

## RESEARCH ARTICLE

# BetaBuddy: An automated end-to-end computer vision pipeline for analysis of calcium fluorescence dynamics in $\beta$ -cells

Anne M. Alsup<sup>1</sup>, Kelli Fowlds<sup>1</sup>, Michael Cho<sup>1\*</sup>, Jacob M. Luber<sup>1,2,3\*</sup>

**1** Department of Bioengineering, University of Texas at Arlington, Arlington, TX, United States of America, **2** Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX, United States of America, **3** Multi-Interprofessional Center for Health Informatics, University of Texas at Arlington, Arlington, TX, United States of America

 These authors contributed equally to this work.

\* [michael.cho@uta.edu](mailto:michael.cho@uta.edu) (MC); [jacob.luber@uta.edu](mailto:jacob.luber@uta.edu) (JML)



## OPEN ACCESS

**Citation:** Alsup AM, Fowlds K, Cho M, Luber JM (2024) BetaBuddy: An automated end-to-end computer vision pipeline for analysis of calcium fluorescence dynamics in  $\beta$ -cells. PLoS ONE 19(3): e0299549. <https://doi.org/10.1371/journal.pone.0299549>

**Editor:** Salman Sadullah Usmani, Albert Einstein College of Medicine, UNITED STATES

**Received:** April 10, 2023

**Accepted:** February 12, 2024

**Published:** March 15, 2024

**Copyright:** © 2024 Alsup et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data is available at Zenodo:<https://doi.org/10.5281/zenodo.10476949> Code is available at: <https://github.com/jacobluber/BetaBuddy>.

**Funding:** This work was supported by Cancer Prevention and Research Institute of Texas grant RR220015 (Jacob Luber), National Heart, Lung, and Blood Institute, T32 Grant HL134613, The University of Texas at Arlington, Alfred R. and Janet H. Potvin Endowment (Michael Cho), and The

## Abstract

Insulin secretion from pancreatic  $\beta$ -cells is integral in maintaining the delicate equilibrium of blood glucose levels. Calcium is known to be a key regulator and triggers the release of insulin. This sub-cellular process can be monitored and tracked through live-cell imaging and subsequent cell segmentation, registration, tracking, and analysis of the calcium level in each cell. Current methods of analysis typically require the manual outlining of  $\beta$ -cells, involve multiple software packages, and necessitate multiple researchers—all of which tend to introduce biases. Utilizing deep learning algorithms, we have therefore created a pipeline to automatically segment and track thousands of cells, which greatly reduces the time required to gather and analyze a large number of sub-cellular images and improve accuracy. Tracking cells over a time-series image stack also allows researchers to isolate specific calcium spiking patterns and spatially identify those of interest, creating an efficient and user-friendly analysis tool. Using our automated pipeline, a previous dataset used to evaluate changes in calcium spiking activity in  $\beta$ -cells post-electric field stimulation was reanalyzed. Changes in spiking activity were found to be underestimated previously with manual segmentation. Moreover, the machine learning pipeline provides a powerful and rapid computational approach to examine, for example, how calcium signaling is regulated by intracellular interactions.

## Introduction

Insulin-secreting  $\beta$ -cells comprise the primary portion of groups of endocrine cells residing in the pancreas known as the Islets of Langerhans, making up approximately 50–70% [1–4]. Their role in the process of glycemic regulation has made them an ongoing source of study for Diabetes Mellitus. Type I Diabetes, specifically, is an autoimmune disorder characterized by the decreased secretion of insulin due to the destruction of  $\beta$ -cells. Complications from diabetes include micro- and macrovascular disorders such as thrombosis, atherosclerosis, and

University of Texas System Rising STARS Award (Jacob Lubner).

**Competing interests:** The authors have declared that no competing interests exist.

declines in cognitive function [5]. Two ongoing challenges to the treatment and/or cure of Type I Diabetes involve the reduced population of  $\beta$ -cells and the autoimmune attack within the body [6]. The main function of  $\beta$ -cells is the production and secretion of insulin. The release of this hormone is triggered by an increase in ATP [7] as the blood becomes oversaturated with glucose. As intense research efforts continue to explore this disease, many of the mechanisms behind insulin secretion are yet to be fully understood.

Insulin secretion is strongly correlated and regulated by calcium signaling. Glucose plays a large role in the stimulus of  $\text{Ca}^{2+}$  movement in the mitochondria, endoplasmic reticulum, nucleus, cytosol, and other areas of the cell and increases the level of cytosolic  $\text{Ca}^{2+}$  [8]. Additionally, an influx of  $\text{Ca}^{2+}$  through voltage-gated channels impacts insulin exocytosis and secretion. Insulin secretion patterns are typically studied by following the movement of  $\text{Ca}^{2+}$  oscillations (or spikes) in  $\beta$ -cells [9]. For example, membrane-permeable and high affinity calcium-sensitive fluorophores can be loaded into the cells. Upon binding free  $\text{Ca}^{2+}$  ion, fluorescent signals are generated and emitted when excited by an appropriate external light source. This process has been routinely captured in real-time subcellular imaging and allowed the investigator to track intracellular calcium levels. While the  $\text{Ca}^{2+}$  dynamics in individual cells has been monitored and tracked [10], challenges remain to examine the  $\text{Ca}^{2+}$  responses in each cell, especially in clusters of cells that are more physiologically representative of islets. Calculating control calcium activity provides a threshold that can be used to quantify significant changes in oscillations due to different forms of stimulation [11]. There is still much unknown regarding the effects of different stimuli on cell spiking activity, changes due to cell-cell interactions, and how insulin secretion fluctuates based on temporal or spatial variables. Moreover, measuring fluorescent intensities of individual cells that are clustered and/or overlap remains a daunting problem during image analysis. Efficient, accurate, and automated computational pipelines to determine potential cell-to-cell communication through  $\text{Ca}^{2+}$  oscillations are expected to enhance the current understanding of endocrine cell behaviors.

Automated image analysis tools are continually being developed and made readily available to investigators. Technological advancements have both decreased the time required to perform wet lab experiments and enabled researchers to obtain more sophisticated data. Open-source software such as ImageJ [12] and its updated architecture, Fiji [13], have been widely utilized for the evaluation of biological images. Fiji's plugin, Trackmate [14], can further be used for the tracking of individual cells. Multiple other tracking procedures have been written in Python for specific research needs [15–17]. These tools are among many that can be combined with customized algorithms for enhanced biological research.

Segmentation is an important tool that allows researchers to computationally identify regions of interest (ROIs) and analyze them using different methodologies [12]. Manual segmentation has in the past been a standard protocol. However, it is time intensive and often subject to human errors and biases (S1 Fig) [18,19]. Typically, analysis of fluorescent images begins by outlining each specific cell to measure fluorescent intensities that might fluctuate with time. With the need for batch analysis of imaging data becoming more prevalent, an automated analysis tool is required to efficiently identify cells through a time-series image stack and eventually predict and validate physiological responses (e.g., insulin secretion). Hand segmentation remains a strong measure of evaluation for overall training accuracy, but automated methods are quickly being established as the new primary method of imaging analysis [20].

In recent years, many tools have been developed to continue improving automated segmentation of cellular microscopy images including watershed algorithms, pre-processing techniques, deep learning methods such as Convolutional Neural Network architectures, and combinations of multiple techniques [21–27]. Before the wide use of deep learning, novel segmentation techniques involved differentiating cells from the background through

