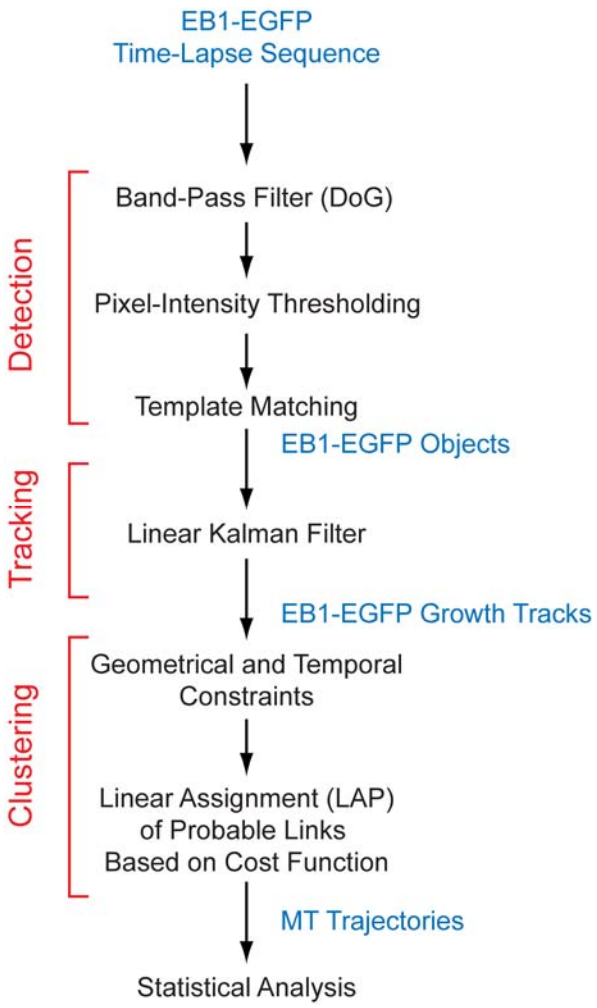
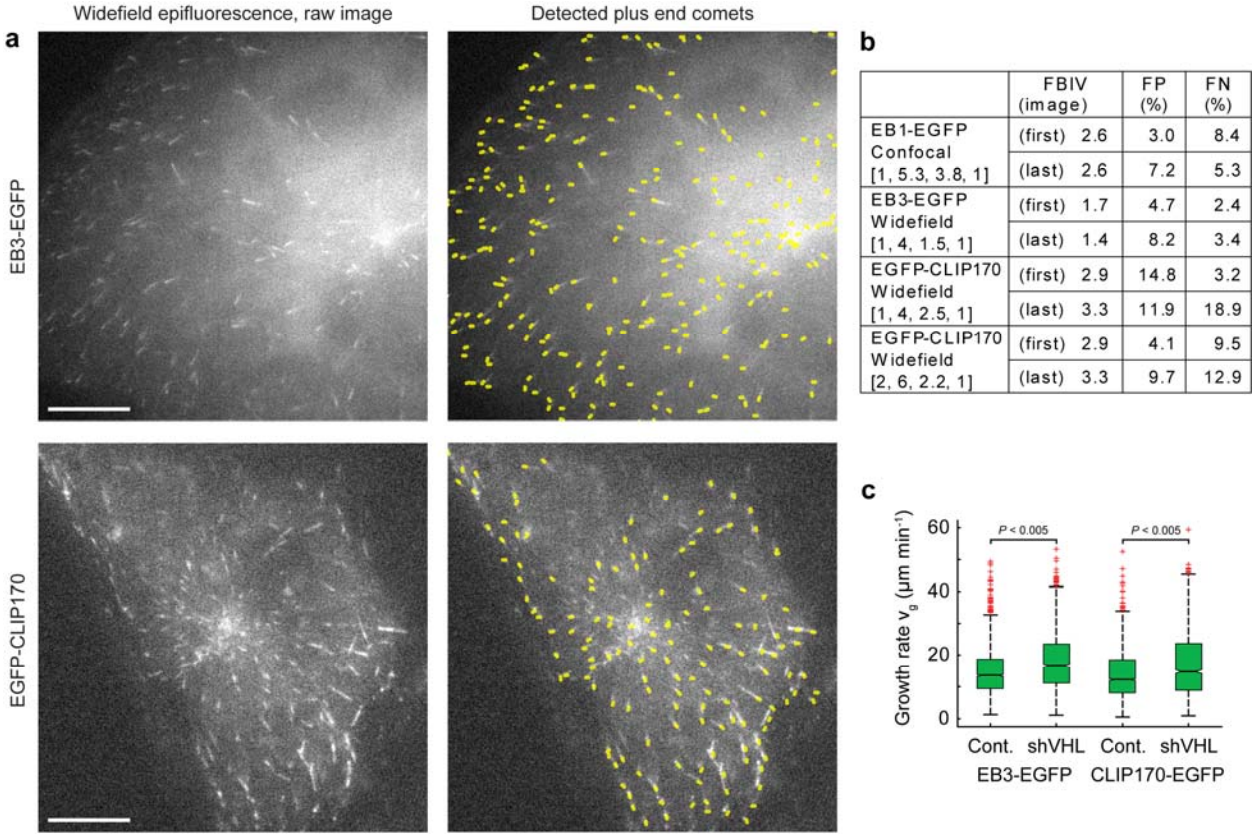
Supplementary Figure 1. Flow diagram of EB1-EGFP comet detection, tracking, and geometrical clustering.



Supplementary Figure 2. Detection in widefield epifluorescence images.

