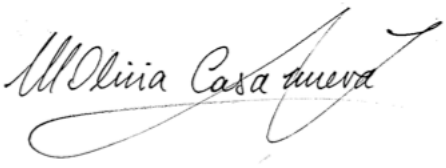
Babraham Institute.
26th/02/2020

Dear Dr. Cao:

We would like to submit a revised JoVE61136 entitled manuscript "Worm-align and Worm_CP, two open-source pipelines for straightening and quantification of fluorescence image data obtained from *Caenorhabditis elegans*". We want to thank the editorial team and the reviewers for bringing up important issues related to both the manuscript and the pipeline. We have made substantial changes to the manuscript and added 2 new figures to address all concerns. We thank reviewers as we believe that our protocol and pipeline have improved compared to the first submission.

Best regards,



Olivia Casanueva

Group leader

Editorial comments:

We have addressed all editorial comments. The changes have been highlighted in the manuscript.

- We have proof-read the manuscript and eliminated all commercial language.
- We have eliminated all personal pronouns, colloquial phrases, we have used imperative tense when appropriate.
- We have revised the text of the notes.
- (4.1.2) We have clarified that after centrifugation, the supernatant eliminated.
- (5.5.2) We have specified the microscopy settings
- We have highlighted the critical steps to be filmed.
- (Figure 4.14) We have added the screen shots in the supplementary material.
- We have sorted materials alphabetically in the Table of materials.

Reviewer#1:
Summary:
In "Worm-align, and Worm_CP, two open-source pipelines for straightening and fluorescence image data obtained from Caenorhabditis elegans.", Okkenhaug et al. describe a FIJI macro that can be used to untangle immobilized or fixed C. elegans individuals without the need for deep-learning based training datasets. In addition, the authors also import their straightened results into Worm_CP, a Cell Profiler pipeline to allow for quantification of fluorescence intensity in individual worms. The results described in this paper should be helpful for high-throughput quantification of worm biology. I believe this paper is suitable for publication in JOVE, however I believe some revisions are required before publication. In particular, I had some concerns about the output of the Worm-align program shown in Figure 15C and Figure 17B and Figure 17C.

Major Concerns:

*1A. I found the straightened worm images in **Figure 17**, panels B and C confusing, and think that some revision to this figure is needed before publication. I had two particular issues. The first issue is the ease of matching up the two-channel (GFP + transmitted light) straightened worm images with the GFP channel straightened images in Panels B and C. For example, in panel B, the middle-left side of the top GFP image appears to show faint pharyngeal labeling from two different animals, while the corresponding top two-channel image only shows one animal. Other parts of panels B and C are similarly hard to interpret. I would suggest outlining the worm body with dashes or some other contrasting marker in the GFP images to make the relationship to the two-channel images clearer.*

We thank the reviewer for noticing this issue. The discrepancies described resulted from a mistake we made when compiling Figure 17: The top (GFP channel + transmitted light) and bottom (GFP

channel only) panels were not produced during the same run of Worm-align. We have reviewed **figure 17** (**now it corresponds to Figure 7**) to best represent the output of the pipeline.

*1B. My second issue with **Figure 17** involves the presence of two apparent head/pharyngeal regions in a single straightened worm. For example, in Panel B, the fourth two-channel (GFP + transmitted light) straightened worm (counting down from top) clearly shows two head regions. It's not clear to me how this result could happen if you are straightening worms using a single nose-to-tail line drawn down an individual animal, and suggests a significant difference between the output of the WormAlign program, and the output the reader expects based on the text. I think the authors need to clarify how this type of output could be produced before publication.*

*2. **Figure 15** has a similar "two head regions" problem for worms 3, 10, and 12 in Panel C (counting down from top).*

