Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

fishROI: A specialized workflow for semi-automated muscle morphometry analysis in teleosts

Yansong Lu
Australian Regenerative Medicine Institute
Michael Pan
University of Melbourne
Vijayishwer Jamwal
Mount Desert Island Biological Laboratory
Putri Halleyana Adrikni Rahman
University of Melbourne
Jake Locop
Australian Regenerative Medicine Institute
Avnika A. Ruparelia
University of Melbourne
Peter D. Currie
peter.currie@monash.edu

Australian Regenerative Medicine Institute

Method Article

Keywords: Muscle, teleost, zebrafish, machine-learning, segmentation, growth, heterogeneity

Posted Date: June 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10009618/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Skeletal Muscle on September 16th, 2026. See the published version at https://doi.org/10.1186/s13395-026-00449-y.

Abstract

Background

Quantitative histological analysis of skeletal muscle morphometry provides critical insights into muscle physiology but remains labor-intensive and technically demanding. While recent developments in machine learning-based image segmentation techniques have facilitated large-scale tissue analysis, existing tools that automate muscle morphometry analysis are largely tailored to mammalian models, with limited applicability to teleosts. Moreover, there is a lack of effective tools for visualizing spatial organization and morphometric variability of teleost muscle fibers, a feature that is important for understanding hyperplastic muscle growth dynamics in teleosts.

Methods

In this study, we benchmarked the performance of existing muscle morphometry analysis workflows and general-purpose cell segmentation algorithms in identifying individual muscle fibers in teleost muscles. We used these findings to inform the development of a new, teleost-optimized analysis tool.

Results

We show that cytoplasmic staining combined with machine learning-based cell segmentation offers a robust and accurate approach for automated muscle morphometry analysis in developing zebrafish. We also introduce fishROI – a FIJI2 plugin implemented in Jython, that streamlines both morphometric analysis and visualization. This tool accommodates shallow and deep learning–based segmentation techniques and incorporates novel quantification and visualization methods suited to teleost-specific muscle features, including mosaic hyperplasia dynamics. The plugin features an intuitive graphical user interface and is designed for flexibility, with minimal constraints regarding species, image quality, or staining protocol. Its modular architecture allows it to be used as a baseline for automated muscle morphometry analysis, while permitting integration with other tools and workflows to permit a diverse range of analytical needs.

Conclusion

This plugin will serve as a foundational platform for semi-automated muscle morphometry analysis in teleosts, enabling a combination of reliable image segmentation, quantification and visualization optimized for teleost muscles.

Background

Quantitative analysis of muscle morphology, or muscle morphometry, is a well-established method for investigating muscle dynamics across processes such as growth, homeostasis, disease progression, regeneration, and aging. While this technique is extremely informative, it historically involved manual quantification of individual muscle fibers for their sizes and shapes, making the analysis tedious and labor-intensive. Recent advances in computational image analysis introduced tools to automate muscle morphometry analysis, particularly in mammalian systems such as humans and mice (1–3). However, when researchers studying non-mammalian muscle attempt to utilize these tools, the results are often poor. This is because existing tools rely on specific staining protocols optimized for mammalian tissue, such as precise extracellular matrix (ECM) or membrane markers (1, 3, 4). These stains are often suboptimal when applied to other valuable but under-supported models such as teleosts, resulting in poor image segmentation. On the other hand, while cytoplasmic staining consistently labels individual myofibers in non-mammalian species, it is not adequately leveraged by existing workflows to enable fiber identification. Beyond the reliance on strong ECM staining, existing algorithms typically assume muscle fibers to be relatively uniform in size to distinguish true fibers from imaging artifacts (1). However, this assumption does not hold true for teleost muscle, because teleosts exhibit extensive hyperplastic muscle growth throughout development, creating a broad distribution of fiber sizes in the developing muscle (5). As a result, algorithms optimized for the structural uniformity of mammalian muscle often misclassify small, newly synthesized muscle fibers as artefacts when applied to teleost tissues. Due to these shortfalls, muscle researchers studying non-mammalian species cannot effectively utilize existing tools, highlighting the pressing need for more adaptable and species-agnostic approaches to automated muscle morphometry analysis that do not depend on narrowly defined staining protocols, sample quality, and morphological assumptions.

In this study, we address the gap in teleost muscle morphometry analysis pipelines by demonstrating that cytoplasmic staining of zebrafish skeletal muscle, combined with general purpose, machine learning-based image segmentation, consistently enables accurate and high-throughput muscle morphometry analysis. We introduce fishROI – a complete, user-friendly workflow implemented as an ImageJ2/FIJI2 plugin (https://imagej.net/software/fiji/). fishROI provides a modular graphical user interface (GUI) that guides users through the entire process – from image segmentation to morphometric quantification and visualization. Users can select from shallow-learning algorithms for faster, less resource-intensive segmentation, deep-learning models for higher accuracy and throughput, or a combination of both depending on experimental needs. All approaches support custom model training using user-specific datasets, allowing for optimization across diverse sample types. Shallow models can be trained within minutes, whereas deep learning models may require several hours. Nevertheless, we show that the pre-trained Cellpose cyto3 model already performs well in segmenting muscle fibers based on cytoplasmic staining (6), which we were able to further improve by training a zebrafish-specific model for fiber segmentation.

Following segmentation, the plugin provides tools for region-of-interest (ROI) generation and filtering, manual refinement of detected ROIs, and detailed quantification and visualization of fiber size distributions and morphometric heterogeneity. These outputs offer both visual and quantitative insights

into the structural organization of teleost muscle, leveraging the large datasets generated by the image segmentation workflow.

We anticipate that this plugin will serve as a foundational platform for automated muscle morphometry analysis in teleosts, while retaining the flexibility needed to accommodate a wide range of species, developmental stages, and experimental designs. Its modular design also ensures future compatibility with evolving segmentation methods and image analysis tools.

Methods

Sample sources

Zebrafish husbandry was conducted according to standard procedures approved by Monash Animal Services Animal Ethics committee, Monash University (ERM41803, ERM29767). All zebrafish belong to the TU strain, including the wildtype and Tg(*actc1b*: GFP) transgenic line. Fish were euthanized by incubation in 1 g/L tricaine methane sulfonate. Muscles were dissected into thin chunks (approximately 10 mm thickness) and fixed in 4% paraformaldehyde (PFA) overnight at 4 degrees. For vibratome sectioning, fixed muscle was embedded in 4% agarose and sectioned into 200 μm slices using Leica VT1200S vibratome, then stored at 4ºC in PBST until staining. For cryosectioning, fixed tissue was washed serially into 10%, 20% and 30% sucrose and embedded in Tissue-Tek® O.C.T. Compound (Sakura), followed by cryo-sectioning into 12 μm slices using Leica CM 1850 cryostats and collected on poly-lysine-coated slides, then stored at -80ºC until staining.

