Biophysical Journal
Article

Biophysical Society

Transport of lysosomes decreases in the perinuclear region: Insights from changepoint analysis

Nathan T. Rayens,[1] Keisha J. Cook,[2] Scott A. McKinley,[2] and Christine K. Payne[1,*]
[1]Thomas Lord Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina and [2]Department of Mathematics, Tulane University, New Orleans, Louisiana

ABSTRACT Lysosomes are membrane-bound organelles that serve as the endpoint for endocytosis, phagocytosis, and autophagy, degrading the molecules, pathogens, and organelles localized within them. These cellular functions require intracellular transport. We use fluorescence microscopy to characterize the motion of lysosomes as a function of intracellular region, perinuclear or periphery, and lysosome diameter. Single-particle tracking data are complemented by changepoint identification and analysis of a mathematical model for state switching. We first classify lysosomal motion as motile or stationary. We then study how lysosome location and diameter affects the proportion of time spent in each state and quantify the speed during motile periods. We find that the proportion of time spent stationary is strongly region dependent, with significantly decreased motility in the perinuclear region. Increased lysosome diameter only slightly decreases speed. Overall, these results demonstrate the importance of decomposing particle trajectories into qualitatively different behaviors before conducting population-wide statistical analysis. Our results suggest that intracellular region is an important factor to consider in studies of intracellular transport.

SIGNIFICANCE Intracellular transport is an essential cellular function. Characterizing and understanding this motion is challenging because of the complex environment of the cell, heterogeneity of transport organelles, and limits inherent to any experimental method. We implement an automated changepoint analysis and, on the basis of the segmented paths, conduct a rigorous statistical and mathematical analysis of lysosomal motion. We quantify the relative impact of intracellular region and lysosome diameter on lysosomal motion. Both factors can be altered as a function of disease. The methods presented here provide a comprehensive template for coupling complex biological data with mathematical analysis. This approach reveals that the intracellular region, perinuclear or periphery, is an important factor in the intracellular motion of lysosomes and possibly other organelles and molecules undergoing transport.

INTRODUCTION

Lysosomes are membrane-bound organelles responsible for processing endocytic molecules, particles, and viruses; phagocytic destruction of pathogens; and the cellular housekeeping of autophagy (1–4). They are distributed throughout the cell, with clustering in the perinuclear region (5,6). As essential intracellular transport vesicles, there has been much work (7–10), including our own (11), examining the relationship between lysosome function, transport, and diameter. In addition to the fundamental knowledge gained by understanding the biophysics of intracellular transport, there are also implications for human health. Lysosome diameter can be increased by disease or exposure to nanoparticles (2,12–14). For example, cystinosis, a lysosomal storage disorder that leads to kidney and eye damage, is associated with enlarged lysosomes (12).

Lysosomal motion is a combination of adenosine triphosphate (ATP)-dependent transport along microtubules and non-ATP-dependent motion, typically attributed to either free or constrained diffusion within the cytosol (3,4,11,15,16). Transport along microtubules is required for efficient processing of endocytic cargo, including delivery of endocytic cargo from late endosomes (15,17–19). Diffusive motion lacks the length scale necessary to

Submitted August 24, 2021, and accepted for publication February 17, 2022.

*Correspondence: christine.payne@duke.edu

Keisha J. Cook's present address is School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA

Editor: William O. Hancock.

https://doi.org/10.1016/j.bpj.2022.02.032

© 2022 Biophysical Society.

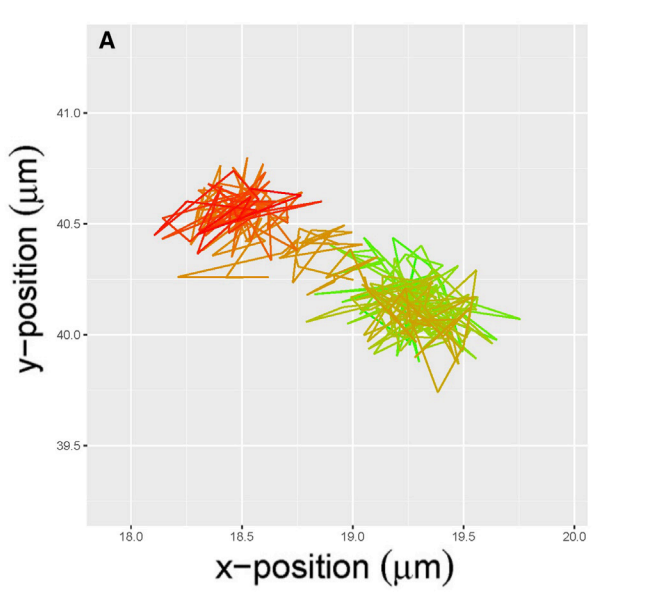
transport lysosomes across a cell on a relevant timescale. Our previous research examined the effect of lysosome diameter on intracellular transport using fluorescence microscopy and single-particle tracking to characterize the motion of individual lysosomes and mean-squared displacement (MSD)-based analysis (11). This research showed that the slope of the ensemble MSD curve was significantly decreased for larger lysosomes. The speed of unidirectional transport was not affected by increased lysosome diameter.

In comparison with our previous study (11), the work described here is based on trajectories recorded with a faster frame rate (20 vs. 3.3 Hz). As a result, distinct states are more apparent, and we can develop a comprehensive mathematical analysis of lysosome transport using a changepoint detection algorithm (Fig. 1). The changes in state can be seen more readily when the evolution of the x and y coordinates are plotted with respect to time (Fig. 1, *B* and *C*; and a summary of the inferred path properties is shown in Fig. 1 *D*). This procedure allows efficient analysis of large populations of trajectories, which enables rigorous comparisons of path properties among different groups. This provides a more detailed approach than methods that rely on comparison of MSD slopes. After establishing a threshold for *motile* and *stationary* behavior, we can compare the time spent in these different states, as well as the distribution of lysosome speeds while motile. With these methods, we are able to show that the intracellular region, perinuclear or periphery, of the lysosomes is an important factor to consider in the study of intracellular transport. Although larger lysosomes have a slightly slower speed of transport, this effect is small relative to the effect of intracellular region.

MATERIALS AND METHODS

Cell culture, fluorescent labeling, and sucrose treatment

BS-C-1 monkey kidney epithelial cells (obtained from the Duke University Cell Culture Facility) were cultured in Dulbecco's modified Eagle medium



D Summary of Segments

Segment	1	2	3
Duration	8.4 s	2.8 s	4.9 s
Speed	16.4 nm/s	418 nm/s	6.79 nm/s
State	Stationary	Motile	Stationary

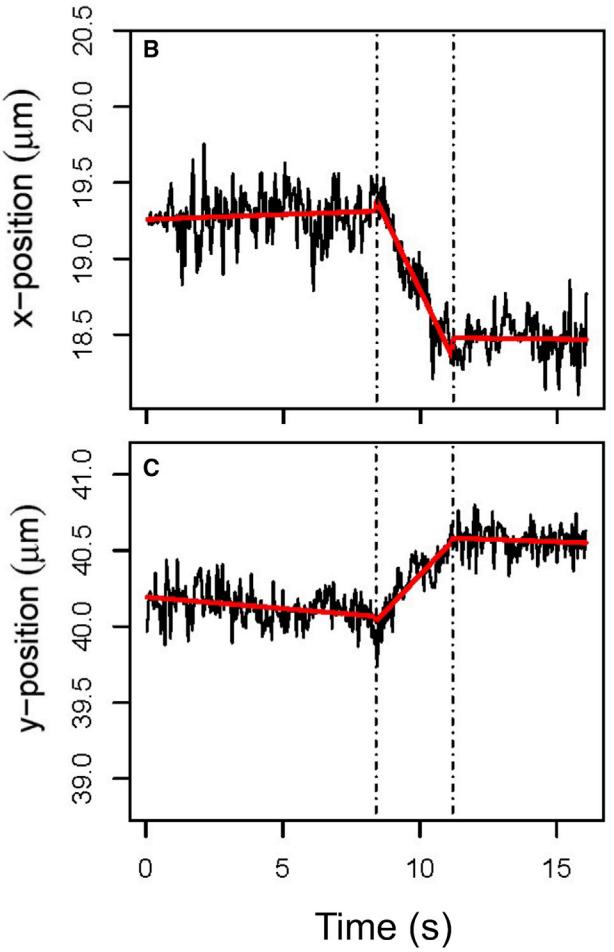


FIGURE 1 Example of the changepoint analysis method applied to lysosomal motion. (*A*) Original trajectory (20 Hz). Green is the start of the trajectory and red is the end of the trajectory. The corresponding time-lapse images and an analysis of localization uncertainty are included in the Supporting material (Fig. S2). (*B* and *C*) The t-versus-x and t-versus-y time series of the original trajectory. The dashed lines represent the changepoints. The result of the segmentation procedure is overlaid in red. (*D*) The estimated duration, speed, and state of each segment in the trajectory. A segment is labeled motile if its estimated speed is greater than 100 nm/s. To see this figure in color, go online.

(catalog no. 12100046; Thermo Fisher Scientific, Waltham, MA). A549 human lung epithelial cells (obtained from the Duke University Cell Culture Facility), described in the Supporting information, were cultured in F-12K medium (catalog no. N3520; Sigma-Aldrich, St. Louis, MO). Besides cell culture medium, all other conditions were identical for the two cell lines. Culture media was supplemented with 10% fetal bovine serum (FBS; catalog no. 10437028; Thermo Fisher Scientific), and cells were incubated at 37°C and 5% carbon dioxide. Cells were passaged with trypsin (catalog no. 25200072; Thermo Fisher Scientific) every 2–3 days. For imaging experiments, 35-mm optical dishes were used (catalog no. 150682; Thermo Fisher Scientific).

For live-cell imaging and single-particle tracking, cells were transduced with CellLight Lysosomes-green emerald fluorescent protein (BacMam 2.0, catalog no. C10596; Thermo Fisher Scientific) to fluorescently label lysosomes. Transduction was carried out according to the manufacturer's instructions, with 7.5 μL found to be the optimal volume. Immunofluorescence was used to label lysosomes for the measurement of diameter and is described below.

Sucrose (50 mM, 24 h; catalog no. JTB-4072-01; J.T. Baker, Phillipsburg, NJ) was used to selectively enlarge lysosomes as described previously (11,20–22).

Immunofluorescence

Immunofluorescence using a Cy3-lysosome associated membrane protein 1 (LAMP1) antibody (catalog no. ab67283; Abcam, Cambridge, UK) was used to label lysosomes for measurement of diameter. Cells were fixed with 4% paraformaldehyde (catalog no. 15710; Electron Microscopy Services, Hatfield, PA) in phosphate-buffered saline (PBS; catalog no. 21300-025; Thermo Fisher Scientific) for 10 min at room temperature. The cells were then washed with PBS and permeabilized (10% FBS, 0.3% bovine serum albumin [catalog no. A2153; Sigma-Aldrich], 0.00005% Triton-X 100 [catalog no. T8787; Sigma-Aldrich]) in PBS for 3–5 min at room temperature. Cells were washed with PBS and blocked (10% FBS, 0.3% body surface area in PBS) for 1 h at 4°C. After 1 h, the blocking buffer was removed and replaced with fresh blocking buffer and the Cy3-LAMP1 antibody (1:750). The cells were stained for 3–18 h at 4°C and washed with PBS. 4′,6-Diamidino-2-phenylindole (DAPI) was used as a nuclear stain (catalog no. 10236276001; Sigma-Aldrich; 50 μM in PBS, 4°C for 30 min).

Immunofluorescence images were recorded with an inverted microscope (IX-71; Olympus, Tokyo, Japan) using a 1.20 N.A., 60×, water-immersion objective (UPlanSApo 60×/1.20w; Olympus) with TRITC (Cy3-LAMP) and DAPI filter cubes and an electron-multiplying charge-coupled device camera (DU-888E; Andor, Belfast, United Kingdom). To minimize selection bias for lysosome diameter measurements, a grid of 750-μm^2 squares was overlaid on each image using ImageJ (23). The lysosome closest to the left of each square was selected (the lower in the event of two left-aligned lysosomes) for analysis and measured along its long axis using the line tool in ImageJ.

Live-cell imaging and single-particle tracking

Live-cell imaging and single-particle tracking were used to observe and characterize lysosome motion. Live-cell imaging was carried out using a spinning disk confocal microscope (IX-81 [Olympus], CSU-X1 confocal scanning unit [Yokogawa, Tokyo, Japan], iXon X3 camera [Andor], and MetaMorph version 7.8.2.0 software [Molecular Devices, San Jose, CA]). Images were collected using the green fluorescent protein excitation filter, with the stream function (MetaMorph) used to control exposure time, field-of-view area, and gain for the desired frame rate. These thin (∼5 μm) epithelial cells, combined with confocal microscopy, provide a two-dimensional framework and are considered isotropic for particle tracking (24). Motion in the z dimension would be observed as the loss of a particle during tracking as it moves out of the focal volume. This was not observed in the trajectories. Time-lapse images were exported to FIJI (version 2.0.0-rc-69./1.52p), noise was removed with the despeckling filter, and the TrackMate macro was used to track lysosomes (25). A comparison of raw and despeckled images confirmed that the despeckling filter did not affect analysis (Fig. S1). To minimize bias, a random number generator was used to select 40 trajectories from each observed cell. Handling of tracker error is described below (Preparation of the data). For each accepted trajectory, lysosome diameter was measured along the long axis using the line tool. These diameter measurements were used to define "large" and "small" lysosomes for comparison of speeds during motile transport. The perinuclear region was defined by identifying the nucleus in the fluorescence microscopy images and defining a border between the densely packed lysosomes and the sparser population in the periphery.

Preparation of the data

Particle-tracking software regularly makes errors, and all trajectories were confirmed manually. Trajectories with unphysical motion, switching to a different particle, or loss of a tracked particle for greater than five consecutive frames were not used for further analysis. Less obvious tracker errors, such as the software switching briefly to a different trajectory or temporarily switching to a high-intensity pixel that does not correspond to a physical object, were removed and replaced using the following procedure. First, because the first or last steps of reported trajectories were so commonly tracker errors, we removed them from all trajectories. Next, we fit a smooth spline to the t-versus-x and t-versus-y components of the lysosome trajectory (R command smooth.spline with 10 degrees of freedom) to set a baseline for comparison. At each time point, t_n, we computed the distance between the reported particle location and the smooth spline projection. A point was labeled an outlier if its distance was greater than the 75th percentile by a factor of 2.5 times the interquartile range. An example of this labeling technique to identify outliers has been described previously (26), but with multiplicative factor 1.5 instead of 2.5. We made the adjustment to account for variation in the accuracy of the spline. If the first or last point of a trajectory (the second or second to last point in the original trajectory) was considered an outlier, we removed it. Otherwise, we replaced the outlier by selecting a position that was normally distributed with a mean located at the spline location associated with the appropriate time and a standard deviation estimated from a subset of typical paths. There were also occasional missing data points. We used the same process to generate artificial data points when obeservations were not evently spaced in time.

