

# Development of an image analysis tool quantifying the efficacy of intra-cell trafficking of the CFTR protein

Catherine Hastings

5th June 2017

## Abstract

Mutations cause defects in the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) protein leading to its reduced function as a chloride ion channel, causing cystic fibrosis (CF). The  $\Delta F508$  mutation leads to reduced intra-cell trafficking, as well as decreased gating function. Wide-field imaging of the CFTR protein tagged with yellow fluorescent protein (YFP) allows for the study of CFTR protein transportation within the cell. Here, an image analysis tool is developed to quantify the amount of CFTR protein at the cell membrane for CFTR-WT (wild type) and CFTR- $\Delta F508$ . Standard image segmentation techniques are applied to images displaying mCherry fluorescence which is expressed throughout the cytosol giving a reference for the cell outline. A simple metric is then applied which demonstrates a ten fold increase in the YFP fluorescence detected at the membrane of cells transfected with CFTR-WT over CFTR- $\Delta F508$ .

# 1 Introduction

## 1.1 Cystic fibrosis and the CFTR protein

The Cystic Fibrosis Transmembrane conductance Regulator (CFTR) is a protein which acts as a chloride ion channel in the membrane of epithelial cells (Lieberman et al., 2009). There are more than 1000 mutations of the gene which encodes for the CFTR protein that have been found to cause the disease cystic fibrosis (CF). The most common CF-causing mutation is CFTR- $\Delta$ F508 (or CFTR-F508del) which accounts for approximately 66% of all CF cases worldwide (Bobadilla et al., 2002).

The CFTR mutation causes an overall decrease in the outflow of chloride ions from epithelial cells. This decrease in chloride ion transport causes, in turn, a decrease in the outflow of water from the cell by osmosis, leading to the epithelial mucus being sticky. The mucus builds up in organs such as the liver, lungs and pancreas. Patients with CF suffer a wide variety of symptoms including malabsorption of certain nutrients due to a decrease in the pancreatic secretion of enzymes into the intestines, and an increase in secondary respiratory infections (Lieberman et al., 2009).

The different mutations vary in the mechanism by which they reduce the outflow of chloride ions. Mutations are split into six classes based upon how they affect the production, processing or functioning of the CFTR protein (Horsley et al., 2015). These six classes are displayed in Figure 1.1.



**Figure 1.1:** A diagram showing some of the defects in the function of the CFTR protein by different CF-causing mutations. From Bush et al. (2016).

The CFTR- $\Delta$ F508 mutation is of class II, meaning that it is transcribed and translated normally, but then misfolds. This results in poor trafficking around the cell, and poor processing in the Golgi. Therefore the majority of the CFTR protein does not reach the membrane. When CFTR- $\Delta$ F508 protein is correctly transported to the membrane, it additionally suffers poor gating function (Langron et al., 2017).

## 1.2 Experimental method

Experiments have been designed which aim to study specific defects caused by different classes of CFTR gene mutation. Langron et al. (2017) describe a method by which a yellow fluorescent protein (YFP) is attached to a specific site on the CFTR protein. The addition of the YFP has does not affect the function of the CFTR protein and allows for its distribution to be studied when imaging.

Following the described methodology, HEK293 (human embryonic kidney) cells were transfected with plasmids encoding for an mCherry protein (a red fluorescing protein) and either YFP-CFTR-WT (wild type) or YFP-CFTR- $\Delta$ F508. After transfection the cells are left to divide in an incubator at 37°C for 24 hours. During the time the cells divide and duplicate the added DNA, producing the specified proteins. The mCherry protein is expressed throughout the cytosol, whereas the CFTR protein should theoretically be transported to the membrane of the cells.

Images were gathered using the ImageXpress, an image-acquisition system equipped with wide-field inverted fluorescence microscope, CMOS camera and fluidics robotics. During imaging, the cell plates were contained in an environmental chamber at 37°C. All images were gathered using a 60x objective and the laser intensity was managed so as to produce a good level of fluorescence without damaging the cells.

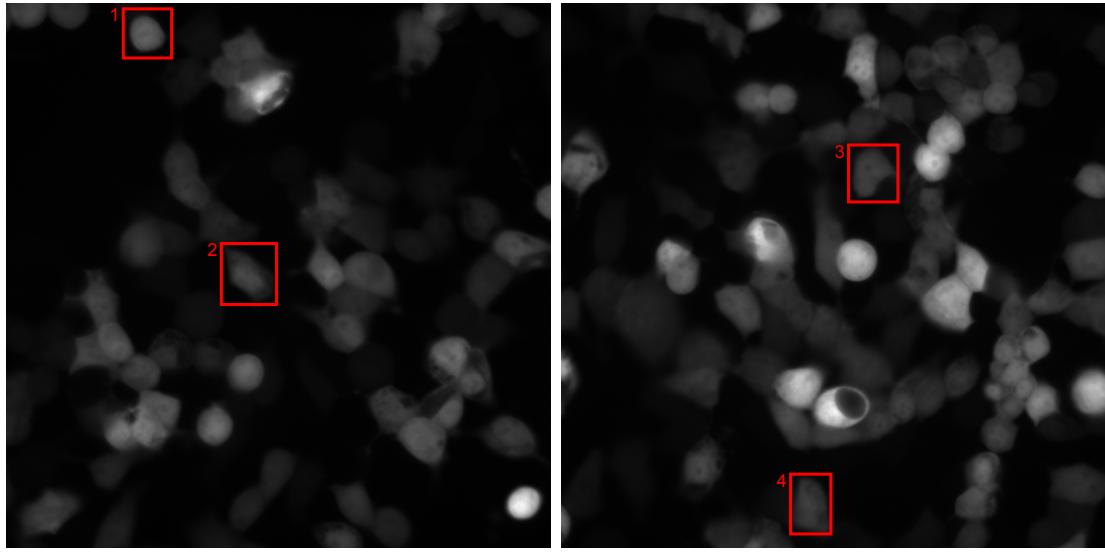
Care was taken during the experimental procedure in order to optimise the cell density. Very high cell densities result in images with very few isolated cells, making the image segmentation process more difficult. Low cell densities risk the HEK cells dying before images can be collected.

