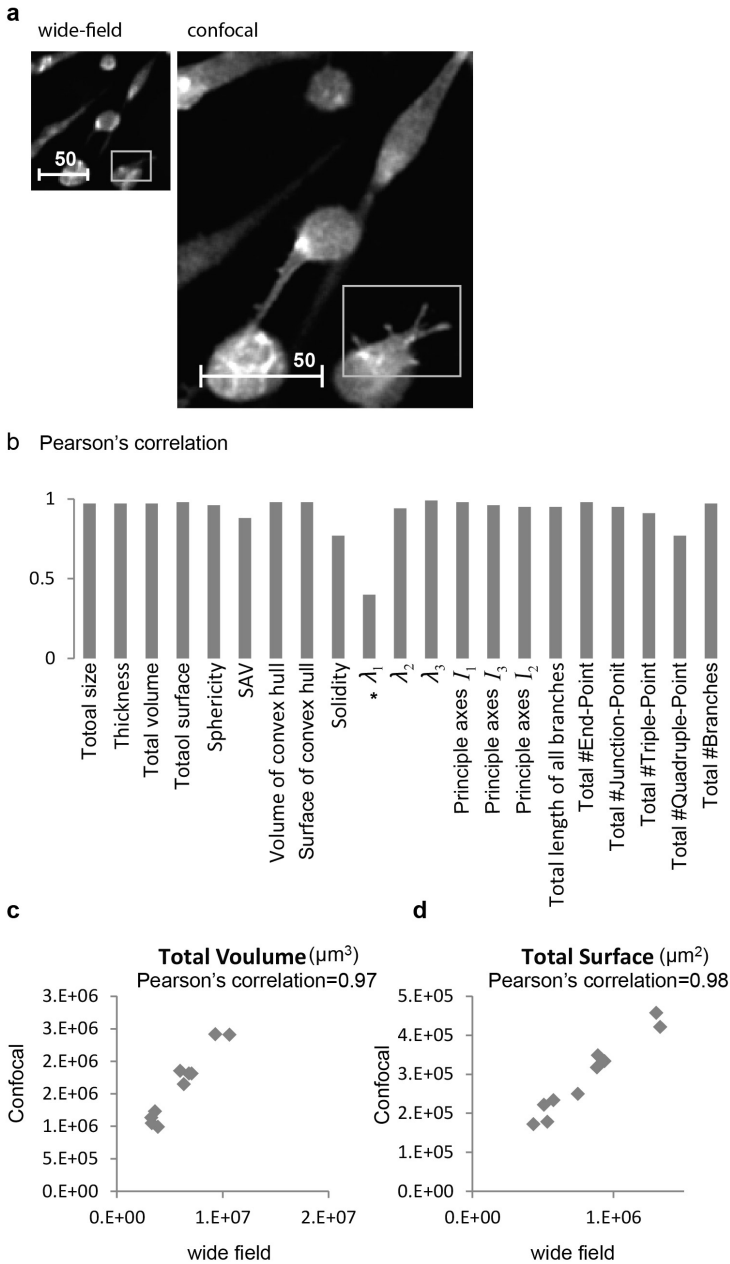
Figure 4.9: Comparison of the quantification result of the micro-tissue networks reconstructed from the wide-field microscope images and the confocal microscope images.
(a) One crop of the same field from a wide-field microscope image slice and a confocal microscope image slice. The white box high lightens the regions with different level of details. The unit of the scale bar is μm. (b) Pearson's correlation coefficient between the quantification result from the wide-field microscope images and from confocal microscope images. * represents $p-value > 0.01$ under the hypothesis that two data samples are not correlated. (c) The scatter plot of total volume and (d) the scatter plot of total surface of micro-tissue network for each image segment. For both (c) and (d), x-axis represents the value from wide-field microscope images and y-axis represents the corresponding value from confocal microscope images.

mode (150 seconds).

4.2.3.4 Assessment of different sampling step sizes in z-direction

Theoretically, increasing the sampling frequency in z-direction would improve the image vertical resolution, however, it would also affect the time efficiency because more image slices are required for imaging the same size of specimen. To find a balance between image quality and time efficiency, we investigated the effect of different z-sampling step sizes on the quantification results. Ten wells of image stacks were resampled with different sampling step sizes in z-direction (10μm, 15μm, 20μm) using the wide-field microscope system settings. The same resampling sizes were also applied to image the 10μm blue fluorescently labeled beads, in order to normalize the dimension of nuclei in z-direction. Next, our image analysis pipeline was applied to reconstruct 3D models for nuclei and micro-tissue network respectively, and finally the quantification results were evaluated by comparison to the quantification results obtained from the confocal microscope images. To our surprise, for the Hoechst stained nuclei channel the optimal performance was not always obtained with the smallest z-step size (5μm). Two-sample KS tests showed (Figure S4.2a) that the z-step size of 10μm and 5μm provided the closest results to the quantification results obtained from the confocal microscope images. The total number of nuclei obtained from the stacks with z-step size of 10μm and 15μm was closer to the number of nuclei from confocal microscope images, than from stacks with z-step size 5μm (Figure S4.2b). Similarly, according to the Pearson's correlation coefficient (Figure S4.2c) z-step size 10μm and 15μm provided higher correlations with the results from confocal microscope images than z-step size 5μm. We presumed this is due to the fact that the segmentation error was enhanced when the sampling frequency is too high and more image slices were analyzed. With the sampling size increased to 20μm, the difference of quantification results between wide-field microscopy and confocal microscopy was enlarged: 8 out of 21 parameters had significant difference based on the two-sample KS test (Figure S4.2a) and 7 parameters had correlation coefficient <0.5 (Figure S4.2b).

For the rhodamine stained F-actin channel, we measured the Pearson's correlation coefficient for different z-step sizes (Figure S4.3). The results showed that when the z-sampling step size was increased from 5μm to 15μm, quantification accuracy was decreased, however, not to a big extent except for the parameter Thickness. When the z-sampling size increased to 20μm, more parameters (SAV and number of quadruple-points) have a much lower Pearson's correlation coefficient. Considering the fact that a z-sampling size of 15μm is much more time efficient than z-sampling sizes of 5μm and 10μm while still providing quantification results comparable to the results of the confocal microscope images for both fluorescence channels, the z-sampling size of 15μm seems optimal for our future high-throughput screening of 3D cultured micro-tissues.

4.3   Discussion and conclusions

In this study, we developed an automated image analysis pipeline for 3D surface reconstruction and phenotype profiling of 3D cultured micro-tissues, suitable for high-content and high-throughput screening. It first applies a deconvolution technique to enhance the image quality by removing out-of-focus effects. Subsequently, two 3D segmentation methods were developed to identify individual nuclei and multi-cellular regions. Based on the segmentation results, a simple and efficient 3D reconstruction method was used to model the 3D structures of nuclei and multi-cellular micro-tissue structures. For nucleus surface structure, we performed a correction of dimension in z-direction to recover the nucleus from the elongation artifacts. Finally, 3D phenotypic parameters were measured directly on the reconstructed structures, including 3D morphological parameters, localization parameters, 3D topological parameters and moments.

This study intends to develop an image analysis pipeline to extract sufficient phenotypic characteristics from the conventional wide-field microscope, in order to achieve a high-content analysis. Granted, using fluorescence microscopy with higher resolution, for example confocal microscopy, would be greatly beneficial, yet due to their complexity, the applicability is often limited in throughput. In this study, we compared the imaging efficiency between a confocal microscope and a wide-field microscope. To image the same field of view in a well, the confocal microscope required up to 40 times more of image acquisition time compared to the time needed for the wide-field mode. Recently, more advanced microscopy techniques such as spinning disc confocal microscopy were made available for high-content screening. However, these techniques are too expensive to be widely available. Therefore, developing image analysis methods that are able to distill information from conventional microscope images would be a reasonable solution.

We have provided statistical evidence that by using this image analysis pipeline the quantification results obtained from the wide-field microscope are not significantly different from the results extracted from the confocal microscope. This achievement does largely rely on the image preprocessing including deconvolution, the segmentation algorithms and the reconstruction method. The most computational expensive part is deconvolution. It takes ~45 seconds to process one image stack on a server which is equipped with 16 Intel(R) Xeon(R) model E5530 processors and 24GB of RAM in total. The rest of our image analysis pipeline was performed on an Intel Core i7-2600 model with 16GB of RAM and a 64-bit Windows 7 operation system. Our complete image analysis of wide-field microscope image stacks for one 384-well plate takes approximately computational time of 1150 minutes. This time is slightly longer than the image acquisition time for one 384-well plate in wide-field mode. Considering the very significant benefit in image acquisition time, wide-field microscopy, in combination with our image analysis pipeline, has a substantial advantage in time efficiency over confocal microscopy, and we do not regard the time efficiency of our image analysis as an important bottleneck in the whole high-content screening pipeline, nor do we regard the minor differences between quantification results from wide-field

microscope and from confocal microscope as a major drawback.

In this study, we also investigated the effect of different z-sampling step sizes on the quantification results, in order to establish a balance between the image analysis accuracy and imaging time efficiency. Surprisingly, we found that the smaller step size is not necessary providing better quantification results, suggesting an oversampling problem which not only increases image acquisition time, computational complexity and image storage capacity, but also degrades the quantification result. Therefore, finding the optimal z-sampling size is crucial for high-content analysis of 3D cultured micro-tissues.

4.4 Methods

4.4.1 Cell culturing

Human prostate cancer cells (PC3) were cultured and exposed to hepatocyte growth factor (HGF) in a mixture of collagen type IV and laminin-rich basement membrane extract (Matrigel) for 4 days in 384-well high content imaging microplates. 72 hours after seeding, the cultured micro-tissues were fixed and stained with Hoechst 33258 and rhodamine-phalloidin to visualize nuclei and F-actin, respectively.

4.4.2 Image acquisition

For each well of a 384-well plate, two stacks of 152 xy epi-fluorescence image slices (16-bit) were collected from two fluorescence channels respectively, using a BD Pathway 855 automated microscope in wide-field mode. The gel was imaged through its entire depth (z-direction) and each image captured approximately 75% of the area of the well.

The confocal microscope images which were used to validate the accuracy of our image pipeline were collected using a Nikon Eclipse Ti confocal laser scanning microscope. For each well, two stacks of 71 xy epifluorescence image slices (16-bit) were generated. In order to capture whole well, each image slice was stitched by 9 images (3 images in a row and 3 images in a column), each of which captured one physical position of the well.

4.4.3 Software

For image analysis, ImageJ plugins (Java) were developed in-house, including a plugin to program a tcl script that can call the Huygens Core (http://www.svi.nl/HuygensCore) to run a batch process of a WideField deconvolution function in 384-or 96-well format, a plugin to perform segmentation, reconstruction and phenotypic quantification, and a plugin to compose obj files that contain vertex coordination of surface triangles. Meshlab (http://meshlab.sourceforge.net/) was used to visualize 3D geometrical models of reconstruction results.

References

[1] Lee, G. Y., Kenny, P. A., Lee, E. H. & Bissell, M. J. Three-dimensional culture models of normal and malignant breast epithelial cells. Nat Methods 4, 359-365 (2007).

[2] Pampaloni, F., Reynaud, E. G. & Stelzer, E. H. K. The third dimension bridges the gap between cell culture and live tissue. Nat Rev Mol Cell Bio 8, 839-845 (2007).

[3] Glory, E. & Murphy, R. F. Automated subcellular location determination and high-throughput microscopy. Dev Cell 12, 7-16 (2007).

[4] Pop, S. et al. Extracting 3D cell parameters from dense tissue environments: application to the development of the mouse heart. Bioinformatics 29, 772-779 (2013).

[5] Korn, K. & Krausz, E. Cell-based high-content screening of small-molecule libraries. Curr Opin Chem Biol 11, 503-510 (2007).

[6] Chi, C. Y. & Chen, W. T. An Adaptive Maximum-Likelihood Deconvolution Algorithm. Signal Process 24, 149-163 (1991).

[7] Sternberg, S. R. Biomedical Image-Processing. Computer 16, 22-34 (1983).

[8] Cao, L., Yan, K., Winkel, L., Graauw, M. D. & Verbeek, F. J. Pattern recognition in high-content cytomics screens for target discovery: case studies in endocytosis. In: PRIB'11 Proceedings of the 6th IAPR international conference on Pattern recognition in bioinformatics (eds Marco Loog, Marcel J. T. Reinders, Dick De Ridder, & Lodewyk Wessels) 330-342 (Springer-Verlag).

[9] Kuan Yan & Verbeek, F. J. Segmentation for high-throughput image analysis: watershed masked clustering. In: SoLA'12 Proceedings of the 5th international conference on Leveraging Applications of Formal Methods, Verification and Validation: applications and case studies. (ed Bernhard Steffen Tiziana Margaria) 25-41 (Springer-Verlag).

[10] Iannuccelli, E. et al. NEMO: a tool for analyzing gene and chromosome territory distributions from 3D-FISH experiments. Bioinformatics 26, 696-697 (2010).

[11] Vincent, L. & Soille, P. Watersheds in Digital Spaces - an Efficient Algorithm Based on Immersion Simulations. IEEE T Pattern Anal 13, 583-598 (1991).

[12] Ridler, T. W. & Calvard, S. Picture Thresholding Using an Iterative Selection Method. IEEE T Syst Man Cyb 8, 630-632 (1978).

[13] MacQueen, J. B. Some Methods for classification and Analysis of Multivariate Observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. (eds L. M. Le Cam & J. Neyman) 281-297 (University of California Press).

[14] Gonzalez, R. C. & Woods, R. E. Digital image processing. 3rd edn, 1-954 (Prentice Hall, 2008).

[15] Cline, H. E., Lorensen, W. E., Ludke, S., Crawford, C. R. & Teeter, B. C. Two algorithms for the three-dimensional reconstruction of tomograms. Med Phys 15, 320-327 (1988).

[16] Barber, C. B., Dobkin, D. P. & Huhdanpaa, H. The Quickhull algorithm for convex hulls. Acm T Math Software 22, 469-483 (1996).

[17] Doube, M. et al. BoneJ: Free and extensible bone image analysis in ImageJ. Bone 47, 1076-1079 (2010).

[18] Paul, B. Kinematics and dynamics of planar machinery. 1-670 (Prentice-Hall, 1979).

[19] Wadell, H. Volume, shape, and roundness of quartz particles. The Journal of Geology 43, 250-280 (1935).

[20] Lee, T. C., Kashyap, R. L. & Chu, C. N. Building Skeleton Models Via 3-D Medial Surface Axis Thinning Algorithms. Cvgip-Graph Model Im 56, 462-478 (1994).

[21] Arganda-Carreras, I., Fernandez-Gonzalez, R., Munoz-Barrutia, A. & Ortiz-De-Solorzano, C. 3D Reconstruction of Histological Sections: Application to Mammary Gland Tissue. Microsc Res Techniq 73, 1019-1029 (2010).

[22] Otsu, N. Threshold Selection Method from Gray-Level Histograms. IEEE T Syst Man Cyb 9, 62-66 (1979).

