Volume segregation programming in a nematode's early embryogenesis

Guoye Guan,[1] Ming-Kin Wong,[2] Zhongying Zhao,[2,3] Lei-Han Tang,[3,4,5,*] and Chao Tang[1,6,7,†]

[1]*Center for Quantitative Biology, Peking University, Beijing 100871, China*
[2]*Department of Biology, Hong Kong Baptist University, Hong Kong, China*
[3]*State Key Laboratory of Environmental and Biological Analysis, Hong Kong Baptist University, Hong Kong, China*
[4]*Department of Physics and Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Hong Kong, China*
[5]*Complex Systems Division, Beijing Computational Science Research Center, Beijing 100094, China*
[6]*Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China*
[7]*School of Physics, Peking University, Beijing 100871, China*

(Received 21 March 2021; accepted 15 October 2021; published 23 November 2021)

Nematode species are well-known for their invariant cell lineage pattern during development. Combining knowledge about the fate specification induced by asymmetric division and the anti-correlation between cell cycle length and cell volume in *Caenorhabditis elegans*, we propose a minimal model to simulate lineage initiation by altering cell volume segregation ratio in each division, and quantify the derived pattern's performance in proliferation speed, fate diversity, and space robustness. The stereotypic pattern in *C. elegans* embryo is found to be one of the most optimal solutions taking minimum time to achieve the cell number before gastrulation, by programming asymmetric divisions as a strategy.

DOI: 10.1103/PhysRevE.104.054409

I. INTRODUCTION

From a fertilized egg to a juvenile, metazoan embryogenesis goes through several distinct stages of development. The maternal-to-zygotic transition (MZT), where maternal gene products are progressively replaced by zygotic ones, separates an initial phase of rapid cell proliferation from gastrulation where morphogenesis begins [1,2]. Over the long history of evolution, organisms have explored many different schemes of pacing cell proliferation with differentiation to optimize their developmental program [3,4]. In nematode *Caenorhabditis elegans*, MZT [5] and gastrulation [6] takes place at the 26-cell stage. However, cell fate specification starts already at the first cleavage [7,8]. The time course of subsequent cleavages is meticulously orchestrated, including reproducible division timing, volume segregation ratio, and migration trajectory of each and every cell [7–10].

Figure 1(a) shows three-dimensional time-lapse images of a wild-type embryo with GFP-marked cell nucleus and mCherry-marked cell membrane taken in our study, where details of the experimental procedure are given in the Supplemental Material (SM) [11]. The somatic lineages AB, EMS, and C each proliferate through a sequence of symmetric and synchronized cleavages where daughter cells acquire nearly the same volume and an interim fate identity [Fig. 1(b)] [7,12]. The germline cells P0-P3, on the other hand, divide asymmetrically with well-defined volume segregation ratios [Fig. 1(c)]. Figure 1(d) presents the cell cycle length against cell volume determined in our experiments. This anticorrelation has been reported in several previous studies of *C. elegans* early embryogenesis [13–15]. While somatic lineages appear to progress at high speed afforded by their larger volume, the smaller germline cells undergo elongated cell cycles that are partially attributed to the disparate volume partitioning in asymmetric divisions. Thus an asymmetric division, following a polarity cue in the mother cell [16], confers not only different fates to the daughter cells but also their respective cell cycle lengths.

The synergistic integration of fate diversification and cell division timings through asymmetric cell division suggests that *C. elegans* may have found a physical way to minimize genetic instructions for robust lineage patterning prior to the MZT. In this work, we explore this programming strategy quantitatively by examining lineage development over a broader class of volume segregation ratios. We evaluate each lineage pattern's performance in proliferation speed, fate diversity, and space robustness. The *C. elegans* lineage pattern is shown to be highly optimized, but alternative patterns with similar or even greater proliferation speed also exist, suggesting untapped programming capacity associated with asymmetric cell division.

