Modeling the ballistic-to-diffusive transition in nematode motility reveals variation in exploratory behavior across species

Stephen J. Helms*, W. Mathijs Rozemuller*, Antonio Carlos Costa*, Leon Avery3, Greg J. Stephens2,4, and Thomas S. Shimizu†

1 AMOLF Institute, Amsterdam, The Netherlands
2 Dept. of Physics & Astronomy, Vrije Universiteit, Amsterdam, The Netherlands
3 Dept. of Physiology and Biophysics, Virginia Commonwealth Univ., Richmond, VA, USA
4 Okinawa Institute of Science and Technology, Onna-son, Okinawa, Japan

SI Materials and Methods

Selection of Strains

A phylogenetic tree with the strains used in this study is shown in Figure 1A. The nematode phylum is classically divided into three major branches—chromadorea, enoplea, and dorylaimia—that are broken into a total of five major B-clades [S2] and twelve minor H-clades [S3]. The chromadorean lineage is the largest, spanning B-clades III-V and H-clades 3-12 [S2,S3]. *C. elegans* is located in clade V9 (the rhabditids), one of the most diverse clades [S4]. In addition to the lab strain N2, we selected three of the most genetically distinct wild isolates of *C. elegans* (CB4856, JU775, and QX1211) to sample intraspecies variation [S5]. From H-clade 9 in order of increasing evolutionary distance, we selected *Caenorhabditis briggsae* JU757, *Rhabditis myriophila*
DF5020, and *Pristionchus pacificus* PS312. The next closest major group, B-clade IV, contains H-clades 10-12. H-clade 12 contains the plant parasitic tylenchs and was thus not included in this study. H-clades 10 and 11 contain many bacterial feeders, of which we selected *Panagrolaimus* sp. PS1159. Finally, from the basal chromadorea, we obtained *Plectus* sp. sjh2, a member of H-clade 6.

*C. elegans* N2, CB4856 and JU775 were provided by the Caenorhabditis Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). *C. elegans* QX1211 was kindly provided by Erik Andersen (Northwestern Univ.). *Plectus* sp. sjh2 was isolated from a soil sample using morphological criteria by Casper Quist and Hans Helder (Wageningen Univ.). SJH then isolated a single species by starting cultures with a single worm. The remaining strains were used in previous studies by Avery [S6].

**Cultivation of Worms**

Worms were grown on NGM-SR plates (3 g NaCl, 24 g agar, 2.5 g peptone, 1 mL 5 mg mL⁻¹ cholesterol in EtOH in 975 mL water, with 1 mL 1 M CaCl₂, 1 mL 1 M MgSO₄, 25 mL 1 M K₂PO₄ pH 6, 1 mL 200 mg mL⁻¹ streptomycin in water, and 0.23 g 5 mL 40 mg mL⁻¹ nystatin in DMSO, added after autoclaving) seeded with *E. coli* HB101, as previously described [S7]. *E. coli* HB101 was first cultured in M9 minimal media (3 g KH₂PO₄, 6 g Na₂HPO₄, 5 g NaCl, 1 mL 1 M MgSO₄ in 1 L water) supplemented with 10% Luria broth and 10 mg mL⁻¹ streptomycin [S8]. Plates were incubated with a light circle of HB101 culture for a day at 37°C and then stored at 4°C. For *Plectus* sp. sjh2, low salt plates (2% agar supplemented with 5 mg L⁻¹ of cholesterol from a 5 mg mL⁻¹ EtOH solution) were used as previously described [S9]. On NGM-SR plates, these worms became shrunken and died. As the plates did not have nutrients for the bacteria to grow, HB101 was grown to high density in Luria broth overnight at 37°C, washed 3X in water, resuspended at 10X concentration, and applied to the plates.

Nematodes were cultured by either transferring a few worms by worm pick or a chunk of agar to a new plate after the worms reached adulthood. The plates were then incubated at 20°C. The growth rate varied considerably among strains, with *Plectus* sp. sjh2 taking nearly two weeks to reach adulthood. We avoided starving the worms at any point during their cultivation, especially in the period before behavioral experiments were performed, as this can induce transgenerational phenotypic changes [S10,S11], and we have observed transient effects on motility lasting at least a couple of generations (data not shown).

**Imaging**

The imaging experiments were done on 3.5 cm plates containing the same media used for cultivation. A 2×2 10 mm repellant grid was made by etching the plate with a tool dipped in 1% sodium dodecyl sulfate, a detergent that *C. elegans* and most other nematodes avoided. (Whereas many *C. elegans* studies have used copper rings as a repellant boundary [S12], we found that it did not
sufficiently repel other nematodes; data not shown). Four young adult, well-fed nematodes were transferred individually by worm pick to a 10 µL drop of M9 (water for *Plectus sp.* sjh2) to remove bacteria stuck to the worms. The worms were then transferred by pipette in a minimal amount of buffer to the imaging plate, and excess buffer was removed as much as possible. The plate was imaged 10-20 minutes after picking the worms, minimizing most transient behaviors. The plate was placed on a custom imaging rig in an inverted, uncovered configuration with illumination by a Schott MEBL-CR50 red LED plate. The behavior was recorded for 30 minutes using a Point Grey Grasshopper Express GX-FW-60S6M-C camera equipped with an Edmund Optics NT54-691 lens (set to a magnification of 0.5X) at a resolution of 2736×2192 (12.5 µm/px) at 11.5 frames/s using a custom National Instruments LabView acquisition program. The video was subsequently compressed using the open-source XViD MPEG-4 compression algorithm using maximal quality settings.