Immunohistochemistry (IHC) protocol

For Tg(*actc1b*: GFP) zebrafish, vibratome sections were mounted directly into 1% low-melt agarose, followed by confocal imaging on Leica SP8 multiphoton microscope (multiphoton functionality disabled). Other zebrafish vibratome sections were incubated with 2% goat serum block (2% goat serum, 1% BSA, 1% DMSO in PBST) for 1 hour, followed by overnight incubation with Alexa Fluor 546 phalloidin (Invitrogen, A22283, 1:400 dilution) and WGA-Fluorescein (VectorLabs, FL1021, 1:300 dilution) diluted in goat serum block at 4ºC, then washed in PBST 6 times before embedding in 1% low melt agarose, followed by imaging on Leica SP8 multiphoton microscope (multiphoton functionality disabled). Cryosections were thawed at room temperature for 10 min, followed by 5 min of PFA fixation and three PBST washes. The tissue was then incubated in 2% goat serum block for 30 min, followed by incubation with WGA-Rhodamine (VectorLabs, RL1022, 1:300 dilution) for 2 hours, followed by 6 PBST washes. Sample was subsequently mounted in 80% glycerol in PBST and imaged on the Leica thunder deconvolution microscope. Additionally, we've also tested the following primary antibodies on cryosections after blocking, including anti-Collagen, type I (Abcam, ab23730), anti-Col2a1a (Sigma, SAB2701915), anti-α-dystroglycan (Sigma, 05-298), anti-β-dystroglycan (Abcam, ab49515) and anti-β-tubulin (Abcam, ab6046). These primary antibodies were diluted in goat serum block at 1:500 dilution

and incubated overnight at 4ºC, followed by 6 PBST washes, and incubation in appropriate Alexa Fluor 488 secondary antibodies (Invitrogen, 1:300 dilution) for 2 hours, then washed again using 6 PBST washes, followed by mounting in 80% glycerol in PBST, followed by imaging on Zeiss Z1 compound microscopes.

Mammalian muscle morphometry analysis workflow benchmarking

For each of the five tools we tested, three cropped early juvenile zebrafish muscle images stained with fluorescently labelled WGA and phalloidin were processed to determine a list of ROIs representing identified fibers. The following parameters were used to execute the workflows: (1) MacroIMRB – Min Area = 20, Max area = 99999, Min circularity = 0.3, Minor diameter = 1.5; (2) MuscleJ2 – Pathophysiology = Damaged, Muscle Types = Limb, Data Analysis by Section = ECM Area Detection, Data Analysis by Fiber = Fiber Morphology, Data Cartographies = Fiber Area; (3) TRUEFAD – Artificially enhance laminin edges = True, label parameters = default; (4) Myosoft – Min area = 20, Max Area = 99999, circularity/solidity/perimeter/ferret/roundness = default. True and false positive rates of generated ROIs were determined by comparing them against manually annotated datasets using custom written Jython scripts described below.

Labkit and Cellpose model training

Labkit models were trained manually on a 2019 MacBook Pro laptop by carefully annotating foreground and background for individual images. Model performance was monitored during training to ensure optimal segmentation quality for each image. The resulting segmentations were exported as binary masks and converted into ROIs using the "Analyze Particles" function in FIJI. For Cellpose, custom models were trained using images and corresponding ROI annotations generated in FIJI, which were converted into segmentation mask PNG files via custom scripts. Training was performed on the MASSIVE-M3 High-Performance Computing (HPC) platform using 6 CPUs, 100 GB RAM, and an NVIDIA T4 GPU (16 GB VRAM), executed through command-line interface (CLI) prompts. Image segmentation with Cellpose was also performed on the same HPC platform, using custom Python scripts.

Model performance validation

A custom Jython script was developed to match ROIs from manual annotations with model predictions based on centroid proximity. For each matched pair, the ROI areas were compared and matches exceeding 65% overlap were classified as True Positives. The true positive rate was calculated as the number of correctly identified ROIs divided by the total number of annotated ROIs. Predicted ROIs that could not be matched to any manual annotation were classified as False Positives, and the false positive rate was calculated as the number of unmatched predicted ROIs divided by the total number of predicted ROIs. To statistically compare model performance across developmental stages, one-way ANOVA was used for comparisons among three groups, followed by post hoc Tukey’s test where appropriate. For

comparisons between two groups, Student's *t*-test was applied. All statistical analyses were performed in R, and visualizations were generated using the ggplot2 package.

Hardware and software requirements

Our FIJI plugin and Julia script have been tested on the following operating systems: macOS Sonoma 14.2.1, Windows 10/11 Home, Ubuntu 22.04.03 LTS (64-bit) with a minimum RAM of 8 GB. However, we recommend up to 64 GB RAM for processing very large images (5–10 GB). While the processing speed of our FIJI plugin primarily depends on the CPU, Labkit and Cellpose both support GPU acceleration to speed up image segmentation. Note that Cellpose requires a GPU that supports CUDA, such as NVIDIA and Intel GPU, but not AMD GPU. We have used the following software versions: FIJI2 v1.54f, Labkit 0.3.11, Julia v1.10.0. Further troubleshooting guide can be found on our GitHub page.

Results

1. Mammalian muscle morphometry analysis workflows are ineffective for teleost muscles

In recent years, muscle morphometry analysis in teleost models has been critical for understanding muscle biology, yet it remains predominantly reliant on manual methods (7–10). Automated tools developed for mammalian systems are often incompatible with teleost tissues, largely due to their dependence on specific antibodies or staining protocols that perform suboptimally in non-mammalian species (Fig. 1A, B). This limitation is especially pronounced in larval and juvenile fish, where extracellular matrix (ECM) and membrane staining is frequently weak or inconsistent (Fig. 1B), rendering most existing segmentation algorithms – which rely on sharp ECM boundaries – ineffective. To further compound this problem, newly synthesized fibers arising from hyperplasia are difficult to resolve from membrane markers, because these small fibers often occupy narrow gaps between ECM structures, resulting in low signal-to-noise ratios and irregular morphologies that pose significant challenges for membrane-based segmentation (Fig. 1C).

