Zoological Studies **49**(3): 289-303 (2010)

Zoological Studies

Symbolic Analysis of Plankton Swimming Trajectories: Case Study of *Strobilidium* sp. (Protista) Helical Walking under Various Food Conditions

Pieter Vandromme[1,2,3,4], François G. Schmitt[1,2,3], Sami Souissi[1,2,3], Edward J. Buskey[5], J. Rudi Strickler[6], Cheng-Han Wu[7], and Jiang-Shiou Hwang[7,*]

[1]*Univ Lille Nord de France, France*
[2]*USTL, LOG, F-62930 Wimereux, France*
[3]*CNRS, UMR 8187, F-62930 Wimereux, France*
[4]*Univ Paris 6, UMR 7093 LOV, Observatoire Océanologique, BP 28, 06234 Villefranche-sur-mer, France*
[5]*University of Texas at Austin, Marine Science Institute, 750 Channel View Drive, Port Aransas, TX 78373-5015 USA*
[6]*Great Lakes WATER Institute, University of Wisconsin - Milwaukee, 600 E. Greenfield Ave., Milwaukee, WI 53204-2944 USA*
[7]*Institute of Marine Biology, National Taiwan Ocean University, Keelung 202, Taiwan*

(Accepted August 20, 2009)

Pieter Vandromme, François G. Schmitt, Sami Souissi, Edward J. Buskey, J. Rudi Strickler, Cheng-Han Wu, and Jiang-Shiou Hwang (2010) Symbolic analysis of plankton swimming trajectories: case study of *Strobilidium* sp. (Protista) helical walking under various food conditions. *Zoological Studies* **49**(3): 289-303. The swimming behavior of the ciliate *Strobilidium* sp. was recorded using cinematographic techniques. A density of 20 ciliates/ml was used under 4 experimental food conditions: 121, 625, 3025, and 15,125 cells/ml of the dinoflagellate *Gymnodinium* sp. In total, 100 trajectories per experiment were recorded and analyzed. We classified this ciliate's swimming trajectories into categories we called "helix", "non-helix", and "break". These swimming states were identified using automated recognition of helices, based on values of swimming trajectory angles. We performed a symbolic analysis of the succession of swimming states which enabled discrimination between food concentration experiments, and provided a more-complete characterization of the swimming behavior. We found that helical swimming patterns first increased with food concentration then decreased with a corresponding increase in the numbers of breaks. Non-helical motions were related to high food concentrations. We further used these results to simulate a ciliate's trajectories using a symbolic dynamic model to generate a sequence series. Helices were reconstructed using a model with 2 inputs: amplitude and period. This study shows that a methodology developed to describe copepod behavior can also be applied to characterize and simulate ciliate helical and non-helical swimming dynamics. http://zoolstud.sinica.edu.tw/Journals/49.3/289.pdf

Key words: Protista, Plankton behavior, Swimming states, Symbolic dynamics, Simulation.

Protists are of major importance to marine ecosystems, especially to microbial food webs (Azam et al. 1983). In such webs, grazing is dominated by phagotrophic protists, yet these organisms are still understudied compared to bacteria and phytoplankton. Protists are ubiquitous grazers, their prey sizes vary from 2 μm flagellates to > 100 μm ciliates and dinoflagellates, and they exhibit a wide array of feeding behaviors (Tillman 2004). It is now well established that protists are a major source of predation on bacteria, and are important grazers of phytoplankton, as well as of other protists and sometimes metazoan eggs and small crustaceans (Sherr and Sherr 2002). In

*To whom correspondence and reprint requests should be addressed. Tel: 886-2-24622192 ext. 5304. Fax: 886-2-24629464.
E-mail:Jshwang@mail.ntou.edu.tw

turn, they constitute a significant source of food for metazooplankton, and contribute to regenerated production and to the export of biogenic carbon in the sea (Legendre and Le Fevre 1995, Mitra et al. 2003). The main grazer-prey interactions (encounter, selection, and capture) occur at individual levels mediated by sensory modalities of mechanoreception and chemoreception. As a result, interactions on micro-scale levels are of major importance, since they may affect the entire dynamic of microbial food webs and ultimately of marine ecosystems. A better understanding of such detailed interactions is now possible for ciliate-copepod interactions (e.g., Broglio 2001, Wu et al. 2009) and for ciliate-protist prey interactions (Jakobsen et al. 2006).

The feeding of ciliates is related to their swimming patterns, and the grazing rates of ciliated protozoan plankters may be larger than those obtained with an average homogeneous prey concentration, because of behavioral adaptations for remaining within areas with higher food concentrations (Buskey and Stoecker 1989). These adaptations are complex, and protists exhibit many different types of swimming patterns (Buskey et al. 1993). One characteristic pattern in protists is helical swimming, which also exists for a wide variety of organisms, including protists, ascidian larvae, and human spermatozoa (Macnab 1977, Buskey and Stoecker 1989, Crenshaw 1989, Strom and Buskey 1993, Crenshaw et al. 2000, Fenchel 2001, McHenry 2001, Farley 2002, Bartumeus et al. 2003, Christensen-Dalsgaard and Fenchel 2004). Fenchel (2001) and Machemer (2001) described the physiological and molecular basis for this swimming pattern. Other authors described it in a more-mathematical way (Buskey and Stoecker 1988 1989, Crenshaw 1989 1993a b, Crenshaw and Edelstein-Keshet 1993, Crenshaw et al. 2000, Bartumeus et al. 2002 2003, Menden-Deuer and Grünbaum 2006). Buskey and Stoecker (1988 1989) studied the swimming behavior of a tintinnid and characterized the swimming patterns using the mean velocity, the density distribution of velocities, the rate of changes in direction, and the net to gross displacement ratio (NGDR), and they observed a decrease in the mean velocity with increasing food concentrations. Crenshaw et al. (2000) characterized helical swimming patterns in 3 dimension (3D) by considering the geometry of the trajectories. With this approach, helices are characterized by their velocity, curvature, and torsion. Those authors developed an algorithm (finite helix fit: FHF) in order to measure these parameters and describe the helices. Other parameters are calculated using these 3 measures based on the trajectories. This algorithm was tested on simulated 3D helices, and the trajectories of flagellates, ciliates, spermatozoa of sea urchins, and ascidian larvae. Bartumeus et al. (2003) and Menden-Deuer and Grünbaum (2006) studied the swimming behavior of the flagellate *Oxyrrhis marina* using different food conditions. They concluded that this species has specific behaviors such as helical motion to exploit spatially structured resources, affecting both the population distribution of *O. marina* as well as their encounter rates with prey. Bartumeus et al. (2003) interpreted these trajectories as being Lévy walks. Lévy walks are considered "super-diffusive" compared to Brownian motion which is considered "diffusive" (Metzler and Klafter 2004). These 2 models describe random motion with the direction between 2 steps being random and the norm of the motion being issued from a Gaussian law in the case of the Brownian walk and in a Lévy law in the case of the Lévy walk.

Characterization of helical patterns and swimming statistics of microorganisms is important to understand the feeding relationships among protists in microbial food webs. We studied the swimming behavior of the ciliate *Strobilidium* sp. in the presence of increasing concentrations of its food, the flagellate *Gymnodinium* sp. Cinematographic and statistical techniques, first used for studying the swimming behavior of metazooplankton especially copepods and cladocerans (Buskey 1984, Schmitt and Seuront 2001 2002, Seuront et al. 2004a b, Schmitt et al. 2006, Moison et al. 2009, Wu et al. 2010), were used to analyze the behaviors. With this approach, we performed a statistical analysis of behavioral dynamics in order to quantitatively characterize behavior activity and perform stochastic simulations.

MATERIALS AND METHODS

Culture of microorganisms

We studied a protozoan species of the genus *Strobilidium* which belongs to the Spirotrichea class of the phylum Ciliophora. *Strobilidium* sp. was isolated from a 20 μm net tow taken from the University of Texas Marine Science Institute pier (27'48.40°N, 97'05.53°W) in July 2004. Ciliates

were cultured in 250 ml polycarbonate flasks with 0.2 μm filtered seawater (at a salinity of 32 psu) and fed a mixture of *Isochrysis galbana* (25%), *Rhodomonas* sp. (25%), and *Heterocapsa* sp. (50%) every other day. Ciliate cultures were grown in incubators set to 20°C with a 12:12 h light/dark photoperiod and transferred once a week. Food cultures were grown in F/2-Si media (Guillard 1975) under the same conditions and transferred monthly.

