UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

High-throughput Quantification and Inference of Bio- Imaging Data

Permalink

https://escholarship.org/uc/item/4tx9j279

Author

Qin, Qin

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

High-throughput Quantification and Inference of Bio-Imaging Data

A Thesis submitted in partial satisfaction of the requirements for the degree Master of
Science

in

Bioengineering

by

Qin Qin

Committee in charge:

> Professor Yingxiao Wang, Chair
> Professor Xiaohua Huang
> Professor Kun Zhang

2015

The Thesis of Qin Qin is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

<div style="text-align: right">Chair</div>

University of California, San Diego

2015

DEDICATION

This thesis is dedicated to my beloved parents, their love and patience help me overcome all challenges.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWDGEMENTS

ABSTRACT OF THE THESIS

High-throughput Quantification and Inference of Bio-Imaging Data

by

Qin Qin

Master of Science in Bioengineering

University of California, San Diego, 2015

Professor Yingxiao Wang, Chair

Molecular interactions and functions in single cells are largely dependent on their subcellular location and local environment. The molecular activities in live cells within a population are usually heterogeneous and dynamic. Therefore, it is crucial to develop high-

throughput live-cell imaging and quantification methods to interpret the comprehensive landscape of molecular events in space and time. Here we present an efficient and user-friendly software package, *Fluocell*, to quantitatively analyze the dynamic sequences of FRET imaging data with subcellular resolutions. In addition to FRET image visualization and high throughput subcellular analysis, *Fluocell* also provides group statistics and kinetic analysis, as well as 3D re-construction functions of live-cell FRET images. Therefore, *Fluocell* provides an integrative environment for live-cell imaging analysis. The goal is to provide a standardized analysis pipeline, toward the ultimate goals of constructing comprehensive molecular maps with subcellular and dynamic resolutions in all cell types.

INTRODUCTION

Microscopic technologies have long been used for cell studies. No matter how sophisticated the optical design is, the resolution of light microscopic images depends on the wavelengths of nature light, which are not small enough to record molecular activation and interaction. In the last decades, scientists have discovered another innovative imaging method based on Fluorescence Resonance Energy Transfer (FRET), which is capable of bypassing the ultimate resolution limit imposed by nature of light[1, 2]. The flourish of FRET imaging technologies provides biologists powerful tool to investigate protein activities and protein-protein interactions, which is critical for cellular functions. FRET essentially is a near-field transfer of energy from an excited fluorophore (donor) and another fluorophore (receptor) through non-radiative dipole-dipole coupling[3, 4]. This phenomenon only occurs when the molecules concerned are in very close proximity, typically in the range of 10 nm and less. In addition, the efficiency of this energy transfer is inversely proportional to the sixth power of the distance between donor and acceptor. The extreme sensitivity of FRET to small distance change makes it possible to track of molecular interactions [5]. The schematic display of FRET principle and construction is illustrated in Figure 1[2]. The overlap between the donor's emission spectrum and acceptor's excitation spectrum, an appropriate distance between two fluorescence proteins and relative orientation between the donor and acceptor are crucial to produce effective FRET. In the Figure 1, the FRET donor and acceptor have been depicted with parallel orientation. Meanwhile, anti-parallel orientation has also been shown to facilitate energy transfer for certain FRET pairs [6].

As a robust platform, the FRET principle allows the design and fabrication of biosensors for the detection of intracellular events. An emerging use of the FRET-based biosensor is to apply them in combination with a second biosensor in order to accomplish task of imaging multiple biochemical molecules in living cell simultaneously. In the Figure 2[7], several types of FRET biosensors have been presented to demonstrate different strategies utilizing FRET. A variety of FRET biosensors allow researchers to investigate diverse cellular events such as stem cell differentiation, cancer metastasis, and immune responses.

Though FRET permits high spatial resolution imaging of molecular activities and interactions, the signals are relatively small. The measurements requires careful interpretation and several control experiments, which makes the later image analysis vital[8]. When the experiments are designed appropriately, effective and precise quantification and analysis of image data facilitate the hidden biological theories to be revealed. Nowadays, it is easy to collect thousands of high resolution images from time-lapse experiments in a single day in diverse scale. However, analysis of tremendous image data becomes bottleneck and limits the boost of biological discoveries[9]. Experienced biologists are capable to examine and interpret several types of data by eyes, such as the image cytometry. This ability is highly appreciated but not suitable for high throughput fluorescence microscope image data analysis.

The fluorescence image data is usually acquired by commercial microscope control software and the final data sets express as format of JPEG, TIFF and others, which are difficult to be tracked in time and quantified by human eyes. Since the acquisition tool is

not intentionally designed for sophisticated image quantification and analysis, it can only handle relative simple and common task without specificity. Effective analysis on the high-resolution imaging becomes an urgent need in biological studies. Molecular interactions and functions in live cells are largely dependent on their subcellular location and environment[10, 11]. Furthermore, the molecular activities in a single cell within a population are usually heterogeneous and dynamics, with the cell-cell variation likely due to stochastic subcellular molecular wiring in structure and function[12, 13]. Therefore, it is crucial to develop high-throughput FRET imaging and mathematical analysis methods to sample and inference the behavior of the whole population.

Many open source software tools have recently been developed in the field of bioimage informatics[14-24]. Some of these software packages are free and allow communication between software programmers and scientific researchers. As a result, these software packages have provided tremendous help for the research community. The CellProfiler[9], which has been supported by NIH and other institutes for around 10 years, can be used to automatically identify cellular features. In addition, the software package ImageJ[23] among others, has assisted imaging processing to many researchers. However, there is currently no project specified for the quantification and analysis of the FRET live-cell image sequences.

Here we introduce the open-source *Fluocell* software package, we hope the development of this software system can benefit the scientific community in the quantitative analysis of live-cell fluorescence and FRET video sequences. *Fluocell* is providing convenient visualization of the ratiometric FRET images, a variety of