To empirically validate the performance of existing algorithms in identifying teleost muscle fibers, we first tested the popular MuscleJ2 (1), MyoVision (3), Myosoft (2), TRUEFAD (11), and MacroIMRB (12) workflows using juvenile zebrafish muscle sections stained for cytoplasm and ECM (Table 1). These juvenile muscles have sharper ECM staining than larval fish muscles but are still weaker than mammalian muscles, allowing us to benchmark mammalian-optimized workflows under a demanding but workable condition. When attempting to execute these algorithms, we immediately encountered several problems. Namely, the download link for MyoVision is no longer available, whereas Myosoft triggers an error halfway through the code, preventing completion of the workflow (Figure S1). While MuscleJ2 and MacroIMRB both remained fully functional, their performances were poor on our images, with MuscleJ2 recognizing only 1–2 fibers per image, and MacroIMRB recognizing around 10% of fibers correctly (Figure S2). Within the five tools we tested, TRUEFAD workflow greatly outperformed others, identifying approximately 40% of fibers correctly, while giving an equal number of false positives.

However, because TRUEFAD is reliant on ECM staining, when we tested it on larval fish that produces extremely faint ECM staining, fibers could no longer be identified. In summary, none of the existing tools reliably perform muscle morphometry analysis in teleosts, largely due to the reliance on ECM staining that is difficult to achieve in young fish.

Table 1
Benchmarking of existing muscle morphometry analysis workflows.

Tools:	Does it still work?	True positive rate	False positive rate	Technical limitations	Misc.
MuscleJ2	Fully functional	< 1%	NA	Model trained exclusively for mammalian muscle using specific antibodies and stains. Does not generalize well to teleosts.	Descriptions are tailored to specific mammalian tissue. For example, choosing "damaged" pathophysiology permits greater fiber size heterogeneity, more suitable to developing fish muscle.
Myosoft	Code error in workflow	< 10%	~ 90%	Model trained exclusively for mammalian muscle using specific antibodies and stains. Also relies exclusively on ECM labelling, which is faint in young fish.	Code successfully creates ROIs, but triggers an error during ROI filtering, likely causing the high false positive rate.
Myovision	No longer available	NA	NA	Relies exclusively on ECM labelling, which is faint in young fish.	
TRUEFAD	Yes	~ 40%	~ 55%	Relies exclusively on ECM labelling, which is faint in young fish.	ReadAndWriteExcel dependency is no longer available in FIJI2, which may be causing the inability to autosave output files.
MacroIMRB	Fully functional	~ 10%*	~ 20%*	Relies on ECM/membrane staining for fiber identification. Does not generalize well beyond the trained human tissue.	Does not output ROI for downstream analysis.

True positive rate is calculated as number of ROI correctly predicted by each tool divided by total fiber count in each image (manually determined). False positive rate is the percentage of predicted ROIs that

cannot be correctly matched to a real fiber. *: as MacroIMRB does not produce output ROIs, we visually estimated the true and false positive rates using the flattened output images.

2. Cytoplasmic staining combined with deep learning segmentation offers robust and accurate muscle morphometry analysis in developing zebrafish

To address the limitations in using ECM or membrane staining to identify individual fibers, we assessed cytoplasmic staining—specifically phalloidin—as an alternative to ECM-based markers such as Wheat Germ Agglutinin (WGA). While the effectiveness of WGA declines markedly in zebrafish at early developmental stages, phalloidin consistently provides strong cytoplasmic labeling across all development stages and offers improved visualization of small, nascent fibers (Fig. 1D). This makes it a more suitable marker for muscles in developing teleosts.

However, even with robust cytoplasmic labeling, segmenting closely apposed fibers remains challenging, particularly in early developmental stages. We evaluated the effectiveness of general-purpose cell segmentation algorithms, including the shallow learning algorithm Labkit (13) and the deep learning algorithm Cellpose (cyto3 model), in identifying phalloidin-stained zebrafish trunk muscle at multiple developmental stages. To optimize Labkit performance, we trained individual models for each image, yet it still frequently failed to distinguish adjacent fibers, especially in early larval tissues where fibers are tightly packed (Fig. 1D). When we compared manually annotated and algorithmic predicted muscle fibers in each image based on centroid proximity and area (Fig. 1E), we found that Labkit performed poorly in early larval samples due to its inability to resolve adjacent fibers, but accuracy improved in late larval and juvenile tissues (Fig. 1F, G). In contrast, the pre-trained Cellpose cyto3 model – despite not being trained specifically on muscle tissue – performed markedly better without the need for retraining. It consistently outperformed Labkit in correctly identifying fibers and generated significantly fewer false positives at both late larval and early juvenile stages. Notably, Cellpose correctly segmented approximately 80% of fibers in late larval and early juvenile samples, with a false positive rate below 20%, highlighting its strong baseline performance for cytoplasmic segmentation of zebrafish muscle. Taken together, these results demonstrate that Cellpose significantly outperforms other existing segmentation workflows in identifying larval and juvenile zebrafish muscle fibers.

3. Custom, sample-specific deep learning model in Cellpose further improves fiber segmentation accuracy

Having established that the baseline cyto3 model in Cellpose can segment zebrafish muscle fibers with reasonable accuracy, we next investigated whether training a custom model based on cyto3 could further improve performance. To this end, we curated a dataset of 54 confocal images of zebrafish trunk muscle cross-sections stained with phalloidin, spanning early larval, late larval, and early juvenile developmental stages. Initial segmentations were generated using a combination of shallow and deep learning algorithms, followed by extensive manual curation to produce a ground truth dataset comprising 19,840 ROIs.

The dataset was randomly split into 70% for training and 30% for testing. New models were trained in Cellpose using the cyto3 model as a base. To evaluate whether a single universal model or stage-specific models would yield superior results, we trained both: a universal zebrafish model using all images (hereafter referred to as the "*rerio*" model), and three stage-specific models trained on images from each developmental stage separately. Model performance was evaluated on the held-out testing set using the same matching criteria described previously: predicted and annotated ROIs were considered true matches if their centroid proximity was acceptable and if the area overlap exceeded 65%. Surprisingly, custom models did not significantly improve the true positive rate for fiber detection across developmental stages, although we observed an upward trend in performance (Fig. 1H). In contrast, custom models led to a significant reduction in false positive rates at all stages (Fig. 1I), indicating that training on zebrafish-specific data improves the algorithm's ability to reject artifacts and non-fiber regions.

