SOP for performing differential protein expression analysis in PIMO +/- regions

# Introduction

Goal is to identify biomarkers of hypoxia. We have GBMs with PIMO administered, and ~15 other IHC markers included in the IMC panel. Our goal is to identify which markers show the strongest association with hypoxia. This will be done through:

* Percent positive counts in PIMO +/- regions for all IHC markers
* Correlation coefficient of PIMO with each IHC marker

Other goals include quantifying vessel density and distance analysis. This will be done through:

* Vessel segmentation of GLUT1 or CD31. Total counts are divided by total viable area
* Distance analysis : change in percent positive and marker intensity relative to vessel distance

We’re also interested in some double positive populations (IBA1+ and ICAM1+ in PIMO +/-, CD68+ and ICAM1+ in PIMO +/-)

# Analysis methodology

## PIMO thresholding for ROI generation

* Create classes “pimo positive” and “pimo negative”
* Pixel classifier > gaussian blur of 10 and pimo threshold of 2.4. Verified w/ pathologist that this is suitable.
  + Name of classifier is 2.4thresh\_10gaussian\_pimo
  + Fill holes less than 300, exclude less than 300
  + Code:
    - Run **annotations from pixel classifier.groovy**

## Cell segmentation

* run **Multimodal StarDist Segmentation.groovy**
  + use channel 20 (Ir193)
  + no preprocessing parameters
  + 10um expansion
  + 0.4 detection threshold
  + 1-99 normalize percentiles
  + Pretrained model dsb2018 heavy augment

## Cell classification

For each of the 16 IHC markers, we identified which compartment (cell or nuclear) and measure (mean or median) to use, and set a single value threshold. We recorded this in an excel sheet and verified each threshold with a pathologist. Although thresholds are present for the vascular markers and nuclear counterstains, those likely won’t be of much use, but are still good to have.

First, we need to create classes for each of the IHC markers. Simple way to do this to go to the classes list, right click and select “populate from image channels”. Keep existing available classes.

First, we need to take those thresholds and create 16 unique single measurement object classifiers.

* Single measurement classifier
  + Detections = all
  + Channel filter = whichever IHC marker you’re setting the threshold for
  + Measurement = cell mean, cell median, nucleus mean, or nucleus median
  + Threshold = value you decided with pathologist
  + Above threshold = *markername*+
  + Below threshold = leave empty
  + Classifier name = *markername+*
* Create composite classifier using all of the above classifiers
* Add shape measurements to cells
  + addShapeMeasurements("AREA", "LENGTH", "CIRCULARITY", "SOLIDITY", "MAX\_DIAMETER", "MIN\_DIAMETER")
* Export measurements. Make sure to save any open projects
  + Export type – cells
  + Separator – comma (.csv)

## Data analysis and visualization

* C:\Users\Mark Zaidi\Documents\Python Scripts\GBM IMC percent positive\ Data Visualization.py
  + Script calculates percent of cells positive in the PIMO +/- regions
    - Region names are currently hardcoded, need to update
    - Variable storing region name is hardcoded
    - Classifications of cells not positive for any marker is also hardcoded
    - Script might break if region is absent OR 0% positive in either region
  + Violin plots of marker intensity of cells in PIMO +/- regions
    - Not all measures are mean, some are median
    - Need to include statistical testing as a p value
* Remaining data analysis
  + Correlation coefficient on density scatterplots between every marker pair
  + Other goals as outlined in the introduction (i.e. vessel density)