quantification methods and statistics tools allowing high-throughput analysis. *Fluocell* was original developed by Drs. Lu and Wang as a research software for live-cell imaging analysis [25, 26]. For my thesis research, I have worked to further develop and improve new functionalities of *Fluocell*. In collaboration with Drs. Lu and Wang, my work has focused on improving the graphic user interface of *Fluocell*, developing the Quanty 1.0 module for high throughput and high resolution multi-cell analysis, and utilizing this software package for specific analysis applications needed by collaborators. For Fyn kinase signaling, Dr. Ouyang developed a new Fyn FRET biosensor and performed imaging experiments. This new Fyn biosensor has been extensively characterized to show that the FRET response is sensitive and specific toward Fyn kinase activity in live cells. I performed data analysis and provide interpretation of the biological significance for in this Fyn kinase signaling study. In this thesis, we do not focus on either the technical details of the software, or computational validation of the published algorithms. Rather, we describe the system, validation of diverse functions, and how it is able to assist researchers to solve real-world biological problems.
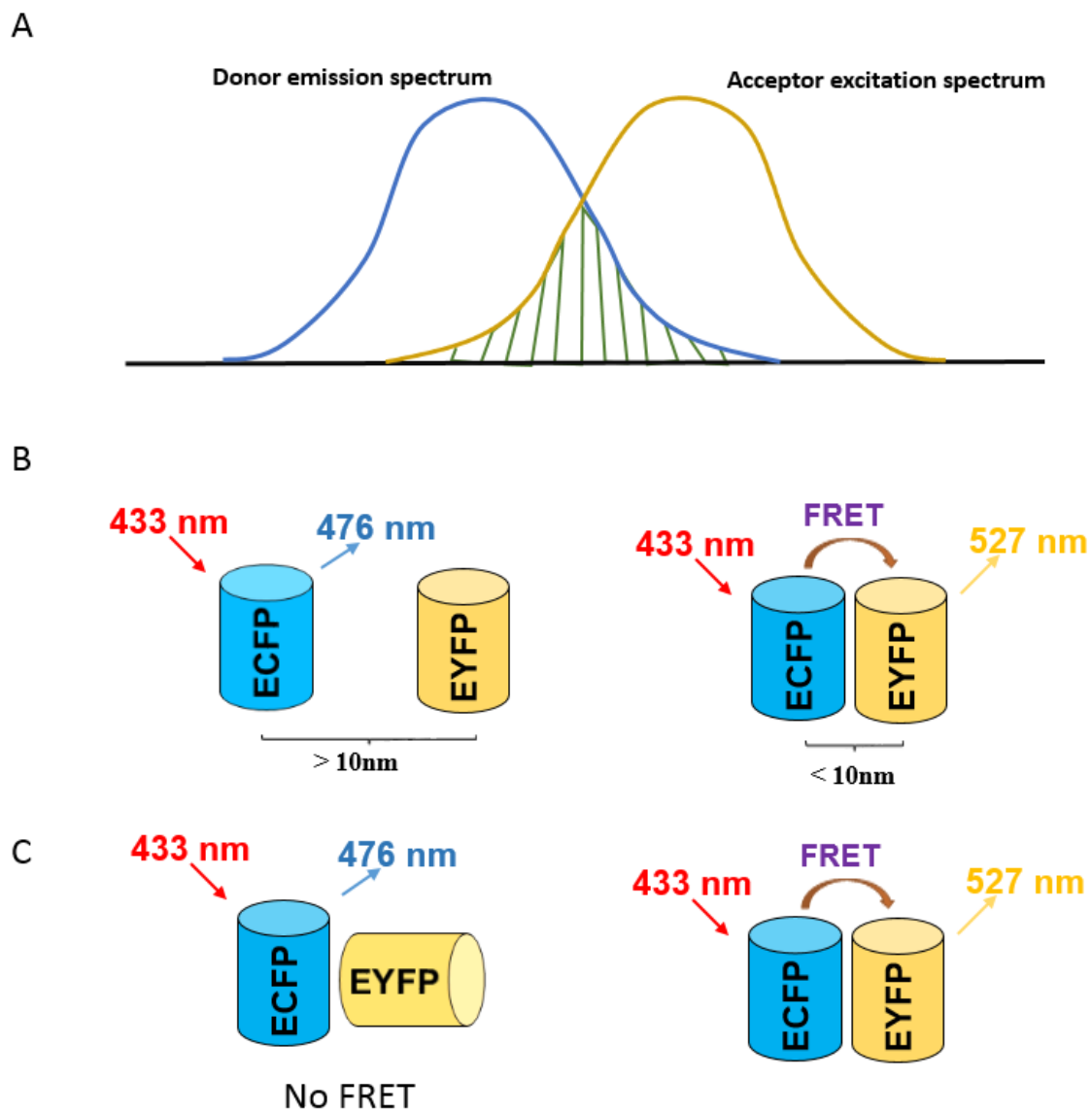
Figure 1. Schematic diagrams depicting the three conditions to produce effective FRET. To generate FRET, there should be **(A)** overlap of donor emission spectrum and acceptor excitation spectrum, **(B)** proper distance between donor and acceptor fluorescence protein and **(C)** appropriate orientation of two fluorescence proteins.
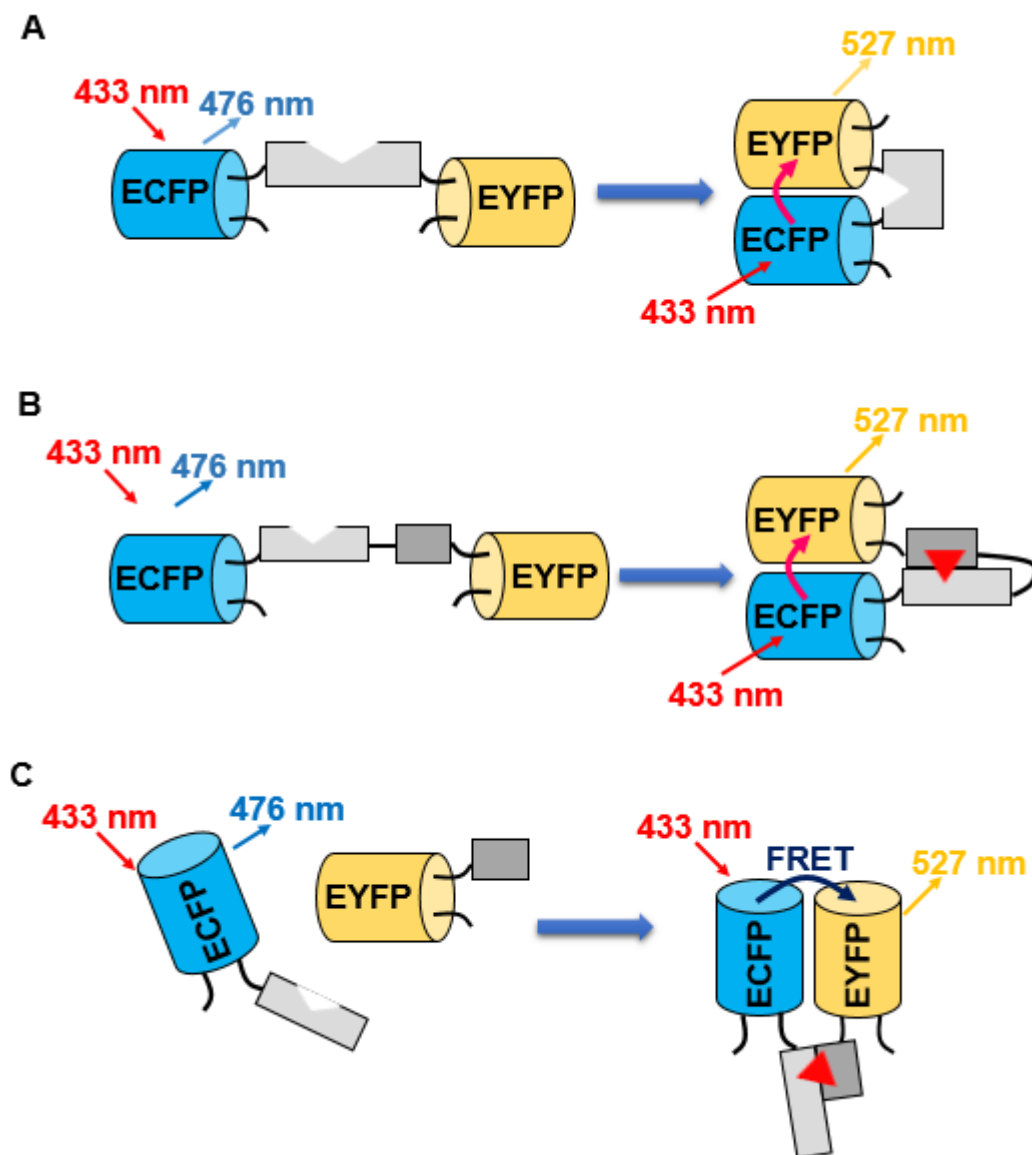
Figure 2. A variety design strategies of fluorescence resonance energy transfer (FRET) biosensors. (A) (B) Intramolecular FRET: conformational change of donor and acceptor or covalent fusing to produce FRET; **(C)** Intermolecular FRET: fluorescence proteins fusing (substrate binding) with each other to produce FRET.

