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

Di, Z.

Chapter 4

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

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

Manuscript in preparation

Abstract

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

4.1 Introduction

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

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

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

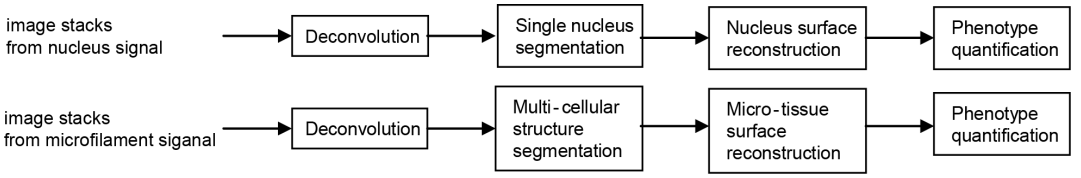

Figure 4.1: Overview of the proposed image analysis pipeline for high-content screening of 3D cultured micro-tissues based on wide-field florescence microscopy.

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

4.2 Results

4.2.1 Method development for 3D surface reconstruction

4.2.1.1 Deconvolution

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

parameters (Table S4.1). First, the signal-to-noise ratio was estimated on randomly selected sample images. This ratio was relatively high (> 80), and considering the time efficiency we used a fast maximum likelihood estimation algorithm [6] for the image restoration (see examples before and after deconvolution in Figure 4.2a-4.2h). A large extent of out-of-focus signals was removed and the the signal-to-noise ratio was enhanced for both Hoechst stained nuclei channel and rhodamine stained F-actin channel.

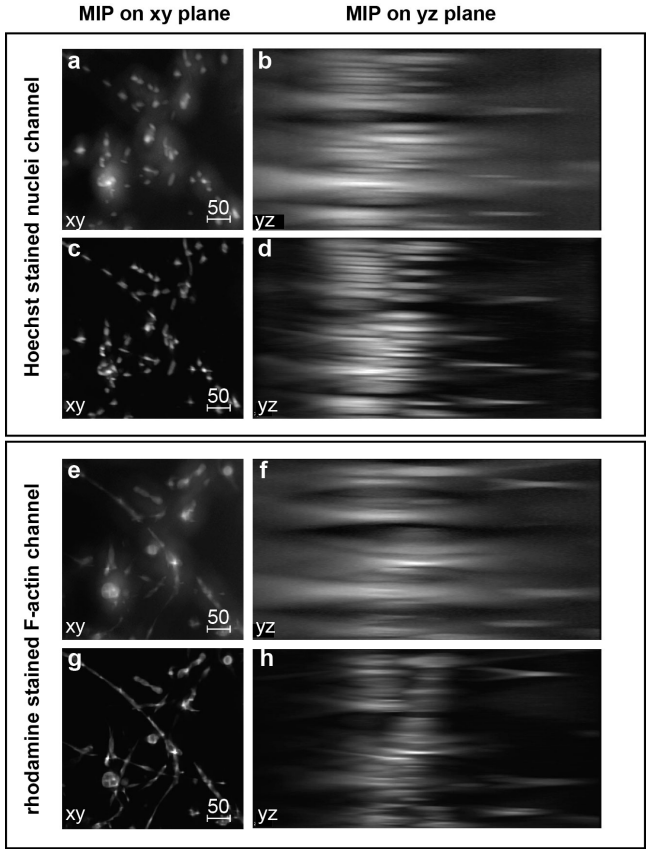

Figure 4.2: An example of deconvolution results for both Hoechst stained nuclei channel and rhodamine stained F-actin channel. For the Hoechst stained nuclei channel, we cropped a part of an original image stack and showed its maximum intensity projection (MIP) on (a) xy-plane and on (b) yz-plane. After deconvolution of this image stack, we showed the MIP of the same part of image stack on (c) xy-plane and on (d) yz-plane. For the rhodamine stained F-actin channel, we showed the MIP of corresponding part of image stack before deconvolution (e-f) and after deconvolution (g-h). The scale bar represents 50μm.

4.2.1.2 Segmentation for each individual nucleus

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

4.2.1.3 Segmentation for multi-cellular micro-tissue structures

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

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

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

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

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

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

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

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

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

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

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

Our 3D surface reconstruction consists of two important steps: 1) 3D labeling which assigns a label to each foreground pixel so that pixels with the same label define one single object; 2) 3D reconstruction which generates a geometrical 3D model for each labeled object. For the Hoechst stained nuclei channel, 3D watershed already created compartments for each single nucleus. Foreground pixels identified in the same compartment were assigned with the same object label. For 3D labeling of the multi-cellular micro-tissue structures, we applied a sequential labeling algorithm based on 18-connected connectivity as follow: suppose one foreground pixel with coordinates (x, y, z) has already been labeled as l, 18 neighboring pixels with coordinates $(x, y, z \pm 1)$, $(x \pm 1, y, z \pm 1)$, $(x, y \pm 1, z \pm 1)$, $(x \pm 1, y, z)$, $(x, y \pm 1, z)$, $(x \pm 1, y \pm 1, z)$ are pushed into a first-in-last-out stack to be assessed later. Every time one pixel is removed from the beginning of the stack. If this pixel is a foreground pixel, label l is assigned to the pixel and its 18 neighbors are pushed into the stack. The labeling of

