Title: A unicellular walker controlled by a microtubule-based finite state machine

Authors: Ben T. Larson[1], Jack Garbus[2], Jordan B. Pollack[2], and Wallace F. Marshall[1*]

Affiliations:

[1] Department of Biochemistry and Biophysics, University of California, San Francisco, California, USA

[2] Computer Science Department, Brandeis University, Waltham, Massachusetts, USA

[*] Corresponding author. Email: wallace.ucsf@gmail.com

Abstract: Cells are complex biochemical systems whose behavior emerges from interactions among myriad molecular components. Computation is often invoked as a general framework for navigating this cellular complexity. However, the manner in which cells might embody computational processes such that theories of computation, including finite state machine models, could be productively applied, remains to be seen. Here we demonstrate finite state machine-like processing embodied in cells using the walking behavior of *Euplotes eurystomus*, a ciliate that walks across surfaces using fourteen motile appendages (cirri). We found that cellular walking entails highly regulated transitions between a discrete set of gait states. The set of observed transitions decomposes into a small group of high-probability, temporally irreversible transitions forming a cycle and a large group of low-probability time-symmetric transitions, thus revealing stereotypy in sequential patterns of state transitions. Taken together, these findings implicate a finite state machine-like process. Cirri are connected by microtubule bundles, and we found that the dynamics of cirri involved in different state transitions are associated with the structure of the microtubule bundle system. Perturbative experiments revealed that the fibers mediate gait coordination, suggesting a mechanical basis of gait control.

One sentence summary: The ciliate *Euplotes* walks across surfaces with an unusual gait, which involves modular subcellular structure and combines elements of stereotypy and variability according to a computational process actively coordinated by a microtubule fiber system.

Main Text:

Cells are complex physical systems controlled by networks of signaling molecules. Single cells can display remarkably sophisticated, animal-like behaviors (*1–3*), orchestrating active processes far from thermodynamic equilibrium in order to properly carry out biological functions (*4*, *5*). Indeed, single cells can make decisions by sensing and responding to diverse cues and signals (*6*), execute coordinated movements (*7*, *8*) and directed motility (*9–12*), and even solve mazes (*13*, *14*) and possibly learn (*15–18*). Such behaviors in animals arise from neural activity and have been studied extensively, but we know comparatively little about the mechanisms of cellular behavior (*19*, *20*). In individual cells, behaviors emerge directly through the joint action of chemical reactions (*21*), cellular architecture (*3*), physical mechanisms and constraints within the cell (*22*, *23*), and interactions of the cell with its local environment (*24*). The links between information processing, decision making, and the physical manifestation of those processes as cell state transitions suggest that cellular behavior might be understood as an embodied computation (*25*, *26*). The theory of computation has often been invoked as a general framework for understanding cellular dynamics (*25*, *27–32*), environmental sensing by bacteria being a deeply studied example (*31–33*), and has been used to engineer programmable cell states (*34*). However, the manner and extent to which cells might embody functional, computational processes as well as the extent to which a computational perspective on cellular behavior might prove productive remains to be seen.

Among the microbial eukaryotes (protists), ciliates display some of the most striking examples of unicellular behavior, including hunting (*3*), sensorimotor navigation (*10*), and predator avoidance (*35*). Spirotrichous ciliates of the genus *Euplotes* are notable for their complex locomotion (*36–38*), using bundles of specialized cilia (cirri) to walk across surfaces (*36*, *37*) (Fig. 1A, Movies S1 and S2). Depending on the species, these cells generally have 14 to 15 ventral cirri arranged in a highly consistent pattern used for walking locomotion (*39*). *Euplotes* live in aquatic environments, and in addition to walking, use their cirri for swimming and rapid escape responses (*40*) (Movie S2). Oral membranelles (Fig. 1B) generate feeding currents to capture bacteria and small protistan prey and are also used for swimming. Early 20th century protistologists were so impressed by the apparent coordination of cirri that they proposed the existence of a rudimentary nervous system, the neuromotor apparatus, to account for their observations (*38*). This theory was motivated by the presence of intracellular fibers connecting various cirri (Fig. 1C), now known to be to be tubulin-based structures (*41*, *42*).

How can a single cell coordinate a gait without a nervous system? Although the walking movements of *Euplotes* are superficially similar to those of animals such as insects, the low Reynolds environment of aquatic microorganisms, where viscous forces dominate over inertial forces, imposes significant physical constraints on all movements that do not impinge on the movements of larger terrestrial animals (*43*). Coordination, to the extent that it exists in the gait of *Euplotes*, would require some kind of dynamical coupling among cirri or between cirri and some shared external influence. Recently, analytical techniques from statistical physics have been used to characterize, understand, and predict mesoscale dynamics in biological systems, including cellular behavior (*4*, *5*, *44*, *45*). These approaches rely on coarse-graining the complexity of biological dynamics into states and analyzing the transitions between states. In particular, a state representation allows us to ask whether forward and reverse transitions

between pairs of states are equal, a condition known as detailed balance (*4*, *46*). Systems that violate detailed balance operate in a non-equilibrium mode and can produce directed cycles in state space (*4*, *47*). Broken detailed balance has been observed in the motility dynamics of cultured mammalian cells as well as the motility dynamics of a freely behaving flagellate protist (*5*, *44*) and implies that non-equilibrium models are most applicable to such systems (*45*). Identification of broken detailed balance, therefore, highlights temporal irreversibility and can indicate active control of biological dynamics.

When information processing drives patterns of state transitions, such a system can be analyzed using automata theory, a fundamental level in the theory of computation (*48–50*). Automata theory can be used to address problems of decision-making and control in complex systems by providing predictive understanding that is independent of the underlying details of how a given process is implemented (*49*). Inspired by work considering cellular behavior in the context of the theory of computation (*25*), we hypothesized that walking cells might be governed by finite state automata with directed, processive movement arising from reproducible patterns of state transitions.

The consistent structure of *Euplotes*, its mode of motility, and its ease of observation make these cells an ideal biological test-bed in which to apply theories of non-equilibrium statistical mechanics and embodied computation, both of which rely on describing a system in terms of discrete state transitions. Here, we use time-lapse microscopy and quantitative analyses to show that *Euplotes eurystomus* walks with a cyclic stochastic gait displaying broken detailed balance and exhibiting elements of stereotypy and variability, in accord with a finite state automaton representation. The observed dynamics are reminiscent of behavioral regulation in some cells and animals (*5*, *51*) but contrast with many other well-characterized examples of cellular and organismal motility (*7*, *9*, *10*, *12*, *52–58*). Our results provide a clear demonstration

of machine-like processes governing cellular state transitions and serve as a framework for investigating the principles of behavioral control and non-equilibrium dynamics in single cells.

Results

A reduced state space is sufficient to describe walking dynamics

In order to ask whether cell behavior is governed by a finite state machine, we analyzed the walking behavior of *Euplotes eurystomus* cells, (*40*), focusing on the simplest case of uninterrupted, linear walking trajectories (Fig. 2A, B, Movie S1). Cells were placed onto coverslips on which free, spontaneous walking behavior was observed by microscopy (imaged at 33 frames/s). A focal plane at the cirrus-coverslip interface was chosen in order to clearly observe cirral dynamics (Fig. 2A). The consistency of the relative spatial positioning of cirri between cells allowed us to give each of the 14 cirri an alphabetic label from a-n (Fig. 2C). In each video frame, the walking state of the cell was encoded as a 14-bit binary vector, with each bit corresponding to a cirrus and receiving a value of "0" if the cirrus was in contact with the coverslip and stationary and a "1" if the cirrus was in motion or had moved in the preceding interval between frames (instances of stationary cirri held above the coverslip for a sustained period of time were not observed). The trajectories of 13 cells were manually tracked and annotated for a total of 2343 time points. This quantitative analysis revealed stepping-like cirral dynamics: cirri tend to undergo rapid movements followed by longer periods of quiescence (Fig. 2D). Cirral dynamics seemed to lack any obvious patterns such as periodicity or repeating sequences of states (e.g. Fig. 2D), implying that the state sequences are generated either by stochastic processes or complex deterministic mechanisms. This lack of periodicity (confirmed by autocorrelation analysis, Fig. S1) or fixed phase relationships between appendage movements is different from those reported for various unicellular flagellates and the gaits of most animals or (*58–60*).

Despite the apparent complexity of cirral dynamics, we wondered whether there might be some underlying structure, which would allow us to obtain a reduced state space that accurately described the dynamics, as has proven successful in behavioral analysis of diverse living systems (*44*, *45*, *61–64*). Dimensionality reduction by non-negative matrix factorization (NMF) (see Materials and Methods and Fig. S2 for more details) revealed the cirral states to be well-described by a three-dimensional NMF space (Fig. 2E-G). The constituent components of the NMF constitute the dimensions of the NMF space and correspond to non-overlapping groups of cirri. These groups constitute spatially distinct partitions of cirri with respect to their positions on the cell body (Fig. 3H). The dimensionality reduction of the gait state space arises at least in part from shared pairwise mutual information between groups of cirri (Fig. 3I).

We next sought to identify individual gait states. The density-based spatial clustering of applications with noise (DBSCAN) algorithm (*65*) was used to group our data into clusters in an unbiased fashion, with members of a given cluster sharing similar patterns of cirral activity. Visual inspection in conjunction with silhouette coefficient (*66*) (a metric of cluster cohesion and separation) analysis revealed that 32 clusters accurately captured the visible structure in the reduced state space without overfitting (Fig. 2E-G, see Materials and Methods and Fig. S3 for more details). These reduced gait states correspond to distinct patterns of cirral activity (Fig. 2J). Taken together, our results reveal stereotypy in the spatiotemporal patterns of cirral activity, consistent with the behavior of a finite state machine. The discrete set of gait states, which exist in the reduced gait state space, demonstrate that cells make use of a subset of the possible patterns of appendage movement during walking locomotion.