We thank the reviewer for highlighting this issue. We have added two new Figures (**Figure 4 and 5**) and a paragraph in the results section (**lines 437-473**) that shows that the output of the pipeline depends on the quality of the lines drawn on top of the images. A complete line from the head to the tip of the tail will produce a properly aligned worm. We have updated **Figure 4** to highlight how worm tracing mistakes during the execution of Worm-align affect the output of the alignment. From the annotated montage now included, it is clear that care should be taken to avoid the following errors, as they could present issues for the proper alignment of worms:

- Tracing of the worm from tail to head (5): This results in the worm being reversed in the alignment
- Incomplete tracing (6): This results in only the part of the worm that was traced to be cropped for the montage
- Including multiple worms in a single trace (7): This results in worms plus their neighbors being inserted in the same panel of the montage
- Adding random lines to the image (8): This results in the insertion of random panels in the montage.

For creation of the montage it should not be an issue if two lines either intersect or join, even if that is over a significant part of their length, as line ROIs are individually selected, cropped and straightened by the workflow. For the CellProfiler analysis however, these occurrences would be problematic, as the Worm_CP pipeline uses the line mask, and not individual ROIs to aid identification of individual worms. We have added a new figure (**Figure 5**) to illustrate this problem. To facilitate identification of "bad worms" we suggest the use of the overlay image and the ROI collection saved in the data subfolder. This image shows the selected lines plus their labels. We have explained this step in the results and discussion sections.

*3. Several of the straightened worm images in the figures (EG **figure 10, figure 17)** show portions of other worms in the straightened images. I would like to see some discussion of how this could affect fluorescence quantitation vs. images with only a single worm - does the additional signal lead to higher measured fluorescence, and is there an easy way to control for this?*

We thank the reviewer for bringing this up. The primary application of Worm-align is to straighten and align worms, and to produce a montage that facilitates visual comparison of different worm populations. The output of the Worm-align macro however also includes line and whole population masks, which in combination with the original raw image data can be used to quantify fluorescence intensities in those worms selected. To avoid problems that arise when incomplete worm masks are segmented, either due to incomplete tracing of worms in Worm-align, or due to segmentation errors in the pipeline itself, it is advised that MeanIntensity values are reported, and not Integrated (Total) Intensities. This is because Mean Intensity values are corrected by the total area measured.

However, to make the pipeline more robust, we have added an optional checkpoint, where the user can observe the outline of the worm mask generated by cell profiler overlaid on the original image, this will show if the worm was properly segmented. The worms are numbered and any incorrect segmented worm can be eliminated from the final output. This has been illustrated in **Figs 4 and 5.**

Minor Concerns:

1. In the Table of Materials, I would list the mouth pipette as individual components, eg Aspirator Tube, Sigma-Aldrich, A5177 (or whatever source the authors used). It could be argued that these are common equipment, but I would suggest that more detail is better here especially for new labs

We have changed the mouth pipette to make the protocol safer (detailed in answer to reviewer #3), and are including a 0.2 μm syringe filter for safety in the making of the mouth micropipette We are detailing in the Materials table, the information about each component of the micropipette. We have also updated the images in supplemental **Figures 1** and **2** with the new mouth micropipette design. The changes have been made to **lines 197 to 206**.

2. In Table of Materials, I would also suggest listing company and catalog number for all commercially available like Bopidy and Levamisole.

We have added the catalog numbers to chemicals, as requested.

3. As described, the WormAlign procedure seems like it may have some issues dealing with different stages of animals (e.g. as L4 and younger animals have a smaller diameter than adults, the user may run into problems if they analyze different stages in the same images.) I would suggest discussing this and noting that the user should stick to analyzing a single stage in each image. This is implied in the text but not explicit.

This is a valid point if an experiment will be run with different developmental stages. Quantitation across genotypes should be done on synchronized animals. If synchronization is not possible, this pipeline has an advantage over fully automated pipelines such as WormTool Box, because Worm_CP allows visual inspection of the samples, so the outlines can be drawn in DIC images of worms at the desired developmental stages. We have added a paragraph in the discussion to highlight this point (**lines 646** to **653**).