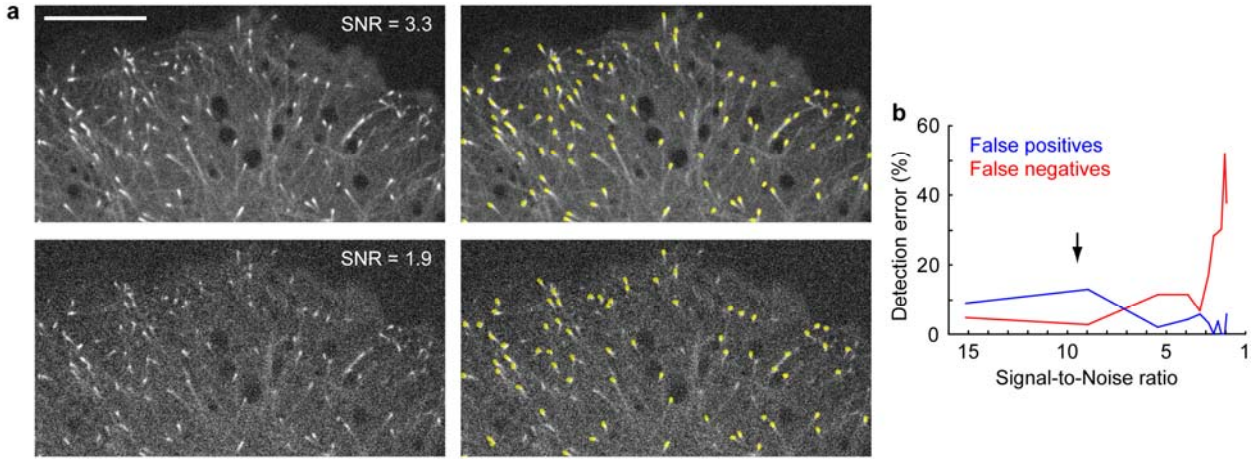


	FBIV (image)	FP (%)	FN (%)
EB1-EGFP Confocal [1, 5.3, 3.8, 1]	(first) 2.6	3.0	8.4
	(last) 2.6	7.2	5.3
EB3-EGFP Widefield [1, 4, 1.5, 1]	(first) 1.7	4.7	2.4
	(last) 1.4	8.2	3.4
EGFP-CLIP170 Widefield [1, 4, 2.5, 1]	(first) 2.9	14.8	3.2
	(last) 3.3	11.9	18.9
EGFP-CLIP170 Widefield [2, 6, 2.2, 1]	(first) 2.9	4.1	9.5
	(last) 3.3	9.7	12.9

(**a**) Robust detection of growing microtubule plus ends in 60x low contrast widefield images of RPE-1 cells expressing either EB3-EGFP or EGFP-tagged CLIP170. Yellow overlays are computer-detected comets. Scale bar, 10 µm. (**b**) Comparison of false positive (FP) and false negative (FN) detections in different imaging modalities in the first and last frame of a time-lapse sequence. The differences in image characteristics are quantified by the Foreground-to-Background Intensity Variation (FBIV), which is calculated as the mean of the difference of detected +TIP intensities and immediate surroundings divided by the standard deviation of the intensity of a cellular region excluding the +TIP signals. It thus compares the local contrast of +TIP comets to the variation of the cellular background, which is a major source of confusion to the detection algorithm. A higher FBIV number indicates less intracellular signal variation relative to the average contrast of a comet and consequently results in a lower percentage of FP detections. Optimized detection parameters used in this table were determined by varying σ_2 and k_1 to achieve the minimum combined error FP + FN, and detection parameters are indicated as $[\sigma_1, \sigma_2, k_1, k_2]$ for each condition. Using the default value of $\sigma_1 = 1$, the detection error of EGFP-CLIP170 was higher because of the larger variability of CLIP170 comets as compared to EBs, but adjustment of this parameter returned the detection errors to levels similar to EBs. This illustrates the need for some parameter adjustments for the detection algorithm with different imaging conditions. However, our experience from analyzing large numbers of movies is that the cell-to-cell variation under one imaging condition or the contrast reduction within one movie due to photobleaching insignificantly affects the detection performance (compare detection error analysis for first and last image). (**c**) Microtubule growth rates derived from cells either expressing EGFP-tagged EB3 or CLIP-170 in

control RPE-1 cells or cells in which a modifier of microtubule polymerization dynamics, VHL, was depleted by small hairpin-mediated RNA interference[4]. Both plus end tracking proteins yielded essentially the same distribution of growth rates, and revealed differences between control and VHL-depleted cells demonstrating that our detection and tracking algorithms are sufficiently self-adaptive and do not depend on the precise shape of the plus end comets.

Supplementary Figure 3. Detection performance in response to high frequency pixel noise.



(**a**) Two examples of images with artificial pixel noise contamination. Increasing noise levels were simulated by adding Gaussian distributed random intensity values to the image in **Fig. 1a**. The image on the bottom had the lowest signal-to-noise ratio used in this analysis. Computer-detected comets are indicated in yellow. Random high frequency noise is a good estimate of camera noise introduced by charged-coupled device cameras. However, cameras using on-chip amplification may introduce more complicated noise patterns. (**b**) The detection accuracy is plotted as a function of the signal-to-noise ratio (SNR = I/σ). The signal I was calculated as the mean of the intensity difference between detected EB1-EGFP comets and their immediate surroundings. The noise σ is defined as the standard deviation of a region outside the cell. Detection parameters were kept constant and the hand-detected set of comets from the original image was used to evaluate detection performance. As expected, increasing noise first obfuscated dim comets and eventually resulted in a drastic increase of false negative detections. This occurred at a signal-to-noise ratio about 5-fold below that of the original image resulting in a sharp increase of false negative detections. The percentage of false positives was relatively independent of high frequency noise, and actually improved somewhat with increasing noise because the dim EB1-EGFP signal along microtubules, which is the main reason for false positive detections, disappeared first with increasing noise levels. The arrow indicates the noise level of the 75th image in this time-lapse sequence (**Video 1**) as a result of photobleaching. Thus, the detection breakpoint is 2-3-fold below the lowest image contrast in this sequence.

Supplementary Figure 4. Magnification dependence of detection and tracking.

