Physics of Complex Systems Master Thesis

Quantitative pattern formation in mammalian stem cell aggregates

Supervisors:
Thomas Gregor
Carla Bosia

Candidate:
Mattia Cascone

June 2025

Abstract

Understanding how biological systems spontaneously generate spatial patterns is fundamental to developmental biology. While quantitative frameworks for positional information have been successful in simple model organisms like Drosophila, extending these approaches to mammalian systems has been limited by technical challenges in imaging dense tissues and extracting single-cell information. This thesis establishes a comprehensive methodological framework for quantifying gene expression patterns in gastruloids, self-organizing stem cell aggregates that recapitulate key aspects of mammalian embryonic development. An optimized optical clearing protocol was developed that enables complete 3D confocal immunofluorescence imaging of these dense structures while preserving morphology and fluorescence. Combined with state-of-the-art deep learning segmentation, this pipeline successfully extracts single-cell information throughout entire gastruloid volumes. Spatial analysis revealed characteristic anterior-posterior expression patterns of key transcription factors, while single-cell resolution uncovered complex expression heterogeneity that would not have been detected with previous approaches. These technical advances provide the foundation for quantitative studies of self-organization in mammalian developmental models, enabling future investigations into how complex spatial patterns emerge spontaneously from initially symmetric conditions.

Contents

1 Introduction

Understanding how biological systems spontaneously generate spatial patterns is a fundamental challenge in developmental biology. This introduction examines the theoretical framework of positional information that has guided quantitative studies of pattern formation, explores how emerging model systems can extend these insights to mammalian development, and outlines the scope of the present work.

1.1 Quantifying positional information in self-organizing systems

Biological self-organization is the term utilized to describe a whole host of phenomena in which biological systems form structured patterns from an initially disordered set of components. These processes arise as the result of natural selection and typically display a complex balance of diverse biological mechanisms, but at the same time show a remarkable capacity for resilience against disturbances. This robustness allows for the same outcome (a biologically functional and evolutionarily motivated one) to be realized in the midst of the disorder of natural life.

The way in which the inherently stochastic biological mechanisms interact with each other to give reproducible outcomes is, for the physicist interested in complex interacting systems, an extremely interesting subject of study.

One of the areas in which this study can be fruitful is the field of developmental biology, concerned with the way in which the genetically identical cells of the developing embryo rearrange themselves in precise spatial patterns of epigenetically different tissues. To achieve such patterning, cells must somehow know where they are within the developing embryo. This requires spatial information to be encoded in the cellular environment (through biochemical signals, mechanical forces, or other cues) and cells to be able to read and interpret these signals. This notion, formalized as "positional information" by Lewis Wolpert, proposes that cells measure local environmental signals to infer their position and then use this information to determine their developmental fate. The positional information framework has pushed developmental biology toward more quantitative approaches, as it allows researchers to measure precisely how much information is encoded in environmental cues and how reliably cells can decode it to guide pattern formation (Tkačik and Gregor, 2021).

These quantitative methods have been very fruitful in simple model systems, particularly in the study of the Drosophila (fruit fly) embryo, where the expression patterns of gap genes provide a clear encoding of positional information along the anterior-posterior axis. Measuring these gene expression domains, and comparing them to the theoretical constraints on the required information, has allowed for precise and mechanism-independent predictions on the mode of interaction between cells and on the flow of information along the process (McGough et al., 2024).

While these studies in Drosophila have validated the positional information framework, they have focused on a system in which spatial information is largely pre-determined by maternal inputs. Mammalian development presents an interesting counterpoint: spatial patterns emerge dynamically through self-organizing processes rather than being imposed by pre-existing morphogen gradients. Successfully extending the theory to such systems would not only test the generality of the principles established in simpler models, but could also reveal new mechanisms by which biological systems spontaneously break symmetry and encode spatial information.

1.2 Gastruloids as a model for self-organization

A fundamental symmetry-breaking process in animal development is the transition from an undifferentiated mass of stem cells to a structured organism defined by distinct tissue layers and axes. This event, known as *gastrulation*, is a highly conserved process across species, including both insects such as Drosophila and mammals. Gastrulation typically involves a simultaneous burst of cell migration, differentiation, and global morphogenetic movements that structure the embryo into three layers of cell populations (the ectodermal, endodermal and mesodermal tissues) organized according to the animal's future body plan.

The context in which gastrulation occurs, however, differs drastically between these systems. In Drosophila, gastrulation follows the establishment of maternal morphogen gradients that pre-pattern the embryo, making the differentiation process essentially a readout of previously encoded symmetry-breaking positional information. In mammals, in contrast, gastrulation represents the first observable instance of symmetry breaking and spatial patterning. The mammalian embryo before this event presents itself as a highly symmetric hollow sphere of largely identical cells, with only subtle and often transient molecular inhomogeneities. It is only with gastrulation that positional information can undeniably be observed to arise through the emergence of structured arrangements of tissues and germ layers.

This makes the quantification of spatial patterns during mammalian gastrulation a powerful experimental setting for understanding how symmetry breaking and positional information can emerge *de novo* in self-organizing systems. Additionally, such measurements may help guide biological investigation into the mechanisms that produce these patterns, revealing the nature of these mechanisms and whether they operate through global gradients or local cell-cell interactions. Gaining a quantitative understanding of these questions requires high-resolution measurements of spatial patterning as it emerges during gastrulation.

This quantitative study faces significant practical obstacles. Mammalian embryos develop within the maternal reproductive tract, making direct observation of the gastrulation process impossible without embryo extraction and culture. Once extracted, embryos are extremely fragile and difficult to maintain in culture conditions that preserve normal development. The small numbers of embryos that can be

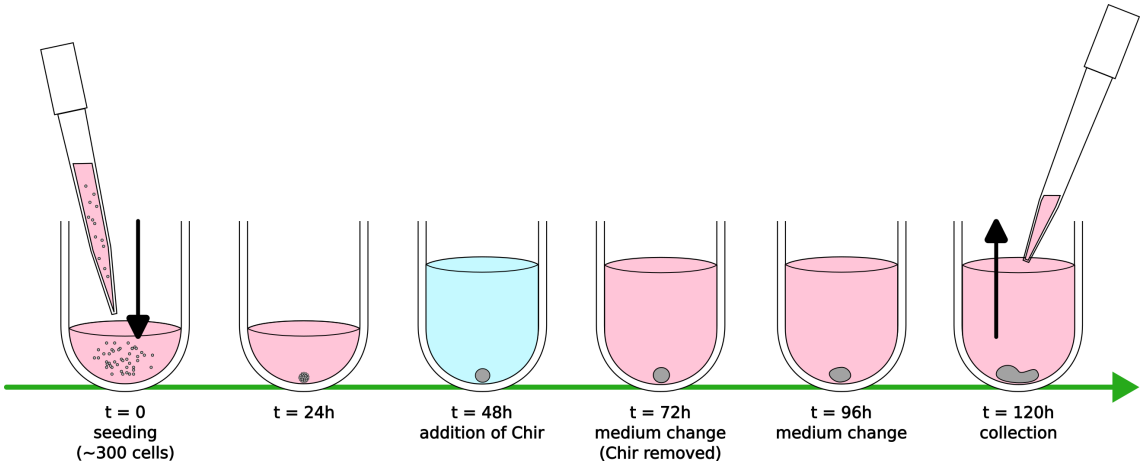


Figure 1: Schematic of the 6-day gastruloid formation protocol. Mouse embryonic stem cells are seeded at t=0 at a concentration allowing for the desired number of cells per well (300 in this study), and spontaneously aggregate. At t=48h, Chir (Wnt pathway activator) is added to stimulate differentiation and patterning. The medium is changed at t=72h to remove Chir, allowing continued development. Gastruloids are collected at t=120h after characteristic elongation has occurred.

obtained from each pregnancy, the time-intensive nature of embryo collection and the associated costs, make it challenging to generate the large sample sizes required for statistical analysis. These practical limitations are compounded by ethical considerations surrounding the use of mammalian embryos in research, which further restrict experimental scope and scale (Shahbazi et al., 2019).

Recent advances in stem cell biology and tissue engineering have opened new possibilities for addressing these challenges through the development of three-dimensional cellular models known as *organoids*. Organoids are synthetic biological systems that recreate key features of specific tissues or organs by providing cells with the appropriate biochemical and physical environment to self-organize into structured, functional units. These systems have revolutionized biological research by enabling controlled, reproducible studies of complex biological processes that would otherwise be difficult to access (Zhao et al., 2022).

For studying mammalian gastrulation specifically, a particularly valuable organoid system has emerged: *gastruloids*. Gastruloids are embryonic stem cell-based aggregates that can be stimulated to undergo a gastrulation-like process, developing the major tissue types found in the posterior part of the embryo and organizing them according to the three major body axes, closely mimicking the spatiotemporal sequence of events observed during posterior axis formation in the mammalian embryo (Beccari et al., 2018). They can be generated reproducibly in large numbers under precisely controlled conditions, enabling the statistical analysis necessary for quantitative studies.

Their production (here focusing on mouse gastruloids, though other species including humans have been used) involves the aggregation of mouse embryonic stem

cells (mESCs), which are originally derived from the inner cell mass of a mouse embryo and maintained under culture conditions that preserve their pluripotency. When seeded into U-bottom ultra-low-adhesion wells, these cells spontaneously aggregate and self-organize into a gastruloid. The size of the resulting structure can be precisely controlled by adjusting the number of mESCs seeded per well, typically ranging from 100 to 600 cells. This straightforward protocol allows for controlled and reproducible generation of high numbers of gastruloids (figure 1), offering a powerful system to model early embryonic development in vitro.