Supplementary figures
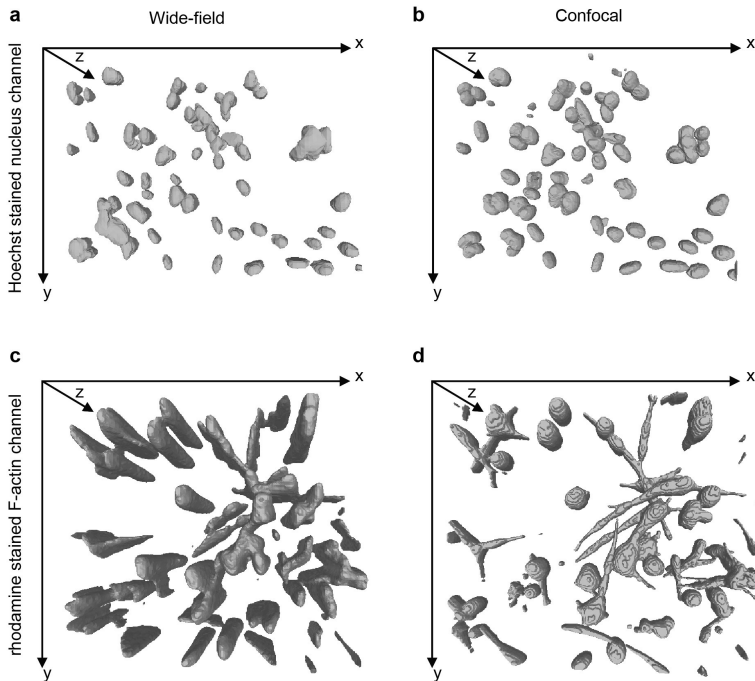
Figure S4.1: Comparison of reconstructed nuclei and micro-tissue structures based on the wide-field microscope images and the confocal microscope images. For the Hoechst stained nuclei channel, the reconstruction results of the same field of a well based on (a) the wide-field microscope image and (b) the confocal microscope image. The corresponding original images are shown in Figure 4.7a and 4.7c. For the rhodamine stained F-actin channel, the reconstruction results from (c) the wide-field microscope image and (d) the confocal microscope image. The corresponding original images are shown in Figure 4.7d and 4.7e.
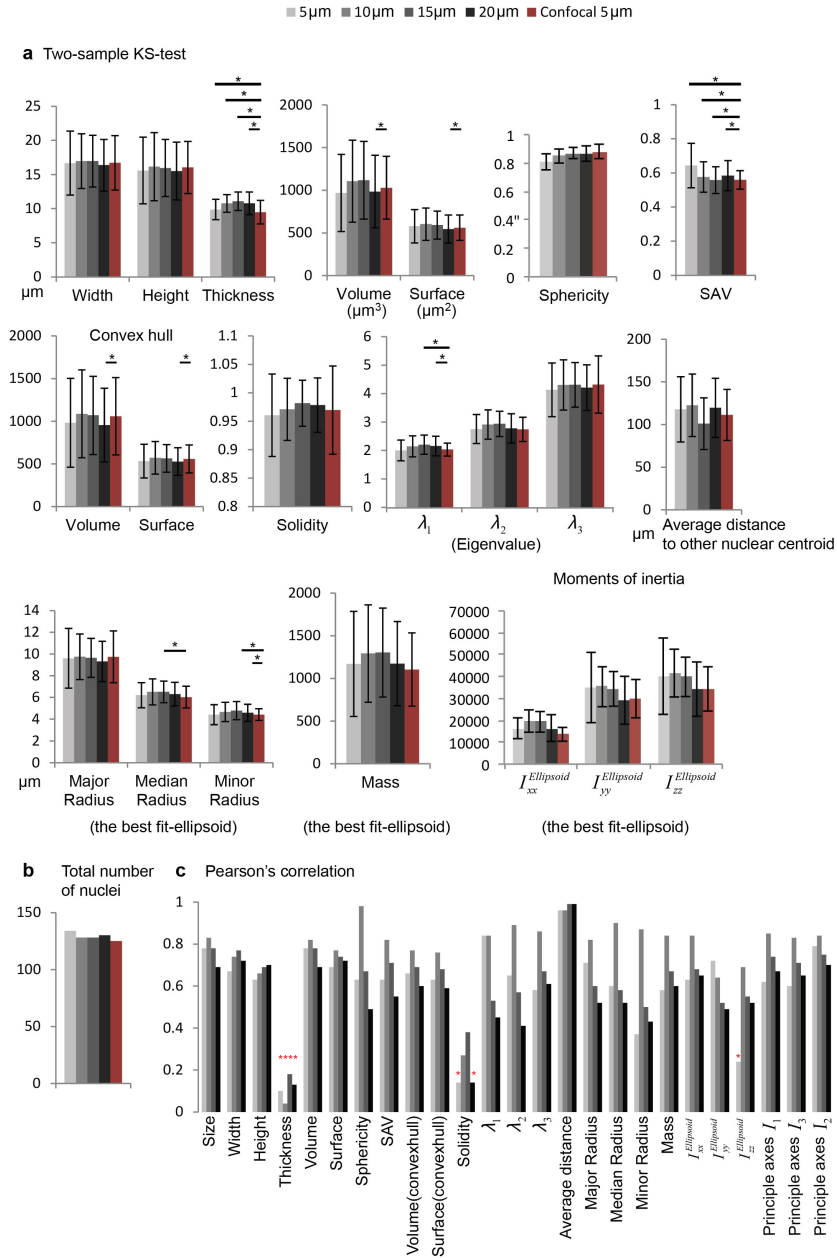
Figure S4.2: Comparison of the quantification results of the nuclei obtained from the different z-sampling step sizes with the results from the confocal microscope images. (a) The result of two-sample KS test, comparing the quantification results of different z-sampling step sizes to the results from the confocal microscope images for the Hoechst stained nuclei channel. For the confocal microscope images, the z-sampling step size is 5μm. For the wide-field microscope images, different sampling sizes are colored in different shades of grey. *: $p - value < 0.01$. (b) Comparison of the total number of nuclei obtained from 5 test segments between the wide-field microscope and the confocal microscope. (c) Pearson's correlation coefficient between each of the quantification result from the wide-field microscope images with different z-sampling step sizes and the results from the confocal microscope images. Red * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated.
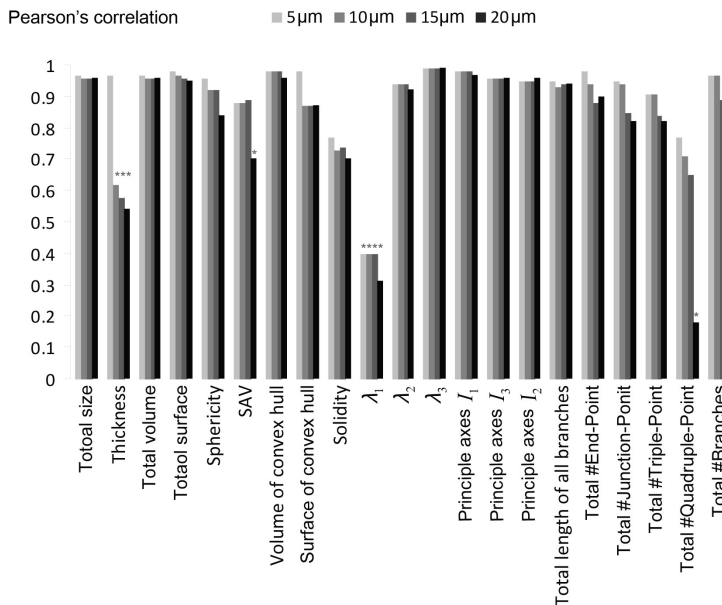
Figure S4.3: Comparison of the quantification results of the micro-tissue network obtained with different z-sampling step sizes and the results from the confocal microscope images. For the rhodamine stained F-actin channel, the Pearson's correlation coefficient between the quantification result from the wide-field microscope images with different z-sampling step sizes and the results from the confocal microscope images. For the confocal microscope, the z-sampling step size is 5μm. For wide-field microscope, different sampling sizes are colored in different shades of grey. * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated.

Supplementary tables

Objective type	BD Pathway Olympus 4XUPLAPO Plan-Apochromat
* Numerical aperture(NA)	0.16
Magnification	4×
* Lens refractive index	1.00 (Air)
* Medium refractive index	1.00 (Air)
* Hoechst excitation wavelength / bandwidth	380 nm/10nm
* Hoechst emission wavelength / bandwidth	435 nm/LP
* Rhodamine excitation wavelength / bandwidth	555 nm/28nm
* Rhodamine emission wavelength / bandwidth	645 nm/75nm
* Sampling size in x-,y-direction	1.60 μm
* Sampling size in z-direction	5 μm
Size of image stack (x,y)	1344 × 1024 pixels

Table S4.1: The parameters of the wide-field microscope to acquire image stacks. LP refers to low-pass filter. The parameters used for Huygens deconvolution software are marked with *.

Objective type	Nikon Plan Fluor 10X DIC L N1
* Numerical aperture(NA)	0.3
Magnification	10×
* Lens refractive index	1.00 (Air)
* Medium refractive index	1.36
* Hoechst excitation wavelength / bandwidth	405 nm
* Hoechst emission wavelength / bandwidth	450 nm/50nm
* Rhodamine excitation wavelength / bandwidth	561 nm
* Rhodamine emission wavelength / bandwidth	595 nm/50nm
* Sampling size in x-,y-direction	0.63 μm
* Sampling size in z-direction	5 μm
Size of image stack (x,y)	5530 × 5530 pixels

Table S4.2: The parameters of the confocal laser scanning microscope.

Chapter 5

Comparison of phenotype profiling of 3D cultured micro-tissues for ultra-high content screening by a 2D projection based method and a 3D method

Zi Di, Maarten J D Klop, , Bob van de Water, Fons J Verbeek,
Leo S Price & John H N Meerman

Abstract:

Due to the increased physiological relevance and often more in vivo like functioning of 3D cell cultures compared to 2D cell cultures, 3D cell cultures have been intensively applied in high-content screening to investigate cellular responses to different genetic or chemical perturbations. In previous work, we investigated mouse breast cancer cells which formed characteristic phenotypes of micro-tissues in 3D cultures after being exposed to a diverse set of anti-cancer drugs, using a novel screening and ultra-high content analysis platform. This platform firstly generated a 2D projection for each image stack that was obtained by imaging micro-tissues in z-direction. Subsequent image analysis was performed on the projections, including segmentation of nuclei and micro-tissue structures, and phenotype quantification. We showed that this platform can successfully retrieve the concentration dependent phenotypic trajectory of each compound that could be used to classify compounds based on their biological target. However, this platform did not take into account tissue development in z-direction and this information might play an important role in phenotype characterization. To fill this gap, we developed a 3D image analysis platform for ultra-high content analysis of 3D cultured micro-tissues. Instead of generating 2D projections of image stacks, this platform reconstructed 3D objects from image stacks and measured phenotypes from the reconstructed objects or directly from the image stacks. Here, we compared the performance of both platforms for the analysis of mouse breast cancer cells cultured in 3D in a 384-well micro plate format, following exposure to 12 compounds at different concentrations. A thorough comparison of those two platforms was performed using criterions of sensitivity, classification accuracy and reproducibility. The results showed that while 2D analysis is more sensitive in detecting a biological effect of the test compounds, 3D analysis has an increased classification accuracy and a significant improvement in reproducibility.

5.1   Introduction

Because 3D cell cultures have a higher physiological relevance than 2D cell cultures, 3D cell cultures have been widely used in high-content screening (HCS) to study the effects of large libraries of compounds, peptides, or RNA interference (RNAi), by investigating cellular phenotypic responses to these factors. Especially for tumor migration and invasion studies where the culture microenvironment, cell-cell, and cell-extracellular matrix (ECM) communications play a crucial role [1, 2], 3D cell cultures mimic the natural organization of tissues more closely than 2D cultures, enabling tumor cells to develop complex micro-tissue phenotypes. However, using 3D cell cultures in high-content and high-throughput screening adds a higher level of complexity in respect of assay preparation, image acquisition and especially image analysis, thus creating numerous technological challenges. In typical 3D HCS, cells are fixed and stained with fluorescence dyes and then z-stack images are acquired using an automated fluorescence microscope. The z-stack images can be used to reconstruct 3D surfaces of cells, or they can be compressed by projecting into single 2D image slices. Subsequent phenotype profiling is performed on the image stacks, 3D surfaces or projected 2D slices depending on which method is used.

For 2D projection based platforms, many algorithms have been developed to generate 2D image slices from image stacks, including maximum intensity projection [3] and in-focus slice extraction [4]. In our previous study (**chapter 3**) we have developed an automated image analysis platform for ultra-high content analysis (uHCA) of 3D cultured micro-tissues, based on 2D projection of image stacks. We generated a composite 2D image slice for each image stack of a fluorescence channel by projecting only in-focus regions from each slice. Subsequently, segmentation and phenotype quantification was applied on the composite 2D images. In **chapter 3**, this analysis platform was applied to investigate mouse breast cancer cells (4T1) in a 384-well micro plate format following exposure to a diverse set of anti-cancer drugs at different concentrations. Based on 598 parameters, we were able to retrieve an unique concentration dependent trajectory for each tested compound, and successfully demonstrated that modulating different biological targets may effect phenotypic development of 4T1 cells in a characteristic way.

One advantage of this projection based method is time efficiency, because it avoids extensive computations such as 3D segmentation and 3D reconstruction. Moreover, because the z-sampling frequency can be much less than required for 3D reconstruction, image acquisition time often is significantly reduced. However, information of tissue development in z-direction was not taken fully into account in the 2D projection based analysis presented in our previous work (**chapter 3**), and this information might play an important role in phenotype characterization. To fill this gap, more recently (**chapter 4**) we developed a new analysis platform which is able to extract information of phenotype in all 3 dimensions with great details. Efficient 3D segmentation and reconstruction methods were developed for both single cell based phenotype profiling and cell cluster (or multi-cellular structure) based phenotype profiling. Many parameters measured in 2D analysis platform were extended to their equiva-

lent 3D version such as morphological parameters, topological parameters, intensity properties, moments, as well as subpopulation related parameters. Furthermore, new 3D parameters were measured to maintain an "ultra" high level of content in our 3D analysis of phenotype.

Because some types of parameters can only be measured in the 2D projection based platform so far due to their complexity, such as Gabor wavelet, this may decrease the performance of our 3D HCA platform despite the theoretical advantages of a full 3D analysis. To discover whether our 3D profiling improves the phenotype characterization compared to the 2D projection based profiling, we evaluated the performance of both platforms in a HCS where 4T1 cells were cultured in ECM-rich hydrogel in 384-well micro plates and then exposed to 12 biologically active compounds at different concentrations. Firstly, we tested the sensitivity of the two platforms according to their ability to identify concentrations that significantly influenced the phenotype development of 4T1 cells. Secondly, we classified compounds according to their biological activity and evaluated the classification accuracy of both platforms for phenotype profiling. In the end, we tested the reproducibly of both platforms based on the concentration dependent phenotypic trajectories. Trajectories of the same compounds in independent duplicate screens should locate more close to each other than to trajectories of different compounds.

5.2 Methods

5.2.1 High-content screening of 4T1 cells

The 3D cell culturing method has been described previously (**chapter 3**). In short, 4T1 cells were cultured in 384-well micro plates (Figure 5.1a) and exposed to 12 compounds (Table 5.1) at 6 different concentrations (0.03μM, 0.1μM, 0.316 μM, 1 μM, 3.16 μM, 10 μM) in quadruplicate. Compounds were dissolved in DMSO such that the final concentration in the cell incubation medium was 0.2%. Control cells were exposed to DMSO (0.2%) only. As a blind test, we repeated the experiment independently including 11 compounds (without erlotinib HCl) and the compound treatments were concealed. The test compounds can be categorized according to biological activities in BCR-ABL inhibitors, VEGFR inhibitors, c-MET inhibitors, autophagy inducer, EGFR inhibitors and HDAC inhibitors. After 72 hours of exposure, cells were fixed and stained with Hoechst 33258 and rhodamine-phalloidin to visualize nuclei and F-actin, respectively. We obtained two channels of epifluorescence image stacks from each well (Figure 5.1b), using a BD Pathway 855 automated microscope fitted with an air objective lens with magnification 4× and numerical aperture (NA) 0.16. The sampling size in x-, y-, z-direction was 1.6μm, 1.6μm and 15μm, respectively.

Compound Index	Compound Name	Biological activity
1	Arq 197	tyrosine kinase inhibitor (c-MET), Positive control
2	AZD 0530	tyrosine kinase inhibitor (SRC family, BCR-ABL)
3	dasatinib	tyrosine kinase inhibitor (SRC family, BCR-ABL)
4	entinostat	HDAC inhibitor
5	imatinib mesylate	tyrosine kinase inhibitor (PDGFR, KIT, BCR-ABL)
6	nilotinib	tyrosine kinase inhibitor (BCR-ABL, KIT, LCK, EPHA3, EPHA8, DDR1, DDR2, PDGFR, MAPK11 and ZAK)
7	sorafenib tosylate	tyrosine kinase inhibitor (VEGFR, PDGFR, Raf kinases)
8	stf-62247	autophagy inducer
9	sunitinib maleate	tyrosine kinase inhibitor (VEGFR, EGFR, RET, KIT)
10	vandetanib	tyrosine kinase inhibitor (VEGFR, EGFR, RET, KIT)
11	erlotinib HCl	tyrosine kinase inhibitor (EGFR)
12	genistein	tyrosine kinase inhibitor (EGFR), Caspase, PPAR gamma and Topoisomerase II inhibitor, Estrogenic activity

Table 5.1: 12 compounds used in the 4T1 cell screening and their corresponding biological activities.