RESULTS AND DISCUSSION

Overview of *Fluocell*

For the living-cell FRET image analysis, most tasks can be accomplished by interacting with *Fluocell*'s graphical user interface (Figure 3A). An analysis pipeline concept is shared by *Fluocell* through connecting each individual functional module (Figure 3B). As shown in the schematics, each module deals with the image data sets at one processing step. All modules are placed in an order to create a typical analysis pipeline: (1) the preprocess of the images, (2) image data visualization, (3) feature detection and tracking, (4) quantification and statistics analysis of the data. In the whole pipeline, though most modules will work in an automatic fashion, for feature detection and tracking, *Fluocell* provides multiple options for users to choose, such as specifying a region of interest in a typical study.

By travelling through such pipeline, the image data will be explored in a comprehensive way. All processes are repeatable and output from each module can be extracted for further quantification on *Fluocell* or new analyses on other computational tools. And all the settings can be saved for reproducing the analysis. Subsections of the *Fluocell* will be introduced in detail subsequently.

Preprocessing of image

It is inevitable to include significant background noise in the raw image data, which imposes obstruction on accessing real quantitative difference. Thus, the subtraction of the

background then becomes one critical step. *Fluocell* software package provides two background subtraction options, manual selection of the background by our users or automatic selection by *Fluocell* (Figure 4A). When dealing with multiple-cell quantification, the users can save a lot of time by using this auto-selection option. In addition to the subtraction of background, another important step in image preprocessing is smoothing by a filter. In the *Fluocell* software system, it has provided a nonlinear digital filtering option, median filter, to remove the noise. The superposition of these two steps will produce a relative clean image which benefits later processing. Sometimes we may have more than one cell in the field of view, which obscures the detection and produce ambiguous result. There is a cropping image option in *Fluocell* to solve the problem. Our user can define a polygon to crop out the target cell for later quantification and analysis, instead of using the whole image.

Image visualization and feature detection/tracking

In addition to the FRET ratio channel, other channels will be shown using the MATLAB function *imshow()* and *imagesc()*, displaying cells in the RGB color space. Since the FRET ratio number is relative small, *Fluocell* maps the ratio value to a pure color scale in the function *get_imd_image()*.The intensity image of donor and acceptor channels is also utilized as a critical aspect in visualization, so that users will have a lighter and more clear FRET ratio image, which we called Intensity Modified Display (IMD). The displayed ratio images can be adjusted by modifying the intensity window and ratio window on *Fluocell*

interface (Figure 3A). Good quality of IMD ratio image will help us better understand cell morphology and the spatial distribution of molecular activities inside cells.

At the next step of the image process pipeline is the feature detection and tracking. Object identification is the most challenging part in image analysis and the accuracy of its measurement determines the accuracy of cellular activities measurements. *Fluocell* utilized a published and well tested algorithm to do this work. The function *my_graythresh()* of *Fluocell* calculates the global image threshold (level) based on Otsu's method[27]. This threshold will help find the cell edge and then our target is finally located. *Fluocell* has provided an option for showing the detected boundary so user is aware which area is being tracked and used for quantification.

Quantification and statistics interpretation

Fluocell measures the fluorescence intensity of each identified cell or subcellular region, such as user-specified region of interest or layers. Since FRET leads to a ratiometric measurement of molecular interactions, *Fluocell* computes the ratio between two FRET pair channels pixel by pixel, then output an average value within the region of interest. The ratio value and time information of each frame will be saved in the MATLAB workspace and can be easily copied to the Microsoft Excel format for external access. Further quantification and analysis can continue in *Fluocell* extension package – quanty 1.0 or other computational tools. Quanty 1.0 has a variety of tools for viewing, plotting data and statistical analysis. Because of its property as a function extension package, all the data format are well compatible. While providing powerful quantification and analysis of image

data, it will save much time without repeat exporting and importing intermediate results of *Fluocell*.

Multiple-cell detection and quantification

With more sophisticated and advanced technique in the field of bioimaging, modern study on cell activities has expanded to average behavior of group of cells instead of individual ones, which is more accurate and representative of cell dynamic kinetics. Because of this, multiple-cell quantification is a necessary and urgent need to meet on imaging analysis. In the *Fluocell* software system, we have developed an extension package - quanty 1.0, fitting such need for biological study. After automatic acquisition of imaging on multiple positions through fluorescence microscopic tools (Figure 3C(i)), *Fluocell* is able to help biologist to sort all the image files to be ordered (Figure 3C(ii)) which will benefit later quantification and analysis of the quanty 1.0 package or other computational tools. All the sorted image files/data travel through the typical process pipeline of *Fluocell* as described or a one designed by users themselves, then the quantification result will be virtualized through the viewing and plotting tools in quanty 1.0, such as the time courses of ECFP/FRET intensity ratio plot (Figure 3C(iii)).
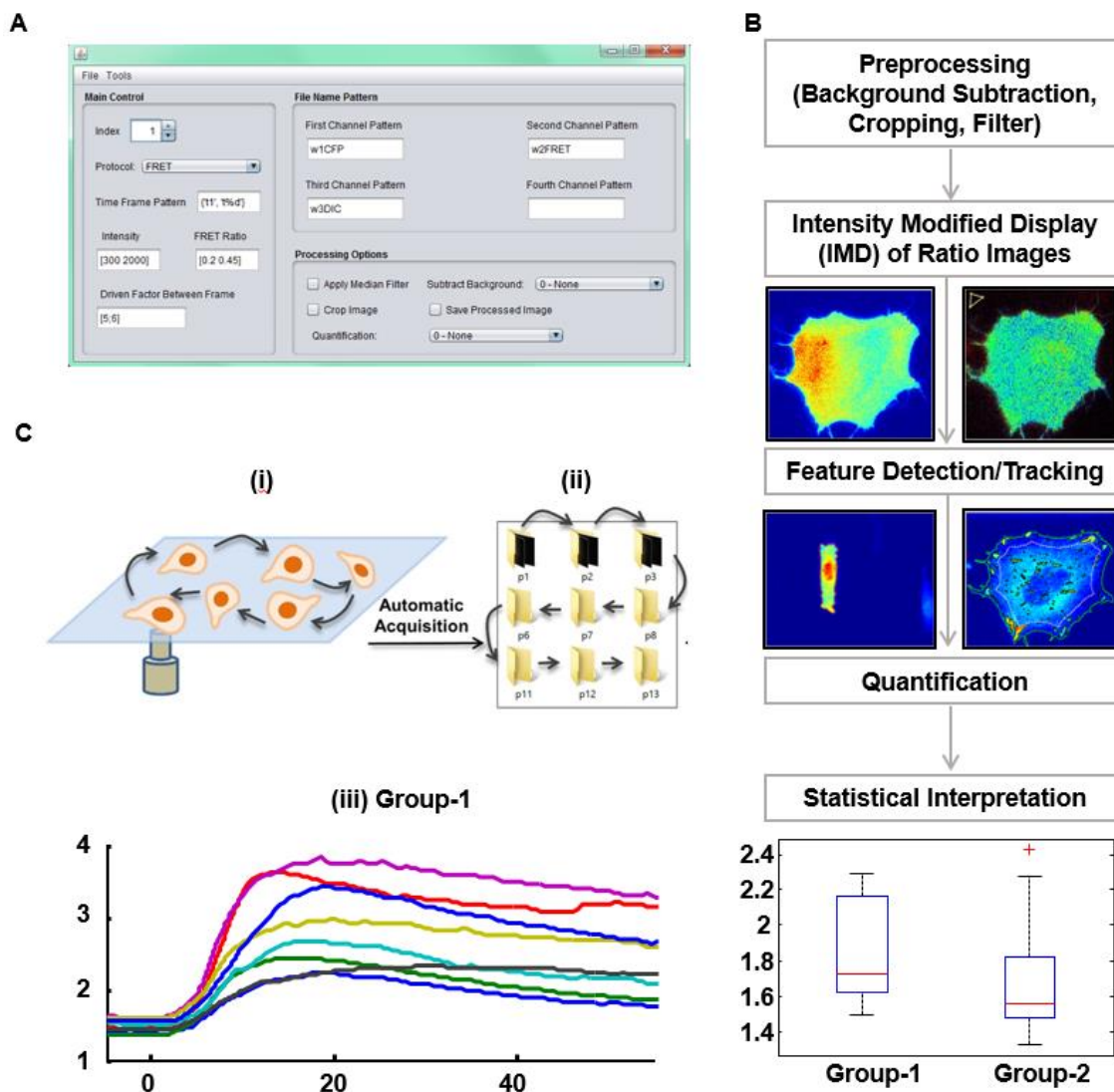
Figure 3. *Fluocell* **overview and features.** (**A**) A screenshot of the main *Fluocell* interface; (**B**) Schematic of a typical image visualization and analysis pipeline of *Fluocell*; (**C**) The multiple position quantification feature (**i**) the time-course image data of multiple cells are obtained by microscope with an automatic stage; (**ii**) the image files for each cell were sorted into a folder; (**iii**) *Fluocell* recognizes the folder pattern and automatic quantify the time course of different cells. The time courses are collected and plotted together.

