Reconstructing the multicellular structure of a developing metazoan embryo with repulsion-attraction model and cell-cell connection atlas *in vivo*

Reconstructing the multicellular structure of a developing metazoan embryo with repulsion-attraction model and cell-cell connection atlas *in vivo*

Guoye Guan[1, 5], Lei-Han Tang[2] and Chao Tang[1, 3, 4]

[1]Center for Quantitative Biology, Peking University, Beijing, 100871, China
[2]Department of Physics, Hong Kong Baptist University, Hong Kong, 999077, China
[3]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, 100871, China
[4]School of Physics, Peking University, Beijing, 100871, China

[5]Email: guanguoye@gmail.com

Abstract. Embryogenesis is a spatio-temporal multicellular evolutional process involved with intracellular biochemical activities and intercellular biophysical interactions. Reproducible and precise multicellular structures contribute to robustness of embryonic development by cell-cell communication, morphogenesis and other significant biological events. Using *Caenorhabditis elegans* as animal model, recently several researches established mechanical models to reconstruct the multicellular structures of this developmental system, in which cells interact via repulsive or attractive potentials inside an ellipsoidal eggshell. However, those models ignored some practical details and lack of test in depth. In this paper, we improved an *in silico* modeling framework based on previous models by revising formulae of interactive force and applying *in vivo* experimental information of eggshell shape, cell volume, cell position and cell-cell connection relationship. Cell pairs with and without empirically repeated connection were regarded to have different types of attractive force, which could help stabilize cells into their experimentally observed locations accompanied by correct neighbour relationships. Both previous models and our revised ones were tested, verified and compared to each other. Our modeling framework not only reproduces the multicellular structure patterns in an artificially compressed embryo with ~50 cells, but also exhibits a potential to uncover active adjustments and controls on cell positioning.

1. Introduction

Metazoan embryogenesis consists of rapid proliferation, fate specification and cell migration, which is genetically programmed to achieve high reproducibility among individuals as well as robustness against perturbation [1, 2]. *Caenorhabditis elegans* (i.e. *C. elegans*) is well-known for its invariant cell lineage and stereotypic cellular behavior during development, including division timing, migration trajectory and fate determination, thus, this model organism has been widely used for research in developmental biology [2, 3]. As *C. elegans* is a kind of eutelic animal with a constant number of nonidentical somatic cells when reaching maturity, while each cell has specific behavior and function, precise cell-arrangement pattern (hereafter referred to as multicellular structure) is highly critical to a lot of delicate developmental processes such as signaling transduction (e.g. fate specification, spindle formation) [4-7] and morphogenesis (e.g. gastrulation and ventral enclosure) [8, 9].

2. Limitations of previous models

To elucidate the underlying mechanisms giving rise to the robust and specific multicellular structures during *C. elegans* embryogenesis, which are basically comprised of both 3D cell positions and cell-cell connection relationships (i.e. direct physical contacts), several studies have laid a foundation by physical modeling, computational simulation and theoretical analysis [10-14]. Kajita et al (2002, 2003) firstly designed a computational model to describe the dynamic process of cell elongation, cell division, cell migration and cell deformation during 1- to 4-cell stage [10-11]. The model used a group of interactive particles to represent the cell membrane, while the cell cleavage was driven by repulsion of two centrosomes inside and invagination of contractile ring. This comprehensive model accurately accounted for the characteristic diamond-shape 4-cell arrangement pattern and explained the key conditions required. However, it considered a lot of details of reality by using dozens of movable objects for each cell and more than 20 system parameters for simulation, leading to low conciseness and high computational cost, in particular at late developmental stage which has a much larger amount of cells. Fickentscher et al (2013, 2016) proposed a limiting-component model by using only one mass point to represent a cell, in which cells interact with only repulsive forces inside an ellipsoid eggshell and divide according to a clock controlled by cell volume (hereafter referred to as 'Repulsion Model') [12, 13]. This simplified model allowed simulation of multicellular system with low computational cost, which successfully reproduced some of the multicellular structures and cell trajectories before 24-cell stage. Based on Fickentscher et al's model, Yamamoto et al (2017) took asymmetric intercellular attraction into account and investigated the diverse possible structures of 4-cell stage by rescaling the shape of eggshell (hereafter referred to as 'Attraction Model') [14], uncovering the biophysical roles of flatness of eggshell to instruct cell packing pattern. Considering requirement to simulate a system with dozens to hundreds of cells, we adopted the coarse-grained methods [12-14]. Although these three works indeed shed light on the mechanical cues coordinating progression of multicellular structure, a few limitations and potential problems still exist, mainly as the following:

1) The eggshell was usually considered as perfect ellipsoid with rotational symmetry along the major axis, but *C. elegans* embryo can actually survive under severe mechanical deformation by slide mounting (compression ratio ≈ 0.5) [15] and exhibits abnormal cell movements which are notably differentiated from the ones in naturally uncompressed embryo [16]. Besides, the shape of compressed eggshell is far away from a rescaled ellipsoid [15].
2) The parameters in models were idealized and assigned arbitrarily. For example, volume segregation ratio was set to be 2:3 for all the asymmetric cell divisions, meanwhile, ratio of major and minor axis was set to be 5:3, which ignored the real situation in a live embryo.
3) The simulation results were not quantitatively compared with experiments for each cell. This may cover the detailed difference between simulation and experiment such as cell location and connection relationship between cells.
4) The models haven't been tested and applied for developmental stages after gastrulation onset (i.e. 26-cell stage), whereas a *C. elegans* embryo would reach over five hundred cells before becoming a larva. A general *in silico* modeling framework is in need to help dissect and decompose the embryogenesis procedure which starts from one cell and ends in hundreds.

3. Materials and methods

To overcome the shortcomings above and provide a universal method, here we first tested the previous models with data obtained from *in vivo* experiments, then optimized the interactive force formulae and system parameters, and further proposed an applicable framework in which the multicellular structures can be precisely reconstructed up to at least ~50-cell stage, by using repulsive and attractive force between cells and an empirical cell-cell connection atlas.