II. LOGISTIC CLEAVAGE MODEL

As shown in Fig. 1(b), each lineage pattern is specified by a set of cleavage timings on the genealogical tree together with fate specification for daughter cells. Taking *C. elegans* embryogenesis as a reference, we developed a mathematical model to explore a much greater space of cell lineage patterns. The cell cycle length T of a cell of volume V, i.e., the time interval between birth and cleavage without cell growth, is

*lhtang@hkbu.edu.hk
†tangc@pku.edu.cn

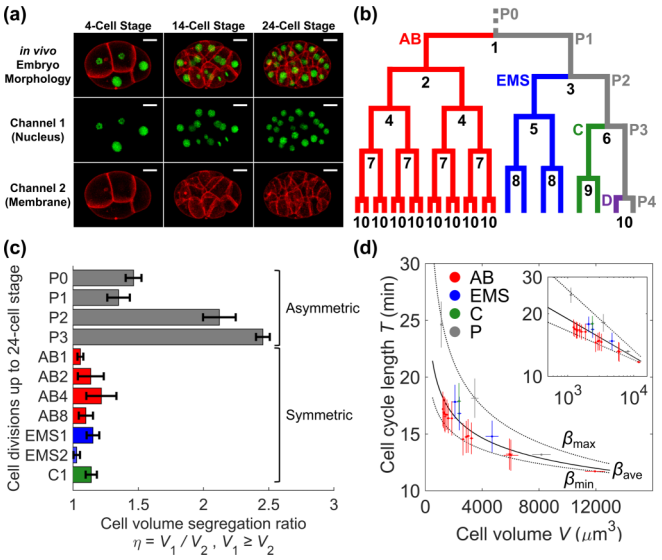

FIG. 1. *C. elegans* embryogenesis up to 24-cell stage. (a) *In vivo C. elegans* embryo morphology illustrated with fluorescence on cell nucleus (GFP, i.e., Green Fluorescent Protein) and cell membrane (mCherry); scale bar, 10 μm. (b) Lineage pattern with invariant division order and fate specification. Numbers under each division indicate the order of synchronized cleavages. (c) Cell volume segregation ratio; error bar, standard deviation. (d) Anticorrelation between cell cycle length and cell volume fitted with power-law curves; inset, data shown on the logarithmic scale.

assumed to follow an empirically observed relation [13]

$$T = T_1 (V/V_1)^{-\beta}, \tag{1}$$

where V_1 the volume of the embryonic egg. From Fig. 1(d), we see that the exponent β has a weak lineage dependence, but generally lies in the range 0.14 to 0.29. Unless otherwise specified, we adopt a common value $\beta = 0.18$ below. The volume and time scales, although not important for our theoretical study, take values $V_1 = 2 \times 10^4 \mu\text{m}^3$ and $T_1 = 10.9$ min, respectively, based on our *in vivo* experiments on *C. elegans* [9,10].

Under Eq. (1), each lineage pattern can be programed by a set of volume segregation ratios η_n at node n of the genealogical tree, starting from the zygote at the top [Fig. 1(b)]. The volume V of a descendent cell is computed from the sequence of cleavages along its ancestral path. The elapsed time to reach N cells coincides with the $(N-1)$-th division t_{N-1}. Hence the proliferation speed or rapidity is given by $P = 1/t_{N-1}$.

III. SPACE ROBUSTNESS

Continuous addition of new cells following sequential cleavages could jeopardize the canonical cell movement during mechanical equilibration inside the egg, leading to defect patterns. In early embryogenesis, *C. elegans* solves this problem by synchronizing cell divisions into temporally clustered events as illustrated in Fig. 2 [10,14,17]. To enforce this property among lineage patterns generated from our model, we introduce two time constants: δ_s for cleavage clustering and δ_a for minimal time gap between clusters. The parameter δ_s needs to be sufficiently small so that new cells created within