MATHEMATICAL MODELS

Numerous methods are available for analyzing particle trajectories and identifying underlying biophysical states, and a number of tools have been developed for studying the MSD of trajectories in particular (see Monnier et al. [27] for one example and Manzo and Garcia-Parajo [28] for a review). However, MSD calculations average over an entire path (or over multiple paths) and short-lived changes of state can be missed. Therefore, some more recent methods have been developed to identify changes of state within individual particle trajectories. Some focus on discovering changes in diffusivity (29,30), while others are focused on discrete-step transitions (31), changes in velocity (32,33), or changes in velocity or diffusivity (34).

On the basis of an analysis conducted by Jensen et al. (33), we selected the bcp algorithm, implemented in R, to partition the paths into segments (see below for further discussion). Once a segmentation was selected for a given path, we modeled the data as piecewise-linear plus stationary noise.

Suppose that a lysosome is observed at times $\{t_i\}_{i=1}^n$, and the observed locations are $\boldsymbol{\varphi} = \{(x_i, y_i)\}_{i=1}^n$. (Vectors in this work are marked in boldface type.) To analyze the trajectories, we assume that they are composed of a sequence of segments in which motion is essentially linear with constant speed but subject to stationary fluctuations that arise from a combination of measurement error and Brownian diffusion while constrained within the cell. In terms of notation, we introduce the following. For each path, there is a finite set of K changepoints, $\boldsymbol{c} = (c_1, \ldots, c_K)$. When necessary, we will adopt the convention $c_0 = 0$ and $c_{K+1} = T$, where T is the final time of observation. Associated with each segment is a velocity vector, and we write this set of velocities $\{(u_k, v_k)\}_{k=1}^{K+1}$. When referring to an entire observed lysosome trajectory, we write $\boldsymbol{\varphi}$; when referring to the piecewise-linear approximation that we derive for it, we write $\widehat{\boldsymbol{\varphi}}$.

Model for lysosome observations

We treat the lysosome observations as being Gaussian fluctuations around a sequence of unobserved *anchor locations*, which are denoted $\{(a_i, b_i)\}_{i=1}^n$. We write

$$x_i = a_i + \sqrt{\beta}\varepsilon_i^{(x)}; \quad y_i = b_i + \sqrt{\beta}\varepsilon_i^{(y)}; \tag{1}$$

Here, $\{(\varepsilon_i^{(x)}, \varepsilon_i^{(y)})\}_{i=1}^n$ is a sequence of independent and identically distributed two-dimensional standard normal random variables with noise magnitude β that is the confluence of physical fluctuations, imprecision in the anchor location, nonlinear microtubules, and experimental error. Because we will not be studying any of these effects, we will not attempt to separate them properly in the model.

As discussed before, we model the anchor trajectories as being piecewise-linear. Using the notation introduced in the previous section, within each segment of constant velocity we have

$$a_i = v_k(t_i - t_{c_{k-1}}) + a_{c_{k-1}}, \tag{2}$$

where $k \in \{1, 2, \ldots, K+1\}$ is the index of the segment, and $i \in \{c_{k-1}+1, \ldots c_k\}$. From this approximation we infer quantities such as segment duration and velocity and construct empirical speed distributions to capture heterogeneous behavior among trajectories for a given population.

Model for transport states and state switching

Our analysis makes no assumptions concerning the specific biological interactions of the lysosome. In comparison, previous mathematical models of transport have involved varying degrees of detail in delineating distinct biophysical states. For example, classic tug-of-war models have been expressed in terms of the number of anterograde and retrograde motors that are simultaneously engaged with the cargo and the microtubule (35–37). Switching rates were defined in terms of the distribution of force among the motors. Other models have been expressed in terms of a finite number of distinct transport speeds, with switching rates that are more phenomenological (38–41). The choice of what type of model to use and what kind of inference to implement depends on the granularity of the data and what is necessary to address the scientific questions of interest. As described in detail in Results and discussion, we do not observe multimodality in the speed distribution. It follows that a finite-set-of-speeds model, such as a hidden Markov model, is inappropriate for this dataset.

Although the underlying motor dynamics, cytoskeleton, and cytosolic crowding are not known, we do find it useful to introduce a post hoc decomposition of the path segments into two descriptive "states": *motile* and *stationary*. What is important from a methods perspective is that the classification is entirely *descriptive*: when we conduct inference on changepoints and associated speeds, we do not make a simultaneous or conditioned inference on motor state. *Motile* refers to the spectrum of biophysical states that result in an average speed of more than 100 nm/s, as inferred by the changepoint algorithm. The remaining segments are labeled *stationary*. We emphasize that these segments are not stationary in the mathematical sense that the distribution of the lysosomes within the state is invariant over time. We use the term in its colloquial sense that the lysosomes are not moving very much (Fig. S3).

With these caveats in mind, we write the sequence of states as a discrete stochastic process $\{J_k\}_{k=1}^{K+1}$. For all k, J_k takes its values either as 0 (stationary) or 1 (motile). We model J as a Markov chain with fixed transition probabilities:

$$\text{Stationary} \rightarrow \text{Motile}: \quad P(J_{k+1} = 1 \mid J_k = 0) = p;$$

$$\text{Stationary} \rightarrow \text{Stationary}: \quad P(J_{k+1} = 0 \mid J_k = 0) = 1 - p;$$

$$\text{Motile} \rightarrow \text{Stationary}: \quad P(J_{k+1} = 0 \mid J_k = 1) = q;$$

$$\text{Motile} \rightarrow \text{Motile}: \quad P(J_{k+1} = 1 \mid J_k = 1) = 1 - q.$$

In an abuse of notation, when we want to emphasize the state as a function of time, we will write the state-switching process as $\{J(t)\}_{t \geq 0}$.

Consistent with the usual assumptions for continuous-time Markov chains (CTMCs), we assume that the random durations of segments are exponentially distributed with different rate parameters for the two different states (Fig. 2):

$$\text{Duration of Motile segments}: \ \tau_k \sim \text{Exp}(\mu);$$
$$\text{Duration of Stationary segments}: \ \sigma_k \sim \text{Exp}(\lambda).$$

In a typical CTMC, it is required that the system change from one state to the other when a segment ends. In this

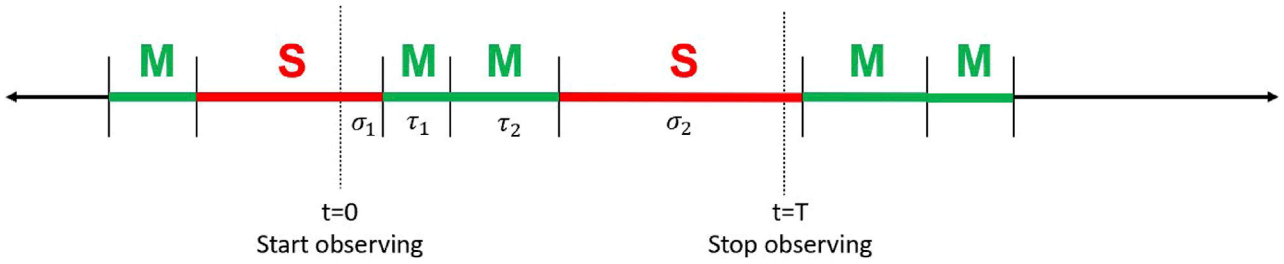


FIGURE 2 Sketch of a hypothetical sequence of states for a lysosome trajectory. Note that the beginning and ending segments are only partially observed. A mathematical model is required to estimate the average duration of stationary and motile segments, and the proportion of time spent in each state. To see this figure in color, go online.

dataset, though, it is common to switch from one kind of motile state to another. For example, molecular motors transporting a lysosome might switch from one microtubule to another. This appears to occur in the trajectory displayed in Fig. S4. This would register as a change in velocity (though the speed along the microtubules might be the same) and the changepoint algorithm will identify this as a change. We record this as a motile $\rightarrow$ motile transition. Similarly, subtle changes in direction while stationary can be inferred as a stationary $\rightarrow$ stationary transition (Fig. S5). This deviation from the usual assumptions for CTMCs will play an important role in the analysis that follows. In Fig. 2, we display a plausible sequence of states with associated state durations.

MATHEMATICAL ANALYSIS AND STATISTICAL METHODS

Velocity estimation for individual segments

Once a trajectory was segmented, we estimated each segment's velocity and starting location using orthogonal least-squares regression and maximum likelihood estimation (MLE). Suppose that we have position observations $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ at times $\{t_1, \ldots, t_n\}$ that are associated with a constant velocity state. Let (u, v) denote the velocity vector, and let (a, b) denote the initial location of the unobserved anchor. Then the likelihood associated with the model defined in Eqs. 1 and 2 is

$$\begin{aligned} &L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\,;\, u, v, a, b, \beta) \\ &\quad = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\beta}} \exp\Big(-\frac{1}{2\beta}\big[(x_i - (a + ut_i))^2 \\ &\qquad\qquad + (y_i - (b + vt_i))^2\big]\Big) \end{aligned} \tag{3}$$

We use standard regression notation,

$$\begin{aligned} \overline{x} &= \frac{1}{n}\sum_i x_i; \overline{t} = \frac{1}{n}\sum_i t_i \\ S_{xx} &= \sum_i x_i^2; \quad S_{xt} = \sum_i x_i t_i; \quad S_{tt} = \sum_i t_i^2, \end{aligned} \tag{4}$$

with analogous definitions for $\overline{y}$, S_{yy}, and S_{yt}. For a given parameter set, the residual sum of squares (RSS) is written

$$\begin{aligned} &\mathrm{RSS}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\,;\, u, v, a, b) \\ &\quad := \sum_{i=1}^{n} \big[(x_i - (a + ut_i))^2 + (y_i - (b + vt_i))^2\big]. \end{aligned} \tag{5}$$

As such, the log likelihood can be written

$$\begin{aligned} &\ell(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\,;\, u, v, a, b, \beta) \\ &\quad = -\frac{n}{2}\ln(2\pi\beta) - \frac{1}{2\beta}\mathrm{RSS}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\,;\, u, v, a, b). \end{aligned} \tag{6}$$

The resulting MLEs are

$$\begin{aligned} \widehat{u} &= \frac{S_{xt} - n\overline{x}\overline{t}}{S_{tt} - n\overline{t}^{\,2}}; \\ \widehat{a} &= \frac{\overline{x}S_{tt} - \overline{t}S_{xt}}{S_{tt} - n\overline{t}^{\,2}}; \\ \widehat{\beta} &= \frac{1}{2n}\mathrm{RSS}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\,;\, \widehat{u}, \widehat{v}, \widehat{a}, \widehat{b}), \end{aligned} \tag{7}$$

where $\widehat{v}$ and $\widehat{b}$ are defined analogously with $\widehat{u}$ and $\widehat{a}$, and we emphasize that the RSS function is evaluated using the MLEs for the other parameters. The linear approximation for a segment is then written

$$\begin{aligned} \widehat{x}_i &= \widehat{u}_i t_i + \widehat{a}; \\ \widehat{y}_i &= \widehat{v}_i t_i + \widehat{b}. \end{aligned} \tag{8}$$

The speed of the associated segment is $\widehat{s}_i := \sqrt{\widehat{u}_i^2 + \widehat{v}_i^2}$.

Piecewise-linear approximations for lysosome trajectories

As a result of the survey conducted by Jensen et al. (33), we selected a Bayesian method, bcp, which is based on work by Barry and Hartigan (42) and has been implemented in the programming language R by Erdman and Emerson (43).

We applied the algorithm to the increments of a nine-step moving average (using rollmean in R). The bcp algorithm assumes a product partition model for the data it is applied to. In this model, the path is partitioned into segments, there is a baseline mean that is inferred, and the mean value of each segment is assumed to be independent of the means in the other segments. The variance of the segment means about the baseline is inversely proportional to the length of the segment. Within each segment, the data points are assumed to be independent and identically distributed. It is assumed that the variance of the steps is constant throughout the whole trajectory. There are examples of paths that appear to have changes in variance, but variation did not appear to be large enough to affect performance (Fig. S6).

To run the algorithm, there are three hyperparameters (parameters set by the practitioner that inform the Bayesian prior distribution but are not themselves inferred; see Robert [44] for a detailed discussion) that must be specified: the assumed rate at which changes occur, γ; the probability that any given step is a changepoint (assuming an exponential distribution for the segment durations, this is $p_0 = 1 - e^{-\gamma\Delta t}$); and a weight parameter w_0, which we set to 0.01, that guides the allowable variation in inferred segment means (42).

Because it is a Bayesian method, the bcp algorithm produces Markov-chain Monte Carlo samples approximately from the posterior distribution of the changepoint vector. For each path, we produced 2,500 samples and discarded the first 1,000 as a "burn-in" set. From among the remaining 1,500, we selected one "representative" changepoint vector for the path to then use to study the speed distribution and state-switching properties. To select a path's representative changepoint vector, we computed the distribution of the number of changepoints among the path's posterior samples and identified the mode of that distribution. From among all posterior changepoint vectors that had the same number of changepoints as the mode, we took a random subsample. For each changepoint vector in the subsample, we constructed a piecewise-linear approximation of the path (as described in the previous subsection) and computed its L_2 distance from the original path. By L_2 distance, we mean for two paths observed at times $\{t_i\}_{i=1}^n$, $||f - g||_{L^2}^2 := \frac{1}{n}\sum_{i=1}^n |f(t_i) - g(t_i)|^2$. Among the changepoint vectors in the subsample, we identified the one whose piecewise-linear approximation was closest to the original path in terms of L_2 distance and chose this as the representative $\widehat{\varphi}$ of the path φ. This method effectively uses the Markov-chain Monte Carlo sampling from the posterior distribution defined by bcp to identify a representative changepoint vector that has likelihood close to the maximum under the piecewise-linear-with-Gaussian-fluctuations model described above. In Fig. S7, we provide a scatterplot that shows the relationship between speeds from one segment to the next.

