Morphological Object Localization: A Novel Image Analysis Pipeline for Quantitative Spatial Localization of Biomolecule Signal from Fluorescence Microscopy Data

Andrew Soltisz[1], Rengasayee Veeraraghavan[2], Vladimir Bogdanov[3] and Sandor Gyorke[1]

[1]The Ohio State University, Columbus, Ohio, United States, [2]The Ohio State University, Ohio, United States, [3]The Ohio State University, United States

The spatial distribution of biomolecules (BMs) within cells and tissues is often a significant determinant of biological / physiological function. Thus, its assessment is not only a ubiquitous feature in the life sciences but also a vital component in the clinical diagnosis and treatment of many pathologies. This typically entails microscopic imaging of fluorescently labeled biomolecules within biological specimens and analysis and interpretation of the resulting images. This is accomplished by methods ranging from qualitative visual assessment of representative images to colocalization analysis, which quantifies the superposition of immunosignals corresponding to co-labeled BMs. These commonly employed approaches have several key limitations. The selection and visual assessment of representative images is subjective and highly susceptible to bias. And, although quantitative, conventional colocalization analysis is susceptible to variations between fluorophores, and oversimplifies complex spatial distributions of BMs into a single numerical index of signal superposition.[1] It provides no information about non-superimposed signals, and thus, lacks the sensitivity and selectivity to capture intra- and inter-individual variability or discern subtle forms of biological remodeling, as occurs in the early stages of disease. Here, we present a novel, high-throughput image analysis pipeline, called Morphological Object Localization (MOL), for comprehensive, quantitative spatial localization of BM signals relative to each other as well as structural landmarks from fluorescence microscopy data. This tool offers a quick and user-friendly alternative to current approaches with unprecedented capabilities for quantitative assessment of cell / tissue structure.

Our pipeline was built around and validated using 3D, high resolution confocal microscopy images of isolated cardiomyocytes with messenger RNA (mRNA) for the $Ca_V1.2$ (Cacna1c) and $Na_V1.5$ (Scn5a) electrogenic membrane proteins fluorescently labeled using in-situ hybridization (RNAScope). The pipeline's <u>initial stage</u> performs segmentation of objects and structures of interest. Here, we used simple, voxel intensity-based thresholding of the nuclear stain (Hoechst, DAPI) to generate cellular landmark masks of the cell's periphery, nuclei, and intercalated discs (IDs). Fluorescent signals for mRNA were similarly segmented to identify bulk BM signal-positive voxels; these masks could be further processed to identify individual BM punctae or clusters if localization of organized or aggregate structures was desired. This segmentation technique can be substituted for any other technique so long as binary masks of the underlying structures and objects are the final output. The pipeline's <u>intermediate stage</u> performs localization of tagged BMs relative to cellular landmarks by identifying the set of shortest distances DL between signal-positive voxels of the BM's mask and the surface or edge of the landmark's mask. This set of distances is defined as the intersection $\cap$ of the BM mask M_{BM} and the landmark mask's distance transformation DT_{LM}, formally expressed as the set of distances $D_L = \{M_{BM} \cap DT_{LM}\}$. A richer version of colocalization analysis can also be performed where BMs are localized relative to each other by finding the intersection of one BM mask M_{BM1} with the other's distance transformation DT_{BM2}, formally expressed as the set of distances $D_{CL} = \{M_{BM1} \cap DT_{BM2}\}$. Additionally, instances of BM superposition themselves can be localized relative to a cellular landmark as the zero-distance colocalization mask's $M_S = M_{BM} \cap \{d \in D_{CL} / d = 0\}$ intersection with a landmark's DT, formally expressed as the set of distances $D_{SL} = \{M_S \cap DT_{LM}\}$. The <u>final stage</u> of the pipeline visualizes distance sets as a set of plots which illustrate different qualities of a BM's spatial distribution (**figure 1**): The simplest of these is the probability density

