

## ARTICLE OPEN



# MorphoSim: an efficient and scalable phase-field framework for accurately simulating multicellular morphologies

Xiangyu Kuang <sup>1,6</sup>, Guoye Guan <sup>1,6</sup>, Chao Tang <sup>1,2,3</sup>✉ and Lei Zhang <sup>1,4,5</sup>✉

The phase field model can accurately simulate the evolution of microstructures with complex morphologies, and it has been widely used for cell modeling in the last two decades. However, compared to other cellular models such as the coarse-grained model and the vertex model, its high computational cost caused by three-dimensional spatial discretization hampered its application and scalability, especially for multicellular organisms. Recently, we built a phase field model coupled with *in vivo* imaging data to accurately reconstruct the embryonic morphogenesis of *Caenorhabditis elegans* from 1- to 8-cell stages. In this work, we propose an improved phase field model by using the stabilized numerical scheme and modified volume constriction. Then we present a scalable phase-field framework, *MorphoSim*, which is 100 times more efficient than the previous one and can simulate over 100 mechanically interacting cells. Finally, we demonstrate how *MorphoSim* can be successfully applied to reproduce the assembly, self-repairing, and dissociation of a synthetic artificial multicellular system - the *synNotch* system.

*npj Systems Biology and Applications* (2023)9:6; <https://doi.org/10.1038/s41540-023-00265-w>

## INTRODUCTION

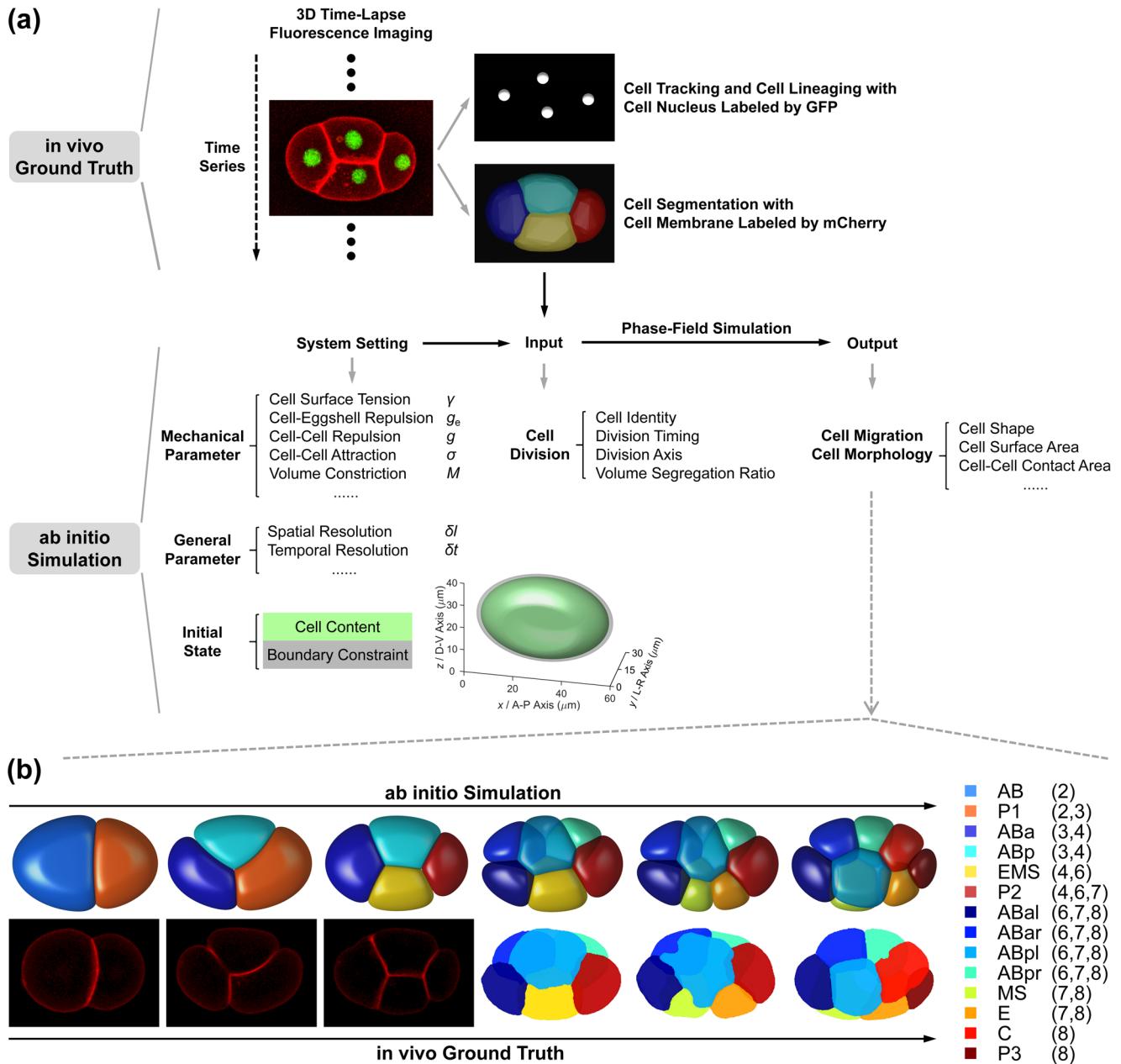
The multicellular systems formed by physically interacting cells are widespread in animals, plants, and microorganisms like fungi and choanoflagellates<sup>1–4</sup>. They usually consist of multiple cell types and characteristic spatial organizations, ranging in scales from tissue, and organ, to an individual, and are often involved with embryogenesis and organogenesis<sup>5,6</sup>. For example, the mouse valley-like small intestine, which contains ~250 cells with six cell types in specific proportions, plays a role in nutrient uptake and has strong architectural robustness and regenerative capacity<sup>7–9</sup>. The mechanics of interacting cell aggregates, as well as their morphological and morphogenetic effects, has attracted increasing attention in the fields of cell biology, developmental biology, cancer biology, etc.<sup>10–12</sup>. Apart from the natural ones, many engineered systems have been constructed *in vitro* artificially, such as the organoid and embryoid used for developmental biology research and high-throughput drug test<sup>13–15</sup>. Furthermore, synthetic biologists are trying to program the cellular interaction *de novo* to build customized multicellular living machines, robots, or patterns, using bottom-up or top-down engineering approaches. The *synNotch* and *Xenobot* systems are two cutting-edge representatives, which are constructed for designed functions using hundreds of animal cells and have broad application areas like medical treatment and synthetic development<sup>16–19</sup>. An urgent need for a computational tool that can efficiently and accurately simulate multicellular morphological behaviors is emerging<sup>20,21</sup>.

Many physical models have been utilized to study the morphological dynamics of multicellular systems, including cellular automaton<sup>22</sup>, cellular Potts model<sup>23</sup>, coarse-grained model<sup>24</sup>, Voronoi tessellation model<sup>25</sup>, vertex model<sup>26</sup>, multi-particle model<sup>27</sup>, phase field model<sup>28–30</sup>, etc. These models have provided much mechanistic knowledge and many insights for understanding biological processes. However, simulation results of the same system may vary from model to model—when a precise

description of the real system is critical the accuracy of the model output becomes an important issue<sup>31</sup>. Besides, most of the existing models often lack detailed validation by quantitative comparison with experiments, especially at three-dimensional (3D) and single-cell levels. Among the existing multicellular models, the phase field model, which uses a continuous diffusion field to describe a cell, is emerging quickly these years for it can comprehensively characterize cell shapes and interface-based cell-cell interactions<sup>32–38</sup> and cell-substrate interactions<sup>39,40</sup> without explicitly tracking the cell interface; besides, it can incorporate intracellular fluid dynamics<sup>41</sup>, biochemical regulatory mechanisms such as the reaction-diffusion of interacting molecules<sup>42</sup>, and the actin flow that mediates cell motility<sup>40</sup>.

Given that the roundworm *Caenorhabditis elegans* embryo has stereotypic cell arrangement patterns among individuals and lots of imaging data about its embryogenesis has been collected, the phase field model has been introduced into this system with quantitative comparisons between *in silico* and *in vivo* morphologies being carried out<sup>43–46</sup>. On the one hand, Seirin-Lee et al. (2022) devised a phase field model that considers cell surface tension, cell–eggshell and cell-cell repulsion, cell-cell attraction, and cell volume constriction; they successfully reproduced the T-reverse-shaped pattern observed in RNAi-treated 4-cell embryos with a slim eggshell and eliminated cell adhesion, which however failed to be reproduced by a coarse-grained model as reported before<sup>46,47</sup>. On the other hand, Kuang et al. (2022) constructed a 3D phase field model with the cell morphology data collected by fluorescence imaging and membrane segmentation (Fig. 1a)<sup>43,48</sup>; the model can accurately regenerate the *in vivo* cell morphologies and the conserved cell-cell contact map and infer the underlying biophysical properties (e.g., intercellular adhesion) from 1- to 8-cell stages (Fig. 1b). Compared to another two mechanical models established in *C. elegans* embryogenesis research, i.e., the multi-particle model (that describes cell shape with many interacting particles and many parameters<sup>27,49</sup>) and the coarse-grained model

<sup>1</sup>Center for Quantitative Biology, Peking University, Beijing 100871, China. <sup>2</sup>Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China. <sup>3</sup>School of Physics, Peking University, Beijing 100871, China. <sup>4</sup>Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China. <sup>5</sup>Center for Machine Learning Research, Peking University, Beijing 100871, China. <sup>✉</sup>These authors contributed equally: Xiangyu Kuang, Guoye Guan. <sup>✉</sup>email: tangc@pku.edu.cn; zhangl@math.pku.edu.cn



**Fig. 1 The phase-field framework for *C. elegans* early embryogenesis. a** Flow chart. **b** Comparison between ab initio simulation and in vivo ground truth of *C. elegans* embryonic morphologies from 2- to 8-cell stages. The embryonic structures at 2-, 3-, and 4-cell stages are two-dimensional and shown by fluorescence images; the ones at 6-, 7-, and 8-cell stages are three-dimensional and shown by segmented images. The cells' identities and corresponding colors and existing stages (represented by the total cell number of the stage) are denoted on the right. All the subfigures, except the schematics for the initial state setting in (a), are adapted from ref. <sup>43</sup> with granted permission.

