

FISHtoFigure Quick-start Tutorial

Calum Bentley-Abbot

University of Glasgow

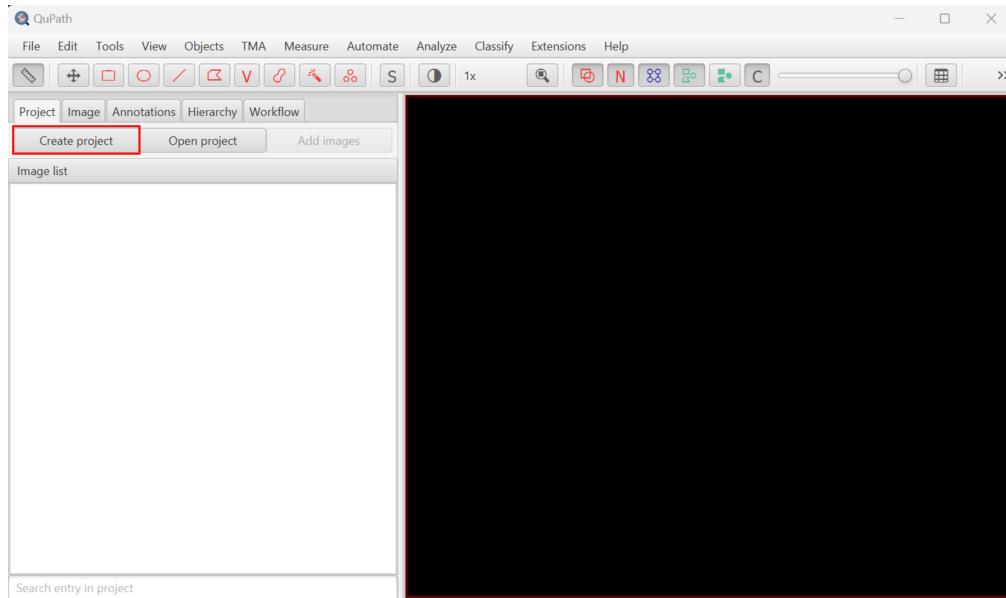
1 Introduction

This document provides a step-by-step guide to using the FISHtoFigure platform to quantify co-expression of fluorescent markers in microscopy data. For troubleshooting and a detailed description of the underlying code see "FISHtoFigure User Information".

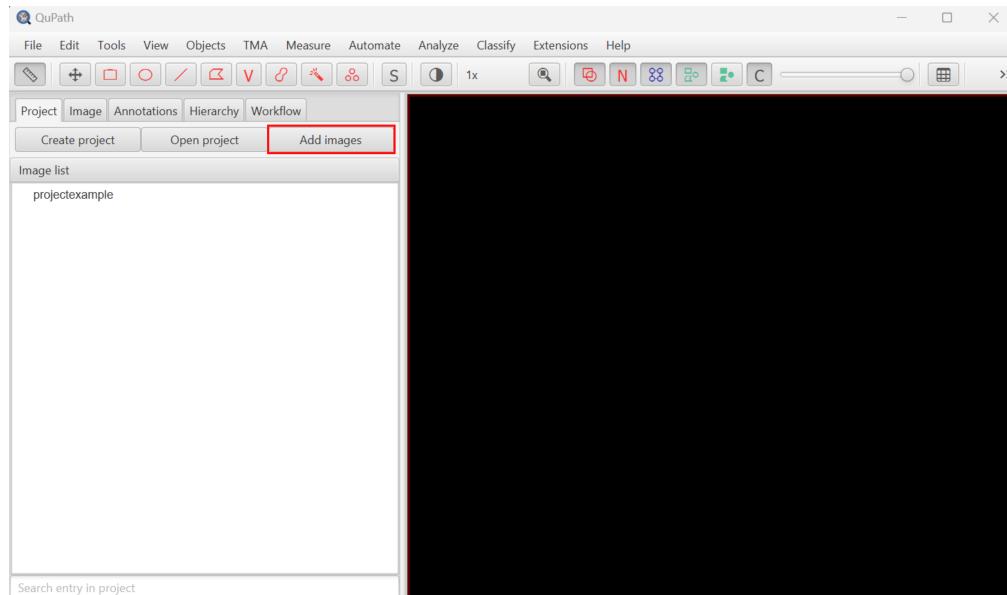
2 QuPath: Cell Boundary and Fluorescent Spot Identification