Broad applicability

The *Fluocell* software system is originally designed to help programming-agnostic biologist to solve image processing and data analysis problems, especially the image data from fluorescence microscope. It is able to convert complex signals into interpretable formats like figures or plots that experimenters can easily understand. To achieve that goal to be a useful and functional tool, efficiency and accuracy are two important aspects addressed by *Fluocell*. The raw image data coming from microscopic system is usually noisy, which hinders later quantification. To accurately extract information out of the image data sets, subtraction of the background becomes one curial step. In the software system, it has provided two methods for users to eliminate the background noise, the manual selection of the background conducted by the user or the auto selection by *Fluocell* (Figure 4A). In the auto background option, the program will calculate the background intensity in the small rectangle at four corners of the image, then determines the one with smallest value as the background. As shown in the comparison plots (Figure 4B), two methods have quite consistent result and it can prove the accuracy of the auto option in *Fluocell*.

When studying molecular interactions and functions in single cells, users can draw a number of region of interest (ROI) on the subcellular region, so later analysis will be able to reveal dynamics of local activities (Figure 4C(Left)). Alternatively, *Fluocell* provided another way to study subcellular activities by separating cells by layers and all the quantification is based on those layers (Figure 4C(Right)). When it relates the cell junctions study, usually the molecular interactions at cell boundary has profound effect. The layer

method and its corresponding quantification on the outmost layer will produce a desired indication of cellular boundary activities. These two ROI selection strategies are able to track subcellular activities at different scale and location in a same cell (Figure 4D).
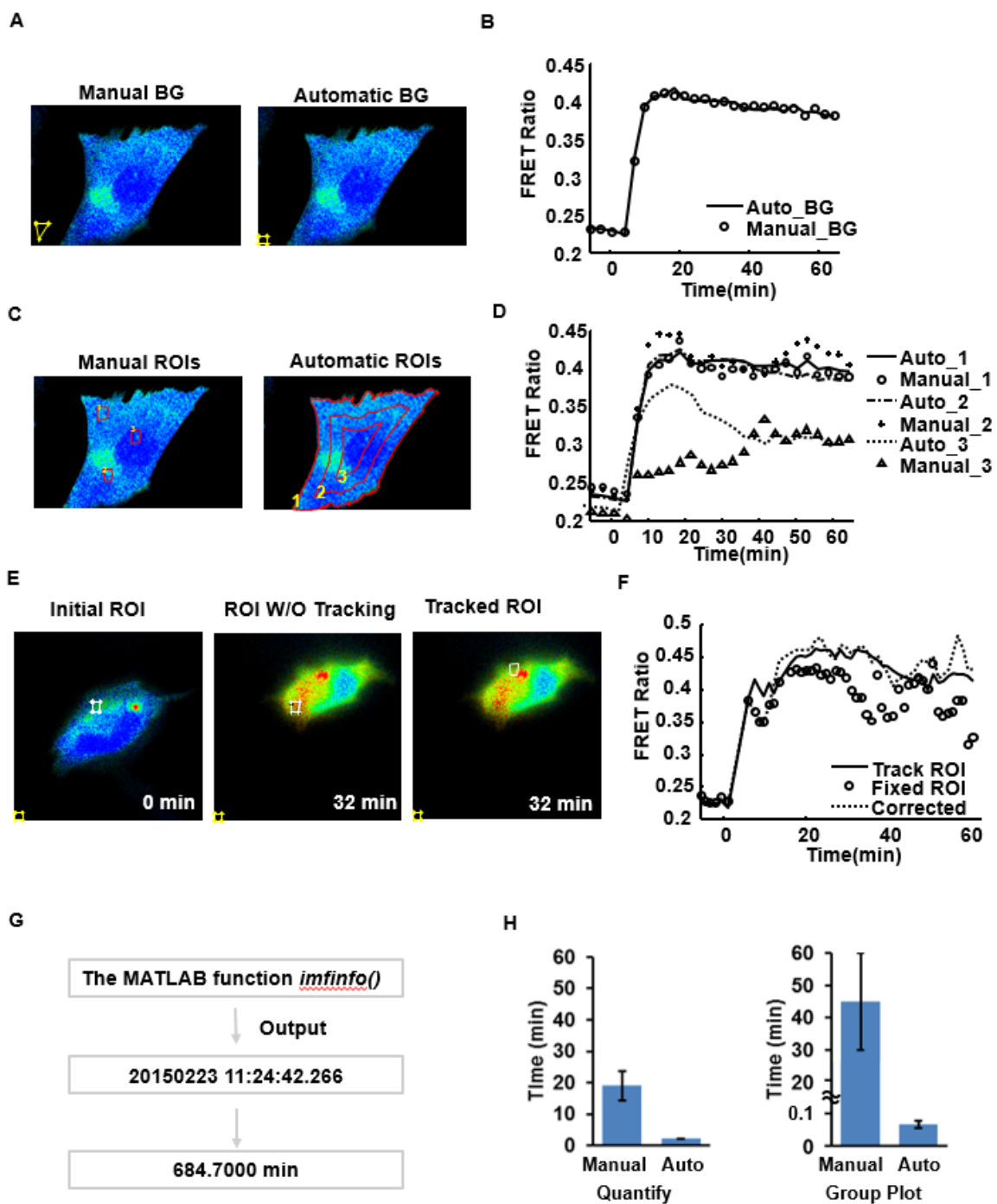
Tracking and time system

In some biological studies, researchers try to simulate the intracellular molecular interactions by adding special stimulus like drugs or traction force, which might introduce migration of cells. In *Fluocell* software package, there is a tracking system which kept adjusting the position of user-specified ROIs during all image frames if the identified cell moves (Figure 4E). With this tracking option, the ROIs biologists first specified will almost be the same regions relatively to the whole cell no matter how cell migrates. The time course of ECFP/FRET ratio plot (Figure 4F) showed the difference of quantification results under tracking ROI, Fixed ROI and Corrected ROI methods (Figure 4E). The Corrected ROI method indicated a manual adjustment of ROI conducted by the user frame by frame, which is quite time-consuming and doesn't guarantee the accuracy. Molecular interactions are complicated and those activities usually have mutual connections and occur in specific orders, where the time cue of each event is extreme important for comprehensive understanding of cell dynamics. By invoking the *imfinfo()* function of MATLAB in the *Fluocell* function *get_time()*, the actual imaging time in minutes from 12 am of the day is retrieved and took into calculation. In the later quantification, all the analysis is based on the real time information and the time resolution of molecular interactions is largely kept.