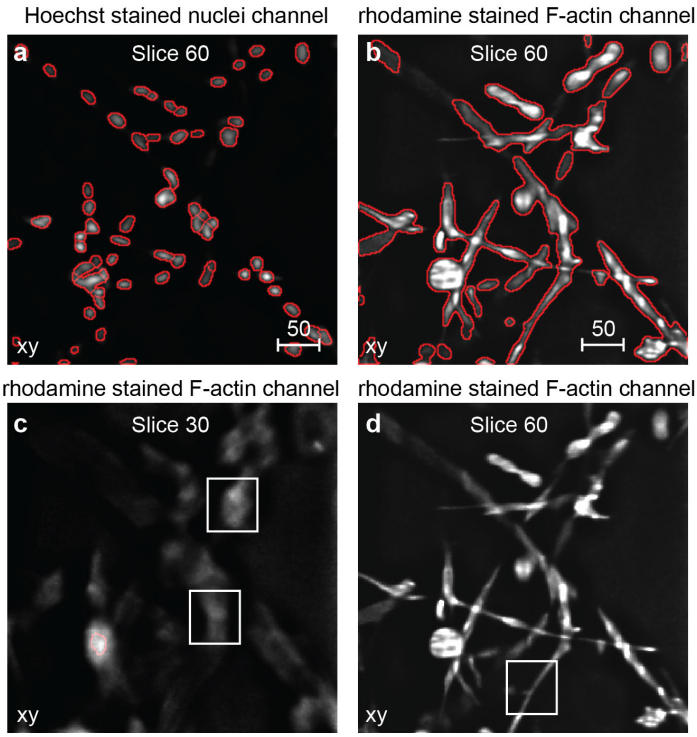

Figure 4.3: Segmentation results for both Hoechst stained nuclei channel and rhodamine stained F-actin channel. (a) Segmentation result of one cropped image slice (deconvolved) for the Hoechst stained nuclei channel. Red marks the contour of the segmentation result. (b) Segmentation result for the corresponding image (deconvolved) from the rhodamine stained F-actin channel. (c) A deconvolved image slice with low sharpness level, from the same stack. White boxes mark the regions which are out-of-focus but with high intensity. Red marks the contour of the segmentation result. (d) A deconvolved image slice with higher sharpness level than (c). White box marks the region which is in-focus but has lower intensity than the regions marked in (c). The segmentation result is shown in (b). The scale bar represents 50μm.

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

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

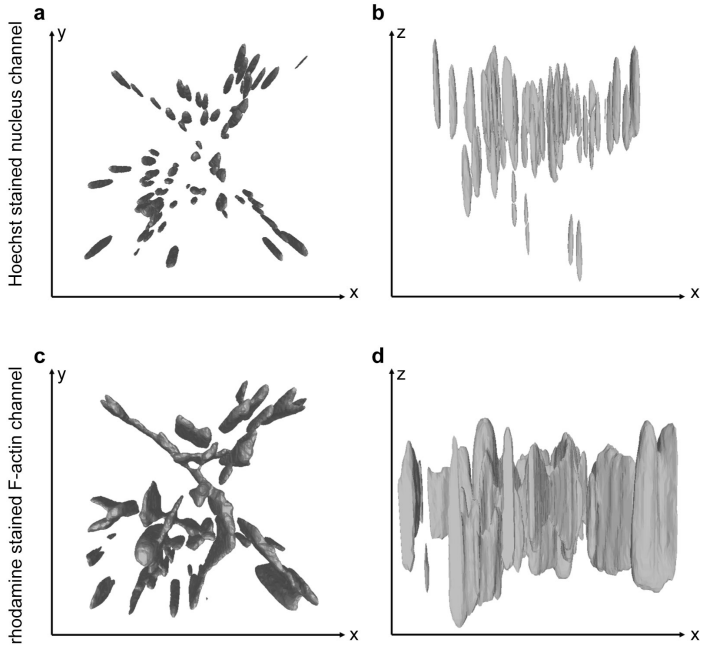

Figure 4.4: 3D reconstruction result for nuclei and multi-cellular micro-tissue structures. (a-b) 3D reconstructed nuclei in xy- and xz-direction, based on the same image stack as shown in Figure 4.3. (c-d) 3D reconstructed multi-cellular structures in xy- and xz-direction.

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

of nuclear foreground voxels as follows:

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

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

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

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

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

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

4.2.2 Phenotype measurement

4.2.2.1 Phenotype measurement for individual nuclei

Parameters to profile the phenotype of each individual nucleus can be categorized in three classes: morphological parameters, localization parameters and image moment parameters (Table 4.1). Morphological parameters include a series of shape properties and are measured from the reconstructed 3D geometrical models. In addition to the basic shape properties such as volume and surface, we computed a convex hull [16] and the best-fit ellipsoid [17] for each nucleus, and relative geometrical parameters were measured, for example the volume of the convex hull and semi-axis of the best-fit ellipsoid. Localization parameters estimate the nuclear density by calculating distance between pairs of nuclear centroids. Moment parameters include centroid coordinates and Eigenvalues calculated by Eigen decomposition of the covariance matrix of voxel coordinates. To calculate Eigen decomposition, all coordinates were normalized by moving centroid to origin according to Equation 4.6. Furthermore, we measured the inertia tensor matrix $\begin{matrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{matrix}$ for each nucleus l [17], according to the parallel axis theorem (Equation 4.7-4.12) [18] :