Cell boundaries and fluorescent spots are identified using the open source software QuPath (<https://github.com/qupath/qupath>). These data are then used by FISHtoFigure to quantify cellular expression profiles.

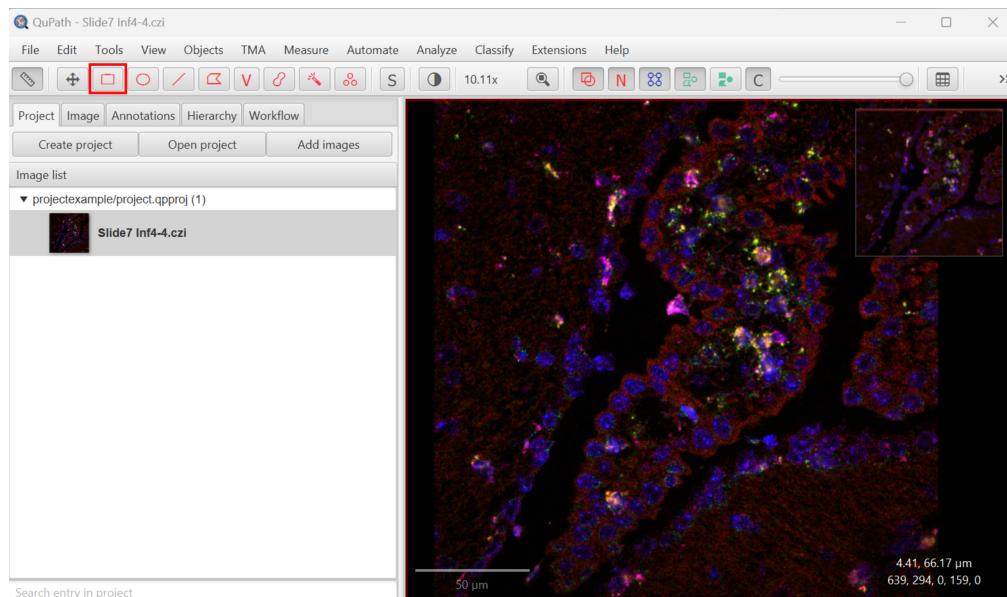
Begin by creating a new project in QuPath:



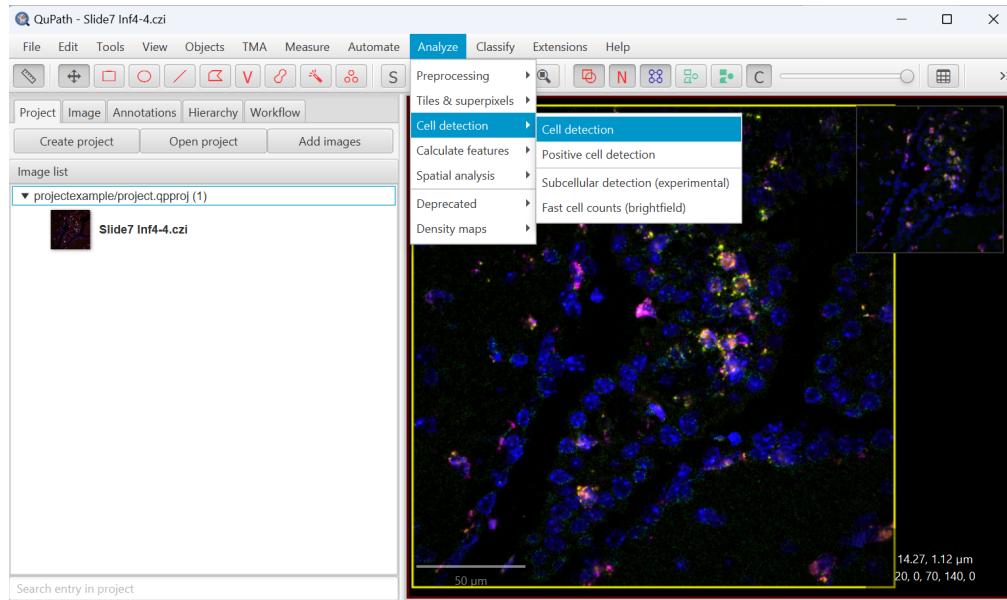
Add images using the button shown below, drag and drop images into the pop-up window and click "Import":