3.1. Acquisition of experimental data

Quantitative parameters of the real embryo system including eggshell shape, cell volume, cell position and cell-cell connection relationship were collected from our previous work and applied to further

computation (Figure 1) [15]. The data was originally obtained from 3-dimensional time-lapse *in vivo* imaging experiments on *C. elegans* embryos from 4- to ~350-cell stage, with temporal resolution of approximate 1.5 minutes. The embryo had two distinguishable fluorescent markers on its nuclei and membrane respectively, which enabled us to reconstruct the morphological information of each cell. We selected the multicellular structures at five developmental stages (4-, 8-, 12-, 26-, 51-cell stages) from imaging experiments, which accordingly and respectively are the last moments before five synchronous division events (AB2, AB4, MS&E, C2, C4 cells) and have totally same cell list in all the embryo samples. Moreover, all the cells have finished their cytokinesis and are separated [3, 15].

We assigned the experimentally measured values to the system parameters including cell radius derived from cell volume ($R_i = (3V_i / 4\pi)^{1/3}$), eggshell shape regarded as a thin elliptical cylinder and original cell positions $\boldsymbol{r}_i$ (Figure 1.a, b and Figure 2.a). Reliable and effective connection relationship between cells was determined under three empirical filter criteria, including contact area (larger than 1/36 of at least one cell's surface area), contact duration (no shorter than 3 minutes) and reproducibility (existing in all the four wild-type embryo samples collected) (Figure 1.c) [15].

3.2. Mathematical expression of Repulsion Model and Attraction Model

Force from the confining eggshell is set up here with linear repulsive force, and will be applied to simulation of all the models mentioned in this work.

$$\boldsymbol{F}_{\text{eggshell}\to i} = k_{\text{eggshell}}\boldsymbol{e}_{\text{eggshell}\to i}\cdot\begin{cases} R_i - d_{\text{eggshell},\,i} & 0 < d_{\text{eggshell},\,i} \le R_i \\ 0 & R_i < d_{\text{eggshell},\,i} \end{cases} \tag{1}$$

$\boldsymbol{F}_{\text{eggshell}\to i}$, the force impacted on cell i from eggshell; k_{eggshell}, spring-like force coefficient which is regarded as constant between eggshell and all the cells; $d_{\text{eggshell},\,i}$, the minimum distance between eggshell and centroid of cell i; $\boldsymbol{e}_{\text{eggshell}\to i}$, unit vector perpendicular to the eggshell's tangent plane at contact point (with minimum distance to the centroid of cell i) and orienting inwards.

Next, we briefly reviewed the two previously proposed coarse-grained models for further test and verification [12-14]. Cells are considered as viscoelastic soft spheres. Mathematically, interaction between cells is simplified as pairwise linear force which could be repulsive or attractive depending on distance between their centroids (Figure 2.b). Both of Repulsion Model and Attraction Model can be formulated with a general expression [14].

$$\boldsymbol{F}_{j\to i} = k_{\text{cell}}\boldsymbol{e}_{j\to i}\cdot\begin{cases} 1 & 0 < d_{j,\,i} \le \min\left\{R_i, R_j\right\} \\ \dfrac{-\left[d_{j,\,i} - \alpha\left(R_i + R_j\right)\right]}{\left[\alpha\left(R_i + R_j\right) - \min\left\{R_i, R_j\right\}\right]} & \min\left\{R_i, R_j\right\} < d_{j,\,i} \le 0.5\left(1+\alpha\right)\left(R_i + R_j\right) \\ \dfrac{\left[d_{j,\,i} - \left(R_i + R_j\right)\right]}{\left[\alpha\left(R_i + R_j\right) - \min\left\{R_i, R_j\right\}\right]} & 0.5\left(1+\alpha\right)\left(R_i + R_j\right) < d_{j,\,i} \le \left(R_i + R_j\right) \\ 0 & \left(R_i + R_j\right) < d_{j,\,i} \end{cases} \tag{2}$$

$\boldsymbol{F}_{j\to i}$, the force impacted on cell i from cell j; k_{cell}, spring-like force coefficient which is regarded as constant in all the cell pairs; R_i and R_j, radii of cell i and cell j; $d_{j,\,i}$, distance between centroids of cell i and cell j; $\boldsymbol{e}_{j\to i}$, unit vector orienting from cell j to cell i; α, attraction factor no larger than 1, which determines the distance between centroids of two cells under mechanical balance, namely, $\alpha\,(R_i + R_j)$. The equation (2) represents Repulsion Model when $\alpha = 1$ and Attraction Model when $\alpha < 1$. It's worth noting that α can be assigned with different values in consideration of different levels of attractive strength, which could be asymmetrically distributed inside an embryo (e.g. E-cadherin protein) [14].

3.3. Model revision with short-range negative-feedback repulsion and cell-cell connection atlas
Two steps of revision were implemented successively here, and both of the assumptions possess specific biological meaning in the actual embryonic system. Note that the necessity of revision will be analysed in details in Results Section, on the basis of the simulation results compared with the ones in Repulsion Model and Attraction Model.