Euplotes walks with a cyclic stochastic gait

In order to relate the gait states identified in our cluster analysis, we asked how changes in the number of active cirri may relate to cell movement. At low Reynolds number, velocity is

proportional to the difference between the net force generated by the cell and the opposing drag (*43*), so cell velocity provides a relative readout of force generation. Naively, one might expect that the force associated with locomotion is roughly proportional to the number of moving appendages (*67*). Alternatively, the velocity might inversely correlate with the net change in cirral activity, which would be expected if stationary cirri were generating a pushing traction force as in crawling or climbing animals (*59*, *68*) or if cirri execute a power stroke just before coming to rest as has been suggested previously (*37*). However, we found that data supported neither expectation: cell velocity was only weakly correlated with number of active cirri (R^2=0.03), and instead, the largest cell velocities corresponded to small-to-moderate positive and negative changes in the number of active cirri (Fig. 3A). We concluded that transitions between gait states must be important to driving the forward progression of walking cells, and thus sought to determine whether such active coordination might manifest in the observed gait dynamics.

To search for evidence of active coordination, we calculated the forward and reverse transition frequencies between states from the 1423 pairwise transitions in our dataset (Fig. 3B, see Materials and Methods for more details). If gait state transitions involved active coordination, we expected to observe associated broken detailed balance (unbalanced forward and reverse transitions). The presence of strongly unbalanced transitions such as from gait state 3 to 17 versus 17 to 3 suggested broken detailed balance, and indeed, a number of forward and reverse transitions were found to be significantly unbalanced by the binomial test (see Materials and Methods). To further understand the degree to which detailed balance was broken, or, similarly, the distance from equilibrium, we calculated the entropy production rate (*5*). Following the procedure detailed by Wan and Goldstein (*5*), we obtained a lower bound estimate for an entropy production rate of 0.4, similar to the value reported for strongly non-equilibrium

gait transitions observed in a flagellate (*5*). Walking *Euplotes* cells, therefore, have a strongly non-equilibrium gait despite a lack of standard gait periodicity.

We next sought to better understand the organization of this unusual gait. First, we noted that gait state transitions are constrained: only 322 of the 1024 possible types of transitions were observed to occur at least once, and within this restricted set, only 173 occurred more than once (Fig. 3C). Also, we found that relatively few transitions corresponded to substantial cellular movement (Fig. 3C). Crucially, the presence of broken detailed balance revealed the existence of directed cycles of cirral activity during locomotion pushing the system away from equilibrium in spite of the lack of standard gait periodicity. To determine the nature of these cycles, we grouped transitions into two categories: balanced transitions, which satisfy detailed balance, and unbalanced transitions, which do not (see Materials and Methods for details). This partitioning allowed us to separately investigate unbalanced, non-equilibrium-like and balanced, equilibrium-like transitions (Materials and Methods, Fig. 3D, E). We found that unbalanced transitions occur at high frequency but only involve a small number of states (Fig. 3D). Of the 32 gait states, only states 2, 3, 4, 7, 17, and 27 are associated with unbalanced transitions, and among these states, three formed a biased, directed cycle following 2→3→17→2. We had expected that unbalanced transitions might be associated with a "power stroke" (in the sense of occurring simultaneously with cell movement) but found instead that high cellular velocities tend to be associated with balanced transitions (Fig. 3D, E). Additionally, we found that, with the exception of transitions between states 1 and 2, transitions occurring at the highest frequencies are unbalanced (Fig. 3D).

Notably, the most frequent balanced transitions were associated with transitions into and out of gait state 1, a unique "rest state" which involves no cirral movement (Fig. 2J). Furthermore, we found by autocorrelation analysis of gait trajectories that gait state 1 uniquely exhibited significant positive autocorrelation (see Materials and Methods, Fig. S4). Investigation

of transitions to and from gait state 1 revealed that although transitions between gait states 1 and 2 are balanced, the most frequent transitions out of gait state 2 are strongly biased toward transitions into gait state 3, from which other strongly biased transitions also frequently occur, including the cycle of biased transitions mentioned above. The presence of high frequency unbalanced transitions does not preclude the existence of highly variable trajectories through state space. The picture of walking trajectories that emerges is one of stochastic excursions from gait state 1 into non-determinate paths through state space involving a mix of balanced and unbalanced transitions. The majority of cell movement occurs during infrequent, equilibrium-like (balanced) transitions. Temporal irreversibility or directedness in the gait of *Euplotes* arises from biased, non-equilibrium-like (unbalanced) transitions, occurring at relatively high frequency from a small subset of states.

Finally, we checked whether gait state transitions were a Markov process, often referred to as "memoryless" (*69*), which entails that transition probabilities are determined completely by the present state, and that previous dynamics contribute no additional predictive information (*70*, *71*). When we compared the gait transition matrix (Fig. 3B) with a computed transition matrix over windows of two adjacent transitions, we observed these matrices to be different from one another, which is inconsistent with the Markov property (see Fig. S5). Further, when we applied a Billingsley test, a chi-squared measure for Markovness (*72*), the null hypothesis that the process was Markov was rejected ($p=0.005$). These analyses showed that *Euplotes* retains some "memory" of the prior sequence of cirral movements during locomotion.

Taken together, our analysis revealed a mixture of unbalanced transitions driving cycles and balanced transitions arranged as networks, for which we propose to apply the term "cyclic stochastic gait". The cyclic stochastic gait of *Euplotes eurystomus* incorporates elements of both stereotypy and variability in gait dynamics, in terms of biased transitions and non-determinate

sequences of gait state transitions respectively. Forward progress of the cell is not produced merely by a physical ratchetting process driven by unpatterned fluctuations in cirral activity, nor is it produced by a highly regular, deterministic process like a clock. It has been argued that significant computation arises in physical systems exhibiting such a mix of stereotypy and variability (*48*, *73–75*) in the sense that the time-evolution of the system is most compactly described by the result of a computation involving state transitions, memory, and decision rules, rather than by a periodic oscillation or a random coin flip.

While our analysis revealed a computational underpinning of gait, we sought to better understand the functional organization of the dynamical patterns driving processive motion of the cell. To do so, we first focused on the highest transition probabilities emanating from each state. Transition probabilities were estimated as N_{ij}/N_i where N_{ij} is the number of transitions from state i to state j, and N_i is the total number of transitions from state i. This allowed us to prune away rare transitions in order to reveal the dominant structure of gait state transitions. Figure 3G displays the pruned transition matrix as a heatmap. We found that relatively few states were the recipients of the majority of high probability transitions, and many states received none. To more clearly visualize the structure of transitions, we grouped together all gait states receiving no more than one unique high probability transition based on the idea that state transitions into this group show little bias in terms of source state, and within the group, transitions between states exhibit low probability, time unbiased, equilibrium-like fluctuations. the majority of cell movement was associated with transitions between states within this group. In contrast to the “cloud” of states linked by low-probability, balanced transitions, nearly all of the states receiving high probability transitions were either the three “cycle” states, or else fed cycle states with their highest probability transitions, with the majority feeding gait state 17.

Focusing on the structure of transitions in this way allowed us to derive a simplified, functional representation of stereotypy in gait dynamics as depicted in Fig. 3H. Although gait state 1 is not the recipient of any individual high probability transitions, we identified it as the unique "start" state from which cells initiate walking. Beginning with this start state, cells transition with high probability to gait state 2, also one of the highest frequency transitions and the first state in the 2→3→17→2 cycle of unbalanced transitions. From this first cycle state, cells transition to gait state 3, the second cycle state, with highest probability and frequency and then similarly on to gait state 17, the third cycle state. This sequence from the start state through the cycle states corresponds to increasing amounts of cirral activity. Although the highest probability transitions from the third cycle state to any single gait state return to the first and second cycle state with equal probability and return to the first cycle state also being unbalanced, cells in fact transition to the equilibrium "cloud" of motility-associated states with overall higher probability. Return to the cycle states tend to occur through various moderately high probability transitions from the motility state cloud or through a restricted set of intermediate states. In conjunction with this set of transitions, we also noted unbalanced transitions stemming from the cycle states to the motility state as well as intermediate state subsequently feeding the next cycle state.

Altogether, the picture that emerges of stereotypical gait dynamics is of biased transitions through cycle states before relatively low probability, unbiased transitions associated with substantial cell movement before returning to the start or cycle states and beginning the sequence again. While this general sequence is repeated during walking, there is variability or apparent stochasticity in the details of gait state transitions with increasingly variable transitions as any given sequence progresses. We propose that the cycle states serve to establish configurations of cirri necessary for cells to later transition to between states from which forward progress of the

cell is generated. Many state transitions along any instance of the stereotyped sequence are unbiased; however, biased, high probability transitions, presumably resulting from active cellular control, give temporal irreversibility to the sequence.

The microtubule fiber system of Euplotes mediates gait coordination

What physical machinery could embody the information processing required to generate the stochastic cyclic state transitions seen during *Euplotes*' walking? We reasoned that there must be some form of coupling or communication between cirri or feedback between gait state and cirral dynamics. Since the early 1900s , the role of the system of cytoskeletal fibers associated with cirri as conduits of information between cirri during cellular locomotion, supported by microsurgical experiments, has been a key hypothesized mechanism of gait coordination (*76*, *77*). We wondered whether the structure of the cytoskeletal fiber system associated with cirri (Fig. 4A) could give some insight into how cirri might be coordinated.

We reconstructed in 3D the tubulin-based fiber system of *Euplotes* associated with cirri and lying just beneath the cell cortex (*38*, *41*, *42*). Upon inspection of our confocal reconstructions of SiR-tubulin labeled cells (Fig. 4A, Fig. S6), we noted the presence of two morphologically distinct classes of fibers, one thicker, linear class and the other thinner, splayed, and less linear, consistent with previous observations (Fig. 1C, (*38*, *41*, *42*)). Additionally, we observed a group of thick linear fibers associated with some of the frontal cirri, which to our knowledge has not been previously reported. Fibers emanate from the base of all cirri, appear to intersect one another, and also connect to the cortex of the cell at various points. Some cirri were found to be associated with only thick fibers, while others have both or only thin fibers. Based on apparent fiber intersections and convergences, we found the fiber system forms a continuous network between all cirri, with the fibers associated with the base of each cirrus intersecting the fiber system of at least one other cirrus (Fig. 4A, B).