In order to stimulate differentiation and patterning, a specific signaling molecule called Chir, that activates the Wnt pathway, is added at 48 hours after seeding and removed 24 hours later. In in-vivo mouse embryos, Wnt signaling is active in one region (what will become the tail end of the embryo) and inhibited in the region that will become the head. In gastruloids, instead, the whole organism is exposed to the same uniform Wnt activation impulse, after which symmetry is spontaneously broken by the beginning of the elongation of the gastruloid, accompanied by the differentiation of cells into distinct germ layer identities (Beccari et al., 2018). This elongation replicates the events happening in the posterior, Wnt-stimulated region of the embryo. The elongation of the gastruloids begins to be visible at the 96 hours-after-seeding timepoint, then accelerates to become quite marked at the 120 hour mark (figure 2).

It must be noted that while gastruloids recapitulate many key features of mammalian gastrulation, they represent a simplified version of the process. Like mammalian embryos, gastruloids begin from a symmetric initial state of pluripotent cells. However, they develop in isolation from the complex signaling environment that emerges after embryo implantation, lacking inputs from maternal tissues and extraembryonic structures that could potentially provide additional patterning cues.

Despite this simplified context, gastruloids develop structured patterns with remarkable precision and reproducibility (Merle et al., 2024). They generate gene expression patterns comparable to those found in the posterior region of mammalian embryos, complete with appropriate spatial organization of the three germ layers. This reproducible self-organization from an initially symmetric state, achieved through what must be predominantly stochastic symmetry-breaking mechanisms, makes gastruloids an ideal system for studying how biological systems can spontaneously generate and encode positional information.

While the work in Merle et al. (2024) demonstrated the reproducibility of gastruloid patterning, their analytical approach did not fully exploit the three-dimensional nature of gastruloids. Their analysis relied on maximum intensity projections to compress fluorescent signals onto a single plane, potentially missing critical spatial information distributed throughout the three-dimensional structure. To fully understand how positional information emerges and flows through these self-organizing systems, we need methodologies that can capture and analyze the complete three-dimensional cellular landscape at single-cell resolution.

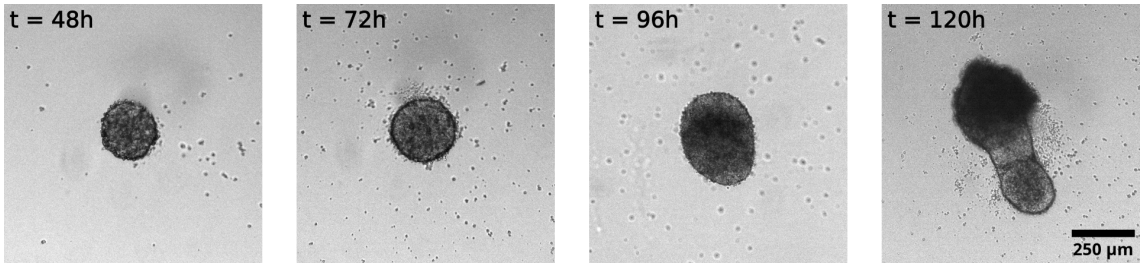


Figure 2: Temporal development of a gastruloid showing spontaneous elongation. Bright-field images taken at 48, 72, 96, and 120 hours after seeding, captured with a 4× objective. The initially spherical aggregate undergoes spontaneous symmetry breaking and elongates along a single axis by 120h, recapitulating key aspects of embryonic body axis formation. Scale bars: 250 μm (shown on t=120h image; scale identical for all timepoints).

1.3 Overview

This work represents an attempt to push forward the development of a methodological framework for the characterization of positional information in mammalian development, using gastruloids as the model system of choice. The goal is to establish the technical capabilities necessary for precisely evaluating the structures formed by these self-organizing systems by addressing the challenges of complete 3D imaging and automated cell identification and phenotyping.

The thesis is structured to reflect the experimental nature of this work, where the development of robust methods constitutes the primary result. Accordingly, the methodological development represents the largest portion of this thesis, as is appropriate for a work focused on establishing new technical capabilities rather than applying existing methods to answer specific biological questions. The Results and Discussion section demonstrates how this methodology enables discovery of complex spatial organization that would be missed by conventional approaches, including the observation of bimodal expression distributions that suggest discrete cell populations rather than continuous gradients. The technical advantages of full 3D analysis, including preservation of native cellular distributions and extraction of true anatomical midlines, represent significant improvements over previous maximum projection approaches.

The Conclusions section considers the broader implications of these methodological advances, discussing how single-cell resolution throughout entire 3D samples now enables quantitative studies of how spatial patterns emerge from initially symmetric conditions. This framework expands the analysis from simple average expression levels to full probabilistic distributions of heterogeneous cell states, opening new avenues for testing quantitative theories about biological self-organization.

2 Methods

All gastruloids in this work were generated following established protocols for mouse gastruloid formation, seeding an $N_0 = 300$ initial number of cells and collecting them at 120 hours after seeding. These conditions were maintained identical to those reported in Merle et al. (2024) to ensure direct comparability with published work. Following collection, gastruloids were fixed, permeabilized and labeled with fluorescent antibodies for immunofluorescent imaging on a confocal microscope.

2.1 In-depth optical clearing for full-gastruloid 3D imaging

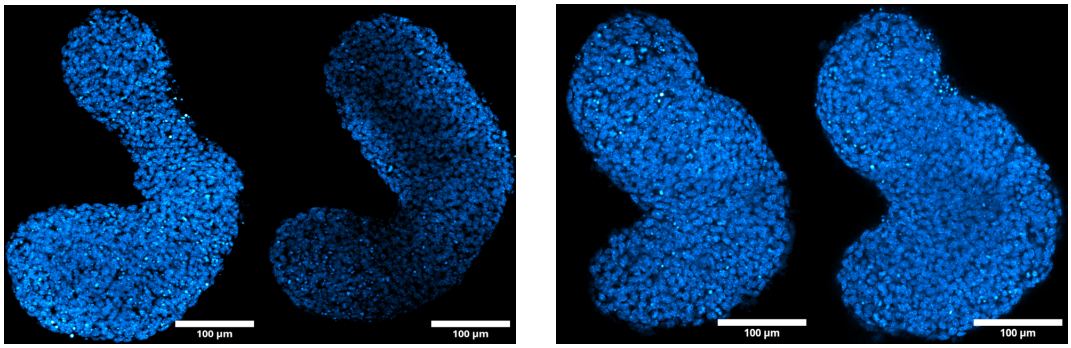
Fluorescence microscopy is a fundamental technique of biological research that allows for the collection of information about specific structures and molecules in a sample by labeling with fluorescent probes. These fluorescent probes are excited by lasers at the appropriate wavelength, and the emitted light is collected by sensors to reconstruct images.

In this context, modern microscopy allows for the three-dimensional imaging of biological samples by leveraging the partial transparency of tissues. The excitation light can be focused to excite specific points at different depths in the sample; the light emitted by each of those points is then collected and associated to a voxel in the three-dimensional digital representation of the sample.

Although the emitted photons may be coming from fluorescent probes anywhere in the excitation path, sectioning in the axial direction (z) can be achieved through the use of a pinhole in the collection path so that only the photons from the point of interest at a given time, the one in focus at that instant, may be collected. This is the principle of *laser scanning confocal microscopy,* by which entire volumes may be imaged one voxel at a time while the sample is displaced by precise piezoelectric actuators (Pawley, 2006).

The imaging of points deep in the sample, however, becomes quickly difficult due to the scattering and absorption of light as it penetrates the heterogeneous mix of molecules in the tissue (Richardson et al., 2021). Both the incoming excitation light and the outgoing emitted light suffer from this limitation, so that the signal-to-noise ratio in slices of tissue deeper in the sample becomes too low to yield useful data.

Gastruloids are characterized by very densely packed cells, which creates size-dependent imaging challenges. While smaller gastruloids (those with lower initial seeding numbers N_0) can be imaged completely, larger ones (including the $N_0 = 300$ gastruloids used in this study) exhibit severe decay of light intensity and detail in deeper sections away from the surface boundaries, precluding complete imaging. This could partially be compensated for by compressing the gastruloids between closely spaced glass slides, but this approach would introduce morphological modifications that would be undesirable given the goal of analyzing naturally arising structures.



(a) (b)

Figure 3: Visual comparison of clearing quality at different depths within gastruloids. **(a)** Sub-optimally cleared gastruloid showing significant signal degradation between surface (left) and deep (right) sections. **(b)** Gastruloid treated with an optimized clearing protocol showing maintained image quality even in deeper sections. For both samples, the left and right images are taken from z-planes spaced $\Delta z = 80\,\mu$m apart.

In contrast, optical clearing addresses these imaging limitations while preserving natural tissue architecture by chemically modifying the refractive index of tissues to reduce light scattering. The difference in imaging quality between cleared and uncleared samples is illustrated in figure 3, which demonstrates the necessity of systematic protocol optimization for achieving complete three-dimensional imaging of gastruloids.

2.1.1 Optimization of the clearing protocol

In order to minimize light loss, the refractive index of all the sections along the optical path must be as uniform as possible. Given the constraint that a glass coverslip, whose refractive index is close to 1.52, must be present to hold the sample in place, the entire optical system must be optimized around this value. This optimization involves three key components: the use of an objective optimized for the desired refractive index, the application of a drop of compound between the objective and the coverslip to fill in the gap that would otherwise be occupied by air, and the suspension of the sample in a solution of similar refractive index, commonly referred to as a mounting medium.

