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

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

Summary

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, it is unclear how cells might embody computational processes such that theories of computation, including finite state machine models, could be productively applied. 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 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 and a large group of low-probability time-symmetric transitions, thus revealing stereotypy in sequential patterns of state transitions. Simulations and experiments suggest that the sequential logic of the gait is functionally important. Taken together, these findings implicate a finite state machine-like process. Cirri are connected by microtubule bundles (fibers), and we found that the dynamics of cirri involved in different state transitions are associated with the structure of the fiber system. Perturbative experiments revealed that the fibers mediate gait coordination, suggesting a mechanical basis of gait control.

[1] Department of Biochemistry and Biophysics, University of California, San Francisco, California, USA
[2] Computer Science Department, Brandeis University, Waltham, Massachusetts, USA
* Correspondence: wallace.ucsf@gmail.com

Keywords

cellular behavior, information processing, computation, decision-making, broken detailed balance, gait coordination, cell motility, cilia, ciliate, tubulin

Introduction

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] (Figure 1A, Videos 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] (Video S2). Oral membranelles (Figure 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 (Figure 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 among states. In particular, a state representation allows us to ask whether forward and reverse transitions among states are equal, a condition known as detailed balance [4,46]. Systems that violate detailed balance operate in a non-equilibrium mode, display net probability flows, 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. Although the behavioral repertoire of *Euplotes* in different environmental conditions represents a rich and complex phenomenology involving information processing [51], for experimental tractability, we chose to focus on the reproducible, spontaneous linear walking behavior of cells. We reasoned that cellular walking might require some form of information processing to properly coordinate the movements of cirri.

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 [5,52] but contrast with many well-characterized examples of cellular and organismal motility [7,9,59,10,12,53–58]. Our results provide a demonstration of state machine-like processes governing cellular state transitions as well as an illustration of how such a computational perspective can drive mechanistic insight 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 (Figure 2A,B, Video 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 (Figure 2A). The consistency of the relative spatial positioning of cirri across cells allowed us to give each of the 14 cirri an alphabetic label from a-n (Figure 2C). During walking, cirri move in a manner analogous to the recovery stroke-power stroke cycle executed by many eukaryotic flagella, first lifting off the substrate and sweeping close to the cell body before extending in roughly the direction of cell orientation before sweeping downward to reestablish contact with the substrate [37,60,61] (Figure 2B, Video S1). In each video frame, the walking state of the cell was encoded as a 14-element binary vector, with each element 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 (Figure 2D). Cirral dynamics seemed to lack any obvious patterns such as periodicity or repeating sequences of states (e.g. Figure 2D), implying that the state sequences are generated either by stochastic processes or complex deterministic mechanisms. This lack of periodicity (confirmed by autocorrelation analysis, Figure S1) or fixed phase relationships between appendage movements is different from those reported for various unicellular flagellates and the gaits of most animals [58,62,63].

Despite the apparent complexity of cirral dynamics, we wondered whether there might be some underlying structure, which would allow us to effectively coarse-grain the dynamics in a principled manner. We first sought to obtain a reduced state space that could accurately describe the dynamics, as has proven successful in behavioral analysis of diverse living systems [44,45,64–67]. Because our ultimate goal was to identify interpretable motifs among the patterns of cirral activity, which entail strictly nonnegative values, we chose to perform dimensionality reduction by non-negative matrix factorization (NMF), a technique that has been used to identify patterns for textual analysis, natural language processing, neural activity analysis, and gene expression analysis [68,69]. NMF revealed the cirral states to be well-described by a three-dimensional NMF space (Figure 2E-G, see Method Details and Figure S1 for more details). Factors of the NMF (H1, H2, H3) constitute the basis of the NMF space and correspond to non-overlapping groups of cirri. These factors indicate features of cirral activity that in various weighted combinations can be used to compactly represent all of the cirral activity measurements. These groups of cirri constitute spatially distinct partitions of cirri with respect to their positions on the cell body (Figure 2H). The dimensionality reduction of the gait state space arises in part from shared pairwise mutual information between groups of cirri (Figure 2I). Here, mutual information

143 quantifies the amount of information about the patterns of activity of one cirrus gained by
144 measuring the activity of another, and in this way acts like a generalized measurement of
145 correlation. Therefore, dimensionality reduction by NMF reflects correlations in cirral activity.

146 Noting the apparent structure of the data visualized in the reduced gait state space in the
147 form of clusters of points in NMF space (Figure 2E-G), we next sought to identify individual gait
148 states defined in terms of similar patterns of cirral activity indicated by these clusters. To do this,
149 we applied the density-based spatial clustering of applications with noise (DBSCAN) algorithm
150 [70] to group the output of NMF into clusters in an unbiased fashion, with members of a given
151 cluster sharing similar patterns of cirral activity (see Method Details and Figure S1 for more
152 details). Although NMF can itself be used to define clusters, the advantage of clustering in two
153 steps, NMF followed by DBSCAN, is that we could use the visual observation of clustering in
154 the NMF output to confirm correct performance of the subsequent clustering step. The effect of
155 this two-step approach to coarse graining, which is a commonly used in clustering of single-cell
156 transcription data, is to reduce noise while retaining features that can accurately capture the
157 complexity of cirral activity. Visual inspection in conjunction with silhouette coefficient (a
158 metric of cluster cohesion and separation) analysis [71] revealed that 32 clusters accurately
159 captured the visible structure in the reduced state space without overfitting (Figure 2E-G,
160 Method Details and Figure S1). These reduced gait states correspond to distinct patterns of cirral
161 activity (Figure 2J).

162 Taken together, our results reveal stereotypy in the spatiotemporal patterns of cirral
163 activity, consistent with the behavior of a finite state machine. The discrete set of gait states,
164 which exist in the reduced gait state space, demonstrate that cells make use of a subset of the
165 possible patterns of appendage movement during walking locomotion.

166 ***Euplotes* walks with a cyclic stochastic gait**

In order to relate the gait states identified in our cluster analysis to cell motility, we asked how changes in the number of active cirri may relate to cell movement. At low Reynolds number, velocity of a suspended particle is proportional to the difference between the net force acting on it and the opposing drag [43], so cell velocity provides a relative readout of net locomotive force. Naively, one might expect that the force associated with locomotion is roughly proportional to the number of moving appendages [72]. 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 [62,73] 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 changes in the number of active cirri (Figure 3A). We hypothesized 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 dictating gait state transitions, we calculated the forward and reverse transition rates between states from the 1423 pairwise transitions in our dataset as N_{ij}/T where N_{ij} is the total number of transitions observed between states i and j, and T is the total observation time (Figure 3B, see Method Details). We note that the transition rate we define here is proportional to the probability current $J_{ij} = P_i k_{ij}$ where P_i is the probability of state i and k_{ij} is the conditional transition probability analogous to a chemical reaction rate. Although walking, and cilia-based motility in general, is fundamentally non-equilibrium in the sense that it requires energy input, a coarse-grained representation of motility such as the one we pursue here need not exhibit broken detailed balance [5]. However, broken detailed balance in gait state transitions (unbalanced forward and reverse transitions) would

indicate a non-equilibrium process dictating temporal directedness (irreversibility) to gait state transitions, thus implicating active coordination in terms of stereotyped sequential changes in cirral activity. Importantly, broken detailed balance in gait state transitions define an identifiable temporal direction in sequences of gait state 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 Method Details). Unbalanced transitions can also arise in equilibrium systems that have not yet reached steady state, in which case transition rates may change in time. To test whether the gait operates at steady state, we checked whether the total number of transitions into each state were balanced by total transitions out of that state: $\sum_j N_{ij} = \sum_j N_{ji}$. Consistent with steady state dynamics, we found that this condition held to within a difference of at most a single transition. Although there are cases where non-equilibrium state transitions can nevertheless exhibit balanced transitions, the presence of unbalanced transitions existing as part of loops in state space unambiguously indicated non-equilibrium dynamics. To illustrate this non-equilibrium or temporally irreversible character of gait state transitions, we apply the Kolmogorov criterion, which specifies the necessary and sufficient condition for reversibility that the product of transition probabilities traversing any closed loop in state space must equal the product of the transition probabilities in the reverse direction [5,74]. Due to the presence of unbalanced transitions, the gait of *Euplotes* clearly violates this condition; for example, $k_{1,2}k_{2,3}k_{3,17}k_{17,1} = 0.003 \neq k_{1,17}k_{17,3}k_{3,2}k_{2,1} = 0.0005$, where each $k_{i,j}$ is the conditional probability of transitioning from state i to state j estimated as N_{ij}/N_i with N_i the total number of transitions from state i. 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 nats, similar to the value reported for strongly non-equilibrium gait transitions observed in a flagellate [5]. Walking *Euplotes* cells, therefore, have a non-equilibrium gait, displaying temporal order in sequences of appendage movement, despite a lack of standard gait periodicity.

We next sought to better understand the organization and sequential logic of this unusual gait. First, we noted that gait state transitions appear 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 (Figure 3C). Crucially, the presence of broken detailed balance revealed directed cycles of cirral activity during locomotion pushing the system away from equilibrium, yielding a temporal directedness or irreversibility to sequences of gait state transitions in spite of the lack of standard gait periodicity. To investigate the nature of these cycles, we grouped transitions into two categories: balanced transitions, which satisfy detailed balance, and unbalanced transitions, which do not (see Method Details). This partitioning allowed us to separately investigate unbalanced, non-equilibrium-like and balanced, equilibrium-like transitions (Method Details, Figure 3D,E). We found that unbalanced transitions occur at relatively high frequency but involve a small number of states (Figure 3D). Of the 32 gait states, only states 2, 3, 4, 7, 17, and 27 (Figure 3F) are associated with unbalanced transitions, and among these states, three formed a 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 (Figure 3D,E) and that relatively few transitions corresponded to substantial cell movement (Figure 3C). Additionally, we found that, with the exception of transitions between states 1 and 2, the highest frequency transitions are unbalanced (Figure 3D).

239 Notably, the most frequent balanced transitions were associated with transitions into and
240 out of gait state 1, a unique "rest state" which involves no cirral movement (Figure 2J).
241 Furthermore, we found by autocorrelation analysis of gait trajectories that gait state 1 uniquely
242 exhibited significant positive autocorrelation (see Method Details, Figure S1). Investigation of
243 transitions to and from gait state 1 revealed that although transitions between gait states 1 and 2
244 are balanced, the most frequent transitions out of gait state 2 are strongly biased toward
245 transitions into gait state 3, from which other strongly biased transitions also frequently occur,
246 including the cycle of biased transitions mentioned above. Despite the presence of high
247 frequency unbalanced transitions, the gait of *Euplotes* involves highly variable trajectories
248 through gait state space. The picture of walking trajectories that emerges is one of stochastic
249 excursions from gait state 1 into non-determinate paths through state space involving a mix of
250 balanced and unbalanced transitions. The majority of cell movement occurs during infrequent,
251 equilibrium-like (balanced) transitions. These balanced transitions are temporally reversible, but
252 we wondered whether there might be some asymmetry in the amount of cell movement due to
253 transitions leading toward versus away from the cycle states. We found that although balanced
254 transitions leading away from the cycle states account for 7.6% more cell movement, the
255 difference in these transitions is not statistically significant ($p=0.09$ by two-sample Kolmogorov-
256 Smirnov test). Temporal irreversibility, or directedness, in sequential changes in cirral activity
257 arises from biased, non-equilibrium-like (unbalanced) transitions, occurring at relatively high
258 frequency from a small subset of states. A natural question arising from the observation of these
259 complex patterns is whether gait state transitions conform to a first-order Markov process, often
260 referred to as "memoryless" [75], which entails that transition probabilities are determined
261 completely by the present state, and that previous dynamics contribute no additional predictive
262 information [76,77]. Does the cell retain some "memory" of past cirral activity that might influence

263 future cirral activity? While not conclusive, several analyses (see Method Details and Figure S1)
264 suggested that *Euplotes* retains some "memory" of the prior sequence of cirral movements
265 during locomotion in that the gait may not conform to a continuous- or discrete-time first-order
266 Markov process.

267 Taken together, our analysis revealed a mixture of unbalanced transitions driving cycles
268 and balanced transitions arranged as networks, for which we propose to apply the term "cyclic
269 stochastic gait". The cyclic stochastic gait of *Euplotes eurystomus* incorporates elements of both
270 stereotypy and variability in gait dynamics, in terms of biased transitions and non-determinate
271 sequences of gait state transitions respectively. Forward progress of the cell is not produced
272 merely by a physical ratchetting process driven by unpatterned fluctuations in cirral activity, nor
273 is it produced by a highly regular, deterministic process like a clock. It has been argued that
274 significant computation arises in physical systems exhibiting such a mix of stereotypy and
275 variability [48,78–80] in the sense that the time-evolution of the system is most compactly described
276 as the result of a computational process involving state transitions, memory, and decision rules,
277 rather than periodic oscillations or random coin flips.

278 Our analysis suggested a computational underpinning of gait, so we sought to better
279 understand the functional organization of the dynamical patterns, the sequential logic of the gait.
280 Focusing on the dominant structure of gait transitions in terms of transition probabilities (Figure
281 3G, Method Details) allowed us to derive a simplified, functional representation of stereotypy in
282 gait dynamics as depicted in Figure 3H. We found that few states were the recipients of the
283 majority of the highest probability transitions and that many received no high probability
284 transitions (Figure 3G). Additionally, we found a "cloud" of states linked by low-probability
285 balanced, equilibrium-like fluctuations. Nearly all of the states receiving high probability
286 transitions were either the three "cycle states" or else fed cycle states with their highest

287 probability transitions, with the majority feeding gait state 17. Although gait state 1 is not the
288 recipient of any individual high probability transitions, we identified it as the unique "start" state
289 from which cells initiate walking. Beginning with this start state, cells transition with high
290 probability to gait state 2, also one of the highest frequency transitions and the first state in the
291 2→3→17→2 cycle of unbalanced transitions. From this first cycle state, cells transition to gait
292 state 3, the second cycle state, with highest probability and frequency and then similarly on to
293 gait state 17, the third cycle state. This sequence from the start state through the cycle states
294 corresponds to increasing amounts of cirral activity. Although the highest probability transitions
295 from the third cycle state to any single gait state tend to return to the first or second cycle state
296 with equal probability, cells in fact transition to the equilibrium "cloud" of motility-associated
297 states with overall higher probability. Return to the cycle states tend to occur through various
298 moderately high probability transitions from the motility state cloud or through a restricted set of
299 intermediate states. In conjunction with this set of transitions, we also noted unbalanced
300 transitions stemming from the cycle states to the motility state as well as the presence of
301 intermediate states from a given cycle state that subsequently feed the next cycle state.

302 Altogether, the picture of stereotypical gait dynamics that emerges is of biased transitions
303 involving cycle states preceding relatively low probability, unbiased transitions associated with
304 substantial cell movement before returning to the start or cycle states and beginning the sequence
305 again. While this general sequence is repeated during walking, there is considerable variability or
306 apparent stochasticity in the details of gait state transitions departing from the start state with
307 increasingly variable transitions as any given sequence progresses. Given that the motion of the
308 cell alternates between periods of highly processive forward motion and other periods in which
309 the cirri seem to flail ineffectually without driving the cell forward, one might assume that the
310 periods of effective motion would correspond to the highly directed cycle of unbalanced

transitions and the periods of flailing would correspond to the cloud of balanced, random transitions, but in fact the opposite is the case. We hypothesized that sequences involving the cycle states serve to establish configurations of cirri necessary for cells to later transition between states from which substantial 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. We note here that although our coarse-graining procedure to identify gait states does not constitute a unique representation of gait structure, we can be confident that our analysis captures structure in gait dynamics, the presence of broken detailed balance in particular. In general, coarse-graining of the state space of a system can obscure broken detailed balance, but the net flux of transitions in the state space of a system should not arise in an illusory fashion based on a coarse-graining procedure or partial observation of a system [81–83]. Additionally, we have chosen a coarse-graining procedure based on the properties of our particular data and demonstrate its performance on simulated cirral dynamics with varying noise and for the case of unpatterned, random cirral fluctuations (Figure S1, Method Details).

