ELSEVIER

Journal of Neuroscience Methods 118 (2002) 9–21

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Using machine vision to analyze and classify *Caenorhabditis elegans* behavioral phenotypes quantitatively

Joong-Hwan Baek [a], Pamela Cosman [b], Zhaoyang Feng [c], Jay Silver [c], William R. Schafer [c,*]

[a] *School of Electronics, Telecommunication and Computer Engineering, Hankuk Aviation University, Koyang City, South Korea*
[b] *Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407, USA*
[c] *Division of Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0349, USA*

Received 26 February 2002; received in revised form 26 April 2002; accepted 26 April 2002

Abstract

Mutants with abnormal patterns of locomotion, also known as uncoordinated (Unc) mutants, have facilitated the genetic dissection of many important aspects of nervous system function and development in the nematode *Caenorhabditis elegans*. Although a large number of distinct classes of Unc mutants can be distinguished by an experienced observer, precise quantitative definitions of these classes have not been available. Here we describe a new approach for using automatically-acquired image data to quantify the locomotion patterns of wild-type and mutant worms. We designed an automated tracking and imaging system capable of following an individual animal for long time periods and saving a time-coded series of digital images representing its motion and body posture over the course of the recording. We have also devised methods for measuring specific features from these image data that can be used by the classification and regression tree classification algorithm to reliably identify the behavioral patterns of specific mutant types. Ultimately, these tools should make it possible to evaluate with quantitative precision the behavioral phenotypes of novel mutants, gene knockout lines, or pharmacological treatments.

Keywords: *Caenorhabditis elegans*; Behavior; Imaging; Classification; Locomotion; Phenotypic analysis

1. Introduction

Understanding the relationship between genes and behavior represents a fundamental challenge in neuroscience. A powerful approach to this problem is to use genetic analysis in simple model organisms, such as the nematode *Caenorhabditis elegans*, to identify genes whose specific behavioral phenotypes reflect a specific role in nervous system function. Since phenotype-driven genetic screens essentially make no prior assumptions about the types of molecules involved in the process being studied, this approach is well suited for identifying previously unknown receptors or signal transduction molecules involved in poorly-understood aspects of neuronal or muscle cell activity. Furthermore, modern molecular genetics provides the ability to manipulate specific gene products in an intact animal, making it possible to assess a particular protein's functions within the context of an intact nervous system. The nematode *C. elegans* has powerful genetics, a well-described nervous system, and a complete genome sequence; thus, it is particularly well suited to genetic analysis of nervous system function and behavior (Riddle et al., 1997).

The genetic analysis of nervous system function in *C. elegans* depends on the availability of reliable assays to detect behavioral abnormalities. Unfortunately, behavioral assays in this organism, particularly in more complex behaviors such as locomotion, are often highly imprecise and subjective. For example, over 100 genes have been described which when mutated lead to abnormal or uncoordinated movement (Brenner, 1974). In the published literature (e.g. Hodgkin, 1983), these uncoordinated ('Unc') mutants are usually classified into a number of descriptive categories, including 'kinky', 'coiled', 'shrinking', 'loopy', 'slow', and 'slug-

* Corresponding author. Tel.: +1-858-822-0508; fax: +1-858-822-2003
E-mail address: wschafer@ucsd.edu (W.R. Schafer).

0165-0270/02/$ - see front matter
PII: S0165-0270(02)00117-6

gish' animals. Since these categories are somewhat vague, and are always scored subjectively by a human observer, it is not uncommon for the same Unc mutant to be described differently by different researchers, or for two mutants with clearly distinguishable mutant phenotypes to be assigned the same classification. Moreover, many mutants with physiologically relevant defects in nervous system function exhibit only subtle alterations in behavior that are difficult for even an expert observer to reliably detect. Among the *C. elegans* genes with extremely subtle knockout phenotypes are the AMPA and NMDA glutamate receptor homologues *glr-1* and *nmr-1* (Zheng et al., 1999), the serotonin biosynthetic and reuptake transporter genes *tph-1* and *mod-5* (Sze et al., 2000; Ranganathan et al., 2001) and the dopamine biosynthetic gene *cat-2* (Lints and Emmons, 1999).

One way these problems have been surmounted is through use of automated video capture and analysis systems. By recording and analyzing the behavior of individual animals, often for long time intervals, it has been possible to rigorously identify and quantify deviations from wild-type behavior that are difficult to discern by eye (Pierce-Shimomura et al., 1999; de Bono et al., 1998). For example, we previously developed a computer-controlled motorized microscope that could record an individual animal's behavior at high magnification (Hardaker et al., 2001). To keep the animal from leaving the field of view, a tracking program was designed to control the movement of a motorized stage and maintain the worm in the center of the field. This system made it possible to follow the position of the animal over long time periods, and by recording the experiments on videotape, to analyze the details of locomotive and egg-laying behavior off-line. In this way, it was possible to obtain precise data on the timing of egg-laying events and to thereby develop a quantitative model for the temporal pattern of egg-laying that could be used to characterize and classify egg-laying-defective mutants (Waggoner et al., 1998; Zhou et al., 1998). Because the tracking system also retained information about the stage position and the position of the animal in the field of view, we could also quantify large-scale features of the animal's locomotor pattern (e.g. speed and directional changes) and investigate their correlation with egg-laying. Comparative analysis of behavioral mutants and animals carrying precise neuronal lesions identified specific serotonergic synapses that were critical for coordinating these behaviors (Hardaker et al., 2001). Taken together, the use of an automated tracking system made it possible to rigorously assay phenotypes that were essentially undetectable by eye and to distinguish functionally relevant phenotypic differences that revealed essential features about underlying molecular and neural mechanisms (Kim et al., 2001).

These preliminary studies suggested that more powerful computer vision tools could be used to generally characterize and classify the locomotor patterns of Unc mutants. Here we describe a new system we have devised for the automated collection and analysis of *C. elegans* locomotion data. Using this system, it has been possible to obtain reliable measurements of key features of an animal's locomotor pattern and to use these features to classify the locomotor patterns of individual mutant types. The behavioral data gathered by this system have many applications for the molecular analysis of the nervous system and behavior in this widely studied model organism.

2. Materials and methods

2.1. Strains and culture methods

Routine culturing of *C. elegans* was performed as described (Brenner, 1974). All worms analyzed in these experiments were young adults; fourth-stage larvae were picked the evening before the experiment and tracked the following morning after cultivation at 22 °C. We observed that animals tended to show higher locomotor activity immediately after being transferred to a fresh plate; thus, experimental animals were allowed to acclimate for at least 1 h before their behavior was analyzed. Plates for tracking experiments were prepared fresh the day of the experiment; a single drop of a saturated LB culture of *E. coli* strain OP50 was spotted onto a fresh NGM agar plate and allowed to dry for 1 h before use. The chromosomal locations of the genes studied in these experiments are as follows: LGI, *unc-38 (x20), goa-1 (n1134)*; LGIII, *unc-36 (e251)*; LGIV, *egl-19 (n582)*; LGX, *nic-1 (lj22)*.