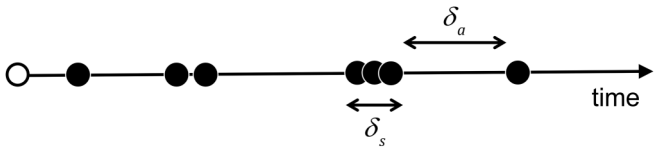

FIG. 2. A sequence of cell cleavage events during early embryogenesis whose timings are illustrated by solid circles along the time axis. Events that fall within a window of size δ_s form a cluster while neighboring clusters or isolated events need to be separated by a minimal time gap $\delta_a > \delta_s$ to ensure punctuated cell movement.

this time interval will move coherently to their equilibrium positions. In contrast, δ_a needs to be longer than the typical completion time of the equilibration process. Based on our experimental cell trajectory analysis [10], we set $\delta_s = 1.5$ min and $\delta_a = 3.0$ min. In the following, we shall use cleavage clustering to classify and compare lineage patterns.

IV. FATE SPECIFICATION

Cell fate determination is a complex process that mostly follows the lineage path but could be affected by other factors such as the local cellular environment. Here, we focus on fate diversification during the early embryogenesis and specify the number of distinct fates F in an N-cell embryo simply from the location of asymmetric divisions in the lineage pattern. Starting from $F = 1$ for the zygote, each symmetric division propagates a given lineage to the next generation without diversification. On the other hand, an asymmetric division creates two new lineages. When the mother cell is the only carrier of the previous lineage, F increases by one. Otherwise, F increases by two. An exception to the last case is when two or more sister cells or cousins in the same lineage divide asymmetrically at the same time, in which case daughters of larger volume share the same fate and the smaller ones another fate. This scheme of fate specification is further explained in the SM.

By analyzing images illustrated in Fig. 1(a) from different embryos, we found that symmetric division in the somatic cell lineages may also yield daughter cells of slightly different sizes, attributable to random fluctuations. Therefore we introduce a division asymmetry threshold η_c such that two daughter cells are considered as founder cells of new lineages only when the volume segregation ratio η of the mother exceeds η_c. Based on the data in Fig. 1(c), we set $\eta_c = 1.28$ to separate symmetric and asymmetric cleavages [9].

V. MODEL EXPLORATION

We devised a two-round sampling scheme to explore the space of lineage patterns, focusing on the 24-cell embryo. To give sufficient weight to symmetric divisions, a truncated Gaussian distribution is adopted for the volume ratio $\eta = \max(1, \xi)$, where ξ is a Gaussian random variable with mean μ and standard deviation σ.

The detailed simulation protocol is presented in the SM, Fig. S3. In the first round, we take $\mu = \sigma = r\eta_{\max}$, where $\eta_{\max} = 3.21$ is the maximum cell volume segregation ratio detected among the first seven generations of cells in the