Create an annotation covering your whole image using this button (only the area within the annotation will be analysed in the next steps):

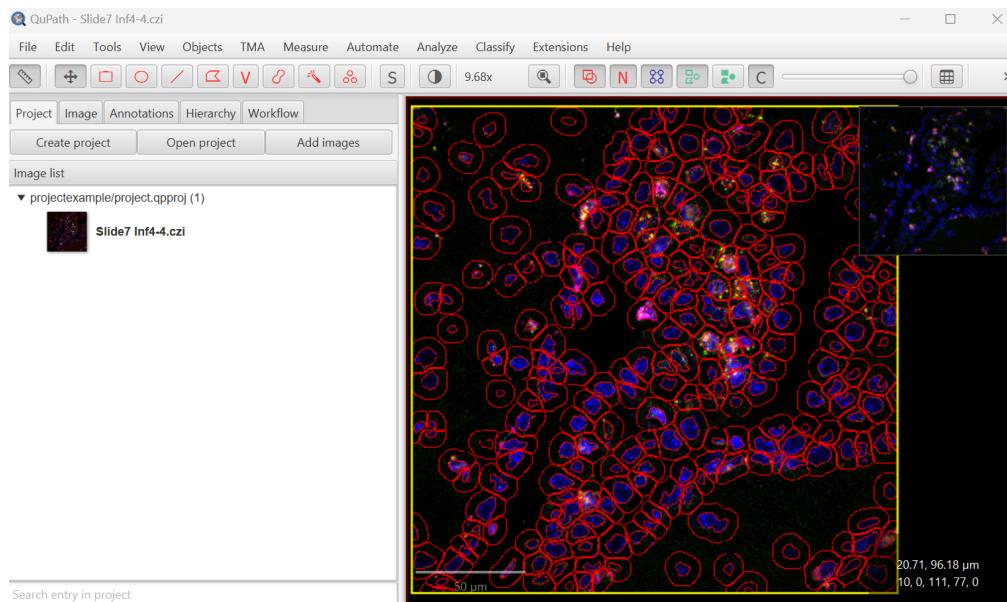


To identify cell boundaries, go to Analyse → Cell detection → Cell detection:

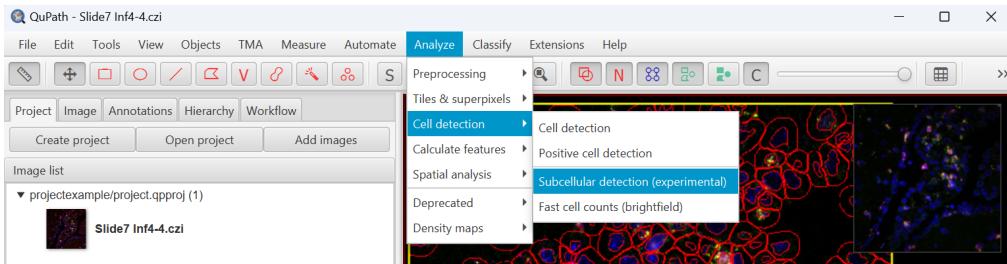


Within the cell detection pop-up, set the detection channel to the channel used for the nuclear stain. "Cell expansion" defines the distance from the nucleus where the cell boundary will be drawn, set this based on the approximate size of cells in your image. Adjust the settings in the pop-up such that each nucleus is recognised and nuclei are not oversegmented.