2.2. Acquisition of image data

C. elegans locomotion was tracked with a Zeiss Stemi 2000-C Stereomicroscope mounted with a Cohu High Performance CCD video camera. A computer-controlled tracker (Parker Automation, SMC-1N) was used to put the worms in the center of the optical field of the stereomicroscope during observation (excluding the microscope, the components for this system cost approximately $10 000). To record the locomotion of an animal, an image frame of the animal was snapped every 0.5 s for at least 5 min. Among those image pixels with values less than or equal to the average value minus three times the S.D., the largest connected component was found. The image was then trimmed to the smallest axis-aligned rectangle that contained this component, and saved as eight-bit grayscale data. The dimensions of each image, and the coordinates of the center of mass of the worm in the tracker field were also saved simulta-

neously as the references for the location of an animal in the tracker field at the corresponding time point when the images are snapped. The stereomicroscope was fixed to its largest magnification ($50\times$) during operation. Depending on the type and the posture of a worm, the number of pixels per image frame varied although the number of pixels per millimeter was fixed at 312.5 pixel/mm for all worms.

2.3. Image pre-processing

To obtain the clean binary image, the background intensity level of the grayscale image was found first by taking the maximum of the values of the four corner points (top-left, top-right, bottom-left, bottom-right) of the trimmed image (at least one of the corner points is always not part of the worm body). After finding the background level (b), a 5×5 moving window was scanned over the trimmed image, and the mean (m) and S.D. (σ) of the pixels inside the window were computed at every pixel position. If the mean was less than $0.7b$ or the σ is larger than $0.3m$, then the pixel was considered to be a pixel of the worm body and is assigned a value 1. In order to clean up the spots inside the worm body, a morphological closing operator (binary dilation followed by erosion) was applied (Gonzalez et al., 2002). Next, the sequential algorithm for component labeling was used to remove unwanted isolated objects (Jain et al., 1995). The connected components were labeled by scanning the image in x and y directions sequentially, and the largest component was selected to guarantee that there will be only one object, the worm, in the image.

2.4. Image feature extraction

All of the software for binarization, skeletonization, and feature extraction was coded in C and implemented on a UNIX machine. Some features (e.g. the area of the worm, that is, the number of pixels which make up the single binary object in the frame) could be computed on a single frame; these were computed for all 600 frames in the sequence. The average value, the maximum value and the minimum value were then computed for these 600 measurements. In some cases, the maximum value or minimum value of a given parameter tended to be a relatively uninformative outlier; in such cases, it was more useful to summarize the group statistics with such quantities as the 90th and 10th percentile values out of the population of 600 numbers. Other features could not be extracted from a single frame, for example, the movement between two frames, or the movement within 10 s (20 frames). Since there are approximately 600 frames total in a sequence, the movement between two frames could be computed 300 times if we take pairs of frames in a non-overlapping fashion, or it could be calculated 599 times taking pairs of frames in a sliding window or overlapping fashion. Likewise, for the movement within 20 frames, we could compute 581 values for overlapping 20-frame intervals. Therefore, quantities of this type were calculated in a sliding window fashion. As before, the average, max, min, and other order statistics can be computed from this set of numbers. Average processing times of a single 600-frame clip are 75.9 and 58.5 s for the digitization and feature extraction, respectively on a personal computer (PC) with 1 GHz CPU.

2.5. Classification and regression tree analysis

The classification and regression tree (CART) algorithm for designing CARTs has its origins in a 1984 monograph by Breiman et al. (1984). Briefly, the CART approach involves recording a set of examples of each worm type (i.e. wild-type or a specific mutant), and measuring features that might in principle be used to distinguish different types. From these measurements, a training vector is generated for each recording consisting of an identifier of worm type along with the values for each feature measurement. Using this learning sample (which in our case consisted of 600 data points-100 of each of the 6 strains), CART produces a binary classification tree in which each binary split of the data involves a splitting question of the form 'Is $x_m \leqslant c$?' where x_m is one of the measurements, and c is a threshold. The root node of the tree contains all the training cases; the worm types are equally mixed together in this node. The goal of CART is to successively subdivide the training set using binary splits in such a way that the data associated with the terminal nodes of the tree do not have a mix of worm types; rather each node should be as 'pure' as possible. We used the Gini index of diversity to measure the impurity of a set of data. A class assignment rule assigned a class to every terminal node. A simple rule is to assign the most popular class for each terminal node; this is called the plurality rule, and is what we used. When two different classes were tied for the most popular class in the node, we arbitrarily chose the lower numbered class as the class for that node.

CART parameters are set as follows. First, we set ERROR CROSS = 10. This means that 10-fold cross validation is used to determine the optimal tree. Second, we set BOPTION SERULE = 0.0. This means that the number of standard errors to be used in the optimal tree selection rule is zero, so the tree with the smallest cost is chosen as the optimal tree. Third, we set BOPTION COMPLEXITY = 0.0. This means that no penalty is given for growing complex trees, so the largest possible tree is grown initially.

3. Results

3.1. A tracking system for the automated acquisition of nematode locomotion data

To classify and quantitatively characterize *C. elegans* Unc mutants, we built a tracking and imaging system that could not only follow an individual animal's position over long time periods, but could also save image data about the animal's body posture and movements. A schematic of this system is shown in Fig. 1. Briefly, the system consisted of (i) a dissecting microscope with a stereoscopic zoom; (ii) a motorized stage controlled by a serial port; (iii) a monochrome analog CCD camera; (iv) a PC with a video acquisition board. A VCR was also included for cross verification of behavioral tracking. During a recording, images from the CCD camera are captured and digitized at a frequency of up to 2 Hz. Next, a tracking program identified the animal in the field of view and saved a grayscale image of the worm, the stage position, the position of the worm in the field of view, and the time of capture. Thus, the system generated a time-coded sequence of images that in principle represented a nearly complete record of the animal's body movements over an indefinitely long time period. To facilitate parameter estimation, the grayscale images in the data structure were subjected to preliminary image processing to generate simplified representations of body position and shape (Fig. 2; see also Section 2).

3.2. Extracting image features for phenotype classification

We have developed tools to extract a wide range of features related to body shape and locomotor patterns from the image data gathered by this system. Methods for measuring some of the most useful features for phenotypic analysis are described below:

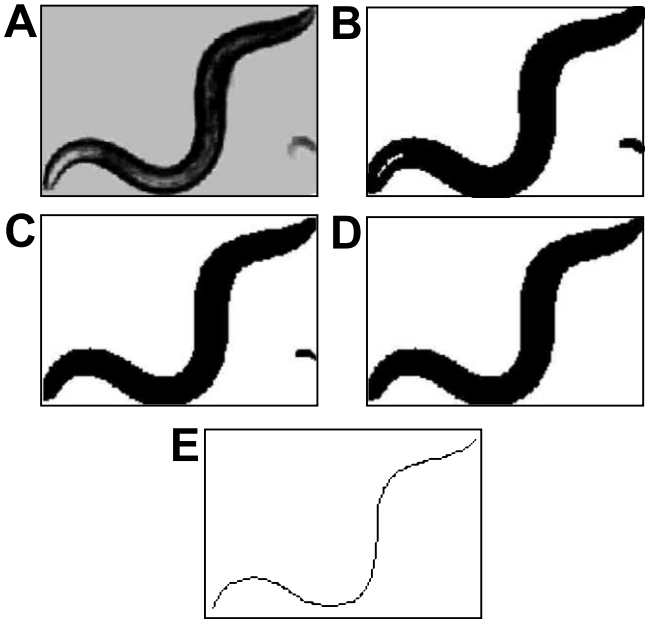

Fig. 2. Acquisition and processing of nematode image data. (A) Original grayscale image. Shown is a grayscale image of a wild-type hermaphrodite acquired by the tracker. Original image was 640 × 480 pixels. (B) Binary image. Image obtained by applying a threshold to the grayscale image in part (A). The threshold level was determined as described in Section 2. (C) Binary image following closing operation. Closing operation (Section 2) was applied to the image in part (B). Note the absence of holes in the processed binary image. (D) Final clean binary image. Small objects external to the worm were removed from the image in part (C) as described in Section 2. (E) Skeleton obtained through thinning and pruning.