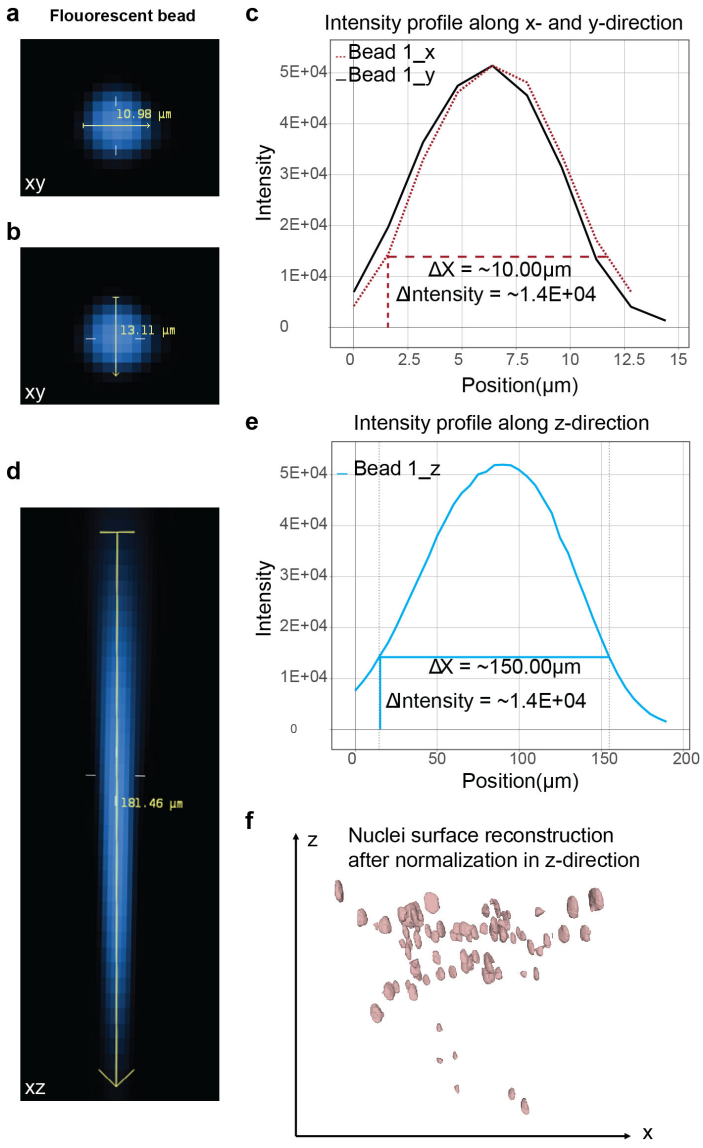

Figure 4.5: Normalization of nuclei in z-direction. (a-b) Image slices through the middle of one fluorescent bead in x- and y-direction. (c) The intensity profiles along the lines indicated in (a) and (b). Red dash line represents the intensity profile in x-direction and black solid line represents the intensity profile in y-direction. (d-e) A xz-plane through the middle of the same bead as in (a) and (b) and corresponding intensity profile. (f) 3D reconstruction for nuclei in xz-direction after normalization. The reconstruction result of the same image stacks before normalization is shown in Figure 4.4b.

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

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

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

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

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

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

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

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

4.2.2.2 Phenotype measurement for multi-cellular micro-tissue network

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

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

Morphological Parameters for each nucleus
Size: Number of foreground voxels assigned to a nucleus
Volume: Volume of a nucleus calculated by summarizing the volume of triangular pyramids that compose the 3D geometrical model.
Surface: Surface of a nucleus calculated by summarizing the area of triangles that compose the surface of the 3D geometrical model.
Width: Width of the 3D bounding rectangular box for a nucleus
Height: Height of the 3D bounding rectangular box for a nucleus
Thickness: Thickness of the 3D bounding rectangular box for a nucleus
Sphericity [19]**:** Compactness measure. Sphericity $= \frac{\pi^{\frac{1}{3}} \times (6 \times \text{Volume})^{\frac{2}{3}}}{\text{Surface}}$
SAV: Surface to volume ratio. SAV $= \frac{\text{Surface}}{\text{Volume}}$
Volume of convex hull: Convex hull was calculated using QuickHull algorithm [16]
Surface of convex hull
Solidity: Solidity $= \frac{\text{Volume}}{\text{Volume of convex hull}}$
Major axis: Length of the longest radius of the best-fit ellipsoid. The best-fit ellipsoid was calculated based on a least-square optimization algorithm according to a ImageJ plugin BoneJ [17]
Median axis: Length of the middle radius of the best-fit ellipsoid
Minor axis: Length of the shortest radius of the best-fit ellipsoid
Mass of the best-fit ellipsoid: Mass $= \frac{4}{3} \times \pi \times \text{Major axis} \times \text{Median axis} \times \text{Minor axis}$
Moments of inertia of the best-fit ellipsoid: $I_{xx}^{ellipsoid} = \frac{1}{5} \times Mass \times (Median\ axis^2 + Minor\ axis^2)$ $I_{yy}^{ellipsoid} = \frac{1}{5} \times Mass \times (Major\ axis^2 + Minor\ axis^2)$ $I_{zz}^{ellipsoid} = \frac{1}{5} \times Mass \times (Major\ axis^2 + Median\ axis^2)$
Localization Parameters
Average distance to other nuclear centroids
Moments Parameters
Centroid(C_x, C_y, C_z), this is used to calculate Eigen decomposition.
Eigenvalue from Eigen decomposition: λ_1, λ_2, λ_3
Principle axes: I_1, I_2, I_3

Table 4.1: Phenotypic parameters measured for each nucleus.