Contrary to the long-standing standing hypothesis from the literature (*76*), the functional modules (groups of co-varying cirri) identified in our dynamical analysis were not exclusively linked by dense fiber intersections (Fig. 4A, B) (*38*, *42*, *76*). In fact, connections between cirri are not generally associated with any statistically significant difference in mutual information (defined in terms of the information that the activation state of one cirrus has concerning the other) compared to unlinked pairs of cirri ($p=0.14$ by Wilcoxon rank sum test, Fig. 4C). However, information flow became apparent when fiber-fiber links were grouped by type (i.e. thick to thick fiber, thick to thin fiber, or thin to thin fiber). Under this grouping, we found that pairs of cirri associated with only thick to thick fiber and only thin to thin fiber links have increased mutual information compared to those without links (Fig. S7). Interestingly, we found that cirri nearby one another and connected by fibers to similar regions of the cell cortex shared the most mutual information (Fig. 2I, 3C, 4D, E), suggesting that if the fibers play a role in cirral coordination, coupling may also be mediated by mechanisms involving the cirrus and fiber-cortex interface. Cirri d, e, h, i, for example, share very little mutual information with any of the other cirri, and fibers emanating from the base of these cirri contact the cell cortex and other fibers at various unique points. On the other hand, cirri g and f, which share more mutual information than any other pair, are associated with both thick and thin fibers terminating at similar regions of the cell cortex. Indeed, distances between pairs of cirri and cross nearest-neighbor distance (*78*) between paired sets of cirrus-cortex contact points both show significant Spearman correlations (-0.49, $p<0.001$ and -0.62, $p<0.001$ respectively) to mutual information (Fig. 4D, E). These correlations indicate that mutual information between pairs of cirri tends to increase with proximity and also tends to increase with similarity between fiber-cortex contact locations. Thus, the cirri with the highest mutual information are those that are close together with similar fiber-cortex connections (Fig. 4D-F).

Together, these observations suggest a mechanism of mechanical coordination in which microtubule bundles allow groups of cirri to influence successive behavior of other groups of cilia. We sought to test this hypothesized mechanism by perturbing the fiber system using drug treatments. We observed that nocodazole, a drug that blocks the polymerization of microtubules, affected walking motility, causing cells to walk along more compact trajectories due to increased turning (Fig. 4G, H, Movie S3, S4) in a manner reminiscent of altered walking due to microdissection of fibers (*76*). Quantifying this effect from video microscopy of cells under darkfield illumination in terms of a scaled path length defined as the total integrated path length walked by cells scaled by the maximum radial distance traversed, where a decrease in linear runs decreases this scaled path length, we found that nocodazole significantly decreased the scaled path length of cells (Fig. 4I) compared to controls. Next, we checked whether nocodazole treatment specifically affected the fiber system. By analyzing 3D reconstructions of confocal images of SiR-tubulin labeled cells, we found that fiber length significantly decreased compared to controls in cells where microtubule polymerization was disrupted by nocodazole (Fig. 4J-L). Further, we were unable to detect the presence of thin fibers in four out of seven nocodazole treated cells used for fiber analysis. Of the three cells with detectable thin fibers, we never observed connections between fibers associated with cirri a, b, and c and any other fibers. Additionally, we did not observe any thin fibers making distal cortical contacts. We did note, however, that when thin fibers were visible, connections appeared to be the same as those in Fig. 4A and B. We then investigated how cirral dynamics were affected. Following the gait annotation procedure detailed previously, we characterized the walking dynamics of 6 nocodazole treated cells for a total of 1133 timepoints. Of those timepoints, 681 corresponded to cirral configurations never observed in untreated cells with a total of 391 new unique configurations. Projecting these new configurations onto the NMF space we obtained previously,

however, revealed that most of the cirral configurations in nocodazole treated cells were near or within the clusters corresponding to the gait states we obtained from untreated cells (Fig. S8A). This allowed us to map new cirral configurations onto the gait states (see Methods for details). We did note that while close to the original gait states, new cirral configurations tended to skew towards more active cirri, and we also noted the presence of a new cluster involving movement in all or nearly all cirri, to which we assigned a new gait state (Fig. S8A). Consistent with this bias toward increased cirral activity, we also found that mutual information between cirri was higher in general, with many pairs of cirri sharing higher mutual information than the highest values obtain in untreated cells (Fig. S8B). This increase in cirral activity and correlations is consistent with the fibers playing a role in conveying inhibitory information during unperturbed walking.

Finally, we investigated how the dynamics of gait state transitions were affected. Following our previous analysis, we evaluated gait state transitions for the presence of broken detailed balance, separating state transitions into balanced and unbalanced transitions and found that the structure of state transitions differed greatly from that of unperturbed cells. Although some gait state such as 1, 2, and 17 were involved in high frequency transitions in both conditions, gait state transitions of cells with perturbed fiber systems exhibited much less broken detailed balance, including the loss of the unbalanced, cyclic transitions (Fig. S8C). Figure 4M summarizes the change in the structure of gait state transitions including changes in broken detailed balance, reduction in transitions toward the states previously involved in cyclic transitions, and the reduction in the occupancy of states associated with the cloud of states involved in substantial cell movement, with only 10 of the original 21 cloud states sampled. We also noted that many of the new highest probability transitions feed gait state 18 (Fig. 4M, S8D), which involves nearly exclusive activity in cirri f and g located close to one another at the edge

of the cell. A persistent bias toward activation of these cirri, which are farthest from the central axis of the cell, may in part explain the increase in turning in trajectories. Taken together, these results are consistent with our hypothesis that the fiber system mechanically mediates gait coordination. Further, these results provide additional evidence that proper, active coordination of cirral dynamics and emergent gait states and gait state transitions is required for proper cellular walking.

Discussion

Faced with the challenge of accounting for the emergence of apparently sophisticated cellular behavior, the directed yet seemingly disordered gait of walking *Euplotes*, we conceptualized the cell as a finite state machine to illuminate underlying principles and mechanisms. Traditionally, studies of computational processes performed by cells have focused primarily on combinatorial logic, where the output of a computational process depends only on the current input, performed by networks of molecules in cells (*25*, *31–33*). We have focused on sequential logic, where outputs depend on the system state as well, an equally important aspect of the theory of computation with notable yet less developed representation in studies of cellular and sub-cellular dynamics (*16*, *20*, *79*). Automata theory, which includes finite state machine models and necessarily involves sequential logic, provides tools for understanding structure and stereotypy in transitions between dynamical states, features that are increasingly appreciated in the behavior of eukaryotic cells. Our approach, leveraging dimensionality reduction to identify dynamical states, revealed modularity in cellular dynamics associated with structural modularity of the cell (Fig. 2, 4) in addition to cyclic patterns of sequential dynamical activity (Fig. 3).

Although there are examples of locomotion reminiscent of the stochastic, non-equilibrium gait dynamics of *Euplotes*, such as gait switching in Kangaroo rats (*51*) or, most saliently, gait switching in an octoflagellate (*5*) and motility in cultured mammalian cells (*44*),

walking locomotion in *Euplotes* represents a departure from many of the best studied appendage-based locomotor systems. For example, limbed locomotion in animals tends to proceed by highly stereotyped, determinate patterns of activity (*58*, *59*), and many small, aquatic animals exhibit periodic movements of appendages, often cilia, during locomotion (*7*, *60*, *80*). Many forms of unicellular locomotion involve such dynamics as well including in sperm cells (*81*), diverse flagellates with various numbers of flagella (*60*), and ciliates (*60*, *82*, *83*). Even in cases where cellular locomotion involves fundamentally stochastic dynamics such as in run-and-tumble motility in *E. coli* (*12*) or analogous behaviors observed in protists (*11*, *84–86*), motility can be described by equilibrium processes (*5*), in contrast to the non-equilibrium character of the gait of *Euplotes*. We propose that broken detailed balance in the gait of *Euplotes* indicates active coordination of motility processes.

Here, broken detailed balance in gait state transitions of *Euplotes* revealed cyclic activity, characterized by transitions into and out of a resting state with a mixture of stereotypy and variability in the intervening steps, in the gait of a single cell (Fig. 3, 4). To explain how these dynamics give rise to directed walking, we propose a mechanism in which biased, actively controlled cyclic transitions serve to establish strain, effectively storing stress, in certain cirri, and the spontaneous release of these cirri from the substrate, during a series of unbiased gait state transitions, allows the cell to move forward. The cloud of unbiased transitions associated with substantial cellular movement is consistent with motility generation not depending on the precise order in which the strained cirri are released from the substrate. Return to the cycle states then are necessary to establish this process anew by winding up the system for continued, proper cell movement. Disruption in this resetting may lead to defects in walking such as we observed with the loss of the cycle states in irregularly walking cells with disrupted fiber systems, which appeared unable to consistently maintain linear trajectories (Fig. 4G-M, Movie S4). We find

additional experimental support for our proposed mechanism in previously reported observations of cyclic velocity fluctuations in the trajectories of walking *Euplotes* (*36*).

Finally, we argue that subcellular processes must be involved in actively coordinating cirri in order to accomplish the observed stereotypy in biased sequential activity. Our analysis of and results of experiments perturbing the tubulin-based cytoskeletal fiber system are consistent with its role in mechanically mediating communication both among cirri and between cirri and cell cortex (Fig. 4). We conjecture that movement of cirri relative to one another can establish tension in the fiber system and that the tension state of fibers associated with each cirrus may then modulate cirral activity in a manner reminiscent of basal coupling in flagellates (*22*). Changes in fiber length, as observed under nocodazole treatment, change the relative positions of cirri, which could lead to different distributions of tension and thereby induce the altered pattern of state transitions and associated walking defects. Although the details of such a mechanism remain to be tested, our results are consistent with the fibers system playing an active role in gait coordination. Thus, by combining information processing to properly dictate patterns of cirral activity and the mechanical actions of cirral movement, walking *Euplotes* embodies the sequential computation of a finite state machine. Furthermore, our approach to understanding a complex cellular behavior, grounded in theory of computation, allowed us to derive mechanistic insight from observational and perturbative experiments.

Among the domains of life, eukaryotes uniquely display remarkable complexity and diversity in cellular behavior (*87*). Our approach, grounded in finite state machine analysis, has revealed modularity and stereotypy underlying complex cellular behavior, implicating a machine-like process. Our results suggest that integrating approaches from theoretical computer science, non-equilibrium statistical physics, and cell biology stands to shed light on the regulation of cellular behavior in eukaryotes more broadly. By revealing principles of cellular

behavior, the line of research established here stands to advance our ability to predict and even one day engineer cellular behavior across diverse eukaryotic systems.