Additionally, no significant differences were observed between the universal *rerio* model and the stage-specific models. However, a slight downward trend in false positive rates for stage-specific models suggests that larger datasets may grant further performance gains from stage-specific-models. Together, these findings demonstrate that training a custom zebrafish model in Cellpose improves segmentation accuracy, primarily by reducing false positive detections. Moreover, a single universal model may be sufficient for robust segmentation across a range of developmental stages, though future work may clarify whether stage-specific models provide additional benefits.

4. Coefficient of variation in fiber area correctly identifies growth zones and quantifies mosaic hyperplasia in zebrafish

To leverage the comprehensive morphometric analysis, we performed across the entire muscle section, we next aimed to quantify a defining feature of teleost muscle growth: mosaic hyperplasia. This process, characterized by the continuous addition of small, newly synthesized fibers throughout the myotome, underlies much of the post-embryonic muscle growth in teleosts (8, 14, 15) (Fig. 2B). While mosaic hyperplasia has traditionally been assessed qualitatively – typically by visually identifying small fibers deep within the myotome (16, 17) – there remains no standardized method for its quantification, which limits comparative analyses across developmental stages, treatments, or species. Because mosaic hyperplasia increases fiber size heterogeneity within the muscle, we developed an approach to quantify and visualize this variability by calculating the coefficient of variation (CoV) in fiber cross-sectional area across localized regions of the myotome (Fig. 2Aiv, Biv). As expected from previous studies, juvenile zebrafish – where mosaic hyperplasia is active – showed elevated CoV values in the deep myotome, indicating substantial local heterogeneity in fiber size. To validate this metric, we compared CoV values deep in the myotome of early larval zebrafish (which do not undergo mosaic hyperplasia) to those in early juveniles. The analysis revealed a significant increase in CoV in juvenile muscle, consistent with hyperplasia dynamics. These findings demonstrate that local CoV provides a robust, quantitative measure of mosaic hyperplasia in zebrafish and may be broadly applicable for assessing muscle growth dynamics in other teleost species.

5. Establishment of a start-to-finish FIJI plugin to streamline muscle morphometry analysis in teleosts

Given the complexity of options available for automated or semi-automated muscle morphometry in teleost samples – particularly the difficulties related to file format compatibility and intermediate processing steps – we developed a FIJI2 plugin designed to streamline the workflow and automate many of these intermediate tasks. In addition, we provide tools to leverage the extensive data generated through this process, including new techniques to visualize and quantify muscle morphometry, many of which have been demonstrated earlier.

As machine learning-based image segmentation continues to advance rapidly, we built the plugin's graphical user interface (GUI) using a modular design. This modularity allows users to run each component in sequence to perform a complete muscle morphometry analysis and visualization pipeline (Fig. 3A, B). Alternatively, users can integrate individual modules from this plugin into other workflow or software tools for improved performance for their sample type. Our modules support both shallow learning (via Labkit) and deep learning (via Cellpose), enabling users to select the segmentation approach best suited to their samples. Additionally, dedicated modules for morphometric quantification and visualization with a range of customizable options provide users with the flexibility to adopt the most effective analytical strategies for their specific use cases.

FIJI2, a free and user-friendly platform, supports a wide range of file formats through the Bio-Formats plugin (https://imagej.net/formats/bio-formats), simplifying the handling of raw microscopy data from diverse imaging systems. Additionally, FIJI2 natively supports the measurement of numerous morphometric parameters – including Feret diameter, circularity, pixel intensity, and spatial coordinates – in addition to area, all of which are incorporated into our workflow to enhance flexibility. The following section outlines the key steps of our workflow; further details are provided in the plugin's GUI and accompanying documentation on GitHub (https://github.com/yansong-lu/FishROI.git).

The core technique underlying automated muscle morphometry analysis is image segmentation – a process that distinguishes true muscle fibers from background noise and designates them as ROIs for downstream quantification and visualization. Early segmentation strategies relied on intensity thresholding, where pixels brighter than a user-defined brightness value were classified as signal and the rest as noise. Although conceptually simple, this method requires uniform signal intensity and high signal-to-noise ratios – conditions rarely met in biological specimens. Recent advances in machine learning-based segmentation overcame this challenge by automatically learning a set of rules from multiple features from the image, based on annotation in the training dataset. This allows significantly improved performance, particularly in biomedical imaging and diagnostics. These algorithms generally fall into two broad categories: shallow learning and deep learning. Deep learning models offer superior accuracy but often require extensive annotated datasets, significant computational resources, and technical expertise. Furthermore, they are often highly specific to the species, tissue type, and staining used during training, limiting their utility for non-model organisms and non-standard sample preparations.

Shallow learning approaches, such as those implemented in Labkit, offer a practical alternative. They allow users to train and apply segmentation models directly on their own images using a standard laptop, often within minutes (13). This makes shallow learning especially useful for small datasets or as a means to generate initial annotations for deep learning workflows. However, as deep learning tools have become more accessible, their advantages in robustness and generalizability are increasingly compelling. As demonstrated in Fig. 1F&G, the default Cellpose model (cyto3) already outperforms Labkit for segmenting cytoplasmic staining in larval and juvenile zebrafish muscle, and our custom-trained *rerio* model further improves performance. Notably, Cellpose also provides a user-friendly interface for both segmentation and model training, making it highly accessible for biologists. We therefore recommend Cellpose for processing larger datasets or for generating reusable models, particularly when working with cytoplasmic staining. For smaller projects or for creating training data, Labkit remains a useful and efficient option.

To run Labkit in our workflow, users can click “Segment with Labkit” in the plugin’s GUI. This launches Labkit in FIJI on the selected image channel. After training the classifier and saving the segmentation output, the plugin provides tools to convert the output into ROIs using the “Generate ROI from…” buttons. Note that shallow learning methods typically yield either: simple (binary) segmentations, where pixels are classified as either foreground or background, or probability maps, where pixel intensities represent the likelihood of being part of a signal. Probability maps allow users to adjust probability thresholds post-segmentation, which is especially useful when batch processing images with variable background noise. Because these segmentation types are standard outputs of shallow learning segmentation algorithms, the plugin also supports importing external segmentation maps – produced via alternative software (e.g. ilastik) – and converting them to ROIs via the same “Generate ROI…” interface (18).

