Extreme-Value Analysis of Intracellular Cargo Transport by Motor Proteins

Takuma Naoi[1], Yuki Kagawa[1], Kimiko Nagino[1], Shinsuke Niwa[2], and Kumiko Hayashi[1,3]*

[1]Department of Applied Physics, Graduate School of Engineering, Tohoku University, Sendai, Japan

[2]Frontier Research Institute for Interdisciplinary Sciences and Graduate School of Life Science, Tohoku University, Sendai, Japan

[3]Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), Tokyo, Japan

*Correspondence to Kumiko Hayashi
kumiko.hayashi.a2@tohoku.ac.jp (K.H.)

ABSTRACT

Extreme-value analysis (EVA) deals with deviations of data from the median of probability distributions. It has been used for various purposes, such as predicting disasters and analyzing sports records. Herein, we extended the use of EVA to investigate nanoscale runners within cells. Motor proteins such as kinesin and dynein run along microtubules, which are protein-containing rails, to deliver cargo (material needed for cells). While the velocity of these runners is markedly affected by large cargos in viscous intracellular environments, the return-level EVA plots showed that the velocity near no-load conditions disclosed the physical difference between the two runners, resulting from their load (force)-velocity relationships. The concave-up force-velocity relationship of dynein hampered the independent and identically distributed conditions for the retrograde velocity dataset and the abnormality of the return-level plot of EVA. Our findings underscore the importance of EVA for assessing the physical properties of proteins *in vivo*.

Introduction

Extreme-value analysis (EVA) [1, 2] is a statistical tool that can retrieve information regarding the extreme values of observed data that deviate from the median of probability distributions. Extreme values are focused on a variety of topics in society, such as disaster prevention [3, 4], finance [5], sports [6, 7], human lifespan [8], and the recent pandemic[9]. Recently, its applications in the biological data analysis have become active [10].

'Motor protein' is a general term for proteins that move and function using energy obtained from adenosine triphosphate (ATP) hydrolysis; these are elaborate nanosized molecular machines that function in our bodies. While myosin swings its lever arm to cause muscle contraction [11, 12], kinesin and dynein walk along the microtubules to transport intracellular materials [13, 14], and F_oF_1 synthase rotates to synthesize ATP molecules [15]. The physical properties of motor proteins, such as force and velocity, have been investigated by *in-vitro* single-molecule experiments, in which the functions of motor proteins consisting of minimal complexes were analyzed in glass chambers [16, 17, 18, 19, 20, 21]. Mechanisms underlying the chemo-mechano coupling of motor proteins have been clarified by manipulating single molecules using optical tweezers [16, 17, 18, 19, 20, 21], magnetic tweezers [22, 23] and electric fields [24, 25]. Furthermore, because these manipulations render motor protein-based systems in a state of non-equilibrium, applications of non-equilibrium statistical mechanics to such systems have proceeded via *in-vitro* single-molecule studies [24, 26]. However, motor proteins function fully in the intracellular environment and are equipped with accessory proteins. Thus, the investigation of motor proteins *in vivo* is as significant as *in-vitro* single-molecule experiments; however, it is difficult to manipulate proteins in complex intracellular environments. Despite this difficulty, we found that EVA successfully provides information regarding the physical properties of motor proteins *in vivo* noninvasively without the need for any manipulation.

In the present study, we extended the use of EVA to investigate nanoscale phenomena associated with the action of motor proteins inside cells, focusing on the *in vivo* velocity of synaptic cargo transport performed by the motor proteins kinesin (UNC-104 [27, 28]) and cytoplasmic dynein [29] in the axons of motor neurons of living *Caenorhabditis elegans* (*C. elegans*), a model organism in neuroscience. Because the axons of these worms are sufficiently long, this *in vivo* system is appropriate for investigating intracellular cargo transport. Synaptic materials packed as cargo are delivered to the synaptic region of the neurons by kinesin-mediated anterograde transport, and unnecessary materials accumulated in the synaptic region are returned to the cell body by dynein-mediated retrograde transport (Fig. 1a). Since the worms' bodies are transparent and their body movement is suppressed by anesthesia, the motion of fluorescently labeled synaptic cargos in living worms can be observed by fluorescent microscopy (Fig. 1b), using a previously described method [30]. Velocities were measured using fluorescence movies. Applying EVA to the velocity data of the intracellular cargo transport, we aimed to investigate the force-velocity relationship for kinesin and dynein, because EVA was considered to show the properties of maximum velocity near no-load conditions, under which performing measurements in cells are generally difficult owing to the existence of a large viscosity load acting on the cargo.

Based on in *in-vitro* single-molecule studies using optical tweezers, the force-velocity relationship of mammalian dynein is known to be concave-up [19, 31], while that of kinesin is concave-down [16]. The difference in convexity reflects the different mechanisms underlying the walking behavior of motor proteins in the microtubules. Thus, the present study aimed to confirm whether such differences could also be measured in an intracellular environment by using EVA, in which the application of optical tweezers is difficult.

Results

Transport velocity of synaptic cargos.

The fluorescence images of green fluorescence protein (GFP)-labeled synaptic cargo transported by motor proteins were captured using a 150× objective lens and an sCMOS camera at 10 frames per second (see methods). The body movement of *C. elegans* worms was suppressed by anesthesia, and synaptic cargo transport was observed in one of several worms. For the motionless worms, kymograph analysis of the recorded images was performed using the 'multi kymograph' module in ImageJ [32] (Fig. 1b). Velocity values were calculated as the slopes of the trajectories of the fluorescently labeled cargo in the kymograph images. Typically, a cargo exhibits moving motion at a constant velocity and pauses, and rarely reverses its direction (Fig. 1c). The histograms of the measured velocities $\{v^i\}$ ($i = 1, \cdots, n$ where n=2091 for anterograde transport and n=1113 for retrograde transport) are shown in Fig. 1d. The mean velocities were 1.57±0.45(SE) μm/s and 1.78±0.72(SE) μm/s for anterograde and retrograde transport, respectively. The retrograde velocity was slightly higher than the anterograde velocity. The velocity values show a large variety, mainly because of the cargo size difference (Fig. 2), whose effect influenced to decrease transport velocity in a high-viscosity intracellular environment. Here, the probability distribution of the cargo size can be estimated from the distribution of the fluorescence intensity (FI) of the cargo (Fig. 2a) through the relation $\mathrm{FI} \propto 4\pi r^2$ (r: radius of cargo), assuming that the fluorescent proteins labeling a cargo are uniformly distributed on its surface. It is observed that cargo sizes are largely distributed, which masks the physical properties of motor proteins from the velocity data. Another load other than the viscous load from a large cargo may be a load from the opposing motor caused by the tug-of-war between the two motors, kinesin and dynein [33]; however, the fact that the directional reversal is a rare event (Fig. 1d) indicated the presence of only a few tug-of-war events, as supported by the "motor coordination model," in which adaptor proteins that connect motors with cargos deactivate opposing motors [29].

Application of EVA to assess transport velocity data.

Under the circumstances where the opposing loads were acting on the moving cargo and they generated a large variety in transport velocity, EVA discloses the differences in velocity between the motor types because it can evaluate velocity values under a condition similar to the no-load condition without cargos. Cargo size differences and the existence of opposing motors are the demerits of physical measurement in complex intracellular environments, unlike the case for a controlled situation in single-molecule experiments; however, in the present study, the unique difficulties observed *in vivo* were overcome through the use of EVA.