Morphological Parameters
Total size: Number of foreground voxels in an image stack
Total volume: Sum of the volume of all micro-tissue structures within an image stack
Total surface: Sum of the surface of all micro-tissue structures of an image stack
Thickness: Thickness of a 3D bounding rectangular box for micro-tissue network
Sphericity: Sphericity(binary image stack) $= \frac{\pi^{\frac{1}{3}} \times (6 \times \text{Total volume})^{\frac{2}{3}}}{\text{Total surface}}$
SAV: SAV(binary image stack) $= \frac{\text{Total surface}}{\text{Total volume}}$
Volume of convex hull of micro-tissue network
Surface of convex hull of micro-tissue network
Solidity: Solidity(binary image stack) $= \frac{\text{Total volume}}{\text{Volume of convex hull}}$
Topological Parameters
Total number of End-point, Junction-point, Triple-point and Quadruple-point: Triple point is one kind of junction point which has 3 skeleton voxels in neighbor . Quadruple point is one kind of junction point which has 4 skeleton voxels in neighbor.
Total number of branches: Branch is defined as the part of skeleton between junction points, end points, or junction point and end point
Total length of all branches
Moments Parameters
Centroid of a micro-tissue network, which is used to calculate Eigen decomposition
Eigenvalue from Eigen decomposition: λ_1, λ_2, λ_3
Principle axes of a micro-tissue network: I_1, I_2, I_3

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

4.2.3 Validation by comparison to the confocal laser scanning microscope results

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

Image analysis on the wide-field microscope images was performed using our pipeline (Figure 4.7b, 4.7d). To define individual nuclei and multi-cellular structures correctly in the confocal microscope image stacks, we performed automated segmentation methods and then the segmentation results were validated by visual inspection.

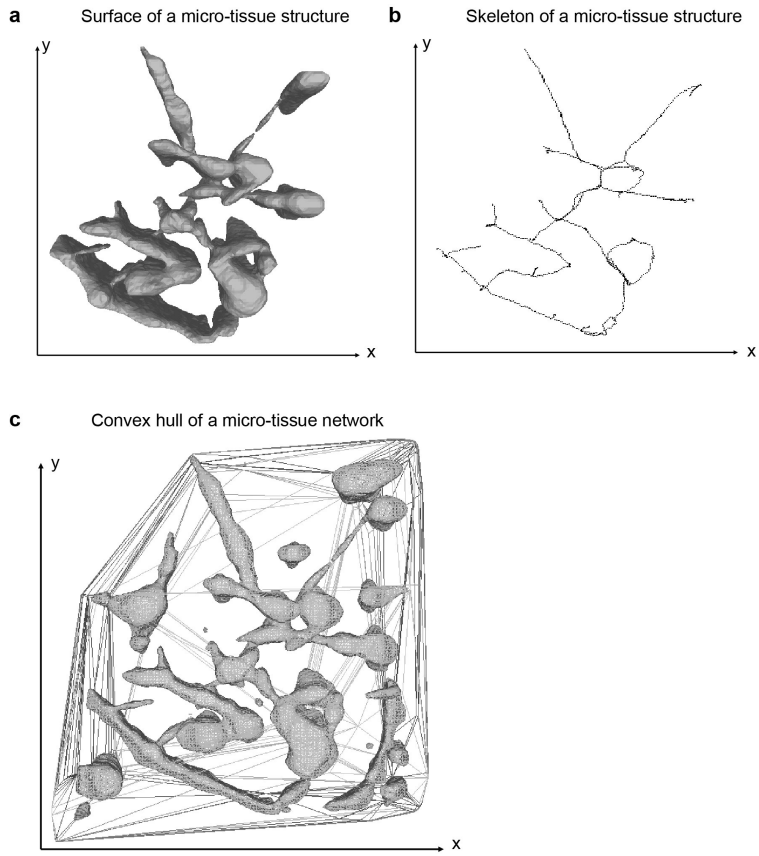

Figure 4.6: Morphological properties for micro-tissue structure and network. (a) Surface of a micro-tissue structure, and (b) corresponding skeleton. (c) Convex hull of a micro-tissue network which contains the structure shown in (a).