If the user wishes to run the deep learning algorithm Cellpose, it is important to note that Cellpose is a standalone Python-based tool and does not operate within FIJI. To support this workflow, we provide all necessary instructions, code, and zebrafish-specific models—accessible via the “Cellpose Instructions” button in the plugin UI. Users will need to execute these commands through a command line interface (CLI) to generate the correct ROI outputs. Once segmentation is complete, the resulting outputs can be imported back into the FIJI plugin for further analysis.

Additionally, for users who wish to train new models in Cellpose, note that Cellpose does not natively accept ROI annotations in the roi.zip format generated by FIJI. To address this limitation, our plugin includes conversion tools that translate FIJI-generated ROI annotations—whether manually drawn or derived from shallow learning algorithms – into a format compatible with Cellpose training. Further details on how to train Cellpose models within our workflow can be found in the user manual available on our GitHub repository (https://github.com/yansong-lu/FishROI.git).

Step 2: manual ROI cleanup

Regardless of the segmentation strategy used, the resulting list of ROIs is rarely perfect. As shown in Fig. 1H&I, even the best-performing models fail to correctly identify some muscle fibers. Manual correction is often necessary to ensure accurate morphometric analysis. To facilitate this step, we have integrated a range of tools that allow users to visualize, manipulate, and manage their ROI sets more effectively. These include a tool that assigns random colors to each fiber, making it easier to detect instances where touching fibers have been incorrectly merged into a single ROI; a bulk ROI removal tool that deletes all ROIs within a user-defined area, avoiding the need to remove them individually; and an autosave function that allows users to periodically save ROI sets and minimize data loss in the event of software crashes – a concern especially relevant when working with large datasets or on lower-specification laptops. Once curated, the ROIs can then be quantified and visualized in the next step of the workflow.

Step 3: Morphometric heatmap generation & quantification

In skeletal muscle, the spatial arrangement of fibers with varying morphometric properties within the myotome can reveal important biological processes. For instance, the localization of smaller fibers may indicate zones of hyperplastic growth or tissue regeneration (Fig. 2A, B). To visualize such patterns, our plugin includes tools that color ROIs based on various morphometric parameters – including area, circularity, and pixel intensity – using a range of preset or custom-designed coloring schemes. Alternatively, users may assign ROIs into a defined number of bins and apply distinct colors to each bin. Regardless of the visualization strategy, a spreadsheet is automatically generated containing the full morphometric profile of each individual ROI (Fig. 2B). These data can then be used for downstream analyses, such as assessing fiber size distributions to study hypertrophy or quantifying fiber numbers to evaluate hyperplasia.

Step 4: Variation heatmap generation & quantification

Morphometric heterogeneity can provide important insights into biological processes such as mosaic hyperplasia. To support this analysis, our workflow includes the tool to quantify the CoV (CoV = standard deviation / mean) among neighboring ROIs and visualize this spatial variation across the myotome (Fig. 2Aiv, Biv). These heatmaps are generated using a sliding window approach, where each pixel is colored based on the CoV between all ROIs within a defined radius (three times the mean Feret diameter) around that pixel. In addition to generating the variation heatmaps for the whole tissue, this tool will also compute the CoV within a user-defined region, allowing quantitative comparison between samples. Note that this step also requires users to run a Julia script outside of FIJI (https://julialang.org/). Our plugin generates the necessary input code by clicking the “Generate Julia Input Code” button. This code must be pasted at the end of the provided Julia script to ensure proper execution on the selected dataset. Additional instructions on installing Julia and running the script are available in our documentation on GitHub (https://github.com/yansong-lu/FishROI.git).

Discussion

In this study, we demonstrate that cytoplasmic staining of muscle fibers is a valuable approach for accurately identifying both small and large fibers in the developing zebrafish myotome. Furthermore, we show that machine learning–based image segmentation techniques – particularly deep learning models such as those implemented in Cellpose – are well-suited for automating the segmentation of zebrafish muscle images. To accommodate the anatomical diversity and variable staining patterns from images across teleost species, we present a flexible workflow that integrates existing machine learning algorithms for muscle image segmentation, followed by novel tools for visualizing and quantifying muscle morphometry and heterogeneity. While this workflow is validated here in zebrafish as a proof of concept, the underlying machine learning strategies are expected to perform robustly across a wide range of teleost and non-teleost species. We hope this workflow serves as a foundation for automated muscle morphometry analysis, allowing users to further expand their analyses using additional tools available within the FIJI platform.

Abbreviations

CLI
command line interface
CoV
coefficient of variation
ECM
extracellular matrix
GUI
graphical user interface
PFA
paraformaldehyde
ROI
region of interest
WGA
wheat germ agglutinin

Declarations

Ethics approval

Ethics clearance was obtained from Monash Animal Services Animal Ethics committee, Monash University (ERM41803, ERM29767).

Consent for publication

Not applicable.

Funding

This work was supported by a NHMRC Investigator Fellowship (2016338) and ARC Discovery Project grant (DP240101647) to PDC and a Medical Research Future Fund grant to AAR and PDC (2015993). MP was supported by a postdoctoral research fellowship from the School of Mathematics and Statistics at the University of Melbourne.

Author Contribution

YL conceptualized the research and wrote the manuscript. YL developed the workflow and FIJI plugin. YL conducted the histological experiments. MP developed the Julia script for heterogeneity analysis. YL and VJ trained the custom deep learning models. YL, PHAR, JL benchmarked alternative workflows and algorithms. AAR and PDC supervised, provided conceptual advice, and contributed to revisions. All authors read and approved the final manuscript.

Acknowledgement

We thank all aquatics technicians at Monash University FishCore for assistance in teleost husbandry. We thank Monash Micro Imaging (MMI) for its aid in microscopy. We thank Dr. Mervyn Dauer for his advice on heatmap designs for visualization.

Data Availability

All plugin, scripts, user manual and all models included in this manuscript is accessible on GitHub (https://github.com/yansong-lu/FishROI.git).