Although we observed 562 worms in total, the velocity values of moving cargo during constant-velocity segments were collected from 232 worms; this reveals that transport velocities were not observed in most of the worms because of body movement and obscurity of fluorescence movies. Approximately 5–20 velocity values were observed for each worm, from which the largest value ($v_{\max}^{i}$) was selected. Using $\{v_{\max}^{i}\}$ $(i = 1, \cdots, N$ where N=228 for anterograde transport and N=217 for retrograde transport), the return-level plot, which is the typical analysis of EVA, defined as $\left(-1/\log\left(i/(N+1)\right), \hat{v}_{\max}^{i}\right)$, was investigated (Figs. 3a, b), where $\{\hat{v}_{\max}^{i}\}$ is the rearranged data of $\{v_{\max}^{i}\}$, such that $\hat{v}_{\max}^{1} \leq \hat{v}_{\max}^{2} \leq \cdots \leq \hat{v}_{\max}^{N}$. The two axes of the return-level plot represent the return period and return level. The return-level plot indicates that the measured value (return level) occurred for the first time in the time period (return period). The strong point of the plot is that values close to the rarely occurring extreme value can be enhanced owing to the logarithmic expression of the *x*-axis. Because the return-level plot for anterograde transport shows a typical convergent behavior (Fig. 3a), a property specific only to a Weibull distribution, $V_{\max}$ was proved to exist and was estimated to be 3.95±0.4 μm/s using the following equation:

$$V_{\max} = \mu - \sigma/\xi, \qquad (1)$$

where ξ, μ, and σ are the parameters of a Weibull distribution

$$W(v_{\max}) = exp\left[-\left\{1 + \xi\left(\frac{v_{\max} - \mu}{\sigma}\right)\right\}^{-1/\xi}\right]. \qquad (2)$$

The error of $V_{\max}$ consisted of the fitting errors of ξ, μ, and σ (see Supplementary Table 1 for the calculation), and $W(v_{\max})$ obtained from the experimental data was fitted using the 'ismev' and 'evd' packages in R [34]. Note that the return-level plot within the black dotted lines (Fig. 3a), representing the reliable section, indicates the validity of the analysis. On the other hand, the return-level plot of retrograde transport

(Fig. 3b) shows a peculiar behavior, with a two-step convergence. This return level plot, showing the peculiar behavior, does not belong to any of the Weibull, Fréchet, and Gumbel types. This abnormality suggests that the assumption imposed during the use of EVA, i.e., that the data were independent and identically distributed (i.i.d.), seemed to be violated. Indeed, when the retrograde data were separated into two datasets with regard to the change point ($i = N_\mathrm{c}$), as shown by the black arrow in Fig. 3b, both $\{\hat{v}^i_\mathrm{max}\}$ ($i = 1, \cdots, N_\mathrm{c}$) and $\{\hat{v}^i_\mathrm{max}\}$ ($i = N_\mathrm{c} + 1, \cdots, N$) showed the typical Weibull-type behavior (Fig. 3c, d). Because each separated data is considered to meet the i.i.d. condition, the retrograde V_max was estimated to be 6.53±0.99 μm/s, based on Eq. (2).

Figures 3e and f show the probability distributions $w(v_\mathrm{max})$ ($= W(v_\mathrm{max})/\mathrm{d}v_\mathrm{max}$) for anterograde and retrograde transport. In the case of retrograde transport, $w(v_\mathrm{max})$ was separated at the change point ($i = N_\mathrm{c}$) by two functions: $w_1(v_\mathrm{max})$ and $w_2(v_\mathrm{max})$. Note that the block sizes (representing the number of worms) from which v^i_max was selected (Supplementary Fig. S1) and bootstrapping analysis of the data $\{v^i_\mathrm{max}\}$ (Supplementary Fig. S2) were investigated to validate the aforementioned results.

Construction of a simulation model using force–velocity relationship.

We considered the abnormal behavior of the return-level plot for retrograde transport (Fig. 3b), showing the violation of the i.i.d. condition of the velocity data, from the viewpoint of the force-velocity relationship of motor proteins. In other words, we would like to find the features of the force-velocity relationship of motor proteins that lead to the violation of the i.i.d. condition.

According to the results of previous studies using single-molecule experiments, two regimes exist in the force-velocity curves of motor proteins: the load-sensitive and load-insensitive regimes [35] (red lines in Figs. 4a, b). In the first case, the velocity changes rapidly with an increase in the load (F), whereas in the other case, the velocity changes only slightly with an increase in the load. *In vitro* single-molecule experiments revealed that the force-velocity curve of kinesin was concave-down (Fig. 4a) [16], whereas that of dynein (except yeast dynein [17, 18]) was concave-up (Fig. 4b) [19, 20, 31]. This mechanical difference in the force-velocity relationship can be explained as follows: kinesin keeps moving at a distance of 8 nm along a microtubule (the interval of the microtubule structural unit) per hydrolysis of single ATP molecule even when a low load is applied, which makes its force-velocity relationship load-insensitive, resulting in a concave-down force-velocity curve. However, dynein, which can assume variable step sizes of 8–40 nm under no-load conditions [17, 19, 21, 31], slows down rapidly by decreasing the step size even when a low load is applied, resulting in a rapid velocity decrease and a concave-up force-velocity curve.

The simplest model of the force-velocity curve can be characterized by the changing point $(F_\mathrm{c}, v_\mathrm{c})$ (Figs. 4a, b) between the load-sensitive and load-insensitive regimes; in the following sections, we aim to find the most probable values (F_c^* and v_c^*) that reproduce the return-level plots shown in Figs. 3a and b by performing numerical simulations. Note that in the following simulation, the axes of the force-velocity relationship are normalized as $(F/F_\mathrm{s},\ v/v(F=0))$, where F_s is the stall force of a motor protein [16, 17, 18, 19, 20, 21], which is the maximum force generated by the motor against an opposing load. Since the absolute values of *in vivo* force could not be measured in living worms, F_s is not yet known.

A simulated value of the transport velocity (v_sim) is generated stochastically using the force-velocity models (Fig. 4). A slope α of the black straight line $v = \alpha F$ (representing Stokes' law) is generated based on the equation $\alpha = c/r$, where c is a constant and $1/r$ is stochastically generated based on the gamma distribution of $1/\sqrt{\mathrm{FI}}$ (Fig. 2b). Note that the stochasticity of the slope α of the line $v = \alpha F$ represents the variety of cargo sizes according to Stokes' law. Here, the proportional constant c was determined such that the distribution of the simulated velocity (v_sim) matches that of the experimentally measured velocity (v) (Fig. 1d). The grey lines in Fig. 4 represent the 95% confidence intervals of α; for a given value of α, a simulated velocity value v_sim can be obtained as the intersection between the black line and force velocity curve (Fig. 4).

$v^i_{\mathrm{sim,max}}$ was chosen from among the 10 values of v_{sim}. This procedure was repeated 200 times (*i.e.*, $i = 1, \cdots, N$ where N=200).

Outline of force-velocity relationships decided by the comparison between the results obtained from the experiment and simulation.

The simulated probability distributions, w_1^{sim} and w_2^{sim}, were calculated from dataset $\{v^i_{\mathrm{sim,max}}\}$ (note that w_2^{sim} did not exist for anterograde transport). Here, the extreme value dataset $\{\hat{v}^i_{\mathrm{sim,max}}\}$ is the rearranged data of $\{v^i_{\mathrm{sim,max}}\}$, such that $\hat{v}^1_{\mathrm{sim,max}} \le \hat{v}^2_{\mathrm{sim,max}} \le \cdots \le \hat{v}^N_{\mathrm{sim,max}}$. Comparing the extreme value distributions w_1^{sim} and w_2^{sim} with the experimentally obtained distributions $w_1$and w_2 (Figs. 3e, f), the error function $E(F_{\mathrm{c}}, v_{\mathrm{c}})$, is defined as follows:

$$E(F_{\mathrm{c}}, v_{\mathrm{c}}) = \frac{1}{N_{\mathrm{c}}}\sum_{i=1}^{N_{\mathrm{c}}} \left\{w_1\left(\hat{v}^i_{\mathrm{sim,max}}\right) - w_1^{\mathrm{sim}}\left(\hat{v}^i_{\mathrm{sim,max}}\right)\right\}^2 + \frac{1}{N - N_{\mathrm{c}}}\sum_{i=N_{\mathrm{c}}+1}^{N} \left\{w_2\left(\hat{v}^i_{\mathrm{sim,max}}\right) - w_2^{\mathrm{sim}}\left(\hat{v}^i_{\mathrm{sim,max}}\right)\right\}^2, \quad (3)$$

where w_2^{sim} does not exist and $N_{\mathrm{c}} = 1$ for anterograde transport. Note that F_{c} and v_{c} are the model parameters and that the datasets $\{\hat{v}^i_{\mathrm{sim,max}}\}$ were generated numerically and repeatedly ten times for each $(F_{\mathrm{c}}, v_{\mathrm{c}})$. Then, $E(F_{\mathrm{c}}, v_{\mathrm{c}})$ was plotted as the mean of the trials (Figs. 5a, b). $E(F_{\mathrm{c}}, v_{\mathrm{c}})$ was calculated using the steepest descent method (‘SciPy’ package in Python). In the anterograde transport, the red region, which indicates the set of $(F_{\mathrm{c}}, v_{\mathrm{c}})$ showing small values of $E(F_{\mathrm{c}}, v_{\mathrm{c}})$, exists above the diagonal line (Fig. 5a); this region exists below the diagonal line in the case of retrograde transport (Fig. 5b). These results indicate that the force-velocity relationship for the anterograde transport was concave-down, and that for the retrograde transport was concave-up.

