An image analysis toolbox for high-throughput *C. elegans* assays

Carolina Wählby[1,2,*], **Lee Kamentsky**[1], **Zihan H. Liu**[1], **Tammy Riklin-Raviv**[3], **Annie L. Conery**[4], **Eyleen J. O'Rourke**[4], **Katherine L. Sokolnicki**[1], **Orane Visvikis**[5], **Vebjorn Ljosa**[1], **Javier E. Irazoqui**[5], **Polina Golland**[3], **Gary Ruvkun**[4], **Frederick M. Ausubel**[4], and **Anne E. Carpenter**[1]

[1]Imaging Platform, Broad Institute of MIT and Harvard, Cambridge, MA

[2]Centre for Image Analysis, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[3]Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA

[4]Dept. of Molecular Biology and Center for Computational and Integrative Biology, Mass. General Hospital, Boston, MA

[5]Developmental Immunology Program, Dept. of Pediatrics, Mass. General Hospital, Boston, MA

Abstract

We present a toolbox for high-throughput screening of image-based *Caenorhabditis elegans* phenotypes. The image analysis algorithms measure morphological phenotypes in individual worms and are effective for a variety of assays and imaging systems. This WormToolbox is available via the open-source CellProfiler project and enables objective scoring of whole-animal high-throughput image-based assays of *C. elegans* for the study of diverse biological pathways relevant to human disease.

Phenotypic assays of intact organisms allow the study of biological pathways and diseases that cannot be reduced to biochemical or cell-based assays. *C. elegans* is a useful model system for studying biological processes shared with humans[1], high-throughput instrumentation and reagent libraries exist for sample preparation and imaging[2], and deviations from wild-type are often readily apparent because worms are visually transparent and follow a stereotypic developmental pattern[3]. Large-scale chemical and RNAi screens using *C. elegans* are widespread[4–6] and can probe complex processes such as metabolism, infection, and behavior, but so far the analysis of such experiments has largely been manual, subjective, and onerous.

Much progress has been made in automating the analysis of particular types of *C. elegans* experiments, such as those involving low-throughput, high-resolution, 3-D, or time-lapse images, or images of embryos[7–11]. However, there is still a strong need to automate the analysis of high-throughput, static images of adult worms in liquid culture, a common screening output. For most assays, the density of worms per microplate well causes the worms to touch or cluster, so that automated analysis has been limited to population-

averaged measurements[12–13], hiding population heterogeneity and prohibiting measurements on individual animals.

An alternative to microscopy is flow systems adapted for worms(e.g., COPAS, Union Biometrica), measuring length, optical density and fluorescence emission at transverse slices along the length of individual worms. However, image-based screens have several benefits: They allow detection of more complex phenotypes by two-dimensional analysis of shape and signal patterns, and do not require re-suspension of worms in additional liquid prior to analysis, allowing smaller sample volumes and closed culture conditions —an important factor when screening large libraries of small molecules and RNAi clones, and when using pathogenic microbes. Also, image based screening allows for visual confirmation of results, the images form a permanent record that can be re-screened for additional phenotypes, and low-throughput experiments require no more equipment than a microscope and a digital camera.

To improve *C. elegans* phenotype scoring from images of adult worms in liquid, we developed an image-analysis toolbox that can detect individual worms regardless of crossing or clustering. It can measure hundreds of phenotypes related to shape, biomarker intensity, and staining pattern in relation to the anatomy of the animals.

A typical workflow starts with bright field images(Fig. 1a). We pre-process to compensate illumination variations, detect well edges, and make the image binary (Fig. 1b). The next step, and the major challenge, is "untangling," i.e., detecting individual worms among clustered worms and debris. To address this, we first construct a model of the variability in worm size and shape from a representative set of training worms(Fig. 1c). The model is then used to untangle and identify individual worms(Fig. 1d). A large number of measurements such as size, shape, intensity, texture, and spot counts can thereafter be made on a per-worm basis using all image channels available, as is common for cell-based assays[14]. Many phenotypes, such as spot area per animal, can be scored directly by such measurements; more complex phenotypes, such as subtle or complex changes in protein expression patterns, can be scored using a combination of measurements and machine learning[15]. If reporter signal location is of interest, we map each worm to a low-resolution atlas allowing quantification correlated to the worm's anatomy.