Experimental conditions

Experiments were performed on 30 July 2004 at the University of Texas at Austin, Marine Science Institute, Port Aransas, TX, USA. Samples of *Strobilidium* sp. were introduced into an experimental vessel of 3.75 (height) × 4 (length) × 1 cm (width) with a cover on the top, containing 15 ml of sample at a salinity of 30 psu. Videos were recorded in a dark room (to avoid phototrophism) at 22°C. The only source of light was a near-infrared LED of 1.45 V, with a peak wavelength of 910 nm. Densities of ciliates and prey (*Gymnodinium* sp.) before and after the experiments are presented in table 1. A low food concentration was used in experiment 1, medium in experiment 2, high in experiment 3, and extremely high in experiment 4. After transferring ciliates and their dinoflagellate food into the filming vessel, an acclimation period of 15 min was allowed prior to filming. Ciliate behavior was recorded using a SONY video camera (model XC-EI50; Tokyo, Japan) with an Angénieux F = 1:1.5, f = 50 mm lens (Paris, France), with a temporal resolution of 1/30 s and a spatial resolution of 60.6 μm.

Acquisition of trajectories

After filming, videos were analyzed using the TrackIt 2 computer program (Iguana Gurus, Milwaukee, WI). Before using this software, it was necessary to transform the movies into image sequences: this was performed using Adobe Premiere software (Adobe systems incorporated). The tracking software displayed images one by one, and by successively clicking on the position of the ciliate in each frame, a trajectory was numerically reconstructed. In total, 100 trajectories were extracted for each experiment. Each trajectory contained approximately 800 frames. In total, 3.2×10^5 frames were analyzed. A resolution of 30 frames/s produced about 3 h of ciliate swimming paths. Representative trajectories for each experiment are shown in figure 1. This digitalization was conducted at the Institute of Marine Biology, National Taiwan Ocean University, Keelung, Taiwan.

RESULTS

All analyses of ciliate swimming trajectories were performed with programs designed by the authors using MatLab (vers. 7.0.1, Mathworks) with statistics and curve-fitting toolboxes.

Global analysis based on swimming velocities

We first considered the velocity of organisms for each experiment. Table 2 shows the mean velocity, its standard deviation (SD), and maximum velocity for all 4 experiments (all values in this table are in μm/s). The mean velocity first increased from experiment 1 to experiment 2 (from 315 to 800 μm/s) and then decreased (to 370 μm/s). Maximum velocities were much higher: from 32.4 (in experiment 1) to 51.2 mm/s (in experiment 3). We then considered the probability density

Table 1. Experimental conditions. This table displays concentrations of *Strobilidium* sp. and *Gymnodinium* sp. before and after the 4 experiments

Food concentration	Density of ciliates (ind./ml)		Density of algae (cells/ml)		Concentration of food (μg C/L)		Exp.
	Before	After	Before	After	Before	After	
Low	20	12	121	113	10	9.3	1
Medium	20	10	625	369	50	29.5	2
High	20	19	3025	2041	250	168.7	3
Extremely high	20	19	15,125	7690	1250	635.5	4

functions (pdfs, i.e., the histogram where the integration from minus infinity to infinity is equal to 1) of velocities for the different experiments (Fig. 2). These pdfs are displayed as log-linear plots and compared to a Gaussian pdf with the same mean and SD: large velocities were much more frequent than with the Gaussian pdf, corresponding to so-called "heavy tails". We then considered the scale dependence of the velocity. We measured it by changing the resolution of the trajectory and estimating the NGDR (Buskey 1984) and the scale dependence of the mean velocity in order to check if the sampling time resolution had an influence. The NGDR is shown in figure 3A in a log-log plot. Scale dependence of the mean velocities is shown in figure 3B. If the sampling resolution is small enough, the velocity should not be scale-dependent. On the other hand, for a particle following Brownian motion, the mean small-scale position increment, $\overline{\Delta X_\tau} = \overline{|X_{t+\tau} - X_t|}$, depends on the time increment, τ, of the form

$$\overline{\Delta X_\tau} = \sqrt{D_{iff}}\,\tau^{1/2}; \quad \text{(1)}$$

where D_{iff} is the diffusivity (see e.g., Berg 1993). The velocity $V_\tau = \Delta X_\tau / \tau$ estimated with time resolution, τ, then has the following scale dependence

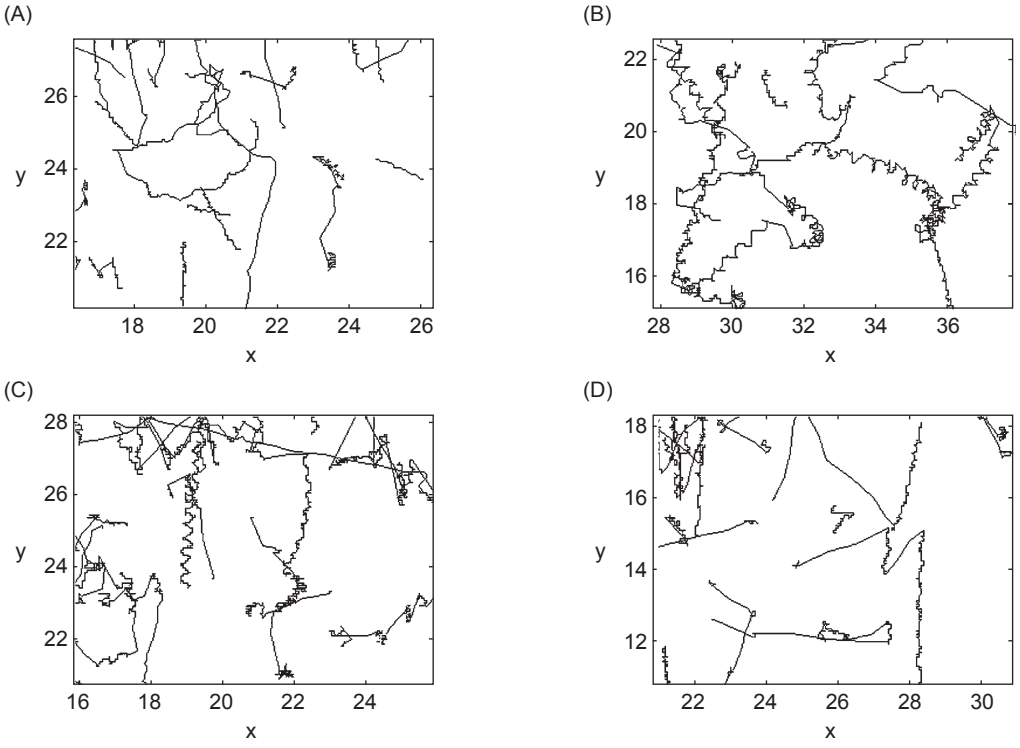

Fig. 1. Examples of trajectories recorded for all experiments. (A) Experiment 1 (low food); (B) experiment 2 (medium food); (C) experiment 3 (high food); and (D) experiment 4 (extremely high food). Positions x and y are in mm.