Before segmentation, we applied the rolling ball algorithm (radius = 50) and a median filter (kernel size 3×3 pixels) to the image stacks, to remove uneven background illumination and image noise. To segment individual nucleus, we performed Otsu segmentation method [22] on each image slice and then used a 2D watershed to divide connected nuclei (Figure 4.7c). Over-segmentation was corrected manually. For the segmentation of the rhodamine channel images, we applied a log K-means clustering algorithm, which firstly transformed images by taking the natural logarithm and then performed K-means clustering to define multi-cellular micro-tissue structures (Figure 4.7e). A sequential labeling algorithm based on 18-connected neighbor pixels was used to label each nucleus and micro-tissue structure after segmentation. Similar to the correction of elongated nuclei in the wide-field microscope images, we imaged blue fluorescent-labeled microspheres with a 10μm diameter for the calculation of a normalization factor using the same confocal microscope settings as for imaging the micro-tissues (Table S4.2). A normalization factor $\frac{1}{5}$ was obtained for the confocal microscope and then used to calibrate dimensions of nuclei in z-direction. Finally, we

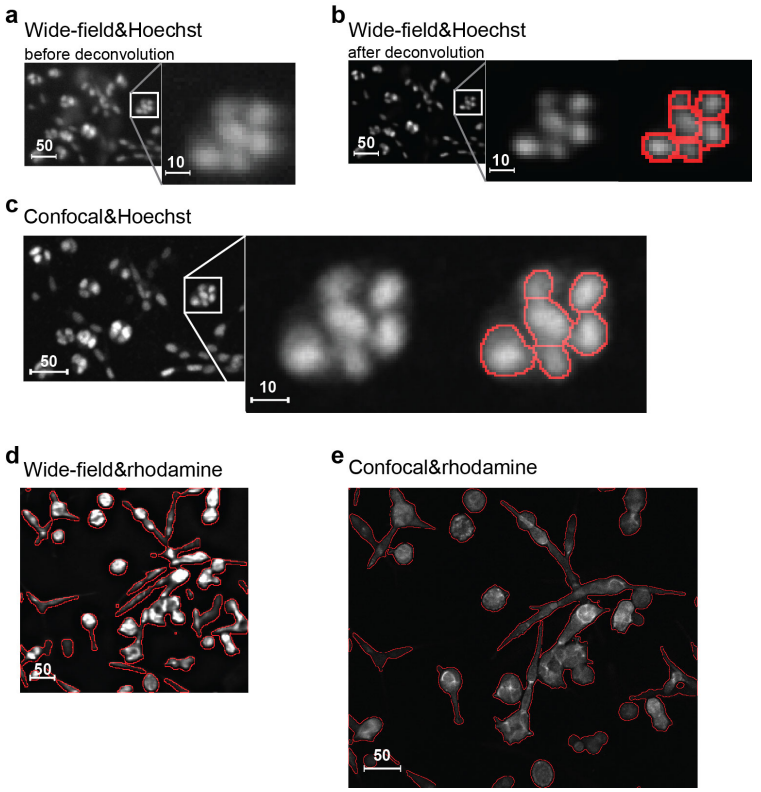

Figure 4.7: Comparison of wide-field microscope images with confocal microscope images. (a) A cropped segment of an original image slice obtained from the wide-field microscopy and Hoechst stained nuclei channel. A part of segment (white box) is magnified 5× to provide a clear perception of resolution. (b) The same image part after deconvolution. Red marks the contour of segmentation result. (c) The same field was cropped from an image slice obtained by the confocal microscope, and the contour of segmentation result is presented in red. (d) A cropped segment of a deconvolved image slice with segmentation result for rhodamine stained F-actin channel. (e) Corresponding region and segmentation result from the confocal microscope image. The unit of the scale bar is μm.

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

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

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

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

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

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

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

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

Next, we calculated Pearson's correlation for each of the parameters that are presented in the Table 4.2. Figure 4.9b shows the coefficients for all parameters of which each has a high value, indicating the quantification results of multi-cellular micro-