High efficiency of *Fluocell*

Like the most new software packages in the laboratory, it may take several days for users to set up the system and get familiar with the analysis pipeline. We have attached a detailed version of instruction on *Fluocell* which provides detailed guidance on all the function. Along with it, a few typical examples and related raw data sets are provided with the package. Additionally, the flexible and friendly point-and-click user interface helps to simplify the learning process for non-programmers. Over time, experienced user is able to set up the analysis pipeline and go through it in only a few hours for their diverse biological study. There is a time comparison between manual method (MetaFluor) and automatic quantification and analysis of *Fluocell* (Figure 4H). Three experienced biological researchers have provided the time length they need to spend on one experiment analysis, where there are 30 cells and 24 frames for each cells. For the quantification on each single cell, they need 19 minutes in average. Then they need another 45 minutes to export the each cell's result out to other analysis tool and make the group quantification and plot. However in *Fluocell*, the quantification part can be finished only in 2.2 minutes, which is almost 9 times less. Additionally, the time spend on group plot part is nearly 650 times less than manual methods. The big difference on imaging processing has demonstrated the efficiency of *Fluocell* as a powerful analysis tool of fluorescence image data. Processing imaging data using *Fluocell* also reduces the amount of energy needed from the researcher the stress imposed.

Figure 4. The efficiency and accuracy of *Fluocell* analysis. (A) The background regions are shown at the bottom left corner (solid yellow polygons) of ECFP/FRET ratio images of a cell. A background region **Left: manually** selected by the user; **Right:** automatically generated by *Fluocell*; **(B)** compare the time courses of ECFP/FRET ratio averaged in the whole cell, based on the background region chosen in (A); **(C)** Regions of interest (ROIs in solid red lines) overlaid with the ECFP/FRET ratio images of the cells. The ROIs are indexed in yellow. **Left:** ROIs manually selected by user; **Right:** ROIs automatically generated by *Fluocell* according to their distance to the extracellular space. **(D)** Compare the time courses of ECFP/FRET ratio averages in the ROIs shown in (C). **(E)** The ECFP/FRET ratio images of a moving cell are shown with different ROIs in solid white quadrilaterals. **Left:** The initial ROI specified by user; **middle:** The same ROI at 32 min when the cell moved; **right:** Tracking ROI at 32 min in this cell; **(F)** Compare the time courses of ECFP/FRET ratio averages in the ROIs shown in (E). The corrected time course was generated by manually moving the ROI following the movement of the cell. **(G)** Automatically retrieve the time of imaging by the *get_time* function. In this function, the output of the MATLAB function *imfinfo* was used to calculate the actual imaging time in minutes from 12 am of the day when the experiments were performed. **(H)** Time comparison between manual (MetaFluor) and automatic quantification by *Fluocell* (30 cells, 24 frames/cell). Left: time used to quantify the ECFP/FRET ratio time course; **right: time used** to plot the time ECFP/FRET ratio time courses based on results from the left panel.

A
Manual BG   Automatic BG

B

C
Manual ROIs   Automatic ROIs

D

E
Initial ROI   ROI W/O Tracking   Tracked ROI
0 min   32 min   32 min

F

G
The MATLAB function *imfinfo()*

↓ Output

20150223 11:24:42.266

↓

684.7000 min

H

Group image visualization

For the multiple-cell experiment of molecular interactions study in live cells, though the numbers and plots can be a big support for biologist to understand the dynamic kinetics of cellular activities, the information conveyed is limited and partial. For example, we don't have any information related to cell morphology, which is also an important feature of those dynamic kinetics. To comprehensively explore the image data, we have included a visualization tool to display the global view of cells at different time points. In the quanty 1.0 extension package of *Fluocell*, there is a function named *group_image_view()* which meets the need for morphological inspection on identified objects. After input the time point user prefers to check (the addition of external stimulus is regarded as the zero minute), the system will generate a group of emission intensity ratio images with pseudo color indicating the FRET ratio value (Figure 5A). If there is no image taking at user specified time point, *Fluocell* will automatically search a nearest time point for checking purpose. As shown, not only the cell morphology can be easily observed, but also the distribution of the activity is easily tracked. In the typical *Fluocell* multiple-cell image analysis pipeline, if there is no abnormality on cell morphology, the next step is usually the single cell quantification through every position. The function *g2p_quantify()* in the system will achieve such goal. The output of this function will be a group of emission intensity FRET ratio plots (Figure 5B), which provides length and scale of the molecular interactions in identified cell of each position.
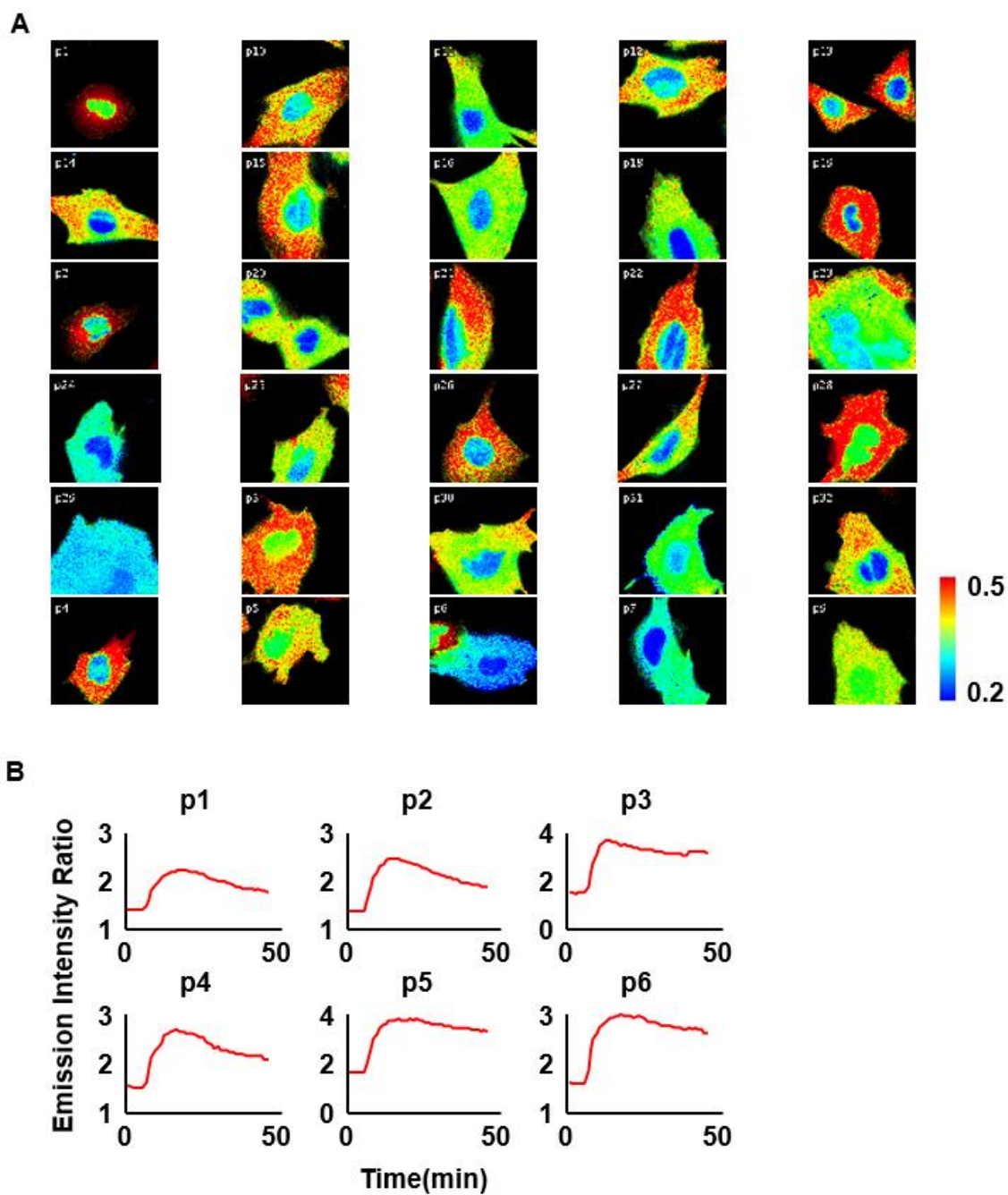
Figure 5. Display the group view of single-cell ratio images and quantified time courses. (A) shows the emission intensity ratio of the cells indicated by pseudo color at a fixed time point; **(B)** shows the emission ratio time courses of six representative cells in the group.

