



HAL
open science

A simple smiFISH pipeline to quantify mRNA at the single-cell level in 3D

Mathilde Bouchet, Séverine Urdy, Wenyue Guan, Chérif Kabir, Steve Garvis,
Jonathan Enriquez

► **To cite this version:**

Mathilde Bouchet, Séverine Urdy, Wenyue Guan, Chérif Kabir, Steve Garvis, et al.. A simple smiFISH pipeline to quantify mRNA at the single-cell level in 3D. *STAR Protocols*, 2023, 4 (2), pp.102316. 10.1016/j.xpro.2023.102316 . hal-04262275

HAL Id: hal-04262275

<https://cnrs.hal.science/hal-04262275v1>

Submitted on 28 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

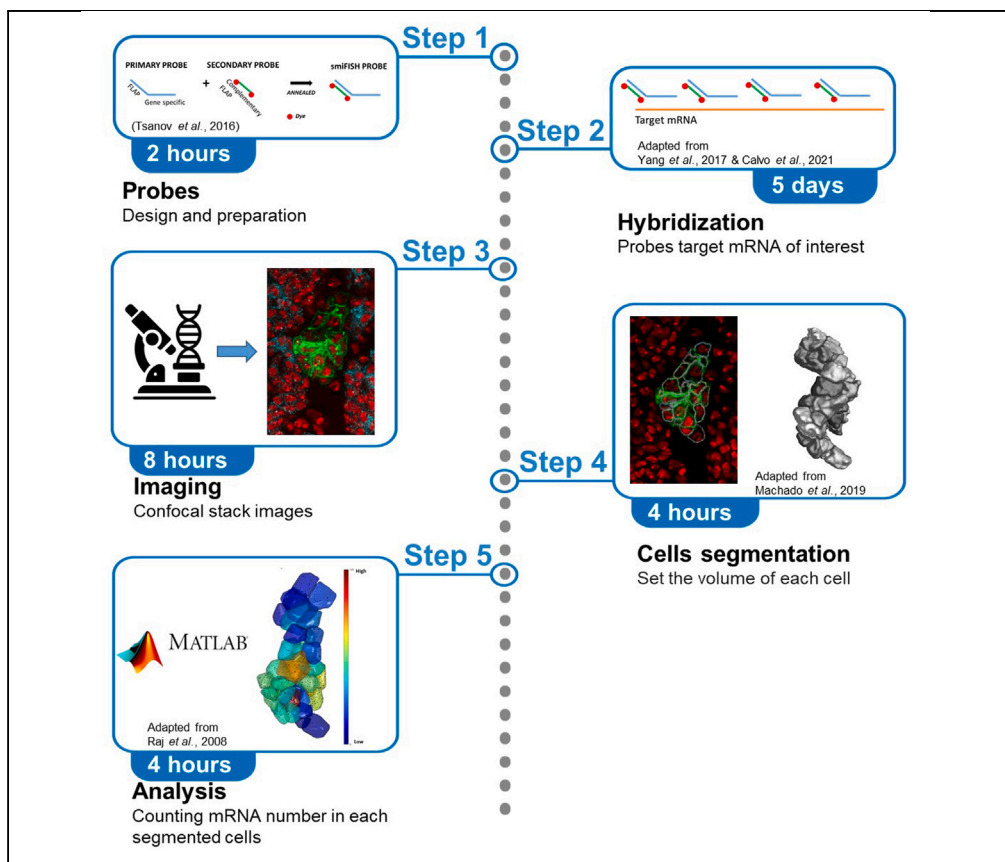
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Protocol

A simple smiFISH pipeline to quantify mRNA at the single-cell level in 3D



Mathilde Bouchet,
S  verine Urdy,
Wenyue Guan,
Ch  rif Kabir, Steve
Garvis, Jonathan
Enriquez

mathilde.bouchet@
ens-lyon.fr (M.B.)
jonathan.enriquez@
ens-lyon.fr (J.E.)

Highlights

The pipeline allows
for the quantification
in 3D of mRNA
labeled by smiFISH

The method
combines cell
segmentation and
mRNA quantification

The number of RNA
spots is accurately
determined only in
the segmented cells

The method uses a
low usage of
computing power

Techniques allowing the precise quantification of mRNA at the cellular level are essential for understanding biological processes. Here, we present a semi-automated smiFISH (single-molecule inexpensive FISH) pipeline enabling quantification of mRNA in a small number of cells (~40) in fixed whole mount tissue. We describe steps for sample preparation, hybridization, image acquisition, cell segmentation, and mRNA quantification. Although the protocol was developed in *Drosophila*, it can be optimized for use in other organisms.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Bouchet et al., STAR Protocols
4, 102316
June 16, 2023    2023 The
Author(s).
[https://doi.org/10.1016/
j.xpro.2023.102316](https://doi.org/10.1016/j.xpro.2023.102316)



Protocol

A simple smiFISH pipeline to quantify mRNA at the single-cell level in 3D

Mathilde Bouchet,^{1,4,*} Séverine Urdy,¹ Wenyue Guan,^{1,3} Chérif Kabir,¹ Steve Garvis,² and Jonathan Enriquez^{1,5,*}

¹Institut de Génomique Fonctionnelle de Lyon, ENS de Lyon, CNRS, University Lyon 1, 46 Allée d'Italie Cedex 07, 69364 Lyon, France

²Laboratoire de Biologie Moléculaire de la Cellule ENS de Lyon, CNRS, University Lyon 1, HCL, 46 Allée d'Italie cedex 07, 69364 Lyon, France

³Present address: Changping Laboratory, Beijing, China

⁴Technical contact

⁵Lead contact

*Correspondence: mathilde.bouchet@ens-lyon.fr (M.B.), jonathan.enriquez@ens-lyon.fr (J.E.)
<https://doi.org/10.1016/j.xpro.2023.102316>

SUMMARY

Techniques allowing the precise quantification of mRNA at the cellular level are essential for understanding biological processes. Here, we present a semi-automated smiFISH (single-molecule inexpensive FISH) pipeline enabling quantification of mRNA in a small number of cells (~40) in fixed whole mount tissue. We describe steps for sample preparation, hybridization, image acquisition, cell segmentation, and mRNA quantification. Although the protocol was developed in *Drosophila*, it can be optimized for use in other organisms.

For complete details on the use and execution of this protocol, please refer to Guan et al.¹

BEFORE YOU BEGIN

The protocol below describes the specific steps to quantify mRNA during development in a neuronal lineage called Lin A/15^{2,3} (Figure 1). The mRNA quantification is performed in the LinA/15 neuronal stem cell (called neuroblast) and its motoneuronal progeny. However, this protocol can be used to quantify mRNA in any cell type in any organism but requires that the cell's cytoplasmic membrane is labelled with an endogenously expressed fluorescent protein or by immunochemistry.

Probe design and preparation

⌚ Timing: 2 h

The first step of smiFISH is to design primary probes that target the genes of interest. The primary probes are flanked by Flap sequences that are recognized by secondary probes labeled by two dyes in their 5' and 3' regions. These secondary probes are used like 'secondary antibodies' in immunochemistry (Figure 2).

1. Primary Probes (Figure 2).
 - a. Design 20 nucleotide (nt) primary probes against mRNA sequences of genes of interest.

Note: Probes should target common sequence of all isoforms of genes of interest, up to 48 probes per gene. Use the Biosearch Technologies stellaris RNA FISH probe designer tool (free with registration, <https://biosearchtech.com>) to design the probes.



