

METHODS & TECHNIQUES

High-resolution, high-throughput analysis of *Drosophila* geotactic behavior

Tijana Canic^{1,2,*}, Juan Lopez¹, Natalie Ortiz-Vega^{2,3}, R. Grace Zhai^{2,3} and Sheyum Syed^{1,†}

ABSTRACT

Drosophila's innate response to gravity, geotaxis, has been used to assess the impact of aging and disease on motor performance. Despite its rich history, fly geotaxis continues to be largely measured manually and assessed through simplistic metrics, limiting analytic insights into the behavior. Here, we have constructed a fully programmable apparatus and developed a multi-object tracking software capable of following sub-second movements of individual flies, thus allowing quantitative analysis of geotaxis. The apparatus monitors 10 fly cohorts simultaneously, with each cohort consisting of up to 7 flies. The software tracks single flies during the entire run with ~97% accuracy, yielding detailed climbing curve, speed and movement direction with 1/30 s resolution. Our tracking permits the construction of multi-variable metrics and the detection of transitory movement phenotypes, such as slips and falls. The platform is therefore poised to advance *Drosophila* geotaxis assay into a comprehensive assessment of locomotor behavior.

KEY WORDS: Geotaxis, *Drosophila*, Multi-object tracking, Slips and falls, Kalman filter

INTRODUCTION

Earth's gravity regulates the shape and size of organisms during development and continues to affect their function throughout life (Narayanan, 2023; Takahashi et al., 2021). Animals use gravity primarily for navigation and orientation, allowing them to locate food and suitable habitat. When startled, the fruit fly *Drosophila melanogaster* displays a natural tendency to move against gravity, resembling an escape response. Flies sense and orient themselves against gravity primarily through Johnston's organ in the antenna along with other sensory structures (Armstrong et al., 2006; Bender and Frye, 2009; Kamikouchi et al., 2009; Kladt and Reiser, 2023 preprint). This innate movement against gravity is called negative geotaxis and is elicited in the laboratory typically by tapping flies to the bottom of a vial and monitoring their subsequent vertical climb. The climbing assay dates back over 60 years and has been coupled with sophisticated genetic tools available in *Drosophila* to elucidate

aging and general decline of motor output in disease states (Hirsch and Erlenmeyer-Kimling, 1961; Zhu et al., 2022).

Despite their widespread utility, current platforms for geotactic experiments have several limitations that can be broadly divided into those related to hardware for initiating movements and those related to software for tracking. Traditionally, the climbing behavior is visually monitored or video recorded and the number of flies reaching a certain height at the end of ~10 s is the measure of their negative geotaxis. Though this approach provides a simple and rapid measure of geotactic response sufficient for detecting gross differences between groups, manual tapping of vials introduces inconsistencies into experiments. To reduce variability, specially designed rigs with automated and uniform force delivery have been demonstrated (Cao et al., 2017; Gargano et al., 2005; Podratz et al., 2013; Spierer et al., 2021). Data from these experiments are either scored manually or quantified using computer programs. Although more reliable than manual scoring, computer programs employed to track geotactic activity generally report population or experiment-wide average metrics such as speed and final percentage of successful climbers (Kohlhoff et al., 2011; Podratz et al., 2013; Spierer et al., 2021). While such analysis can be valuable, it overlooks individual behavioral attributes that can offer critical insights into understanding geotaxis.

Here, we present a new platform for measuring and interpreting *Drosophila* geotaxis at a single-fly, sub-second resolution, addressing most issues with currently available methods. The platform combines the strengths of a fully automated apparatus with a powerful computer vision tracking algorithm, enabling generation of data that are reproducible within the natural variations in fly behavior, while tracking provides detailed locomotor trajectories of single flies with ~97% accuracy and a 1/30 s temporal resolution. We leverage the high-resolution tracking data to go beyond population-averaged quantities and construct multivariable analytics gleaned from individual flies and video frames to help explain geotactic behavioral differences seen in several wildtype laboratory strains.

MATERIALS AND METHODS

Drosophila husbandry

We utilized three laboratory strains in these studies: *white* (*w¹¹¹⁸*), Canton-S (CS) and *yellow, white* (*yw*). Virgin females were separated and maintained at 25°C on standard lab food on a 12 h:12 h light:dark (LD) cycle for 5 or 6 days. The flies were then briefly anesthetized with CO₂ and loaded into experimental vials. Experiments were conducted after flies acclimated for an hour at 25°C and 75% relative humidity.

Hardware and video recording

We used a custom-designed acrylic vial mount that holds 10 plastic vials (Uline, Pleasant Prairie, WI, USA; model S-21972) and has an opaque plastic backing connected to a plastic rod. A metal lever, connected to a programmable stepper motor (Trinamic, model

¹Department of Physics, University of Miami, Coral Gables, FL 33146, USA.

²Department of Molecular and Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, FL 33136, USA. ³Department of Neurology, University of Chicago, Chicago, IL 60637, USA.

*Present address: Graduate Program in Computational Neuroscience, University of Chicago, Chicago, IL 60637, USA.

†Author for correspondence (s.syed@miami.edu)

 R.G.Z., 0000-0002-7599-1430; S.S., 0000-0002-4642-6678

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

PD86-3-12978), lifts the mounted vials ~6.5 cm by pushing up on the acrylic rod during half a rotation then lets the vials fall for the second half. A mechanical sensor (Omron Electronics Inc.-EMC Div, model SS-10GL2T) at the bottom of the lever rotation sets the start and stop positions of the motor. By default, the vials are vertically moved 4 times in 3 s, but this setting can be changed by reprogramming the motor. Each vial is normally loaded with up to 7 flies, which we found allows for high throughput as well as accurate tracking. We have tested up to 10 flies per vial with acceptable accuracy but found that in general having more than 7 flies decreases individual tracking accuracy.

White light-emitting diodes (LEDs, JOYLIT model A2835-240-60K-V24-NWP) are mounted to a metal plate 4.5 cm behind the mounted vials. Two sheets of diffuser (Rosco model R3000) are placed over the LEDs. A digital camera (ImagingSource LLC, model DMK23U445) is placed on a separate table 38 cm from the center of the opaque plastic backing. The camera is adjusted using the focus knobs until the edges of the mounted vials are at the edge of the frame. Both the camera and the motor are connected to the computer via MATLAB packages (Computer Vision, Image Acquisition, Image Processing, Instrument Control and Statistics and Machine Learning Toolboxes; third-party Instrument Control Toolbox Support Package for Keysight IO Libraries and VISA Interface).

The user initiates an experiment from the MATLAB command line by specifying the number of successive videos/trials (default, 10), duration of each video (13 s), rate of video capture (30 Hz), wait time between videos (15 s) and video names ('exp1_'). The default values used in our experiments are included in parentheses. Once a new experiment is started, the computer engages the motor, waits 3 s for the fly vials to come to a complete stop, then begins recording the first video. Once complete, the program saves the video as an '.mp4' with the appended video number, then waits the specified time to start the next trial.

Video processing

Processing of the video recordings starts with defining the rectangular regions of interest (ROI) around the vials for each individual video. The ROI is saved in the working folder and can be utilized for recursive video processing. The saved videos are subsequently pushed through a sequence of steps that utilize the calibration data to localize, label and track each fly over time (Movie 3). Finally, the tracked fly coordinate data are used for assessing geotaxis through several analytic outputs. This pipeline of simultaneously measuring multiple groups and interpreting each group video separately offers flexibility for meta and comparative studies across trials, number of flies, treatments and genotypes.

Calibration

The calibration step is recommended whenever a fly of significantly different size or morphology is first used, a new camera is introduced or the platform-camera distance is altered. The calibration program (Fig. S2B) presents a video frame in which boxes have been drawn around individual 'objects' and asks user to classify each object as 'Is one fly', 'Multiple flies' (two or more flies together) or 'Not one fly' (noise). The program saves two pieces of information for every classified object: pixel area and major/minor axis lengths (Fig. S2B, enlarged panel). Axis lengths are defined as the shortest and longest lines, respectively, that can be drawn from one side of an object to another. The area characterizes the size of the objects that the tracking function will look for while the axis lengths together characterize the shape of the objects. For every

video recording loaded, the program automatically presents to the user every 15th frame for assessment. The calibration table helps discriminate against objects with shape or size outside the bounds of those expected for fly strains in use. Because it informs the program about the general physical characteristics of the objects of interest, calibration also adds to the versatility of our algorithm. Simply by creating a new set of calibration data, a user can examine insect strains of drastically different characteristics or under different apparatus set ups.

Manual accuracy assessment

The tracked ROI videos and the associated coordinate matrices are used to check whether the same fly is associated with the correct track number over time. For every 15th frame, each fly is manually assessed as correct if the fly is within the appropriate bounding box. If correct, the button 'correct' is clicked and will add the instance to a tally. If incorrect, the button 'incorrect' is clicked and the program waits for the mouse click with the accurate fly location from which the Euclidean distance to the saved track is calculated and reported in a spreadsheet file.

Calculation of metrics

Climbing curves

Climbing curves are constructed for a population by noting the time of the first instance each fly crosses a given height (e.g. 10 cm above vial bottom). The cumulative sum of these data generates step-like patterns with a step size of 1 when counting individuals in a single trial (e.g. Fig. 3C, left) or a step size of $1/N_{\text{trials}}$ when counting average number of individuals from N_{trials} (e.g. Fig. 3C, right). A fly that has crossed above a target height (maximum target=14 cm from vial bottom) is counted as having climbed to the height even if the fly displays positive geotaxis later on and goes below the target.