For each fluorescing protein, 45 images were collected. Images were taken from 9 sites for each of 5 wells in a 96-well plate (see Figure 2.1 for example images). The images are 2160 by 2160 pixels in dimension and no pixel binning was used in the image gathering process. Mathematica automatically scales the pixel intensities between 0 and 1, though the image intensity values are initially of `uint16` type meaning they range between 0 and 65535. No adjustment of the images was performed prior to analysis. All example images here displayed have had some image adjustment: without at least rescaling of the image data, very little is visible to the eye.

## 2 Method

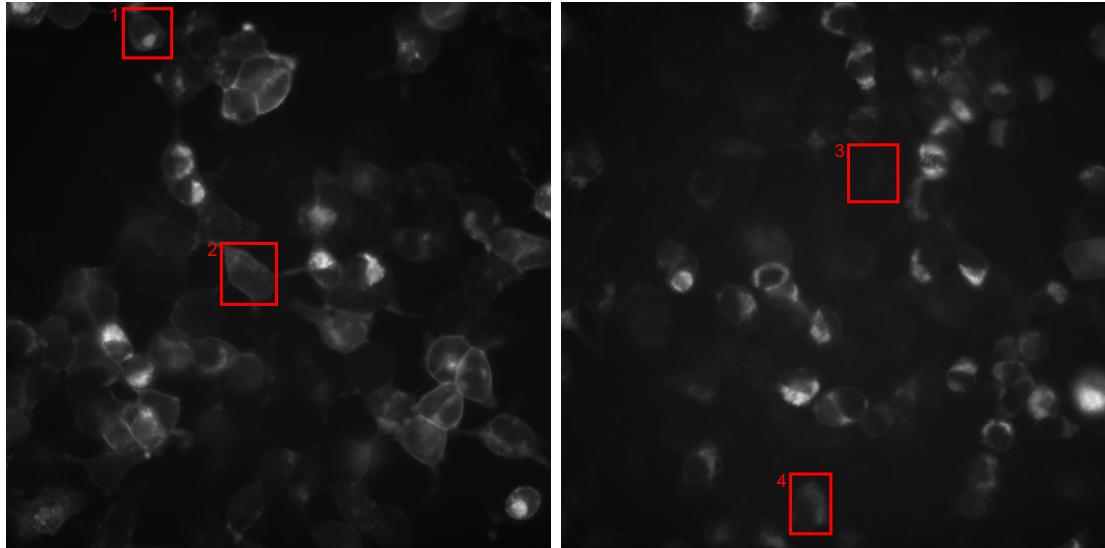
Once images have been obtained by the experimental method outlined in Section 1.2, standard techniques for image processing techniques (such as those discussed in Gonzalez and Woods (2008)) are used to segment and analyse the images.

Figure 2.1 shows examples of the images obtained by the experimental procedure. Throughout the description of the segmentation, processing and analysis, intermediate results will be displayed using these images (Table 2.2), and using the individual cells marked therein



(a) mCherry and CFTR-WT.

(b) mCherry and CFTR- $\Delta$ F508.



(c) YFP and CFTR-WT.

(d) YFP and CFTR- $\Delta$ F508.

**Figure 2.1:** Examples of unprocessed images obtained from the experimental procedure, with fluorescence intensity adjusted so as to be visible. Marked are four cells which will be used to display example results of the segmentation and analysis.

(Table 2.1). These example cells have been chosen because they show variety in the possible distributions of the CFTR protein. It is worth noting that while two of the example cells are of wild type and two have the CFTR- $\Delta$ F508 mutation, the YFP distributions shown occur in cells of both conditions.

## 2.1 Image processing

### 2.1.1 Segmentation

As the mCherry is expressed throughout the cytosol, a good approximation of the cells border can be found by analysing these images. Another benefit of performing the seg-

mentation upon the mCherry images is that they are very similar for cells expressing both CFTR-WT and CFTR- $\Delta$ F508 as can be seen in Figures 2.1a and 2.1b. The steps followed here follow the ideas shown in Gonzalez and Woods (2008), whilst the practical coding advice is by Mathematica (2017).