3.2.1. Large-scale movement

Previous work indicated that speed and reversal frequency represent important features of many mutants' behavioral phenotypes. These features can be measured most simply by following the trajectory of the

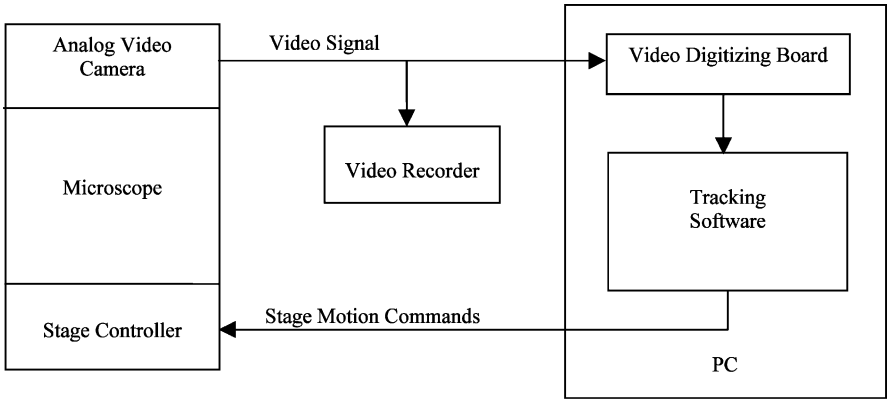

Fig. 1. Tracking and imaging system. The CCD video camera is fitted to the microscope and outputs analog video images to the digitizing board on the PC. The same analog video can also be recorded on a VCR. Customized tracking software computes the centroid of the worm and sends commands to the stage controller to re-center the field of view on the worm. The software has additional modules to collect a variety of behavioral data. Details on the components of the system are described in Section 2.

animal's centroid over time. To measure speed, the centroid position data are sampled over a constant time interval, and the worm's displacement is proportional to its average speed during that interval. Interval durations used in our experiments ranged from 0.5 s (1 frame), to 5 min (the total time of observation). To measure reversals, the trajectory of the centroid is sampled at intervals of constant distance (typically 30 pixels, which is one-tenth of the normal worm length). The turning angle at every vertex is computed; if the angle is greater than 120°, then the position is considered to be a reversal (Fig. 3A). A previous study implementing this approach found that over 90% of reversal events detectable by eye were identified using this method, and that greater than 99% of these events were in fact reversals rather than large turns (Hardaker et al., 2001).

3.2.2. Body size

The worm's area was obtained by determining the total number of 'on' pixels in the binary image. Likewise, the animal's length could be obtained by determining the number of pixels in the image skeleton. The worm thickness was measured at the center and head/tail positions of the worm skeleton (the center position was the value at the center of the skeleton pixel list; the head/tail position was defined as the position which is 7 pixels away from each end of the worm body). In order to measure the center thickness, we first took a 9-pixel-long segment from the skeleton list, and computed the best fit line for the segment by a line fitting algorithm. Then we rotated the line by 90° to get a perpendicular line to it (Fig. 3B). We traversed the perpendicular line in both directions from the center position until we reached the edges of the worm body, and then computed the distance between the two edges. We also rotated the perpendicular line by $-5°$ and $+5°$, and measured the thickness in those two directions. The minimum value of the three measurements was considered to be the center thickness. Similarly, in order to measure the head/tail thickness, we took two 9-pixel-long segments from each end of the skeleton list. After getting the best fit lines for the segments, we found the designated head/tail position by going back 7 pixels from the end of the worm body along the best fit line. Then we computed the thickness at these two measuring positions (one at each end) by traversing the perpendicular lines to the best fit lines. The minimum value of the two measurements was considered to be the head/tail thickness. We also define the worm's fatness as the ratio of worm area to length.

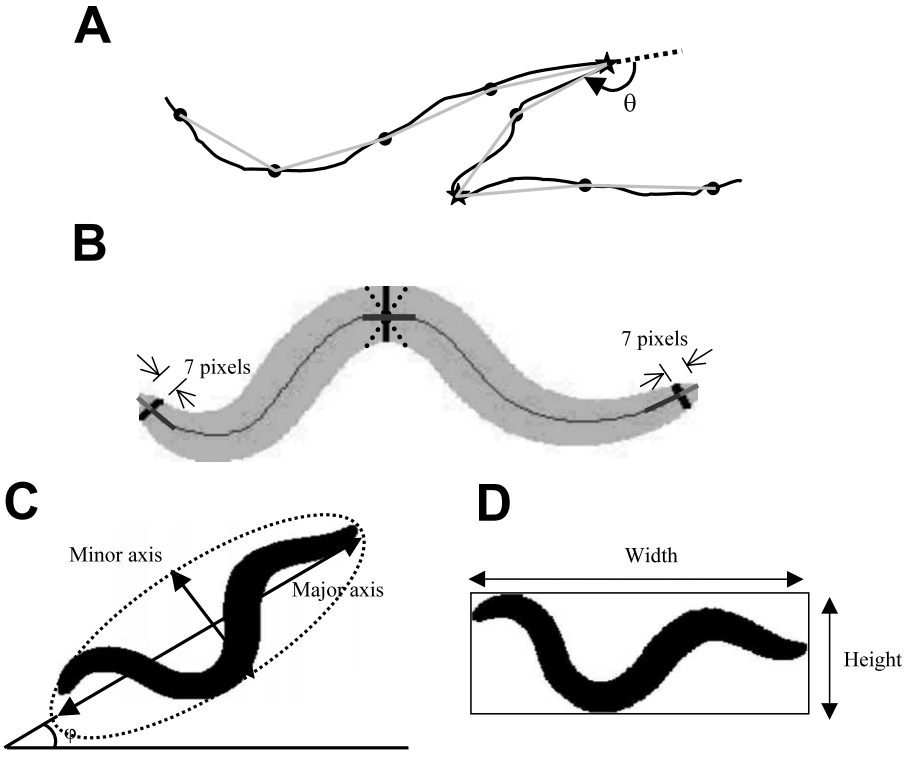

Fig. 3. Measurement of features based on large-scale movement and shape. (A) Directional change detection method. The trajectory of the worm's centroid (black solid line) is sampled at intervals of 30 pixels. The directional change position (mark with a star) is found by computing the angle deviation at every vertex of the polygon (gray line). If the angle (θ) is greater than 120°, then the position is considered to be a reversal. (B) Thickness measuring method. The length of a perpendicular cross-section of the binary image was computed at the center (i.e., the midpoint of the skeleton), head and tail (i.e. 7 pixels from the respective ends of the skeleton) as described in the text. (C) Best fit ellipse, and its associated parameters. (D) MER. The methods for deriving the best fit ellipse and MER are described in the text.