References

1. Danckaert A, Trignol A, Le Loher G, Loubens S, Staels B, Duez H, et al. MuscleJ2: a rebuilding of MuscleJ with new features for high-content analysis of skeletal muscle immunofluorescence slides. Skelet Muscle. 2023;13(1):14. 2023/08/23;.
2. Encarnacion-Rivera L, Foltz S, Hartzell HC, Choo H, Myosoft. An automated muscle histology analysis tool using machine learning algorithm utilizing FIJI/ImageJ software. PLoS ONE. 2020;15(3):e0229041.
3. Wen Y, Murach KA, Vechetti IJ Jr., Fry CS, Vickery C, Peterson CA, et al. MyoVision: software for automated high-content analysis of skeletal muscle immunohistochemistry. J Appl Physiol (1985). 2018;124(1):40–51.
4. Mayeuf-Louchart A, Hardy D, Thorel Q, Roux P, Gueniot L, Briand D, et al. MuscleJ: a high-content analysis method to study skeletal muscle with a new Fiji tool. Skelet Muscle. 2018;8(1):25.

2018/08/06.

5. Mommsen TP. Paradigms of growth in fish. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 2001 2001/06/01/;129(2):207–19.
6. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods. 2021 2021/01/01;18(1):100–6.
7. Wei Y, Li B, Xu H, Liang M. Fish protein hydrolysate in diets of turbot affects muscle fibre morphometry, and the expression of muscle growth-related genes. Aquacult Nutr. 2020;26(5):1780–91. 2020/10/01.
8. Nguyen PD, Gurevich DB, Sonntag C, Hersey L, Alaei S, Nim HT, et al. Muscle Stem Cells Undergo Extensive Clonal Drift during Tissue Growth via Meox1-Mediated Induction of G2 Cell-Cycle Arrest. Cell Stem Cell. 2017;21(1):107–19. e6.
9. Aguiar DH, Bock C, Padovani CR, Pai-Silva MD. MyoD, myogenin and proliferating cell nuclear antigen expression in growing Nile tilapia (Oreochromis niloticus L). Aquaculture Res 2008 2008/11/01;39(15):1673–9.
10. Silva P, Valente LMP, Olmedo M, Galante MH, Monteiro RAF, Rocha E. Hyperplastic and hypertrophic growth of lateral muscle in blackspot seabream Pagellus bogaraveo from hatching to juvenile. J Fish Biol. 2009;74(1):37–53. 2009/01/01;.
11. Brun A, Mougeot G, Denis P, Collin ML, Pouchin P, Montaurier C, et al. A new bio imagery user-friendly tool for automatic morphometry measurement on muscle cell cultures and histological sections. Sci Rep. 2024;14(1):3108. 2024/02/07;.
12. Reyes-Fernandez PC, Periou B, Decrouy X, Relaix F, Authier FJ. Automated image-analysis method for the quantification of fiber morphometry and fiber type population in human skeletal muscle. Skelet Muscle. 2019;9(1):15. 2019/05/27.
13. Arzt M, Deschamps J, Schmied C, Pietzsch T, Schmidt D, Tomancak P et al. LABKIT: Labeling and Segmentation Toolkit for Big Image Data. Frontiers in Computer Science. [Methods]. 2022 2022-February-10;Volume 4–2022.
14. Keenan SR, Currie PD. The Developmental Phases of Zebrafish Myogenesis. J Dev Biol. 2019;7(2).
15. Lee H-T. Cellular and molecular studies of postembryonic muscle fibre recruitment in zebrafish. Danio rerio L.): University of St Andrews; 2010.
16. Biga PR, Goetz FW. Zebrafish and giant danio as models for muscle growth: determinate vs. indeterminate growth as determined by morphometric analysis. Am J Physiology-Regulatory Integr Comp Physiol. 2006;291(5):R1327–37.
17. Perez ÉS, Duran BOS, Zanella BTT, Dal-Pai-Silva M, Review. Understanding fish muscle biology in the indeterminate growth species pacu (Piaractus mesopotamicus). Comp Biochem Physiol A: Mol Integr Physiol. 2023 2023/11/01/;285:111502.
18. Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C et al. ilastik: interactive machine learning for (bio)image analysis. Nat Methods. 2019 2019/12/01;16(12):1226–32.