Table 2. Mean velocity, standard deviation, and maximum velocity during the 4 experiments. All values are in μm/s

Exp.	Mean velocity	Standard deviation	Maximum velocity
1	315	1320	32,370
2	800	2010	42,170
3	540	1860	51,200
4	370	1835	40,370

$$V_\tau \equiv \frac{\overline{\Delta X_\tau}}{\tau} = \sqrt{D_{iff}}\,\tau^{-1/2}; \quad \text{..... (2)}$$

which shows that the velocity increases for τ with smaller and smaller resolution. In order to emphasize such possible scale dependence and to compare it with Brownian motion, a log-log velocity plot is shown in figure 3B. In this figure, we show the velocity estimated with resolution τ, for all experiments (normalized in each case by the velocity at the smallest scale, i.e., the mean velocities in table 2) compared to the Brownian case, $V_\tau / V_{\tau_0} = (\tau_0/\tau)^{1/2}$, where τ_0 is the smallest scale. The faster this decrease, the more the trajectory is irregular. The curves obtained in figure 3A and B could be fitted by power-law functions. The NGDR at scale τ could be fitted for a large range of scales by a power-law relation of the form:

$$NGDR(\tau) = NGDR(T) \cdot \left(\frac{\tau}{T}\right)^{D-1}; \quad \text{..... (3)}$$

where T is the large scale and D is interpreted as a fractal dimension (Mandelbrot 1984). The same type of law can be written for the scale dependence of the velocity V_τ:

$$V\tau = V\tau_0 \left(\frac{\tau}{\tau_0}\right)^{1-D}; \quad \text{..... (4)}$$

In figure 3B, 2 groups are visible: velocities for experiments 2 and 3, and velocities for experiments 1 and 4. Fractal dimensions were $D \approx 1.21$ for experiments 2 and 3 and $D \approx 1.15$ for the others. These plots illustrate that in all cases, velocity and the NGDR did not reach a plateau at small scales. This means that a sampling rate of 30 frames/s is not large enough to accurately estimate the ciliate's true velocity. The latter is then expected to be in fact larger than our present estimate. Only for experiment 4 did it seem that a small-scale plateau was almost reached. The fractal dimension introduced above is an indicator of changes in the shapes of the trajectories. An important pattern that was visually identified was

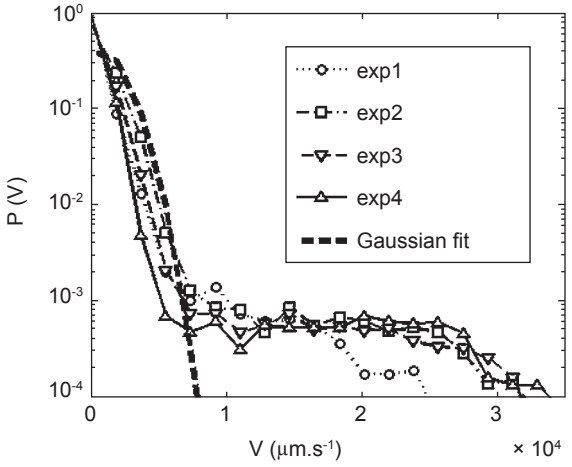

Fig. 2. Velocity probability density functions (pdfs) estimated for all 4 experiments. A Gaussian fit is presented for comparison. Heavy tails are visible: large velocities have a much larger pdf than a Gaussian distribution would predict.

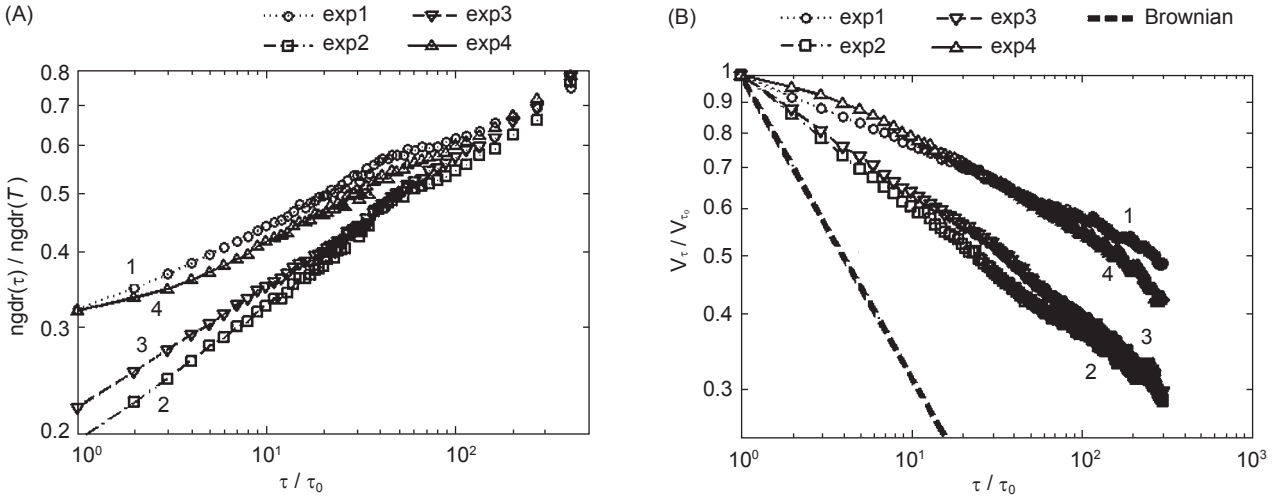

Fig. 3. Scale dependence of the net to gross displacement ratio (NGDR) (A) and velocity (B). For comparison, velocity curves were normalized to the velocity at the smallest scale (τ_0 = 1/30 s). The horizontal scale is the time resolution normalized to the smallest time resolution, τ_0 . The curves are presented as log-log plots to emphasize the power-law trends. A dotted line of equation $(\tau/\tau_0)^{-1/2}$ corresponding to Brownian motion is shown for comparison (B).

the helical pattern, which is important for these organisms. We then focused in the next step on automatically recognizing helices and the succession of swimming states.

Identification of helices

We propose here a method to clearly identify helices using the values of angles between 2 consecutive segments of a ciliate's swimming path. We first faced a problem of resolution: coordinates x and y are integers, so the smallest displacement recordable was 1 pixel. Since a pixel is about 60 μm, which is more than a ciliate's size, we observed jerky motion with a staircase appearance, providing a discretization effect with many angles of 90°. We then chose to smooth the trajectories with a running average using a window of 7/30 s corresponding to a running average involving 7 consecutive values, i.e., 3 steps before and 3 after the current position. This value is the minimum necessary to smooth small-scale discretized angles. An example of this smoothing is shown in figure 4, where we show visual helical trajectories before and after this treatment. After this, it became possible to perform angle analyses and use the angle estimation to characterize the helices. The angle distribution for each experiment and for 6 trajectories with observed important helical patterns are shown in figure 5. The main difference between the 2 cases was obtained for angles between 10° and 30°. This suggests that there was a large probability of these angles being associated with a helical pattern. On this basis, we proposed an algorithm to identify helices. For each point at time t, we first considered a window taking into account the last T points and the current point. We defined $N_1(t)$ as the number of angles, θ, between 10° and 30° in this window, and $N_2(t)$ as the number of angles between 40° and 100°:

$$N_1(t) = \sum_{i=t-T}^{t} \{\Theta_i \in [10°, 30°]\} \text{ and} \quad (5)$$

$$N_2(t) = \sum_{i=t-T}^{t} \{\Theta_i \in [40°, 100°]\} \quad (6)$$

Then we proposed the following helix criterion:

$$\text{if } a \cdot N_1(t) - b \cdot N_2(t) > c, \quad (7)$$

then the state at time t is called a helix, where T, a, b, and c are parameters that are determined below. If the conditions above are not verified, there are 2 possibilities: (i) it is impossible to calculate an angle, so the ciliate is in a "break" event, and (ii) the ciliate is not in a break event. In that case, it is moving, but the movement is not a helix: we defined this state as a non-helix.

With the above algorithm, we had a procedure to transform quantitative trajectory information (x and y positions) into a succession of symbols possessing 3 states: H (helix), B (break), and NH (non-helix). Because testing the helical pattern

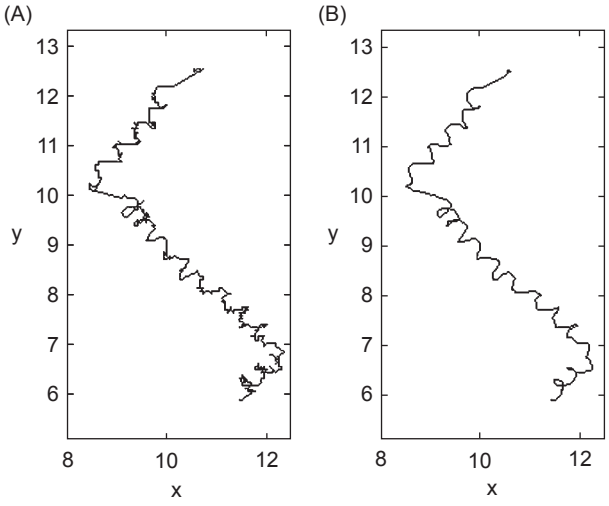

Fig. 4. Effects of smoothing on a trajectory. (A) Trajectory before smoothing and (B) after smoothing. The helical pattern of this trajectory clearly appears and enables the automatic recognition of helices. Positions x and y are in mm.