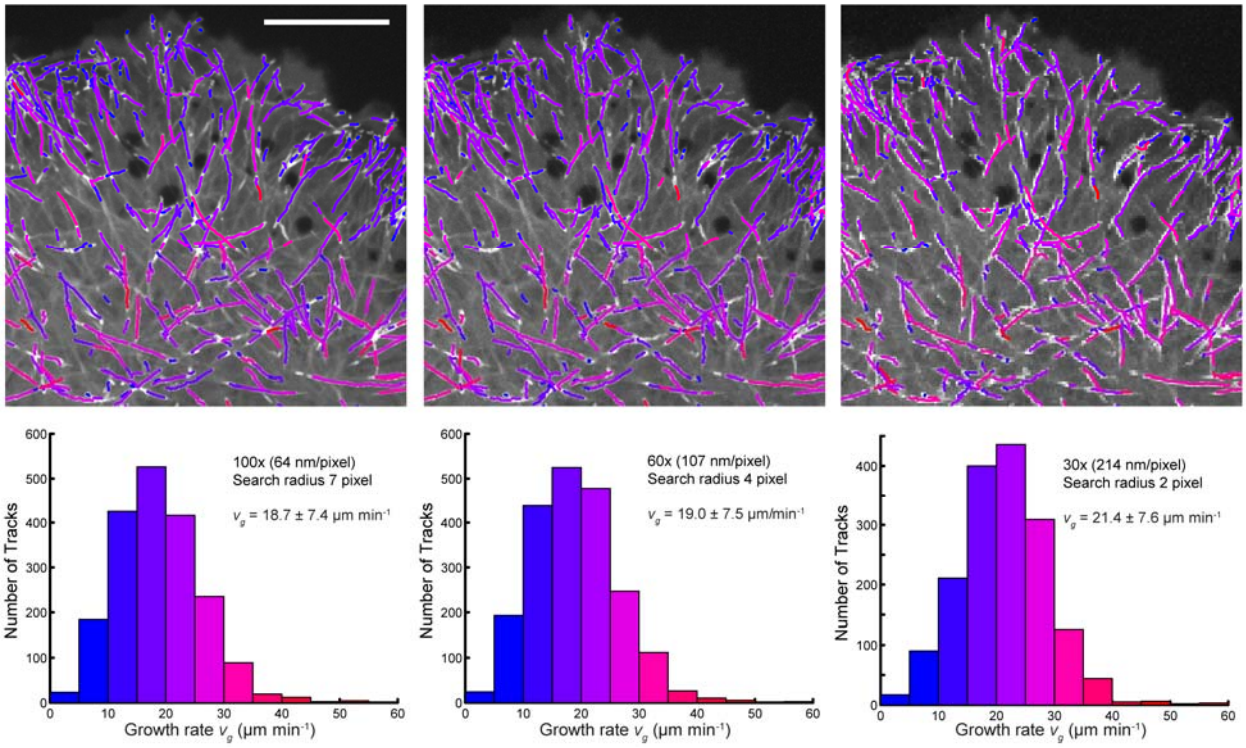


Image sequences of different spatial resolution were generated from an original 100x sequence by linear downsampling. This simulates images with pixel sizes as if they were acquired at 60x or 60x with 2x2 binning (= 30x) on the same microscope system. Effective pixel sizes are indicated on the growth rate histograms. During tracking the search radius was adjusted such that it represented approximately the same maximally allowed growth rate of ~65 µm min^{-1}. Images show computer-generated EB1-EGFP tracks overlaid on maximum intensity projections of the entire sequence (75 images, 0.4 s frame^{-1}) in a region of the analyzed cell. The number of tracks, distribution of growth rates, and the average growth rate were very similar at 100x and 60x magnification, which is at or above the Nyquist sampling criterion demonstrating that our detection algorithm is scalable and only limited by the optical resolution of the microscope. As expected, the number of detected comets and thus computer-generated tracks decreased significantly at the simulated 30x magnification, which represents undersampling of the optical resolution, and thus EB1-EGFP comets in close proximity are no longer resolved. Because this mainly impacts EB1-EGFP detection and tracking in the cell periphery (top edge of the image), where microtubule ends are more densely packed yet slower (**Fig. 5**), the average growth rate at 30x is significantly increased. Scale bar, 10 µm.

Supplementary Figure 5. Frame rate dependence of detection and tracking.