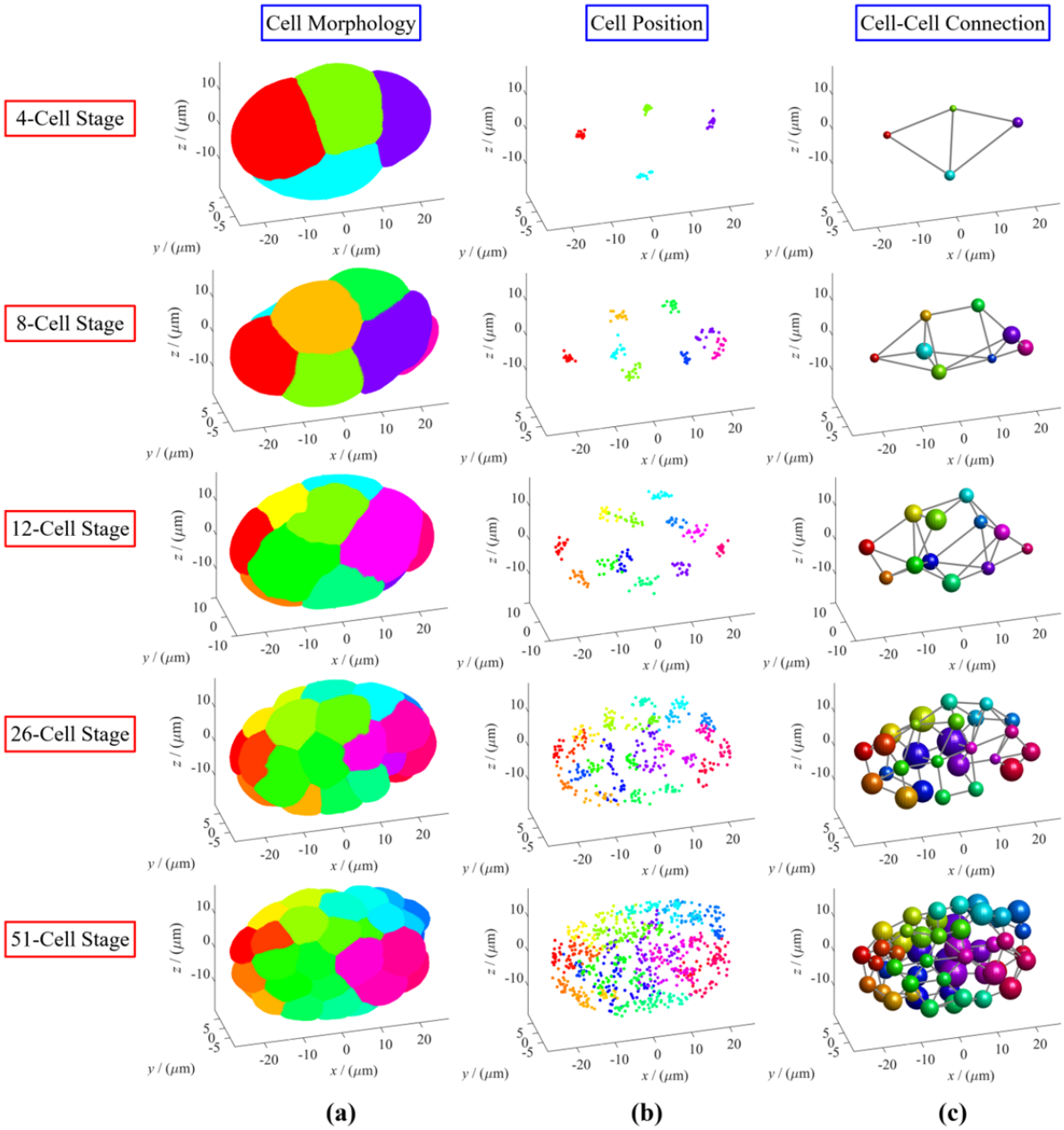


Figure 1. Quantitative 4-dimensional experimental data of multicellular structures during 4- to 51-cell stage of *C. elegans* embryogenesis obtained from Ref. [15], including (a) cell morphology, reconstructed with fluorescent signal on cell membrane; (b) cell position, automatically traced with nucleus position; (c) cell-cell connection relationship, identified under a set of arbitrary filter criteria including contact area, contact duration and reproducibility among wild-type embryo samples; each color represents a cell with specific identity.

Firstly, the previous models used a constant repulsive force between cells when they are very close to each other ($d_{j,i} < \min\{R_i,R_j\}$), partly in order to describe the cell elongation during cytokinesis [14]. However, the constant force was implemented between any two cells. Considering that all the cells are completely divided and separated in the five multicellular structures selected, repulsion between cells should be larger when their centroid distance decreases, so that they can repel each other dramatically and wouldn't be squeezed together unnaturally. Therefore, we replaced the expression with linearly negative-feedback force for the short range $d_{j,i} < \min\{R_i,R_j\}$, i.e., replaced the $\min\{R_i,R_j\}$ with 0 in equation (2).

Secondly, in both Repulsion Model and Attraction Model, no mechanical interaction maintains as long as the distance between two cells' centroids is longer than the sum of their radii. In the real case, cells are deformable, which means if a pair of cells are tightly attractive to each other and meanwhile receive external forces that tend to separate them apart, the mechanical link could still exist even though they are a little distant, for that cells can be stretched and elongated with remained physical contact and connection. This effect could be raised in real embryo by high-density adhesive membrane protein or gap-junction bridge between cells.

Here we took the passive cell elongation into account and introduced an elongation factor β ($\beta \geq 1$), the cells would be physically separated only when $\beta(R_i + R_j) < d_{j,i}$. After replacing $\min\{R_i,R_j\}$ with 0, ($R_i + R_j$) with $\beta(R_i + R_j)$ and α with $\alpha' = \alpha/\beta$ in equation (2), the final version of our revised model was obtained as below.

$$\boldsymbol{F}_{j\to i} = k_{\text{cell}}\boldsymbol{e}_{j\to i}\cdot\begin{cases}\dfrac{-\left[d_{j,i}-\alpha'\beta\left(R_i+R_j\right)\right]}{\alpha'\beta\left(R_i+R_j\right)} & 0 < d_{j,i} \le 0.5\beta\left(1+\alpha'\right)\left(R_i+R_j\right)\\ \dfrac{\left[d_{j,i}-\beta\left(R_i+R_j\right)\right]}{\alpha'\beta\left(R_i+R_j\right)} & 0.5\beta\left(1+\alpha'\right)\left(R_i+R_j\right) < d_{j,i} \le \beta\left(R_i+R_j\right)\\ 0 & \beta\left(R_i+R_j\right) < d_{j,i}\end{cases} \quad (3)$$

Note that there are two kinds of cell-cell interaction here. The first one is the same to Attraction Model with attraction factor $\alpha < 1$ and elongation factor $\beta = 1$, while the second ones with attraction factor $\alpha' = \alpha/\beta < 1$ and elongation factor $\beta > 1$ which consequently lead to strengthened attraction force and wider interactive range up to centroid distance $\beta(R_i + R_j)$. Both sets of connections share the same balanced distance $\alpha(R_i + R_j)$. In the final model, the cell pairs with and without repeated physical contact in experiments correspond to the latter and former interaction respectively.