**Tracking and Image Analysis**

The behavioral videos were analyzed using a custom automated analysis program in MathWorks Matlab. The average background was calculated from 50 frames evenly sampled across the entire video. The background was then subtracted from each frame and a global threshold was applied. The thresholded image was cleaned by applying a series of morphological operations: Incomplete thresholding of the worm was smoothed by applying morphological closing with a disk with a similar radius as the worm. Any remaining holes were filled in using a hole-filling algorithm. Small holes or ones with a low perimeter to area ratio were excluded as they sometimes fill in worms undergoing an omega turn, as described in [S13]. Finally, regions in which the worm was just barely touching itself were split by sequentially applying open, diagonal fill, and majority morphological operations. The worm was then identified as the largest connected component with an area within 2-fold of the expected value. The centroid was tracked across frames to obtain \( \mathbf{x}(t) \). In addition, the image skeleton was calculated. Sample images from each of the processing steps are shown in Figure S11.

The head of the worm was automatically identified using two statistical properties of the worm’s behavior, namely (i) on average, the head of the worm moves more than the tail, and (ii) on average, worms spend more time moving forward (in the direction of their head) than they do moving in reverse. The procedure is based on skeletonization and centroid detection of the worm image, which can fail in situations where image contrast is low (e.g. due to non-uniform background), so trajectories were first divided into segments that contain no more than 3 frames missing the skeleton and centroid information, and the head orientation was assigned within each segment based on local behavioral statistics. Finding statistical criteria that allow unambiguous assignment of head orientation across all strains studied here was challenging because of the diversity in their behavior, but the following procedure was found to work well empirically. The identity of the two ends of the skeleton across image frames
were accounted for by a simple tracking algorithm based on minimizing the total distance between skeleton points. For segments longer than 150 frames (with no more than ten consecutive missing skeletons), we found that we could apply property (i) by computing the variance in body angles within 10% of the body length from the ends, and assigning the head to the end with the greater summed variance. However, manual inspection revealed that this sporadically resulted in misassignment of the head, identifiable as long reversals interrupted by short forward runs. Therefore, in addition, for segments longer than 200 frames (with no more than five consecutive missing centroids), we used property (ii), defining the head as the end of the skeleton that spent the majority of the trajectory at the leading edge of movement. Segments shorter than 150 frames were discarded from further analysis.

The velocity $\vec{v}(t)$ was calculated from the centroid position $\vec{x}(t)$ using the derivative of a cubic polynomial fit to a sliding 1 s window. The direct estimation of the velocity using a symmetrized derivative had a large $\delta$-correlated component that interfered with later analysis. The use of the cubic polynomial did not noticeably distort the correlation functions (Figure S12). When the worm’s speed $s(t) = |\vec{v}(t)|$ is very low, its projections on the lab-frame x- and y-axes $v_x = \vec{v}(t) \cdot \hat{x}$ and $v_y = \vec{v}(t) \cdot \hat{y}$ become dominated by discretization (pixelation) noise, and the bearing $\phi(t) = \tan^{-1}(v_y/v_x)$ is poorly defined. This in turn leads to large fluctuations in $\Delta \psi(t) = \phi(t) - \psi(t)$, which can introduce a large number of false reversal events, noticeable as a steep decrease in the autocorrelation $C_{\Delta \psi}(\tau) = \langle \cos(\Delta \psi(t + \tau) - \Delta \psi(t)) \rangle$ at small values of the delay $\tau$. We therefore exclude segments of the trajectories corresponding to run intervals shorter than six frames (less than half a second). When these artifacts are filtered out in this manner, the $\Delta \psi$ autocorrelation functions were well described by single exponentials (Figure S7C). We note that the exclusion of short runs effectively excludes segments of data in which the worm remains stopped (or at a very low speed) — a feature that is more pronounced in some strains than others — and this leads to a systematic bias for simulated model trajectories to have a higher effective diffusivity $D_{\text{eff}}$ than the data for the corresponding strain (as can be seen in Figure 5C).

Calculation of Behavioral Statistics

The worm’s behavior fluctuated or sometimes drifted over long times (Figure S4), but the average statistics over 100 s windows were approximately stationary. In order to focus on dynamics within the 100 s timescale, the mean-squared displacement and all auto- and cross-correlation functions were calculated for 100 s windows and then averaged. This reduced the influence of longer timescale fluctuations in the speed and reversal rate. For all calculations, observations near the boundaries and pairs of points between which the worm approached the boundary were excluded. The uncertainty of each individual’s phenotype projection on the principal behavioral mode was computed by projecting the motility parameters after bootstrapping over the 100 s windows of each individual’s trajectory. The standard deviation of the bootstrapped projections is used
as uncertainty.