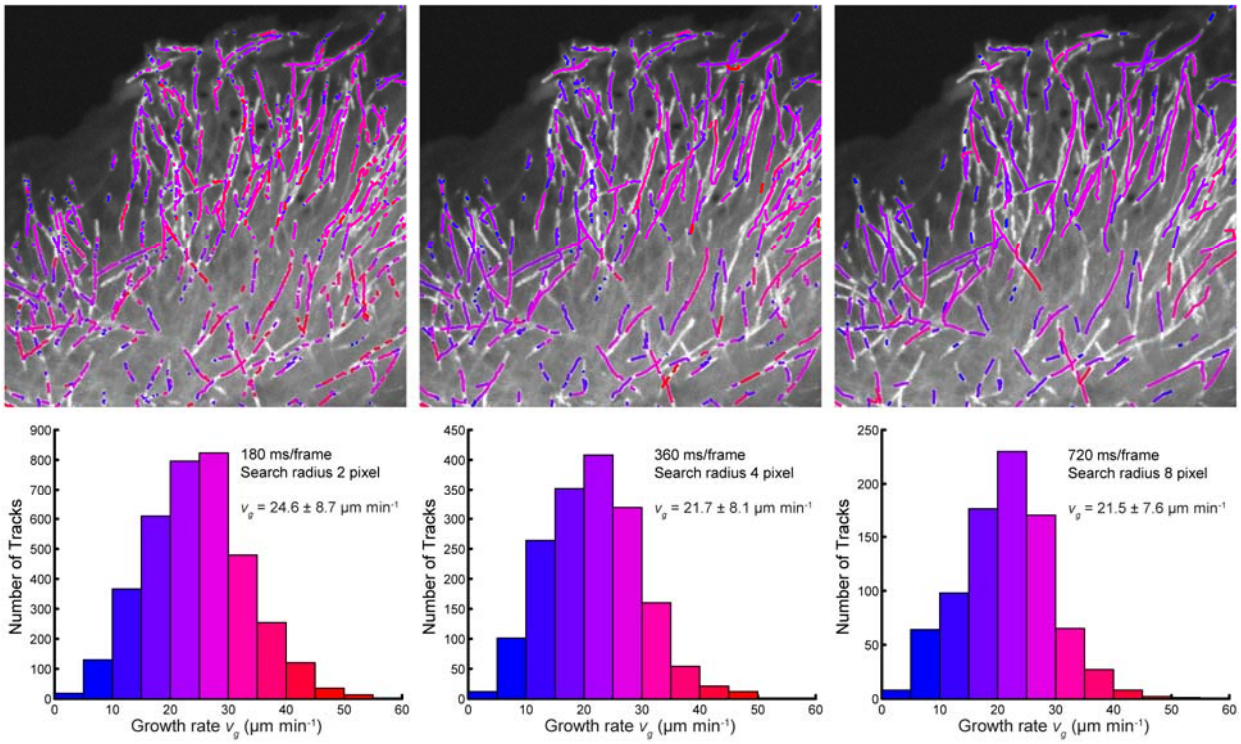


Image sequences of different temporal resolution were generated from an original 60x sequence that was acquired at 180 ms frame^{-1}, the maximum that could be achieved on our microscope system, by only using each second or fourth acquired image for tracking. The resulting effective frame rates are indicated on the growth rate histograms. During tracking the search radius was adjusted such that it represented approximately the same maximally allowed growth rate of ~70 µm min^{-1}. Images show computer-generated EB1-EGFP tracks overlaid on maximum intensity projections of the entire sequence (124 images in the original sequence) in a region of the analyzed cell. Although the number of detected tracks decreased with decreasing frame rate, the distribution of growth rates and the calculated average growth rates were almost identical for 360 ms frame^{-1} and 720 ms frame^{-1}. This indicates that there is no systematic loss of data in this frame rate window and demonstrates that frame rates of ~500 ms frame^{-1} that we used throughout this study are optimal for our tracking approach. At 180 ms frame^{-1} we obtained significantly higher growth rates, which may be due to multiple factors. Frequent disappearance of EB1-EGFP comets due to rapid fluctuations of microtubule growth rates that are more frequently missed at slower frame rates may result in a larger population of short and fast tracks[5]. Moreover, at such high frame rates false negative EB1-EGFP detections as well as the positional error of comet detection also become more relevant. To achieve 180 ms frame^{-1} acquisition it was also necessary to image a cell expressing high levels of EB1-EGFP. Thus, the slightly higher growth rate determined in this experiments compared to other cells analyzed may reflect the effect of EB1 overexpression because EB1 enhances microtubule growth rates at high concentrations. Scale bar, 10 µm.

Supplementary Figure 6. Demonstration of tracking fidelity in the cell interior.

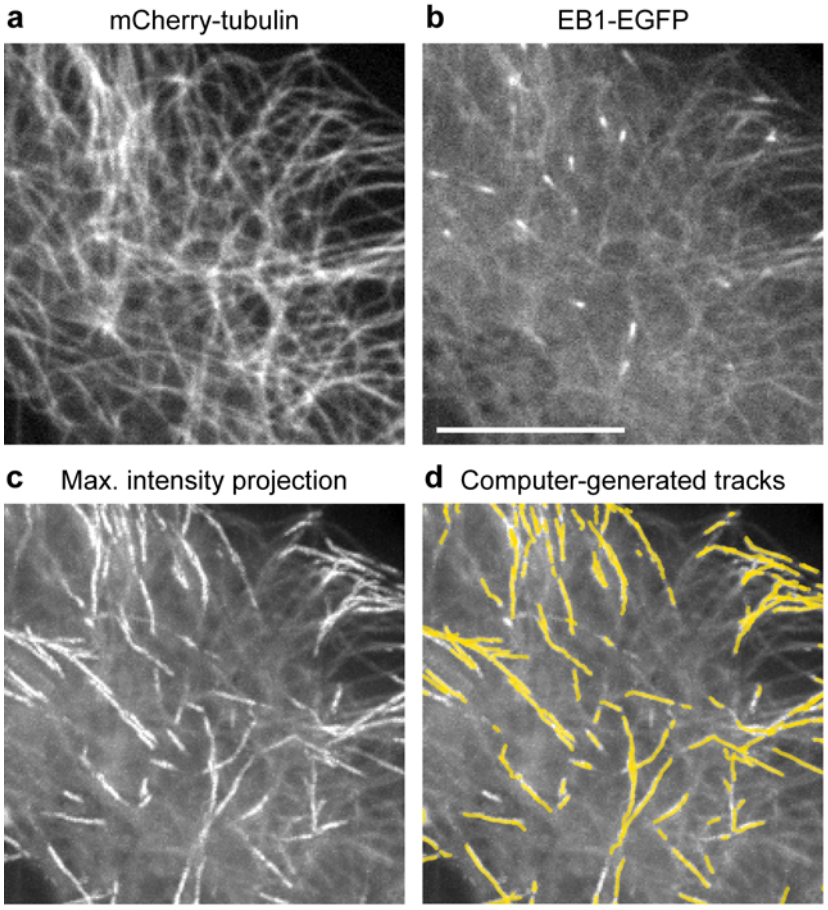


Image region 3 of the cell in **Fig. 3e** expressing both mCherry-tubulin and EB1-EGFP is shown. (**a**) Because of the high density of microtubules in the cell interior it is basically impossible to identify the ends of individual microtubules. (**b**) Nevertheless, growing microtubule ends are clearly visible in the EB1-EGFP channel in the same image region. Overlay of the EB1-EGFP maximum intensity projection of the entire sequence (**c**) and computer-generated tracks (**d**, yellow) shows almost complete overlap demonstrating that by EB1-EGFP tracking microtubule dynamics can be followed in cell areas not amenable to analysis by continuously labeled microtubules. Scale bar, 10 µm.

Supplementary Table 1. Control parameters for comet detection, tracking, and clustering.

Input parameter	**Default** (useful range)	**Notes**
Detection		
σ_1: Standard deviation of Gaussian filter kernel for high spatial frequencies	**1.25**	Defined by the spread (standard deviation) of the Gaussian approximation of the point spread function (**Supplementary Note 1**).
σ_2: Standard deviation of Gaussian filter kernel for low spatial frequencies	**4.2** **(**1.5 10)	Needs adjustment to the average length of the comet depending on magnification and experimental conditions (**Supplementary Note 1**).
k_1: Noise thresholding coefficient	**1** (0.5 4)	Needs adjustment depending on the comet to comet pixel intensity variation. Low contrast images (widefield) have a low variation, allowing k_1 close to the default value. High-contrast (confocal) or drug perturbation images may have high variation, which may require larger k_1. Higher values result in more stringent detection.
k_2: Feature template thresholdingcoefficient	**1** **(**0.5 2)	Needs occasional adjustment in conditions that change the signal to background characteristic of the comets. Higher values result in less stringent detection.
Tracking		
Maximum search angle (not discussed in paper)	**45°**	Robust, hard-coded parameter; no adjustment necessary.
Maximum direction change (not discussed in paper)	**90°**	Robust, hard-coded parameter; no adjustment necessary.
Search radius / Linking distance threshold (not discussed in paper)	**no default**	Maximal displacement expected for comets between consecutive frames in pixels. This parameter needs to be adjusted for every experiment. A good strategy is to manually determine the approximate velocity of a few comets. Usually, a 3-fold larger value is sufficient