5.2.2 Automated phenotype profiling platform based on 2D projection

An ultra-high content 3D phenotype profiling platform based on 2D projection has been described previously (**chapter 3**). In short, an in-house ImageJ plugin was developed in Java to perform 2D projection, segmentation and phenotype quantification. For each image stack, a single image plane was firstly composed by projection of only in-focus regions from each slice (Figure 5.1c). Segmentation (Figure 5.1d) and quantification (Figure 5.1e) were then applied to the 2D projected images. For segmentation, we applied a 2D watershed masked clustering [5,6] to segment individual nuclear regions and a local Niblack algorithm [7] to segment cell clusters. After segmentation, quantification was carried out on the segmentation results or on the 2D projected images, as described in **chapter 3**. Measured parameters included morphological parameters, topological parameters, intensity parameters, texture parameters, moments and subpopulation parameters. In this study, we extended the ability of this platform to measure additional morphological and topological parameters compared to our earlier work (**chapter 3**) (see Table S5.1 and Supplementary note for the detailed description). Moreover, the correlation between the two fluorescence channels was also taken into account (Table S5.2). In total, 794 parameters were measured to profile each well.

5.2.3 Automated 3D phenotype profiling platform

We have developed an in-house ImageJ plugin in Java to automatically profile the phenotype of 3D cultured micro-tissues for uHCA based on the reconstructed 3D structures. Firstly, a deconvolution method (Figure 5.1f) was applied to the stacks

to restore images which were degraded due to light scattering. This was achieved by running a batch process script on the Huygens Core to call a WideField Deconvolution function (**chapter 4**). Next, we used a 3D watershed masked clustering to segment each individual nucleus region (Figure 5.1g) on the Hoechst stained nuclei channel. For cell clusters segmentation on the rhodamine stained F-actin channel, we applied a background variation based method which is described in **chapter 4** (Figure 5.1g). 3D surface reconstruction (Figure 5.1h) was carried out using a marching cube algorithm [8] for both channels. To recover each reconstructed 3D nucleus structure from the elongated artifact (caused by low NA and low vertical resolution), a normalization method was applied to calibrate the dimension of nuclei in z-direction, as described in **chapter 4**. However, this normalization method is not suitable for the cell cluster structures because of their irregular shapes and sizes.

Quantitative 3D phenotyping was carried out on the 3D reconstructed objects or on the image stacks directly (Figure 5.1i). Similar to the 2D projection based profiling platform, parameters measured by the 3D profiling platform can be categorized as morphological parameters, topological parameters, intensity parameters, 3D moments, Haralick co-occurrence parameters and subpopulation parameters. In total 290 parameters were calculated for each well:

1. **Morphological and topological parameters:** Morphological parameters (Table S5.3) include a series of shape properties and are measured directly from each reconstructed 3D object. We firstly extended 2D morphological parameters to their equivalents 3D version. such as changing area to volume. In addition, many parameters were defined as described in **chapter 4** (Table 4.1 and 4.2). To measure the topological parameters (Table S5.3), a topological skeleton of each object was firstly computed using a 3D-thinning algorithm [9] and relative parameters (Table S5.3) were measured from the obtained skeleton (see **chapter 4** for details). In the end, the mean, sum and standard deviation of each parameter over all 3D objects were quantified for the phenotypic profiling of each well.

2. **Intensity parameters:** Intensity parameters (Table S5.4) were measured on the image stacks which were already processed by deconvolution. Except for the basic intensity properties such as average intensity, maximum intensity of whole image stack, intensity gradient information was also taken into account by using a Sobel filter [10] to calculate the gradient magnitude for each pixel. Since image sampling size in z-direction (15μm,) is different from that in the horizontal direction (1.6μm), the Sobel filter was only applied in the horizontal direction. In the end, the mean, maximum and standard deviation of gradient magnitude were calculated over all pixels in one image stack.

3. **3D Zernike moments** [11]: We extended 2D Zernike moments to their comparable 3D versions (Table S5.5). For each binary image stack (after segmentation) which consists of pixels assigned with 1 (foreground) and 0 (background), we firstly applied normalization as follows so that the center of each image stack
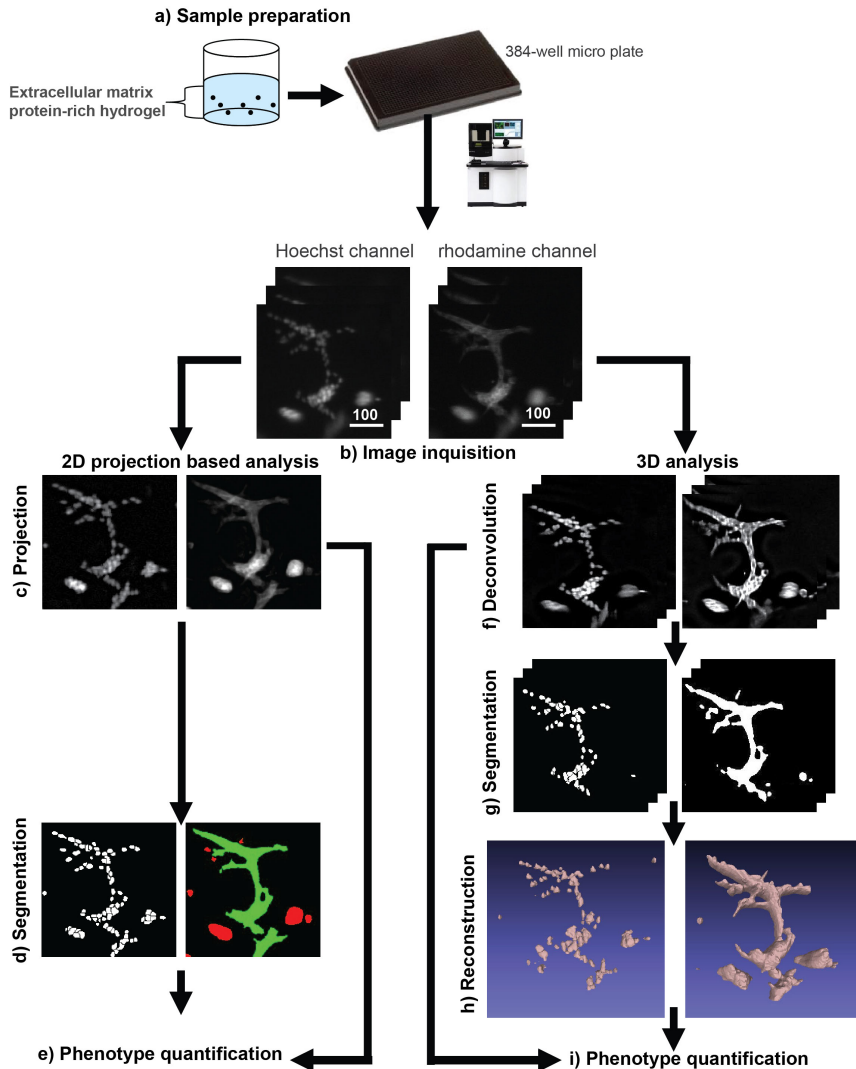
Figure 5.1: Workflow of uHCS of 3D cultured micro-tissues for both 2D projection based analysis platform and 3D analysis platform. (a) Cancer cells, for example 4T1 cells, are grown in the 3D cell culture and then transferred to a 384-well micro plate. (b) Two channels of image stacks are produced by a BD pathway 855 microscope or equivalent apparatus in the wide-field mode with $4\times$ magnification of air objective with 0.16 NA. Scale bar represents 100μm. (c) In-focus 2D images are projected from the stacks of the Hoechst stained nuclei channel and the rhodamine stained F-actin channel. (d) Nuclear masks are obtained by 2D Watershed Masked Clustering, and masks of cell clusters are obtained using local Niblack segmentation. The green masks represent the subpopulation of branched and interconnected multi-cellular structures. The red masks represent the subpopulation of spherical cell clusters. (e) Quantitative parameters are derived for each well from segmentation results and projected images. (f) Deconvolution is applied on each channel of image stack, separately. (g) Nuclear masks are obtained by 3D Watershed Masked Clustering, and masks of cell clusters are obtained using the background variation based segmentation. (h) 3D surface reconstruction results. (i) Quantitative parameters are derived for each well from reconstructed objects and deconvolved image stacks.

(x_c, y_c, z_c) was transformed to the origin:

$$(x_c, y_c, z_c) \quad = \quad (\frac{imageWidth}{2}), (\frac{imageHeight}{2}), (\frac{imageThickness}{2}) \quad (5.1)$$

$$x_i' = (x_i - x_c) \tag{5.2}$$

$$y_i' = (y_i - y_c) \tag{5.3}$$

$$z_i' = (z_i - z_c) \tag{5.4}$$

where (x_i, y_i, z_i) is the original coordinate of pixel i and (x_i', y_i', z_i') is the corresponding coordinate after transforming. $imageWidth$ and $imageHeight$ represent the number of pixels in each row and column of an image slice, respectively. $imageThickness$ is the number of slice per image stack. Subsequently, the image was scaled as follows so that whole image stack could be mapped into the unit 3D ball:

$$x_i'' = \frac{x_i'}{\frac{imageWidth}{2}} \times \frac{1}{\sqrt{3}} \tag{5.5}$$

$$y_i'' = \frac{y_i'}{\frac{imageHeight}{2}} \times \frac{1}{\sqrt{3}} \tag{5.6}$$

$$z_i'' = \frac{z_i'}{\frac{imageThickness}{2}} \times \frac{1}{\sqrt{3}} \tag{5.7}$$

where x_i'', y_i'', z_i'' is the coordination of pixel i after scaling. In the second step, we calculated all geometrical moments of each scaled image stack up to degree of n:

$$M_{rst} = \sum_{x_i} \sum_{y_i} \sum_{z_i} (x_i'')^r (y_i'')^s (z_i'')^t f(x_i, y_i, z_i) \tag{5.8}$$

where $r, s, t \geq 0$ and $r + s + t \leq n$. $f(x_i, y_i, z_i) = 0$ when pixel (x_i, y_i, z_i) is a background pixel and $f(x_i, y_i, z_i) = 1$ when pixel (x_i, y_i, z_i) is identified as a foreground pixel. In the third step, we calculated all Ω_{nl}^m according to the literature [11]. Here $n \in [0, N], l \in [0, n]$ such that $(n - l)$ be an even number and $m \in [-l, l]$.

$$\Omega_{nl}^m = \frac{3}{4\pi} \sum_{r+s+t<n} \overline{\chi_{nlm}^{rst}} M_{rst} \tag{5.9}$$

Since the calculation of χ is independent of a particular image stack, it only needs to be calculated once for all stacks to save computational time. Finally, all Zernike moments up to degree of n were calculated according to:

$$F_{nl} = \| (\Omega_{nl}^l, \Omega_{nl}^{l-1}, \Omega_{nl}^{l-2}, ..., \Omega_{nl}^{-l})^t \| \tag{5.10}$$

In this project, 9 3D Zernike moments($n = 4$) were calculated for each image stack. Later, F_{00} was not included for profiling because it is equal to the number of foreground pixels which is already counted in the category of morphological parameters.

4. **Haralick texture parameters**: The Haralick texture parameters [12–14] measure the co-occurrence distribution based on deconvolved grey level images. Simply speaking, the co-occurrence distribution measures the frequency that a particular gray level is found adjacent to another gray level. Many researches [14, 15] have already shown the importance of the Haralick parameters in phenotype recognition, therefore we also included them in this project. We firstly normalized the grey value of each deconvolved image stack (32-bit) to $[0, 255]$. Secondly, co-occurrence matrices were generated according to the literature [14]. In an image stack, adjacency can occur in each of 13 directions, thus 13 co-occurrence matrices were calculated corresponding to each direction. Finally, 13 statistics (Table S5.5) were extracted from each co-occurrence matrix, and two types of texture features were generated for phenotype profiling: the average and the range between maximum and minimum of each statistics over all 13 directions.

5. **Subpopulation parameters**: In our earlier work (**chapter 3**) which was based on 2D projection profiling analysis, we have determined that the implementation of subpopulation parameters increased the accuracy of phenotype classification. Therefore, in this platform, we embedded a simple classifier to distinguish spherical cell clusters and branched cell clusters according to the sphericity and number of branches (see **chapter 4** for the description of sphericity and number of branches). The cell clusters with sphericity higher than 0.5 and number of branches smaller than 2 were classified as spherical cell clusters, and the rest of the clusters were classified as branched structures. Finally, relevant information was collected for both subpopulations (Table S5.6).

5.3 Results

To evaluate the performance of 2D projection profiling platform versus the 3D profiling platform, three criterions were used based on hits identification, phenotype classification and reproducibility of concentration dependent trajectories. To achieve this, quality control was firstly carried out to exclude wells that appear to be outliers compared to other replicates. It was determined by visual validation of the original image stacks that they were due to errors occurring at the image acquisition stage. Next, the robust $z - score$ [16] was calculated for each parameter in order to implement cross-plate comparison.

5.3.1 Hit identification

For high-content screening, the primary goal is to identify compounds and their concentrations which significantly influence biologically relevant phenotypes compared to control condition. Those identified compounds or concentrations are usually referred to as "hits". Many multi-parametric tests have been applied for this purpose including: Mahalanobis distance [17, 18], Chi-square [19] and Wilks' lambda test [20].

In our previous study where we investigated the phenotypic response of 4T1 cells to different anti-cancer compounds in 3D cell cultures (**chapter 3**), Mahalanobis distance yielded the closest result to visual scoring a treatment as having an effect on phenotype. Therefore, we chose to measure the Mahalanobis distance of each well to the DMSO control in this screening for hits identification. Firstly, a principle component analysis (PCA) was performed based on the $z - scores$ of all measured parameters, resulting in 15 principle components (PCs) obtained from 2D projection analysis and 13 PCs from 3D analysis respectively, which preserved 90% of the data variation. Next, Mahalanobis distance was measured based on the PCA result, in order to quantify the similarity of phenotype between each treated well and control condition and next, the median Mahalanobis distance was calculated over all replicates. To identify hits, we used a significance level $\alpha = 0.05$. Table 5.2 shows all 12 test compounds and their corresponding concentrations with $p - values$ smaller than 0.05, in either 2D projection analysis or 3D analysis. From a total of 23 compounds (12 test compounds in the first experiment and 11 compounds in the repeated, blind test experiment) $\times$ 6 concentrations = 138 conditions, 99 conditions were identified as hits using 2D projection analysis platform, while 84 conditions were identified as hits using 3D analysis platform. For the 11 compounds which were retested in the blind experiment, 2D projection analysis platform detected 50 conditions as hits in the first experiment and 40 of them (80%) were confirmed in the blind test experiment. When we used 3D analysis, 43 conditions of 11 compounds were identified as hits in the first experiment and 36 of them (83.7%) were confirmed in the blind test experiment. For both experiments, more conditions were identified as hits using 2D projection analysis platform than using 3D analysis, suggesting that 2D projection analysis might be more sensitive in terms of hits identification. However, this higher sensitivity of 2D projection analysis may be the consequence of measuring more parameters than the 3D analysis. This is further substantiated by comparing this 2D projection based platform with the previous 2D projection based platform which measured only 598 parameters for each well (**chapter 3**). We noticed that the relatively low concentration of compounds entinostat (0.03μ**M**), genistein (0.316 μ**M**), imatinib mesylate (0.316 μ**M**) and sunitinib maleate (0.316 μ**M**) were not detected using the 598 parameters based platform (Table S3.8), but were identified as hits when 794 parameters were used in this study.