Parameter estimation for state-switching model

Given a set of piecewise-linear path approximations, we used a Bayesian approach to estimate the parameters of the state-switching model $\{J_k\}_{k\in\mathbb{N}}$. To proceed with estimating the parameter set $\boldsymbol{\theta} = (\lambda, \mu, p, q)$ given a set of piecewise-linear approximations for N trajectories, $\{\widehat{\boldsymbol{\varphi}}_{\boldsymbol{n}}\}_{n=1}^N$, we articulated the likelihood function $L(\boldsymbol{\theta}\,;\,\{\widehat{\boldsymbol{\varphi}}_{\boldsymbol{n}}\})$ and the prior distributions π_λ, π_μ, π_p, and π_q of the parameters, which together form the total prior distribution $\pi(\boldsymbol{\theta}) = \pi_\lambda\pi_\mu\pi_p\pi_q$. The posterior distribution is then given by the Bayes formulation

$$\pi(\boldsymbol{\theta} \mid \{\widehat{\boldsymbol{\varphi}}_{\boldsymbol{n}}\}) \overset{c}{=} L(\boldsymbol{\theta}\,;\,\{\widehat{\boldsymbol{\varphi}}_{\boldsymbol{n}}\})\,\pi(\boldsymbol{\theta}), \tag{9}$$

where $\overset{c}{=}$ means "equal up to a constant that does not depend on $\boldsymbol{\theta}$." The likelihood function that we used is derived in the Appendix. For the priors, we chose invariant Jeffreys priors, which are improper distributions (they do not integrate to one), but they yield proper posterior distributions and the associated estimators here are unbiased (see Robert [44] for one introduction on the topic).

In the Appendix, we show that the posterior distributions for the model parameters are independent and have the form

$$\begin{aligned} \lambda|_{\widehat{\boldsymbol{\varphi}}} &\sim \text{Gamma}(n_S, \sigma), \quad & \mu|_{\widehat{\boldsymbol{\varphi}}} &\sim \text{Gamma}(n_M, \tau), \\ p|_{\widehat{\boldsymbol{\varphi}}} &\sim \text{Beta}(n_{SM}, n_{SS}), \quad & q|_{\widehat{\boldsymbol{\varphi}}} &\sim \text{Beta}(n_{MS}, n_{MM}), \end{aligned} \tag{10}$$

where σ and τ are the total times that the lysosomes spent in the stationary and motile states; n_S and n_M are the number of stationary and motile states in the collections of piecewise-linear approximations $\{\widehat{\boldsymbol{\varphi}}_{\boldsymbol{n}}\}$; and n_{SM}, n_{SS}, n_{MS}, n_{MM} are the numbers of switches of the observed types (e.g., n_{SM} is the number of stationary to motile switches). The means of these distributions give reasonable point estimates for each of the associated parameters. For the switch rates, the posterior-mean point estimates are $\widehat{\lambda} = n_S/\sigma$ (number of stationary states divided by the total time spent stationary) and $\widehat{\mu} = n_M/\tau$. For the switch probabilities, the posterior-mean point estimates for stationary to motile and motile to stationary are $\widehat{p} = n_{SM}/(n_{SM} + n_{SS})$ and $\widehat{q} = n_{MS}/(n_{MS} + n_{MM})$, respectively.

Bayesian uncertainty quantification for trajectory properties

We can use the posterior distributions for the parameters and the mathematical model to estimate (with quantified uncertainty) various properties of lysosome trajectories. In particular, we will derive an estimate for the proportion of time spent in the motile state (to show consistency with a direct estimation method described in Results and discussion) and will estimate the model prediction for the duration of motile and stationary periods. In each

case, the uncertainty quantification (UQ) method proceeds as follows:

1. We analyze the model to determine its prediction for the quantity of interest given a set of parameters (λ, μ, p, q).
2. We draw 2,000 samples from the posterior distributions of the parameters given in (10).
3. For each parameter sample, we compute the associated model prediction for the quantity of interest. We call this set of predictions the quantity's posterior UQ distribution.
4. Our point estimate for the quantity of interest is the mean of the posterior UQ distribution.
5. The 95% credible region for the quantity of interest is the range of posterior UQ distribution values that fall between the 2.5th and 97.5th percentiles.

If the posterior UQ distribution is approximately symmetric, then we report these values in the form of the mean plus or minus half the width of the credible region. In the two following subsections we describe our model's predictions for proportion of time in the motile state and the average duration of motile/stationary periods, which are used in step 1 of the UQ method.

Model prediction for asymptotic proportion of time spent in motile state

Recall that in our state-switching model, the process $\{J(t)\}_{t\geq 0}$ takes the value 1 if the lysosome is in a motile segment at time t and 0 otherwise. The long-term proportion of time in the motile state, denoted α_∞, can then be written

$$\alpha_\infty := \lim_{T\to\infty} \tfrac{1}{T}\int_0^T J(t)\mathrm{d}t. \tag{11}$$

The quantity λp serves as the effective rate of transition from stationary to motile; λ is the rate at which a stationary segment ends and is multiplied by the probability p that a switch occurs. Similarly, μq is an effective switch rate from motile to stationary. As a two-state system, the long-term probability of being in the motile state is

$$\alpha_\infty(\lambda, \mu, p, q) = \frac{\lambda p}{\lambda p + \mu q}. \tag{12}$$

A mathematically rigorous derivation of this calculation can be accomplished through use of the renewal-reward theorem (45).

Model prediction for length of motile/stationary periods

The model described above assumes that path segments have durations that are exponentially distributed in time (rate parameter λ for stationary segments and μ for motile segments). Then when a segment is complete, there is a probability transition matrix to determine what the next state will be. The transition probabilities are denoted p for stationary → motile, $1-p$ for stationary → stationary, q for motile → stationary, and $1-q$ for motile → motile. In this way, the model allows motile → motile transitions in which a lysosome either changes speeds (say, from 200 to 500 nm/s) or changes direction. In our model, each switch is considered to be an independent event.

When there are consecutive motile (respectively, stationary) segments, we call the union of these segments motile (stationary) *periods*. Suppose that the initial state is motile, and let N be the number of consecutive motile states, including the initial state. The switches can be viewed as a sequence of independent weighted coin flips in which the "success probability" is q, the probability of switching from motile to stationary and ending the motile period. As such, the expected length of a motile period can be computed using Wald's equation. Recalling that the lengths of motile segments are denoted τ_i, each with mean $1/\mu$, the expected length of a motile period is

$$E\left(\sum_{i=1}^{N}\tau_i\right) = E(N)E(\tau_1) = \frac{1}{\mu q}. \tag{13}$$

Similarly the expected length of a stationary period in the model is $1/\lambda p$.

Bootstrap method for comparing speed distributions

For a given dataset $\boldsymbol{\xi} := \{\xi_i\}_{i=1}^N$, its empirical cumulative distribution (ecdf) is the function

$$F(x\,;\,\boldsymbol{\xi}) = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\xi_i \leq x\}}. \tag{14}$$

That is, for any value x, it is the fraction of the dataset that is less than x. In comparing probability distributions based on finite samples, the ecdf is a more stable mathematical object than a histogram or kernel density estimator. The classical Kolmogorov-Smirnov test, which can be used to test the null hypothesis that a dataset is drawn from a given probability distribution, uses the ecdf of the data and defines its distance to the true distribution by taking the maximum over all absolute differences between the ecdf and the theoretical cumulative distribution. For our application, there is no theoretical distribution for the motile state lysosome speeds. Nevertheless, we would like to test whether the two distributions may have come from the same distribution. We used the following bootstrapping approach.

First, we define a distance between the ecdfs of two different samples $\boldsymbol{\xi} := \{\xi_i\}_{i=1}^N$ and $\boldsymbol{\zeta} := \{\zeta_i\}_{i=1}^M$:

$$D(\boldsymbol{\xi}, \boldsymbol{\zeta}) := \sup_{0\leq x<\infty} \left| F(x\,;\,\boldsymbol{\xi}) - F(x\,;\,\boldsymbol{\zeta}) \right|. \tag{15}$$

The question of interest is whether a given distance is statistically significantly large. In our application, the comparison of ecdfs is further complicated by the fact that our datasets have different sample sizes. To account for both sources of noise—intrinsic randomness in each dataset and extrinsic randomness due to differences in sample size—we constructed a bootstrap distribution of ecdfs for the larger of the datasets in which each bootstrap sample has the size of the smaller dataset. That is to say, suppose that the size of ξ is N and the size of ζ is $M<N$. Then we create a bootstrap sample (we denote the jth subsample $\xi^{(j)}$ from ξ by taking M independent draws from its members with replacement after each draw). We created 10,000 such bootstrap samples and for each recorded the distance

$$D_j := D\left(\xi, \xi^{(j)}\right). \tag{16}$$

To construct a 95% confidence interval for M-sized subsamples from the N-sized dataset ξ, we extracted the 95th percentile value from the bootstrapped set of D statistics, $\{D_j\}_{j=1}^{J}$. If the distance between the two observed datasets $D(\xi, \zeta)$ is greater than the 95th percentile of bootstrapped D statistics, then we say there is sufficient evidence to reject the null hypothesis that the two datasets were drawn from the same probability distribution.

RESULTS AND DISCUSSION

Path-by-path analysis of lysosomal motion shows switching behavior

Lysosomes were fluorescently labeled with LAMP1-emerald green fluorescent protein to track their intracellular motion. This fluorescently labeled protein is expressed on the outer membrane of the lysosome. Images were recorded at a frame rate of 20 Hz over 30-s durations. This frame rate allowed us to observe individual periods of motile transport, roughly linear segments in which the lysosome traveled at a near constant speed, and stationary motion, with lysosomes fluctuating in a restricted region over long segments of time. This confined motion could be due to the lysosome's being tethered to a microtubule or endoplasmic reticulum or local crowding in the cytoplasm. In Materials and methods we described our approach to partitioning individual lysosome paths into segments of constant velocity and estimating the speed within each segment (Fig. 1, as described in the Introduction, and see also Figs. S3–S6 for more examples.).

Often, in many mathematical models of molecular-motor-based transport it is assumed that there are a finite number of distinct biophysical states, each with an associated characteristic velocity (35–37,40). The present analysis did not include direction of lysosome transport, and we did not observe multimodality in the speed distribution. Rather, we observed a smoothly spread-out distribution of speeds and segment durations (Fig. S8). These plots show many short, fast-moving segments; many long, slow-moving segments; and many short, slow-moving segments. (Long, fast-moving segments are not observed in these 30-s trajectories.) In light of this observation, we adopted a phenomenological two-state labeling scheme. Lysosomes are considered *motile* if the estimated speed in a segment is 100 nm/s or greater and *stationary* otherwise. We are forced to establish a strictly positive threshold because, within each segment, there is always some displacement and the speed estimation method always returns positive values. For example, in Fig. 1, the initial segment is considered stationary, but it is clear that there is some fluctuating movement in both the positive-x and positive-y directions. This movement could be due to diffusion of the lysosome while it is not attached to a microtubule or, if the lysosome is bound, by diffusional motion of the microtubule. Similar fluctuations underlying in stationary segments can be seen in Fig. S6 as well. The particular choice we made follows from the scatterplots of segment duration versus speed (Fig. S8; very few long, slow paths had an inferred speed greater than 100 nm/s) and from a numerical experiment that is described in the Supporting material and summarized in Fig. S9. To test for robustness, we also ran our analysis using 50 and 200 nm/s as the motile/stationary threshold. For example, in Fig. S5 we display a path with a segment that has an inferred speed of 75 nm/s. It is stationary in our primary analysis, but could reasonably be considered slow and motile. As expected, changing the threshold modifies the estimated proportion of time spent motile, but the important qualitative conclusions concerning regional differences remained true (Table S1).

Lysosomes in the perinuclear region spend less time in motile segments

Fluorescence microscopy images of lysosomes in cells lead to two immediate impressions. First, lysosomes are more concentrated in the perinuclear region and dispersed throughout the periphery. Second, as noted above, intracellular motion is varied with periods of unidirectional transport and a range of speeds. In comparing transport properties as a function of intracellular region, we found that lysosomes in the perinuclear region spent less time in motile transport. We used two methods to establish this conclusion. The first was a straightforward statistical comparison in which we computed the proportion of time spent motile within each path and pooled these values to estimate the average proportion within each experimental population. We quantified the uncertainty using standard normal error. Using this method, we found that lysosomes in the perinuclear region spent 5.3% $\pm$ 1.5% of their time motile, while those in the periphery were 14.2% $\pm$ 1.8% motile (Table 1).

Although the path-by-path assessment of the proportion of time spent motile can detect significant differences among groups, it does not provide insight about which aspects of transport are most responsible for the change.

TABLE 1 Values estimated from lysosome tracking data

Groups	*n* M, S	Motile proportion (direct estimate)	Motile proportion (model estimate)	Average stationary period (model estimate)	Average motile period (model estimate)
Control cells	209, 347	**.053 ± .015**	**.059**	**72.5s**	4.59 s
Perinuclear			**(.044, .080)**	**(58.3, 90.1)**	(3.60, 5.76)
Control cells	1,268, 947	.142 ± .018	.130	36.4 s	5.46 s
Periphery			(.116, .148)	(32.9, 40.2)	(4.95, 6.04)
Sucrose-treated cells	992, 751	.139 ± .020	.136	34.9 s	5.51 s
Periphery, small lysosomes			(.118, .157	(31.3, 39.0)	(4.90, 6.17)
Sucrose-treated cells	643, 565	.121 ± .021	.116	35.6 s	4.67 s
Periphery, large lysosomes			(.099, .136)	(31.4, 40.1)	(4.09, 5.32)

The samples sizes (n) are the number of motile (M) and stationary (S) segments in each group. The total number of trajectories were equal for each group. Proportion of time spent motile ("Motile proportion") and average duration of stationary and motile periods ("Average period") are defined in Mathematical analysis and statistical methods. "Direct estimate" was computed by calculating the proportion of time spent motile on a path-by-path basis and averaging over all paths. "Model estimates" were derived using our state-switching model, which is described in Parameter estimation for state-switching model. Uncertainty for these quantities is expressed in terms of Bayesian 95% credible regions, provided in parentheses below each estimate. We consider a difference to be significant if the credible regions associated with the two samples do not overlap. Boldface values are significantly different from comparison groups. Results comparing perinuclear and peripheral transport in human lung cells (A549) (Table S3) and in sucrose-treated BS-C-1 cells (Table S5) were similar.