		to accommodate the full range of velocities. An appropriate setting is confirmed when the last bin of the velocity histogram contains a negligible population of values.
Minimum growth track lifetime considered for geometric clustering	**4 frames** (> 2)	The parameter should be >2, as a significant population of growth tracks with a lifetime = 2 are associated with false positives of the detection step. As a rule of thumb the parameter should be set in the range of the latency of comet signal formation.
Clustering		
φ: Forward cone opening	**±45°** (±15° ±45°)	The parameter reflects the variation in lateral displacement of a microtubule end during pauses. For most of our analyses we use the default setting. Narrowing the cone increases stringency as documented in the Validation section.
ρ: Backward cone opening	**±10°**	The parameter reflects the variation in lateral displacement of a rescued microtubule from its original axis of growth. The backward cone opening is set narrower than the forward cone for two reasons: First, rescued microtubules tend to follow the previous axis of growth relatively strictly, more so than microtubules re-growing after pauses. Second, the clustering of microtubules in the backward direction is less robust as a microtubule undergoing catastrophe could be linked to a growth track of another following microtubule. To minimize such errors, the search space is chosen conservatively.
α_{max}: Maximal angle for directional changes between track pairs	**60°** (15° 60°)	The parameter is set by observations of microtubule bending. It limits the change in microtubule growth direction over phases of invisible growth or shortening. For most of our analyses we use the default setting. Effects of lowering the cut off angle on false positives and false negatives are documented in the Validation section.
T_{max}: Maximum time window of linking	**30 frames**	Usually, this parameter is robust when set to allow forward and backward links over ~20 s (thus, adjustment is needed for different frame rates).

		The setting can be validated based on the lifetime histogram of forward gaps. An appropriate setting is confirmed when the last bin of by the lifetime histogram of forward gaps contains a negligible population of values.
C_{fwd}: Distance threshold for forward gaps	Self-adjusted for each candidate gap	
C_{bwd}: Distance threshold for backward gaps	Self-adjusted for each candidate gap	
γ: Multiplication factor for microtubule shortening rates	**2** (1.1 3)	Robust, usually hard-coded parameter. An adjustment is necessary; when the last few bins of the histogram of shortening rates contain a significant population of values (γ should be increased to accommodate faster rates); or when visual inspection of track groups indicate a large number of nonsensical backward links in cell regions of high comet density (γ should be decreased).

Supplementary Table 2. Quantification of false positive events assigned by growth track clustering.

Clustering stringency (low/**high**)	Growth	Pause	Shortening	Catastrophe	Rescue
False positive	2%/**1%**	14%/**5%**	20%/**10%**	13%/**8%**	11%/**7%**
Unconfirmed	3%/**2%**	33%/**24%**	48%/**58%**	49%/**57%**	53%/**62%**
True positive	95%/**97%**	53%/**71%**	32%/**32%**	38%/**35%**	36%/**31%**
True positive (of confirmed events)	98%/**99%**	79%/**94%**	61%/**77%**	74%/**82%**	77%/**83%**
Number of events	331/**224**	101/**42**	111/**92**	111/**92**	111/**92**

Each cell shows the percentage of events as verified by visual examination of the clustered trajectory compared to the original EB1-EGFP and mCherry-tubulin image sequence in region 2 of **Fig. 3e**. The percentage of false positive events is shown for clustering at low ($\varphi = 45°$, $\alpha = 60°$), and high stringency ($\varphi = 15°$, $\alpha = 15°$).

Supplementary Note 1: Difference of Gaussian transformation and σ_1 and σ_2 selection

The difference of two Gaussian transformation (*DoG*) is calculated as

$$DoG = \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{\sqrt{2\pi\sigma_2^2}} \cdot e^{-\frac{x^2+y^2}{2\sigma_2^2}}$$

to enhance the band of image frequencies associated with EB1-EGFP comets, while attenuating frequencies associated with image noise ($> 1/\sigma_1$), or with long range variations of the diffuse background ($<1/\sigma_2$). Accordingly, the value of σ_1 is set to match the point spread function of the microscope[1], which corresponds to the expected width of the image of an EB1 comet:

$$\sigma_1 = \frac{0.21\lambda}{NA \cdot P_{xy}}$$

λ is the emission wavelength of the fluorescent protein, *NA* the numerical aperture, and P_{xy} the pixel size in object space. The value of σ_2 depends on the length of the average image of an EB1 comet. Although most comets are elongated, the eccentricity of detectable EB1-EGFP comets is influenced by multiple factors including magnification, EB1-EGFP expression level, and microtubule growth dynamics. In addition, low signal-to-noise ratio can result in less eccentric EB1-EGFP comets because only the intensity maxima of the EB1-EGFP comets are sufficiently above the noise floor. To determine σ_2 for our imaging conditions, we measured the average length l_e of a typical EGFP blob along its longitudinal axis in pixels. This one-dimensional profile is modeled as:

$$f_{l_e}(x) = \begin{cases} 1 \text{ if } x \in \left[-\frac{l_e}{2}, \frac{l_e}{2}\right] \\ 0 \text{ otherwise} \end{cases}$$

An optimal value for σ_2 is obtained by solving

$$\hat{\sigma}_2 = \arg\max_{\sigma_2}(f_{l_e} \cdot \nabla^2 G_{\sigma_2}(0))$$

which yields the solution

$$\sigma_2 = \sqrt{2}\frac{l_e}{2}$$

Supplementary Note 2: Unimodal thresholding to select EB1-EGFP pixels

The unimodal thresholding algorithm determines a threshold at the histogram index that maximizes the perpendicular distance to a line connecting the largest bin to the first empty bin[2]. In a normal distribution, this selection cuts the population at $\mu+3\sigma$, i.e. at the 99th percentile. Unimodal thresholding works best for images with a single distribution of bright objects with relatively uniform intensity that is hidden in a much larger distribution of dim objects detected due to noise. Thus, the threshold estimation is sensitive to very bright features leading to exceedingly high-valued bins as well as to the presence of additional intensity distributions such as, for example, additional EB1-EGFP signal along microtubules. We therefore introduced the coefficient k_1 as a user-defined input to adjust the threshold, if necessary for certain imaging conditions.