5.3.2 Phenotype classification

In our previous study (**chapter 3**), we have observed that compounds with different biological targets, each induced a typical phenotype (recapitulated in Figure S5.1, left panel). Here, we observed the same using 3D analysis (Figure S5.1, right panel). We also observed in our previous study that compounds which share the same biological target influenced the phenotypic development of 3D cultured 4T1 cells in a similar pattern. Based on this observation we successfully classified the tested compounds according to their biological activity. Here, we evaluated the classification accuracy for 2D projection analysis and 3D analysis respectively. To obtain classification train-

Compound	0.03μM	0.1μM	0.316μM	1μM	3.16μM	10μM
Arq 197	-	-	2D	2D/3D	2D/3D	2D/3D
Arq 197(blind test)	-	-	2D/3D	2D/3D	2D/3D	2D/3D
AZD 0530	-	2D	2D/3D	2D/3D	2D/3D	2D/3D
AZD 0530 (blind test)	2D	-	2D/3D	-	-	2D/3D
dasatinib	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D
dasatinib (blind test)	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D
entinostat	2D	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D
entinostat (blind test)	-	-	2D	2D/3D	2D/3D	2D/3D
imatinib mesylate	-	-	2D/3D	2D	2D/3D	2D/3D
imatinib mesylate (blind test)	-	-	3D	2D	2D/3D	2D/3D
nilotinib	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D
nilotinib (blind test)	2D/3D	2D	2D/3D	2D/3D	2D/3D	2D/3D
sorafenib tosylate	2D	2D	2D/3D	2D/3D	2D/3D	2D/3D
sorafenib tosylate (blind test)	3D	2D/3D	2D/3D	2D/3D	2D/3D	2D/3D
stf-62247	-	-	-	2D/3D	2D/3D	2D/3D
stf-62247 (blind test)	-	2D	2D	2D/3D	2D/3D	2D/3D
sunitinib maleate			2D	2D/3D	2D/3D	2D/3D
sunitinib maleate (blind test)	2D	2D	-	2D/3D	2D/3D	2D/3D
vandetanib	-	-	-	2D/3D	2D/3D	2D/3D
vandetanib (blind test)	-	2D	-	2D/3D	2D/3D	2D/3D
erlotinib HCl	-	-	-	2D	2D/3D	2D/3D
genistein	-	-	2D	3D	2D/3D	2D/3D
genistein (blind test)	-	-	-	-	-	2D/3D

Table 5.2: List of compounds and corresponding concentrations identified as hits in either 2D projection analysis or 3D analysis, using Mahalanobis distance ($\alpha = 0.05$). "2D" indicates that the condition (per compound per concentration) was identified as a hit in our 2D projection based analysis. "3D" indicates that the condition was identified as a hit in our 3D analysis. "2D/3D" indicates that the condition was identified as a hit in both 2D projection analysis and 3D analysis. "-" indicates not a hit in neither of the analysis.

ing data, we firstly collected the robust $z - scores$ of all measured parameters from the conditions that were identified as hits by the corresponding analysis platform (Table 5.2). Four classes of compounds were defined according to the biological target (Figure 5.2a and 5.2b). The class of c-MET inhibitor (Arq 197) and the class of autophagy inducer (stf-62247) contained less than 15 data points respectively, thus they were not included in the classification, in order to avoid the curse of dimensionality. As the result, 46 conditions were selected from 2D projection analysis and 39 conditions were selected from 3D analysis for the classification training process, respectively. The blind test experiment was not included in the classification training process but was used to test the obtained classifiers later. For each platform, we firstly performed forward feature selection with the criterion of Mahalanobis distance and then evaluated several classification algorithms using 10-fold cross-validation (see

Supplementary note for pseudo code). The tested classification methods include k-nearest neighbor classification, linear Bayes normal classification [5,21,22], quadratic Bayes normal classification [21,22], nearest mean classification, support vector machine classification (SVC) [23,24], and Fisher linear classification [21,22,25]. For each subsample test of 10-fold cross-validation, a classification error rate (%) was defined as:

$$\sum \frac{number\ of\ erroneously\ classified\ objects\ per\ class}{size\ of\ class} \times prior\ probability\ of\ class$$

$$(5.11)$$

where the prior probability is equal for each class (25%). For the 2D projection analysis, the best performance of classification was obtained when we used quadratic Bayes normal classification with 6 selected parameters (Figure 5.2c), and the corresponding error rate was 20.2% (Table 5.3). For 3D analysis platform, the error rate was 12.8% when we applied linear Bayes normal classification with 14 parameters selected (Figure 5.2d; Table 5.3). This indicates that the error rate for 3D analysis is lower than for 2D projection analysis.

In the previous section, we noticed that the 2D projection analysis platform has a higher sensitivity of hits identification compared to the 3D analysis platform. The consequence was that relatively low concentrations were selected as hits and thus included for classification, which might affect the classification accuracy. To validate whether the relatively lower classification accuracy of 2D projection analysis platform was caused entirely by including those low concentrations for the training process, we repeated the classification for two platforms, including only the conditions (38 conditions) that were identified as hits in both 2D projection analysis and 3D analysis (Table 5.2; Figure S5.2a). As the result, the lowest classification error rate of 2D projection analysis was decreased from 20.2% to 15.8% using quadratic Bayes normal classification (Table 5.3; Figure S5.2b), however, this is still higher than the lowest classification error rate of 3D analysis which was 12.4% (Table 5.3; Figure S5.2c), indicating that the 3D analysis platform preserved more characteristic features of phenotype and those features played an important role in phenotype classification.

For both analysis platforms, we investigated the parameters that contributed most to the classification. In 2D projection analysis, the classification result (Figure 5.2c) showed that the best performance was obtained when quadratic Bayes normal classification was used and 6 parameters were selected. Table S5.7 shows a list of 6 parameters that were most frequently selected as discriminative parameters using the feature selection algorithm in 10-fold cross-validation. The selected parameters included morphological parameters and intensity attributes measured from total population and from subpopulations. For the 3D analysis platform, the most frequently selected 14 discriminative parameters are shown in Table S5.8. Similar to 2D projection analysis, we noticed that not only morphological parameters, but also intensity parameters and subpopulation information played an important role for phenotype classification.

To further establish the value of analyzing different categories of parameters, we firstly repeated feature selection and classification excluding subpopulation parame-

a Classification training data set obtained from the 2D projection analysis platform

Class index	Target	Compounds	Number of data points
1	BCR-ABL	AZD 0530, dasatinib, nilotinib, imatinib mesylate	14 24 24 16
2	VEGFR	sunitinib maleate, vandetanib, sorafenib tosylate	15 11 24
3	EGFR	erlotinib HCL, genistein	11 12
4	HDAC	entinostat	24

b Classification training data set obtained from the 3D analysis platform

Class index	Target	Compounds	Number of data points
1	BCR-ABL	AZD 0530, dasatinib, nilotinib, imatinib mesylate	11 24 24 12
2	VEGFR	sunitinib maleate, vandetanib, sorafenib tosylate	11 12 16
3	EGFR	erlotinib HCL, genistein	8 12
4	HDAC	entinostat	19

c

d

Figure 5.2: Comparison of classification accuracy between our 2D projection analysis platform and our 3D analysis platform. (a-b) Four classes of compounds were defined according to their biological activity. Two training data sets for classification were collected respectively from our 2D projection analysis (a) and 3D analysis (b). These training data sets included the conditions which were identified as biologically active by the corresponding platforms. The corresponding compounds and number of data points are also shown in the tables. (c-d) Classification result using multiple classification methods for (c) 2D projection based analysis, and (d) 3D analysis. Feature selection with search algorithm "forward" and criterion "Mahalanobis distance" was applied to detect the optimal number of features. For each classification method and each number of selected features, 10-fold cross-validation was carried out resulting in 10 error rates. The average error rates are shown in the chart with standard deviation as error bar. SVC means support vector machine classification.

ters, resulting in only 460 $z - scores$ of 2D projection analysis and 191 $z - scores$ of 3D analysis that were used as classification training data. The classification results are shown in Figure 5.3a, 5.3b and Table 5.3. For both 2D projection analysis and 3D analysis, decreased classification accuracy was obtained when subpopulation parameters were not taken into account. In 2D projection analysis, quadratic Bayes normal classification still performed relatively better than other classification methods when 5 parameters were selected, but the error rate increased to 24.6% compared to the

error rate of 20.2% when subpopulation parameters were used. In 3D analysis, linear Bayes normal classification performed the best with an error rate of 14.4% when 10 parameters were selected for classification. This is higher than 12.8% which is the lowest error rate of linear Bayes normal classification when subpopulation parameters were included for classification. Next, we repeated the analysis without moments, wavelet and intensity parameters. Only morphological parameters were measured from the whole object population and the two subpopulations, resulting in 224 $z-scores$ of 2D projection analysis and 210 $z-scores$ of 3D analysis that were measured from each well and used for classification. The classification results (Table 5.3; Figure 5.3c, 5.3d) showed that omission of moments and intensity parameters increased the classification error from 20.2% to 21.8% for 2D projection analysis, from 12.8% to 14.4% for 3D analysis.

Previously, we have analyzed the phenotypic response of 3D cultured 4T1 cells to the same classes of compounds (Table S3.8), using a 2D projection based profiling system which extracted 598 parameters from each well. Forward feature selection with the criterion of Mahalanobis distance was performed, and several classification methods were evaluated by 10-fold cross-validations. The best performance of classification was obtained using quadratic Bayes normal classification with 8 selected parameters and the corresponding error rate was 15.1%. In this chapter, we extended this 2D projection analysis platform so that it can measure 794 parameters for each well. However, the classification error increased to 20.2% using the same feature selection and classification method (Table 5.3). We presume that this is due to the increased sensitivity of hits identification as more parameters were used for PCA. It led to the identification as hits of more relatively low concentrations (entinostat 0.03 μM; genistein 0.316μM; imatinib mesylate 0.316μM; sunitinib maleate 0.316μM) which were subsequently included in the classification training process. At these concentrations, the differences in phenotypes between different compound classes are not yet very significant and consequentially, the classification accuracy might be affected. To confirm this hypothesis, we removed these four concentrations from the data set and repeated the feature selection and classification based on the phenotype profiles of 794 $z-scores$. As expected, the smallest classification error reduced significantly to 14.3% (Table 5.3) when we applied quadratic Bayes normal classification. This error rate was also slightly lower than the error rate of 15.1% which was obtained using the 598 parameters based platform and using the same concentrations for classification training process, indicating that the additional parameters in fact contributed to the classification of compounds.

Finally, we used the phenotype profiling from the blind test experiment to evaluate the classifier built on the training data (Figure 5.2a and 5.2b). To obtain test data, we firstly performed hit identification and selected the conditions which were identified as hits in both 2D and 3D analysis. Next, for each $z-score$ the average over replicates was calculated and used as test data. In 10-fold cross-validation, quadratic Bayes normal classification with 6 selected parameters showed the highest classification accuracy for 2D projection analysis, while linear Bayes normal classification with 14 selected parameters performed relatively better than other classification methods for

The profiling system used to generate training data for classification	Number of conditions (hits) used for classification training process	The classification method(s) with the smallest error rate	
3D with 290 parameters	38 conditions (2D projection with 794 parameters and 3D with 290 parameters)	Linear Bayes normal classification	12.4%
3D with 290 parameters	39 conditions (3D with 290 parameters)	Linear Bayes normal classification	12.8%
2D projection with 794 parameters	42 conditions (2D projection with 598 parameters)	Quadratic Bayes normal classification	14.3%
3D with 210 parameters excluding the intensity, moments and texture parameters	39 conditions (3D with 290 parameters)	Linear Bayes normal classification	14.4%
3D with 191 parameters excluding the subpopulation parameters	39 conditions (3D with 290 parameters)	Linear Bayes normal classification	14.4%
2D projection with 598 parameters	42 conditions (2D projection with 598 parameters)	Quadratic Bayes normal classification	15.1%
2D projection with 794 parameters	38 conditions (2D projection with 794 parameters and 3D with 290 parameters)	Quadratic Bayes normal classification 1-nearest neighbor classification	15.8%; 15.7%
2D projection with 794 parameters	46 conditions (2D projection with 794 parameters)	Quadratic Bayes normal classification	20.2%
2D projection with 224 parameters excluding the intensity, moments and wavelet parameters	46 conditions (2D projection with 794 parameters)	Quadratic Bayes normal classification	21.8%
2D projection with 460 parameters excluding the subpopulation parameters	46 conditions (2D projection with 794 parameters)	Quadratic Bayes normal classification 3-nearest neighbor classification	24.6%; 23.7%

Table 5.3: Comparison of the smallest classification error rates obtained using different conditions for training and different platforms for phenotype profiling. The profiling system used to identified the active conditions are given in brackets.
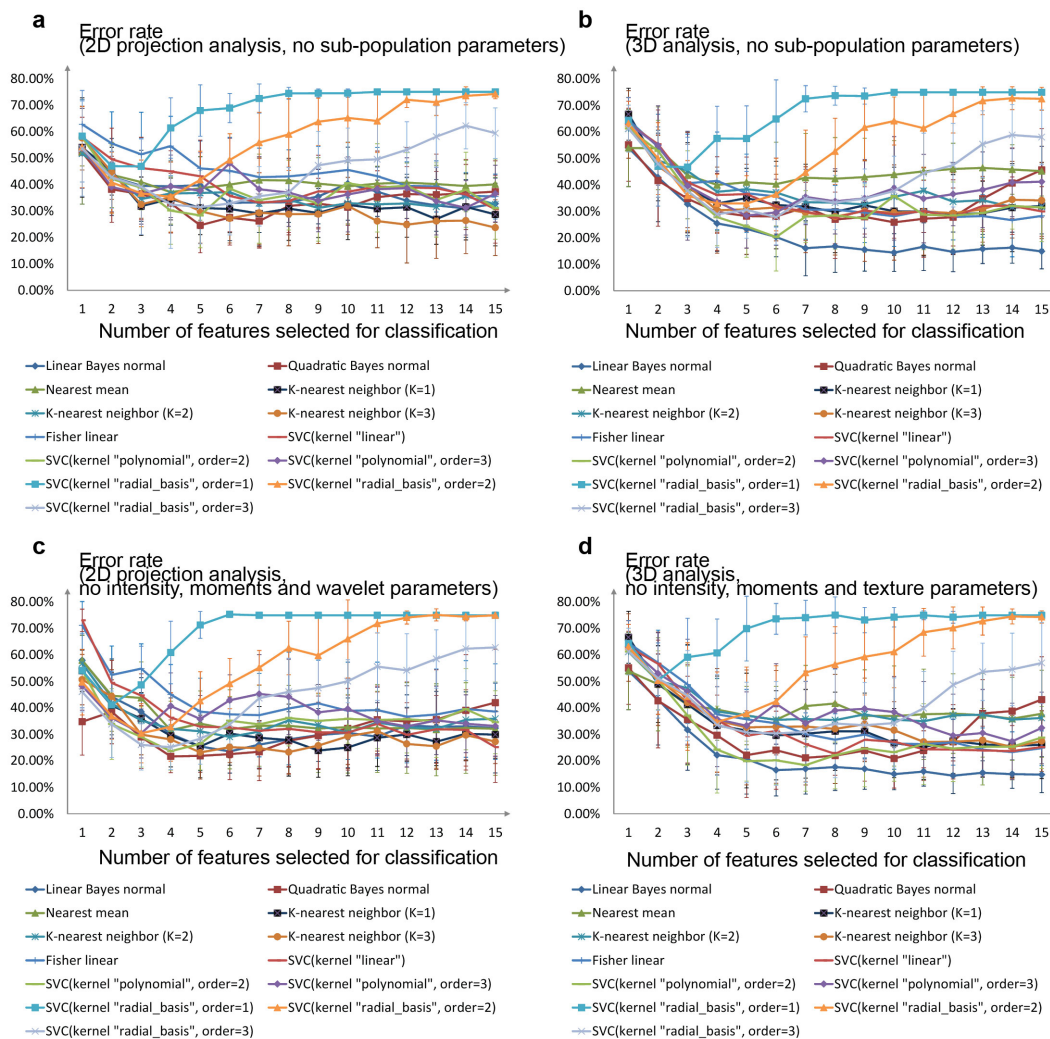
Figure 5.3: The contribution of different parameters to the phenotype classification for both 2D projection analysis platform and 3D analysis platform. (a-d) Classification result using multiple classification methods for (a) 2D projection based analysis without subpopulation information, (b) 3D analysis without subpopulation information, (c) 2D projection based analysis without intensity parameters, moments and wavelet parameters, (d) 3D analysis without intensity parameters, moments and texture parameters. The classes were defined according to the Figure 5.2a, 5.2b. Feature selection and classification algorithms are the same as in the Figure 5.2.

3D analysis, respectively. According to this result, we trained a quadratic Bayes normal classifier for 2D projection analysis based on the parameters shown in Table S5.7. For 3D analysis, a linear Bayes normal classifier was trained based on the parameters shown in Table S5.8. We tested these 2D and 3D classifiers with the obtained test data set, and the results are shown in Table 5.4 and Table 5.5. The 3D classification accuracy of BCR-ABL inhibitors is decreased compared to the 2D results, but the classification accuracy of VEGFR is improved. The overall error rate

True Class	Output of classifier			
	BCR-ABL	VEGFR	EGFR	HDAC
BCR-ABL	13	2	0	0
VEGFR	3	7	0	1
EGFR	0	0	3	0
HDAC	0	0	0	1

Table 5.4: **Classification result for 2D analysis platform using 6 parameters.** The quadratic Bayes normal classifier achieved an overall error rate of 20.0%. The prior probability of test data set was defined as class frequencies.