**Calculation of Effective Diffusivity, \( D_{\text{eff}} \)**

To estimate the effective diffusivity \( D_{\text{eff}} \), we fit the mean-squared displacement \( \langle [\Delta x(\tau)]^2 \rangle \) over the diffusive regime. For this purpose, we defined the diffusive regime as the time-lag interval after which the normalized velocity autocorrelation \( C_v(\tau)/C_v(0) \) decayed to below 0.1. We note that in some cases (especially for fast-moving strains such as CB4856, JU775 and sjh2) the fit to \( \langle [\Delta x(\tau)]^2 \rangle = 4D_{\text{eff}}\tau \) in this regime was poor due to boundary effects arising from the finite size of the behavioral arena. For these strains, \( D_{\text{eff}} \) should be regarded as a lower bound for the true diffusivity.

**Reversal Analysis**

The reversal state was assigned as described in the main text by analysis of \( \Delta \psi(t) \). Assuming a random telegraph process that generates states \( \Delta \psi = 0 \) (forward) and \( \Delta \psi = \pi \) (reverse) with probabilities \( 1 - f_{\text{rev}} \) and \( f_{\text{rev}} \), respectively, the autocorrelation at long time lags is \( C_{\Delta \psi}(\tau \to \infty) = (1 - 2f_{\text{rev}})^2 \). For the proposed telegraph process, each state has an exponentially distributed lifetime \( (\tau_{\text{fwd}}, \tau_{\text{rev}}) \) and therefore \( f_{\text{rev}} = \tau_{\text{rev}}/\tau_{\text{rev}} + \tau_{\text{fwd}} \). The expected correlation timescale for the mixture of the two states is \( \tau_{\text{RT}}(\tau_{\text{rev}}, \tau_{\text{fwd}}) = (1 + \frac{1}{f_{\text{rev}}})^{-1} \). The \( \Delta \psi \) autocorrelation function was therefore fit to

\[
C_{\Delta \psi}(\tau) = [1 - C_{\Delta \psi \infty}(\tau_{\text{rev}}, \tau_{\text{fwd}})] \exp \left[ -\frac{\tau}{\tau_{\text{RT}}(\tau_{\text{rev}}, \tau_{\text{fwd}})} \right] + C_{\Delta \psi \infty}(\tau_{\text{rev}}, \tau_{\text{fwd}})
\]

where \( C_{\Delta \psi \infty}(\tau_{\text{rev}}, \tau_{\text{fwd}}) = \frac{(\tau_{\text{fwd}} - \tau_{\text{rev}})^2}{\tau_{\text{rev}} + \tau_{\text{fwd}}} \). The fraction of time spent reversing is: \( f_{\text{rev}} = 0.5 - \sqrt{C_{\Delta \psi \infty}(\tau_{\text{rev}}, \tau_{\text{fwd}})/4}, \) where \( f_{\text{rev}} \in [0, 0.5] \). The transition time constants are then \( \tau_{\text{rev}} = \frac{\tau_{\text{RT}}(\tau_{\text{rev}}, \tau_{\text{fwd}})}{1 - f_{\text{rev}}} \) and \( \tau_{\text{fwd}} = \frac{\tau_{\text{RT}}(\tau_{\text{rev}}, \tau_{\text{fwd}})}{f_{\text{rev}}} \).

To validate our approach, we compared the parameters obtained with our fitting procedure with those obtained from the distribution of time intervals between detected switching events (Figure S7). For both forward and reverse states, the distribution of time intervals between detected switching events (Figure S7B) were well-fit by a biexponential distribution \( P(T_{\text{run}} > t) = C_{\Delta \psi \infty} \exp(-t/\tau_{\text{short}}) + (1 - C_{\Delta \psi \infty}) \exp(-t/\tau_{\text{long}}) \) with the time constants \( \tau_{\text{short}} \) and \( \tau_{\text{long}} \) typically separated by > 10-fold, and the fraction of short intervals \( C_{\Delta \psi \infty} \) varying broadly over its full range, \( 0 \leq C_{\Delta \psi \infty} \leq 1.0 \) (Figure S7D,E). Values for \( \tau_{\text{short}} \) were typically below 1 s (Figure S7D). While some fraction of these short intervals might represent true runs, they could also arise from spurious detection of switches in velocity bearing due to noise in estimating the centroid (see legend of Figure S7D) and in any event, contribute little to the overall dynamics of bearing decorrelation.
Values for $\tau_{\text{fwd}}$ and $\tau_{\text{rev}}$ obtained by fitting equation S1 to the measured autocorrelation functions correlated well with $\tau_{\text{long}}$ (Figure S7E), thus confirming that $\tau_{\text{long}}$ contributes to bearing randomization. We conclude that the forward/reverse switching dynamics are well described by equations (6) and (7), with parameters $\tau_{\text{fwd}}$, and $\tau_{\text{rev}}$.