The most probable value of F_{c}^* (or v_{c}^*) was defined as the mean of $\{F_{\mathrm{c}}\}$ (or $\{v_{\mathrm{c}}\}$) belonging to the top 10% minimum values of $E(F_{\mathrm{c}}, v_{\mathrm{c}})$. $(F_{\mathrm{c}}^*, v_{\mathrm{c}}^*)$=$(0.61 \pm 0.20(\mathrm{SE}), 0.96 \pm 0.05(\mathrm{SE}))$ for anterograde transport and by (F_c^*, v_c^*)=$(0.63 \pm 0.1(\mathrm{SE}), 0.25 \pm 0.0(\mathrm{SE}))$ for retrograde transport. With (F_c^*, v_c^*), the return-level plots of $\{\hat{v}_{\mathrm{sim,max}}^i\}$ for both anterograde and retrograde transports were reproduced using numerical simulations (Figs. 5c, d). Thus, the outline of *in vivo* force-velocity relationships was estimated using the experimental results of the EVA.

The part of the velocity data for the retrograde transport in Fig. 5d (violet symbols) corresponds to the velocity values in the load-sensitive regime in Fig. 5f. The steep velocity decreases under the low-load condition for the retrograde transport, causing a major variation in the larger values of $v_{\mathrm{sim,max}}^i$. In fact, the velocity data obtained by the anterograde force-velocity model also hampered the i.i.d. condition because small amounts of data were generated from the load-insensitive regime of the force-velocity relationship, whereas the majority of the data were generated from the load-sensitive regime (Fig. 5e). However, the velocity values in the load-insensitive regime resulted in a minor variation in the large values of $v_{\mathrm{sim,max}}^i$, and the return-level plot showed Weibull-type convergence for the anterograde transport.

Effects of cooperative transport by multiple motors.

It has been suggested that a single cargo can be transported by multiple motors. The effects of multiple motor transports on the results of the present study were investigated (Fig. 6). Previously, we used a non-invasive force measurement technique [36, 37, 38] developed by our research group for examining the neurons of *C. elegans,* and estimated that the number of motors carrying synaptic cargo was 1–3 [30]. The frequency $P(N_{\mathrm{m}})$ of the number of motors ($N_{\mathrm{m}} = 1,2,3$) carrying cargo was approximately $P(1){:}P(2){:}P(3) = 1{:}2{:}1$, based on previous observations [30]. After choosing N_{m} according to $P(N_{\mathrm{m}})$, $v(F/N_{\mathrm{m}})$ was used instead of $v(F)$ to determine v_{sim} from the intersection between $v(F/N_{\mathrm{m}})$ and the line $v = \alpha F$ (Fig. 6a, b). The EVA was applied to $\{v_{\mathrm{sim}}\}$ using the same procedure to obtain the results depicted in Fig. 5. By calculating $E(F_{\mathrm{c}}, v_{\mathrm{c}})$ (Eq. 3), we found that multiple motor transports did not affect the outlines of the force-velocity relationship for anterograde/retrograde transport.

Chemo-mechanical coupling models of the force-velocity relationship.

Finally, we referred to the force-velocity relationship of kinesin and dynein, theoretically derived based on the mechanisms underlying ATP hydrolysis by motor proteins. Force-velocity relationships are derived from the one-state [19] and three-state models [39] of motor proteins, as well as an energy-landscape model [16]. In particular, the force-velocity relationship for the three-state model for ATP hydrolysis is represented as follows:

$$v_{\mathrm{three}}(F) = (k_{01} - k_{02})l,$$

$$\frac{1}{k_{01}} = \frac{1}{\kappa_1} + \frac{1}{\lambda_1} e^{d_1 F / k_{\mathrm{B}} T},$$

$$\frac{1}{k_{02}} = \frac{1}{\lambda_2} e^{-d_2 F / k_{\mathrm{B}} T}, \qquad (4)$$

(see a reference[39] for the definitions of parameters for both anterograde and retrograde transport.) The simple force-velocity relationships $v(F)$ depicted in Fig. 4a and b are replaced with $v_{\mathrm{three}}(F)$ (Fig. 7a for kinesin and Fig. 7c for dynein). The difference in the model parameters (Eq. (4)) resulted in the different convexities of the force-velocity relationship in the case of the three-state model); we numerically obtained the return-level plots for the model by using a simulation method similar to that used to obtain the results in Fig. 5. In Fig. 7a and b, the red points represent the $\{v_{\mathrm{sim,max}}^{i}\}$ values obtained from the simulation (Fig. 7a, b). We found that the tendencies (Fig. 7b and d) shown by the three-state model were similar to those observed using the simple force-velocity relationships (Fig. 5), i.e., the i.i.d. condition was easily hampered in the case of the concave-up force-velocity relationship used for retrograde transport (Fig. 7c). The results for the one-state and energy-landscape models of motor proteins are summarized in the supplementary material (Supplementary Fig. S3, S4).

Discussion

In the present study, to the best of our knowledge, we applied EVA, for the first time, to assess cargo transport by the motor proteins kinesin and dynein in the neurons of living worms, as observed by high-resolution fluorescence microscopy. We investigated the velocities of the transport and found that the return-level plots of the extreme values of velocity revealed the differences between the motor protein types. Specifically, the experimental velocity data of dynein-mediated retrograde transport hampered the independent and identically distributed conditions, and the return-level plots became abnormal. Indeed, when the retrograde data were separated into two datasets, they showed the typical Weibull-type behavior. Using the simulation, the abnormality of the return-level plot that appeared only for the retrograde velocity data was attributed to the fact that the force-velocity relationship for the retrograde transport was concave-up, while that for its anterograde counterpart was concave-down. The steep velocity decrease in the low-load condition for the retrograde transport caused a major variation in the larger values of velocity, and this behavior tends to violate the i.i.d. conditions for $\{v_{\max}^{i}\}$ and causes an abnormality in the return-level plot. Although the mean values of the velocity data did not clearly show the difference between the two motor proteins, EVA was able to reveal the extreme values of the velocity data. While the mean values of the velocity were affected by loads resulting from various cargo sizes and the differences in the velocities of the two motors diminished, the differences in the extreme values of the velocity showed a larger difference; this highlights the property of the velocities under the no-load condition. We successfully obtained information regarding the force-velocity relationships of intracellular cargo transport.

Recent results of *in vitro* single-molecule experiments have suggested a concave-up force-velocity relationship for ciliary [20] and mammalian dynein [19, 31], while yeast dynein exhibits a concave-down (kinesin-like) force-velocity relationship [17, 18]. In the present study, we found a concave-up force-velocity relationship for cytoplasmic dynein in *C. elegans*. For the investigation of mammalian dynein, EVA was also applied to examine the data of synaptic cargo transport in mouse hippocampal neurons; this was originally reported in a previous study [36]. A two-step convergence in the return-level plot was also observed for retrograde transport (Supplementary Fig. S5), which corresponds to the concave-up force-velocity relationship reported in previous studies [19, 31]. Interestingly, several dynein motors showed a concave-up force velocity curve. The biological significance of collective cargo transport by multiple motor proteins is explained below, and was first introduced in a previous study [31]. When multiple motors work together, the leading dynein decreases its velocity rapidly in the presence of a low load, so that the trailing dynein can catch up. This allows the trailing dynein to share the load with the leading dynein, thereby preventing the detachment of the leading dynein from the microtubules. In other words, the rapid decrease of velocity in the load-sensitive region results in the self-correction of the position of dynein molecules, allowing them to move as a loosely bunched group [31]. However, leading kinesin does not slow down, with regard to a concave-down force-velocity relationship, in the presence of a low load. As a result, trailing kinesin cannot catch up with the leading kinesin, causing it to easily detach from the microtubules [31].