True Class	Output of classifier			
	BCR-ABL	VEGFR	EGFR	HDAC
BCR-ABL	11	4	0	0
VEGFR	0	11	0	0
EGFR	0	0	3	0
HDAC	0	0	0	1

Table 5.5: **Classification result for 3D analysis platform using 14 parameters.** The linear Bayes normal classifier achieved an overall error rate of 13.3%. The prior probability of test data set was defined as class frequencies.

of the 3D classification was slightly lower than of the 2D classification.

5.3.3 Reproducibility of concentration dependent phenotypic trajectories

For the 2D projection profiling platform, we were able to model the concentration dependent trajectory in multidimensional principle component space for each compound, using 2nd order polynomial regression models (**chapter 3**). Here we applied the same method to model a trajectory of each compound for both 2D projection based analysis and 3D analysis, based on the first two PCs. Next, we measured for each compound in the blind test experiment the difference between its trajectories and the trajectory of every compound in the first experiment, based on the coefficient of determination R^2 (see Supplementary note of **chapter 3** for details). In theory, good reproducibility means that the same compound in two different experiments should have a closer distance than the distance between different compounds. Therefore, we predicted the name of each compound in the blind test experiment as the compound based on its closest neighbor in the first experiment. In 2D analysis, only 5 out of 11 compounds were correctly predicted, while 9 out 11 compounds were correctly predicted in 3D analysis.

5.4   Discussion and conclusions

In this study, we firstly extended our previous 2D projection based uHCA platform to measure 794 parameters for the phenotype profiling of 3D cultured micro-tissues. As the number of parameters increased from 598 to 794, an enhanced sensitivity of hits identification was obtained, however, this compromised the classification accuracy of compounds with different biological targets. We found that this is due to the low compound concentrations which were now additionally identified as hits. The corresponding phenotypes however, were not yet representative. When we removed these lower concentrations, the classification accuracy increased significantly and even surpassed the best classification performance obtained using the 598 parameters based platform, indicating that the additional parameters contributed to the separation and characterization of the phenotypes.

Next, we developed a new high-content analysis platform for the phenotype profiling of 3D cultured micro-tissues, based on image stacks and reconstructed 3D structures. Novel segmentation algorithms and a reconstruction method were applied, and 290 parameters were automatically measured from each well. Although it managed to extract phenotype properties in numerous aspects, including morphology, topology, intensity, moments, texture, as well as subpopulations, compared to 794 parameters that were measured by the 2D profiling platform, much less information was retrieved from the 3D phenotype profiles. This was reflected in the sensitivity of hits identification. We applied both analysis platforms for the identification of biologically active concentrations of 12 compounds that significantly influenced the phenotypic development of the 4T1 cells. More concentrations were identified as hits using our 2D projection based platform, compared to the active concentrations detected by the 3D analysis platform.

However, the 3D analysis platform showed some clear advantages in phenotype characterization. We compared the classification accuracy between our 2D projection analysis and our 3D analysis based on the same list of concentrations for training and test. Moreover, we evaluated the reproducibility of phenotype for each platform, by comparing the trajectory models of compounds in the blind experiment with the trajectories of compounds in the first experiment. Both tests showed that the 3D analysis provided better performance, indicating that to systematically study phenotypes associated with modulation of different cellular pathways, it is necessary to extract information of phenotype in all 3 dimensions with great details.

However, 3D phenotype profiling costs much more computational time and memory than 2D projection profiling. For our 2D projection profiling, analyzing 384 wells on a Windows 7 64-bits operation system with an Intel core i7-2600 processor and 16 GB of RAM took 280 minutes. To analyze the same wells with our 3D profiling platform, 1150 minutes were required using the same hardware. One of the most computational expensive procedures is deconvolution and 3D watershed masked clustering, each of which took approximately 1.2 minutes to process one well. Considering both computational expenses and hits identification ability of the 3D analysis platform, we suggest that our 2D projection based platform might already be sufficient

for primary screens where a large amount of compounds or even whole libraries of RNAi's need to be tested for activity identification. For screens which aim to systematically study the effect of biologically active compounds on cellular phenotypes, the 3D analysis platform is more recommended. Importantly, relative high concentrations should be tested in this type of screen, for example concentrations at the IC50s and higher, to achieve the maximum separation of phenotypes between different classes.

At this stage we compared our 2D projection and 3D phenotype profiling platform so far in only one type of cell screen, with only a limited set of biologically active compounds. We will further investigate the advantages and disadvantages for both 2D and 3D phenotype classification in other cancer cell types. This will further streamline strategies for candidate drug discovery.

References

[1] Petersen, O. W., Ronnov-Jessen, L., Howlett, A. R. & Bissell, M. J. Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. Proc Natl Acad Sci U S A 89, 9064-9068 (1992).

[2] Wang, F. et al. Reciprocal interactions between beta1-integrin and epidermal growth factor receptor in three-dimensional basement membrane breast cultures: a different perspective in epithelial biology. Proc Natl Acad Sci U S A 95, 14821-14826 (1998).

[3] LaBarbera, D. V., Reid, B. G. & Yoo, B. H. The multicellular tumor spheroid model for high-throughput cancer drug discovery. Expert Opin Drug Discov 7, 819-830 (2012).

[4] Krausz, E. et al. Translation of a tumor microenvironment mimicking 3D tumor growth co-culture assay platform to high-content screening. J Biomol Screen 18, 54-66 (2013).

[5] Liu, C. & Wechsler, H. Robust coding schemes for indexing and retrieval from large face databases. IEEE Trans Image Process 9, 132-137 (2000).

[6] Kuan Yan & Verbeek, F. J. Segmentation for high-throughput image analysis: watershed masked clustering. In: SoLA'12 Proceedings of the 5th international conference on Leveraging Applications of Formal Methods, Verification and Validation: applications and case studies. (eds T. Margaria & B. Steffen) 25-41 (Springer-Verlag).

[7] Niblack, W. An introduction to digital image processing. 1-215 (Prentice-Hall International, 1986).

[8] Cline, H. E., Lorensen, W. E., Ludke, S., Crawford, C. R. & Teeter, B. C. Two algorithms for the three-dimensional reconstruction of tomograms. Med Phys 15, 320-327 (1988).

[9] Lee, T. C., Kashyap, R. L. & Chu, C. N. Building Skeleton Models Via 3-D Medial Surface Axis Thinning Algorithms. Cvgip-Graph Model Im 56, 462-478 (1994).

[10] Gonzalez, R. C. & Woods, R. E. Digital image processing. 3rd edn, 1-954 (Prentice Hall, 2008).

[11] Novotni, M. & Klein, R. Shape retrieval using 3D Zernike descriptors. Comput Aided Design 36, 1047-1062 (2004).

[12] Haralick, R. M., Shanmuga.K & Dinstein, I. Textural Features for Image Classification. IEEE T Syst Man Cyb Smc3, 610-621 (1973).

[13] Haralick, R. M. Statistical and Structural Approaches to Texture. P IEEE 67, 786-804 (1979).

[14] Chen, X. & Murphy, R. F. Robust classification of subcellular location patterns in high resolution 3D fluorescence microscope images. 2007/02/03 edn, Vol. 3 (2004).

[15] Wang, J. et al. Cellular phenotype recognition for high-content RNA interference genome-wide screening. J Biomol Screen 13, 29-39 (2008).

[16] Birmingham, A. et al. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods 6, 569-575 (2009).

[17] McLachlan, G. J. Discriminant analysis and statistical pattern recognition. Vol. 235 52-54 (John Wiley & Son, 1992).

[18] Mahalanobis, P. C. On the generalised distance in statistics. Proceedings of the National Institute of Sciences of India 2, 49-55 (1936).

[19] Greenwood, P. E. & Nikulin, M. S. A guide to chi-squared testing. Vol. 280 1-280 (John Wiley & Son, 1996).

[20] Mardia, K. V., Kent, J. T. & Bibby, J. M. Multivariate analysis. 1-521 (Academic Press, 1979).

[21] Webb, A. R. Statistical Pattern Recognition (2nd Edition) 123-163 (John Wiley & Sons, 2002).

[22] Duda, R. O., Hart, P. E. & Stork, D. G. Pattern Classification (2nd Edition) 215-268 (John Wiley and Sons, 2001).

[23] Theodoridis, S. & Koutroumbas, K. Pattern recognition. 3rd edn, 93-118 (Academic Press, 2006).

[24] Cortes, C. & Vapnik, V. Support-Vector Networks. Mach Learn 20, 273-297 (1995).

[25] Raudys, S. & Duin, R. P. W. Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix. Pattern Recogn Lett 19, 385-392 (1998).

Supplementary Note

1. Additional parameters measured for 2D projected analysis

Previously, we developed an automated image analysis platform to automatically profile micro-tissue phenotypes on 2D projected images with 598 parameters (Table S3.2-S3.5). Here, we extended this platform so that it can measure more parameters.

- (*M30-M36*)**More classic morphological parameters** in addition to the classic morphological parameters represented in the **Chapter 3** (*M1-M29*)

    - Rectangular bounding box of an object: For each binary object (either nucleus or cell cluster), we measured the bounding rectangle which is the smallest rectangle enclosing the object. Relative properties were measured for each bounding box including area (*M30*) and $\frac{\text{Area of object}}{\text{Area of bounding box}}$ (*M31*).
    - Axis angle of an object (*M32*): Axis angle measures the angle between the major axis and a line parallel to the x-axis of the image.
    - Feret's angle (*M33*): Feret angle measures the angle between the Feret's diameter and a line parallel to the x-axis of the image.
    - Circularity of an object (*M34*): Circularity of each segmented object (either nucleus or cell cluster) is defined as $4\pi \times \frac{\text{Area}}{\text{Perimeter}^2}$.
    - Eccentricity of object (*M35*): Eccentricity of each segmented object (either nucleus or cell cluster) is defined as $\sqrt{1 - (\frac{\text{Minor axis}}{\text{Major axis}})^2}$.

- Accumulated intensity (*I141*) calculated for each binary object masked region on the projected grey value images, in addition to the intensity-based parameters represented in the **Chapter 3** (*I1-I140*)

- (*O1-O5*)The correlation between two fluorescence channels.

    - For each cell cluster, we calculated the number of nuclei it contains (*O1*).
    - For each cell cluster masked region, we calculated the Pearson correlation coefficients of the pixel intensity between the projected image from the Hoechst stained nuclei channel and the corresponding projected image from the rhodamine stained F-actin channel (*O2*).
    - For each cell cluster masked region ,we calculated the slop of the linear correlation of the pixel intensity between the projected image from the Hoechst stained nuclei channel and the corresponding projected image from the rhodamine stained F-actin channel (*O3*).
    - The Pearson correlation coefficients of intensity of all pixels between the projected images of two channels (*O4*).
    - The slop of the linear correlation of intensity of all pixels intensity between the projected images of two channels (*O5*).

2. Pseudo code for 10-fold cross-validation

```
Define a dataset A containing training data of each well and its class label: parameters(robust z-scores) are in column and wells are in row
Randomly assigned each well to one of 10 subsamples so that the size of each subsample is equal
for int i = 1 to 10 subsamples
            trainingDataset  = all samples from A except i'th subsample;
            testDataset  = i'th subsample of A;
            ParameterList[i] = ForwardFeatureSelection(trainingDataset)  //using the criterion Mahalanobis distance
                                                                         //to selects 15 most discriminative parameters;

        for int d = 1 to 15 parameters
                    training = extract the parameters (the first d parameters in ParameterList[i]) from trainingDataset;
                    test = extract the parameters (the first d parameters in ParameterList[i]) from testDataset;
                    classifier1 =  linearBayesClassifier(training);
                    classifier2 =  quadraticBayesClassifier(training);
                    classifier3 =  nearestMeanClassifier(training);
                    classifier4 =  k-nearestNeighborClassifier(training, k=1);
                    classifier5 =  k-nearestNeighborClassifier(training, k=2);
                    classifier6 =  k-nearestNeighborClassifier(training, k=3);
                    classifier7 =  fisherLinearClassifier(training);
                    classifier8 =  SVC(training, kernel = 'linear');
                    classifier9 =  SVC(training, kernel = 'polynomial', order = 2);
                    classifier10 =  SVC(training, kernel = 'polynomial', order = 3);
                    classifier11 =  SVC(training, kernel = 'radial basis', order = 1);
                    classifier12 =  SVC(training, kernel = 'radial basis', order = 2);
                    classifier13  =  SVC(training, kernel = 'radial basis', order = 3);
                    error1 = testClassifier(classifier1, test);
                    error2 = testClassifier(classifier2, test);
                    error3 = testClassifier(classifier3, test);
                    error4 = testClassifier(classifier4, test);
                    error5 = testClassifier(classifier5, test);
                    error6 = testClassifier(classifier6, test);
                    error7 = testClassifier(classifier7, test);
                    error8 = testClassifier(classifier8, test);
                    error9 = testClassifier(classifier9, test);
                    error10 = testClassifier(classifier10, test);
                    error11 = testClassifier(classifier11, test);
                    error12 = testClassifier(classifier12, test);
                    error13 = testClassifier(classifier13, test);
        end for int d = 1 to 15 parameters
for int i = 1 to 10 subsamples

// error is defined as equation 11.
```

Supplementary figures
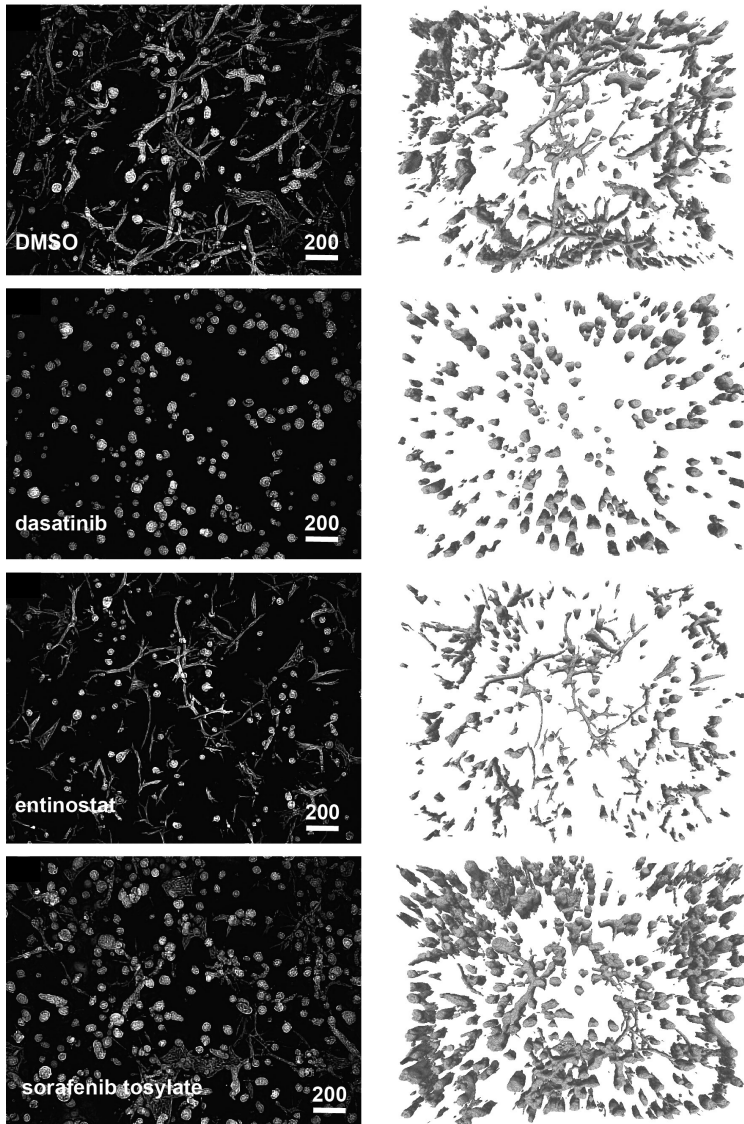
Figure S5.1: Phenotypes of 4T1 3D micro-tissues after exposure to compounds with different biological targets. Composite, projected 2D image slices of micro-tissues after being exposed to three compounds with different biological targets and DMSO control. The image stacks were obtained from the rhodamine stained F-actin channel (left panel) and corresponding reconstructed 3D cell cluster structures retrieved from the same image stacks (right panel). Concentration of all compounds shown here was 0.316 μM. Compared to the phenotype of DMSO control, dasatinib inhibited branching but not proliferation so that bigger cell clusters were formed, while entinostat induced much thinner branches. Sorafenib tosylate caused formation of much shorter branches, indicative of inhibition of invasion. Scale bar represents 200μm.

a

Classification training data set obtained from the conditions
which were identified as hits in both 2D projection analysis
and 3D analysis

Class index	Target	Compounds	Number of data points
1	BCR-ABL	AZD 0530, dasatinib, nilotinib, imatinib mesylate	11 24 24 12
2	VEGFR	sunitinib maleate, vandetanib, sorafenib tosylate	11 11 16
3	EGFR	erlotinib HCL, genistein	8 8
4	HDAC	entinostat	19

b

c

Figure S5.2: Comparison of classification accuracy between our 2D projection analysis platform and 3D analysis platform. (a) Four classes of compounds were defined according to their biological activity. Only concentrations that were identified as an active concentration in both analysis platforms were included for classification training. The corresponding compounds and number of data points are also shown in the tables. (b-c) Classification result using multiple classification methods for (b) 2D projection based analysis, (c) 3D analysis. Feature selection with search algorithm "forward" and criterion "Mahalanobis distance" was applied to detect the optimal number of features. For each classification method and each number of selected features, 10-fold cross-validation was carried out resulting in 10 error rates. The average error rates are shown in the chart with standard deviation as error bar. SVC means support vector machine classification.