**Speed Analysis**

Transitions between forward and reverse runs tended to be excluded from the analysis because the speed crosses zero, rendering $\phi$ a noisy variable generating many short runs below our exclusion threshold of 6 frames (see above). The speed set point $\mu_s$ was fit by taking the mean. The remaining parameters of the speed dynamics (3) were fit by its analytical autocorrelation function: $C_s(\tau) = D_s \tau_s \exp(-\tau/\tau_s)$.

**Orientation Analysis**

Changes in orientation during runs (i.e. intervals between reversal events) were analyzed with respect to their mean-squared angular displacements (MSAD) over time, corresponding to a model for angular diffusion with drift

For an object lying on a two-dimensional plane, rotational diffusion about an axis normal to the plane leads to fluctuations in the orientation (an angle measured in the lab frame) $\psi(t)$ over time according to:

$$d\psi(t) = \sqrt{2D_\psi}dW_t,$$

where $D_\psi$ is the rotational diffusion coefficient, and $dW_t$ represents increments of a Wiener process. Bias in these fluctuations over time can be captured, to first order, by adding a linear drift term so that

$$d\psi(t) = k_\psi dt + \sqrt{2D_\psi}dW_t,$$

with $k_\psi$ the drift coefficient.

If $k_\psi$ and $D_\psi$ are constant in time, the mean-squared angular displacement $\text{MSAD}(\tau) = \langle [\psi(t+\tau) - \psi(t)]^2 \rangle$, is a quadratic function of the time delay $\tau$:

$$\langle [\psi(t+\tau) - \psi(t)]^2 \rangle = \langle [k_\psi \tau + \sqrt{2D_\psi}(W_{t+\tau} - W_t)]^2 \rangle = k_\psi^2 \tau^2 + 2D_\psi \tau,$$

where $\langle \cdot \rangle$ denotes averaging over all time pairs separated by $\tau$ and the last equality follows from the Wiener process properties $\langle W_{t+\tau} - W_t \rangle = 0$ and $\langle [W_{t+\tau} - W_t]^2 \rangle = \tau$.

More generally, if $k_\psi(t)$ and $D_\psi(t)$ are time-varying quantities, we can still approximate within a finite time window (centered about time $t_w$) the “local” values $k_{\psi,w} \approx k_\psi(t_w)$ and $D_{\psi,w} \approx D_\psi(t_w)$. In this study, we extract estimates of these (possibly time varying) parameters from fits to the averaged MSAD computed over time windows:
\[ W^{-1} \sum_{w=1}^{w=W} \langle (\psi(t + \tau) - \psi(t))^2 \rangle = \langle k_{\psi}^2 \rangle w \tau^2 + 2\langle D_{\psi} \rangle w \tau, \]  
(S5)

where \( W \) is the number of windows and \( \langle x \rangle_w = W^{-1} \sum_{w=1}^{w=W} x_w \) represents averages over windows. By fitting this averaged MSAD by a quadratic function \( a\tau + b\tau^2 \), we thus obtain the estimates \( a/2 = \langle D_{\psi} \rangle w \) and \( \sqrt{b} = \langle k_{\psi}^2 \rangle w^{1/2} \). Note that \( a/2 \) obtained by this procedure yields an estimate of the mean value for \( D_{\psi} \), but \( \sqrt{b} \) corresponds to an estimate not of the mean value, but the root-mean-square (rms) value for \( k_{\psi} \). Throughout the text, we therefore explicitly refer to the latter estimate as \( k_{\psi \text{rms}} \) (and refer to the former simply as \( D_{\psi} \)).

**Simulations**

Reversals, orientation, and speed dynamics were all simulated independently using the model described. Forward and reverse run durations were chosen according to equations (5) and (6) by drawing exponential random numbers with mean value \( \tau_{\text{fwd}} \) or \( \tau_{\text{rev}} \). During reverse runs, \( \Delta \psi \) was set to \( \pi \). The orientation (4) and speed (3) dynamics were simulated using the Euler-Maruyama method [S14] with a time step that matched the frame rate. To prevent negative speeds, a reflective boundary condition was imposed by taking the absolute value of the speed at each simulation step. The velocity was then calculated from the decomposition in (1) and trapezoidally integrated to give the centroid position \( \vec{x}(t) \).

**Behavioral Mode Analysis**

The model parameters were fit to each trajectory to give a phenotypic matrix \( T \). The phenotypic matrix was centered by subtracting the mean phenotype, \( \hat{T} = T - \langle T \rangle_{\text{indiv.}} \). The correlation matrix was then calculated, \( C_T = \text{corr} \hat{T} \), and decomposed into eigenvalues \( \lambda \) and eigenvectors (behavioral modes) \( b \), \( C_T b = \lambda b \). To reduce any bias coming from a single trajectory, this calculation was bootstrapped 1000 times. The significance of the \( k \)-th top mode is assessed by a comparison with the expected variance explained of the \( k \)-th top mode of randomly chosen directions in the behavioral space. We use the explained variance of the \( k \)-th mode of a newly created set of modes where the first \( k - 1 \) modes are equal to the top behavioral modes and the remaining modes are pointing in randomly chosen orthogonal directions. This process is repeated 1000 times.