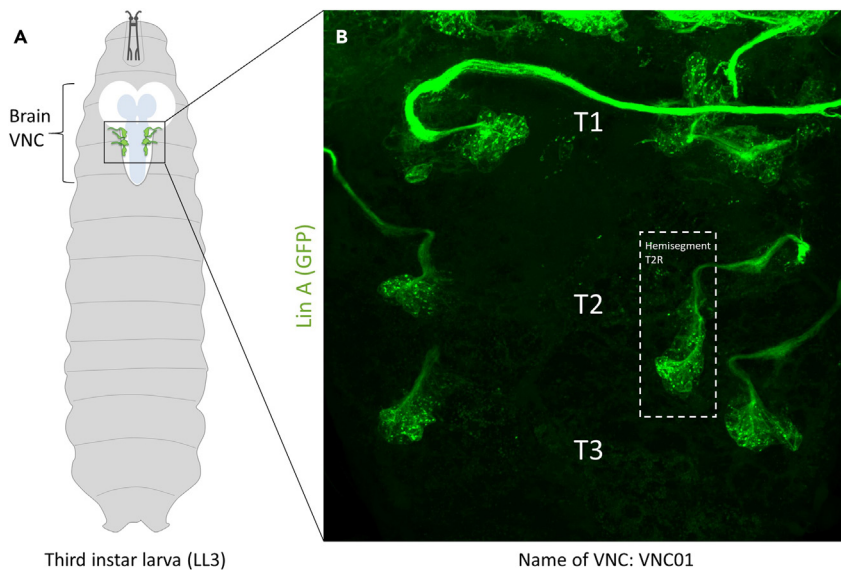


Figure 1. Membrane labeling of Lin A/15

(A) Drawing of an L3 larva showing the CNS and Lin A/15.

(B) Maximum projection of confocal images of the six thoracic hemisegments in an L3 larva where Lin A/15 are labeled with myr:GFP (green). The dotted box indicates a T2R LinA/15 chosen for mRNA quantification in this article. Field of view : 223 × 223 μm.

Note: Only probes with GC content >40% are used.

- b. Add the X Flap sequence (sequence recognized by labeled secondary probes) to the 5' end of each 20 nt primary probe.

Note: X Flap sequence: CCTCCTAAGTTTCGAGCTGGACTCAGTG (Figure 2).⁴

- c. Prepare an equimolar mixture of primary probes flanked by the X Flap sequence.

Note: The 5'-extended primary probe sets are synthesized in 96 well plates by Integrated DNA Technologies (IDT : www.idtdna.com), using 25 nmol synthesis scale, standard desalting, and at 100 μM in IDTE Buffer pH 8.0 (10 mM Tris-HCl/0.1 mM EDTA).

Note: This equimolar mixture of primary probes can be store at -20°C.

2. Synthesize secondary probes labelled with a dye in the 5' and 3' regions. The secondary probes are the reverse complement of the X flap sequence: CACTGAGTCCAGCTCGAACTTAGGAGG.

Note: We end-labeled the 5' and 3' regions with Quasar 570. Here you can use your favorite dyes however in our hands Quasar 570 gave the best signal.

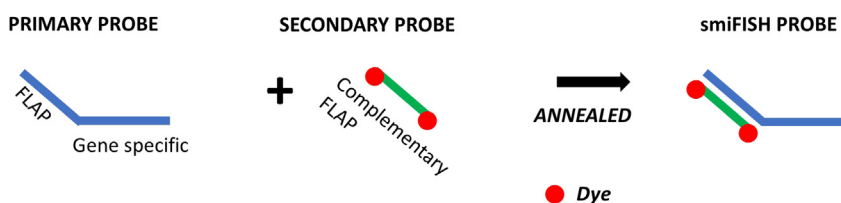


Figure 2. Schematic of the annealing reaction

Primary probes are pre-hybridized *in vitro* with the secondary probe via the FLAP sequence.^{4,5}

Note: The secondary probes were synthesized by Biosearch Technologies (purification Dual-HPLC). The fluorescent FLAP is delivered lyophilized. Resuspend the FLAP in TE buffer at final concentration of 100 μ M and store at -20° C.

Note: Flap annealing reaction.

Prepare Fluorophore-labeled probe described in Tsanov et al.,⁴ by hybridizing the fluorophore-labeled X Flap sequence and 5'-extended primary probe sets. The final concentration of the annealing reaction is 20 μ M. In short the reaction mix is:

Probe-mix: 2 μ l (200pmol).

100 μ M Flap: 2,5 μ l (250pmol).

10XNEB(buffer 3): 1 μ l.

H₂O: 4,5 μ l.

Incubate reaction mix in a thermocycler:

1 cycle: 85 $^{\circ}$ C 3min.

1 cycle: 65 $^{\circ}$ C 3min.

1 cycle: 25 $^{\circ}$ C 5min.

Lid : 99 $^{\circ}$ C.

Put sample tubes on ice. For long-term storage, keep hybridized duplexes at -20° C.

Preparation of research materials

⌚ Timing: 1 week

3. Prepare all the critical reagents mentioned in the [key resources table](#).
4. Prepare your fly stock, or any other animal model stock, with your selected genetic profile.

Note: In our study, we labeled the cytoplasmic membrane of a stem cell and its neuronal progeny, during development. The lineage produce by this stem cell is called Lin A/15 and produces 29 adult motoneurons (MNs).^{2,3,6} The following genotypes were used: *Dpn>KDRT>Cre//; act>LoxP > LexA, LexA-myr::GFP//; UAS-Kd//* crossed to *R10c12-GAL4*.^{7,8}

5. Use a hybridization oven allowing temperature control with rocking option.

Note: in our protocol, the smiFISH experiments were performed with Big Shot III Hybridization Oven (230402).

6. Use a confocal microscope to perform high spatial resolution imaging.

Note: in our protocol confocal stacks were taken with a confocal microscope (Leica, SP8) using a 40 \times objective (1.4 NA).

△ **CRITICAL:** One essential step before performing the smiFISH experiment is to label specifically the cytoplasmic membrane without labeling the nuclear membrane. The cytoplasmic membrane GFP signal is used for single cell segmentation. While we use genetically encoded markers to label the cytoplasmic membrane, any other technique such as membrane immunostaining can be used to label the cytoplasmic membrane.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Dextran sulfate 50%	Sigma	S4030
Formaldehyde 16%	Thermo Scientific	Cat#28906
Formamide	Sigma	Cat#F9037
NEB Buffer 3	NEB	B7003S
PBS	Dutscher	Cat#X0515-500
Schneider's medium	Thermo Fisher	21720024
SSC 20X	Sigma	S6639-1L
Triton-X100	Sigma	Cat#T8787
Tween20	Sigma	Cat#P7949
SuperFrost Plus slides	Fischer Scientific	J1800AMNZ
Coverslip #1,5	Fischer Scientific	15787572
Vectashield mounting medium	Vector Laboratories	Cat#H1000
Vectashield Vibrance Antifade Mounting Medium	Vector Laboratories	Cat#H1700
Vectashield Antifade Mounting Medium with DAPI	Vector Laboratories	Cat#H1200
Pyrex plates	Euromedex	Cat# EM-71560-06
Experimental models: Organisms/strains		
R10c12-GAL4 (3rd chromosome, attP2)	Lacin and Truman ⁸	N/A
UAD-KD (attP2[68A4])	Awasaki et al. ⁷	N/A
dpn>KDRT-stop-KDRT>CRE (su(Hw)attP8 [8E10])	Awasaki et al. ⁷	N/A
act>loxP-stop-loxP > LexA::P65 (attP40 [25C7])	Lacin and Truman ⁸	Tzumin Lee
Oligonucleotides		
Reverse complement of the X flap sequence	Tsanov et al. ⁴	CCTCCTAAGTTTCGAGCTGGACTCAGTG
Software and algorithms		
ImageJ plugin Limeseg	Machado et al. ⁹	https://imagej.net/LimeSeg
MATLAB R2020	SCR_001622	http://www.mathworks.com/
smFish analysis source code	This study	https://doi.org/10.5281/zenodo.7674721
Remap source code	This study	https://doi.org/10.5281/zenodo.7456891
List of smiFISH probes	This study	https://doi.org/10.5281/zenodo.7674721