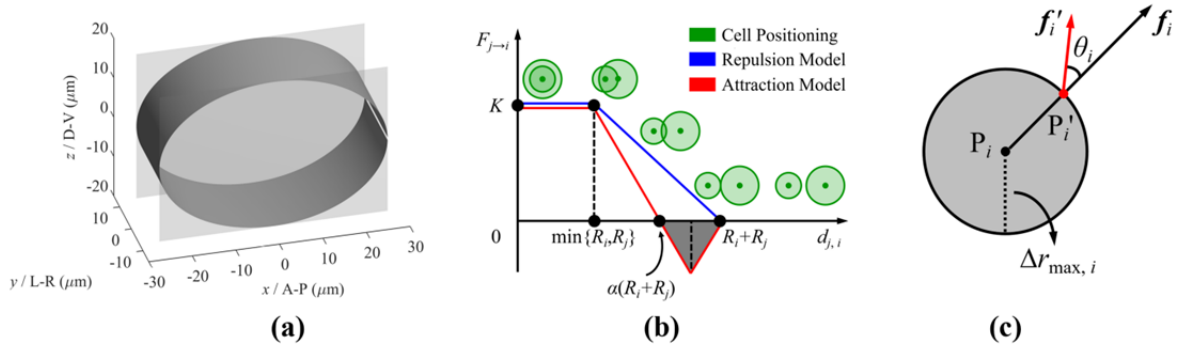


Figure 2. Settings of the mechanical system including (a) eggshell shape; (b) cell-cell interaction force in Repulsion Model and Attraction Model; (c) preliminary estimation on mechanical stability.

3.4. Parameter refinement and fitting

So far, all the previous models and the newly revised ones have been presented with mathematical formulas, with only four system parameters: $k_{eggshell}$, k_{cell}, α and β. As we considered the environment inside embryo perfectly overdamped and adopted quasi-static assumption to calculate the cell positions under mechanical equilibrium, the two force coefficients $k_{eggshell}$ and k_{cell} can be combined into one, that is, $K = k_{cell} / k_{eggshell}$, representing the physical difference of elasticity or stiffness between cells and eggshell. Briefly, a small value of K indicates that the eggshell is harder than the cells, *vice versa*. The other two independent parameters α and β determine the turning points of attractive force, so eventually there are a total of three parameters in the models.

To select appropriate values for system parameters, we preliminarily estimated if a cell has stable state at its experimental location by inspecting if it receives a restoring force with component vector pointing at its original position when getting perturbed (Figure 2.c). Here we defined $\Delta r_{max,\,i}$ as the boundary of available spatial variation for cell i and used it for this estimation, which is the minimum value between 3 times the RMSD of cell positions among 17 wild-type embryo samples (Figure 1.b, c) and the radius of cell i.

$$\Delta r_{max,\,i} = \min\left\{3\sqrt{\sum_{i=1}^{N}\left|\boldsymbol{r}_i - \overline{\boldsymbol{r}_i}\right|^2 \Big/ N},\, R_i\right\} \tag{4}$$

Here, N denotes the total number of embryo samples ($N = 17$); $\overline{\boldsymbol{r}_i}$ denotes the average position of cell i. For the given K (i.e. $k_{cell} / k_{eggshell}$), α and β, we put all the cells into their experimental positions at the beginning. Next, for each cell i located at P_i, we calculated its resultant force $\boldsymbol{f}_i$ using equations (1)(2) or equations (1)(3) and marked the new location P_i', which shifted from P_i along the direction of $\boldsymbol{f}_i$ with distance of $\Delta r_{max,\,i}$. Then, we calculated the new resultant force $\boldsymbol{f}_i$' impacted on cell i at P_i'. If there's a restoring component factor in $\boldsymbol{f}_i$', in other words, the intersection angle θ_i between $\boldsymbol{f}_i$ and $\boldsymbol{f}_i$' is larger than 90° (obtuse), the cell would be considered with considerable mechanical stability. If all the cells acquire $\theta_i > 90°$, the values of K, α, β will pass the test and be adopted as valid ones. Note that computation on mechanical analysis is much more efficient and time-saving than that on cell motion, however, this rough calculation ignored the effect of tangential force, change of force field within the boundary as well as the displacement of other cells. Thus, we used this estimation method only for preliminary and large-scale parameter searching, and then simulation on 3D cell movement over time would be carried out to make a further verification.

3.5. Evaluation on positional variation under mechanical equilibrium

As the parameter-searching method above only approximately estimate the stability of a given multicellular structure, one needs to confirm that, with the assigned values filtered by equation (4), the cells could indeed migrate into locations close to the experimental ones as their final states. To this end, we let the cell system relax from their original positions, namely, the cells are free to move via interaction with eggshell and other cells by equations (1)(2) or equations (1)(3), until all the cells reaching their mechanical equilibrium state ($\boldsymbol{f}_i = \boldsymbol{0}$, time step = 0.05, at least 300000 steps). The positional variation between experiment and simulation of each cell was evaluated by comparing the spatial deviation and cell radius.

$$\eta_i = \frac{\left|\boldsymbol{r}_{exp,\,i} - \boldsymbol{r}_{sim,\,i}\right| - R_i}{R_i} \tag{5}$$

Simply, $\eta_i = 0$ when the cell's terminal location was exactly on its boundary surface (i.e. $|\boldsymbol{r}_{exp,\,i} - \boldsymbol{r}_{sim,\,i}| = R_i$); $\eta_i = -1$ when the terminal location perfectly coincides with experiment measurement (i.e. $|\boldsymbol{r}_{exp,\,i} - \boldsymbol{r}_{sim,\,i}| = 0$). The cells with $\eta_i > 0$ would be regarded as extreme outlier in the mechanical model, which may be influenced by other unconsidered potential factors, for example, severe cell deformation and active cell motion [17, 18].

3.6. Comparison of performance between models

All the operations mentioned would be carried out for the previous models (Repulsion Model and Attraction Model) as well as the two revised ones. The two proposed revisions were added one by one, so that these two models would be termed as 'Revised Model 1' and 'Revised Model 2' for simplicity and clarity. As the models were improved progressively, we first used equations (1)(2) to test Repulsion Model ($\alpha = 1$) and Attraction Model ($\alpha < 1$) and analysed their performance. Then we used equations (1)(3) to test the two implemented corrections successively, compared the results with previous models using cellular positional variation defined by equation (5), and investigated the underlying biophysical meanings of the new model components.

4. Results

Here, a total of four models (i.e. Repulsion Model, Attraction Model, Revised Model 1, Revised Model 2) would be simulated under the best combination of system parameters. Both their advantages and limitations would be analysed and discussed deeply.