Speed and angle

We calculate the Euclidean distance moved by a fly between frames i and $i+1$ according to:

$$\Delta r_i = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}, \quad (1)$$

where (x_i, y_i) and (x_{i+1}, y_{i+1}) are the position coordinates in the two corresponding frames. The related speed s_i is defined as $\Delta r_i/\Delta t$, where, throughout this work $\Delta t=1/30$ s. $\Delta r_i \geq 4$ mm are excluded, as they are considered abnormally large frame-frame locomotion. This cut-off in frame-frame displacement values prevents artifactual jumps in position, slips, falls and occasional flight within the vial from being considered in the speed-angle analysis. It is because of this constraint that our fly speed data have an upper limit of ~ 100 mm s^{-1} . The movement direction between frames i and $i+1$ is calculated according to:

$$\theta_i = \tan^{-1} \left[\frac{y_{i+1} - y_i}{|(x_{i+1} - x_i)|} \right]. \quad (2)$$

Taking the absolute value of Δx makes a rightward movement indistinguishable from a leftward movement, keeping the focus on whether a movement is in the $+y$ or $-y$ direction. θ_i is undefined if a fly is stationary between two video frames; that is, if $\Delta r_i=0$ between frames i and $i+1$. By counting the number of non-zero angular changes $\Delta\theta \neq 0$, the θ_i can also yield a trajectory turning rate for a cohort. For example, we find that a w^{1118} fly makes ~ 24 turns s^{-1} , a CS fly makes ~ 22 turns s^{-1} and a yw fly makes ~ 20 turns s^{-1} . It is important to note, however, that the turning rate calculated this way

relates to changes in the track direction and not necessarily the heading direction. Finally, the average speed and angle are calculated from the frame–frame s_i and θ_i , respectively. The average metrics also readily lend themselves to other quantities as well. For instance, the average speed multiplied by the duration of video gives the distance traveled while the average angle provides a rough snapshot of a cohort's net travel direction.

Bootstrapping speed–angle correlations

To estimate the mean and standard error of correlations between instantaneous speed and movement angle, we followed the basic bootstrapping procedure of randomly sampling with replacement speed–angle pairs. Each sample size, N_{sample} , was set equal to the size of the data. For instance, $N_{\text{sample}}=41,330$ for CS flies. Speed–angle pairs within each random sample set were separated in terms of $\theta < 0$ and $\theta > 0$ and the corresponding Spearman correlation coefficients ρ_- and ρ_+ and their ratio ρ_+/ρ_- were calculated. This procedure was repeated 1000 times, yielding a bootstrapped probability distribution of ρ_+/ρ_- , as shown in Fig. 4C (right). The distribution mean and standard deviation are the ratio mean and error.

Extraction of analytics

Detailed guidelines on analysis of tracked videos are provided with the code library at <https://github.com/sheyums/Insect-Geotaxis>.

Some basic familiarity with the Matlab command line is required. Briefly, the function ‘Geotaxis_analytics’ uses fly location results from tracking to calculate the metrics discussed above. The output of Geotaxis_analytics is a structure with one field – Group – itself a structure containing user-provided names and all associated data of input datasets. For example, from the analysis of four groups of data, the user can access the name and the climbing data of the 3rd group by typing ‘output.Group(3).Names’ and ‘output.Groups(3).ClimbCurves’, respectively, at the command line. Entering ‘output.Group(1).Fly_no(2).AveSpeed’ will return the average speed of the 2nd fly in the 1st group. Other results, for example the paired frame–frame speed and angular movement data of fly#19 in the 4th group, can be accessed by typing ‘output.Group(4).Fly_no(19).SpdAngle’. For additional quantification, such as determination of the number of turns made by fly#19 in group#4, the indexing feature in Matlab can again be utilized to calculate the number of non-zero angle changes between successive frames.

RESULTS

Mechanical design

To reduce commonly found systematic errors in manual assessment of fly geotaxis, we built an apparatus (Fig. 1A) that delivers a fixed number of taps of identical force over the same duration across experiments. The core of the apparatus consists of a

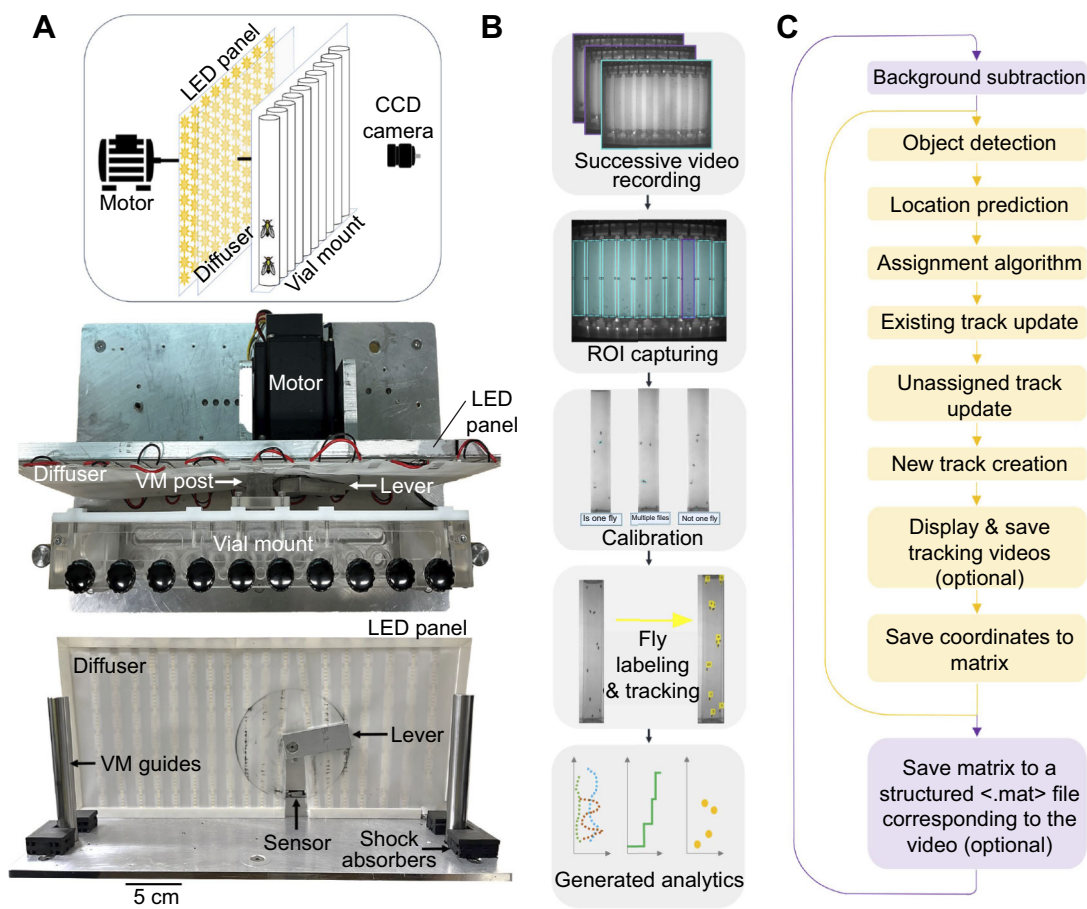


Fig. 1. Overview of mechanics and software for the automated system. (A) Schematic diagram (top), overhead view (middle) and frontal view (bottom) of the mechanical device with essential components labeled on the images. VM, vial mount. (B) Flow diagram of the main steps for this assay. (C) Flow diagram detailing internal steps of the motion-based, multi-object tracking algorithm. Steps in purple are repeated for each vial and steps in yellow are repeated for each frame within a vial video.

custom-designed acrylic vial mount (VM) that holds up to 10 plastic vials with slots on the side that allow it to slide vertically along metal guides (VM guides). When the motor is activated, a metal lever connected to it contacts the VM post, lifting the VM for the first half of the lever's rotation and allowing the VM to fall during the second half. A mechanical sensor (Fig. 1A, bottom panel) detects lever motion and provides an appropriate signal to initiate or terminate motor control. Normally, the VM is vertically moved 4 times in 3 s, ensuring all flies are at the bottom of their respective vials before video recording starts. A panel of LEDs provides diffuse illumination from behind the VM.

Video recording and image processing

A CCD camera acquires grayscale images at a rate of 30 Hz once the vial mount comes to rest. A 10 s recording (Movie 1) constitutes one trial. The process of mechanical fly perturbation and subsequent movement recording is usually repeated 10 times per experiment (Fig. 1B). Each video is cropped by ROI and saved as separate videos (Fig. 1B; Movie 2). A subset of ROI videos can be used for calibration, where the user manually classifies detections into categories of single and multiple flies in different orientations and distances from the camera. The calibration step is critical as it yields shape and size distributions of flies as imaged by the CCD and helps the tracking algorithm (see Materials and Methods) reject any non-fly objects captured.

Tracking summary

Tracking individual flies (Movie 3) over time is the most challenging task for our method and we delve into its details over the next three figures. In brief, the first step in our tracking is the creation of a background frame (Fig. 1C). Objects are found by thresholding background-subtracted individual frames. Among detected objects, only those falling within shape and size distributions from calibration data are accepted as flies. A series of specialized algorithms next predicts the flies' locations and assigns each a track. These steps of prediction and updated assignment are repeated for each CCD frame and the position information is saved in a matrix.