For this purpose it is natural to try to estimate the average duration of motile or stationary segments, but the shortness of the observation window causes a problem. Consider the path in Fig. 1, for example. We completely observe a motile segment in the middle of the trajectory, but the two stationary segments are only partially observed. For the first segment, we do not know how long it was stationary before the observation window began. Similarly, the observation window ends during the final stationary segment, but we do not know how long it continues thereafter.

To address this problem of truncated segment durations, we built a two-state Markov-chain model for the state-switching dynamics of lysosomes along with an associated Bayesian inference analysis to infer model parameters. We then use the posterior distributions of the model parameters to quantify uncertainty for different model predictions, which then allows us to make statements about statistically significant differences.

First, to demonstrate consistency with the path-by-path estimation approach, we report on the model's estimation for the proportion of time spent motile. In the perinuclear region, motile state percentage was 5.9% (95% credible region [CR]: [4.4%, 8.0%]), while in the periphery the motile percentage was found to be 13.0% (95% CR: [11.6%, 14.8%]). To understand what accounts for this difference in percentage of time spent motile, we used the model to estimate the average duration of periods of motile transport (defined in Materials and methods) versus periods of stationary transport. While the average length of motile periods were not significantly different in the two regions, the average length of stationary periods changed by a factor of two. We found that average stationary period of perinuclear-localized lysosomes was 72.5 s (95% CR: [58.3 s, 90.1 s]), while in the periphery the value is 36.4 s (95% CR: [32.9 s, 40.2 s]). Because the credible regions are not overlapping, we consider the difference to be statistically significant. Table 1 also reports these quantities for lysosomes in the periphery of the sucrose-treated cells, described below, and we find similar results as those in the periphery of the untreated control cells. In the Supporting material we provide estimates and credible regions for the parameters of our Markov-chain model for state switching (Table S2). We also include state-duration estimates (Table S3) and associated model parameter estimates (Table S4) for lysosome trajectories in human lung cells (A549). Model parameter estimations and state-duration estimates for a regional comparison in sucrose-treated BS-C-1 cells can be found in Tables S2 and S5, respectively.

The piecewise-linear approximations of the paths also provide a tool for quantifying movement while the lysosomes are motile. For each study group, we pooled all estimated segment speeds larger than 100 nm/s and conducted a bootstrap bias-corrected and accelerated comparison of samples means. We found that lysosomes in the perinuclear region had an average motile-state speed of 449 nm/s (95% CI: [407 mm/s, 498 nm/s]), while lysosomes in the periphery had an average motile-state speed of 482 nm/s (95% CI: [461 mm/s, 505 nm/s]). As we argue in a later subsection, even if this difference turned out to be significant with a greater number of replicates, such a difference in velocity is small compared with the difference in proportion of time spent motile.

Large lysosomes move more slowly during motile segments than small lysosomes

Previous work (7–10), including our own (11), examined the relationship between lysosome function, transport, and diameter. Our previous work showed that diffusion was slower for larger lysosomes, but unidirectional transport was not affected (11). This previous research did not include intracellular region. On the basis of the results described above showing that lysosomal motion is a function of intracellular region, we revisited this previous work, but now including intracellular region as a factor in the analysis.

Sucrose-mediated enlargement of lysosomes

The use of sucrose as a lysosome-specific osmotic agent to increase the diameter of lysosomes has been described previously (11,20–22). Sucrose dissolved in the cell culture medium (50 mM, 24 h) is internalized pinocytically accumulating in and enlarging the lysosomes osmotically. It does not alter the cytosol. Immunofluorescence microscopy was used to image lysosomes (Cy3-LAMP1 antibody), and the diameter was measured using ImageJ (Fig. 3). In untreated cells, the lysosomes are visible as punctate spots with a mean diameter and standard error of 763 ± 14 nm (n = 101 cells, n = 501 lysosomes). This value is just above the point spread function of the microscope (550 nm), allowing the possibility that smaller lysosomes are not resolved. Lysosomes in the sucrose-treated cells appear circular with a mean diameter and standard error of 1.07 ± 0.026 μm (n = 85 cells, n = 389 lysosomes) greater than the diffraction limit of the microscope. Previous work has shown that increased concentrations of sucrose (100 mM) lead to a greater percentage of enlarged lysosomes but no further increase of diameter (11). Lysosome diameter has previously been reported, with values ranging from 50 to 500 nm (1), as a function of cell type. For example, mouse macrophage lysosomes were reported with an average diameter of 136 nm (n = 4,850 lysosomes) and rat kidney fibroblast lysosomes with a diameter of 300–400 nm, both characterized using transmission electron microscopy (46,47). Early centrifugal sedimentation experiments using rat livers led to a calculated diameter of 400 nm (48). Our own experiments with the same BS-C-1 cells reported a median lysosome diameter of 520 nm, likely an overestimate, as smaller lysosomes were not resolved (11).

In comparison with the static immunofluorescence microscopy images (Fig. 3), live-cell imaging at 20 Hz had a decreased signal-to-noise ratio, although the size difference between untreated and sucrose-treated cells is still discernible. Untreated (n = 26) and sucrose-treated (n = 32) cells were imaged in different dishes and at different passage numbers, and 40 lysosomes/cell were randomly selected for analysis (n = 988 trajectories for untreated cells and n = 1,240 trajectories for sucrose-treated cells). For analysis of single lysosomes trajectories, the threshold value between "large" and "small" lysosomes chosen was 687 nm, which is the median value of the sucrose-treated lysosome diameters resolvable with the 20-Hz data (control median = 631 nm, sucrose-treated median = 687 nm; $p < 0.0001$, Kolmogorov-Smirnov test). Considering the dependence on intracellular region described above, we compared the motion of small and large lysosomes, all located at the periphery of sucrose-treated cells, in the main text, below. A separate analysis comparing transport in the perinuclear and peripheral regions of sucrose-treated cells yielded the same trends as untreated cells (Tables S5 and S7).

We found that large lysosomes were slower in motile transport than small lysosomes. For a quantitative comparison, we studied the speed distributions for segments with 100 nm/s or larger inferred speed for large and small lysosome populations, all located in the cell periphery. A comparison of the means yielded a statistically significant difference, though the difference was only 45 nm/s (Table 2). The distributions are not normally distributed. The mean speeds of the populations were between 400 and 500 nm/s with large sample standard deviations, between 300 and 400 nm/s. As such, the speed distributions were found to be skewed, with a mode around 200 nm/s, but large tails featuring speeds as high as 2,000 nm/s. To visualize the densities we used kernel density estimation (density in R with default bandwidth parameter) and show the results for our two comparison groups in Fig. 4. In Fig. 4 *A*, we compare regions (lysosomes in control cells in the perinuclear region

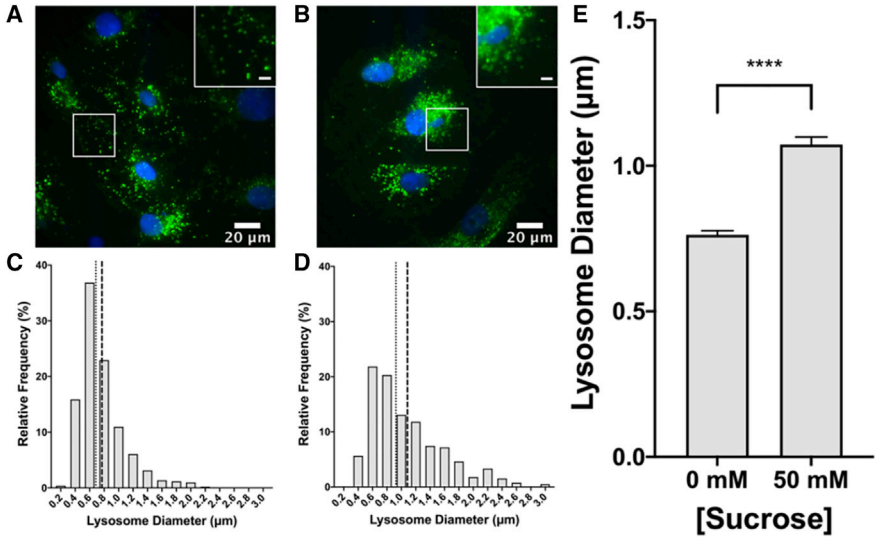


FIGURE 3 Sucrose-mediated enlargement of lysosomes. (*A*) Immunofluorescence image of untreated cells with Cy3-LAMP1 labeled lysosomes (green) and DAPI-labeled nuclei (blue). The inset has a scale bar of 5 μm. (*B*) Sucrose-treated cells (50 mM, 24 h). (*C*) Diameter of individual lysosomes in untreated cells (n = 101 cells, n = 510 lysosomes). The mean is shown with a dashed line and the median with a dotted line. The distribution of the lysosomes in the untreated cells is likely an overestimate as small lysosomes are below the diffraction limit of the microscope. (*D*) Diameter of individual lysosomes in sucrose-treated cells (n = 85 cells, n = 389 lysosomes). (*E*) Mean and standard error of lysosome diameters in untreated (0 mM) and sucrose-treated (50 mM, 24 h) cells. ****$p < 0.0001$, Kolmogorov-Smirnov test. To see this figure in color, go online.

TABLE 2 Within each of the groups, we pooled all inferred speeds from segments labeled as motile (100 nm/s or greater)

Group	Average motile speed
Control cells	449 nm/s (407, 498)
Perinuclear region	
Control cells	482 nm/s (461, 505)
Peripheral region	
Sucrose-treated cells	**470 nm/s (449, 494)**
Periphery, small lysosomes	
Sucrose-treated cells	**425 nm/s (401, 452)**
Periphery, large lysosomes	

These are the estimated average speeds of each group, with 95% bias-corrected and accelerated bootstrap confidence intervals in parentheses (49). We used a permutation test for two-sample comparisons. The p value for the comparison by region was 0.257. The comparison for large/small in the periphery of the sucrose-treated cells had a p value of 0.0112, indicating a significant difference between the groups' averages. The effect of size is small, 50 nm/s. Results comparing perinuclear and peripheral transport in human lung cells (A549) (Table S6) and comparing region in sucrose-treated BS-C-1 cells (Table S7) were similar.

[black] versus periphery [blue]), and in Fig. 4 *B*, we compare diameters (large [black] versus small [blue]) within the periphery. All density estimations exhibit the skewness described above. The means are seen as the vertical dashed lines in Fig. 4, *C* and *D*.

Although kernel density estimation provides a qualitative comparison, ecdfs provide a quantitative comparison and a notion of significant difference. To this end, for each trajectory group, we computed the ecdf F so that $F(y)$ is the proportion of segments whose associated lysosome speeds are less than y μm/s. We used the Kolmogorov-Smirnov D distance to measure the difference between two distributions and a nonparametric bootstrap method for assessing statistically significant difference (described in Materials and methods). In Fig. 4 *D*, the shaded region around the black ecdf represents a "confidence band," and the fact that the black curve is outside the confidence band indicates that there is a significant difference in the speed distributions of large and small lysosomes in the cellular periphery. The 95% one-sided bootstrap confidence interval for the null hypothesis that the black ecdf has the same generating distribution as the blue ecdf was $D<0.051$. The distance between the two data samples was $D_{\text{samples}} = 0.088$ with an implied p value of less than 0.001. As a benchmark, we ran this analysis for the comparison by region as well. As the black curve (control group, perinuclear

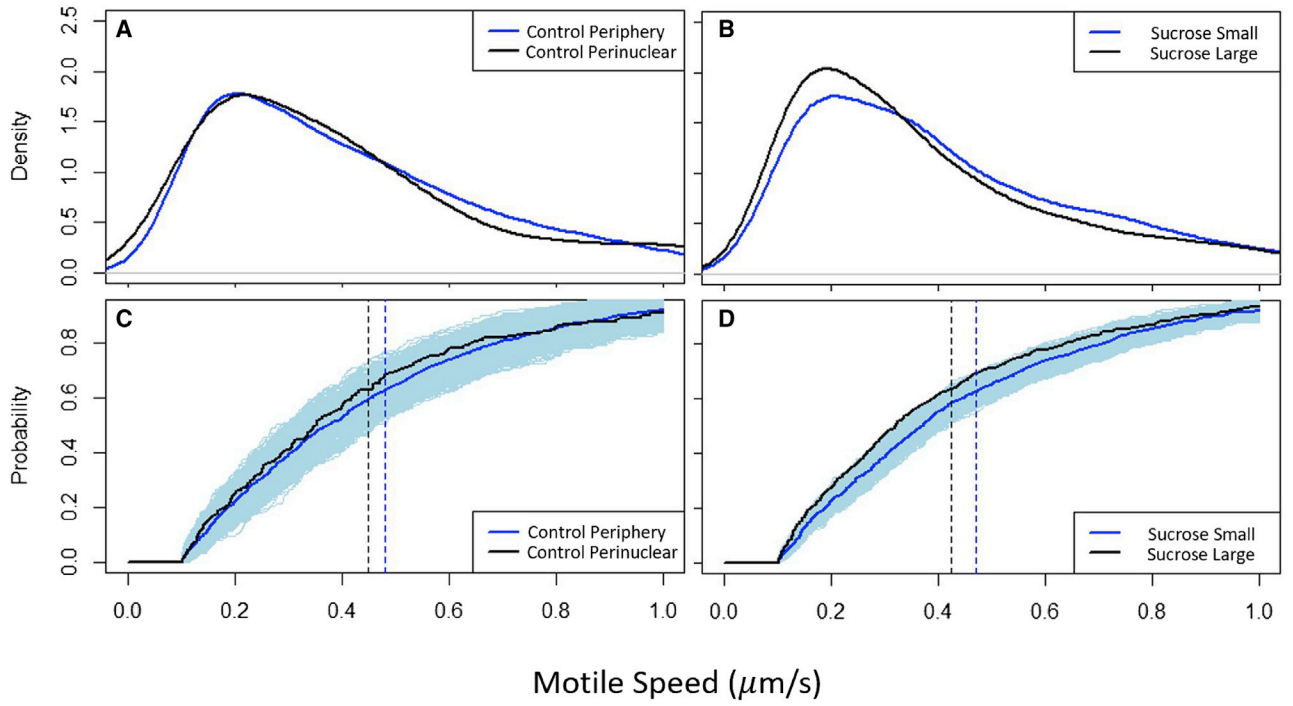


FIGURE 4 Visualizations of the differences among motile speed distributions for different group comparisons. Sample sizes for the number of motile segments are provided in Table 1 and described below. (*A* and *B*) Empirical kernel density estimations for the inferred speeds of the indicated groups. Note that all speeds included in these samples are larger than 100 nm/s and the "overhang" observed from 0–100 nm/s is an unavoidable artifact of the kernel density estimation technique. (*C* and *D*) Associated ecdfs for these populations (solid curves) and the ecdfs for bootstrapped subsamples of the larger sample (light blue curves). The subsample ecdfs create a shaded region (light blue) that shows the range of variability that arises purely because of subsampling. The means of these speed distributions are displayed as vertical dashed lines. The sucrose-treated comparison in (*B*) and (*D*) are restricted to the periphery. Sample sizes are as follows: (*A* and *C*) control cells; periphery (blue, n =1,268) and perinuclear (black, n = 209); (*B* and *D*) sucrose-treated cells; small (blue, n = 992) and large (black, n = 643) lysosomes in the periphery of sucrose-treated cells. See Table 2 and Materials and methods for further details. Box-and-whisker plots for these samples can be found in Fig. S10. Results comparing perinuclear and peripheral transport in human lung cells (A549) and in sucrose-treated BS-C-1 cells were similar (A549: Fig. S11 and Table S6; sucrose-treated BS-C-1: Fig. S12 and Table S7). To see this figure in color, go online.

region) in Fig. 4 *C* falls within the confidence band of the blue curve (control group, periphery), the difference is not significant (95% CI: $D<0.092$, $D_{\text{samples}} = 0.056$, $p = 0.49$). This result is consistent with the comparison of means reported above for the periphery and perinuclear regions (Table 2). For validation of the described bootstrap method, we also implemented a permutation test in the form described in Efron and Tibrashani (chapter 16) (49). For the small/large lysosome comparison, the permutation test one-sided 95% CI was $D<0.067$, implying a p value of 0.0033. Meanwhile, the regional comparison yielded a 95% CI of $D<0.101$, implying a p value of 0.580.