References

1. H. S. Jennings, *Behavior of the lower organisms.* (Columbia University Press, New York, NY, 1906).
2. C. R. Reid, H. MacDonald, R. P. Mann, J. A. R. Marshall, T. Latty, S. Garnier, Decision-making without a brain: how an amoeboid organism solves the two-armed bandit. *J. R. Soc. Interface*. **13**, 20160030 (2016).
3. S. M. Coyle, E. M. Flaum, H. Li, D. Krishnamurthy, M. Prakash, Coupled Active Systems Encode an Emergent Hunting Behavior in the Unicellular Predator Lacrymaria olor. *Curr. Biol.* **29**, 3838-3850.e3 (2019).
4. C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, F. C. MacKintosh, Broken detailed balance at mesoscopic scales in active biological systems. *Science*. **352**, 604–7 (2016).
5. K. Y. Wan, R. E. Goldstein, Time Irreversibility and Criticality in the Motility of a Flagellate Microorganism. *Phys. Rev. Lett.* **121**, 58103 (2018).
6. A. Woznica, A. M. Cantley, C. Beemelmanns, E. Freinkman, J. Clardy, N. King, Bacterial lipids activate, synergize, and inhibit a developmental switch in choanoflagellates. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7894–7899 (2016).
7. K. Y. Wan, Coordination of eukaryotic cilia and flagella. *Essays Biochem.* **62**, 829–838 (2018).
8. M. Nomura, K. I. Ishida, Fine-structural Observations on Siliceous Scale Production and Shell Assembly in the Testate Amoeba Paulinella chromatophora. *Protist*. **167**, 303–318

(2016).

9. R. R. Bennett, R. Golestanian, A steering mechanism for phototaxis in Chlamydomonas. *J. R. Soc. Interface*. **12**, 20141164–20141164 (2015).

10. H. Machemer, The swimming cell and its world: Structures and mechanisms of orientation in protists. *Eur. J. Protistol.* **37**, 3–14 (2001).

11. M. Almagor, A. Ron, J. Bar-Tana, Chemotaxis in Tetrahymena thermophila. *Cell Motil.* **1**, 261–268 (1981).

12. H. C. Berg, D. A. Brown, Chemotaxis in Escherichia coli analysed by three-dimensional tracking. *Nature*. **239**, 500–504 (1972).

13. T. Nakagaki, H. Yamada, Á. Tóth, Maze-solving by an amoeboid organism. *Nature*. **407**, 470–470 (2000).

14. L. Tweedy, P. A. Thomason, P. I. Paschke, K. Martin, L. M. Machesky, M. Zagnoni, R. H. Insall, Seeing around corners: Cells solve mazes and respond at a distance using attractant breakdown. *Science (80-.).* **369** (2020), doi:10.1126/science.aay9792.

15. P. B. Applewhite, in *Biochemistry and Physiology of Protozoa* (Elsevier, 1979), pp. 341–355.

16. J. P. Dexter, S. Prabakaran, J. Gunawardena, A Complex Hierarchy of Avoidance Behaviors in a Single-Cell Eukaryote. *Curr. Biol.* **29**, 4323-4329.e2 (2019).

17. H. S. Jennings, Studies on reactions to stimuli in unicellular organisms. IX—On the behavior of fixed infusoria (Stentor and Vorticella), with special reference to the modifiability of protozoan reactions. *Am. J. Physiol. Content*. **8**, 23–60 (1902).

18. S. K. Y. Tang, W. F. Marshall, Cell learning. *Curr. Biol.* **28** (2018), pp. R1180–R1184.

19. P. Lyon, The cognitive cell: Bacterial behavior reconsidered. *Front. Microbiol.* **6** (2015), , doi:10.3389/fmicb.2015.00264.

20. W. F. Marshall, Cellular Cognition: Sequential Logic in a Giant Protist. *Curr. Biol.* **29** (2019), pp. R1303–R1305.

21. D. Greenfield, A. L. McEvoy, H. Shroff, G. E. Crooks, N. S. Wingreen, E. Betzig, J. Liphardt, Self-Organization of the Escherichia coli Chemotaxis Network Imaged with Super-Resolution Light Microscopy. *PLoS Biol.* **7**, e1000137 (2009).

22. K. Y. Wan, R. E. Goldstein, Coordinated beating of algal flagella is mediated by basal coupling. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E2784-93 (2016).

23. T. Höfer, J. A. Sherratt, P. K. Maini, Dictyostelium discoideum : cellular self-organization in an excitable biological medium. *Proc. R. Soc. London. Ser. B Biol. Sci.* **259**, 249–257 (1995).

24. V. Kantsler, J. Dunkel, M. Polin, R. E. Goldstein, Ciliary contact interactions dominate surface scattering of swimming eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1187–1192 (2013).

25. D. Bray, *Wetware* (Yale University Press, New Haven, CT, 2009).

26. R. Pfeifer, M. Lungarella, F. Iida, Self-organization, embodiment, and biologically inspired robotics. *Science (80-.).* **318** (2007), pp. 1088–1093.

27. A. Regev, E. Shapiro, Cellular abstractions: Cells as computation. *Nature*. **419**, 343–343 (2002).

28. M. Mitchell, Biological Computation. *Comput. J.* **55**, 852–855 (2012).

29. S. Navlakha, Z. Bar-Joseph, Algorithms in nature: the convergence of systems biology and computational thinking. *Mol. Syst. Biol.* **7**, 546 (2011).

30. T. Savin, N. A. Kurpios, A. E. Shyer, P. Florescu, H. Liang, L. Mahadevan, C. J. Tabin, On the growth and form of the gut. *Nature*. **476**, 57–62 (2011).

31. P. Mehta, D. J. Schwab, Energetic costs of cellular computation. *Proc. Natl. Acad. Sci. U.*

S. A. **109**, 17978–17982 (2012).

32. H. C. Berg, E. M. Purcell, Physics of chemoreception. *Biophys. J.* **20**, 193–219 (1977).

33. W. Bialek, S. Setayeshgar, Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10040–5 (2005).

34. L. B. Andrews, A. A. K. Nielsen, C. A. Voigt, Cellular checkpoint control using programmable sequential logic. *Science (80-.).* **361** (2018), doi:10.1126/science.aap8987.

35. J. J. Gilbert, Jumping behavior in the oligotrich ciliates Strobilidium velox and Halteria grandinella, and its significance as a defense against rotifer predators. *Microb. Ecol.* **27**, 189–200 (1994).

36. W. Lueken, N. Ricci, T. Krüppel, Rhythmic spontaneous depolarizations determine a slow-and-fast rhythm in walking of the marine hypotrich Euplotes vannus. *Eur. J. Protistol.* **32**, 47–54 (1996).

37. F. Erra, A. Iervasi, N. Ricci, R. Banchetti, Movement of the cirri during the creeping of Euplotes crassus (Ciliata, Hypotrichida). *Can. J. Zool.* **79**, 1353–1362 (2001).

38. C. V. Taylor, The neuromotor system of Euplotes. *Exp. Biol. Med.* **16**, 101–103 (1919).

39. M. J. Syberg-Olsen, N. A. T. Irwin, C. Vannini, F. Erra, G. Di Giuseppe, V. Boscaro, P. J. Keeling, Biogeography and character evolution of the ciliate genus Euplotes (Spirotrichea, Euplotia), with description of Euplotes curdsi sp. nov. *PLoS One*. **11** (2016), doi:10.1371/journal.pone.0165442.

40. N. Ricci, R. Giannetti, C. Miceli, The ethogram of Euplotes crassus (ciliata, hypotrichida): I. The wild type. *Eur. J. Protistol.* **23**, 129–140 (1988).

41. Y. Li, L. Shi, F. Gu, Comparison of ciliature microtubule organelles in three hypotrichous ciliate species*. *Chinese J. Oceanol. Limnol.* **28**, 443–450 (2010).

42. A. Fleury, Dynamics of the cytoskeleton during morphogenesis in the ciliate Euplotes I.

Basal bodies related microtubular system. *Eur. J. Protistol.* **27**, 99–114 (1991).

43. E. M. Purcell, Life at low Reynolds number. *Am. J. Phys.* **45**, 3–11 (1977).

44. J. C. Kimmel, A. Y. Chang, A. S. Brack, W. F. Marshall, Inferring cell state by quantitative motility analysis reveals a dynamic state system and broken detailed balance. *PLoS Comput. Biol.* **14**, e1005927 (2018).

45. A. Y. Chang, W. F. Marshall, Dynamics of living cells in a cytomorphological state space. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21556–21562 (2019).

46. L. Boltzmann, Weirere Studien uber das warmegleich-gewich unter gasmolekulen. *Sitzungsberichte der Akad. der Wissenschaften, Vienna, part II*. **66**, 275–370 (1872).

47. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular biology of the cell* (Garland Science, New York, ed. 4th, 2002).

48. C. R. Shalizi, J. P. Crutchfield, Computational Mechanics: Pattern and Prediction, Structure and Simplicity. *J. Stat. Phys.* **104**, 817–879 (2001).

49. J. P. Crutchfield, The calculi of emergence: computation, dynamics and induction. *Phys. D Nonlinear Phenom.* **75**, 11–54 (1994).

50. J. E. Hopcroft, R. Motwani, J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation* (Pearson Education, New York, NY, ed. 3rd, 2008).

51. T. Y. Moore, K. L. Cooper, A. A. Biewener, R. Vasudevan, Unpredictability of escape trajectory explains predator evasion ability and microhabitat preference of desert rodents, doi:10.1038/s41467-017-00373-2.

52. E. L. Barnhart, G. M. Allen, F. Jülicher, J. A. Theriot, Bipedal locomotion in crawling cells. *Biophys. J.* **98**, 933–942 (2010).

53. R. E. Goldstein, Green Algae as Model Organisms for Biological Fluid Dynamics. *Annu. Rev. Fluid Mech.* **47**, 343–375 (2015).

54. M. Simon, H. Plattner, in *International Review of Cell and Molecular Biology* (Elsevier Inc., 2014), vol. 309, pp. 141–198.