We evaluated the untangling performance using images from our prior work[8], where 15 worms were placed in each well of a 384-well plate. Approximately 1500 worms from 100 wells were manually delineated, revealing that 46% of the worms were clustered or touching other worms (Supplementary Fig. 1). Compared to manual delineation, 51% of the worms were correctly detected with automated foreground-background segmentation followed by connected component labeling. When applying the untangling algorithms of the WormToolbox the performance increased to 81%, which proved sufficient for the assays presented here. The major source of error was poor image contrast close to well edges; performance improved to 94% when the foreground-background segmentation was manually corrected, decoupling errors caused by untangling from errors in the initial segmentation. We also tested the performance of the untangling in relation to the size of the training set, and found that performance plateaus using a worm model constructed from 50 randomly selected training worms. This means that training can be done on a relatively small number of samples representing the phenotypic variation of a given experiment (Supplementary Fig. 2).

We first evaluated the toolbox on data from a different laboratory and imaging system[13]. The challenge was to detect individual adult worms that were partly clustered and mixed with eggs and progeny. We trained the worm model on L4 and adult worms only, and

observed that untangling improved the accuracy of finding individual adult worms as compared to thresholding and size-sorting alone (Fig. 1e and Supplementary Fig. 3). The model efficiently excluded smaller larvae (L1, L2, and L3) and eggs, and performance was relatively robust in the presence of up to 6-fold more progeny than adults (Supplementary Fig. 4). We also evaluated the performance of worm untangling as the number of worms per well increased. Wells contained either L1, L3, or adult worms at increasing concentrations, and we created a separate worm model for each developmental stage. As expected, the performance was higher for the slightly smaller L3 worms as more space between worms leads to less clustering, but untangling became unstable when the worms were so small(L1) that the image resolution only allowed a few pixels per worm (Supplementary Fig. 5 and 6).

In the second assay, we evaluated the toolbox for scoring viability, which can be read out as a morphological phenotype in bright field images alone, without the need for a viability stain. Worms in liquid culture tend to be curved and evenly opaque when alive but become rod-shaped and textured when they die (Fig. 1f). We untangled high-throughput images of worms infected with *Enterococcus faecalis* and either mock-treated with DMSO or treated with ampicillin[12]. After making shape, intensity and texture measurements of each untangled worm, we manually selected 150 live and dead training examples from one 384-well plate. We thereafter used the gentle-boosting classifier of CellProfiler Analyst[15] (Supplementary Fig. 7) to identify a combination of measurements that discriminates live and dead worms. Finally, we applied the classifier to 1,500 worms from a different 384-well plate, and verified that it distinguished live and dead worms as well as humans can (Fig. 1g). To evaluate the performance of the viability scoring on more heterogeneous data from a real high throughput experiment we selected 1,766 random images and 200 hits from a 37,200 compound screen[12] and compared the automated scoring with that of visual scoring based on bright field images (Supplementary Fig. 8). We achieved an accuracy of 97%, and a precision of 83%, indicating that morphology-based viability screening could be a feasible alternative to the viability stains(SYTOX)used in the original screen.

In the third assay, we evaluated how well the toolbox could differentiate between a positive and a negative control from an RNAi screen for regulators of fat accumulation[16]. The positive control down-regulates *daf-2*, and the negative control was an empty vector. We compared two different approaches for pattern quantification: per-well measurements(using the basic functionality of CellProfiler), where no effort was made to assign fatty regions to individual worms, yielded a false discovery rate (FDR) of 22.2% (Supplementary Fig. 9); and per-worm measurements(using the untangling functionality of the Worm Toolbox), yielded an FDR of 4.5% (Supplementary Fig. 10). The per-worm measurements were superior because they captured the heterogeneity of the population, which was lost in the population averages from per-well measurements.