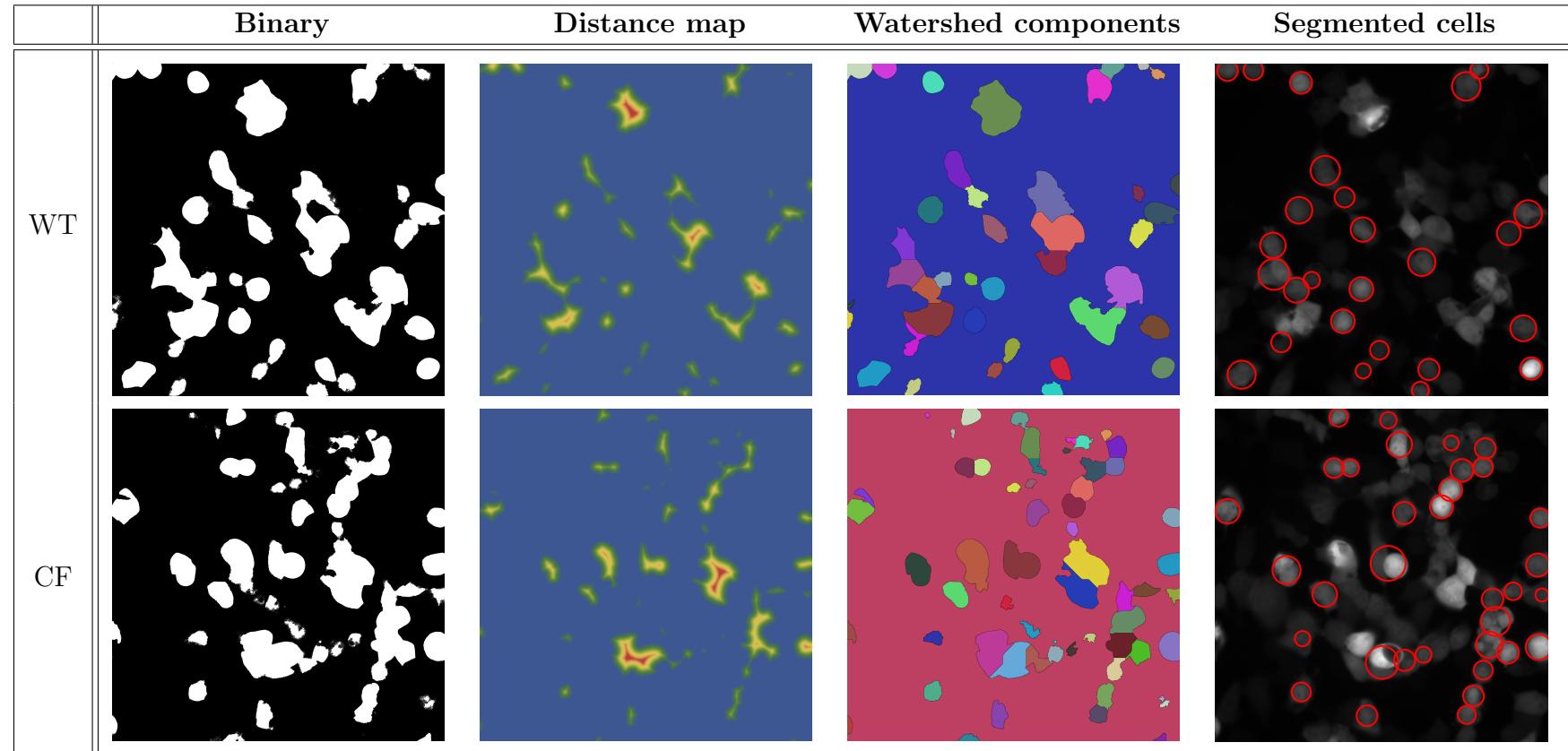
The first step of the segmentation process is to ‘binarise’ the raw image. This means that all pixel intensity values above a threshold are given a value of 1, and all values below the threshold are given a value of 0. Here the threshold is calculated as 90% of Mathematica’s `FindThreshold` function. The reduced threshold value is used as the cell borders found with the automatic threshold value appeared by eye to be universally within the actual border. Any small holes in the binary images are filled using the `FillingTransform`. Example binary images produced are shown in the first column of Table 2.2.

The next step is create a distance map based upon the binary image. For each pixel with a value of 1 in the binary image, the distance transform finds the distance in pixels to a nearest pixel with value 0. Pixels with value 0 are not changed. The result of the distance transform function can be seen in the second column of Table 2.2.

The local maxima of the distance map are located in order to create a marker for each cell (as shown in Table 2.1). A `GradientFilter` of the binary image is performed in order to blur the boundaries between the black and white areas. This blurred image is overlaid with the cell markers and this becomes the foundation image upon which to perform the segmentation.

	mCherry	Binary	Dist. map	Marker	Segment	YFP
1						
2						
3						
4						

**Table 2.1:** Images demonstrating the stages of the automated segmentation upon the four example cells highlighted in Figure 2.1. Segmentation is based solely upon the mCherry image and the YFP image is included only for reference.



**Table 2.2:** Displaying intermediate results of the segmentation process.

Mathematica has several in-built functions for performing image segmentation. Here the `WatershedComponents` function was used and each cell is now labelled as a ‘component’ of the segmentation process. These components have been coloured and are shown in Table 2.2, column three. However, studying these images shows that the software has found it very difficult to separate cells which are very close together, creating large components which represent multiple cells. Therefore, a first filtering of automatically identified components is performed. Cells are only included in the analysis if their lengths are between 100 and 300 pixels and their widths are between 80 and 200 pixels. This means that very small components are also discarded. The full results of the segmentation are given in the final column of Table 2.2.

### 2.1.2 Filtering cells

Studying the last column of Table 2.2 shows that filtering the segmentation components by size has excluded many, but not all of the abutting cells from the analysis. Also, Table 2.1 shows that the bounding box of selected components appears to be too small, and cutting off the edges of the cells. This is particularly evident for example cells 2 and 4. In order to solve these problems further processing and filtering is performed.

When performing segmentation, Mathematica automatically defines a ‘bounding box’ for each component. These automatic bounding boxes are those used as the component image borders in Table 2.1. Each component bounding box is now enlarged by a set amount of 20 pixels, to ensure that the entire cell is included.

The next segmentation components to be excluded from the analysis are those bordering the edge of the image. These components are excluded as it cannot be ensured that there is data for the entire cell. Abutting cells are also excluded from the analysis. This is because the automated segmentation process has difficulty locating an accurate border between touching cells. Abutting cells are found by studying the coloured component image shown in column five of Table 2.1. If an individual component image contains more than two distinct colours (the black boundary line is not counted), this component is excluded from the analysis.

### 2.1.3 Clarifying cell boundaries

The segmentation process in Section 2.1.1 has identified components and the filtering process in Section 2.1.2 has removed components which are not ideal for analysis. However, the borders for the components so far identified are not very distinct, as can be seen in the second column of Table 2.1. Therefore, further image processing is carried out in order to find a smooth and distinct cell border.

During the segmentation process a single binarising threshold was identified for each full raw image. Now, a new threshold for binarising the mCherry image is found for each enlarged single component image, and the results are displayed in the third column of Table 2.3.

In order to smooth the border of these binary images we apply Mathematica’s `GaussianFilter` and `FillingTransform` functions. This smoothed border is then also

morphologically dilated by a fixed number of pixels so as to visually match the binary and actual cell borders as closely as possible. It should be noted that there is some trading between the parameters for binarising the component image and for morphologically dilating the component. However, they do have slightly different effects and some tweaking of these parameters has been carried out in order to find the optimal values.