Mounting media are solutions of various composition specifically formulated not only to achieve optimal refractive index matching, but also to remove unwanted light-scattering molecules from the sample, to preserve samples over longer periods of time, and to increase the photostability of fluorescent probes to avoid, or at least delay, the onset of photobleaching (Ravikumar et al., 2014).

Due to all these effects, and to the varying degrees of efficacy with which different solutions can penetrate the sample to clear it effectively, it was not obvious that the mounting medium with the highest refractive index would perform the best. Therefore, four different mounting media formulations were compared:

- VectaShield (Vector Laboratories), a widely used commercial mounting medium with a refractive index of 1.45
- A laboratory-prepared solution of water, glycerol and fructose, prepared according to Dekkers et al. (2019), with a measured refractive index of 1.471
- FOCM (Fast Optical Clearing Method), prepared according to Zhu et al. (2019), containing DMSO, urea, glycerol and sorbitol, with a measured refractive index of 1.496
- RapiClear (SunJin Lab), a commercial clearing solution specifically formulated to achieve a refractive index of 1.52

The RapiClear solution was observed to display significant batch-to-batch variability in formulation and efficacy, with initial experiments revealing that an older batch showed suboptimal clearing performance compared to a fresh batch under identical conditions.

Moreover, it was observed that mounting medium performance in general was critically dependent on the degree of tissue permeabilization. Permeabilization is the step of the immunostaining protocol in which a detergent (in our case, Triton X-100) is applied to the sample to solubilize the lipids of the cell membranes, creating pores that allow better penetration of both the fluorescently labeled antibodies and the mounting medium. While established protocols recommended Triton X-100 concentrations between 0.03% and 0.2%, I found that using 0.5% Triton during the initial permeabilization step, followed by 0.2% throughout subsequent steps, significantly enhanced clearing medium penetration and imaging quality without compromising tissue integrity.

To systematically evaluate all these variables (mounting medium formulation, batch effects, permeabilization degree), my comparative analysis comprised six experimental conditions in total:

- Fresh RapiClear batch with enhanced permeabilization (0.5% Triton X-100)
- Fresh RapiClear with standard permeabilization (0.1% Triton X-100)
- Original RapiClear batch with enhanced permeabilization
- VectaShield with enhanced permeabilization
- Glycerol-fructose solution with enhanced permeabilization
- FOCM with enhanced permeabilization

Each of these conditions was evaluated using the quantitative metrics described below, with $n = 26$ gastruloids tested for the fresh RapiClear with enhanced permeabilization condition, and $n = 5$ gastruloids for each of the other conditions.

2.1.2 Microscopy setup and imaging parameters

For the objective lens, I chose a 25x oil immersion objective rather than the 40x objective previously used in the laboratory for this type of sample. Oil immersion objectives are designed to work with compounds that have a refractive index very close to that of the glass coverslip, and thus provide the best optical conditions. The choice of 25x over 40x magnification was motivated by the significantly longer working distance that the lower magnification provides (570 μm as compared to the 40x objective's 130 μm), which allowed me to use thicker spacers between the glass slides thus avoiding compression of the samples. The modest reduction in numerical aperture (0.8 as compared to the 40x objective's 1.4) was acceptable given sufficient detail for single-cell analysis could still be achieved. All final imaging was performed at a voxel size of $0.404 \times 0.404 \times 0.800\,\mu\text{m}^3$.

Image acquisition employed sequential four-channel scanning to minimize spectral bleed-through between fluorophores. The channels were configured according to the following fluorescent probes: DAPI nuclear staining, AlexaFluor 488, AlexaFluor 546, and AlexaFluor 647. Laser powers were optimized individually for each channel to maximize signal-to-noise ratio while avoiding photobleaching, typically ranging from 0.5-2% of maximum laser output.

2.1.3 Quantitative evaluation of clearing quality

For the quantitative analysis of clearing quality, I focused specifically on the DAPI channel of the fluorescence images. DAPI (4',6-diamidino-2-phenylindole) is a fluorescent stain that binds specifically to DNA, thereby labeling the nucleus of every cell in the sample with blue fluorescence when excited with UV light ($\lambda = 405\,\text{nm}$). This compound is added to make all cells visible, regardless of whether they express the specific genes being labeled with fluorescent antibodies.

The fact that it binds everywhere in the sample makes DAPI an ideal reference for assessing image quality throughout the three-dimensional structure of the gastruloids. Moreover, DAPI fluorescence is particularly sensitive to clearing inefficiences, as its short excitation wavelength is more prone to scattering than the longer wavelengths used for other fluorescent markers (Jacques, 2013). Consequently, if the DAPI channel demonstrates good penetration and image quality at depth, it can be assumed that the other fluorescent channels (using green, red and far-red wavelengths) will perform even better.

The ideal clearing protocol should achieve three goals: maintain high image resolution, preserve fluorescent signal intensity, and retain sufficient contrast to distinguish individual cells all throughout the sample depth. Therefore, I employed three complementary metrics to objectively assess clearing quality across different protocols:

1. Fourier Ring Correlation Quality Estimate (FRC-QE), developed by Preusser et al. (2021) specifically for evaluating clearing protocols in 3D organoid imag-

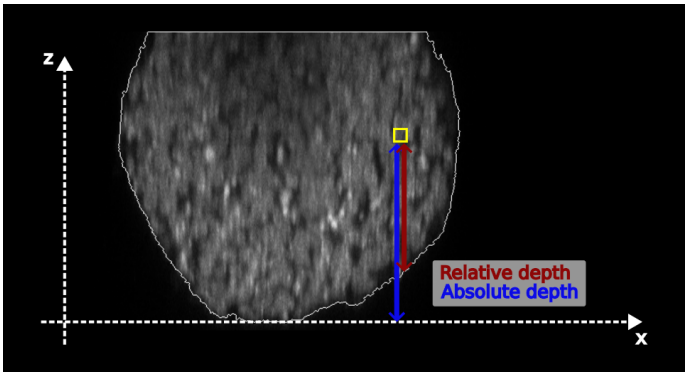


Figure 4: Illustration of absolute versus relative depth classification for clearing analysis. Cross-sectional view of a gastruloid (gray) with the full 3D mask boundary shown as a white outline. For the highlighted pixel (yellow square), absolute depth (blue arrow) is measured from the minimum z-position of the entire gastruloid mask, while relative depth (red arrow) is measured from the minimum z-position at the same xy coordinate. This distinction accounts for the variable thickness of the gastruloid: pixels at the same absolute depth may experience different amounts of light attenuation depending on their local tissue environment.

ing. This metric quantifies image quality by measuring the correlation between successive z-slices, based on the principle that well-cleared samples should show high consistency between adjacent optical sections. In poorly cleared samples, scattering and aberrations introduce random noise that reduces this correlation, particularly at greater depths. The FRC-QE values have arbitrary units and therefore no direct interpretation when in isolation, but the metric enables direct comparison between different clearing conditions when imaging parameters (particularly voxel size and z-step) are held constant. The metric was computed using the Fiji plugin provided by the original authors.

2. Depth-dependent average intensity analysis. This approach aimed to quantify the intuitive observation that better-cleared samples should maintain brighter fluorescence at greater depths. To account for the irregular shape of gastruloids, I developed a masking algorithm that excluded background pixels and analyzed intensity values within the sample based on both absolute depth (z-position from the minimum z-coordinate of the entire gastruloid mask) and relative depth (z-position from the minimum z-coordinate at each xy location, accounting for local sample thickness at that position, figure 4). This dual analysis was motivated by the observation that signal degradation typically occurs first in the thickest regions of the sample, where light must traverse the most tissue.

   The metric itself consisted of computing the mean DAPI intensity of all pixels at each depth level, after normalizing the intensity values of each individual gastruloid to its maximum intensity to account for sample-to-sample variation in absolute brightness (due to e.g. variations in staining efficiency or

fluctuations in laser power).

3. Coefficient of variation analysis. The ratio of standard deviation to mean intensity at each depth level was computed to assess contrast preservation, using the same absolute and relative depth organization as before. This metric was motivated by the hypothesis that image quality depends not just on absolute signal intensity, but on the preservation of contrast as a measure of the amount of detail in an image. In well-cleared samples, individual nuclei should remain distinguishable at depth (high contrast), whereas poor clearing could lead to a uniform haze that reduces local intensity variations and obscures detail (low contrast).

For each clearing condition, these three metrics were computed individually for each gastruloid in the dataset, then averaged across samples to produce summary statistics. Based on the evaluation of these metrics, the combination of fresh RapiClear with enhanced permeabilization (0.5% Triton X-100) demonstrated the most consistently good performance, maintaining image quality throughout the full depth of gastruloid samples. This protocol was therefore selected as the standard clearing method for all subsequent experiments.

2.2 Automated cell segmentation with machine learning models

The quantification of positional information in gastruloids requires identifying and characterizing individual cells throughout the three-dimensional structure. Computationally, this is a significant challenge: converting large three-dimensional arrays of fluorescent intensity values into discrete cellular objects with defined boundaries. In the terms of the computer vision community, this is an *instance segmentation* problem, where each cell must not only be classified as distinct from the background, but also separated from its neighbors and assigned a unique identifier (Hafiz and Bhat, 2020).

The classical methods to solve this problem are based on treating pixel intensities as elevation maps, iteratively detecting object contours through energy minimization, or clustering pixels based on similarity of features like intensity or color (Bachani et al., 2024). While these can work for simple images with well separated objects, they face significant limitations in dense biological tissues. They typically require manual parameter tuning for each dataset, struggle with overlapping or touching cells, and cannot easily incorporate complex contextual information about cell morphology. Most critically, they lack the ability to learn from examples, meaning they cannot adapt to the specific characteristics of different imaging conditions or cell types.