MATERIALS AND EQUIPMENT

The two buffers following must be prepared on the same day of the experiment and cannot be stored.

smiFISH Wash Buffer¹⁰

Reagent	Final concentration	Amount
Formamide	10%	1 mL
Saline Sodium Citrate (SSC) 20X	2X	1 mL
DEPC H ₂ O	N/A	8 mL
Total	N/A	10 mL

smiFISH Hybridization Buffer¹⁰

Reagent	Final concentration	Amount
Formamide	10%	1 mL
SSC 20X	2X	1 mL

(Continued on next page)

Continued

Reagent	Final concentration	Amount
Dextran Sulfate 50%	5%	1 mL
smiFISH probes	between 0,4μM–2μM	
DEPC H ₂ O	N/A	Qsp
Total	N/A	10 mL

Note: Freshly prepared 2 X SSC solution is essential for achieving a good signal-to-noise ratio of the smiFISH experiments.

Note: 20x SSC solution and Dextran Sulfate 50% can be stored at room temperature (between 20°C–25°C). Formamide can be stored at –80°C. The smiFISH probes can be stored at –20°C.

△ CRITICAL: Formamide is carcinogenic necessitating the use of protective clothing and fume hood.

- 4% formaldehyde: Mix 1 mL lightbulb of formaldehyde 16%, 400 μL of PBS 10X, 12 μL of Triton-X100 and add 2,6 mL MilliQ H₂O to prepare 4 mL buffer.

Note on storage conditions: keep on ice and protected from light, maximum 4 h.

- 0,3% PBST: Mix 10 mL of PBS 10X, 300 μL Tween-20 and add 90 mL MilliQ H₂O to prepare 100 mL buffer.

Note on storage conditions: keep at room temperature, maximum 6 months.

STEP-BY-STEP METHOD DETAILS

Sample dissection and fixation

⌚ Timing: 2 h

This protocol has been developed in *Drosophila* to quantify mRNA in Lin A/15 during the third larval instar (LL3). This neuronal lineage is localized in the ventral nerve cord (VNC) (Figure 1).

Note: All the steps of the protocol are done in pyrex nine well plates. The volume of washes is approximately 150 μL. The plate is covered with parafilm to prevent drying.

△ CRITICAL: Our tissue is optically transparent enough for microscopy. Other tissues might need an optical clarification protocol.

1. Dissect Larval VNC from third instar larvae in Schneider medium (or appropriate cell culture medium).
2. Fix in 4% formaldehyde (in PBS with 0.3% Triton X-100) for 20 min at room temperature with rocking.
3. Quickly wash the samples 3 times with PBST (PBS with 0.3% Tween-20).
4. Wash samples 3 times for 15 min with PBST.

Note: Samples can be stored in PBST at 4°C for a few h before the hybridization steps.

SmiFISH

⌚ Timing: 5 days

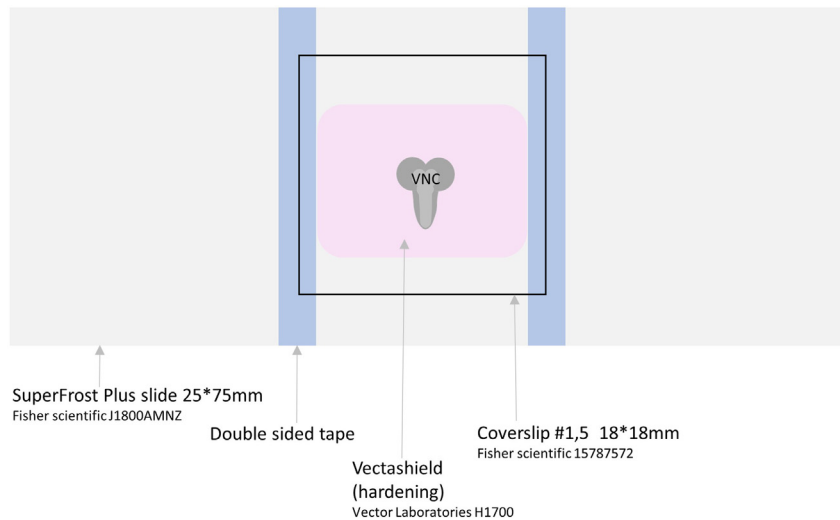


Figure 3. Schematic of the mounting setup

VNC larvae is immersed in Vectashield (pink) on a slide. A bridge is made by placing double sided tapes (blue) on each side of the VNC. A coverslip is placed of the top of the VNC (black box).

The smiFISH procedure is adapted from Yang et al.¹⁰ and Calvo et al.⁵ All hybridization steps are performed in Big Shot III Hybridization Oven with rocking.

5. Hybridization steps – Day 1.

- a. Perform a pre-hybridization wash in smiFISH wash buffer at 37°C for 30 min with rocking.
- b. Incubate the samples with smiFISH probe sets (Figure 2) diluted in smiFISH hybridization buffer at 37°C for 12–14 h with rocking.

Note: The working concentration of individual probe sets are: jim (400 nM), chinmo (800 nM), brC (1 μM), nvy (2 μM), mammo (2 μM) and oli (400 nM). Sequences are available at <https://doi.org/10.5281/zenodo.7674721>. These concentrations were established by performing a dilution series for each probe set in order to choose the best one. Cross-reactions between probes may occur (see [troubleshooting](#)).

6. Washes – Day 2

- a. Wash the samples for 40 min at 37°C in smiFISH wash buffer with rocking.
- b. Perform 3 additional washes of 15 min at room temperature with smiFISH wash buffer with rocking.
- c. Wash 10 min in PBST at room temperature with rocking.

Note: Store samples for 2 days minimum in pyrex plates at 4°C in Vectashield (H1200) with Dapi to label the nucleus.

△ **CRITICAL:** 2 days is necessary to allow Dapi to penetrate the tissue.

7. Mounting and storage (Figure 3).

- a. Place Larval VNC in the center of the slide (SuperFrost Plus 25mm*75mm) with double-sided tape at each side.
- b. Replace Vectashield by 25μL of Vectashield (H1700) to allow the sample to be better contained.
- c. Place a coverslip (#1,5) over the sample (18mm*18mm). After 1 day at 4°C, seal the coverslip with nail polish.

- d. Store the slides at -20°C to maintain the signal.

Note: Slides can be stored for several weeks.

Image acquisition of larval VNC

⌚ Timing: 8 h

At this step use a confocal microscope and acquire images with the optimal optical resolution of your system.

8. Acquire images on a Leica TCS SP8 microscope using a 40x/1.4 glycerol objective.

Note: this results in bright and high-resolution images due to the well-matched refractive index with the Vectashield.