The analysis relies upon comparing pixel intensity values at different distances from the border of the cell. Using Mathematica's `DistanceTransform` function gives each pixel with a value of 1 in a binary image a value which corresponds to the minimum distance to any pixel with value 0. This function does not require its result to be an integer, therefore the distance transform is applied and the results are rounded down to the nearest integer. The results of this transform are shown in Table 2.3, column 5. The analysis also requires that the relative pixel intensity values for outside of the cell are studied. Therefore the `DistanceTransform` function is also performed on the negative of the smoothed binary image, with the results being shown in column six of Table 2.3. The positive and negative distance maps can be added together (pointwise), and the resulting distance map used as a basis of all image intensity analysis.

	mCherry	YFP	Rough binary	Smooth binary	Positive dist. map	Negative dist. map
1						
2						
3						
4						

**Table 2.3:** Showing the stages of the process to smooth the component borders and create a distance map using the smoothed cell border as the zero point.

## 2.2 Image analysis

The image segmentation and processing described in Section 2.1 have resulted in a distance map for each individual component giving the position of each pixel relative to the border of the cell. Using this to label each pixel in the YFP intensity image, information can

now be collated about the YFP intensity value as it relates to the distance from the cell border.

Thus far, the edge of the cell has here been referred to as the component (or cell) ‘border’. This has been used to mean the line between pixels labeled in the distance map as 1 and -1. (It is a slight eccentricity of this image analysis tool that no pixels are labeled with distance 0 relative to the border. This has lead to some slight programming impracticalities, but is clearer when defining which pixels are inside the cell.)

The term component or cell ‘boundary’ is now here defined to be the area of an image (YFP or mCherry) corresponding to distance map values between 1 and 10 inclusive. The boundary width of 10 has been chosen to correspond with an approximate component diameter of 130 pixels.

It is then assumed that any increase in the intensity measurement of YFP fluorescence within the newly defined boundary to correspond to an increased accumulation of the CFTR protein on the membrane. In order to quantify the quantity of CFTR protein within the membrane an answer is sought to the question,

What is the mean YFP intensity per pixel within the boundary and how does this compare with the mean YFP intensity per pixel within the entire cell?

In answering this questions, it is hoped that a simple quantification of a cell’s ability to transport CFTR protein to the cell membrane can be obtained.

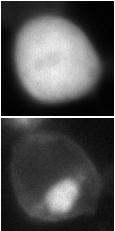
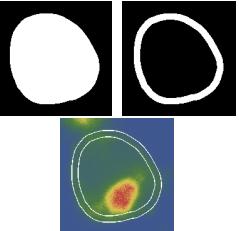
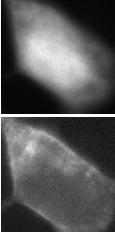
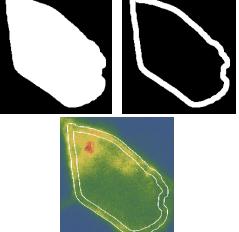
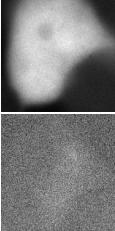
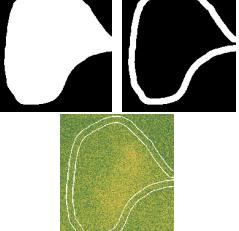
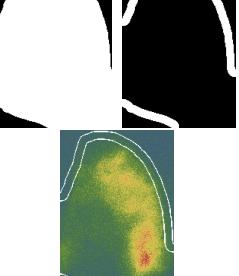
## 3 Results

### 3.1 Studying the intra-cell CFTR distribution

Table 3.1 gives the results for each example component marked in Figure 2.1. The mean intensity data for each component is given, with separate values calculated for inside and outside the cell and within the cell boundary. Across the four component shown in Table 3.1 these values range between 0.00506 and 0.00923. These values show that there can be a significant difference between mean intensity values in different regions of an image.

The value of most importance given in Table 3.1 is that highlighted in red: the ratio of the mean YFP intensity inside the cell boundary to the mean YFP intensity inside the entire cell (including the boundary). This value gives the most telling information on how the CFTR protein is distributed throughout the cell, particularly how much CFTR has accumulated on the cell membrane. When this value is close to or above 1 it is a good indication that there is more CFTR protein on the cell membrane than there is distributed throughout the rest of the cell. Example component 2 is a perfect example of this, as can be verified by studying the corresponding YFP image.

However, example component 3 also gives a high value for the boundary to cell intensity

	Raw data	Processed data	
1			<p>Mean intensity outside cell = 0.00506      Mean intensity inside cell = 0.00856      Inside/outside intensity ratio = 0.591</p> <p>Mean intensity inside boundary = 0.00724      Mean intensity inside whole cell = 0.00856      Boundary/cell intensity ratio = <b>0.845</b></p>
2			<p>Mean intensity outside cell = 0.00603      Mean intensity inside cell = 0.00873      Inside/outside intensity ratio = 0.691</p> <p>Mean intensity inside boundary = 0.00923      Mean intensity inside whole cell = 0.00873      Boundary/cell intensity ratio = <b>1.06</b></p>
3			<p>Mean intensity outside cell = 0.00523      Mean intensity inside cell = 0.00538      Inside/outside intensity ratio = <b>0.972</b></p> <p>Mean intensity inside boundary = 0.00529      Mean intensity inside whole cell = 0.00538      Boundary/cell intensity ratio = <b>0.982</b></p>
4			<p>Mean intensity outside cell = 0.00506      Mean intensity inside cell = 0.00640      Inside/outside intensity ratio = 0.790</p> <p>Mean intensity inside boundary = 0.00554      Mean intensity inside whole cell = 0.00640      Boundary/cell intensity ratio = <b>0.865</b></p>

**Table 3.1:** Question 1 results. For each component, the raw YFP and mCherry images are given, along with the binary images showing the whole cell and the area labeled the cell boundary. The additional image is a coloured version of the YFP image (for greater clarity) imposed with a ring corresponding to the component boundary. Mean intensity data is given for inside and outside the cell, and within the cell boundary. Mean ratios of these values are also given.

ratio. This is not because there is on average a higher YFP intensity within the boundary than in the rest of the cell, but because the distribution of the YFP intensity is very even throughout the cell, and in fact, throughout the entire image. This can be seen in that the inside to outside cell mean intensities ratio is also very close to 1 (highlighted in blue). Many components from the segmentation seemed to display this behaviour, and by this metric read as a false positive.

