学位論文　（要約）

Label-free detection of drug responses of cancer cells by machine-learning-assisted high-throughput bright-field imaging

（機械学習を用いたハイスループット明視野顕微法による
がん細胞薬剤応答性の無標識検出）

平成 29 年 12 月博士（理学）申請

東京大学大学院理学系研究科
化学専攻

小林　博文

Label-free detection of drug responses of cancer cells by machine-learning-assisted high-throughput bright-field imaging

by

Hirofumi Kobayashi

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
at
The University of Tokyo

Abstract

Cellular drug responses are the most fundamental reactions in biology and medicine, which provides valuable information for understanding the fundamental mechanisms of biology and the development of new drugs. To investigate the drug responses efficiently and massively, high-throughput screening has played a major role over the past decades. This technique is powerful in exploring a large number of samples, but is often limited by the fact that it only acquires a single variate from each sample, for example, live or dead. Owing to the technological advances of imaging and robotics, image-based high-throughput screening, namely high-content screening, has been developed to acquire a significantly larger number of variables in each sample to capture more information in the cellular drug responses. Such a multivariate image-based technique has been proven effective in a broad spectrum of quantitative analysis not simply in drug responses but also associated with genome and proteome analysis.

Despite the utility of image-based screening, it falls short in addressing the full needs of biological applications as it inherently requires fluorescent labeling which has several drawbacks. First, fluorescent probes are not available for all target molecules and may interfere with natural cellular functions. Although a wide range of immunofluorescent probes are commercially available for cellular imaging, they are costly and require time-consuming labeling processes including cell fixation which kills the cells, and thus hinders live-cell and large-scale assays. In addition, while fluorescently tagged cell lines can offer live-cell assays without such a labeling process, the development of such cell lines requires even more efforts than immunofluorescent labeling. Therefore, a method for image-based high-content screening without using fluorophores is desired for easy manipulation and economical assays.

In my doctoral work, I studied cellular drug responses in a label-free manner by high-throughput bright-field imaging with the aid of machine-learning techniques in various settings. Specifically, small morphological changes present in the cellular images, which are too subtle to be observed by human eyes, were consistently detected. Hence, dose dependence of the drug responses was visualized. In addition to the proof-of-concept

demonstration, I also applied the method to detection of drug responses in whole blood as well as in an ex vivo setting.

Thesis Supervisor: Keisuke Goda
Tile: Professor of Chemistry

Acknowledgement

I would like to express my deepest gratitude to everyone who have made this thesis work possible. First of all, I am deeply grateful to my supervisor Keisuke Goda for accepting me in his group as a PhD student. His generosity and kindness allowed me to learn from a variety of works including setting up new laboratory rooms and writing grant applications. These works have been my precious experiences and greatly enriched my knowledge and perspectives. He also gave me the freedom to work on what I wanted while providing the necessary resources to help me succeeded in my work. His leadership and mentorship were demanding but instructive. I have to admit that I wouldn't be able to come so far without his mentorship.

Second of all, it has been my great fortune to work with Cheng Lei, my team leader. His greatness and mercifulness always encouraged and backed me to take on a challenge while helping me solve problems patiently. This thesis work will never be completed without his dedicated support. His humor entertained everyone in the laboratory. Thanks to him, I could also improve my Chinese dramatically and make many Chinese friends. I wish I could take him with me as I move onto the next stage of my professional career.

I would also like to thank another team leader of mine Hideharu Mikami for his dedicated teaching and supervision. His superb technical skills and expertise about building optical setups, electronics and programming impressed me greatly. Most of my knowledge and skills about optics are inherited from him. His attitude toward research also influenced me greatly.

Yuko Kanda has been an incredible supporter as well. Being a seasoned technician, she has been the only one who could help me for my biological experiments. She has also been one of the few people with whom I can discuss my results of biological experiment. My thesis work would not finish timely without her devoted support.I am also grateful to Kotaro Hiramatsu who has kindly been my supervisor for animal experiments. His advices substantially improved the quality of for my presentation at the preliminary thesis report. Akihiro Isozaki has also been a great supporter for teaching me and helping me perform microfluidic experiments. I enjoyed every discussion with him.

I am thankful to my research collaborators. First, I appreciate Haku Kyo, Ryohei Katayama and Naoya Fujia for providing me knowledge and resources for gene modification and animal experiments. I also thank Kazumichi Nagasawa and Hiroshi Watarai for providing FACS expertise. The advices given by Ming Li were invaluable, and accelerated my work significantly. My team members, especially Baoshan Guo and Yiyue Jiang have always been cheerful. I will never forget the days I spent with them. Finally, and most importantly, I am deeply grateful to my family for being supportive and letting me peruse my PhD degree.

This work was funded by ImPACT Program of the CSTI (Cabinet Office, Government of Japan) and partly by Noguchi Shitagau Research Grant, New Technology Development Foundation, Konica Minolta Imaging Science Encouragement Award, JSPS KAKENHI Grant Numbers 15J02613, 25702024 and 25560190, JGC-S Scholarship Foundation, Mitsubishi Foundation, TOBIRA Award, and Takeda Science Foundation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cellular drug responses in medicine

Cellular drug responses are the most fundamental reactions in biology and medicine. Historically, many drugs have been developed by the investigation of drug responses between cells and natural products. A recent analysis of new drugs approved by the US Food and Drug Administration (FDA) between 1999 and 2008 has revealed that among the discovery of first-in-class drugs[1] small-molecule drugs[2], 28 originated from phenotypic screening in which drug responses are monitored to select the leading compounds, whereas 17 were discovered through target-based drug discovery where leading compounds were screened to act on specific hypothetical targets. Although most large molecule drugs have been isolated from target-based screening, phenotypic screening is also employed to isolate antibodies as new medicines against cancer and infectious diseases [1]. Moreover, recent rapid advances in biotechnology including the generation of induced pluripotent stem (iPS) cell [2], gene-editing technology such as CRISPR-Cas [3], and three-dimension cell culture technology such as spheroid [4], organ-on-chip [5] and organoids [6] opens up new fields for phenotypic screening.

Looking into the cellular drug responses also helps understanding of the biology behind the pharmaceutical reactions. Coupled with genetic modification, studying cellular drug responses can elucidate the biology of diseases at a genome level. For example, by treating Drosophila primary neurons that expresses pathogenic Huntingtin protein with small molecule drugs and RNAi, a suppressing protein and four novel drugs of Huntington

[1] Drugs that have mechanisms of action different from those of preceding drugs. First-in-class drug is one indicator to scale the innovation of drug discovery.

[2] Small molecule drugs refer to drugs often with molecular weight lower than 900, which can easily enter cells through cell membranes. In contrast, large molecule drugs such as monoclonal antibodies are not able to diffuse across cell membranes.

disease were identified [7]. Another study found discrete functions of cell wall components in budding yeast by analyses of morphological changes after drug treatment [8]. By employing multi-variate methods, investigation in cellular drug responses can elucidate and characterize cellular heterogeneity [9]. As new therapeutic agents are rapidly gaining their momenta such as monoclonal antibodies, stem-cell therapy and gene-editing technology, the definition of "drug" is becoming more and more ambiguous. With the definition of drug in a broad sense, cellular drug responses can refer any perturbation that acts on cells including small molecule drugs, antibodies, RNAi and gene modification. In this context, studying cellular drug responses especially with multi-variate methods holds a great promise to uncover new biological findings.

1.2 High-throughput screening

High-throughput screening (HTS) is a platform typically designed for drug discovery and related biological research. It consists of multiple automated technologies including liquid handling, cell culturing and optical measurement collective implementation of automated technologies, which replaces former manual processes greatly improving the capacity and speed of drug discovery. The development of HTS can be traced back to the late 1980s in which microbes were used to screen antibiotics [10]. In the early 1990s HTS was applied to anticancer drug screening in which a panel of 60 human tumor cell lines were used to represent leukemia, melanoma, as well as cancers of colon, lung, ovary, kidney, and central nervous system [11]. In the first demonstration of HTS, cells treated with five 10-fold dilutions of test compounds for 48 hours were measured in an end-point assay of cell viability at a throughput of 400 compounds per week [11]. Owing to the advances of robotics, biochemistry and nanotechnology, a variety of assays is now compatible with HTS enabling the throughput up to 100,000 compounds per day [12].

HTS compatible assays can be categorized into two types: biochemical assays and cell-based assays. Biochemical assays are designed to directly investigate specific molecules, and have a longer history in the pharmaceutical industry. Such assays include the measurement of enzyme activity (such as kinase [13], protease [14], and transferase [15]), receptor-ligand binding (such as G-protein coupled receptor [16], nuclear receptor [17], and ion-channel [18]) and protein-protein interaction [19]. Although

this type of assays is readily miniaturized with less variation, the output activities do not always represent the actual ones in a cellular context [20]. In addition, not all target molecules can be purified for this type of assays. Cell-based assays, on the other hand, are designed to assess an entire pathway as a whole, and do not require prior knowledge about target molecules. Such assays include functional assays [21,22], reporter gene assays [23,24], and phenotypic assays [25,26].

Diverse types of detection modality can be incorporated in HTS including LC-MS (liquid chromatography-mass spectrometry) [27], GC-MS (gas chromatography-mass spectrometry) [28], genome sequencing [29], RNA sequencing [30] and optical measurement such as detection of fluorescence, luminescence, absorbance, and spectroscopic approach. In the recent HTS, fluorescence detection is arguably the most broadly used method in HTS due to its high sensitivity, high specificity, a variety of fluorophores, and relatively simple operation comparing to other modalities. Fluorescence-based assays can be categorized into two types: bulk detection and single-molecule detection. The first type includes detection of total fluorescence intensity, fluorescence polarization [31], fluorescence resonance energy transfer (FRET) [22], fluorescence lifetime [32], time-resolved fluorescence [33], and combinations of these parameters such as time-resolved fluorescence polarization [34]. The other type includes fluorescence correlation spectroscopy [35] and fluorescence intensity distribution analysis [36].

A major limitation in conventional HTS is that it only provides single variate readout from each well, which represents the average value of all the cells in the well. This becomes problematic particularly in cell-based assays because the cell populations that are used in assays often do not uniformly respond to the assay. The variation in the population in each well may be large enough to mislead the interpretation of assay results [37]. Not only overcome the limitation of single-variate methods, can access to multi-variate measurements also provide additional advantages [38]:

- One can find a best variable to select the hits after running the whole assay.
- More variables from the same samples can offer richer information to characterize the hits.
- Variables from each cell in the same well can offer multiple statistical representations other than mean (e.g. median, variation, percentage of cells above a threshold).

- Per-cell data may be useful in combination with other single-cell analysis methods in downstream.
- Certain variables may be useful to remove artifacts or unwanted cells (e.g. dead cells). Multivariated high-throughput screening is often referred to as high-content screening. This method will be discussed in the next section.

1.3 High-content screening

High-content screening (HCS) refers to a platform or technique by which more than one measurement are obtained from each well. In comparison with conventional HTS in which typically only one variable per well is obtained, HCS can offer richer information from the same cells simultaneously. In conventional single-variate HTS, the more variates one desires to obtain, the more wells one has to prepare, which dramatically increases the experimental costs, and thus in many cases, HTS is carried out with only a few or single drug concentration despite the sample number may be large [39,40]. To efficiently and quickly obtain a large number of variables, imaging modality is employed in HCS. This is why HCS normally refers to image-based screening methods. In the rest of this thesis, unless specially noted, HTS refers to single-variate screening, and HCS refers to image-based screening.

The first HCS was created in the late 1990s, in which each cell was labeled with multiple fluorescent reagents and analyzed at a subcellular resolution [41]. HCS was developed because HTS at that time could not fully meet the needs from pharmaceutical industry. This was because HTS can only deal with relatively simple assays such as mix-and-read in spite of the complexity of drug responses, and traditionally the development of HTS focused to improve the handling and reading speed of plates [42]. Indeed, despite increasing investment in research and development, the productivity of pharmaceutical industry had decreased by the time of early 2000s [43].

On the path of the development in HCS, there are two fundamental factors limiting the performance of HCS: imaging speed and the use of fluorescence. The former will be discussed in the next section, and hence we cover the limitation of fluorescence labeling in this section. As we discussed in the previous section, fluorescence detection is the most widely used method in HTS. This trend also applies to HCS because the HTS and HCS

share the same advantages. However, fluorescence labeling has several drawbacks. First of all, fluorescent probes are not available for all target molecules and may interfere with natural cellular functions [44]. While a wide range of immunofluorescent probes are commonly used in single-cell imaging for multivariate profiling, they are costly and require time-consuming labeling processes, including cell fixation which kills the cells, hindering large-scale assays [45]. Fluorescently tagged cell lines can offer live-cell assays without the labeling process, but the development of such cell lines requires more effort than immunofluorescent labeling [46]. Therefore, an alternative for image-based high-content screening without the need for fluorophores is clearly needed for easy manipulation and economical assays. To address this limitation, this thesis describes a study of investigating cellular drug responses with an image-based method but without using fluorescence labeling.

1.4 High-speed imaging modality in high-content screening

High-speed imaging is one of the key factors in image-based screening. Increase in frame rate can directly enhance the throughput of screening. Currently all commercially available instruments for image-based screening use charge-coupled device (CCD) or complementary metal–oxide–semiconductor (CMOS) (including EMCCD (electron multiplied CCD) and sCMOS (scientific CMOS)) cameras for image acquisition [47]. CCD or CMOS cameras consist of an array of photodiode sensors in which each sensor convert photons to electric signals which are later digitized and output as each pixel of an image. The CCD camera has a single digitizer to digitize the sequential signal from all the pixels, while the CMOS camera uses multiple digitizers on each sensor to speed up the charge download [48]. However, regardless of CMOS or CCD, a fundamental limitation in a sensor array that hampers high-speed image acquisition is the time required to transmit signals from each sensor. Therefore, the number of pixels in each image must be reduced in exchange for increase frame rates. The penalty is that pixel resolution has to be compromised at high frame rates. As the maximum frame rate of commercially available state-of-art sCMOS is 100 fps with full pixel number (2560×2160 pixels) [49], the throughput of all commercially available HCS instruments is limited by the frame rate of camera.

Another scheme of image acquisition uses beam scanning and a single-pixel photodetector. In this scheme, the target object is scanned point-by-point with a moving illumination beam, and the resultant optical signal is detected by the single-pixel photodetector. This type of imaging methods is suitable for sensitive detection where only small number of photons can reach the photodetector. Accordingly, this imaging scheme is employed in confocal microscopy in which spatial resolution can be improved but acquirable photon number is lower than those of wide-field microscopy. The drawback of this scheme is the low frame rate as the imaging speed is determined by galvanometric mirrors. The maximum scan rate of galvanometric mirrors is typically up to a few kHz in one direction, thus the frame rate for a 1000-by-1000 pixels image is at most a few frames per second (fps). Besides, the field of view (FOV) in confocal microscopy is limited so that only one or a few cells can fit in. This is a significant drawback in the context of HCS, thus the manufactures of HCS instruments employ spinning disk confocal microscopy to circumvent the low frame-rate problem. Spinning disk microscopy uses "spinning pin hole array" to achieve high spatial resolution as the conventional confocal microscopy does but acquires photons with a sensor array (i.e. CCD or CMOS) to raise imaging speed. However, the frame rate spinning disk microscopy is ultimately limited by the frame rate of the sensor array.