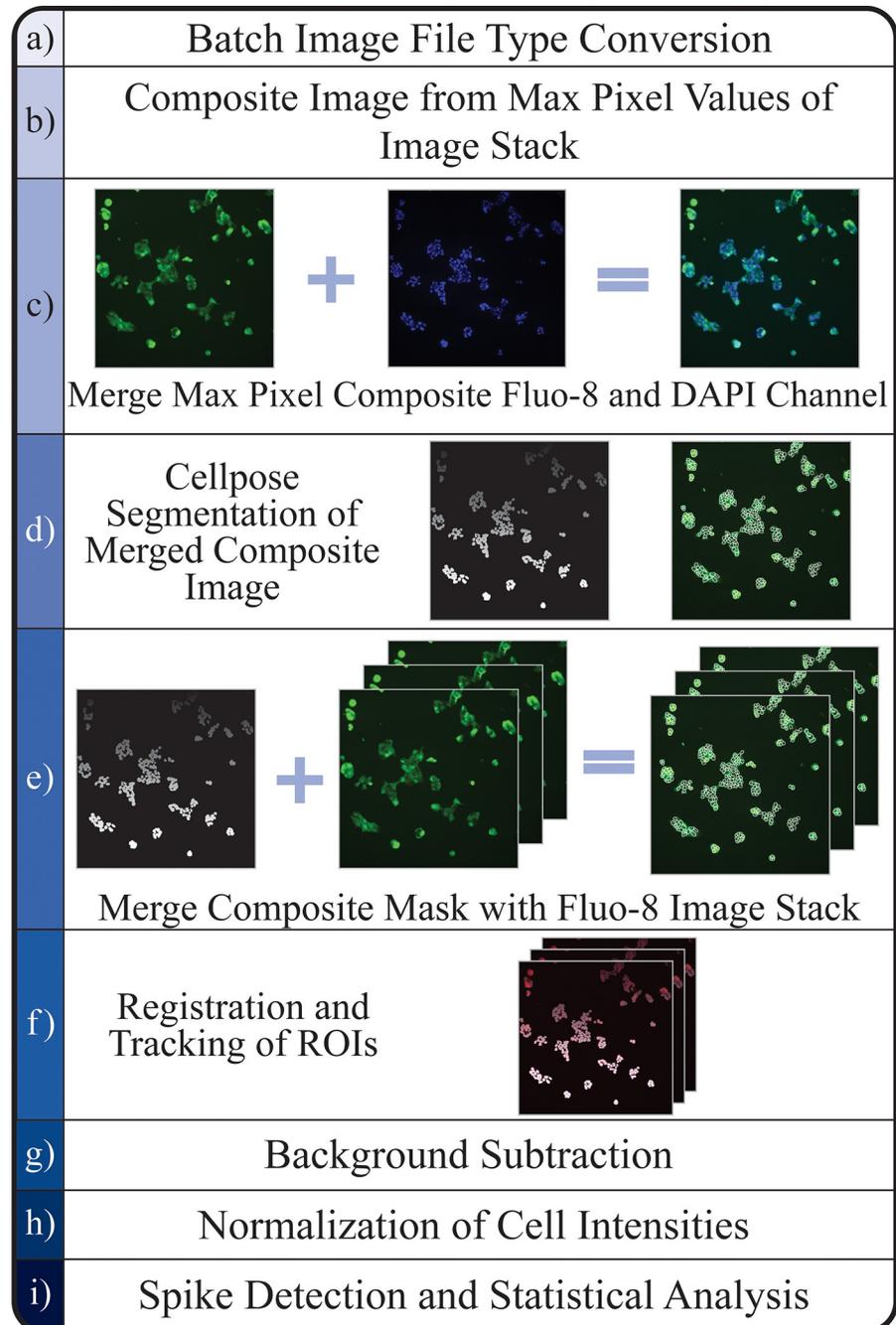
thresholding, masking general cell outlines through complex mathematical operations, redefining cell shapes with the watershed algorithm, incorporation of active contours to identify rounded cell shapes, or a combination of various methods to create a more accurate segmentation of clustered cells [28–31]. While these previous techniques are still applicable and useful in conjunction with newer methods, the incorporation of deep neural networks has been able to overcome common issues such as heterogeneous distribution of the fluorescent markers for watershed and uniform shapes making masking difficult. Deep learning-based methods are able to utilize large hand annotated data sets to accurately segment cells without relying on predefined shapes, distinct cell boundaries, or an evenly distributed cell culture with minimal clustering [32–36]. 3D segmentation of sub-organelles within islets has allowed for more definitive quantification of intracellular interactions [37,38]. More novel deep learning networks, such as Cellpose, aim to create generalist models that increase flexibility of image analysis by training with a variety of cell types.

Cellpose [32] is a generalist, deep-learning algorithm designed for accurate segmentation of a wide variation of cell types. Its model is optimal for different types of biological research including 2D and 3D image sets, phase-contrast, and fluorescence microscopy as it provides segmentation without the need for retraining [32]. A variety of pre-trained models, named the ‘model zoo’, allows researchers to find the best segmentation styles for their image based on the large pre-segmented data sets (Cellpose, TissueNet, and LiveCell). The algorithm is user-friendly and customizable for specific experimental needs, and its newest updates offer human-in-the-loop pipelines and methods for retraining [39]. Since its inception, several studies have already utilized its model. Saad et al. [40] developed a procedure using Cellpose and Fiji to characterize ice crystals, and Hoeren et al. [41] adapted Cellpose into their pipeline analyzing cells adhering to PMP-ECMO membranes as an alternative segmentation process to Fiji. Reinbigler et al. [42] modified the model’s base framework and used it in conjunction with QuPath [43] to create a pipeline for the segmentation of Hematoxylin-Eosin (HE) stained histopathological whole slide images of myofibers. Waisman et al. [44] similarly used Cellpose for myofiber segmentation, outputting the images to ImageJ for ROI analysis. Fisch et al. [45] incorporated StarDist [33] for nuclear segmentation and Cellpose for nuclear segmentation in their automated workflow for the analysis of host-pathogen interactions. While Cellpose is a powerful tool, improvements are needed to obtain an end-to-end pipeline as segmentation is simply the first step of a long analysis process. BetaBuddy allows for complete analysis of image sets by automatically linking ROIs across frames, tracking fluorescence intensity, extracting cellular data (size, shape, intensity, etc.), and conducting initial statistical tests.

Our aim in this study was to develop and apply computational methods to obtain a faster, more accurate and in-depth analysis of calcium signaling in  $\beta$ -cells. We have developed a pipeline intended to decrease the time required for batch analysis of large image datasets and created a series of algorithms to automatically detect and calculate significant changes in fluorescent intensities.

## Materials and methods

Our primary goal is to create an automated system that outlines cell boundaries (segmentation) and track  $\beta$ -cells over a sequence of different time points (Fig 1). The system includes and performs statistical analyses on the calcium intensity patterns from individual cells. The completed pipeline can be used through a Jupyter Notebook and all package dependencies are described in the GitHub repository. Referred to as the BetaBuddy GitHub, it also contains installation and usage instructions for inexperienced programmers to easily follow and create a more accessible tool for all researchers to implement in their experimental procedures. The



**Fig 1. Pipeline flow chart.** (a) Original  $\text{Ca}^{2+}$  microscopy images of  $\beta$ -cells (20x objective) converted to TIFF file from the Nikon-specific format. (b) A composite image is made by identifying the maximum value of each pixel location from the image stack. The resulting image is the brightest pixel out of all images in the stack, ensuring each cell is fully fluorescing in one image. This will allow for a more accurate segmentation without dropout due to non-fluorescing cells. (c) Composite Fluo-8 stained  $\beta$ -cell image and DAPI stain of nuclei of  $\beta$ -cells showing only the nuclei of each cell merged with both images in a separate channel, assisting in a more accurate segmentation. Detailed description of Cellpose is outlined in the next section. (d) ROI mask after composite segmentation and overlaid predicted outlines from after Cellpose segmentation. (e) Mask of segmented composite image is merged with all images in the image stack. Due to this step only image stacks with exceptionally minimal movement should be used. (f) Visualization of tracks created after registration and tracking. This figure includes the original fluorescent image, mask image, ROI outlines, and predicted tracks cells followed throughout the imaging process. (g-i) Statistical analysis is incorporated to generate automated results, which can be compared with manually analyzed results (see Fig 6 below).

<https://doi.org/10.1371/journal.pone.0299549.g001>

automated pipeline is expected to improve the accuracy of data and decrease the overall time necessary for future studies. The entire procedure was run using an NVIDIA DGX A100 High Memory GPU Supercomputer, containing 8 NVIDIA A100 high memory graphics cards with 80GB of GPU RAM each for a total of 640GB of GPU RAM for training large Deep Bayesian model architectures. The system also has 2 TB of HBM2 RAM and 128 AMD EPYC CPU Cores, and it connects to the internet via 10 GB/S connection. The process from image conversion, segmentation, tracking, and finally statistical analysis was successfully performed with 1 GPU, 12 CPUs, and 32GB of memory allocated.

## Calcium imaging and deep learning-based segmentation

The images used in our automated pipeline were sourced from previous experimental images that monitored and recorded  $\text{Ca}^{2+}$  dynamics in real-time [46]. Mouse derived  $\beta\text{TC-6}$  insulinoma cells were cultured and allowed to grow for several days. A calcium indicator dye (Fluo-8) was used to measure changes in intracellular calcium levels. The cells were stained with 0.8  $\mu\text{M}$  Fluo-8, and 1 drop/mL Nucblue to visualize the nuclei, for 30 min at 37°C and loaded onto a custom-designed electric field exposure chamber [46]. Non-invasive electric field stimulation (EFS; 15 min exposure) was chosen to manipulate the voltage-gated calcium channels and therefore modulate intracellular  $\text{Ca}^{2+}$  dynamics. A Nikon microscope was used to image the cells at 5 s intervals over a period of 2 min before and after EFS. This method provided a baseline and changes in the intracellular  $\text{Ca}^{2+}$  levels were measured. Data for such experiments consisted of stacked 2048x2048 pixel image sets of 25 frames before and after EFS. In total, 56 image stacks were analyzed. The Fluo-8 fluorescent image data sets consisted of pre- and post-exposure images of non-invasive 1 to 3 V/cm EFS. The analysis pipeline was used to analyze approximately 9–10 image sets for each experimental condition. Pre-exposure cells served as their own control.