Although the outlines of the *in vivo* force-velocity relationships could be estimated using the EVA analysis, the stall force values, regarding the maximum forces of the motors, could not be estimated from this analysis. Many *in-vitro* single-molecule studies have provided the stall force values of kinesin and dynein using optical tweezers [16, 17, 18, 19, 20, 21]. Another progress in the field of *in-vitro* single-molecule study is the stall force measurement of myosin VI using DNA origami nanospring [40]; however, *in-vivo* force measurement needs to be developed.

Interpretation of return-level plots based on the force-velocity relationship is a promising tool for future research regarding neuronal diseases, particularly, KIF1A-associated neurological disorders [41, 42]. KIF1A is a type of kinesin-transporting synaptic vesicle precursor cargo, and the force and velocity of pathogenic mutant KIF1A was reported to be impaired [41, 42]. Because *in vivo* force measurement is difficult, the estimation of physical properties using EVA can be helpful for understanding the *in vivo* behavior of motor proteins. Thus, we believe that the findings of the present study represent a step forward in broadening the scope of EVA applications.

Methods

Sample preparation.

In our study, we used *C. elegans* stains *wyIs251*[Pmig-13::gfp::rab-3; Podr-1::gfp]; *wyIs251* has been previously described [43, 44].

Culture. *C. elegans* was maintained on OP50 feeder bacteria on nematode agar plates (NGM) agar plates, as per the standard protocol [43, 44]. The strains were maintained at 20ºC. All animal experiments complied with the protocols approved by the Institutional Animal Care and Use Committee of Tohoku University (2018EngLMO-008-01, 2018EngLMO-008-02).

Fluorescence microscopy observations.

A cover glass (32 mm × 24 mm, Matsunami Glass Ind., Ltd., Tokyo, Japan) was coated with 10% agar (Wako, Osaka, Japan). A volume of 20 μL of 25 mM levamisole mixed with 5 μL 200-nm-sized polystyrene beads (Polysciences Inc., Warrington, PA, USA) was dropped onto the cover glass. The polystyrene beads increased the friction and inhibited the movement of worms; levamisole paralyzed the worms. Ten to twenty worms were transferred from the culture dish to the medium on the cover glass. A second cover glass was placed over the first cover glass forming a chamber, thereby confining the worms. The worms in the chamber were observed under a fluorescence microscope (IX83, Olympus, Tokyo, Japan) at room temperature. Images of a GFP (green fluorescence protein)-labelled synaptic cargos in the DA9 motor neuron were obtained using a 150× objective lens (UApoN 150x/1.45, Olympus) and an sCMOS camera (OLCA-Flash4.0 V2, Hamamatsu Photonics, Hamamatsu, Japan) at 10 frames per second.

Cargo size distribution

The fluorescence images of various synaptic cargos are shown in Fig. 2. The difference in fluorescence intensity (FI) represents the difference in the sizes of cargos. This is because fluorescence proteins are uniformly distributed on the surface of a cargo and FI $\propto 4\pi r^2$ (r: radius of a cargo). Anterograde and retrograde transports were not distinguished in the analysis of FI. The distribution of $x = 1/\sqrt{\mathrm{FI}}$ is well fitted by a Gamma distribution

$$f(x; k, \lambda) = \frac{\lambda^k}{\Gamma(k, \lambda)} x^{k-1} e^{-\lambda x}$$

where Γ is a Gamma function, $k = 6.54 \pm 0.78$ and $\lambda = 201 \pm 25$ (the errors are fitting errors).

Error of V_{max}

The mean ($\overline{V_{\mathrm{max}}}$) and error ($\delta V_{\mathrm{max}}$) of V_{max} were estimated from the fitting parameters $\mu\,(= \bar{\mu} \pm \delta\mu)$,

σ ($= \bar{\sigma} \pm \delta\sigma$) and ξ ($= \bar{\xi} \pm \delta\xi$) of the Weibull distributions defined by Eq. 1 in the main text as follows:

$$V_{\max} = \left(\bar{\mu} - \frac{\bar{\sigma}}{\bar{\xi}}\right) \pm \sqrt{(\delta\mu)^2 + \left(\frac{\delta\sigma}{\bar{\xi}}\right)^2 + \left(\frac{\bar{\sigma}\delta\xi}{\bar{\xi}^2}\right)^2} \quad . \tag{S1}$$

The values of μ ($= \bar{\mu} \pm \delta\mu$), σ ($= \bar{\sigma} \pm \delta\sigma$) and ξ ($= \bar{\xi} \pm \delta\xi$) for anterograde and retrograde transport are summarized in Supplementary Table S1.

Block size

We investigated the dependence of the fitting parameters (μ, σ and ξ) on the number (s) of worms (block size), from which $v_{\max}^i$ was chosen, as shown in Supplementary Fig. S1. All values converge as s increases.

Bootstrapping method

A part (r: ratio) of the data $\{v_{\max}^i\}$ $(i = 1, \cdots, N)$ was randomly selected, and then the fitting parameters of μ, σ and ξ of the Weibull distributions were calculated for the part. Here, duplication of the same data was allowed if it was selected. This procedure was repeated 100 times, and then the errors of μ, σ and ξ were calculated. In Supplementary Fig. S2, each parameter is plotted as a function of r in the cases of anterograde and retrograde transport. The values were stable for a wide range of r $(0.4 \leq r \leq 1)$.

Code availability

https://github.com/kumikohayashia2/eva

Acknowledgments

We acknowledge Dr. T. Saigo for facilitating useful discussions regarding EVA and Dr. K. Sasaki for facilitating useful discussions regarding motor proteins. We would like to thank Editage (www.editage.com) for their assistance with English language editing. This work was supported by the JST PRESTO (Grant No. JPMJPR1877), and the FRIS Creative Interdisciplinary Research Program, Tohoku University to K. H., and by JSPS KAKENHI (grant nos. 19H04738, 20H03247, and 20K21378) to S. N.

Author contributions

K.H. conceived the project with the help of S.N and wrote the paper. T.N. analyzed the data and performed the simulation. Y.K. and K.N. performed the experiments. S.N. provided the sample.

Competing financial interests

The authors have no competing interests to declare.

Data availability

Data supporting the findings of this study are available within the article and its Supplementary Information files, and also from the corresponding author on reasonable request.

References

1. Gilleland E, Katz RW. extRemes 2.0: An Extreme Value Analysis Package. *Journal of Software* 2016, **72**(8)**:** 1-39.
2. Coles S. An Introduction to Statistical Modeling of Extreme Values. *Springer, London* 2001.
3. de Haan L, Ferreira A. Extreme Value Theory. *Springer* 2006.
4. Tippett MK, Lepore C, Cohen JE. More tornadoes in the most extreme U.S. tornado outbreaks. *Science* 2016, **354**(6318)**:** 1419-1423.
5. Kratz M. Introduction to Extreme Value Theory. Applications Risk Analysis & Manegament. *Matrix annals* 2017**:** pp.591-636.
6. Einmahl JH, Magnus JR. Records in Athletics Through Extreme-Value Theory. *Journal of the American Statistical Association* 2008, **103:** pp. 1382-1391.
7. Gembris D, Taylor JG, Suter D. Trends and random fluctuations in athletics. *Nature* 2002, **417**(6888)**:** 506.
8. Rootzen H, Zolud D. Human life is unlimited. but short. *Extremes* 2017, **20:** pp. 713-728.
9. Wong F, Collins JJ. Evidence that coronavirus superspreading is fat-tailed. *Proc Natl Acad Sci U S A* 2020, **117**(47)**:** 29416-29418.
10. Basnayake K, Mazaud D, Bemelmans A, Rouach N, Korkotian E, Holcman D. Fast calcium transients in dendritic spines driven by extreme statistics. *PLoS Biol* 2019, **17**(6)**:** e2006202.
11. Huxley H, Hanson J. Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature* 1954, **173**(4412)**:** 973-976.
12. Huxley AF, Niedergerke R. Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature* 1954, **173**(4412)**:** 971-973.
13. Hirokawa N, Noda Y, Tanaka Y, Niwa S. Kinesin superfamily motor proteins and intracellular transport. *Nat Rev Mol Cell Biol* 2009, **10**(10)**:** 682-696.
14. Vale RD. The molecular motor toolbox for intracellular transport. *Cell* 2003, **112**(4)**:** 467-480.
15. Okuno D, Iino R, Noji H. Rotation and structure of FoF1-ATP synthase. *J Biochem* 2011, **149**(6)**:** 655-664.
16. Schnitzer MJ, Visscher K, Block SM. Force production by single kinesin motors. *Nat Cell Biol* 2000, **2**(10)**:** 718-723.
17. Gennerich A, Carter AP, Reck-Peterson SL, Vale RD. Force-induced bidirectional stepping of cytoplasmic dynein. *Cell* 2007, **131**(5)**:** 952-965.
18. Toba S, Watanabe TM, Yamaguchi-Okimoto L, Toyoshima YY, Higuchi H. Overlapping hand-over-hand mechanism of single molecular motility of cytoplasmic dynein. *Proc Natl Acad Sci U S A* 2006,

103(15)**:** 5741-5745.