4. The text suggests that manual alignment is faster than generating training models for deep learning-based approaches. I would like to see some more quantification or discussion of this - is there a limit to manual analysis where using training datasets are just as fast?

Cell profiler is an open source highly resourceful program that is ideal for high-content screening. Proficiency on the use of the software requires, however, a significant time investment. Tutorials are available that are very helpful in this process. However, the advantage of Worm-align/worm CP, is that it is very easy to use. For simple visualization and quantitation, we think that it is preferable to use this pipeline, however, WormTool Box presents many options for users and if high content screening is required, the choice would be to get training on WormTool Box. We have added a paragraph in the introduction and in the discussion to guide users to the right choice (**lines 65-73** and **666** to **670)**.

5. Clarify where to download the Worm-Align Macro during the install process. As written, the paper seems to suggest that the Worm-Align macro is included in the FIJI macros folder with the standard FIJI distribution. I would suggest a more explicit statement to the effect of "Download the Worm Align Macro from our included link, then paste into the FIJI startup macros file."

Worm-align is not included as a FIJI package, we have added the Macros as supplementary material and also added them to GITHUB. https://github.com/hannekeo/Worm-align.

Reviewer#2:

Summary:

Okkenhaug et al present a workflow to take images of a population of worms on a slide, which may be curved and overlapping, segment and computationally straighten individuals, and quantify fluorescence. The workflow appears to be a useful tool for quantifying data from small-to-medium size datasets, but some more details would improve the manuscript.

Major_Concerns:

In the introduction, the authors outline the rationale for developing their worm-align/worm_CP pipeline. An existing pipeline, WormToolbox, uses a pre-trained model to segment worms, and may require retraining when the experimental setup is modified. The worm-align/worm_CP pipeline in turn does not depend on a trained model, but the experimenter must draw the skeletons of each worm so that they can be segmented. A quantification of the time taken to process small, medium and large datasets, as well as a comparison to the time taken to process the same data using WormToolbox (including retraining) is required for readers to judge which tool best suits their needs. The WormToolbox authors suggest 50 training worms were sufficient to train the model, which seems quite a manageable investment when setting up a large experiment

The reviewer is correct in that the worm-align/worm_CP is not automatic and that it requires the user to draw a line of the worms. This process is time consuming but it halves the time compared to manually drawing an outline around each animal to quantify fluorescence using FIJI alone.

We have changed the text to reflect this point in the discussion and in the introduction.

Compared to Worm_CP, WormToolbox is an automatic solution because in principle, all that it takes from the user is to pre-train the system with a subset of the samples. Although this is correct, in practice the community -unfortunately-does not use it so often. In PubMed only 50 studies have cited the open-source software WormtoolBox in the 8 years of availability. WormTool Box provides a great number of features and capabilities that allow high-content screening. The use of WormTool or CellProfiler, however, requires an investment of time to learn how to use it. Although there are tutorials available online and anyone who needs an automated high-content phenotyping capability should consider investing this time, Worm_align/Worm_CP does only require opening the macro on Fiji or CellProfiler and following simple instructions. Worm align was developed to visualize worms to produce quickly a montage of straightened worms for improved visualization. The analysis side of the pipeline is secondary. The simplicity of the pipeline may suit some users better, but it does not provide the same number of capabilities accessible through WormToolbox.

We have added a paragraph in both the introduction and in the discussion to guide users in their choice of pipeline (**lines 65-73 and 666 to 670**).

Due to imperfect drawing of the worm skeleton, the straightened worms are often distorted. Also, some parts of the worm are being stretched or compressed during straightening. Whether this could affect the resulting quantification should be discussed. For example, if a fluorescent patch lies along the bend of the animal and the corresponding pixels are stretched or distorted during straightening, would it affect the quantification of fluorescence intensity? To what degree? This could be measured by comparing measurements from individual worms before and after straightening.