Once the cells are imaged, multiple single or stacked images can be imported directly to the script and begin segmentation after format conversion. The Java library, Bio-Formats, is first used to convert the Nikon-specific file types to TIFF for ease of use in Linux ImageJ macro functions. The stack of images is then made into a single composite image with every pixel representing the maximum value of its pixel location. The final image will have the largest pixel value from all images in the stack. This was done due to the oscillating nature of  $\text{Ca}^{2+}$  signaling causing cells to be completely dark in certain frames. By creating an image with only the brightest pixels, we are able to ensure all cells able to uptake the fluorescence  $\text{Ca}^{2+}$  tag will be segmented and negate the issue of ROI drop-out caused by cells not fluorescing in all frames. Due to the clustering and static, non-migratory nature of  $\beta$ -cells, one mask image was determined sufficient in assisting segmentation for all frames. The composite image is then merged with an image of only the nuclei (labeled with DAPI) from the first frame of the exposure period (Fig 1C). It was determined that adding a DAPI channel greatly improved the segmentation process.

A validated deep learning-based generalist cellular segmentation algorithm, Cellpose, is used to automatically outline the boundary of each cell and create an ROI with user-defined parameters [32]. Users may specify if their images contain a DAPI channel, which channel contains the entire cell, the Cellpose model for segmentation, and thresholds. These parameters can be tested on one image to ensure quality segmentation. The generalist algorithm has a better performance with the addition of a DAPI channel due to the clustering nature of  $\beta$ -cells and difficulty differentiating neighboring cells with only one calcium indicator dye. After segmentation, an image defining the ROI of every cell is saved. The 'mask' image differentiates cells by assigning every ROI a different grayscale color (Fig 1D). The unique color can be used as a label during cell registration and tracking.

The mask image is replicated and merged with each frame of the original image stack without the DAPI channel in order to analyze the correct fluorescent intensity. Merging is performed by sorting the newly made mask image stack and original image stack into two different color channels (Fig 1E). This format will allow for the Fiji plugin, TrackMate, to begin cell registration by identifying the ROIs labeled in the channel containing the mask images [47]. The cell track detector analyzes the mask channel and registers every uniquely colored shape as a cell. Due to limitations of hand segmentation, including time of analysis and difficulty identifying cell boundaries across frames, 56–185 cells were identified in each trial. Manual segmentation is an established method of analysis in the field, and it is common practice to label a representative sample of cells that can be tracked over time for consistency (S1 Fig). As the automated pipeline is able to segment and register nearly every cell within the experimental population, 70–680 cells were labeled in each trial using BetaBuddy.

Next, each cell is linked across multiple frames, creating a “track” (Fig 1F). Lastly, fluorescent intensity for each frame in a cell’s track is obtained by referencing the first channel, which holds the original fluorescent image. The intensity values are taken from within the cell’s complementary ROI. This process is completely automated using ImageJ’s macro and python scripting features. This step will create an XML file that can be loaded into TrackMate and create a visualization of the tracks produced. Lastly, a CSV file is saved containing fluorescent intensity values, area, and spatial coordinate positions by grouping each cell into a “Track ID” over time.

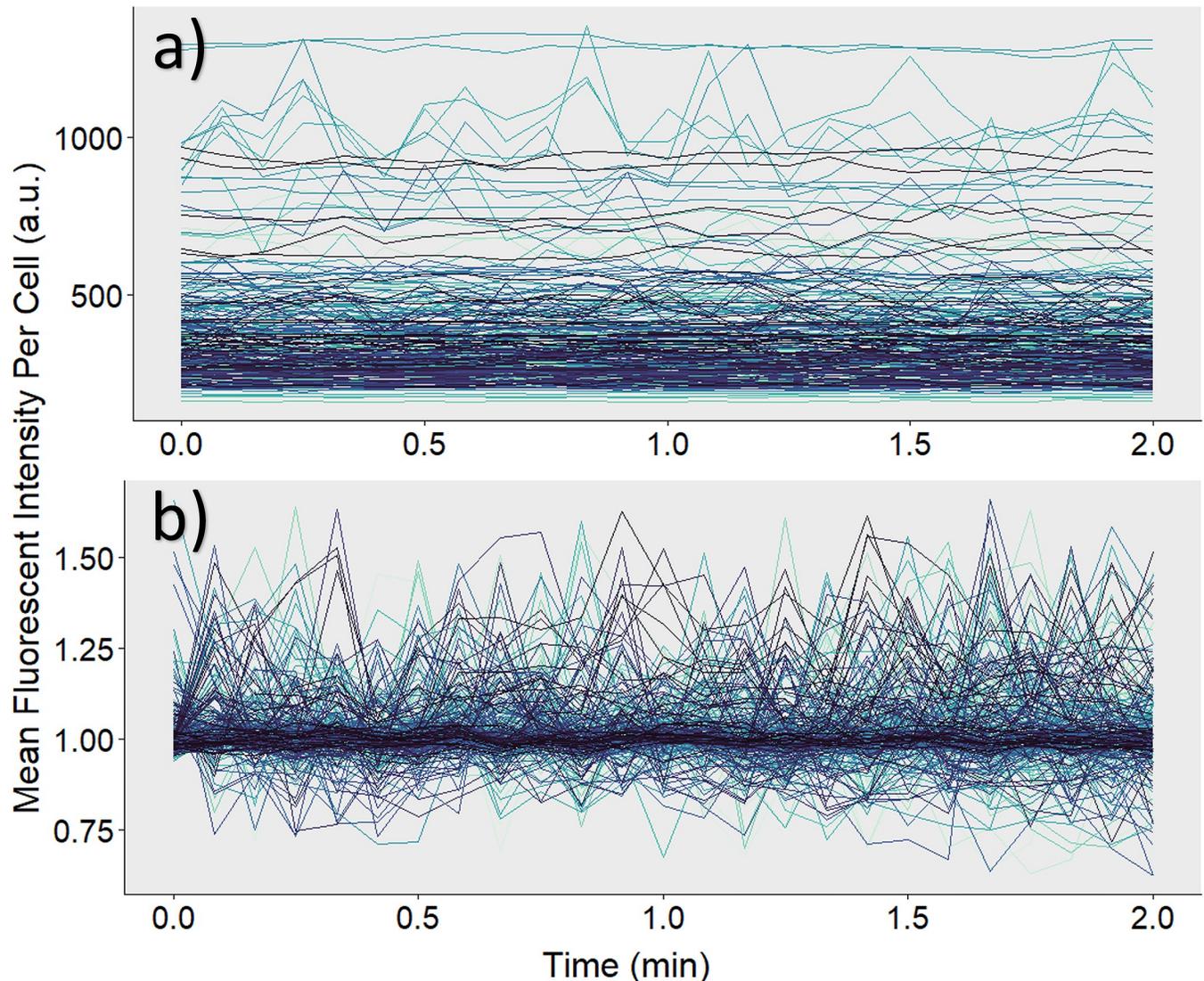
### Background subtraction

Initially, the raw data CSV created after registration and tracking are very saturated with external background noise (Fig 2A). Fluorescent images typically have background noise associated with excessive dye or thermal electronic fluctuation. Therefore, raw data should undergo a background subtraction process. To obtain a well-distributed sample of background intensities, 100 pixel locations outside of the ROIs are randomly selected from the original image stack. The randomly selected pixel is compared to the mask image stack to ensure the pixel intensity has a value of 0, verifying it is not within an ROI.

Evaluation of all acquired data is analyzed through a procedure written in R and called upon through a terminal command in the Jupyter Notebook. The working directory is set to the user’s current operating space, and the CSV files containing the raw tracking information and background points are automatically imported. First, the average of all 100 randomly sampled background pixels is calculated for each unique frame. This mean value is then merged into the cell tracking information to create a singular data set. The average background intensity for each time point is then subtracted from the mean intensity of each cell within that frame. Next, to establish individual baseline activity, mean intensities were normalized specific to each individual cell.