Finally, we sought to investigate the functional significance of the sequential logic of the gait in driving processive cell movement. To do so, we developed a simple model based on a coarse-grained physical picture of cellular walking (Figure 3I). Briefly, we consider a 2D system where a cell walks using its 14 cirri, which can exist in two states: moving or not moving. As we are agnostic to the details by which cirri generate forces involved in motility, we model cirri as producing constant force in the direction of cell orientation while moving and resisting displacement by acting as linear springs when not moving. Cell position and orientation is defined in terms of the equilibrium positions of the cirri. For simulations, we use relative positions of cirri taken from stationary cell measurements (Figure 3J). In each timestep of the

335 simulation, we calculate a cell displacement based on the sum of forces due to the cirri and
336 changes in cell orientation from the sum of torques (see Method Details). We found that this
337 simple model was sufficient to qualitatively and semi-quantitatively reproduce the linear
338 trajectories of walking cells when we ran simulations using the actual patterns of cirral activity
339 from walking cells (Figure 3K, Figure S2). When we ran simulations with either the same gait
340 states as those from actual cells but in a shuffled order or random cirral activity with the same
341 average cirral activity as actual cells, we found that path straightness significantly decreased in
342 the case of shuffled transitions (p=0.04 by Wilcoxon rank sum test) and scaled cell velocity
343 significantly decreased in the case of random activity (p=0.003 by Wilcoxon rank sum, Figure
344 3K-O). Note that in both the case of shuffled transitions and random cirral activity, gait state
345 transitions satisfy detailed balance (Figure S2). These results suggest that the sequential logic of
346 the gait is functionally significant in a manner consistent with our biophysical picture of walking.

347 **The microtubule-based fiber system mediates gait coordination**

348 The complex yet sequentially structured gait patterns in conjunction with our simulation
349 results are consistent with the existence of some form of nontrivial gait coordination. What
350 physical machinery could embody the information processing required to generate the stochastic
351 cyclic state transitions seen during *Euplotes*' walking? We reasoned that there must be some
352 form of coupling or communication between cirri or feedback between gait states and cirral
353 dynamics. Since the early 1900s , the role of the system of cytoskeletal fibers associated with
354 cirri as conduits of information between cirri during cellular locomotion, supported by
355 microsurgical experiments, has been a dominant yet contentious hypothesized mechanism of gait
356 coordination [84–86]. We wondered whether the structure of the cytoskeletal fiber system associated
357 with cirri (Figure 4A) could give some insight into how cirri might be coordinated.

358 We reconstructed in 3D the microtubule-based fiber system of *Euplotes* associated with
359 cirri and lying just beneath the cell cortex [38,41,42]. Upon inspection of our confocal
360 reconstructions of SiR-tubulin labeled cells (Figure 4A, Figure S3), we noted the presence of two
361 morphologically distinct classes of fibers, one thicker and linear and the other thinner, splayed,
362 and less linear, consistent with previous observations (Figure 1C, [38,41,42]). Fibers emanate from
363 the base of all cirri, appear to intersect one another, and also connect to the cortex of the cell at
364 various points. Some cirri were found to be associated with only thick fibers, while others have
365 both or only thin fibers. Based on apparent fiber intersections and convergences, we found the
366 fiber system forms a continuous network between all cirri, with the fibers associated with the
367 base of each cirrus intersecting the fiber system of at least one other cirrus (Figure 4A,B).
368 Contrary to the long-standing standing hypothesis from the literature [84], the functional
369 modules (groups of co-varying cirri) identified in our dynamical analysis were not exclusively
370 linked by dense fiber intersections (Figure 4A,B) [38,42,84]. In fact, connections between cirri are
371 not generally associated with any statistically significant difference in mutual information
372 (defined in terms of the information that the activation state of one cirrus has concerning the
373 other) compared to unlinked pairs of cirri ($p=0.14$ by Wilcoxon rank sum test, Figure 4C).
374 However, information flow became apparent when fiber-fiber links were grouped by type (i.e.
375 thick to thick fiber, thick to thin fiber, or thin to thin fiber). Under this grouping, we found that
376 pairs of cirri associated with only thick to thick fiber and only thin to thin fiber links have
377 increased mutual information compared to those without links (Figure S3). Interestingly, we
378 found that cirri nearby one another and connected by fibers to similar regions of the cell cortex
379 shared the most mutual information (Figure 2C,G, 4D,E), suggesting that if the fibers play a role
380 in cirral coordination, coupling may also be mediated by mechanisms involving the cirrus and
381 fiber-cortex interface. Cirri d, e, h, i, for example, share very little mutual information with any

382 of the other cirri, and fibers emanating from the base of these cirri contact the cell cortex and
383 other fibers at various unique points. On the other hand, cirri g and f, which share more mutual
384 information than any other pair, are associated with both thick and thin fibers terminating at
385 similar regions of the cell cortex. Indeed, distances between pairs of cirri and cross nearest-
386 neighbor distance [87] between paired sets of cirrus-cortex contact points both show significant
387 Spearman correlations (-0.49, $p<0.001$ and -0.62, $p<0.001$ respectively) to mutual information
388 (Figure 4D,E). These correlations indicate that mutual information between pairs of cirri tends to
389 increase with proximity and also tends to increase with similarity between fiber-cortex contact
390 locations. Thus, the cirri with the highest mutual information are those that are close together
391 with similar fiber-cortex connections (Figure 4D-F).

392 Together, these observations suggest a mechanism of mechanical coordination in which
393 microtubule bundles allow groups of cirri to influence successive behavior of other groups of
394 cirri. We first sought to test this hypothesized mechanism by perturbing the fiber system using
395 drug treatments. We observed that nocodazole, a drug that blocks the polymerization of
396 microtubules, affected walking motility, causing cells to walk along more compact trajectories
397 due to increased turning (Figure 5A,B, Videos S3 and S4) in a manner reminiscent of historical
398 reports of altered walking due to microsurgery in ciliate *Stylonychia* [88]. In contrast to the effect
399 of nocodazole, we found that the microtubule stabilizer paclitaxel caused cells to walk along less
400 convoluted trajectories compared to controls (Figure 5C). Quantifying these effect from video
401 microscopy of cells under darkfield illumination in terms of a scaled path length defined as the
402 total integrated path length walked by cells scaled by the maximum radial distance traversed,
403 where a decrease in linear runs decreases this scaled path length, we found that nocodazole
404 significantly decreased and that paclitaxel significantly increased the scaled path length of cells
405 compared to controls (Figure 5D). We further noted that nocodazole acted in a dose dependent

406 and reversible manner (Figure 5D and Figure S4). To test whether the actin cytoskeleton might
407 also regulate motility, we treated cells with actin inhibitors including latrunculin, cytochalasin,
408 and jasplakinolide and observed no effect on motility (Figure S4). The specific and opposing
409 effects of a microtubule polymerization inhibiting drug (nocodazole) versus a microtubule
410 stabilizing drug (paclitaxel) were consistent with the microtubule cytoskeleton playing a key role
411 in coordinating walking motility. Next, we checked whether nocodazole treatment had an
412 observable effect on the fiber system. By analyzing 3D reconstructions of confocal images of
413 SiR-tubulin labeled cells, we found that fiber length significantly decreased compared to controls
414 in cells where microtubule polymerization was disrupted by nocodazole (Figure 5E-G). Further,
415 we were unable to detect the presence of any thin fibers in four out of seven nocodazole treated
416 cells used for fiber analysis. Of the three cells with detectable thin fibers, we never observed
417 connections between fibers associated with cirri a, b, and c and any other fibers. Additionally, we
418 did not observe any thin fibers making distal cortical contacts. We did note, however, that when
419 thin fibers were visible, connections appeared to be the same as those in Figure 4A,B. We then
420 investigated how cirral dynamics were affected. Following the gait annotation procedure detailed
421 previously, we characterized the walking dynamics of 6 nocodazole treated cells for a total of
422 1133 timepoints. Of those timepoints, 681 corresponded to cirral configurations never observed
423 in untreated cells with a total of 391 new unique configurations. Projecting these new
424 configurations onto the NMF space we obtained previously, however, revealed that most of the
425 cirral configurations in nocodazole treated cells were near or within the clusters corresponding to
426 the gait states we obtained from untreated cells (Figure S4). This allowed us to map new cirral
427 configurations onto the gait states (see Method Details). We did note that while close to the
428 original gait states, new cirral configurations tended to skew towards more active cirri, and we
429 also noted the presence of a new cluster involving movement in all or nearly all cirri, to which

430 we assigned a new gait state (Figure S4). We also found that mutual information between cirri
431 was higher in general, with many pairs of cirri sharing higher mutual information than the
432 highest values obtain in untreated cells (Figure 5H, Figure S4). This increase in cirral activity
433 and correlations is consistent with the fibers playing a role in conveying inhibitory information
434 during unperturbed walking. In further support of this inhibitory role, we found that paclitaxel
435 treatment also caused an overall increase in mutual information of pairs of cirri (Figure 5H).
436 Next, we investigated how the dynamics of gait state transitions were affected. Following
437 our previous analysis, we evaluated gait state transitions for the presence of broken detailed
438 balance, separating state transitions into balanced and unbalanced transitions and found that the
439 structure of state transitions differed greatly from that of unperturbed cells. Although some gait
440 state such as 1, 2, and 17 were involved in high frequency transitions in both conditions, gait
441 state transitions of cells with perturbed fiber systems exhibited less broken detailed balance and
442 closer to equilibrium-like dynamics as indicated by an entropy production rate of 0.1 nats
443 (compared to 0.4 nats in the unperturbed case), including the loss of the unbalanced, cyclic
444 transitions (Figure S4). Figure 5I summarizes the change in the structure of gait state transitions
445 including changes in broken detailed balance, reduction in transitions toward the states
446 previously involved in cyclic transitions, and the reduction in the occupancy of states associated
447 with the cloud of states involved in substantial cell movement, with only 10 of the original 21
448 cloud states sampled. We also noted that many of the new highest probability transitions feed
449 gait state 18 (Figure 5I, Figure S4), which involves nearly exclusive activity in cirri f and g
450 located close to one another at the edge of the cell. A persistent bias toward activation of these
451 cirri, which are farthest from the central axis of the cell, may in part explain the increase in
452 turning in trajectories. These results, in conjunction with the fact that cell velocities were
453 indistinguishable control conditions (Figure S4) are consistent with predictions from simulations

454 in which shuffling the order of gait state transitions reduced processivity without affecting speed
455 (Figure 3N,O).

456 Although the result that disrupted gait coordination involves increased mutual
457 information between cirri, note that this increase is due to more frequent joint activation or
458 synchronous movements among cirri and so stems from reduced complexity of gait dynamics.
459 For a complementary perspective linking these changes in gait dynamics to an underlying
460 computational process, we applied the Causal State Splitting Reconstruction (CSSR) algorithm
461 to construct ϵ-machines corresponding to walking cells [48,89]. ϵ-machines are automaton models
462 consisting of a set of causal states with transitions between them and represent the minimal
463 model consistent with accurate prediction of a stochastic process [48]. The causal states of an ϵ-
464 machine indicate how the process from which it is constructed stores information, and state
465 transitions indicate how the process transforms information [48]. We found that ϵ-machines
466 constructed from the cirral activity of untreated cells tended to be similar to one another and
467 were more complex in terms of having more causal states and transitions than those constructed
468 from nocodazole treated cells, which also tended to be similar to one another (Figure S5, Method
469 Details). This reduction in complexity may reflect reduced computational capacity of the
470 nocodazole treated cells. Further, in conjunction with the decrease in entropy production under
471 nocodazole treatment, these results are consistent with recent findings linking increases in
472 broken detailed balance to increases in information processing in biological systems [90].

473 As a final additional test of the role of the fibers in mediating gait coordination, we
474 revisited historical microdissection experiments. Although Taylor first reported in 1921 results of
475 microdissection experiments indicating a role of the fiber system in mediating gait coordination
476 [84], a reproduction of these experiments in 1966 with improved physiological conditions and
477 quantitative analysis failed to observe disrupted coordination among cirri upon bisection of cells

478 [85]. Importantly, neither of these reports involved analysis of walking behavior. We performed
479 microdissections on cells using pulled quartz microneedles, severing cells transversely just in
480 front of cirrus h, ensuring that we had severed all fibers associated with cirri j-n (Figure 5J,K).
481 Similar to previous reports, we found that cell fragments regained spontaneous cirral activity
482 after a brief recovery period. After 24 hr, we found that anterior portions of cells began to exhibit
483 spontaneous walking activity (Figure 5K,L, Video S6), which persisted for up to 72 hr
484 subsequently and has not been previously reported. Although these fragments were able to walk
485 across surfaces, in all cases, they walked along curling or circular trajectories reminiscent of
486 nocodazole treated cells except in the reverse direction (Figure 5L, Video S6). We analyzed the
487 cirral dynamics of 10 fragments and found significantly higher mutual information shared
488 between cirri indicating more synchronous cirral activity, similarly microtubule inhibitor treated
489 cells, compared to cells that were merely wounded and displaying apparently normal motility
490 (Video S7) or unperturbed ($p<0.001$ in both cases by Wilcoxon rank sum test), which displayed
491 values consistent with one another (Figure 5M). As with the microtubule inhibitor experiments,
492 these results are consistent the fibers playing an inhibitory role in gait coordination.

493 Taken together, these results are consistent with our hypothesis that the fiber system
494 mechanically mediates gait coordination. Further, these results provide additional evidence that
495 proper, active coordination of cirral dynamics and emergent gait states and gait state transitions
496 is required for proper cellular walking.

497 Discussion

498 Faced with the challenge of accounting for the emergence of apparently sophisticated
499 cellular behavior, the directed walking of *Euplotes* driven by a seemingly disordered gait, we
500 conceptualized the cell as a finite state machine. Traditionally, studies of computational
501 processes performed by cells have focused on combinatorial logic, where the output of a

502 computation depends only on the current input, performed by networks of molecules in cells
503 [25,31–33]. We have focused on sequential logic, where outputs depend on the system state as well,
504 an equally important aspect of the theory of computation with notable yet less developed
505 representation in studies of cellular dynamics [16,20,91]. Automata theory, which includes finite state
506 machine models and necessarily involves sequential logic, provides tools for understanding
507 structure and stereotypy in transitions between dynamical states, increasingly appreciated as
508 features of the behavior of eukaryotic cells. Our approach revealed modularity in cellular
509 dynamics associated with structural modularity of the cell (Figure 2, 4) in addition to stereotyped
510 patterns of sequential activity (Figure 3) and yielded new insight into regulatory mechanisms
511 (Figure 3, 5).

512 Related approaches involving the coarse-graining of complex dynamics and subsequent
513 analysis have proved useful in revealing simplicity and stereotypy in the behavior and movement
514 patterns of various organisms including flies, worms, and protists[3,5,65,66,92]. Reducing the
515 complexity of biological dynamics in this way can be an important step toward making
516 quantitative predictions and for interpreting the effects of perturbations. In general, particular
517 approaches to coarse-graining will depend on details of the biological system, type of data
518 collected, and underlying patterns in the data. The specific steps we have taken in the present
519 study are best suited to biological processes that can be characterized in terms of discrete events.
520 The analyses following coarse graining stand to provide the most insight into processes
521 involving information processing where the sequential order of events may be important to the
522 resultant function. Additionally, our perturbative experiments demonstrated how these
523 approaches can drive mechanistic insight by helping connect function and computation to
524 underlying mechanics and physiology.