The projections of each trajectory on these behavioral modes were calculated by \( P = T b \). The uncertainty in the locus of each individual phenotype along the behavioral mode was computed by projecting the motility parameters after bootstrapping over the 100 second windows and taking the standard deviation.
Statistics

Unless otherwise indicated, errorbars and confidence intervals represent the 2.5% and 97.5% percentiles (spanning the 95% confidence interval) estimated from 1000 bootstrap samples. All probability distributions were empirically estimated using kernel density methods in Python’s Seaborn package with a bandwidth automatically selected using Scott’s rule of thumb [S15]. Tabulated mean values of the effective diffusivity model and the motility model (Table S1-S4) represent geometric rather than arithmetic means was used as the parameters varied log-normally.

Derivation of the Velocity Autocorrelation Function Under the Assumption of Independence

The velocity autocorrelation function can be written in terms of the motility components,

\[ C_v(\tau) = \langle \vec{v}(0) \cdot \vec{v}(\tau) \rangle \]

\[ = \langle s(0) [\cos \psi(0) + \Delta \psi(0)] , \sin \psi(0) + \Delta \psi(0)] \times \langle s(\tau) [\cos \psi(\tau) + \Delta \psi(\tau)], \sin \psi(\tau) + \Delta \psi(\tau) \rangle \]  

(S6)

The expected value of the product of independent random variables is the product of the expected value of each variable, i.e., \( \langle xy \rangle = \langle x \rangle \langle y \rangle \). Therefore we can factor out \( C_s = \langle s(0) s(\tau) \rangle \), leaving the vector product with \( \psi \) and \( \Delta \psi \). The expanded vector product is:

\[ C_v(\tau) = C_s(\tau) \times \langle \cos [\psi(0) + \Delta \psi(0)] \cos \psi(\tau) + \Delta \psi(\tau) \]  

+ \sin \psi(0) + \Delta \psi(0) \sin \psi(\tau) + \Delta \psi(\tau) \rangle \]  

(S7)

The trigonometric functions on \( \psi(t) + \Delta \psi(t) \) can be rewritten as products of trigonometric functions of the terms:

\[ \cos [\psi(t) + \Delta \psi(t)] = \cos \psi(t) \cos \Delta \psi(t) - \sin \psi(t) \sin \Delta \psi(t) \]

\[ \sin [\psi(t) + \Delta \psi(t)] = \sin \psi(t) \cos \Delta \psi(t) + \cos \psi(t) \sin \Delta \psi(t) \]

However, since \( \Delta \psi(t) = \{0, \pi\} \), \( \sin \Delta \psi(t) = 0 \):

\[ \cos [\psi(t) + \Delta \psi(t)] = \cos \psi(t) \cos \Delta \psi(t) \]

\[ \sin [\psi(t) + \Delta \psi(t)] = \sin \psi(t) \cos \Delta \psi(t) \]

Substituting into (S7),

\[ C_v(\tau) = C_s(\tau) \times \langle \cos \psi(0) \cos \psi(\tau) \cos \Delta \psi(0) \cos \Delta \psi(\tau) + \sin \psi(0) \sin \psi(\tau) \cos \Delta \psi(0) \cos \Delta \psi(\tau) \rangle \]  

(S8)

We can now factor out \( C_\psi(\tau) = \langle \cos \psi(\tau) - \psi(0) \rangle = \langle \cos \psi(0) \cos \psi(\tau) + \sin \psi(0) \sin \psi(\tau) \rangle \) to get:

\[ C_v(\tau) = C_s(\tau) C_\psi(\tau) \langle \cos \Delta \psi(0) \cos \Delta \psi(\tau) \rangle \]
Figure S1: Confinement by the boundary affects the mean-squared displacement (MSD) at long times, but does not impair resolution of the ballistic to diffusive transition. We compare the statistical behavior of *C. elegans* N2 in the experiments presented here within small (1-cm) arenas (black) and a previously reported dataset that used larger (5-cm) arenas [S1] (red). The MSD (A), defined as $\langle [\Delta x(\tau)]^2 \rangle \equiv \langle [\vec{x}(t + \tau) - \vec{x}(t)]^2 \rangle$, of our small-arena dataset is similar to that of the large-arena dataset at short times, but does show mild effects of confinement at long times ($\gtrsim 100\text{ s}$). The ballistic to diffusive transition can be more closely studied by examining decay of the velocity autocorrelation function (VACF), defined as $C_v(\tau) \equiv \langle \vec{v}(t) \cdot \vec{v}(t + \tau) \rangle$ (B), which is related to MSD (i.e. $[\Delta x(\tau)]^2$) by $(d/d\tau)\langle [\Delta x(\tau)]^2 \rangle = 2 \int_0^\tau d\tau' C_v(\tau')$ [S16]. The decay of the VACF to zero, which indicates orientation randomization and hence the transition from the ballistic to diffusive regime, is not significantly affected by the presence of the confining boundary.