(that describes a cell as a single particle and neglects the cell shape)<sup>47,50–53</sup>, the phase field model can both simulate the cell shape and shape-associated interaction and has a considerable number of biophysically-significant parameters. Despite its outstanding performance, the model is still limited by the high computational cost caused by spatial discretization and increasing cell numbers, especially when the cell number reaches dozens and hundreds.

In this work, we propose an efficient and scalable phase-field framework that can accurately simulate multicellular morphologies. We first develop a stabilized numerical scheme that allows for large-time steps. Next, we enhance the precision of cell volume control to avoid “cell disappearance”. Finally, a phase-field framework, *MorphoSim*, is established along with Matlab-based

software. By testing on the simulations of *C. elegans* embryogenesis, it can achieve a computational efficiency of more than 100 times the previous one and is capable of computing over 100 cells. As an application, the *MorphoSim* framework can successfully reproduce the assembly, self-repairing, and dissociation of the *synNotch* system reported in ref. <sup>16</sup>.

## RESULTS

### Review of the phase field model

The original phase field model adopted by this study considered the surface tension  $F_{\text{ten}}$  and volume constriction  $F_{\text{vol}}$  imposed on a cell, and the repulsion  $F_{\text{rep}}$  and attraction  $F_{\text{atr}}$  between cells<sup>29,30,43,54</sup>. The spatial constraint from the eggshell is also

repulsive to a cell toward the embryo's center and thus is included in the term  $\mathbf{F}_{\text{rep}}$ ; the eggshell boundary is set as a truncated ellipsoid fitted with experimental data (Fig. 1a and Supplementary Note 1)<sup>43,48</sup>. For a system composed of  $N$  cells, the  $i$ -th cell with a prescribed volume  $V_i(t)$  is represented by a phase field  $\phi_i(\mathbf{r}, t)$  and assumed to deform and migrate in an overdamped domain  $\Omega$ , following the governing equations:

$$\mathbf{F}_{\text{ten}} = -\gamma(\Delta\phi_i - cW'(\phi_i)) \frac{\nabla\phi_i}{|\nabla\phi_i|^2}, \quad (1)$$

$$\mathbf{F}_{\text{vol}} = M \left( \int_{\Omega} \phi_i d\mathbf{r} - V_i(t) \right) \hat{\mathbf{n}}, \quad (2)$$

$$\mathbf{F}_{\text{rep}} = \left( g_e \phi_i \phi_e^2 + g \phi_i \sum_{j \neq i}^N \phi_j^2 \right) \frac{\nabla\phi_i}{|\nabla\phi_i|^2}, \quad (3)$$

$$\mathbf{F}_{\text{atr}} = \sum_{j \neq i}^N \sigma_{ij} \nabla\phi_j, \quad (4)$$

$$\frac{d\phi_i}{dt} = -\frac{1}{\tau} (\mathbf{F}_{\text{ten}} + \mathbf{F}_{\text{vol}} + \mathbf{F}_{\text{rep}} + \mathbf{F}_{\text{atr}}) \cdot \nabla\phi_i, \quad (5)$$

Here  $\Delta$  and  $\nabla$  are Laplacian and gradient operators respectively;  $\gamma$  denotes the cell surface tension that drives the cell shape to be spherical;  $c$  controls the cell boundary thickness;  $W(\phi) = \phi^2(\phi - 1)^2$  separates two phases at  $\phi = 0$  and  $\phi = 1$ , corresponding to the exterior and interior of a cell; on the contrary,  $\phi_e$ , the phase field of the eggshell, describes its outer space with 1 and its inner space with 0, consequently constraining the phase fields of cells and keeping the multicellular system evolving inside the eggshell;  $g_e$  and  $g$  represent the cell–eggshell and cell–cell repulsions respectively;  $\sigma_{ij}$  represents the attraction strength between the  $i$ -th and  $j$ -th cells;  $V_i(t)$  is the target volume of the  $i$ -th cell at the specific time point  $t$  and the designated volume can be obtained from experimental measurement or set automatically;  $M$  is the volume constriction strength;  $\hat{\mathbf{n}}$  is the unit normal vector at the cell surface and orients inward;  $\tau$  is the ambient viscosity.

Division of the  $i$ -th cell is simplified as an instant splitting of the phase field denoted as  $\phi_i^*$  into two new regions, according to a certain splitting plane. The plane is normal to the cell division axis  $\mathbf{n}$  and is located in  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_c) - b = 0$ , where  $\mathbf{r}_c = \frac{\int_{\Omega} \mathbf{r} \phi_i^* d\mathbf{r}}{\int_{\Omega} \phi_i^* d\mathbf{r}}$  is the center of  $\phi_i^*$  and  $b$  is uniquely determined by setting a designated volume ratio to the two separate regions. Subsequently, the division of the phase field of the  $i$ -th cell is implemented as follows<sup>55</sup>:

$$\phi_{N+1} = \phi_i^* \left( \tanh \frac{\frac{\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_c) - b}{\epsilon} + 1}{2} \right), \quad (6)$$

$$\phi_i = \phi_i^* \left( \tanh \frac{\frac{b - \mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_c)}{\epsilon} + 1}{2} \right). \quad (7)$$

where  $\phi_{N+1}$  and  $\phi_i$  are the initial state of the two phase fields generated by cell division;  $\epsilon$  represents the width of the splitting interface.

During *C. elegans* embryogenesis, every cell has unique and identifiable developmental behavior and is systematically named based on its cell type, lineal origin, and spatial location<sup>44</sup>. Remarkably, the confocal 3D time-lapse fluorescence imaging on the cell nucleus and cell membrane has enabled cell-resolved monitoring for *C. elegans* embryogenesis at ~1.5 intervals (Fig. 1(a))<sup>48</sup>. On the one hand, the cell tracking and cell lineageing based on the GFP-labeled cell nucleus provides information on cell identity, cell division timing and order, and cell division axis;

on the other hand, the cell segmentation based on the mCherry-labeled cell membrane provides the volume of each cell and the volume segregation ratio in each cell division. Besides, fluorescence labeling on adhesive protein HMR-1 has shown that it's rarely accumulated in the newly-formed membrane between sister cells and in some specific cell–cell contacts, including EMS-P2 at the 4-cell stage and ABpl-E at the 8-cell stage<sup>47,56</sup>; for the simplicity, the cell–cell attraction in the phase field model is binarized into relatively weak ( $\sigma_w$ ) for the abovementioned contacts and strong ( $\sigma_s$ ) for the others<sup>43</sup>. With the input of cell division order and axis and volume segregation ratio measured experimentally (Supplementary Table 1), the previous phase-field framework based on Eqs. (1–7) successfully reproduced the typical embryonic morphologies seen in the experiment up to the 8-cell stage (Fig. 1b)<sup>43</sup>. The conserved cell–cell contact map observed in vivo was fully reproduced and the regulatory programs on cell–cell adhesion were reversely inferred, including the relatively weak adhesion in EMS-P2 and ABpl-E contacts reported before<sup>47,56</sup>. Those results have validated the applicability of the mathematical forms of the governing equations. The simulation pipeline about how the cell–cell adhesion is assigned and inferred and how the cell division timing is set for each stage is summarized in Supplementary Note 1. To distinguish the time scales with different meanings, hereafter the time in the experiment and computer are referred to as “in vivo time” and “in silico time” (in silico time = step number  $\times$  step size =  $n_t \delta t$ ) respectively, while the time cost for simulation is termed “computing time” and is one of the optimization targets in this work.

### Stabilized numerical scheme

The high computational cost of phase-field simulation is largely attributed to its 3D spatial discretization and cell number increase. One direct attempt to reduce the computational cost is to minimize the spatial and temporal resolutions while guaranteeing the results are consistent with the previous biological findings<sup>43</sup>. When adopting the parameter assignments from the original framework<sup>43</sup> (Supplementary Table 2) and inputting the cell division order and axis and volume segregation ratio from in vivo *C. elegans* embryos<sup>48</sup> (Supplementary Table 1), the requirements that judge if the framework works precisely enough from the 1- to 8-cell stages mainly include three parts: (1) the cell–cell contact maps are the same as the ones conserved between individual embryos (Supplementary Fig. 1a–d); (2) the previously reported cell–cell adhesion programs (i.e., relatively weak adhesion between EMS and P2 cells at the 4-cell stage and between ABpl and E cells at 8-cell stage) can be inferred by morphological comparison to experiment and parameter scanning (Fig. 1b and Supplementary Figs. 2a, b, 3a–c)<sup>47,56</sup>; (3) the embryonic morphologies resemble the ones in vivo (Fig. 1b and Supplementary Fig. 3a, b). The details of the simulation procedure and quantitative criteria are introduced in Supplementary Note 1. Given the original spatial grid size  $\delta l = 0.25 \mu\text{m}$  and time step size  $\delta t = 0.10$  (corresponding to 0.0036 s in vivo) from ref. <sup>43</sup>, we scan  $\delta l$  from 0.25 to 2.00  $\mu\text{m}$  and  $\delta t$  from 0.1 to 0.4 by repeating the simulation procedure from 1- to 8-cell stages as established in Supplementary Note 1 and Supplementary Fig. 4a. In each simulation, the root-mean-square velocity of all cells' mass centers (i.e.,  $\bar{v}$ ) is calculated at each time step and used for triggering the cell division(s) in an experimentally-observed order (Table 1). For 1-, 2-, 3-, and 4-cell stages, the next round of cell division(s) takes place when the system reaches its steady state (defined by  $\bar{v} < 1 \times 10^{-4}$ ); for 6- and 7-cell stages, the system is allowed to evolve to its first quasi-steady state (defined by  $\frac{d\bar{v}}{dt}|_{t=t_q} = 0, \frac{d^2\bar{v}}{dt^2}|_{t=t_q} > 0$ ) and then cell division(s) takes place; finally, the 8-cell stage lasts for a constant in silico time = 15,000 (Supplementary Note 1)<sup>43</sup>. As a result, we

find that the three requirements can be satisfied by a coarser spatial grid size  $\delta l = 0.50 \mu\text{m}$  (grid nodes in  $x$ ,  $y$ , and  $z$  axes:  $120 \times 60 \times 80$ ) and time step size  $\delta t = 0.30$  (Fig. 4b, c). However, such improvement is very limited.