Time delay integration or TDI is another technique particularly used for CCD camera to improve its imaging speed by using a 2D image sensor as a 1D line-scanning sensor [50]. TDI is employed in the systems where the targets are moving linearly toward the same direction, and thus it has been utilized in both flow-cytometric or plate-wise HCS systems. When CCD transmit charges from photodiodes along each column as shown in Figure 1-1, TDI delays the timing of signal transfer and adds the charges from the photodiodes in the upper row to the next row synchronously with the moving target [51]. This technique allows the light from a given point consecutively expose on each pixel of the same column so that signal-to-noise ratio can be improved by virtually increase the exposure time without slowing the moving target. Despite this technique is useful in CCD cameras, it is incompatible with CMOS cameras because light is converted to digital signals at each pixel and no charge is passed to the photodiodes of next row [52].

The trade-off between pixel resolution and frame rate is a fundamental limitation. To overcome this limitation, an entirely new imaging modality has to be employed. In this thesis, we employed optofluidic time-stretch (OTS) microscopy to overcome this limitation. In short, this method takes the advantage of single-pixel photodetector that has higher sampling rate than that of sensor arrays. By changing the scanning mechanism from galvanometric mirror to an all-optical manner, the frame rate has dramatically increased. The principles and details will be discussed in the next chapter.

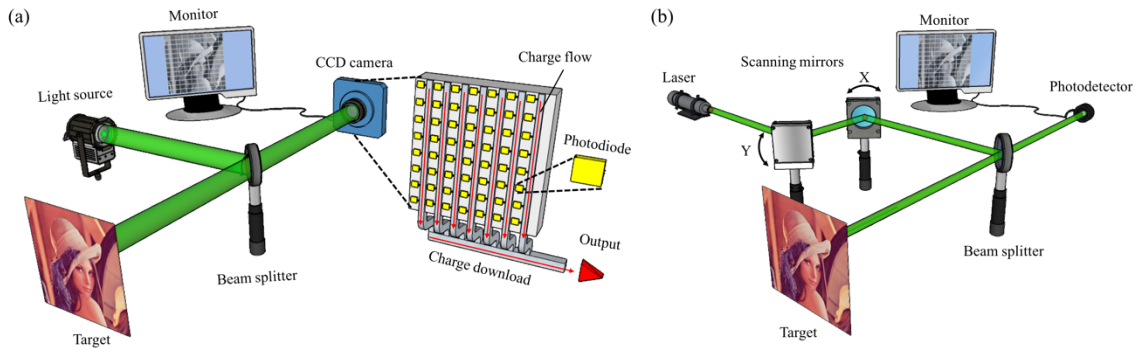

Figure 1-1 Principles of charge-coupled device (CCD) camera and point scanning imaging system.

1.5 Flow cytometry

Flow cytometry is a technology to investigate each cell in a suspension by detecting the scatter light and fluorescence of the cell while it is flowing in a stream of fluid. The two types of scatter light are used to assess cell size (forward scatter) and granularity (side scatter). Fluorescence is employed as indicators from the cell for biomarker detection. A short introduction of flow cytometry is given in this section to offer a better comprehension of this thesis as the new method developed in this thesis also falls in the category of flow cytometry.

In conventional flow cytometry, that is to say non-imaging flow cytometry, the signal collected from each cell is a waveform representing the total amount of light, whether it is scatter light or fluorescence. Three parameters can be extracted from the waveform: peak intensity (height), width, and the area under the curve (area). Among witch, width is used for calibration, height or area are used as the final readout, and thus only single variate is obtained for one observation per cell. Multivariate profiling can be done only by increasing the number of wavelengths of fluorescence. Although the currently commercially available

instruments can obtain more than 30 parameters simultaneously, labeling cells with more than 10 different wavelengths of fluorophores becomes extremely difficult for sample preparation due to heavy overlap in wavelength.

Numerous efforts have been made to apply flow cytometry to clinical diagnosis. Today, flow cytometry has been approved for clinical diagnosis especially for leukemia [53,54]. Flow cytometric methods offer some advantages in diagnosis over other genetic methods. First, flow cytometry provides relatively quick assays with low costs in contrast to most molecular methods [55]. Second, it is more robust in targeting marker when compared with genetic methods such as quantitative PCR [56]. For example, leukemic blast populations often contain multiple subpopulations which carries varying number of different mutations [57]. Moreover, chemotherapy can also change the landscape of mutations or even generate new mutations, making it difficult to target a specific marker [58,59].

Given the advantages of using flow cytometry for clinical diagnosis, I also explored the feasibility of applying the method developed in this thesis to more practical situations such as using whole blood and xenografts.

1.6 Scope and structure of thesis

This thesis explores the possibility of only using high-throughput bright-field imaging (i.e. label-free) to investigate drug responses in cancer cells. The work focuses on the investigation of dose-dependent morphological changes from which cellular drug responses can be interpreted. Not only demonstrating the proof-of-concept, does this thesis also explore the feasibility of using high-throughput bright-field imaging in various applications such as the investigation of cells spiked in undiluted whole blood and cells extracted from xenografts in mice.

Chapter 2 discusses the principles of the major components that were employed in my study including optofluidic time-stretch (OTS) microscopy, microfluidic design, image processing and machine learning. Chapter 3 describes the experimental demonstration of distinguishing cellular drug responses through morphological changes as a proof-of-concept. Specifically, I treated MCF-7, a human breast cancer cell with different concentrations of paclitaxel, a conventional anticancer drug to induce morphological changes on cells. While those morphological changes are too subtle to be observed by

human eyes, they were consistently detected on images through the method I developed in my thesis. Chapter 4 describes a challenging application of label-free detection of drug responses with whole blood. In this chapter, K562, a chronic myeloid leukemia cell line, and its drug resistant strain K562/ADM were treated with adriamycin, another conventional anticancer drug. which drug responses of cancer cells were evaluated as they were spiked in undiluted blood. The drug treated cells were further spiked into whole blood to mimic a practical situation in blood diagnosis, and their morphological changes were successfully captured by our system. Chapter 5 describes another challenging application where the drug responses of cancer cells harvested from mouse xenografts were evaluated. In this chapter, A549/EGFP and PC-9/mRuby (two lung cancer cell lines into which fluorescence proteins, mEGFP and mRuby2 were genetically inserted respectively) were generated to discriminate them from somatic murine cells. These cells were transplanted into athymic nude mice, and harvested after grown. The collected cells were further treated with gefitinib, an anticancer drug as targeted medicine for the evaluation of drug responses. Finally, Chapter 6 discusses the advantages and limitations in label-free detection of drug responses as well as possible future work.

Chapter 2

Label-free detection of drug responses of cancer cells: principles

2.1 Overview of optofluidic time-stretch microscopy

OTS microscope is the key experimental setup that allows us to acquire high-speed images of flowing cells. The whole schematic illustration is shown in Figure 2-1. In this section, an overview of OTS microscope will be briefly introduced. Detailed discussions about each key component will follow. The whole setup can be divided into two parts: bright-field imaging part (Figure 2-1a) and non-imaging fluorescence signal acquisition part (Figure 2-1b). The results described in Chapter 3 were acquired only with bright-field imaging part. In Chapter 4, fluorescence detection with 488 nm excitation was utilized to efficiently capture target cells that were spiked in whole blood. In Chapter 5, fluorescence detection with 405, 488 and 561 nm excitations were utilized to distinguish transplanted human cells from somatic murine cells as well as examine cell viability. Note that although fluorescence labeling was utilized, it did not influence the label-free analysis of cellular image. Fluorescence labeling in this thesis only works as “ground truth” indicator for the validation of the label-free analysis.

The bright-field imaging part of OTS microscope consists of multiple independently evolved technologies including optical time stretch, spatial-temporal mapping, and microfluidic focusing. Optical time stretch is the essential technology enabling high-speed imaging, and microfluidic focusing technology ensures cells are adequately align on the focal plane while offers high flow speed enabling high-throughput capability. These two technologies will be discussed before we get in to the detail of other components. The fluorescence detection part, on the other hand, only functions as a part of conventional flow cytometer, and will not be further discussed in this chapter.

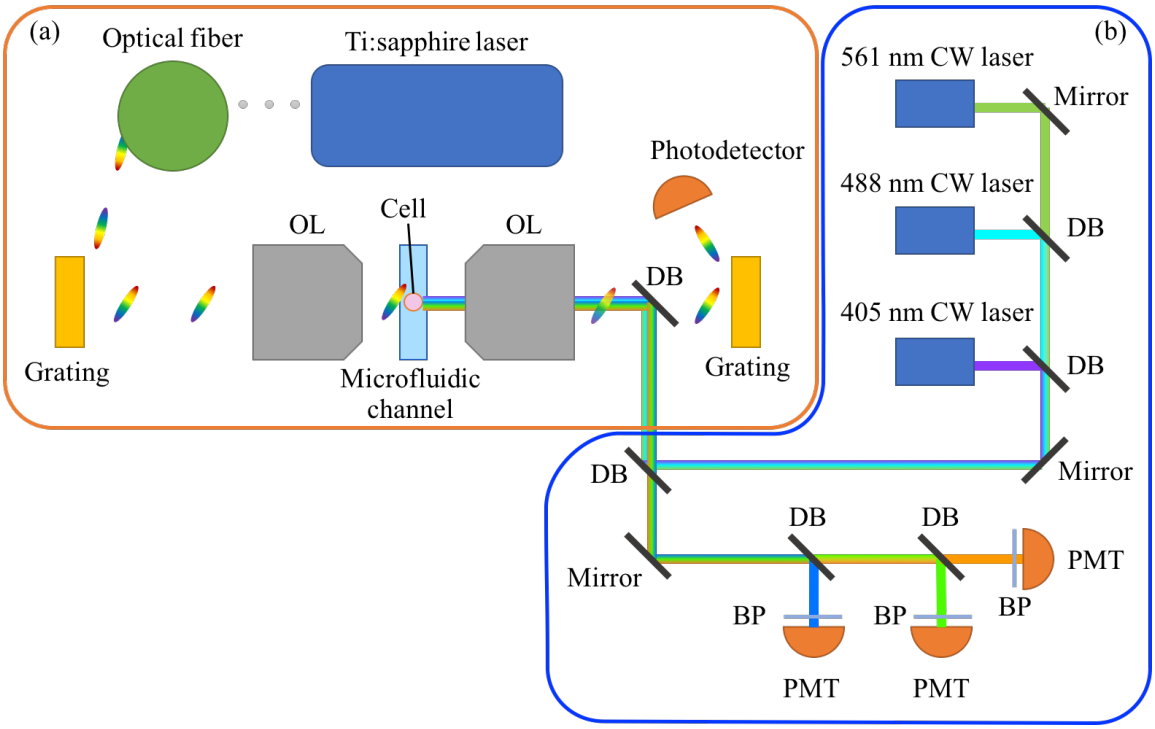

Figure 2-1. Schematic illustration of optofluidic time-stretch microscope utilized in this thesis. (a) bright-field imaging part, (b) non-imaging fluorescence excitation and detection part. OL: objective lens, DB: dichroic beam splitter, PMT: photomultiplier, CW: continuous wave, BP: band pass filter.

2.2 Principles of optical time stretch

Time stretch is a basic component of OTS microscope, and thus it is important to understand the working mechanism of time stretch as well as why time stretch has been employed. In this section, we introduce the principle and the motivation of using times stretch.

The term "time stretch" refers to a technique that stretches very fast optical signals in time domain before they are detected by an electric detector or digitized by an analog-to-digital converter (ADC). The time stretching is realized by a process called dispersive Fourier transformation (DFT) in which the spectrum of the pulse is mapped to a temporal waveform by the large group velocity dispersion[3] (GVD) in the dispersive element [60]. In the applications where an optical signal is at a very high speed, the ADC cannot precisely

[3] Group velocity dispersion refers to a characteristic of an optical element that the group velocity of light depends on the optical wavelength.

restore the waveform of optical signal to a digital signal due to the ambiguity in quantizing the amplitude and a jitter of the sampling clock (Figure 2-2). The former is caused by the gain bandwidth of the transistor, and the latter is caused by the phase noise of the sampling clock [61]. In such case, time stretch technique is useful to alleviate the speed limitation of ADC [62,63] making the signal acquisition closer to its optimum condition. In our OTS microscope, the original illumination light is a 30-fs pulse, and thus detecting the spatial information encoded in the pulse without using time-stretch technique is not possible at current technology.

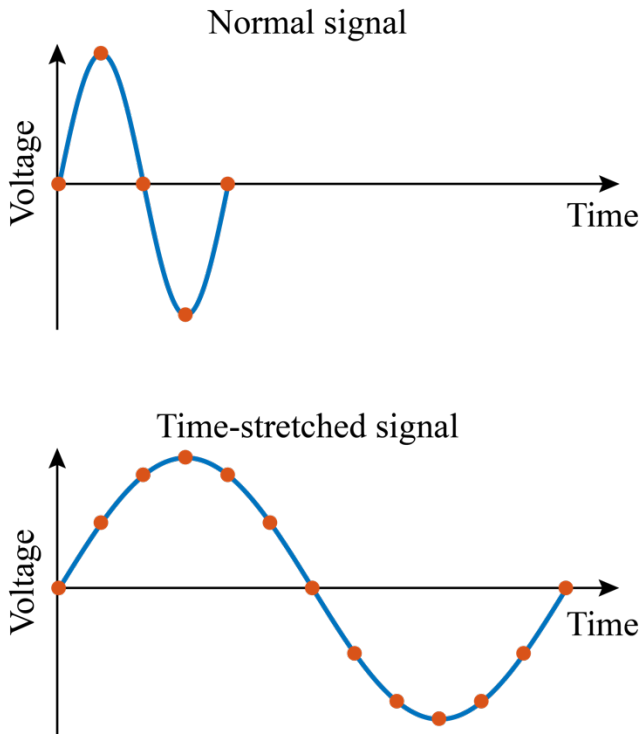

Figure 2-2. Using time-stretch technique can increase the number of sampling points with fixed sampling rate of ADC.

2.3 Microfluidic technology for optofluidic time-stretch (OTS) microscopy

Microfluidic device plays an important role in the OTS microscopy to 1) flow cells at a very high speed, 2) ensure a constant linear flow at a uniform velocity, and 3) physically focus and align cells in the optical focal plane and field-of-view (FOV) to ensure the image is in focus, and covers the whole cell. Since our OTS microscope does not have an active

scanning device, the imaging target has to be moving by itself. This is why the microfluidic device which controls the movement of cells directly affects the quality of the images.

1) The flow speed in our OTS microscope determines the imaging speed, or in other words frame rate. As we discussed in the previous section, imaging speed determines the throughput. Consequently, the throughput of our OTS microscope is determined by the flow speed.
2) A constant unidirectional flow at a uniform velocity is very important because cells are sequentially scanned while they are moving. If the flow direction or velocity randomly changes, every image will be distorted in a random way which cannot be compensated in the post-processing.
3) Physical focusing of the cells in the microfluidic channel also determines the optical focusing. As the cells flow at more than 10 m/s in our OTS microscope, it is impossible to mechanically auto-focus the objective lens to each cell. Therefore, each cell has to be physically or hydrodynamically focused in the optical focal plane as well as FOV so that every cell can be entirely imaged in focus.