Although there is a significant difference in speed while motile, we did not find any significant differences in considering percentage of time in motile transport, average motile period, or average stationary period (Table 1). This observation suggests avenues for future biophysical studies. For example, it could suggest that although the stepping rate of motors responsible for carrying out lysosome transport is susceptible to decrease with drag force, the on-rate/off-rate properties are unaffected by this degree of difference in load.

Intracellular location, not diameter, is a key factor in lysosomal motion

We have established that intracellular region and lysosome diameter can affect different aspects of lysosome transport. In this section we show that the regional differences have a greater impact on overall transport. There are two prominent measures of overall transport that we might consider, effective velocity and effective diffusivity, and the choice for which measure is appropriate follows from the geometry of the cells and microtubules. For example, intracellular transport in axons is strongly biased in one direction: kinesin-based transport on a predominantly plus-end-out microtubule architecture results in most movement toward the distal end of the cell. (See Blum and Reed [38] and Reed et al. [39] for a classic mathematical treatment of the problem.) In our case, over the experimentally observed timescale and within the regions we observed, there was no such overall transport bias, making effective diffusivity a natural first tool for quantifying large-scale transport.

There are numerous tools for estimating effective diffusivity. The most direct and most commonly used method is to calculate an ensemble-average MSD and use regression to infer its slope. This method assumes that the paths are diffusive and the MSD is linear. Our dataset presents two challenges to this method: first, the paths are a mixture of straight-line segments, which would be parabolic in the MSD, mixed with stationary periods, which would be constant. One expects, through ergodicity, that with a large enough population and long enough paths the ensemble average would be essentially linear, permitting the use of the MSD slope. However, our paths are relatively short compared with the switch rate (only a few switches are observed in each path), and the paths have highly variable lengths. We observed that this led to biases and noisiness in MSD curves that prevent an unambiguous estimate for the slope.

There are multiple methods for estimating MSD through studying segment properties. Some of these, for example Berg's formula for the "run-and-tumble" motion of *E. coli* (50), assume that the speeds of the segments are independent of the segment durations. As reported in Fig. S8, segment speeds and durations are highly (anti)correlated in this dataset. Fortunately, the renewal-reward framework presented in Hughes et al. (51) and Ciocanel et al. (41) (Eq. 8 in the latter with effective velocity set to zero) permits a simple and intuitive formula for estimating effective diffusivity (defined in the sense of Mogilner et al. [52]) with the speed-duration relationship in mind. To this end, let

$$D_{\text{eff}} = \frac{\langle (\Delta X)^2 \rangle}{4\langle \Delta T \rangle}, \tag{17}$$

where the numerator is the ensemble-average squared displacement of the segments and the denominator is twice the average time of a segment. The results are reported for each study group in Table 3, and we note that although the MSD-slope technique was noisy, those estimates did agree in order of magnitude with values given by Eq. 18.

CONCLUSIONS

We have introduced a series of methods for studying the transport of lysosomes in live cells. After applying a standard changepoint algorithm, we studied the resulting trajectory segments to determine the impact of intracellular region and lysosome diameter on transport speeds, as well as the proportion of time lysosomes spent motile, defined as an average speed of more than 100 nm/s in segments, inferred by our segmentation method. This protocol is broadly applicable to any transport in which a particle switches between motile and stationary phases. What we present is particularly useful when the speed distribution is not multimodal. Moreover, in contrast to MSD-based methods, which rely

TABLE 3 Predicted effective diffusivity for each of the comparison groups

Group	D_{eff}
Control cells *Perinuclear region*	$0.00135\ \frac{\mu m^2}{s}$
Control cells *Peripheral region*	$0.00408\ \frac{\mu m^2}{s}$
Sucrose-treated cells *Periphery, small lysosomes*	$0.00438\ \frac{\mu m^2}{s}$
Sucrose-treated cells *Periphery, large lysosomes*	$0.00286\ \frac{\mu m^2}{s}$

The difference in proportion of time spent motile has a greater impact on transport than the effect due to size. Results comparing perinuclear and peripheral transport in human lung cells (A549) were similar (Table S6).

on averaging over all states, and often over all particles, our approach allows a more refined understanding of changes in distinct behaviors.

Using this analysis, we found that the location of the lysosome in the cell is an important factor in lysosomal transport. Although the speed distribution during motile segments is not affected, the proportion of time spent in the motile state is significantly decreased in the perinuclear region. This is due almost entirely to a 100% increase in the duration of stationary periods. This decreased transport in the perinuclear region could point toward cytoskeletal differences in the perinuclear region, a lysosomal maturation process, or the importance of the endoplasmic reticulum as a tethering site for lysosomes (5,53,54). These possibilities will be examined in future work. Our previous work (11), as well as many single-particle tracking studies, did not consider intracellular region as a factor in intracellular transport, instead focusing lysosome diameter. The more detailed mathematical analysis described above shows that the diameter of the lysosome does have some impact on speed in the motile segments of the trajectories. But by one measure, effective diffusivity, the impact on transport due to size is smaller than the impact due to region. Although our present work does not examine the underlying mechanism that leads to this regional dependence on intracellular transport, it does point toward including intracellular region as a factor in future studies of intracellular transport for lysosomes, as well as other organelles and large molecules.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2022.02.032.

APPENDIX

Derivation: posterior distributions for state-switching model parameters

To minimize bias in the Bayesian estimation of the state-switching model, we used scale-invariant Jeffreys priors (44). These have the form

$$\begin{aligned}\pi_\lambda(\lambda) &= \frac{1}{\lambda};\ \pi_\mu(\mu) = \frac{1}{\mu};\\ \pi_p(p) &= p^{-1}(1-p)^{-1};\ \pi_q(q) = q^{-1}(1-q)^{-1}.\end{aligned} \tag{18}$$

For the likelihood function, before giving the full form, we present an example. Ignoring velocities, suppose that a piecewise-linear path estimate $\widehat{\varphi}$ consists of the following sequence of (state, duration) pairs:

$$\widehat{\varphi} = \{(S,\sigma_1),(M,\tau_2),(M,\tau_3),(S,\sigma_4),(M,\tau_5)\}. \tag{19}$$

The first segment is stationary and lasts for time σ_1, which has likelihood $\lambda e^{-\lambda\sigma_1}$. It then transitions from stationary to motile, an event that has probability p, and then the likelihood of that motile state lasting for time τ_2 is $\mu e^{-\mu\tau_2}$. These considerations continue until the final state. Note that there is not a transition at the end of the path; the observation stops while the particle is still in a motile state. The likelihood that this occurs is the probability that the state is longer than the time τ_5 (in other words, the probability that there is no transition in this amount of time). This "survival" probability for the exponential is $\int_{\tau_5}^{\infty}\mu e^{-\mu t}dt = e^{-\mu\tau_5}$. The likelihood for the path given in (19) can therefore be written

$$\begin{aligned}L(\boldsymbol{\theta};\widehat{\varphi}) = &\left(\lambda e^{-\lambda\sigma_1}\right)(p)(\mu e^{-\mu\tau_2})(1-q)(\mu e^{-\mu\tau_3})\\ &\times(q)\left(\lambda e^{-\lambda\sigma_4}\right)(p)(e^{-\mu\tau_5}).\end{aligned} \tag{20}$$

When taking a population of trajectories, which are assumed to be independent, the likelihood given the whole population is a product of the likelihoods given the individual paths. With this product structure, there is a significant simplification that follows. Let $\widehat{\Phi}$ be the collection of all piecewise-linear approximations of lysosome paths; then

$$L(\boldsymbol{\theta};\widehat{\Phi}) = \lambda^{n_S}e^{-\lambda\sigma}\mu^{n_M}e^{-\mu\tau}p^{n_{SM}}(1-p)^{n_{SS}}q^{n_{MS}}(1-q)^{n_{MM}}, \tag{21}$$

where n_S and n_M are the number of stationary or motile segments that are completed during an observation period; τ and σ are the total time in the motile state and stationary state, respectively; and n_{MM}, n_{MS}, n_{SS}, n_{SM} are the numbers of observed transitions of each type.

The posterior distribution can then be written

$$\begin{aligned}\pi(\boldsymbol{\theta}\mid\widehat{\Phi}) &\overset{c}{=} \lambda^{n_S-1}e^{-\lambda\sigma}\mu^{n_M-1}e^{-\mu\tau}\\ &\times p^{n_{SM}-1}(1-p)^{n_{SS}-1}q^{n_{MS}-1}(1-q)^{n_{MM}-1}.\end{aligned} \tag{22}$$

By inspection we recognize that the posterior distributions are independent and have the gamma and beta distributions displayed in Henry et al. (10).

AUTHOR CONTRIBUTIONS

N.T.R., K.J.C., S.A.M., and C.K.P. conceived the experiments. N.T.R. conducted the experiments. N.T.R., K.J.C., and S.A.M. analyzed the results. N.T.R., K.J.C., S.A.M., and C.K.P. wrote the manuscript. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

This research was supported by the NSF-Simons Southeast Center for Mathematics and Biology (SCMB) through grant NSF-DMS1764406 and Simons Foundation-SFARI 594594. We thank Riley Juenemann (Pratt Grand Challenges REU 2019; NSF-1659615) and Trinity Campagna (University of Chicago Odyssey Scholar, 2019) for assistance.

REFERENCES

1. Alberts, B. 2008. Molecular Biology of the Cell. Garland Science.
2. Saftig, P., and J. Klumperman. 2009. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat. Rev. Mol. Cell Biol.* 10:623–635.
3. Luzio, J. P., P. R. Pryor, and N. A. Bright. 2007. Lysosomes: fusion and function. *Nat. Rev. Mol. Cell Biol.* 8:622–632.
4. Xu, H., and D. Ren. 2015. Lysosomal physiology. *Annu. Rev. Physiol.* 77:57–80.
5. Pu, J., C. M. Guardia, …, J. S. Bonifacino. 2016. Mechanisms and functions of lysosome positioning. *J. Cell Sci.* 129:4329–4339.
6. Heuser, J. 1989. Changes in lysosome shape and distribution correlated with changes in cytoplasmic pH. *J. Cell Biol.* 108:855–864.