**Table 1.** Computing time in GPU and CPU using the previous and current frameworks.

Cell Stage	Computing time			
	Previous framework <sup>43</sup>		Current framework ( <i>MorphoSim</i> )	
	GPU <sup>a</sup>	CPU <sup>b</sup>	GPU <sup>a</sup>	CPU <sup>b</sup>
6-Cell Stage	35.3 min	14280.0 min	0.3 min	4.1 min
7-Cell Stage	30.9 min	11161.0 min	0.1 min	2.0 min
8-Cell Stage	229.8 min	7767.8 min	2.6 min	39.5 min
25-Cell Stage	854.3 min	26067.4 min	17.2 min	186.6 min
50-Cell Stage	— <sup>c</sup>	67629.4 min	29.8 min	356.3 min
100-Cell Stage	— <sup>c</sup>	— <sup>c</sup>	93.8 min	706.1 min

<sup>a</sup>GPU, NVIDIA Tesla P100.

<sup>b</sup>CPU, Inter Xeno E5-2680 v3.

<sup>c</sup>“—” represents “out of memory”.

The stability restriction on the time step size impedes the further acceleration of computation. In order to allow a much larger time step size than the explicit schemes, we first adopt a first-order semi-implicit scheme<sup>57</sup>:

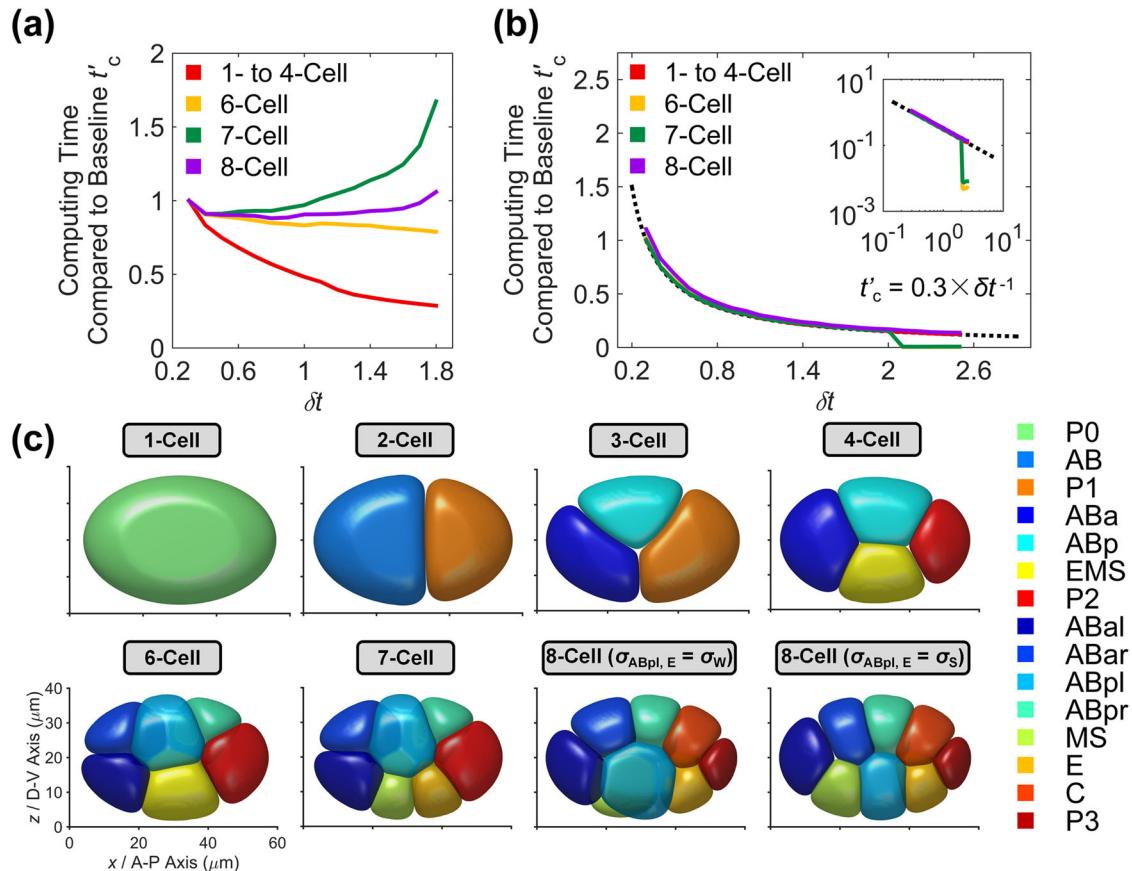
$$\frac{\tau}{\delta t} (\phi_i^{n+1} - \phi_i^n) = \gamma \Delta \phi_i^{n+1} + F_i^n, \quad (8)$$

where  $n$  denotes the  $n$ -th time step and the Fourier spectral method is used for the spatial discretization throughout this work. The linear term  $\gamma \Delta \phi_i^{n+1}$  was treated implicitly while the nonlinear term  $F_i^n$  is treated explicitly and expressed by:

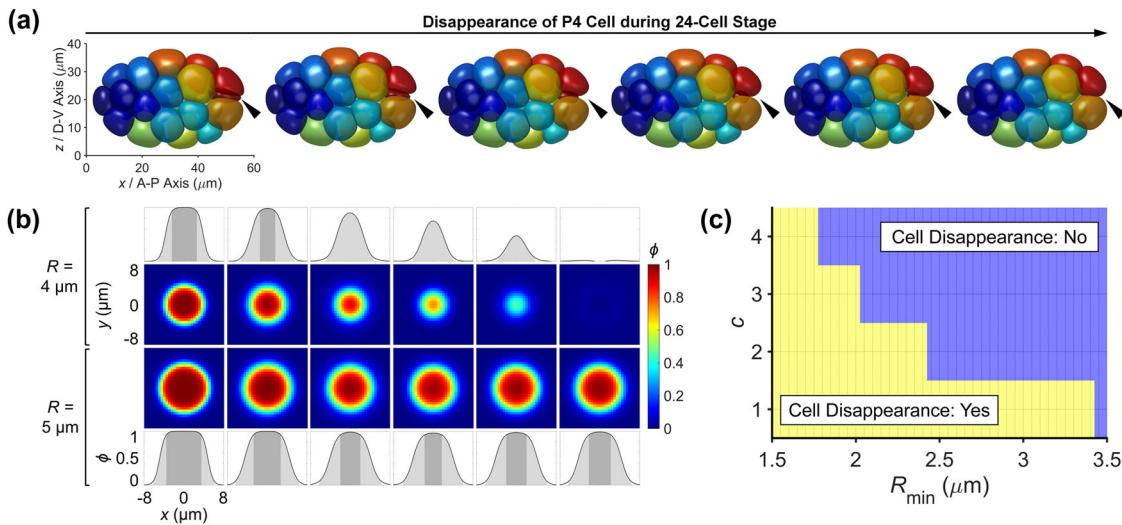
$$F_i^n = -\gamma c \left( 4(\phi_i^n)^3 - 6(\phi_i^n)^2 + 2\phi_i^n \right) - g_e \phi_i^n \phi_e^2 - g \phi_i^n \sum_{j \neq i}^N (\phi_j^n)^2 - \nabla \phi_i^n \cdot \sum_{j \neq i}^N \sigma_{ij} \nabla \phi_j^n + M \left( V_i(t) - \int_{\Omega} \phi_i^n d\mathbf{r} \right) |\nabla \phi_i^n|, \quad (9)$$

Next, we relax the restriction by the stabilization method. The main idea is to add an artificial stabilization term that has a dissipative effect to balance the instability caused by the explicit treatment of the nonlinear term<sup>58,59</sup>. Here, we first introduce a first-order stabilization term  $-S(\phi_i^{n+1} - \phi_i^n)$  to alleviate the strict constraint on temporal evolution:

$$\frac{\tau}{\delta t} (\phi_i^{n+1} - \phi_i^n) = \gamma \Delta \phi_i^{n+1} + F_i^n - S(\phi_i^{n+1} - \phi_i^n), \quad (10)$$



**Fig. 2 Framework improvement by the addition of stabilization term.** **a** Computing time compared to the baseline at different stages in the first-order scheme. **b** Computing time compared to the baseline at different stages in the second-order scheme. In both (a) and (b), the baseline is set as the computing time when  $\delta t = 0.3$  and no stabilization term is added. In b, the dashed line is the curve  $t'_c = 0.3 \times \delta t^{-1}$  and the inset is the same figure presented in a dual logarithmic coordinate system. **c** Embryonic morphologies from 1- to 8-cell stages in the second-order scheme with  $\delta t = 2.0$ . The cells' identities and corresponding colors are denoted on the right;  $\sigma_S$  and  $\sigma_W$  denote the relatively strong (i.e.,  $\sigma_S = 0.5$ ) and weak (i.e.,  $\sigma_W = 0$ ) cell-cell attraction inferred at the 4-cell stage respectively (Supplementary Fig. 2a, b and Supplementary Note 1).