Setting $k_{eggshell} = 1$, $K = k_{cell} = 0.01 \sim 100$ with equal interval of 0.01 and $\alpha = 0.70 \sim 1.00$ with equal interval of 0.05, we used the 4-, 8- ,12-, 26-cell stages for preliminary estimation on cell positional stability with equation (1)(2)(4), so as to search the best value combination of K and α. We chose these four stages for that robust onset of maternal-zygotic transition and gastrulation just starts to be activated since 26-cell stage, so that the embryo system could be comparatively simple [8, 12, 13, 19]. Besides, the screened range of α was determined between 0.70 and 1.00, for that it was measured to be around $0.75 \sim 0.90$ at 4-cell stage (0.90 for EMS-P2 cell pair and 0.75 for the others) [14]. Due to lack of high-quality 3D time-lapse imaging on adhesive protein distribution (e.g. E-cadherin protein) with single-cell resolution, it's difficult to figure out the real value of α for each cell pair, thus, we neglected the attraction asymmetry and searched a universal value to represent this type of interaction for all the cell pairs without any bias.

4.1. The Attraction Model has better performance than Repulsion Model

Here, we would test and compare the previously proposed Repulsion Model ($\alpha = 1$) and Attraction Model ($\alpha < 1$). For all the α values inspected (i.e. 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00), the intersection ranges of valid K values among the four stages were always $0.01 \sim 0.04$ in both models, indicating that the eggshell is much harder than cells and help shape the multicellular structure inside an embryo. With the K values filtered (i.e. $K = 0.01, 0.02, 0.03, 0.04$) and α values inspected, movement of cells as well as their terminal steady positions were computed in both models, so as to ensure that the system could indeed reach the experimental pattern with high stability and little deviation (Figure 3.a, b). Repulsion Model possesses positional variation lower than 0 at both 4- and 8-cell stages. Nevertheless, there are more and more cells with $\eta > 0$ at 12- and 26-cell stages and the highest value could reach around 0.43, in other words, the reconstructed structure begins to severely deviate from the experimental one as cell number increases. On the other hand, Attraction Model succeeds in reproducing the experimental structure and its best performance was achieved when $K = 0.03 \sim 0.04$ and $\alpha = 0.90$ (Figure 3.a, c). As the multicellular structures under two sets of value of repulsion factor $K = 0.03$ and 0.04 are similar to each other, hereafter we would use the former one to represent Attraction Model and give further analysis. It's worth pointing out that the well-tuned value of α is in line with the one obtained from *in vitro* experiment before, which measured the balanced distance $\alpha (R_i + R_j)$ between cells at 4-cell stage after removal of eggshell, and found that 4 pairs of cells share the same level of attraction ($\alpha \approx 0.75$) and the other cell pair is weaker ($\alpha \approx 0.90$, less accumulation of E-cadherin protein) [14]. Notably, Attraction Model has much lower positional variation than Repulsion Model, with the highest values of η always smaller than 0, except the MS cell at 8-cell stage. This observations revealed the significant role of cell-cell attraction in robustness and stability of multicellular structure, which also helps shape the topology of developing embryo. At 4-cell stage, both models exhibit same performance, which means that the final balanced diamond-shape structure possesses high stability and is easy to form.