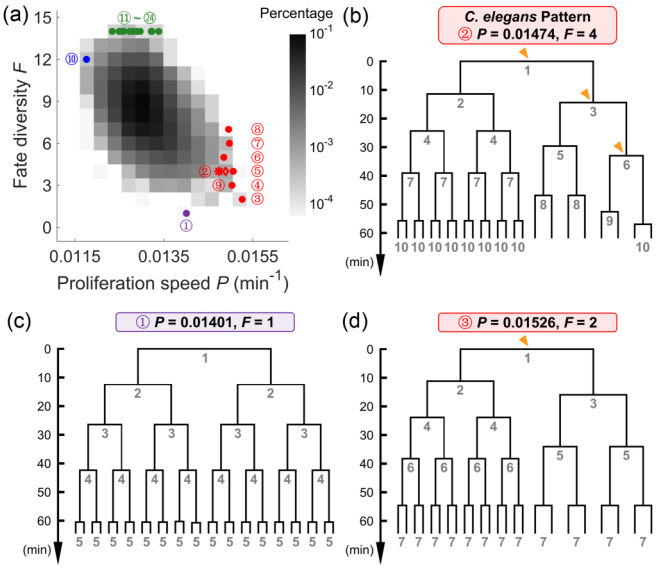

FIG. 3. Lineage solutions at $N = 24$ from random sampling. (a) Distribution of the 19 654 nonidentical solutions in P-F space; lineage patterns with the lowest and highest fate diversity (Patterns No. 1 and Nos. 11–24), the lowest and highest proliferation speed (Patterns No. 10 and Nos. 3–8) are indicated with purple, green, blue, and red points, respectively; the *C. elegans* pattern is denoted by a red asterisk (Pattern No. 2). A complete list of the labeled lineage patterns can be found in the SM, Figs. S4 and S5. (b) The *C. elegans* pattern. (c) The fully symmetric solution (Pattern No. 1). (d) The fastest solution (Pattern No. 3). Orange triangles denote the asymmetric divisions with $\eta > \eta_c$.

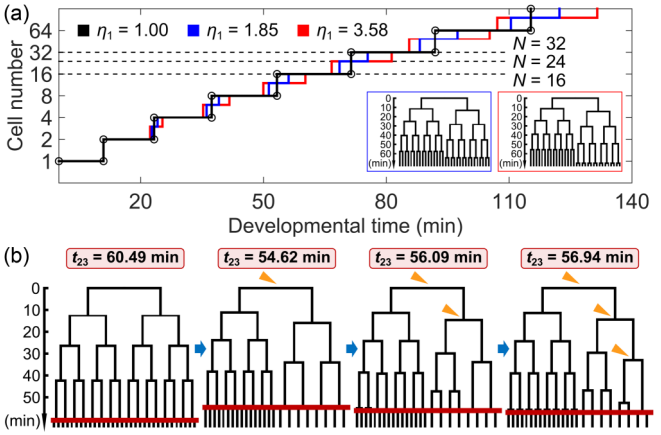

FIG. 4. Proliferation speed under asymmetric divisions. (a) Cell number against developmental time with a single asymmetric division of volume ratio η_1 at the root. (b) A path from the fully symmetric pattern to the *C. elegans* pattern through successive addition of asymmetric divisions (orange triangles); red lines indicate the time to reach 24 cells.

experiment [9]. Here r is a fraction number taking values from $1/6$ to 1 in steps of $1/6$. For each r, a total of $Q = 5 \times 10^5$ independent sets of $\{\eta_n\}$'s are generated and then filtered under the clustering criterion shown in Fig. 2. In the second round, we sample the space around the first round solutions $\{\eta_n\}$ by substituting $\eta_n \to \eta_n' = \max(1, \eta_n + \xi')$, where ξ' is another Gaussian random variable with mean 0 and standard deviation $r\eta_{\max}$. The procedure is repeated Q times for each solution from the first round. Results for different r values are then merged to yield a final set of 19 654 viable lineage patterns.

Figure 3(a) presents our lineage solutions in the P-F plane, which exhibit considerable spread. The *C. elegans* lineage pattern shown in Fig. 3(b) is located in the lower right corner of the plot (Pattern No. 2 indicated by the asterisk). Its proliferation speed ranks among the top 0.2%. The remarkable performance persists when Eq. (1) is replaced by other functional forms, supporting the general validity of this conclusion (SM, Table S2).

Fate diversity can be tuned extensively in our model. The lineage pattern with the least diversity ($F = 1$) consists of only symmetric divisions [Pattern No. 1 in Fig. 3(c)]. However, some can accommodate 14 fates by introducing 11—13 asymmetric divisions during development [green dots in Fig. 3(a) and SM, Fig. S5]. Also seen is a tiny fraction of solutions that proliferate faster than the *C. elegans* pattern [red dots in Fig. 3(a)]. An example is shown in Fig. 3(d) (Pattern No. 3). Their fate diversities are limited to the range 2–7 (see SM, Fig. S4). Overall, the data exhibits a statistical tradeoff between the proliferation speed P and fate diversity F.