**Fig. 3 Cell disappearance that happens when the cell size is too small.** **a** Gradual disappearance of the P4 cell during the 24-cell stage; the P4 cell is colored dark red and pointed by an arrowhead. **b** Phase field distribution in 1D ( $x$  axis) and 2D ( $xy$  plane) of a free cell with a radius of  $4 \mu\text{m}$  (upper two rows) and  $5 \mu\text{m}$  (lower two rows). The single-cell simulation is conducted with open boundaries and illustrated with an in silico interval of 20 from left to right. In the 1D distribution, the boundary of a cell ( $0.07 \leq \phi < 0.93$ ) is painted with light gray while the interior ( $\phi \geq 0.93$ ) and exterior ( $\phi < 0.07$ ) of a cell are painted with dark gray and white respectively. **c** A scanning on parameter  $c$  with different cell radii  $R_{\min}$ ; yellow and blue indicate if the cell disappears or not in a single-cell simulation.

where  $S$  is a positive coefficient proportional to the dissipative effect. Increasing  $S$  will make computation stabler but also introduce extra numerical error. Thus, it is necessary to keep a balance between stability and accuracy. Therefore, we optimize  $S$  by finding its minimal value that can stabilize a simulation with  $\delta t > 0.3$  while preserving enough accuracy, i.e., still reconstructing the morphogenetic dynamics observed experimentally (Fig. 1b). For each  $\delta t$  from 0.3 to 2.0 in a step of 0.1, we perform simulation from 1- to 8-cell stages to search for the optimal  $S$  value in a step of 0.1. Here, the time scale of each simulation is proportionally fitted to the one without stabilization term ( $\delta t = 0.3$ ; baseline), using the duration of 6- and 7-cell stages determined by their first quasi-steady states; then the in silico time for the 8-cell stage is calculated with the linear relationship given the base value set as 15,000. It's shown that the system bifurcates into another 8-cell topology when  $\delta t$  exceeds 1.8 (Supplementary Fig. 5a). Besides, with the increment of  $\delta t$ , the computing time compared to the baseline (i.e.,  $t_c'$ ) is not always monotonously decreasing (Fig. 2a and Supplementary Table 3), revealing that the in silico time  $n_t \delta t$  for the same stage/process also increases (Supplementary Fig. 5b and Supplementary Table 4). This overdamping-like effect, i.e., the increase of in silico time  $n_t \delta t$ , was also shown in previous research and may be caused by the numerical error of the stabilization term in the first-order scheme, which then limits the computing time reduction gained from a larger  $\delta t$ <sup>60</sup>.

To solve the problem above, we further introduce a second-order semi-implicit scheme with a stabilization term to achieve less numerical error<sup>41,61</sup>:

$$\frac{T}{2\delta t} (3\phi_i^{n+1} - 4\phi_i^n + \phi_i^{n-1}) = \gamma \Delta \phi_i^{n+1} + 2F_i^n - F_i^{n-1} - S(\phi_i^{n+1} - 2\phi_i^n + \phi_i^{n-1}). \quad (11)$$

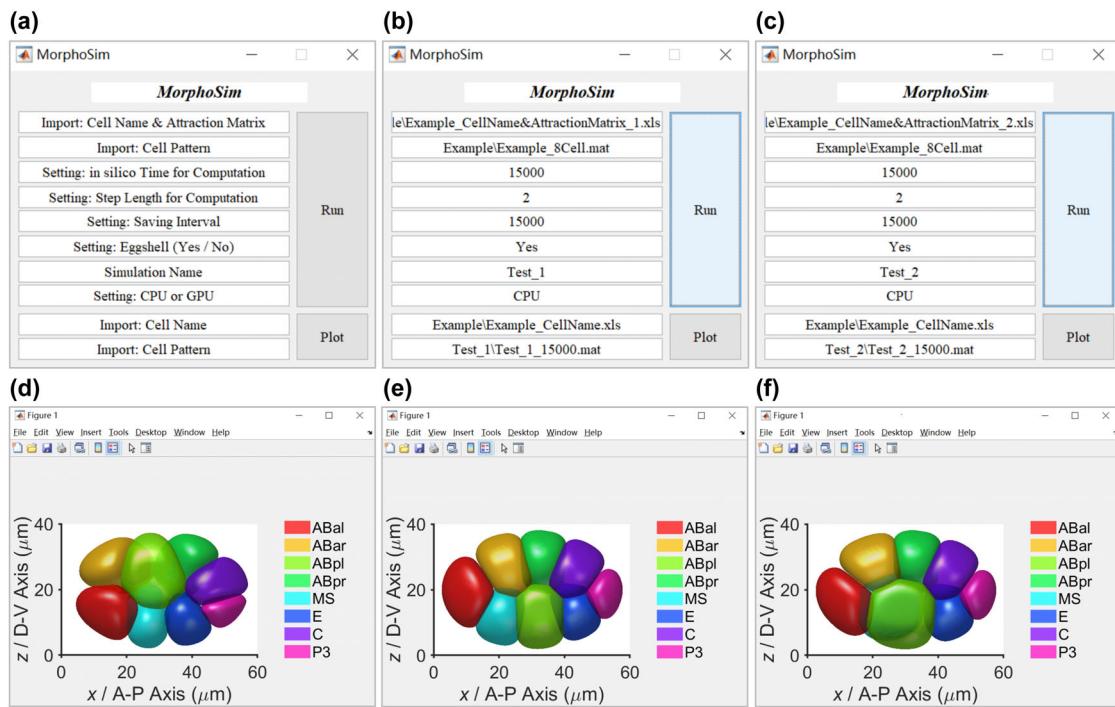
We extend the scanning range of  $\delta t$  to 0.3–2.5 and adopt the same optimization procedure for finding the optimal  $S$  value. The second-order scheme takes full advantage of the larger  $\delta t$  by avoiding the overdamping-like effect, therefore, substantially reducing the computing time and keeping the simulations with different  $\delta t$  values scalable (Fig. 2b and Supplementary Table 3). The perfect scalability, which faithfully follows the relationship  $t_c' = 0.3 \times \delta t^{-1}$  with a goodness of fit larger than 0.95 for all the stages, is suddenly broken when  $\delta t$  exceeds 2.0, probably due to

numerical error accumulation. Thus, we choose the time step size  $\delta t = 2.0$  with  $S = 12$  as the optimal condition for further simulation, which can still recapitulate the 1- to 8-cell morphogenesis of *C. elegans* embryo (Fig. 2c). Although the second-order scheme gains an edge over the first-order one in computing time, it's worth pointing out that the first-order scheme is better at maintaining numerical stability<sup>59</sup>. In other words, the first-order scheme allows larger values assigned to the parameters. Note that the  $\delta t$  and  $S$  values presented here are optimized to minimize the computing time, and one may select other values according to the actual problems or parameter settings.

#### A new formation of volume constriction to avoid “cell disappearance”

After the improvement of the numerical scheme, we then perform simulations for the later stages of *C. elegans* embryogenesis to find the maximum cell number afforded by the current framework. With the experimentally-measured cell division order and axis and volume segregation ratio inputted (Supplementary Table 1), the simulation proceeds properly until the 24-cell stage, when the whole phase field of the P4 cell,  $\phi_{P4}$ , shrinks to zero erroneously (Fig. 3a and Supplementary Movie 1). Hereafter, such a phenomenon caused by numerical error is referred to as “cell disappearance” in this work. In the simulations, cell disappearance is quantitatively defined when a cell's phase field is globally smaller than 0.5.

When cell disappearance occurs, the volume constriction (Eq. (2)) fails to prevent a cell's phase field from dropping to the homogeneous state  $\phi \equiv 0$ . Further simulation on a single cell with different sizes and without any external force reveals that a threshold of cell size exists to determine if cell disappearance happens (Fig. 3b). During the relaxation of a free cell, its interior always shrinks, while the width of the diffusing interface remains nearly constant. If the cell size is below the threshold ( $R \leq 4 \mu\text{m}$  or  $V \leq 268 \mu\text{m}^3$ ), the interior of the phase field shrinks to disappear and its boundaries gradually overlap, making the phase field eventually converge to  $\phi \equiv 0$ , namely, cell disappearance (Fig. 3b). As a small cell size appears in all kinds of biological processes in vivo, like blastomere cleavage and cell apoptosis, solving the



**Fig. 4** The graphical user interface of **MorphoSim**. **a** The interface and instruction of inputs required. **b, c** The simulation inputs for 8-cell *C. elegans* embryogenesis with relatively strong (i.e.,  $\sigma_{ABpl,E} = \sigma_S = 0.5$ ) and weak (i.e.,  $\sigma_{ABpl,E} = \sigma_W = 0.5$ ) adhesion in ABpl-E contact respectively. **(d)** The initial state (in silico time = 0) of the 8-cell embryo. **e, f** The final state (in silico time = 15,000) of the 8-cell embryo with relatively strong (i.e.,  $\sigma_{ABpl,E} = \sigma_S = 0.5$ ) and weak (i.e.,  $\sigma_{ABpl,E} = \sigma_W = 0.5$ ) adhesion in ABpl-E contact respectively.

problem of cell disappearance is vital for simulations of those scenes<sup>62,63</sup>.

The phenomenon of cell disappearance can be avoided in two ways. First, we replace the volume constriction (Eq. (2)) with a new formulation based on the relative error of volume, instead of the absolute error controlled in the previous one:

$$F_{vol} = M' \left( \frac{\int_{\Omega} \phi_i d\mathbf{r}}{V_i(t)} - 1 \right) \hat{n}. \quad (12)$$

where  $M'$  is the volume constriction strength. By limiting the relative error, the new formulation can achieve a more accurate simulated volume, consistent with the designated one measured experimentally or set arbitrarily. Apart, the stronger volume constriction for the cells with small size ( $\sim 10^2 \mu\text{m}^3$ ) lowers the cell size threshold of cell disappearance. The new formation also maintains numerical stability when handling the relatively larger cells ( $10^3 \sim 10^4 \mu\text{m}^3$ ), while strengthening volume constriction in the previous formation (i.e., by amplifying  $M$ ) would cause numerical instability easily. The second way to resist cell disappearance is to amplify the value of parameter  $c$ , the positive coefficient of double-well potential  $W(\phi)$ , so that the ability of a phase field to separate into two phases can be enhanced (Eq. (1)) (Fig. 3c). Here, we select  $c = 2$  and  $M' = 8$  (with a similar error level to the previous formulation at 4-cell stage) considering both numerical stability and accuracy for simulations of later stages, where the derived binarized relatively strong and weak cell–cell attraction is  $\sigma_S = 0.5$  and  $\sigma_W = 0$  respectively. The final iterative process with the stabilization term added and the volume constriction modified is detailed in Supplementary Note 2.