### Normalization

Calcium intensities were normalized to account for different dye loading conditions that can be attributed to variations in cell size, cluster size, and potential differences in how each cell takes up the dye. This process allowed us to examine changes in fluorescent intensity in a way that could be fairly compared across all data points (Fig 2B). Normalization was performed by obtaining a “baseline” individual to each cell. In keeping with the original experiment’s methodology, a change in fluorescent intensity  $\geq 10\%$  was notated as spiking activity [46]. We kept this 10% threshold for the purpose of comparing the accuracy of either manual or automated data analysis of calcium dynamics in  $\beta$ -cells. To find a baseline for each cell, the number of data points until each specific Track ID reached a change of  $\geq 10\%$  from its first recorded



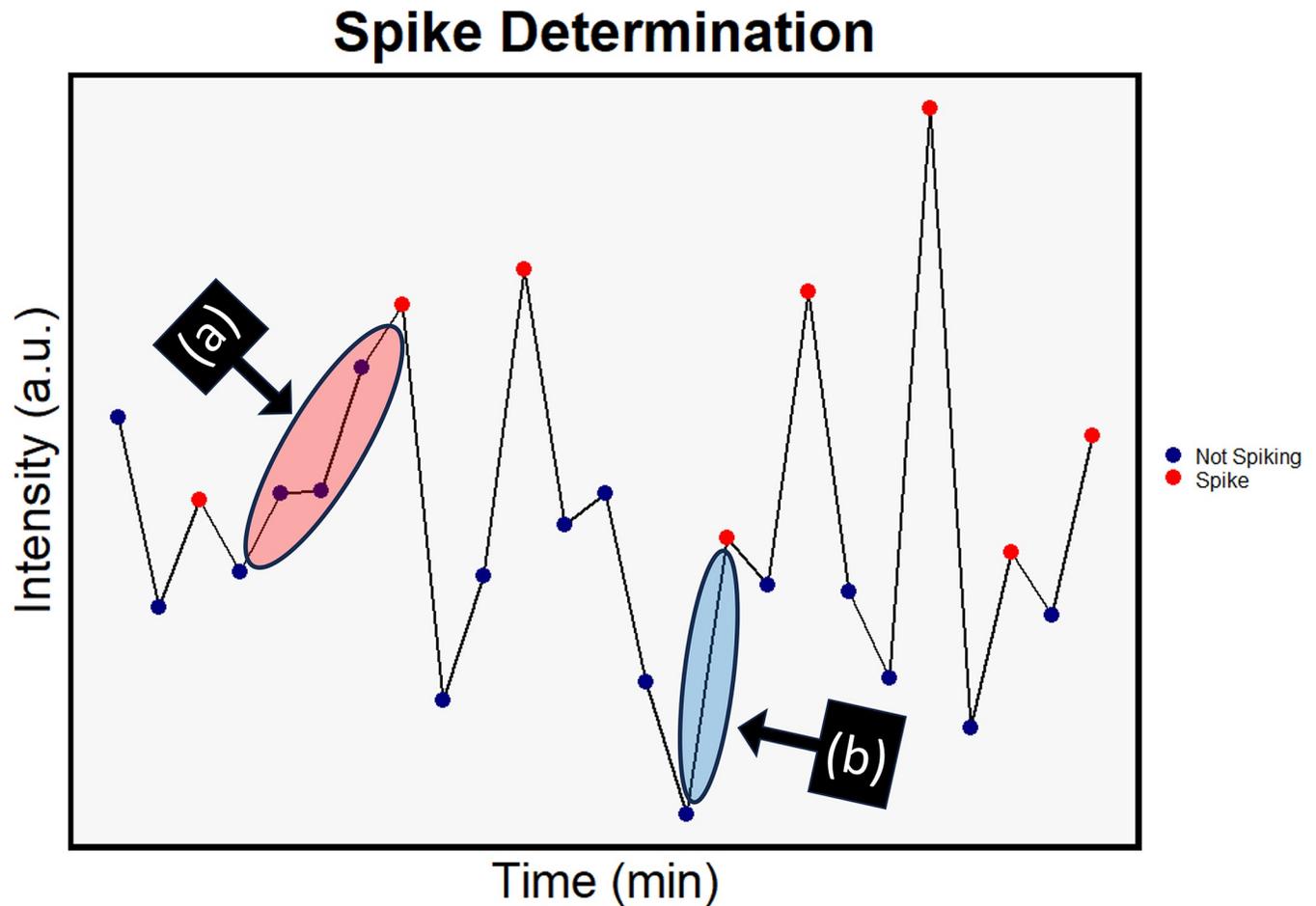
**Fig 2. Intensity of experimental data.** Each line represents an individual cell and its fluorescent intensity tracked over a period of 2 minutes. (a) Initial data is saturated with background noise. (b) Raw intensity values undergo background subtraction and normalization for a more accurate depiction of calcium activity.

<https://doi.org/10.1371/journal.pone.0299549.g002>

intensity was calculated, and the mean of these first  $n$  points was divided from each background-subtracted intensity within the group. If a cell never reached an intensity level greater than 10% of its first value, then its baseline was determined using the mean intensity of all recorded data points for that cell.

### Spike determination

The purpose of this experiment was to repeat the analysis of spiking activity within the original data set that utilized hand segmentation methods. Our goal was to determine if there was any significant loss of data in the original methodology. For the automated analysis, we defined a calcium spike using the 10% threshold, as previously stated. The percent change in normalized intensity from its previous value was calculated for each  $\beta$ TC-6 cell. Based on the value changes, a spike was defined in two ways: either (a) the sum of the cumulative percent changes



**Fig 3. Calcium spike determination.** Computational methods were used to automatically designate calcium spikes at a 10% threshold. This threshold was met through either (a) the sum of percent changes between two designated peaks or (b) a singular percent change of  $\geq 10\%$  immediately preceding a defined peak. The line graph represents fluctuations in fluorescent intensity over time for a single cell in a trial.

<https://doi.org/10.1371/journal.pone.0299549.g003>

between two peaks was  $\geq 10\%$  or (b) the percent change from the data point immediately preceding any peak in intensity was  $\geq 10\%$  (Fig 3).

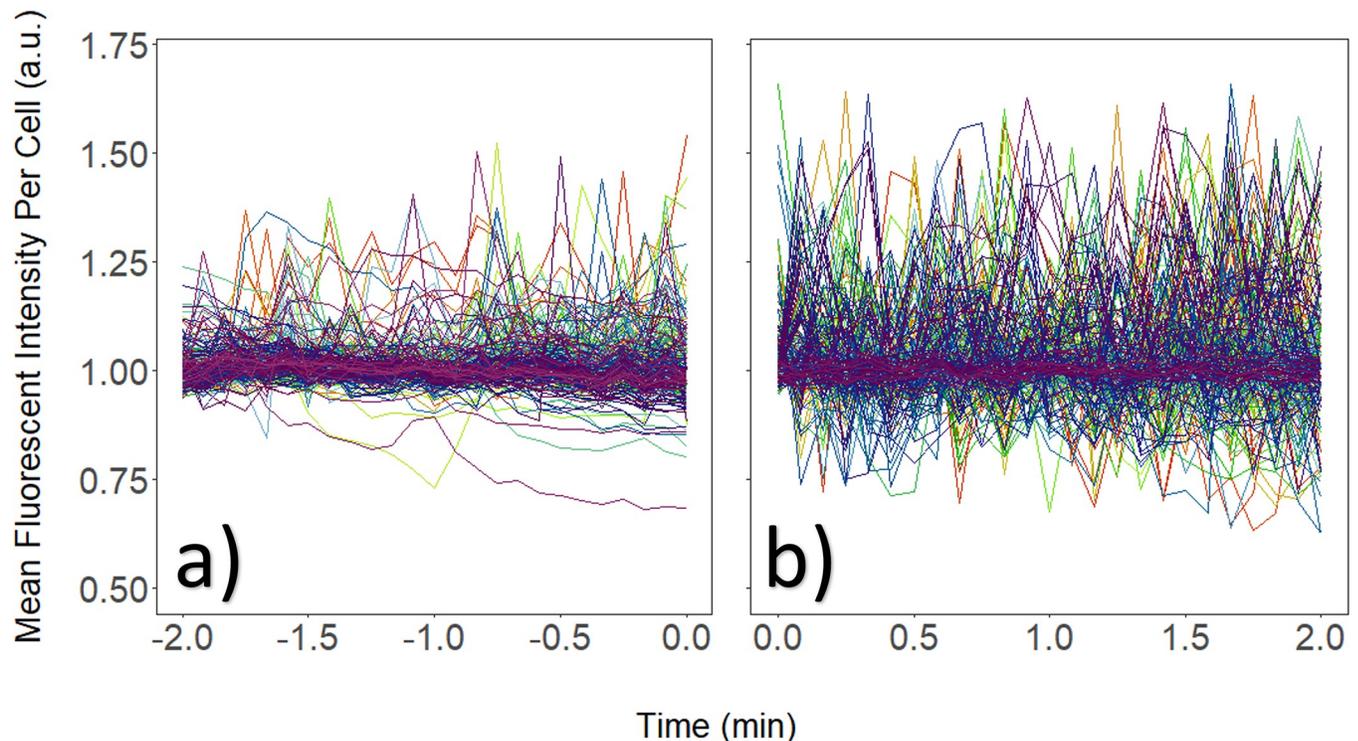
Using these parameters, we quantified each instance individual cells spiked over the course of the imaging period before and after external stimulation. Mean activity for each trial was calculated from the aggregate of all spikes within the population.

## Results

### Statistical analysis

Prior to statistical tests, outliers for singular trials were detected using a standard Z score. Track IDs with any intensity values greater than a Z score of 3.0 or less than -3.0 were automatically removed. Mean spiking activity was calculated as the total number of spikes recorded for each trial divided by total number of cells over time.

Fig 4 illustrates one experiment using 1 V/cm stimulation. Normalized calcium activity was plotted from -2 to 0 min (i.e., control before EFS was applied) and from 0 to 2 min following 15 min EFS. Calcium recording during 15 min EFS was not performed to minimize fluorophore degradation and/or bleaching.