Finally, we evaluated the toolbox's ability to detect worms with a change in the location of GFP expression(Fig. 2a). We used a *C. elegans* strain where GFP expression in the intestine is under the control of a promoter that responds to *Staphylococcus aureus* infection[17]. A pharyngeal stain (mCherry) served as an internal control. The assay could not be scored using simple approaches, such as measuring the total intensity of GFP expression per well or per worm, or counting the number of GFP spots (Supplementary Fig. 11). However, using worm straightening (Fig. 2b) and our atlas-mapping (Fig. 2c), we were able to quantitatively detect elevated expression of *clec-60::GFP* in the anterior intestine (Fig. 2d) and separate positive and negative controls with a Z'-factor of 0.21. Here we focused on location of signal along the length of the worm, but asymmetric signal distribution across the width of the worm (e.g. fluorescence in full worm as compared to only eggs, or only gut) could also be discerned, using the outline of the worm as a spatial reference for the atlas.

The WormToolbox is the first system to automatically, quantitatively, and objectively score a variety of phenotypes in individual *C. elegans* in static, high-throughput images. The toolbox is implemented as modules for the open-source CellProfiler[14,18] software, emphasizing ease-of-use, is compatible with cluster computing to speed analysis, and is flexible to new assays developed by the scientific community. Training the worm model takes less than an hour, and once an image analysis pipeline is set up for an assay, a typical analysis takes 10–30s per image; much less if a computing cluster is available.

The performance of the WormToolbox depends on the contrast between worms and the surrounding background, making it sensitive to large variations in background illumination and to the worm-like tracks sometimes formed when growing worms on agar medium. The WormToolbox can handle images of worms on agar in large plates, but further optimization is needed for worms on solid medium in 384 well plates. In liquid culture, the untangling can handle up to 20 adult worms per well in 384-well format, and is designed to detect worms of the size and shape range of the training worms used to create the worm model. Unexpected phenotypes are likely to be discarded as debris, but wells with a low fraction of correctly detected worms may be flagged for visual examination. In future work we will extend the WormToolbox by adding further worm-specific measurements based on their unique anatomy and better handling of mixed worms at various stages of development.

Online methods

The open-source code of the CellProfiler WormToolbox algorithms described here is available as Supplementary Software 1 and for download at http://www.cellprofiler.org. Example pipelines for worm model training, worm untangling, and feature extraction available as Supplementary Software 2, and on the CellProfiler website. Compiled version of the code and updates are available at http://www.cellprofiler.org. Here we describe the steps of the workflow as well as our four sample assays. Instructions for how to get started using the WormToolbox are provided in Supplementary Methods 1.

Compensating for uneven illumination

Uneven illumination often distorts bright field microscopy images of the multi-well plates typically used in high throughput chemical and genetic screens, making foreground-background intensity thresholding difficult. Our novel approach for approximating background illumination and well edge position is based on the convexity of both the well and the illumination field (Supplementary Fig. 12). The algorithm is as follows: Choose 256 evenly spaced intensity levels between the minimum and maximum intensity for the image. Starting from the lowest intensity, for each intensity, find all pixels with equal or higher intensity. Find the convex hull that encloses those pixels, set the pixels of the output image within the convex hull to the current intensity, and continue to the next intensity level. If the well edges are dark and the well has a convex shape, this approach removes the well edge and compensates uneven illumination without the need of any input parameters, making it robust to the variations often present in high-throughput experiments. The final result is thresholded using Otsu's[19] method, resulting in a binary image that serves as input for the worm untangling step.

Worm detection

Following illumination correction and thresholding, we create a mathematical description of each worm cluster (Supplementary Fig. 13). We reduce each binary object to its morphological skeleton and let each segment of the skeleton represents a worm segment, and each branch point represents a point where worms touch or intersect. This way, the segments and branch points comprising the worm cluster can be described as a mathematical

graph, and untangling becomes a search for paths through the graph that are likely to represent complete worms. More precisely, we search for the ensemble of paths through a cluster that best represents the true worms as compared to a worm model, limiting worm overlap, and maximizing cluster coverage. Our advancements as compared to our previous work[20–21] are described in the next three sections.