#### A Matlab-based GUI for automatic computation and structural illustration

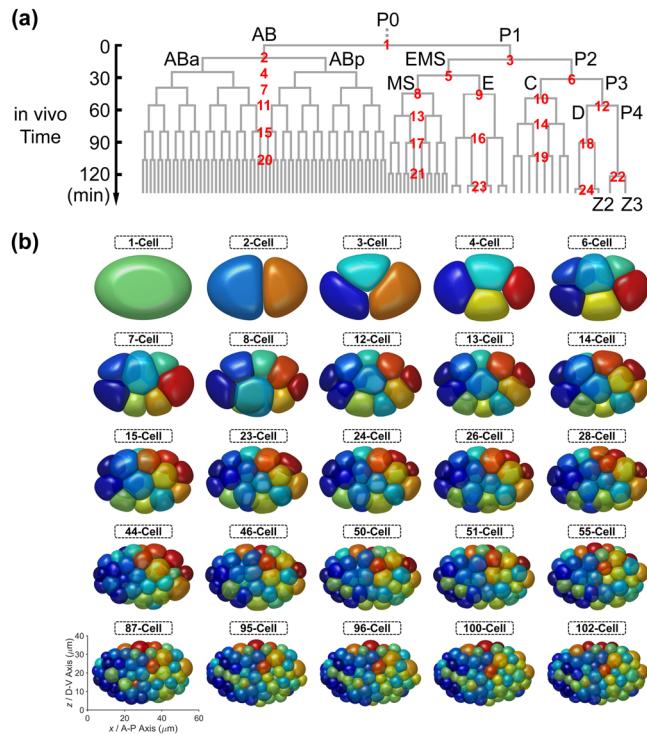
Given the phase field model and numerical methods, we achieve the final framework named *MorphoSim* (**M**orphology **S**imulator). We further pack it into an open-source graphical user interface

(GUI) using the software Matlab (Fig. 4a)<sup>64</sup>. One can input the binary distribution of cells, assign a cell–cell attraction matrix arbitrarily, and set up the in silico time, step length (i.e., time step size  $\delta t$ ), and saving interval for simulation. It should be pointed out that the eggshell boundary can be removed as an open-boundary condition and the computation can be performed on a CPU or GPU according to user requirements. When the simulation is over, the user can import the output file of a specific time point and plot the 3D structure automatically. The GUI and a detailed user guidebook are accessible at <https://github.com/XiangyuKuang/MorphoSim>.

During the 8-cell *C. elegans* embryogenesis, the relatively weak adhesion in ABpl-E contact was reported to be critical for the robust formation of 3D embryonic structure, which serves as a criterion to check if the framework preserves its precision (Fig. 2c)<sup>43</sup>. The inputs for 8-cell simulations with relatively strong (i.e.,  $\sigma_{ABpl,E} = \sigma_S = 0.5$ ) and weak (i.e.,  $\sigma_{ABpl,E} = \sigma_W = 0$ ) adhesion in ABpl-E contact are shown in Fig. 4b, c, and the initial state of phase fields is shown in Fig. 4d. The simulation results are in line with previous findings (Fig. 4e, f)<sup>43,56</sup>. Using a personal computer, this simulation lasts for less than 1.5 h on CPU (Intel(R) Core(TM) i5-10210U CPU @ 1.60 GHz 2.11 GHz) and less than 12.5 min on GPU (NVIDIA GeForce GTX 1060) for an in silico time = 15,000.

#### Simulation of the *C. elegans* embryogenesis from 1- to 102-cell stages

To test the availability of *MorphoSim*, we first perform the simulations of *C. elegans* embryo from 1- to over 100-cell stages according to the cell division order observed in vivo (Supplementary Table 1). For simplicity, here we consider the 6 founder cells (i.e., AB, MS, E, C, D, and P4) as well as their ancestors (i.e., P0, P1, P2, P3, and EMS) and progenies (Fig. 5a). The cell divisions in the same generation of a founder cell are pseudo-synchronous with a slight variation in reality and are idealized as a whole division group with the same cell cycle length in simulation. We use the



**Fig. 5 The 1- to 102-cell *C. elegans* embryogenesis. a** The cell lineage tree reconstituted from *in vivo* experiment. The cell divisions from the same founder cell (i.e., AB, MS, E, C, D, or P4) and in the same generation are regarded as synchronous; the cell cycle length of each division group is approximated with the shortest one among the cells obtained from experimental measurement<sup>65</sup>. The *in vivo* time axis is placed on the left and the cell division order is labeled near each cell division group. **b** The embryonic morphologies generated by phase-field simulation, using the cell division order and axis and volume segregation ratio measured experimentally as input. At each stage defined by the cell division order in (a), the *in silico* time (i.e., simulation time) from 1- to 8-cell stages is introduced in Supplementary Note 1, while the one after the 8-cell stage is set to reach the equilibrium ( $\bar{v} < 1 \times 10^{-4}$ ).

experimentally-measured shortest cell cycle length within the group as the common value<sup>65</sup>; finally, we obtain the cell division order formed by 24 independent division groups as shown in Fig. 5a. Besides, the inputted cell division axis and volume segregation ratio are acquired from<sup>48</sup> (Supplementary Table 1). The simulation from 1- to 8-cell stages is performed following the pipeline in Supplementary Note 1, while the one after the 8-cell stage is set to reach mechanical equilibrium (defined by  $\bar{v} < 1 \times 10^{-4}$ ) or last for a long enough duration (set as *in silico* time = 10,000) at each stage and then the next cell division(s) would be activated; after 8-cell stage, the cell–cell attraction matrix is still simplified as binary and follows the rule that the attraction between non-sister cells and between sister cells is relatively strong ( $\sigma_S = 0.5$ ) and weak ( $\sigma_W = 0$ ) respectively<sup>43</sup>. With the inputs above and the time step size  $\delta t = 1.5$ , the phase-field framework successfully simulates a multicellular system with up to 102 cells, without any cell disappearance (Fig. 5b and Supplementary Movie 2).

Next, we test its computational efficiency in GPU and CPU compared to the previous phase-field framework, in which the *in silico* time scale has been well fitted with the *in vivo* one<sup>43</sup>. Here, we perform the simulation from 6- to 8-cell stages with the previous and current frameworks, whose time step size  $\delta t$  is 0.1 and 2 respectively (Fig. 2b, c)<sup>43</sup>. The time scales between the two frameworks are proportionally fitted using the duration of 6- and 7-cell stages determined by their first quasi-steady states so that

the simulation process of each stage is equivalent in both frameworks and is comparable with the *in vivo* time scale. It shows that *MorphoSim* reaches a computational efficiency of more than  $10^3$  times the previous one averagely, where the computing time in GPU is around the 0.1- to 1-min levels (Table 1). Besides, for all the 6-, 7-, and 8-cell stages, the mean positional shift of the cells at each time step  $\delta t$  in *MorphoSim* simulation is at least one magnitude larger than the one in the previous framework (Supplementary Fig. 6). Further, we increase the cell number to 25, 50, and 100 and estimate the time cost for the system to evolve for an *in silico* time corresponding to 5 min in reality. For the original framework, the computation is out of memory in both GPU and CPU when the cell number reaches 100. The problem is well solved in *MorphoSim*, allowing full utilization of computational resources and numerical experiments coupled with parameter scanning (Table 1).

### Reconstruction of the assembly, self-repairing, and dissociation of the *synNotch* systems

The *synNotch* system is the state-of-the-art methodology to generate self-organizing multicellular living machines, which consist of multiple cell types (illustrated by different colors in Fig. 6) with genetically programmed differential adhesion<sup>16</sup>. The topology of cell aggregate is dependent on the combinatorial cell–cell adhesion strengths and can self-repair after cleavage and dissociate after eliminating the adhesive protein. Here, we employ *MorphoSim* to reproduce the stereotypic tricomponent and biocomponent topologies reported before, which can form spherically asymmetric and symmetric patterns corresponding to two different sets of cell–cell adhesion programs.

In the simulation of the *synNotch* systems, we add the Gaussian white noise  $\xi_i(t)$  onto each cell's motion to model the stochasticity in reality. Hence, the evolution equation turns into:

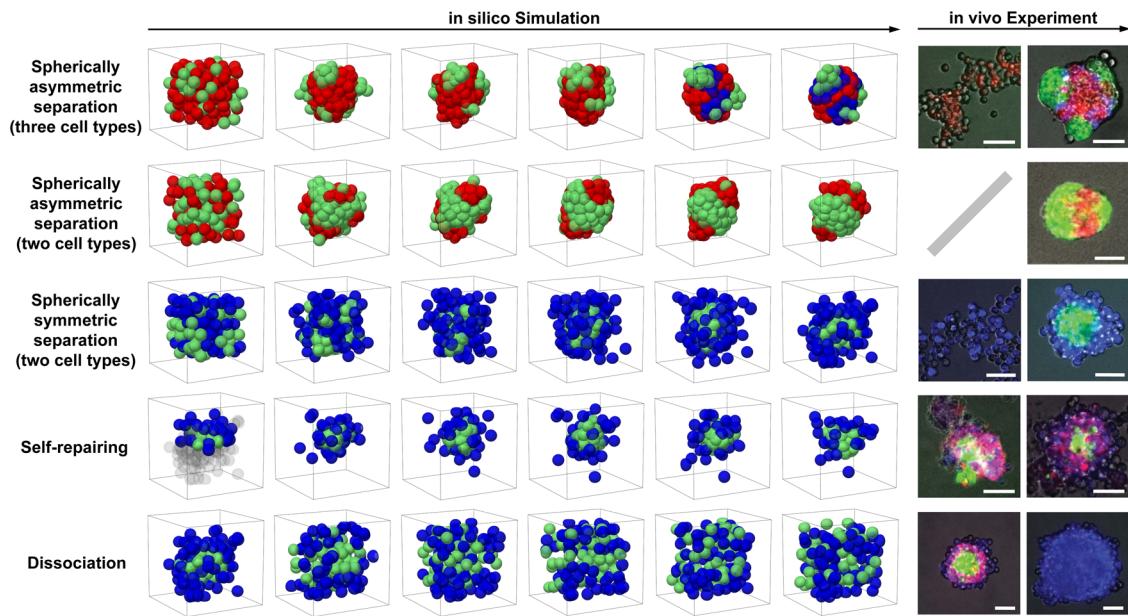
$$\frac{\partial \phi_i}{\partial t} = -\frac{1}{\tau} (\mathbf{F}_{\text{ten}} + \mathbf{F}_{\text{vol}} + \mathbf{F}_{\text{rep}} + \mathbf{F}_{\text{atr}} + \kappa \xi_i(t)) \cdot \nabla \phi_i, \quad (13)$$