**Fig 4. Signaling changes post EFS stimulation.** Variations in activity, including changes in the fluorescent intensity and spiking frequency, increased significantly following a 15 min exposure to 1 V/cm EFS. (a) Fluorescent activity prior to stimulation. (b) Fluorescent activity post EFS.

<https://doi.org/10.1371/journal.pone.0299549.g004>

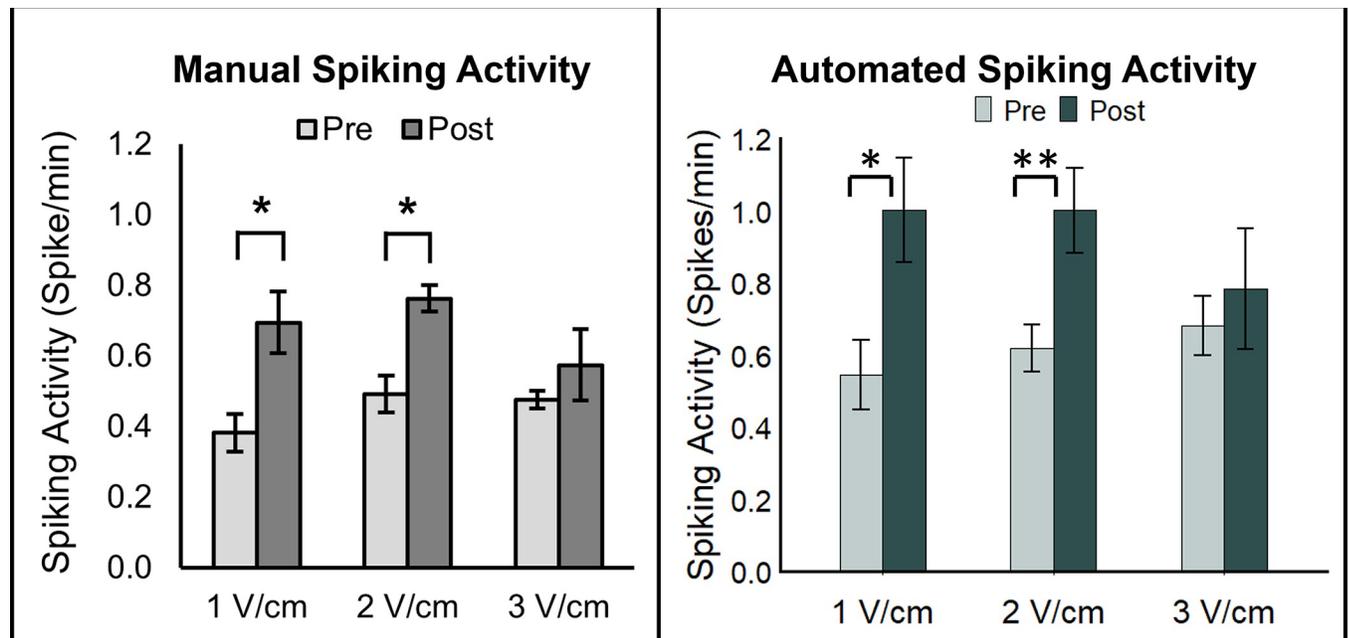
A two-tailed paired *t* test was performed to evaluate significant changes in the calcium signaling of the  $\beta$ -cells in response to EFS of 1 to 3 V/cm. The numbers of spikes calculated through the automated pipeline were plotted and compared with published results (Fig 5).

Results obtained through automated methods validated the findings previously recorded using manual analysis. EFS of 1 and 2 V/cm significantly increased the calcium spiking activity compared to their respective controls. There was no significant increase in calcium activity found at 3 V/cm. In further comparison of the manual vs. automated analysis, it was noted that the number of calcium spikes calculated using hand segmentation were underestimated by ~40% in all control and experimental conditions with respect to their counterparts in the automated results. This trend is indicative of systematic errors that may have been caused by human biases. For example, one factor that could account for the discrepancy in these calculations could be attributed to a limited number of ROIs in the manual segmentation (Fig 6).

ROI identification is often constrained with manual segmentation practices. Images may only have active cells and/or cells with clear boundaries registered due to visual impairment and time constraints. Which ROIs are identified can also be a source of bias. Thus, calculations of spiking activity may only be sourced from a limited number of cells. In contrast, automated segmentation using BetaBuddy resulted in near complete registration of all cells present within each frame.

## Discussion

The application of electric field stimulation to  $\beta$ TC-6 cells was found to successfully increase calcium spiking activity. This indicates that EFS within 1 to 2 V/cm serves as a potentially beneficial treatment for pre-conditioning  $\beta$ -cells for enhanced insulin secretion. It is interesting to



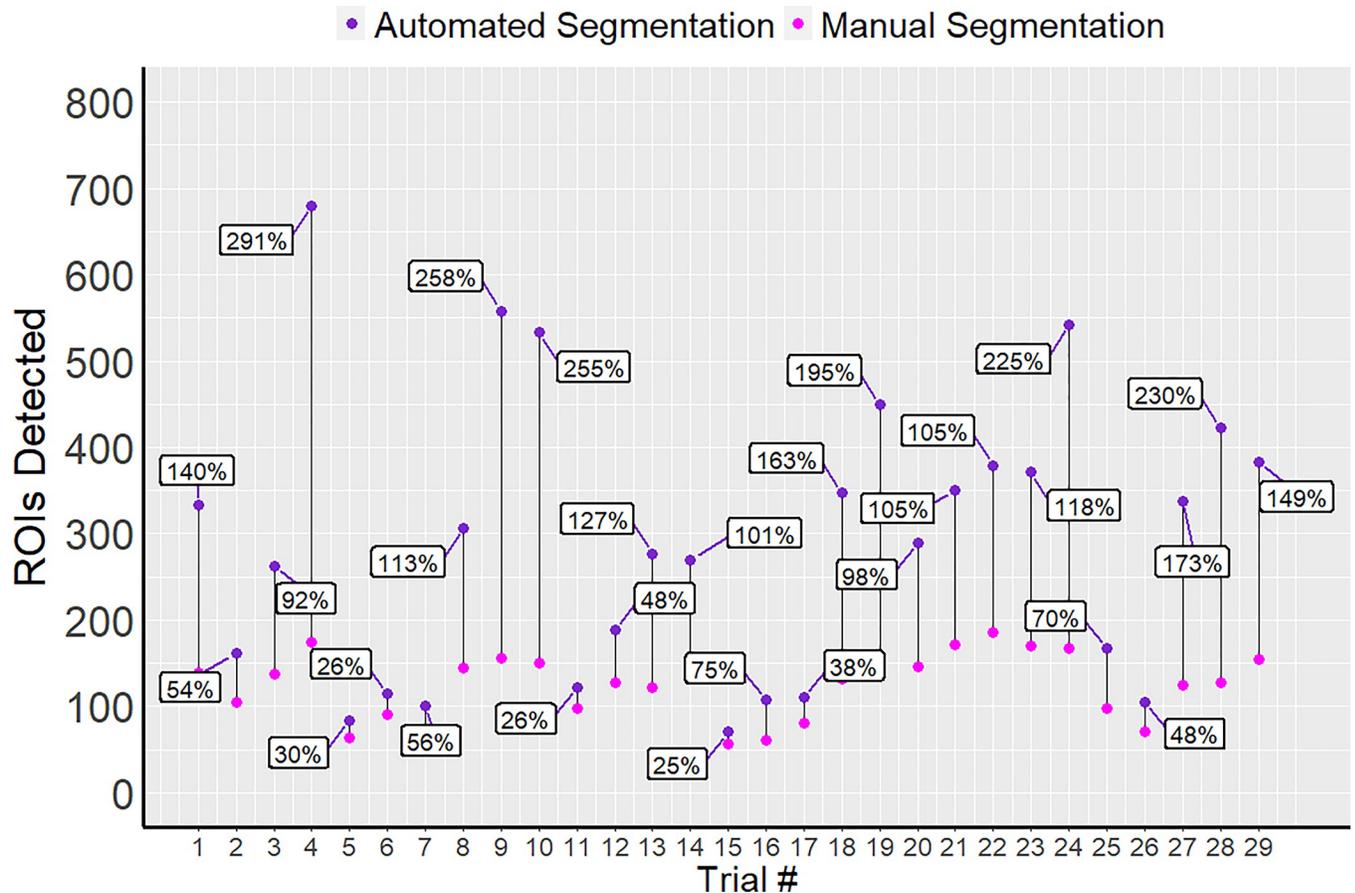
**Fig 5. Comparison of results.** The number of calcium spikes determined by hand are shown in the left panel. Using the same images, the results from the automated analysis are shown in the right panel. \* denotes  $p < 0.05$  and \*\*  $p < 0.01$  using a two-tailed paired t-test.