55. W. M. Durham, E. Climent, M. Barry, F. De Lillo, G. Boffetta, M. Cencini, R. Stocker, Turbulence drives microscale patches of motile phytoplankton. *Nat. Commun.* **4** (2013), doi:10.1038/ncomms3148.

56. A. Mogilner, G. Oster, Cell motility driven by actin polymerization. *Biophys. J.* **71**, 3030–3045 (1996).

57. J. J. Collins, S. A. Richmond, "Biological cybemet Hard-wired central pattern generators for quadrupedal locomotion" (1994).

58. J. J. Collins, I. N. Stewart, Coupled nonlinear oscillators and the symmetries of animal gaits. *J. Nonlinear Sci.* **3**, 349–392 (1993).

59. R. McNeill Alexander, *Principles of Animal Locomotion* (Princeton University Press, Princeton, NJ, 2003).

60. K. Y. Wan, G. Jékely, On the unity and diversity of cilia. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **375**, 20190148 (2020).

61. T. Ahamed, A. C. Costa, G. J. Stephens, Capturing the Continuous Complexity of Behavior in C. elegans, doi:10.1101/827535.

62. G. J. Stephens, B. Johnson-Kerner, W. Bialek, W. S. Ryu, Dimensionality and Dynamics in the Behavior of C. elegans. *PLoS Comput. Biol.* **4**, e1000028 (2008).

63. G. J. Berman, D. M. Choi, W. Bialek, J. W. Shaevitz, Mapping the stereotyped behaviour of freely moving fruit flies. *J. R. Soc. Interface*. **11**, 20140672 (2014).

64. G. J. Stephens, L. C. Osborne, W. Bialek, Searching for simplicity: Approaches to the analysis of neurons and behavior. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15565–15571 (2010).

65. M. Ester, H.-P. Kriegel, J. Sander, X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise" (1996).

66. L. Kauffman, P. J. Rousseeuw, *Finding Groups in Data: and introduction to cluster analysis* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 1990), *Wiley Series in Probability and Statistics*.

67. J. B. Kirkegaard, A. O. Marron, R. E. Goldstein, Motility of Colonial Choanoflagellates and the Statistics of Aggregate Random Walkers. *Phys. Rev. Lett.* **116**, 038102 (2016).

68. M Cartmill, in *Functional Vertebrate Morphology* (Harvard University Press, Cambridge, MA, 1985), pp. 73–88.

69. P. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation* (Wiley, New York, 2017).

70. A. A. Markov, Extension of the Limit Theorems of Probability Theory to a Sum of Variables Connected in a Chain. *Notes Imp. Acad. Sci. St. Petersbg. VIII Ser. Physio-Mathematical Coll.* **22** (1907).

71. R. A. Howard, *Dynamic and Probabilistic Systems, Volume 1: Markov models* (Dover Publications, Mineola, NY, 1971).

72. P. Billingsley, *Statistical Methods in Markov Chains* (Institute of Mathematical Statistics, 1961), vol. 32.

73. J. P. Crutchfield, K. Young, in *Complexity, Entropy, and Physics of Information*, W. Zurek, Ed. (Addison-Wesley, Reading, MA, 1989).

74. J. P. Crutchfield, K. Young, Inferring statistical complexity. *Phys. Rev. Lett.* **63**, 105–108 (1989).

75. C. Moore, Dynamical recognizers: Real-time language recognition by analog computers. *Theor. Comput. Sci.* (1998), doi:10.1016/S0304-3975(97)00028-5.

76. C. V. Taylor, Demonstration of the function of the neuromotor apparatus in Euplotes by the method of microdissection. *Naturwissenschaften*. **9**, 910 (1921).

77. A. Okajima, H. Kinosita, Ciliary activity and coordination in Euplotes eurystomus-I. Effect of microdissection of neuromotor fibres. *Comp. Biochem. Physiol.* **19**, 115–131 (1966).

78. A. Okabe, F. Miki, A Conditional Nearest-Neighbor Spatial-Association Measure for the Analysis of Conditional Locational Interdependence. *Environ. Plan. A Econ. Sp.* **16**, 163–171 (1984).

79. R. Bar-Ziv, T. Tlusty, A. Libchaber, Protein-DNA computation by stochastic assembly cascade. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11589–11592 (2002).

80. J. S. Guasto, R. Rusconi, R. Stocker, Fluid Mechanics of Planktonic Microorganisms. *Annu. Rev. Fluid Mech.* **44**, 373–400 (2012).

81. R. E. Goldstein, M. Polin, I. Tuval, Noise and Synchronization in Pairs of Beating Eukaryotic Flagella. *Phys. Rev. Lett.* **103**, 168103 (2009).

82. B. Párducz, Ciliary Movement and Coordination in Ciliates. *Int. Rev. Cytol.* **21**, 91–128 (1967).

83. A. Funfak, C. Fisch, H. T. Abdel Motaal, J. Diener, L. Combettes, C. N. Baroud, P. Dupuis-Williams, Paramecium swimming and ciliary beating patterns: A study on four RNA interference mutations. *Integr. Biol. (United Kingdom)*. **7**, 90–100 (2015).

84. J. B. Kirkegaard, A. Bouillant, A. O. Marron, K. C. Leptos, R. E. Goldstein, Aerotaxis in the closest relatives of animals. *Elife*. **5**, e18109 (2016).

85. G. L. Miño, M. A. R. Koehl, N. King, R. Stocker, Finding patches in a heterogeneous aquatic environment: pH-taxis by the dispersal stage of choanoflagellates. *Limnol. Oceanogr. Lett.* **2**, 37–46 (2017).

86. M. Polin, I. Tuval, K. Drescher, J. P. Gollub, R. E. Goldstein, Chlamydomonas Swims with Two "Gears" in a Eukaryotic Version of Run-and-Tumble Locomotion. *Science (80-.).* **325**, 487–490 (2009).

87. P. J. Keeling, Combining morphology, behaviour and genomics to understand the evolution and ecology of microbial eukaryotes. *Philos. Trans. R. Soc. B Biol. Sci.* **374** (2019), doi:10.1098/rstb.2019.0085.

88. N. King, S. L. Young, M. Abedin, M. Carr, B. S. C. C. Leadbeater, *Cold Spring Harb. Protoc.*, in press, doi:10.1101/pdb.prot5148.

89. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis. *Nat. Methods*. **9**, 676–682 (2012).

90. J. Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, S. Y. Bednarek, S. L. Shorte, K. W. Eliceiri, TrackMate: An open and extensible platform for single-particle tracking. *Methods*. **115**, 80–90 (2017).

91. B. Kanagal, V. Sindhwani, in *Advances in Neural Information Processing Systems* (2010), vol. 23.

92. A. B. Owen, P. O. Perry, Bi-Cross-Validation of the SVD and the Nonnegative Matrix Factorization. *Source Ann. Appl. Stat.* **3**, 564–594 (2009).

93. S. Wold, Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models. *Technometrics*. **20**, 397–405 (1978).

94. T. Ronan, Z. Qi, K. M. Naegle, Avoiding common pitfalls when clustering biological data. *Sci. Signal.* **9** (2016), pp. re6–re6.

95. T. Kawamoto, Y. Kabashima, Cross-validation estimate of the number of clusters in a

network. *Sci. Rep.* **7**, 1–17 (2017).

96. P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: An explanation of the 1/f noise. *Phys. Rev. Lett.* **59**, 381–384 (1987).

97. R. Sánchez, D. E. Newman, B. A. Carreras, Waiting-Time Statistics of Self-Organized-Criticality Systems. *Phys. Rev. Lett.* **88**, 4 (2002).

98. R. N. Hiscott, "Chi-Square Tests for Markov Chain Analysis" (1981).

99. A. Cardona, S. Saalfeld, J. Schindelin, I. Arganda-Carreras, S. Preibisch, M. Longair, P. Tomancak, V. Hartenstein, R. J. Douglas, TrakEM2 Software for Neural Circuit Reconstruction. *PLoS One*. **7**, e38011 (2012).

100. P. Thévenaz, U. E. Ruttimann, M. Unser, A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).

101. A. Kahl, *Urtiere oder Protozoa.* (Fischer, Jena, 1935).

102. H. B. Yocom, thesis, University of California, Berkeley (1917).

Figures:

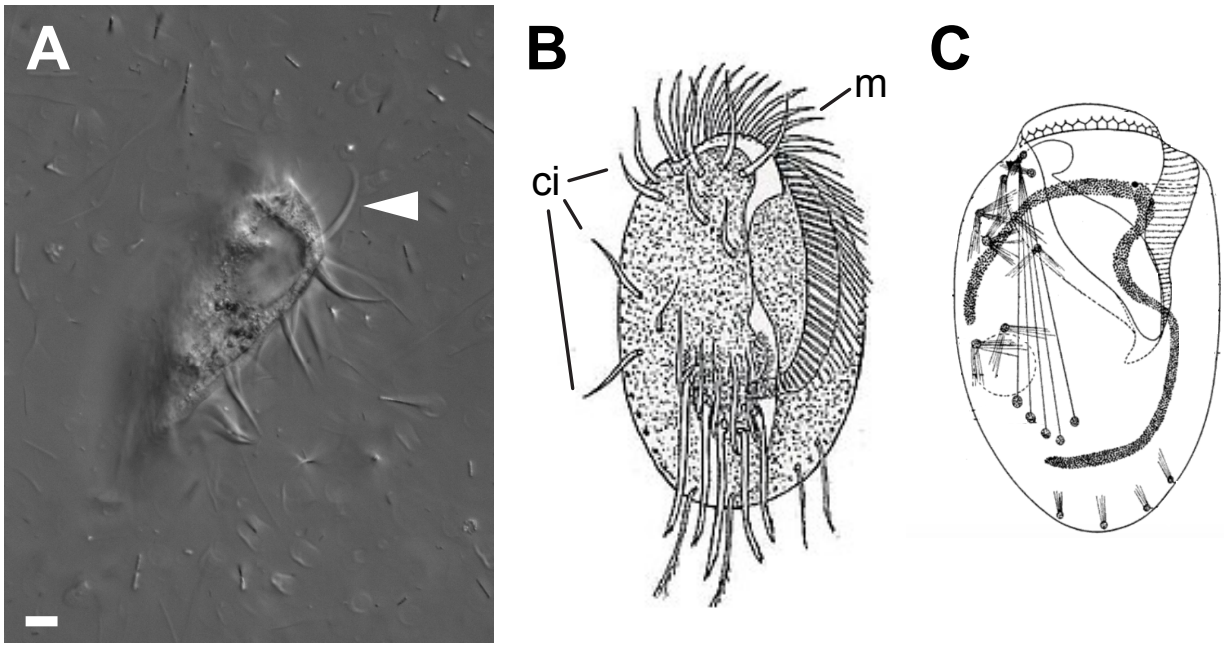


Fig. 1. *Euplotes* exhibits highly polarized, complex cellular architecture and walks across surfaces using microtubule-based organelles called cirri, some of which are physically linked. (**A**) A single *Euplotes eurystomus* cell in profile displays its ventral cirri, which are used for walking locomotion across surfaces (arrowhead indicates a single cirrus stretching out from the cell). Scale bar is 10 µm. (**B**) A drawing of a *Euplotes* cell, viewed from the ventral surface, highlighting the complex, asymmetric structure of the cell. Notable features include the cirri (ci) and the membranellar band (m), wrapping from the top of the cell to the center, which is used to generate a feeding current to draw in prey items. Drawing adapted and obtained from Wikimedia Commons, from original source (*101*). (**C**) A drawing of a *Euplotes* cell, highlighting the fiber system associated with the cirri, historically referred to as the neuromotor apparatus. Drawing adapted from (*102*).