Make movie through *Fluocell*

Image visualization expresses in diverse ways like figures, plots or movies. As one important media broadcasting format, more effectively than equivalent text, rigid image or figure with legends, animation provides a precious way to communicate complex sequences of biological processes (signal transduction)[28-30]. For many biological researchers, they have a preference on showing the study of molecular interactions in animation while performing the presentation, which is more representative of dynamics process. In the system, the *make_movie()* function expands the visualization format of *Fluocell* to movies more than figures and plots. By means of invoking this function, the software system is capable of creating animations for image data manifestation and inspection.

In the graphical user interface (Figure 3A), there is a tab called 'saved processed image', which facilitates our user save all the processed ratio images. After checking this option, there will be ratio images generated and saved in time order during the analysis. In the later task of making movie, all the processed ratio images will be connected together to produce an animation in AVI format with a default name FRET.avi (Figure 6), attached with related time information in the every frame. What is more, the *make_movie()* is available for flexible input, users is able to customize the movie by specifying related parameters like an accurate movie name, stimulus information or a color bar indicating the distribution and level of cellular activities in the movie. In the extension package quanty 1.0 of *Fluocell*, we proposed an advanced version of movie-making function called *group_make_movie()*. After specifying which position we want to make movie on of the

multiple-position data set, this function is able to create a movie in corresponding position

output folder.
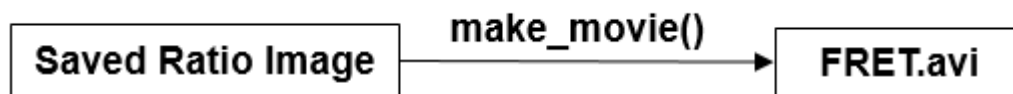
Figure 6. Movie-making function of *Fluocell*. All the saved ratio image will be connected to an animation through make_movie() of *Fluocell*.

Case study: difference behavior of cytosolic and membrane-targeted Fyn biosensor

Fyn is a 59kDa member of the Src family of tyrosine kinases[31]. The structure of Fyn is similar to that of the other Src family members. At the end of C-terminus, there is a tyrosine kinase domain (Src Homology 1, or SH1). Two modulatory Src Homology (SH) domains locate at the center mediating the protein-protein interactions. One of them, the SH2 domain, binds to phosphotyrosine residues and the other one, which is called SH3 domain, can recognize polyproline-rich PXXP-like sequence. By myristate and palmitate, the fourth Src Homology domain, SH4, containing the 'Met-Gly-Cys' motif, dictates dual acylation and membrane binding. There is also a unique sequence between SH3 and SH4 domains[32-34]. Biologists have found that after synthesized and N-myristoylated on the cytosolic polysomes, Fyn is able to rapidly target the plasma membrane and get palmitoylated[31].

As a member of SrcA subfamily[35], Fyn involved in a variety of different signaling pathways and cellular activities. Two decades ago, Fyn tyrosine kinase signaling has been found to play an important role in the initiation of myelination, which is a critical process for neuro networking communication[36]. Also, Fyn involved in the process of T cell receptor signaling, organization of cytoskeleton, cell migration, integrin signaling and so on[37-40]. Because of its importance in cell biology, Fyn becomes one of the most studied Src family kinase member. Biological researchers found the Fyn activity was primarily membrane-bound[41-44]. However, our recent observation from the image data may suggest that beside location on membranes, there is a subpopulation of Fyn activity in cytosol.

To study Fyn kinase signals in live cells and compare the contribution from these two populations, Dr. Ouyang in our group has built two types of Fyn FRET biosensors. One of them is a cytosolic Fyn biosensor and the other one is a membrane-targeted version. In the experiments, Dr. Ouyang has used the Platelet-derived growth factor (PDGF) as the external stimulus and taken the fluorescence images on more than 30 positions in each group, where one position indicates a cell. Later all the quantification and statistical analysis are conducted in the *Fluocell* software system.

Since it is a multiple-cell experiment, we have utilized a typical multiple-position quantification pipeline in data analysis. First, the raw image data sets are sorted into corresponding position folder. In the graphical user interface of *Fluocell* (Figure 3A), there is an item called 'Driven Factor Between Frame', where we input the time index when the external stimulus (PDGF) is added. Another important information *Fluocell* needs for later quantification is the number of images taken in each position, which will help the system quantify through the right frames. At this point, all the pre-required information is input via the graphical user interface and the rest of the analysis will be finished in MATLAB command window.

First, the function *g2p_init_data()* will pass all the parameters from *Fluocell* interface to MATLAB workspace. Then, the program will calculate a zero time point for subsequent plotting, which is based on the information cue we input in the 'Driven Factor Between Frame'. Then the function *g2p_quantify()* will be invoked to quantify through every frame of each cell. Plots similar to Figure 5B will be created in the MATLAB. In addition, the quantification result of each position is generated and saved in corresponding position output folder. By means of *group_plot()*, *Fluocell* is capable of plotting the

emission ratio time course of all cells tracked in one experiment by using one MATLAB command (Figure 7A). In fact, the output is two sets of plot, one is the original ratio versus time and the other is the normalized ratio versus time. Here we show the normalized time course of emission ratio, which displays the group behavior that can be compared among different experiments.

The statistical analysis is embedded in the function *group_compare()* of the quanty 1.0 package. It can be used to perform several basic statistics on the image data, such as averaging and T test. Given necessary parameters for the statistical analysis, *group_compare()* will generate a group of plots as virtualization of the results (Figure 7B). In Figure 7B (i), it showed the average time course of ECFP/FRET emission ratio of two groups with error bars on them. The error bar is calculated from standard error of mean, which indicates how accurate the estimate of the mean is likely to be. As we can see, the response of cytosolic Fyn biosensor is greater than that of the membrane-targeted version.

The violin plot is an alternative method for plotting statistical data. It is shown as a combination of box plot and kernel density plot, where the kernel density plot is rotated and added at each side of the box plot. The violin plot not only contains the basic information like median, minimum and maximum from the box plot, but also manifest the distribution of the data points (kernel density estimation). Thus, this additional violin plot provides rich information of the application. The rest of Figure 7B are plotted based on three parameters used to compare between the cytosolic and membrane-targeted Fyn biosensor responses. They are the average ratio value between time point 20 – 35 minute, the peak ratio value, and the time length to reach the peak ratio value in each cell within two groups, respectively. It is obvious that the cytosolic group has a stronger response and

it reaches the maximal signal in a shorter period of time. According to the probability distribution in the violin plot, the data points in membrane group have less variability.

