γH2AX Foci Counting Workflow

Author: Albert Wang
Last updated: 2025-07-17

Summary

This document describes a semi-automated MATLAB pipeline for quantifying γH2AX foci within individual nuclei. The workflow covers image import, preprocessing, segmentation, foci detection, per-cell metric computation, and results export.

γH2AX is the phosphorylated form of the histone variant H2AX, rapidly formed at sites of DNA double-strand breaks as an early cellular response to genotoxic stress. Quantifying γH2AX foci provides a sensitive and reliable indicator of DNA double-strand breaks, allowing researchers to assess both the extent of DNA damage and the cellular response to genotoxic stress, such as radiation therapy. Quantifying γH2AX foci offers a sensitive and reliable measure of DNA double-strand breaks, enabling researchers to evaluate both the extent of DNA damage and the cellular response to genotoxic stress, such as radiation therapy.

This script reads nucleus and foci TIFF images from user selection, converts each to a summed-intensity grayscale map, applies either dynamic or universal thresholding, filters out small nuclei and removes their associated foci, computes per-nucleus intensity statistics and foci counts, writes results to an Excel file, and saves an annotated segmentation JPEG with colored nucleus indices.

Example usage:

https://doi.org/10.1158/1535-7163.MCT-17-0897

Elsaid, M., Shahi De, A., Wang, A., Baiu, D., Li, C., Werner, L., . . . Otto, M. (2018). Enhanced Radiosensitivity in Solid Tumors using a Tumor-Selective Alkyl Phosphocholine Analog. Mol Cancer Ther, 17, molcanther.0897.2017.

Workflow Steps

MCF10A_CLR_1h_1 is used as example image here

1. Image Import

- Use uigetfile to select two TIFF files: first the nucleus channel, then the γH2AX (foci) channel. Read both images with imread
- Convert each RGB image into a single-channel grayscale intensity map by summing the red, green, and blue channels (I_nucleus, I_foci).
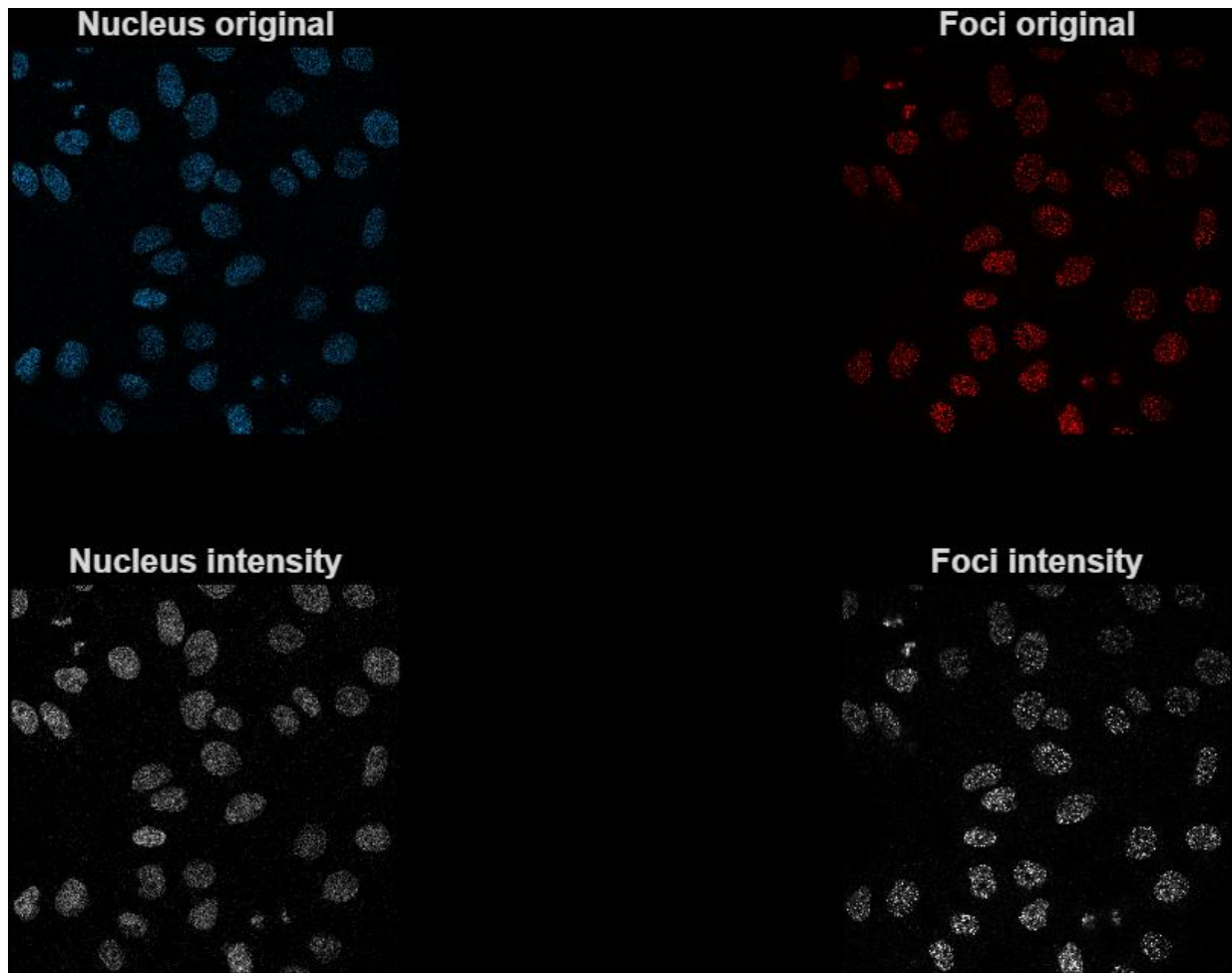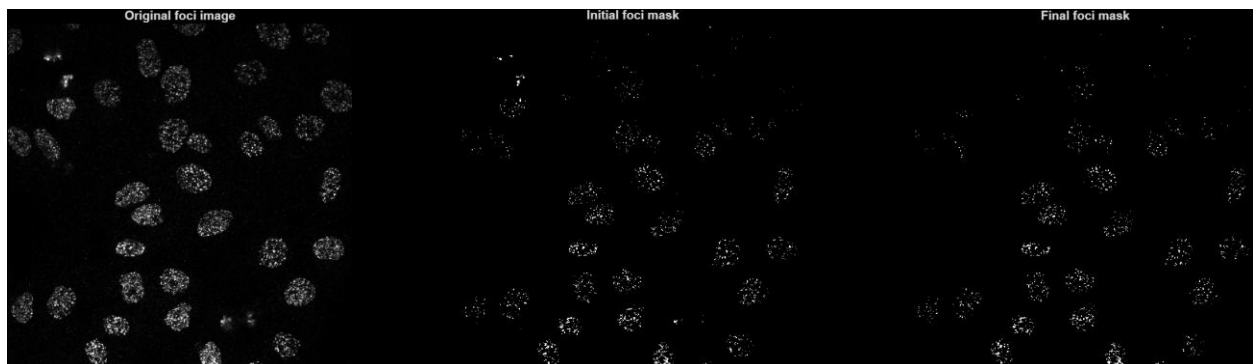