Figure 3.1 shows results of the metric as applied to all components. These results show that on average the mean intensity ratio for inside to outside the cell is much lower than the ratio for mean intensity for inside the boundary to inside the entire cell. This is encouraging as one would expect there to be more CFTR protein within cells than without. It also suggests that background noise from overlapping and abutting cells has not had a large impact on the results.

```
Mean ratio of intensity inside and outside cell for WT = 0.653589
Mean ratio of intensity inside and outside cell for CF = 0.770234

Mean ratio of intensity within boundary and within cell for WT = 0.900883
Mean ratio of intensity within boundary and within cell for CF = 0.871407

For all cells:
Number of cells with boundary/cell ratio over 1 for WT = 56 of 219
Number of cells with boundary/cell ratio over 1 for CF = 22 of 277

When excluding cells with high inside/outside ratio:
Number of cells with boundary/cell ratio over 1 for WT = 48 of 207
Number of cells with boundary/cell ratio over 1 for CF = 5 of 195
```

**Figure 3.1:** Print-out of the results of analysing the difference in mean YFP intensity per pixel inside and outside of the cell, and within the cell boundary. A more complete read-out can be found in Section A.

It can also be seen that the mean ratio of intensities inside and outside the cell is higher for cells with the CFTR- $\Delta$ F508 mutation than those with the CFTR-WT gene. This possibly is due to a higher number of cells with a very even distribution of the YFP fluorescence in the CF-type cells; cells like example component 3.

The ratio for inside the boundary to inside the cell is larger for WT cells than for those of CF type. This is perhaps due to a higher number of cells like example component 2 within the WT cell population, and would support the idea that WT cells should show a high level of YFP within the component boundary.

However, the difference in both of these ratios between wild type and cystic fibrosis cells is not very large. It seems likely that this is because for both cell type there is a large variation in CFTR distribution behaviours. The mean ratio value therefore does not show significant differences.

It is for this reason that a threshold is set for the mean ratio of YFP intensities within the boundary and within the entire cell. Counting the number of cells above a given threshold, provided that the threshold is chosen sensibly, will give an idea of the number of cells qualitatively mimicking the behaviour displayed by example component 2. The threshold has been set to 1 because this ensures that the average pixel intensity within the boundary is higher than the average pixel intensity within the entire cell. Such a high threshold provides a very clear signal of there being a large amount of YFP fluorescence within the boundary, and therefore a probable large amount of CFTR protein at the membrane.

The results show that when a threshold of 1 is set, and the number of cells with either CFTR variant and with mean boundary/cell ratio values of greater than the threshold are counted, 25.6% of wild type cells satisfy the criteria versus only 7.94% of cells with CFTR- $\Delta$ F508. This proportion is more than 3 times greater for wild type cells than those with the cystic fibrosis causing protein, showing a significant difference in cells with different CFTR proteins.

When studying the example components, it was shown that the distribution of the YFP fluorescence was very even in component 3. It is unclear whether cells showing this type of spread of the CFTR protein occur as a consequence of the experimental method, the image gathering process, or whether they are a feature of the CFTR protein itself. If the lattermost is true, the results ought not to be excluded from the analysis, and thus far they have not been.

However, by this metric, an even distribution of the YFP fluorescence both inside and outside of the cell produces a false positive. That is, these cells visibly do not have more CFTR protein at the membrane of the cell. We can exclude them from the analysis by asserting that no cell should be counted if the mean YFP pixel intensity outside of the cell is greater than 90% of that inside the cell. When this exclusion is performed, 23.2% of WT cells have a mean boundary/cell ratio over 1 versus 2.56% for CF type.

## 3.2 Testing the efficacy of the image analysis tool

The results in Section 3.1 show that this image analysis tool gives simple results which demonstrate a significant difference between WT and CF cells. However, there are several tests that can be performed to solidify the validity of these results.

### 3.2.1 Similarity of WT and CF mCherry images

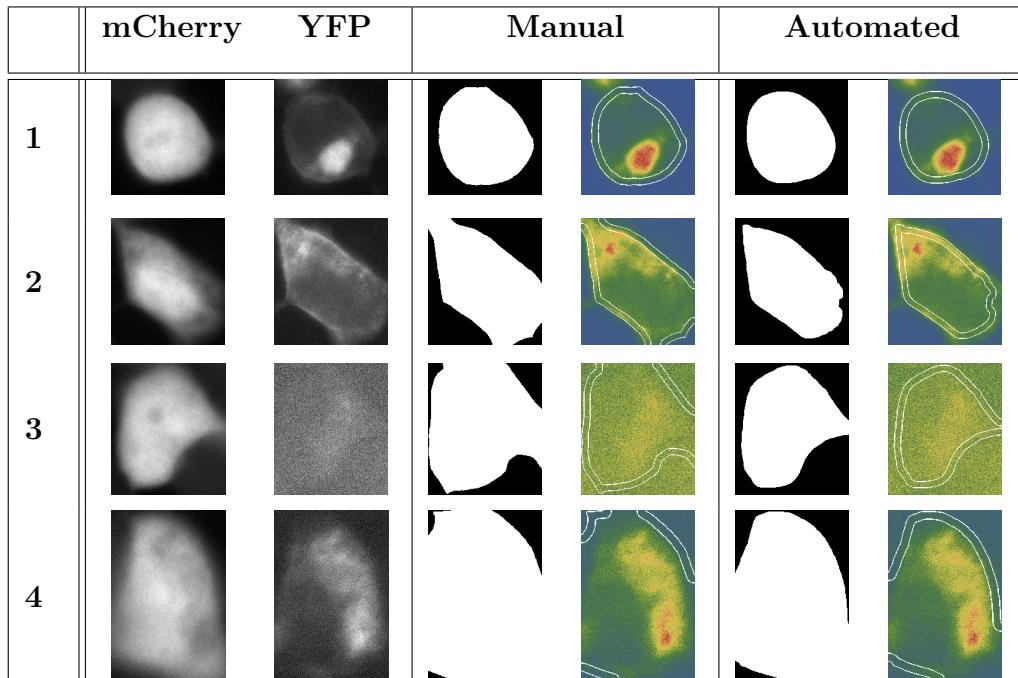
Studying the raw intensity data from the images has two benefits. Firstly, some quantitative information about the images here processed allows future users to verify the differences between the images data they have collected and the image data the tool was tested on. Secondly, Table 3.2 shows that the data for the mCherry images are not significantly different for WT and CF images. It is important to verify this as the segmentation process relies on this similarity.