Furthermore, the results of t-test between these three statistics are displayed in the MATLAB command window. The test results include the p-value, number of samples in each group and whether a significant difference has been found between two groups. In this particular example, all three statistics are significantly different between the two groups. Thus, the results indicate that the cytosolic Fyn kinase activity responds stronger and faster to PDGF stimulation than the membrane population. Our results suggest that upon PDGF stimulation, there are more active Fyn kinase located in the cytosol, which is reported by the cyto-Fyn biosensor, than that in the intracellular membrane compartment as reported by the Lyn-Fyn biosensor. This result is consistent with earlier reports that Fyn kinase has been found to colocalize with centrosomal and mitotic structures[41, 45]. In addition, Fyn has been shown to promote phosphorylation of the transmembrane adaptor molecule PAG, which recruits Csk, a known inhibitor of Fyn kinase[46]. Therefore, Fyn kinase activation can trigger a negative feedback loop at cell membrane and down regulate its own activity, which may contribute to the relative low Fyn signal observed using our Lyn-Fyn biosensor. Thus, our results provide a clear view of the subcellular localization of Fyn kinase activity: more active Fyn are localized in cytosol of fibroblasts which responds faster and stronger to PDGF stimulation than that in the cellular membrane.
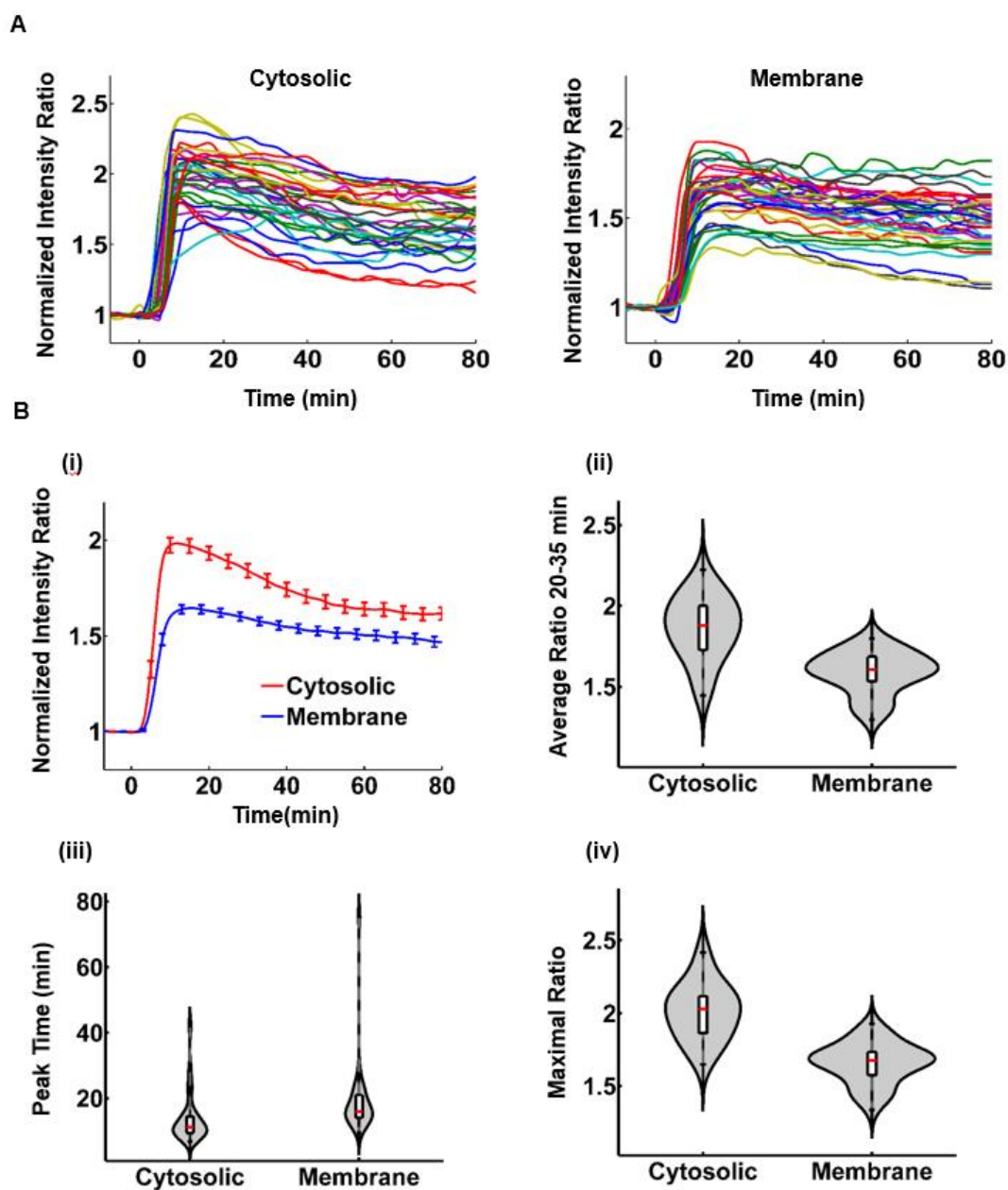
Figure 7. Multiple-cell quantification and statistical interpretation for the cytosolic and membrane-tagged Fyn biosensors in MEF cells. (A) The time courses of normalized emission intensity ratio of the Fyn biosensors in different cells under growth-factor stimulation. **Left:** the cytosolic Fyn biosensor; **Right:** the membrane-tagged Lyn-Fyn biosensor. **(B)** Statistical interpretation of the different results between the cytosolic and membrane groups: **(i)** The average time courses of the emission ratio in each group. Error bars: standard error of mean (SEM). Other results can be compared based on violin plots: **(ii)** The normalized ratio during a selected period of time; **(iii)** The maximal ratio; **(iv)** Time to reach the maximal ratio.

Future development

Currently, *Fluocell* is designed to track the time course of a single cell. We hope in the near future to improve the detection and tracking functions of *Fluocell* so that it is able to identify more than one cell in a sequence of images. For example, a few type of cell experiments have very complicated condition and biologists always try to get as much as possible information in one imaging process or cells grow in groups and we cannot differentiate them in individual ones. In those situations, there are usually more than one target in a single image but current detection will only stick on one of them. Later we will try to modify current single detection method of the software system so that it is capable of fitting a multiple identification purpose in one image. Combining with more applicable data structure for intermediate parameter and result storage, the multiple detection problem should be resolved.

For now, a number of *Fluocell* functions are executed by directly typing commands in MATLAB command window, which are nor convenient to our programming-agnostic users. Sometimes a slight mistake in the commands will produce error messages which are incomprehensible for non-programmers. To solve this problem, we plan to implement most of the *Fluocell* functions into the graphical user interface for suitable usage. For functions with flexible inputs, our users just need to fill the parameters into corresponding boxes, which are then saved and communicated with MATLAB.

With the successful application of sophisticated image analysis methods, the FRET imaging based cell study is expanding to reconstruct more intricate dynamic model of molecular activations and interactions in living cells. By means of appropriate data

visualization, quantification and statistics analysis, the hidden biological cellular activity

network will be revealed gradually.

REFERENCE

1. Born, M. and E. Wolf, *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light*. 7th expanded ed1999, Cambridge ; New York: Cambridge University Press. xxxiii, 952 p.

2. Jares-Erijman, E.A. and T.M. Jovin, *FRET imaging.* Nat Biotechnol, 2003. **21**(11): p. 1387-95.

3. Cheng, P.-C., *The contrast formation in optical microscopy.* Handbook Of Biological Confocal Microscopy.2006: Springer. 162-206.

4. Clegg, R.M., *Fluorescence resonance energy transfer.* Current opinion in biotechnology 6.1, 1995: p. 103-110.

5. Harris, D.C., *Quantitative chemical analysis*. 8th ed2010, New York: W.H. Freeman and Co.

6. Seong, J.H., *Visualization of Src Activity at Different Compartments of the Plasma Membrane by FRET Imaging.* Chemistry & Biology, 2009. **16**(1): p. 48-57.