The introduction of deep learning has revolutionized biological image segmentation. Modern approaches differ fundamentally not just in model architecture, but

also in how they represent the segmentation problem: some methods predict segmentation masks directly, while others generate intermediate representations that are subsequently processed to identify individual objects. For example, StarDist (Schmidt et al., 2018) takes the latter approach by predicting star-convex polygons, learning to output radial distances from object centers to their boundaries at fixed angular intervals, making it particularly effective for roughly spherical objects like nuclei.

2.2.1 Cellpose-SAM architecture and implementation

For this work, I mainly employed Cellpose (Stringer et al., 2021), which takes a distinctive approach by decomposing segmentation into two complementary tasks. Rather than predicting masks directly, Cellpose trains a neural network to output two intermediate representations that are then combined through classical post-processing to identify individual cells.

The first output is a *cell probability* map that assigns to each pixel the likelihood of it belonging to a cell rather than background. The second is a *gradient field* that encodes directional information pointing from cell boundaries toward cell centers. To generate the final segmentation, the algorithm follows these gradients from each high-probability pixel until a singular point is reached. All pixels whose gradient paths converge to the same singular point are assigned to the same cell instance. In this way, the complex instance segmentation problem is decomposed into two simpler tasks that neural networks can learn more effectively. During training, manually annotated cell masks are converted into the corresponding probability maps (straightforward binary conversion) and gradient fields (generated by simulating heat diffusion from each cell centroid to its boundaries). The network then learns to reconstruct these target representations from raw microscopy images.

This representation has been demonstrated to enable well-trained networks with superior performance across diverse real-world segmentation tasks (Stringer et al., 2021). Moreover, it generalizes quite naturally to three-dimensional image segmentation, without requiring a separate architecture. In fact, the same neural network is used to predict two-dimensional gradients in all three orthogonal directions by slicing the volume along each axis. These orthogonal gradient predictions are then combined to construct 3D gradient fields, enabling the same gradient-following algorithm to operate in three dimensions and identify cells throughout the entire volume.

Previous releases of Cellpose mainly employed a U-Net architecture with additional skip connections. The models could be trained on datasets of a given type of image (e.g. fluorescent microscopy, brightfield microscopy, nucleus-only segmentation) to produce a variety of pre-trained models that the users could choose from depending on their intended application. However, with the recent release of version 4 of Cellpose, also called Cellpose-SAM, the whole architecture has been simplified to only providing a single, extremely well-generalizing model for all applications (Pachitariu et al., 2025).

Cellpose-SAM fully replaces the U-Net architecture with the image encoder from Meta's Segment Anything Model (SAM). SAM is a transformer-based vision foundation model trained on over one billion segmentation masks across diverse image domains, designed to segment any object in an image given a user prompt in the form of a bounding box, a set of points, a rough outline, or a text prompt (Kirillov et al., 2023). The input image and the prompt are embedded into high-dimensional vectors by two separate embedding models; the embeddings are then combined and interpreted by a fast, efficient *lightweight mask decoder* which predicts the segmentation masks directly.

Cellpose-SAM only leverages the image embedding component of the SAM architecture, utilizing its pretrained weights that have learned general visual features across diverse image domains. This image encoder is then retrained to predict the cell probabilities and vector gradients required by the Cellpose framework. This hybrid approach yields a system that substantially outperforms the previous U-Net-based models and can generalize to diverse imaging datasets without any specific fine-tuning.

For this application, this new framework offered two critical advantages over previous segmentation models. First, its native adaptability to diverse cell diameters allowed me to work at my chosen imaging resolution ($0.404 \times 0.404 \times 0.8\,\mu\text{m}^3$ voxel size) without requiring oversampling, which would have produced prohibitively large image files that were computationally intractable. Second, unlike previous models that typically accepted only a single channel (either nuclear staining or membrane staining), Cellpose-SAM can effectively utilize information from up to three fluorescent channels simultaneously, providing much richer context for accurate cell boundary determination.

2.2.2 Experimental approach to gastruloid segmentation

To evaluate segmentation performance and optimize parameters, I created a ground-truth dataset by manually annotating two high-resolution volumes ($0.202 \times 0.202 \times 0.800\,\mu\text{m}^3$ voxel size) using the SAMJ plugin in Fiji. This plugin employs the same Segment Anything Model that forms the backbone of Cellpose-SAM, but in its original interactive form, enabling efficient manual annotation through point-and-click interfaces. Nine orthogonal slices (three per axis) were annotated from each volume to capture the full three-dimensional complexity of nuclear morphologies. One volume served as a training set for parameter optimization, while the other functioned as an independent test set. The final test set comprised 1703 manually segmented masks, providing a substantial ground truth dataset for validation.

I compared the results from two segmentation frameworks: Stardist-Tapenade (Gros et al., 2024), a StarDist model specifically fine-tuned on organoid data, to benchmark against existing literature and assess the transferability of published models to our clearing protocol and imaging conditions; and Cellpose-SAM, which I selected for further optimization. While this model captured essentially all cells,

it exhibited two systematic biases compared to my manual annotations: a tendency toward oversegmentation (splitting individual nuclei into multiple objects) and the production of more compact masks that left fewer background pixels between adjacent cells.

To address these issues, I first explored fine-tuning the Cellpose-SAM model on my manually annotated training data to potentially improve performance on this specific dataset. However, the performance of trained models degraded dramatically after as little as 15 epochs, with segmentation quality dropping to zero. This behavior could indicate rapid overfitting due to the limited size of my training dataset, or may reflect instabilities in the still young Cellpose-SAM training interface that have also been reported by other users on GitHub. Given the strong performance of the base model with optimized parameters, I chose to proceed with the pre-trained model rather than pursue further fine-tuning experiments.

Given these technical difficulties, I reverted to the base model and instead focused on systematically optimizing the post-processing hyperparameters that control how Cellpose converts predicted gradients and cell probabilities into final segmentation masks. These are:

- cellprob_threshold: sets the minimum probability value for pixel inclusion in cells. This parameter controls the boundary tightness of segmented masks, with higher values producing more conservative, compact masks by excluding lower-confidence boundary pixels.

- flow3D_smooth: controls the width of Gaussian blur applied to gradient fields before the gradient-following algorithm. Higher smoothing values reduce sensitivity to small gradient fluctuations that might otherwise cause oversegmentation of individual nuclei.

The optimization process involved systematically varying these parameters across a grid of values and evaluating the resulting segmentations against the manual ground truth using standard metrics.

2.2.3 Final segmentation pipeline

Based on the parameter optimization analysis, I selected the following configuration for all final segmentations:

- cellprob_threshold: 1.0

- flow3D_smooth: 3.0

These parameters were chosen to balance completeness and accuracy while producing masks with biologically reasonable morphology and consistent boundaries across the dataset.

Finally, to further refine segmentation quality, I implemented a series of morphological operations applied to the masks output by Cellpose-SAM:

1. Small object removal: masks smaller than 30 voxels were eliminated to remove noise and spurious detections that likely represent imaging artifacts rather than genuine nuclei.

2. Binary hole filling: internal gaps within masks were filled to ensure solid nuclear representations, as true nuclei should not contain large internal voids.

3. Morphological closing: a combination of dilation followed by erosion using a spherical structuring element (radius = 3 voxels) to smooth mask boundaries and close small gaps that might arise from imaging noise.

The complete pipeline, comprising Cellpose-SAM segmentation with optimized parameters followed by morphological post-processing, was applied uniformly to all gastruloids in the dataset. The pipeline was designed to be fully automated, requiring no manual intervention or sample-specific parameter adjustments, ensuring consistent processing across the entire experimental dataset. The resulting segmentation masks provided the foundation for all subsequent single-cell analysis and spatial pattern quantification.

2.3 Spatial analysis of gene expression patterns

Once a robust method for imaging gastruloids in full and isolating single cells from the 3D stacks was developed, quantitative analysis of gene expression patterns required extracting fluorescence intensities and spatial coordinates for each identified cell. For each segmented cell, mean fluorescence intensity was calculated across four fluorescence channels, providing quantitative measures of expression for each marker. Combined with the 3D spatial coordinates of each cell's centroid, this approach yields a dataset with seven features per cell.

Beyond the baseline DAPI nuclear staining that labels all cells uniformly, expression analysis focused on the intensity values from antibody staining for three key transcription factors selected for their roles in gastruloid patterning:

- Sox2, a posterior neural ectoderm marker
- Foxc1, an anterior mesoderm marker
- Bra (Brachyury/T), a posterior mesoderm marker

These transcription factors are proteins that bind to DNA and regulate the expression of other genes, thus controlling cell fate decisions. By visualizing their spatial distribution, researchers can map how different cell identities emerge and organize themselves within the gastruloid.