Finally, we substitute (again dropping $\sin \Delta \psi(t)$ terms):

$$C_{\Delta \psi}(\tau) = \langle \cos [\Delta \psi(0) - \Delta \psi(\tau)] \rangle = \langle \cos \Delta \psi(0) \cos \Delta \psi(\tau) \rangle$$

to get:

$$C_{\vec{v},\text{indep}}(\tau) = C_s(\tau)C_v(\tau)C_{\Delta \psi}(\tau)$$
Figure S2: An overview of the dataset. Trajectories of all worm included in the study. Each box represents a 10 mm by 10 mm chamber. In blue, we highlight points excluded from the analysis because they were influenced by the boundary.

Figure S3: The ballistic to diffusive transition for all strains. We show the average mean-squared displacement (MSD), calculated across individual trajectories, for each strain (black). The expected ballistic (blue) and diffusive MSD curves (red), as in Figure 1B.
Figure S4: The worms’ behavior was approximately stationary. For each strain, we show the average speed (top) and fraction of time spent reversing (bottom) calculated over 100 s sliding windows and averaged across individuals.

Figure S5: Characterization of speed statistics across strains. (A) The speed autocorrelation (black) of each strain decays exponentially (red). (B) The speed distribution (black) of each strain is closely reproduced by model simulations (red).
Figure S6: Characterization of orientation statistics across strains. (A) The mean squared angular displacement of body orientation (black) was fit to a quadratic function (red) in all strains. (B) The orientation correlation (black) decays non-exponentially for many strains.
Figure S7: Characterization of reversal statistics across strains. (A) Distribution of $\Delta \psi$ for each strain shows two prominent peaks at 0° and 180°. (B) Cumulative distributions of the forward and reverse run durations ($T_{\text{fwd}}, T_{\text{rev}}$) for an individual worm from each strain (black), fit to a biexponential function (red). (C) The autocorrelation function of $\Delta \psi$ for each strain (black) along with exponential fit (red). (D, left) The fraction of short runs measured by the biexponential fits of the transition time distributions (as in B) was inversely correlated with the average speed of the worm. At low speed, the bearing (and therefore also $\Delta \psi$, which is used to identify runs) is expected to be dominated by noise (e.g. pixelation artifacts). (D, right) The fitted time constants for short forward and reverse intervals were uncorrelated (unlike those for long runs, see E and also Figure 5A), and typically below the timescale of smoothing filter for velocity data (1 s), further motivating the exclusion of short intervals in modeling reversal dynamics. (E) $\tau_{\text{long}}$, extracted from fits to the transition time distributions, were correlated with $\tau_{\text{fwd}}$ and $\tau_{\text{rev}}$, estimated from $C_{\Delta \psi}$ (panel C).
Figure S8: Worm run speeds are similar during forward and reverse runs for individual trajectories (dots, colored by strain). We define run speed as the top speed during runs (rather than the mean speed, to avoid biases due to run-length differences). The top speed is computed as the 95th percentile of the speed distribution (rather than the maximum, to avoid outlier effects).
Figure S9: Cross-correlation analysis of motility dynamics. The cross-correlation between (top) speed and bearing changes, (middle) speed and velocity alignment, and (bottom) bearing changes and velocity alignment are shown for each strain. There is very little cross-correlation among the motility variables in any of the strains. All cross-correlations were normalized to unit variance by dividing by the product of the standard deviation ($\sigma$) of the two components.
Figure S10: The top behavioral mode effectively captures changes in diffusivity compared to random projections. (A) The effect of variation on the top behavioral mode (black, as in Figure 5C) compared with a sampling of 100 random modes (red) on the diffusivity of simulated trajectories. For random modes, the sign of the mode was chosen such that the diffusivity increased with the projection along the mode. (B) For each random mode we compute the relative change in diffusivity between mode values $\Delta D_{\text{eff}} = D_{\text{eff}}(2)/D_{\text{eff}}(-2)$ and compare to the same relative diffusivity computed from the top behavioral mode. The kernel density distribution of the observed change is shown for the 100 samples (ticks). The black line indicates a ratio of 1 (no difference) and most random projections exhibit less range in $\Delta D_{\text{eff}}$.