Tracking errors

The relatively unconstrained movement of flies produces four major types of tracking errors (Fig. S1) which the algorithm must rectify to achieve high accuracy. (1) Occlusion by another fly or temporary immobility can cause a fly to become invisible to tracking only to reappear frames later (Fig. S1A). This 'lost and found error' is addressed by updating the unassigned track with the last known coordinates and utilizing the assignment algorithm to reconnect the correct fly with its existing track once it reappears. (2) Two or more flies can appear as a single detected object by virtue of proximity. In this type of 'grouping error' (Fig. S1B), the enlarged size of the detected object combined with the multiple fly characteristics from calibration would be used to create overlapping tracks from a single detection. (3) Overlap of multiple flies can also lead to switching of individual assignments (Movie 3). To estimate the frequency of this error, we analyzed 79,800 video frames each with 7 flies and detected 2762 instances of fly collisions (3.5% or ~7 overlaps per 200 frames). But as swapping of labels in a collision is a random event, we assume about half of collisions lead to identity switching. In other words, we incur an identity switch in <2% of video frames (<6 switches per 300 frame video). Importantly, identity switch does not cause loss of single-fly resolution. Because our analyses rely on anonymized single-fly data, the less than 0.6% error (<12

incorrect values within a 7×300 location matrix) contributed by these events is negligible in the inherently heterogeneous behavioral metrics we compute. (4) Another source of assignment-related error is aberrant movement – stationarity and unexpected movement direction. For instance, two flies that are separate from each other can erroneously exchange their assignments (Fig. S1C), or preference is incorrectly given to an aberration in illumination instead of a fly, if the aberrant object is more mobile than the fly (Fig. S1D). However, the large unphysical jumps in positions can be used to correct both types of error.

Background subtraction and detection calibration

Detection of flies in each image frame can be broadly divided into three steps (Fig. 2A): (a) background subtraction, (b) applying size and shape discrimination to identify true flies among detected objects, and (c) determining whether one or more flies comprise each detection. A background frame is generated for each vial by calculating the median pixel value from all frames. Each frame is subtracted from the background and thresholded by setting pixel values ≤ 0.06 to 0 and the rest to 1 (Fig. S2A). Because several flies are moving in three dimensions through an imperfectly illuminated space, the detections after background subtraction can include single and multiple flies and noise objects such as shadows and streaks. To correctly account for all flies, the user can opt to go through calibration (see Materials and Methods).

Fig. 2B,C shows calibration data for a set of several commonly used wildtype strains (w^{1118} , CS, γw) as examples. Objects with areas larger than the maximum of the multi-fly areas and smaller than the lowest 16% of single-fly areas are classified as noise in our algorithm (Fig. 2B, shaded regions). There is some overlap between the single- and multi-fly area distributions and to remove ambiguity we define the maximum single-fly area to be the threshold for multiple flies (Fig. 2B, dashed vertical line). Therefore, objects are accepted as multiple fly detections if their areas are larger than this threshold and smaller than the maximum multi-fly area. We arrived at these cut-offs through trial and error aimed at maximizing fly identification accuracy. However, the extreme cut-offs alone do not eliminate all false positives as the left tail end of the single-fly distribution can still have some overlap with noise. Shape information, specifically the ratio of minimum and maximum axis lengths, is used to eliminate remaining incorrect detections (Fig. 2C). Objects with axis-length ratios not corresponding to single or multi-fly objects (Fig. 2C, shaded regions) are marked as noise detections and removed from further analysis.

The detections that remain after size- and shape-based discrimination are highly likely to be only flies. To account for the total number of flies, we determine the number of individual flies in the multi-fly classifications by dividing each multi-fly area by the maximum single-fly area found in calibration then rounding to the bigger whole number. Individuals found in a multi-fly object are given a common set of coordinates and appended to the list of detections. Though these steps do not guarantee every fly in the frame is detected, they do ensure that every detection is an individual fly.

As our goal is not limited to detecting flies but also involves following them over time, the algorithm next associates the detected fly locations in one frame with fly locations from the previous frame. The association is done by assigning each detection to the most likely track, through a scheme called the Hungarian matching algorithm (Munkres, 1957). In our method, we complete the detection-to-track matching in three steps (Fig. 2D). The first task is constructing the cost matrix (Fig. S3) whose columns represent

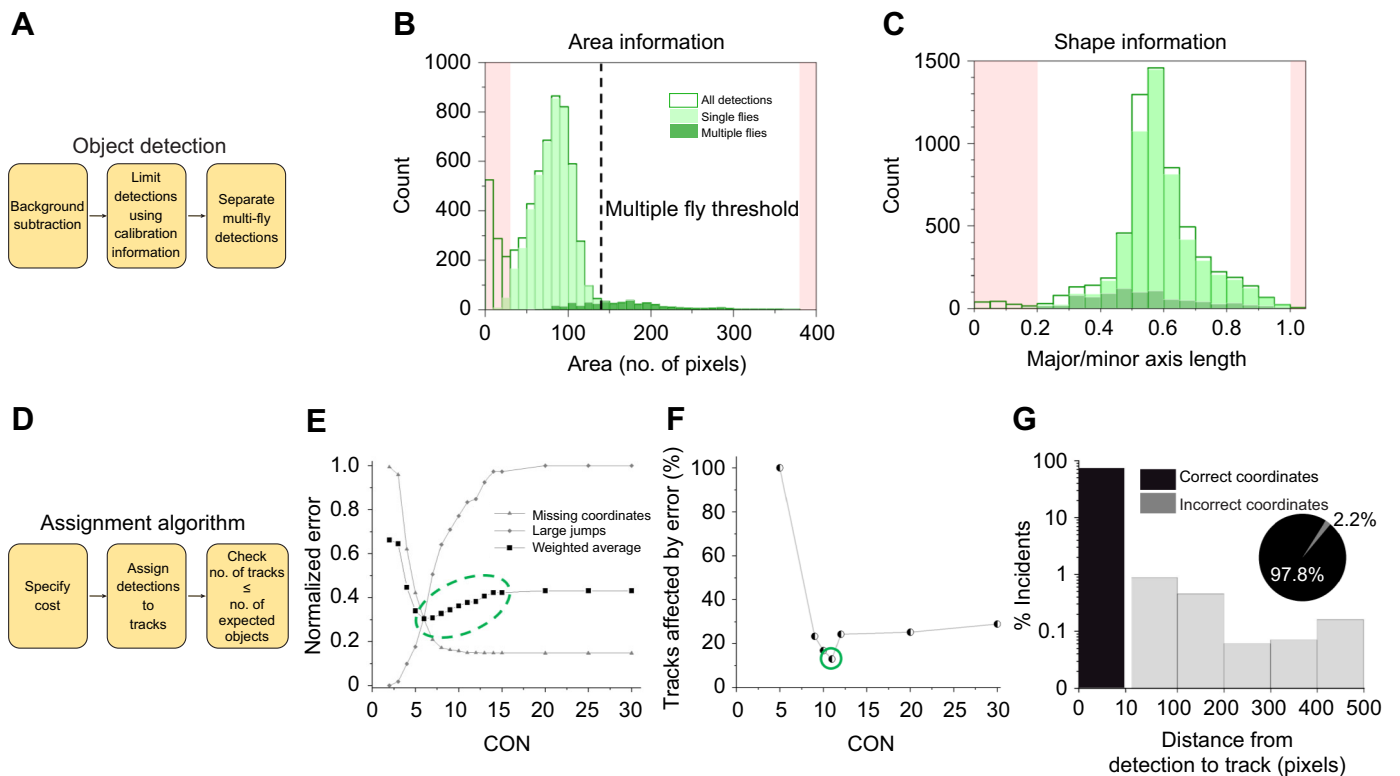


Fig. 2. Using calibration information and cost calculation for optimizing assignment. (A) Flowchart of the main detection steps. (B,C) Histograms of pixel areas (B) and the ratio of minor to major axis lengths (C) of all detected objects. Shaded regions at histogram extrema correspond to noise detection. The legend in B applies to B and C. (D) The main steps in detection-to-track assignment. (E) Computational assessment over 17 cost of non-assignment (CON) values. The average number of untracked frames is normalized to the maximum average found (triangles). The top 1% of the largest frame-to-frame displacements of each track are averaged and normalized to vial length (diamonds). The weighted average (squares) uses coefficients of 1 and 0.1, respectively. The green oval highlights a CON range yielding greater accuracy. (F) Manual assessment of tracks affected by one or more errors over a narrower CON range identifies CON=11 as the most optimal (green circle). (G) Histogram of fly location error (in pixels). Inset: percentage of flies localized to <10 pixels. The same 50 videos (256 flies) were used in B,C and F,G.

locations in the current frame (Detections) and whose rows represent locations from the previous frame (Tracks). For each element of the cost matrix, the Euclidean distance between a current detection and its predicted location is calculated, d_{pred} . For predictions, we use the constant velocity Kalman filter. The filter first makes an initial guess about the future location of a fly using default inputs and updates the guess after comparing it with the final assigned location. However, we found that a cost matrix relying only on Kalman filter predictions is unreliable for our experiments because of idiosyncrasies in fly movement. We therefore define matrix elements as:

$$\sqrt{d_{\text{pred}} \sqrt{d_{\text{prev}}}}$$

where d_{prev} is the Euclidean distance between current and previous detection coordinates. As detection relies on motion, stationary individuals could be lost, or it may take several frames to recognize a fly. The solution for either is the creation of ‘phantom’ tracks and detections within the cost matrix to allow for lost or new flies (Fig. S3, tracks 1 and 3).

In creating phantom tracks, the cost matrix is padded with an experimentally determined value called the cost of non-assignment (CON). The CON is a threshold value that determines whether assigning a detection to a track is too costly. If the CON is too low, it will not assign detections to tracks, resulting in missing coordinates or ‘lost and found errors’ and if the CON is too high, it will assign

detections to the wrong tracks, resulting in swap and jump errors. To visualize the effect of the CON, we recorded the number of missing coordinates and the number of abnormally large jumps in coordinates recorded, as we ran the tracking algorithm on a set of 50 videos at 17 different CON settings (Fig. 2E). As expected, increasing CON minimizes the number of missing coordinates but increases bias towards detection of large location jumps. The optimal CON produces a minimum in their average. As large jumps included in this estimation do not distinguish between actual falls and erroneous jumps in location, we calculated a weighted average by assigning the missing coordinates error twice the weight of the large jumps error. The weighted error gave us a narrow range of ~6–13 to optimize the CON (Fig. 2E). Using a set of 50 videos tracked at 7 different CON, we manually assessed errors within each individual track (Fig. 2F). An entire track was considered incorrect if any significant error was found, regardless of how accurate the rest of the track was. For ease of comparison across groups, we normalized the error to the number of total tracks assessed. These studies found a clear minimum for CON=11, which we used as the default in our tracking algorithm.