19. Elshenawy MM, Canty JT, Oster L, Ferro LS, Zhou Z, Blanchard SC*, et al.* Cargo adaptors regulate stepping and force generation of mammalian dynein-dynactin. *Nat Chem Biol* 2019, **15**(11)**:** 1093-1101.
20. Hirakawa E, Higuchi H, Toyoshima YY. Processive movement of single 22S dynein molecules occurs only at low ATP concentrations. *Proc Natl Acad Sci U S A* 2000, **97**(6)**:** 2533-2537.
21. Mallik R, Carter BC, Lex SA, King SJ, Gross SP. Cytoplasmic dynein functions as a gear in response to load. *Nature* 2004, **427**(6975)**:** 649-652.
22. Rondelez Y, Tresset G, Nakashima T, Kato-Yamada Y, Fujita H, Takeuchi S*, et al.* Highly coupled ATP synthesis by F1-ATPase single molecules. *Nature* 2005, **433**(7027)**:** 773-777.
23. Watanabe R, Iino R, Noji H. Phosphate release in F1-ATPase catalytic cycle follows ADP release. *Nat Chem Biol* 2010, **6**(11)**:** 814-820.
24. Toyabe S, Okamoto T, Watanabe-Nakayama T, Taketani H, Kudo S, Muneyuki E. Nonequilibrium energetics of a single F1-ATPase molecule. *Phys Rev Lett* 2010, **104**(19)**:** 198103.
25. Toyabe S, Watanabe-Nakayama T, Okamoto T, Kudo S, Muneyuki E. Thermodynamic efficiency and mechanochemical coupling of F1-ATPase. *Proc Natl Acad Sci U S A* 2011, **108**(44)**:** 17951-17956.
26. Ariga T, Tomishige M, Mizuno D. Nonequilibrium Energetics of Molecular Motor Kinesin. *Phys Rev Lett* 2018, **121**(21)**:** 218101.
27. Hall DH, Hedgecock EM. Kinesin-related gene unc-104 is required for axonal transport of synaptic vesicles in C. elegans. *Cell* 1991, **65**(5)**:** 837-847.
28. Otsuka AJ, Jeyaprakash A, Garcia-Anoveros J, Tang LZ, Fisk G, Hartshorne T*, et al.* The C. elegans unc-104 gene encodes a putative kinesin heavy chain-like protein. *Neuron* 1991, **6**(1)**:** 113-122.
29. Chen CW, Peng YF, Yen YC, Bhan P, Muthaiyan Shanmugam M, Klopfenstein DR*, et al.* Insights on UNC-104-dynein/dynactin interactions and their implications on axonal transport in Caenorhabditis elegans. *J Neurosci Res* 2019, **97**(2)**:** 185-201.
30. Hayashi K, Hasegawa S, Sagawa T, Tasaki S, Niwa S. Non-invasive force measurement reveals the number of active kinesins on a synaptic vesicle precursor in axonal transport regulated by ARL-8. *Phys Chem Chem Phys* 2018, **20**(5)**:** 3403-3410.
31. Rai AK, Rai A, Ramaiya AJ, Jha R, Mallik R. Molecular adaptations allow dynein to generate large collective forces inside cells. *Cell* 2013, **152**(1-2)**:** 172-182.
32. Rasband WS. ImageJ. *National Institute of Health, Bethesda, Maryland, USA* 1997, **http://imagej.nih.gov/ij/**.
33. Gross SP. Hither and yon: a review of bi-directional microtubule-based transport. *Phys Biol* 2004, **1**(1-2)**:** R1-11.
34. R Core Team. R: Alanguage and environment for statistical computinh. R Foundation for Statistical

Computing, Vienna, Austria. URL https://www.R-project.org/. 2018.

35. Longoria RA, Shubeita GT. Cargo transport by cytoplasmic Dynein can center embryonic centrosomes. *PLoS One* 2013, **8**(7)**:** e67710.
36. Hayashi K, Miyamoto MG, Niwa S. Effects of dynein inhibitor on the number of motor proteins transporting synaptic cargos. *Biophys J* 2021, **120**(9)**:** 1605-1614.
37. Hayashi K, Tsuchizawa Y, Iwaki M, Okada Y. Application of the fluctuation theorem for non-invasive force measurement in living neuronal axons. *Mol Biol Cell* 2018**:** mbcE18010022.
38. Hasegawa S, Sagawa T, Ikeda K, Okada Y, Hayashi K. Investigation of multiple-dynein transport of melanosomes by non-invasive force measurement using fluctuation unit chi. *Sci Rep* 2019, **9**(1)**:** 5099.
39. Sasaki K, Kaya M, Higuchi H. A Unified Walking Model for Dimeric Motor Proteins. *Biophys J* 2018, **115**(10)**:** 1981-1992.
40. Iwaki M, Wickham SF, Ikezaki K, Yanagida T, Shih WM. A programmable DNA origami nanospring that reveals force-induced adjacent binding of myosin VI heads. *Nat Commun* 2016, **7:** 13715.
41. Anazawa Y, Kita T, Hayashi K, Niwa S. De novo disease-associated mutations in KIF1A dominant negatively inhibit axonal trasnport of synaptic vesicle precursors. *Biorxiv* 2021, **https://www.biorxiv.org/content/10.1101/2021.07.22.453457v1**.
42. Gabrych DR, Lau VZ, Niwa S, Silverman MA. Going Too Far Is the Same as Falling Short(dagger): Kinesin-3 Family Members in Hereditary Spastic Paraplegia. *Front Cell Neurosci* 2019, **13:** 419.
43. Niwa S, Lipton DM, Morikawa M, Zhao C, Hirokawa N, Lu H*, et al.* Autoinhibition of a Neuronal Kinesin UNC-104/KIF1A Regulates the Size and Density of Synapses. *Cell Rep* 2016, **16**(8)**:** 2129-2141.
44. Wu YE, Huo L, Maeder CI, Feng W, Shen K. The balance between capture and dissociation of presynaptic proteins controls the spatial distribution of synapses. *Neuron* 2013, **78**(6)**:** 994-1011.