VI. SPEED OPTIMIZATION

To better understand the oval-shaped distribution seen in Fig. 3(a), we start with Pattern No. 3 at the lower right corner which has only two lineages following the first cleavage. The timing of subsequent cell proliferation events can be easily computed in our model. Figure 4(a) shows our results for selected values of the zygotic volume segregation ratio η_1. Two sample lineages with smaller and larger values of η_1 are shown in the inset. At specific time points, the number of cells in the symmetric case ($\eta_1 = 1$) changes from n to $2n$, where n is a power of 2. Asymmetric division creates a larger and a smaller blastomere which become founder cells of respective lineages. In each generation, $n/2$ cells have a slightly greater volume than the other half and hence divide at an earlier time. Consequently, the embryo reaches $3n/2$ cells faster than the fully symmetric case. The latter series includes the 24-cell embryo. Tuning up η_1 shortens the time for the faster lineage to reach 16 cells until it overlaps with the previous event where the slower lineage turns into 8 cells. At this point, we arrive at Pattern No. 3 that leads all other lineage patterns in speed and grows approximately 6 min faster than the fully symmetric pattern [Fig. 3(d)].

Likewise, the *C. elegans* pattern can be considered to have evolved from the fully symmetric case by adding three asymmetric divisions onto its smaller blastomere sequentially [Fig. 4(b)]. In theory, any additional asymmetric division prolongs the time for the smaller blastomere to reach 8 cells, thus decreasing the proliferation speed. Nevertheless, the total time loss between Pattern No. 3 and the *C. elegans* pattern is less than 2.5 min. It is worth pointing out that, when the target cell number is a power of 2, the fully symmetric pattern always has the highest P value. These findings together explain the counterintuitive phenomenon that meticulously programed asymmetric divisions make the embryo grow faster than the

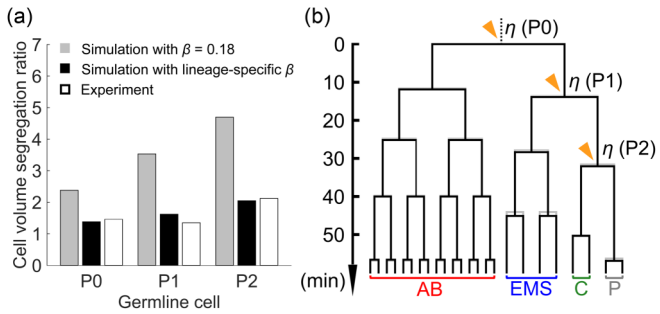

FIG. 5. Division timings up to 24 cells with constant and lineage-specific β values. (a) The cell volume segregation ratio of germline cells P0-P2 derived by simulations and measured in experiment. (b) The *C. elegans* patterns plotted with constant (gray) and lineage-specific (black) β, using the η values shown in (a).

fully symmetric one, while the *P-F* tradeoff is maintained statistically.

We have checked that the above scheme applies to all fast-proliferating lineage patterns at the lower right corner of Fig. 3(a) (see also SM, Fig. S4). Their zygote divides asymmetrically and generates two fate-specific blastomeres; then the larger one undergoes symmetric and synchronous divisions to produce identical descendants, while the asymmetric divisions occur in another blastomere (Patterns Nos. 2–5 and 9). Alternatively, asymmetric divisions can be programed in the larger blastomere as well to allow more cell fates with only slightly reduced proliferation speed (Patterns Nos. 6–8).

VII. LINEAGE INTERFERENCE

The fast-proliferating solutions have a relatively small number of distinct lineages. Within each lineage, cell divisions are synchronized, hence only a few temporal clusters of cleavage events are needed to reach the target cell number N. On the other hand, for a solution with many lineages or equivalently a large F, downstream temporal clustering and spacing of cleavage events as shown in Fig. 2 become more intertwined under the minimal separation requirement. We name this phenomenon "lineage interference". In the SM, we present simulation results to show that the number of distinct lineage patterns reaches a maximum around $N = 24$ under the model parameter values derived from *C. elegans* embryogenesis. Beyond this cell number, the solution space for lineage optimization shrinks rapidly.