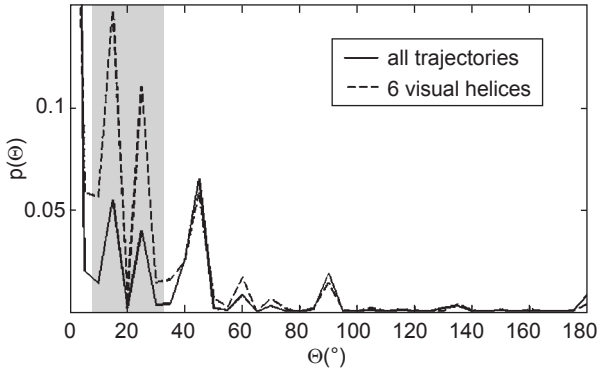

Fig. 5. Comparison of the probability density function (pdf) of the angle between consecutive segments for all trajectories (continuous curve) and for 6 extracted trajectories corresponding to visually identified helices (dotted curve). The main difference lies in the 10°-30° range (highlighted in light gray).

came first, we localized null velocities inside. This prevented the misidentification of short breaks between successive helices that could correspond to an effect of using 2D projections.

We parameterized the algorithm using simulated helices (with amplitudes and periods provided by the experiments) and simulations of persistent Brownian motion (Brownian motion with a constant direction) by changing the parameters *T*, *a*, *b*, and *c* to find the best ones, i.e., the ones which recognized the most helices and the least Brownian motion. The periods of the helices were provided by computing the Fourier spectral energy densities of the trajectories (Fig. 6). We found that periods fell within a range of 0.8-1.7 s (24/30-51/30 s). To estimate the amplitude, we smoothed the helices with a large window running average, and then by calculating the distance between this curve and the original curve, we found that amplitudes were close to 10 pixels. The parameters *T*, *a*, *b*, and *c* were estimated by exploring the parameter space and choosing the values that maximized the recognition of simulated helices and minimized the false recognition of persistent Brownian motion as a "helix". For $T = 25$, $a = 2.1$, $b = 0.9$, and $c = 3.4$, we found that 99.76% of the simulated helices were recognized as helices, and 1.99% of cases of persistent Brownian motion were recognized as helices (for 50 generated helices of 1000 points and the same for persistent Brownian motion). With non-persistent Brownian motion, about 10% of the helices were recognized. We noted that the parameterization gave a value of *T* which was close to the maximum half period of the helices. To validate this algorithm and check its robustness with respect to the noise level, we used Monte-Carlo-like degradation of the data, following the method of Bernard and Souissi (2004) applied to identify significant extrema of time-series data. A trajectory was simulated; then 1%-60% of the data were randomly eliminated. After that, the algorithm was applied to the degraded data. This procedure was repeated 2000 times, and statistics of helix identification were calculated (Table 3). We found that 87% of helices were still recognized with a degradation rate of 10%, emphasizing the good robustness of the recognition algorithm. We also tested the robustness of the algorithm adding Gaussian noise to simulate helices, as done by Crenshaw et al. (2000). Since the size of a ciliate is about 1 pixel, we chose Gaussian noise with a standard deviation equal to 'x' pixels, with x varying from 1 to 30 (Table 4). When noise was

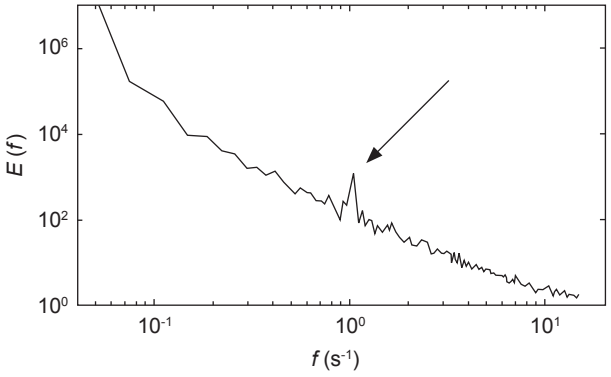

Fig. 6. Power spectrum on logarithmic scales of a trajectory visually corresponding to a helix. It was separately computed here on x and y coordinates and then summed. The bold arrow shows the peak corresponding to the frequency of the helical waves. This frequency of 1/s corresponds to a period of 1 s or 30 τ_0. Computed on other helices, this period varied from 0.8 to 1.7 s.

Table 3. Results of the test of the robustness of the algorithm to identify helices. One can see the percentage of part of the trajectories recognized as helices under different degradations of trajectories. This test was performed on 2000 trajectories of 1000 points each for every percentage of degradation

Percent (%) degradation	0	1	5	10	15	20	30	40	50	60
Percent (%) of helices recognized	99.8	99.5	96.3	87.1	73.8	59.4	33.6	15.9	6.5	2.4

Table 4. Results of the test of the robustness of the algorithm for identifying helices. One can see the percentage of part of the trajectories recognized as helices under different levels of noise addition. The noise was Gaussian noise with a standard deviation (SD) equal to 'x' times the size of the ciliate. This test was performed on 2000 trajectories of 1000 points each for every percent of degradation

SD of the noise (x)	0	1	1.5	2	3	4	5	10	20	30
Percent (%) of helices recognized	99.	99.4	96.4	87.0	56.6	36.2	26.9	10.2	8.5	7.4

added with an SD of 2 pixels, 87% of the helices were recognized, indicating the good robustness with noise addition. An example of the recorded trajectories recognized by this algorithm is given in figure 7.

Symbolic analysis and simulations

Symbolic analysis

Symbolic analysis is a method often used to study non-linear and chaotic phenomena, when classical statistics are not sufficient, by transforming the observed phenomena into a sequence of symbols and considering its dynamic properties (Ebeling and Nicolis 1992, Nicolis et al. 1997). We applied the algorithm presented above to all experimental trajectories belonging to experiments 1-4. With this procedure, trajectories were transformed into a series of swimming sequences, represented by the 3 symbols, H, B, and NH. We then studied the statistics and dynamics of the sequences of these symbols.

Several types of information can be extracted from symbolic data. We first considered the relative frequency in each state, represented as a percentage (Table 5). Experiment 2 with a medium food concentration contained 45% helices, while experiment 4 with a extremely high food concentration only contained 5% helices, which was less than non-persistent Brownian motion. The basic measure is the residence time probability density. Such an approach was previously undertaken by Buskey (1984) when considering histograms of residence times for copepod swimming states (histograms and probability densities differed by only a multiplicative factor). Other useful information was provided by the transition probabilities, i.e., the probability (denoted P_{ij}) to go from state i to state j. With this type of modeling, we characterized the symbolic dynamics by probability densities in each state, $q_i(x)$, and transition probabilities, P_{ij}. Such modeling was done to study copepod behavior using 2 states where the partition was based on swimming velocities (Schmitt et al. 2006) and more recently using 4 states of slow swimming, fast swimming, breaking, and grooming (Moison et al. 2009). Herein, the partition differed, since it was based on swimming types as determined by the above procedure. Figures 8A-L shows the $q_i(x)$ in all states, for all 4 experimental conditions. This was estimated using the following number of residence time events: 6489 for case 1; 6181 for case 2; 9522 for case 3; and 9364 for case 4. These residence times are represented in log-log plots, and some of them were fitted by a distribution emphasizing the hyperbolic tails in the form:

$$p(x) = p_0 x^{-(\mu+1)} \quad x \geq x_0 ; \quad \text{..... (8)}$$

where x_0 is a threshold, p_0 is a constant, and μ is an important parameter since it determines the relative strengths of all extreme events. We observed in figure 8 that such fitting was only possible for residence times in the helical state for experiments 1 and 2. For other pdfs, we considered exponential fits associated with a test of the Markovian nature of the symbolic dynamics (Dynkin 1965). Consider q_{ii}, the probability to be in state i at time $t+1$ knowing that at time t it was in the same state i. This is a 1 step conditional probability. For a Markovian process, there is