Supplementary tables

Parameter index	Parameter Description
From 1-6	The mean of *M30-M35* of all cell clusters in one image
From 7-12	The standard deviation of *M30-M35* of all cell clusters in one image
From 13-18	The sum of *M14-M19* of cell clusters
From 19-32	The mean of *M1, M2, M5-M8, M12, M13 and M30-M35* of all nuclei in one image
From 33-46	The standard deviation of *M1, M2, M5-M8, M12, M13 and M30-M35* of all nuclei in one image
From 47-95	The mean of *S1-S49* of all nuclei in one image
From 96-144	The standard deviation of *S1-S49* of all nuclei in one image
From 145-149	The mean of *I137-I141* of all cell clusters in one image
From 150-154	The standard deviation of *I137-I141* of all cell clusters in one image
From 155-159	The mean of *I137-I141* of all nuclei in one image
From 160-164	The standard deviation of *I137-I141* of all nuclei in one image
From 165-170	The mean of *M30-M35* of cell clusters which are classified to spherical object
From 171-176	The standard deviation of *M30-M35* of cell clusters which are classified to spherical object
From 177-182	The mean of *M30-M35* of cell clusters which are classified to branched object
From 183-188	The standard deviation of *M30-M35* of cell clusters which are classified to branched object

Table S5.1: Profiling based on the 2D projection: List of parameters, in additional to the parameters described in Table S3.2-S3.5. The definition of parameters *M1-M29*, *S1-S49* and *I1-I140* is in the **Chapter 3**-Methods section.

Parameter Index	Parameter Description
From 1-3	The mean of *O1-O3* of all cell clusters in one image
From 4-6	The standard deviation of *O1-O3* of all cell clusters in one image
7	*O4*
8	*O5*

Table S5.2: Profiling based on the 2D projection: List of parameters measured on the correlation between two fluorescence channels. The definition of parameters *O1-O5* is in Supplementary note.

Parameter Index	Parameter Description
1	The number of cell clusters
2	The total number of pixels assigned to the segmented cell clusters regions
From 3-4	The total volume and total surface of cell clusters
5	Sphericity (Table 4.2)
6	The total surface to the total volume ratio (SAV) (Table4.2)
7	The thickness of the 3D bounding rectangular box for the micro-tissue network
From 8-9	The average and standard deviation of the number pixels assigned to each cell cluster
From 10-19	The average and standard deviation of the volume, surface, sphericity, solidity and SAV of cell clusters (Similar to the measurement of nuclei in Table 4.1)
From 20-25	The average and standard deviation of the width, height and thickness of 3D rectangular bounding boxes of cell clusters (Similar to the measurement of nuclei in Table 4.1)
From 26-29	The average and standard deviation of the volume and surface of the convex hull of cell clusters: Convex hull was calculated using the QuickHull algorithm as described in Table 4.1
From 30-35	The average and standard deviation of eigenvalue λ_1 (the shortest), λ_2 (the median), λ_3 (the longest), of cell clusters. (Similar to the measurement of nuclei in Table 4.1)
From 36-41	The average and standard deviation of principle axis I_1, I_2, I_3 of cell clusters. (Similar to the measurement of nuclei in Table 4.1)
From 42-44	Eigenvalues (the shortest, the median and the longest) measured when considering all binary cellular regions as one object (Table 4.2).
From 45-47	Principle axises measured when considering all binary cellular regions as one object (Table 4.2).
From 48-53	The sum of #end point, #junction point, #slab point, #triple point, #quadruple point and #single branch of cell clusters (see **Chapter 4** for details).
54	The sum of the length of all single branches
From 55-66	The average and standard deviation of #end point, #junction point, #slab point, #triple point, #quadruple point and #single branch of cell clusters.
67	The average branch length of cell clusters
From 68-69	The average and standard deviation of the accumulated branch length of cell clusters
From 70-72	The maximum average branch length, maximum branch length and maximum accumulated branch length of cell clusters
73	The number of nuclei
74	The total number of pixels assigned to the segmented nuclear regions
From 75-76	The total volume and total surface of nuclei
From 77-110	Parameter index from 8-41 but measured for nuclei (see Table 4.1 for details).
111	Average density of nuclei: average peer-to-peer distance between nuclei. (see Table 4.1 for details).

Table S5.3: Profiling based on the 3D analysis: List of morphological parameters and topological parameters calculated on the reconstructed cell clusters and nuclei for the whole population. "#" means "number of"

Parameter Index	Parameter Description
1	The average intensity of deconvolved image stack from the rhodamine channel
2	The intensity standard deviation of deconvolved image stack from the rhodamine channel
3	The maximum intensity of deconvolved image stack from the rhodamine channel
4	The average intensity of deconvolved image stack from the Hoechst channel
5	The intensity standard deviation of deconvolved image stack from the Hoechst channel
6	The maximum intensity of deconvolved image stack from the Hoechst channel
7	The average intensity gradient magnitude of deconvolved image stack from the rhodamine channel
8	The standard deviation of intensity gradient magnitude of deconvolved image stack from the rhodamine channel
9	The maximum intensity gradient magnitude of deconvolved image stack from the rhodamine channel
10	The average intensity gradient magnitude of deconvolved image stack from the Hoechst channel
11	The standard deviation of intensity gradient magnitude of deconvolved image stack from the Hoechst channel
12	The maximum intensity gradient magnitude of deconvolved image stack from the Hoechst channel

Table S5.4: Profiling based on the 3D analysis: List of intensity parameters calculated on the deconvolved image stack for two fluorescence channels.

Parameter Index	Parameter Description
From 1-16	8 Zernike moments measured for the rhodamine channel and Hoechst channel
17	Haralick parameter measured for the rhodamine channel: Average of angular second moment
18	Haralick parameter measured for the rhodamine channel: Range of angular second moment
19	Haralick parameter measured for the rhodamine channel: Average of contrast
20	Haralick parameter measured for the rhodamine channel: Range of contrast
21	Haralick parameter measured for the rhodamine channel: Average of correlation
22	Haralick parameter measured for the rhodamine channel: Range of correlation
23	Haralick parameter measured for the rhodamine channel: Average of sum of square of variance
24	Haralick parameter measured for the rhodamine channel: Range of sum of square of variance
25	Haralick parameter measured for the rhodamine channel: Average of inverse difference moment
26	Haralick parameter measured for the rhodamine channel: Range of inverse difference moment
27	Haralick parameter measured for the rhodamine channel: Average of sum average
28	Haralick parameter measured for the rhodamine channel: Range of sum average
29	Haralick parameter measured for the rhodamine channel: Average of sum variance
30	Haralick parameter measured for the rhodamine channel: Range of sum variance
31	Haralick parameter measured for the rhodamine channel: Average of sum entropy
32	Haralick parameter measured for the rhodamine channel: Range of sum entropy
33	Haralick parameter measured for the rhodamine channel: Average of entropy
34	Haralick parameter measured for the rhodamine channel: Range of entropy
35	Haralick parameter measured for the rhodamine channel: Average of difference variance
36	Haralick parameter measured for the rhodamine channel: Range of difference variance
37	Haralick parameter measured for the rhodamine channel: Average of difference entropy
38	Haralick parameter measured for the rhodamine channel: Range of difference entropy
39	Haralick parameter measured for the rhodamine channel: Average of info measure of correlation 1
40	Haralick parameter measured for the rhodamine channel: Range of info measure of correlation 1
41	Haralick parameter measured for the rhodamine channel: Average of info measure of correlation 2
42	Haralick parameter measured for the rhodamine channel: Range of info measure of correlation 2
From 43-68	Parameter index from 17-42 but for the Hoechst channel

Table S5.5: Profiling based on the 3D analysis: 3D Zernike moments and Haralick texture parameters calculated for two fluorescence channels.

141

Parameter index	Parameter Description
1	The percentage of cell clusters classified as spherical objects
From 2-35	The responding morphological parameters(Table S5.3 parameter index from 8 to 41) calculated from the cell clusters which are classified as spherical objects
From 36-69	The responding morphological parameters(Table S5.3 parameter index from 8 to 41) calculated from the cell clusters which are classified as branched object
From 70-84	The responding topological parameters(Table S5.3 parameter index from 55 to 69) calculated from the cell clusters which are classified as spherical object
From 85-99	The responding topological parameters(Table S5.3 parameter index from 55 to 69) calculated from the cell clusters which are classified as branched object

Table S5.6: Profiling based on the 3D analysis: List of parameters calculated for the subpopulations of cell clusters.

Parameter index	Parameter Description
1	The average minor axis of cell clusters that are classified as spherical objects
2	The standard deviation of the intensity standard deviation of cell clusters that are classified as spherical objects
3	The average solidity of cell clusters that are classified as branched objects
4	The average of cell clusters' minimum intensity
5	The sum of number of quadruple point of cell clusters
6	The average Zernike moment (order 34) of cell clusters

Table S5.7: A list of 2D projection profiling parameters that were most frequently selected as discriminative parameters for the compounds classification, using feature selection algorithm in each round of 10-fold cross-validation.

Parameter index	Parameter Description
1	The standard deviation of convex hull's volume of cell clusters
2	The total number of end point
3	The standard deviation of eigenvalue (the longest) of cell clusters that are classified as spherical objects
4	The standard deviation of eigenvalue (the shortest) of cell clusters that are classified as spherical objects
5	The standard deviation of solidity of cell clusters
6	The standard deviation of branch length of cell clusters that are classified as branched objects
7	The average volume of convex hull of cell clusters that are classified as branched objects
8	The average number of junction points of cell clusters that are classified as branched objects
9	The standard deviation of number of triple points of cell clusters that are classified as branched objects
10	The standard deviation of bounding box's thickness of cell clusters that are classified as branched objects
11	The average surface to volume ratio(SAV) of cell clusters
12	The number of cell clusters
13	The average intensity of deconvolved image stack from the rhodamine channel
14	The standard deviation of solidity of cell clusters that are classified as branched objects

Table S5.8: A list of 3D profiling parameters that were most frequently selected as discriminative parameters for the compounds classification, using feature selection algorithm in each round of 10-fold cross-validation.

Chapter 6

Discussion and conclusions

6.1 Developing ultra-high content analysis platforms for high-throughput screening which do not require high resolution microscopes

Over the past years, high-throughput and high-content screening has been developed and successfully applied in various screens to identify the functional role of small molecules, peptides or RNA interference (RNAi) molecules by investigating cellular phenotypic development after exposure to these agents. With advanced laboratory robotics and automated microscopy systems, high-throughput and high-content screening enables thousands of experiments to be performed simultaneously with large volumes of microscope images generated automatically. Therefore the term "high-throughput" is used. The term "high-content" refers to the information of cellular phenotypic changes and their dynamics contained in the microscope images. With an increasing interest of observing more subtle changes for more complex cellular phenotypes, several research groups have made efforts to improve microscopy techniques or implement high resolution microscopes such as spinning disc confocal microscopy [1,2] or super resolution microscopy [3,4] for the purpose of high-throughput and high-content screening, thus providing improved sensitivity and image quality and resolution. However, these technologies are often too expensive and require major technical modifications for wider availability. Therefore many laboratories cannot afford these techniques.

In this thesis, we aimed to develop robust image analysis platforms that do not require high resolution microscopes for high-throughput and high-content screening. In **chapter 2**, we investigated the NF-κB nuclear translocation dynamics based on a confocal fluorescence microscope. In order to study a sufficient number of cells from each siRNA treatment, a dry Plan Apochromat objective with relatively low magnification (20$\times$) and low numerical aperture (0.75 NA) was used to acquire time-lapse image series. The biggest challenge was to resolve individual cells, especially in the regions where cells were touching each other or even overlap with each other. To solve this problem, we developed a novel segmentation method to estimate single cellular

area based on the topology of cells. We firstly applied watershed masked clustering (WMC) [5, 6] to detect single nuclear regions. Subsequently, the Voronoi diagram was generated to estimate the edges of cells. Finally, based on the convention that cells are ellipsoid-like objects, the best-fit ellipse in each Voronoi cell was calculated to refine single cell regions.

In **chapters 3, 4 and 5**, two images analysis platforms were developed to extract phenotype characteristics of 3D cultured micro-tissues from wide-field microscope image stacks, which contain both in-focus and out-of-focus signals due to the limited depth of field. One platform presented in **chapter 3** is based on the 2D projection. It started with compositing a single image slice by projecting only in-focus regions from each slice of image stacks. Next, watershed masked clustering was applied on the projected 2D slice to segment individual nuclei, while a local Niblack algorithm was used to define multi-cellular regions. Finally, quantification was carried out on the segmentation results and projected images, resulting in 598 parameters measured for the phenotypic profiling of each well. In **chapter 4**, we extended the 2D projection based analysis platform to enable 3D phenotypic profiling of micro-tissues. A deconvolution technique was applied to each image stack to remove out-of-focus signal by computing a point spread function (PSF) according to the optical principles. After deconvolution, we applied 3D watershed masked segmentation to detect individual nuclei. To define multi-cellular micro-tissue regions, we developed a novel segmentation method based on a sharpness level metric. By incorporating this metric in a K-means clustering method, correct intensity variation of background was estimated and a threshold was calculated dynamically for each slice. To reduce the elongated effect of nuclei in z-direction, resulting from low NA and low vertical resolution, we introduced a simple normalization method using the image of fluorescent beads with known size. Subsequently, 3D geometric models of nuclei and multi-cellular structures (also refers to as cell clusters) were reconstructed to perform phenotype quantification.

6.2   Balance between "high-content" and "high-throughput"

Image quality directly depends on several factors such as imaging modality, resolution and magnification of objective, and imaging sampling. To obtain higher quality of images, more advanced imaging techniques, higher magnification and resolution, or smaller sampling size are often required, however, this limits the throughput in high-content screening applications due to increased imaging time and computational complexity. In order to find the balance between "high-content" and "high throughput", this thesis optimized the imaging techniques, sampling sizes and image analysis modalities for different screening purposes.

6.2.1   Imaging techniques: Confocal or wide-field microscope

Compared to wide-field microscopy, confocal microscopy provides much higher resolution image stacks for 3D imaging, as each generated image slice contains only in-focus

information from the focal plane. However, the slow scanning process limits their application in high-throughput screening. Especially when the cellular structure under investigation is at a micro-tissue level, a large amount of imaging time is required to process a whole plate and consequently bleaching of the fluorescence of the specimen is caused. Therefore, wide-field microscopy is an alternative solution for the high-throughput and high-content screening of 3D cultured micro-tissues. Though the images generated by wide-field microscopes contain out-of-focus regions due to the limited depth of field, by developing an efficient image analysis method for wide-field microscope, we retrieved quantification results comparable to results generated with confocal microscope. In **chapter 4**, we developed an automated image analysis platform to reconstruct 3D micro-tissue structures from wide-field microscope images, and then profile the reconstructed structures with phenotypic parameters. We evaluated these parameters by comparing with the same parameters measured from the confocal microscope images. No significant difference was obtained, indicating that with advanced image analysis techniques, wide-field microscopy might already be sufficient for the high-content screening of 3D cultured micro-tissues to establish changes in cellular phenotypes.