7. Carlson, H.J. and R.E. Campbell, *Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging.* Curr Opin Biotechnol, 2009. **20**(1): p. 19-27.

8. Piston, D.W. and G.J. Kremers, *Fluorescent protein FRET: the good, the bad and the ugly.* Trends Biochem Sci, 2007. **32**(9): p. 407-14.

9. Carpenter, A.E., *CellProfiler: image analysis software for identifying and quantifying cell phenotypes.* Genome Biol, 2006. **7**(10): p. R100.

10. Berdeaux, R.L., *Active Rho is localized to podosomes induced by oncogenic Src and is required for their assembly and function.* Journal of Cell Biology, 2004. **166**(3): p. 317-323.

11. Thomas, S.M. and J.S. Brugge, *Cellular functions regulated by Src family kinases.* Annual Review of Cell and Developmental Biology, 1997. **13**: p. 513-609.

12. Cohen, A.A., *Dynamic Proteomics of Individual Cancer Cells in Response to a Drug.* Science, 2008. **322**(5907): p. 1511-1516.

13. Sun, J., *Antagonism between binding site affinity and conformational dynamics tunes alternative cis-interactions within Shp2.* Nature Communications, 2013. **4**.

14. Peng, H.C., *Bioimage informatics: a new area of engineering biology.* Bioinformatics, 2008. **24**(17): p. 1827-1836.

15. Marceglia, S., *Education in biomedical informatics: Learning by doing bioimage archiving.* 2007 Annual International Conference of the Ieee Engineering in Medicine and Biology Society, Vols 1-16, 2007: p. 5924-5928.

16. de Alarcon, P.A., A. Gupta, and J.M. Carazo, *A framework for querying a database for structural information on 3D images of macromolecules: A Web-based query-by-content prototype on the BioImage macromolecular server.* Journal of Structural Biology, 1999. **125**(2-3): p. 112-122.

17. Carazo, J.M. and E.H.K. Stelzer, *The BioImage Database Project: Organizing multidimensional biological images in an object-relational database.* Journal of Structural Biology, 1999. **125**(2-3): p. 97-102.

18. Carazo, J.M., *Organising multi-dimensional biological image information: The BioImage Database.* Nucleic Acids Research, 1999. **27**(1): p. 280-283.

19. Diaspro, A., *Studies on the structure of sperm heads of Eledone cirrhosa by means of CLSM linked to bioimage-oriented devices.* Microscopy Research and Technique, 1997. **36**(3): p. 159-164.

20. Schiffmann, D.A., *Open Microscopy Environment and FindSpots: integrating image informatics with quantitative multidimensional image analysis.* Biotechniques, 2006. **41**(2): p. 199-208.

21. Goldberg, I.G., *The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging.* Genome Biol, 2005. **6**(5).

22. Eliceiri, K.W. and C. Rueden, *Tools for visualizing multidimensional images from living specimens.* Photochemistry and Photobiology, 2005. **81**(5): p. 1116-1122.

23. Collins, T.J., *ImageJ for microscopy.* Biotechniques, 2007. **43**(1): p. 25-+.

24. Lamprecht, M.R., D.M. Sabatini, and A.E. Carpenter, *CellProfiler: free, versatile software for automated biological image analysis.* Biotechniques, 2007. **42**(1): p. 71-5.

25. Lu, S.Y., *The Spatiotemporal Pattern of Src Activation at Lipid Rafts Revealed by Diffusion-Corrected FRET Imaging.* Plos Computational Biology, 2008. **4**(7).

26. Lu, S.Y., *Computational Analysis of the Spatiotemporal Coordination of Polarized PI3K and Rac1 Activities in Micro-Patterned Live Cells.* Plos One, 2011. **6**(6).

27. Otsu, N., *Threshold Selection Method from Gray-Level Histograms.* Ieee Transactions on Systems Man and Cybernetics, 1979. **9**(1): p. 62-66.

28. Stith, B.J., *Use of animation in teaching cell biology.* Cell Biol Educ, 2004. **3**(3): p. 181-8.

29. Tversky, B., J.B. Morrison, and M. Betrancourt, *Animation: can it facilitate?* International Journal of Human-Computer Studies, 2002. **57**(4): p. 247-262.

30. Williamson, V.M. and M.R. Abraham, *The Effects of Computer Animation on the Particulate Mental Models of College Chemistry Students.* Journal of Research in Science Teaching, 1995. **32**(5): p. 521-534.

31. Resh, M.D., *Fyn, a Src family tyrosine kinase.* International Journal of Biochemistry & Cell Biology, 1998. **30**(11): p. 1159-1162.

32. Noble, M.E.M., *Crystal-Structure of the Sh3 Domain in Human Fyn - Comparison of the 3-Dimensional Structures of Sh3 Domains in Tyrosine Kinases and Spectrin.* Embo Journal, 1993. **12**(7): p. 2617-2624.

33. Boggon, T.J. and M.J. Eck, *Structure and regulation of Src family kinases.* Oncogene, 2004. **23**(48): p. 7918-7927.

34. Xia, D. and J. Gootz, *Premature lethality, hyperactivity, and aberrant phosphorylation in transgenic mice expressing a constitutively active form of Fyn.* Frontiers in Molecular Neuroscience, 2014. **7**.

35. Roskoski, R., *Src protein-tyrosine kinase structure and regulation.* Biochem Biophys Res Commun, 2004. **324**(4): p. 1155-1164.

36. Umemori, H., *Initial Events of Myelination Involve Fyn Tyrosine Kinase Signaling.* Nature, 1994. **367**(6463): p. 572-576.

37. Thomas, S.M., P. Soriano, and A. Imamoto, *Specific and Redundant Roles of Src and Fyn in Organizing the Cytoskeleton.* Nature, 1995. **376**(6537): p. 267-271.

38. Cary, L.A., J.F. Chang, and J.L. Guan, *Stimulation of cell migration by overexpression of focal adhesion kinase and its association with Src and Fyn.* J Cell Sci, 1996. **109**: p. 1787-1794.

39. Wary, K.K., *A requirement for caveolin-1 and associated kinase Fyn in integrin signaling and anchorage-dependent cell growth.* Cell, 1998. **94**(5): p. 625-634.

40. Hanke, J.H., *Discovery of a novel, potent, and Src family-selective tyrosine kinase inhibitor - Study of Lck- and FynT-dependent T cell activation.* Journal of Biological Chemistry, 1996. **271**(2): p. 695-701.

41. Lang, V., *Fyn membrane localization is necessary to induce the constitutive tyrosine phosphorylation of Sam68 in the nucleus of T lymphocytes.* Journal of Immunology, 1999. **162**(12): p. 7224-7232.

42. Stuermer, C.A.O., *Glycosylphosphatidyl inositol-anchored proteins and fyn kinase assemble in noncaveolar plasma membrane microdomains defined by reggie-1 and-2.* Molecular Biology of the Cell, 2001. **12**(10): p. 3031-3045.

43. Liang, X.Q., *The N-terminal SH4 region of the Src family kinase Fyn is modified by methylation and heterogeneous fatty acylation - Role in membrane targeting, cell adhesion, and spreading.* Journal of Biological Chemistry, 2004. **279**(9): p. 8133-8139.

44. Stenberg, P.E., *The Src family kinases, Fgr, Fyn, Lck, and Lyn, colocalize with coated membranes in platelets.* Blood, 1997. **89**(7): p. 2384-2393.

45. Palacios, E.H. and A. Weiss, *Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation.* Oncogene, 2004. **23**(48): p. 7990-8000.

46. Salmond, R.J., *T-cell receptor proximal signaling via the Src-family kinases, Lck and Fyn, influences T-cell activation, differentiation, and tolerance.* Immunological Reviews, 2009. **228**: p. 9-22.