In order to meet above-mentioned requirements, two types of microfluidic device have been used in imaging flow cytometry, i.e. hydrodynamic focusing (Figure 1-1a) and inertial focusing (Figure 1-1b). Hydrodynamic focusing technique [64] historically have been the mainstay in commercial flow cytometers for continuous cell positioning. In this technique, a slower flowing sample stream is encompassed by a faster flowing sheath stream from one or more sides. Once the sample stream and sheath stream begin flowing side-by-side and in the same direction in a laminar flow condition, the center stream (sample fluid) is surrounded by the second streams (sheath fluids) and thus cells are focused in the center. The cross-sectional area of the focused flows can be controlled by manipulating the relative flow rates of sample stream and sheath stream. Commonly, the sheath fluid is introduced at a far higher flow rate than that of sample fluid to occupy a larger portion of the channel, forcing the sample fluid into a small cross-sectional area in the center of the channel. On the way of this confinement, hydrodynamic focusing also aligns cells one-by-one into the core stream to prevent cells from overlapping each other as shown in Figure 1-1a. While this technique is relatively easy to implement, the use of a cell-free sheath flow with high flow rate will dilute the sample and limit the overall sample throughput.

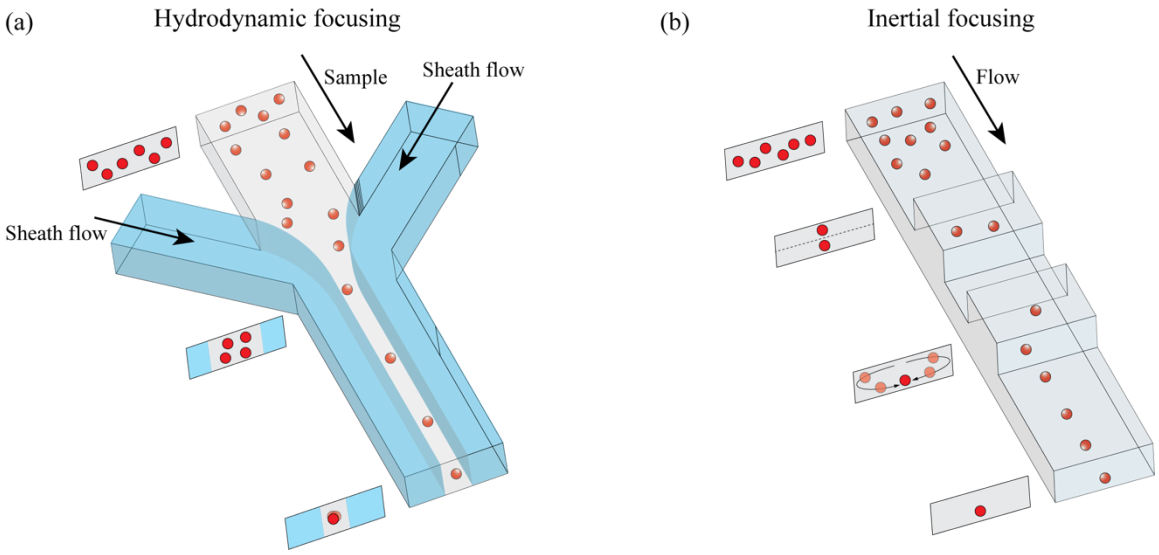

Figure 2-3. Schematic illustration and working principle of cell focusing in microfluidic devices. (a) Two-dimensional hydrodynamic focusing technique. (b) Two-dimensional inertial focusing technique

Inertial focusing [65] is a passive technique (i.e. no sheath stream required) for cell focusing in confined channel flows at moderate Reynolds number (~1 < Re < ~100) (Figure 1-1b). Reynolds number (Re) is a dimensionless quantity in fluid mechanics to estimate flow conditions. It expresses the ratio of inertial forces to viscous forces and is defined by

$$\mathrm{Re} = \frac{\rho u L}{\mu} \tag{1}$$

where ρ is the density of the fluid (kg/m^3), u is the velocity of the fluid (m/s), L is the characteristic length defined by the channel geometry, μ is the dynamic viscosity of the fluid (kg/m・s). In the inertial focusing channel, two inertial lift forces (i.e. shear-gradient lift and wall-effect lift forces) are formed perpendicular to the main flow direction. These forces make the cells migrate across streamlines, focus, and order deterministically at equilibrium positions between the centerline and walls of the channel. The number of equilibrium positions in a channel depends on the symmetry of the channel (e.g. a rectangular channel yields 2-4 equilibrium positions depending on Re). For instance, single-stream focusing of mammalian cells [66] and microalgae cells [67] have been demonstrated using a stepped channel, in which a low aspect ratio straight rectangular channel equipped with a series of steps in the channel upper wall induced local helical

secondary flows allowing cells to migrate to the focal plane as shown in Figure 1-1b. The focusing process begins with cell migration to the vertical center from random positions under the effect of inertial lift forces, and is followed by a lateral migration in response to the pair of helical secondary flows induced by the step. Finally, cells are directed to a single equilibrium position. Compared to hydrodynamic focusing, inertial focusing only requires a single stream of fluid, which does not suffer from sample dilution due to sheath fluid, and thus in this sense it can achieve higher throughput. However, it may be problematic when the sample is highly heterogeneous in size and shape because the inertial lift forces and resultant focal positions are dependent on these properties. While several applications in imaging flow cytometry using inertial focusing technique have been reported [68-71], we did not employ this technique in this thesis mainly due to the limitation in the flow speed. Inertial focusing works effectively only at moderate Reynolds number ($\sim 1 < Re < \sim 100$) whereas flow speed at more than 10 m/s usually exceeds this range of Reynolds number. In addition, the inherent drawback of hydrodynamic focusing that sample fluid is diluted by sheath fluid can actually benefit us in the application of using whole blood. This application will be described in Chapter 4.

2.4 Introduction of OTS microscopy

2.4.1 Working mechanism of OTS microscopy

Time-stretch imaging including optofluidic time-stretch microscopy is one of the most successful application of time-stretch technique and dispersive Fourier transformation (DFT). The optical time-stretch imaging is characterized by:

1) Generation of an illumination light with a spatially dispersed broadband pulse. (Figure 2-4a)
2) Optical mapping of the spatial information of imaging target into the optical spectrum as the process of imaging. (Figure 2-4b)
3) Conversion of the spatial information into a one-dimensional temporal data stream by optical time-stretch. (Figure 2-4c)
4) Optical detection of the data stream with a single-pixel photodetector. (Figure 2-4d)

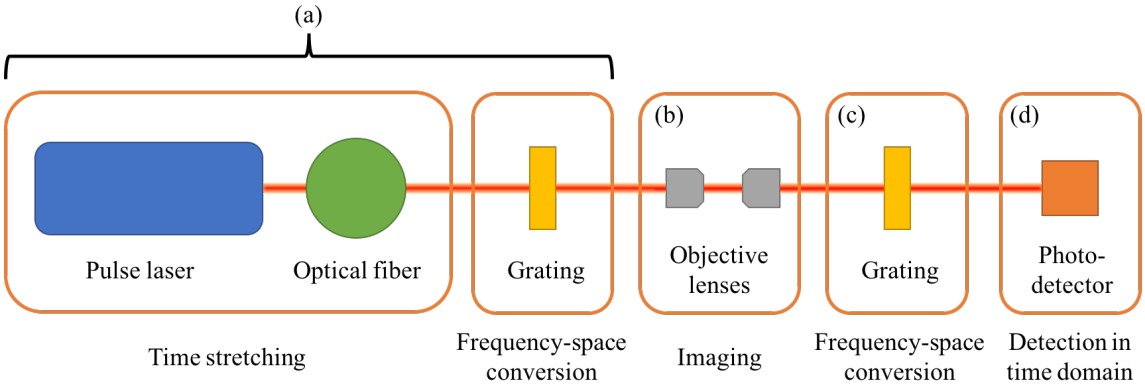

Figure 2-4. Schematic illustration of time-stretch imaging. (a) Generation of an illumination light with a spatially dispersed broadband pulse. (b) Optical mapping of the spatial information of imaging target into the optical spectrum as the process of imaging. (c) Conversion of the spatial information into a one-dimensional temporal data stream by optical time-stretch. (d) Optical detection of the data stream with a single-pixel photodetector.

The entire process is schematically shown in Figure 2-4. The whole system consists of five essential components including generation of time-stretched illumination light, time-space conversion of illumination light, imaging, time-space conversion of transmitted light, and photo detection in time domain. The pulse light generated from the femtosecond laser has a broadband in spectrum. While a femtosecond laser is not strictly required and a picosecond laser can also support OTS imaging [72,73], the spectral bandwidth becomes narrower when the pulse width is longer. The broadband pulse light is temporally stretched in the optical fiber as longer wavelength of the light transmits faster than the shorter wavelength part. For this reason, the optical fiber is also called dispersive fiber. As a result, the pulse light comes out from the dispersive fiber is spectrally aligned in the time domain as Figure 2-5 shows. This pulse is further diffracted by the diffraction grating to various angles depending on the wavelength. With this, spatially dispersed broadband pulse is formed and ready for imaging.

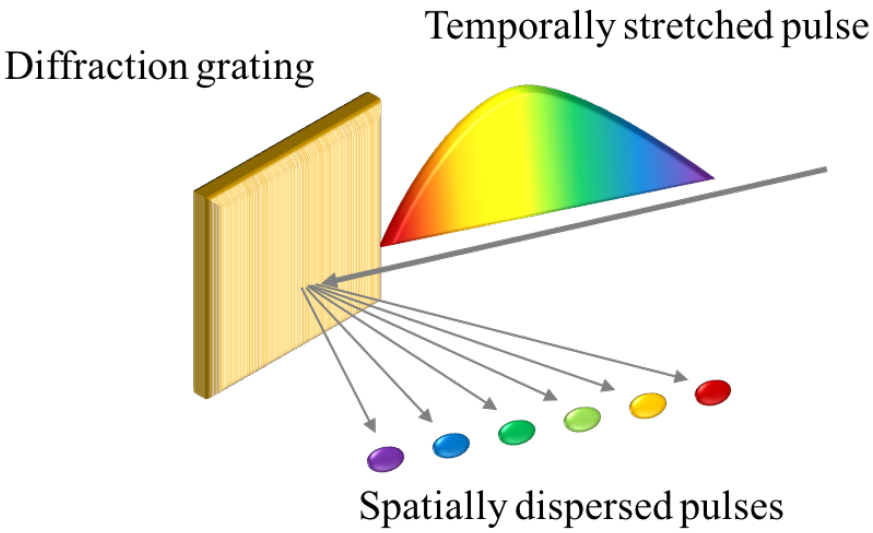

Figure 2-5. Frequency-space conversion at a diffraction grating. The temporally stretched broadband pulse is mapped into space domain.

After the diffraction grating, spatially dispersed pulse is focused at the imaging target through an objective lens. Each spectral component is focused at different locations as shown in Figure 2-6. The spatial information of the imaging target is mapped into the spectrum of the illumination light as it passes through the target. This transmitted light is collected with another objective lens and converted from space domain back to time domain for the detection at the single pixel photodetector (Figure 2-7).

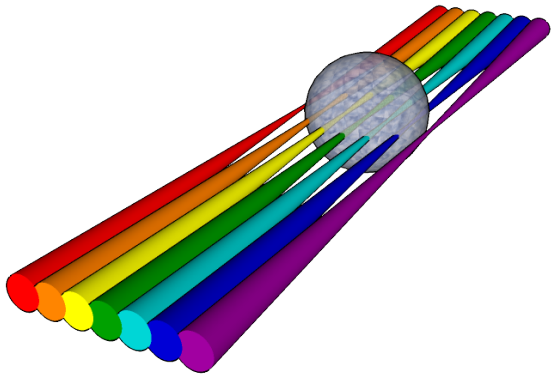

Figure 2-6. Imaging of target by mapping the spatial information into the spectrum of the illumination light.

Optically, the second time-space conversion is identical to the first one, except that the dispersed light now bears the spatial information of the target. The two sets of time-space conversion work as point-scanning system in which spatial information is sequentially scanned and collected by a single-pixel photodetector. However, in comparison with conventional mechanical point-scanning system, OTS imaging scans target with a passive optics system which can achieve exceptionally higher scan rate. In our OTS microscope, the broadband pulse is stretched to approximately 9.2 ns in which about 4 ns is used for illuminating the FOV. This means that a single line scan is performed in 4 ns which is impossible with mechanical devices. The scan rate for each line is determined by the repetition rate of the pulse laser. In our system, the repetition rate of our femtosecond laser is 75 MHz, which makes the line scan every 13 ns. This is also not possible with a mechanical scanner.

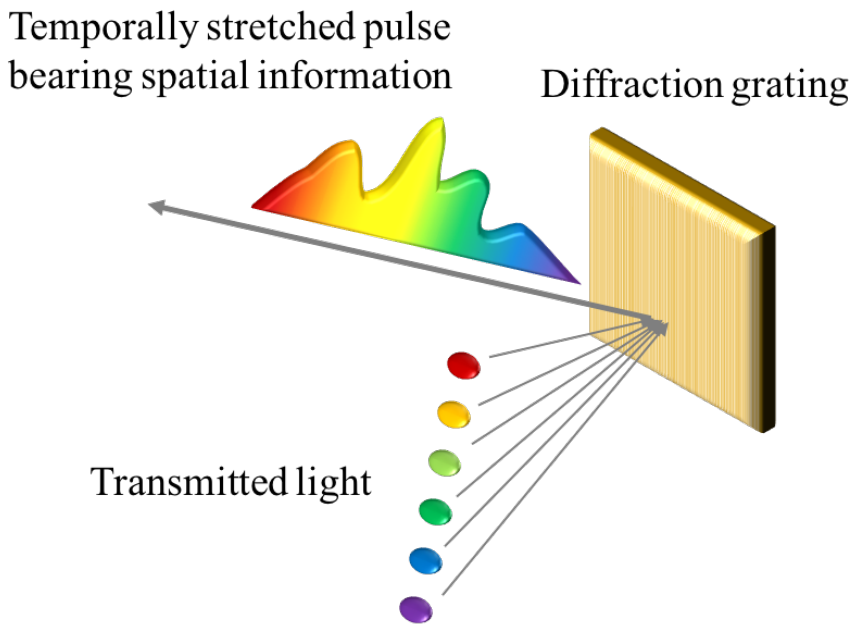

Figure 2-7. Frequency-space conversion at a diffraction grating. The transmitted light from sample is converted back to a single stream of stretched pulse, but now bearing the spatial information of the target.

The recorded raw data of OTS imaging is a continuous waveform. To construct a two-dimensional image out of the raw data, signal processing has to be done by digitally stacking each line as shown in Figure 2-8 [74]. The number of pixel in lateral direction is determined by the sampling rate of ADC. For example, in our system, each lateral line is measured in approximately 4 ns with a sampling rate of 50 GS/s (Giga samples per second). This setting produces about 200 pixels in each lateral line. Given that the FOV in lateral direction is approximately 80 μm, the pixel resolution is about 400 nm/pixel. This is slightly lower than the diffraction limit (~660 nm) of our system where the actual spatial resolution is 780nm [75]. The number of pixel in longitudinal direction is determined by the recording time in our OTS microscope because the imaging target is scanned as the target moves through illumination area. In principle, the image can be as long as the memory in the oscilloscope can afford. The pixel resolution depends on the repetition rate of the pulse and how fast the target moves. Given the repetition rate of 75 MHz and the flow speed of 10 m/s for a cell in our system, the cell moves 133 nm between each line scan. This is about 1/5 of the theoretical diffraction limit and 1/6 of the actual spatial

resolution, which indicates that our system is capable of acquiring cellular image moving (flowing) at 50-60 m/s with diffraction-limited resolution.