525 Although there are examples of locomotor coordination reminiscent of the stochastic,
526 non-equilibrium gait dynamics of *Euplotes*, such as gait switching in Kangaroo rats [52] or, most
527 saliently, gait switching in an octoflagellate [5] and motility in cultured mammalian cells [44],
528 walking locomotion in *Euplotes* represents a departure from many of the best studied appendage-
529 based locomotor systems. For example, limbed locomotion in animals tends to proceed by highly
530 stereotyped, determinate patterns of activity [58,62], and many small, aquatic animals exhibit
531 periodic movements of appendages, often cilia, during locomotion [7,63,93]. Many forms of
532 unicellular locomotion involve such dynamics as well including in sperm cells [94], diverse
533 flagellates with various numbers of flagella [63], and ciliates [63,95,96]. Even in cases where cellular
534 locomotion involves fundamentally stochastic dynamics such as in run-and-tumble motility in *E.*
535 *coli* [12] or analogous behaviors observed in protists [11,97–99], motility can be described by
536 equilibrium processes [5], in contrast to the non-equilibrium character of the gait of *Euplotes*.
537 Here, broken detailed balance in gait state transitions revealed cyclic activity with a
538 mixture of stereotypy and variability in the gait of a single cell (Figure 3). To explain how these
539 dynamics give rise to directed walking, we propose a mechanism, captured by our biophysical
540 model, in which biased, actively controlled cyclic transitions serve to establish strain, effectively
541 storing stress, in certain cirri, and the spontaneous release of these cirri from the substrate, during
542 a series of unbiased gait state transitions, allows the cell to move forward. The cloud of unbiased
543 transitions associated with substantial cellular movement is consistent with motility generation
544 not depending on the precise order in which the strained cirri are released from the substrate.
545 Return to the cycle states then are necessary to establish this process anew by winding up the
546 system for continued, proper cell movement. Disruption in this resetting may lead to defects in
547 walking as predicted by simulations (Figure 3K-O, 5) and with the reduction in entropy
548 production and loss of the cycle states in irregularly walking cells with disrupted fiber systems,

549 which appeared unable to consistently maintain linear trajectories (Figure 5B,D,I, Video S4). We
550 find additional experimental support for our proposed mechanism in previously reported
551 observations of cyclic velocity fluctuations in the trajectories of walking *Euplotes* [36].
552 We argue that subcellular processes are involved in actively coordinating cirri and
553 propose that broken detailed balance in the gait of *Euplotes* indicates this active coordination.
554 The results of experiments perturbing the tubulin-based cytoskeletal fiber system are consistent
555 with its role in mechanically mediating communication both among cirri and between cirri and
556 the cell cortex (Figure 4,5). We conjecture that movement of cirri relative to one another can
557 establish tension in the fiber system and that the tension state of fibers associated with each
558 cirrus may then modulate cirral activity in a manner reminiscent of basal coupling in flagellates
559 [22]. Further, It is possible that microtubules may mechanically mediate the coordination of cellular
560 processes in other eukaryotes, and microtubules in cells have recently been shown to respond
561 directly to mechanical forces [100]. Microtubules have also been shown and proposed to be
562 involved in more complex signal transduction pathways[101,102]. Changes in fiber length, as
563 observed under nocodazole treatment, change the relative positions of cirri, which could lead to
564 different distributions of tension and thereby induce the altered pattern of state transitions and
565 associated walking defects. Our results show that perturbation of the fiber system shift the gait of
566 *Euplotes* from a regime of asynchronous yet coordinated movement to a dysregulated regime
567 with synchronous yet improperly coordinated movement. It should be noted, however, that the
568 fact that cells with perturbed fiber systems, including bisected cells, display some walking ability
569 implicates additional robustness and complexity in mechanisms of gait coordination and cell
570 motility. Although details of the mechanism remain to be fully explored and tested, our results
571 are consistent with the fibers system playing an active and inhibitory role in gait coordination.
572 Thus, by combining information processing to properly dictate patterns of cirral activity and the

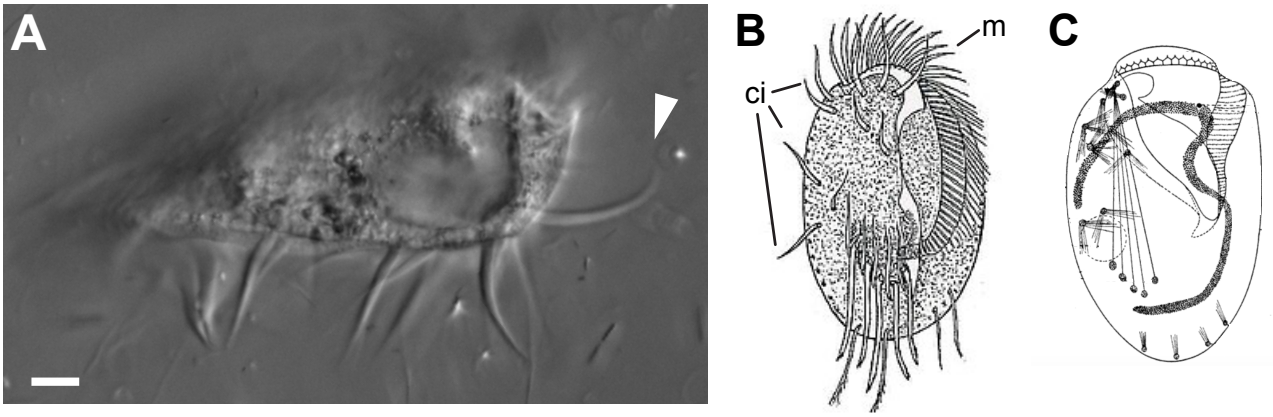
573 mechanical actions of cirral movement, walking *Euplotes* embodies the sequential computation
574 of a finite state machine. Furthermore, our approach to understanding a complex cellular
575 behavior, grounded in theory of computation, allowed us to derive mechanistic insight from
576 observational and perturbative experiments. Because reproducible biological function often
577 emerges from the productive management of stochastic fluctuations, we expect our conceptual
578 and analytical approaches may be useful in studying other living systems.

579 Our work lays a foundation for studies of sensorimotor behavior in *Euplotes*, which we
580 believe will shed light on principles of cellular behavior. We have focused here on a single
581 behavior (spontaneous, linear walking), but *Euplotes* exhibits a complex behavioral repertoire in
582 response to diverse environmental cues, contexts, and physiological states[39,40,103104]. Many of
583 these behaviors involve different patterns of cell motility. Our biophysical model of cell walking
584 could help predict and interpret how changes in gait state transitions or the emergence of new
585 gait states might give rise to these different patterns of movement. We speculate that the unusual
586 gait of *Euplotes* and associated asymmetric cell structure, contrasting sharply with deterministic
587 appendage movements and symmetric structure in animals and many swimming flagellates and
588 ciliates, may reflect physical and phylogenetic constraints. The framework we have developed
589 integrates across scales of space, time, and biological organization to link mechanics and
590 subcellular dynamics and structure to overall cell movement patterns. Therefore, our work
591 provides a way to begin to address the functional significance of cell morphology, gait patterns,
592 and mechanisms of gait control, and in conjunction with comparative approaches, could provide
593 new insight into how cellular behavior evolves.

594 Among the domains of life, eukaryotes uniquely display remarkable complexity and
595 diversity in cellular behavior [105]. Our approach, grounded in finite state machine analysis, has
596 revealed modularity and stereotypy underlying complex cellular behavior and has provided

insight into regulatory mechanisms across scales of biological organization. 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, understand, and even one day engineer cellular behavior across diverse eukaryotic systems.

605 **Figures**



606

607 **Figure 1. *Euplotes* exhibits highly polarized, complex cellular architecture and walks across**
608 **surfaces using microtubule-based organelles called cirri, some of which are physically**
609 **linked.** (A) A single *Euplotes eurystomus* cell in profile displays its ventral cirri, which are used
610 for walking locomotion across surfaces (arrowhead indicates a single cirrus stretching out from
611 the cell). Scale bar is 10 µm. (B) A drawing of a *Euplotes* cell, viewed from the ventral surface,
612 highlighting the complex, asymmetric structure of the cell. Notable features include the cirri (ci)
613 and the membranellar band (m), wrapping from the top of the cell to the center, which is used to
614 generate a feeding current to draw in prey items. Drawing adapted and obtained from Wikimedia
615 Commons, from original source [106]. (C) A drawing of a *Euplotes* cell, highlighting the fiber
616 system associated with the cirri, historically referred to as the neuromotor apparatus. Drawing
617 adapted from [107].

618

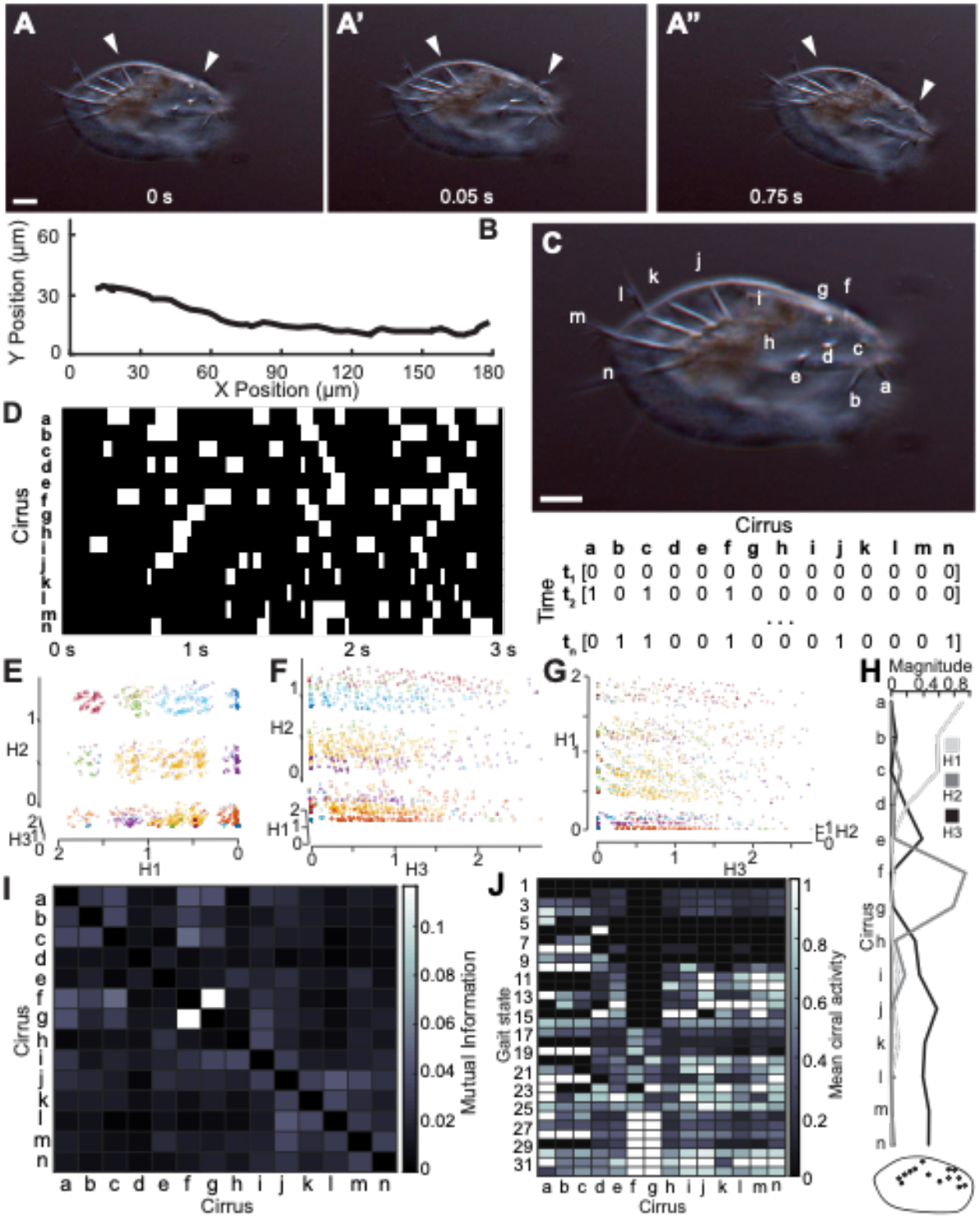


Figure 2. *Euplotes* walks across surfaces according to a complex gait, which can be described in a discrete, reduced state space with gait states corresponding to identifiable patterns of cirral activity. (A-A''), The movements of cirri during walking locomotion are

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 panel E. 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 panel F. 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 panel E, 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 [70]. We refer to these clusters as gait states, and they correspond to unique configurations of cirral activity during walking locomotion. (H) A plot of the magnitudes associated with each cirrus corresponding to the components of the NMF of cirral dynamics shows distinct contributions

647 from spatially distinct groups of cirri. Component H1, for example, is associated with activity in
648 cirri a, b, and c. The tracing of a cell including the position of cirri has the same color map as the
649 plot above and shows the grouping of the cirri corresponding to each component. (I) A heatmap
650 of mutual information between all pairs of cirri shows that correlations in cirral activity
651 correspond to the NMF components displayed in panel K. For example, cirri a, b, and c share
652 mutual information with one another and are the cirri contributing to component H1. (J) A
653 heatmap representation of the cirral activity associated with each of the 32 gait states. Values for
654 each cirrus are the mean over all instances of the gait state. Note that each gait state has a unique
655 signature of cirral activity. See also Figure S1, Video S1, and Video S2.

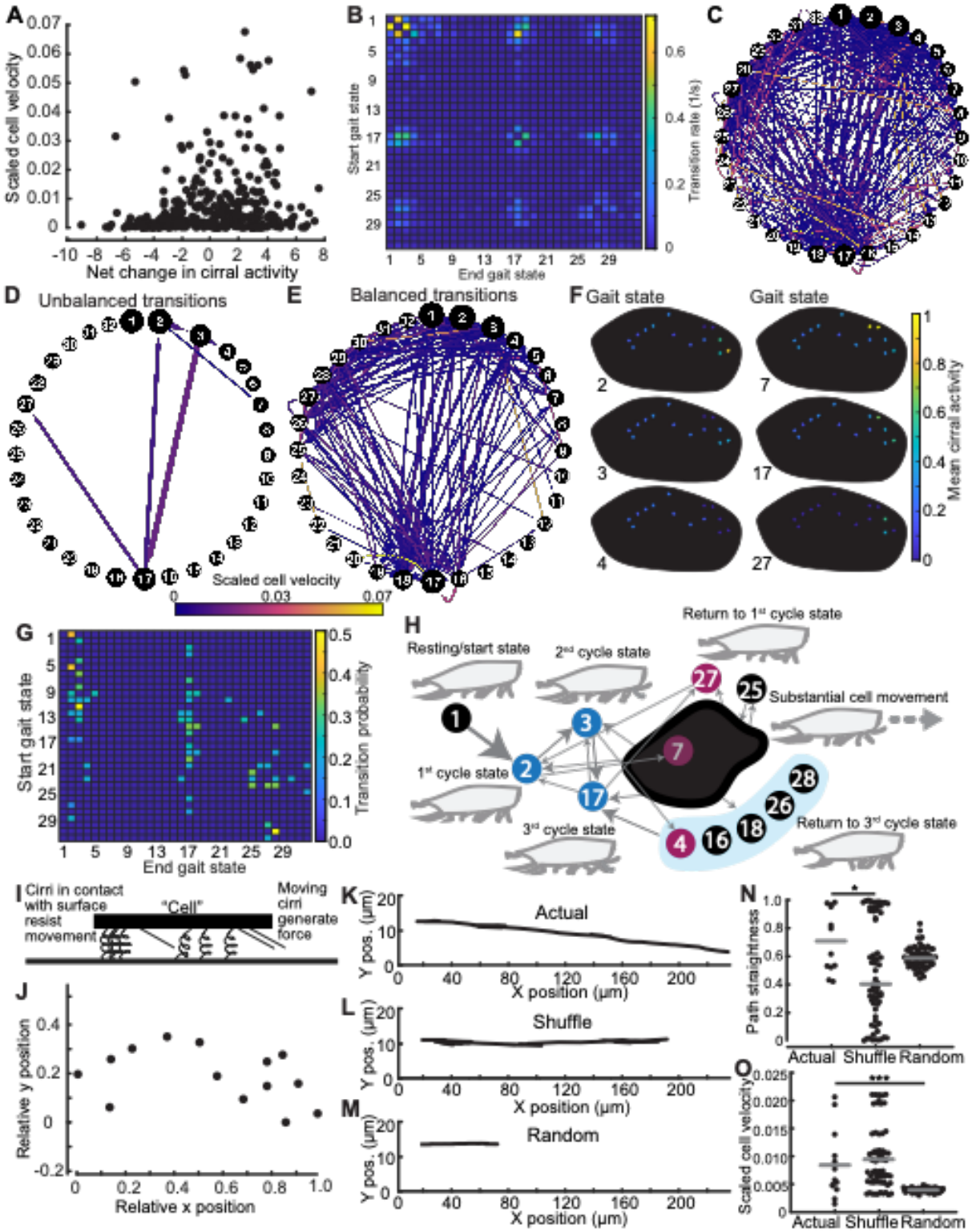


Figure 3. ***Euplotes*** **walks with a cyclic stochastic gait exhibiting broken detailed balance, stereotypy, and state machine-like dynamics.** All data for panels A-H are pooled