2. Thresholding and Mask Creation

- Prompt the user for:
    - A nucleus threshold value (0–255, used with imbinarize).
    - A foci threshold (0–255 intensity units).
    - Thresholding mode:
        - Dynamic (1): Determines a per-nucleus foci threshold.
        - Universal (2): Uses a single global foci threshold.

- Nucleus mask:
  - Binarize I_nucleus using imbinarize.
  - Clean the binary mask with bwareaopen and imfill.
  - Use bwconncomp to identify connected nuclei.
  - Use regionprops to extract nucleus area and centroid.
  - Discard nuclei smaller than a defined area threshold (e.g., 1,000 pixels) and record remaining indices for downstream analysis.
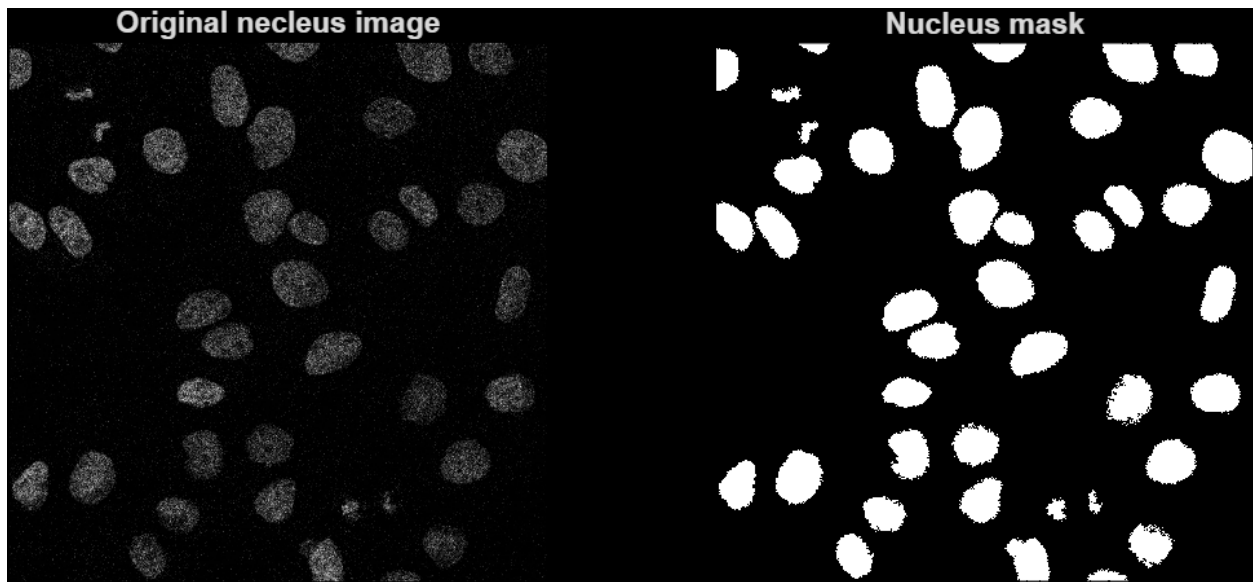
- Foci mask:
  - Create an initial foci mask by thresholding I_foci.
  - For each retained nucleus:
    - Apply the selected thresholding mode to define foci regions.
    - Remove small specks (< 2 px) with bwareaopen.
    - Aggregate per-nucleus foci masks into a global foci mask.
  - For each discarded (small) nucleus:
    - Remove surrounding foci in a square region (default: 90 px half-width) to eliminate false positives.

- Can also use MATLAB's Color Thresholder App for auto thresholding (https://www.mathworks.com/help/images/image-segmentation-using-the-color-thesholder-app.html)

3. Per-Nucleus Metric Computation

For each retained nucleus:

- Extract the grayscale intensity values from I_foci.
- Compute summary statistics:
  - Mean, median, mode, standard deviation.
- Apply per-nucleus foci thresholding and count foci:
  - Use bwconncomp and regionprops to exclude small objects.
- Store results as a row of:
  - Index, Area, MeanInt, MedianInt, ModeInt, StdInt, FociCount

4. Results Export

- Store all per-nucleus results in a MATLAB table.
- Write the table to an Excel file using writetable.
- Generate a 2×2 segmentation figure including:
  - Nucleus segmentation overlay
  - Foci segmentation overlay
  - Color-coded nucleus index map (with overlaid numbers)
  - Original foci image
- Save the figure as a high-resolution JPEG (print -djpeg -r600).
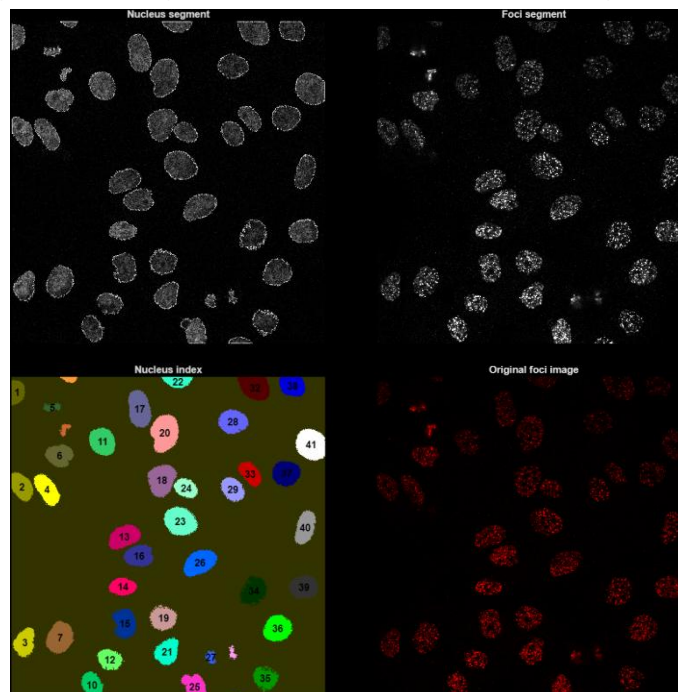
Note: Cells with their nuclei extremely close to each other (CHLA-20 in particular) would be difficult for the program to distinguish individual nucleus (see example below). Discard these data to avoid skewing the overall trend.