<https://doi.org/10.1371/journal.pone.0299549.g005>

note that a larger EFS (3 V/cm) did not induce statistically significant changes in the calcium spiking determined either by hand or by automated pipeline. We speculate that, since the induced electrical potential depends on the cluster size, larger EFS causes non-negligible and perhaps irreversible changes in the cell membrane potential. Generally, automated analysis found greater overall spiking activity in the  $\beta$ TC-6 cells. While it remains to be further elucidated whether one-time treatment with EFS is sufficient to pre-condition pancreatic cells, it does offer a non-biologic approach to regulate insulin secretion and/or trafficking.

Furthermore, results that were previously reported using manual segmentation (see Fig 5) were compared with those determined by BetaBuddy. Both sets of results were comparable, suggesting analysis by hand was able to quantify the calcium dynamics in  $\beta$ -cells but underestimated the calcium spiking rate. Automated segmentation thus appears to be a viable and more accurate method for obtaining reliable results without the tedious and laborious practices. Development of the BetaBuddy pipeline therefore enables us to predict the optimal application of exogenous stimuli that can be validated experimentally and provides a pathway to fully utilize the machine learning-based discoveries in the  $\beta$ -cell physiology. One novel method of data processing includes isolating active cells from the general population. Quantification of spikes from individual cells provides specialized data that allows us to track the active population between the control and experimental sets (S2 Fig). Active cells can be defined as any cell that spikes at least once during an imaging period (see Fig 7). These groups can be compared prior to and following external stimuli to determine any notable changes in calcium activity.

There are many methods through which insulin and  $\text{Ca}^{2+}$  activity can be stimulated. Pharmacological stimulation has been shown to increase calcium activity through drugs such as tolbutamide, which closes the  $\text{K}_{\text{ATP}}$  channels of the  $\beta$ -cells [48,49]. In addition to EFS, non-invasive stimulation can occur through applications such as photobiomodulation (PBM) therapy [7]. Using this pipeline, the direct comparison of changes in calcium dynamics between different stimuli can now readily be explored.



**Fig 6. Manual vs Automated ROI identification.** The total number of ROIs identified using manual and automated segmentation can be compared. The percent change between the two methods is displayed for each trial. The difference in labeled ROIs between the two segmentation methods ranged from 25–291% per trial.

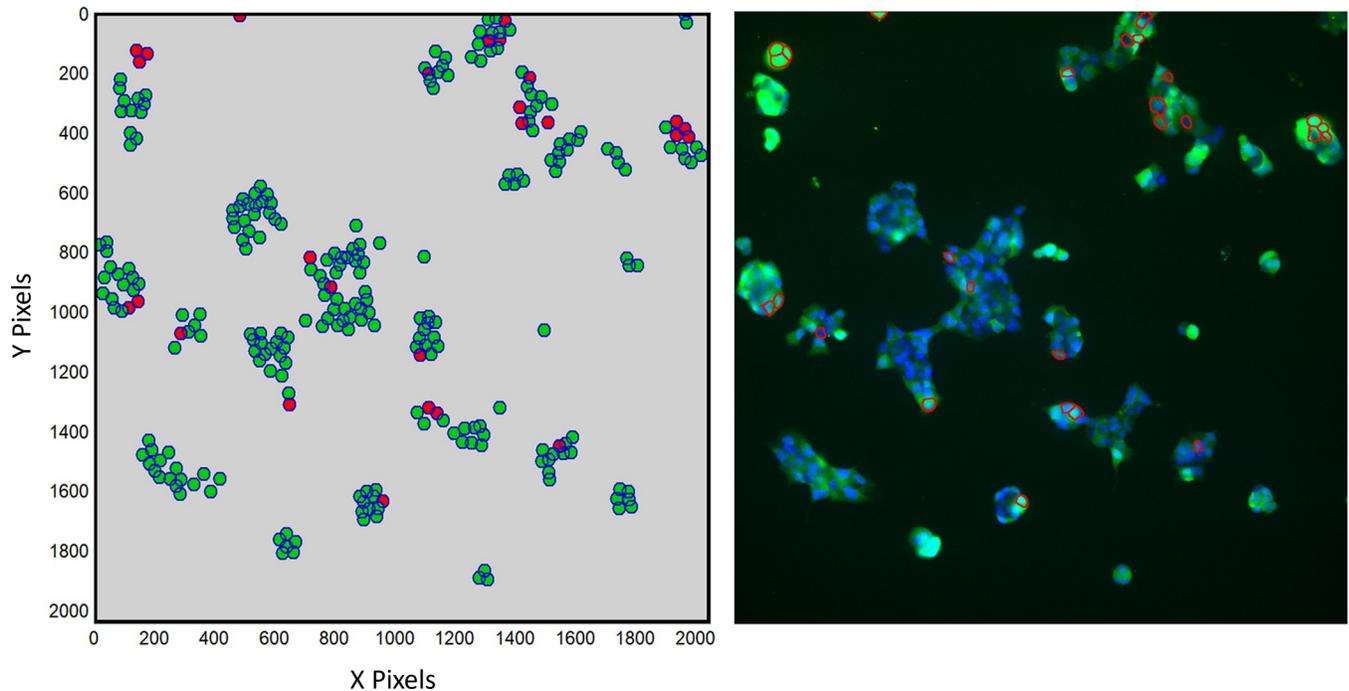
<https://doi.org/10.1371/journal.pone.0299549.g006>

### Future work

We plan to continue expanding the accessibility of our pipeline. A cleaner, more interactive user interface within the Jupyter Notebook is being developed along with an Anaconda package for researchers to easily create an environment with all the necessary packages pre-installed. Additionally, we are developing more user-friendly options for running different data analyses. This includes options for various statistical tests depending on researcher preference and the types of biological questions being asked. With these new features, users will be able to run desired tests within the pipeline along with having access to both the raw and normalized data.

We intend to use this improved process to first reanalyze a previously hand-segmented data set exploring the effects of PBM therapy on the  $\text{Ca}^{2+}$  oscillations in  $\beta$ - and  $\alpha$ -cells [7]. PBM is believed to increase insulin and glucagon secretion through accelerated ATP synthesis caused absorption of near-infrared light by the cytochrome c oxidase enzyme in the mitochondria [50–52]. We will also be developing  $\beta/\alpha$ -cell co-cultures to explore how calcium activity is affected through intercellular modulation and determine any changes in signaling responses post-PBM.

In future studies, spatially resolved interpretation of data offers an alternative to traditional statistical methods of analyzing calcium signals. Cell-specific spatial and fluorescent intensity



**Fig 7. Spatial visualization.** a) Cartesian coordinates were used to visualize the locations of spiking cells. Only cells currently spiking are outlined at a given time. The cells highlighted in red color, for example, exhibit calcium spiking activity at  $t = 25$  s. The visual analysis does not correlate to the actual size of cell clusters. b) The generated spatial data can be used to trace back and identify cells of interest in the original experimental images.

<https://doi.org/10.1371/journal.pone.0299549.g007>

data can be utilized, for example, to evaluate varying levels of insulin secretion in given locations (Fig 6). For example,  $\beta$ -cell location and population within an islet has been shown to strongly impact the overall islet dynamics [53]. Researchers can use these methods to not only look at current spiking patterns, but also begin to analyze how calcium signaling is affected by intercellular interactions.

Questions regarding calcium spiking frequency and cell-to-cell interaction through calcium propagation are now feasible for detailed analysis. This is perhaps one of the major advantages that were made possible by the development of automated segmentation and data evaluation. For example, intercellular communication often relies on calcium waves through gap junctions [54–56]. Since EFS we used in this work does not penetrate the cell, it may be hypothesized that the outer cells in a cluster of cells are likely affected by the EFS and generate calcium activities first. Through calcium intercellular communication, the inner cells in the cluster are affected at a later time. However, the cluster is collectively activated by EFS that leads to an increase in insulin production. Experiments are underway to validate this hypothesis and the results will again be used to train the model to recognize the calcium-dependent cell-to-cell communication that may be relevant to the islet physiology.

## Conclusions

Our current pipeline successfully isolates  $\beta$ -cells, tracks calcium signaling patterns over time, and produces preliminary data values that can be used in future statistical analyses. Automated segmentation reduces biases and decreases time required for analysis compared with manual segmentation. With this information, we are able to pinpoint individual cells with the most biologically significant spiking patterns. Additionally, individual calcium activities can be monitored over time, offering critical spatial and temporal information.

BetaBuddy has a multitude of benefits for potential research. Its customizable platform allows it to be adapted for use with multiple cell types. Its base structure has already been utilized for the automated analysis of mouse brain endothelial cells (MBECs) and human mesenchymal stem cells (hMSCs). Furthermore, the pipeline is designed in a way that researchers from diverse backgrounds can use it regardless of past coding experience.