Worm model construction and shape cost

The worm model is created from a comprehensive set of non-touching training worms essentially as in our prior work[20], here with the shape descriptor based on angles rather than spatial coordinates. We sample equidistant control points along the morphological skeleton of each training worm using cubic spline interpolation. Each of the control points other than the first and last is at the vertex of an angle formed by the lines from its predecessor and successor. These angles and the path's length form a feature vector functioning as our shape descriptor. Worm width, length, and area are also extracted and we make the training set symmetric by mirroring all samples along the x- and y-axis. The shape cost of a path potentially representing a worm in a cluster is given by the dot product of the feature vector describing the path and a cross-correlation matrix derived from the training data. Note that we are only looking at over-all body shape when training the model, which does not vary as much as other features of worms such as fluorescent markers or bright field stains and texture. The training worms should represent the worms in the data set, and the variation in size and shape must be within certain limits, for example, a variation in length of a factor 2 might cause the untangling step to divide some long worms in half or exclude some short worms as debris. Any worms that deviate in shape and posture from that expected by the model will be discarded as debris. It is therefore feasible to flag wells with few detected worms as compared to foreground pixels so they can be screened visually (or with an improved worm model), to detect unexpected body shape or size phenotypes (or failures in worm detection due to large amounts of debris or other problems). It is also worth noting, that due to similarity in worm size and shape, we were able to re-use the worm model created for the second assay in both the third and fourth assay, which consisted of images captured over several years and on different microscope systems.

Preprocessing of cluster skeletons

Artifacts appear when two or more adjacent worms form regions wider than the worm-width, resulting in a skeleton no longer centered on a true worm. In the two-worm case (Supplementary Fig. 14), the skeleton is composed of the two segments that enter the area where the worms touch, the two segments that leave the area, and a single segment running the length of the area where the worms touch. To improve alignment of the segmentation result with the true worms we introduce a preprocessing: Touching areas are defined by a circular structuring element whose diameter is the maximum width of a worm. All skeleton ends adjacent to the area are connected with new paths and the best paths are selected by the path search described below. To improve worm detection in cases where two worms touch end-to-end without producing a branch point, we add branch points at an average worm's length starting from each endpoint of every skeleton segment longer than the longest training worm.

Worm untangling by path search

Once the skeleton has been preprocessed we consider the combined cost of different ensembles of paths representing worms. Conceptually, the algorithm is composed of three steps: enumeration of paths, calculation of costs of individual paths, and calculation of costs of ensembles of paths. The first step is to generate all paths whose lengths are between the minimum and maximum acceptable length, as defined by the worm model, and discard a path if the shape cost function exceeds the maximum acceptable cost. There are three parts

to the cost of a particular ensemble of paths: the sum of costs of the individual paths in the ensemble, a penalty cost that is proportional to the length of all segments that are shared by paths in the ensemble and a penalty cost that is proportional to the length of all segments that do not appear in any path in the ensemble. Details on the algorithm are described in Supplementary Methods 2, and the open source of the code (www.cellprofiler.org).

Straightening and atlas-based feature extraction

To extract reporter signal location, worms are transformed to a straight shape by re-sampling the image data along lines perpendicular to the central axis of the worm, much as previously described[21]. However, here we make the processing much faster by re-using the spline function describing the path through the worm during untangling. If a head- or tail-specific marker is available, worms flipped so that they all have an intensity distribution skewed in the same direction. Our low-resolution worm atlas consists of a user-defined number of transversal and longitudinal segments. Intensity mean and standard deviation is extracted from each sub-segment in any number of image channels.

Evaluation of worm detection

We evaluated the performance of the untangling on 100 bright field images from our published high-throughput experiment[12]. Each image of a well from a 384-well plate contains approximately 15 worms. Ground-truth was created by manually delineating all worms and saving them as individual binary masks to enable evaluation of worm detection, clustering, and overlap (provided though www.broadinstitute.org/bbbc). In this data set, 46% of the worms touch or overlap, with most of the worms in clusters of two (Supplementary Fig. 1). We calculated accuracy, precision, recall and F-factor for individual worms, and a threshold on F-factor of 0.8 yields 81% correctly segmented worms by automated foreground-background segmentation followed by untangling, and 94% correct segmentation if the foreground-background segmentation was manually corrected before untangling. If worms are defined by conventional intensity thresholding and connected component labeling, only 51% of the worms are correctly segmented. The performance is generally higher for smaller clusters (Supplementary Fig. 2). Performance is also affected by the size of the training set, and plateaus at about 50 training worms. Processing speed has been reduced by about ten-fold as compared to our previously published implementation[17], and all steps combined, including image preprocessing, worm untangling and straightening, typically take less than 10s.