	mCherry		YFP	
	WT	CF	WT	CF
Max	0.296	0.248	0.105	0.0587
Median	0.00488	0.00560	0.00497	0.00475
Min	0.00226	0.00218	0.00220	0.00200
Mean	0.0112	0.0119	0.00590	0.00511
STD	0.0146	0.0136	0.00318	0.00192

**Table 3.2:** Averaged pixel intensity data over all images.

### 3.2.2 Manual versus automated segmentation

The next test to validate the function of the image analysis tool is to compare the automated segmentation process with results provided by manual segmentation. Table 3.3 shows that the component identified by manual segmentation is universally larger than that provided by the automated segmentation.



**Table 3.3:** Comparing manual and automated segmentation. The raw component images are given for the mCherry and YFP fluorescence. For both manual and automated segmentation, the binary images are given along with a coloured version of the YFP image with the area defined as the boundary marked.

Table 3.3 also seems to show that the by eye the automated segmentation is better than the manual segmentation. Manual segmentation was performed only upon the mCherry images, and the manual segmentation procedure proved that these images to do not provide a very clear boundary for the cell. The border appears to extend much further than would be guessed from studying the YFP images.

### 3.2.3 Testing the analysis metric

We can test the efficacy of the metric used to analyse the distribution of the CFTR protein throughout the cell by performing the same analysis on the mCherry images for each component.

```
Red mean ratio of intensity inside and outside cell for WT = 0.268745
Red mean ratio of intensity inside and outside cell for CF = 0.311721

Red mean ratio of intensity within boundary and within cell for WT = 0.647801
Red mean ratio of intensity within boundary and within cell for CF = 0.659027

For all cells:
Number of red cells with boundary/cell ratio over 1 for WT = 0 of 219
Number of red cells with boundary/cell ratio over 1 for CF = 0 of 277

When excluding cells with high inside/outside ratio:
Number of red cells with boundary/cell ratio over 1 for WT = 0 of 219
Number of red cells with boundary/cell ratio over 1 for CF = 0 of 277
```

**Figure 3.2:** Similarly to Figure 3.1, a print-out showing the results of analysing the difference in mean mCherry intensity per pixel inside and outside of the cell, and within the cell boundary. A more complete read-out can be found in Section A.

Comparing the results in Figures 3.1 and 3.2 shows that the image analysis tool gives very different results dependent upon the input. The fact the analysing the mCherry data in the same way as the YFP data gives no results for either WT or CF cells is very encouraging. It shows unequivocally that no mCherry fluorescence accumulates within the cell boundary, as would be expected given that this protein is expressed throughout the cytosol.

The results for the WT and CF cells are also very similar, providing more evidence that the mCherry images for both conditions are similar enough to perform a fair segmentation upon.

It can also be noted that the mean ratio of mCherry fluorescence intensity in the boundary to the intensity within the cell is far below 1. This means that the threshold set is high, and that surpassing it is a strong criterion to fulfil.

## 4 Discussion

An image analysis tool has been developed to quantify the effect that a mutation in the CFTR gene has on the trafficking of the CFTR protein to the cell membrane. Here, this tool has been used to compare the difference in CFTR distribution of the wild type protein against that with the  $\Delta F508$  mutation. It has been shown that a higher proportion of

cells transfected with the wild type CFTR protein results display a significant amount of CFTR-associated fluorescence within the cell boundary.

This result aligns well with the understanding that the CFTR- $\Delta$ F508 mutation results in poor intra-cell trafficking meaning that this ion channel protein has a low success rate in reaching the cell membrane (Horsley et al., 2015).

The results of this tool have been shown in detail on some example cells. The good correlation between the output of the tool and the observed features of the individual cells provides a proof-of-concept in the tool's efficacy. However, the fact that the tool identifies cells with an even distribution of YFP inside and outside of the cell as a positive result could be deemed as some cause for concern. Though, there is still a very noticeable difference in the proportion of cells giving a positive result even when these false positives are included. Any future user has the ability to observe the results with and without these cells' inclusion. It is interesting to observe that many more cells transfected with the mutated CFTR gene display this 'evenly distributed' behaviour. This is perhaps evidence that this behaviour is a result of the malfunctioning protein rather than an artefact of the method.

Compared here were the function of the CFTR-WT and CFTR- $\Delta$ F508 protein. However, there are many other CF-causing mutations which could be studied using this image analysis tool. In addition, drugs have been designed which attempt to modulate CFTR function (such as ivacaftor as described by Eckford et al. (2012)). These drugs must be designed to effect a single class of CFTR mutation; for example ivacaftor improve the function of CFTR with class III mutations (?). It is sometimes the case that a drug can resolve the defective function of the CFTR protein, but simultaneously impair its function at another part of the cell. In these situations, this image analysis tool could be used in the development of CFTR modulators for treating cystic fibrosis.

## 4.1 Improvements

During the development of this tool two other metrics were tested. For example, it was asked whether there were a higher proportion of cells with a local maximum in the mean YFP fluorescence intensity within the boundary for CFTR-WT over CFTR- $\Delta$ F508 (see Table A.2). The results were very similar to those found for the metric here presented. It is possible that having multiple metrics would provide more solid evidence in the difference in the trafficking of the CFTR protein, and multiple metrics could be included with future versions of the tool. However, the simplicity of having a single number as an output is a benefit of the tool as it is.

In providing only a single number, it is then very important that the accuracy of this number is assured. It would be worth testing the impact of different cell densities on the tool's efficacy. Even, if future users were to follow this methodology and plate a similar density of cells there is likely to be some variation in the conditions.

Currently the tool only analyses isolated cells. It should be verified that isolated cells are

not qualitatively or quantitatively different. This could be done via manually segmenting some clustered cells or improving the tool’s segmentation method so that clustered cells could be included automatically.