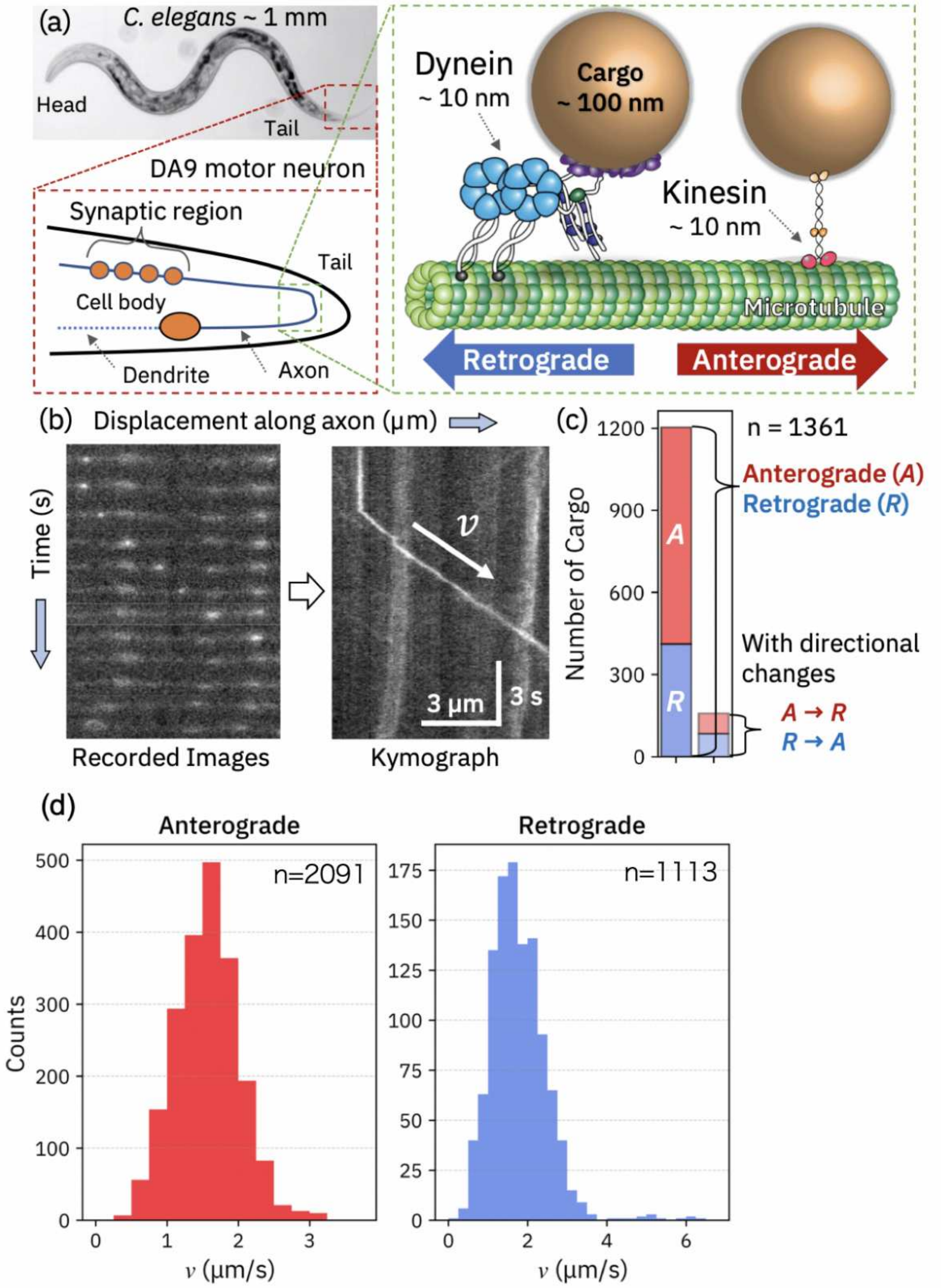

Fig. 1 (a) Schematics of anterograde and retrograde synaptic cargo transport by kinesin and dynein, respectively, in the DA9 motor neurons of *C. elegans*. (b) Kymograph analysis. Velocity (v) was measured as the slope of the trajectory of a fluorescently labeled cargo. (c) Number of synaptic cargos moving anterogradely (A), retrogradely (R), and exhibiting direction change (A→R and R→A). (d) Histogram of the velocity $\{v^i\}$ of synaptic cargos for anterograde (red) and retrograde (blue) transport.

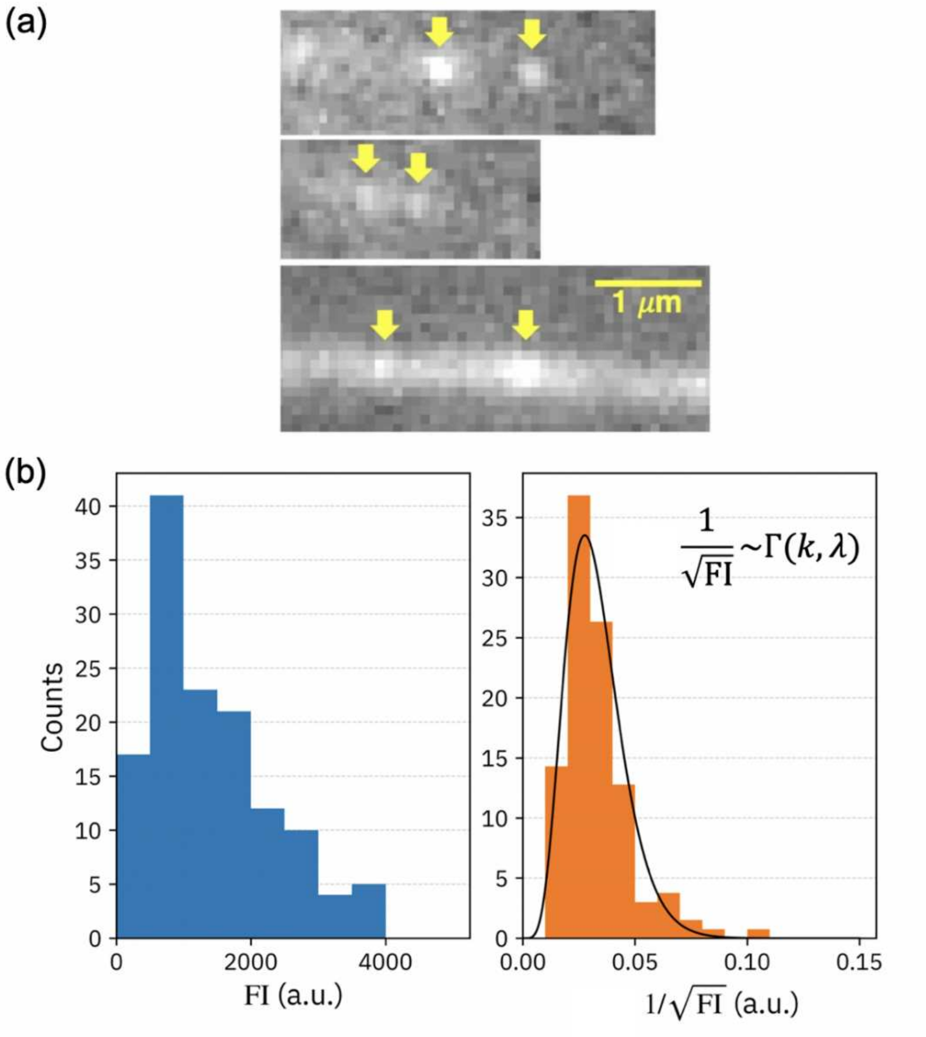

Fig. 2. (a) Fluorescence micrographs of synaptic cargos transported by motor proteins in the axons of DA9 neurons in *C. elegans* worms. (b) Fluorescence intensity (FI) of synaptic cargos ($n = 133$) (left panel). The distribution of $1/\sqrt{\mathrm{FI}}$ was fitted using a Gamma distribution (right panel).