As explained to Reviewer 1, the straightening of the worms does not affect quantitation. Moreover, we have added an image that shows all lines superimposed with the montage, to any worm that is incorrectly traced should be eliminated from the final quantitation. (**Figures 4 and 5, lines 437-473)**

Minor_Comments:
We have updated the text to reflect changes related to all concerns by reviewer 2.

- We have added Carpenter, 2006; Rueden 2017 and Schindelin 2012 papers (refs 13,14,15)
- Lines 165-6: We have used Beaker instead of Becher and conical flask instead of Erlenmeyer.

- Lines 168-9: 'negatively impacts the pad's thickness' - I know what you mean, but the way it is written I had to think about it. Could be more precise, such as 'makes it harder to achieve the desired pad thickness'
- Line 171: we have changed 'agar' for agarose
- Line 183-4: We have written 'Mount the worms within 5 min' - followed by 'to prevent the pad from drying out before use'
- Line 190-1: instead of 'try to aspirate H2O with it', could make it be clearer by saying 'test whether the end of the capillary is open by trying to aspirate water'
- Lines 220-1: might be useful to give rationale for detergent - presumably to prevent worms from sticking to the sides of the plastic pipette tip
- Line 221: 'with a scissor' should be with a pair of scissors
- Line 226: English: 'shake a little bit the slide' should be 'gently shake the slide'
- We agree that Protocol G is not easy to follow by reading, however, we think that the video will be very helpful. We have made the manuscript open source, so that everyone can access the video.
- Unfortunately, the misplaced numbers that indicate individual worms cannot be changed, as they are generated automatically from FIJI, however, it is quite obvious which number belongs to which worm.
- We have italicized *dbl-1* in **Figure 16 (now figure 6)**.
- We have updated **Figure 16 (now figure 6)** and **Figure 17 (now figure 7)** and their respective legends.
- We have updated the table of materials. Table of materials: in the comments/description column, references need to be inserted ('prepared in the lab according to ref').
- We are providing an example dataset for potential users to test the workflow on their own computer setup. We have added it to GitHub.

Reviewer#3:

Summary:

The manuscript "Worm-align and Worm_CP, two open-source pipelines for straightening and quantification of fluorescence image data obtained from Caenorhabditis elegans" by Okkenhaug et al. presents computation tools for manipulation and quantitation of fluorescent images. The tools should be useful for the C. elegans community, which heavily uses fluorescent reporters to monitor proteostasis and other biological phenomena. Overall, this manuscript will be a nice addition to collection.

Major_Concerns:

In the abstract, the authors claim that their methodology represents a high-throughput solution that can be adapted to high-content screening with additional modules. However, I am concerned that this is overstated since the method requires manually drawn lines for every worm in each image to be analyzed. I would classify this approach as low-throughput and I do not anticipate that it would be suitable for high-content screening.

We apologize and thank the reviewer for pointing this out, the high-throughput comment was related to WormTool Box, but we agree that this can cause confusion to the readers. We have changed the text in the abstract.

I have serious reservations about the mouth pipetting described in Protocol C. Mouth pipetting is prohibited in most laboratories due to safety concerns and readily available alternatives. At a minimum, the authors must explain why other methodology is unsuitable and must describe safety precautions for this technique. Furthermore, the technique itself is not clearly described. It is not clear whether the capillary is broken after extending in the flame and the nature of the adaptor, silicone tubing, and the assembly of the pipette are all unclear.

We thank the reviewer for taking into consideration safety. We have updated the preparation of custom-made mouth micropipettes, according to the safety standards in mouse labs, which use this type of micropipettes routinely to isolate mouth embryos[12]. In particular, a 0.2 μm syringe filter is now being inserted along the silicone tube as a safety check. We have added a picture of this in **Figure 1C**. We also propose an alternative method for mounting the worms, if the experimenter does not want to use a mouth micropipette (**lines 211-216**).

The heat shock experiments described and shown in Figure 17 have two issues. First, in order to compare the reporter induction under the two different conditions, it would be useful to include a control set of images with worms that are not heat shocked.