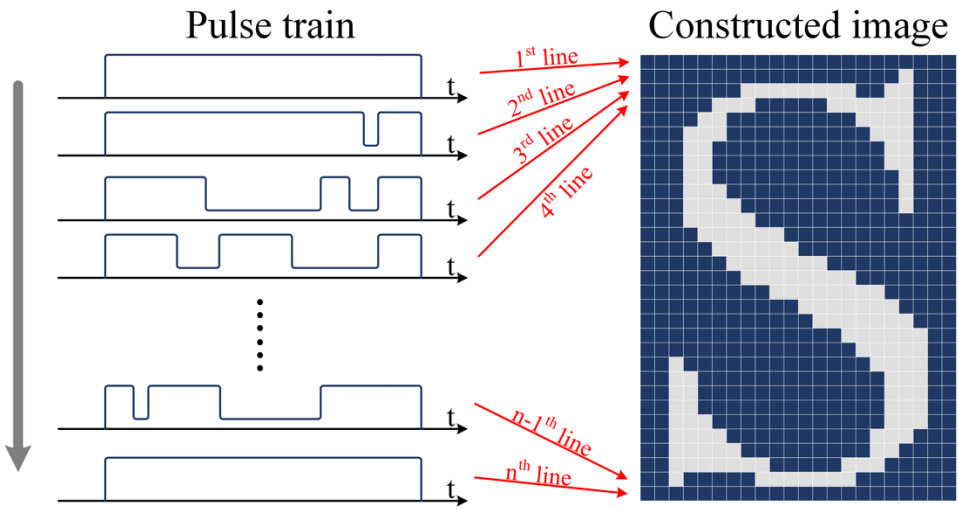

Figure 2-8. Image construction in OTS imaging. Each pulse is digitally stacked to create a two-dimensional image.

2.4.2 Specification of OTS microscope used in this thesis

The specification of OTS microscope used in this thesis is descried in the following table.

Table 2–1. Specifications of components of the OTS microscope used in this thesis.

Item	Value
Center wavelength of the laser	790 nm
Bandwidth of pulse	40 nm
Repetition rate of the laser	75 MHz
Average output power of the laser	650-700 mW
Group-velocity dispersion of the optical fiber	-120 ps/nm/km
Length of the optical fiber	2 km
Groove density of diffraction gratings	1200 lines/mm
Magnification of the objective lenses	40
Numerical aperture of the objective lenses	0.6
Detection bandwidth of the photodetector	12 GHz
Detection bandwidth of the oscilloscope	16 GHz

Sampling rate of the oscilloscope	50 GS/s

2.5 Theory of image feature extraction

In order to perform single-cell analysis, the acquired images need to be transformed into a numerical feature space that preserves relevant information of the cell. Although each pixel in the image can be considered as a numerical feature, the pixels only contain local information and therefore a very complex and nonlinear model [76] is required to capture the geometry or high-level cellular features from the pixel level. And complex nonlinear models, such as convolutional neural networks (CNNs) [77], are generally more difficult to train since they have a large number of trainable parameters (require more training samples) and are more likely to overfit on training data. For this reason, feature extraction is recommended in this protocol. If the goal is to extract features that are interpretable and have biological meaning, CellProfiler, an open-source software for cellular image analysis, can be employed to extract human-interpretable numerical features [78,79]. On the other hand, if interpretability is not the primary concern, then CNNs pre-trained on natural images, such as the VGG net [80] or ResNet [81], can be used as a feature extractor [82].

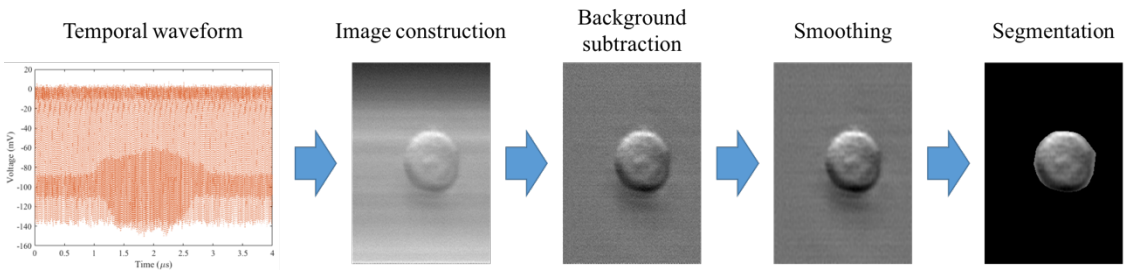

Figure 2-9. The process of image processing and segmentation.

In this thesis, we used CellProfiler for feature extraction. CellProfiler is convenient for feature extraction but limited in the functionality of segmentation. Due to the discrepancy in the mechanism of image generation in optofluidic time-stretch microscopy and conventional cameras, some types of noises (e.g. horizontal lines) that can be often seen on time-stretch images may not be frequently appear in the images taken by conventional cameras. For this reason, in order to perform more accurate and robust segmentation to all images, cell segmentation is performed on Matlab [83,84]. The quality of segmentation depends on the image quality. Noise reduction prior to segmentation can

improve the accuracy of segmentation. Segmentation may not be properly completed when multiple cells are captured in the same image or the cell is not completely captured in the field of view. In such case, these images were discard those images.

2.6 Theory of support vector machine classification

In this thesis, we propose and demonstrate a method using multivariate data to visualize minute changes expressed in the cellular morphology which were induced by drug treatment. We employed machine learning technique (particularly supervised learning) to analyze multivariate data rather than using multivariate analysis because we focus on extracting a predictive model from data, which will be useful for screening and diagnostic application. Nevertheless, our method is certainly compatible with any type of multivariate analysis such as linear discriminant analysis (LDA) [85]. In fact, we employed maximum mean discrepancy (MMD) as an alternative approach to validate the analysis based on machine learning techniques. In this thesis, we mainly employed support vector machine or SVM, which is a type of machine learning technique used for classification or regression by setting up a hyperplane to separate two populations.

The working mechanism of SVM can be described as following:

1) SVM maps the input data into a high dimensional feature space through a mapping operation called kernel methods [86]. In this space, two populations are expected to be linearly "separable." (Figure 2-10a)
2) Among the infinite possible hyperplanes (Figure 2-10b), SVM sets up a unique one by maximizing the margin between the nearest sample points (Figure 2-10c).
3) If two populations are nonseparable, penalty is introduced according to the distance between the hyperplane and each sample. A unique hyperplane is determined by minimizing the summation of penalty (Figure 2-10). The margin defined in such way is called soft margin.

While a number of derivations of SVM have been developed to extend its functionality, in this thesis, we only perform binary classification which is the most fundamental function of SVM. The classifier we choose is the soft-margin support vector machine, which aims to find large-margin separating hyperplane [87]

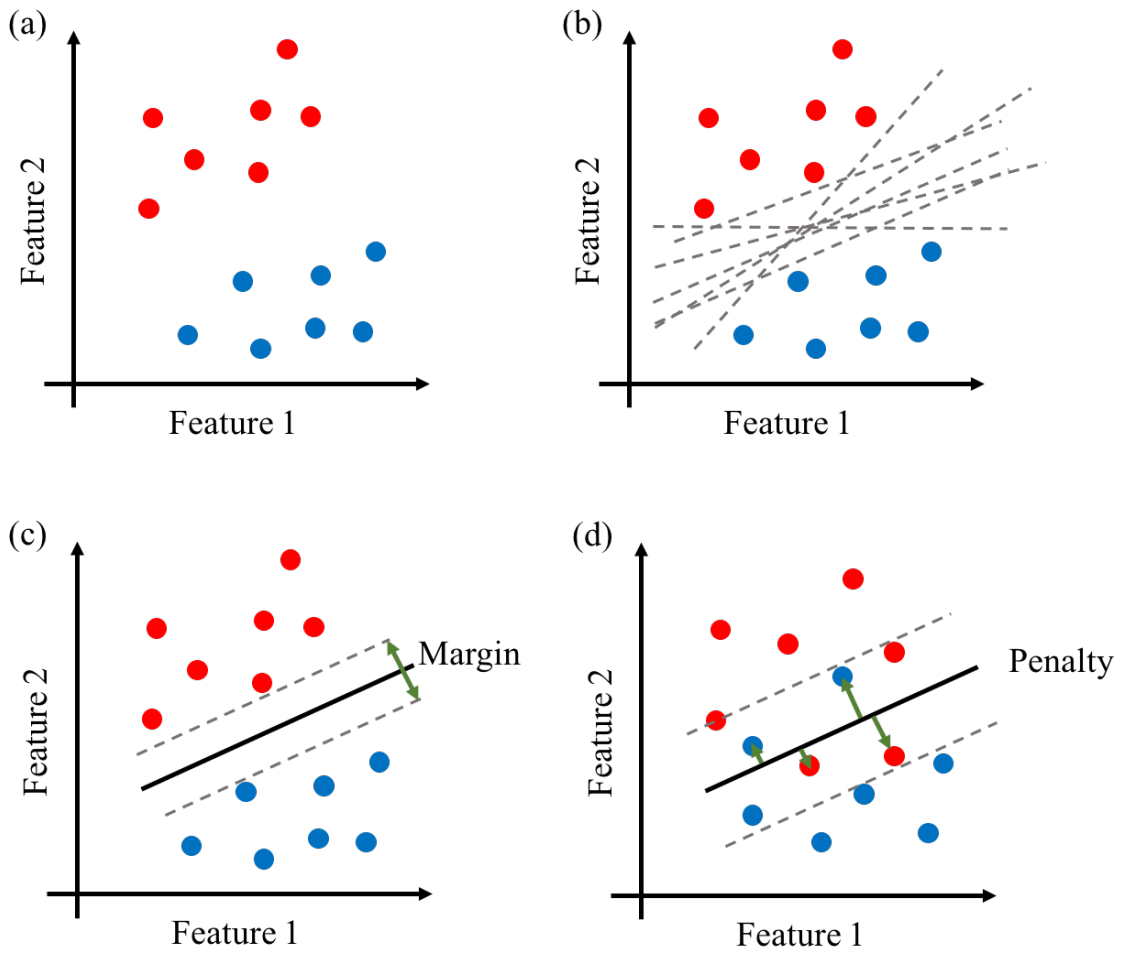

Figure 2-10. Illustration of how SVM classifier sets up a hyperplane to separate two populations. (a) Data points are distributed to a high dimensional feature space. (b) Multiple hyperplanes can be set to separate two populations. (c) SVM sets up a unique hyperplane in the feature space by maximizing the margin between the nearest samples. (d) In the cases two populations are nonseparable, SVM classifier assigns penalty to each sample point according to the distance to the hyperplane. SVM classifier sets up the hyperplane by minimizing the sum of penalties.

Chapter 3

Label-free detection of drug responses of cancer cells: experimental demonstration

3.1 Introduction

As we discussed in the Chapter 1, label-free method has been desired to overcome the disadvantages that fluorescence techniques have posed. In this chapter, we describe the experimental demonstration of using optofluidic time-stretch (OTS) microscopy to probe drug responses in a label-free manner. While a few recent researches demonstrated that bright-field images in combination with dark-field [45] or phase images [88] can be used to distinguish different types of cells or cells in different cell cycles, the evaluation of drug responses via solely bright-field images has yet to be investigated. Here, we present a method for evaluating cellular drug responses only by high-throughput bright-field imaging with the aid of machine learning. We also quantitatively analyzed morphological change of cells which inferred from SVM classification accuracy, suggesting that the morphological change observed from bright-field images can be utilized as an indicator for drug discovery.

The workflow can be separated as four parts as illustrated in Figure 3-1. Specifically, Drug-treated and -untreated cells were imaged by the OTS microscope, followed by feature extraction on computer. The extracted numerical features were then subject to binary classification between drug-treated and -untreated populations. Higher classification accuracy obtained from the classification suggests that more morphological changes were captured through images and the feature extraction.

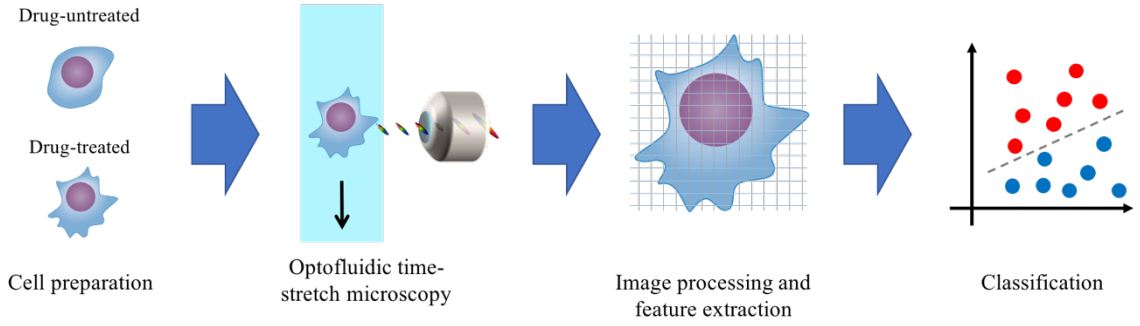

Figure 3-1. Workflow of label-free detection of drug responses of cancer cells. Drug-treated and -untreated cells were imaged by the optofluidic time-stretch microscope. Then, the morphological features were extracted from the acquired images, based on which the two types of cells were classified into two groups.

3.2 Materials and methods

3.2.1 Cell culture

MCF-7 (DS Pharma Biomedical), a human breast cancer cell line, was employed as model cells. The cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Wako Chemicals) supplemented with 10% fetal bovine serum (MP Biomedicals) and 1% penicillin streptomycin (Wako Chemicals), at 37°C in 5% CO_2. One day prior to drug treatment, cells were seeded to a 12-well plate at 10^5 cells/mL with 1mL/well culture medium.

3.2.2 Drug treatment

Paclitaxel, an FDA-approved and classic anti-cancer drug, was employed as a model drug. Paclitaxel (Cayman Chemical) in a powder form was dissolved in dimethyl sulfoxide (DMSO, Wako Chemicals) to make a concentration of 1 mM as a stock solution. Paclitaxel was serially diluted with culture medium in 10-fold and applied to the cells to make the final concentration ranging from 1 nM to 10 μM including negative controls in which no paclitaxel was added.

3.2.3 Sample preparation

Paclitaxel-treated cells were harvested, and suspended in the culture medium by trypsinization at two intervals (12, 24 hours). The final concentration of cell suspension was adjusted to approximately 10^5 – 10^6 cells/mL to ensure reliable single-cell image

acquisition in each image frame. The cell suspensions were filtered with 30-μm nylon mesh (Sysmex) to break up aggregated cells and remove debris before taken up by 1-mL syringes and ready for image acquisition with the OTS microscope.