where  $\kappa$  is the noise strength. For all the simulations, the computational domain is set as a  $128 \times 128 \times 128$  cubic grid with grid size  $\delta l = 0.5 \mu\text{m}$  and time step size  $\delta t = 1.25$ , and  $\kappa$  is assigned  $\frac{\sqrt{5}}{10}$ . It's worth noting that the spatial scale of the simulation can be adjusted to the experimental one by rescaling parameters. Here we provide a rescaled parameter setting with the spatial scale of simulations approximating the experimental scales of the *synNotch* system (Supplementary Table 5)<sup>16</sup>. To set up the initial state, we uniformly randomize the cell positions inside the cubic domain, while each cell is assigned a sphere with a radius  $R = 5 \mu\text{m}$ ; then we impose an intercellular repulsion  $\mathbf{F}_{ij}$  and a repulsion  $\mathbf{F}_{\text{boundary},i}$  between cell and boundary to eliminate the possible overlap between cellular regions:

$$\mathbf{F}_{ij} = \begin{cases} \mathbf{k}_1(\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j), & |\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j| < 2R \\ \mathbf{0}, & |\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j| \geq 2R \end{cases}, \quad (14)$$

$$\mathbf{F}_{\text{boundary},i} = \begin{cases} k_2(\mathbf{d}_{\text{boundary},i} - R), & |\mathbf{d}_{\text{boundary},i}| < R \\ \mathbf{0}, & |\mathbf{d}_{\text{boundary},i}| \geq R \end{cases}. \quad (15)$$

Here,  $\bar{\mathbf{r}}_i$  and  $\bar{\mathbf{r}}_j$  are the centroids of the  $i$ -th and  $j$ -th cells;  $\mathbf{d}_{\text{boundary},i} = (d_{i,x}, d_{i,y}, d_{i,z}) = \min(L - \bar{\mathbf{r}}_i, \bar{\mathbf{r}}_i)$  is the nearest distance between the centroid of the  $i$ -th cell and the cubic boundary in three orthogonal directions;  $L$  is the side length of the cubic domain;  $k_1 = 0.05$ ;  $k_2 = 0.1$ . The iteration  $\bar{\mathbf{r}}_i^{n+1} = \bar{\mathbf{r}}_i^n + \mathbf{F}_{\text{boundary},i} + \sum_{i \neq j} \mathbf{F}_{ij}$  continues until all the cells have a distance no less than  $2R$  from the others (roughly 200–300 time steps), then their distribution will be used as the initial condition for the *synNotch* simulation. The simulation scenarios along with their parameter settings and time scales are described in Table 2, where all the



**Fig. 6 In silico simulation (left panel) and in vivo experiment (right panel) of the synNotch systems.** In each row, the in silico structures are shown from time points 0 to 50,000 in a step of 10,000; the computational domain is outlined with a black cube. In the first row, since in silico time = 40,000 (the fifth column), a part of red cells, which contact at least one green cell, are painted blue considering that the red cells (Type 1) receive contact-dependent signaling from the green cells (Type 2) and undergo subsequent differentiation into the blue cells (Type 3), in accordance to the experimental description<sup>16</sup>. In both the fourth and fifth rows, the first in silico structure is adopted from the last one in the third row; for the former one, the cells with  $z < 0$  or without contact with the cell aggregate are removed and shown by gray shadow. For the in silico structures, the cell types and corresponding colors and value assignments on cell-cell adhesion are listed in Table 2. For the in vivo structures, the scale bars represent 100  $\mu\text{m}$  in reality; the fluorescence colors are in line with the ones used for the in silico structures, except that the red cells in the fourth and fifth rows also correspond to the blue ones in silico as they are the same in the cell-cell adhesion program. The images of in vivo experiment are from ref. <sup>16</sup>, reprinted with permission from AAAS.

simulations last for an in silico time = 50,000 and no cell division is executed. It is noteworthy that the adhesion within and between cell types is set once for all and only the simulation for the tricomponent spherically asymmetric separation incorporates cell differentiation, described by a cell type transition since in silico time = 40,000 based on contact-dependent cell-cell signaling.

The first condition is to program two types of cells, which have strong adhesion within either cell type and weak adhesion between the two cell types (Table 2). Since the cells of either cell type tend to be near each other, the system would separate into two major regions corresponding to the two cell types (second row in Fig. 6 and Supplementary Movie 3). If a cell in Type 1 is additionally programmed to differentiate into another fate (Type 3) after a specific time point (i.e., in silico time = 40,000) but maintains its original adhesion properties when contacting a cell in Type 2, a nested structure formed by three cell types appears as seen in vivo (first row in Fig. 6 and Supplementary Movie 4). The second condition is to program two types of cells as well, but the cells in Type 1 and Type 2 have strong and weak adhesion respectively (Table 2), which generates a layered structure in line with the experimental one (third row in Fig. 6 and Supplementary Movie 5). Furthermore, we adopt the final state of the layered structure and test if the self-repairing (by removal of half of the cells) and dissociation (by eliminating cell-cell adhesion) take place when the experimental conditions are mimicked in silico. As expected, both morphogenetic phenomena are successfully reproduced in simulation (fourth and fifth rows in Fig. 6 and Supplementary Movies 6, 7).

### Simulating and exploring more biological processes extensively with MorphoSim

As exemplified by simulations on the nematode embryogenesis and synNotch systems, MorphoSim is capable of modeling organic

and embryonic morphologies precisely. Despite that only the cell-cell attraction is studied as a variable, all physical parameters can be changed and adapted to specific biological processes. The key physical parameters and their biological significance are listed in Table 3, in which we explain how they should be customized for different biological scenes.

## DISCUSSION

Multicellular morphology is a fascinating topic and a long-term focus of biological research. Extending on our previous phase field model established with high-quality in vivo data<sup>43,48</sup>, in this work we developed an efficient and scalable framework, MorphoSim, for multicellular systems. We proposed a stabilized numerical scheme and new volume constriction to lower the model's computational cost for simulating a large number of cells simultaneously with considerable accuracy. MorphoSim provides an efficient and powerful tool that not only affords large-scale simulations, but also allows high-dimensional parameter scanning on GPU/CPU clusters. Moreover, the morphodynamics of the tricomponent and biocomponent synNotch systems can be reproduced by MorphoSim in approximately half a day, demonstrating the framework's validity, efficiency, and applicability.

In synthetic biology, more and more attentions are being paid to cell-based morphology synthesis including organoid, embryo, bio-robot, etc. Taking the synNotch system as an example, they are genetically programmed with differential cell-cell adhesion to achieve different modes of structure and function. However, it still faces some challenges: e.g., (1) How do biological parameters like cell stiffness and surface tension affect the dynamics? (Supplementary Fig. 7a-c and Supplementary Table 6); and (2) How to choose the parameter combination to optimize a specific function? It would be very helpful if the parameter space can be efficiently explored<sup>66</sup>. MorphoSim can regenerate the synNotch

**Table 2.** Parameter settings and time scales of the *synNotch* simulations.

Topological dynamic	Stage	Cell type 1 <sup>a</sup>	Cell type 2 <sup>a</sup>	Cell type 3 <sup>a</sup>	In silico time	Computing time (h)	Value assignment on cell-cell adhesion
Spherically asymmetric separation (three cell types) <sup>b</sup>	Start	80 (R)	40 (G)	0 (B)	50,000	12.2	$\sigma_{1,1} = 0.9$
	End	40 (R)	40 (G)	40 (B)			$\sigma_{1,2} = 0.5$
Spherically asymmetric separation (two cell types)		62 (R)	62 (G)	/	50,000	11.9	$\sigma_{2,2} = 0.9$
Spherically symmetric separation	From start to end	50 (G)	75 (B)	/	50,000	13.3	$\sigma_{1,1} = 0.9$
Self-repairing		24 (G)	32 (B)	/	50,000	4.4	$\sigma_{1,2} = 0.5$
Dissociation		50 (G)	75 (B)	/	50,000	12.1	$\sigma_{2,2} = 0.3$

<sup>a</sup>The painting color of each cell type in Fig. 6 is labeled in the columns “Cell type 1”, “Cell type 2”, and “Cell type 3”, where “R”, “G”, and “B” represent red, green, and blue, respectively.

<sup>b</sup>For the “Spherically asymmetric separation (three cell types)”, “Cell type 3” is defined as the cell that originally belongs to “Cell type 1” but contacts at least one cell in “Cell type 2” since in silico time = 40,000, so the mechanical properties of “Cell type 3” is same to the ones of “Cell type 1”.

**Table 3.** The key physical parameters and their biological significance.

Physical parameter	Biological significance	Remark
$\gamma$	Cell surface tension	The cell surface tension is mainly attributed to cortical contractility <sup>72</sup> . As cell surface tension minimizes the cell surface area and pushes a cell to be spherical instead of being deformed, it also characterizes the effect of cell stiffness (hard or soft) if the cell’s natural shape is spherical (Supplementary Fig. 7a–c and Supplementary Table S6). In reality, the cell stiffness is contributed by the cytoskeleton and nucleus and varies from cell type to cell type. For instance, the tumor cells are usually softer than their coexisting benign counterparts <sup>42,73,74</sup> .
$\phi_e$	Eggshell/boundary	The size and shape of the eggshell or boundary can be customized according to the system of interest (e.g., ellipsoidal or cylindroid for <i>C. elegans</i> embryo and boundless for <i>synNotch</i> system) <sup>16,52,53</sup> .
$\sigma$	Cell–cell attraction	The attraction between cells often comes from cell–cell adhesion (e.g., cadherin family), gap junction (e.g., connexin family), and ligand–receptor binding <sup>75–77</sup> . Asymmetric intercellular attraction ( $\sigma_{i,j} \neq \sigma_{j,i}$ ) could be employed to model the biological phenomena in which one cell is unidirectionally attracted by another, such as the chemotactic cell movement induced by cell–cell signaling and engulfment during apoptosis <sup>78–80</sup> .
$V_i(t)$	Cell volume	The cell volume can be time-dependent according to specific situations. For example, it decreases during cell apoptosis and increases during cell growth <sup>63,81</sup> . The cell volume can also be affected by extrinsic factors like osmotic pressure and substrate stiffness <sup>82,83</sup> .
$\tau$	Ambient viscosity	The viscosity is dependent on the real scenario <sup>84</sup> . When the environment disobeys the overdamping condition, one can rewrite Eq. (5) accordingly <sup>85,86</sup> .
$\kappa$	Stochastic noise	The stochasticity in cell movement can be raised by intrinsic (e.g., thermal effect and fluctuation) and extrinsic (e.g., stirring during incubation of the <i>synNotch</i> spheroid) reasons <sup>16,87,88</sup> .

dynamics quickly and large-scale simulations can also be carried out on GPU and CPU clusters for both mechanistic studies and parameter optimization. Note that the *MorphoSim* parameters (e.g., adhesion strength, motional noise, cell number, spatial resolution, and temporal resolution) should be readjusted to the real experimental condition for better simulation performance.