The manual segmentation should be carried out again, perhaps allowing the use of both the YFP and mCherry images.

Table A.2 shows plots of the mean YFP intensity against the distance is pixels relative to the cell border. For these plots the cell border was determined by intensity thresholding. However, it can be observed that there seems to be a strong correlation in the location of the maximal gradient of the mCherry intensity and the location of the cell border. If the average distance between the maximum mCherry gradient and the intensity thresholded border were equal for WT and CF cells, it would provide strong evidence that the segmentation procedure works similarly for both cell types.

## 4.2 Extensions

A different metric for analysing how much CFTR protein has been could be implemented. For each cell, a fixed number of linear projections could be made from the centre of the cell (the point of maximum distance to the border) outward in various directions. Then measuring the YFP intensity along each projection, it could be ascertained what proportion of these projections have their maximum value within the boundary. This would perhaps give a more reliable result than the current metric, in that it is less likely to be subject to identifying false positives.

The tool was only tested on images of cells at 60x magnification and with no binning of pixels. It might be useful to broaden the scope of the tool so that it can be applied to images of varying quality.

## Acknowledgements

The experimental procedures and image gathering were done by Stella Prins, for whose help throughout this project I am extremely grateful. I woud also like to thank my supervisors Lewis Griffin and Paola Vergani.

## References

Joseph L. Bobadilla, Milan Macek, Jason P. Fine, and Philip M. Farrell. Cystic fibrosis: A worldwide analysis of cftr mutations—correlation with incidence data and application to screening. *Human Mutation*, 19(6):575–606, 2002.

Andrew Bush, Margaret Hodson, and Diana Bilton. *Hodson and Geddes’ cystic fibrosis*. CRC Press, fourth edition, 2016.

Paul DW Eckford, Canhui Li, Mohabir Ramjeesingh, and Christine E Bear. Cystic fibrosis transmembrane conductance regulator (cftr) potentiator vx-770 (ivacaftor) opens the defective channel gate of mutant cftr in a phosphorylation-dependent but atp-independent manner. *Journal of Biological Chemistry*, 287(44):36639–36649, 2012.

Rafael C. Gonzalez and Richard E. Woods. *Digital image processing*. Pearson Prentice Hall, third edition, 2008.

Alex Horsley, Steve Cunningham, and Alistair Innes. *ORML Cystic Fibrosis*. Oxford University Press, Oxford, UK, second edition, 2015. ISBN 9780199582709.

Emily Langron, Michela I Simone, Clémence Delalande, Jean-Louis Reymond, David L Selwood, and Paola Vergani. Improved fluorescence assays to measure the defects associated with f508del-cftr allow identification of new active compounds. *British journal of pharmacology*, 174(7):525–539, 2017.

Michael Lieberman, Allan D. Marks, and Matthew Chansky. *Marks' basic medical biochemistry: a clinical approach*. Wolters Kluwer/Lippincott Williams Wilkins Health, third edition, 2009.

Mathematica. Analyze segmented cells in an image, Mathematica Documentation.

<http://reference.wolfram.com/language/example/AnalyzeSegmentedCellsInAnImage.html>, 2017. [Accessed 28-May-2017].