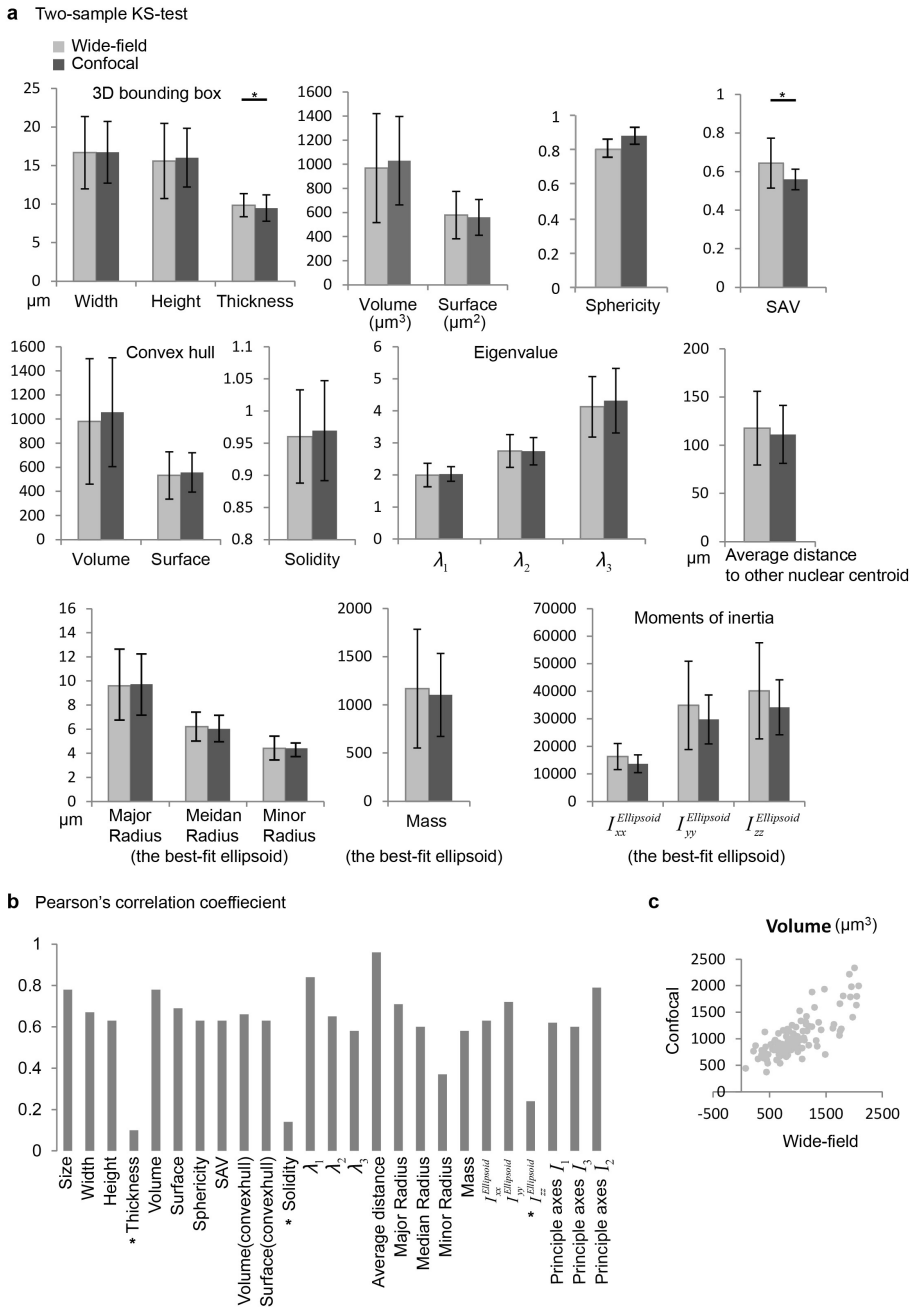

Figure 4.8: Comparison of the quantification results of the nuclei obtained from the wide-field microscope images and the confocal microscope images. (a) The result of the two-sample KS tests, comparing the quantification result between wide-field microscope and confocal microscope for the Hoechst stained nuclei channel. (b) Pearson's correlation coefficient between the quantification result from the wide-field microscope images and the results from the confocal microscope images. * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated. (c) The scatter plot of volume for each nucleus. x-axis represents the value from wide-field microscope images and y-axis represents the corresponding value from confocal microscope images.

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

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

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

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

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

We also recorded the image acquisition time for the BD Pathway microscope system with the same settings as in Table S4.1, but now with laser scanning confocal mode: 105 minutes were required to collect a two channels of image stack from one well which compares very unfavorable to the image acquisition time in the wide-field