659 from the walking trajectories of 13 different cells over 2343 timepoints and 1423 pairwise gait
660 state transitions. (A) A plot of the mean net change in cirral activity versus the net scaled cell
661 velocity associated with all transitions between the 32 gait states identified in Figure 2 shows
662 that the change in number of active cirri is not strongly correlated with cell velocity (R^2=0.03).
663 Cell velocities were obtained from manually tracked walking trajectories and then scaled by
664 dividing frame to frame displacements for each trajectory by the length of the cell being tracked
665 and also dividing by the average frequency of cirral inactivity. Scaling provided a non-
666 dimensional velocity scaled by natural units of the system. Note that at low Reynolds number,
667 velocity should be proportional to force [43], so this plot also reflects the net walking force
668 generated by the cell. Net change in cirral activity was computed using the data presented in
669 Figure 2. Note that the largest velocities are associated with small negative and small to
670 moderate positive net changes in cirral activity. (B) The transition matrix of all gait state
671 transitions, with rows representing the starting state and columns indicating the ending state,
672 exhibits broken detailed balance. Rates were estimated by dividing the total number of observed
673 transitions between each state pair and dividing by the total time observed. Under detailed
674 balance or equilibrium conditions, transitions from one state to another should be balanced by
675 reverse transitions. Lack of this kind of reversibility, as seen by the lack of symmetry of the
676 heatmap across the diagonal, indicates broken detailed balance and non-equilibrium dynamics.
677 (C) A directed graph representation of all gait state transitions. Nodes correspond to the 32 gait
678 states, and node sizes are scaled by the proportion of total time cells spent in each state. Directed
679 edges are represented by arrows between nodes and signify state transitions. The size of the
680 arrows is scaled by transition rates as in panel b. Edge color represents scaled cell velocity as in
681 panel A according to the indicated color scale. (D) A subset of transitions visualized as in panel
682 C shows the restricted and relatively high frequency nature of unbalanced, non-equilibrium-like

683 transitions. Only transitions that were observed to happen more than one time and exhibiting a
684 significant difference between forward and reverse transitions ($p<0.05$ by binomial test, see
685 Method Details) are displayed. (E) A subset of transitions, similarly to panel d, except that only
686 the balanced transitions, lacking a significant difference between forward and reverse transitions
687 ($p<0.05$ by binomial test) are displayed, also show a complex and widespread structure, this time
688 of balanced, equilibrium transitions. Note that the majority of transitions associated with high
689 cell velocity involve equilibrium-like dynamics. (F) Examples illustrating the spatial
690 organization of cirral activity corresponding to gait states. Some states, such as 7, correspond to
691 activity in spatially discrete groups of cirri, while others, such as 17, correspond to cirral activity
692 across the cell. The gait states displayed here are those involved in unbalanced transitions. (G) A
693 heatmap of transition probabilities between states, showing only the most probable transitions
694 from a given state with all others set to zero, shows distinct structure. In cases were multiple
695 state transitions from a state were tied for the highest probability, all of these transitions are
696 displayed. Fewer than half of the total states are recipients of multiple high probability
697 transitions, and many states are the recipients of no high probability transitions. (H) A
698 representation of functional states and transitions between them highlights the machine-like
699 nature of the gait of *Euplotes*. Gait states are represented as circles with numerical labels. Blue
700 circles represent states that are both recipients and sources of unbalanced transitions as identified
701 in panel D and constitute the three cycle states. Red circles represent states that are recipients but
702 not sources of unbalanced transitions as identified in panel D. Black circles correspond to gait
703 states that are associated only with balanced transitions as in panel E. States receiving no more
704 than one unique high probability transition from states with only a single highest one as
705 identified in panel g were grouped together into a compound state represented by the dark gray
706 blob. The blue background behind states 4, 16, 18, 26, and 28 indicates that these states all share

707 the same highest probability transitions between states identified in this panel, and thus, the
708 group constitutes a single compound functional state. Arrows represent the highest probability
709 transitions between the states, including compound states composed of multiple gait states as
710 identified in Figure 2 (dark gray blob and blue background). Gait state 1 is also depicted, as it is
711 the state in which cells spent the most time over all walking trajectories and also is uniquely the
712 state from which cells begin walking. Cells also frequently return to the state during walking.
713 Further, transitions from gait state 2 from gait state 1 constitute the single highest frequency
714 transition. Together, all identified states in this panel constitute functional states. Arrows
715 represent the most probable transitions between functional states, and all unbalanced transitions
716 are also represented with size scaled by their proportional probability compared to all other
717 transitions emanating from the source functional state. Cartoons are a walking cell in profile with
718 cirri in a configuration representative of the corresponding functional state. Labels refer to the
719 apparent functional role of states and their associated transitions. Beginning from gait state 1, the
720 resting/start state, cells are most likely to follow transitions from gait state 2 to 3 to 17 at which
721 point cells are likely to enter the functional state associated with substantial cell movement
722 involving variable balanced transitions between a number of gait states. Transitions are then
723 likely to lead back toward the cycle states. Note that while this representation of gait dynamics
724 highlights the most probable transitions, substantial variability, primarily involving reversible
725 transitions, occurs during walking trajectories. (I) A cartoon depicts the basic features of our
726 model of a walking cell. Cirri generate a constant force when moving and resist displacement in
727 a spring-like fashion when in contact with the substrate. Force and torque balance yield cell
728 displacement and orientation in each simulation timestep. (J) A plot of the relative average
729 resting cirral-surface contact positions measured from the 13 cells used for gait analysis, which
730 were used for all simulations and define simulated cell position and orientation. (K) A

representative trajectory of a simulation of a walking cell using experimentally recorded cirral activity illustrates the linear, processive motion, which matches experimentally measured cell walking dynamics qualitatively and semi-quantitatively. (L) A representative trajectory of a simulation of a walking cell using the same experimentally recorded cirral activity as in panel K but in a shuffled order illustrates a decrease in processivity of cell movement. (M) A representative trajectory of a simulation of a walking cell with random patterns of cirral activity with the same average level of cirral activity and the same number of timesteps as in panels k and l illustrates the decrease in cell speed. (N) A plot of path straightness measured from simulations using actual cirral pattern measured experimentally (Actual), actual cirral patterns in a randomly shuffled order (Shuffle), and randomly generated cirral activity with the same average cirral activity as actual patterns (Random) suggest that the order of gait state transitions matters for processivity of cell movement due to a predicted decrease in path straightness for Shuffle trajectories. Path straightness is the distance between the start and end of the cell trajectory divided by the sum total path length. The trajectories of the 13 cells analyzed for Figure 2 were used for the Actual simulations. For the Shuffle simulations, cirral activity patterns from each of the 13 cells were shuffled to generate a random order five times. For the Random simulations, five sets of patterns of lengths equal to each of the 13 cell trajectories were generated. The asterisk indicates $p=0.04$ by the Wilcoxon rank sum test. (O) A plot of scaled cell velocity for the simulations described for panel k predict that the gait states are responsible for generating cell velocities as indicated by decreased scaled cell velocity for random patterns of cirral activity. Scaled velocity was computed as described for panel a and averaged over the entire cell trajectory. Three asterisks indicates $p<0.001$ by the Wilcoxon rank sum test. See also Figure S1 and Figure S2.

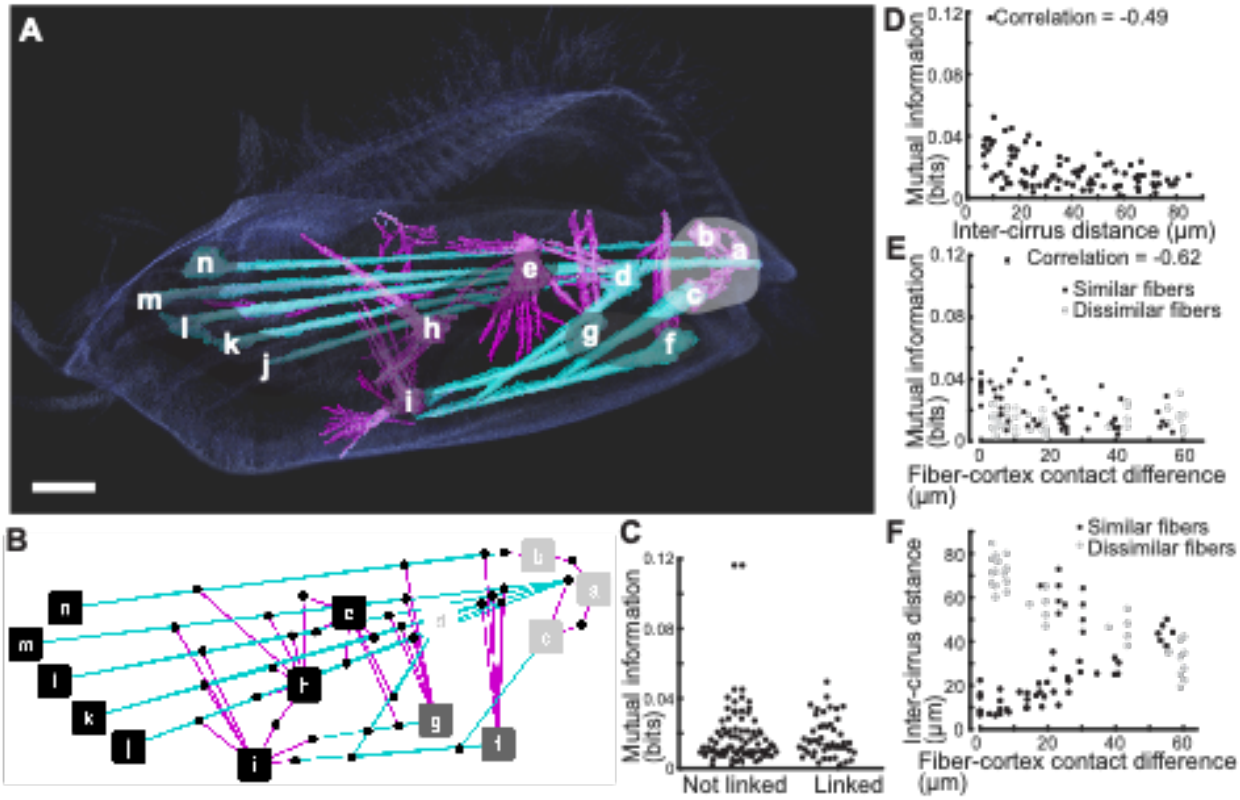


Figure 4. The structure of the complex, interconnected fiber system of *Euplotes* correlates with dynamical associations between cirri. (A) The SiR-tubulin labeled cell (faint, dark blue) was imaged by confocal microscopy, and a 3D reconstruction as obtained from serial confocal slices. Fibers were manually 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 Figure S3 for more 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 Figure 2C). Gray shading indicates the dynamical groups identified by dimensionality reduction and follows the same color scheme as in Figure 2H. 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 panel 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 Figure S3). (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 panel A. Fiber-cortex contact difference is measured by the mean cross nearest neighbor distance (see Method Details) for all fiber-cortex contact points associated with each cirrus. The negative correlation values from the data plotted in panels 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 panels D and E)

792 illustrates that nearby cirri tend to have similar associated fiber-cortex contacts, highlighting that
793 nearby cirri with similar fiber-cortex contacts share the most mutual information. See also Figure
794 S3.

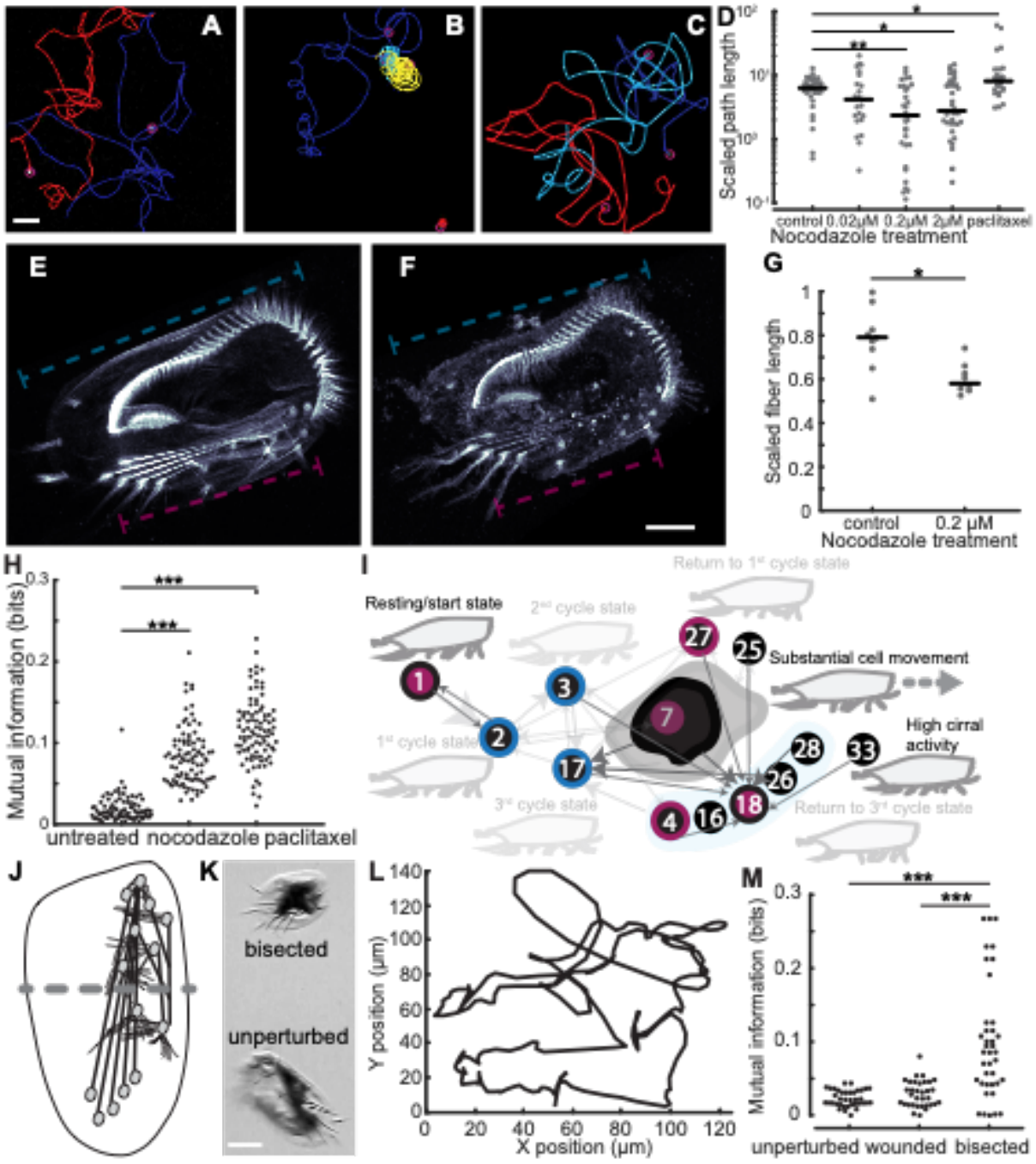


Figure 5. The fiber system of *Euplotes* mediates gait coordination. (A-C) Representative cell motility trajectories of cells imaged under darkfield illumination and tracked using the TrackMate plugin in FIJI [108] in control (A), nocodazole treated (B), and paclitaxel treated (C) cells highlight the increased turning causing curved and curled trajectories in cells treated with the microtubule polymerization inhibitor nocodazole and decrease in turns leading to smoother,