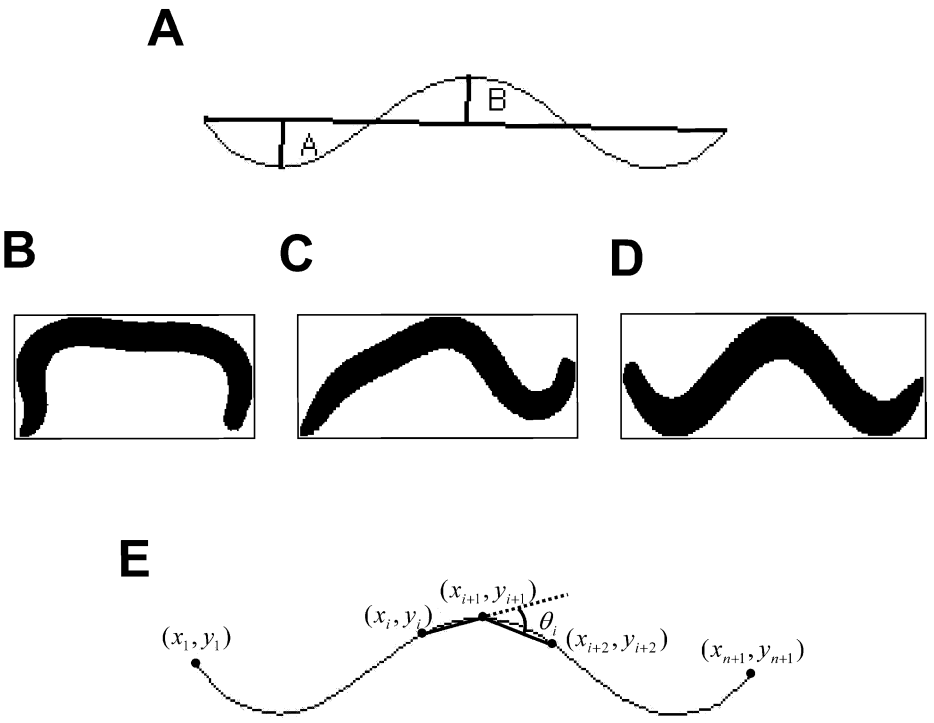

Fig. 4. Measurement of body curvature features. (A) Morphological skeleton with lower (A) and upper (B) maximum distance points along the straight line connecting two end points. The sum of the peak A and B distances is designated the animal's amplitude, and the ratio of min (A, B) to max (A, B) is designated the amplitude ratio. (B)–(D) Sample images and their amplitude ratios (B: $P = 0$, C: $P = 0.4$, D: $P = 0.97$). (E) Measurement of the angle change rate. As described in the text, the angle change rate is calculated by segmenting the skeleton using a constant distance of 10 pixels, and dividing the average angle difference between each two consecutive segments along the skeleton by the total worm length. Thus, a larger angle change rate means that a worm is more wavy.

3.2.3. Best fit ellipse and minimum enclosing rectangle

Several other parameters that correlated with both body size and body posture could be obtained by finding the best fit ellipse to the binary shape, and then using the length of the major and minor axes of the ellipse as shape features (Fig. 3C). An eccentricity variable was then computed as the ratio of the distance between the foci of the ellipse and its major axis length; this value (which is between 0 for a circle and 1 for a line) provided a measure of the elongation of the worm. By rotating the image according to the orientation of the best-fit ellipse's major axis, we could also obtain the minimum enclosing rectangle (MER) of the shape (Fig. 3D). The height and width of the MER, as well as the ratio

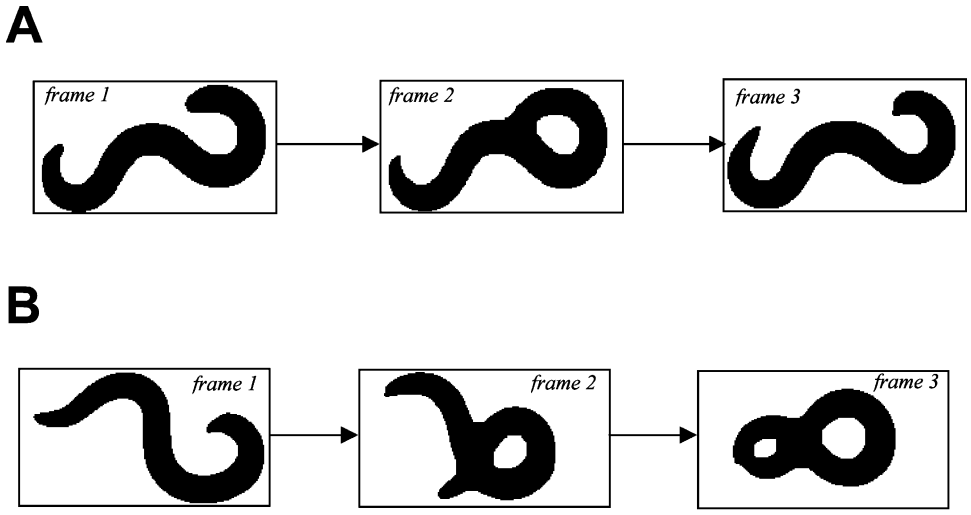

Fig. 5. Detection of coiled body postures. (A) Successive image frames showing an animal that coiled briefly. (B) Successive image frames from an animal that made multiple coils.

Table 1
All feature variables used in CART analysis (total 94 variables)

CART variable name	Description
MVHLFMIN, MVHLFMAX, MVHLFAVG	Min, max, average distance moved in 0.5 s
PRP10MIN, PRP10MAX, PRP10AVG	Min, max, average distance moved in 5 s or 10 frames
PRP20MIN, PRP20MAX, PRP20AVG	Min, max, average distance moved in 10 s or 20 frames
PRP30MIN, PRP30MAX, PRP30AVG	Min, max, average distance moved in 15 s or 30 frames
PRP40MIN, PRP40MAX, PRP40AVG	Min, max, average distance moved in 20 s or 40 frames
PRP50MIN, PRP50MAX, PRP50AVG	Min, max, average distance moved in 25 s or 50 frames
PRP60MIN, PRP60MAX, PRP60AVG	Min, max, average distance moved in 30 s or 60 frames
TOTMOVE	Total amount of movement in 5 min
RV80MIN, RV80MAX, RV80AVG	Min, max, average number of reversals in 40 s/80 frames
TOTRV	Total number of reversals in 5 min
AREAMIN, AREAMAX, AREAAVG	Min, max, average area of the worm
LENGTHMIN, LENGTHMAX, LENGTHAVG	Min, max, average length of the worm
CNTHKMIN, CNTHKMAX, CNTHKAVG	Min, max, average center thickness
HTTHKMIN, HTTHKMAX, HTTHKAVG	Min, max, average head/tail thickness
CNTLRMIN, CNTLRMAX, CNTLRAVG	Min, max, average ratio of center thickness to length
HTTLRMIN, HTTLRMAX, HTTLRAVG	Min, max, average ratio of head/tail thickness to length
FATMIN, FATMAX, FATAVG	Min, max, average fatness of the worm
MAJORMIN, MAJORMAX, MAJORAVG	Min, max, average length of best-fit ellipse's major axis
MINORMIN, MINORMAX, MINORAVG	Min, max, average length of best-fit ellipse's minor axis
ECCTYMIN, ECCTYMAX, ECCTYAVG	Min, max, average best-fit ellipse's eccentricity
LNECRMIN, LNECRMAX, LNECRAVG	Min, max, average ratio of worm length to eccentricity
WDTHMIN, WDTHMAX, WDAVG	Min, max, average width of MER
HGHTMIN, HGHTMAX, HGHTAVG	Min, max, average height of MER
WHRATMIN, WHRATMAX, WHRATAVG	Min, max, average width-to-height ratio of MER
MERFLMIN, MERFLMAX, MERFLAVG	Min, max, average ratio of worm area to MER area
LNMFRMIN, LNMFRMAX, LNMFRAVG	Min, max, average ratio of worm length to MER fill
AMPMIN, AMPMAX, AMPAVG	Min, max, average amplitude of worm skeleton wave
AMPRMIN, AMPRMAX, AMPRAVG	Min, max, average amplitude ratio
ANCHRMIN, ANCHRMAX, ANCHRAVG	Min, max, average angle change rate
ANCHSMIN, ANCHSMAX, ANCHSAVG	Min, max, average S.D. of angle change
LPRUNMIN, LPRUNMAX, LPRUNAVG	Min, max, average length of time the worm remains coiled
LPSUM	Total number of frames the worm is looped
LPMLTSUM	Total number of frames the worm has multiple loops
LPRUNS	Total number of times the worm starts to coil
LPBRIEF	Total number of times the worm briefly loops
LPBRFRUN	Percentage of brief runs to total runs