Supplementary Note 3: Selection of EB1-EGFP comets based on average comet shape

To discriminate against objects that are not true EB1-EGFP comets, but may represent protein aggregates or other artifacts, we selected only objects that conform to the average shape of EB1-EGFP comets in the image. The original image intensities $I_i(x,y)$ of all extracted objects were aligned and averaged into a template to obtain the mean EB1 comet image I_{EB1} in each frame of the time-lapse sequence. To exclude objects that are not a good fit to the average EB1-EGFP comet, we again used unimodal thresholding of the histogram of the sum of the squared intensity differences

$$LS = \sum_{i=1}^{n} (I_{EB1} - I_i(x,y))^2$$

between I_{EB1} and the intensity distribution in each region $I_i(x,y)$ (**Fig. 1 E**, right panel). The coefficient k_2 was introduced as a user-defined input to adjust the automatically detected threshold. An increased value for k_2 allows for a larger heterogeneity of EB1-EGFP comet shapes, and may be required in conditions strongly affecting microtubule growth.

Supplementary Note 4: Estimation of growth rate measurement error

For the purpose of this study we were primarily concerned about the variation in comet morphology which could introduce imprecision in growth rate measurements due to a variation of the distance between true and detected microtubule end. While the mean difference does not affect tracking quality, the variation of comet positions around the mean may introduce apparent changes in measured growth rates. The standard deviation we determined for this distance for highly elongated EB1-EGFP comets was in the order of ~1.5 pixel. To estimate the magnitude of this error we assumed the worst case that variation across a comet population equals the variation of an individual comet over time. The standard deviation of the displacement reflects the combined uncertainties of the manual identification of the microtubule end, of the noise-perturbed computationally-defined centroid position, and the error introduced by temporal changes of comet shape. Given a standard deviation of ~1.5 pixels and assuming 1 pixel for the uncertainty in manual localization of the microtubule end, the positional variation of the computer-detected centroid is also in the order of 1 pixel or less.

This uncertainty imposes limits on the resolution of instantaneous growth rates and growth rate changes between consecutive frames. At a pixel size of 64 nm and a sampling rate of 1-2 frames s^{-1}, a 1 pixel error yields an error in the instantaneous, i.e. frame-by-frame, growth rate of ~4-8 μm min^{-1}. In cases where time-averaging is affordable, we thus calculated growth rates rather as an average head-to-tail rate of the whole growth track. In this case, the precision in the growth rate is improved approximately by $N^{-1/2}$. N is the number of frames an individual track covers. As we consider tracks lasting 4 frames or more the expected error in the growth rate is less than 2-4 μm min^{-1}. Thus, for velocity measurements over several frames a localization uncertainty of 1 pixel has only a small effect on the analysis of microtubule dynamics. Of note, this performance could be improved further for data in which microtubules move slower or in cell systems with less dense microtubules, which would allow slower sampling.

Supplementary Note 5: EB1-EGFP Comet Tracking

Detected EB1-EGFP comets were tracked using a Kalman filter-based multi-object tracking algorithm[3]. In brief, this tracker uses a Kalman filter for each EB1-EGFP comet to project the continuation of its track into the next frame, given its history of motion. The assignment of the projected comet positions to the positions of detected comets in that frame is then accomplished by global minimization of the joint cost of all assignments. The cost of an individual assignment is defined by the Mahalanobis distance between projected and extracted comets. The uncertainty of the projection is defined by Kalman filter propagation of projection adjustments in the previous time-points of a track. In addition to the positional cues, we used the long axis of each comet to direct the projection. The comet eccentricity determined the level of certainty of this directional cue. This allowed us to reduce the notorious latency of Kalman filters in following abrupt changes in object motion, and allowed tracking of bending microtubules that changed direction (**Fig. 2b** and **Video 1**). To limit the combinatorial complexity of the track assignment, the user needs to define a search radius based on *a priori* knowledge of the maximal microtubule growth rates (**Supplementary Table 1**). An important validation of this setting is that the histogram of tracked velocities has no values near the allowed maximum microtubule growth rate (**Fig. 2c**).

Supplementary Note 6: Latency of EB1-EGFP comet formation

The latency Δt_{lat} for a pause event can be inferred from $v_g \Delta t_{lat} = v_{fwd} \Delta t_{fwd}$, which assumes that the entire distance a microtubule end translocates over the forward gap duration Δt_{fwd} is associated with the formation of a detectable EB1-EGFP comet, and that the microtubule grows with the average growth rate v_g during this period. Depending on imaging conditions, Δt_{lat} was between 2 and 3 s. A similar delay may of EB1-EGFP comet disappearance may also occur when a microtubule stops growing because of the time it takes for EB1 dissociation from the stalled microtubule.

Supplementary Note 7: Adjustments to hand-tracked trajectories

Hand-tracking relies on frame-to-frame measurements of microtubule end position, which overestimates both growth and shortening rates due to the uncertainty with which microtubule end positions can be manually defined. At 0.6 s frame^{-1} and 107 nm effective pixel size with a 60x objective, translocation of more than 1 pixel, which is the minimum that can be manually detected, corresponds to a minimal measurable rate of ~10 μm min^{-1}. This also results in significant overestimation of pause events by hand-tracking at such high frame rates. In contrast, the clustering algorithm defines a pause as EB1-EGFP dissociation from the microtubule end and subsequent re-association. Because of the latency of comet formation, most pauses inferred by track clustering have to be longer than 2–3 s, or 5–6 frames at our sampling rate.

For these reasons, we grouped series of growth or pause phases into longer segments if they were interrupted by 1 or 2 frames. Due to the latency of comet disappearance and reappearance, our tracking algorithm cannot detect such frequent changes. Moreover, the very short phases in hand-tracking are likely due to the inaccuracy of the coordinate definition, which is constrained by the pixel grid. We then calculated head-to-tail velocities of these corrected segments. We also eliminated shortening phases in the beginning and at the end of hand-tracked microtubules because

such events are not accessible to the clustering of EB1-EGFP growth tracks. Among these non-rescued shortening events, i.e. those at the end of hand-tracked microtubule trajectories, a small number (<5%) exhibited very fast velocities (up to 100 μm/min), which indicates that very fast shortening likely results in complete microtubule depolymerization.