- a. Settings for image stacks:
 - i. format 4400x4400 px (223,56 μm * 223,56 μm (XY) and 44,95 μm (Z)),
 - ii. speed 700 Hz,
 - iii. unidirectional,
 - iv. sequential line scanning,
 - v. line averaging 8,
 - vi. pinhole 1 airy unit.
 - vii. Image stacks are acquired with a 350 nm z interval.
- b. Use a 405nm laser (laser power 2%, detection 412–496nm) to excite and image DAPI, in the cell nucleus.
- c. Use a 488nm laser (laser power 4.5%, detection 494–530nm) to excite and image the cell membrane labeled with GFP.
- d. Use a 548nm (laser power 5%, detection 555–605nm) to excite and image the mRNA foci labeled with Quasar-570.

⚠ CRITICAL: Do not use too much power of the 548 laser to avoid signal bleaching.

Post-imaging processing: Lightning

⌚ Timing: 1 h

A step of post-imaging processing (deconvolution) is necessary to maximize detection and quantification of signal in MATLAB. Deconvolution is an image processing technique that attempts to improve contrast and resolution, minimizing the effects of any –out-of-focus light in fluorescence images. It uses the metadata of the acquisitions (objective, immersion, wavelengths, etc.) to characterize these effects and minimize them, proceeding with successive iterations in order to optimize image quality. Here, we use the Lightning option of the Leica software.

9. Select adaptive mode in the Strategy box.
 - a. Chose the refractive Index of mounting medium: 1,46 Vectashield.

Note: With the adaptive function, the deconvolution parameters automatically vary according to the local XYZ quality of the fluorescent signal (on small areas of a few XYZ voxels) channel by channel. In this way, the system deconvolutes the areas of good or poor quality in an optimal way. Use the default deconvolution settings:

- b. Number of Iterations: Auto.

Note: number of times that the deconvolution process is repeated.

c. Optimization: 1.

Note: The software deconvolves by default with the slider on 1, which corresponds to a fairly conservative deconvolution without risk of artifacts.

d. Regularization Parameter: 0.05.

Note: This number correspond to the degree of denoising of the image during deconvolution.

e. Smoothing: Very high.

Note: smoothing option before deconvolution.

f. Contrast enhancement, Cut Off and deblurring are disabled.

Note: After this process, one new confocal stack is created with the lightning correction.

Note: Other methods of deconvolution can be used at this step. However, we found the Lightning extremely powerful compared to other methods.

△ CRITICAL: LAS_X_4.5.0 or higher version of the Leica software must be used. Older versions cannot correctly process the large data sets which are created.

10. Save the processed confocal stack.

△ CRITICAL: Name the processed confocal stack as followed: *Stage_Transcriptionfactor_VNC*.

In our experiment, we named it: *LL3_Jim_VNC01*. We performed our experiment in a third instar larva (*LL3*), we analysed the mRNA expression of *jim* (*Jim*) in a VNC called 01 (*VNC01*).

Save the processed confocal stack in the following path:

C:\Count_mRNA_Urdy\Stage_Transcriptionfactor\VNC.

Note: For example in our study we wrote:

C:\Count_mRNA_Urdy\LL3_Jim\VNC01.

△ CRITICAL: Name the path exactly as written above since the MATLAB script will use this path to quantify the mRNA expression. This protocol was developed using Windows 10.

Pipeline for cell segmentation with limeseg: A Fiji plugin

⌚ **Timing:** 4 h

We segmented Lin A/15 confocal stacks using the Limeseg plugin⁹ in ImageJ/Fiji software.^{11–13} The Limeseg plugin uses a particle-based active contour method allowing 3D object segmentation on objects labelled on their outline from image stacks, for example confocal stacks containing cells labeled on their membrane (Figure 4). The Limeseg plugin does not require a powerful PC. We could run it on a Windows 10, Intel Core i5-6600 CPU 3.31GHz, 16Go RAM.

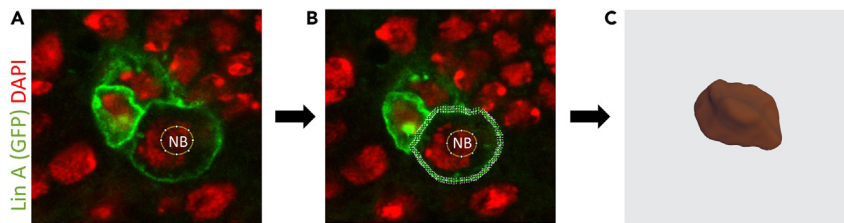


Figure 4. Schematic of the Limeseg pipeline

(A and B) Confocal section of the LinA/15 neuroblast (NB) labelled with a membrane GFP (*myr::GFP*). (A) A ROI is placed in the center of the cell. (B) After running Limeseg, the border of the NB is marked by dots (green) and (C) a 3D reconstruction of the NB is generated.

11. Identify the cells of interest with the Cell Counter plugin of the ImageJ/Fiji software^{11–13} (Figure 5B).

Note: The Cell counter plugin in Fiji allows the counting of the number of cells present in Lin A/15 and their localizations.

- a. Open the confocal image stack and the Cell Counter plugin.

Note: Plugins\Analyze\Cell Counter\Cell Counter

- b. Go through the confocal stacks placing a number in the center of each cell.

Note: the GFP and DAPI signal are used to find the center of the cell. At this step, the coordinates of each point are managed in the measurement tab.

- c. Save the cell counter results table.

△ CRITICAL: To allow a good synchronization between the image stack and the cell counter when they will open again, do not modify the name proposed by the cell counter.

12. Smooth the GFP signal with a median filter of the ImageJ/Fiji software^{11–13} (optional).

Note: The Median filter of ImageJ will smooth discontinuity of the contours of the cell membranes. This step helps during the cell segmentation which is based on membrane signal.

- a. Open the original confocal stack (not processed with the Lightning) with the Fiji software.
- b. Apply a median filter only on the channel with the membrane signal: *Process\Filter\Median*.

Note: We used the GFP signal that was not processed with the Lightning Settings of the Leica software.

13. Place a ROI (region of interest) on each cell with the ROI manager of the ImageJ/Fiji software.^{11–13}
 - a. Open two windows (/ images).
 - i. the GFP membrane smoothing and DAPI signals (composite stack).
 - ii. the cell counter stack.
 - b. Place an ROI on the center of each cell in the composite stack (GFP membrane smoothing and DAPI signals).

Note: Use the cell counter stack as a reference in order to not forget any cell or to avoid placing two ROI in the same cell (Figure 5B and 5C).