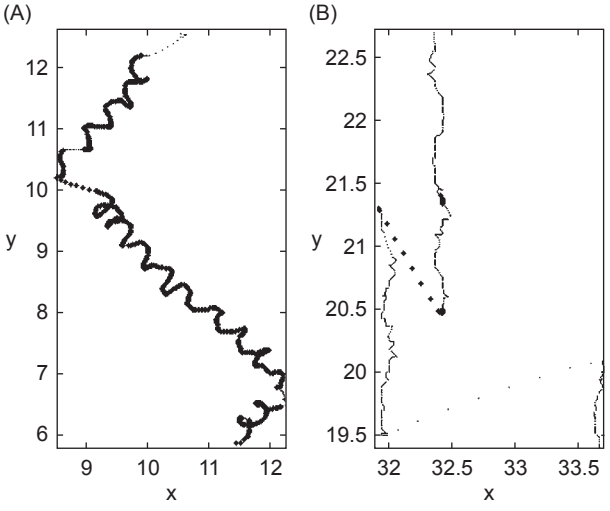

Fig. 7. Results of our helix identification algorithm, applied to 2 sample trajectories: a trajectory which could visually be identified as an helix (A) and a trajectory which could not (B). Parts of the trajectories recognized as helices are in black. Positions x and y are in mm.

Table 5. Percentage of time spent in each state as estimated from the experimental data

Exp.	Break	Helix	Non-helix
1	61%	13%	26%
2	27%	46%	27%
3	35%	17%	48%
4	48%	6%	46%

no memory, so q_{ii} is sufficient to estimate the residence probability, $q_i(x)$, corresponding to x consecutive values in the i-state:

$$q_i(x) = q_{ii}^{\,x} = e^{-x/T_0}, \quad (9)$$

with $T_0 = -1/\log(q_{ii}) > 0$. This relation corresponds to an exponential decrease in residence time probabilities with a characteristic time determined by q_{ii}. Values of q_{ii} can be directly estimated for each state (Table 6), and the resulting fits given by equation (9) are displayed in figure 8, together with the Pearson determination coefficients (R^2) between experimental distributions and exponential ones (more precisely between the log of experimental values and straight lines). One can see that in state B of experiments 2 and 4 and state H of experiments 1 and 2, experimental residence times were the most distant from a Markovian model.

Other important parameters in the symbolic analysis are transition probabilities, P_{ij}, that indicate to which state $j \neq i$ the ciliate swimming type symbolic dynamics is going when leaving state i (not to be confused with q_{ij}). These transitions are shown in table 7. We see that the food concentration also changed these probabilities. For example, if a ciliate in experiment 2 left state H, it had only a 16% probability of going to state B. In the same situation, a ciliate in experiment 1 had a 36% probability of going to state B.

This model can be summed up as follows

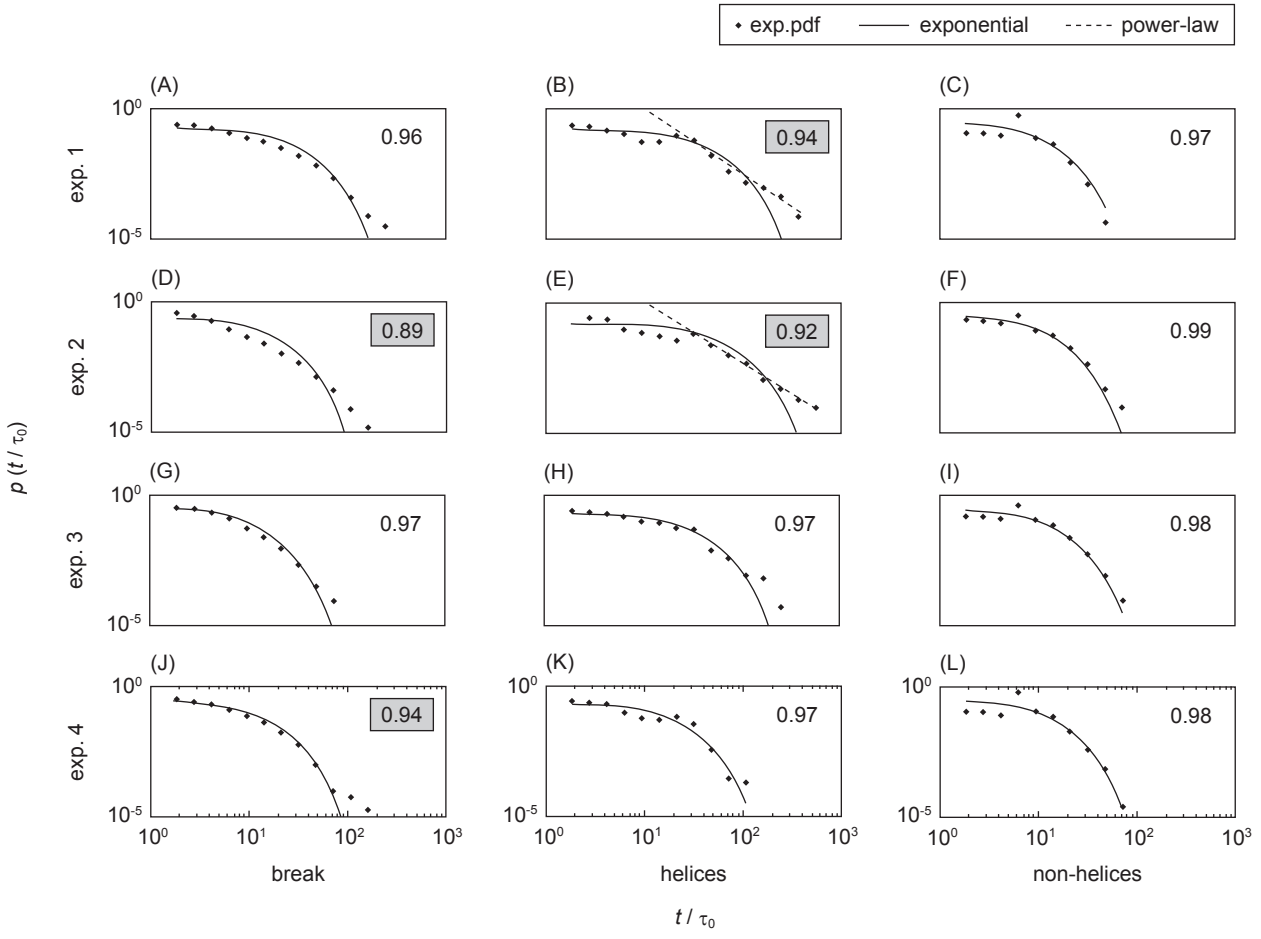

Fig. 8. Log-log plot of residence times of probability densities for all states in each experiment and comparison with the exponential (Markovian) model (gray lines) and power-law fits (black dotted lines) when possible. These densities are presented using a logarithmic binning (presented by Sims et al. 2007, as the best method to show the probability density function (pdf) of flight/residence times) of $1.5^{\tau/\tau_0}$. (A-C) Experiment 1; (D-F) experiment 2; (G-I) experiment 3; and (J-L) experiment 4. (A, D, G, J) The break state; (B, E, H, K) the helix state; and (C, F, I, L) the non-helix state. Pearson determination coefficients (R^2) between exponential and experimental pdfs are shown in the right upper part of each plot. The ones which were < 0.95 are highlighted. In (B) and (E), we adjusted a power-law (formula 8) for the distribution tails, and the exponent μ was 1.5 in case B and 1.3 in case E.