We thank the reviewer for pointing out an important missing control. We have now included in **Figure 16 (now figure 6 A-B)**, images of strain AM446 carrying *hsp-70*p::GFP in the absence of heat-shock. We have also quantified them in **Figure 6E**.

Second, it has been documented in other systems that the amount of temperature change is an important determinant of the degree of induction, but that lower growth temperatures result in stronger heat shock responses (see Abravaya et al. Attenuation of the heat shock response in HeLa cells is mediated by the release of bound heat shock transcription factor and is modulated by changes in growth and in heat shock temperatures, Genes and Development, 1991). Therefore, the result that the authors present is quite unexpected, and thus the experiment warrants a better description, requires motivation/justification, and also it needs to be put into the proper context. Specifically, why was this experiment done? What do these results mean? When would a reader of the manuscript need this information? Finally, do the error bars represent SD or SEM, what is the N value, and how was the p-value calculated?

We have added the missing information regarding errors bars, n value, and statistics in the legend of **Figure 6.** We have performed this experiment in the context of a research project in the lab, investigating the influence of temperature acclimation on the heat shock response (data unpublished yet). This experiment provides an example of slight change in a transcriptional reporter, which can be quantified using the Worm_CP method. We report data from one experiment here, but we have observed this result several times.

We thank Reviewer #3 for putting this experiment in a broader context, and it is true that this result seems to contradict the conclusion of Abravaya *et* al, 1991, where HeLa cells that were incubated at a slightly lower growth temperature exhibited stronger activation of HSF-1. It should be noted that in this experiment, the heat shock is sustained (4h at 42°C degrees), whereas in our case, it is a short heat shock (30 min at 34°C degrees). We have been looking for similar experiments in a wide range of organisms, but did not find many reports on the consequences of temperature acclimation on the heat shock response. There is another report in mice showing that the heat shock response is stronger if the housing temperature is lower (Eng *et al*., housing temperature influences the pattern of heat shock protein induction in mice following mild whole-body hyperthermia, 2014). This report seems to corroborate Abravaya *et al*. results in mammalian cells.

However, this is less clear in ectotherms. Although there is evidence that lines of *D. melanogaster* evolved to grow at slightly higher temperatures (still within the range of acceptable temperatures) for several years are more resistant to thermal stress (Gilchrist *et al*, thermal sensitivity of *Drosophila melanogaster:* evolutionary responses of adults and eggs to laboratory selection at different temperatures, 1997), we did not find reports in *D. melanogaster* of temperature acclimation experiments on the scale or one generation. In a study on heat shock response in intertidal gastropods, 2 out of 3 species investigated (*T. funebralis* and *T. montereyi*) exhibited higher *hsp-70* heat-induction after acclimation at higher temperature (Tomanek, Variation in the heat shock response and its implication for predicting the effects of global climate change on species biogeographical distribution ranges and metabolic costs, 2009). We wonder whether the link between temperature acclimatization and heat shock response could be different in endotherms such as mammals versus in ectotherms organisms such as *C. elegans*. We have not discussed this point of this manuscript, since it is not relevant to this publication *per se* and we have only kept animals grown at 25°C, eliminated the animals grown at 15°C, to avoid any confusion.

Minor_Concerns:

- Line 93 - We have changed fixated to fixed.
- Line 107-108 We have added a note regarding the following point: “why are 1000 worms needed for fixing if only 20-30 are analyzed by immobilization. This is probably due to large loss of worms when handling but should be more clearly indicated for someone who is not familiar with this type of manipulation”.
- Line 105 -W have checked the diameter of our plates (9cm). “standard sizes for petri dishes are 6cm and 10cm, please verify that your plates are 9cm”.
- Line 111 - We have changed the term falcon tube by conical tube.
- Line 113 - We have changed "repeat wash one more time" to just "repeat wash".
- Line 115- The tube has been described as a microcentrifuge tube.
- Line 121-122 and throughout the protocol - We have replaced Bodipy by BODIPY. It is true that two kinds of BODIPY dyes exist: BODIPY 493/503 and BODIPY 581/591 C11. We have mentioned BODIPY 493/503 in the Materials Table. In the text, we refer to BODIPY, which is the abbreviation of *4,4-difluoro-4-bora-3a, 4a-diaza-s-indacene.* We cannot refer to in the text as BODIPY 493/503 as it is a trademark. On the other hand, we did not want to be too vague by referring to it as lipid droplets binding dye, as several dyes of that sort exist (e.g. Nile red, ORO) and it could be confusing.