6.2.2   Sampling sizes: undersampling and oversampling

For imaging, ideal sampling size is defined as the largest distance between two signal-recording points that enables to reconstruct the original continuous signal without any information loss. According to the Nyquist-Shannon sampling theorem, the conventional fluorescence wide-field microscope that was used in **chapters 4 and 5** should have the ideal axial sampling size (also referred to as pixel size) of 679 nm and vertical sampling size (also referred to as z-sampling step size) of 16882 nm. However, due to the limitation of the microscope system, which is that the smallest sampling size in axial direction is 16000 nm, almost 24 times of the ideal axial sampling size, an undersampling problem was caused. As a consequence, the performance of the deconvolution was degraded as it lacked necessary information to compute the correct PSF, and the segmentation and reconstruction results were not optimal. In this thesis, we could not intensively evaluate the influence of undersampling on deconvolution, segmentation and reconstruction due to the limitation of the microscope system. Nevertheless, it should be taken into account in future research.

Oversampling is the effect of having an actual sampling size that is smaller than the ideal sampling size. Excessive oversampling in z-direction would significantly increase the imaging time and number of image slices that need to be processed. However, the impact of oversampling on the image analysis results is still unknown. To assess this impact, in **chapter 4** we compared the phenotypic parameters measured from the wide-field microscope image stacks acquired using different z-sampling sizes (5000nm, 10000nm, 15000nm and 20000nm) with the same parameters measured from the confocal microscope images with z-sampling size of 5000nm. To our surprise, oversampling by using z-sampling size 5000nm did not help us to get more comparable results to the confocal microscope. Compared with the z-sampling size 15000nm, the

differences of parameters between the wide-field microscope images with z-sampling size 5000nm and confocal microscope images are bigger, indicating that oversampling does not only increase the imaging time, computational complexity and image storage capacity, but also impairs the quantification results.

6.2.3 Image analysis modalities: 2D projection or 3D reconstruction

To investigate the phenotypic organization of 3D cultured micro-tissues for high-throughput screening, we developed two image analysis platforms to perform high-content analysis. The platform described in **chapter 3** started with a 2D projection process to collapse each image stack into a single image slice, and followed by phenotypic profiling based on this single slice. The second platform described in **chapter 5** was able to reconstruct 3D structures from image stacks. Subsequently, phenotypic profiling was performed either on the reconstructed structures or directly on the image stacks. The advantage of the 2D projection platform is firstly reflected in the imaging time and image size. 2D projection does not require ideal z-sampling size and therefore we set up a z-sampling size (50000nm) much bigger than the ideal z-sampling size (16882nm), which significantly reduced the imaging time, as well as the image stack size. Another advantage is computational simplicity. Without computational expensive procedures such as deconvolution, 3D segmentation and 3D reconstruction, the whole analysis process is fast and not limited by available computer memory or processor speed. We tested this platform on a high-content screening of mouse breast cancer cells (4T1 cells) in a 384-well plate format using an Intel i7 2600 model with 16 GB of RAM and Windows 7 64-bits operation system. Our complete method took approximately 280 minutes of computational time for processing 384 wells, this is much faster than the analysis performed with our 3D analysis platform which took 1150 minutes for the same number of wells. The disadvantage of this platform is the fact that it discards the depth information of tissue development, which may significantly influence the phenotype characterization and classification. Our second, 3D analysis platform fills this gap by establishing a real 3D analysis platform that is able to reconstruct 3D micro-tissue structures and measure phenotypic parameters based on 3D information. However, this achievement is at the cost of almost 5 times the computational time that is needed to complete the 2D projection based platform for processing 384 wells. Furthermore, how crucial the information in z-direction that this 3D platform delivers in phenotype characterization and classification of mouse breast cancer 4T1 cells, is unknown. To answer this question, in **chapter 5** we statistically assessed the performance of the two image analysis platforms in a 4T1 cells screen, where the 4T1 cells were cultured in 384-well high-content imaging micro-plates and exposed to 12 compounds at 6 different concentrations in quadruplicate. We firstly compared the sensitivity of both platforms according to their ability to identify active concentrations of compounds (also referred to as "hits") that significantly affect the invasion of the 4T1 cells. As a result, more concentrations were identified as effective using the platform based on 2D projection, reflecting its higher sensitivity in terms

of hits identification. Secondly, we labeled the compounds according to their biological activity and applied various classification techniques based on the phenotypic parameters. The classification accuracy of the 3D platform was slightly higher than the 2D projection platform, indicating a benefit of incorporating spatial information in z-direction for phenotype characterization and classification. In the end, we tested both platforms on the reproducibility of concentration dependent phenotypic trajectory for each compound. The method to model the phenotypic trajectory is described in **chapter 3**. The 3D analysis platform provided much higher reproducibility than the 2D projection platform, further confirming the superiority of the 3D analysis in respect to phenotype characterization.

6.3 Ultra-high content analysis

Advances in high throughput technology have enabled collecting thousands of images from large-scale screens in a single day [7–10]. However, the image analysis technique is still a bottleneck. Several pioneering screens have either relied on visual scoring by experts [11,12], or developed their in-house automated analysis software to measure single or at most few phenotypic parameters [13]. Those parameters are often too specific for a certain type of phenotype and hence not suitable for other screens. Moreover, those methods largely depend on the expertise of biologists, but often even biologists are not certain of all possible effects. Consequently, many details were lost and subtle changes may not have been detected. In this thesis, our goal was to develop ultra-high content analysis platforms which should be able to collect maximum information of phenotypes from images, and applied for various types of screening assays. Compared to other existing high-content analysis platforms, the novelties of our platforms were presented in its ability to extract subpopulation information and its ultra-high level in information content.

6.3.1 Subpopulations

Many researchers have found that responses of cells are inhomogeneous [14–16]. For example, in **chapter 2** we showed that not all HepG2 cells respond to the TNFα stimulation synchronously. Therefore, subpopulations should be taken into account and this requires single cell measurement. In this thesis, all developed image analysis platforms are able to perform single-object measurements in order to extract inhomogeneous, cell subpopulation information. In **chapter 2**, we firstly measured the time course of NF-κB nuclear translocation for each single HepG2 cell. Afterwards, the cells were categorized into different subgroups according to the number of translocation peaks, and finally cell subpopulations were investigated before and after TNFα stimulation. In **chapter 3**, we embed an automated classification system in the developed platform to automatically distinguish the spherical cell clusters and branched cell networks. Subsequently, the subpopulation related parameters were collected and included as a part of phenotypic profiling. In order to assess the value of these subpopulations information, we performed compounds clustering and classification with

and without subpopulation parameters on the 3D cultured 4T1 cells screen, where cells were exposed to 29 compounds that can be categorized according to their biological target. The result showed that excluding subpopulation parameters resulted in a failure to co-cluster compounds with the same biological activity and decreased classification accuracy. Similarly, the 3D analysis platform developed in **chapter 5** classified cell clusters to either spherical structure or branched structure according to the sphericity and number of branches. Extracting subpopulation related parameters and using them for compounds classification increased the classification accuracy, further confirming the importance of subpopulation information and which should be taken into account for phenotype characterization.

6.3.2 Ultra-high level in information content

Most of NF-κB translocation studies only investigate the dynamics of NF-κB nuclear translocation through the time course of translocation profiles. We have extended our system so that it can automatically quantify 26 analogue parameters for each individual translocation profile, such as the number of translocation peaks and time between consecutive peaks [5]. These parameters can be used to distinguish inhomogeneous cell populations and might be a powerful tool to study heterogeneous cell behavior in the future. Both image analysis platforms we developed in **chapter 3** and **chapter 5** can measure not only classic morphological parameters, but also topological parameters, intensity properties, texture, moments, as well as subpopulation information, therefore providing the full spectrum of phenotypic information. This is not only beneficial for detecting subtle changes that would otherwise be missed, but also enables system-level studies of the full range of phenotype characteristics. For example, we have showed in **chapter 3** that the concentration dependent phenotypic trajectory for each compound tested in the 4T1 cells screen could be modeled based on the full spectrum of phenotypic characteristics. We also used the phenotypic trajectories to successfully co-cluster compounds with the same biological activity. For both the 2D projection based platform and 3D analysis platform, the compounds classification according to their biological activity was achieved with high classification accuracy.

Another advantage of measuring the full spectrum of phenotypic information is robustness. It enables us to analyze 3D-cellular phenotypes under a wide variety of conditions. In **chapter 3**, we tested our analysis platform on a high-content screening of 44 known human breast cancer cell lines that have been categorized as basal-A, basal-B, luminal or luminal/ERBB2+ based on their gene expression profiles. The result showed that those cell lines can be correctly classified based on the phenotypic profiling with a classification error rater less than 6%. Although initially the use of all parameters may seem redundant, our feature selection methodologies automatically identify those features that contribute most to the separation and characterization of the particular phenotypes under study. In addition to the initial application of our analysis platforms in the 4T1 cell screen and classification of 44 human breast cancer cell lines, described in **chapter 3 and 4**, they have been successfully used in various other screens, such as compound screens in invasive prostate cancer cells and

invasive lung cancer cells (not yet published), confirming their robustness and their wider applicability.

6.4   Future perspective

With the successful development of advanced image analysis platforms for high-throughput and high-content screens, we are able to extract a wealth of information from screening images. The platform developed for analyzing NF-κB nuclear translocation in **chapter 2** can provide a time course profile of translocation dynamics and 26 analogue parameters for each individual cell, as well as for the whole population. In **chapter 5**, the 2D projection based platform we developed to profile the phenotype of 3D cultured micro-tissues can measure up to 794 parameters, while the 3D analysis platform can measure up to 290 parameters. Now the bottleneck is moving downstream to the data analysis and data mining. How to fully explore these rich datasets to reveal cellular signaling networks, is a challenge. Most of the data analysis methods for high-throughput and high-content screening still remain at the stage of hit identification. Quality control such as evaluating *Z'-factor* based on a single parameter is already out of date. For data mining, we showed a few applications in this thesis. For example, we applied phenotype clustering and classification on the 3D cultured 4T1 cell screen to characterize the compounds according to their biological target. For the NF-κB nuclear translocation assay, we applied this method to identify novel regulators of TNFα-induced apoptosis in human HepG2 cells [17]. However, more powerful tools and systematical methodologies are still needed to relate phenotype characteristics to cellular pathways.

In order to extract reliable information from microscope images in an efficient way, more efforts are required to improve current image analysis methods for high-throughput and high-content screening. For example, the platform developed in **chapter 3** still requires human interaction to train an embedded classifier for recognizing different subpopulations, though human interaction is needed only once for the screening of all similar phenotype structures. A built-in clustering which is able to automatically group subpopulations would improve the automation level so that human labor is not required in the whole image analysis process. Moreover, potential subpopulations which are not identified by human vision would be detected. In **chapter 4**, a simple method to correct nuclear size in z-direction was performed using images of fluorescent beads with known size. However, more advanced methods to normalize the dimension of multi-cellular structures despite of their irregular size and shapes are still needed. Another direction to improve our image analysis methods is to increase the analysis time efficiency, where enabling multi-thread processing on a computer cluster would greatly contribute.

Although the impact of imaging resolution on the quantification result was not extensively investigated in this thesis, we should bear in mind that different imaging modalities and image resolutions would generate different quantification results and therefore may lead to a different conclusion. In **chapter 4**, we showed that different z-

sampling sizes influenced the morphological parameters, topological parameters and moments obtained from our image analysis platform, and therefore an optimal z-sampling size was defined for high-content screening of 3D cultured micro-tissues. In the future, sampling in axial direction should also be optimized, as well as other imaging setting.

The integration of high-throughput and high-content screening with other omics technologies is getting more and more attention. For example, a study recently integrated a high-content RNAi screen with phosphoproteomics and transcriptomics to unravel DNA damage response signaling processes in mouse embryonic stem cells treated with cisplatin [18]. Another example is the integration of our NF-κB nuclear translocation assay in a RNAi screen to identify novel regulators of TNFα-induced apoptosis in human HepG2 cells [18]. This study used an in vitro liver cell model and methods such as high-content imaging and functional genomics, which brought us closer to understanding the molecular mechanisms of chronic inflammation and drug-induced liver injury. All those studies showed that combining different omics data with high-throughput and high-content screening would piece together a more comprehensive picture of signaling networks. Furthermore, fueled with mathematical modeling, integration of high-throughput and high-content screening with other omics technologies would open a new era for systems biology.

References

[1] Chan, G. K., Richards, G. R., Peters, M. & Simpson, P. B. High content kinetic assays of neuronal signaling implemented on BD pathway HT. Assay Drug Dev Technol 3, 623-636 (2005).

[2] Krausz, E. High-content siRNA screening. Mol Biosyst 3, 232-240 (2007).

[3] Gitai, Z. New fluorescence microscopy methods for microbiology: sharper, faster, and quantitative. Curr Opin Microbiol 12, 341-346 (2009).

[4] Ries, J., Kaplan, C., Platonova, E., Eghlidi, H. & Ewers, H. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat Methods 9, 582-584 (2012).

[5] Di, Z. et al. Automated analysis of NF-kappaB nuclear translocation kinetics in high-throughput screening. PloS one 7, e52337 (2012).

[6] Kuan Yan & Verbeek, F. J. Segmentation for high-throughput image analysis: watershed masked clustering. In: SoLA'12 Proceedings of the 5th international conference on Leveraging Applications of Formal Methods, Verification and Validation: applications and case studies. (eds T. Margaria & B. Steffen) 25-41 (Springer-Verlag).

[7] Carpenter, A. E. & Sabatini, D. M. Systematic genome-wide screens of gene function. Nat Rev Genet 5, 11-22 (2004).

[8] Dasgupta, R. & Perrimon, N. Using RNAi to catch Drosophila genes in a web of interactions: insights into cancer research. Oncogene 23, 8359-8365 (2004).

[9] Moffat, J. et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell 124, 1283-1298 (2006).

[10] Vanhecke, D. & Janitz, M. Functional genomics using high-throughput RNA interference. Drug Discov Today 10, 205-212 (2005).

[11] Kiger, A. A. et al. A functional genomic analysis of cell morphology using RNA interference. J Biol 2, 27 (2003).

[12] Kim, J. K. et al. Functional genomic analysis of RNA interference in C. elegans. Science 308, 1164-1167 (2005).

[13] Korn, K. & Krausz, E. Cell-based high-content screening of small-molecule libraries. Curr Opin Chem Biol 11, 503-510 (2007).

[14] Covert, M. W., Leung, T. H., Gaston, J. E. & Baltimore, D. Achieving stability of lipopolysaccharide-induced NF-kappaB activation. Science 309, 1854-1857 (2005).

[15] Lee, T. K. et al. A noisy paracrine signal determines the cellular NF-kappaB response to lipopolysaccharide. Sci Signal 2, ra65 (2009).

[16] Paszek, P. et al. Population robustness arising from cellular heterogeneity. Proc Natl Acad Sci U S A 107, 11644-11649 (2010).

[17] Fredriksson, L. TNFα-signalling in drug induced liver injury. 113-135 (2013).

[18] CPuigvert, J. C. et al. Systems Biology Approach Identifies the Kinase Csnk1a1 as a Regulator of the DNA Damage Response in Embryonic Stem Cells. Sci Signal 6 (2013).

Nederlandse samenvatting

High-throughput en high-content beeldanalyse is een krachtige methode om fenotypische responsen van cellen en multicellulaire structuren te associëren met veranderingen in hun functioneren die worden veroorzaakt door kleine moleculen, peptiden of interferentie RNA (RNAi), wat uiteindelijk de identificatie van nieuwe geneesmiddel targets en/of nieuwe geneesmiddelen mogelijk maakt. Echter, een aantal beperkingen van de huidige beeldanalyse methoden beperken de toepassing van high-throughput en high-content screening tot alleen de identificatie van biologisch actieve verbindingen of genetische componenten. Het volledige potentieel van high-throughput en high-content screening om systematisch fenotypes te bestuderen, bijvoorbeeld om fenotypes te karakteriseren en te associëren met modulatie van cellulaire signaal mechanismen, wordt nu nog onderschat. Een belangrijke beperking is dat huidige beeldanalyse methoden allen een aantal typische eigenschappen meten die gebaseerd zijn op bekende fenotypische veranderingen die plaats vinden onder vergelijkbare experimentele omstandigheden. Daarom worden nog onbekende fenotypes en subtiele veranderingen niet herkend en is een complete karakterisering van fenotypes niet mogelijk. Een andere belangrijke beperking van de huidige analyse methoden is het gebrek aan een meer algemene toepassing omdat veel van deze methoden alleen zijn ontwikkeld voor bepaalde fenotypes en cellen. Nog belangrijker is dat deze methoden sterk afhankelijk zijn van de resolutie van de beelden. Wanneer de beeld resolutie en vergroting relatief laag zijn, kunnen deze methoden niet goed individuele cellen fenotyperen. Als gevolg is heterogeen gedrag van cellen niet waarneembaar.