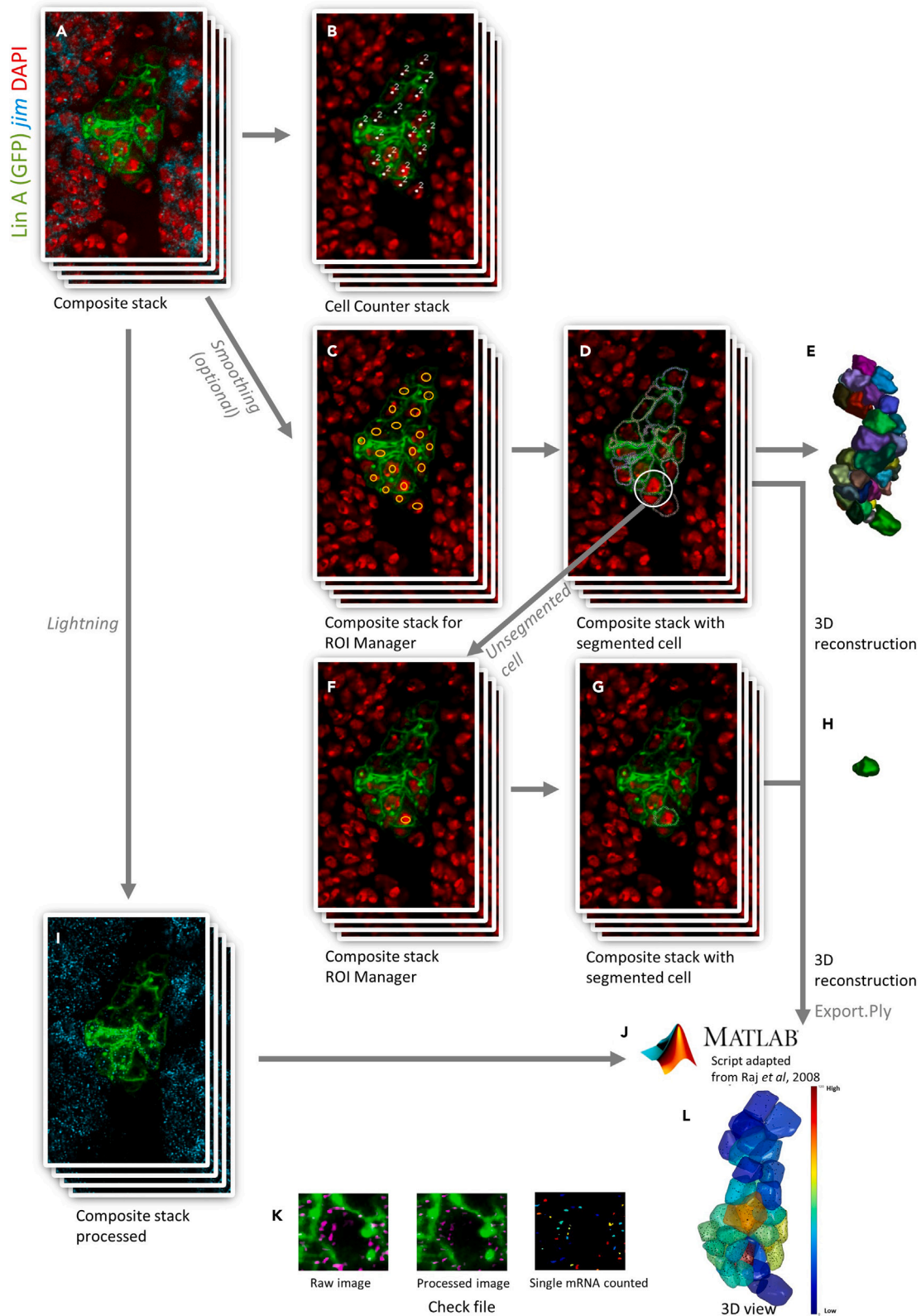


Figure 5. Schematic of the smiFISH Pipeline

- (A) Confocal stack images (composite stack) where the cell membranes are labelled with GFP (green), nucleus with DAPI (red) and *jim* mRNA dots with Quasar 570 (cyan).
 (B) Same confocal stack as in (A), here called Cell counter stack, where each cell is identified by a number (2).
 (C) Composite stack as in (A) processed by a median filter. ROI are placed in each cell using the cell counter stack (B) as reference.
 (D) Composite stack as in (A) processed by a median filter. ROI are placed in each cell using the cell counter stack (B) as reference.
 (E) Composite stack as in (C) after running Limeseg where the border of the cells are marked by dots.
 (F) 3D reconstruction of Lin A/15 cells after running Limeseg.
 (F–H) Same picture as (C–E) respectively where unsegmented cells have been processed individually.
 (I) Confocal stack images (composite stack) from (A) processed with Lightning.
 (J) The composite stack processed with Lightning (I) and the ply files generated with Limeseg (C–H) are imported in MATLAB.
 (K) Check files are generated by MATLAB for each cell, here is the example of one of them. left: original image, center: image after Gaussian filter, and right: mRNA dots detected color-labelled.
 (L) MATLAB 3D view showing *jim* mRNA expression (dots), the border of each cell is color-coded according to the mRNA expression level (the color code is indicated on the right).

- c. Save all the coordinates (X,Y,Z from the center of the circle), More/Save.
14. Segment each cell using the Limeseg⁹ plugin of the ImageJ/Fiji software.^{11–13}
 - a. Open the Sphere Seg : Pluggin\Limeseg\Sphere Seg.
 - b. Choose the parameters according to the explanations: (<https://imagej.net/plugins/limeseg>).

Note: The images raw data are available in Metadata in ImageJ/Fiji.

Note: for our study¹, the chosen parameters are :

D_0 ~4–6, the smallest feature size that you want LimeSeg to detect (in pixels); F pressure = 0.01, the "default pressure" exerted on the surface; Z_scale = 6.8, the physical ratio of the spacing; Range in d0 unit ~4–6, the size over which each particle will look for a local maximum (in pixels); Optimisation step = -1, number of integration step; real XY pixel size = 50, the size of a XY pixel in physical units.

- c. Click on OK to run *Sphere Seg*, and wait for the 3D segmentation ([Methods video S1](#)).

Note: After clicking ok there are two possibilities.

- i. The segmentation stops by itself as it has completed the job. In this case, go through the stack and verify that the cell segmentation is correct ([Figure 5D](#)).
- ii. The segmentation is unable to finish the job because some cells are not correctly segmented ([Figure 5D](#)). In this case, go to Pluggin\Limeseg\Stop optimization. Change the parameters (see <https://imagej.net/plugins/limeseg>) and start the process again until you find the optimal parameters. Users may find it helpful to increase the signal intensity of the GFP channel. After this step, most of the cells are well segmented, though sometimes unsegmented cells are still present. In this case, remove the ROI of the unsegmented cells and identify them on a new duplicated stack with cell counter. Run Limeseg again without these unsegmented cells and save the data (see next paragraph for saving procedure). Then run Limeseg again with different parameters on the individual cells that were not segmented (do it one by one) ([Figure 5F–5H](#)).
- d. Save the data.
 - i. Save the table: *File*\save as.

Note: Once the segmentation is performed on multiple cells or on individual cells, a results table will be created. This table contains the coordinates and the volumes of segmented cells.

- ii. Save also the .ply files: *Plugin\Limeseg\show GUI*. In the 'write' box of the 'show GUI' window write where to save the .ply file and click *SaveStateToXmlPly*.

Note: A folder containing a .ply file and a 3D file of the cell is created and saved for each cell. These files will be used in MATLAB to quantify mRNA in each cell.

△ **CRITICAL:** As written above, sometimes several rounds of segmentation are needed to segment all the cells of a lineage A. In this case, compile all the folders containing the .ply file in the same folder. Before compiling all the .ply folders, it is essential that each folder is manually renamed as followed: *'Developmental time_ gene name _ sample name _localisation_cell_number of the cell'*

Note: For example, in our study, we wrote: LL3_jim_VNC01_T2R_cell_1. In this experiment, we were observing a third instar larva (LL3), we have analysed the mRNA expression of jim (jim) in a VNC called 01 (VNC01) in the Lin A/15 localized in the right thoracic hemisegment number 2 (T2R) (Figure 1). The number of the cell is arbitrary. In our study, these numbers vary from 35 to 45 since Lin A/15 contained 37 cells on average at this stage.

- iii. Export all the folders in the following path: C:\Count_mRNA_Urdy\Stage_Transcriptionfactor\VNC\Hemisegment\Limeseg.

Note: In our example the path is: C:\Count_mRNA_Urdy\LL3_Jim\VNC01\T2R\Limeseg.

△ **CRITICAL:** Rename each folder containing the .ply files before putting them in the Limeseg folder. If you don't rename them at this step you might overwrite a folder since they can have the same name.