Figures

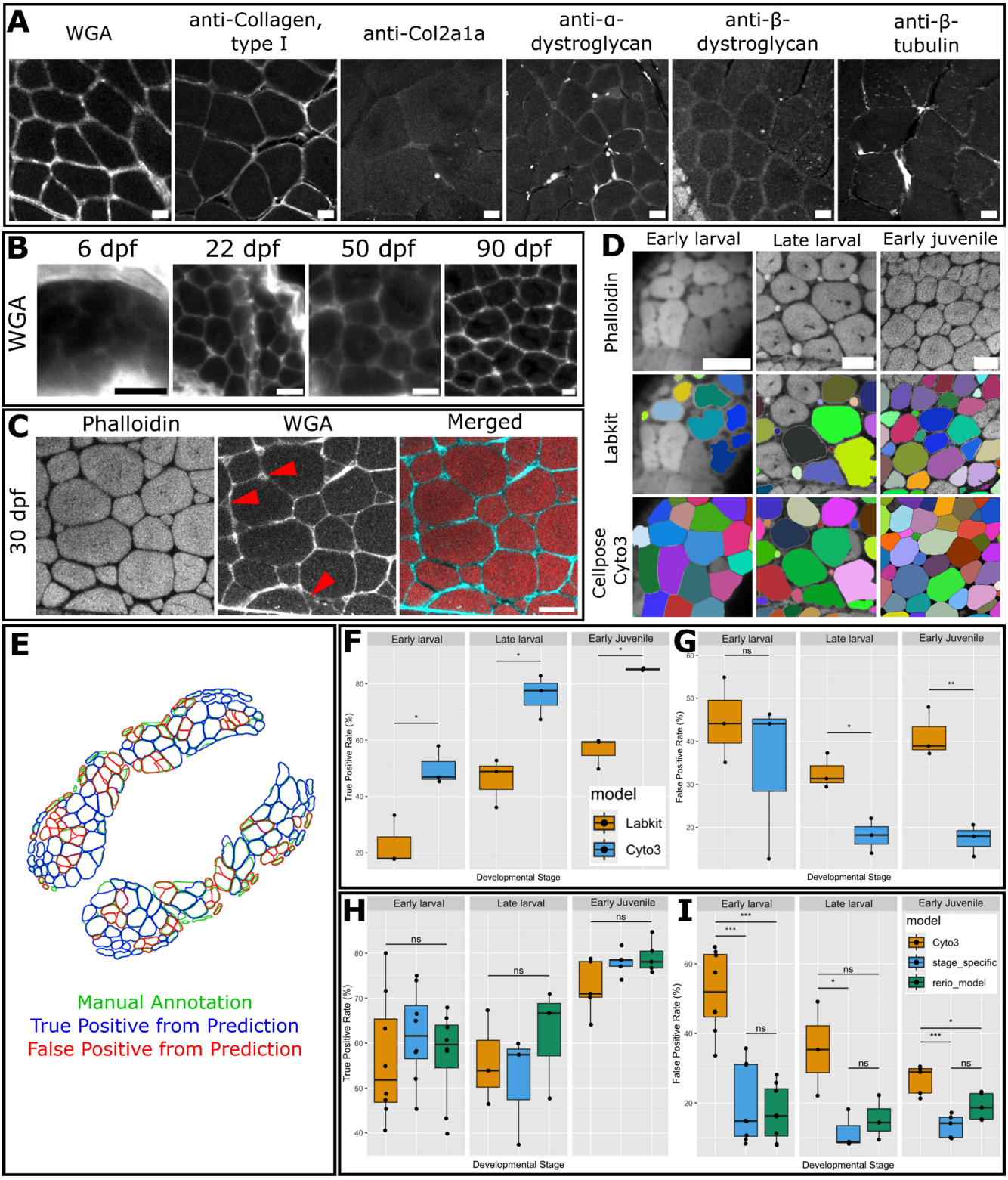


Figure 1

Deep learning segmentation of cytoplasmic staining enables robust muscle morphometry analysis in developing zebrafish. (A) ECM and membrane antibodies or stains often perform suboptimally in teleost muscle, with WGA giving sharpest signals. Images from adult zebrafish cryosections. (B) WGA signal is

markedly reduced in younger zebrafish, hindering segmentation. Images from zebrafish cryosections. (C) Narrow ECM gaps between larger fibers frequently contain small, nascent fibers, but not all such spaces do. These gaps can be difficult to resolve due to low signal-to-noise ratios. Red arrows indicate ECM gaps, with and without nascent fibers. Images from vibratome sections. (D) Phalloidin consistently provides sharp cytoplasmic staining at all developmental stages, but segmentations using the shallow learning algorithm Labkit performed poorly, especially at younger stages, whereas the deep learning algorithm Cellpose (cyto3 model) was better at segmenting fibers based on cytoplasmic signal. Colored ROIs represent segmentation results. (E) To evaluate segmentation accuracy, we overlaid manually annotated ROIs (green) with predicted ROIs. Predicted ROIs that matched the annotated area by more than 65% were considered true positives (blue); others were classified as false positives (red). (F, G) Across all developmental stages, Cellpose cyto3 model consistently outperformed Labkit in correctly identifying fibers. Stats: Student’s t-test. (H, I) A custom Cellpose model trained on zebrafish muscle images (rerio model), as well as stage-specific models, further improved performance by reducing the false positive rate. However, no significant difference was observed between the rerio model and the stage-specific models. Statistics: one-way ANOVA followed by Tukey's post hoc test. Scale bars = 20 μm in all panels.