If the number of tracks in a frame is found to exceed the number of flies after completion of the detection-track assignments, we recheck the distance between the detected coordinates and the previous coordinates of the unassigned tracks and remove the detection with the maximum distance. Though this accounting method is somewhat arbitrary, extra tracks are rare and do not lead to

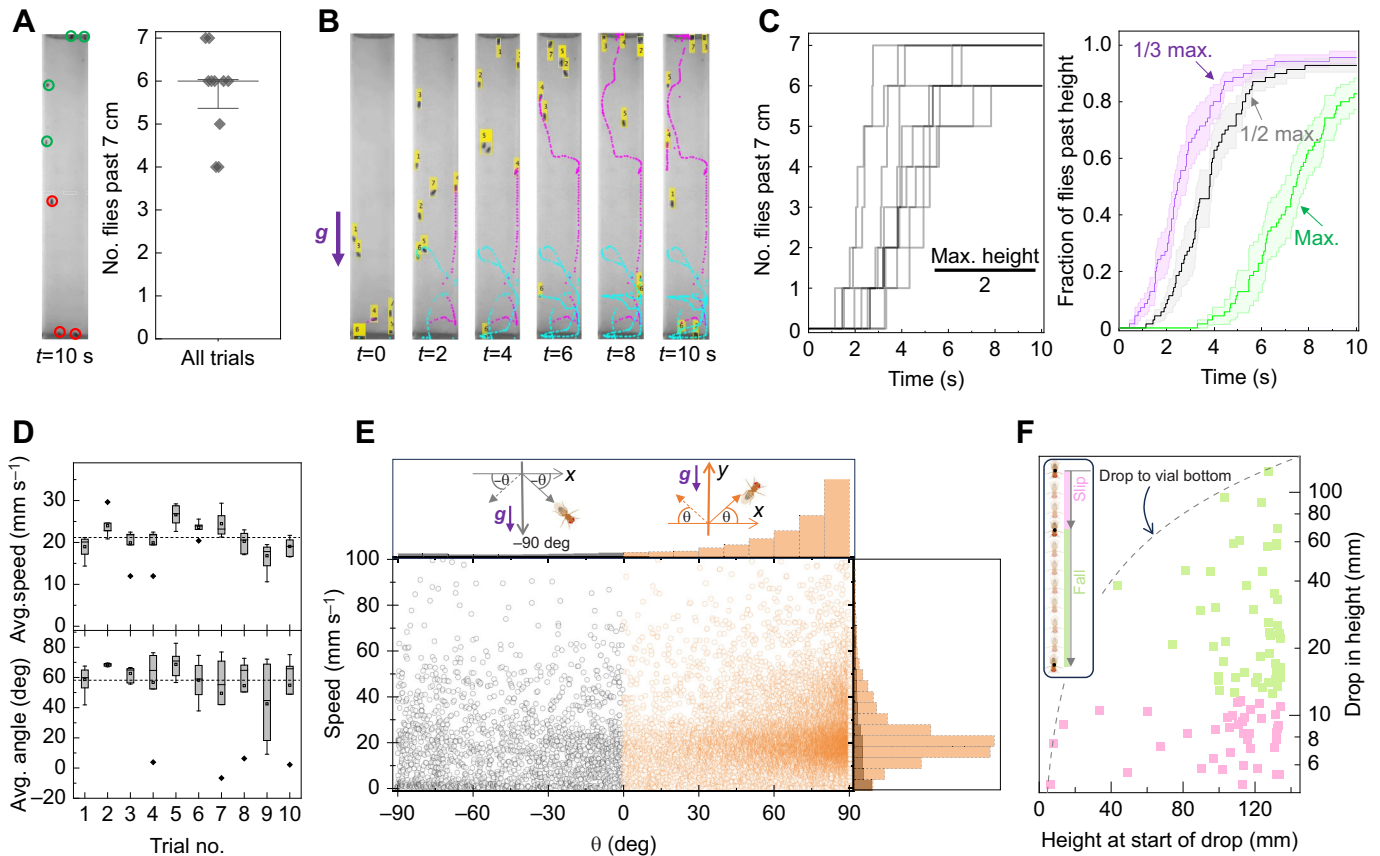


Fig. 3. High-resolution quantification of w^{1118} geotactic behavior. (A) Left: vial image showing fly ($N_{\text{flies}}=7$) locations at $t=10$ s. Flies past the 7 cm height mark are circled in green; remaining flies are circled in red. Right: final fly count (median \pm s.e.) above 7 cm from each trial ($N_{\text{trials}}=10$). (B) Time-lapse images showing tracked flies every 2 s in one trial. Yellow flags represent numbers assigned to each fly at that instant by the tracking program. Example tracks 4 and 6 over time are shown in magenta and cyan, respectively. (C) Left: individual w^{1118} climbing past half the vial height. Each gray staircase represents a single trial. Right: average climbing curves for 1/3, 1/2 and maximum height. Mean (solid line) and s.e.m. (shaded band) for each height are from 10 trials. (D) Per-trial average speed (top) and average angle of movement (bottom). Box plots show mean (open squares), median, upper and lower quartiles, 1.5 \times interquartile range and outliers (filled diamonds). Dashed horizontal line denotes average across trials. (E) Scatter plot of speed and angle of movement (θ) between consecutive video frames, from all flies and trials (open circles, $N_{\text{total}}=15,422$), and relative frequency histograms of speed (right) and angle (top). Data for $\theta>0$ are in orange and those for $\theta<0$ are in gray. Upper quadrant angles are defined as $\theta>0$ and lower quadrant angles are $\theta<0$ (top, cartoons). (F) Height of a fly just prior to a drop versus the size of the vertical drop (squares). Dashed line: maximum possible drop from a given height. Note: the y-axis is on a log scale. Inset: a 'slip' or 'fall' in terms of body length. In B,E, ' $g\downarrow$ ' indicates the direction of gravitational force.

major information loss. If the number of tracks is less than the known number of flies, nothing is done as some flies remain stationary. At the end of the tracking, a random coordinate at the bottom of the vial is selected and appended to the final coordinate matrix to account for the missing fly.

To gauge the overall accuracy of our algorithm, two individuals separately evaluated tracking results of each fly in a set of 50 videos. If a tracking result was deemed inaccurate, the individual clicked on the actual fly coordinate and the distance between the actual location and the tracked location was recorded. These assessments showed that our algorithm correctly tracks flies $\sim 97\%$ of the time (Fig. 2G, inset) with a positional error of <9.7 pixels ≈ 2 mm (Fig. 2G).

Introduction of new metrics

Traditionally, geotaxis experiments report the percentage of flies that successfully pass a threshold height after a specified time. In an experiment with w^{1118} flies, simple counting at $t=10$ s shows that in 70% of trials, six or more flies are past the 7 cm mark, half of the full vial height (Fig. 3A). Our detailed tracking provides a more comprehensive picture of the events that led up to the end of one trial (Fig. 3B–F). Select video frames from one trial shows fly #1 (purple

trajectory) reaching the top by $t=8$ s but falling to $\sim 2/3$ of the vial height thereafter, while fly #6 (cyan trajectory) reaches $<1/3$ of the vial height by the end (Fig. 3B). Manual analysis would incorrectly exclude fly #1 from having displayed robust negative geotaxis before falling. Our method avoids such errors by monitoring fly positions throughout the experiment. We can construct a climbing curve for each trial with discrete steps representing the passage of single flies for a height of 7 cm (Fig. 3C, left). These data show that 100% of trials ended with 6 or 7 flies successfully passing the 7 cm height threshold, in notable contrast to traditional assessment (Fig. 3A). This discrepancy underscores a critical shortcoming of assessing climbing from the final video frame. Without knowledge of entire trajectories, recorded behavior is susceptible to brief events such as falls that can introduce significant errors. Our results show that detailed automatic tracking not only is convenient but also leads to more accurate assessment of climbing behavior.

The single-fly resolution of the staircase data also highlights individual variability generally disregarded in existing analysis methods. Although all trials ended with similar fly numbers (6 or 7) above threshold, each took a distinct path largely owing to behavioral variability in individuals. To assess the degree of