3.2.4 Microfluidic device fabrication

The structure of microfluidic device was designed on AutoCAD, a CAD (computer aided-design) software. The channel design is shown in Figure 3-1. The design was printed on a transparent film as a photomask. The whole fabrication process is described in Figure 3-2. Specifically, KMPR1035 (MicroChem), a negative photoresist, was spin-coated on a silicon wafer. Spin-coated silicon wafer was soft baked at 100 °C for 15 min, then cooled down to room temperature. Spin-coated wafer was then exposed to UV light for 60 sec, followed by another bake at 100 °C for 3 min. After cooled down to room temperature, photoresist was then developed by submerging the wafer in SU-8 developer (Nippon Kayaku) to completely remove unexposed photoresist. After removing residual SU-8 developer by rinsing with isopropanol and deionized water, the photoresist on the wafer can be used as a master mold. A 10:1 mix of PDMS (polydimethylsiloxane) base and curing agent (Dow Corning) was poured onto the master mold followed by degas to purge trapped air. PDMS was cured at 80 °C for 1 hour. The cured PDMS was removed from the master mold, and punched holes at inlets and outlet. Trimmed PDMS and a slide glass were treated with O2 plasma for 3 min to form irreversible bonding between them. A microfluidic device is completed after PEEK (Polyetheretherketone) tubes were inserted in inlets and outlet. The final dimension of the micro channel at the observation are was 100 μm wide and 44 μm high.

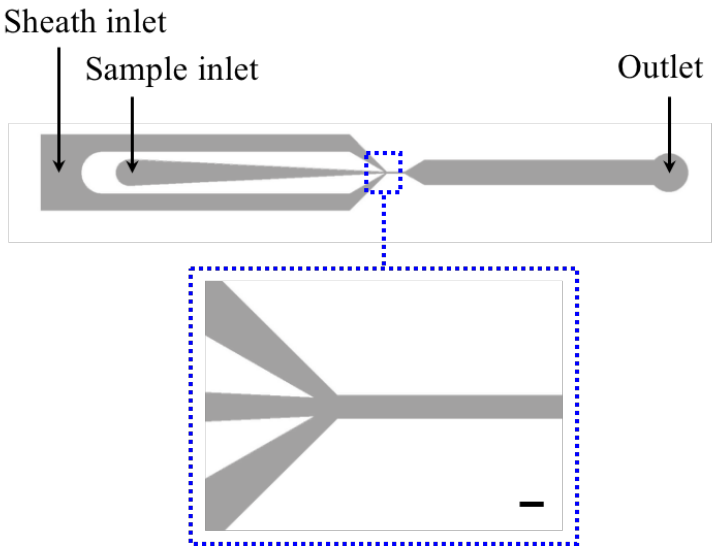

Figure 3-2. Design of microfluidic channel. Water flows in the dark area. Scale bar: 100μm.

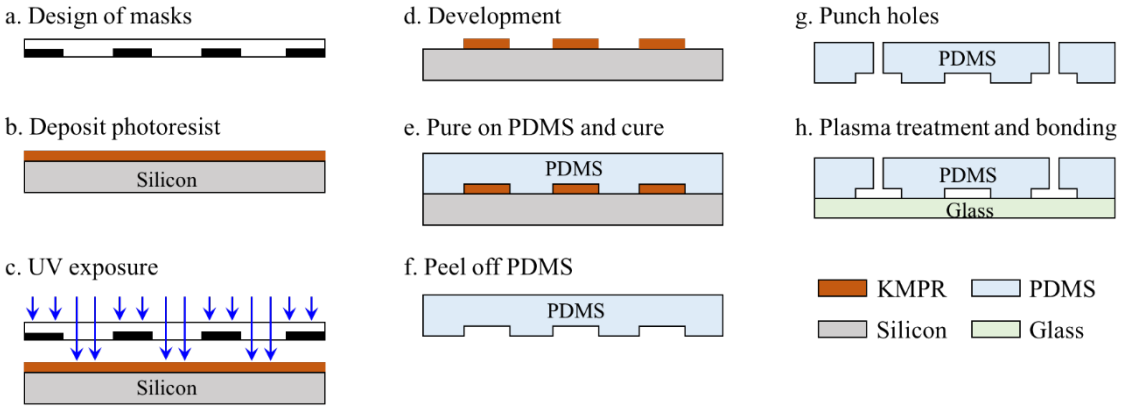

Figure 3-3. Process of microfluidic device fabrication. (a) Print the channel design on a transparent film as a photomask. (b) Photoresist is spin-coated on a silicon wafer. (c) Photoresist coated silicon is exposed to UV light. (d) Photoresist is then developed by SU-8 developer to remove unexposed photoresist. (e) PDMS mixture is poured on top of silicon wafer and cured by baking. (f) PDMS is peeled off and cut into a device shape. (g) Inlets and outlet holes are punched. (h) PDMS and a slide glass were bonded by a plasma treatment.

3.2.5 OTS microscopic imaging

The principle and instrumentation of optofluidic time-stretch microscopy is described in Chapter 2. In the experiment described in this chapter, a flow rate of 2.75 mL/min was achieved resulting a flow speed of approximately 10 m/s. Consequently, each frame was acquired in 4 μs corresponding to 250,000 frames/s. That is to say, the theoretical

maximum throughput is 250,000 cells/s. However, due to a low concentration of cell suspension, the actual throughput would lie between 10,000 and 100,000 cells/s. All images were formatted to 192 × 192 pixels for downstream analysis.

3.2.6 Conventional bright-field microscopic imaging

For a comparison between time-stretch images and conventional bright-field microscopic images, a conventional inverted microscope (Olympus) and a CMOS camera (DCC1545M, Thorlabs) were used to acquire static images of suspended MCF-7 cells. Cells were prepared following the procedure described in 3.2.1 through 3.2.3. Each static image was cropped to 200 × 200 pixels.

3.3 Results and discussion

3.3.1 Optofluidic time-stretch imaging

To underscore the imaging quality of optofluidic time-stretch microscope in taking bright-field images of flowing cells at a very high speed, two image libraries of drug-treated and -untreated MCF-7 cells that were taken under a conventional microscope and our optofluidic time-stretch microscope is shown in Figure 3-4. The fine structures in the cells can be clearly seen in the optofluidic time-stretch images with sufficient contrast, at an equivalent level in the conventional bright-field microscopic images. In addition to the imaging quality, both static and optofluidic time-stretch images (i.e. cells are in the flow) demonstrate consistent cellular morphology, indicating no significant morphological changes are induced by the flow. It is noteworthy that all blur-free optofluidic time-stretch image were acquired in 4 μs (corresponding to a frame rate of 250,000 frames/s), while maintaining the same imaging quality and pixel resolution as those taken by a CMOS camera on a conventional bright-field microscope. In contrast to most of the conventional high-speed cameras such as CCD or sCMOS where pixel resolution will be deteriorated when used at such high frame rate, the high pixel resolution of our optofluidic time-stretch microscope retain as much cellular information as conventional bright-field images do. The high frame rate of our system also enables high-throughput imaging, which favors the application of machine learning techniques to mine the cellular information in the images since these techniques often require a large number of training data for creating an accurate

prediction model. With the above-mentioned advantages of optofluidic time-stretch microscope, we evaluated the cellular drug responses only by bright-field imaging, which had been challenging otherwise.

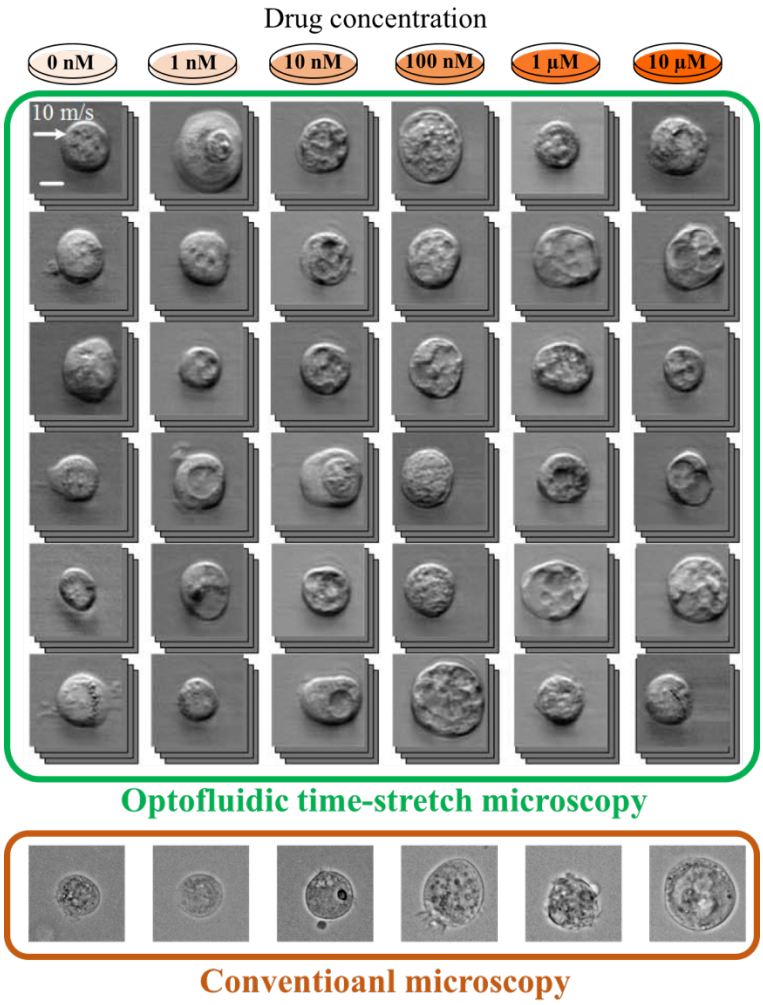

Figure 3-4. Image libraries of drug-treated and -untreated MCF-7 cells under optofluidic time-stretch microscope (flowing at a speed of 10 m/s) and conventional microscope (static). Despite the high flow speed, the optofluidic time-stretch microscope acquires blur-free cellular images with equivalent image quality as the static images obtained by a conventional microscope. Scale bar: 10 μm.

3.3.2 Classification of drug-treated and -untreated cells

To quantitatively evaluate the impact of an anti-cancer drug concentration on the morphological change of cancer cells, we treated MCF-7 cells with various concentrations of paclitaxel for 24 hours. We performed binary SVM classification between the negative control and each drug-treated population. Figure 3-5a shows the distribution of

classification scores between the two populations. As we use a linear kernel for the SVM classification, the classification score is given by

$$Y = \boldsymbol{w} \cdot \boldsymbol{x} + b \tag{2}$$

where w is the normal vector to the hyperplane, representing the weight assigned to each feature, x is the test data, and b is the bias. The separation between the two populations becomes larger as the drug concentration increases up to 1 μM, indicating that features corresponding to each class (drug-treated or negative control) become more distinct. Given that extracted features correspond to morphological change, the SVM classification suggests a dose-dependent drug-induced morphological change. This dose-dependent change was further supported by four trials of SVM classification as shown in Fig. 3b, where higher classification accuracy equates to larger separation between the two populations. The classification accuracy here is defined by $A = (X_1 + X_2)/N$, where X_1 and X_2 are the are the numbers of correctly assigned incidences and N is the total number of test data points. The accuracy range is from 50% (random) to 100% (perfect). We note that classification accuracy reaches its maximum at 1 μM for 24-hour drug-treated cells and then drops at 10 μM. The drop at 10 μM is possibly due to the presence of the high concentration of DMSO (1% v/v). Next, we evaluated the impact of drug-treatment time on morphological change. We performed the same dose-ranging experiment with a 12-hour drug treatment whose classification accuracy evolution is shown in Figure 3-5b. The classification accuracy curve shows a relatively monotonic increase, but lower accuracy at each drug concentration than that of the cells treated for 24 hours. This result illustrates that shorter treatment time induces less discrepancy in the features of each class, hence making it more difficult to identify drug-induced morphological change. Accordingly, we use the data acquired from the 24-hour treatment on the quantitative analysis shown below.

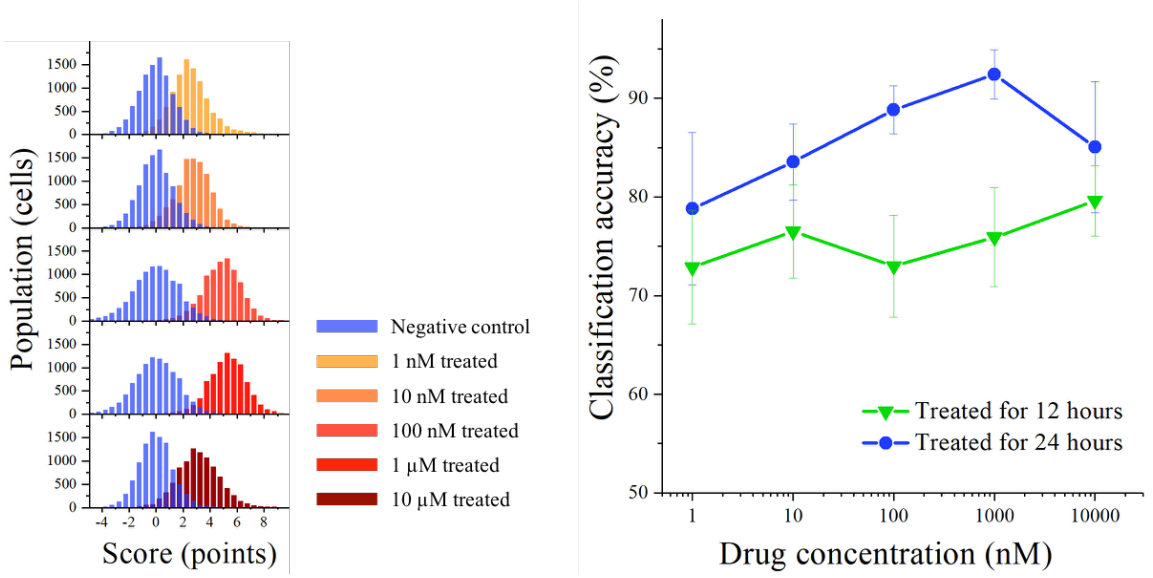

Figure 3-5. Classification of drug-treated and -untreated cancer cells. (a) Histograms of SVM classification scores for MCF-7 cells treated with various concentrations of paclitaxel for 24 hours. Each population consists of up to 10,000 cells. (b) Classification accuracy at various drug concentrations and incubation times. The error bars represent standard errors of the cross-validation estimation of average classification accuracy (n = 4).