VIII. FINE TUNING OF VOLUME SEGREGATION RATIOS

To achieve optimal proliferation speed under the three biophysical constraints considered in our study, the respective η values at asymmetric cell divisions need to be adjusted carefully. We performed simulations of the *C. elegans* pattern using both a constant $\beta = 0.18$ and lineage-specific values given in Table S3 (SM), which provide a more accurate description of the respective cell cycles [see Fig. 1(d)]. The interval between AB and P1 divisions is set to 2 min according to previous reports [18,19]. Figure 5(a) gives the η values for the three asymmetric divisions that optimize the proliferation speed up to 24 cells under constant and lineage-dependent β values, respectively. Huge disparity in size between somatic and germline cells is seen when a constant β is adopted in the simulation (gray bars). In comparison, the optimal η values in the lineage-specific case (black bars) are much closer to the experimental ones (open bars). The simulated lineage patterns for the two cases are shown in Fig. 5(b). Due to the compensating effect of β and η values, the lineage patterns are essentially indistinguishable from each other, and from that of experiments, giving further credence that proliferation speed is optimized during the organism's evolution.

IX. DISCUSSION

Eutelic organisms have stereotypic developmental programs down to the single-cell level, especially for its cell lineage pattern. Why a nematode embryo programs cell divisions with specific order and segregation ratios has been elusive. In this work, a simple model is proposed to evaluate a lineage pattern's performance in proliferation speed, fate diversity and space robustness, on the basis of two assumptions derived from *C. elegans* early embryogenesis. The first is fate specification induced by asymmetric segregation of cell volume and its molecular contents; the second is the anticorrelation between cell cycle length and cell volume. Given that space robustness is essential for developmental accuracy at cellular resolution, we simulate and compare lineage patterns that exhibit punctuated cell movement. The *C. elegans* pattern is well reproduced by our model with an outstanding proliferation speed, suggesting a minimum time principle at work. This principle was also found in the development of mouse intestinal crypts, which leads to a sharp transition between symmetric and asymmetric stem cell divisions [20]. The fitness gain from speed optimization may arise from competition among offsprings, threats by predators or by environment hostile to the cell or embryo [2].

Solutions (e.g., Patterns Nos. 3–8) that proliferate as fast as the *C. elegans* pattern might be used by other nematode species [21]. We identified a pattern that has one more asymmetric division compared to the *C. elegans* pattern in the first 4 cells, leading to significant asynchrony in their daughters, which is observed in a free-living marine nematode *Plectus sambesii* [22] [Pattern No. 9 in Fig. 3(a) (red diamond) and SM, Fig. S4]. Besides, other biological, biophysical or environmental factors may affect lineage pattern selection as well. For instance, Pattern No. 5 has the same fate diversity as the *C. elegans* pattern, and even higher proliferation speed [SM, Fig. S4]. Why the *C. elegans* pattern is selected among those solutions remains to be answered. One possibility is that topological features not considered in our model, such as polarity establishment and reorientation and rotational motion of cells [16,23], introduce additional physical constraints that merit further study.

ACKNOWLEDGMENTS

We thank F. Liu, X. Yang, X. Kuang, K. Kong, and C. Luo for helpful discussions and comments. This work was supported by the National Natural Science Foundation of China (Grants No. 11635002,

No. 12090053, No. 32088101, No. U1430237, and No. U1530401), the Hong Kong Research Grants Council (Grants No. HKBU12100118, No. HKBU12100917, No. HKBU12123716, No. HKBU12324716, and No. HKBU12301514), and the HKBU Interdisciplinary Research Cluster Fund. Computation was performed partly on the High-Performance Computing Platform at Peking University.