801 less convoluted trajectories in cells treated with the microtubule stabilizer paclitaxel. Different
802 colors represent the trajectories of different cells. Scale bar is 500 μm. (D) Nocodazole affects
803 cell motility in a dose dependent and specific manner. Motility was quantified by the scaled path
804 length, which is the total integrated distance walked by a cell scaled by the maximum distance
805 the cell travelled from its starting point. This scaled path length decreases with a decrease in
806 long, straight segments of trajectories corresponding to normal cell walking. Scaled path length
807 decreased with increased nocodazole concentration, becoming significantly less than controls by
808 0.2 μM nocodazole. Note that at 2 μM nocodazole, cells were often observed to swim instead of
809 walk, which may account for the lack of decrease in scaled path length compared to 0.2 μM
810 treatment. In contrast to nocodazole, treatment of cells with 0.02 μM paclitaxel, which stabilizes
811 microtubules, increases scaled path length compared to the control. The trajectories of at least 20
812 cells were analyzed for each condition. The black bars are median values. A single asterisk
813 indicates $p<0.05$, and a double asterisk indicates $p<0.005$, as computed by a Wilcoxon rank-sum
814 test. (E,F) Representative images illustrating the effect of the inhibition of microtubule
815 polymerization by nocodazole on the fiber system and approach to quantifying this effect.
816 Images are maximum intensity projections of confocal z-stack images of cells labeled by SiR-
817 tubulin. Scale bar is 20 μm. (G) A plot showing that nocodazole treatment shortens fibers
818 compared to controls. Scaled fiber length was measured by dividing the length of the longest
819 fiber by the length of the entire cell. Nine cells were analyzed for each condition. The black bars
820 are median values. A single asterisk indicates $p<0.05$ as computed by a Wilcoxon rank-sum test.
821 (H) A plot of mutual information of all pairs of cirri shows that treatment of cells with the
822 microtubule inhibitors nocodazole and paclitaxel change the distribution of mutual information
823 due to an overall increase in cirral synchrony compared to untreated cells. The data for untreated
824 cells is that from Figures 2 and 3. Nocodazole treatment was 0.2 μM, and 6 cells over 1133

825 timepoints were analyzed. Paclitaxel treatment was 0.02 μM, and 6 cells over 441 timepoints
826 were analyzed. Three asterisks indicates $p<0.001$ as computed by a two sample Kolmogorov-
827 Smirnov test. (I) A representation of the changes in the nature and organization of functional
828 states as well as transitions between them highlights the effects of nocodazole treatment. This
829 panel is partially adapted from Figure 3H and was produced by following the same analysis
830 procedure. Any functional states and transitions depicted in Figure 3H no longer observed under
831 nocodazole treatment appear faded. Arrows that are not faded represent the new highest
832 probability transitions associated with the states. Outer colors of the nodes corresponding to the
833 numbered gait states denote the type or absence of broken detailed balance associated with the
834 gait state for untreated cells while the core color represents that for the treated cells. As in Figure
835 3H, black denotes the lack of unbalanced transitions, red denotes the state is only a source or
836 only a recipient of unbalanced transitions, while blue denotes states which are both sources and
837 recipients of unbalanced transitions. Note the loss of the cycle states as well as the functional
838 states directing return to the cycle states, the emergence of the high cirral activity state involving
839 activity in nearly all cirri, the reduction in gait states associated with the cloud of states of
840 substantial cell movement (indicated by the reduction in the size of the black blob), and the
841 emergence of many high probability transitions to gait state 18. (J) A cartoon illustrates the
842 location of the cuts made for microsurgery experiments (dashed gray line). (K) A representative
843 image shows a recovered bisected cell and a fully intact, unperturbed cell for reference. Scale bar
844 is 50 μm. (L) A plot of the manually tracked trajectory of a walking, bisected cell illustrates the
845 unsteady, curving manner of cell movement. (M) A plot of mutual information of all pairs of
846 cirri h-n for unperturbed, wounded, and bisected cells demonstrating a change in the distribution
847 of mutual information only in the case of bisected cells where fibers associated with h-n have
848 been severed. Three asterisks indicates $p<0.001$ as computed by a two sample Kolmogorov-

849 Smirnov test. The data for unperturbed cells is that from Figure 2, and data for wounded and
850 bisected conditions involved 10 cells over 1308 timepoints in the case of the wounded cells and
851 1815 timepoints for bisected cells. See also Figure S4, Figure S5, and Videos S3-S7.
852
853

854 # STAR Methods

855 ## Resource Availability

856 ***Lead contact***

857 Further information and requests for resources, data, and code should be directed to and will be
858 fulfilled by the lead contact, Wallace F. Marshall (wallace.ucsf@gmail.com).

859 ***Materials availability***

860 This study did not generate new unique reagents.

861 ***Data and code availability***

862 All data reported in this paper will be shared by the lead contact upon request.
863 All original code has been deposited at GitHub and is publicly available as of the date of
864 publication. DOIs are listed in the key resources table.
865 Any additional information required to reanalyze the data reported in this paper is available from
866 the lead contact upon request.

867

868 ## Experimental Model and Subject Details

869 Cell lines

870 Cultures of *Euplotes eurystomus* were obtained from Carolina Biological Supply Company (Item
871 #131480) and were kept at room temperature under ambient light conditions.

872

873 ## Method Details

874 Cell husbandry

875 Individual cells were isolated from cultures, which contained other protists and
876 meiofauna, by pipetting and placed in non-treated 6-well plates (Thermo Fischer Scientific 08-
877 772-49) containing spring water taken from cultures. Cells were kept in wells for no longer than

878 five days before imaging, and if cells were to be kept for longer than 48 hours, wells containing
879 cells were supplemented with 1% Cereal Grass Medium [109] (Thermo Fischer Scientific S25242)
880 to prevent depletion of prey bacteria and otherwise maintain *Euplotes* under constant growth
881 conditions.
882
883 Live cell brightfield microscopy
884 Cells were concentrated by centrifugation (500×g for 5 min) and resuspended either in
885 0.5 mL of spring water in coverglass bottomed FluoroDishes (World Precision Instruments
886 FD35-100) or in 0.2 mL spring water on a coverslip (FisherScientific, 12-545-D) for imaging.
887 No more than three cells were kept in 0.5 mL imaging samples and only one cell was ever kept
888 in 0.2 mL imaging samples in order to minimize cell-cell interactions. Cells were observed to
889 exhibit spontaneous walking activity on coverglass. Walking cells in FluoroDishes were imaged
890 under brightfield illumination using a Zeiss Z.1 Observer and Hamamatsu Orca Flash 4.0 V2
891 CMOS camera (C11440-22CU) with a 20x, 0.8 NA Plan-Apochromat (Zeiss) objective. Cells on
892 coverslips were imaged under brightfield illumination with coverslips inverted over a well
893 containing a small amount of distilled water to reduce evaporation using a Zeiss Axio Zoom.V16
894 and a PCO pco.dimax S1 camera. Importantly, in both imaging systems, the focal plane was set
895 at the interface between cirri of walking cells and the glass surface upon which they were
896 walking. Images were acquired at 0.033 seconds per frame with a 0.005 second exposure in order
897 to capture all cirral dynamics during walking with minimal blur.
898
899 Quantification of walking dynamics
900 Movies of walking cells were viewed using FIJI [110]. Movement of cirri, or lack thereof
901 was clearly visible in each movie frame (see Figure 2A and Video S1). The dynamical state of

902 each cirrus in each movie frame was manually annotated. For each frame, each cirrus received a
903 label of "1" if the cirrus was in motion and "0" if the cirrus was not moving and in contact with
904 the coverslip. Motion of cirri was evident in terms of a change in cirrus shape or tip position
905 often in addition to blur due to motion during image acquisition or position out of the focal plane
906 (see Figure 2A and Video S1). While only slowly walking cells were recorded, sometimes cells
907 nevertheless exhibit brief, spontaneous departures from slow walking during the course of movie
908 acquisition. Any frame in which the movement of the cell and/or cirri were too fast to be
909 resolved, such as during spontaneous escape responses [40] (Video S2), was excluded from
910 analysis such that some videos were split into a number of separate continuous sequences. Thus,
911 each movie frame associated with a particular time point in the walking trajectory, with the
912 exception of those excluded from analysis as described, yielded a corresponding 14-element
913 binary vector encoding the motility state of the cell in terms of the movement of cirri. Cell
914 movement was tracked using the manual tracking feature of the TrackMate plugin in FIJI [108].
915 The center of each cell was used as the reference feature for tracking. We analyzed the walking
916 dynamics of 13 different cells.
917
918 Dimensionality reduction
919 Dimensionality reduction was performed by non-negative matrix factorization (NMF)
920 implemented in MATLAB release 2019b (Mathworks, Natick). NMF was chosen as a
921 dimensionality reduction technique to allow us to obtain a reduced, sparse, and interpretable
922 representation of walking dynamics. Because NMF derives non-negative factors, the basis
923 vectors in NMF space correspond straightforwardly to patterns of cirral activity. NMF involves
924 factoring data, A, an n by m matrix, into non-negative factors W, an n by k matrix, and H, a k by
925 m matrix where the product $W*H$ approximates A. To determine the appropriate number of

926 dimensions or rank, k, that are necessary to accurately represent the data without overfitting, we
927 performed cross-validation by imputation with random holdouts [111,112], also implemented in
928 MATLAB. We randomly held out 15% of our walking dynamics data, performed NMF for a
929 given k, and then used the NMF reconstruction $W*H$, to update the missing data entries. This
930 process of updating is known as imputation, and we repeated the imputation process 50 times, by
931 which point the imputed values had stabilized, to obtain a final NMF reconstruction. We then
932 computed the root mean squared residual (RMSR) between the final NMF reconstruction, $W*H$,
933 and our dataset, A. We performed this entire process 100 times for each value of k. As is
934 generally the case for NMF, we observed a monotonic decrease in reconstruction error with
935 increasing k without performing the imputation procedure [113] (Figure S1). In contrast to this
936 trend, we observed an increase in RMSR of imputed values with increasing k indicating
937 overfitting [111] (Figure S1). We chose k=3 because this value was the highest value before a
938 notable increase in imputation error (Figure S1), which would indicate overfitting [111,112]. Thus,
939 our choice of rank 3 selects the lowest rank approximation that captures structure of the dataset
940 without overfitting that structure. Further, our choice facilitated the visual inspection of the
941 structure of data in the reduced dimensional reconstruction.

942 Finally, we noted that for our chosen value of k, due to the stochastic nature of the NMF
943 algorithm, which involves a random initialization step, we obtained slightly different solutions
944 for different iterations [111]. In order to choose the best reduced dimensional approximation,
945 therefore, we performed NMF 500 times and chose the particular solution corresponding to the
946 lowest RMSR compared to our dataset.

947

948 <u>Clustering</u>

949 Clustering on the dataset obtained using NMF was performed by density-based spatial
950 clustering of applications with noise (DBSCAN) algorithm [70] implemented in MATLAB release
951 2019b (Mathworks, Natick). Structure in NMF space was clearly visible (Figure 2E-G), and
952 DBSCAN using a Euclidean distance metric, was initially chosen as a clustering method because
953 it yielded qualitatively good partitioning of the data. The DBSCAN algorithm involves stochastic
954 search within neighborhoods of a given radius ε around datapoints, and points with a minimum
955 number of neighbors, *n*, within their neighborhood are grouped as belonging to the same cluster,
956 leaving two free parameters to determine. We set ε by first using the
957 clusterDBSCAN.estimateEpsilon function in MATLAB (release 2020b, Phased Array System
958 Toolbox), which yielded a value of 0.15. We next set about determining the minimum neighbor
959 number, *n*. To do so, we computed the average Silhouette coefficient, a commonly used measure
960 of clustering quality that indicates how well-separated clusters are [114], for various values of *n*.
961 The results of this analysis are plotted in Figure S1. Higher Silhouette coefficients indicate better
962 clustering, and we found that a value of *n*=8 maximized the mean Silhouette coefficient (Figure
963 S1). We also noted, however, that for this value, many datapoints were found to be outliers, not
964 belonging to any cluster due to having too few points within a distance of ε. Figure S1 displays
965 percentage of datapoints found to be outliers as a function of *n*. In order to avoid categorizing
966 more than 5% of datapoints as outliers, we chose to settle on *n*=4, which does not have a
967 significantly different mean Silhouette coefficient compared to any of the others in the range
968 *n*=2-7. This choice was further supported by the fact that major clusters involving more than 5
969 datapoints identified with *n*=8 were also identified with *n*=4.
970 Although this set of parameters gave qualitatively and quantitatively reasonable
971 clustering results, we sought to further refine our clusters and to further reduce the outlier
972 datapoints. We noted the obvious partitioning of the NMF dataset into three groups along the H2

axis (Figure 2E). We found the previously determined parameter values to yield good clustering for the top and middle partitions (H2≤1.1 and 0.2<H2<1.1), with no outliers. For the lower partition (H2≤0.2), however, we found that we were able to improve clustering by using ε=0.1182. With this updated value, we found no statistically significant change in Silhouette coefficient and reduced outliers to 0%. The clusters obtained by this process constituted the identification of the 32 gait states. We note here that the problem of determining the true or optimal number of clusters is an unresolved problem [115], and we note that we have followed standard methods to determine cluster number, and we found that our key results do not depend sensitively on the precise number of clusters identified (see following section and Figure S1 for more details).

State transition analysis

Following dimensionality reduction and clustering to identify gait states, we proceeded to characterize state transition dynamics. For each cell trajectory, we identified all unique gait state transitions for a total of 1423 unique pairwise transitions over the cumulative 2343 video frames for 77.14 s of recording. We computed empirical transition rates between states as the total number of observed transitions divided by the total time of observation. In order to determine which transitions were balanced and which were unbalanced, we followed Chang and Marshall [45], and performed binomial tests of statistical significance. Assuming a system at equilibrium, with all transitions obeying detailed balance, we expect to observe some deviation from exactly reciprocal transitions and can calculate the probability of observing a given set of ratios given underlying probabilities of forward and reverse transitions. The binomial probability of observing a set of transitions with known forward and reverse probabilities is given by

$$P(X = f) = \binom{n}{f} p_{forward}^{f} p_{reverse}^{n-f}$$

where $\binom{n}{f} = \frac{n!}{f!(n-f)!}$ is the number of combinations, f is the number of forward transitions, n is the total number of transitions (such that n-f is the number of reverse transitions), and the probabilities $p_{forward}$ and $p_{reverse}$ are the forward and reverse probabilities. Considering only the set of transitions involving a specific pair of states, and calculating the probability that a transition between those states is either in the forward or reverse direction, the values of forward and reverse probabilities in the balanced case must be equal such that $p_{forward} = p_{reverse} = 0.5$. With an α level of 0.05, we then considered reciprocal transition pairs with binomial probabilities less than 0.05 to be significantly unbalanced. Figure S1 displays the binomial probabilities associated with all transitions.

