Topological data analysis of pattern formation in 2D multichannel microscopy images*

Daniel A. Cruz†

Current methods for quantifying multicellular patterning tend to capture the spatial properties of cell colonies at a fixed scale and typically rely on human annotation. In [1], we present a computational pipeline that utilizes topological data analysis (TDA) to generate quantitative, multiscale descriptors which capture the shape of data extracted from 2D multichannel microscopy images. Because of its broad applicability to immunofluorescence microscopy images, our pipeline is well-positioned to serve as a general-purpose tool for the quantitative study of multicellular pattern formation.

In the figure, we provide an overview of how our pipeline extracts topological descriptors from microscopy images of multicellular colonies. Figure 1a - An input microscopy image of cells with a nucleus signal in blue and two other signals: red (R) and green (G) (scale bar, $440\mu m$). This image is processed through the segmentation module of our pipeline to identify cell locations and associate a signal intensity to each cell. Figure 1b - A discretized version of the microscopy image in which cells are represented as points in the Euclidean plane and categorized into one of four cell types based on signal intensities by the cell type identification module. For the upper right patch, points in each cell type are shown. In general, there will be 2^n cell types identified for n signals based on signal intensity. Figure 1c - The points of a single cell type (R^+G^-) are connected using an increasing sequence of structures called the Delaunay filtration; the structures (simplicial complexes) at one radius are shown (left). The resulting persistence diagram (center) is a topological summary that encodes the radii at which holes appear and disappear. The corresponding persistence landscape (right) is the primary output of our pipeline and can be combined with statistical and machine learning tools. For simplicity, the $(\text{birth} + \text{death})/2$ axis is labeled 'Radius' and the $(\text{death} - \text{birth})/2$ axis is labeled 'Persistence'.

References

[1] Hartsock, I., Park, E., Toppen, J. Bubenik, P., Dimitrova, E.S., Kemp, M.L., and Cruz, D.A. Topological data analysis of pattern formation of human induced pluripotent stem cell colonies. *Scientific Reports* **15**, 11544 (2025).

*Associated publication can be accessed at `https://doi.org/10.1038/s41598-025-90592-1`

†Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, University of Florida

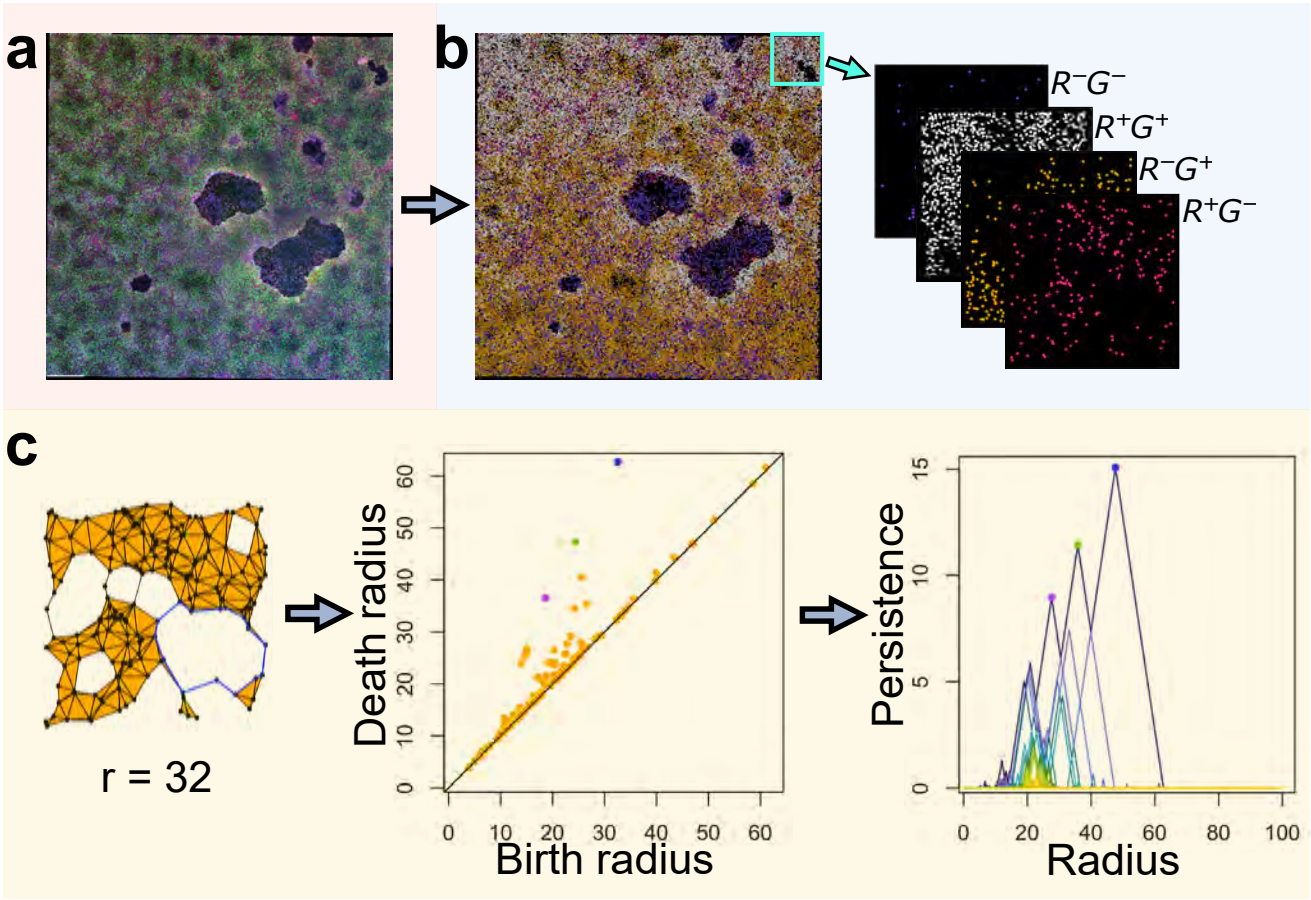

Figure 1: Overview of our computational pipeline [1].