(Fig. 9): the symbolic dynamics are defined by a state and a residence time corresponding to this state. The fluxes between states are controlled by transition probabilities. The distribution of residence times and values of transition probabilities are the fundamental parameters of this symbolic analysis.

Simulations

In this section, we present simulations done using the following algorithm: (0) start from a given initial state i; (1) generate a random variable the pdf of which is close to the experimental pdf of residence times in state i; (2) generate a uniform random variable and, depending on its value, choose the next state j according to transition probabilities from state i to other states $j \neq i$; this determines the next state. Then return to step (1). This algorithm generates a succession of symbols according to a stochastic model. When there are only 2 states, such a model is reduced to an alternate renewal process (Feller 1971); see e.g., Schmitt et al. (1998) for an application to rainfall and Schmitt et al. (2006) for an application to copepod behavior.

To go from a succession of symbols to a trajectory, another set of hypotheses is needed. We considered here a simple choice: there are 3 states, two of which are associated with movements (the H and NH states). For the H state, we chose to reproduce deterministic helices with the amplitude and period given by random variables. These random variables were estimated only once for each helix sequence, according to the experimental pdfs estimated from the data. For the NH state, a simple random-walk approach was

Table 6. Probability q_{ii} (in percent) of staying in the same state during 2 consecutive times (1/30 s). q_{11} corresponds to the probability of staying in the break state, q_{22} of staying in the helix state, and q_{33} of staying in the non-helix state. Values are presented for all 4 experiments

exp.	q_{11}	q_{22}	q_{33}
1	94	96	85
2	90	97	86
3	86	95	88
4	89	92	88

Table 7. Transition probabilities (P_{ij}; in %) of going to state j after leaving state $i \neq j$ for all 4 experiments. States are break (B), helix (H), and non-helix (NH)

	Experiment 1			Experiment 2		
	B	H	NH	B	H	NH
B	-	36	89	-	16	68
H	2	-	11	3	-	32
NH	98	64	-	97	84	-

	Experiment 3			Experiment 4		
	B	H	NH	B	H	NH
B	-	22	86	-	24	93
H	1	-	14	1	-	7
NH	99	78	-	99	76	-

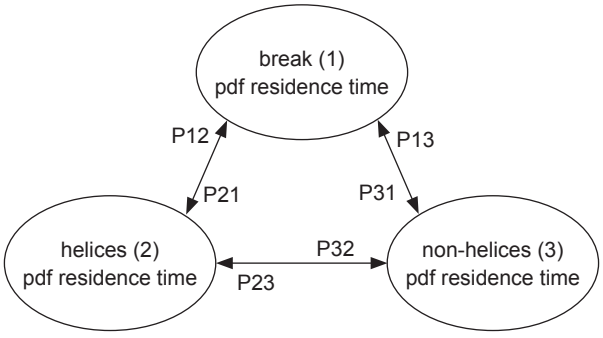

Fig. 9. Conceptual scheme of the symbolic analysis. Double arrows correspond to transition probabilities among the 3 states.

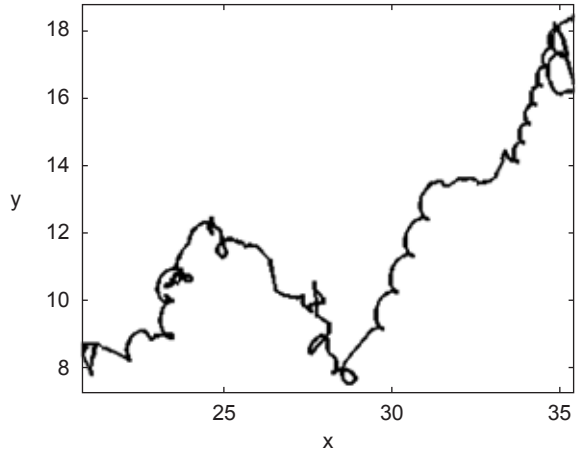

Fig. 10. Example of a symbolic trajectory using a model to reconstruct the helices. This model was associated with a period, an amplitude, and a direction to each helix. Positions x and y are in mm.

chosen: a velocity and an angle were randomly estimated for each time step and then generated from experimental pdfs.

To generate residence times, 2 different methods were adopted: (i) when the experimental pdf was close to the exponential (Fig. 8), the following generation formulae for exponential random variables were used (Evans et al. 2000): $T = (1/log(q_{ii})) \times log(R)$; where R is a uniform random variable between 0 and 1 and q_{ii} is an experimentally determined parameter; and (ii) when the experimental pdf was far from the exponential (state B for experiments 2 and 4 and state H for experiment 1 and 2), we adopted the rejection method using experimental pdfs (Devroye 1986).

A trajectory simulated using this algorithm is shown in figure 10. This trajectory is visually realistic. To quantitatively check the degree of realism of the simulations, we calculated the percentage of each state for each concentration of food (Table 8). A comparison of this table with table 5 indicates that the percentages of each state in simulated trajectories were close to the recorded ones. This was estimated using 100 trajectories of 1000 points for each food condition. This quantitatively showed that the algorithm which was chosen generates sufficiently realistic trajectories.

DISCUSSION

Methods

We considered ciliate behavior using different classical and non-classical tools. The NGDR is one of the classical measures widely used to characterize the behavior of zooplankters (Buskey 1984). Authors generally compute the NGDR at the smallest scale, but, as shown in figure 3A, this measure is scale dependent. To provide only 1 value for an NGDR is therefore not enough for comparison with other studies, since this indicator depends on the resolution, which might not be the same for each experiment (e.g., the resolution of the camera can differ). It is much more general to compute the NGDR at different scales and provide a more-general equation corresponding to the experimental curve.

Table 8. Percentage of each state simulated using results from the symbolic analysis for the 4 experiments. This table can be compared to table 5

Exp.	B	H	NH
1	62	12	26
2	25	44	30
3	37	15	48
4	47	6	47

Bartumeus et al. (2003) recorded ciliate trajectories using camera recordings and image analysis software to extract 2D trajectories. Their analysis was mainly based on tumbling angles at the smallest available scale. They chose 100° as the cut-off angle between helical paths and reorientation leaps. The hypothesis herein was to consider the helix state as fully characterized by the small-scale angle. This approach may also be considered a symbolic analysis: there are only 2 symbols, helix (H) when the angle is < 100°, and turning (T) when the latter is larger. They found that the probability density of the residence time in the helix state has a power law tail of the form $\tau^{-(\mu+1)}$ with μ = 1.1 or 1.2 depending on the experiment. They interpreted this as a "Lévy walk" (Metzler and Klafter 2004). This assumes that, during the helix state, the motion of the ciliate is deterministic with a constant velocity and direction. However, this is not realistic since the organism exhibits random behavior, with variable velocity and angles between 0 and 99°, which allows large freedom in wandering behavior. The Lévy walk model proposed in this paper therefore captures only 1 aspect of the angle symbolic dynamics, but the complexity of random displacements associated to angles in the range of 0°-100° was not taken into account.

In order to overcome these limitations, we used another approach. We proposed a partition into 3 states according to a non-local criterion. This way, we characterized the helices in a consistent way, and showed that the residence time probability density in the H state has a power-law tail with a slope of the form $\tau^{-(\mu+1)}$ with $\mu = 1.5 \pm 0.5$. This exponent is somewhat larger than the exponent recorded by Bartumeus et al. (2003), but we must note that the definitions of helix state were not identical. We do not propose here the term "Lévy walk" to characterize such animal behavior, for the following reasons: 1st, we partitioned the behavior into 3 symbols, whereas the Lévy walk theoretical framework corresponds to 2 states (break and constant velocity; see Shlesinger et al. 1987), and 2nd, during the helical state, the velocity is not constant.

The main motivation of our approach here was to propose a way to characterize the helical/other mixture of an animal's motion, which enabled the reconstruction of realistic synthetic trajectories. This symbolic analysis provided parameters that can be used to perform simulations (i.e., transitions probabilities and pdfs of residence times). Figure 10 and a quantitative analysis showed that the simulated trajectories were very close to observed ones. Performing realistic simulations can help validate an analysis, i.e., show that the main parameters of the process were identified and were able to reproduce it. This last point is very important in the context of individual-based models (e.g., Caparroy 2004). The software used by Caparroy (2004), called PASTIS, enables the reconstruction of a virtual reality in 3D by giving behavior to each numerical animal. The main objective of the present study was to test and validate some methods of analyzing swimming behavior, and provide some insights into the ecology of ciliates.