The results:

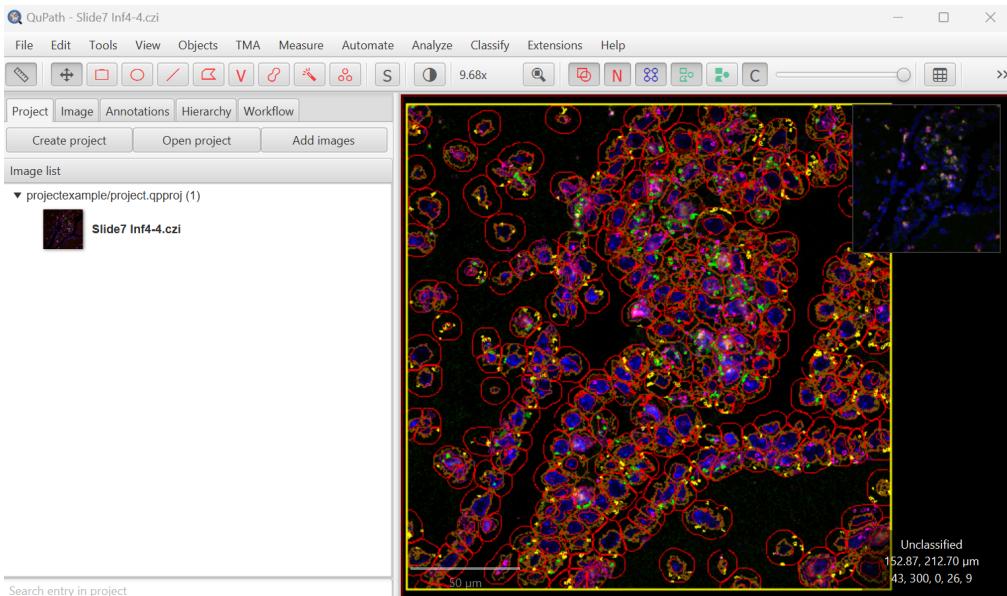


Then, to identify fluorescent spots, go to Analyse → Cell detection → Subcellular detection:

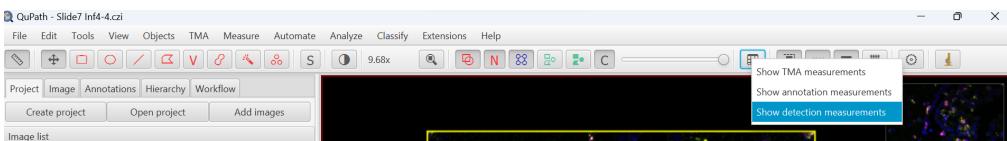


Within the pop-up, first set the detection threshold, this is the fluorescent brightness a spot must have to be detected and should be based on a negative control. A value of -1 means this channel will not be used for detecting spots, use this value in the channel for the nuclear stain and any channels which do not show any fluorescent spots.

The results:



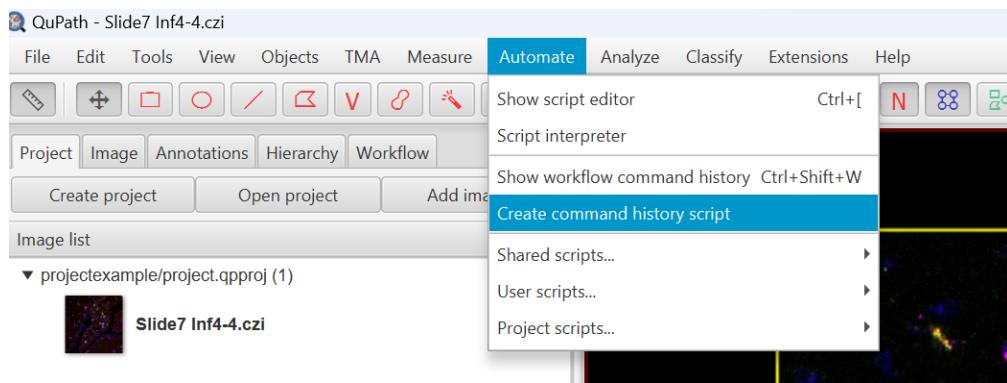
Clicking the button in the top right of the QuPath window and selecting "Show detection measurements" will show you all the identified cells and fluorescent spots as a spreadsheet, this is the data which we will process with FISHtoFigure:



Save this data as a textfile using the "Save" button in the pop-up window.

2.1 QuPath: Batch processing

If multiple images from the same dataset are being analysed, the commands described above can be turned into a script for batch processing in Automate → Create command history script:



This script will contain all commands which have been run so far, if you have attempted to run any detection commands multiple times to refine parameters, make sure to remove these duplicate lines from the script.

Once you finish editing the script, use the three dots in the bottom right and "Run for project" to batch process all images.

3 FISHtoFigure: Analysing differential mRNA expression

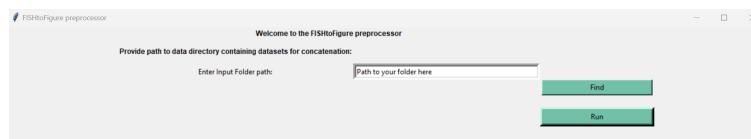
Before beginning, there are a few things to note: FISHtoFigure can only conduct differential analysis between datasets which have the same markers, make sure all datasets being analysed have the same markers and that fluorescent channels associated with these markers appear in the same order (this will be the case already if they were processed the same way in QuPath). Before beginning, create a new folder for all your datasheet text files.

3.1 Preprocessing of batch data

If you have processed multiple images in QuPath and wish to analyse these together as part of a single dataset (i.e. multiple images from a given treatment group), these data can be combined using the FISHtoFigure preprocessing tool available in the GitHub repository.

Place all datasheets from QuPath into a folder and rename these in ascending order; "Dataset1", "Dataset2", etc. Open the preprocessing tool and copy the path to your folder into the tool. The path can be found navigating to your folder in a terminal and using the "pwd" command, or by right clicking and selecting "Copy as Path" on Windows machines.

Clicking "Find" followed by "Run" will create a new text file in your folder containing the data from all your images. This can then be used with the main FISHtoFigure program:

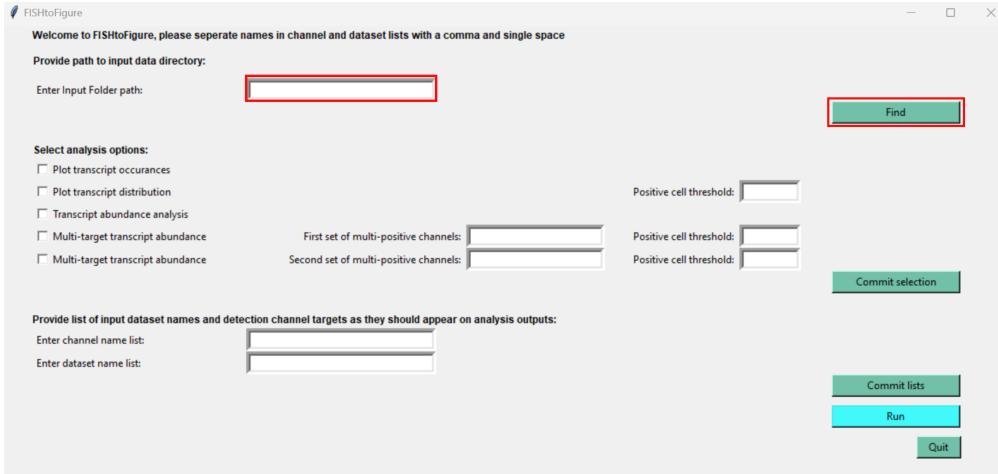


3.2 Main FISHtoFigure tool

Prior to any analysis, create a new folder. This folder should contain all datasheets which are to be analysed by FISHtoFigure, these can either be datasheets directly from QuPath for a single image, or datasheets containing information from multiple images (If you are analysing datasheets created using the preprocessing