[1] W. Tadroz and H. D. Lipshitz, Development **136**, 3033 (2009).
[2] P. H. O'Farrell, Cold Spring Harb. Perspect. Biol. **7**, a019042 (2015).
[3] P. Heyn, M. Kircher, A. Dahl, J. Kelso, P. Tomancak, A. T. Kalinka, and K. M. Neugebauer, Cell Rep. **6**, 285 (2014).
[4] M. T. Lee, A. R. Bonneau, and A. J. Giraldez, Annu. Rev. Cell Dev. Biol. **30**, 581 (2014).
[5] M. K. Wong, D. Guan, K. H. C. Ng, V. W. S. Ho, X. An, R. Li, X. Ren, and Z. Zhao, J. Biol. Chem. **291**, 12501 (2016).
[6] J. Nance and J. R. Priess, Development **129**, 387 (2002).
[7] J. E. Sulston, E. Schierenberg, J. G. White, and J. N. Thomson, Dev. Biol. **100**, 64 (1983).
[8] R. Schnabel and J. R. Priess, in *C. elegans II*, edited by D. R. Riddle *et al.* (Cold Spring Harbor Laboratory Press, New York, 1997), pp. 361–382.
[9] J. Cao, G. Guan, M. K. Wong, L. Y. Chan, C. Tang, Z. Zhao, and H. Yan, bioRxiv (preprint) (2019), https://doi.org/10.1101/797688.
[10] G. Guan, M. K. Wong, V. W. S. Ho, X. An, L. Y. Chan, B. Tian, Z. Li, L.-H. Tang, Z. Zhao, and C. Tang, bioRxiv (preprint) (2019), https://doi.org/10.1101/776062.
[11] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.104.054409 for the details of experimental procedures, modeling methods, and analytical results, which includes Refs. [5,13–15,24–28].
[12] Z. Bao, Z. Zhao, T. J. Boyle, J. I. Murray, and R. H. Waterston, Dev. Biol. **318**, 65 (2008).
[13] Y. Arata, H. Takagi, Y. Sako, and H. Sawa, Front. Physiol. **5**, 529 (2015).
[14] R. Fickentscher, P. Struntz, and M. Weiss, Phys. Rev. Lett. **117**, 188101 (2016).
[15] R. Fickentscher, S. W. Krauss, and M. Weiss, New J. Phys. **20**, 113001 (2018).
[16] L. Hubatsch, F. Peglion, J. D. Reich, N. T. L. Rodrigues, N. Hirani, R. Illukkumbura, and N. W. Goehring, Nat. Phys. **15**, 1078 (2019).
[17] B. Tian, G. Guan, L.-H. Tang, and C. Tang, Phys. Biol. **17**, 026001 (2020).
[18] M. Brauchle, K. Baumer, and P. Gönczy, Curr. Biol. **13**, 819 (2003).
[19] B. Schumacher, A. Alpi, and A. Garter, Curr. Biol. **13**, R560 (2003).
[20] S. Itzkovitz, I. C. Blat, T. Jacks, H. Clevers, and A. V. Oudenaarden, Cell **148**, 608 (2012).
[21] J. Schulze and E. Schierenberg, EvoDevo **2**, 18 (2011).
[22] J. Schulze, W. Houthoofd, J. Uenk, S. Vangestel, and E. Schierenberg, EvoDevo **3**, 13 (2012).
[23] P. Gross, K. V. Kumar, N. W. Goehring, J. S. Bois, C. Hoege, F. Jülicher, and S. W. Grill, Nat. Phys. **15**, 293 (2019).
[24] J. Cao, G. Guan, V. W. S. Ho, M. K. Wong, L. Y. Chan, C. Tang, Z. Zhao, and H. Yan, Nat. Commun. **11**, 6254 (2020).
[25] L. Chen, V. W. S. Ho, M. K. Wong, X. Huang, L. Y. Chan, H. C. K. Ng, X. Ren, H. Yan, and Z. Zhao, Genetics **209**, 37 (2018).
[26] J. I. Murray and Z. Bao, Cold Spring Harb. Protoc. **2012**, 887 (2012).
[27] C. C. Mello, B. W. Draper, and J. R. Priess, Cell **77**, 95 (1994).
[28] C. J. Thorpe, A. Schlesinger, J. C. Carter, and B. Bowerman, Cell **90**, 695 (1997).