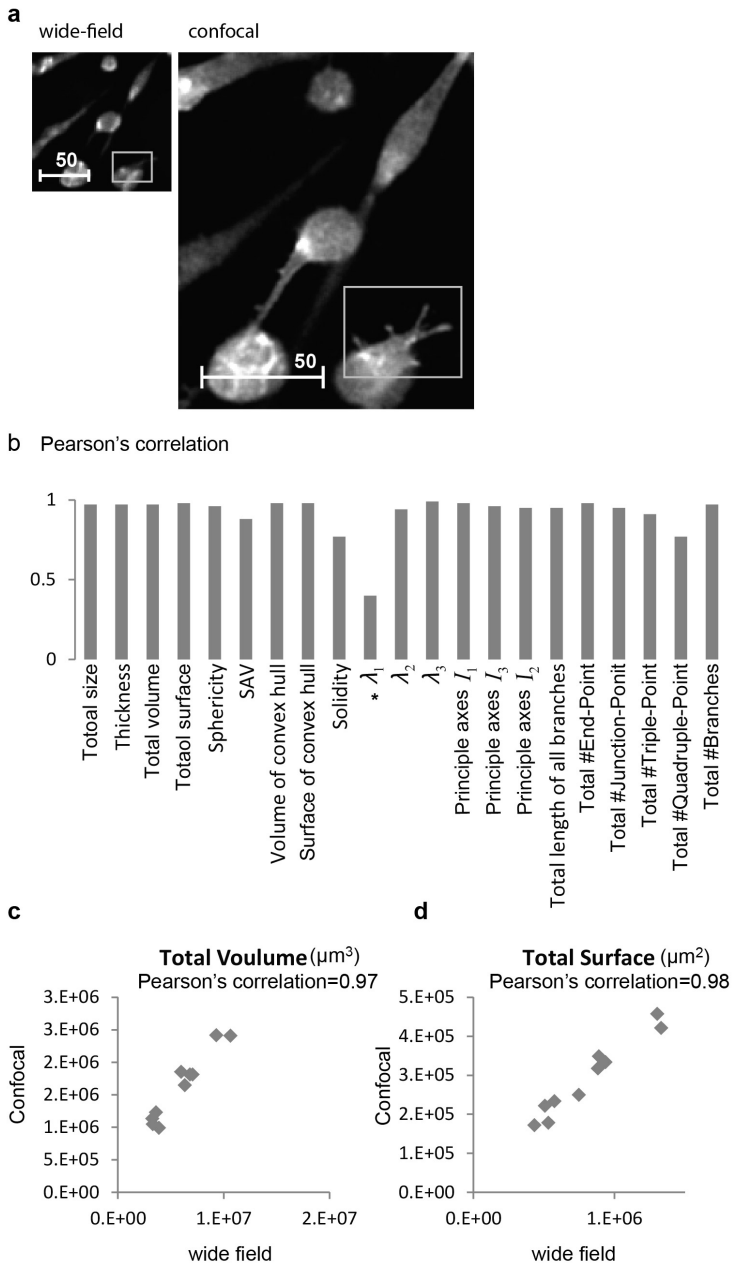

Figure 4.9: Comparison of the quantification result of the micro-tissue networks reconstructed from the wide-field microscope images and the confocal microscope images. (a) One crop of the same field from a wide-field microscope image slice and a confocal microscope image slice. The white box high lightens the regions with different level of details. The unit of the scale bar is μm. (b) Pearson's correlation coefficient between the quantification result from the wide-field microscope images and from confocal microscope images. * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated. (c) The scatter plot of total volume and (d) the scatter plot of total surface of micro-tissue network for each image segment. For both (c) and (d), x-axis represents the value from wide-field microscope images and y-axis represents the corresponding value from confocal microscope images.

mode (150 seconds).

4.2.3.4 Assessment of different sampling step sizes in z-direction

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

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

4.3 Discussion and conclusions

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

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

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

microscope and from confocal microscope as a major drawback.

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

4.4 Methods

4.4.1 Cell culturing

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

4.4.2 Image acquisition

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

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

4.4.3 Software

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Supplementary figures

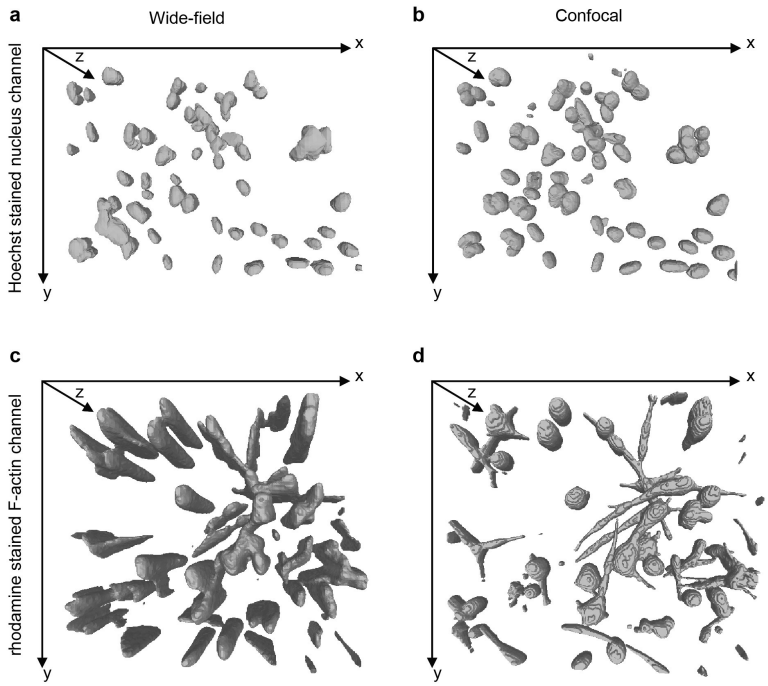

Figure S4.1: Comparison of reconstructed nuclei and micro-tissue structures based on the wide-field microscope images and the confocal microscope images. For the Hoechst stained nuclei channel, the reconstruction results of the same field of a well based on (a) the wide-field microscope image and (b) the confocal microscope image. The corresponding original images are shown in Figure 4.7a and 4.7c. For the rhodamine stained F-actin channel, the reconstruction results from (c) the wide-field microscope image and (d) the confocal microscope image. The corresponding original images are shown in Figure 4.7d and 4.7e.