- Line 123 - Isopropanol is not capitalized.
- Line 128-129 -has been grammatically fixed.
- Line 139 - the temperature has been specified to be about 40 degrees.
- Line 152 - the time and force of the spin has been indicated to be 1500 rpm
- Line 154 - We have removed this note. “this note should be removed because it is better described in the line 136 note”.
- Line 171-177 - We have used the technique based on tape on flanking slides to make pads. In our experience, we obtained better results with the technique described here. We mention now in the protocol the tape on flanking sides as an alternative method. 2.2 - “it is typical to use a piece of tape as a spacer when forming agarose pads - was tape used here?”
- Line 482 -we have replaced GFP with fluorescent dye that incorporates into LDs. “GFP is incorrect here”.
- Line 497 - The macro can be easily updated to include more channels, but we 2 channels plus DIC is what is used most commonly in labs, so we have kept the first version simpler.
- We have corrected this - “Figure 12 is incorrectly labelled as Figure 11”
- We have corrected this: “Figure 16 is incorrectly labelled as Figure 14”

Figure 16 (now figure 6) - it would be helpful to know how a 12% decrease in signal compares to the published literature and/or other methods that have been used to quantitate this effect, especially since this seems like a modest effect.

In ref 18 (Clark *et al*., 2018), the decrease observed in lipid droplet content is stronger (about 35% decrease). However, the lipid droplet content is monitored with another dye in this study (ORO staining) and on another allele of *dbl-1(wk70),* whereas we are investigating *dbl-1(nk3)*. We observe the same trend as Clark and colleagues, however the discrepancies in the amplitude of the decrease could be due to methods or genetic background differences. Therefore, we did not compare directly the amplitude of our result to the one observed by Clark and colleagues.

Reviewer#4:
Summary:
This paper describes software to easily quantify fluorescence of C. elegans. Fluorescence quantification is such a regularly used technique for C. elegans, that this protocol will provide a valuable reference for essentially all labs doing C. elegans research. The protocol is very well described and illustrated.

Major Concerns:
The text that refers to figure 16 mentions measuring GFP fluorescence, but I think this should be BODIPY fluorescence. If GFP is correct, the figure legend will have to be modified. Figure 16 is labeled figure 14.

We thank the reviewer for pointing this out, we have altered the text to reflect that wavelength used.

Minor_Concerns:
Section 2.2: If uniform thickness is important, why not use flanking layers of tape on the lower slide to provide pads of a reproducible and uniform thickness? This is the standard technique.

We have used the standard technique based on tape on flanking slides to make pads. In our experience, we obtained better results with the technique described here. However, we also mention in the protocol the tape on flanking sides as an alternative method (**lines 171-177**).

If you modify section 2.2 as I have suggested, you can then simplify section 5.1.

We have kept section 5.1, as we kept the description as our pad method along the description of the standard technique using tapes on flanking slides.

In the first note under section 6.3, you should indicate acceptable file formats.

The protocol now indicates acceptable file formats, nd2, czi, tif, rgb, jpg and png (**lines 286-287**).

section 76.3 was disorienting. You state that something is second, but nowhere do you state what is first. If this is second, you need to explicitly state what is first. [You may be referring to a menu item that is second in a list? If so, write as if you are navigating a menu.

We have rewritten parts of section 7.3 (**lines 369-370**).

Please see more specific comments/corrections in the attached pdf.

We have addressed all the comments mentioned in the pdf document.