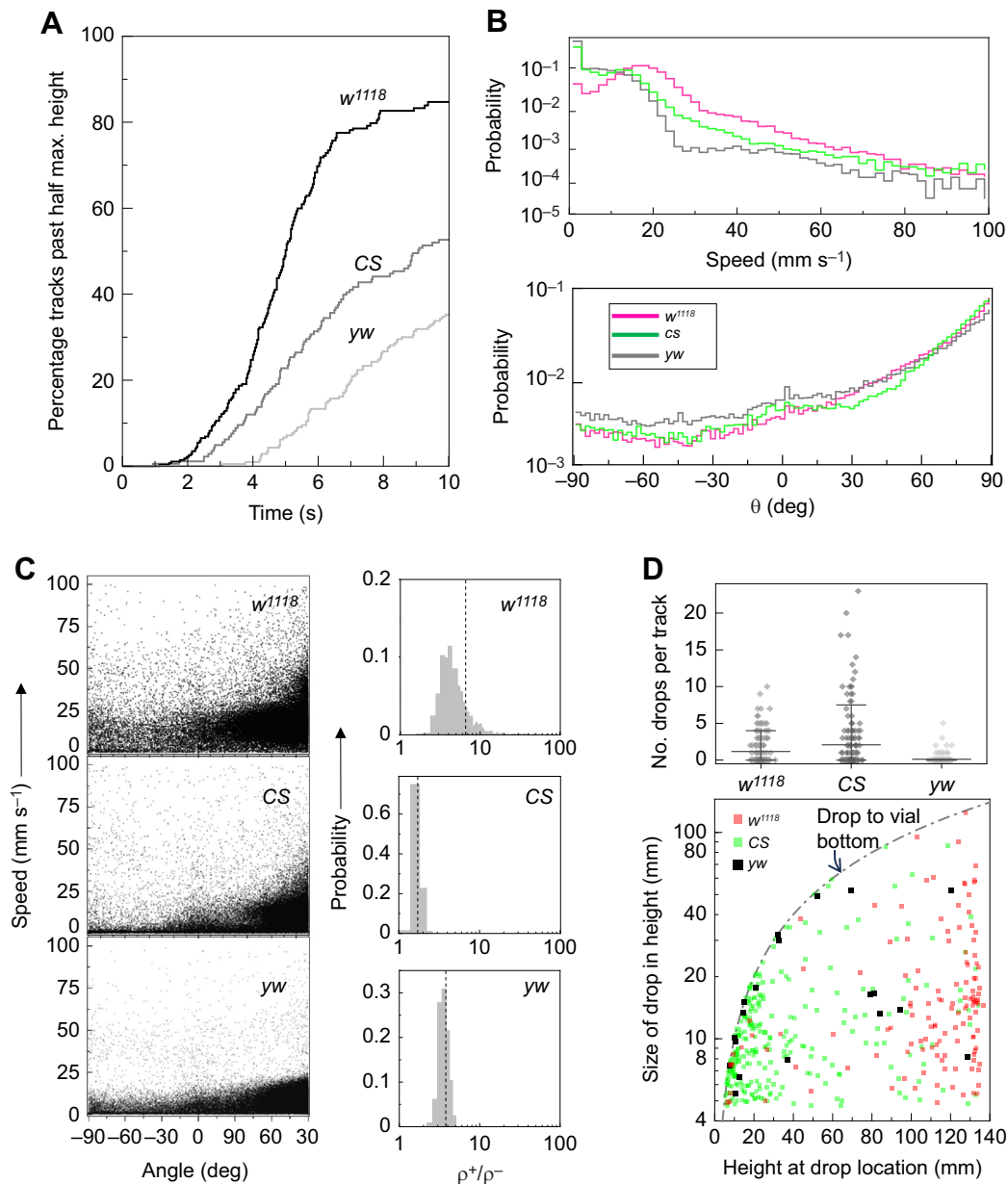


Fig. 4. New method identifies locomotor features underlying geotactic differences among *w¹¹¹⁸*, CS and *yw* strains. (A) Average climbing curves of 10 trials for *w¹¹¹⁸* ($N_{\text{flies}}=18$), CS ($N_{\text{flies}}=17$) and *yw* ($N_{\text{flies}}=21$) female flies. (B) Probability distributions of speed (top) and movement angle (bottom). Mean speed is 17.5 mm s⁻¹ (from $n=48,186$ values), 9.5 mm s⁻¹ (from $n=48,182$ values) and 6.9 mm s⁻¹ (from $n=59,845$ values) for *w¹¹¹⁸*, CS and *yw* flies, respectively; $P < 10^{-200}$ in all three pair-wise comparisons. Mean angle for *w¹¹¹⁸* and CS flies is 53.6 deg while that of *yw* is 43.6 deg. The *yw* distribution differs significantly from that of the other two groups ($P=2 \times 10^{-10}$). (C) Speed versus angle scatter plot (left) and ratio ρ^+/ρ^- of Spearman correlations (right) for each genotype. Bootstrapped histograms yield mean values (dashed vertical lines) of 6.6 ± 21.1 , 1.7 ± 0.17 and 3.7 ± 0.55 for *w¹¹¹⁸*, CS and *yw* flies, respectively; $P < 10^{-200}$ in all three pair-wise comparisons. (D) Top: number of slips or falls detected per track (circles, from all trials). *yw* versus *w¹¹¹⁸*, $P=6 \times 10^{-10}$; *yw* versus CS, $P=1.2 \times 10^{-14}$; *w¹¹¹⁸* versus CS, $P=0.29$. Mean, 10% and 90% of data are indicated. Bottom: height of fly versus drop size. P -values in B–D are from Kruskal–Wallis null hypothesis test followed by a *post hoc* test.

variability, we averaged staircase data for three different heights across 10 trials (Fig. 3C, right). We found that the mean variation is 13%, 18% and 28% for climbing to 1/3, 2/3 and the full 14 cm height, respectively. The between-trial variability increases with height because fewer flies climb to greater heights. To assess whether there are any general trends in performance, such as degradation in climbing ability due to fatigue (Bazzell et al., 2013; Damschroder et al., 2018), we next examined the average speed and angular direction of the group in each trial (Fig. 3D). These data, which include movements in all directions, show no particular trend

from the first to the last trial, indicating that fly movement kinetics do not change dramatically over the course of an experiment. We measured the angle of movement θ relative to the horizontal direction ($\pm x$) where θ is considered positive if it is in the upper quadrants and negative if it is in the lower quadrants (Fig. 3E, top).

While useful for learning about overall trends in geotactic kinetics, population averages can mask short time-scale dynamics and the relationship between speed and angle between consecutive frames (Fig. 3E). For the experiments on *w¹¹¹⁸* flies, we found $\sim 89\%$ of events have $\theta > 0$ (Fig. 3E, top histogram, orange) with an

associated speed of $21.9 \pm 12.7 \text{ mm s}^{-1}$ (mean \pm s.d.). For the remaining $\sim 11\%$ of movements with $\theta < 0$, the associated speed distribution has a smaller average and a larger variance, $19.2 \pm 17.8 \text{ mm s}^{-1}$ (Fig. 3E, right histogram). Though they yield similar mean values, speed histograms in the two directions assume different shapes – a peaked distribution for $\theta > 0$ and a relatively even distribution for $\theta < 0$. Differences in distribution shapes suggest the possibility of differences in how movement speed and angle are correlated, with the narrower distribution in the $+y$ direction implying a bias towards negative geotaxis, consistent with the strong climbing behavior displayed by these flies (Fig. 3C). Calculation of the Spearman correlation further supports the idea of a directional bias, with a correlation coefficient of $\rho_+ = 0.145$ ($P < 10^{-60}$) in the $+y$ direction and a correlation coefficient of $\rho_- = 0.039$ ($P = 0.09$) in the $-y$ direction. Robustness of negative geotaxis can depend on such directional bias in movement, and these analyses indicate we can use the ratio ρ_+/ρ_- as a new metric to assess the behavior.

Close examination of videos revealed that flies can occasionally lose traction and consequently slip or fall ('drop') during climbing. How often and how far they drop can impact their geotactic performance. To better characterize these disruptions in climbing, we defined a vertical drop $-\Delta y$ between two consecutive frames of more than a body length and up to three body lengths ($4.5 \text{ mm} \leq -\Delta y \leq 12 \text{ mm}$) as a slip. A larger vertical drop, $-\Delta y > 12 \text{ mm}$, is defined as a fall. In the w^{1118} cohort, we detected a total of 41 slips and 52 falls in 10 trials, indicating that $< 0.5\%$ of movements involved these large drops in height. Both slips and falls occurred at various heights, with a higher incidence rate above a height of $\sim 8 \text{ cm}$ (Fig. 3F). The increase in the incidence of slips and falls with height is consistent with the idea that the longer flies are engaged in climbing, the greater their chances of losing traction. Importantly, these data suggest that similar to the ratio ρ_+/ρ_- , slip and fall characterization can be used as an independent measure of locomotor coordination in geotaxis.

Comparison of genotypes

To further exploit the analytic capabilities of our platform, we next compare geotaxis of three widely used laboratory strains w^{1118} , *CS* and *yw*. Each genotype was measured in three independent experiments (Fig. S4). Our experiments showed that the strains differ in their geotactic performance, with w^{1118} on average showing the best and *yw* the worst climbing ability (Fig. 4A). We wondered whether our analyses could reveal kinetic details behind these differences. Speed distributions (Fig. 4B, top) show that w^{1118} flies have a significantly higher probability ($P = 0.31$) of moving at speeds $> 20 \text{ mm s}^{-1}$ compared with *CS* ($P = 0.09$) and *yw* ($P = 0.03$) flies. While *CS* and *yw* distributions roughly overlap for speeds $< 20 \text{ mm s}^{-1}$, *CS* flies are more likely than *yw* flies to move at higher speeds. These speed data yield median values of 16.3, 7.8 and 5.2 mm s^{-1} for w^{1118} , *CS* and *yw*, respectively, in accordance with their relative ranking in climbing rates. Movement angle distributions (Fig. 4B, bottom) showed that all three genotypes generally move in the $+y$ direction ($P \geq 0.81$). However, *yw* flies have a slightly higher tendency to display positive geotaxis, as their probability for all movement angles $\theta \leq 0$ is consistently higher than that of the other two genotypes. In addition to speed, the number of turns during walking can serve as an alternative window into locomotor activity (see Materials and Methods).