The immunostaining protocol employed a standard two-step approach. First, gastruloids were incubated with primary antibodies from different animal species

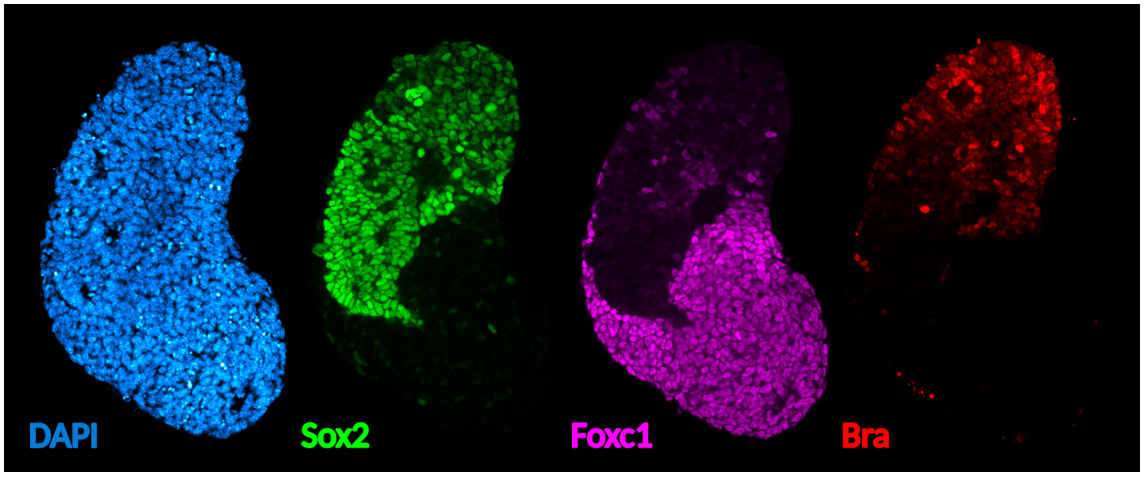


Figure 5: Four-channel fluorescence imaging of a representative gastruloid showing the characteristic DAPI (blue), Sox2 (green), Foxc1 (magenta), and Bra (red) expression patterns.

(rat anti-Sox2, rabbit anti-Foxc1, and goat anti-Bra) to specifically bind each transcription factor. Second, fluorophore-conjugated secondary antibodies were applied to detect each primary antibody independently:

- Anti-rat antibody conjugated to AlexaFluor 488 ($\lambda_{ex} = 488\,\text{nm}$ excitation, green emission)
- Anti-rabbit antibody conjugated to AlexaFluor 546 ($\lambda_{ex} = 546\,\text{nm}$ excitation, red emission)
- Anti-goat antibody conjugated to AlexaFluor 647 ($\lambda_{ex} = 647\,\text{nm}$ excitation, far-red emission)

This approach enables simultaneous visualization of all three markers plus DAPI in each gastruloid (figure 5).

The relationship between measured fluorescence intensity and actual molecular content requires careful consideration. In principle, the detected fluorescence intensity is proportional to the number of emitted photons, which depends on several factors: the number of target molecules present in the cell, the number of antibody-fluorophore complexes bound to these targets, the quantum efficiency of the fluorophore, and the excitation power used. Each fluorophore has different quantum efficiencies and requires wavelength-specific excitation powers optimized for signal detection rather than absolute quantification. Furthermore, different transcription factors may be produced at vastly different absolute concentrations within cells.

However, for the purposes of analyzing spatial patterns and cell type distributions, absolute molecular counts are not necessary. The key biological questions concern relative changes in expression, i.e. which cells express more or less of a given marker compared to their neighbors, and how these relative expression levels

vary across the tissue. Therefore, throughout this analysis, fluorescence intensities are treated as relative measurements and normalized appropriately when comparing across samples. This approach avoids the complex calibration that would be required for absolute quantification while preserving all information relevant to understanding spatial patterning and cell fate decisions.

2.3.1 Medial axis projection of the intensity values

The most well-characterized feature of gastruloids is the differentiation of cells into anterior and posterior tissues along the elongation axis, making this the natural starting point for spatial gene expression analysis. This anterior-posterior patterning represents a fundamental aspect of embryonic development, where cells adopt different identities based on their position along the body axis. To quantitatively analyze this patterning, robust methods for extracting expression patterns along the gastruloid's primary elongation axis are essential.

While previous published work (Merle et al., 2024) has analyzed this axis using the midline of 2D images obtained through maximum projections of the fluorescent signals on a single plane, the 3D imaging capability developed here enables extraction of the true three-dimensional curvilinear medial axis, directly improving upon previous approaches. This methodology provides a more accurate representation of the actual tissue geometry and cellular organization.

Computing the 3D medial axis requires several geometric processing steps. First, all segmented cells are pooled to create a binary mask distinguishing gastruloid tissue from background. The binary volume is then converted to a triangulated surface mesh using the marching cubes algorithm (Lewiner et al., 2003), which creates a smooth surface representation by connecting points of equal value (the tissue boundary) with triangles. This mesh is refined using Laplacian smoothing (Vollmer et al., 1999), an iterative process that moves each vertex toward the average position of its neighbors, reducing surface roughness while preserving overall shape. The smoothed volume is then re-converted to voxels and reduced to a one-dimensional skeleton using morphological thinning operations (Lee et al., 1994), which iteratively remove voxels from the surface while preserving connectivity, ultimately yielding a graph structure representing the core geometry. Finally, the medial axis is identified as the diameter (the longest of the shortest paths) of the skeletal graph, then smoothed and extended to reach the tissue boundaries (figure 6).

To analyze expression patterns along the anterior-posterior (A-P) axis, each cell's 3D centroid position is reduced to a single coordinate representing its position along the axis. For each cell, this medial axis projection involves finding the nearest point on the medial axis using a k-d tree spatial data structure (Maneewongvatana and Mount, 1999). The arc length along the medial axis from one end to this nearest point defines the curvilinear coordinate of that cell along the medial axis. The curvilinear coordinates were normalized by the total length of the axis to lie in a $[0, 1]$ range, because it has been shown (Merle et al., 2024) that gastruloids of

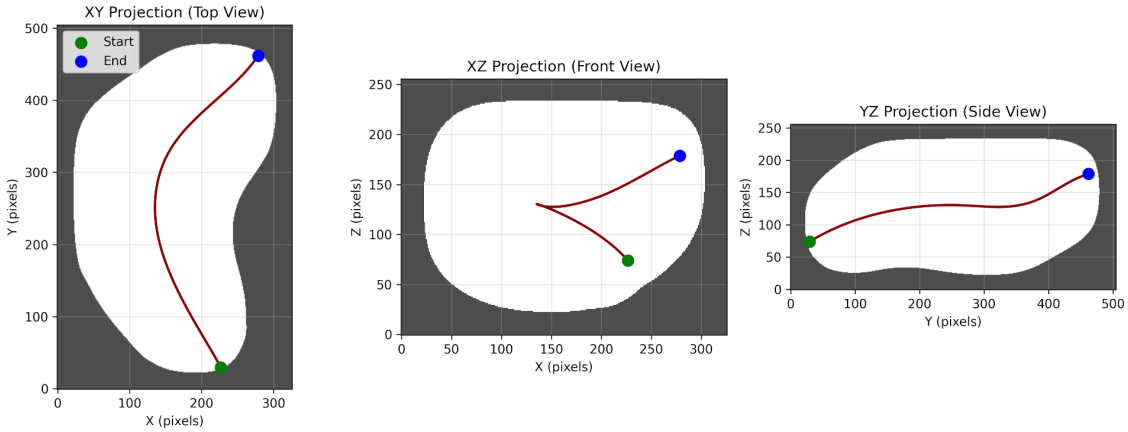


Figure 6: Three-dimensional medial axis extraction illustrated for a representative gastruloid. Orthogonal projections (XY, XZ, YZ views) showing the smoothed tissue mask (white) and the computed medial axis (red curve). The axis captures the primary anterior-posterior elongation direction while adapting to the curved 3D morphology of the sample.

different absolute sizes show similar relative patterning.

To identify the anterior and posterior ends of what would otherwise be an adirectional axis, the expression pattern of Bra, an exclusively posterior marker, is used as a reference. The average gradient of Bra expression along the curve is computed, and the coordinate system is oriented so that Bra expression consistently increases from anterior to posterior, with 0 corresponding to the anterior end and 1 to the posterior end.

Having characterized depth-dependent signal attenuation in the DAPI channel as a measure of clearing efficiency, it is important to consider whether similar effects might compromise the antibody channel measurements. However, I chose not to apply depth corrections to the antibody channel intensities for several reasons. First, the longer excitation wavelengths used for the antibody channels (488 nm, 546 nm, 647 nm) are significantly less prone to scattering than the shorter DAPI excitation wavelength (405 nm), as are the emitted wavelengths from each fluorophore, making the depth effect less pronounced. Second, and more critically, unlike the uniformly expressed DAPI signal, the antibody channels exhibit highly structured spatial expression patterns that are intrinsically heterogeneous.

Attempting to decouple systematic technical artifacts from genuine biological expression gradients would require *a priori* assumptions about the expected expression patterns, which could introduce bias into the analysis. In fact, examination of the z-dependence of the expression patterns revealed instances where cells deep within the sample exhibited higher antibody intensities than cells near the surface, a pattern that would be impossible if depth-dependent attenuation were the dominant effect. This observation suggests that biological expression heterogeneity, rather than technical limitations, is the primary source of intensity variation in the antibody channels. Therefore, the uncorrected intensity values were used directly for

this analysis, with the understanding that any residual depth effects are likely small compared to the biological signal of interest. Moreover, since the primary focus of this analysis is on relative differences along the anterior-posterior axis, which is largely orthogonal to the depth axis, any systematic depth-dependent effects would be unlikely to significantly alter the observed spatial expression patterns.

2.3.2 Average profile extraction

To enable comparison with previously published approaches, a binning strategy was employed to compute smoothed expression profiles along the normalized medial axis coordinate. For each gastruloid and each channel, cells were grouped into 50 bins spanning the full anterior-posterior extent, and mean intensities were calculated within each bin. This yielded smooth expression profiles for individual gastruloids, which could then be averaged across samples to obtain population-level expression patterns.