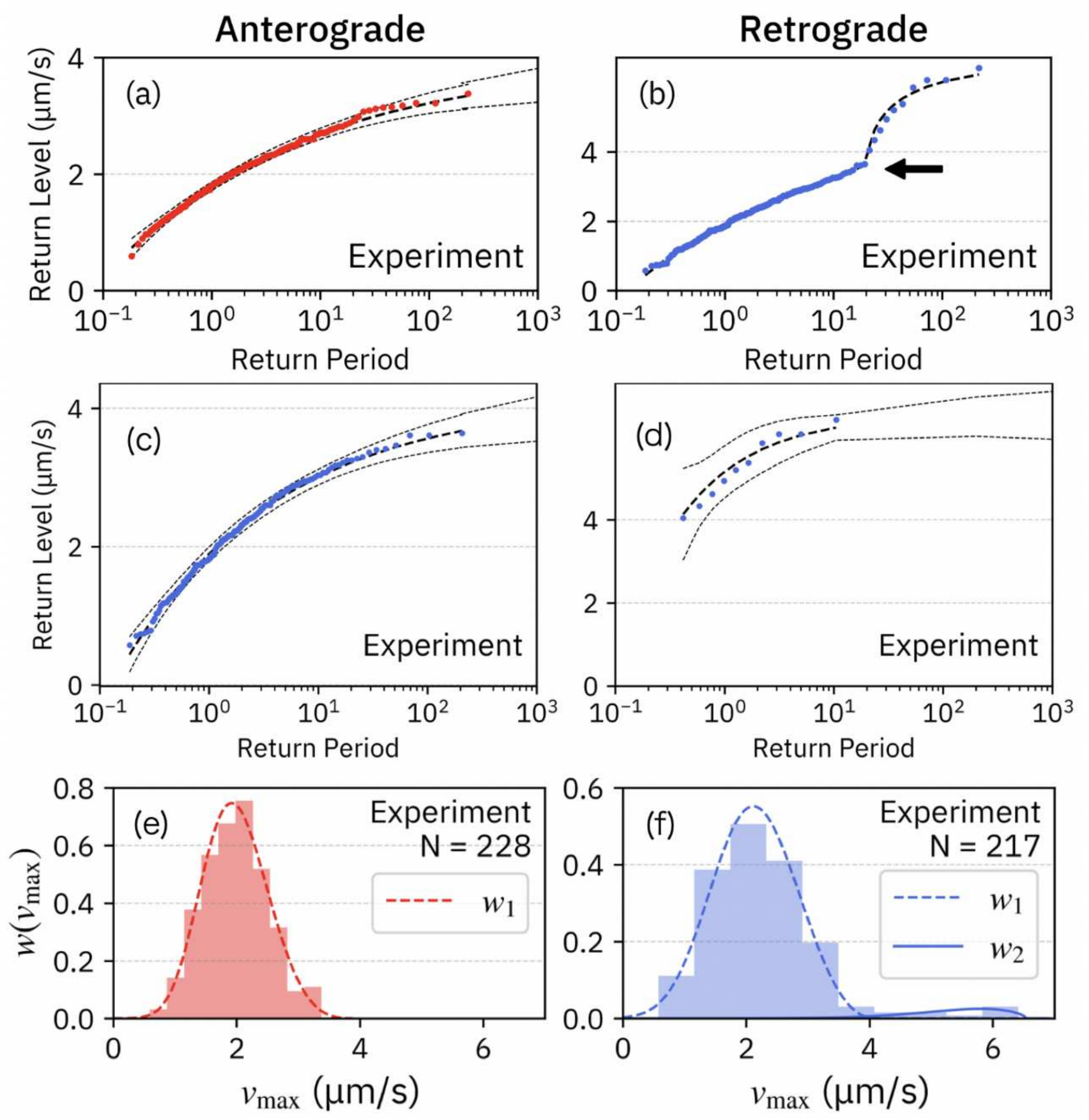

Fig. 3. Return-level plots experimentally measured for anterograde (a) and retrograde transport (b). (c, d) Retrograde data (b) were separated by the change point (the black arrow in (b)). Probability distributions of transport velocity for anterograde (e) and retrograde transport (f).

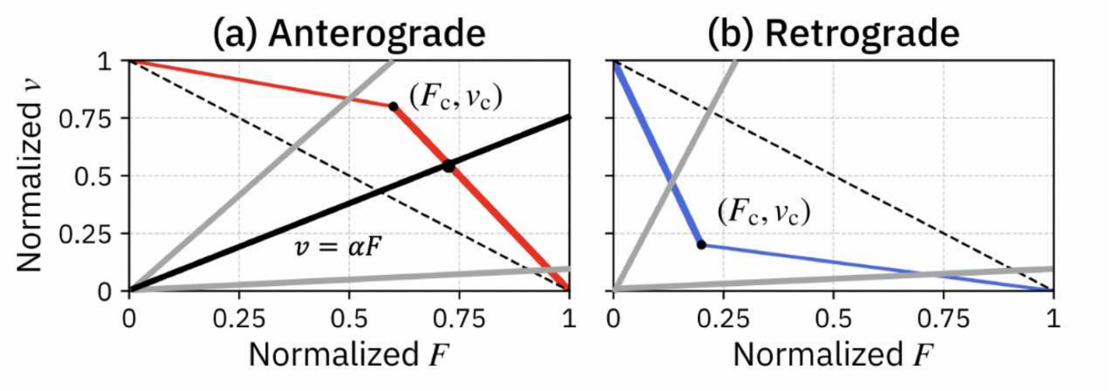

Fig. 4. Concave-down (a) and concave-up (b) force-velocity curves of a motor protein. Two regimes, load-sensitive (thick colored line) and load-insensitive (thin colored line), are represented. The thick black line represents the Stokes' law $v = \alpha F$. The grey lines in each graph represent the 95% confidence intervals of α. Normalized F: F/F_s and v: $v/v(F = 0)$, where F_s is the stall force of a motor protein.

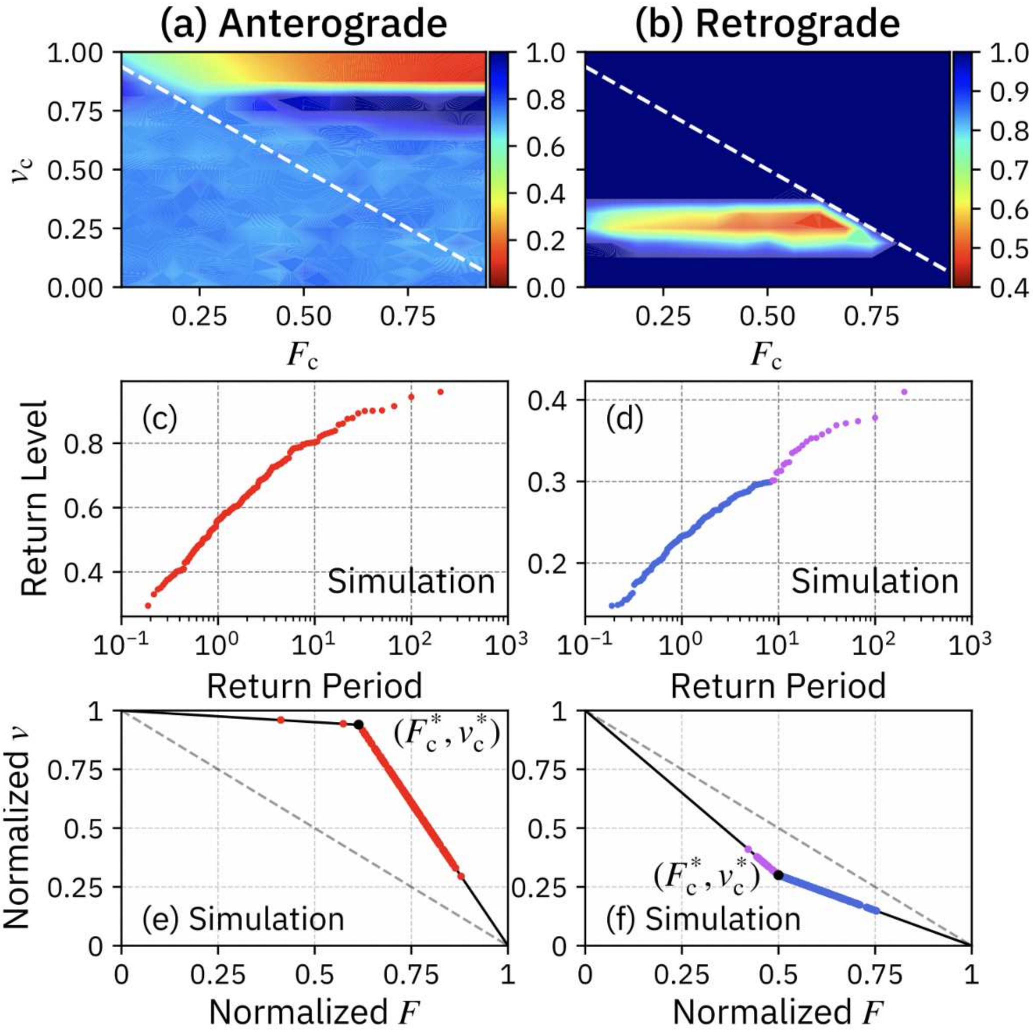

Fig. 5. Contour of $E(F_c, v_c)$ simulated return-level plots and force-velocity relationship obtained by using the most probable value of (F_c^*, v_c^*) for anterograde ((a), (c), and (e)) and retrograde ((b), (d), and (f)) transport. Note that the *y*-axis in panels (c) and (d) are normalized as $v/V_{\max}$.