3.3.3 Dose-dependent changes in feature space

In order to analyze the amount and type of morphological changes in the feature space, we calculated the maximum mean discrepancy (MMD) between the populations of drug-treated and -untreated cells at each drug concentration. Here the MMD represents the distance between the mean embeddings of distributions in a reproducing kernel Hilbert space (RKHS), in our case, defined by a Gaussian kernel11 (Figure 3-6). If the extent of morphological change is reflected in the extracted numerical features, then one would expect a larger MMD score at concentrations with greater change in morphology, since the two classes should be more distinguishable. Figure 3-7 shows the change in the MMD against drug concentration for two experimental trials (see Methods), in which the pillars for both experimental results are similar. The trend of the MMD shown in Figure 3-7 is also consistent with the trend of classification accuracy in Figure 3-5b, demonstrating that the measured dose dependence in the SVM results is supported by the MMD score of the feature space. Note that in contrast to the histograms shown in Figure 3-5a, in which the separation of two classes is made possible by supervised learning with respect to class

labels, the MMD (the distance between the distributions of two classes) is computed in a closed form and therefore does not require supervised learning. As a larger MMD indicates a larger morphological change present in the distribution, it is more likely that the SVM model at the drug concentration giving the largest MMD assigns large weights to features that reflect this drug-induced morphological change. Accordingly, we also quantitatively analyzed the MMD for each feature at the drug concentrations in which the overall MMD between the two class distributions is largest (1 μM of the first experiment and 100 nM of the second experiment as shown in Figure 3-7). We computed the MMD of each feature to examine the variation between the two experiments. Figure 3-8 shows that the features giving larger MMD scores are highly correlated between the two experiments, indicating that the significant features are consistent in both experiments. It is also observed in Figure 3-8 that features with large MMD score represent various types of information of cell images, such as geometry, granularity, intensity, and texture, indicating that the multivariate data provided by single-cell images is effective for identifying the cellular response to the drug.

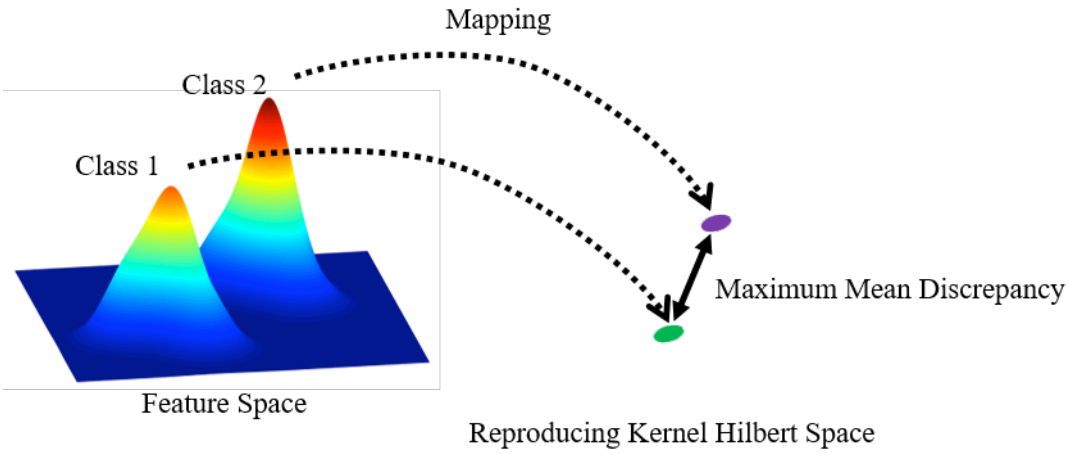

Figure 3-6. Calculating maximum mean discrepancy (MMD) between the negative control and drug-treated cell population. Illustration of the MMD.

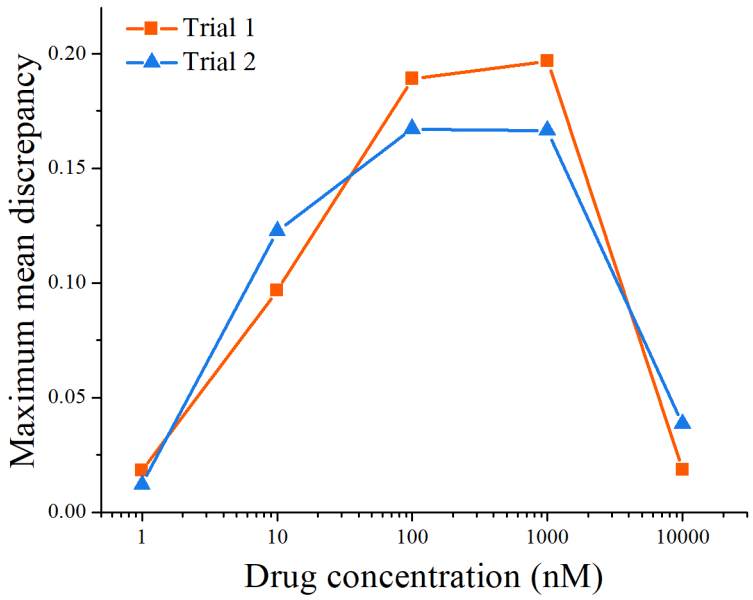

Figure 3-7. MMD between the negative control and drug-treated cell population at each drug concentration. Trial 1: data from the first experiment. Trial 2: data from the second experiment.

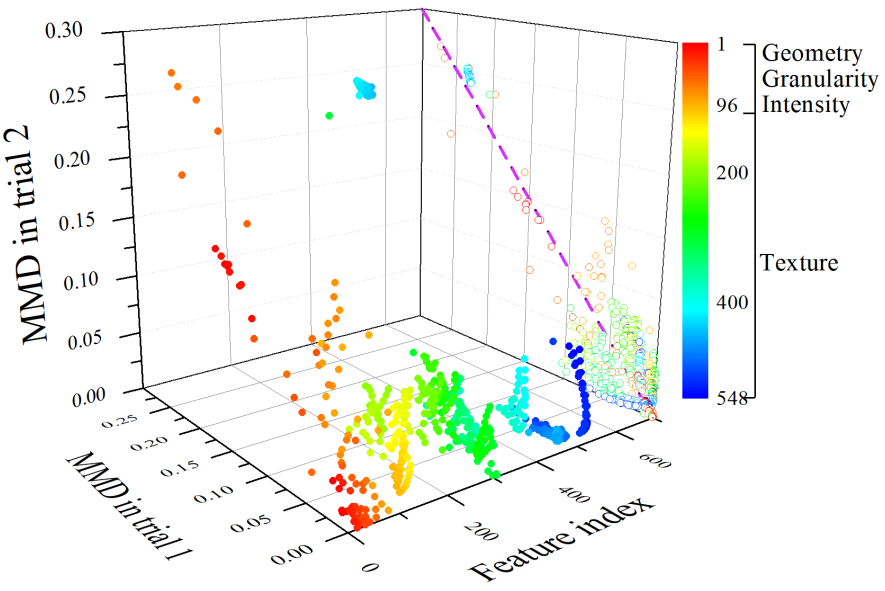

Figure 3-8. MMD of each feature in trial 1 at 1 μM and trial 2 at 100 nM. At these concentrations, the MMD in the whole feature space is the largest in each experiment. Features with a higher score of the MMD in both trials are highly correlated, indicating that the significant features were consistent in both experiments. The color scale represents feature index, showing types of morphological changes that undergo larger scores of the MMD.

We further investigated how the number of features contributes to the classification accuracy. We iteratively performed SVM classification between negative control and the data at 1 μM in the first experiment by removing the features with lower MMD (Figure 3-9). Consequently, we found that an accuracy of 90% (dotted line in Figure 3-9) was maintained when more than 400 features were removed, suggesting that with approximately 100 features, the performance of our label-free method is comparable to that obtained with fluorescence imaging techniques [89]. This property is noteworthy because fluorescence imaging requires fluorescent labeling with several inherent drawbacks as mentioned above. In other words, the high specificity of fluorescence imaging can also be provided by the combination of high-throughput bright-field imaging and machine learning, which is highly beneficial for pharmaceutical industry in which the cost of drug discovery is one of the major limiting factors.

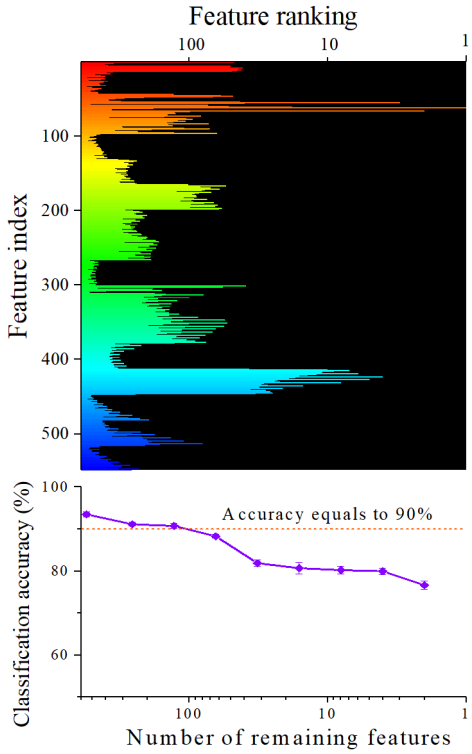

Figure 3-9. Classification accuracy dependence on the number of feature. (a) Feature ranking of trial 1 based on MMD (b) Classification accuracy with a reduced number of features. Lower MMD features were removed based on the ranking of the MMD for each feature in the classification between the negative control and the dataset at 1 μM in the first experiment (top). The classification accuracy was

maintained over 90% with more than 100 features (bottom). The color scale is consistent with that in Figure 3-8. Feature ranking and the number of remaining features are illustrated in logarithmic scale. The error bars represent standard errors of the cross-validation estimation of average classification accuracy (n = 10).

3.3.4 Cell classification of different experimental trials with a single model

We tested whether a single SVM model with a linear kernel can represent the dose dependence through the classification accuracy at all concentrations. We applied the SVM model trained at one specific concentration to all the other concentrations, and show the result in Figure 3-10a and Figure 3-10b. The result shows that the SVM models of those concentrations with larger MMD between two class distributions (such as 100 nM and 1 μM) can better preserve the tendency in classification accuracy as shown in Figure 3-5b, whereas the SVM models of those concentrations with lower MMD (such as 1 nM) fail to demonstrate such trend. Therefore, it is reasonable to conclude that our approach can provide a single classification model exhibiting the dose dependence of drug-induced morphological change. In addition, we further tested whether the single SVM models can be applied to different experimental trials. Specifically, we applied the model trained at the concentration with the largest MMD in the first experiment to the dataset of the second experiment, and vice versa (Figure 3-11). For the sake of comparison, the classifications in which both training and testing data are from the same trial of experiment are also included in the Figure 3-11. The result shows a consistent tendency in classification accuracy regardless of whether the training and testing data are from different trials of experiment, suggesting that the SVM models trained at the concentration giving the largest MMD can demonstrate the dose dependence across multiple experiments. This is a significant property in comparison with most of the previous work where a new training is required for a new dataset of images [90].

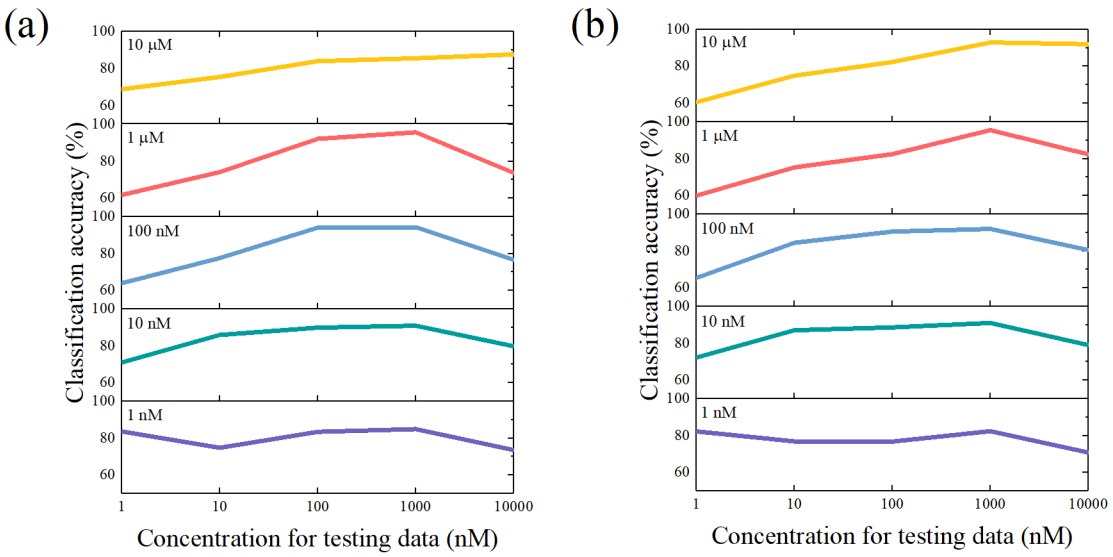

Figure 3-10. Classification accuracy using single SVM models. Classification accuracy produced by the SVM models in the first experiment. (a) and the second experiment (b). Each row demonstrates the classification accuracy at each drug concentration produced by a single SVM model.

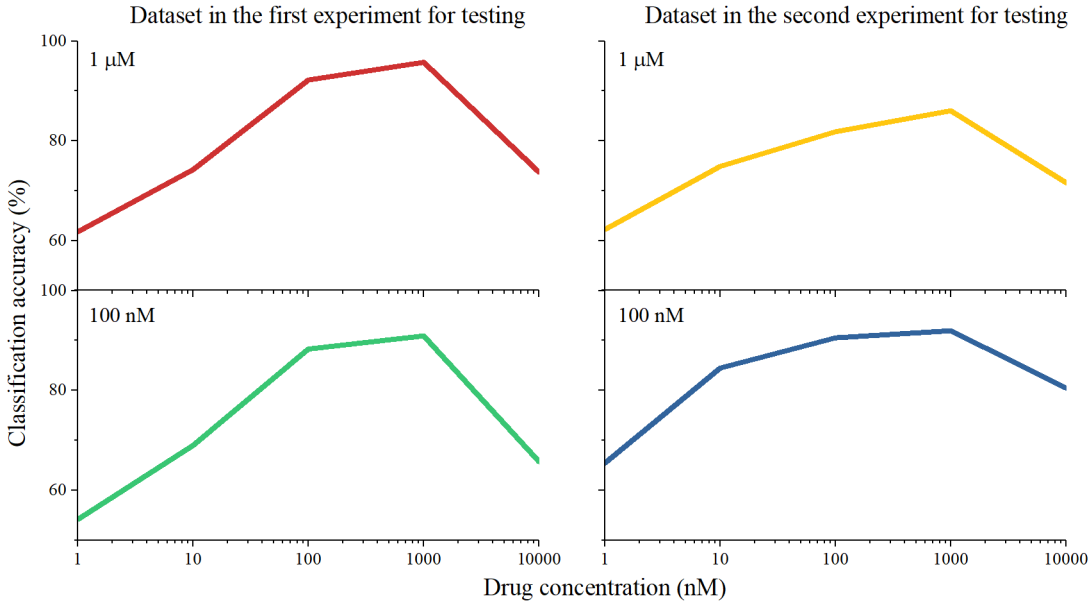

Figure 3-11. Evaluation of single SVM models across different experiments. Each row demonstrates the classification accuracy at each drug concentration produced by a single SVM model trained with the data from 1 µM in the first experiment (upper) and 100 nM in the second experiment (lower). Each column demonstrates the testing data from different trial of experiments.

Chapter 4

Label-free detection of drug responses of cancer cells in human whole blood

本章については、5年以内に雑誌等で刊行予定のため、非公開。

Chapter 5

Label-free detection of drug responses of *ex vivo* cancer cells in mice

本章については、5年以内に雑誌等で刊行予定のため、非公開。

Chapter 6

Summary and outlook

In this thesis, we proposed and experimentally demonstrated a method to detect cellular drug responses via drug-induced morphological change which are inarguably too subtle for the human eye to identify, but are identifiable with the combination of numerous bright-field cell images and machine learning. By leveraging machine learning techniques, we captured these subtle changes and successfully distinguished drug-treated and -untreated cells only by the use of bright-field images at a high classification accuracy of over 90% without the need for any labeling techniques. Furthermore, we evaluated the feasibility of this method in various perspectives.