Supplementary Note 8: Determination of thresholds for the selection of significant growth and shortening events in computer-generated growth tracks and backward gaps

The velocity distribution of forward gaps comprises both pauses with low values and a sub-population of fast gaps that most likely corresponds to growth events in which the EB1-EGFP comet is lost due to detection errors, occlusion, or temporary out-of-focus movement of the microtubule plus end. We separated these two event classes by unimodal thresholding[2]. Only events slower than the 9.3 μm/min threshold were considered pause events (**Fig. 3c**). Of note, the value 9.3 μm/min is close to 1 pixel per frame, v_{fwd} below this value is likely caused by comet localization errors. The remaining population of slow forward gap events had a pause speed of 4.6 μm/min. This rate corresponded well with the hand-tracked data sets (**Table 1**).

Backward gaps also contain pauses in addition to events of microtubule shortening. We assumed that noise-induced coordinate shifts are symmetric in forward and backward direction. Therefore, we used the same 9.3 μm/min threshold to separate pauses from true shortening events (**Fig. 3d**). As a result, the mean rate of computer-inferred shortening events increased from 17.6 μm/min to 24.3 μm/min, producing excellent agreement with the data from hand-tracked shortening events (**Table 1**).

The rate of ~12.5% of growth tracks was below the threshold of 9.3 μm/min. Although these events were molecularly marked as growth events by the presence of EB1-EGFP comets, in a hand-tracked data set they would be classified as pauses. Therefore, we eliminated these very slow events from the growth track population. This resulted in an increase of the mean rate of computer-inferred growth events from 16.8 μm/min to 18.7 μm/min, again producing excellent agreement with the data from hand-tracked growth events (**Table 1**).

Supplementary Note 9: Practical guidelines for applying the method

The software used for this paper can be downloaded from the website lccb.hms.harvard.edu, together with instruction for installation and use. The package served our labs to establish a proof-of-principle for the growth track clustering and thus is limited to primitive user-interface and data management tools. A more user friendly package relying on the core of this work, yet with much extended bioinformatics facilities for the analysis of microtubule polymerization dynamics in large data sets will be released in parallel. A description of the new features of this tool will be published elsewhere.

The quality of image sequences of cells expressing fluorescent fusions to microtubule plus-end tracking proteins varies significantly between cell lines, in dependence of the expression levels of the fusion protein, and between different perturbations of microtubule dynamics. Moreover, the image acquisition parameters may have to be adjusted to the speed level of microtubule growth and shortening, to the frequency of switching between the phases, and to the density of microtubules. This, in turn, may again affect the duration of the movies that can be collected and the signal-to-noise ratio. For these reasons, we emphasize that users will have to carefully validate on an

experiment-by-experiment basis the performance of the method on their data sets and possibly adjust control parameters (**Supplementary Table 1**). In this manuscript, we have outlined several procedures for validation, relying on; visual verification of detection, tracking, and clustering results against the raw image sequences of plus-end tracking markers; verification of the results against hand-tracked microtubule dynamics in movies with continuously-labeled tubulin in a second spectral channel; and quantification of parameter shifts in response to perturbations with known effects on microtubule dynamics. We have also outlined a schema for the conversion of computer-inferred rates of growth and shortening into hand-tracked rates of microtubule polymerization and depolymerization. The compound trajectories of multiple linked growth tracks permit the calculation of several probabilistic parameters that reflect the frequencies of catastrophe and rescue. These make use of the inferred times spent in growth, pause and shortening phases and the relative occurrence of forward and backward gaps. Importantly, these parameters introduce a novel metric for the dynamics of microtubule polymerization that relies on the presence and absence of a detectable plus-tip tracking protein that cannot be directly converted into conventional catastrophe and rescue frequencies as defined by classification of hand-tracked plus-tip movements. Ongoing work in our labs is directed towards establishing a bi-directional mapping between these two sets of parameters, in account of the kinetics of comet formation and dissolution. However, our experience with the here proposed method in several applications provides us with strong evidence that shifts in catastrophe and rescue frequency are robustly reproduced by shifts in the probabilistic parameters.

In the following we give guidelines for implementing a reliable assay of microtubule dynamics using the proposed method (see also **Supplementary Table 1** for a summary of control parameters and their sensitivity).

- *Detection:* This step is best validated visually. The detection parameters are robust, i.e. small variations in the parameter value have mild effects on the detection results. The quality of the detection is readily appreciated by eye, but can also be quantified with affordable effort as discussed in **Fig. 1**. Importantly, the validation should be repeated for frames at the beginning and the end of time-lapse sequences to account for the effect of photo-bleaching. If necessary, the analysis has to be limited to shorter movies. Overall, the detection is quite robust because of the unimodal thresholding for accepting pixels and comets (**Fig. 1c,e**) is applied frame-by-frame and thus it is adaptive to changes in the signal-to-noise ratio.
- *Tracking:* In images with a relatively low density of microtubules, low background fluorescence, and sufficiently fast sampling the growth tracking can be visually validated by overlay of growth tracks to the maximal intensity projections of the image sequence (**Fig. 2a,b**; **Supplementary Figs 4 – 6**). Under less optimal imaging conditions, the tracking may be validated visually for individual comets. However, quite often such image conditions tend to confuse the human eye as well, rendering the generation of a visual ground truth difficult. In either situation, it is highly recommended that for a new cell line and/or a new construct of a fluorescent protein fusion to a microtubule plus-end marker the tracking is validated by application of low concentrations of nocodazole and taxol, which both reduce the growth rate (**Fig. 4**). Ideally, the results are cross-validated between multiple plus-end markers (**Supplementary Fig. 2**).
  The key parameter for controlling the tracking is the search radius (**Supplementary Table 1**), which needs to be adjusted to the expected microtubule growth speed and the frame rate of the movie. A good quality control for the setting of the search radius is the histogram of

frame-by-frame displacements[6]. If the last 2 – 3 bins below the search radius contain significant data (~5% or more), then the search radius is too narrow, leading to truncation of the data. If an increased radius does not eliminate the truncation, then the sampling rate is too slow for the density of microtubule ends, i.e. the tracking algorithm will produce random links between comets, many of which are still near the limit set by the search radius. Of note, in our implementation the search radius is mostly relevant for the tracking of comets during the first 2 – 3 frames after initiation. After this period the Kalman filter locks in with a projection of the growing microtubule end between consecutive frames[3].