between these two, are useful features that can indicate whether the worm tends to take on elongated positions with low amplitude waves, or, on the contrary, tends to have deeper body bends or looped body positions. The approximate amplitude of the worm skeleton wave is also found by computing the perpendicular distance from every skeleton point to the line connecting the two end points of the skeleton. Let the coordinates of a point on the skeleton be (x_i, y_i) and the equation of the line $ax + by + c = 0$. Then the perpendicular distance (d) from the point to the line can be obtained by the following equation:

$$d = \frac{|ax_i + by_i + c|}{\sqrt{a^2 + b^2}}.$$

The sum of the two maximal distances (A and B) is considered to be the amplitude of the skeleton (Fig. 4A). Then amplitude ratio (P) is defined by the following equation:

$$P = \frac{\min(A, B)}{\max(A, B)}.$$

Some examples of worms with different amplitude ratios are shown in Fig. 4. As shown in Fig. 4B–D, this variable was a quantitative identifier of various atypical body postures, including the so-called omega turn, which has an amplitude ratio of zero.

The angle change rate and its S.D. are also computed from the skeleton of the worm (Fig. 4E). The skeleton points are spaced apart by 10 pixels. Then the angle change rate (R) is defined as the ratio of the average angle difference between every two consecutive segments along the skeleton to the worm length, which can be represented by the following equations:

$$R = \left\{ \frac{1}{n-1} \sum_{i=1}^{n-1} \theta_i \right\} / L,$$

$$\theta_i = \arctan \frac{y_{i+2} - y_{i+1}}{x_{i+2} - x_{i+1}} - \arctan \frac{y_{i+1} - y_i}{x_{i+1} - x_i},$$

where n is the number of segments and L is the worm

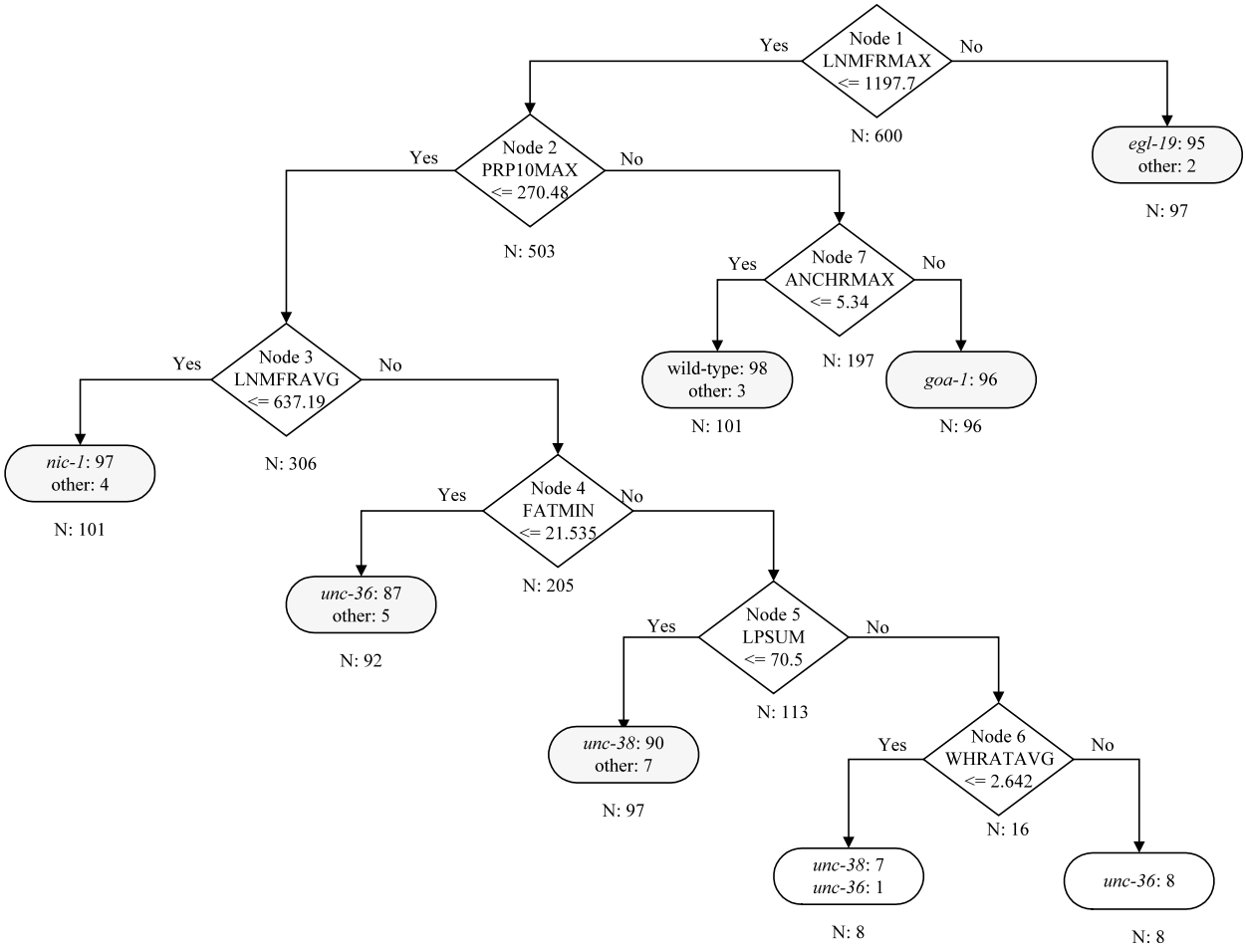

Fig. 6. Optimal classification tree. The tree was constructed using the CART algorithm as described. The number of total animals in each node (N:) is indicated below the respective node; the number of animals of a particular type in that node is indicated within the node.

length. Note that a larger angle change rate means that a worm is more wavy.

The amount of time a worm spent in a coil as well as how often it coiled are unique behavioral characteristics of several types of worms. A coiled body posture creates a 'hole' in the image where the worm loops or touches itself. To identify coiled postures, we searched for 'holes' in the worm image by performing connected component labeling on the inverted image (Jain et al., 1995). Counting up the number of connected objects will always give a value of at least one for the background; thus the number of holes is equal to the number of connected components minus one. In our subsequent analysis, we counted the number of frames the worm was in a coiled posture as well as the number of times the worm switched from a non-coiled to a coiled posture (i.e. the number of 'runs'). The length of time the worm remained coiled was charactrized by finding the minimum, maximum, and average of the run lengths. We also counted the total number of times the worm briefly looped (Fig. 5A) and the total number of frames the worm had multiple loops (Fig. 5B).