In order to calculate the estimated entropy production rate, we followed Wan and Goldstein [5], where the entropy production rate is defined as

$$\dot{S} = \frac{1}{2}\sum_{i \neq j} J_{ij} A_{ij}$$

with conjugate fluxes $J_{ij} = p_i k_{ij} - p_j k_{ji}$ and forces $A_{ij} = \ln\left(\frac{p_i k_{ij}}{p_j k_{ji}}\right)$ where the p_l are the probabilities of being in state l at steady state and the k_{ij} are the transition probabilities from states i to j. We estimate the state occupancy probabilities p_l as $\frac{T_l}{T_{Total}}$, where T_l is the amount of time spent in state l over all trajectories and T_{Total} is the total recorded time, and the transition probabilities k_{ij} as $\frac{N_{ij}}{N_i}$, where N_{ij} is the total number of observed transitions from state i to state j and N_i is the total number of transitions emanating from state i. To avoid $k_{ji} = 0$ for pairs of states for which we did not observe any transitions during our experiments, we let $k_{ji} = \frac{1}{p_j T_{max}}$ where $T_{max} = 238$ is the maximum number of observed transitions for any single recorded walking trajectory.

In the course of our state transition analysis, we also checked whether the waiting times between instances of each state might be non-exponentially distributed, with exponential distributions indicative of an embedded Markov process or possibly self-organized criticality [116,117]. Using the Lilliefors test implemented in MATLAB, we found that in general, waiting times were not exponentially distributed, although states 2, 3, 6, 16, 17, 18, 25, 27, 28, 32 were found to have waiting times consistent with exponential distributions with Benjamini-Hochburg corrected p-values of 0.046, 0.046, 0.022, 0.008, 0.046, 0.017, 0.046, 0.0081, 0.0046, 0.0046 respectively. Interestingly, none of the waiting times between the movements of individual cirri were found to be consistent with exponential distributions. These results are consistent with mechanisms constraining the temporal dynamics of cirri and state transitions.

In order to begin evaluating whether state transitions obeyed the Markov property for a discrete-time, first-order Markov process, where the transition probabilities from one state to the next are completely determined by current state [75,76], we estimated the transition matrix for walking dynamics, consisting of the transition probabilities between all states. We estimate the transition probability from state i to state j as $k_{ij} = \frac{N_{ij}}{\sum_k N_{ik}}$ such that $\sum_i k_{ij} = 1$. The entries of the transition matrix, P, are these transition probabilities with indices i for rows and j for columns. If gait state transitions obeyed the Markov property, we expect that the product of the transition matrix with itself, P^2, would be equivalent to the two-step transition matrix where transition probabilities are computed as before except that state j is the state to which i has transitioned after an intervening transition. Figure S1 displays the results of this analysis showing that the two matrices show some quantitative and qualitative differences. Although these results strongly suggest violation of the Markov property, we applied the Billingsley test for a more statistically rigorous evaluation [118,119]. This test was implemented and performed in MATLAB. The Billingsley test gives a χ^2 metric with M^2-2M degrees of freedom given by

$$\sum_{i=1}^{M}\sum_{j=1}^{M}\frac{\left(N_{ij}-R_{ij}\sum_{j=1}^{M}N_{ij}\right)^2}{R_{ij}\sum_{j=1}^{M}N_{ij}},$$

where R_{ij}, the independent trials probability matrix, is given by $R_{ij}=\sum_{k=1}^{M}N_{kj}/(\sum_{h=1}^{M}\sum_{l=1}^{M}N_{hl}-\sum_{k=1}^{M}N_{ik})$. Applying this test to our gait state transition data, we found that the null hypothesis that the gait conforms to a first-order discrete time Markov process was rejected (p=0.005).

Importantly, we also noted that the key qualitative results of our state transition analysis are robust to the details of clustering results. In particular, we find that strongly unbalanced transitions and violation of the Markov property exist for a range of clustering parameters. Figure S1 displays the transition matrices for different clustering results.

To arrive at the simplified, state machine representation of the gait, we focused on the highest probability transitions emanating from each state. Transition probabilities were estimated as k_{ij} (as defined above). 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.

<u>Biophysical model and simulations</u>

For our simple biophysical model, we consider a 2D system in which a *Euplotes* cell is walking across a surface in a low Reynolds number environment [43]. The cell has 14 cirri, which

exist in one of two states: actively moving or not actively moving, following our quantitative gait characterization. Cell position and orientation is defined in terms of equilibrium position of the cirri. Cirri can generate a motive force to drive cell motility when moving and resist displacement when not moving and in contact with the surface. For our model, we remain agnostic to the details by which cirri produce generate force noting only that in our experiments, no cell displacement was observed when cirri were not moving and in contact with the coverslip. We therefore let cirri generate a constant force in the direction dictated by cell orientation when moving. We conceptualize the resistance to displacement of unmoving cirri as stemming from the adhesive interaction between the cirrus and the substrate on which the cell is walking and the energy required to bend or deflect a cirrus. Consistent with experimental observations, we do not allow for translation of a cirrus-substrate contact point while a cirrus is not actively moving.

For a particle moving through a fluid at low Reynolds number, such as our cell, velocity $\vec{v}$ will be given by

$$\vec{v} = \frac{\sum \vec{F}}{\gamma},$$

where $\sum \vec{F}$ is the sum of the external forces acting on the particle, and γ is constant related to the geometry of the particle and the viscosity of the fluid [120,121] accounting for drag. In our model, cirri are responsible for the forces involved in motility, so $\sum \vec{F} = \overrightarrow{F_a} + \overrightarrow{F_\iota}$ with the motive force $\overrightarrow{F_a} = nF_a\vec{o}$ where n is the number of active cirri, F_a is the magnitude of the constant force generated by active cirri, and $\vec{o}$ is the unit vector in the direction of cell orientation, and the resistive force $\overrightarrow{F_\iota} = \sum_i k\overrightarrow{d_\iota}$ is a sum over the inactive cirri where k is a constant controlling the resistance of a cirrus to deformation and $\overrightarrow{d_\iota}$ is the displacement vector of inactive cirrus i. We note that the forces driving cell motility in *Euplotes* stem from complex mechanical interactions, but for our model, we have chosen simple, first order expression to capture very basic features.

Similar to the expression for velocity above, angular velocity of a walking cell in our model is given by

$$\vec{\omega} = \frac{\sum \vec{F} \times \vec{r}}{\beta},$$

where β is a constant related to the geometry of the cell and viscosity of the fluid, $\vec{F_i}$ is the force due to cirrus i ($F_a\vec{o}$ for active cirri and $k\vec{d}$ for inactive cirri), and $\vec{r_i}$ is the vector pointing from the center of the cell to cirrus i.

In addition to the relative positions of the cirri and patterns of cirral activity, the four parameters γ, β, F_a and k govern cell motility in our model. From these four parameters, we obtain three related dimensionless parameters:$\alpha = \beta/\gamma l$ where l is the maximum distance between cirri, which can be thought of as characterizing the unsteadiness of the cell the intrinsic susceptibility of the cell to turning due to cell geometry; $\mathrm{M} = F_a/kl$, which can be thought of as characterizing degree to which cirral activity will tend to induce cell movement in opposition to inactive cirri; and $\mathcal{F} = F_a t/\gamma l$ where t is the duration of a timestep in simulations, which can be thought of as the strength of the cirral motive force relative to the viscous drag experienced by the cell due to its fluid environment.

For all simulations, relative equilibrium cirral positions calculated from the average cirral positions in a video frame with no cirral activity over the 13 cells used for gait analysis (Figure 2) were used. In each simulation timestep, defined by the timestep used for recording videos for gait analysis, velocity and angular velocity are calculated based on the positions and activity of cirri, and the positions of all cirri are updated accordingly before proceeding. To calibrate the parameters of the model, we used the cirral patterns recorded from the cells used for gait analysis. We swept parameter space and found that simulations qualitatively and semi-quantitatively recapitulated experimentally measured cell motility with $\alpha = 0.001$, $\mathrm{M} = 0.26$,

1111 and $\mathcal{F} = 0.008$ (Figure 3K, Figure S1). These parameter values were subsequently used for all
1112 simulations.

1113 For simulations with shuffled gait state transitions, we used MATLAB's shuffle
1114 command on the cirral dynamics of actual cells to obtain sequences of gait state transitions of the
1115 same length as those that were experimentally obtained except in a random order. To obtain
1116 random patterns of cirral activity similar to those measured experimentally, we generated cirral
1117 activity according to a process defined by two probabilities: p_a, the probability of transitioning
1118 from inactive to active at each timestep and p_i, the probability of transitioning from inactive to
1119 active at each timestep. We initialized sequences with no cirral activity and then updated cirral
1120 activity according to these probabilities for each timestep in the sequence. We found that setting
1121 p_a=0.3 and p_i=0.1 yielded the same average cirral activity as that recorded experimentally, 0.23 ±
1122 0.42 per frame and 0.24 ± 0.43 per frame respectively where the values are mean ± standard
1123 deviation. All simulations were performed in MATLAB.

1124

1125 Confocal microscopy

1126 Cells were prepared for imaging and placed into a FluoroDish as described in the Live
1127 Cell Brightfield Microscopy section. Cells were then labeled with SiR-tubulin (Spirochrome
1128 provided by Cytoskeleton, Inc, CY-SC002) at 1 μM concentration. Cells were imaged using a
1129 Zeiss LSM 880 AxioExaminer and a 40x, 1.2 NA C-Apochromat water immersion objective
1130 (Zeiss) and excitation provided by a 633 nm laser (Zeiss). Only one full confocal z-stack of a
1131 complete cell was obtained during imaging to avoid effects of photodamage.

1132

1133 Fiber reconstruction and analysis

1134 The image stack resulting from confocal imaging was first aligned in FIJI using the
1135 StackReg plugin [110]. Next, fibers were manually segmented in each of the aligned z-stack images
1136 using the TrakEM2 plugin in FIJI [122,123]. Thick and thin fibers (Figure 4A) were morphologically
1137 distinguished, with thick fibers having a diameter of no less than 5 μm at the thinnest point.
1138 Fibers were traced from their distal termini to their convergences at the base of the cirri with
1139 which they were associated. Following segmentation, 3D surfaces were reconstructed in
1140 TrakEM2. Inter-fiber contacts were then found by inspection of 3D reconstructions and verified
1141 by examining individual z-stack frames to confirm intersections between fibers.

1142
1143
1144 Drug treatment experiments and analysis

1145 For all cytoskeleton inhibitor treatment experiments, 1 mL of cells in culture were placed
1146 in wells of 12-well plates (Thermo-Scientific, 12-565-321). Nocodazole (Sigma-Aldrich,
1147 M1404) as a stock solution of 6.64 mM in DMSO diluted further in distilled water, was added to
1148 achieve appropriate concentrations, with no more than 1 μL of additional volume added, and 1
1149 μL of distilled water with equivalent DMSO concentration to nocodazole treatments added to
1150 controls. Cells were incubated for 1 hr before initiating experiments. No cell death was observed
1151 at any concentration of nocodazole in the 6 hours following nocodazole treatment. Paclitaxel
1152 (Sigma-Aldrich, T7191) as a stock solution of 2.23 mM in DMSO diluted further to 20 μM in
1153 distilled water, was added to achieve a final concentration of 20 nM to cells in solution.
1154 Latrunculin B (Thomas Scientific, C834E37) as a stock solution of 1.1 mM in ethanol was
1155 further diluted in distilled water and added to achieve a final concentration of 10 μM with cells
1156 in solution. Cytochalasin B (Fisher Scientific 1493-96-2) as a stock solution of 2.1 mM in
1157 ethanol was further diluted in distilled water was added to achieve a final concentration of 50
1158 μM with cells in solution. Jasplakinolide (Fisher Scientific 42-012-750UG) as a stock solution of

1159 1 mM in DMSO was further diluted in distilled water to achieve a final concentration of 10 μM
1160 with cells in solution. No cell death was observed in the 6 hours following treatment with any of
1161 the actin inhibitors. For the control condition for actin inhibitor experiments, both DMSO and
1162 ethanol was added to match the concentrations added in the cytochalasin B and jasplakinolide
1163 conditions.

1164 *Washout experiments*

1165 For nocodazole washout experiments, motility assays (described below) were also
1166 performed after placement of cells into well plates and before nocodazole treatment. Following
1167 0.2 μM nocodazole treatment and another motility analysis, cells were picked in 5 μL of media
1168 and placed into wells of 6 well plates (Corning, CLS3736) each containing 2mL of fresh media.
1169 No more than 20 μL of nocodazole treatment media was added to any well so that the resultant
1170 nocodazole concentration in the washed condition was no more than 2 nM. Cells were allowed to
1171 recover in this condition for 4 hr, and then a final motility assay was performed.

1172 *Motility analysis*

1173 For motility analysis, cells were picked from well-plates and placed onto well-slides
1174 created by using a paper hole punch tool to punch a hole in 0.25 mm thick silicone spacer
1175 material (CultureWell, 664475), and adhering the spacer material to a glass slide (Corning, 2947-
1176 75X25). A total volume of 20 μL including up to 4 cells was added to well slides. A glass
1177 coverslip (FisherScientific, 12-545-D) was placed atop the well slide, sealing the well and
1178 creating an imaging chamber. After creating the imaging chamber, cells were allowed to
1179 acclimate to their new environment for 10 min. Cells were then imaged on a Zeiss Axio
1180 Zoom.V16 microscope under darkfield illumination with a Canon EOS T5i DSLR camera
1181 recording at 30 fps for 2 min. Videos were then processed using FIJI [110]. First, images were
1182 cropped to remove extraneous parts of the field of view that did not contain the imaging

1183 chamber, and then background subtraction was performed by creating an image composed of
1184 mean pixel intensity values over all frames of the video and subtracting this mean image from all
1185 frames of the video. A mean filter with a four-pixel radius was then applied to each frame of the
1186 video for the purpose of smoothing. After processing, tracking was performed using the
1187 TrackMate FIJI plugin [108]. For detection of objects (cells), a Laplacian of Gaussian filter was
1188 applied with an estimated blob diameter of 25 pixels and threshold of 0.1. A quality threshold
1189 was set manually when necessary to filter out any detected objects that were not cells. The
1190 Linear Assignment Tracker with a linking max distance of 15 pixels, gap-closing max distance
1191 of 150 pixels, and a gap-closing max frame gap of 200 frames was then used to generate linked
1192 tracks (trajectories) of detected cells. Trajectories were analyzed in MATLAB. Scaled path
1193 length for each tracked cell was calculated by summing the length of all segments of the track
1194 and dividing by the maximum distance the cell traveled from its starting point.

1195 *Fiber length analysis*

1196 Just prior to confocal imaging, cells were washed by picking up to five cells in 10 μL
1197 volume each and placing into 1 mL fresh media in a 12-well plate (Thermo-Scientific, 12-565-
1198 321). Cells were then prepared for imaging as described in the Confocal Imaging section of
1199 Method Details. Confocal z-stacks were then loaded into FIJI and aligned using the StackReg
1200 plugin [110]. Cell lengths were determined by finding the maximal distance between two points on
1201 the front and rear ends of the cell. Because of variability in the detectable fibers in nocodazole
1202 treated cells, only fibers associated with the rear cirri (j-n), which were visible in all cells, were
1203 used for analysis. All of these rear fibers were measured, and the reported scaled fiber length was
1204 obtained by dividing the length of the longest fiber by the corresponding cell length. In all cases,
1205 the fiber associated with cirrus m was the longest fiber.