Note: Renaming manually the folders can be -time-consuming. Optionally you can use our Python script: <https://doi.org/10.5281/zenodo.7456891> . Before using it download and install Python 3: <https://www.python.org/downloads/release/python-3111/>.

The Script will rename the folders containing the .ply file and pool them in a single Limeseg folder.

MATLAB

⌚ Timing: 4 h

MATLAB (Figure 5) is used to quantify mRNA in each Lin A/15 Cell. Two types of files are imported in MATLAB: 1/ the .ply files containing the segmented cells and their coordinates and 2/ the composite stack with the smiFISH signal processed with the Lightning. These stacks are cropped around each cell using the point clouds generated by individual cell segmentation with LimeSeg.

The mRNA spots are detected in 3D, in the mRNA channel of these cropped stacks, using the method described by Raj et al.¹⁴ In short, the spots are identified computationally by running a MATLAB image processing script that runs the raw data through a filter, a Laplacian of a Gaussian, a mathematical operator commonly used to enhance spots of the correct size and shape while removing the varying background. On the filtered stacks an optimal threshold is used to remove remaining background noise. In order to choose an optimal threshold, all possible thresholds are computed. The thresholds are always chosen manually and close to the plateau after the first peak (Figure 6).¹⁴

Note: At this stage of the protocol, the processed confocal stack and the .ply folders should be located respectively as follows:

C:\Count_mRNA_Urdy\LL3_Jim\VNC01\LL3_Jim_VNC01.

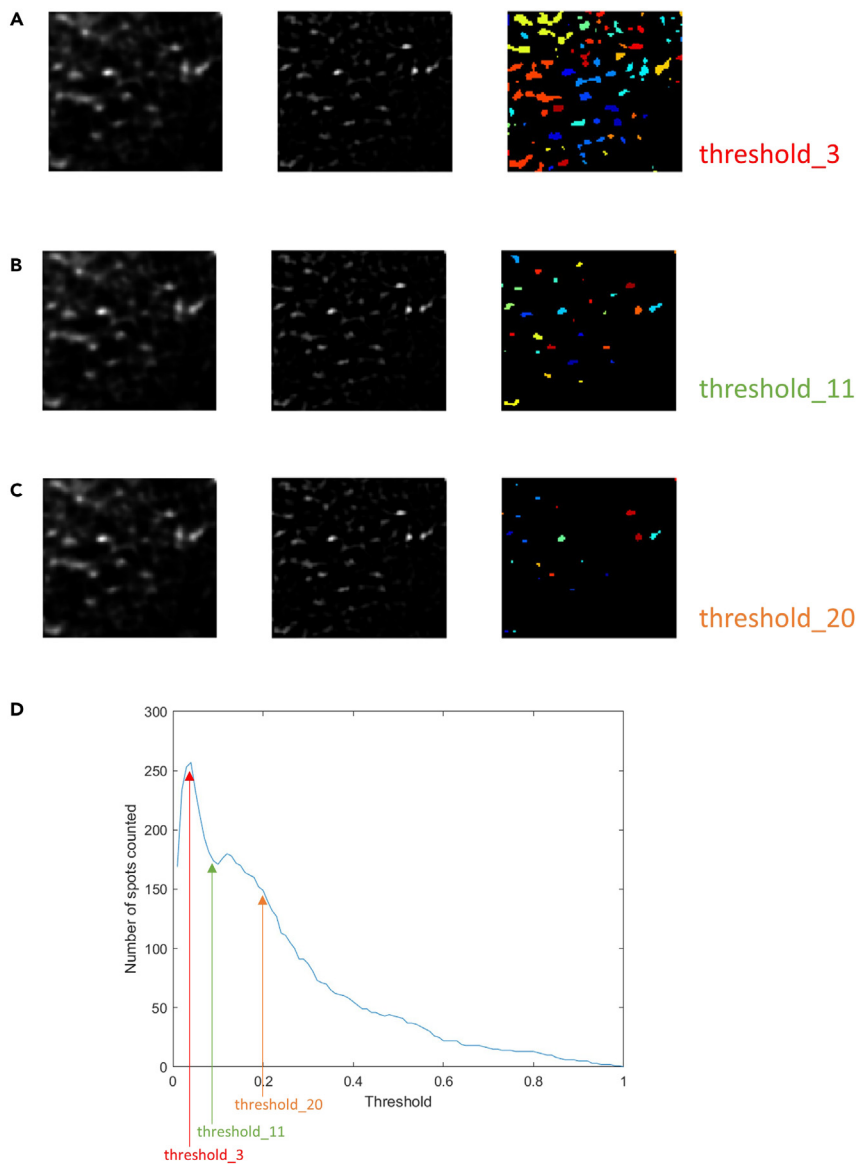


Figure 6. Threshold decision

(A–C) Example of Check file figures generated by MATLAB for *jim* mRNA. Note: MATLAB generates check file figures for each cell, left: original image, center: image after Gaussian filter, and right: mRNA dots detected are color-labelled. The applied threshold is different in (A–C): (A) threshold is 3, (B) threshold is 11 and (C) threshold is 20. (D) Threshold curve generated by MATLAB for one cell. MATLAB generates Threshold curve for each cell. The thresholds are always chosen manually close to the plateau after the first peak, in our example with select threshold 11. In most of our case, the chosen for one cell is valid for all the cells of the experiment.

C:\Count_mRNA_Urdy\ LL3_Jim\VNC01\T2R\Limeseg\ LL3_jim_VNC01_T2R_cell 1 to N.

15. Download MATLAB : <https://fr.mathworks.com/>.
16. Download the source_code: <https://doi.org/10.5281/zenodo.7674721>.
17. Place this folder in C:\Count_mRNA_Urdy.
18. Open MATLAB:

Editor\Open\Count_mRNA_Urdy\Source_code\Count_mrna_modified_Urdy_upgraded_5.
(Figure 7).