Data on speed–angle correlations indicate that w^{1118} flies have the largest bias in the $+y$ direction with a mean ρ_+/ρ_- of 6.6, followed by *yw* flies with a ρ_+/ρ_- of 3.7 and, lastly, *CS* with a ρ_+/ρ_- of 1.7

(Fig. 4C). The large correlation bias for w^{1118} flies is consistent with their superior climbing rate. But a larger *yw* bias compared with *CS* flies is surprising, as *yw* flies perform poorly in overall climbing. The *CS*–*yw* difference suggests that a poor ρ_+/ρ_- ratio is not a major impediment to climbing ability in *CS* relative to *yw*. We also compared the incidence rate of slips and falls in the three groups. The slips and falls data show that *yw* flies have the lowest rate, ~ 11 per 100 tracks on average, compared with ~ 117 per 100 tracks in w^{1118} flies and ~ 200 per 100 tracks in *CS* flies (Fig. 4D, top). Thus, based solely on the drop rates, *yw* flies would be expected to be the best climbers and *CS* or w^{1118} flies the worst. Examination of the drop locations and magnitudes provides additional insights (Fig. 4D, bottom). These data show that the majority of w^{1118} drops occur $> 110 \text{ mm}$ above the bottom, while the majority of *CS* and *yw* drops occur below $\sim 50 \text{ mm}$ from the bottom. As w^{1118} flies drop from greater heights, it is reasonable to expect the size of their drops to be larger. Yet, we find that w^{1118} and *yw* flies have on average the same drop size of $\sim 20 \text{ mm}$ while those of *CS* flies are statistically smaller, $\sim 12.6 \text{ mm}$ (Fig. S5A). When the size of a drop is compared with the maximum possible size of the drop (this ratio is 1 along the curved dashed line in Fig. 4D), we find that it is 0.22, 0.53 and 0.62 for w^{1118} , *CS* and *yw* flies, respectively; that is, a larger fraction of *yw* drops bring the flies to the bottom of their vial (Fig. S5B).

In short, we find that the two best predictors of w^{1118} , *CS* and *yw* relative climbing performance are speed and the size of drop compared with its maximum possible size. The other metrics we introduced – such as movement angle, ρ_+/ρ_- ratio and the frequency of drops – only partially capture relative geotaxis ability of the three strains. For instance, ρ_+/ρ_- results correctly predict that w^{1118} flies are the best climbers but incorrectly predict that *yw* flies are superior to *CS* flies. The incongruent conclusions drawn from the different metrics indicate that they capture different features of locomotion and are, at least partly, independent of each other. Having several independent metrics is a powerful analytic resource whose potential for uncovering geotactic aberrations will become clearer through future quantitative studies (Lobato et al., 2024).

DISCUSSION

The custom tracking and analyses programs and detailed instructions on how to use them are freely available from GitHub (<https://github.com/sheyums/Insect-Geotaxis>). This work joins several recent efforts to develop platforms for robust examination of fruit fly geotactic behavior. Our new apparatus builds upon prior experimental procedures by incorporating mostly off-the-shelf components in an intuitive design that allows stable, high-throughput and reproducible measurements. Computer control of the apparatus permits changes in experimental parameters. We have shown that single-fly trajectory reconstructions yield both accurate climbing curves and a wealth of kinetic metrics, enabling a deeper understanding of geotactic behavior than what has been possible thus far.

Tracking multiple freely moving flies over time is the most challenging part of single-fly trajectory reconstructions. We addressed the challenge with the application of two noteworthy computer vision tools. One is a predictive algorithm, called the Kalman filter, which provides an estimate of the most likely match between a new detection with existing tracks. We use the Kalman estimate together with a separate estimate based on previous detection coordinates. The joint estimate provides a key input for the second important computer vision algorithm, called the Hungarian method, that optimizes detection-to-track association.

The number of flies that climb to a given height within an allotted time is the most used measure of geotaxis. Current platforms with limited tracking ability report final fly count, low time resolution intermediate counts or climbing curves for only single flies per vial (Gargano et al., 2005; Podratz et al., 2013; Willenbrink et al., 2016). Other methods with ostensibly better tracking still do not present climbing curves but instead rely on other population-wide metrics (Cao et al., 2017; Damschroder et al., 2018; Kohlhoff et al., 2011; Spierer et al., 2021). Our algorithm can track multiple flies simultaneously ~97% of the time with millimeter accuracy. The climbing curves in our studies provide a dynamic picture, revealing single-fly movements with a 1/30 s temporal resolution. The high-resolution tracking allows us to dissect geotactic behavior at the level of single video frames, demonstrated by quantifying instantaneous movement angle and speed and by conducting meta-analysis on their correlations across three different laboratory strains. Flies are known to lose traction occasionally during climbing, but prior *Drosophila* studies overlooked these events. Quantification of such falls is critical, especially in fly disease models (Chaudhuri et al., 2007; Zhu et al., 2019), as several human neurological disorders associated with loss of balance show an increased frequency of falls (Bloem et al., 2001; Stolze et al., 2004). We demonstrate that accounting for these rare events can uncover the nature of movement deficiencies underlying defective geotaxis.

Acknowledgements

We thank Manuel Collazo for assisting with construction of the apparatus, Amanda Lobato and Jiaqi Liu for providing fly culture support, and Camille Free and Samantha Lattanze with helping with manual calibration of videos.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: T.C., J.L., G.Z., S.S.; Methodology: T.C., J.L., S.S.; Software: T.C., S.S.; Validation: T.C., J.L., N.O.-V.; Formal analysis: T.C., S.S.; Investigation: T.C.; Resources: N.O.-V., G.Z., S.S.; Writing - original draft: T.C., S.S.; Writing - review & editing: G.Z., S.S.; Supervision: S.S.; Project administration: S.S.; Funding acquisition: G.Z., S.S.

Funding

This work was supported in part by the National Science Foundation under grant no. 2131037 to S.S. and National Institutes of Health grant no. R33AT010408 to R.G.Z. Open Access funding provided by University of Miami. Deposited in PMC for immediate release.

Data availability

All custom-written Matlab software and example data for tracking and analyses is available at <https://github.com/sheyums/Insect-Geotaxis>.

References

- Armstrong, J. D., Texada, M. J., Munjaal, R., Baker, D. A. and Beckingham, K. M. (2006). Gravitaxis in *Drosophila melanogaster*: a forward genetic screen. *Genes Brain Behav.* **5**, 222-239. doi:10.1111/j.1601-183X.2005.00154.x
- Bazzell, B., Ginzberg, S., Healy, L. and Wessells, R. J. (2013). Dietary composition regulates *Drosophila* mobility and cardiac physiology. *J. Exp. Biol.* **216**, 859-868. doi:10.1242/jeb.078758
- Bender, J. A. and Frye, M. A. (2009). Invertebrate solutions for sensing gravity. *Curr. Biol.* **19**, R186-R190. doi:10.1016/j.cub.2008.12.024
- Bloem, B. R., Grimbergen, Y. A. M., Cramer, M., Willemsen, M. and Zwinderman, A. H. (2001). Prospective assessment of falls in Parkinson's disease. *J. Neurol.* **248**, 950-958. doi:10.1007/s004150170047
- Cao, W., Song, L., Cheng, J., Yi, N., Cai, L., Huang, F.-De. and Ho, M. (2017). An automated rapid iterative negative geotaxis assay for analyzing adult climbing behavior in a *Drosophila* model of neurodegeneration. *J. Vis. Exp.* **2017**, 56507.
- Chaudhuri, A., Bowling, K., Funderburk, C., Lawal, H., Inamdar, A., Wang, Z. and O'Donnell, J. M. (2007). Interaction of genetic and environmental factors in a *Drosophila* parkinsonism model. *J. Neurosci.* **27**, 2457. doi:10.1523/JNEUROSCI.4239-06.2007
- Damschroder, D., Cobb, T., Sujkowski, A. and Wessells, R. (2018). *Drosophila* endurance training and assessment of its effects on systemic adaptations. *BIO-Protocol* **8**, e3037. doi:10.21769/BioProtoc.3037
- Gargano, J., Martin, I., Bhandari, P. and Grotewiel, M. (2005). Rapid iterative negative geotaxis (RING): a new method for assessing age-related locomotor decline in. *Exp. Gerontol.* **40**, 386-395. doi:10.1016/j.exger.2005.02.005
- Hirsch, J. and Erlenmeyer-Kimling, L. (1961). Sign of taxis as a property of the genotype. *Science* **134**, 835-836. doi:10.1126/science.134.3482.835
- Kamikouchi, A., Inagaki, H. K., Effertz, T., Hendrich, O., Fiala, A., Göpfert, M. C. and Ito, K. (2009). The neural basis of *Drosophila* gravity-sensing and hearing. *Nature* **458**, 165-171. doi:10.1038/nature07810
- Kladt, N. and Reiser, M. B. (2023). *Drosophila* antennae are dispensable for gravity orientation. *bioRxiv* 2023.03.08.531317. doi:10.1101/2023.03.08.531317
- Kohlhoff, K. J., Jahn, T. R., Lomas, D. A., Dobson, C. M., Crowther, D. C. and Vendruscolo, M. (2011). The iFly tracking system for an automated locomotor and behavioural analysis of *Drosophila melanogaster*. *Integr. Biol.* **3**, 755-760. doi:10.1039/c0ib00149j
- Lobato, A. G., Ortiz-Vega, N., Canic, T., Tao, X., Bucan, N., Ruan, K., Rebelo, A. P., Schule, R., Zuchner, S., Syed, S. et al. (2024). Loss of Fic causes progressive neurodegeneration in a *Drosophila* model of hereditary spastic paraplegia. *Biochim. Biophys. Acta Mol. Basis Dis.* **1870**, 167348. doi:10.1016/j.bbdis.2024.167348
- Munkres, J. (1957). Algorithms for the assignment and transportation problems. *J. Soc. Ind. Appl. Math.* **5**, 32-38. doi:10.1137/0105003
- Narayanan, S. A. (2023). Gravity's effect on biology. *Front. Physiol.* **14**, 1199175. doi:10.3389/fphys.2023.1199175
- Podratz, J. L., Staff, N. P., Boesche, J. B., Giorno, N. J., Hainy, M. E., Herring, S. A., Klennert, M. T., Milaster, C., Nowakowski, S. E., Krug, R. G. et al. (2013). An automated climbing apparatus to measure chemotherapy-induced neurotoxicity in *Drosophila melanogaster*. *Fly* **7**, 187-192. doi:10.4161/fly.24789
- Spierer, A. N., Yoon, D., Zhu, C. T. and Rand, D. M. (2021). FreeClimber: automated quantification of climbing performance in *Drosophila*. *J. Exp. Biol.* **224**, jeb229377. doi:10.1242/jeb.229377
- Stolze, H., Klebe, S., Zechlin, C., Baecker, C., Friege, L. and Deuschl, G. (2004). Falls in frequent neurological diseases: Prevalence, risk factors and aetiology. *J. Neurol.* **251**, 79-84. doi:10.1007/s00415-004-0276-8
- Takahashi, K., Takahashi, H., Furuichi, T., Toyota, M., Furutani-Seiki, M., Kobayashi, T., Watanabe-Takano, H., Shinohara, M., Numaga-Tomita, T., Sakaue-Sawano, A. et al. (2021). Gravity sensing in plant and animal cells. *Npj Microgravity* **7**, 1-10.
- Willenbrink, A. M., Gronauer, M. K., Toebben, L. F., Kick, D. R., Wells, M. and Zhang, B. (2016). The Hillary Climber trumps manual testing: an automatic system for studying *Drosophila* climbing. *J. Neurogenet.* **30**, 205-211. doi:10.1080/01677063.2016.1255211
- Zhu, Y., Li, C., Tao, X., Brazill, J. M., Park, J., Diaz-Perez, Z. and Grace Zhai, R. (2019). Nmnat restores neuronal integrity by neutralizing mutant Huntingtin aggregate-induced progressive toxicity. *Proc. Natl. Acad. Sci. USA* **116**, 19165-19175. doi:10.1073/pnas.1904563116
- Zhu, Y., Lazopulo, S., Syed, S. and Zhai, R. G. (2022). Behavioral studies in *Drosophila* models of human diseases. In *Encyclopedia of Behavioral Neuroscience*, Vol. 1-3, 2nd edn, pp. 13-23. Elsevier.