3.3. Classification of representative mutant types using image features

Based on qualitative descriptions of Unc mutant phenotypes, we expected the features measured by our system would provide useful quantitative definitions for specific mutant types. To assess the ability of these image features to provide effective characterization of *C. elegans* mutant phenotypes, we tested the ability of the CART algorithm, a classifier used in machine vision and medical diagnostic applications, to distinguish data obtained from different mutant types from one another and from wild-type.

In our application of this algorithm, the measurement vector consisted of 94 measured features (Table 1) associated with an image sequence of a single worm. For our initial test of the CART approach, we compared data obtained from wild-type worms with five different Unc mutants with distinct locomotor patterns. Thus, the training set contained a total of six classes: wild-type, and loss-of-function mutants defective in the genes *unc-36* (encoding a calcium channel α-2

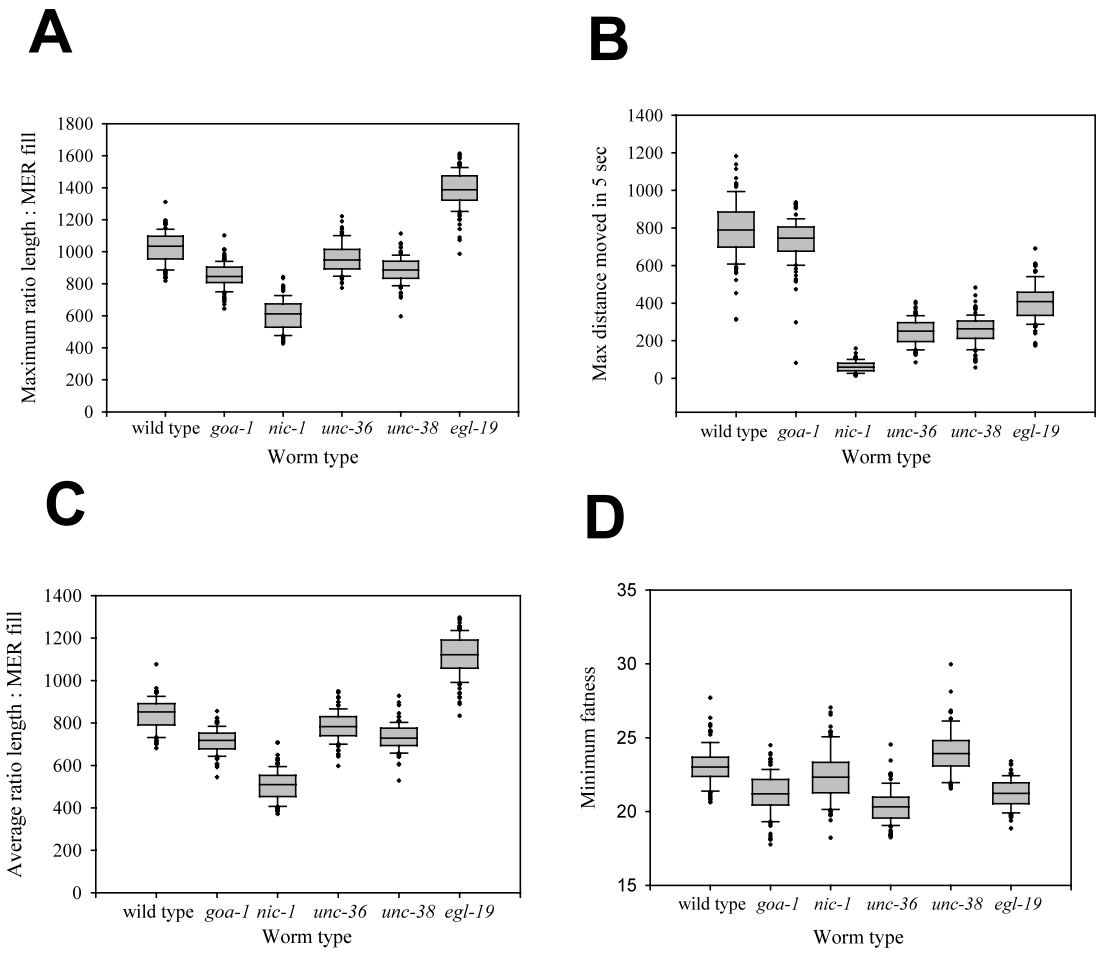

Fig. 7. Characterization of Unc mutant phenotypes using image feature parameters. Shown are the feature measurement distributions for each of the six initial worm types analyzed in this study. (A) Maximum length-to-MER fill ratio. (B) Maximum distance moved in 5 s. (C) Average length-to-MER fill ratio. (D) Minimum fatness. In all cases, the box extends from the first quartile (25th percentiles) to the third quartile (75th percentiles), and the horizontal line within the box indicates the median. The lower and upper error bars indicate 10th and 90th percentiles, respectively; each outlier is indicated with a dot.

subunit), *unc-38* (encoding a nicotinic receptor subunit), *goa-1* (encoding the Go α subunit), *egl-19* (encoding the L-type calcium channel α-1 subunit) and *nic-1* (encoding a type-1 glysosyltransferase) were analyzed. For each strain, the training set consisted of 100 5-min recordings, with images captured every 0.5 s.

We found that it was possible to generate an optimal classification tree that could reliably identify the type of a given worm using only 7 parameters (Fig. 6). These included measures of body size or shape (minimum fatness), large scale locomotion (maximum distance moved in 5 s or 10 frames), and correlates of body posture (maximum and average ratio of length/MER fill, average width/height ratio of MER, total frames in looped posture, and maximum angle change rate for skeleton). Not only could these features be used to identify individual animals of a given type, but the distinctive pattern in which these features varied between Unc mutants also provided a characteristic behavioral signature for each mutant type (Figs. 7 and 8; Table 2).

3.4. Reliability of automated phenotype identification

To assess the reliability of our system at correctly identifying animals of a given mutant type, we performed cross-validation analysis. A measure of classifier success is the 'true misclassification rate,' that is, the fraction of new cases that would be misclassified by this classifier. If a classifier is designed using the samples in L, the resubstitution estimate of misclassification error is obtained by running the samples in L through the classifier to see how many of them get misclassified. Growing a larger and larger tree will reduce this resubstitution estimate of misclassification error, until such time as each terminal node is completely pure and the resubstitution estimate is zero. However, such a tree is unlikely to perform well on other data. Therefore, we