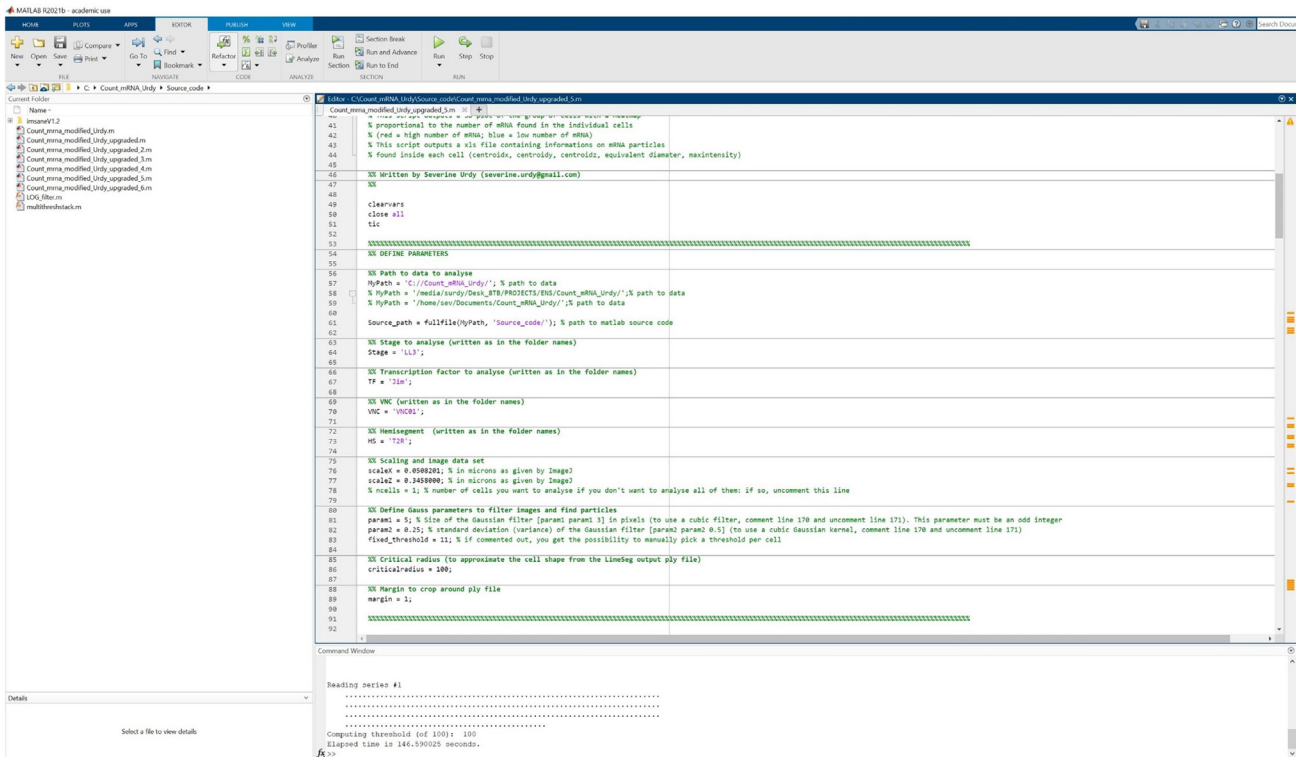


Figure 7. Screenshot of MATLAB software

In line 64, 67, 70 and 73 the different information needed by MATLAB to import the data. In line 76, 77, 81, 82 and 83 the different parameters used for RNA spot detection.

19. Define the parameters in lines 57, 64, 67, 70 and 73.

Note: These parameters are used to import the correct .ply file and confocal stacks (Figure 5J). Figure 7 illustrates the parameters we used.

20. Define the parameters in lines 77, 78, 82, 83 and 84.

- Define the voxel size of your confocal stack in lines 77 and 78.
- Define the Gaussian parameters to filter images and find particles in line 82 to 84. Chose the parameters that give the best visual detection of mRNA spots.

Note: Our datasets are as follow: width of the Gaussian filter = 5 (unit: number of pixels, in our case 1 pixel = 50nm); variance of the Gaussian filter = 0.25; threshold (filter to remove the background) = 0.11 for jim, 0.08 for chinmo, 0.15 for mamo, 0.09 for br, 0.29 for nvj and 0.15 for oli. The detected spots in each cell have a normal distribution of diameters with a mean of 0.25 +/- 0.05 micrometers and their maximum intensities displayed a unimodal distribution, arguing that the detected spots are mostly individual molecules.

21. Run the script *Count_mrna_modified_Urdy_upgraded_5* in the command window.

- When the script is done, a 3D view (Figure 5L) (see outcomes) and Check files (Figure 5K) are generated. The check File stack are used to visualize the accuracy of the spot detection for a given threshold.

Note: In most of our samples, common thresholds are chosen for all the cells of a given Lin A/15. However, specific thresholds are occasionally chosen for some cells.

	A	B	C	D	E	F	G	H	I	J	K
1	Cellid	CellCentroidX	CellCentroidY	CellCentroidZ	CellSurfaceArea	CellVolume	CellSurfaceArea_mu	CellVolume_mu	number_of_mrna	Number_mRNA_inside	
2	Cell_1	3055,803393	3206,88965	231,2764389	73975,13477	1652345,897	191,0542907	216,8740113	68	6	
3	Cell_2	3078,817778	2789,368707	158,52322	50240,22901	947170,1179	129,7545635	124,3181487	99	23	
4	Cell_3	2927,368532	3117,56041	229,2992314	30160,93316	427610,7751	77,89611619	56,12484906	131	50	
5	Cell_4	2867,58899	3111,744662	197,5654877	31477,5701	442166,4949	81,29657145	58,0353191	154	52	
6	Cell_5	2905,437507	3011,55614	194,9719899	28989,33588	385757,4501	74,87025232	50,63150865	176	44	
7	Cell_6	2822,685706	2986,895139	180,7807071	31093,08001	429499,4785	80,30355562	56,37274551	169	53	
8	Cell_7	2799,433119	2919,027163	217,5980747	33212,30183	496162,3041	85,77683283	65,12238711	167	59	
9	Cell_8	2894,025981	2942,367867	197,0689929	35898,26521	512088,0521	92,71382364	67,21267634	207	52	
10	Cell_9	2957,171393	2759,282961	222,0045935	64409,70992	1176011,611	166,3498348	154,3540949	105	44	
11	Cell_10	3002,902433	2676,227264	209,3046842	58422,38998	1052937,134	150,8864879	138,2003008	45	14	
12	Cell_11	3096,961006	2699,995498	212,3877295	54911,29809	1001386,025	141,8184521	131,434105	123	24	
13	Cell_12	2827,994577	3023,064022	235,679246	36719,71502	537990,8738	94,83536774	70,61247832	228	60	
14	Cell_13	3044,683644	2760,115176	276,7563378	81068,86202	1728276,879	209,3751364	226,8401187	183	55	
15	Cell_14	3087,941428	2610,289465	286,5200113	59471,88634	1118614,017	153,5970039	146,8205352	81	28	
16	Cell_15	3070,188171	2872,803725	258,0625306	53814,28279	946692,6031	138,9852099	124,2554738	288	47	
17	Cell_16	3034,865396	2954,530747	242,563648	37847,13886	592269,4332	97,74714564	77,73665791	213	57	
18	Cell_17	3071,485057	3049,369598	275,4566836	38806,26124	618891,6863	100,2242543	81,23088683	178	68	
19	Cell_18	2875,128648	2867,587404	270,3663833	50795,39446	960908,4098	131,1883796	126,1213295	152	54	
20	Cell_19	2850,803495	3069,448256	277,6487343	27305,38245	367203,0399	70,52113515	48,19620176	101	36	
21	Cell_20	2979,412508	3172,404186	279,9175982	24439,02694	298490,6788	63,11824876	39,17755414	20	3	
22	Cell_21	3109,799124	2935,611843	281,8296492	42398,74091	601880,9803	109,5024889	78,99819448	195	71	
23	Cell_22	2856,241492	2948,190056	279,5925589	30866,90784	419746,5151	79,71942468	55,09264774	174	58	
24	Cell_23	2986,413261	2860,413213	328,8908943	74822,50905	1539534,242	193,2427895	202,0672349	317	77	
25	Cell_24	3041,647195	2951,336427	327,5114032	41109,5119	623884,3248	106,1728196	81,88618154	206	89	
26	Cell_25	2942,323253	3038,427946	288,774205	66283,43016	1318865,503	171,1890594	173,1039805	308	120	
27	Cell_26	2900,488222	2979,856132	309,4049922	43537,61733	665058,8269	112,4438452	87,29042496	330	73	
28	Cell_27	3003,495682	2604,350602	374,8158888	77068,42257	1581540,229	199,0432712	207,5806114	138	45	
29	Cell_28	3009,184459	2485,496911	427,1065013	57690,03005	1021307,925	148,9950347	134,0488979	43	14	
30	Cell_29	3159,064668	2488,509137	436,5105033	75779,138	1477289,346	195,7134584	193,8974563	86	21	
31	Cell_30	2939,082923	2567,060459	445,2140167	50774,5451	911215,5292	131,1345323	119,5990303	42	11	
32	Cell_31	3027,876241	2507,985431	509,3534034	59223,45287	1124843,689	152,9553791	147,6381932	46	14	
33	Cell_32	3086,948473	2407,30002	501,3017165	67658,57257	1324120,675	174,7406157	173,7937333	70	16	
34	Cell_33	2983,710919	3116,323589	272,8097579	34644,32343	519211,8574	89,47529007	68,14769137	73	18	
35	Cell_34	2998,824961	3024,838533	306,2489228	42307,77945	534066,9025	109,2675643	70,09744851	153	25	
36	Cell_35	3053,175103	2556,425157	359,3048793	72141,08567	1458264,646	186,3175241	191,4004229	99	26	
37	Cell_36	3023,707732	3111,870465	314,0043095	9782,684106	80352,47821	25,26556767	10,54643843	13	2	
38	Cell_37	2973,621188	2858,762713	201,811001	114309,3804	2841120,505	295,2248436	372,9032775	535	89	
39	Cell_38	2981,413976	2993,317415	340,1669817	26015,19886	317966,5714	67,1890005	41,73380763	90	27	
40											