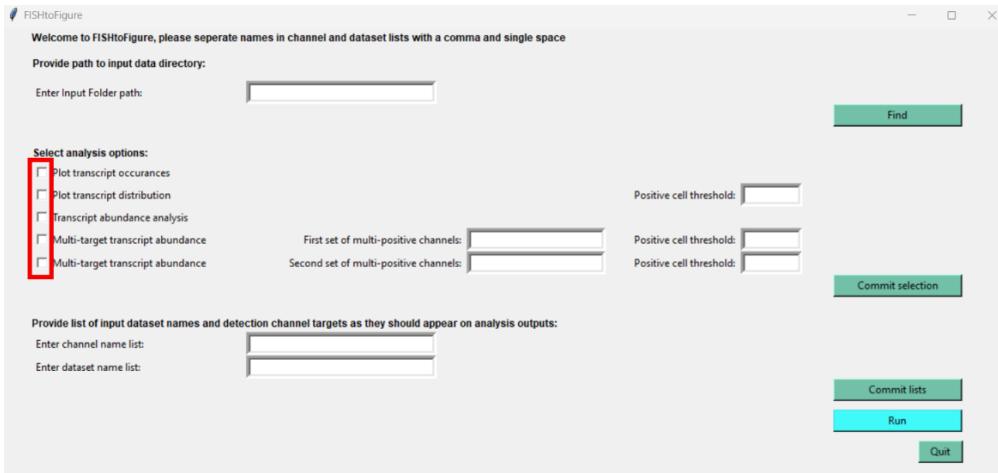
tool DO NOT include the original QuPath files used to create these tools, only the final datasheets being analysed). Rename these files in ascending order; "Dataset1", "Dataset2", etc.

Open FISHtoFigure and provide the path to your folder in the first text box, click "Find" to confirm:

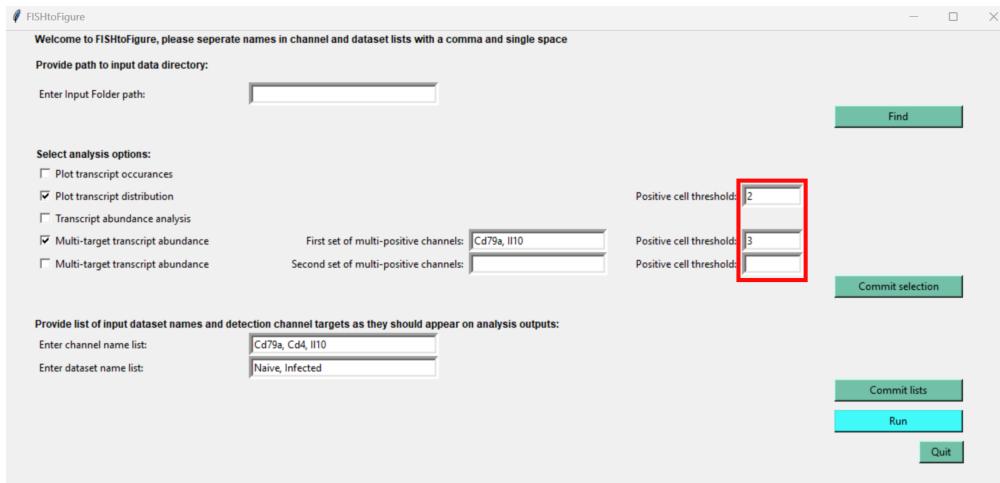


In the next section, select which analyses you would like to perform on your data. Briefly, "Plot Transcript Occurrences" plots the number of times each transcript occurs across a dataset. "Plot Transcript Distribution" produces a scatter plot of the fluorescent spots across your image (note that this will not work for datasheets containing data from multiple images). "Transcript abundance analysis" produces multiple graphs relating to the expression of each of your fluorescent markers in each of your datasets. "Multi-target transcript abundance" produces graphs relating to the co-expression of a specific combination of markers across each dataset (i.e. you pick a set of markers in your dataset, only cells expressing all of these will be analysed), two sets of markers can be analysed simultaneously, hence this option appears twice. More details of each option are available on GitHub and in the "FISHtoFigure User Guide".

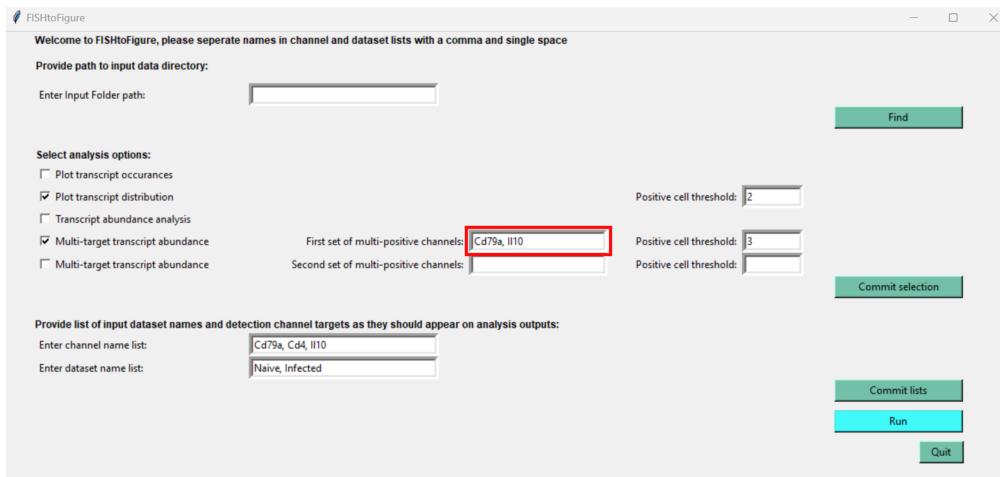
Select your chosen analysis option by checking the boxes:



If you select "Plot transcript distribution" or "Multi-target abundance" you can define a "Positive cell threshold". This allows you to define how many transcripts from each target cells must express to be registered during analysis. If left blank, no threshold is applied. This is useful for removing cells which are miscategorised or noisy, see the preprint @ <https://www.biorxiv.org/content/10.1101/2023.06.28.546871v1> for details:

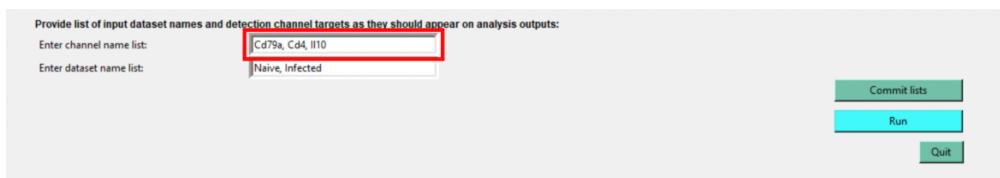


When selecting "Multi-target abundance" you need to define which markers you wish to analyse. This should be entered as a list, with each markers name being seperated by a comma and a space. The names in this list should be exactly the same as the markers name in the "Enter channel names list" textbox below:



Once all the above options are selected and relevant fields filled out, click "Commit" to confirm.