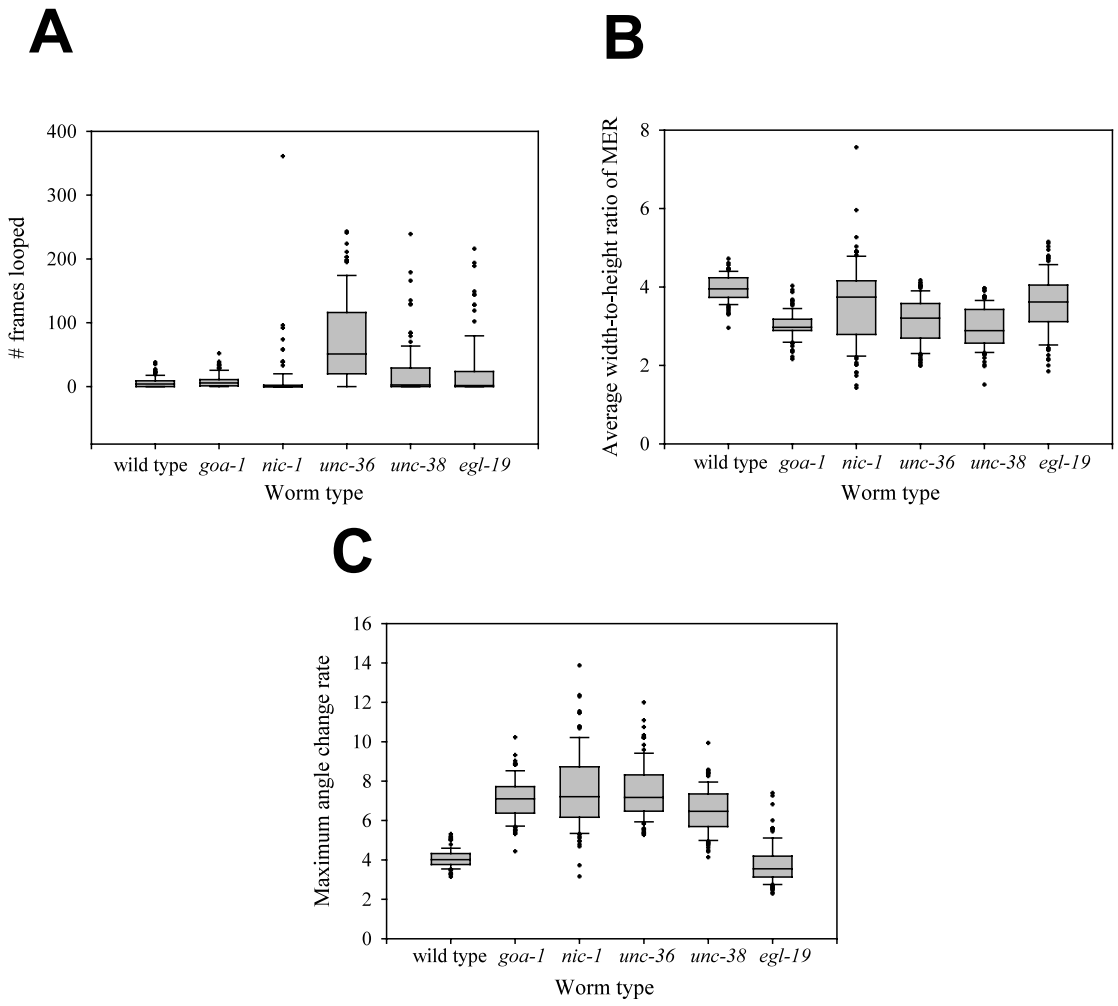

Fig. 8. Characterization of Unc mutant phenotypes using image feature parameters. Shown are the feature measurement distributions for each of the six initial worm types analyzed in this study. (A) Total number of 'looped' frames. (B) Average width-to-height ratio of the MER. (C) Maximum angle change rate of the image skeleton. In all cases, the box extends from the first quartile (25th percentiles) to the third quartile (75th percentiles), and the horizontal line within the box indicates the median. The lower and upper error bars indicate 10th and 90th percentiles, respectively; each outlier is indicated with a dot.

instead used 10-fold cross-validation to design a tree of the right size and to estimate the misclassification rate.

To perform 10-fold cross validation, the entire learning sample was divided into 10 roughly equal parts, each containing a similar distribution of the dependent variable (i.e. the worm type). Nine tenths of the data were used to construct a very large tree, and the remaining tenth of the data was used to estimate the error rate of selected sub-trees. The process was repeated until each part of the data had been held in reserve one time as a test sample. The results of the 10 mini-test samples were then combined to estimate error rates for trees of each possible size. The cross-validated relative cost and resubstitution relative cost versus the number of terminal nodes are shown in Fig. 9. Since we consider the cost of misclassifying a worm of one type as being another type to be the same for all types, the term 'cost' simply means 'misclassification rate' in our application. Note that the cross-validated relative cost is an error measure based on a test sample, while the resubstitution relative cost is an error measure based on the learning sample. The cross validation classification probability for each type is given in Table 3. The success rates are listed along the diagonal while the off-diagonal entries represent the misclassification error rates. From this, we can see that for all types, the prediction success rates of the optimal classification tree were approximately 90% or higher for each sample animal. Thus, the classifier showed a high degree of success at identifying the correct mutant type even if presented with a single example recording. The success of the classification would be even higher if one based the classification on multiple recordings. For example, if a single worm is classified correctly 93% of the time, then if one takes a group of three or five worms of the same type, and uses a majority vote of their classification results, the correct

Table 2
Statistics of the variables used in the classification tree

Variable	Statistics	Worm type					
		Wild-type	*goa-1*	*nic-1*	*unc-36*	*unc-38*	*egl-19*
LNMFRMAX	Min	818.50	644.74	426.39	774.81	596.49	987.42
	Max	1311.54	1102.03	843.08	1221.75	1113.80	1613.84
	Mean	1023.96	846.62	605.26	961.08	884.38	1386.09
	S.D.	95.34	78.25	96.98	93.14	80.86	115.57
PRP10MAX	Min	312.49	82.47	14.37	84.98	57.84	174.29
	Max	1182.38	936.56	159.04	407.21	483.21	691.21
	Mean	794.56	731.48	61.03	246.87	257.44	402.00
	S.D.	156.00	126.31	28.50	69.44	78.37	98.23
LNMFRAVG	Min	681.73	545.07	371.71	597.96	528.68	834.15
	Max	1076.61	856.50	708.20	949.29	927.93	1296.60
	Mean	842.36	715.39	505.86	784.62	733.44	1116.23
	S.D.	72.39	56.19	71.26	67.73	62.33	96.67
FATMIN	Min	20.63	17.77	18.23	18.25	21.55	18.86
	Max	27.71	24.49	27.05	24.54	29.97	23.40
	Mean	23.07	21.21	22.40	20.37	24.08	21.25
	S.D.	1.31	1.40	1.72	1.14	1.50	0.93
LPSUM	Min	0.00	0.00	0.00	0.00	0.00	0.00
	Max	38.00	52.00	361.00	243.00	239.00	216.00
	Mean	6.27	8.43	9.46	72.36	22.23	23.38
	S.D.	7.86	10.21	39.44	65.37	40.91	45.20
WHRATAVG	Min	2.96	2.16	1.43	1.99	1.51	1.85
	Max	4.72	4.03	7.56	4.17	3.97	5.15
	Mean	3.96	3.02	3.56	3.15	2.96	3.59
	S.D.	0.33	0.33	1.02	0.57	0.52	0.72
ANCHRMAX	Min	3.14	4.44	3.16	5.27	4.14	2.29
	Max	5.32	10.23	13.88	12.00	9.94	7.40
	Mean	4.07	7.09	7.55	7.49	6.52	3.78
	S.D.	0.45	1.03	1.99	1.39	1.10	1.01