function (PDF) where the fraction of total BM-positive voxels are plotted as a function of distance away from a landmark which can be used to assess BM distribution. This function can be cumulatively summed to generate the cumulative distribution function (CDF), which is amenable to statistical analysis with either simple comparisons of central tendency (Student's t-test, Wilcoxon's test) or more comprehensively with distribution-wide comparisons (Kolmogorov-Smirnov test) to assess landmark-relative BM aggregation. Finally, the set of distances can be plotted as the BM's volume-normalized concentration as a function of normalized distance away from the landmark. Assessing the deviation from 0 of this plot's linear regression slope can be used to assess the uniformity of a BM's distribution and could provide clues regarding trafficking mechanisms underlying the BM's observed distribution. Additionally, more general measurements of BM signal, such as volume and density, can easily be calculated to provide further context and aid in output interpretation (**figure 2**).

Our pipeline offers significant improvements over current biological image analysis methodologies due to its quantitative nature, richness and visual palatability of its outputs, modularity, and high throughput potential. The pipeline's outputs may also provide a direct source of parameter values for computational models which describe the nature of the underlying biology to further probe the experimental design space.
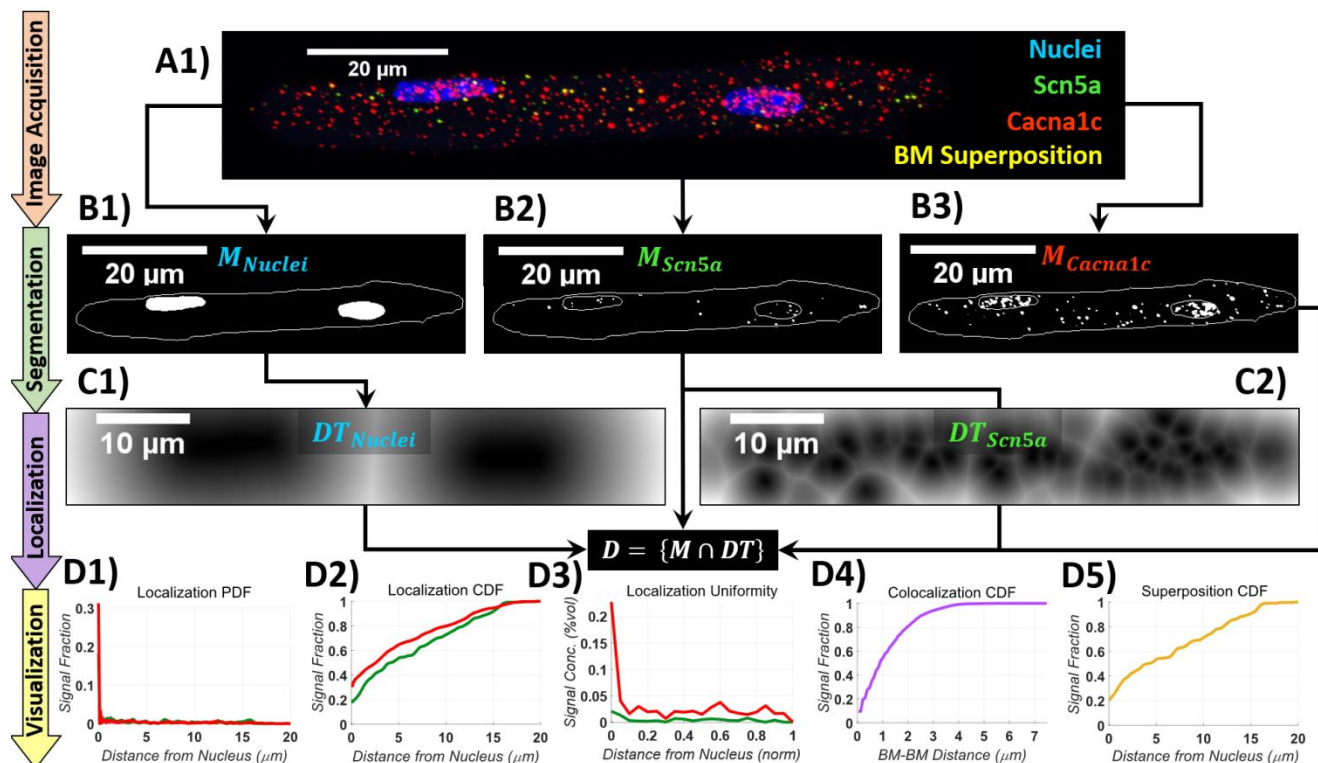