In Chapter 3, we have demonstrated the proof-of-concept of our method, and verified that our approach is capable of robustly capturing invariant and distinctive features in drug-induced morphological change, such that a one-time trained classifier model can be used for different datasets. This is an important property for screening applications because the system does not require a new retraining for a new trial of screening, which can significantly reduce the computation costs.

In Chapter 4, we have examined our method with different types of cells including drug susceptible and resistant strains. We have successfully demonstrated that the difference in drug responses inherited in the cell line were clearly expressed through morphological changes and machine learning. This result suggests that our method can be applied to detect the drug resistant cells or evaluate the drug susceptibility in a label-free manner. This potential capability is particularly beneficial for blood diagnosis where drug susceptibility test is desired to be done in a high-throughput manner. To this end, we evaluated the feasibility of using whole blood as a sample to evaluate the drug induced morphological changes. We have successfully demonstrated that our optofluidic time-stretch microscope achieved a flow speed of 15 m/s. , enabling a frame rate up to 380,000 frames/s, corresponding to an unprecedented throughput of 380,000 cells/s, which is over 50 times higher than that of commercial imaging flow cytometers [107]. Despite its high

throughput, the spatial resolution of our microscope is maintained at the diffraction limit of 780 nm [75], which is comparable to that of conventional optical microscopes. To the best of our knowledge, this is the first experimental demonstration of an imaging flow cytometer capable of taking diffraction-limited images with a throughput of more than 100,000 cells/s, which is challenging even for conventional non-imaging flow cytometry [108-110]. Moreover, by flowing whole blood, which contains blood cells in a very high concentration, we further achieved an extreme throughput of 10^6 cells/s. Although the throughput of 10^6 cells/s was achieved previously by parallelizing inertia-focusing channels [97], the pixel resolution was sacrificed, and the sample blood has to be diluted down to 1% due to its low flowing speed. In addition, in inertia-focusing channel, the focusing efficiency is greatly affected by the cell size which makes it unsuitable for blood diagnosis. With our method, the combination of such high-throughput and high-resolution capabilities allows us to use only bright-field images and simple machine learning algorithms to identify the miniscule drug-induced morphological change of cells with high accuracy.

We further explored other combination of cell types and drug to examine the applicability in various settings. In Chapter 5, we have demonstrated that the drug responses of two lung cancer cell line, A549/EGFP and PC-9/mRuby (both were genetically modified to express fluorescent proteins) against gefinitib, a FDA-approved anticancer drug for targeted therapy were also expressed through our method. This combination is essentially distinct from the drugs used in previous chapters because gefitinib only inhibits a certain type of molecules with a very high specificity [111]. By demonstrating the feasibility of identifying the drug responses induced by gefitinib, it is suggested that a molecular specific reaction occurring inside a cell may be detectable through label-free bright-field images. We have also examined the drug responses of cells that were dissociated from xenograft of nude mice. Although the results show differences in drug responses of these dissociated cells, we found that the experimental processes are highly influenced by the efficiency of dissociation where some types of cells are easier to become single cells while some are not. Further optimization for dissociation and the overall experimental process will improve the quality of the results.

Although some of the cells we used in this thesis are adhesive cells whose morphological change is typically observed when they are adhered on a surface [112,113], we chose to observe their morphological change in suspension despite the risk of losing their morphological change during the process of trypsinization, because higher throughput can be achieved in the flow-cytometric manner. Nonetheless, we demonstrated that morphological change can also be investigated for the adhesive cells in a suspended condition with our approach, suggesting that it can be applied in distinguishing dissociated tissue samples via their morphological variations. We also chose linear SVM classifier in this study for the proof-of-concept because it is a simple machine-learning algorithm and is unlikely to overfit. More advanced machine-learning algorithms can also be applied to achieve higher accuracy; for instance, ensemble methods [45] or deep architectures [88] can be used for classification. In addition, instead of hand-coding features, convolutional networks can be used to extract features that are problem-specific and therefore further improve classification accuracy.

Table of Acronyms

Acronyms	Full Name
ADC	analog-to-digital converter
Cas	CRISPR associated protein
CCD	charge-coupled device
CMOS	complementary metal–oxide–semiconductor
CNN	convolutional neural network
CRISPR	clustered regularly interspaced short palindromic repeats
DFT	dispersive Fourier transformation
DMEM	Dulbecco's modified eagle medium
DMSO	dimethyl sulfoxide
EGFP	enhanced green fluorescent protein
EGFR	epidermal growth factor receptor
EMCCD	Electron multiplying charge-coupled device
FDA	US Food and Drug Administration
fps	frames per second
FOV	field of view
FRET	fluorescence resonance energy transfer
GC-MS	gas chromatography-mass spectrometry
GF-AFC	glycylphenylalanyl-amino uorocoumarin
GVD	group velocity dispersion
HCS	high-content screening

HTS	high-throughput screening
iPS	induced pluripotent stem
LC-MS	liquid chromatography-mass spectrometry
LDA	linear discriminant analysis
OTS	optofluidic time-stretch
PCR	polymerase chain reaction
PEEK	polyetheretherketone
PDMS	polydimethylsiloxane
Re	Reynolds number
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
sCMOS	scientific complementary metal–oxide–semiconductor
TDI	time delay integration
UV	ultraviole

Publications

During my Ph.D., I have worked on a couple of different projects, some of which are related to OTS microscope and others are not. Listed below are my publication in peer-reviewed journals.

Publication related to this thesis

1. **Label-free detection of cellular drug responses by high-throughput bright-field imaging and machine learning**, Kobayashi H, Lei C, Wu Y, Mao A, Jiang Y, Guo B, Ozeki Y, Goda K, Scientific Reports. 2017;**7**(1):12454.
2. **High-throughput imaging flow cytometry by optofluidic time-stretch microscopy**, Lei C, Kobayashi H, Wu Y, Li M, Isozaki A, Yasumoto A, Mikami H, Ito T, Nitta N, Sugimura T, Yamada M, Yatomi Y, Di Carlo D, Ozeki Y, Goda K, Nature Protocols. 2018. In press.

Publication not related to this thesis

3. **Ultrafast confocal fluorescence microscopy beyond the fluorescence lifetime limit**, Mikami H, Harmon J, Kobayashi H, Hamad S, Wang Y, Iwata O, Suzuki K, Ito T, Aisaka Y, Katsuna N, Nagasawa K, Watarai H, Ozeki Y, Goda K, Optica. 2018. **5**(2) 117-126
4. **Label-free detection of aggregated platelets in blood by machine-learning-aided optofluidic time-stretch microscopy**, Jiang Y, Lei C, Yasumoto A, Kobayashi H, Aisaka Y, Ito T, Guo B, Nitta N, Kutsuna N, Ozeki Y, Goda K, Lab on a Chip. 2017;**17**(14):2426-34.
5. **Optofluidic time-stretch quantitative phase microscopy**, Guo B, Lei C, Wu Y, Kobayashi H, Ito T, Yalikun Y, Lee S, Isozaki A, Li M, Jiang Y, Yasumoto A, Di Carlo D, Tanaka Y, Yatomi Y, Ozeki Y, Goda K, Methods. 2017. In press
6. **High-throughput, label-free, single-cell, microalgal lipid screening by machine-learning-equipped optofluidic time-stretch quantitative phase microscopy**, Guo B, Lei C, Kobayashi H, Ito T, Yalikun Y, Jiang Y, Tanaka Y, Ozeki Y, Goda K, Cytometry A. 2017;**91**(5):494-502.

7. **GHz Optical Time-Stretch Microscopy by Compressive Sensing**, Lei C, Wu Y, Sankaranarayanan AC, Chang SM, Guo B, Sasaki N, Kobayashi H, Sun W.C, Ozeki Y, Goda K, IEEE Photonics Journal. 2017;**9**(2):1-8.
8. **High-throughput label-free image cytometry and image-based classification of live Euglena gracilis**. Lei C, Ito T, Ugawa M, Nozawa T, Iwata O, Maki M, Okada G, Kobayashi H, Sun X, Tiamsak P, Tsumura N, Suzuki K, Di Carlo D, Ozeki Y, Goda K, Biomed Opt Express. 2016;**7**(7):2703-8
9. **Ultrafast Confocal Fluorescence Microscopy by Frequency-Division-Multiplexed Multi-Line Focusing**, Mikami H, Kobayashi H, Hamad S, Ozeki Y, Goda K, 2016 Conference on Lasers and Electro-Optics (CLEO); 2016 5-10 June 2016.

Bibliography

1 Gonzalez-Munoz, A. L., Minter, R. R. & Rust, S. J. Phenotypic screening: the future of antibody discovery. *Drug Discov Today* **21**, 150-156; (2016).
2 Shi, Y., Inoue, H., Wu, J. C. & Yamanaka, S. Induced pluripotent stem cell technology: a decade of progress. *Nat. Rev. Drug Discov.* **16**, 115-130; (2017).
3 Fellmann, C., Cowen, B. C., Lin, P. C., Doudna, J. A. & Corn, J. E. Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat. Rev. Drug Discov.* **16**, 89-100; (2017).
4 Fang, Y. & Eglen, R. M. Three-Dimensional Cell Cultures in Drug Discovery and Development. *Slas Discov* **22**, 456-472; (2017).
5 Skardal, A., Shupe, T. & Atala, A. Organoid-on-a-chip and body-on-a-chip systems for drug screening and disease modeling. *Drug Discov Today* **21**, 1399-1411; (2016).
6 Mills, R. J. *et al.* Functional screening in human cardiac organoids reveals a metabolic mechanism for cardiomyocyte cell cycle arrest. *Proc Natl Acad Sci U S A* **114**, E8372-E8381; (2017).
7 Schulte, J., Sepp, K. J., Wu, C., Hong, P. & Littleton, J. T. High-content chemical and RNAi screens for suppressors of neurotoxicity in a Huntington's disease model. *PLoS One* **6**, e23841; (2011).
8 Okada, H., Ohnuki, S., Roncero, C., Konopka, J. B. & Ohya, Y. Distinct roles of cell wall biogenesis in yeast morphogenesis as revealed by multivariate analysis of high-dimensional morphometric data. *Mol Biol Cell* **25**, 222-233; (2014).
9 Slacka, M. D., Martineza, E. D., Wu, L. F. & Altschuler, S. J. Characterizing heterogeneous cellular responses to perturbations. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 19306-19311; (2008).
10 Pereira, D. A. & Williams, J. A. Origin and evolution of high throughput screening. *Br. J. Pharmacol.* **152**, 53-61; (2007).
11 Monks, A. *et al.* Feasibility of a High-Flux Anticancer Drug Screen Using a Diverse Panel of Cultured Human Tumor Cell Lines. *J. Natl. Cancer Inst.* **83**, 757-766; (1991).
12 Szymański, P., Markowicz, M. & Mikiciuk-Olasik, E. Adaptation of high-throughput screening in drug discovery—toxicological screening tests. *Int. J. Mol. Sci.* **13**, 427-452; (2011).
13 Burns, S. *et al.* Identification of Small-Molecule Inhibitors of Protein Kinase B (PKB/AKT) in an AlphaScreen™ High-Throughput Screen. *J. Biomol. Screen.* **11**, 822-827; (2006).
14 Sudo, K. *et al.* High-Throughput Screening of Low Molecular Weight NS3-NS4A Protease Inhibitors Using a Fluorescence Resonance Energy Transfer Substrate. *Antiviral Chem. Chemother.* **16**, 385-392; (2005).
15 Swaney, S. *et al.* Characterization of a High-Throughput Screening Assay for Inhibitors of Elongation Factor P and Ribosomal Peptidyl Transferase Activity. *J. Biomol. Screen.* **11**, 736-742; (2006).
16 Allen, M., Reeves, J. & Mellor, G. High Throughput Fluorescence Polarization: A Homogeneous Alternative to Radioligand Binding for Cell Surface Receptors. *J. Biomol. Screen.* **5**, 63-69; (2000).

17 Parker, G. J., Law, T. L., Lenoch, F. J. & Bolger, R. E. Development of High Throughput Screening Assays Using Fluorescence Polarization: Nuclear Receptor-Ligand-Binding and Kinase/Phosphatase Assays. *J. Biomol. Screen.* **5**, 77-88; (2000).
18 Xu, J. *et al.* Ion-channel assay technologies: quo vadis? *Drug Discov. Today* **6**, 1278-1287; (2001).
19 Kenny, C. H. *et al.* Development of a fluorescence polarization assay to screen for inhibitors of the FtsZ/ZipA interaction. *Anal. Biochem.* **323**, 224-233; (2003).
20 An, W. F. & Tolliday, N. Cell-Based Assays for High-Throughput Screening. *Mol. Biotechnol.* **45**, 180-186; (2010).
21 Chris, C. *et al.* Measuring Intracellular Calcium Fluxes in High Throughput Mode. *Comb. Chem. High Throughput Screen.* **6**, 355-362; (2003).
22 Kariv, I., Stevens, M. E., Behrens, D. L. & Oldenburg, K. R. High Throughput Quantitation of cAMP Production Mediated by Activation of Seven Transmembrane Domain Receptors. *J. Biomol. Screen.* **4**, 27-32; (1999).
23 Li, X. *et al.* Functional characterization of cell lines for high-throughput screening of human neuromedin U receptor subtype 2 specific agonists using a luciferase reporter gene assay. *Eur. J. Pharm. Biopharm.* **67**, 284-292; (2007).
24 Beck, V., Pfitscher, A. & Jungbauer, A. GFP-reporter for a high throughput assay to monitor estrogenic compounds. *J. Biochem. Biophys. Methods* **64**, 19-37; (2005).
25 Yarrow, J. C., Totsukawa, G., Charras, G. T. & Mitchison, T. J. Screening for Cell Migration Inhibitors via Automated Microscopy Reveals a Rho-Kinase Inhibitor. *Chem. Biol.* **12**, 385-395.
26 Eggert, U. S. *et al.* Parallel Chemical Genetic and Genome-Wide RNAi Screens Identify Cytokinesis Inhibitors and Targets. *PLoS Biol.* **2**, e379; (2004).
27 Imaduwage, K. P., Lakbub, J., Go, E. P. & Desaire, H. Rapid LC-MS Based High-Throughput Screening Method, Affording No False Positives or False Negatives, Identifies a New Inhibitor for Carbonic Anhydrase. *Sci. Rep.* **7**, 10324; (2017).
28 Smit, B. A. *et al.* Development of a high throughput screening method to test flavour-forming capabilities of anaerobic micro-organisms. *J. Appl. Microbiol.* **97**, 306-313; (2004).
29 Thomas, R. K. *et al.* High-throughput oncogene mutation profiling in human cancer. *Nat. Genet.* **39**, 347; (2007).
30 Moyerbrailean, G. A. *et al.* A high-throughput RNA-seq approach to profile transcriptional responses. *Sci. Rep.* **5**, 14976; (2015).
31 Hall, M. D. *et al.* Fluorescence polarization assays in high-throughput screening and drug discovery: a review. *Methods Appl. Fluoresc.* **4**, 022001-022001; (2016).
32 Moger, J., Gribbon, P., Sewing, A. & Winlove, C. P. The Application of Fluorescence Lifetime Readouts in High-Throughput Screening. *J. Biomol. Screen.* **11**, 765-772; (2006).
33 Du, Y. *et al.* A Time-Resolved Fluorescence Resonance Energy Transfer Assay for High-Throughput Screening of 14-3-3 Protein–Protein Interaction Inhibitors. *Assay Drug Dev. Technol.* **11**, 367-381; (2013).
34 Jeong, E.-m., Lee, M. Y., Lee, J. H., Lee, B. H. & Oh, K.-S. A Dual Readout Assay Based on Fluorescence Polarization and Time-Resolved Fluorescence Resonance Energy Transfer to Screen for RSK1 Inhibitors. *Biol. Pharm. Bull.* **39**, 547-555; (2016).