7. Durchfort, N., S. Verhoef, …, D. M. Ward. 2012. The enlarged lysosomes in beigej cells result from decreased lysosome fission and not increased lysosome fusion. *Traffic.* 13:108–119.
8. de Araujo, M. E., G. Liebscher, …, L. A. Huber. 2020. Lysosomal size matters. *Traffic.* 21:60–75.
9. Cao, Q., Y. Yang, …, X.-P. Dong. 2017. The lysosomal Ca2+ release channel TRPML1 regulates lysosome size by activating calmodulin. *J. Biol. Chem.* 292:8424–8435.
10. Henry, A. G., S. Aghamohammadzadeh, …, W. D. Hirst. 2015. Pathogenic LRRK2 mutations, through increased kinase activity, produce enlarged lysosomes with reduced degradative capacity and increase ATP13A2 expression. *Hum. Mol. Genet.* 24:6013–6028.
11. Bandyopadhyay, D., A. Cyphersmith, …, C. K. Payne. 2014. Lysosome transport as a function of lysosome diameter. *PLoS One.* 9:e86847.
12. Cherqui, S., and P. J. Courtoy. 2017. The renal Fanconi syndrome in cystinosis: pathogenic insights and therapeutic perspectives. *Nat. Rev. Nephrol.* 13:115.
13. Lin, J., S.-S. Shi, …, L.-P. Wen. 2016. Giant cellular vacuoles induced by rare earth oxide nanoparticles are abnormally enlarged endo/lysosomes and promote mTOR-dependent TFEB nucleus translocation. *Small.* 12:5759–5768.
14. Jayaram, D. T., A. Kumar, …, C. K. Payne. 2019. TiO2 nanoparticles generate superoxide and alter gene expression in human lung cells. *RSC Adv.* 9:25039–25047.
15. Bálint, Š., I. V. Vilanova, …, M. Lakadamyali. 2013. Correlative live-cell and superresolution microscopy reveals cargo transport dynamics at microtubule intersections. *Proc. Natl. Acad. Sci. U. S. A.* 110:3375–3380.
16. Caviston, J. P., and E. L. Holzbaur. 2006. Microtubule motors at the intersection of trafficking and transport. *Trends Cell Biol.* 16:530–537.
17. Humphries IV, W. H., N. C. Fay, and C. K. Payne. 2010. Intracellular degradation of low-density lipoprotein probed with two-color fluorescence microscopy. *Integr. Biol.* 2:536–544.
18. Humphries, W. H., IV, C. J. Szymanski, and C. K. Payne. 2011. Endo-lysosomal vesicles positive for Rab7 and LAMP1 are terminal vesicles for the transport of dextran. *PLoS One.* 6:e26626.
19. Humphries, W. H., IV, and C. K. Payne. 2012. Imaging lysosomal enzyme activity in live cells using self-quenched substrates. *Anal. Biochem.* 424:178–183.
20. Cohn, Z. A., and B. A. Ehrenreich. 1969. The uptake, storage, and intracellular hydrolysis of carbohydrates by macrophages. *J. Exp. Med.* 129:201–225.
21. Ferris, A. L., J. C. Brown, …, B. Storrie. 1987. Chinese hamster ovary cell lysosomes rapidly exchange contents. *J. Cell Biol.* 105:2703–2712.
22. DeCourcy, K., and B. Storrie. 1991. Osmotic swelling of endocytic compartments induced by internalized sucrose is restricted to mature lysosomes in cultured mammalian cells. *Exp. Cell Res.* 192:52–60.
23. Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods.* 9:671–675.
24. Suh, J., M. Dawson, and J. Hanes. 2005. Real-time multiple-particle tracking: applications to drug and gene delivery. *Adv. Drug Deliv. Rev.* 57:63–78.
25. Tinevez, J.-Y., N. Perry, …, K. W. Eliceiri. 2017. TrackMate: an open and extensible platform for single-particle tracking. *Methods.* 115:80–90.
26. Hawkins, D. M. 1980. Identification of outliers. *In* Monographs on Applied Probability and Statistics. Springer, Dordrecht.
27. Monnier, N., S.-M. Guo, …, M. Bathe. 2012. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophys. J.* 103:616–626.
28. Manzo, C., and M. F. Garcia-Parajo. 2015. A review of progress in single particle tracking: from methods to biophysical insights. *Rep. Prog. Phys.* 78:124601.
29. Persson, F., M. Lindén, …, J. Elf. 2013. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat. Methods.* 10:265–269.
30. Karslake, J. D., E. D. Donarski, …, J. S. Biteen. 2020. SMAUG: analyzing single-molecule tracks with nonparametric Bayesian statistics. *Methods.* 193:16–26.
31. Shuang, B., D. Cooper, …, C. F. Landes. 2014. Fast step transition and state identification (STaSI) for discrete single-molecule data analysis. *J. Phys. Chem. Lett.* 5:3157–3161.
32. Neumann, S., R. Chassefeyre, …, S. E. Encalada. 2017. KymoAnalyzer: a software tool for the quantitative analysis of intracellular transport in neurons. *Traffic.* 18:71–88.
33. Jensen, M. A., Q. Feng, …, S. A. McKinley. 2021. A change point analysis protocol for comparing intracellular transport by different molecular motor combinations. *Math. Biosci. Eng.* 18:8962–8996.
34. Yin, S., N. Song, and H. Yang. 2018. Detection of velocity and diffusion coefficient change points in single-particle trajectories. *Biophys. J.* 115:217–229.
35. Klumpp, S., and R. Lipowsky. 2005. Cooperative cargo transport by several molecular motors. *Proc. Natl. Acad. Sci. U S A.* 102:17284–17289.
36. Müller, M. J., S. Klumpp, and R. Lipowsky. 2008. Tug-of-war as a cooperative mechanism for bidirectional cargo transport by molecular motors. *Proc. Natl. Acad. Sci. U. S. A.* 105:4609–4614.
37. Müller, M. J., S. Klumpp, and R. Lipowsky. 2010. Bidirectional transport by molecular motors: enhanced processivity and response to external forces. *Biophys. J.* 98:2610–2618.
38. Blum, J., and M. Reed. 1989. A model for slow axonal transport and its application to neurofilamentous neuropathies. *Cell Motil. Cytoskelet.* 12:53–65.
39. Reed, M. C., S. Venakides, and J. J. Blum. 1990. Approximate traveling waves in linear reaction-hyperbolic equations. *SIAM J. Appl. Math.* 50:167–180.
40. Bressloff, P. C., and J. M. Newby. 2013. Stochastic models of intracellular transport. *Rev. Mod. Phys.* 85:135–196.
41. Ciocanel, M.-V., J. Fricks, …, S. A. McKinley. 2020. Renewal reward perspective on linear switching diffusion systems in models of intracellular transport. *Bull. Math. Biol.* 82:1–36.
42. Barry, D., and J. A. Hartigan. 1993. A Bayesian analysis for change point problems. *J. Am. Stat. Assoc.* 88:309–319.
43. Erdman, C., and J. W. Emerson. 2007. bcp: an R package for performing a Bayesian analysis of change point problems. *J. Stat. Softw.* 23:1–13.
44. Robert, C. 2007. The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation. Springer Science & Business Media.
45. Serfozo, R. 2009. Basics of Applied Stochastic Processes. Springer Science & Business Media.
46. Steinman, R. M., S. E. Brodie, and Z. A. Cohn. 1976. Membrane flow during pinocytosis. A stereologic analysis. *J. Cell Biol.* 68:665–687.
47. Bakker, A. C., P. Webster, …, N. W. Andrews. 1997. Homotypic fusion between aggregated lysosomes triggered by elevated (Ca2+) i in fibroblasts. *J. Cell Sci.* 110:2227–2238.
48. Novikoff, A. B. 1961. Lysosomes and related particles. *In* The Cell. Elsevier, pp. 423–488.
49. Efron, B., and R. J. Tibshirani. 1994. An Introduction to the Bootstrap. CRC press.
50. Berg, H. C. 1984. Random Walks in Biology. Princeton University Press.
51. Hughes, J., W. O. Hancock, and J. Fricks. 2011. A matrix computational approach to kinesin neck linker extension. *J. Theor. Biol.* 269:181–194.
52. Mogilner, A., T. C. Elston, …, G. Oster. 2002. Molecular motors: theory. *In* Computational Cell Biology. Springer, pp. 320–353.
53. Cabukusta, B., and J. Neefjes. 2018. Mechanisms of lysosomal positioning and movement. *Traffic.* 19:761–769.
54. Jongsma, M. L., I. Berlin, …, J. Neefjes. 2016. An ER-associated pathway defines endosomal architecture for controlled cargo transport. *Cell.* 166:152–166.

Biophysical Journal, Volume *121*

Supplemental information

Transport of lysosomes decreases in the perinuclear region: Insights from changepoint analysis

Nathan T. Rayens, Keisha J. Cook, Scott A. McKinley, and Christine K. Payne

Supporting Material

Transport of lysosomes decreases in the perinuclear region: Insights from changepoint analysis

Nathan T. Rayens[1], Keisha J. Cook[2], Scott A. McKinley[2], Christine K. Payne[1]

[1] Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA

[2] Department of Mathematics, Tulane University, New Orleans, LA 70118, USA

Comparison of analysis using trajectories obtained from raw or despeckled images

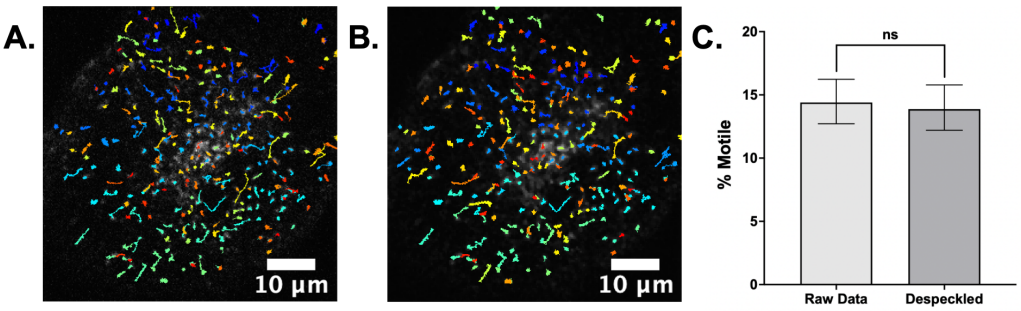


Figure S1: Comparison of analysis using trajectories obtained from raw or despeckled images. A. Fluorescence microscopy image of a BS-C-1 cell overlaid with lysosome trajectories obtained using TrackMate in the absence of a filter. B. Same cell with trajectories obtained following despeckling. C. A comparison of percent motile trajectories obtained from raw or despeckled images (n=600 lysosomes from 6 cells) shows no significant difference (ns, overlapping confidence intervals).

Figure S2 shows snapshots of the trajectory displayed in Fig. 1 and analysis of localization uncertainty. Localization uncertainty of tracking was determined by tracking (TrackMate) fluorescent polystyrene beads (200 nm, #F8764, ThermoFisher Scientific) or emGFP-labeled lysosomes (as described in Experimental Methods) with controlled motion of the microscope stage (H117P2IX, Prior Scientific, Rockland, MA) used to provide a known position. This method was adapted from D. Nong, et al., Biomedical Optics Express, 12, 3253, 2021. The fluorescent beads, adhered on a glass coverslip, provide the uncertainty inherent to the microscope and analysis method. The emGFP-labeled lysosomes in BS-C-1 cells, identical to our experimental system, report on the uncertainty due to motion in the cellular environment. In both systems, the plateaus of the displacement

curves were determined using 1.5 s intervals (30 frames), confirmed to be statistically indistinguishable from zero slope through individual linear regressions. The values of each of these intervals were scaled to center around 0 using the best fit y-intercept as an offset. These zero-adjusted plateau intervals were combined and analyzed through a linear regression. The localization uncertainty was then calculated from the standard deviation of the y-intercept of that final fit. For the immobilized fluorescent beads, the localization uncertainty was 6.3 nm. For the live cell emGFP-labeled lysosomes, the localization uncertainty was 63.5 nm.

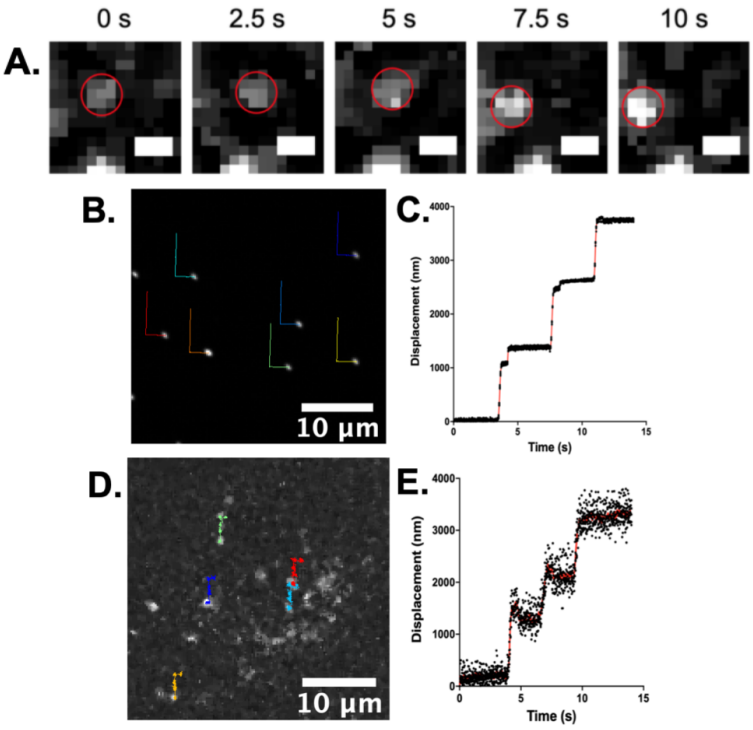
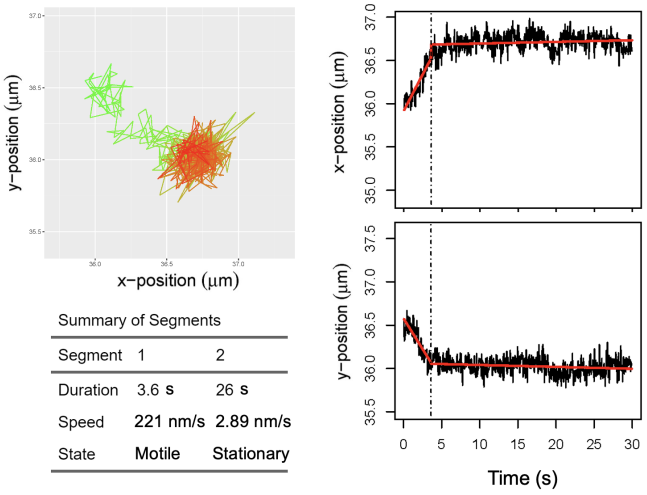


Figure S2: A. Snapshots of the trajectory shown in Fig. 1. The scale bar is 1 μm. B. To measure localization uncertainty, fluorescent beads (200 nm) were immobilized on an optical dish by evaporation. The position of the beads was controlled by the microscope stage. C. Plot of displacement from origin for the resulting trajectories ($n = 7$). The red curve is the centerline of all seven trajectory displacements and the black dots are individual subpixel localization coordinates obtained from TrackMate. D and E. An identical approach was used for emGFP-labeled lysosomes, as described in Experimental Methods ($n = 5$).

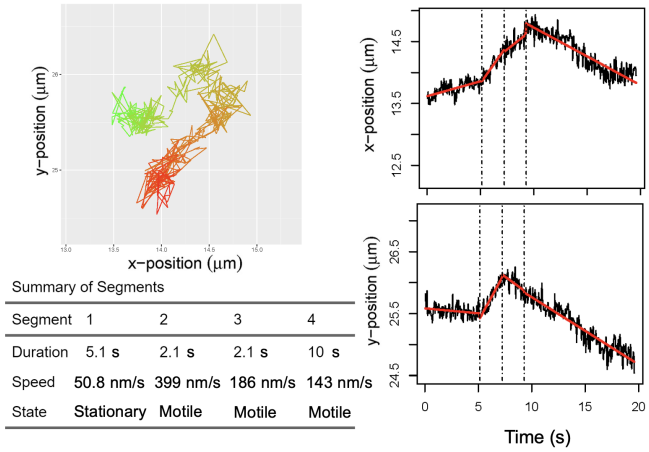
Further examples of segmented trajectories



Summary of Segments

Segment	1	2
Duration	3.6 s	26 s
Speed	221 nm/s	2.89 nm/s
State	Motile	Stationary

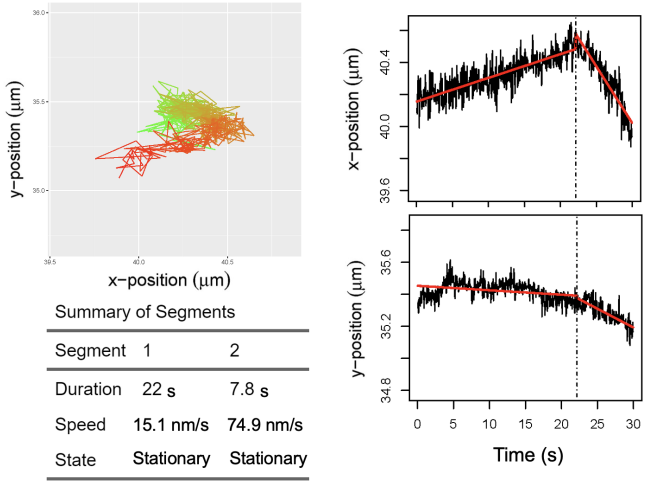
Figure S3: Trajectory from a BS-C-1 cell. Many of the trajectories feature short-lived Motile segments and long, sustained periods without much movement. In these "stationary" phases, there appear two types of fluctuations: rapid fluctuations about an "anchor" and slow "wobbles" where the anchor itself seems to move.