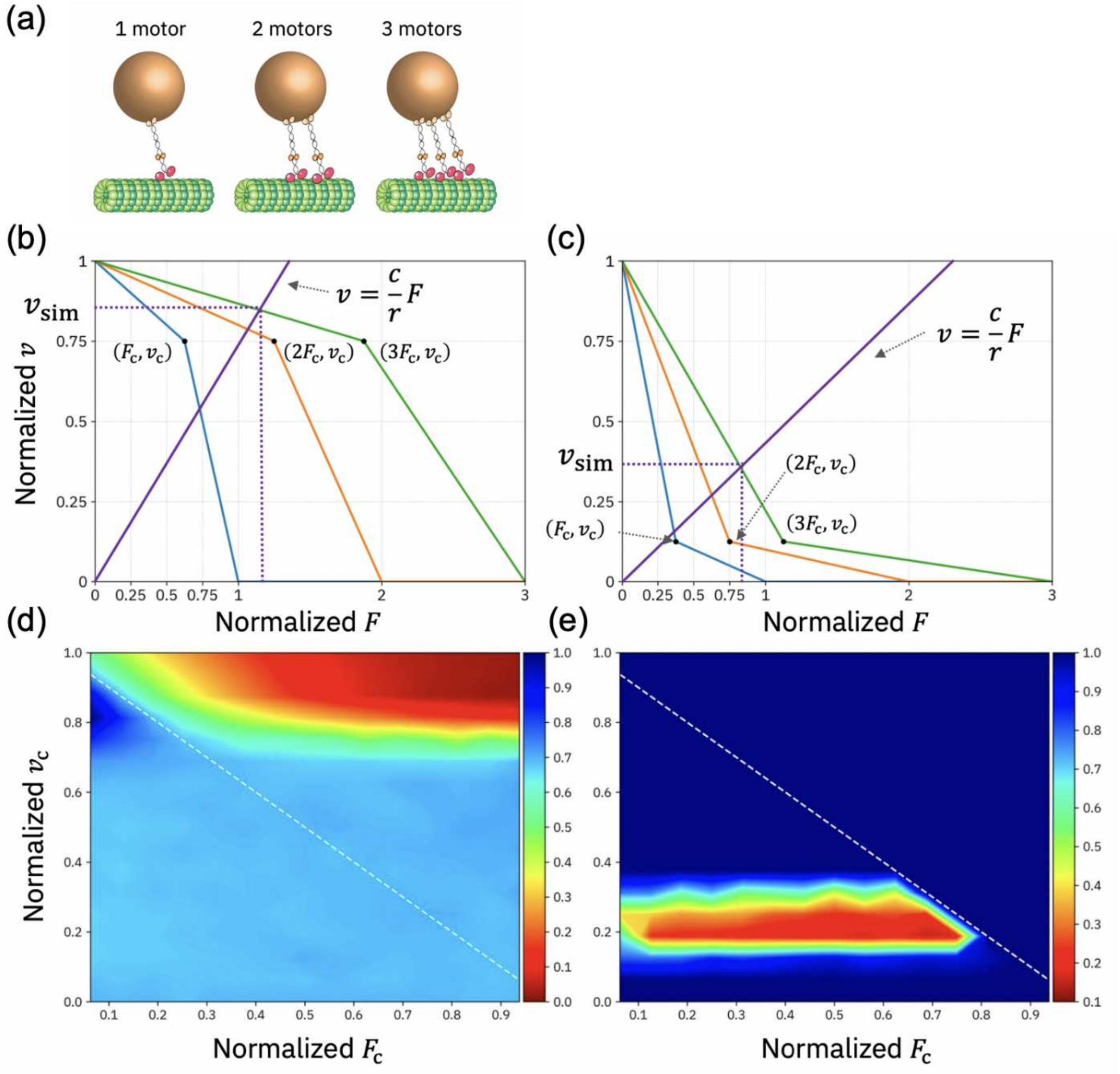

Fig. 6. (a) Schematics of cargo transport by multiple motor proteins. Force-velocity relationship and corresponding contour of $E(F_c, v_c)$ for multiple motor proteins ($N_m = 1,2,3$) during anterograde ((b) and (d)) and retrograde ((c) and (e)) transport.

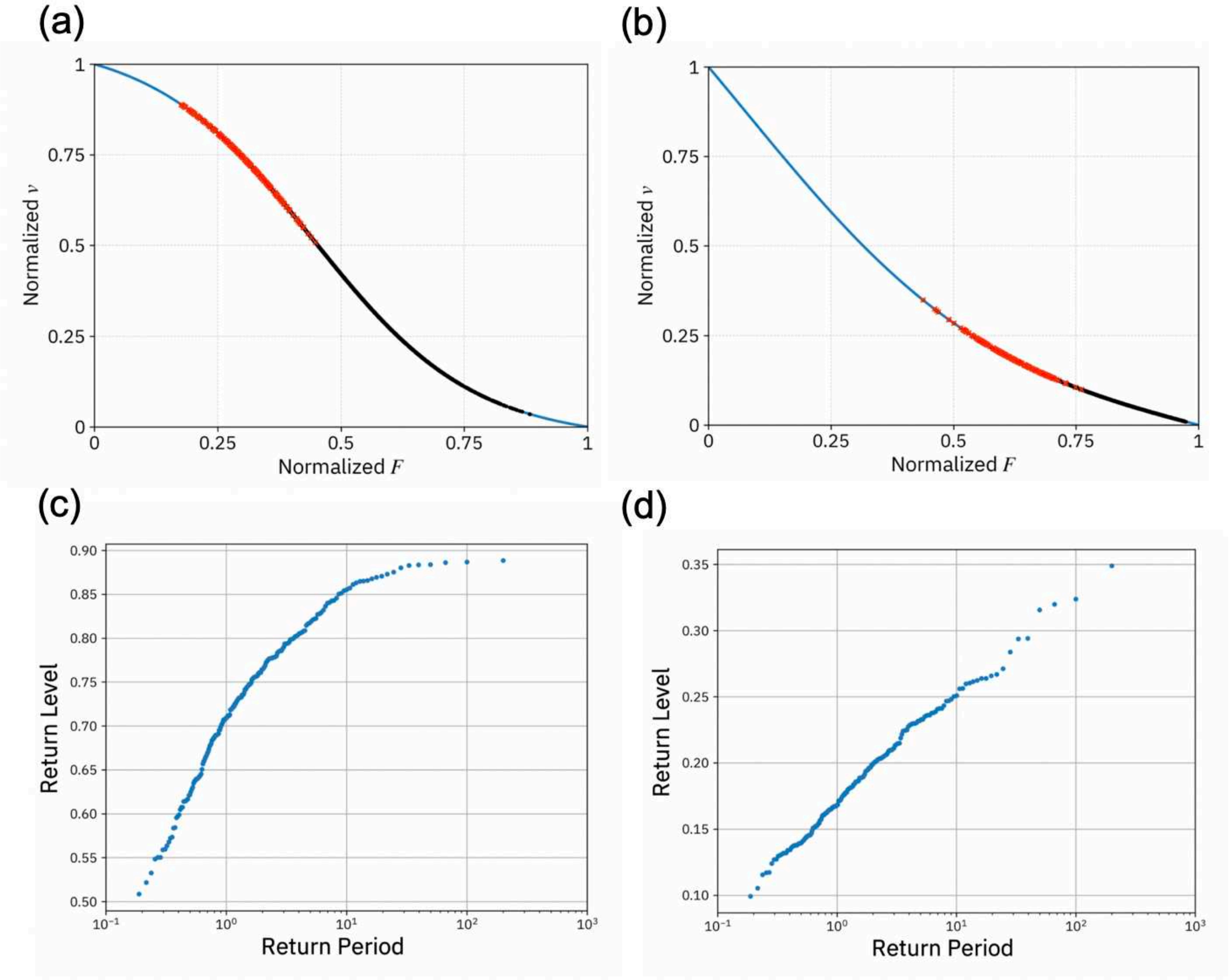

Fig. 7. Force-velocity relationship and corresponding return-level plots of the three-state model for anterograde ((a) and (c)) and retrograde ((b) and (d)) transport.