Eén van de doelstellingen van het onderzoek dat in dit proefschrift wordt gepresenteerd is de ontwikkeling van robuuste en efficiënte beeldanalyse methoden die geschikt zijn voor studies naar individuele cellen die niet zijn beperkt tot specifieke cel types, cel dichtheden of een specifieke beeldkwaliteit. Bovendien moet er rekening gehouden worden met heterogeniteit van subpopulaties van cellen en moet relevante informatie kunnen worden verzameld voor fenotypische karakterisering. In hoofdstuk 2 hebben we een dergelijke methode ontwikkeld om de dynamiek van translocatie van NF-κB naar de kern te bestuderen in enkelvoudige cellen na stimulatie van de NF-κB route na externe stimulatie. We hebben aangetoond dat deze methode kan worden toegepast bij confocale fluorescentie microscopie met relatief lage vergroting (20$\times$) en lage numerieke apertuur (0.75NA). Alhoewel de cellijn die we hebben gebruikt om de methode te valideren, samengepakte en op elkaar groeiende cellen vertoonde wat de mogelijkheid om losse cellen te analyseren moeilijker maakte, hebben we door twee

nieuwe segmentatie methoden BEVC en WMC te introduceren toch enkelvoudige cellen en plaatsen van kernen kunnen definiëren. In dit onderzoek werden subpopulaties van cellen gedefinieerd op grond van het aantal translocaties van NF-κB naar de kern. Een vergelijking van de grootte van de verschillende subpopulaties liet zien dat BMS-345541 (een IKK-remmer die de translocatie van NF-κB naar de kern voorkomt) een effect heeft op de translocaties van NF-κB naar de kern door het percentage van cellen met 3 translocatie pieken te verminderen, en het percentage van cellen met maar één piek te vergroten. Dit effect zou niet te zien zijn geweest in de gemiddelde respons van de hele cel populatie.

Een andere belangrijke doelstelling van het onderzoek dat in dit proefschrift wordt gepresenteerd is de ontwikkeling van een beeldanalyse systeem voor de gedetailleerde en complete karakterisering van cellulaire fenotypes uit beelden met lage resolutie. In **hoofdstukken 3, 4 en 5** zijn hiervoor twee beeldanalyse methoden ontwikkeld om de fenotypes te karakteriseren van in 3D geweekt microweefsel uit planparallelle beelden die waren verkregen met conventionele wide-field fluorescentie microscopie. Eén system dat beschreven wordt in **hoofdstuk 3** is gebaseerd op de projectie van de planparallelle beelden naar een 2D beeld. Dit begint met het samenstellen van een enkel beeld door projectie van alleen die gedeelten van de planparallelle beelden die in het scherptevlak liggen. Vervolgens werd WMC toegepast op het geprojecteerde 2D beeld om de individuele kernen te segmenteren, terwijl een lokaal toegepast Niblack algoritme werd gebruikt om multicellulaire structuren te definiëren. Tenslotte werd fenotypering uitgevoerd op de gesegmenteerde resultaten en geprojecteerde beelden. In **hoofdstuk 4** hebben we de op 2D projectie gebaseerde methode uitgebreid om 3D fenotypering van microweefsel mogelijk te maken. Eerst werd een deconvolutie techniek toegepast op iedere opeenvolgende reeks van beelden om informatie die niet in het scherptevlak lag te verwijderen. Vervolgens werd een 3D WMC methode geïntroduceerd om iedere afzonderlijke kern te definiëren, en een nieuwe methode om de scherpte te schatten werd ingebouwd in de methode om multicellulaire gebieden te segmenteren. Na corrigeren van een effect dat de kernen in de z-richting te lang maakt, werden 3D geometrische modellen gemaakt van kernen en multicellulaire gebieden om fenoypes te typeren. Zowel in de op 2D projectie gebaseerde methode als in de 3D fenotyperings methode (**hoofdstukken 3 en 5**) is een automatische classificering ingebouwd om ronde celclusters en vertakte celnetwerken te onderscheiden. We hebben statistisch aangetoond dat het gebruik van parameters die kenmerkend zijn voor de subpopulaties in de fenotypische classificering de nauwkeurigheid van classificeren verhoogd, wat het belang van informatie van de subpopulaties voor fenotypering bevestigd. Om de bredere toepassing van onze methode te onderzoeken, hebben we de methode beschreven in **hoofdstuk 2** gebruikt om 44 bekende humane borstkanker cellijnen te classificeren gebaseerd op het profileren van hun fenotypes. Er werd een hoge nauwkeurigheid van classificatie verkregen, waarmee een bredere toepassing van onze methode wordt gesuggereerd. Wij hebben beide methoden ook toegepast in verschillende andere screens, zoals een screen van chemische verbindingen in invasieve PC3 cellen en invasieve longkanker cellen. Dit toont ook aan dat de methoden breder toegepast kunnen worden.

Nog een ander doel van het onderzoek beschreven in dit proefschrift is het ontwikkelen van een 'ultra-high content' beeldanalyse methode die in staat moet zijn het maximum aan informatie te verzamelen uit fenotypes. In **hoofdstuk 2** hebben we onze methode zodanig uitgebreid dat deze in staat is om 26 analoge parameters te kwantificeren voor ieder translocatie profiel van NF-κB naar de kern zoals aantal translocatie pieken, en tijd tussen twee opeenvolgende pieken. Deze parameters konden worden gebruikt om verschillende subpopulaties te onderscheiden binnen een niet-homogene populatie, waarmee dit een belangrijke methode is die in de toekomst gebruikt kan worden om heterogene celpopulaties te bestuderen. Beide methoden die in **hoofdstuk 3** en in **hoofdstuk 5** zijn ontwikkeld kunnen niet alleen klassieke morfologische parameters meten, maar ook topologische parameters, intensiteit parameters, texturen, momenten, en informatie van subpopulaties, waarmee het volledige spectrum van fenotypische informatie wordt aangeboden. Wij hebben de bijdrage bepaald van de verschillende parameter klassen aan de typering van fenotypes van 4T1 cellen blootgesteld aan verschillende tyrosine kinase remmers. De resultaten lieten zien dat weglaten van een bepaalde parameter klasse uit de analyse, de nauwkeurigheid van classificatie verminderde waarmee het belang van kwantificering van beelden met het volledige spectrum van fenotypische informatie werd aangetoond.

Omdat met onze beeldanalyse methoden een schat aan informatie kan worden verkregen, kunnen meer systematische benaderingen worden toegepast om fenotypes te karakteriseren of om fenotypes te associëren met verschillende cellulaire signaal routes. In **hoofdstuk 3** hebben we een regressiemoduleringsmethode geïntroduceerd om de concentratie-afhankelijke trajecten te construeren in multifenotypische parameter ruimte voor biologische actieve verbindingen die het cellulaire fenotype beïnvloeden. Met behulp van een nieuwe clusteringsmethode die gebaseerd is op de afstand tussen twee trajecten, hebben we gegevens verkregen die laten zien dat fenotypische responsen van 4T1 cellen op kinase remmers specifiek zijn voor het biologische doelwit dat word geremd. Om dit resultaat te valideren, hebben we de geteste verbindingen in verschillende klassen ingedeeld naar gelang hun biologische doelwit. Er konden vervolgens classificeerders worden getraind die de verbindingen op grond van hun fenotypische profiel (2D projectie gebaseerd of 3D gebaseerd) met succes konden classificeren. Dit resultaat opent de weg naar een nieuwe strategie om activiteit te onderzoeken van farmacologisch actieve verbindingen of om bijwerkingen van bestaande verbindingen te identificeren.

English summary

High-throughput and high-content image analysis is a powerful tool to associate phenotypic responses of cells and multi-cellular structures, with modulation of their functional components by small molecules, peptides or RNA interference (RNAi), ultimately enabling the identification of novel drug targets and/or novel drug molecules. However, several limitations of current image analysis methods restrict the application of high-throughput and high-content screening to the identification of biologically active compounds or genetic players only. The full potential of high-throughput and high-content screening to systematically study phenotypes, for example to characterize phenotypes and associate them with modulation of cellular signaling pathways, is underestimated as yet. A major limitation is that current image analysis methods only measure a few representative features which are based on known phenotypic changes that occur under similar experimental conditions. Consequently, novel phenotypes and subtle changes remain undetected and a comprehensive characterization of phenotypes cannot be achieved. Another major limitation of current image analysis methods is lack of robustness as many of them are designed only for specific phenotypes and cell types. More importantly, these methods are highly dependent on the resolution of microscopy images. When image resolution and magnification are relatively low, these methods cannot profile phenotypes at a single cell level. As a consequence, heterogeneous cell behavior is obscured.

One of the goals of the research presented in this thesis is to develop robust and efficient image analysis methods suitable for single-cell studies that should not be limited to specific cell types, cell densities or image quality. More importantly, heterogeneity between cell subpopulations should be taken into account and relevant information should be collected for phenotype characterization. In **chapter 2**, we developed such a methodology to investigate the NF-κB nuclear translocation dynamics in single cells after activating the NF-κB pathway with external stimuli. It is shown that this methodology can be applied to confocal fluorescence microscopy with relatively low magnification (20$\times$) and low numerical aperture (0.75NA) objective. Although the cell line (HepG2) we used for the method validation showed clustered and stacked cell growth which increased the difficulty to separate individual cells, by introducing two novel segmentation methods BEVC and WMC we can still define single cellular and nuclear regions accurately. In this research, different cell subpopulations were defined according to the number of NF-κB nuclear translocation peaks. Comparing the size of different subpopulations revealed that BMS-345541 (an

IKK-inhibitor which prevents NF-κB nuclear translocation) has an impact on NF-κB nuclear translocation by decreasing the percentage of cells with three translocation peaks, and increasing the percentage of cells with only one translocation peak. This impact would not have been resolved in the average nuclear translocation response over the whole cell population.

Another important goal of this research is the development of image analysis platforms for the detailed and comprehensive characterization of cellular phenotypes from low-resolution images. In **chapters 3, 4, and 5**, two image analysis platforms were developed to profile phenotype of 3D cultured micro-tissues from image stacks which were generated by conventional wide-field microscopes. One platform presented in **chapter 3** is based on the 2D projection of image stacks. It starts with composing a single image slice by projecting only in-focus regions from each slice of image stacks. Next, WMC was applied on the projected 2D slice to segment individual nuclei, while a local Niblack algorithm was used to define multi-cellular regions. Finally, phenotype profiling was carried out on the segmentation results and the projected images. In **chapter 4**, we extended the 2D projection based analysis platform to enable 3D phenotypic profiling of micro-tissues. Firstly, a deconvolution technique was applied to each image stack to remove out-of-focus information. Secondly, a 3D WMC method was introduced to define individual nuclei, and a novel sharpness estimation method was incorporated in the segmentation of multi-cellular regions. After correcting the elongation effect of nuclei in z-direction, 3D geometric models of nuclei and multi-cellular structures were reconstructed to perform phenotype profiling. Both the 2D projection based analysis platform and the 3D phenotypic profiling platform (**chapters 3 and 5**) embed an automated classification system to automatically distinguish the spherical cell clusters and branched cell networks. We statistically proved that extracting subpopulation related parameters and including them for phenotype classification increased the classification accuracy, confirming the importance of subpopulation information in the phenotype characterization. To investigate the robustness of our image analysis platforms, we applied the platform presented in **chapter 2** to classify 44 known human breast cancer cell lines based on the phenotype profiles. The result showed high classification accuracy, indicating a wider applicability of our method. We also successfully applied both platforms in various other screens, such as compound screens in invasive PC3 cells and invasive lung cancer cells. This further confirms the robustness of our methods.

In this research, we aim to develop ultra-high content analysis platforms which should be able to collect maximum information of phenotypes from images. In **chapter 2**, we have extended our image analysis so that it can automatically quantify 26 analogue parameters for each individual NF-κB nuclear translocation profile, such as the number of translocation peaks and time between consecutive peaks. These parameters were used to distinguish different cell subpopulations within an inhomogeneous population, representing a powerful tool to study heterogeneous cell behavior in the future. Both image analysis platforms we developed in **chapter 3** and **chapter 5** can measure not only classic morphological parameters, but also topological parameters, intensity properties, texture, moments, as well as subpopulation informa-

tion, therefore providing the full spectrum of phenotypic information. We analyzed the contribution of different parameter classes to the phenotype characterization of the 4T1 cell cultures exposed to various protein kinase inhibitors. The result showed that excluding any parameter class decreased the accuracy of phenotype classification, confirming the importance of quantifying images with a full spectrum of phenotypic information.

Due to a wealth of information that can be extracted by our image analysis methods, more systematical approaches can be used to characterize phenotypes or associate phenotypes with modulation of different cellular pathways. In **chapter 3**, we introduced a regression modeling method to construct the concentration dependent trajectories in multi-phenotypic parameter space for biologically active compounds that affect cellular phenotypes. By means of a novel clustering method which is based on the distance between two trajectories, we obtained data suggesting that the phenotypic responses of 4T1 cells to protein kinase inhibitors are specific for the biological target that is inhibited. To validate this result, we categorized tested compounds into different classes according to their biological target. Classifiers were trained to successfully classify those compounds based on their corresponding phenotype profiles (2D projection based profiles or 3D profiles). This result may lead to a new strategy to study activity of new pharmacologically active compounds or identify off-target effects of existing compounds.

Curriculum vitae

Zi Di was born on May the 24th 1984 in Chongqing, P. R. China. In 2002, she graduated from Yongrong Secondary School of Chongqing, China. From 2002 and 2006, she studied at the Beijing Language and Culture University, China. In the summer of 2006, she received her BSc degree in Computer Science. Directly after her bachelor study, she started her Master education in Leiden University, The Netherlands. During her Master education, she specialized in Bioinformatics which is a track of the Computer Science study and is organized by Leiden University and Delft University. In 2009, she finished her Master project "Evaluation of systems for analysis of high-throughput screening of static focal adhesions" under the supervision of Dr. Fons Verbeek in the Section Imaging & Bioinformatics of the Leiden Institute of Advanced Computer Science (LIACS). Later in the same year, she received her MSc degree in Bioinformatics.

In 2009, Zi Di started her PhD research through a joint project between the Division of Toxicology of the Leiden Academic Centre for Drug Research (LACDR) and the Section Imaging & BioInformatics of LIACS in the Netherlands. Under the supervision of Dr. John Meerman, Dr. Fons Verbeek, and Prof. Dr. Bob van de Water, her research focused on developing novel image analysis algorithms and performing advanced data analysis methods, especially for high-throughput and high-content screening. She was also involved in a data-integration project in collaboration with other bioinformaticians to provide bioinformatics support to PhD students in the context of Netherlands Toxicogenomics Centre (NTC).

List of publications

Automated Analysis of NF-kB Nuclear Translocation Kinetics in High-Throughput Screening
Di Z., Herpers B., Fredriksson L.E., Yan K., van de Water B., Verbeek F.J., Meerman J.H., PLOS One, vol. 7, no. 12, pp. e52337.

Automated imaging, quantification and classification of NF-kB and Nrf2 responses to identify adverse drug reactions in HepG2 cells
Herpers B., Di Z., Fredriksson L.E., van de Water B., Drug Metab Rev, vol. 42, pp. 156-156.

Automation in Cytomics: A Modern RDBMS Based Platform for Image Analysis and Management in High-Throughput Screening Experiments
Larios E., Zhang Y., Yan K., Di Z., Le Dévédec S.E., Groffen F., Verbeek F.J., Proceedings of Int. Conf. on Health Information Science, vol. 7231, Beijing, Springer Berlin / Heidelberg, pp. 76-87.

Ultra high-content analysis and phenotype profiling of 3D cultured micro-tissues
Di Z., Klop M.J.D., Rogkoti V., Le Dévédec S.E., van de Water B., Verbeek F.J., Meerman J.H., Price L.S., Submitted to Bioinformatics, November 2013

Automated analysis pipeline for 3D surface reconstruction and phenotype profiling of 3D cultured micro-tissues suitable for high-content and high-throughput screening
Di Z., Klop M.J.D., van de Water B., Price L.S., Meerman J.H., Verbeek F.J., Manuscript in preparation, 2013