Figure S11: An Overview of the image processing steps. The video frames were processed by (1) subtracting the average of 50 frames evenly sampled from the entire movie and (2) cropping to each of the SDS-enclosed regions. (3) The largest worm-sized object was identified following several image morphology operations, and (4) the centroid and image skeleton were measured.
Figure S12: Comparison of velocity calculation methods. Velocity autocorrelation functions for the three example strains with and without filtering of the data and without averaging over 100 s windows. The unfiltered velocity (black), estimated using a symmetrized derivative, contained a δ-correlated short-timescale component in all strains that was particularly prominent in slow-moving strains such as PS312. The velocity calculated using a 1 s cubic polynomial filter (red) does not contain this δ-correlated component.
Table S1: The geometric mean of the effective diffusivity for each strain. For each trajectory, an effective diffusivity ($D_{\text{eff}}$) was extracted by analysis of mean-squared displacements and the velocity autocorrelation function.

| Strain | $D_{\text{eff}} \times 10^2$ ($\mu m^2/s$) |
|--------|-----------------------------------------|
|        | Mean | 2.5% | 97.5% |
| N2     | 140  | 105  | 185   |
| CB4856 | 429  | 307  | 620   |
| JU775  | 448  | 360  | 558   |
| QX1211 | 36   | 12   | 98    |
| JU757  | 210  | 123  | 327   |
| DF5020 | 128  | 55   | 255   |
| PS312  | 8    | 5    | 13    |
| PS1159 | 81   | 32   | 183   |
| sjh2   | 425  | 314  | 553   |

Table S2: The model parameters related to the speed dynamics are listed for each strain. For each worm in a strain, time-averaged parameters were calculated.

| Strains | $\mu_s$ ($\mu m/s$) | $\tau_s$ (s) | $D_s \times 10^2$ ((\mu m/s)^2/s) |
|---------|---------------------|--------------|-----------------------------------|
|         | Mean | 2.5% | 97.5% | Mean | 2.5% | 97.5% | Mean | 2.5% | 97.5% |
| N2      | 77   | 68   | 85    | 1.9 | 1.5 | 2.4 | 5.8 | 4.3 | 7.9 |
| CB4856  | 108  | 91   | 128   | 1.8 | 1.5 | 2.3 | 4.0 | 3.1 | 5.2 |
| JU775   | 112  | 98   | 127   | 2.1 | 1.6 | 2.7 | 4.3 | 3.3 | 5.5 |
| QX1211  | 40   | 26   | 66    | 0.7 | 0.4 | 1.1 | 7.5 | 4.6 | 11.6 |
| JU757   | 97   | 72   | 120   | 4.2 | 3.3 | 5.2 | 3.8 | 2.8 | 5.0 |
| DF5020  | 65   | 50   | 83    | 1.1 | 0.8 | 1.3 | 11.7 | 9.1 | 14.4 |
| PS312   | 27   | 23   | 32    | 0.7 | 0.6 | 0.8 | 5.3 | 3.7 | 7.4 |
| PS1159  | 38   | 26   | 53    | 3.1 | 1.4 | 6.2 | 0.6 | 0.3 | 1.1 |
| sjh2    | 159  | 138  | 184   | 3.3 | 2.5 | 4.7 | 8.8 | 5.8 | 13.2 |
Table S3: The model parameters related to the orientation dynamics are listed for each strain. For each worm in a strain, time-averaged parameters were calculated.

| Strains | \( k_{\text{rms}} \) (rad/s) | \( D_\psi \) (rad²/s) |
|---------|-------------------------------|------------------------|
|         | Mean 2.5% 97.5%               | Mean 2.5% 97.5%        |
| N2      | 0.036 0.026 0.048            | 0.034 0.017 0.054      |
| CB4856  | 0.029 0.018 0.041            | 0.024 0.018 0.033      |
| JU775   | 0.038 0.026 0.053            | 0.021 0.016 0.030      |
| QX1211  | 0.040 0.028 0.056            | 0.017 0.009 0.036      |
| JU757   | 0.039 0.030 0.052            | 0.036 0.021 0.054      |
| DF5020  | 0.037 0.032 0.042            | 0.033 0.026 0.041      |
| PS312   | 0.017 0.011 0.029            | 0.024 0.018 0.033      |
| PS1159  | 0.023 0.015 0.031            | 0.009 0.005 0.017      |
| sjh2    | 0.066 0.057 0.077            | 0.090 0.065 0.127      |

Table S4: The model parameters related to the reversal state dynamics are listed for each strain. For each worm in a strain, time-averaged parameters were calculated.

| Strains | \( \tau_{\text{fwd}} \) (s) | \( \tau_{\text{rev}} \) (s) |
|---------|-----------------------------|-----------------------------|
|         | Mean 2.5% 97.5%             | Mean 2.5% 97.5%             |
| N2      | 23.8 13.9 41.1              | 4.1 3.0 5.7                |
| CB4856  | 78.6 56.5 109.5             | 8.1 4.2 16.3               |
| JU775   | 85.3 51.5 144.0             | 8.0 4.2 16.3               |
| QX1211  | 26.5 12.5 63.0              | 3.7 2.7 5.2                |
| JU757   | 32.3 20.5 50.5              | 6.4 4.1 10.3               |
| DF5020  | 32.7 15.2 63.2              | 6.4 4.1 10.3               |
| PS312   | 5.6 4.2 7.4                | 3.3 2.9 3.9                |
| PS1159  | 80.8 38.0 174.6             | 8.8 5.2 16.2               |
| sjh2    | 155.5 75.2 419.9           | 4.9 2.3 10.8               |