The methodology presented here provides a robust framework for quantitative analysis of spatial gene expression patterns in three-dimensional self-organizing systems, enabling direct comparison with existing literature while preserving the capability for more detailed single-cell analyses.

3 Results and Discussion

3.1 Enhanced optical clearing enables complete 3D imaging of gastruloids

The systematic evaluation of clearing protocols revealed dramatic differences in imaging depth and quality between conditions. The optimal protocol was found to be the combination of a fresh batch of RapiClear with enhanced permeabilization (0.5% Triton X-100).

The Fourier Ring Correlation Quality Estimate (FRC-QE) analysis provided the most discriminating metric for clearing quality (figure 7). While suboptimal conditions showed rapid decay of image quality with increasing depth into the sample, the optimal protocol maintained high FRC-QE values throughout the entire gastruloid volume. This metric, which quantifies the consistency between adjacent optical sections, effectively captured the visual impression of maintained resolution at depth.

While average fluorescence intensity might seem like a natural metric for assessing clearing quality, it failed to differentiate clearly between clearing conditions and to highlight the best one (figure 8). This counterintuitive result could be explained by considering that poorly cleared samples may show uniformly elevated background fluorescence due to scattered light, which elevates the average intensity even as useful signal is lost.

The coefficient of variation analysis proved more informative, revealing that the

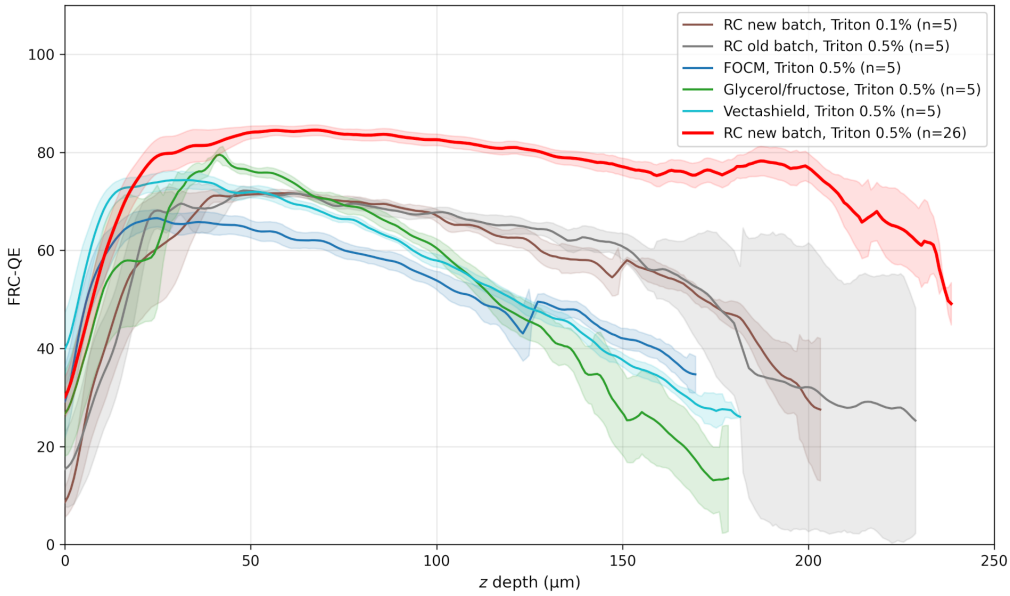


Figure 7: Fourier Ring Correlation Quality Estimate mean curves for six different clearing protocols. For each condition, the FRC-QE was computed individually for each gastruloid, then averaged to produce the mean curve shown (shading indicates standard error of the mean across gastruloids within each condition). The optimal protocol, utilizing a fresh RapiClear (RC) batch and enhanced permeabilization, was evaluated on $n = 26$ gastruloids and is highlighted in red; the other five conditions were each evaluated on $n = 5$ gastruloids.

optimal protocol better preserved image contrast at depth (figure 9). This metric, which measures the relative standard deviation of pixel intensities, effectively captured the preservation of cellular detail that enables accurate segmentation. This metric aligned well with both visual assessment and the FRC-QE results, confirming that the optimal protocol provides significantly better optical clearing throughout the three-dimensional structure of the gastruloids.

The practical impact of these improvements was substantial. Prior to optimization, complete imaging of $N_0 = 300$ gastruloids was impossible without physical compression, which would have introduced unacceptable morphological artifacts. The optimized protocol enabled routine imaging of entire gastruloids while preserving their natural three-dimensional architecture, providing the foundation for comprehensive single-cell analysis throughout the complete sample volume.

3.2 Automated segmentation achieves reliable single-cell identification

Developing an effective segmentation pipeline for gastruloid imaging data involved exploring diverse computational approaches, as described in the Methods section.

Despite being fine-tuned on organoid data, StarDist-Tapenade (Gros et al., 2024) only produced sparse, incomplete segmentations on my data (figure 10). The poor performance likely stems from differences in imaging conditions: my enhanced clear-

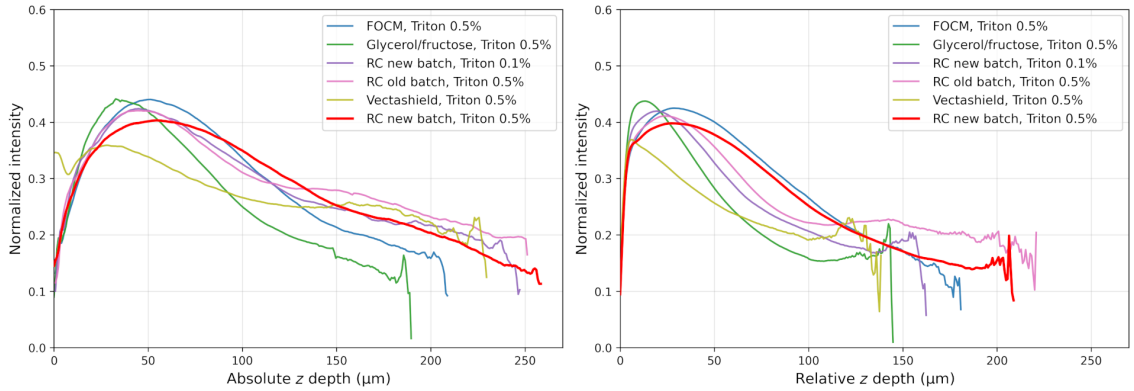


Figure 8: Mean intensity profiles as a function of depth for six clearing protocols. Left: absolute depth measured from the global minimum z-position of each gastruloid. Right: relative depth measured from the local minimum z-position at each xy coordinate. Intensity values were normalized to the maximum intensity within each gastruloid before averaging. Optimal protocol ($n = 26$ gastruloids) shown in red; other conditions evaluated on $n = 5$ gastruloids each. Note that the optimal protocol does not show the highest intensity at depth, illustrating why average intensity is a misleading metric for clearing quality assessment.

ing protocol may have produced image characteristics (contrast profiles, signal-to-noise ratios) that differ significantly from the training data used for StarDist-Tapenade. Additionally, StarDist models are limited to working on single-channel nuclear staining, which may limit their robustness when the nuclear channel alone provides insufficient information for reliable segmentation.

In contrast, Cellpose-SAM's multi-channel capability allowed it to leverage information from all available fluorescence markers, providing additional context for resolving ambiguous boundaries and ultimately yielding more complete and accurate segmentations. This stark performance difference illustrates the advantage of foundation model approaches over domain-specific fine-tuning. While StarDist-Tapenade relies on a smaller network fine-tuned within a narrow organoid domain, Cellpose-SAM employs a transformer backbone trained on billions of images from diverse and unrelated domains, enabling the model to leverage vastly broader visual knowledge when encountering novel conditions. Furthermore, Cellpose's gradient-based mask representation may be better suited than StarDist's star-convex polygons for capturing the irregular nuclear shapes common in dense tissues.

The systematic optimization of Cellpose's post-processing parameters (final parameter selection: cellprob_threshold = 1.0, flow3D_smooth = 3.0) yielded segmentations that captured the essential cellular organization while maintaining consistency across the dataset.

The optimized parameters still produced masks that differed visually from my manual annotations, but offered a valid alternative segmentation of the same cells. In this optimization process, I reached a fundamental limit of what constitutes "correct" segmentation. As discussed by Pachitariu et al. (2025), manual masks

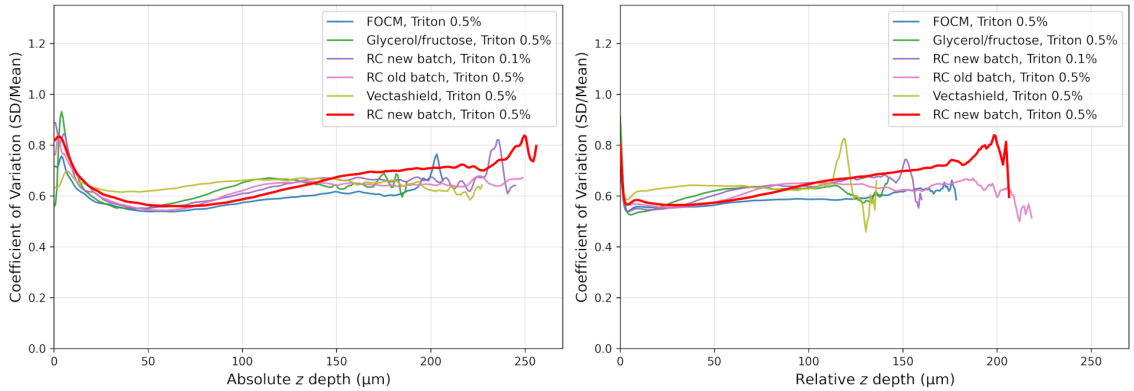


Figure 9: Coefficient of variation profiles as a function of depth for six clearing protocols, measuring contrast preservation as an indicator of clearing quality. Left: absolute depth; Right: relative depth. At a given depth, higher coefficient of variation indicates better preservation of cellular detail and contrast. The optimal protocol (red, $n = 26$ gastruloids) exhibits superior performance with higher contrast values, especially at intermediate and deep regions, compared to other conditions ($n = 5$ gastruloids each).

from different human annotators can vary significantly due to subjective preferences and segmentation styles, while still representing equally valid interpretations of the underlying biological structures. In this context, my manual annotations cannot be considered definitive ground truth, but rather one valid interpretation among many possible ones, and the automated predictions from Cellpose-SAM represent another equally legitimate interpretation.