35 Wachsmuth, M. *et al.* High-throughput fluorescence correlation spectroscopy enables analysis of proteome dynamics in living cells. *Nat. Biotechnol.* **33**, 384; (2015).
36 Haupts, U. *et al.* Single-Molecule Detection Technologies in Miniaturized High-Throughput Screening: Fluorescence Intensity Distribution Analysis. *J. Biomol. Screen.* **8**, 19-33; (2003).
37 Lu, Y. *et al.* High-Throughput Secretomic Analysis of Single Cells to Assess Functional Cellular Heterogeneity. *Anal. Chem.* **85**, 2548-2556; (2013).
38 Carpenter, A. E. in *Book Extracting Rich Information from Images* Vol. 486 193-211 (Humana Press, 2009).
39 Giaever, G. *et al.* Chemogenomic profiling: Identifying the functional interactions of small molecules in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 793-798; (2004).
40 Josiah, S. in *Book Interpretation of Uniform-Well Readouts* 177-192 (Humana Press, 2009).
41 Giuliano, K. A. *et al.* High-Content Screening: A New Approach to Easing Key Bottlenecks in the Drug Discovery Process. *J. Biomol. Screen.* **2**, 249-259; (1997).
42 Taylor, D. L. in *Book Past, Present, and Future of High Content Screening and the Field of Cellomics* 3-18 (Humana Press, 2006).
43 Posner, B. High-throughput screening-driven lead discovery: meeting the challenges of finding new therapeutics. *Curr. Opin. Drug Discov. Devel.* **8**, 487-494; (2005).
44 Wojcik, K. & Dobrucki, J. W. Interaction of a DNA intercalator DRAQ5, and a minor groove binder SYTO17, with chromatin in live cells-influence on chromatin organization and histone-DNA interactions. *Cytometry A* **73a**, 555-562; (2008).
45 Blasi, T. *et al.* Label-free cell cycle analysis for high-throughput imaging flow cytometry. *Nat. Commun.* **7**, 10256; 10.1038/ncomms10256 (2016).
46 Kang, J. *et al.* Improving drug discovery with high-content phenotypic screens by systematic selection of reporter cell lines. *Nat. Biotechnol.* **34**, 70-77; (2016).
47 Hale, C. M. & Nojima, D. in *Book Approaches to High Content Imaging and Multi-feature Analysis High Throughput Screening Methods: Evolution and Refinement* Ch. 8, 162-180 (The Royal Society of Chemistry, 2017).
48 Holst, G. C. & Lomheim, T. S. *CMOS/CCD sensors and camera systems*. Vol. 408 (JCD publishing USA, 2007).
49 Beier, H. T. & Ibey, B. L. Experimental comparison of the high-speed imaging performance of an EM-CCD and sCMOS camera in a dynamic live-cell imaging test case. *PLoS One* **9**, e84614; (2014).
50 Wong, H. S., Yao, Y. L. & Schlig, E. S. TDI charge-coupled devices: Design and applications. *IBM Journal of Research and Development* **36**, 83-106; (1992).
51 Johnson, J. F. Modeling imager deterministic and statistical modulation transfer functions. *Appl. Opt.* **32**, 6503-6513; (1993).
52 Lepage, G., Bogaerts, J. & Meynants, G. Time-Delay-Integration Architectures in CMOS Image Sensors. *IEEE Transactions on Electron Devices* **56**, 2524-2533; (2009).
53 Wang, X. M. Advances and issues in flow cytometric detection of immunophenotypic changes and genomic rearrangements in acute pediatric leukemia. *Translational Pediatrics* **3**, 149-155; (2014).
54 Peters, J. M. & Ansari, M. Q. Multiparameter flow cytometry in the diagnosis and management of acute leukemia. *Arch Pathol Lab Med* **135**, 44-54; (2011).

55 Hokland, P. & Ommen, H. B. Towards individualized follow-up in adult acute myeloid leukemia in remission. *Blood* **117**, 2577-2584; (2011).
56 Coustan-Smith, E. & Campana, D. Should evaluation for minimal residual disease be routine in acute myeloid leukemia? *Curr. Opin. Hematol.* **20**, 86-92; (2013).
57 Jaso, J. M., Wang, S. A., Jorgensen, J. L. & Lin, P. Multi-color flow cytometric immunophenotyping for detection of minimal residual disease in AML: past, present and future. *Bone Marrow Transplant.* **49**, 1129-1138; (2014).
58 Hourigan, C. S. & Karp, J. E. Minimal residual disease in acute myeloid leukaemia. *Nat Rev Clin Oncol* **10**, 460-471; (2013).
59 Jorgensen, J. L. & Chen, S. S. Monitoring of minimal residual disease in acute myeloid leukemia: methods and best applications. *Clin Lymphoma Myeloma Leuk* **11 Suppl 1**, S49-53; (2011).
60 Goda, K. & Jalali, B. Dispersive Fourier transformation for fast continuous single-shot measurements. *Nat. Photon.* **7**, 102-112; (2013).
61 Mahjoubfar, A. *et al.* Time stretch and its applications. *Nat. Photon.* **11**, 341-351; (2017).
62 Bhushan, A. S., Coppinger, F. & Jalali, B. Time-stretched analogue-to-digital conversion. *Electron. Lett.* **34**, 839-841; (1998).
63 Coppinger, F., Bhushan, A. S. & Jalali, B. Photonic time stretch and its application to analog-to-digital conversion. *IEEE Trans. Microwave Theory Tech.* **47**, 1309-1314; (1999).
64 Golden, J. P., Justin, G. A., Nasir, M. & Ligler, F. S. Hydrodynamic focusing—a versatile tool. *Anal. Bioanal. Chem.* **402**, 325-335; (2012).
65 Di Carlo, D. Inertial microfluidics. *Lab Chip* **9**, 3038-3046; (2009).
66 Chung, A. J., Gossett, D. R. & Di Carlo, D. Three dimensional, sheathless, and high-throughput microparticle inertial focusing through geometry-induced secondary flows. *Small* **9**, 685-690; (2013).
67 Li, M. *et al.* Inertial focusing of ellipsoidal *Euglena gracilis* cells in a stepped microchannel. *Lab Chip* **16**, 4458-4465; (2016).
68 Li, M. *et al.* Inertial focusing of ellipsoidal Euglena gracilis cells in a stepped microchannel. *Lab Chip* **16**, 4458-4465; (2016).
69 Goda, K. *et al.* High-throughput single-microparticle imaging flow analyzer. *Proc. Natl. Acad. Sci. USA* **109**, 11630-11635; (2012).
70 Di Carlo, D., Irimia, D., Tompkins, R. G. & Toner, M. Continuous inertial focusing, ordering, and separation of particles in microchannels. *Proceedings of the National Academy of Sciences* **104**, 18892-18897; (2007).
71 Gossett, D. R. *et al.* Hydrodynamic stretching of single cells for large population mechanical phenotyping. *Proceedings of the National Academy of Sciences* **109**, 7630-7635; (2012).
72 Buckley, B. W., Madni, A. M. & Jalali, B. Coherent time-stretch transformation for real-time capture of wideband signals. *Opt. Express* **21**, 21618-21627; (2013).
73 Wei, X. *et al.* Coherent Laser Source for High Frame-Rate Optical Time-Stretch Microscopy at 1.0 μm. *IEEE J. Sel. Topics Quantum Electron.* **20**, 384-389; (2014).
74 Lei, C., Guo, B., Cheng, Z. & Goda, K. Optical time-stretch imaging: Principles and applications. *Appl. Phys. Rev.* **3**, 011102; (2016).

75 Ugawa, M. *et al.* High-throughput optofluidic particle profiling with morphological and chemical specificity. *Opt. Lett.* **40**, 4803-4806; (2015).
76 Bengio, Y. Learning Deep Architectures for AI. *Found. Trends Netw.* **2**, 1-127; (2009).
77 Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks in Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1.1097-1105 (Curran Associates Inc., 2012 Published).
78 Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100; (2006).
79 Kamentsky, L. *et al.* Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* **27**, 1179-1180; (2011).
80 Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*; (2014).
81 He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition in Proceedings of the IEEE conference on computer vision and pattern recognition.770-778 (2016 Published).
82 Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. How transferable are features in deep neural networks? in Advances in neural information processing systems.3320-3328 (2014 Published).
83 Kobayashi, H. *et al.* Label-free detection of cellular drug responses by high-throughput bright-field imaging and machine learning. *Sci. Rep.* **7**, 12454; (2017).
84 Jiang, Y. *et al.* Label-free detection of aggregated platelets in blood by machine-learning-aided optofluidic time-stretch microscopy. *Lab Chip* **17**, 2426-2434; (2017).
85 Lei, C. *et al.* High-throughput label-free image cytometry and image-based classification of live Euglena gracilis. *Biomed. Opt. Express* **7**, 2703-2708; (2016).
86 Boser, B. E., Guyon, I. M. & Vapnik, V. N. A training algorithm for optimal margin classifiers in Proceedings of the fifth annual workshop on Computational learning theory.144-152 (ACM, 1992 Published).
87 Cortes, C. & Vapnik, V. Support-vector networks. *Mach Learn.* **20**, 273-297; (1995).
88 Chen, C. L. *et al.* Deep learning in label-free cell classification. *Sci. Rep.* **6**, 21471; (2016).
89 Loo, L. H., Wu, L. F. & Altschuler, S. J. Image-based multivariate profiling of drug responses from single cells. *Nat. Methods* **4**, 445-453; (2007).
90 Grys, B. T. *et al.* Machine learning and computer vision approaches for phenotypic profiling. *J. Cell Biol.* **216**, 65-71; (2016).
91 Nishio, K. *et al.* Cross-resistance to tumour promoters in human cancer cell lines resistant to adriamycin or cisplatin. *Br. J. Cancer* **62**, 415-419; (1990).
92 Asano, T., Zhi, C. L., Hayakawa, J. & Fukunaga, Y. Analysis of Resistance-Related Gene Expression in Doxorubicin-Resistant Leukemia Cell Line by DNA Microarray. *Jpn. J. Pediatr. Hematol.* **17**, 340-345; (2003).
93 Earnshaw, W. C., Martins, L. M. & Kaufmann, S. H. Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu. Rev. Biochem.* **68**, 383-424; (1999).

94 Binkowski, B. F. & Butler, B. L. *A real-time assay of caspase-3/7 activity in cell culture and animal models.*, <http://www.promega.jp/resources/pubhub/a-real-time-assay-of-caspase-activity-in-cell-culture-and-animal-models/> (2014).
95 Fan, F. & Wood, K. V. Bioluminescent assays for high-throughput screening. *Assay Drug Dev. Technol.* **5**, 127-136; (2007).
96 Kratz, A., Ferraro, M., Sluss, P. M. & Lewandrowski, K. B. Case records of the Massachusetts General Hospital. Weekly clinicopathological exercises. Laboratory reference values. *N Engl J Med* **351**, 1548-1563; (2004).
97 Hur, S. C., Tse, H. T. & Di Carlo, D. Sheathless inertial cell ordering for extreme throughput flow cytometry. *Lab Chip* **10**, 274-280; (2010).
98 Segovia-Mendoza, M., Gonzalez-Gonzalez, M. E., Barrera, D., Diaz, L. & Garcia-Becerra, R. Efficacy and mechanism of action of the tyrosine kinase inhibitors gefitinib, lapatinib and neratinib in the treatment of HER2-positive breast cancer: preclinical and clinical evidence. *Am J Cancer Res* **5**, 2531-2561; (2015).
99 Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497-1500; (2004).
100 Kawahara, A. *et al.* Molecular Diagnosis of Activating EGFR Mutations in Non-Small Cell Lung Cancer Using Mutation-Specific Antibodies for Immunohistochemical Analysis. *Clin. Cancer. Res.* **16**, 3163-3170; (2010).
101 Noro, R. *et al.* Gefitinib (IRESSA) sensitive lung cancer cell lines show phosphorylation of Akt without ligand stimulation. *BMC Cancer* **6**; (2006).
102 Simonetti, S. *et al.* Detection of EGFR mutations with mutation-specific antibodies in stage IV non-small-cell lung cancer. *J Transl Med* **8**; (2010).
103 Valenzuela, D. M. & Groffen, J. Four human carcinoma cell lines with novel mutations in position 12 of c-K-ras oncogene. *Nucleic Acids Res* **14**, 843-852; (1986).
104 Ono, M. *et al.* Sensitivity to gefitinib (Iressa, ZD1839) in non-small cell lung cancer cell lines correlates with dependence on the epidermal growth factor (EGF) receptor/extracellular signal-regulated kinase 1/2 and EGF receptor/Akt pathway for proliferation. *Mol. Cancer Ther.* **3**, 465-472; (2004).
105 Janmaat, M. L., Rodriguez, J. A., Gallegos-Ruiz, M., Kruyt, F. A. & Giaccone, G. Enhanced cytotoxicity induced by gefitinib and specific inhibitors of the Ras or phosphatidyl inositol-3 kinase pathways in non-small cell lung cancer cells. *Int. J. Cancer* **118**, 209-214; (2006).
106 Tomayko, M. M. & Reynolds, C. P. Determination of subcutaneous tumor size in athymic (nude) mice. *Cancer Chemother. Pharmacol.* **24**, 148-154; (1989).
107 Basiji, D. A., Ortyn, W. E., Liang, L., Venkatachalam, V. & Morrissey, P. Cellular image analysis and imaging by flow cytometry. *Clin. Lab. Med.* **27**, 653-670; (2007).
108 Black, C. B., Duensing, T. D., Trinkle, L. S. & Dunlay, R. T. Cell-based screening using high-throughput flow cytometry. *Assay Drug Dev. Technol.* **9**, 13-20; (2011).
109 Yang, Y. S. S. *et al.* High-throughput quantitation of inorganic nanoparticle biodistribution at the single-cell level using mass cytometry. *Nat. Commun.* **8**, 14069; 10.1038/ncomms14069 (2017).
110 Tang, A. H. L. *et al.* Time-stretch microscopy on a DVD for high-throughput imaging cell-based assay. *Biomed. Opt. Express* **8**, 640-652; (2017).

111 Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* **305**, 1163-1167; (2004).
112 Chalut, K. J., Ostrander, J. H., Giacomelli, M. G. & Wax, A. Light scattering measurements of subcellular structure provide noninvasive early detection of chemotherapy-induced apoptosis. *Cancer Res.* **69**, 1199-1204; (2009).
113 Saunders, D. E. *et al.* Paclitaxel-induced apoptosis in MCF-7 breast-cancer cells. *Int. J. Cancer* **70**, 214-220; (1997).