The proposed phase-field framework is also practical to model natural systems such as a developing embryo with large cell numbers. By comparing the in silico and in vivo morphologies, the low-cost parameter scanning permits the inference of a multicellular system’s mechanical state, which is hard to measure directly or infer by the previous phase field model<sup>43</sup>. As the phase field model describes a cell on a dense mesh, more space-related biochemical and biophysical details (e.g., polarity and cytoskeleton) can be added to reconstruct the real system comprehensively.

Despite that *MorphoSim* can simulate over one hundred interacting cells, there is still a great need from the field to keep increasing the cell number and decreasing the computational cost. For example, the early *Drosophila* embryogenesis involves thousands of cells and was usually simulated using the

coarse-grained model or vertex model<sup>67,68</sup>. On the one hand, the limited GPU memory hampers mesh enlargement/densification when considering a large number of cells. Algorithms like parallel computation<sup>69</sup>, adaptive mesh refinement<sup>70</sup>, and moving mesh method<sup>71</sup>, can be employed to reduce the computational cost and improve efficiency.

In addition to the algorithmic improvements above, the *MorphoSim* framework can be further developed based on the multiscale method. Instead of using a phase field variable to represent a single cell, we may use a macroscopic phase field variable to represent a group of cells that have the same biophysical properties and microscopic phase field variables to mimic the cells with distinct properties, which could greatly reduce the computational cost and allow for simulations with a much larger number of cells. Besides, for an in vivo system with known morphological information (e.g., *C. elegans* embryo)<sup>48</sup>, it is feasible to only simulate a special group of cells or the regions of interest instead of the entire system while assigning the cell morphology measured experimentally onto the other cells as a boundary constraint.

## DATA AVAILABILITY

All relevant data are within the paper and its Supplementary Information files.

## CODE AVAILABILITY

The scripts of *MorphoSim* and its Matlab-based GUI are deposited at <https://github.com/XiangyuKuang/MorphoSim>. In addition to *MorphoSim*, all other source codes used in this paper are available on reasonable request.

Received: 16 April 2022; Accepted: 4 January 2023;

Published online: 17 February 2023

## REFERENCES

- Heisenberg, C. P. & Bellaïche, Y. Forces in tissue morphogenesis and patterning. *Cell* **153**, 948–962 (2013).
- Qi, J. et al. Mechanical regulation of organ asymmetry in leaves. *Nat. Plants* **3**, 724–733 (2017).
- Riquelme, M. et al. Fungal morphogenesis, from the polarized growth of hyphae to complex reproduction and infection structures. *Microbiol. Mol. Biol. R.* **82**, e00068–17 (2018).
- Brunet, T. et al. Light-regulated collective contractility in a multicellular choanoflagellate. *Science* **366**, 326–334 (2019).
- Bryant, D. M. & Mostov, K. E. From cells to organs: building polarized tissue. *Nat. Rev. Mol. Cell Biol.* **9**, 887–901 (2008).
- Palovaara, J., Zeeuw, T. D. & Weijers, D. Tissue and organ initiation in the plant embryo: a first time for everything. *Annu. Rev. Cell Dev. Biol.* **32**, 47–75 (2016).
- Potten, C. S. Regeneration in epithelial proliferative units as exemplified by small intestinal crypts. *Ciba Found. Symp.* **160**, 54–71 (1991).
- Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274–284 (2013).
- Nelson, C. M. The mechanics of crypt morphogenesis. *Nat. Cell Biol.* **23**, 678–679 (2021).
- Roca-Cusachs, P., Conte, V. & Trepat, X. Quantifying forces in cell biology. *Nat. Cell Biol.* **19**, 742–751 (2017).
- Dance, A. The secret forces that squeeze and pull life into shape. *Nature* **589**, 186–188 (2021).
- Jonietz, E. Mechanics: the forces of cancer. *Nature* **491**, S56–S57 (2012).
- Kim, J., Koo, B. K. & Knoblich, J. A. Human organoids: model systems for human biology and medicine. *Nat. Rev. Mol. Cell Biol.* **21**, 571–584 (2020).
- Moris, N. et al. An in vitro model of early anteroposterior organization during human development. *Nature* **528**, 410–415 (2020).
- Rosner, M., Reithofer, M., Fink, D. & Hengstschläger, M. Human embryo models and drug discovery. *Int. J. Mol. Sci.* **22**, 637 (2021).
- Toda, S., Blauch, L. R., Tang, S. K. Y., Morsut, L. & Lim, W. A. Programming self-organizing multicellular structures with synthetic cell-cell signaling. *Science* **361**, 156–162 (2018).
- Toda, S., Brunger, J. M. & Lim, W. A. Synthetic development: learning to program multicellular self-organization. *Curr. Opin. Syst. Biol.* **14**, 41–49 (2019).
- Kriegman, S., Blackiston, D., Levin, M. & Bongard, J. A scalable pipeline for designing reconfigurable organisms. *Proc. Natl. Acad. Sci. USA* **117**, 1853–1859 (2020).
- Blackiston, D. et al. A cellular platform for the development of synthetic living machines. *Sci. Robot.* **6**, eabf1571 (2021).
- Shinbrot, T., Chun, Y., Caicedo-Carvajal, C. & Foty, R. Cellular morphogenesis in silico. *Biophys. J.* **97**, 958–967 (2009).
- Boekhorst, V. T., Preziosi, L. & Friedl, P. Plasticity of cell migration in vivo and in silico. *Annu. Rev. Cell Dev. Biol.* **32**, 491–526 (2016).
- Deutsch, A., Nava-Sedeño, J. M., Syga, S. & Hatzikirou, H. BIO-LGCA: a cellular automaton modelling class for analysing collective cell migration. *PLoS Comput. Biol.* **17**, e1009066 (2021).
- Turner, S. & Sherratt, J. A. Intercellular adhesion and cancer invasion: a discrete simulation using the extended Potts model. *J. Theor. Biol.* **216**, 85–100 (2002).
- Nielsen, B. F., Nissen, S. B., Sneppen, K., Mathiesen, J. & Trusina, A. Model to link cell shape and polarity with organogenesis. *iScience* **23**, 100830 (2020).
- Bock, M., Tyagi, A. K., Kreft, J. U. & Alt, W. Generalized Voronoi tessellation as a model of two-dimensional cell tissue dynamics. *Bull. Math. Biol.* **72**, 1696–1731 (2010).
- Alt, S., Ganguly, P. & Salbreux, G. Vertex models: from cell mechanics to tissue morphogenesis. *Philos. Trans. R. Soc. B* **372**, 20150520 (2016).
- Kajita, A., Yamamura, M. & Kohara, Y. Computer simulation of the cellular arrangement using physical model in early cleavage of the nematode *Cae-norhabditis elegans*. *Bioinformatics* **19**, 704–716 (2003).
- An, Y. et al. Apical constriction is driven by a pulsatile apical myosin network in delaminating *Drosophila* neuroblasts. *Development* **144**, 2153–2164 (2017).
- Nonomura, M. Study on multicellular systems using a phase field model. *PLoS ONE* **7**, e33501 (2012).
- Löber, J., Ziebert, F. & Aranson, I. S. Collisions of deformable cells lead to collective migration. *Sci. Rep.* **5**, 9172 (2015).
- Osborne, J. M., Fletcher, A. G., Pitt-Francis, J. M., Maini, P. K. & Gavaghan, D. J. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. *PLoS Comput. Biol.* **13**, e1005387 (2017).
- Camley, B. A. et al. Polarity mechanisms such as contact inhibition of locomotion regulate persistent rotational motion of mammalian cells on micropatterns. *Proc. Natl. Acad. Sci. USA* **111**, 14770–14775 (2014).
- Palmeri, B., Bresler, Y., Wirtz, D. & Grant, M. Multiple scale model for cell migration in monolayers: elastic mismatch between cells enhances motility. *Sci. Rep.* **5**, 11745 (2015).
- Kulawiak, D. A. & Camley, B. A. & Rappel, W.-J. Modeling contact inhibition of locomotion of colliding cells migrating on micropatterned substrates. *PLoS Comput. Biol.* **12**, e1005239 (2016).
- Mueller, R., Yeomans, J. M. & Doostmohammadi, A. Emergence of active nematic behavior in monolayers of isotropic cells. *Phys. Rev. Lett.* **122**, 048004 (2019).
- Zhang, G., Mueller, R., Doostmohammadi, A. & Yeomans, J. M. Active inter-cellular forces in collective cell motility. *J. R. Soc. Interface* **17**, 20200312 (2020).
- Wenzel, D. & Voigt, A. Multiphase field models for collective cell migration. *Phys. Rev. E* **104**, 054410 (2021).
- Moure, A. & Gomez, H. Phase-field modeling of individual and collective cell migration. *Arch. Comput. Methods Eng.* **28**, 311–344 (2021).
- Moure, A. & Gomez, H. Phase-field model of cellular migration: three-dimensional simulations in fibrous networks. *Comput. Methods Appl. Mech. Eng.* **320**, 162–197 (2017).
- Ghabache, E. et al. Coupling traction force patterns and actomyosin wave dynamics reveals mechanics of cell motion. *Mol. Syst. Biol.* **17**, e10505 (2021).
- Kim, J. Phase-field models for multi-component fluid flows. *Commun. Comput. Phys.* **12**, 613–661 (2012).
- Tao, K. et al. Tuning cell motility via cell tension with a mechanochemical cell migration model. *Biophys. J.* **118**, 2894–2904 (2020).
- Kuang, X. et al. Computable early *Caenorhabditis elegans* embryo with a phase field model. *PLoS Comput. Biol.* **18**, e1009755 (2022).
- Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
- Jiang, J., Garikipati, K. & Rudraraju, S. A diffuse interface framework for modeling the evolution of multi-cell aggregates as a soft packing problem driven by the growth and division of cells. *Bull. Math. Biol.* **81**, 3282–3300 (2019).
- Seirin-Lee, S., Yamamoto, K. & Kimura, A. The extra-embryonic space and the local contour are crucial geometric constraints regulating cell arrangement. *Development* **149**, dev200401 (2022).
- Yamamoto, K. & Kimura, A. An asymmetric attraction model for the diversity and robustness of cell arrangement in nematodes. *Development* **144**, 4437–4449 (2017).
- Cao, J. et al. Establishment of a morphological atlas of the *Caenorhabditis elegans* embryo using deep-learning-based 4D segmentation. *Nat. Commun.* **11**, 6254 (2020).
- Kajita, A., Yamamura, M. & Kohara, Y. Physical modeling of the cellular arrangement in *C. elegans* early embryo: effect of rounding and stiffening of the cells. *Genome Inf.* **13**, 224–232 (2002).
- Fickentscher, R., Strutz, P. & Weiss, M. Mechanical cues in the early embryogenesis of *Caenorhabditis elegans*. *Biophys. J.* **105**, 1805–1811 (2013).
- Fickentscher, R., Strutz, P. & Weiss, M. Setting the clock for fail-safe early embryogenesis. *Phys. Rev. Lett.* **117**, 188101 (2016).
- Tian, B., Guan, G., Tang, L. H. & Tang, C. Why and how the nematode's early embryogenesis can be precise and robust: a mechanical perspective. *Phys. Biol.* **17**, 026001 (2022).
- Guan, G., Tang, L. H. & Tang, C. Reconstructing the multicellular structure of a developing metazoan embryo with repulsion-attraction model and cell-cell connection atlas in vivo. *J. Phys.* **1592**, 012020 (2020).
- Shao, D., Rappel, W.-J. & Levine, H. Computational model for cell morphodynamics. *Phys. Rev. Lett.* **105**, 108104 (2010).
- Akiyama, M., Nonomura, M., Tero, A. & Kobayashi, R. Numerical study on spindle positioning using phase field method. *Phys. Biol.* **16**, 016005 (2018).
- Dutta, P., Odedra, D. & Pohl, C. Planar asymmetries in the *C. elegans* embryo emerge by differential retention of aPARs at cell-cell contacts. *Front. Cell Dev. Biol.* **7**, 209 (2019).
- Chen, L. Q. & Shen, J. Applications of semi-implicit Fourier-spectral method to phase field equations. *Commun. Comput. Phys.* **108**, 147–158 (1998).
- Zhu, J., Chen, L. Q., Shen, J. & Tikare, V. Coarsening kinetics from a variable-mobility Cahn-Hilliard equation: application of a semi-implicit Fourier spectral method. *Phys. Rev. E* **60**, 3564–3572 (1999).