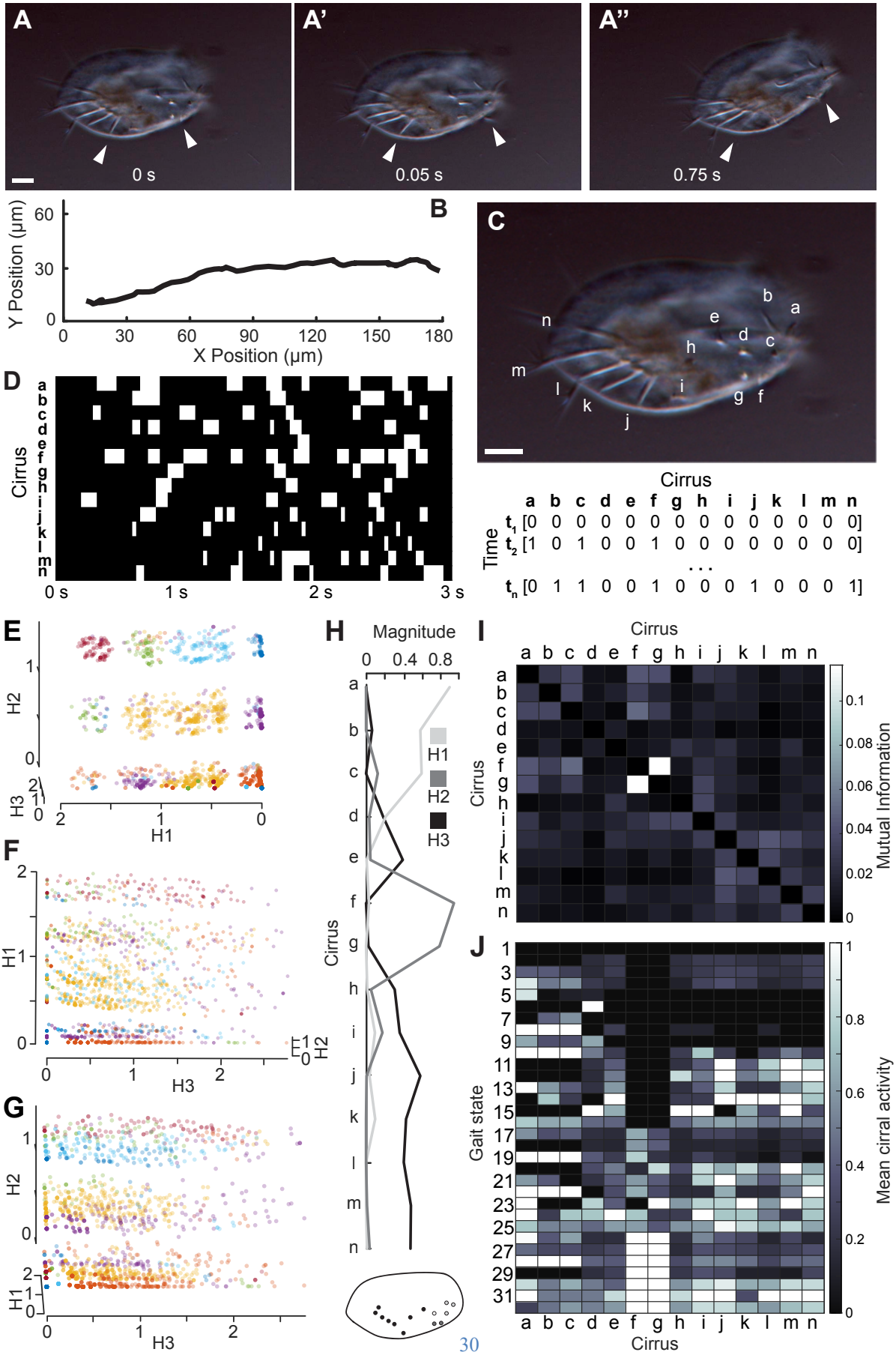

A
A'
A"
0 s
0.05 s
0.75 s
B
Y Position (μm)
X Position (μm)
C
D
Cirrus
0 s
1 s
2 s
3 s
Cirrus
Time
E
H2
H3
H1
F
H1
H3
H2
G
H2
H1
H3
H
Magnitude
Cirrus
H1
H2
H3
I
Cirrus
Mutual Information
J
Gait state
Mean cirral activity
Cirrus

Fig. 2. **The gait of *Euplotes* can be described in a discrete, reduced state space with gait states corresponding to identifiable patterns of cirral activity.** (**A-A''**), The movements of cirri during walking locomotion are clearly visible by brightfield microscopy by focusing on a plane at the surface of the coverslip on which cells are walking. Three snapshots depict different time points during a single walking trajectory, and white arrowheads indicate cirri. In the panels from left to right, the cirrus indicated by the arrowhead on the left is stationary, stationary, and then moving, and the cirrus indicated by the arrowhead on the right is stationary, moving, and then stationary. Scale bar is 15 µm. (**B**) The trajectory of a cell during a single recorded trajectory as the cell walked across a coverslip from left to right. The cell position was manually tracked in each frame. (**C**) The scheme for encoding cirral dynamics during walking involved labeling each of the 14 distinguishable ventral cirri (a-n), and recording cirral activity in each frame, corresponding to timepoints (t_1,...,t_n), of recordings of walking cells as a 14-bit binary vector. Each entry in each vector is given a value of either 0 if the cirrus is not moving and in contact with the coverslip or 1 if the cirrus is moving. Scale bar is 15 µm. (**D**) Representative visualization of cirral dynamics for a single trajectory of a walking cell. These dynamics correspond to the walking trajectory in (B). Each row corresponds to a cirrus and each column is a single video frame. White denotes cirral activity, a value of 1, in the vector encoding of dynamics from (C). Note the dynamical complexity and discrete, stepping-like nature of cirral movements. (**E-G**) Three roughly orthogonal views of a plot displays the structure of all recorded cirral dynamics encoded as in Figure 2C from 13 cells over 2343 timepoints in a reduced state space obtained by non-negative matrix factorization (NMF). Axes correspond to the components of the NMF (H1, H2, H3), and each point is a single timepoint. Randomized colors highlight the 32 clusters identified using the density-based spatial clustering of applications with noise (DBSCAN) algorithm (*65*). We refer to these clusters as gait states, and

they correspond to unique configurations of cirral activity during walking locomotion (see (F)). (**H**) A plot of the magnitudes associated with each cirrus corresponding to the components of the NMF of cirral dynamics shows distinct contributions from spatially distinct groups of cirri. Component H1, for example, is associated with activity in cirri a, b, and c. The tracing of a cell including the position of cirri has the same color map as the plot above and shows the grouping of the cirri corresponding to each component. (**I**) A heatmap of mutual information between all pairs of cirri shows that correlations in cirral activity correspond to the NMF components displayed in (D). For example, cirri a, b, and c share mutual information with one another and are the cirri contributing to component H1. (**J**) A heatmap representation of the cirral activity associated with each of the 32 gait states. Values for each cirrus are the mean over all instances of the gait state. Note that each gait state has a unique signature of cirral activity.