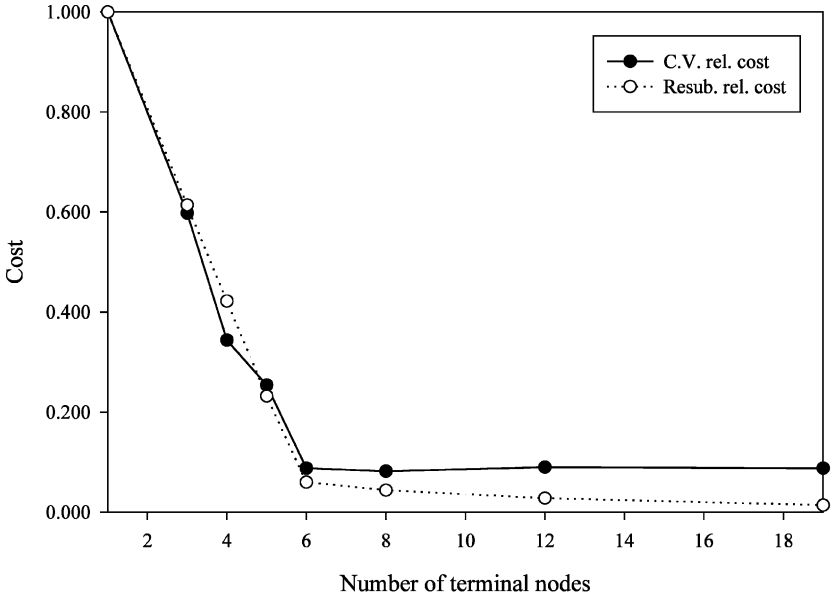

Fig. 9. Cross-validated relative cost and resubstitution relative cost verses the number of terminal nodes. The cross-validated relative cost is an error measure based on a test sample, while the resubstitution relative cost is an error measure based on the learning sample. Unlike the cross-validated relative cost, the resubstitution relative cost always decreases as the number of terminal nodes increases. A tree with eight terminal nodes is optimal, because the cross-validated relative cost becomes worse again as the tree grows.

Table 3
Cross validation classification probability table

		Predicted worm types					
		Wild-type	*goa-1*	*nic-1*	*unc-36*	*unc-38*	*egl-19*
Actual worm types	Wild-type	0.96	0.01	0	0	0.02	0.02
	goa-1	0.03	0.95	0.02	0.01	0.01	0
	nic-1	0	0	0.92	0.01	0.05	0.01
	unc-36	0	0	0.02	0.89	0.08	0.02
	unc-38	0.01	0	0.04	0.02	0.93	0
	egl-19	0.02	0	0	0.03	0.01	0.94

classification would be achieved 98.6% (for three) or 99.7% (for five) of the time.

4. Discussion

4.1. Applications for machine vision-based behavioral quantification in C. elegans

This study represents the first step toward developing computer vision methods for characterizing different classes of Unc mutants and distinguishing them from one another. Starting with a small but representative set of mutants exhibiting distinct Unc phenotypes, we have shown that individual animals of a particular mutant type can be reliably identified on the basis of a small number of image features collected by an automated tracking and image processing system.

Computer vision methods offer a number of clear advantages over real-time observation for the characterization of behavioral phenotypes. First, these approaches provide a specific, quantitative definition of a particular mutant phenotype, facilitating quantitative comparisons between different mutant strains. Second, a computerized imaging system has the potential to be much more reliable at detecting abnormalities that are subtle or manifested over long time scales. Finally, a computerized system makes it possible to comprehensively assay multiple aspects of behavior simultaneously, yielding a complex phenotypic signature that can be highly diagnostic of a specific molecular defect. Although this study focused on the analysis of phenotypes associated with abnormal locomotion, it should be noted that many of the parameters used in the CART approach will allow us to obtain automated phenotypic data on other aspects of nematode behavior and development, in particular those that affect morphology. We expect that the development of automated tools for phenotypic analysis will provide reliable, comprehensive analysis of behavioral abnormalities that would normally require the efforts of a battery of expert human observers. Such information would not only allow a more precise understanding of the relationship between genes and behavior in this organism, but also would make it possible to identify genes affecting the activity of common molecular targets in the nervous system.

Another potentially important application of quantitative image analysis to *C. elegans* neurobiology is to investigate molecular mechanisms of drug response. One way to use worm genetics to identify drug targets is to isolate a mutant with altered response to a given compound. Because such mutants have alterations in nervous system function, they frequently display abnormal behavioral patterns in the absence of drug. By quantitatively characterizing these patterns of behavior and then searching for additional mutants with a similar behavioral pattern, it may be possible to identify additional genes whose products function in the same process. A second, complementary approach would be to treat wild-type animals with a given compound and characterize the behavioral pattern caused by drug treatment itself. In principle, if the drug inhibits the activity of a specific gene product, it should induce a similar behavioral abnormality to that caused by a loss-of-function mutation in the gene encoding the target molecule. By comparing the behavioral patterns of drug-treated animals with those of known mutants, it should therefore be possible to make informed initial hypotheses about a drug's target. To successfully apply this approach, it will be critical to accumulate behavioral data on a wider range of Unc mutants and to generate a database correlating each mutant gene with its characteristic behavioral pattern. A comprehensive behavioral database would be extremely powerful for identifying groups of mutants and pharmacological treatments that have similar effects on behavior or development. With the accumulation of increasing phenotypic data on known mutants, it should ultimately be possible to record from mutant or drug-treated animals and quickly gain insight into the molecular pathways affected by a given gene or pharmacological agent.

References

Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Belmont, CA: Wadsworth, 1984.

Brenner S. The genetics of *Caenorhabditis elegans*. Genetics 1974;77:71–94.

de Bono M, Bargmann CI. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. Cell 1998;94:679–89.

Gonzalez R, Woods R. Digital image processing, 2nd ed.. Prentice-Hall, 2002.

Hardaker LA, et al. Serotonin modulates locomotory behavior and coordinates egg-laying and movement in *Caenorhabditis elegans*. J Neurobiol 2001;49:303–13.

Hodgkin J. Male phenotypes and mating efficiency in *Caenorhabditis elegans*. Genetics 1983;103:43–64.

Jain R, Kasturi R, Schunck BG. Machine vision. McGraw-Hill Inc, 1995.

Kim J, et al. Genes affecting the activity of nicotinic receptors involved in *C. elegans* egg-laying behavior. Genetics 2001:1599–610.

Lints R, Emmons SW. Patterning of dopaminergic neurotransmitter identity among *Caenorhabditis elegans* ray sensory neurons by a TGFB family signaling pathway and a Hox gene. Development 1999;126:5819–31.

Pierce-Shimomura JT, Morse TM, Lockery SR. The fundamental role of pirouettes in *Caenorhabditis elegans* chemotaxis. J Neurosci 1999;19:9557–69.

Ranganathan R, et al. Mutations in the *Caenorhabditis elegans* serotonin reuptake transporter MOD-5 reveal serotonin-dependent and -independent activities of fluoxetine. J Neurosci 2001;21:5871–84.

Riddle DL, et al, editor. *C. elegans* II. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1997.

Sze JY, et al. Food and metabolic signaling defects in a *Caenorhabditis elegans* serotonin-synthesis mutant. Nature 2000;403:560–4.

Waggoner L, et al. Control of behavioral states by serotonin in *Caenorhabditis elegans*. Neuron 1998;21:203–14.

Zheng Y, et al. Neuronal control of locomotion in *C. elegans* is modified by a dominant mutation in the *GLR-1* ionotropic glutamate receptor. Neuron 1999;24:347–61.

Zhou GT, Schafer WR, Schafer RW. A three-state biological point process model and its parameter estimation. IEEE Trans On Signal Processing 1998;46:2698–707.