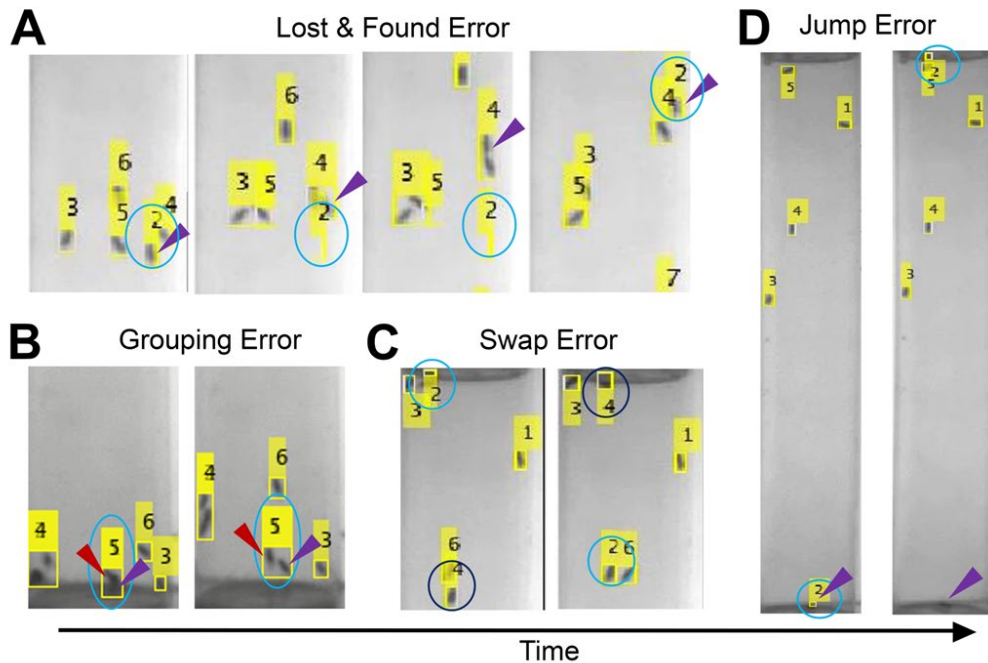


Fig. S1. Frequently occurring tracking error examples. (A) Four frames of the same video captured within 0.5 second period show how a fly disappears owing to proximity with another and reappearing once they separated. The cyan circle indicates where the track is located, while the purple arrowhead indicates the true location of the fly at each timepoint. (B) Two frames of the same video captured within 0.1 second show how two flies can be mistaken as one detection if they are too close in proximity. The cyan oval indicates where the singular track is located, the red arrow points to fly A, the purple arrow points to fly B at each timepoint. (C) Two frames of the same video captured within 0.3 second show how two tracks can switch between two flies even if they are not proximal. The cyan circle indicates track 2, the navy circle indicates track 4 at each timepoint. Note, the flies are not moving very much between these two frames. (D) Two frames of the same video captured within 0.1 seconds show how noise can create a false detection and force the track to jump from a stationary fly at the bottom to a moving shadow at the top. The cyan circle indicates the track while the purple arrow indicates the fly associated with that track at each timepoint.

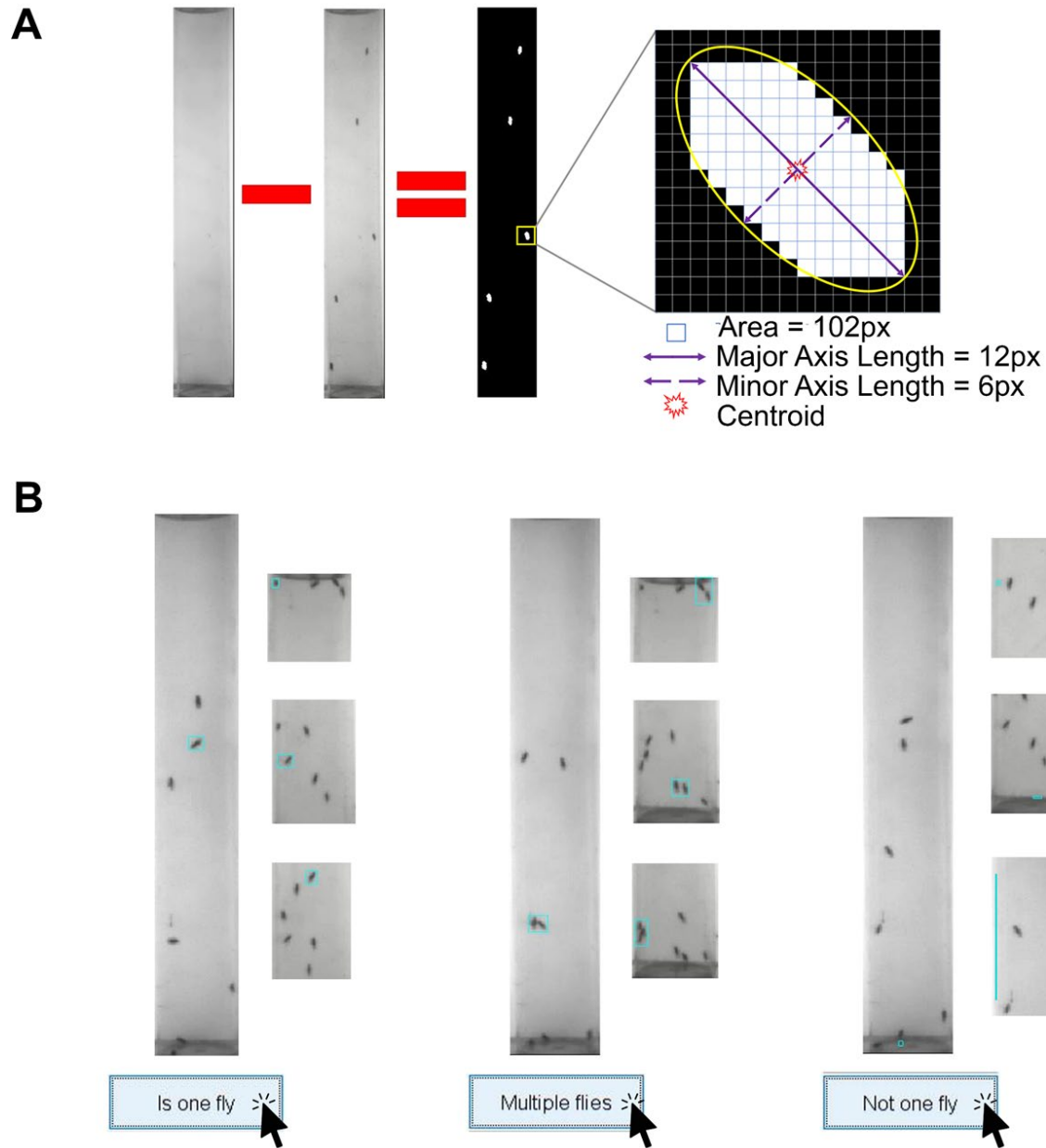


Fig. S2. An elaboration of calibration classifications for more precise object detection. (A) Visual representation of background subtraction with a zoomed in depiction of important information regarding each detection. Typical properties of a detected fly are noted below zoom. (B) Real images showcasing possible detections sorted by manual calibration. From left to right shows two columns each of singular flies per detection, multiple flies per detection, and incorrect detections (noise). Detected objects are inside cyan rectangles.