The algorithmic approach may actually offer advantages over manual annotation in terms of consistency and reproducibility. While human annotators introduce variability based on fatigue, subjective interpretation, and inconsistent application of criteria, Cellpose-SAM applies the same learned priors uniformly across all cells. The systematic nature of the differences (with Cellpose producing more regular, compact masks) suggests that the model has learned biologically plausible constraints on nuclear morphology.

Given that my downstream analysis required consistent identification of cell positions and reasonable morphology rather than pixel-perfect agreement with any particular annotation style, the algorithmic segmentation proved entirely suitable for quantifying positional information in gastruloids.

Visual inspection (figure 11) confirmed that the segmentation captured the essential spatial organization of cells throughout the three-dimensional gastruloid structure. The complete pipeline successfully segmented an average of 18,829 $\pm$ 2,132 cells per gastruloid across 26 samples, yielding remarkably complete cell identification despite the algorithm having no specific training on this tissue type. This provided a suitable foundation for subsequent structural analysis.

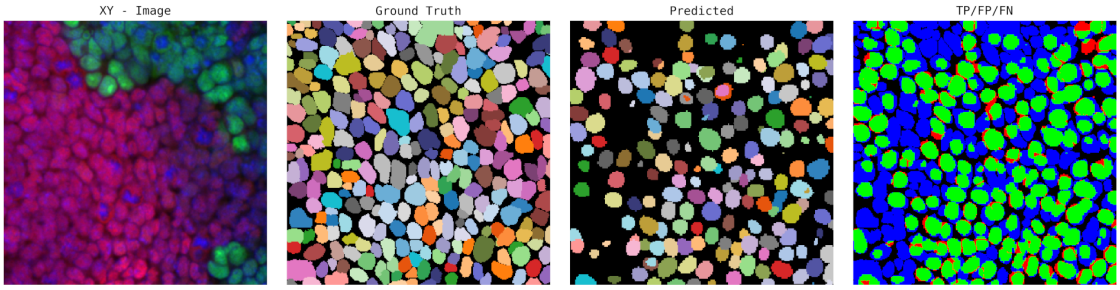


Figure 10: Comparison of StarDist-Tapenade segmentation performance against manual ground truth annotations, for one slice of the test volume. From left to right: original fluorescence image, manually annotated ground truth masks (each cell assigned a unique color), StarDist-Tapenade predictions, and performance evaluation (green = true positives, blue = false negatives, red = false positives, black = true negatives). The model detected only a small fraction of cells present in the manual annotations, with most cells appearing as false negatives (blue), demonstrating poor performance.

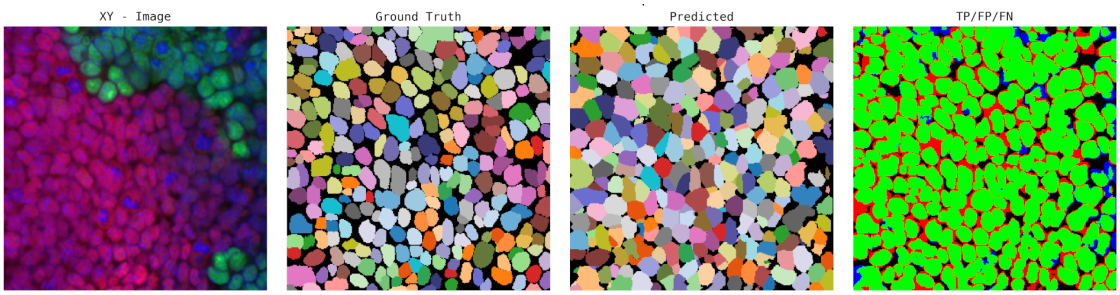


Figure 11: Segmentation performance of the complete segmentation pipeline against manual ground truth annotations, for one slice of the test volume. From left to right: original fluorescence image, manually annotated ground truth masks (each cell assigned a unique color), post-processed predictions, and performance evaluation (green = true positives, blue = false negatives, red = false positives, black = true negatives).

3.3 Spatial gene expression patterns reveal complex tissue organization

The analysis of gene expression patterns along the anterior-posterior axis revealed structured spatial organization consistent with known gastruloid biology. The population-averaged expression profiles (figure 12) showed characteristic patterns for each marker which align qualitatively with those reported by Merle et al. (2024), validating our imaging and analysis pipeline against established results.

Importantly, our approach offers significant methodological advantages over previous studies. Unlike analyses based on maximum intensity projections, our full 3D approach preserves the native three-dimensional distribution of cells and avoids the information loss inherent to flattening complex 3D structures onto 2D images. Furthermore, our medial axis extraction operates directly on the 3D gastruloid geometry, providing a true anatomical midline rather than the distorted midline that results from 2D projection artifacts. These improvements enable more accurate

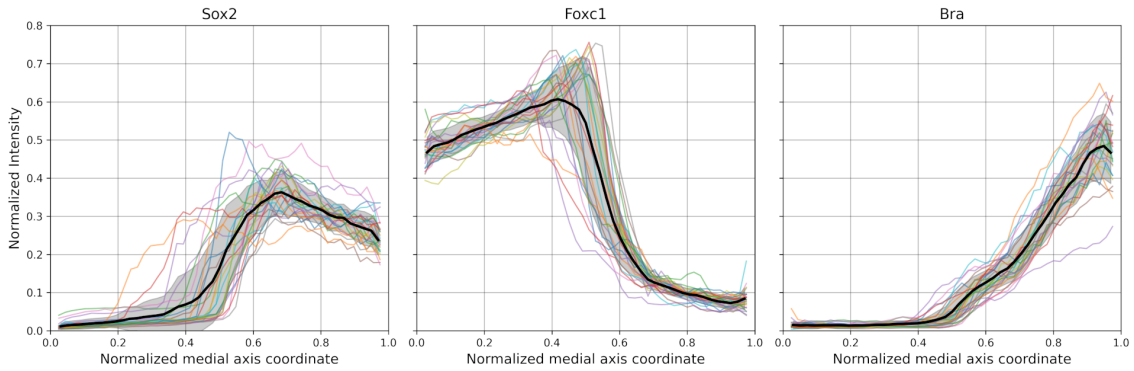


Figure 12: Population-averaged expression profiles along the anterior-posterior axis. Solid lines show mean intensity across $n = 25$ gastruloids; shaded regions indicate standard error. The profiles show characteristic spatial patterns for each marker along the gastruloid axis.

quantification of spatial gene expression patterns and better preservation of the cellular-level information that drives self-organization.

However, the true power of single-cell resolution becomes apparent when examining the full distribution of expression values rather than just averages. The averaging approach, while useful for comparison with the literature, necessarily discards rich information present in the single-cell data. The scatter plots of individual cell intensities (figure 13) reveal heterogeneities that would be completely masked by population averaging. Most notably, two markers (Sox2 and Foxc1) show clear bimodal distributions at specific positions along the axis, indicating the coexistence of distinct cell populations rather than a continuous gradient of expression levels.

This bimodality was observed reproducibly across the majority of analyzed gastruloids. Given these heterogeneous distributions, taking simple averages produces intermediate values that may not correspond to any actual cell state. This consideration is important for interpreting averaged expression profiles and suggests that more sophisticated analytical approaches incorporating cellular heterogeneity represent an important direction for future development.

Visual inspection of the antibody channels showed little apparent relationship between imaging depth and fluorescence intensity for the gene expression markers, in contrast to the clear depth dependence visible in the DAPI channel. This observation supports our decision not to apply depth correction to the antibody channels, as the biological heterogeneity in gene expression appears to dominate over any technical attenuation effects.

4 Conclusions

Through systematic optimization and the application of state-of-the-art computational tools, this work has established a robust pipeline for quantifying gene expression patterns in gastruloids at single-cell resolution throughout entire 3D samples.