Summary of Segments

Segment	1	2	3	4
Duration	5.1 s	2.1 s	2.1 s	10 s
Speed	50.8 nm/s	399 nm/s	186 nm/s	143 nm/s
State	Stationary	Motile	Motile	Motile

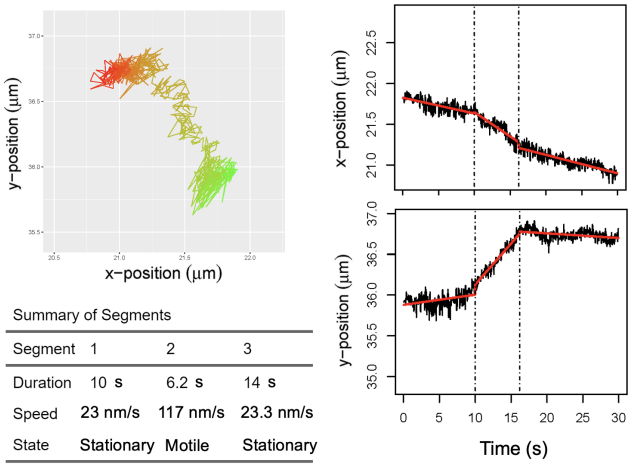
Figure S4: Trajectory from a BS-C-1 sucrose-treated cell. There can be consecutive motile segments with similar speeds (i.e. Segments 3 and 4) because of a change in direction.



Summary of Segments

Segment	1	2
Duration	22 s	7.8 s
Speed	15.1 nm/s	74.9 nm/s
State	Stationary	Stationary

Figure S5: Trajectory from a BS-C-1 cell. This trajectory features two consecutive stationary segments. This a good example of how some stationary states seem to have a slow, but well-defined, direction. The second segment is an example of a stationary segment that would qualify as motile if the threshold was reduced to 50 nm/s.



Summary of Segments

Segment	1	2	3
Duration	10 s	6.2 s	14 s
Speed	23 nm/s	117 nm/s	23.3 nm/s
State	Stationary	Motile	Stationary

Figure S6: Trajectory from a BS-C-1 sucrose-treated cell. We include this path for two reasons. First, the final stationary segment exemplifies the tendency of lysosomes to have some slight steady drift, not due to the microscope, during what we label "stationary" states. Second, the initial stationary segment appears to have a higher variability in the y-coordinate than in the x-coordinate. This appears to be some kind of tracking effect. Despite the increased variation, the algorithm identifies this as a stationary state and ignores the spurious tracker-induced feature.

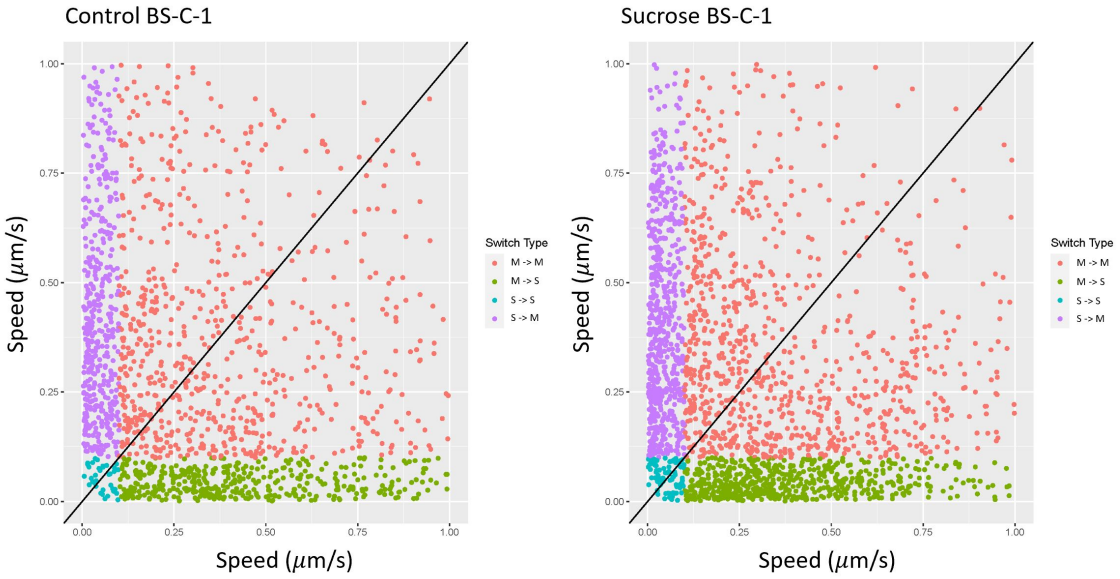


Figure S7: BS-C-1 cells. Scatter plot of the speed before and after each state switch in each trajectory. The x-axis represents the speed of the segment before a switch and the y-axis represents the speed of the segment after the switch.

Discussion concerning the choice of the Motile/Stationary speed threshold.

In order to set a threshold between motile and stationary segments, we pooled all estimated speeds from each comparison group and created a scatter plot with the segment durations (Fig. S8). While there are many short segments (10 seconds or less) with speeds of 100 nm/s and higher, there are very few long segments that attain such a speed.

To understand the distribution of inferred speeds better, we developed two null models for stationary periods. In the first, we used the model articulated in the Mathematical Models section, restated here for convenience. Let $\{(x_i, y_i)\}_{i=1}^n$ be a sequence of lysosome positions that can be observed at times $\{t_i\}_{i=1}^n$. In the first model, we assume that the anchor positions proceed deterministically with a fixed velocity (u, v), starting from position (a, b) at time 0. Assuming that the cargo fluctuates about the anchor position with magnitude β, we have for $i \in \{1, \ldots, n\}$

$$\begin{aligned} x_i &= a_i + \beta\epsilon_i^x, \\ a_i &= ut_i + a, \\ y_i &= b_i + \beta\epsilon_i^y, \\ b_i &= vt_i + b, \end{aligned} \tag{S1}$$

where $\{(\epsilon_i^x, \epsilon_i^y)\}_{i=1}^n$ is a sequence of 2d iid standard normal random variables. A "station-

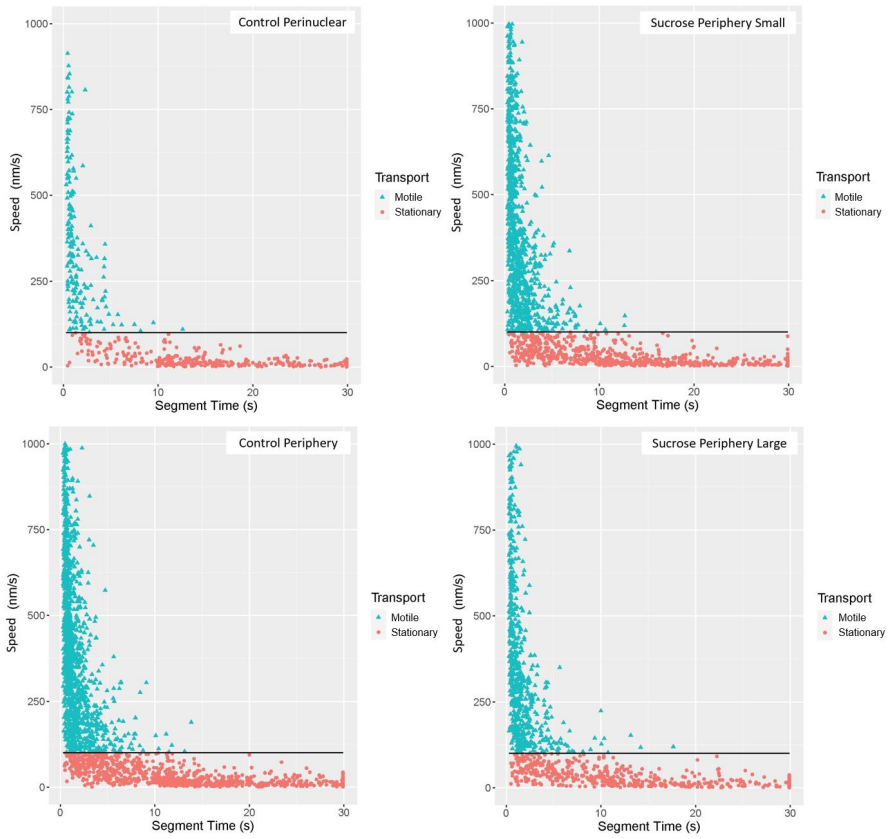


Figure S8: Visualization of the apparent inverse relationship between segment duration and speed. Each point corresponds to the duration and inferred speed during a segment of lysosomal trajectories from the indicated groups. Using a threshold of 100 nm/s, the segments were labeled Motile or Stationary and colored green or red accordingly.

ary" state would correspond the choice $(u, v) = (0, 0)$.

As we observe in the main text, there are constant fluctuations in the data, even in the paths with the least amount of movement. See, for example, the "stationary" first segment in the path displayed in Figure S5. The fluctuations are not consistent among multiple paths. There are multiple possible sources of this movement. For example, the lysosome may be detached from a microtubule and diffusing in a crowded environment. It might instead be attached to multiple motors engaged in a tug-of-war, or to a single motor that is not stepping, and the attached microtubule may be moving. In any case, we can model this as a second source of noise, which in this case appears in the anchor position. In contrast to the anchor-reverting noise term in the motor equations, this noise

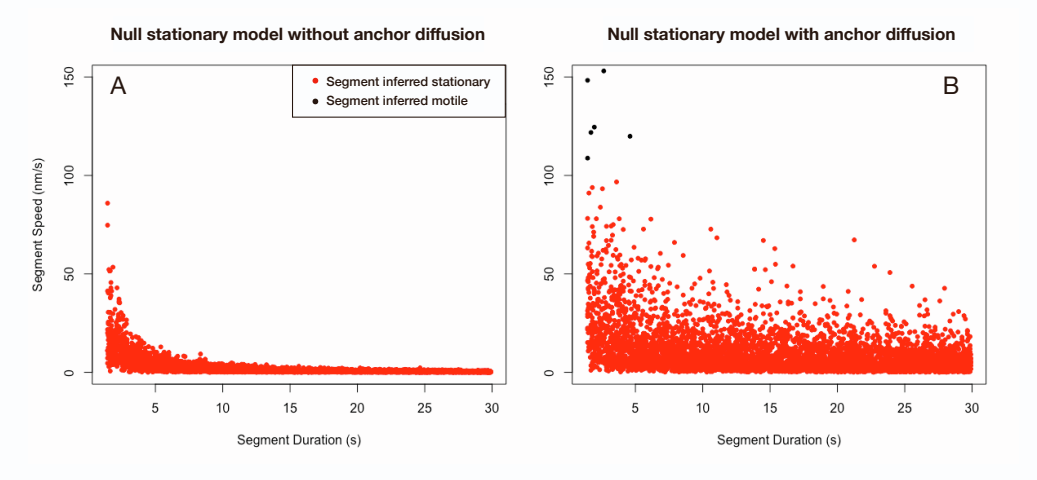


Figure S9: A visualization of the relationship between segment duration and estimated speed when there is no underlying motor stepping. A. Result of 10,000 simulations of the model displayed in Equations S1. B. Result of 10,000 simulations of model displayed in Equations S2. Very few segments are inferred to have a speed of greater than 100 nm/s.

accumulates. When $(u, v) = (0, 0)$, the anchor is then a Brownian motion. We have for $i \in \{1, \ldots, n\}$

$$\begin{aligned} x_i &= a_i + \beta\epsilon_i^x, \\ a_i &= a_{i-1} + u(t_i - t_{i-1}) + \rho\zeta_i^x, \\ y_i &= b_i + \beta\epsilon_i^y, \\ b_i &= b_{i-1} + v(t_i - t_{i-1}) + \rho\zeta_i^y, \end{aligned} \tag{S2}$$

where we employ the convention that $t_0 = 0$, $(a_0, b_0) = (a, b)$, ρ is a small positive constant, and $\{(\zeta_i^x, \zeta_i^y)\}_{i=1}^n$ is a sequence of 2d iid standard normal random variables.

We ran 10,000 simulations of each stationary null model with the velocity (u, v) set to $(0, 0)$ and with the magnitude of the fluctuations about the anchor to be $\beta = 0.1$. This value was inferred from a subset of the "most clearly stationary" paths in the lysosome data. (To be specific, we looked at the overall variance of all paths and looked studied the subset with the smallest values.) For the anchor diffusion coefficient we chose $\rho = 0$ for the first trial group and $\rho = 0.001$ for the second trial group. The segment durations were drawn uniformly at random from the interval $[0, 30]$ seconds. We display the results of the numerical experiment in Figure S9. If we use the anchor-with-diffusion model as our null model, then we can say that for any segment with estimated speed of greater than 100 nm/s, we can reject the null hypothesis with a p-value of less than 0.001.

Comparison for different Motile/Stationary thresholds

Group	Proportion of time spent motile		
	50 nm/s	**100 nm/s**	**200 nm/s**
BS-C-1 Control *periphery*	21.2%	13.0%	7.6%
BS-C-1 Control *perinuclear*	10.2%	5.9%	2.9%
BS-C-1 Sucrose *periphery*	22.5%	12.9%	7.1%
BS-C-1 Sucrose *perinuclear*	14.1%	6.4%	3.5%
A549 *periphery*	15.2%	8.5%	4.5%
A549 *perinuclear*	5.6%	3.4%	1.6%

Table S1: Comparison between different Motile/Stationary thresholds. To confirm that our threshold value of 100 nm/s was not affecting our results, we re-ran the analysis using 50 nm/s and 200 nm/s as the threshold values. As expected, the estimates for proportion of time spent motile changed, but the ratio of difference between the periphery and perinuclear regions remained largely the same.