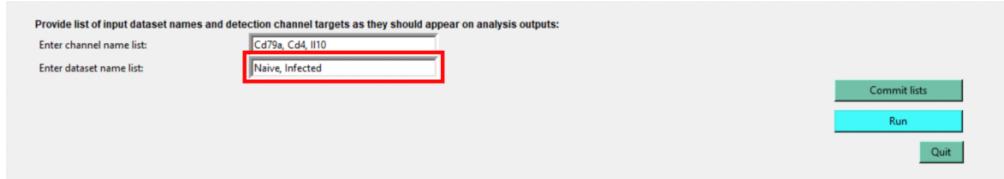
Next, you need to define the names of the markers in your datasets, this is how markers will be named in plots generated by FISHtoFigure so call them whatever you like. These markers must appear in the SAME ORDER as they were quantified in QuPath, an easy wa to check this is to check which order they appear in the pop-up window when identifying fluorescent spots in QuPath. Enter names as a list separating each entry by a comma and space:



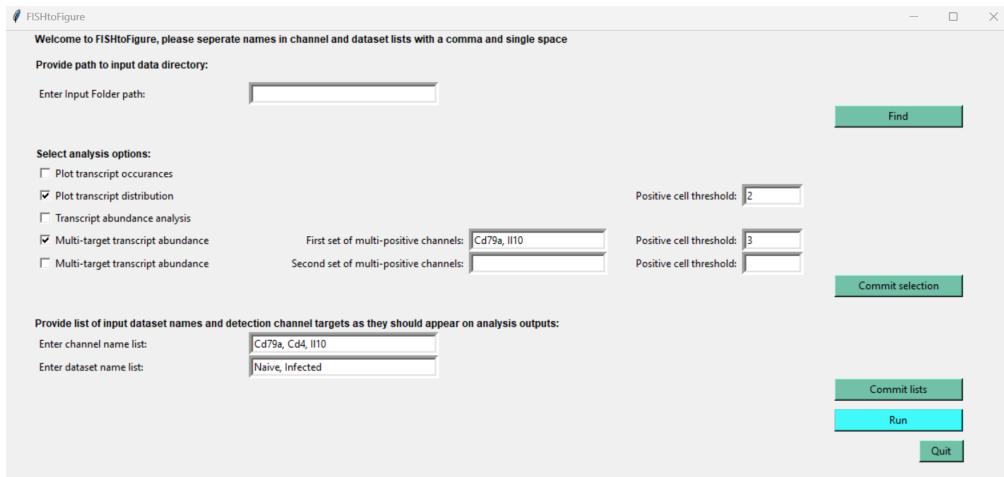
Note: Make sure the number of marker names is the same as the number of markers in the dataset. If one of your markers has no spots, this marker will be ignored by QuPath. This can cause users to provide

one more marker name than the QuPath dataset actually has, causing FISHtoFigure to hang forever. This is not common, but if you run into this problem, check your dataset for the number of fluorescent markers (look for columns named "Channel 1: Num spots", "Channel 2: Num spots" etc. to work this out).

Now, define the names of the datasets being analysed in the same fashion, again, make sure the number of names entered is equal to the number of datasheets in your folder and that the first name in the list corresponds to "Dataset1" in your folder, etc.:



Your window should now look something like this, click the final "Commit" button and "Run" to start analysis:



FISHtoFigure will generate a new folder within your folder called "F2F Output", in here you'll find all graphs for your analysis options, along with a Output script detailing the running of the code, the datasets analysed and details of specific marker profiles for each cell in your dataset.

FISHtoFigure doesn't take very long to run, since we are just working with the datasheets, not the images themselves. In our larger experiments, running analysis on 2 datasets containing data from over 30 images took less than 5 minutes on a consumer grade laptop. If you find that FISHtoFigure hangs for a long time check to see if the any folder within "F2F Output" contains an error message, if it does, see the FISHtoFigure User Information" document on GitHub for details.

Thanks for using FISHtoFigure, if you have any questions please email me at:
c.bentley-abbot.1@research.gla.ac.uk