1206 *Analysis of cirral dynamics in nocodazole and paclitaxel treated cells*

1207 Cells were prepared for imaging and imaged as described in the Live Cell Brightfield
1208 Microscopy section of the Method Details with the exception that a Canon EOS T5i DSLR
1209 mounted on a Zeiss Axio Zoom.V16 microscope was used to record movements. Additionally,
1210 video was recorded at 0.066 seconds per frame to avoid blurring and then videos were
1211 downsampled to 0.033 seconds per frame for analysis. Cirral dynamics were quantified as
1212 described in the Quantification of Walking Dynamics section of the Method Details.
1213 To assign cirral configurations of nocodazole treated cells to previously identified gait
1214 states, we first matched any cirral configurations with known gait state identity. Next, due to
1215 proximity of new cirral configurations not observed in untreated cells (Figure S1), we were able
1216 to map the new cirral configurations onto the clusters defining the gait states by determining the
1217 nearest cluster to the new cirral configuration. Distance between new cirral configurations and
1218 clusters were determined by finding the shortest distances to points defining convex hulls of each
1219 cluster. The shortest of all these distances then indicated the nearest cluster and corresponding
1220 gait state to which the new cirral configuration was assigned. We noted an obvious dense cluster
1221 of points corresponding to activation of nearly all cirri, and we identified a new cluster and
1222 corresponding gait state by applying the DBSCAN algorithm as described in the Clustering
1223 section of the Method Details. Evaluation of transition dynamics was performed as described in
1224 the State Transition Analysis section of the Method Details. This analysis was all performed in
1225 MATLAB.
1226
1227 <u>ϵ-machine construction</u>
1228 Our representation of the Euplotes at timestep t takes the form of a length 14
1229 binary string $x^t = [x_1^t, x_2^t, \ldots, x_{14}^t]$ where $x_i^t \in \{0, 1\}$. For reducing dimensions,

we found that bigrams of the 14-dimensional strings, yielded more consistent, interpretable results than unigrams, so $x'^t = [x_1^t, x_2^t, \ldots, x_{14}^t, x_1^{t+1}, x_2^{t+1}, \ldots, x_{14}^{t+1}] \in \{0,1\}^{28}$.

In order to learn latent states, we used Variational AutoEncoders (VAEs) [124] to reduce each fourteen dimensional timestep to three dimensions. VAEs in particular are used for their ability to learn an interpolatable latent space where high-dimensional training data points are mapped to points in low-dimensional space that mimic a normal distribution centered around the origin. This processes yields a smooth latent space with dimensions that represent core aspects of the data. We used a very minimal VAE with an Adam optimizer [125], consisting of one dense layer for the generating encoder means and one dense layer for generating encoder variance. This creates a 3-dimensional normal-like distribution, which we sample from and decode with a one layer decoder.

$\text{VariationalEncoder}(x'^t) \rightarrow z^t \in \mathbb{R}^3$

After reducing the dimensionality of our data, we perform spectral clustering [126] using scikit-learn [127] with 100 k-means runs and 4 clusters, where the cluster count was decided by inspection of the resulting classes.

$\text{SpectralClustering}(z^t) \rightarrow c^t \in \{1,2,3,4\}$

Applying this procedure to the data from the cells used for gait analysis (the same as used for Figure 2) with bigram training parameters lr=0.0011, weight_decay=0.000727, k=3, hidden_dim=5, beta1=0.5, and beta2=0.900, we found the following four clusters corresponding to particular patterns of cirral ativity: 1) Front cirral activity 2) High overall cirral activity 3) Rear cirral activity and 4) Little cirral activity (Figure S5). When we applied this procedure to the data from cells treated with 0.2 μM nocodazole (the same data used for Figure 5) with bigram training parameters lr=0.0026, weight_decay=0.000106, k=3, hidden_dim=3, beta1=0.5, and beta2=0.999, we obtained only three distinct clusters with similar cirral activity patterns to

1254 those of untreated cells except that the rear cirral activity cluster (3 above) disappeared (Figure
1255 S5). These clusters represent an independent, alternative coarse-graining of patterns of cirral
1256 activity into gait states to that presented in Figure 2. Although independent, this procedure
1257 identified the same underlying structure in the data. In particular, the alternative gait states here
1258 involve similar patterns of cirral activity to those defining the NMF components depicted in
1259 Figure 2H. Whereas the spatially distinct groups of cirri sharing the most mutual information
1260 define the space of gait states under the NMF-based approach, under this alternative, activity or
1261 complete lack thereof in those same groups of cirri defines the gait states. Furthermore, the loss
1262 of cluster 4 under nocodazole treatment is consistent with the loss of gait states as depicted in
1263 Figure 5I.

1264 To construct ϵ-machines [79], characterized by a set of causal-states and transitions between
1265 them, we feed a stream of behavioral symbols $c_0, c_1, \ldots, c_T$ into a Causal-State Splitting
1266 Reconstruction (CSSR) [89,128] algorithm. Due to the nature of the CSSR algorithm, we can only
1267 construct ϵ-machines from unbroken data streams, so we constructed ϵ-machines for each cell
1268 from a given dataset. Performing this process over gait state data streams derived from cirral
1269 activity recordings as described above for each cell from the untreated and nocodazole datasets
1270 yielded a set of ϵ machines. Representative examples of ϵ-machines obtained from each dataset
1271 are depicted in Figure S5.

1272

1273 Microsurgery experiments

1274 All microsurgeries were performed by hand under observation with a Zeiss Axio
1275 Zoom.V16 using quartz microneedles pulled from quarts rods (Sutter Instrument, QR100-10)
1276 using a Sutter Instrument P-2000 Laser-based micropipette puller. Individual cells were picked
1277 and placed in a 5 µL droplet at the center of an imaging chamber as described in the Motility

1278 Analysis section of the Method Details. For bisections, microneedles were quickly and firmly
1279 pressed onto cells just anterior to cirrus h, cleanly severing the cell including the fibers
1280 associated with cirri j-n and sealing the cell membrane. Cells were allowed to recover for at least
1281 5 minutes until some motility resumed, and then 15 µL of fresh culture media was added to the
1282 imaging chamber, which was subsequently sealed with a glass coverslip (FisherScientific, 12-
1283 545-D) to minimize evaporation. After 24 hr, once cell fragments resumed walking motility, gait
1284 analysis was performed as described in the quantification of walking dynamics section of the
1285 Method Details, except that cell fragments only had 7 cirri (j-n) instead of the full 14.
1286 Experiments with wounded cells were conducted in the same fashion as described for bisections
1287 except that instead of cutting cells in half, wounding was performed by stabbing with a
1288 microneedle a portion of the cell that does not have any fibers and was confirmed visually by the
1289 loss of some amount of cytoplasm.
1290

1291 **Quantifications and statistical analysis**

1292 Statistical details of the experiments can be found in the figure legends, the main text, or the
1293 Method Details section. Statistical details include exact value of n, what n represents (generally
1294 the number of cells measured), definitions of center, and dispersion and precision measures.
1295 Statistical tests and fits were performed as described in the Method Details section using
1296 MATLAB release 2019b or 2020b (Mathworks, Natick).
1297

1298 Supplemental Video Captions

1299 **Video S1**. **A walking cell.** A representative example of a single *Euplotes eurystomus* cell
1300 walking across a coverslip using motile cirri illustrates the walking gait. Note that the movement
1301 of cirri, or lack thereof is clearly visible in each frame of the movie. This cell was imaged using a

1302 Zeiss Axio Zoom.V16 microscope under brightfield illumination. The video is slowed down by a
1303 factor of 4 from real time to show that the movements of cirri are clearly visible.
1304
1305 **Video S2**. **Walking motility including escape responses.** A representative example of a single
1306 *Euplotes eurystomus* cell walking across a coverslip using motile cirri including escape
1307 responses, which were excluded from gait analysis. Escape responses occur at t= 18s, 20s, and
1308 25s. This cell was imaged under brightfield illumination using a Zeiss Z.1 Observer with a 20x,
1309 0.8 NA Plan-Apochromat (Zeiss) objective. The video is slowed down by a factor of 4 from real
1310 time so that the movements are clearly perceivable.
1311
1312 **Video S3. Cell tracking for motility analysis.** A representative example of the trajectories of
1313 cells imaged at low magnification under darkfield illumination using a Zeiss Axio Zoom.V16
1314 microscope. Cells appear as bright objects with magenta circles around them. Different colored
1315 lines correspond to the tracks of different cells tracked using the FIJI plugin TrackMate.
1316
1317 **Video S4. Motility of nocodazole treated cells.** A representative example of the trajectories of
1318 cells treated with 200 nM nocodazole imaged at low magnification under darkfield illumination.
1319 Cells appear as bright objects with magenta circles around them. Different colored lines
1320 correspond to the tracks of different cells tracked using the FIJI plugin TrackMate. Note the
1321 confined movements of cells due to a decrease in long, linear runs. Compare to Movie 3 for the
1322 control condition.
1323
1324 **Video S5. Motility of paclitaxel cells.** A representative example of the trajectories of cells
1325 treated with 20 nM paclitaxel imaged at low magnification under darkfield illumination. Cells

appear as bright objects with magenta circles around them. Different colored lines correspond to the tracks of different cells tracked using the FIJI plugin TrackMate. Note the increase in long, linear runs and decrease in abrupt changes in direction and tight turns. Compare to Movie 3 for the control condition.

Video S6. A walking, bisected cell. A representative example of a bisected cell analyzed in microsurgery experiments displaying the characteristic backwards, spiral walking motility defect. This cell fragment was imaged using a Zeiss Axio Zoom.V16 microscope under brightfield illumination. The video is slowed down by a factor of 4 from real time to clearly show the movements of cirri.

Video S7. A walking, wounded cell. A representative example of a wounded cell analyzed in microsurgery experiments. This cell was imaged using a Zeiss Axio Zoom.V16 microscope under brightfield illumination. The video is slowed down by a factor of 4 from real time to clearly show the movements of cirri.

Acknowledgements

We would like to thank Nicole King for the use of the confocal microscope and Adair Oesterle for assistance with microsurgery. 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. This work was funded by the I2CELL Seed Award of the Fourmentin-Guilbert Scientific

Foundation (WFM, JBP). Additional funding was provided by NIH grant R35 GM130327 (WFM), NSF grant MCB- 2012647 (WFM), and a Merck Fellowship of the Jane Coffin Childs Memorial Fund for Medical Research (BTL).

Author Contributions

BTL, JDP, and WFM conceived of the study. BTL, JG, and WFM developed methodology. BTL performed the experiments, and all authors analyzed and interpreted data. BTL and WFM originally drafted the manuscript, and all authors reviewed and edited the work.

Declaration of Interests

The authors declare no competing interests.

References

1. Jennings, H.S. (1906). Behavior of the lower organisms. (Columbia University Press).
2. Reid, C.R., MacDonald, H., Mann, R.P., Marshall, J.A.R., Latty, T., and Garnier, S. (2016). Decision-making without a brain: how an amoeboid organism solves the two-armed bandit. J. R. Soc. Interface *13*, 20160030.
3. Coyle, S.M., Flaum, E.M., Li, H., Krishnamurthy, D., and Prakash, M. (2019). Coupled Active Systems Encode an Emergent Hunting Behavior in the Unicellular Predator Lacrymaria olor. Curr. Biol. *29*, 3838-3850.e3.
4. Battle, C., Broedersz, C.P., Fakhri, N., Geyer, V.F., Howard, J., Schmidt, C.F., and MacKintosh, F.C. (2016). Broken detailed balance at mesoscopic scales in active biological systems. Science *352*, 604–7.

1373 5. Wan, K.Y., and Goldstein, R.E. (2018). Time Irreversibility and Criticality in the Motility
1374 of a Flagellate Microorganism. Phys. Rev. Lett. *121*, 58103.
1375 6. Woznica, A., Cantley, A.M., Beemelmanns, C., Freinkman, E., Clardy, J., and King, N.
1376 (2016). Bacterial lipids activate, synergize, and inhibit a developmental switch in
1377 choanoflagellates. Proc. Natl. Acad. Sci. U. S. A. *113*, 7894–7899.
1378 7. Wan, K.Y. (2018). Coordination of eukaryotic cilia and flagella. Essays Biochem. *62*,
1379 829–838.
1380 8. Nomura, M., and Ishida, K.I. (2016). Fine-structural Observations on Siliceous Scale
1381 Production and Shell Assembly in the Testate Amoeba Paulinella chromatophora. Protist
1382 *167*, 303–318.
1383 9. Bennett, R.R., and Golestanian, R. (2015). A steering mechanism for phototaxis in
1384 Chlamydomonas. J. R. Soc. Interface *12*, 20141164–20141164.
1385 10. Machemer, H. (2001). The swimming cell and its world: Structures and mechanisms of
1386 orientation in protists. Eur. J. Protistol. *37*, 3–14.
1387 11. Almagor, M., Ron, A., and Bar-Tana, J. (1981). Chemotaxis in Tetrahymena thermophila.
1388 Cell Motil. *1*, 261–268.
1389 12. Berg, H.C., and Brown, D.A. (1972). Chemotaxis in Escherichia coli analysed by three-
1390 dimensional tracking. Nature *239*, 500–504.
1391 13. Nakagaki, T., Yamada, H., and Tóth, Á. (2000). Maze-solving by an amoeboid organism.
1392 Nature *407*, 470–470.
1393 14. Tweedy, L., Thomason, P.A., Paschke, P.I., Martin, K., Machesky, L.M., Zagnoni, M.,
1394 and Insall, R.H. (2020). Seeing around corners: Cells solve mazes and respond at a
1395 distance using attractant breakdown. Science (80-.). *369*.
1396 15. Applewhite, P.B. (1979). Learning in Protozoa. In Biochemistry and Physiology of

1397 Protozoa (Elsevier), pp. 341–355.
1398 16. Dexter, J.P., Prabakaran, S., and Gunawardena, J. (2019). A Complex Hierarchy of
1399 Avoidance Behaviors in a Single-Cell Eukaryote. Curr. Biol. *29*, 4323-4329.e2.
1400 17. Jennings, H.S. (1902). Studies on reactions to stimuli in unicellular organisms. IX—On
1401 the behavior of fixed infusoria (Stentor and Vorticella), with special reference to the
1402 modifiability of protozoan reactions. Am. J. Physiol. Content *8*, 23–60.
1403 18. Tang, S.K.Y., and Marshall, W.F. (2018). Cell learning. Curr. Biol. *28*, R1180–R1184.
1404 19. Lyon, P. (2015). The cognitive cell: Bacterial behavior reconsidered. Front. Microbiol. *6*.
1405 20. Marshall, W.F. (2019). Cellular Cognition: Sequential Logic in a Giant Protist. Curr. Biol.
1406 *29*, R1303–R1305.
1407 21. Greenfield, D., McEvoy, A.L., Shroff, H., Crooks, G.E., Wingreen, N.S., Betzig, E., and
1408 Liphardt, J. (2009). Self-Organization of the Escherichia coli Chemotaxis Network
1409 Imaged with Super-Resolution Light Microscopy. PLoS Biol. *7*, e1000137.
1410 22. Wan, K.Y., and Goldstein, R.E. (2016). Coordinated beating of algal flagella is mediated
1411 by basal coupling. Proc. Natl. Acad. Sci. U. S. A. *113*, E2784-93.
1412 23. Höfer, T., Sherratt, J.A., and Maini, P.K. (1995). Dictyostelium discoideum : cellular self-
1413 organization in an excitable biological medium. Proc. R. Soc. London. Ser. B Biol. Sci.
1414 *259*, 249–257.
1415 24. Kantsler, V., Dunkel, J., Polin, M., and Goldstein, R.E. (2013). Ciliary contact interactions
1416 dominate surface scattering of swimming eukaryotes. Proc. Natl. Acad. Sci. U. S. A. *110*,
1417 1187–1192.
1418 25. Bray, D. (2009). Wetware (Yale University Press).
1419 26. Pfeifer, R., Lungarella, M., and Iida, F. (2007). Self-organization, embodiment, and
1420 biologically inspired robotics. Science (80-.). *318*, 1088–1093.