- *Growth track clustering:* Without doubt, the growth track clustering is the most delicate component of the proposed method and thus requires careful validation. High-confidence data will require a validation against continuously labeled microtubules (**Fig. 3**). However, these experiments are tedious. More practically, the settings of control parameters can be validated again by avoiding truncation effects in the velocity histograms of forward and backward gaps, as well as in their respective lifetime histograms.
  A second test for internal data consistency will utilize image sequences sampled at two-fold higher rate, generally, at the expense of shorter movie lengths. If the base sampling rate is sufficiently faster than the latency in comet formation and dissolution (so far, we have found the latency to be consistently 1 – 2 s over several cell lines), then the higher sampling should not significantly affect the two velocity distributions. Of note, a duplication of the sampling rate will require an extension of the maximal time window for linking; and it will likely affect the lifetime distributions of the forward and backward gaps. The confidence in the similarity of forward and backward gap velocities between sampling rates should be determined by a non-parametric test of the differences between the respective distributions.
  To gain confidence in the parameters of image acquisition and growth track clustering experiments with known stabilizers and destabilizers of microtubule dynamics are obligatory. As illustrated in **Fig. 4**, low concentrations of nocodazole and taxol are suitable and there are several microtubule-associated proteins with well-characterized effects on the polymerization dynamics.
  Finally, in a recent application of our software to an analysis of microtubule stabilizing tumor suppressor genes, we discovered that some microtubule-associated proteins (MAPs) alter the frequencies of catastrophe and rescue by shifting the rates of GTP-hydrolysis. In cells, this shift can be measured by the location and density of fluorescently-tagged antibodies that recognize specifically a conformational feature of GTP-bound tubulin in the lattice[7,8]. Thus, shifts in the catastrophe and rescue frequencies, reported by growth track clustering as shifts in the catastrophe probability (**Fig. 5h**) and shortening time (**Fig. 5j**), could be validated orthogonally by differences in the GTP-tubulin distribution. For several MAPs and MAP mutants we found that decreased catastrophe probabilities corresponded quantitatively with an increased number of microtubule ends with GTP-tubulin, while a decreased shortening time, i.e. higher probability for rescue, corresponded quantitatively with shorter distances between GTP-remnants along the lattice of microtubules[8].
- *Establishing the confidence level in parameter shifts between different experimental conditions:* A key question in the practical application of the proposed method is how small detectable shifts in a particular parameter of microtubule dynamics can be, especially in view of the high number of false negative events the method produces. It is not possible to provide strict theoretical bounds for the sensitivity as they depend on many experimental variables. In our applications we have relied on the reproducibility of results between cells

under identical experimental conditions to define the confidence levels of parameter changes between different experimental conditions, a test that is common practice in single cell experiments. The test does not only account for variation of the parameter outputs due to imaging noise and uncertainty in the detection, tracking, and linking models, but also due to molecular heterogeneity between cells. We recommend two statistical procedures to establish the level of confidence in parameter shifts, dependent on the level of cell-to-cell variation:

1) Because of the large number of measurements the method extracts from single cell movies, each parameter can be tested as to whether its intrinsic variation or the cell-to-cell variation dominates the total range of parameter values for a particular condition. Generally, none of the parameters is normally-distributed. Therefore, the parameter distributions from individual cells should be compared by a non-parametric test of the distribution shifts, such as the Kolmogoroff-Smirnov test. Systematic rejection of these tests in pairwise comparisons of cells will indicate that the parameter distributions significantly alter between cells. If the test is rejected for a specific subset of cells, one may consider eliminating these cells as outliers. If the majority of parameter distributions agree between cells the data can be pooled. This applied to all data presented in this paper. In this case the significance in parameter shifts between conditions can be determined using a permutation t-test (see Methods) that accounts for the non-normality of pooled parameter distributions and for their intrinsic spread within an experimental condition. The spread is partially due to the measurement uncertainty, but may also reflect systematic spatial variations in the microtubule dynamics within individual cells. Hence, a refined analysis with comparisons of specific sub-cellular regions may be appropriate (**Fig. 5f - j**).

2) If the majority of parameter distributions disagree between cells it is indicated that the data spread mostly originates in cell-to-cell variation. In this scenario the parameter shifts between conditions can be determined by a t-test of the distributions of parameter mean values (or any other characteristic descriptors) of individual cells. Of note, if the data from different cells can not be pooled it is critical to generate a representative sample of the cell-to-cell variation. As a rule of thumb, at least 12 – 15 cells per condition are required to apply a t-test of the means.

To conclude, the proposed statistical tests will not be suitable to detect apparent parameter shifts between experimental conditions with vastly different comet densities, potentially causing a differential bias in the tracking and clustering, especially if the respective control parameters are not adjusted. Therefore, before applying the proposed tests, it is critical that the confidence in the parameter extraction be established at the level of the image analysis, using approaches discussed with the validation of tracking and clustering.

Supplementary References

1. Thomann,D., Rines,D.R., Sorger,P.K., & Danuser,G. Automatic fluorescent tag detection in 3D with super-resolution: application to the analysis of chromosome movement. *J. Microsc.* **208**, 49-64 (2002).

2. Rosin,P.L. Unimodal Thresholding. *Pattern Recognition* **34**, 2083-2096 (2001).

3. Yang,G., Matov,A., & Danuser,G. Reliable Tracking of Large Scale Dense Antiparallel Particle Motion for Fluorescence Live Cell Imaging. *IEEE Conf. Computer Vision and Pattern Recognition*(2005).

4. Thoma,C.R. *et al.* pVHL and GSK3beta are components of a primary cilium-maintenance signalling network. *Nat. Cell Biol.* **9**, 588-595 (2007).

5. Howard,J. & Hyman,A.A. Growth, fluctuation and switching at microtubule plus ends. *Nat. Rev. Mol. Cell Biol.* **10**, 569-574 (2009).

6. Jaqaman K. and Danuser G. Computational image analysis of cellular dynamics: a case study based on particle tracking. *Cold Spring Harbor Protoc*. Eds Goldman R.D., Swedlow J.R., Spector D.L. (2009).

7. Dimitorov, A., *et al.* Detection of GTP-Tubulin Conformation in Vivo Reveals a Role for GTP Remnants in Microtubule Rescues. *Science*. 322:1353-1356 (2008).

8. Thoma C., Matov A., Gutbrodt K.L., Hoerner C.R., Krek W., Danuser G. Quantitative Image Analysis Identifies pVHL as a Key Regulator of Microtubule Dynamic Instability. *J. Cell Biology.* Accepted for publication.