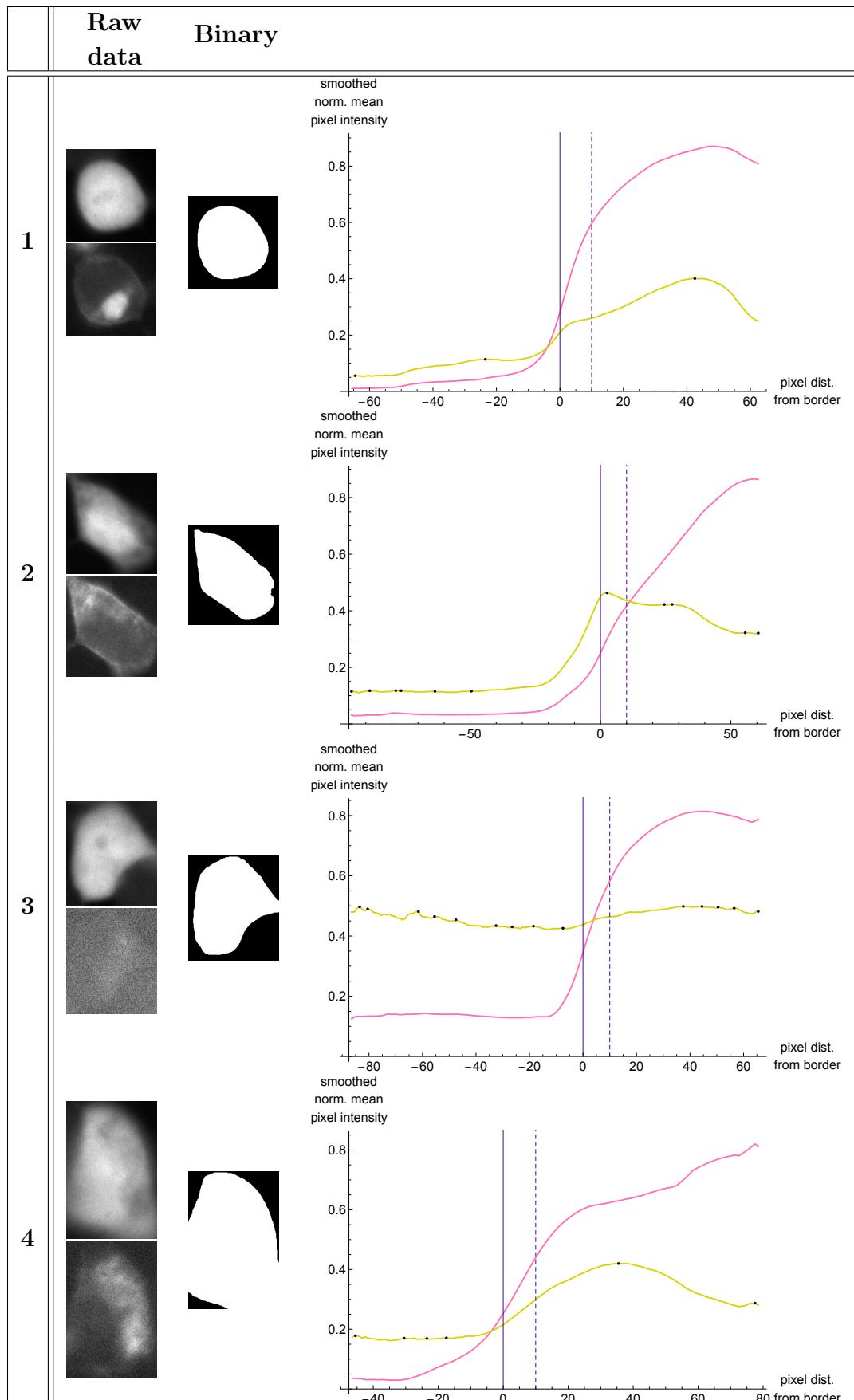
## A Additional results

		WT	CF
Number of cells per image after cell size filter	Min	1	17
	Max	45	40
	Mean	24.3	29.4
	<b>Total</b>	<b>1095</b>	<b>1325</b>
Number of cells per image after full filter	Min	0	1
	Max	10	12
	Mean	4.97	6.15
	<b>Total</b>	<b>219</b>	<b>277</b>

**Table A.1:** Numbers of cells after segmentation (Section 2.1.1) and after filtering (Section 2.1.2).

### Caption for Table A.2

Plots showing how the mean YFP (yellow) and mCherry (pink) varies in relation to the distance in pixels from the border of the cell. The intensity values have been smoothed (with a moving average, parameter 5) and normalised, so that the absolute intensity should not be compared. The solid vertical line demonstrates the cell border, and the dashed line shows the inner edge of the boundary. Local maxima of the smoothed curves have also been plotted (points).



**Table A.2:** See above for caption.

Mean intensity outside cell for WT = 0.00584673

Mean intensity outside cell for CF = 0.00493774

Mean intensity inside boundary for WT = 0.00847265

Mean intensity inside boundary for CF = 0.00578909

Mean intensity inside cell for WT = 0.00979922

Mean intensity inside cell for CF = 0.00695744

Mean ratio of intensity inside and outside cell for WT = 0.653589

Mean ratio of intensity inside and outside cell for CF = 0.770234

Mean ratio of intensity within boundary and within cell for WT = 0.900883

Mean ratio of intensity within boundary and within cell for CF = 0.871407

For all cells:

Number of cells with boundary/cell ratio over 1 for WT = 56 of 219

Number of cells with boundary/cell ratio over 1 for CF = 22 of 277

When excluding cells with high inside/outside ratio:

Number of cells with boundary/cell ratio over 1 for WT = 48 of 207

Number of cells with boundary/cell ratio over 1 for CF = 5 of 195

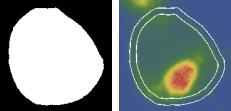
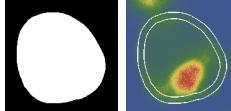
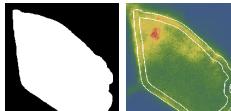
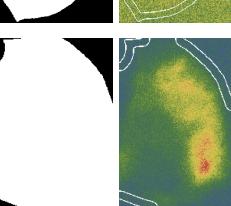
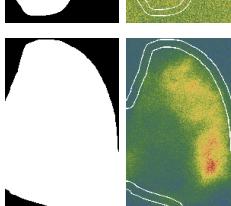
Number of cells for WT = 219

Number of cells for CF = 277

Number of images for WT = 45

Number of images for CF = 45

**Figure A.1:** More complete print-out for the CFTR distribution metric, analysing the YFP images.

	Manual segmentation	Automated segmentation
1	 Inside/outside intens. ratio = 0.596 Boundary/cell intens. ratio = <b>0.804</b>	 Inside/outside intens. ratio = 0.591 Boundary/cell intens. ratio = <b>0.845</b>
2	 Inside/outside intens. ratio = 0.642 Boundary/cell intens. ratio = <b>0.971</b>	 Inside/outside intens. ratio = 0.691 Boundary/cell intens. ratio = <b>1.06</b>
3	 Inside/outside intens. ratio = 0.980 Boundary/cell intens. ratio = <b>0.978</b>	 Inside/outside intens. ratio = 0.972 Boundary/cell intens. ratio = <b>0.982</b>
4	 Inside/outside intens. ratio = 0.794 Boundary/cell intens. ratio = <b>0.820</b>	 Inside/outside intens. ratio = 0.790 Boundary/cell intens. ratio = <b>0.865</b>

**Table A.3:** Comparing metric results for automated and manual segmentation. The binary images are given along with a coloured version of the YFP image with the area defined as the boundary marked. The ratios between the mean intensity data for inside and outside the cell are given along with ratios for within the cell boundary and the entire cell.

```
Red mean intensity outside cell for WT = 0.00898248  
Red mean intensity outside cell for CF = 0.00892116
```

```
Red mean intensity inside boundary for WT = 0.0243857  
Red mean intensity inside boundary for CF = 0.0213
```

```
Red mean intensity inside cell for WT = 0.0377836  
Red mean intensity inside cell for CF = 0.0326398
```

```
Red mean ratio of intensity inside and outside cell for WT = 0.268745  
Red mean ratio of intensity inside and outside cell for CF = 0.311721
```

```
Red mean ratio of intensity within boundary and within cell for WT = 0.647801  
Red mean ratio of intensity within boundary and within cell for CF = 0.659027
```

For all cells:

```
Number of red cells with boundary/cell ratio over 1 for WT = 0 of 219  
Number of red cells with boundary/cell ratio over 1 for CF = 0 of 277
```

When excluding cells with high inside/outside ratio:

```
Number of red cells with boundary/cell ratio over 1 for WT = 0 of 219  
Number of red cells with boundary/cell ratio over 1 for CF = 0 of 277
```

Number of cells for WT = 219

Number of cells for CF = 277

Number of images for WT = 45

Number of images for CF = 45

**Figure A.2:** More complete print-out for the CFTR distribution metric, analysing the mCherry images.