Assay 1: Finding individual adult worms in the presence of eggs and progeny

Images were kindly provided by Gosai et al.[13] Briefly, *C. elegans* were cultured at 22 °C on nematode growth medium (NGM) plates seeded with *E. coli* strain OP50. Next, 36 animals, with a predetermined percentage (0, 25, 50, 75 and 100%) of adult worms, were dispensed into each well of a 384-well plate. Images were acquired on the ArrayScan V[TI] HCS Reader (Cellomics, ThermoFisher) fitted with a 2.5x objective and a 0.63x coupler. We compensated for uneven background illumination using the convex-hull approach and identified objects by automated intensity thresholding. We constructed a worm model from the non-touching, adult worms, which we identified based on their areas and maximum widths. We thereafter untangled all worms (Supplementary Fig. 3) and counted adult worms per image (Fig. 1e). Next we tested the limit of adult worm detection at increasing concentration of progeny (Supplementary Fig. 4), concluding that the untangling is stable to about six-fold more progeny than adults. Finally we found the limits in worms per well using a COPAS worm sorter to seed 1–96 L1, L3, or adult worms per well in a 384 well plate, in four replicates (Supplementary Fig. 5). For adult worms, the untangling works well until reaching about 20 worms per well, while for L3 worms, the limit is reached at about 30 worms per well, which is to be expected as the worms are smaller and further apart in the

well. For the small L1 worms, the image resolution only allows a few pixels per worm, making the model-based worm untangling unstable, particularly in the presence of small bubbles, which can be confused with small worms at this resolution. In order to explore the modes of failure, we examined a subset of segmentation results visually, and saw that the initial illumination correction and intensity thresholding have a large effect on the resulting segmentation (Supplementary Fig. 6).

Assay 2: Live/dead scoring based on bright field morphology

We cultured *C. elegans* (*glp-4(bn2);sek-1(km4)* mutant) on plates seeded with *E. coli* strain HB101. We infected sterile adult worms by pipetting them onto a lawn of *E. faecalis* and incubating for 15 h at 15 °C. Using a COPAS worm sorter, we transferred 15 of the infected worms to each well of a 384- well plate. As a positive control, we added 21 μg/ml Ampicillin to 192 wells and mock-treated 192 wells with an equal volume of DMSO[12]. We captured bright field transmitted-light images showing the entire well using a Molecular Devices Discovery-1 microscope with a transmitted light module, a 2x low-magnification objective, and MetaXpress software. We manually delineated 60 worms from positive and negative control wells and used them to construct a worm model. After untangling, we measured worm shape, intensity and texture. Using CellProfiler Analyst[15], we trained a classifier to distinguish the live and dead phenotype based on 150 training examples. The classifier used 8 shape-, intensity-, and texture features (Supplementary Fig. 7). We applied the classifier to a set of images that did not include the training examples, classifying each worm as live or dead. Finally, we scored wells by the fraction of live worms and compared to the majority vote of three *C. elegans* specialists scoring by visual inspection (Fig. 1g). We also compared automated and visual scoring of images randomly selected from a full-scale HTS experiment (Supplementary Fig. 8).

Assay 3: Fat accumulation scoring based on staining pattern

We treated animals with either an empty vector (L4440) or RNAi against the insulin receptor (*daf-2*) according to standard procedures. We stained them with the fat-specific stain oil red O. Using an upgraded Axioscope microscope (Zeiss) with automated hardware (Biovision Inc.) and Surveyor software (Objective Imaging Ltd.), we acquired six bright field color images per well. The original images were 2782×3091 pixels, but we scaled them to 690×765 pixels to speed analysis. Before detecting and untangling worms, we combined the color channels into a single gray scale image. We used the same worm model as for Assay 2, and achieved satisfactory segmentation results without any adjustments. We defined fat regions by intensity thresholding and quantified fat patterns by measuring the extent of the fatty regions (Supplementary Fig. 9). We thereafter compared per-well, per-cluster, and per-worm measurements (Supplementary Fig. 10), finding the latter to be ideal for the assay.