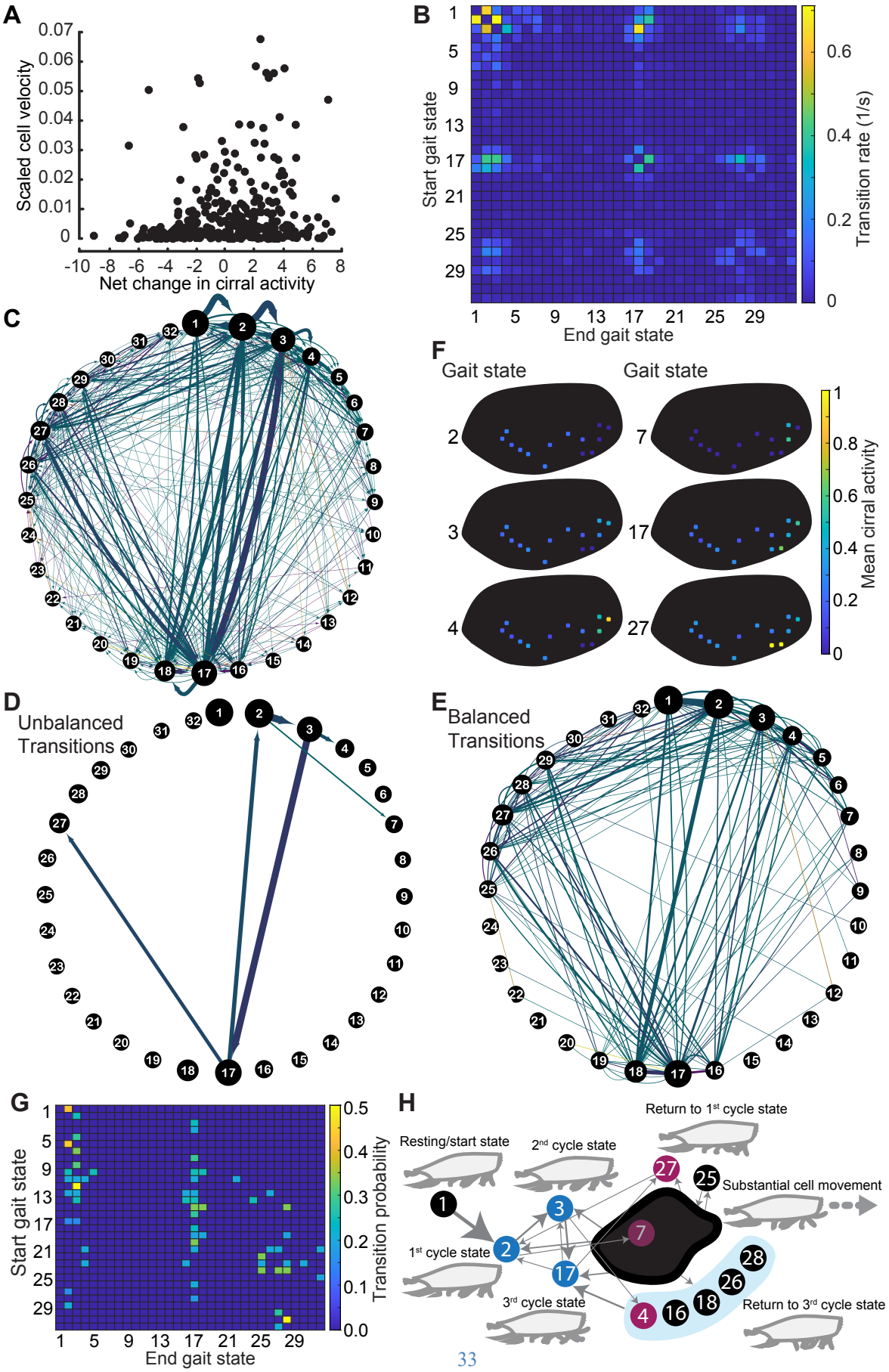
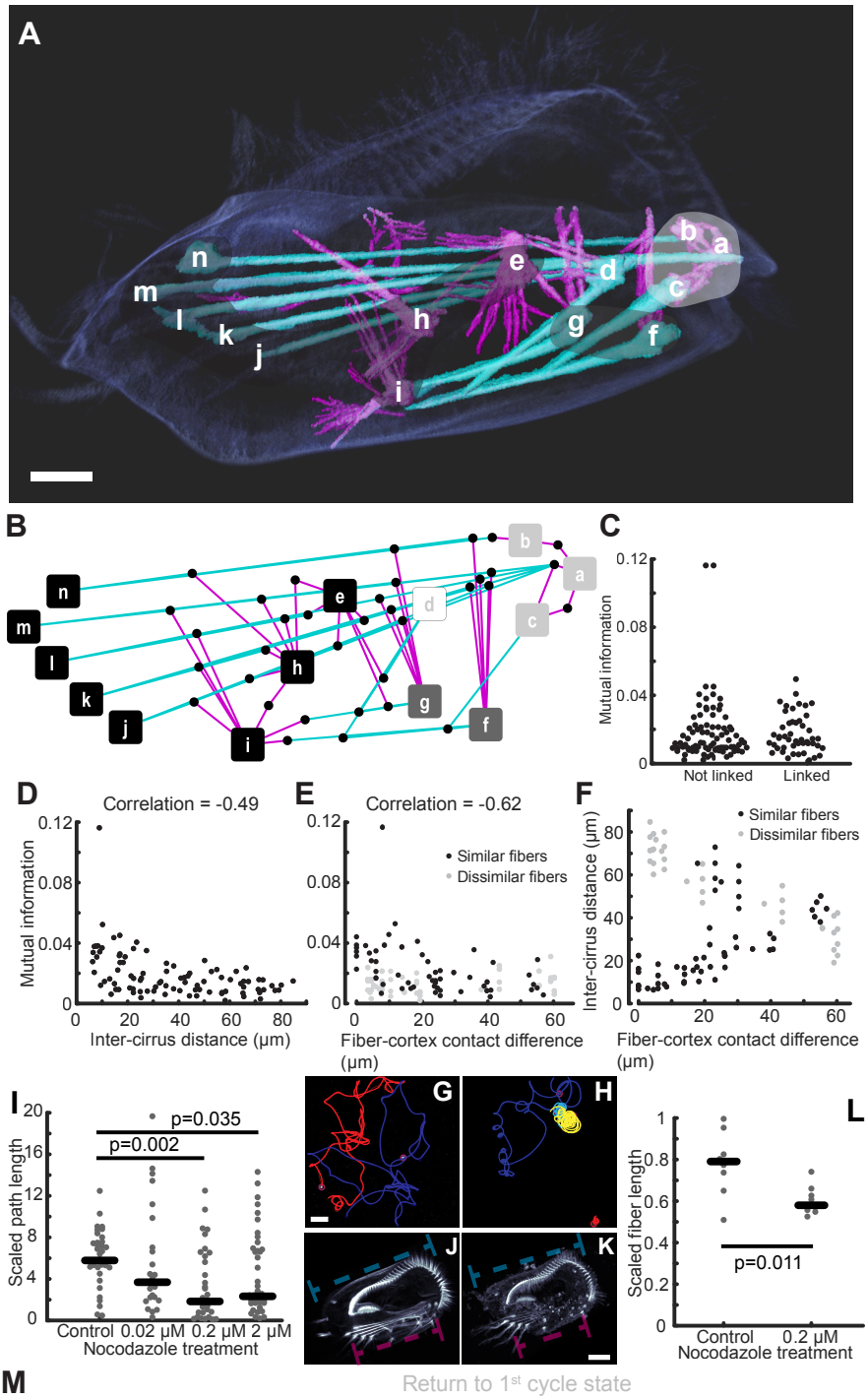

A
Scaled cell velocity
Net change in cirral activity
B
Start gait state
End gait state
Transition rate (1/s)
C
D
Unbalanced Transitions
E
Balanced Transitions
F
Gait state
Gait state
Mean cirral activity
G
Start gait state
End gait state
Transition probability
H
Resting/start state
2nd cycle state
Return to 1st cycle state
Substantial cell movement
1st cycle state
3rd cycle state
Return to 3rd cycle state

Fig. 3. *Euplotes* walks with a cyclic stochastic gait exhibiting broken detailed balance, stereotypy, and state machine-like dynamics. All data are pooled from the walking trajectories of 13 different cells over 2343 timepoints and 1423 pairwise gait state transitions. (**A**) A plot of the mean net change in cirral activity versus the net scaled cell velocity associated with all transitions between the 32 gait states identified in Fig. 3 shows that the change in number of active cirri is not strongly correlated with cell velocity (R^2=0.03). Cell velocities were obtained from manually tracked walking trajectories and then scaled by dividing frame to frame displacements for each trajectory by the length of the cell being tracked and also dividing by the average frequency of cirral inactivity. Scaling provided a non-dimensional velocity scaled by natural units of the system. Note that at low Reynolds number, velocity should be proportional to force (*43*), so this plot also reflects the net walking force generated by the cell. Net change in cirral activity was computed using the data presented in Fig. 3F. Note that the largest velocities are associated with small negative and small to moderate positive net changes in cirral activity. (**B**) The transition matrix of all gait state transitions, with rows representing the starting state and columns indicating the ending state, exhibits broken detailed balance. Rates were estimated by dividing the total number of observed transitions between each state pair and dividing by the total time observed. Under detailed balance or equilibrium conditions, transitions from one state to another should be balanced by reverse transitions. Lack of this kind of reversibility, as seen by the lack of symmetry of the heatmap across the diagonal, indicates broken detailed balance and non-equilibrium dynamics. (**C**) A directed graph representation of all gait state transitions. Nodes correspond to the 32 gait states, and node sizes are scaled by the proportion of total time cells spent in each state. Directed edges are represented by arrows between nodes and signify state transitions. The size of the arrows is scaled by transition rates as in (B). Edge color represents scaled cell velocity as in (A), with cool colors (more cyan) representing lower velocity, and

warmer colors (more magenta) representing higher velocity. (**D**) A subset of transitions visualized as in **c** shows the restricted and relatively high frequency nature of unbalanced, non-equilibrium-like transitions. Only transitions that were observed to happen more than one time and exhibiting a significant difference between forward and reverse transitions ($p<0.05$ by binomial test, see Methods for more details) are displayed. (**E**) A subset of transitions, similarly to panel E, except that only the balanced transitions, lacking a significant difference between forward and reverse transitions ($p<0.05$ by binomial test) are displayed, also show a complex and widespread structure, this time of balanced, equilibrium transitions. Note that the majority of transitions associated with high cell velocity involve equilibrium-like dynamics. (**F**) Examples illustrating the spatial organization of cirral activity corresponding to gait states. Some states, such as 7, correspond to activity in spatially discrete groups of cirri, while others, such as 17, correspond to cirral activity across the cell. The gait states displayed here are those involved in unbalanced transitions. (**G**) A heatmap of transition probabilities between states, showing only the most probable transitions from a given state with all others set to zero, shows distinct structure. In cases were multiple state transitions from a state were tied for the highest probability, all of these transitions are displayed. Fewer than half of the total states are recipients of multiple high probability transitions, and many states are the recipients of no high probability transitions. (**H**) A representation of functional states and transitions between them highlights the machine-like nature of the gait of *Euplotes*. Gait states are represented as circles with numerical labels. Blue circles represent states that are both recipients and sources of unbalanced transitions as identified in (D) and constitute the three cycle states. Red circles represent states that are recipients but not sources of unbalanced transitions as identified in (D). Black circles correspond to gait states that are associated only with balanced transitions as in (E). States receiving no more than one unique high probability transitions from states with only a single highest as identified in

(G) were grouped together into a compound state represented by the dark gray blob. The blue background behind states 4, 16, 18, 26, and 28 indicates that these states all share the same highest probability transitions between states identified in this panel, and thus, the group constitutes a single compound functional state. Arrows represent the highest probability transitions between the states, including compound states composed of multiple gait states as identified in Fig. 2 and 3 (dark gray blob and blue background). Gait state 1 is also depicted, as it is the state in which cells spent the most time over all walking trajectories and also is uniquely the state from which cells begin walking. Cells also frequently return to the state during walking. Further, transitions from gait state 2 from gait state 1 constitute the single highest frequency transition. Together, all identified states in this panel constitute functional states. Arrows represent the most probable transitions between functional states, and all unbalanced transitions are also represented with size scaled by their proportional probability compared to all other transitions emanating from the source functional state. Cartoons are a walking cell in profile with cirri in a configuration representative of the corresponding functional state. Labels refer to the apparent functional role of states and their associated transitions. Beginning from gait state 1, the resting/start state, cells are most likely to follow transitions from gait state 2 to 3 to 17 at which point cells are likely to enter the functional state associated with substantial cell movement involving variable balanced transitions between a number of gait states. Transitions are then likely to lead back toward the cycle states. Note that while this representation of gait dynamics highlights the most probable transitions, substantial variability, primarily involving reversible transitions, occurs during walking trajectories.