Figure 8. MATLAB result table

The table contains the Cell centroid in X, Y Z, Cell surface, Cell volume and the number of mRNA in the crop or in the segmented cell. The table contains the Cell centroid in X, Y Z, Cell surface, Cell volume and the number of mRNA in the crop or in the segmented cell.

- b. Change the parameters until you get good accuracy, it means the number of mRNAs detected by the script match with the reality (Figure 6A–6C).

Note: For our example in Figure 6, the validated threshold is 11: the number of spots correlates with the stack and the threshold is after the first peak.

Note: An Excel file with Count mRNA number for each cell are generated and automatically saved in C:\Count_mRNA_Urdy\LL3_Jim\VNCO1\T2R.

EXPECTED OUTCOMES

The expected outcome is the precise quantification of single molecules of mRNA at the single cell level in 3D. The MATLAB script generates two types of files reflecting the number of mRNA in each segmented cell. The first file (Figure 8) is a spreadsheet containing the following: Cell centroid in X, Y Z, Cell surface, Cell volume and the number of mRNA in the crop or in the segmented cell. The second file (Figure 5L) is a 3D representation of the segmented cells where the border of each cell is color-coded according to the mRNA expression level and each mRNA spot is represented by spheres that are visible by transparency through the cells.

LIMITATIONS

The main limitation of the protocol is that the pipeline uses many manual steps. We are currently improving this pipeline to make it more automatic.

TROUBLESHOOTING

Problem 1

We have noticed cross-reactions in some experiments (related to step 2) where two types of mRNA are detected in the same experiment with two different dyes. This cross interaction is characterized by the detection of one RNA species into two different channels. We have hypothesized that this is due to non hybridized secondary probes and free flap sequence (that has not been hybridized by secondary probes) in our mixture of hybridization.

Potential solution

- We have solved the problem by performing sequential hybridization.
- We are also designing new flap sequences that will avoid cross hybridization between flap sequences and secondary probes.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jonathan Enriquez (jonathan.enriquez@ens-lyon.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The codes generated during this study are available at zenodo.

- Remap source code: <https://doi.org/10.5281/zenodo.7456891>
- smiFISH analysis source code: <https://doi.org/10.5281/zenodo.7674721>
- List of smiFISH probes : <https://doi.org/10.5281/zenodo.7674721>

All data generated and analyzed using this pipeline are available in this study.¹

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xpro.2023.102316>.

ACKNOWLEDGMENTS

This work was funded by the Atip-Avenir Program and the Equipex+Spatial Cell-ID (ANR-21-ESRE-0016). We acknowledge the contribution of SFR Biosciences (Arthro-tool facility, UAR3444/CNRS, US8/INSERM, ENS de Lyon, UCBL) and the IGFL microscopy platform.

AUTHOR CONTRIBUTIONS

Conceptualization, J.E. and S.U.; methodology, J.E. and S.U.; software, S.U., C.K., and J.E.; investigation, W.G., J.E., M.B., C.K., and S.U.; writing – original draft, J.E.; writing – review & editing, J.E., S.G., and M.B.; funding acquisition, J.E.; resources, J.E.; supervision, J.E.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Guan, W., Bellemin, S., Bouchet, M., Venkatasubramanian, L., Guillermin, C., Laurençon, A., Kabir, C., Darnas, A., Godin, C., Urdy, S., et al. (2022). Post-transcriptional regulation of transcription factor codes in immature neurons drives neuronal diversity. *Cell Rep.* 39, 110992. <https://doi.org/10.1016/j.celrep.2022.110992>.
- Baek, M., and Mann, R.S. (2009). Lineage and birth date specify motor neuron targeting and dendritic architecture in adult *Drosophila*. *J. Neurosci.* 29, 6904–6916. <https://doi.org/10.1523/JNEUROSCI.1585-09.2009>.
- Enriquez, J., Rio, L.Q., Blazeski, R., Bellemin, S., Godement, P., Mason, C., and Mann, R.S. (2018). Differing strategies despite shared lineages of motor neurons and glia to achieve robust development of an adult neuropil in *Drosophila*. *Neuron* 97, 538–554.e5. <https://doi.org/10.1016/j.neuron.2018.01.007>.
- Tsanov, N., Samacoits, A., Chouaib, R., Traboulsi, A.-M., Gostan, T., Weber, C., Zimmer, C., Zibara, K., Walter, T., Peter, M., et al. (2016). smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res.* 44, e165. <https://doi.org/10.1093/nar/gkw784>.
- Calvo, L., Ronshaugen, M., and Pettini, T. (2021). smiFISH and embryo segmentation for single-cell multi-gene RNA quantification in arthropods. *Commun. Biol.* 4, 352–412. <https://doi.org/10.1038/s42003-021-01803-0>.
- Truman, J.W., Schuppe, H., Shepherd, D., and Williams, D.W. (2004). Developmental architecture of adult-specific lineages in the ventral CNS of *Drosophila*. *Development* 131, 5167–5184. <https://doi.org/10.1242/dev.01371>.
- Awasaki, T., Kao, C.-F., Lee, Y.-J., Yang, C.-P., Huang, Y., Pfeiffer, B.D., Luan, H., Jing, X., Huang, Y.-F., He, Y., et al. (2014). Making *Drosophila* lineage-restricted drivers via patterned recombination in neuroblasts. *Nat. Neurosci.* 17, 631–637. <https://doi.org/10.1038/nn.3654>.
- Lacin, H., and Truman, J.W. (2016). Lineage mapping identifies molecular and architectural similarities between the larval and adult *Drosophila* central nervous system. *Elife* 5, e13399. <https://doi.org/10.7554/eLife.13399>.
- Machado, S., Mercier, V., and Chiaruttini, N. (2019). LimeSeg: a coarse-grained lipid membrane simulation for 3D image segmentation. *BMC Bioinf.* 20, 2. <https://doi.org/10.1186/s12859-018-2471-0>.
- Yang, L., Titlow, J., Ennis, D., Smith, C., Mitchell, J., Young, F.L., Waddell, S., Ish-Horowicz, D., and Davis, I. (2017). Single molecule fluorescence in situ hybridisation for quantitating post-transcriptional regulation in *Drosophila* brains. *Methods* 126, 166–176. <https://doi.org/10.1016/j.jymeth.2017.06.025>.
- Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T., and Eliceiri, K.W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinf.* 18, 529. <https://doi.org/10.1186/s12859-017-1934-z>.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. <https://doi.org/10.1038/nmeth.2019>.
- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
- Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* 5, 877–879. <https://doi.org/10.1038/nmeth.1253>.