This first analysis serves as a reliable test case for BetaBuddy's ability to batch analyze data. From here we will seek to apply the developed pipeline to evaluating other forms of stimuli (e.g. chemical/physical). Information obtained from this data may greatly assist in treatments for Type I Diabetes. The first step to developing a more quantifiable method for determining viable cells is obtaining a better understanding of the calcium dynamics of both  $\beta$ - and  $\alpha$ -cells. Co-culturing these cell types together will provide insight into how different forms of stimulation might improve cell viability in an in vivo setting. Moreover, our automated pipeline can decrease evaluation times for multiple treatment alternatives and even serve as a drug repurposing tool.

## Supporting information

**S1 Fig. Manual vs. Automated segmentation.** Deep learning algorithms have been found to greatly improve segmentation accuracy. (a) Previous hand segmentation methods often under-segmented the total ROIs present, often ignoring highly clustered areas. (b) Our automated system comprising merging DAPI with the targeted fluorescein channel, segmentation, and subsequent ROI tracking was able to consistently identify more cells at a higher accuracy and track localized signals.

(TIF)

**S2 Fig. Effects of EFS on active cell population.** Each point represents the number of active cells within a specific trial. Active cells were defined as any cell that spiked at least once during the trial's imaging period. The change in population following EFS stimulation can be tracked across each connecting line. The mean, upper quartile, and lower quartile of the active populations pre- and post-EFS are represented with their respective boxplots.

(TIFF)

## Acknowledgments

We would like to thank Dr. Caleb Liebman for his dedicated work culturing insulinoma ( $\beta$ TC-6) mouse cells and providing the calcium images.

## Author Contributions

**Conceptualization:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Data curation:** Anne M. Alsup, Kelli Fowlds.

**Formal analysis:** Anne M. Alsup, Kelli Fowlds.

**Funding acquisition:** Michael Cho, Jacob M. Lubner.

**Investigation:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Methodology:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Project administration:** Michael Cho, Jacob M. Lubner.

**Resources:** Michael Cho, Jacob M. Lubner.

**Software:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Supervision:** Michael Cho, Jacob M. Lubner.

**Validation:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Visualization:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Writing – original draft:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

**Writing – review & editing:** Anne M. Alsup, Kelli Fowlds, Michael Cho, Jacob M. Lubner.

## References

1. Over Cabrera, Berman Dora M., Kenyon Norma S., Camillo Ricordi, Per-Olof Berggren, Alejandro Caidedo. The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc Natl Acad Sci*. 2006 Feb 14; 103(7):2334–9. <https://doi.org/10.1073/pnas.0510790103> PMID: 16461897
2. Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC.  $\beta$ -Cell Deficit and Increased  $\beta$ -Cell Apoptosis in Humans With Type 2 Diabetes. *Diabetes*. 2003 Jan 1; 52(1):102–10.
3. Pisanía A, Weir GC, O'Neil JJ, Omer A, Tchishopashvili V, Lei J, et al. Quantitative analysis of cell composition and purity of human pancreatic islet preparations. *Lab Invest*. 2010 Nov 1; 90(11):1661–75. <https://doi.org/10.1038/labinvest.2010.124> PMID: 20697378
4. Yoon KH, Ko SH, Cho JH, Lee JM, Ahn YB, Song KH, et al. Selective  $\beta$ -Cell Loss and  $\alpha$ -Cell Expansion in Patients with Type 2 Diabetes Mellitus in Korea. *J Clin Endocrinol Metab*. 2003 May 1; 88(5):2300–8.
5. DiMeglio LA, Evans-Molina C, Oram RA. Type 1 diabetes. *The Lancet*. 2018 Jun 16; 391(10138):2449–62.
6. Melton D. The promise of stem cell-derived islet replacement therapy. *Diabetologia*. 2021 May 1;64. <https://doi.org/10.1007/s00125-020-05367-2> PMID: 33454830
7. Liebman C, Loya S, Lawrence M, Bashoo N, Cho M. Stimulatory responses in  $\alpha$ - and  $\beta$ -cells by near-infrared (810 nm) photobiomodulation. *J Biophotonics*. 2022 Mar 1; 15(3):e202100257.
8. Klec C, Ziomek G, Pichler M, Malli R, Graier WF. Calcium Signaling in  $\beta$ -cell Physiology and Pathology: A Revisit. *Int J Mol Sci*. 2019 Jan; 20(24):6110.
9. Idevall-Hagren O, Tengholm A. Metabolic regulation of calcium signaling in beta cells. *Adv Pancreat B Cell Funct Dysfunct*. 2020 Jul 1; 103:20–30. <https://doi.org/10.1016/j.semcd.2020.01.008> PMID: 32085965
10. Ashcroft FM, Rorsman P. KATP channels and islet hormone secretion: new insights and controversies. *Nat Rev Endocrinol*. 2013 Nov 1; 9(11):660–9. <https://doi.org/10.1038/nrendo.2013.166> PMID: 24042324
11. Mezo M, González-Warleta M, Castro-Hermida JA, Martínez-Sernández V, Ubeira FM. Field evaluation of the enhanced MM3-COPRO ELISA test for the diagnosis of Fasciola hepatica infection in sheep. *PloS One*. 2022; 17(3):e0265569. <https://doi.org/10.1371/journal.pone.0265569> PMID: 35324959
12. Schroeder AB, Dobson ETA, Rueden CT, Tomancak P, Jug F, Eliceiri KW. The ImageJ ecosystem: Open-source software for image visualization, processing, and analysis. *Protein Sci*. 2021 Jan 1; 30(1):234–49. <https://doi.org/10.1002/pro.3993> PMID: 33166005
13. Schindelin J, Arganda-carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012 Jul; 9(7):676–82. <https://doi.org/10.1038/nmeth.2019> PMID: 22743772
14. Tinevez JY, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, et al. TrackMate: An open and extensible platform for single-particle tracking. *Image Process Biol*. 2017 Feb 15; 115:80–90. <https://doi.org/10.1016/j.ymeth.2016.09.016> PMID: 27713081
15. Schwanbeck J, Oehmig I, Dretzke J, Zautner AE, Groß U, Bohne W. YSMR: a video tracking and analysis program for bacterial motility. *BMC Bioinformatics*. 2020 Apr 29; 21(1):166. <https://doi.org/10.1186/s12859-020-3495-9> PMID: 32349658
16. Kok RNU, Hebert L, Huelsz-Prince G, Goos YJ, Zheng X, Bozek K, et al. OrganoidTracker: Efficient cell tracking using machine learning and manual error correction. *PLOS ONE*. 2020 Oct 22; 15(10):e0240802. <https://doi.org/10.1371/journal.pone.0240802> PMID: 33091031
17. Mosby LS, Polin M, Köster DV. A Python based automated tracking routine for myosin II filaments. *J Phys Appl Phys*. 2020 May; 53(30):304002.