Assay 4: Reporter pattern detection by worm straightening and atlas mapping

We used a transgenic strain of *C. elegans* that expresses GFP from the promoter of the gene *clec-60* and *myo-2*::mCherry for labeling the pharynx. We cultured the worms on NGM plates seeded with *E. coli* OP50 at 15–20°C according to standard procedures, and sorted 15 worms into each well of a 384-well plate. Of these, 48 wells received wild-type worms (L4440) expressing *clec-60::GFP* and 48 wells received *pmk-1(km25)* mutants. Using a Discovery-1 microscope (Molecular Devices), we acquired bright field images as well as fluorescence images at two wavelengths (for GFP and mCherry). The images were 696×520 pixels[17]. We tested three approaches to phenotype scoring (Supplementary Fig. 11). First, we defined spots of GFP signal by intensity thresholding, and likewise approximated worm count by thresholding intensities in the image channel showing the pharynx (mCherry). Because large variations in GFP expression and touching worm heads lead us to under-

estimate the number of worms, this approach was not successful. Instead, we proceeded to untangle the worms using the same worm model as for Assay 2, and achieved satisfactory segmentation results without any adjustments. The mean and standard deviation of the GFP expression in individual worms was also insufficient to separate the two phenotypes, so we continued the analysis by straightening the worms, aligning them to the low-resolution worm atlas and measuring mean and standard deviation of GFP expression from each of six transversal sub-segments evenly spread along the length of the worm. Instead of examining each measurement separately, we trained CellProfiler Analyst software[15] to distinguish the phenotypes by presenting it with examples of mutant and wild-type worms. The resulting classifier relied primarily on the difference in standard deviation of GFP fluorescence in transversal segment number two (second from head, T2 of 6) to distinguish wild-type and mutant worms. Based on this, we labeled as mutant worms those with a standard deviation in GFP expression in T2 of 6 greater than 0.4. Finally, we scored each well by the percentage of mutant worms, and achieved a Z'-factor of 0.21.

A list of general aspects of assay design and error handling for image based screening is presented in Supplementary Table 1, and details regarding image settings are given in Supplementary Note 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Harris TW, et al. Nucleic Acids Res. 2004; 32:D411–417. [PubMed: 14681445]