Estimates for the parameters of the state-switching model

In the main text, we included the predictive intervals for state-switching model outcomes like proportion of time spent motile and average duration of motile and stationary periods. Those UQ intervals were constructed using the posterior distribution for the parameters μ, λ, q, and p. In Table S2 we provide point estimates and 95% credible regions for each of these parameters based on the trajectories within each displayed group.

BS-C-1 Groups	μ : **Motile Switch Rate**	λ : **Stationary Switch Rate**	q : **M** $\to$ **S Switch Prob.**	p : **S** $\to$ **M Switch Prob.**
Control cells	.564 s^{-1}	**.015** s^{-1}	.392	.941
perinuclear	(.480, .654)	(**.012**, **.018**)	(.323, .466)	(.880, .981)
Control cells	.548 s^{-1}	.030 s^{-1}	.335	.913
periphery	(.517, .582)	(.027, .033)	(.308, .362)	(.884, .938)
Sucrose-treated cells	.569 s^{-1}	.031 s^{-1}	.320	.931
periphery, small lysosomes	(.534, .606)	(.028, .034)	(.292, .351)	(.901, .957)
Sucrose-treated cells	.545 s^{-1}	.033 s^{-1}	.394	.860
periphery, large lysosomes	(.503, .592)	(.029, .037)	(.355, .436)	(.814, .899)
Sucrose-treated cells	.555 s^{-1}	**.018**s^{-1}	.402	.856
perinuclear	(.492, .619)	(**.015**, **.021**)	(.345, .463)	(.791, .912)
Sucrose-treated cells	.572 s^{-1}	.033 s^{-1}	.350	.900
periphery	(.543, .602)	(.030, .035)	(.325, .375)	(.873, .922)

Table S2: Point estimates and 95% credible regions for the parameters of the state-switching model described in the main text (BS-C-1 cells). The bolded interval marks a parameter that is significantly different than comparison groups. The top two sections are the study groups included in the main text (comparison by region in control cells and comparison by size in the periphery of sucrose-treated cells). The bottom group is a comparison by region in the sucrose-treated cells to show robustness of the by-region results.

Results from other comparison groups

To ensure that the observed differences in perinuclear and peripheral transport of lysosomes were not unique to the monkey kidney cells (BS-C-1) described in the main text, experiments and analysis were repeated with human lung cells (A549). Also, in the main text we provided a comparison of small and large lysosomes in the periphery of sucrose-treated BS-C-1 cells. Here we include results for the regional comparison (perinuclear vs. periphery) in sucrose-treated BS-C-1 cells. Results described below are in agreement with results presented in the main text.

A549 Groups	**n M, S**	**Motile prop. (Direct est.)**	**Motile prop. (Model)**	**Avg. Stationary period (Model)**	**Avg. Motile period (Model)**
A549 cells *perinuclear*	57, 202	**.028 ± .014**	**.034 (.020, .061)**	**141.7 s (96.5, 205.4)**	5.07 s (3.30, 7.79)
A549 cells *periphery*	584, 625	.092 ± .018	.085 (.069, .103)	62.2 s (53.6, 72.95)	5.7 s (4.95, 6.64)

Table S3: Regional comparison for A549 cells. The "direct estimate (Direct est.)" was computed by calculating the proportion of time spent motile on a path-by-path basis and averaging over all paths. The "model estimates (Model)" were derived using our state-switching model, which is described in the main text, "Parameter estimation for state-switching model". Uncertainty for these quantities is expressed in terms of Bayesian 95% credible regions in parentheses below each estimate. We consider a difference to be significant if the credible regions associated with the two samples do not overlap. Bolded values are significantly different from comparison groups.

A549 Groups	μ : **Motile Switch Rate**	λ : **Stationary Switch Rate**	$q : \mathbf{M} \to \mathbf{S}$ **Switch Prob.**	$p : \mathbf{S} \to \mathbf{M}$ **Switch Prob.**
A549 cells *perinuclear*	.424 s^{-1} (.311, .553)	**.009**s^{-1} (**.006**, **.012**)	.488 (.345, .631)	.821 (.674, .926)
A549 cells *periphery*	.523 s^{-1} (.479, .569)	.018 s^{-1} (.015, .020)	.337 (.297, .375)	.916 (.870, .951)

Table S4: Model parameter estimates for lysosome trajectories in human lung cells (A549).

BS-C-1 Groups	n M, S	Motile prop. (Direct est.)	Motile prop. (Model)	Avg. Stationary period (Model)	Avg. Motile period (Model)
Sucrose-treated *perinuclear*	57, 672	**.068 ± .018**	**.064 (.050, .082)**	**65.91 s (54.27, 79.62)**	4.53 s (3.76, 5.42)
Sucrose-treated *periphery*	584, 2397	.129 ± .015	.129 (.116, .143)	33.98 s (31.32, 36.94)	5.01 s (4.61, 5.46)

Table S5: Regional analysis of sucrose-treated BS-C-1 cells.

A549 Group	Average Motile Speed	Effective Diffusivity
A549 cells *perinuclear region*	399 $\frac{\text{nm}}{\text{s}}$; (332, 501)	8.81×10^{-4} $\frac{\mu\text{m}^2}{\text{s}}$
A549 cells *peripheral region*	498 $\frac{\text{nm}}{\text{s}}$; (467, 534)	3.48×10^{-3} $\frac{\mu\text{m}^2}{\text{s}}$

Table S6: Average motile speeds from A549 cells with 95% BCa bootstrap confidence intervals in parentheses. We pooled all inferred speeds from segments labeled as motile (100 nm/s or greater). These are the estimates for average speed in each of the groups with associated BCa bootstrap confidence intervals. In the last column we report the estimated effective diffusivity for each group.

BS-C-1 Group	Average Motile Speed	Effective Diffusivity
Sucrose-treated cells *perinuclear region*	459 nm/s; (424, 503)	.00196 $\frac{\mu\text{m}^2}{\text{s}}$
Sucrose-treated cells *peripheral region*	455 nm/s; (439, 474)	.00374 $\frac{\mu\text{m}^2}{\text{s}}$

Table S7: Average motile speeds of sucrose-treated BS-C-1 cells.

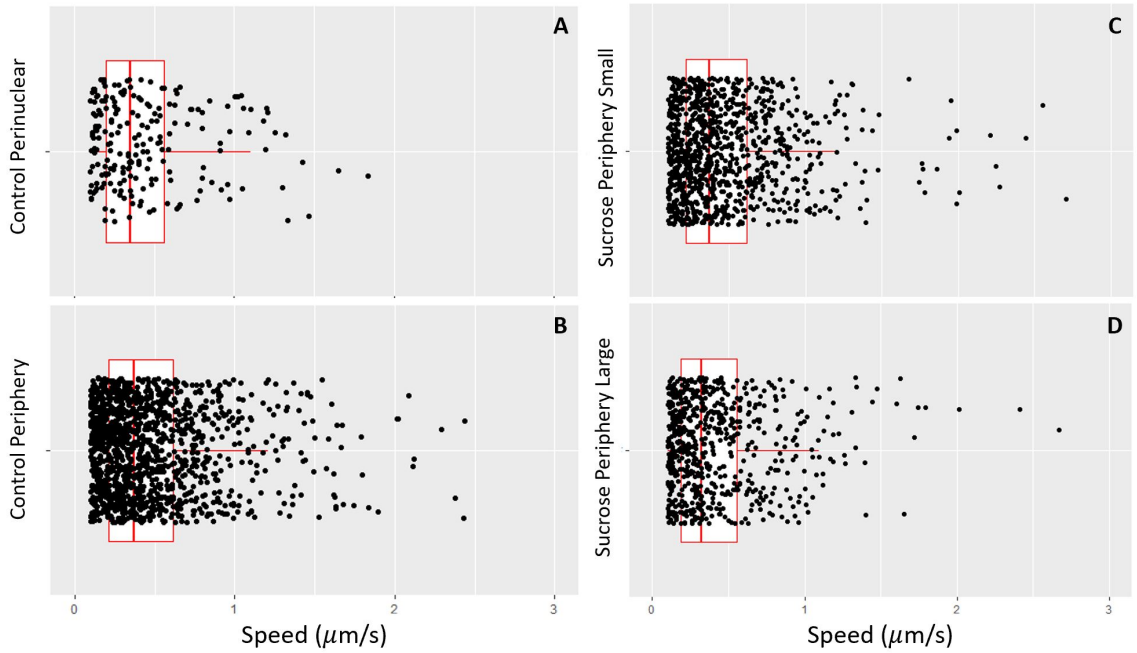


Figure S10: Scatter plots of the inferred segment speeds that are summarized in Fig. 4 in the main text. The data points are vertically positioned to display multiple data points with similar speed values. The comparison on the left (A and B) shows the value of using the ecdf to test differences. While there are fewer large values (say, greater than 1 μm/s) in the perinuclear group of the BS-C-1 control cells, the blue confidence band in Fig. 4 shows that it would not be unusual for a subsample of the periphery group to look like the perinuclear group. By contrast, the difference between the groups on the right (C and D) are significantly different.

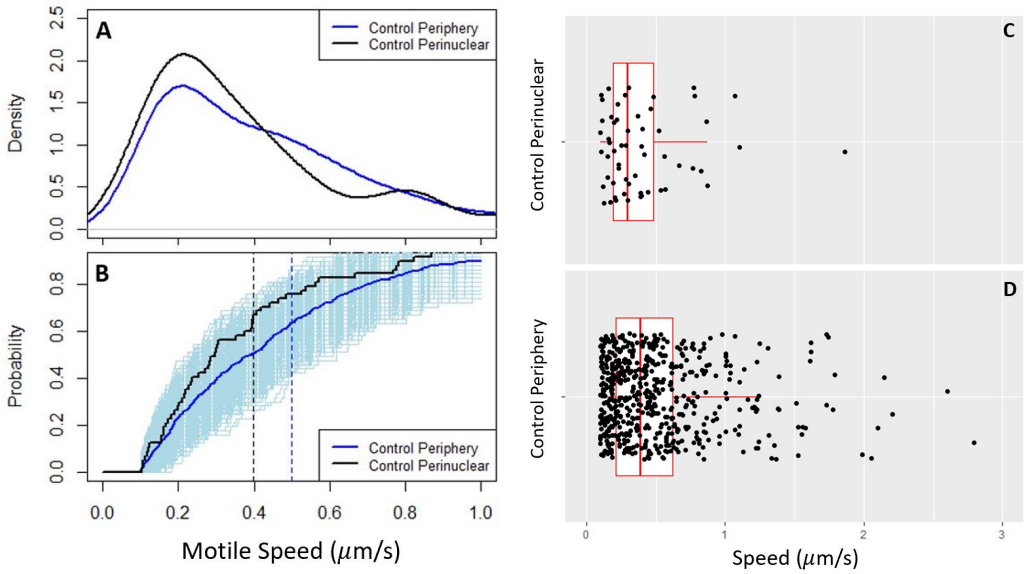


Figure S11: Visualizations of the differences among motile speed distributions for different group comparisons in A549 cells. A. Empirical kernel density estimations for the inferred speeds. B. Associated empirical cumulative distribution functions (ecdfs) for these populations (solid curves) and the ecdfs for bootstrapped subsamples of the larger sample (light blue curves). The subsample ecdfs create a shaded region (light blue) that shows the range of variability that arises purely due to subsampling. The means of these speed distributions are displayed as vertical dashed lines. C. and D. Corresponding scatter plots of the motile speeds overlaying its associated box-and-whisker plot.

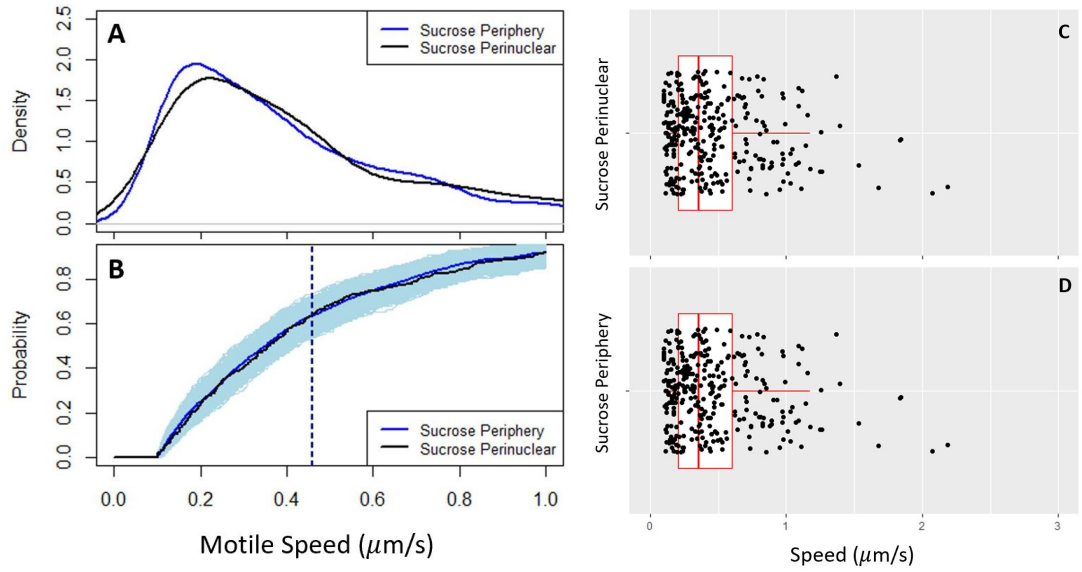


Figure S12: Visualizations of the differences among motile speed distributions for different group comparisons in sucrose-treated BS-C-1 cells. A. Empirical kernel density estimations for the inferred speeds of the indicated groups. B. Associated empirical cumulative distribution functions (ecdfs) for these populations (solid curves) and the ecdfs for bootstrapped subsamples of the larger sample (light blue curves). The subsample ecdfs create a shaded region (light blue) that shows the range of variability that arises purely due to subsampling. The means of these speed distributions are displayed as vertical dashed lines. C. and D. Corresponding scatter plots of the motile speeds overlaying its associated box-and-whisker plot.