Ecological implications

The mean velocity estimated through 100 trajectories showed an interesting functional response in relation to the amount of food, i.e., an increase from experiments 1 to 2, and a decrease from experiments 2 to 4. This evolution is similar to that recorded by Buskey and Stoecker (1988 1989) for the tintinnid *Favella* sp., where velocity decreased with increasing food concentrations. Furthermore, we also observed a strong decrease in velocity at low food concentrations. The degradation of data (Fig. 3) discriminated experiments 2 and 3 from 1 and 4, and the latter 2 showed a less-ramified pattern. It seems to be approximately the same functional response as for the mean velocities: there were 2 groups, experiments 2 and 3 and experiments 1 and 4. Medium and high food concentrations (experiments 2 and 3) showed the same influence on swimming behavior but with higher values for the medium food level. Low and extremely high food concentrations (experiments 1 and 4) seemed to have the same effect on ciliate's behavior and were much more difficult to discriminate with these basic measures. We noted differences in the density distributions of velocities (Fig. 2), with experiment 4 showing more-frequent extreme values. The main differences among these experiments were obtained using more-advanced analyses, showing the importance of symbolic dynamics for the characterization of swimming states. The most important pattern which visually appeared was helical motion. Some authors (Buskey and Stoecker 1989, Crenshaw 1989 1993a b, Crenshaw and Edelstein-Keshet 1993, Crenshaw et al. 2000, Bartumeus et al. 2002 2003) showed the importance of this motion and its widespread distribution. To study this motion, we proposed a robust algorithm to automatically recognize it in 2D. This enabled us to statistically analyze it using symbolic tools. Table 4 shows the relative contribution of each state for each experiment. This also corresponds to a functional response. Experiment 4 showed the smallest helical motion percentage, followed by experiments 1, 3, and finally 2, which showed the largest helical motion percentage. This situation was the opposite to that of the B state. For the NH state, we had 2 groups: experiments 4 and 3 showed the largest percentage and experiments 1 and 2 showed the smallest. This percentage analysis discriminated all experiments. Experiment 1 was characterized by long breaks and few helical and non-helical motions, experiment 2 by few breaks and non-helical motions, and long times spent in helices, experiment 3 by medium times spent in helices and breaks, and a lot of time spent in non-helical motion, and experiment 4 by very few helices, quite-long breaks, and a lot of time spent in non-helical motion.

The proximity between experiments 1 and 4 obtained from previous analyses was not evident for the state residence statistics. The short time spent in non-helical motion seemed to reveal stress. Long rests, low complexity, and small mean velocities seemed to be caused by starvation stress in experiment 1, since the ciliates had to conserve energy. For experiment 4, the same symptom could be interpreted as being associated with satiety. In experiment 2, there was only a little time spent in non-helical motion; thus, there was some starvation stress, but less than in the 1st experiment. Ciliates performed many helices in order to scan their environment to enhance the encounter rate with *Gymnodinium* sp. Experiment 3 seemed to be halfway between experiments 2 and 4. The concentrations of ciliates before and after the experiments (Table 1) in experiments 1 and 2 had significantly decreased (from 20 to 10 ind./ml); while in experiments 3 and 4, the concentrations were similar (20 to 19 ind./ml). This high mortality observed in experiments 1 and 2 confirms that the animals were under stress.

The transformation of trajectories into

a sequence of states highlights some deep differences between experiments associated with dynamic analysis, whereas the percentage in each state is a statistical analysis. For state H, the process was the most distant from an exponential decrease (Markovian, Fig 8). The same pattern was seen for B states in experiments 2 and 4 but not for NH states, which were always close to an exponential decrease. Power-law probability densities for the B and H states revealed more-complex behavior associated with long-range memory and scale invariance. This type of pdf was associated with Lévy walks (Cole et al. 1995, Viswanathan et al. 1996, Harnos et al. 2000, Bartumeus et al. 2003, Ramos-Fernández et al. 2004, Brown and Liebovitch 2007, Sims et al. 2007), but as discussed in Schmitt et al. (2006), this is a much more-general property which is frequently found in animal behavior. Transition probabilities also revealed complex behavior. They were affected by both the food concentration and time spent in the previous state, and this influenced the dynamics of the process.

Our modeling approach involved a mixture of deterministic and stochastic frameworks. Helical patterns were mainly locally deterministic since they were associated with a rotating frequency and a translation velocity. This is a behavioral strategy which appears following an evolutive process. The cellular arrangement of protists makes it easy for them to perform such swimming motions (Machemer 2001). But the succession of swimming state sequences was still modeled in a stochastic framework. Such sequences may be associated with ciliate's responses to local stimuli; with the approach adopted here, we did not attempt to characterize each individual response, but instead the global response statistics.

In conclusion, we isolated helical motion from recorded trajectories and identified different states (break, helix, and non-helix). This allowed us to perform a symbolic analysis that provided characteristic parameters (percentage of each state, transition probabilities, and residence times), which discriminated the different experiments, and enabled, as a second step, stochastic simulations. All these behavioral parameters (and some other basic measures) highlighted the functional responses which may be widespread in protist behavior and of major importance in understanding their ecology. As highlighted by Sherr and Sherr (2002), predation by protists is a key to understanding fluxes inside microbial food webs. We need to understand processes at the individual (small) level and their consequences at higher scales (i.e., population) prior to developing predictive tools (Wu et al. 2010). Moreover, methods developed herein can be used to study widespread varieties of trajectories, from protists to humans and monkeys (Ramos-Fernández et al. 2004, Brown and Liebovitch 2007).

Acknowledgments: Infrastructure support was provided by the Univ. of Lille 1 and the French National Center for Scientific Research (CNRS). This work was partially financed within the bilateral CNRS-National Science Council of Taiwan (grant no. NSC96-2621-B-019-001 and NSC97-2621-B-019-001) cooperation project entitled "Analyses of zooplankton diversity using new cinematographical technologies, behaviour analyses and numerical modeling". An NSF grant (OCE-0452159) to Buskey and Strickler is acknowledged for supporting the experimental part of the work.

REFERENCES

Azam F, T Fenchel, JG Field, JS Graf, LA Meyer-Reil, F Thingstad. 1983. The ecological role of water-column microbes in the sea. Mar. Ecol.-Prog. Ser. **10:** 257-263.

Bartumeus F, J Catalan, UL Fulco, ML Lyra, GM Viswanathan. 2002. Optimizing the encounter rate in biological interactions: Lévy versus Brownian strategies. Phys. Rev. Lett. **88:** 097901.

Bartumeus F, F Peters, S Pueyo, C Marrase, J Catalan. 2003. Helical Lévy walks: adjusting searching statistics to resource availability in microzooplankton. Proc. Natl. Acad. Sci. USA **100:** 12771-12775.

Berg HC. 1993. Random walks in biology. Princeton, NJ: Princeton Univ. Press.

Bernard O, S Souissi. 2004. Identification of interactions in copepod populations using a qualitative study of stage-structured population models. *In* L Seuront, PG Strutton, eds. Handbook of scaling methods in aquatic ecology: measurement, analysis, simulation. Boca Raton, FL: CRC Press, pp. 361-379.

Broglio E, M Johansson, PR Jonsson. 2001. Trophic interactions between copepods and ciliates: effects of prey swimming behavior on predation risk. Mar. Ecol.-Prog. Ser. **220:** 179-186.

Brown CT, LS Liebovitch. 2007. Lévy flights in Dobe Ju/'hoansi foraging patterns. Hum. Ecol. **35:** 129-138.

Buskey EJ. 1984. Swimming pattern as an indicator of the roles of copepod sensory systems in the recognition of food. Mar. Biol. **79:** 165-175

Buskey EJ, CJ Coulter, SL Strom. 1993. Locomotory patterns of microzooplankton: potential effects on selectivity of larval fish. Bull. Mar. Sci. **53:** 29-43.