2. O'Rourke EJ, Conery AL, Moy TI. Methods Mol Biol. 2009; 486:57–75. [PubMed: 19347616]

3. Sulston JE, Horvitz HR. Dev Biol. 1977; 56(1):110–156. [PubMed: 838129]

4. Artal-Sanz M, de Jong L, Tavernarakis N. Biotechnol J. 2006; 1(12):1405–1418. [PubMed: 17109493]

5. Kamath RS, Ahringer J. Methods. 2003; 30:313–321. [PubMed: 12828945]

6. Labbé JC, Roy R. Clin Genet. 2006; 69:306–314. [PubMed: 16630163]

7. Buckingham SD, Sattelle DB. Invert Neurosci. 2008; 8(3)

8. Long F, Peng H, Liu X, Kim SK, Myers E. Nat Meth. 2009; 6(9):667–672. http://www.ncbi.nlm.nih.gov/pubmed/19684595.

9. Murray JI, et al. Nat Meth. 2008; 5(8):703–709. http://www.ncbi.nlm.nih.gov/pubmed/18587405.

10. Sönnichsen B, et al. Nature. 2005; 434(7032):462–469. [PubMed: 15791247]

11. Green RA, et al. Cell. 2011; 145(3):470–482. [PubMed: 21529718]

12. Moy TI, et al. ACS Chem Biol. 2009; 4(7):527–533. [PubMed: 19572548]

13. Gosai S, et al. PLoS One. 2010; 5(11):e15460. [PubMed: 21103396]

14. Carpenter AE, et al. Genome Biol. 2006; 7(10):R100. [PubMed: 17076895]

15. Jones TR, et al. Proc Natl Acad Sci USA. 2009; 106(6):1826–1831. [PubMed: 19188593]

16. O'Rourke EJ, Soukas AA, Carr CE, Ruvkun G. Cell Metab. 2009; 10(5):430–435. [PubMed: 19883620]

17. Irazoqui JE, Urbach JM, Ausubel FM. Nat Rev Immunol. 2010; 10(1):47–58. [PubMed: 20029447]

18. Kamentsky L, et al. Bioinformatics. 2011; 27(8):1179–1180. [PubMed: 21349861]

19. Otsu N. IEEE Trans Syst Man Cybern. 1979; 9(1):62–66. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4310076.

20. Raviv TR, et al. Med Image Comput Comput Assist Interv. 2010:634–641. [PubMed: 20879454]

21. Wählby C, et al. Proc IEEE Int Symp Biomed Imaging. 2010

22. Peng H, Long F, Liu X, Kim SK, Myers EW. Bioinformatics. 2008; 24(2):234–242. [PubMed: 18025002]