A
n
m
l
k
j
i
h
e
d
g
f
c
b
a
B
C
Mutual information
Not linked
Linked
D
Correlation = -0.49
Mutual information
Inter-cirrus distance (μm)
E
Correlation = -0.62
Similar fibers
Dissimilar fibers
Fiber-cortex contact difference (μm)
F
Inter-cirrus distance (μm)
Similar fibers
Dissimilar fibers
Fiber-cortex contact difference (μm)
I
Scaled path length
p=0.035
p=0.002
Control 0.02 μM 0.2 μM 2 μM
Nocodazole treatment
G
H
J
K
L
Scaled fiber length
p=0.011
Control 0.2 μM
Nocodazole treatment

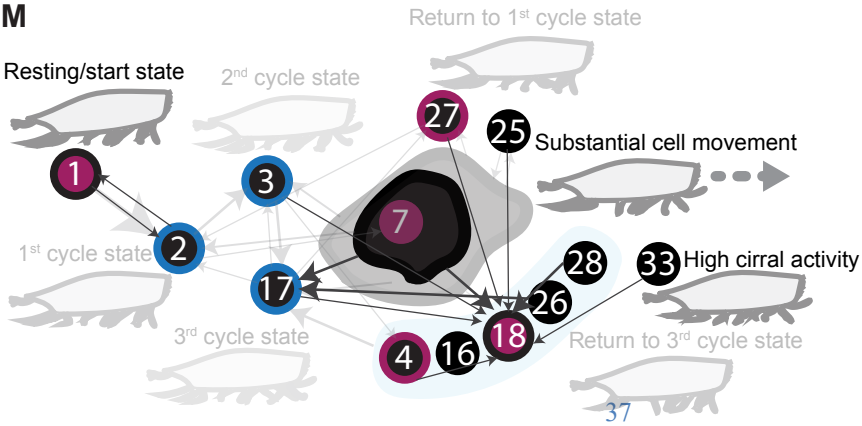

M
Return to 1st cycle state
Resting/start state
2nd cycle state
Substantial cell movement
1st cycle state
High cirral activity
3rd cycle state
Return to 3rd cycle state
1
2
3
4
7
16
17
18
25
26
27
28
33

Fig. 4. The complex, interconnected fiber system of *Euplotes* mediates gait coordination. (**A**) The SiR-tubulin labeled cell (faint, dark blue) was imaged by confocal microscopy, and a 3D reconstruction as obtained from serial confocal slices. Fibers were manually traced in each slice using TrakEM2 in FIJI. Two morphologically distinct classes of fibers were observed and are indicated as follows: thick, linear fibers are cyan and thinner, filamentous fibers are magenta (see Fig. S6 for raw image data). Fibers emanate from the base of each cirrus and form a connected network between all cirri. The base of each cirrus is indicated by corresponding letters (as in Fig. 2c). Gray shading indicates the dynamical groups identified by dimensionality reduction and follows the same color scheme as in Fig. 3d. Scale bar is 10 µm. (**B**) A graph representation of fiber-fiber connections illustrates the complex and interconnected nature of cirrus associated fiber topology. Nodes correspond to the cirri to which each fiber system is associated, and edges indicate connections between fiber systems. Colors of nodes indicate the same groups as in (A), and colors of edges indicate the types of fibers connecting to one another, cyan for thick fiber connections, magenta for thin fiber connections, and purple for thick to thin fiber connections. (**C**) Pairs of cirri that are linked by fiber-fiber contacts show no statistically significant difference in mutual information compared to those lacking fiber-fiber contacts. The plot displays mutual information between all pairs of cirri grouped by the absence (Not linked) or presence (Linked) of associated fiber-fiber connections. Statistical significance was evaluated by the Wilcoxon rank sum test. Note that when pairs of cirri were grouped by fiber-fiber connection type, we did observe an increase in mutual information for cirri associated with only thin fiber-fiber connection and only thick fiber-fiber connections compared to those lacking fiber-fiber connections (see Fig. S6). (**D**) A plot of mutual information as a function of inter-cirrus distance displays negative correlation, with a Spearman correlation coefficient of -0.49 ($p<0.001$). Plotted values are defined with respect to pairs of cirri. (**E**) A plot of mutual

information as a function of fiber-cortex contact distance grouped by fiber type similarity and lack thereof displays negative correlation, with a Spearman correlation coefficient of -0.62 ($p<0.001$) for pairs of cirri with similar fiber types and no significant correlation for those with dissimilar fiber types. Similarity of fiber types is defined in terms of sharing at least some fiber types as defined in (A). Fiber-cortex contact difference is measured by the mean cross nearest neighbor distance (see Methods) for all fiber-cortex contact points associated with each cirrus. The negative correlation values from the data plotted in (D) and (E) indicate that cirri that are closer to one another and also cirri with fiber-cortex contacts in nearby regions of the cell tend to have higher mutual information, and indeed cirri that are both close to one another and with similar patterns of fiber-cortex contacts display the highest mutual information. (**F**) A plot of fiber-cortex contact difference versus inter-cirrus difference (as in (D) and (E)) illustrates that nearby cirri tend to have similar associated fiber-cortex contacts, highlighting that nearby cirri with similar fiber-cortex contacts share the most mutual information. (**G, H**) Representative cell motility trajectories of cells imaged under darkfield illumination and tracked using the TrackMate plugin in FIJI (*90*) in control (G) and nocodazole treated (H) cells highlight the increased turning leading to curved trajectories in cells treated with the inhibitor of microtubule polymerization nocodazole. Different colors represent the trajectories of different cells. Cells were imaged by dark field microscopy and images were processed using FIJI. Tracking was performed using the TrackMate FIJI plugin (*90*). Scale bar is 500 μm. (**I**) Nocodazole affects cell motility in a dose dependent manner. Motility was quantified by the scaled path length, which is the total integrated distance walked by a cell scaled by the maximum distance the cell travelled from its starting point. This scaled path length decreases with a decrease in long, straight segments of trajectories corresponding to normal cell walking. Scaled path length decreased with increased nocodazole concentration, becoming significantly less than controls by 0.2 μM

nocodazole. Note that at 2 μM nocodazole, cells were often observed to swim instead of walk, which may account for the lack of decrease in scaled path length compared to 0.2 μM treatment. The trajectories of at least 20 cells were analyzed for each condition. The black bars are median values. The p-values were calculated by Wilcoxon rank-sum test. (**J, K**) Representative images illustrating the effect of the inhibition of microtubule polymerization by nocodazole on the fiber system and approach to quantifying this effect. Images are maximum intensity projections of confocal z-stack images of cells labeled by SiR-tubulin. Scale bar is 20 μm. (**L**) A plot showing that nocodazole treatment shortens fibers compared to controls. Scaled fiber length was measured by dividing the length of the longest fiber by the length of the entire cell. Nine cells were analyzed for each condition. The black bars are median values. The p-value was calculated by Wilcoxon rank-sum test. (**M**) A representation of the changes in the nature and organization of functional states as well as transitions between them highlights the effects of nocodazole treatment. This panel is partially adapted from Fig. 3H and was produced by following the same analysis procedure. Any functional states and transitions depicted in Fig. 3H no longer observed under nocodazole treatment appear faded. Arrows that are not faded represent the new highest probability transitions associated with the states. Outer colors of the nodes corresponding to the numbered gait states denote the type or absence of broken detailed balance associated with the gait state for untreated cells while the core color represents that for the treated cells. As in Fig. 3H, black denotes the lack of unbalanced transitions, red denotes the state is only a source or only a recipient of unbalanced transitions, while blue denotes states which are both sources and recipients of unbalanced transitions. Note the loss of the cycle states as well as the functional states directing return to the cycle states, the emergence of the high cirral activity state involving activity in nearly all cirri, the reduction in gait states associated with the cloud of states of substantial cell movement (indicated by the reduction in the size of the black blob), and the

emergence of many high probability transitions to gait state 18.

Acknowledgements: We would like to thank Nicole King for the use of the confocal microscope. We would like to thank the students and faculty of the 2016 Marine Biological Laboratory Physiology Course and current and former members of the Marshall Lab and King Lab for comments, critiques, and encouragement during the development of this project. We would also like to thank David Booth, Greyson Lewis, Dennis Bray, and members of the Fourmentin-Guilbert Scientific Foundation for comments on the manuscript.

Funding:

I2CELL Seed Award of the Fourmentin-Guilbert Scientific Foundation (WFM)

NIH grant R35 GM130327 (WFM)

NSF grant MCB- 2012647 (WFM)

Merck Fellowship of the Jane Coffin Childs Memorial Fund for Medical Research (BTL).

Author contributions:

Conceptualization: BTL, JDP, WFM

Methodology: BTL, JG, WFM

Investigation: BTL

Formal analysis: BTL, JG, WFM

Visualization: BTL

Funding acquisition: WFM, JDP

Project administration: WFM, BTL, JDP

Writing – original draft: BTL, WFM

Writing – review & editing: BTL, JG, JDP, WFM

Competing interests: The authors declare no competing interests.

Supplementary Materials:

Materials and Methods

Figures S1-S10

Movies S1-S4.