Table S5: The loadings of each parameter on the top behavioral mode are listed.

| Parameter | Loading | Mean 2.5% 97.5% |
|-----------|---------|-----------------|
| \( \log_{10} \mu_s \) | 0.50    | 0.25 0.58      |
| \( \log_{10} \tau_s \) | 0.51    | 0.40 0.54      |
| \( \log_{10} D_s \) | -0.19  -0.43 0.08 |
| \( \log_{10} k_\psi \) | 0.24    | -0.04 0.43     |
| \( \log_{10} D_\psi \) | 0.15    | -0.22 0.40     |
| \( \log_{10} \tau_{\text{fwd}} \) | 0.50    | 0.36 0.55      |
| \( \log_{10} \tau_{\text{rev}} \) | 0.35    | 0.15 0.48      |
Table S6: The phenotypic projection along the first behavioral mode is listed for each strain. For each worm in a strain, a time-averaged projection was calculated.

| Strain | Projection |
|--------|------------|
|        | Mean  2.5% 97.5% |
| N2     | -0.26 -0.48 0.49 |
| CB4856 | 0.42 -0.35 0.76 |
| JU775  | 0.86 -0.06 1.04 |
| QX1211 | -1.45 -1.82 0.32 |
| JU757  | 0.78 -0.07 1.23 |
| DF5020 | -0.65 -0.70 1.19 |
| PS312  | -2.42 -2.66 -0.13 |
| PS1159 | 0.45 -2.64 1.01 |
| sjh2   | 1.61 0.80 2.41 |
SI references

[S1] Stephens, G. J, Johnson-Kerner, B, Bialek, W, & Ryu, W. S. (2010) From modes to movement in the behavior of Caenorhabditis elegans. *PLoS One* **5**, e13914.

[S2] Blaxter, M. L, De Ley, P, Garey, J. R, Liu, L. X, Scheldeman, P, Vierstraete, a, Vanfleteren, J. R, Mackey, L. Y, Dorris, M, Frisse, L. M, Vida, J. T, & Thomas, W. K. (1998) A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**, 71–5.

[S3] Holovachov, O, van Megen, H, Bongers, T, Bakker, J, Helder, J, van den Elsen, S, Holterman, M, Karssen, G, & Mooymen, P. (2009) A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* **11**, 927–950.

[S4] Kiontke, K & Fitch, D. (2005) The phylogenetic relationships of Caenorhabditis and other rhabditids. *WormBook*.

[S5] Andersen, E. C, Gerke, J. P, Shapiro, J. A, Crissman, J. R, Ghosh, R, Bloom, J. S, Félix, M.-A, & Kruglyak, L. (2012) Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. *Nat. Genet.* **44**, 285–90.

[S6] Chiang, J.-T. A, Steciuk, M, Shtonda, B, & Avery, L. (2006) Evolution of pharyngeal behaviors and neuronal functions in free-living soil nematodes. *J. Exp. Biol.* **209**, 1859–73.

[S7] Davis, M. W, Somerville, D, Lee, R. Y, Lockery, S, Avery, L, & Fambrrou, D. M. (1995) Mutations in the Caenorhabditis elegans Na,K-ATPase alpha-subunit gene, eat-6, disrupt excitable cell function. *J. Neurosci.* **15**, 8408–18.

[S8] Stiernagle, T. (2006) Maintenance of C. elegans. *WormBook*.

[S9] Lahl, V, Halama, C, & Schierenberg, E. (2003) Comparative and experimental embryogenesis of Plectidae (Nematoda). *Dev. Genes Evol.* **213**, 18–27.

[S10] Hall, S. E, Beverly, M, Russ, C, Nusbaum, C, & Sengupta, P. (2010) A cellular memory of developmental history generates phenotypic diversity in C. elegans. *Curr. Biol.* **20**, 149–55.

[S11] Rechavi, O, Houri-Ze’evi, L, Anava, S, Goh, W. S. S, Kerk, S. Y, Hannon, G. J, & Hobert, O. (2014) Starvation-Induced Transgenerational Inheritance of Small RNAs in C. elegans. *Cell* **158**, 277–87.

[S12] Hart, A. (2006) Behavior. *WormBook*.
[S13] Huang, K.-M, Cosman, P, & Schafer, W. R. (2006) Machine vision based
detection of omega bends and reversals in C. elegans. J. Neurosci. Methods
158, 323–36.

[S14] Kloeden, P. E & Platen, E. (1992) Numerical Solution of Stochastic Dif-
ferential Equations. (Springer Berlin Heidelberg, Berlin, Heidelberg).

[S15] Scott, D. W. (1992) Multivariate density estimation: theory, practice, and
visualization. (John Wiley & Sons).

[S16] Chaikin, P. M., Lubensky, T. C & Witten, T. A. (1995) Principles of
condensed matter physics.. Cambridge University Press).