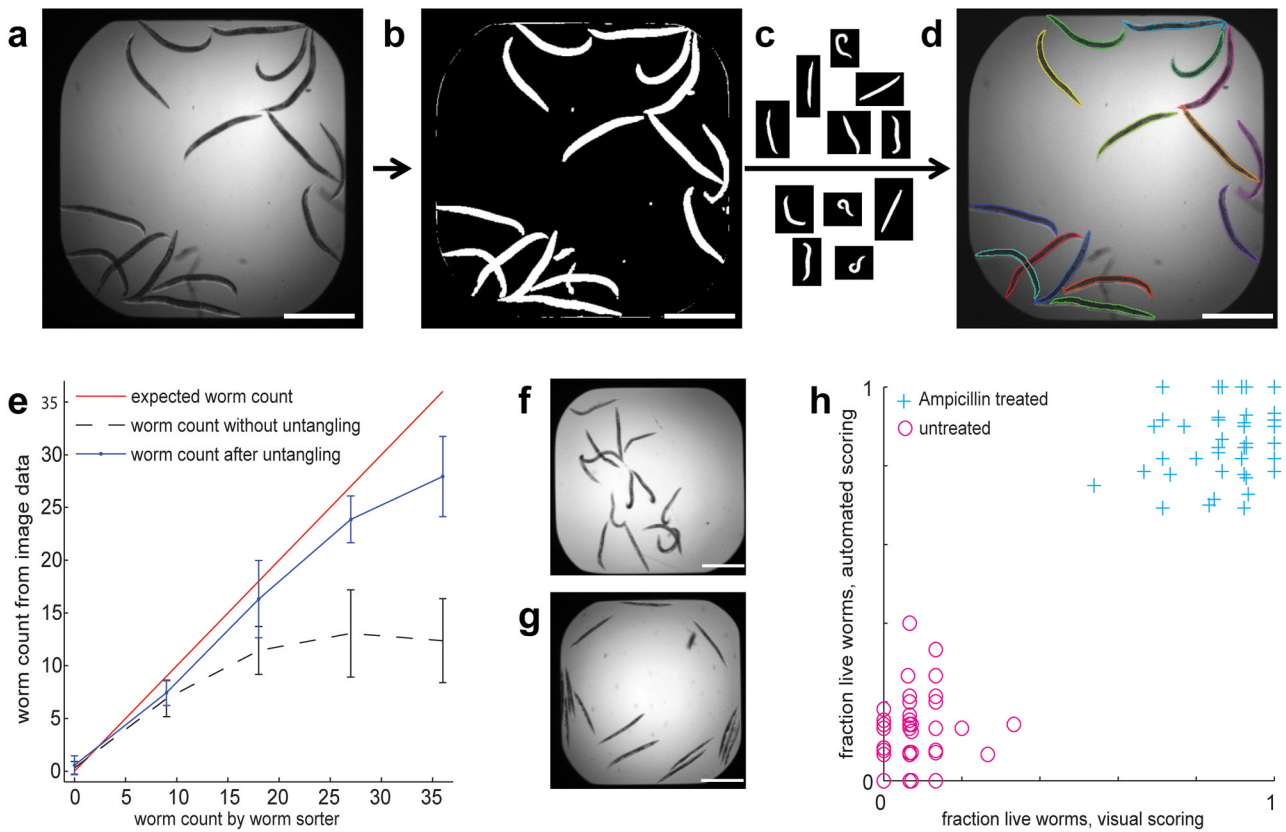
Figure 1.
Workflow and performance of the WormToolbox. (**a**) Starting from a bright field image we (**b**) remove variations in illumination and separate objects from background. (**c**) We create a worm model from non-touching training worms and (**d**) untangle individual worms from clusters and debris using the model. Each worm is given a uniquely colored outline. (**e**) Automated detection of adult worms in the presence of eggs and progeny in images from a 384-well plate, each well containing a predetermined percentage of adults, and counted with or without untangling. The number of adult worms counted by image analysis is plotted against the known number of adult worms sorted into each well. Error bars, s.d. ($n = 16$ images).(**f**) The images show the morphology of *E. faecalis*-infected worms without (top; worms are dead) and with (bottom; worms are alive) treatement with Ampicillin. (**h**) Automated viability scoring based on shape and texture measurements from individual worms compared to visual scoring of positive (Ampicillin treated) and negative (untreated) control wells. Scale bars, 1mm.
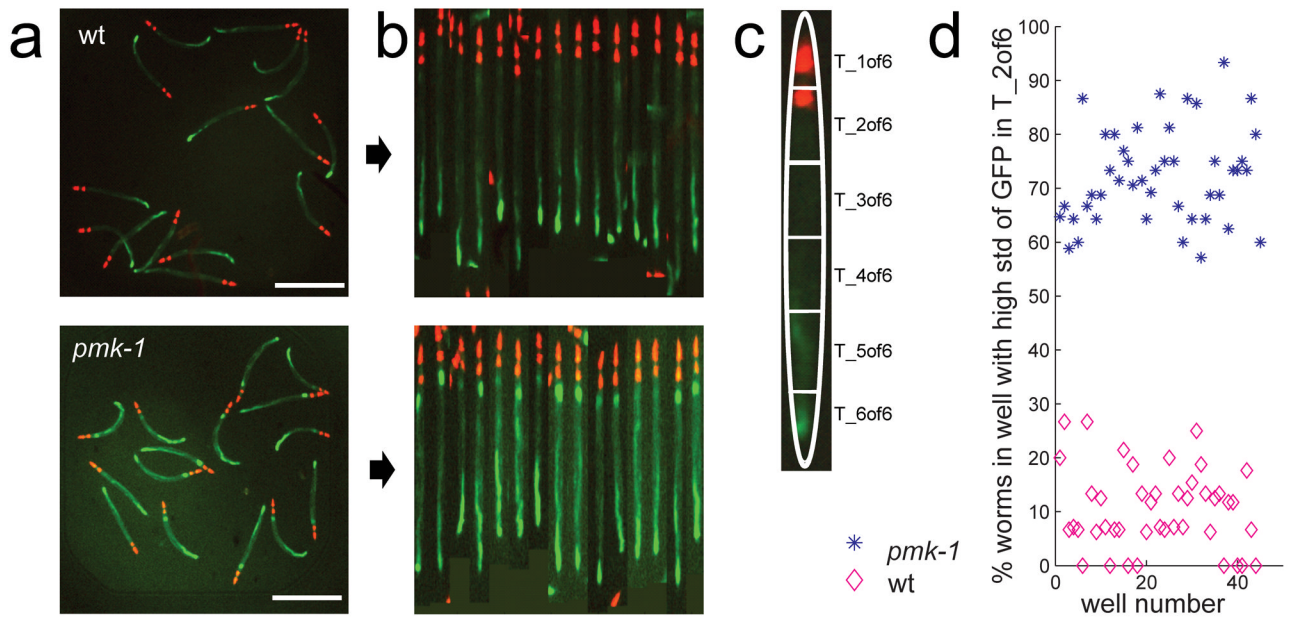
Figure 2.
Scoring fluorescence signal distribution with the WormToolbox. (**a**) The images show *clec-60::GFP* expression (green) in wild type worms (top) and in *pmk-1(km25)* mutants (bottom). The pharynx is marked with *myo-2*::mCherry (red). (**b**) Automated worm detection and straightening of the worms in (a) enables feature measurements in relation to (**c**) a low-resolution worm atlas. (**d**) Comparison of standard deviation of GFP fluorescence intensity in transversal T_2 of 6 of worms in 48 wells containing either positive (*pmk-1* mutant, blue stars) or negative controls (wild type, magenta diamonds).