18. Rizwan I Haque I, Neubert J. Deep learning approaches to biomedical image segmentation. *Inform Med Unlocked*. 2020 Jan 1; 18:100297.
19. Laukamp KR, Pennig L, Thiele F, Reimer R, Görtz L, Shakirin G, et al. Automated Meningioma Segmentation in Multiparametric MRI. *Clin Neuroradiol*. 2021 Jun 1; 31(2):357–66.
20. Voigt SP, Ravikumar K, Basu B, Kalidindi SR. Automated Image Processing Workflow for Morphological Analysis of Fluorescence Microscopy Cell Images. *JOM*. 2021 Aug 1; 73(8):2356–65.
21. Gamarra M, Zurek E, Escalante HJ, Hurtado L, San-Juan-Vergara H. Split and merge watershed: A two-step method for cell segmentation in fluorescence microscopy images. *Biomed Signal Process Control*. 2019 Aug 1; 53:101575. <https://doi.org/10.1016/j.bspc.2019.101575> PMID: 33719364
22. Li G, Liu T, Nie J, Guo L, Chen J, Zhu J, et al. Segmentation of touching cell nuclei using gradient flow tracking. *J Microsc*. 2008; 231(1):47–58. <https://doi.org/10.1111/j.1365-2818.2008.02016.x> PMID: 18638189
23. Malpica N, de Solórzano CO, Vaquero JJ, Santos A, Vallcorba I, García-Sagredo JM, et al. Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry*. 1997 Aug 1; 28(4):289–97. [https://doi.org/10.1002/\(sici\)1097-0320\(19970801\)28:4<289::aid-cyto3>3.0.co;2-7](https://doi.org/10.1002/(sici)1097-0320(19970801)28:4<289::aid-cyto3>3.0.co;2-7) PMID: 9266748
24. Valen DAV, Kudo T, Lane KM, Macklin DN, Quach NT, DeFelice MM, et al. Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments. *PLOS Comput Biol*. 2016 Nov 4; 12(11):e1005177. <https://doi.org/10.1371/journal.pcbi.1005177> PMID: 27814364
25. Akram SU, Kannala J, Eklund L, Heikkilä J. Cell Segmentation Proposal Network for Microscopy Image Analysis. In: Carneiro G, Mateus D, Peter L, Bradley A, Tavares JMRS, Belagiannis V, et al., editors. *Deep Learning and Data Labeling for Medical Applications*. Cham: Springer International Publishing; 2016. p. 21–9.
26. Ahmed Raza S. E., Cheung L., Epstein D., Pelengaris S., Khan M., Rajpoot N. M. MIMO-Net: A multi-input multi-output convolutional neural network for cell segmentation in fluorescence microscopy images. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). 2017. p. 337–40.
27. Eschweiler D., Spina T. V., Choudhury R. C., Meyerowitz E., Cunha A., Stegmaier J. CNN-Based Pre-processing to Optimize Watershed-Based Cell Segmentation in 3D Confocal Microscopy Images. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). 2019. p. 223–7.
28. Sharif J. M., Miswan M. F., Ngadi M. A., Salam M. S. H., bin Abdul Jamil M. M. Red blood cell segmentation using masking and watershed algorithm: A preliminary study. In: 2012 International Conference on Biomedical Engineering (ICoBE). 2012. p. 258–62.
29. Koyuncu CF, Arslan S, Durmaz I, Cetin-Atalay R, Gunduz-Demir C. Smart Markers for Watershed-Based Cell Segmentation. *PLOS ONE*. 2012 Nov 12; 7(11):e48664. <https://doi.org/10.1371/journal.pone.0048664> PMID: 23152792
30. Zimmer C., Labruyere E., Meas-Yedid V., Guillen N., J. -C. Olivo-Marin. Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing. *IEEE Trans Med Imaging*. 2002 Oct; 21(10):1212–21. <https://doi.org/10.1109/TMI.2002.806292> PMID: 12585703
31. Ersoy I., Palaniappan K. Multi-feature contour evolution for automatic live cell segmentation in time lapse imagery. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2008. p. 371–4.
32. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: A generalist algorithm for cellular segmentation. *Nat Methods*. 2021 Jan 1; 18(1):100–6. <https://doi.org/10.1038/s41592-020-01018-x> PMID: 33318659
33. Schmidt U, Weigert M, Broaddus C, Myers G. Cell Detection with Star-Convex Polygons. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Cham: Springer International Publishing; 2018. p. 265–73.
34. Apthorpe N, Riordan A, Aguilar R, Homann J, Gu Y, Tank D, et al. Automatic Neuron Detection in Calcium Imaging Data Using Convolutional Networks. In: *Advances in Neural Information Processing Systems* [Internet]. Curran Associates, Inc.; 2016 [cited 2023 Mar 30]. Available from: <https://proceedings.neurips.cc/paper/2016/hash/0771fc6f0f4b1d7d1bb73bbbe14e0e31-Abstract.html>
35. Liu Z, Jin L, Chen J, Fang Q, Ablameyko S, Yin Z, et al. A survey on applications of deep learning in microscopy image analysis. *Comput Biol Med*. 2021 Jul 1; 134:104523. <https://doi.org/10.1016/j.combiomed.2021.104523> PMID: 34091383
36. Xu J, Zhou D, Deng D, Li J, Chen C, Liao X, et al. Deep Learning in Cell Image Analysis. *Intell Comput* [Internet]. [cited 2023 Mar 30];2022. Available from: <https://doi.org/10.34133/2022/9861263>

37. Müller A, Schmidt D, Xu CS, Pang S, D'Costa JV, Kretschmar S, et al. 3D FIB-SEM reconstruction of microtubule-organelle interaction in whole primary mouse  $\beta$  cells. *J Cell Biol.* 2021 Feb 1; 220(2): e202010039.
38. Viana MP, Chen J, Knijnenburg TA, Vasani R, Yan C, Arakaki JE, et al. Integrated intracellular organization and its variations in human iPS cells. *Nature.* 2023 Jan; 613(7943):345–54. <https://doi.org/10.1038/s41586-022-05563-7> PMID: 36599983
39. Pachitariu M, Stringer C. Cellpose 2.0: how to train your own model. *Nat Methods.* 2022 Dec 1; 19(12):1634–41. <https://doi.org/10.1038/s41592-022-01663-4> PMID: 36344832
40. Saad J, Fomich M, Díaz VP, Wang T. A novel automated protocol for ice crystal segmentation analysis using Cellpose and Fiji. *Cryobiology.* 2023 Jun 1; 111:1–8. <https://doi.org/10.1016/j.cryobiol.2023.02.002> PMID: 36773632
41. Hoeren F, Görmez Z, Richter M, Troidl K. Deetect: A Deep Learning-Based Image Analysis Tool for Quantification of Adherent Cell Populations on Oxygenator Membranes after Extracorporeal Membrane Oxygenation Therapy. *Biomolecules.* 2022 Dec 3; 12(12):1810. <https://doi.org/10.3390/biom12121810> PMID: 36551238
42. Reinbigler M, Cosette J, Guesmia Z, Jimenez S, Fetita C, Brunet E, et al. Artificial intelligence workflow quantifying muscle features on Hematoxylin–Eosin stained sections reveals dystrophic phenotype amelioration upon treatment. *Sci Rep.* 2022 Nov 19; 12:19913. <https://doi.org/10.1038/s41598-022-24139-z> PMID: 36402802
43. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep.* 2017 Dec 4; 7(1):16878. <https://doi.org/10.1038/s41598-017-17204-5> PMID: 29203879
44. Waisman A, Norris AM, Elías Costa M, Kopinke D. Automatic and unbiased segmentation and quantification of myofibers in skeletal muscle. *Sci Rep.* 2021 Jun 3; 11:11793. <https://doi.org/10.1038/s41598-021-91191-6> PMID: 34083673
45. Fisch D, Evans R, Clough B, Byrne SK, Channell WM, Dockterman J, et al. HRMan 2.0: Next-generation artificial intelligence–driven analysis for broad host–pathogen interactions. *Cell Microbiol.* 2021; 23(7):e13349. <https://doi.org/10.1111/cmi.13349> PMID: 33930228
46. Liebman C, Vu TM, Phillips A, Chen B, Cho M. Altered  $\beta$ -cell calcium dynamics via electric field exposure. *Ann Biomed Eng.* 2021 Jan; 49(1):106–14.
47. Ershov D, Phan MS, Pylvänäinen JW, Rigaud SU, Le Blanc L, Charles-Orszag A, et al. Bringing TrackMate into the era of machine-learning and deep-learning. *bioRxiv.* 2021 Jan 1; 2021.09.03.458852.
48. Szollosi A, Nenquin M, Henquin J. Pharmacological stimulation and inhibition of insulin secretion in mouse islets lacking ATP-sensitive  $K^+$  channels. *Br J Pharmacol.* 2010 Feb; 159(3):669–77. <https://doi.org/10.1111/j.1476-5381.2009.00588.x> PMID: 20128805
49. Henquin JC. Glucose-induced insulin secretion in isolated human islets: Does it truly reflect  $\beta$ -cell function in vivo? *Mol Metab.* 2021 Jun 1; 48:101212.
50. Hamblin MR. Mechanisms and Mitochondrial Redox Signaling in Photobiomodulation. *Photochem Photobiol.* 2018 Mar; 94(2):199–212. <https://doi.org/10.1111/php.12864> PMID: 29164625
51. Wang X, Tian F, Soni SS, Gonzalez-Lima F, Liu H. Interplay between up-regulation of cytochrome-c-oxidase and hemoglobin oxygenation induced by near-infrared laser. *Sci Rep.* 2016 Aug 3; 6(1):30540. <https://doi.org/10.1038/srep30540> PMID: 27484673
52. Castro KMR, de Paiva Carvalho RL, Junior GMR, Tavares BA, Simionato LH, Bortoluci CHF, et al. Can photobiomodulation therapy (PBMT) control blood glucose levels and alter muscle glycogen synthesis? *J Photochem Photobiol B.* 2020 Jun 1; 207:111877. <https://doi.org/10.1016/j.jphotobiol.2020.111877> PMID: 32298941
53. Johnston NR, Mitchell RK, Haythorne E, Pessoa MP, Semplici F, Ferrer J, et al. Beta Cell Hubs Dictate Pancreatic Islet Responses to Glucose. *Cell Metab.* 2016 Sep 13; 24(3):389–401.
54. Benninger RKP, Head WS, Zhang M, Satin LS, Piston DW. Gap junctions and other mechanisms of cell–cell communication regulate basal insulin secretion in the pancreatic islet. *J Physiol.* 2011 Nov 1; 589(22):5453–66. <https://doi.org/10.1113/jphysiol.2011.218909> PMID: 21930600
55. Evans WH, Martin PEM. Gap junctions: structure and function (Review). *Mol Membr Biol.* 2002 Jan 1; 19(2):121–36. <https://doi.org/10.1080/09687680210139839> PMID: 12126230
56. Benninger RKP, Zhang M, Head WS, Satin LS, Piston DW. Gap Junction Coupling and Calcium Waves in the Pancreatic Islet. *Biophys J.* 2008 Dec 1; 95(11):5048–61. <https://doi.org/10.1529/biophysj.108.140863> PMID: 18805925