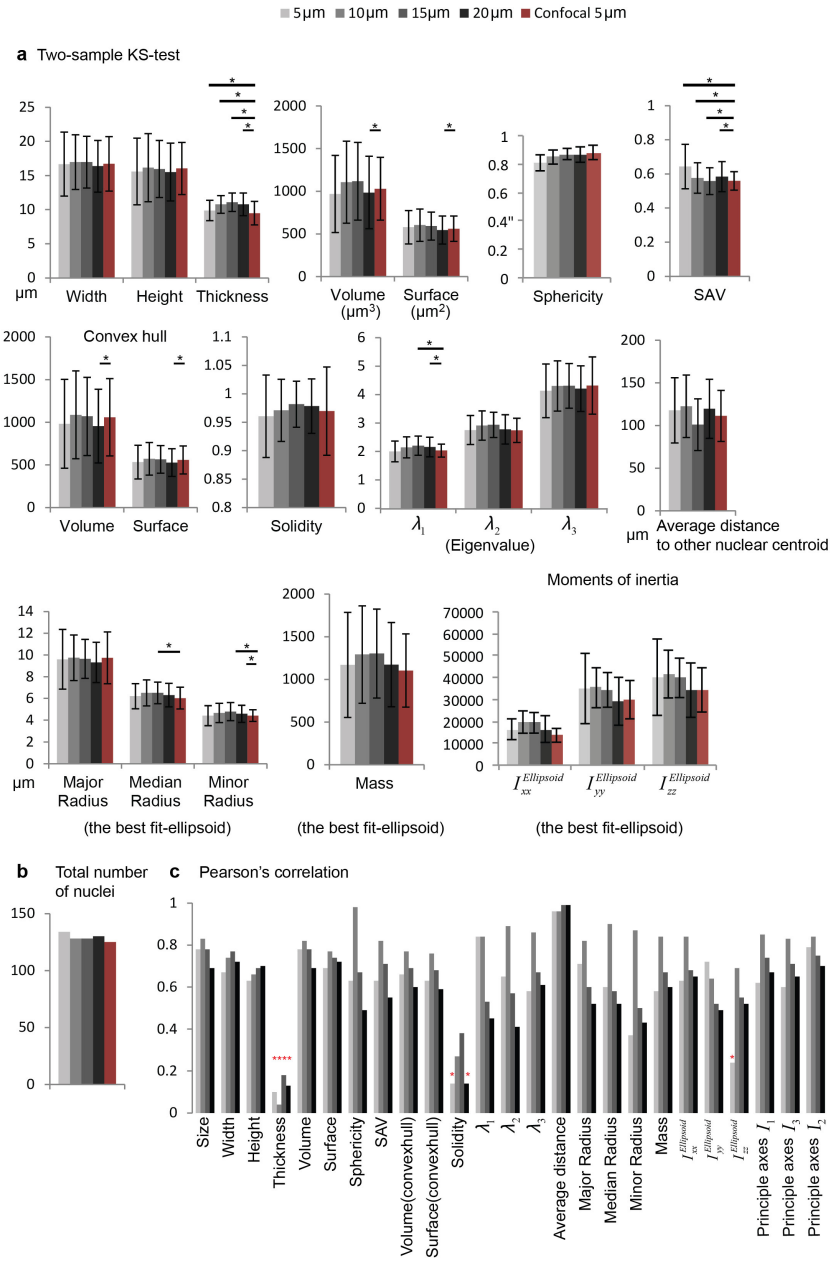

Figure S4.2: Comparison of the quantification results of the nuclei obtained from the different z-sampling step sizes with the results from the confocal microscope images. (a) The result of two-sample KS test, comparing the quantification results of different z-sampling step sizes to the results from the confocal microscope images for the Hoechst stained nuclei channel. For the confocal microscope images, the z-sampling step size is 5μm. For the wide-field microscope images, different sampling sizes are colored in different shades of grey. *: $p - value < 0.01$. (b) Comparison of the total number of nuclei obtained from 5 test segments between the wide-field microscope and the confocal microscope. (c) Pearson's correlation coefficient between each of the quantification result from the wide-field microscope images with different z-sampling step sizes and the results from the confocal microscope images. Red * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated.

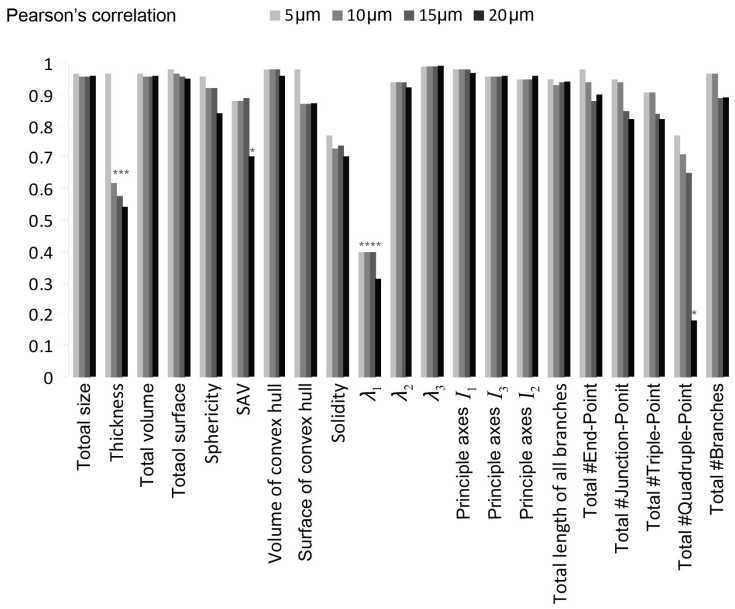

Figure S4.3: Comparison of the quantification results of the micro-tissue network obtained with different z-sampling step sizes and the results from the confocal microscope images. For the rhodamine stained F-actin channel, the Pearson's correlation coefficient between the quantification result from the wide-field microscope images with different z-sampling step sizes and the results from the confocal microscope images. For the confocal microscope, the z-sampling step size is 5μm. For wide-field microscope, different sampling sizes are colored in different shades of grey. * represents $p - value > 0.01$ under the hypothesis that two data samples are not correlated.

Supplementary tables

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

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

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

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