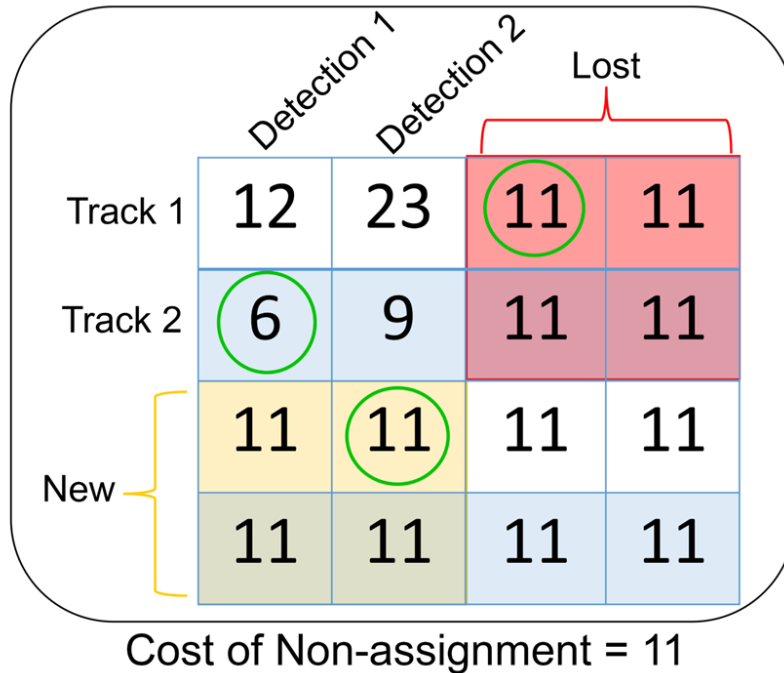


Fig. S3. An example of a padded cost matrix. Rows are tracks that existed in the previous frame of the video. Columns are detections found in the current frame of the same video. The cost matrix is padded by the cost of non-assignment to allow for new tracks to be created or lost, indicated by the yellow and red regions, respectively. Circled elements indicate the lowest total cost, thus the final assignment of the detection to track.

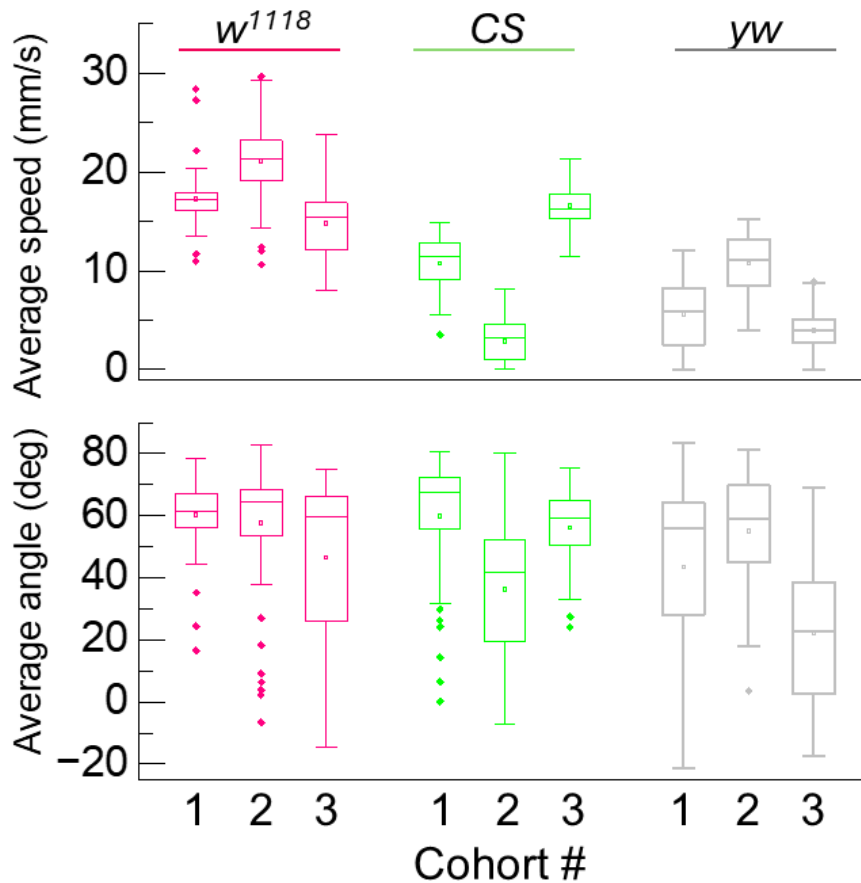


Fig. S4. Speed and angular movement data from Figure 4 averaged to illustrate experimental variability between cohorts of the same genotype. Each box plot was constructed with average speed (upper) and average angle (lower) of individual flies of a cohort. Individual averages were calculated from each trial and collated from 10 trials per cohort. Box-and-whisker plots have different representations in the literature. Here: top and bottom of box indicate 1st and 3rd quartiles, horizontal line within box indicates the median, square represents the mean, whiskers include 1.5 times the inter-quartile range, and symbols (when present) indicate outliers.

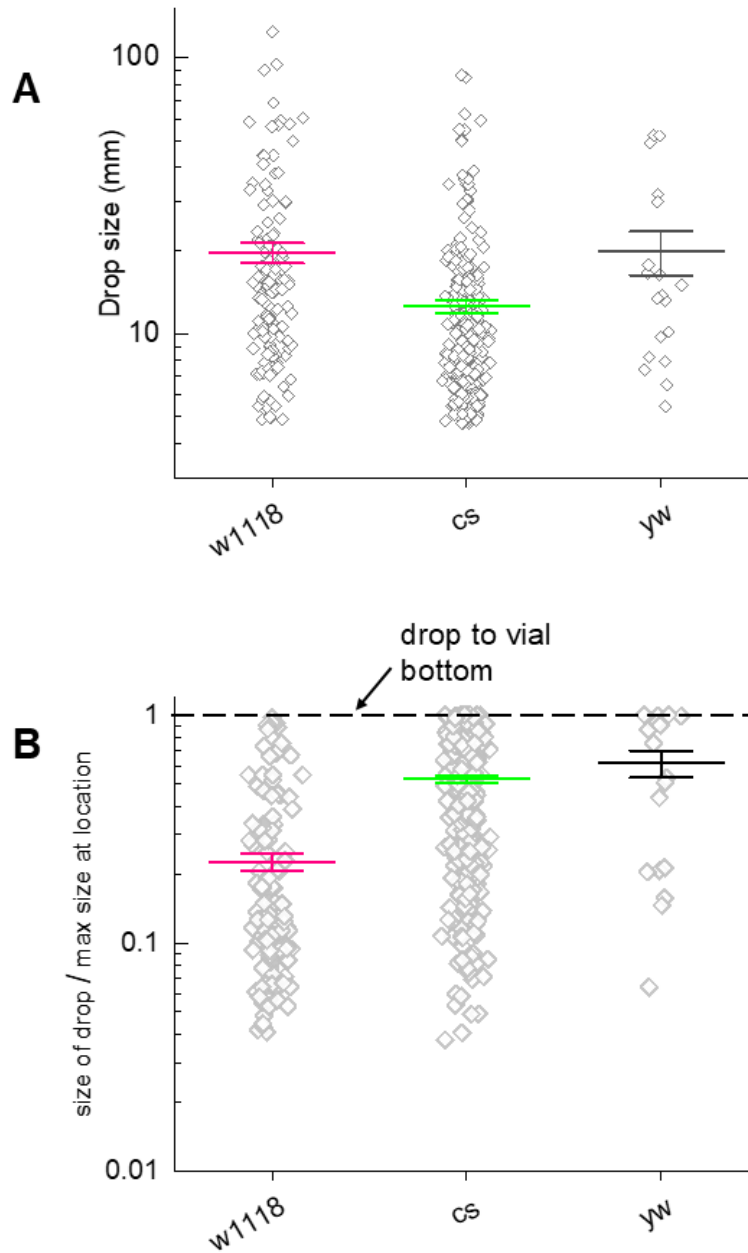


Fig. S5. A closer examination of drop statistics. $N = 18 w^{1118}$, 17 *CS*, and 21 *yw* from Figure 4. (A) Size of slips and falls of the three strains. w^{1118} and *yw* drop sizes are similar ($P=0.92$) but larger than *CS* ($P=2 \times 10^{-6}$ vs. w^{1118} and $P=0.024$ vs. *yw*). (B) Ratio of drop size to maximum possible drop from that height. The fractional drop size of w^{1118} flies is on average the smallest ($P < 7 \times 10^{-6}$) while that of *CS* are moderately smaller than *yw* ($P=0.5$). Mean \pm standard errors are shown in both panels. P values are from Kruskal-Wallis null hypothesis test followed by a post hoc test.



Movie 1. An example of a typical experimental trial video. A forward-facing 12 second recording of the vial mount from the automated apparatus showcasing 10 wild type cohorts. Vials 1-3 contain *yw* females (N=7 flies/vial), vial 4 contains CS females (N=6), vial 5 contains *w¹¹¹⁸* females (N=3), vial 6 contains *w¹¹¹⁸* males (N=5), vial 7 contains CS males (N=5), vials 8-10 contain *yw* males (N=4,6,5 respectively).



Movie 2. An example of a ROI video after video processing. From the original trial video (Movie 1), the user defined Region of Interest (ROI) is used to cut the 12 second video into individual vial videos. This example is vial 2, female yw flies (N = 7).



Movie 3. A real time visual representation of the output from the tracking algorithm.

Using the same ROI video from Movie 2 (vial 2, *yw* females, N=7), per frame each fly is detected, shown by the yellow bounding boxes, and assigned a corresponding number attached by yellow flags. The centers of the bounding boxes in each video frame correspond to the coordinates saved in the algorithms output matrix. The video shows several collisions between flies. Analysis of a larger dataset of 79800 frames (266 flies x 10 seconds x 30 frames s⁻¹) detected 2762 collisions, amounting to ~ 3.5% of possible instances. Not every collision leads to switching in identities. As an upper estimate, track label swapping occurs in < 2% of the time.