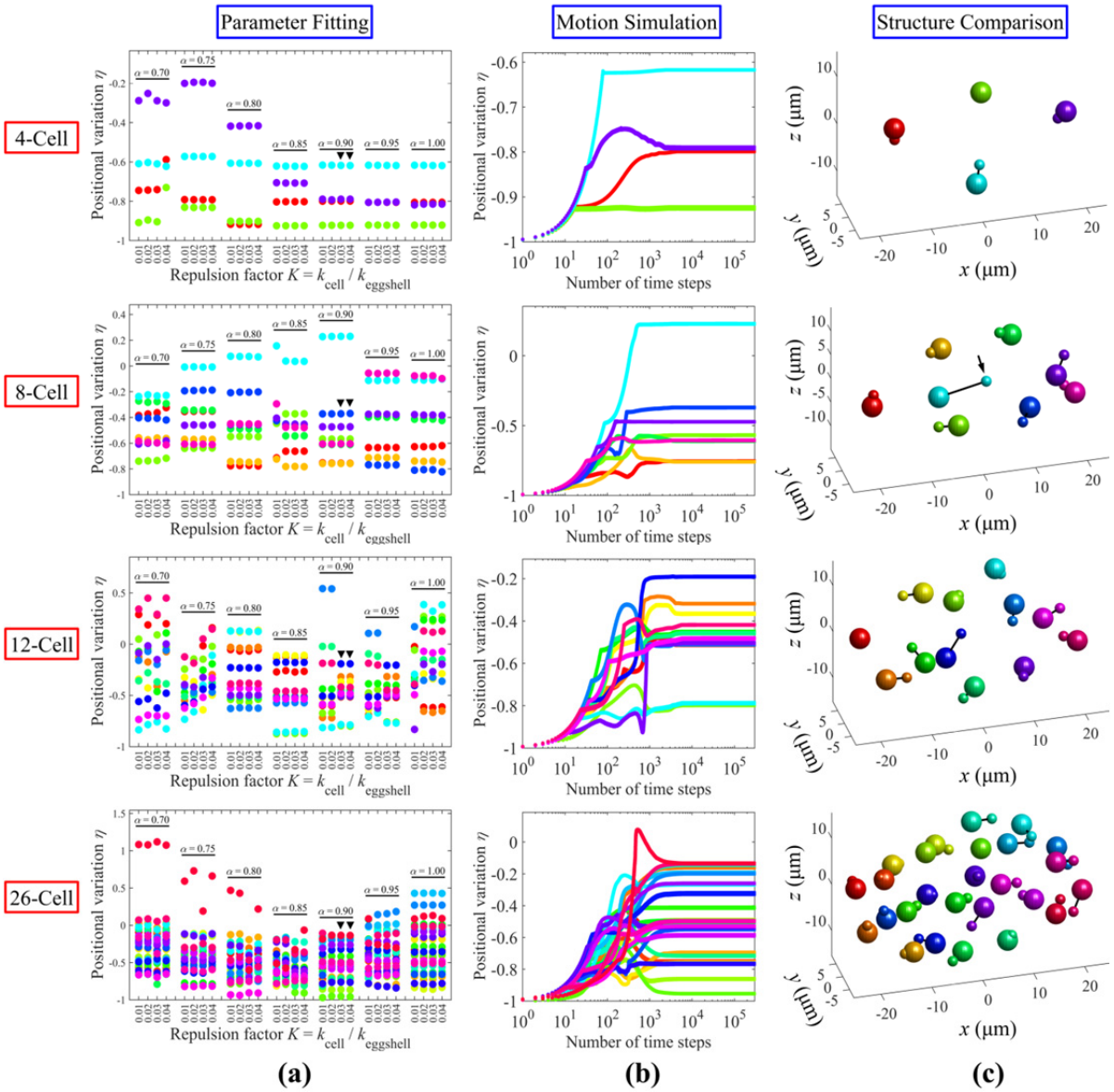
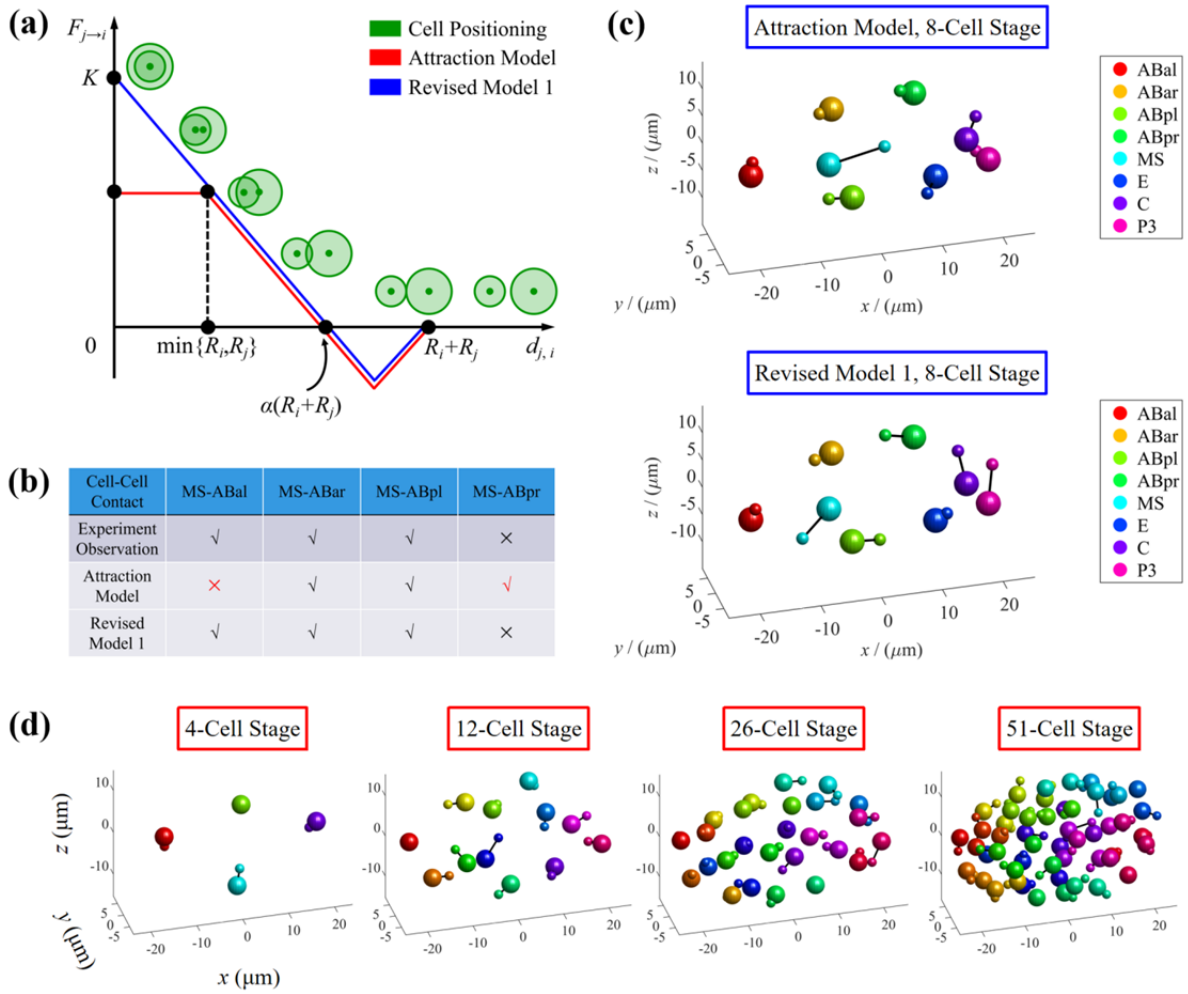


Figure 3. Positional variation between cell positions from experiment and simulation at 4-, 8-, 12-, 26-cell stages, using Repulsion Model (α = 1) and Attraction Model (α < 1). (a) Positional variation under different value assignments of K and α, revealing the optimal parameters as K = 0.03 ~ 0.04 and α = 0.90 for Attraction Model, indicated with black triangles; (b) positional variation over time during simulation process in Attraction Model (K = 0.03, α = 0.90), which allowed the system to ultimately reach mechanical equilibrium; (c) cell-arrangement patterns of experiments (larger spheres with radius = 2 μm) and simulations (smaller spheres with radius = 1 μm) in Attraction Model (K = 0.03 and α = 0.90); the mislocation of MS cell is indicated with black arrow; black lines connecting cell positions in experiment and simulation are plotted to help visual comparison; each color represents a cell with specific identity.

4.2. A short-range negative-feedback force can rescue incorrect cell location in Attraction Model

In spite of the high performance of Attraction Model, there's one cell named MS, the founder cell of pharynx and muscle [1], obviously and mistakenly shifting into the center of embryo in simulation at 8-cell stage ($\eta \approx 0.23$). Even though the mislocation was corrected in later stages, it still leads to

incorrect connection relationships between MS and AB cells at this moment, while the connection map among them has been reported to have decisive roles on cell pattern progression and cell fate specification. In brief, MS cell has to contact with ABal to regulate its division orientation through Latrophilin signaling, making its spindle orient toward MS. This guidance on cell division further builds up the precise locations of AB blastomere, where MS contacts with ABalp and ABara instead of ABala and ABarp so as to transduct the second Notch signaling for cell differentiation [7, 20-22]. One possible reason accounting for this error is that, the cell-cell repulsion was set to be constant when they are close to each other ($d_{j,i} < \min\{R_i,R_j\}$), so the cells could not repel each other strongly enough and then consequently migrate into some unnatural positions. To test if the severe mislocation of MS is related to the constant short-range force and correct it as possible, we executed the first formula correction, namely a short-range negative-feedback force, which has linear expression as described in Section 3.3 (Figure 4.a).