59. Shen, J., Xu, J. & Yang, J. A new class of efficient and robust energy stable schemes for gradient flows. *SIAM Rev.* **61**, 474–506 (2019).
60. Yang, X. Error analysis of stabilized semi-implicit method of Allen-Cahn equation. *Discret. Contin. Dyn. Syst. Ser. B* **11**, 1057–1070 (2009).
61. Shen, J., Yang, X. & Yu, H. Efficient energy stable numerical schemes for a phase field moving contact line model. *J. Comput. Phys.* **284**, 617–630 (2015).
62. Suzuki, A. S. & Miki, K. Cellular basis of neuralization of induced neurectoderm in amphibian embryogenesis: changes of cell shape, cell size, and cytodifferentiation of the neurectoderm after neural induction. *Dev. Growth Differ.* **25**, 289–297 (1983).
63. Bortner, C. D. & Cidlowski, J. A. Uncoupling cell shrinkage from apoptosis reveals that Na<sup>+</sup> influx is required for volume loss during programmed cell death. *J. Biol. Chem.* **278**, 39176–39184 (2003).
64. Matlab: version 9.5.0.944444 (R2018b). Natick, Massachusetts: The MathWorks Inc. (2018).
65. Guan, G. et al. System-level quantification and phenotyping of early embryonic morphogenesis of *Caenorhabditis elegans*. Preprint at <https://doi.org/10.1101/776062> (2019).
66. Ma, W., Trusina, A., El-Samad, H., Lim, W. A. & Tang, C. Defining network topologies that can achieve biochemical adaptation. *Cell* **138**, 760–773 (2009).
67. Wu, X., Kong, K., Xiao, W. & Liu, F. Attractive internuclear force drives the collective behavior of nuclear arrays in *Drosophila* embryos. *PLoS Comput. Biol.* **17**, e1009605 (2021).
68. Spahn, P. & Reuter, R. A vertex model of *Drosophila* ventral furrow formation. *PLoS ONE* **8**, e75051 (2013).
69. Hötzter, J. et al. The parallel multi-physics phase-field framework Pace3D. *J. Comput. Sci.* **26**, 1–12 (2018).
70. Guo, Z. & Xiong, S. M. On solving the 3-D phase field equations by employing a parallel-adaptive mesh refinement (Para-AMR) algorithm. *Comput. Phys. Commun.* **190**, 89–97 (2015).
71. Shen, J. & Yang, X. An efficient moving mesh spectral method for the phase-field model of two-phase flows. *J. Comput. Phys.* **228**, 2978–2992 (2009).
72. Fagotto, F. The cellular basis of tissue separation. *Development* **141**, 3303–3318 (2014).
73. Luo, Q., Kuang, D., Zhang, B. & Song, G. Cell stiffness determined by atomic force microscopy and its correlation with cell motility. *Biochim. Biophys. Acta Gen. Subj.* **1860**, 1953–1960 (2016).
74. Rheinlaender, J. et al. Cortical cell stiffness is independent of substrate mechanics. *Nat. Mater.* **19**, 1019–1025 (2020).
75. Roy, F. V. & Berx, G. The cell-cell adhesion molecule E-cadherin. *Cell Mol. Life Sci.* **65**, 3756–3788 (2008).
76. Elias, L. A. B., Wang, D. D. & Kriegstein, A. R. Gap junction adhesion is necessary for radial migration in the neocortex. *Nature* **448**, 901–907 (2007).
77. Coombs, D., Dembo, M., Wofsy, C. & Goldstein, B. Equilibrium thermodynamics of cell-cell adhesion mediated by multiple ligand-receptor pairs. *Biophys. J.* **86**, 1408–1423 (2004).
78. Pani, A. M. & Goldstein, B. Direct visualization of a native Wnt in vivo reveals that a long-range Wnt gradient forms by extracellular dispersal. *eLife* **7**, e38325 (2018).
79. Humphries, A. C. & Mlodzik, M. From instruction to output: Wnt/PCP signaling in development and cancer. *Curr. Opin. Cell Biol.* **51**, 110–116 (2018).
80. Reddien, P. W. & Horwitz, H. R. The engulfment process of programmed cell death in *Caenorhabditis elegans*. *Annu. Rev. Cell Dev. Biol.* **20**, 193–221 (2004).
81. Bryan, A. K., Goranov, A., Amon, A. & Manalis, S. R. Measurement of mass, density, and volume during the cell cycle of yeast. *Proc. Natl Acad. Sci. USA* **107**, 999–1004 (2010).
82. Guo, M. et al. Cell volume change through water efflux impacts cell stiffness and stem cell fate. *Proc. Natl Acad. Sci. USA* **114**, E8618–E8627 (2017).
83. Gonzalez, N. P. et al. Cell tension and mechanical regulation of cell volume. *Mol. Biol. Cell* **29**, 2591–2600 (2018).
84. Gonzalez-Molina, J. et al. Extracellular fluid viscosity enhances liver cancer cell mechanosensing and migration. *Biomaterials* **177**, 113–124 (2018).
85. Brückner, D. B., Ronceray, P. & Broedersz, C. P. Inferring the dynamics of under-damped stochastic systems. *Phys. Rev. Lett.* **125**, 058103 (2020).
86. Brückner, D. B. et al. Learning the dynamics of cell-cell interactions in confined cell migration. *Proc. Natl Acad. Sci. USA* **118**, e2016602118 (2021).
87. Peskin, C. S., Odell, G. M. & Oster, G. F. Cellular motions and thermal fluctuations: the Brownian ratchet. *Biophys. J.* **65**, 316–324 (1993).
88. Kießling, T. R., Stange, R., Käs, J. A. & Fritsch, A. W. Thermorheology of living cells – impact of temperature variations on cell mechanics. *N. J. Phys.* **15**, 045026 (2013).

## ACKNOWLEDGEMENTS

We thank Prof. Lei-Han Tang, Prof. Zhongying Zhao, Prof. Xiaojing Yang, and Dr. Zhen Xu for their helpful discussions and comments. This work was supported by the National Key R&D Program of China (Grant No. 2021YFF1200500) and the National Natural Science Foundation of China (Grant Nos. 12225102, 12050002, 12226316, 12090053, and 32088101).

## AUTHOR CONTRIBUTIONS

X.K. and G.G. designed and implemented the *MorphoSim* framework, performed simulation, and data analysis, and wrote the paper and are considered to be the co-first authors. L.Z. and C.T. conceived and coordinated the study, provided theoretical guidance, and revised the paper.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41540-023-00265-w>.

**Correspondence** and requests for materials should be addressed to Chao Tang or Lei Zhang.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023