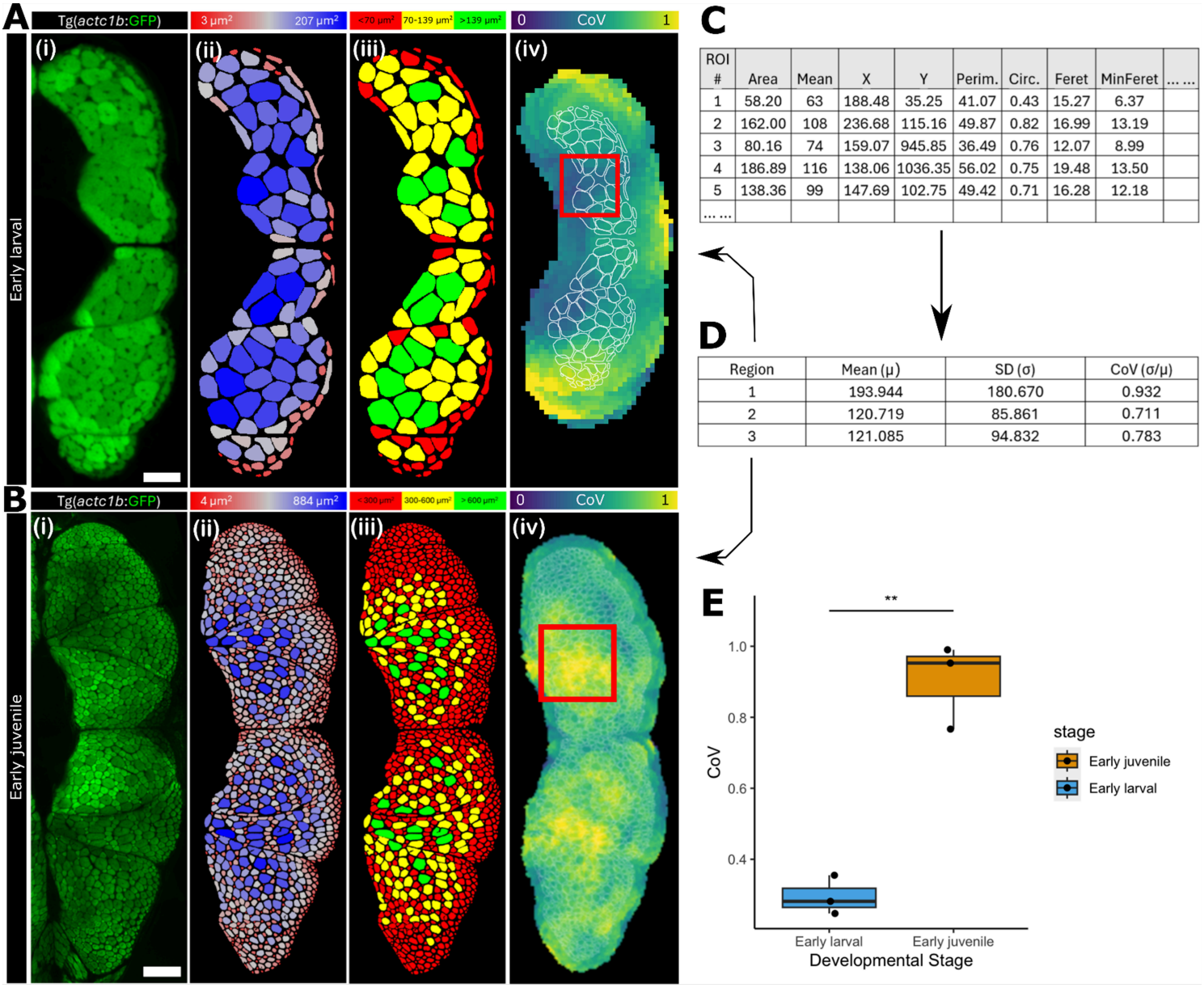


Figure 2

Variations heatmaps offer powerful assessment of morphometric heterogeneity during muscle hyperplasia. (A, B) (i) Confocal images of vibratome-sectioned early larval and juvenile zebrafish showing endogenous GFP signals. (ii, iii) Identified ROIs were pseudo-colored by fiber area to generate heatmaps, facilitating visual identification of small, nascent fibers. (iv) Fiber size heterogeneity was further quantified and visualized by calculating the CoV across the myotome. (C) Heatmaps were generated based on morphometric properties of individual ROIs. (D) These morphometric parameters were used to compute local CoV across the muscle section. (E) By comparing CoV in the deep myotome of larval and juvenile fish, we demonstrate that fiber area CoV is an effective metric for quantifying mosaic hyperplasia in zebrafish.

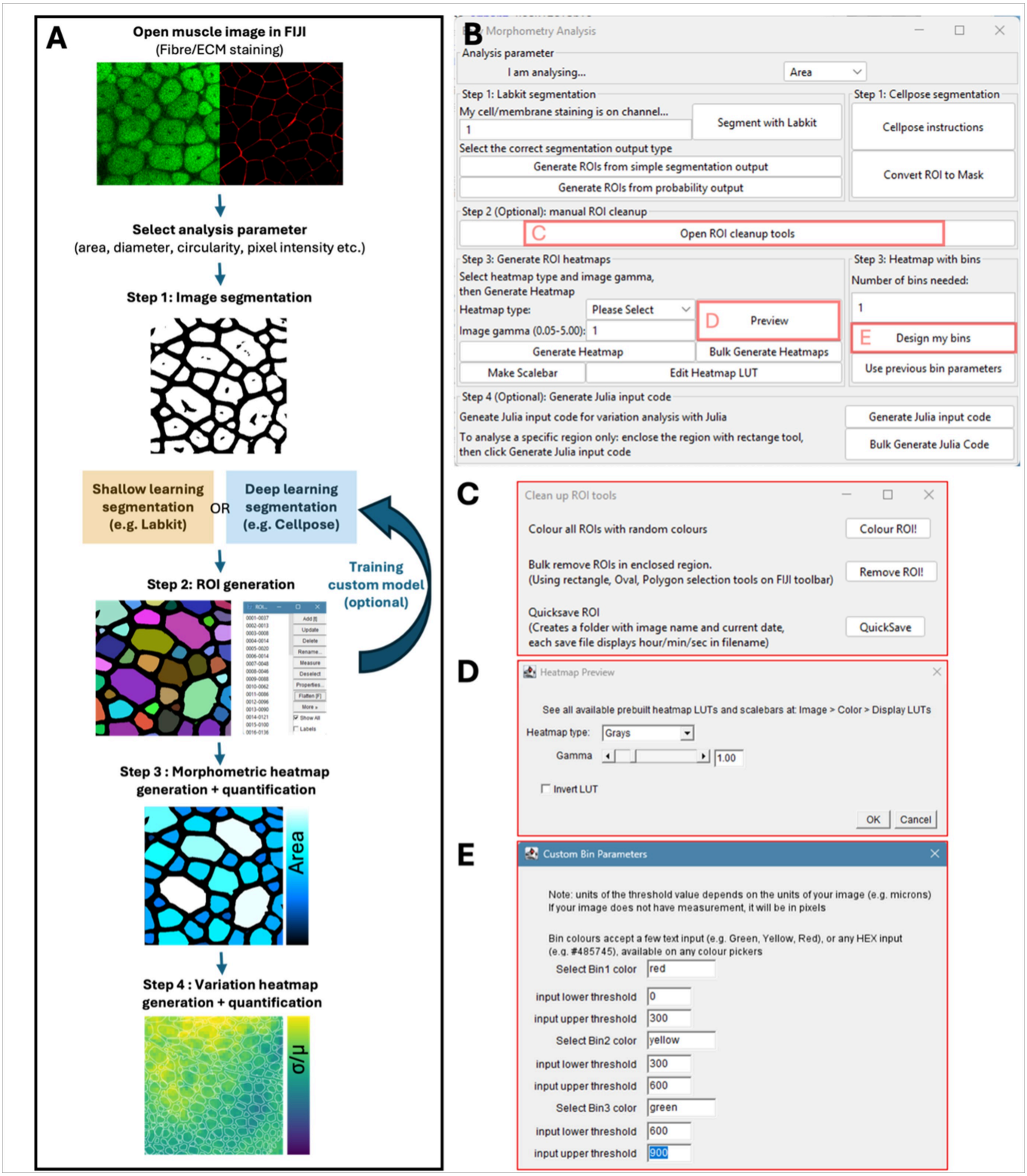


Figure 3

Muscle morphometry analysis workflow & UI. (A) Overview of the analysis pipeline. Our modular workflow supports both shallow and deep learning-based segmentation, enabling flexibility across sample types and staining strategies. We recommend deep learning algorithms for segmenting cytoplasmic staining, using our custom *rerio* model (if working with zebrafish). Alternatively, users could train their own custom models, using either shallow or deep learning algorithms to draft annotation

datasets, which can then be manually corrected. (B) The plugin's main user interface (UI) displays modules arranged in workflow order, each initiated by clicking its corresponding button. Some modules trigger additional dialog windows (indicated by red boxes) or initiate additional instructions to be followed. (C) Selecting "Open ROI cleanup tools" provides access to accessory tools for refining segmentations, including random ROI coloring, bulk deletion of ROIs circled in the image, and a quick-save button for ROIs. (D) The "Preview" function allows users to visualize heatmaps in real time under different color schemes and gamma settings, facilitating optimal visualization for their data. (E) The "Design my bins" function enables users to define custom binning criteria for ROI classification. When "Step 3: Heatmap generation" is executed, a spreadsheet containing all extracted morphometric parameters is automatically generated.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- fishROIS1.docx
- fishROIS2.png