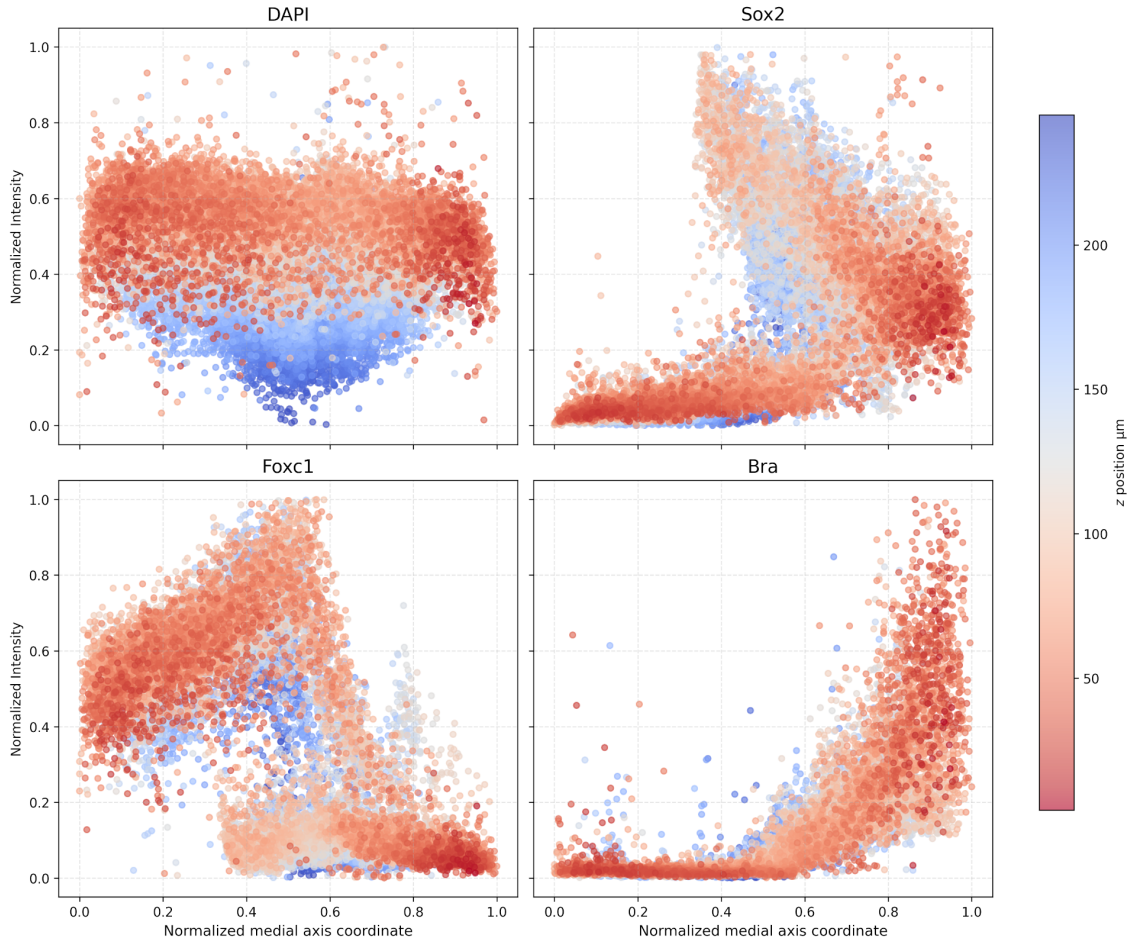


Figure 13: Single-cell fluorescence intensities along the medial axis of a representative gastruloid. Each dot represents one cell, positioned according to its normalized coordinate along the anterior-posterior axis. The four panels show DAPI, Sox2, Foxc1, and Bra expression respectively. Datapoints are color-coded by their absolute z-depth within the sample (red = shallow, blue = deep). While systematic intensity decay with depth is clearly visible in the DAPI channel, the relationship is much more complex for the antibody channels, where biological expression patterns dominate over technical artifacts.

This represents a significant methodological advance for studying self-organization in mammalian developmental models, where previous approaches were limited by difficulties in imaging dense tissues and reconstructing complete spatial patterns.

The key technical achievements include the development of an effective clearing protocol that enables deep tissue imaging while preserving morphology and fluorescence, the construction of an automated segmentation pipeline capable of accurately extracting single-cell information over thousands of cells per sample, and the extension of spatial analysis approaches to enable gene expression pattern analysis along the anterior-posterior axis in a fully three-dimensional setting. These methodological improvements enable more accurate quantification of spatial gene expression patterns and better preservation of the cellular-level information that drives self-organization.

The preliminary analysis I performed demonstrates the existence of complex spatial organization within this self-organizing system. The bimodal distributions observed for multiple markers at specific axial positions suggest that positional information may be encoded not in simple gradients of mean expression, but rather in the spatially-varying probability distributions of discrete cell states. This probabilistic encoding could provide a more robust and information-rich mechanism for pattern formation than classical morphogen gradient models suggest.

The technical capabilities established here will be crucial for understanding how these structures emerge spontaneously and how this process relates to the generation of positional information. Furthermore, applying this analytical framework to different developmental timepoints and initial seeding densities will enable evaluation of how structures emerge over time and how they scale with system size. This represents a critical step toward understanding the fundamental principles governing biological self-organization.

By extending quantitative frameworks previously limited to simpler model systems like Drosophila to more complex self-organizing mammalian stem cell systems, this work provides the flexibility to probe self-organization under controlled yet biologically relevant conditions. With the ability to measure spatial patterns precisely and comprehensively, researchers can now test quantitative theories about the fundamental limits and mechanisms of biological self-organization.

References

Bachani, V., Roychowdhury, A., Dasgupta, A., and EI-Alfy, H. (2024). Image Segmentation Survey: Classical and Deep Learning Methods. In *2024 International Conference on Electrical, Computer and Energy Technologies (ICECET*, pages 1–6.

Beccari, L., Moris, N., Girgin, M., Turner, D. A., Baillie-Johnson, P., Cossy, A.-C., Lutolf, M. P., Duboule, D., and Arias, A. M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature*, 562(7726):272–276.

Dekkers, J. F., Alieva, M., Wellens, L. M., Ariese, H. C. R., Jamieson, P. R., Vonk, A. M., Amatngalim, G. D., Hu, H., Oost, K. C., Snippert, H. J. G., Beekman, J. M., Wehrens, E. J., Visvader, J. E., Clevers, H., and Rios, A. C. (2019). High-resolution 3D imaging of fixed and cleared organoids. *Nature Protocols*, 14(6):1756–1771.

Gros, A., Vanaret, J., Dunsing-Eichenauer, V., Rostan, A., Roudot, P., Lenne, P.-F., Guignard, L., and Tlili, S. (2024). A quantitative pipeline for whole-mount deep imaging and multiscale analysis of gastruloids.

Hafiz, A. M. and Bhat, G. M. (2020). A survey on instance segmentation: State of the art. *International Journal of Multimedia Information Retrieval*, 9(3):171–189.

Jacques, S. L. (2013). Optical properties of biological tissues: A review. *Physics in Medicine & Biology*, 58(11):R37.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., and Girshick, R. (2023). Segment Anything.

Lee, T. C., Kashyap, R. L., and Chu, C. N. (1994). Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms. *CVGIP: Graphical Models and Image Processing*, 56(6):462–478.

Lewiner, T., , Hélio, L., , Antônio Wilson, V., and and Tavares, G. (2003). Efficient Implementation of Marching Cubes' Cases with Topological Guarantees. *Journal of Graphics Tools*, 8(2):1–15.

Maneewongvatana, S. and Mount, D. M. (1999). Analysis of approximate nearest neighbor searching with clustered point sets.

McGough, L., Casademunt, H., Nikolić, M., Aridor, Z., Petkova, M. D., Gregor, T., and Bialek, W. (2024). Finding the Last Bits of Positional Information. *PRX Life*, 2(1):013016.

Merle, M., Friedman, L., Chureau, C., Shoushtarizadeh, A., and Gregor, T. (2024). Precise and scalable self-organization in mammalian pseudo-embryos. *Nature Structural & Molecular Biology*, 31(6):896–902.

Pachitariu, M., Rariden, M., and Stringer, C. (2025). Cellpose-SAM: Superhuman generalization for cellular segmentation.

Pawley, J. (2006). *Handbook of Biological Confocal Microscopy*. Springer Science & Business Media.

Preusser, F., dos Santos, N., Contzen, J., Stachelscheid, H., Costa, É. T., Mergenthaler, P., and Preibisch, S. (2021). FRC-QE: A robust and comparable 3D microscopy image quality metric for cleared organoids. *Bioinformatics*, 37(18):3088–3090.

Ravikumar, S., Surekha, R., and Thavarajah, R. (2014). Mounting media: An overview. *Journal of Dr. YSR University of Health Sciences*, 3(Suppl 1):S1.

Richardson, D. S., Guan, W., Matsumoto, K., Pan, C., Chung, K., Ertürk, A., Ueda, H. R., and Lichtman, J. W. (2021). Tissue clearing. *Nature Reviews Methods Primers*, 1(1):84.

Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). Cell detection with star-convex polygons. In Frangi, A. F., Schnabel, J. A., Davatzikos, C., Alberola-López, C., and Fichtinger, G., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 265–273, Cham. Springer International Publishing.

Shahbazi, M. N., Siggia, E. D., and Zernicka-Goetz, M. (2019). Self-organization of stem cells into embryos: A window on early mammalian development. *Science*, 364(6444):948–951.

Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: A generalist algorithm for cellular segmentation. *Nature Methods*, 18(1):100–106.

Tkačik, G. and Gregor, T. (2021). The many bits of positional information. *Development*, 148(2):dev176065.

Vollmer, J., Mencl, R., and Müller, H. (1999). Improved Laplacian Smoothing of Noisy Surface Meshes. *Computer Graphics Forum*, 18(3):131–138.

Zhao, Z., Chen, X., Dowbaj, A. M., Sljukic, A., Bratlie, K., Lin, L., Fong, E. L. S., Balachander, G. M., Chen, Z., Soragni, A., Huch, M., Zeng, Y. A., Wang, Q., and Yu, H. (2022). Organoids. *Nature reviews. Methods primers*, 2:94.

Zhu, X., Huang, L., Zheng, Y., Song, Y., Xu, Q., Wang, J., Si, K., Duan, S., and Gong, W. (2019). Ultrafast optical clearing method for three-dimensional imaging with cellular resolution. *Proceedings of the National Academy of Sciences*, 116(23):11480–11489.