1421 27. Regev, A., and Shapiro, E. (2002). Cellular abstractions: Cells as computation. Nature
1422 *419*, 343–343.
1423 28. Mitchell, M. (2012). Biological Computation. Comput. J. *55*, 852–855.
1424 29. Navlakha, S., and Bar-Joseph, Z. (2011). Algorithms in nature: the convergence of
1425 systems biology and computational thinking. Mol. Syst. Biol. *7*, 546.
1426 30. Savin, T., Kurpios, N.A., Shyer, A.E., Florescu, P., Liang, H., Mahadevan, L., and Tabin,
1427 C.J. (2011). On the growth and form of the gut. Nature *476*, 57–62.
1428 31. Mehta, P., and Schwab, D.J. (2012). Energetic costs of cellular computation. Proc. Natl.
1429 Acad. Sci. U. S. A. *109*, 17978–17982.
1430 32. Berg, H.C., and Purcell, E.M. (1977). Physics of chemoreception. Biophys. J. *20*, 193–
1431 219.
1432 33. Bialek, W., and Setayeshgar, S. (2005). Physical limits to biochemical signaling. Proc.
1433 Natl. Acad. Sci. U. S. A. *102*, 10040–5.
1434 34. Andrews, L.B., Nielsen, A.A.K., and Voigt, C.A. (2018). Cellular checkpoint control
1435 using programmable sequential logic. Science (80-.). *361*.
1436 35. Gilbert, J.J. (1994). Jumping behavior in the oligotrich ciliates Strobilidium velox and
1437 Halteria grandinella, and its significance as a defense against rotifer predators. Microb.
1438 Ecol. *27*, 189–200.
1439 36. Lueken, W., Ricci, N., and Krüppel, T. (1996). Rhythmic spontaneous depolarizations
1440 determine a slow-and-fast rhythm in walking of the marine hypotrich Euplotes vannus.
1441 Eur. J. Protistol. *32*, 47–54.
1442 37. Erra, F., Iervasi, A., Ricci, N., and Banchetti, R. (2001). Movement of the cirri during the
1443 creeping of Euplotes crassus (Ciliata, Hypotrichida). Can. J. Zool. *79*, 1353–1362.
1444 38. Taylor, C. V. (1919). The neuromotor system of Euplotes. Exp. Biol. Med. *16*, 101–103.

1445 39. Syberg-Olsen, M.J., Irwin, N.A.T., Vannini, C., Erra, F., Di Giuseppe, G., Boscaro, V.,
1446 and Keeling, P.J. (2016). Biogeography and character evolution of the ciliate genus
1447 Euplotes (Spirotrichea, Euplotia), with description of Euplotes curdsi sp. nov. PLoS One
1448 *11*.
1449 40. Ricci, N., Giannetti, R., and Miceli, C. (1988). The ethogram of Euplotes crassus (ciliata,
1450 hypotrichida): I. The wild type. Eur. J. Protistol. *23*, 129–140.
1451 41. Li, Y., Shi, L., and Gu, F. (2010). Comparison of ciliature microtubule organelles in three
1452 hypotrichous ciliate species*. Chinese J. Oceanol. Limnol. *28*, 443–450.
1453 42. Fleury, A. (1991). Dynamics of the cytoskeleton during morphogenesis in the ciliate
1454 Euplotes I. Basal bodies related microtubular system. Eur. J. Protistol. *27*, 99–114.
1455 43. Purcell, E.M. (1977). Life at low Reynolds number. Am. J. Phys. *45*, 3–11.
1456 44. Kimmel, J.C., Chang, A.Y., Brack, A.S., and Marshall, W.F. (2018). Inferring cell state by
1457 quantitative motility analysis reveals a dynamic state system and broken detailed balance.
1458 PLoS Comput. Biol. *14*, e1005927.
1459 45. Chang, A.Y., and Marshall, W.F. (2019). Dynamics of living cells in a cytomorphological
1460 state space. Proc. Natl. Acad. Sci. U. S. A. *116*, 21556–21562.
1461 46. Boltzmann, L. (1872). Weirere Studien uber das warmegleich-gewich unter gasmolekulen.
1462 Sitzungsberichte der Akad. der Wissenschaften, Vienna, part II *66*, 275–370.
1463 47. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002).
1464 Molecular biology of the cell 4th ed. (Garland Science).
1465 48. Shalizi, C.R., and Crutchfield, J.P. (2001). Computational Mechanics: Pattern and
1466 Prediction, Structure and Simplicity. J. Stat. Phys. *104*, 817–879.
1467 49. Crutchfield, J.P. (1994). The calculi of emergence: computation, dynamics and induction.
1468 Phys. D Nonlinear Phenom. *75*, 11–54.

1469 50. Hopcroft, J.E., Motwani, R., and Ullman, J.D. (2008). Introduction to Automata Theory,
1470 Languages, and Computation 3rd ed. (Pearson Education).
1471 51. Ricci, N., Giannetti, R., and Miceli, C. (1988). The ethogram of Euplotes crassus (ciliata,
1472 hypotrichida): I. The wild type. Eur. J. Protistol. *23*, 129–140.
1473 52. Moore, T.Y., Cooper, K.L., Biewener, A.A., and Vasudevan, R. Unpredictability of
1474 escape trajectory explains predator evasion ability and microhabitat preference of desert
1475 rodents.
1476 53. Goldstein, R.E. (2015). Green Algae as Model Organisms for Biological Fluid Dynamics.
1477 Annu. Rev. Fluid Mech. *47*, 343–375.
1478 54. Simon, M., and Plattner, H. (2014). Unicellular Eukaryotes as Models in Cell and
1479 Molecular Biology: Critical Appraisal of Their Past and Future Value. In International
1480 Review of Cell and Molecular Biology (Elsevier Inc.), pp. 141–198.
1481 55. Durham, W.M., Climent, E., Barry, M., De Lillo, F., Boffetta, G., Cencini, M., and
1482 Stocker, R. (2013). Turbulence drives microscale patches of motile phytoplankton. Nat.
1483 Commun. *4*.
1484 56. Mogilner, A., and Oster, G. (1996). Cell motility driven by actin polymerization. Biophys.
1485 J. *71*, 3030–3045.
1486 57. Collins, J.J., and Richmond, S.A. (1994). Biological cybemet Hard-wired central pattern
1487 generators for quadrupedal locomotion.
1488 58. Collins, J.J., and Stewart, I.N. (1993). Coupled nonlinear oscillators and the symmetries of
1489 animal gaits. J. Nonlinear Sci. *3*, 349–392.
1490 59. Barnhart, E.L., Allen, G.M., Jülicher, F., and Theriot, J.A. (2010). Bipedal locomotion in
1491 crawling cells. Biophys. J. *98*, 933–942.
1492 60. Machemer, H., and Sugino, K. ELECTROPHYSIOLOGICAL CONTROL OF CILIARY

1493 BEATING: A BASIS OF MOTILE BEHAVIOUR IN CILIATED PROTOZOA. Comp.
1494 Biochem. Physiol *94*, 365.
1495 61. Mogami, Y., Pernberg, J., and Machemer, H. (1992). Ciliary beating in three dimensions:
1496 Steps of a quantitative description. J. Math. Biol. 1992 303 *30*, 215–249.
1497 62. McNeill Alexander, R. (2003). Principles of Animal Locomotion (Princeton University
1498 Press).
1499 63. Wan, K.Y., and Jékely, G. (2020). On the unity and diversity of cilia. Philos. Trans. R.
1500 Soc. Lond. B. Biol. Sci. *375*, 20190148.
1501 64. Ahamed, T., Costa, A.C., and Stephens, G.J. Capturing the Continuous Complexity of
1502 Behavior in C. elegans.
1503 65. Stephens, G.J., Johnson-Kerner, B., Bialek, W., and Ryu, W.S. (2008). Dimensionality
1504 and Dynamics in the Behavior of C. elegans. PLoS Comput. Biol. *4*, e1000028.
1505 66. Berman, G.J., Choi, D.M., Bialek, W., and Shaevitz, J.W. (2014). Mapping the
1506 stereotyped behaviour of freely moving fruit flies. J. R. Soc. Interface *11*, 20140672.
1507 67. Stephens, G.J., Osborne, L.C., and Bialek, W. (2010). Searching for simplicity:
1508 Approaches to the analysis of neurons and behavior. Proc. Natl. Acad. Sci. U. S. A. *108*,
1509 15565–15571.
1510 68. Devarajan, K. (2008). Nonnegative Matrix Factorization: An Analytical and Interpretive
1511 Tool in Computational Biology. PLoS Comput. Biol. *4*.
1512 69. Lee, D.D., and Seung, H.S. (1999). Learning the parts of objects by non-negative matrix
1513 factorization. Nat. 1999 4016755 *401*, 788–791.
1514 70. Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A Density-Based Algorithm for
1515 Discovering Clusters in Large Spatial Databases with Noise.
1516 71. Kauffman, L., and Rousseeuw, P.J. (1990). Finding Groups in Data: and introduction to

cluster analysis (John Wiley & Sons, Inc.).

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

73. M Cartmill (1985). Climbing. In Functional Vertebrate Morphology (Harvard University Press), pp. 73–88.

74. Kolmogoroff, A. (1936). Zur Theorie der Markoffschen Ketten. Math. Ann. *112*, 155–160.

75. Gagniuc, P. (2017). Markov Chains: From Theory to Implementation and Experimentation (Wiley).

76. Markov, A.A. (1907). 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*.

77. Howard, R.A. (1971). Dynamic and Probabilistic Systems, Volume 1: Markov models (Dover Publications).

78. Crutchfield, J.P., and Young, K. (1989). Computation at the Onset of Chaos. In Complexity, Entropy, and Physics of Information, W. Zurek, ed. (Addison-Wesley).

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

80. Moore, C. (1998). Dynamical recognizers: Real-time language recognition by analog computers. Theor. Comput. Sci.

81. Skinner, D.J., and Dunkel, J. (2021). Improved bounds on entropy production in living systems. Proc. Natl. Acad. Sci. U. S. A. *118*.

82. Martínez, I.A., Bisker, G., Horowitz, J.M., and Parrondo, J.M.R. Inferring broken detailed

1541 balance in the absence of observable currents.
1542 83. Esposito, M. (2012). Stochastic thermodynamics under coarse graining. Phys. Rev. E -
1543 Stat. Nonlinear, Soft Matter Phys. *85*, 041125.
1544 84. Taylor, C.V. (1921). Demonstration of the function of the neuromotor apparatus in
1545 Euplotes by the method of microdissection. Naturwissenschaften *9*, 910.
1546 85. Okajima, A., and Kinosita, H. (1966). Ciliary activity and coordination in Euplotes
1547 eurystomus-I. Effect of microdissection of neuromotor fibres. Comp. Biochem. Physiol.
1548 *19*, 115–131.
1549 86. Naitoh, Y., and Eckert, R. (1969). Ciliary Orientation: Controlled by Cell Membrane or by
1550 Intracellular Fibrils? Science (80-.). *166*, 1633–1635.
1551 87. Okabe, A., and Miki, F. (1984). A Conditional Nearest-Neighbor Spatial-Association
1552 Measure for the Analysis of Conditional Locational Interdependence. Environ. Plan. A
1553 Econ. Sp. *16*, 163–171.
1554 88. Jennings, H.S., and Jamieson, C. (1902). Studies on reactions to stimuli in unicellular
1555 organisms. X. The movements and reactions of pieces of ciliate infusoria. Biol. Bull. *3*,
1556 225–234.
1557 89. Shalizi, C.R., Shalizi, K.L., and Crutchfield, J.P. (2002). An Algorithm for Pattern
1558 Discovery in Time Series. ArXiv.
1559 90. Lynn, C.W., Cornblath, E.J., Papadopoulos, L., Bertolero, M.A., and Bassett, D.S. (2021).
1560 Broken detailed balance and entropy production in the human brain. Proc. Natl. Acad. Sci.
1561 U. S. A. *118*.
1562 91. Bar-Ziv, R., Tlusty, T., and Libchaber, A. (2002). Protein-DNA computation by stochastic
1563 assembly cascade. Proc. Natl. Acad. Sci. U. S. A. *99*, 11589–11592.
1564 92. Berman, G.J., Bialek, W., and Shaevitz, J.W. (2016). Predictability and hierarchy in

Drosophila behavior. Proc. Natl. Acad. Sci. U. S. A. *113*, 11943–11948.

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

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

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

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

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

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

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

100. Li, Y., Kučera, O., Cuvelier, D., Rutkowski, D.M., Deygas, M., Rai, D., Pavlovič, T., Vicente, F.N., Piel, M., Giannone, G., et al. (2022). Compressive forces stabilise microtubules in living cells. bioRxiv, 2022.02.07.479347.

101. Maniotis, A.J., Chen, C.S., and Ingber, D.E. (1997). Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure. Proc. Natl. Acad. Sci. U. S. A. *94*, 849–854.

1589 102. Craddock, T.J.A., and Tuszynski, J.A. (2010). A critical assessment of the information
1590 processing capabilities of neuronal microtubules using coherent excitations. J. Biol. Phys.
1591 *36*, 53–70.
1592 103. Stock, C., Krüppel, T., Key, G., and Lueken, W. (1999). Sexual behaviour in Euplotes
1593 raikovi is accompanied by pheromone-induced modifications of ionic currents. J. Exp.
1594 Biol. *202*.
1595 104. Tuffrau, M., Fryd-Versavel, G., Tuffrau, H., and Génermont, J. (2000). Description of
1596 Euplotes versatilis n. sp., a marine tropical ciliate exhibiting an unusually extensive
1597 phenotypic plasticity. Eur. J. Protistol. *36*, 355–366.
1598 105. Keeling, P.J. (2019). Combining morphology, behaviour and genomics to understand the
1599 evolution and ecology of microbial eukaryotes. Philos. Trans. R. Soc. B Biol. Sci. *374*.
1600 106. Kahl, A. (1935). Urtiere oder Protozoa. (Fischer).
1601 107. Yocom, H.B. (1917). The neuromotor apparatus of Euplotes patella.
1602 108. Tinevez, J.Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E.,
1603 Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: An open and
1604 extensible platform for single-particle tracking. Methods *115*, 80–90.
1605 109. King, N., Young, S.L., Abedin, M., Carr, M., and Leadbeater, B.S.C.C. (2009). Starting
1606 and maintaining Monosiga brevicollis cultures. Cold Spring Harb. Protoc. *4*,
1607 pdb.prot5148.
1608 110. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
1609 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source
1610 platform for biological-image analysis. Nat. Methods *9*, 676–682.
1611 111. Kanagal, B., and Sindhwani, V. (2010). Rank Selection in Low-rank Matrix
1612 Approximations: A Study of Cross-Validation for NMFs. In Advances in Neural

Information Processing Systems.

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

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

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

115. Kawamoto, T., and Kabashima, Y. (2017). Cross-validation estimate of the number of clusters in a network. Sci. Rep. *7*, 1–17.

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

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

118. Billingsley, P. (1961). Statistical Methods in Markov Chains (Institute of Mathematical Statistics).

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

120. Happel, J., and Brenner, H. (1981). Low Reynolds number hydrodynamics. *1*.

121. Lighthill, J. (1976). Flagellar Hydrodynamics. SIAM Rev. *18*, 161–230.

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

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

124. Kingma, D.P., and Welling, M. (2013). Auto-Encoding Variational Bayes.

125. Kingma, D.P., and Ba, L.J. (2015). Adam: A Method for Stochastic Optimization.

126. Bach, F.R., and Jordan, M.I. (2003). Learning Spectral Clustering. Adv. Neural Inf. Process. Syst. *16*.

127. Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., et al. (2011). Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

128. Darmon, D. (2020). transCSSR. GitHub Repos.