Figure 1. Pipeline Workflow Diagram: (A1) The pipeline receives 2- or 3-D microscopy images containing any number of fluorescence channels as input, then uses a segmentation method of the user's choice to produce masks which identify pixels composing the landmarks (B1) and biomolecules (B2-3) of interest. Next, the distance transformation DT of each mask is generated (C1-2) whose intersection with a biomolecule mask MBM produces the set of shortest distances D between the landmark and biomolecules. Finally, this distance set is visualized using an array of plots which illustrate different features of a biomolecule's spatial distribution (D1-3), colocalization with other biomolecules (D4), and spatial distribution of the superposition of biomolecule pairs (D5).
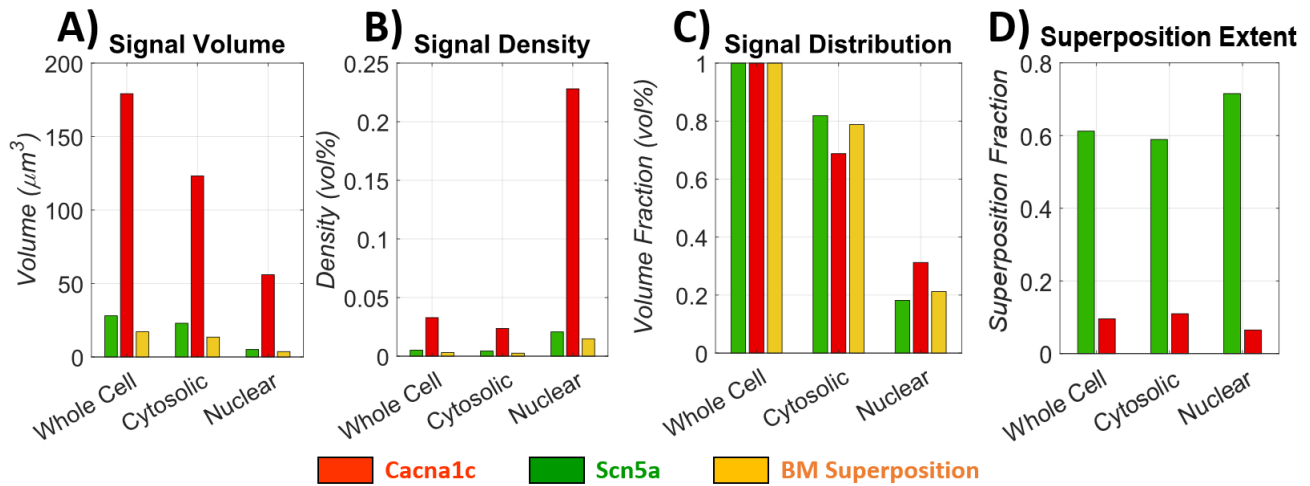
Figure 2. Accessory Plots: Additional plots which visualize more general characteristics of biomolecules and their spatial distributions can be used to provide context to and aid in interpretation of localization plots. Here, a sample of such parameters are plotted for different cellular regions, depicting the raw biomolecule or superposition signal volume (A), the volume fraction of each region which a signal occupies (B), the fraction of signal volume localized to each region (C), and the fraction of biomolecule signal in superposition with the other biomolecule species (D).

References
Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).