Cell-Cell Contact	MS-ABal	MS-ABar	MS-ABpl	MS-ABpr
Experiment Observation	√	√	√	×
Attraction Model	×	√	√	√
Revised Model 1	√	√	√	×

Figure 4. Comparison between Attraction Model ($K = 0.03$, $\alpha = 0.90$) and Revised Model 1 ($K = 0.01$, $\alpha = 0.90$). (a) Cell-cell interaction force in Attraction Model and Revised Model 1; (b) contact relationships between MS and AB cells at 8-cell stage; (c) cell-arrangement patterns at 8-cell stage of experiments (larger spheres with radius = 2 μm) and simulations (smaller spheres with radius = 1 μm) in both models; (d) cell-arrangement patterns of experiments (larger spheres with radius = 2 μm) and simulations (smaller spheres with radius = 1 μm) in Revised Model 1; black lines connecting cell positions in experiment and simulation are plotted to help visual comparison; each color represents a cell with specific identity.

Both Attraction Model and Revised Model 1 were simulated for all the five developmental stages to make a comprehensive comparison. The optimal fitting result of system parameters in Revised Model 1 is $K = 0.01$ and $\alpha = 0.90$. Interestingly, the severe deviation in MS was rescued with limited positional variation $\eta \approx -0.11$ and completely correct connection relationships to the four AB cells (Figure 4.b, c). In addition, using Revised Model 1, simulations for all the 4-, 8-, 12-, 26-cell stages were well in line with experiment results for each cell, i.e., no cell shifted farther than its radius in simulation (Figure 4.c, d). The same level of precision could be also achieved in Attraction Model except the extreme outlier MS cell. For the 51-cell stage, there are 3 extreme outliers with $\eta > 0$ in both models, the highest one in Attraction Model was 0.37, while the one in Revised Model 1 was 0.11 (data not shown). Although Revised Model 1 appears to be a bit better than Attraction Model in performance of cell-position prediction at this stage, it raised our concerns that if the model can be further improve to recapture more features of the multicellular structure, which is comprised of both 3D cell position and cell-cell connection relationship. Therefore, we carried out the second formula correction and test the performance of Revised Model 2.

4.3. An elongated range of attractive force can further improve precision of multicellular structure

Regarding the possible deformation and elongation of cell caused by intense external force, it's reasonable to believe that a pair of cells with strong attraction to each other can be stretched and maintain their mechanical link at the same time, even though the distance of their centroids is larger than the sum of their radii ($R_i + R_j < d_{j,i}$). Here we introduced the elongation factor β to extend the interactive range between cells, as described in Section 3.3 (Figure 5.a). We proposed a new assumption that, the connections between specific cells with large contact area, long contact duration and high reproducibility observed empirically, are coordinated by genetically programmed strong attraction such as high-density adhesive membrane protein or gap-junction bridge between cells [14, 23]. As a result, those cell pairs can be deformed and stretched with physical contact retained, and should have a wider interactive range than the ones without consistent and effective connection in real embryo. The cell pairs without consistent connection observed from repeated experiments would be assigned with interaction form completely same to Attraction Model ($\alpha' = \alpha$, $\beta = 1$), while the ones with consistent connection relationships have an elongated interactive range ($\alpha' = \alpha/\beta$, $\beta > 1$), and both kinds of interactions obey equation (3).

Inspired by the mislocation and misconnection of MS in Attraction Model which was successfully rescued in Revised Model 1, here we also took the cell-cell connection into account and check if the elongated range of attractive force help reproduce the experimentally observed cell-cell connection atlas. We used the percentage of incorrect connection pairs γ to quantify variation of cell-cell connection atlas. As the total number of all the possible connection pairs is combination number C_N^2 where N denotes cell number of embryo, γ would be calculated as (Number of Lost Connection + Number of Extra Connection) / C_N^2. Note that the cells in simulation are regarded as physically connected only if $d_{j,i} < R_i + R_j$.

Setting $\beta = 1.00 \sim 2.00$ with equal interval of 0.05, Revised Model 2 was first simulated for the 51-cell stage with parameter values of K and α adopted from Revised Model 1 (i.e. $K = 0.01$, $\alpha = 0.90$), and the optimized fitting result of elongation factor was $\beta = 1.20$, which could approximately lower the average positional variation η from -0.40 to -0.47 and the percentage of incorrect connection γ from 0.072 to 0.046 (Figure 5.b).