Buskey EJ, DK Stoecker. 1988. Locomotory patterns of the planktonic ciliate *Favella* sp.: adaptation for remaining within food patches. Bull. Mar. Sci. **43:** 783-796.

Buskey EJ, DK Stoecker. 1989. Behavioral responses of the marine tintinnid *Favella* sp. to phytoplankton: influence of chemical, mechanical and photic stimuli. J. Exp. Mar. Biol. Ecol. **132:** 1-16.

Caparroy P. 2004. Discrete events-based Lagrangian approach as a tool for modelling predator-prey interactions in the plankton. *In* L Seuront, PG Strutton, eds. Handbook of scaling methods in aquatic ecology: measurement, analysis, simulation. Boca Raton, FL: CRC Press, pp. 559-573.

Christensen-Dalsgaard KK, T Fenchel. 2004. Complex flagellar motions and swimming patterns of the flagellates *Paraphysomonas vestita* and *Pteridomonas danica*. Protist **155:** 79-87.

Cole BJ. 1995. Fractal time in animal behavior: the movement activity of Drosophila. Anim. Behav. **50:** 1317-1324.

Crenshaw HC. 1989. Kinematics of the helical motion of microorganisms capable of motion with four degrees of freedom. Biophys. J. **56:** 1029-1035.

Crenshaw HC. 1993a. Orientation by helical motion. I. Kinematics of the helical motion of microorganisms with up to six degrees of freedom. Bull. Math. Biol. **55:** 197-212.

Crenshaw HC. 1993b. Orientation by helical motion. III. Microorganisms can orient to stimuli by changing the direction of their rotational velocity. Bull. Math. Biol. **55:** 231-255.

Crenshaw HC, C Ciampaglio, M McHenry. 2000. Analysis of the three-dimensional trajectories of organisms: estimates of velocities, curvature and torsion from positional information. J. Exp. Biol. **203:** 961-982.

Crenshaw HC, L Edelstein-Keshet. 1993. Orientation by helical motion. II. Changing the direction of the axis of motion. Bull. Math. Biol. **55:** 213-230.

Devroye L. 1986. Non-uniform random variate generation. New York: Springer.

Dynkin EB. 1965. Markov processes. Vol. 122. New York: Academic Press.

Ebeling W, G Nicolis. 1992. Word frequency and entropy of symbolic sequences: a dynamical perspective. Chaos Solitons Fractals **2:** 635-650.

Evans M, N Hastings, B Peacock. 2000. Statistical distributions. 3rd ed. New York: J Wiley.

Farley GS. 2002. Helical nature of sperm swimming affects the fit of fertilization-kinetics models to empirical data. Biol. Bull. **203:** 51-57.

Feller W. 1971. An introduction to probability theory and its applications. Vols. 1 and 2. New York: J Wiley.

Fenchel T. 2001. How dinoflagellates swim. Protist **152:** 329-338.

Guillard RRL. 1975. Culture of phytoplankton for feeding marine invertebrates. *In* W Smith, MH Chaney, eds. Culture of marine invertebrate animals. New York: Plenum Press, pp. 26-60.

Harnos A, G Horváth, AB Lawrence, G Vattay. 2000. Scaling and intermittency in animal behavior. Physica A **286:** 312-320.

Jakobsen HH, LM Everett, SL Strom. 2006. Hydromechanical signaling between the ciliate *Mesodinium pulex* and motile protist prey. Aquat. Microb. Ecol. **44:** 197-206.

Legendre L, J Le Fevre. 1995. Microbial food webs and the export of biogenic carbon in oceans. Aquat. Microb. Ecol. **9:** 69-77.

Machemer H. 2001. The swimming cell and its world: structures and mechanisms of orientation in protists. Eur. J. Protistol. **37:** 3-14.

Macnab R. 1977. Bacterial flagella rotating in bundles: a study in helical geometry. Proc. Natl. Acad. Sci. USA **74:** 221-225.

Mandelbrot B. 1984. Les Objets Fractals. 2nd ed. Paris: Flammarion.

McHenry MJ. 2001. Mechanisms of helical swimming: asymmetries in the morphology, movement and mechanics of larvae of the ascidian *Distaplia occidentalis*. J. Exp. Biol. **204:** 2959-2973.

Menden-Deuer S, D Grünbaum. 2006. Individual foraging behaviors and population distributions of a planktonic predator aggregating to phytoplankton thin layers. Limnol. Oceanogr. **51:** 109-116.

Metzler R, J Klafter. 2004. The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. J. Phys. A-Math. Gen. **37:** R161-R208.

Mitra A, K Davidson, KJ Flynn. 2003. The influence of changes in predation rates on marine microbial predator/prey interactions: a modelling study. Acta Oecolog. **24:** S359-S367.

Moison M, FG Schmitt, S Souissi, L Seuront, JS Hwang. 2009. Symbolic dynamics and entropies of copepod behaviour under non-turbulent and turbulent conditions. J. Marine Syst. **77:** 388-396.

Nicolis C, W Ebeling, C Baraldi. 1997. Markov processes, dynamic entropies and the statistical prediction of mesoscale weather regimes. Tellus **49a:** 108-118.

Ramos-Fernández G, JL Mateos, O Miramontes, G Cocho, H Larralde, B Ayala-Orozco. 2004. Lévy walk patterns in the foraging movements of spider monkeys (*Ateles geoffroyi*). Behav. Ecol. Sociobiol. **55:** 223-230.

Schmitt FG, L Seuront. 2001. Multifractal random walk in copepod behavior. Physica A **301:** 375-396.

Schmitt FG, L Seuront. 2002. Diffusion anormale multifractale dans le comportement natatoire d'organismes marins. *In* Y Pomeau, R Ribotta, eds. Proc. 5th Rencontre du Non-linéaire. Paris: Non-linéaire Publications, pp. 237-242.

Schmitt FG, L Seuront, JS Hwang, S Souissi, LC Tseng. 2006. Scaling of swimming sequences in copepod behavior: data analysis and simulation. Physica A **364:** 287-296.

Schmitt FG, S Vannitsem, A Barbosa. 1998. Modeling of rainfall time series using two-state renewal processes and multifractals. J. Geophys. Res. **103:** 23181-23194.

Seuront L, JS Hwang, LC Tseng, FG Schmitt, S Souissi, CK Wong. 2004a. Individual variability in the swimming behavior of the sub-tropical copepod *Oncacea venusta* (Copepoda: Poecilostomatoida). Mar. Ecol.-Prog. Ser. **283:** 199-217.

Seuront L, FG Schmitt, MC Brewer, JR Strickler, S Souissi. 2004b. From random walk to multifractal random walk in zooplankton swimming behavior. Zool. Stud. **43:** 498-510.

Sherr E, B Sherr. 2002. Significance of predation by protists in aquatic microbial food webs. Antonie Leeuwenhoek **81:** 293-308.

Shlesinger MF, BJ West, J Klafter. 1987. Lévy dynamics of enhanced diffusion: application to turbulence. Phys. Rev. Lett. **58:** 1100-1103.

Sims DW, D Righton, JW Pitchford. 2007. Minimizing errors in identifying Lévy flight behavior of organisms. J. Anim. Ecol. **76:** 222-229.

Strom SL, EJ Buskey. 1993. Feeding, growth, and behavior

of the thecate heterotrophic dinoflagellate *Oblea rotunda*. Limnol. Oceanogr. **38:** 965-977.

Tillmann U. 2004. Interactions between planktonic microalgae and protozoan grazers. J. Eukaryot. Microbiol. **51:** 156-168.

Viswanathan GM, V Afanasyev, SV Buldyrev, EJ Murphy, PA Prince, HE Stanley. 1996. Lévy flight search patterns of wandering albatrosses. Nature **381:** 413-415.

Wu CH, HU Dahms, EJ Buskey, JR Strickler, JS Hwang. 2010. Behavioral interactions of the copepod *Temora turbinata* with potential ciliate prey. Zool. Stud. **49:** 157-168.