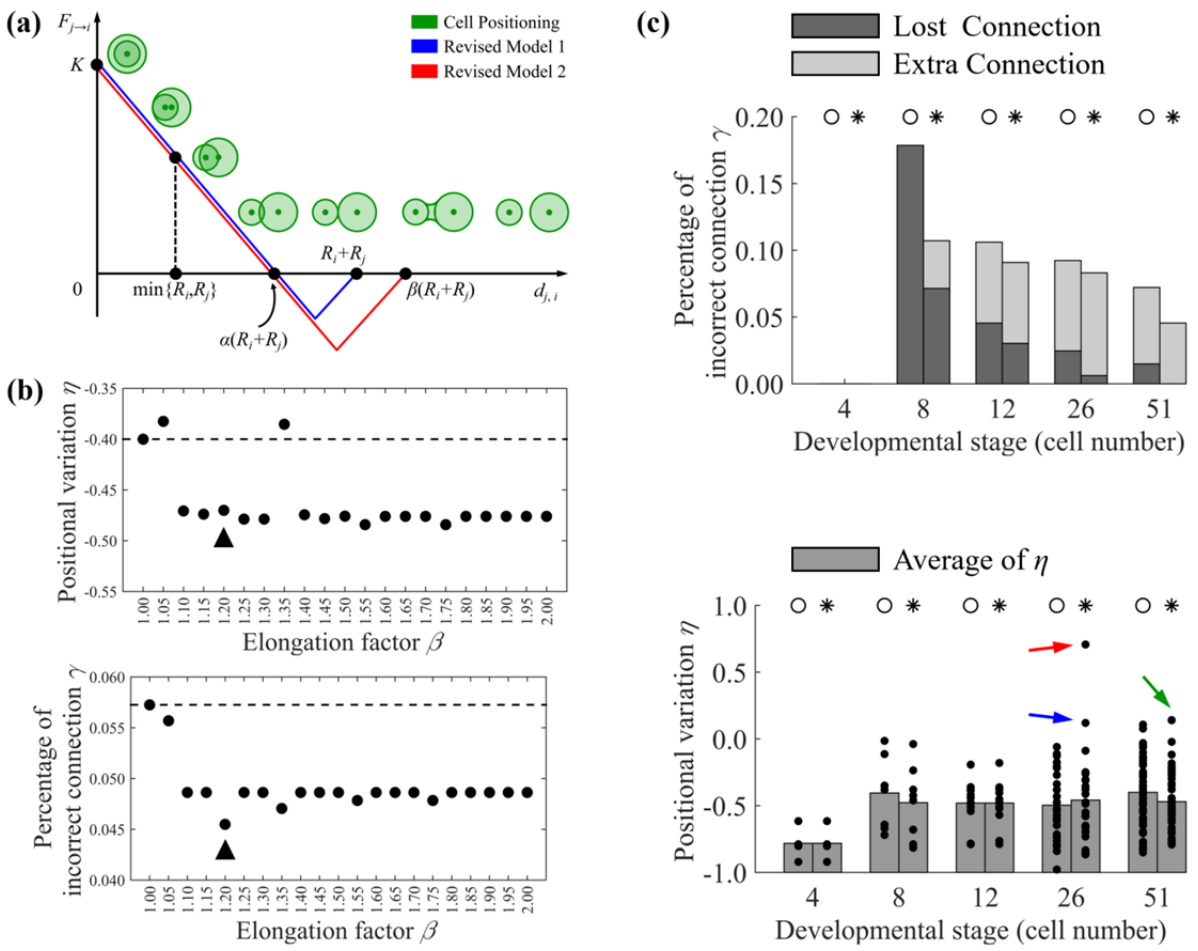


Figure 5. Comparison between Revised Model 1 ($K = 0.01$, $\alpha = 0.90$, $\beta = 1.00$) and Revised Model 2 ($K = 0.01$, $\alpha = 0.90$, $\beta = 1.20$). (a) Cell-cell interaction force in Revised Model 1 and Revised Model 2; (b) average positional variation and percentage of incorrect connection under different value assignments of β, revealing the optimal value assignment as $\beta = 1.20$, indicated with black triangles; (c) positional variation and percentage of incorrect connection including the lost ones and extra ones in Revised Model 1 (marked with circle □) and Revised Model 2 (marked with asterisk *); the 3 extreme outliers with high positional variation in Revised Model 2 are indicated with red arrow for P4 ($\eta \approx 0.70$), blue arrow for MSpp ($\eta \approx 0.12$) and green arrow for MSapa ($\eta \approx 0.14$).

This step of model correction mainly contributes to reconstruction of cell-cell connection atlas, as lower error rate was achieved at all the five developmental stages, especially for the lost connections (Figure 5.c). In brief, the extensive mechanical link can help achievement of effective connection between specific cells observed in real *C. elegans* embryos repeatedly, which is probably required for specific cell-cell communications and subsequent fate specification or spindle reorientation. In the case of positional variation, although the average positional variation was constant or lower at 4-, 8-, 12-, 51-cell stages and there's only one cell left (MSapa) with $\eta > 0$ at 51-cell stage, two more cells deviated dramatically at 26-cell stage, that is MSpp and P4. Intriguingly, all the three cells are involved with an critical active developmental events in metazoan animal called gastrulation, which produces ectoderm, mesoderm and endoderm by cell ingression [8, 24-27]. Both MSpp and P4 contact with Ea and Ep respectively, while Ea and Ep are the first ingressing cells and reported to be able to influence the passive motion of neighbouring cells dominantly [26]. Apart, the MSap descendants would also enter the body cavity and have been proposed to be potential Notch signaling centers inside the embryo [27].

The process of gastrulation is actively regulated by actomyosin contractility and asymmetric adhesion [24, 25], which is not considered in these models. Nevertheless, as the cell number increases and most of cells still remain low positional variation $\eta < 0$, the mechanical system consisting of repulsion between cell and eggshell and repulsion-attraction interaction between cells still predominately controls the basic multicellular structures of the developing *C. elegan*s embryo.

5. Conclusion and discussion

Robustness and precision of multicellular structure progression is required for a series of biological procedures during metazoan embryogenesis. For *C. elegans*, how cells interact with each other and subsequently form stereotypic multicellular structures has been investigated by several studies using physical model and computational simulation, but with a few limitations and in lack of test in depth. To overcome the shortcomings, verify the validation of previous models and provide a more applicable framework, we collected experimental data of eggshell shape, cell volume, cell position to refine system parameters, which ultimately left no more than three independent parameters: repulsion factor $K = k_{\text{cell}} / k_{\text{eggshell}}$, attraction factor α and elongation factor β. After parameter searching, we compared the four models successively, including Repulsion Model [12, 13], Attraction Model [14] and two revised models (i.e. Revised Model 1 and Revised Model 2). We found that Repulsion Model can only well fit the early developmental stages while the Attraction Model succeeds in reproducing the majority of cell locations up to ~50-cell stage. Mistaken cell position in Attraction Model could be rescued by introducing a short-range negative-feedback force to avoid cells from being squeezed too much toward each other. Combining information of empirically observed cell-cell connection atlas with the revised model, prediction on cell position and contact mapping was further improved, but meanwhile failed to predict few cells which all coincide to be involved with gastrulation. This could be potentially used to infer the cell-level regulated biophysical activity during embryogenesis and morphogenesis.

Despite that we collected real information from *in vivo* imaging experiment and improved the framework, there are still a few of potential limitations that have to be explained. Firstly, the experimentally observed cell position is not perfectly accurate for that the nucleus is not always located at the center of cell, and positional variation exists among individuals [3]. Secondly, the cell-cell contact classification in simulation is based on distance between cell centers, ignoring the 3D deformable morphology of cell. Thirdly, the interactive force between cells is simplified, other types of actually existing force such as surface friction and internal cytoskeleton growth are ignored. Last but not least, the introduction of elongated range of attractive force based on empirically consistent cell-cell contact atlas is hypothesized, one needs to conduct high-quality 4D live imaging on distribution of adhesion protein and gap junction in embryo, and further reconstruct the embryo *in silico* under guidance of solid experimental evidence [14, 23].

C. elegans embryo can survive under compression or removal of the eggshell [16, 28], which means that the necessary cell-cell physical connection and signaling transduction network is somehow preserved and robust against mechanical perturbations. This could probably be achieved by both universally passive cell packing and genetically programmed adhesion or junction between cells, so as to instruct the accurate progression of global multicellular structure as well as cell-cell connection network. The simulations in this work tested, verified and improved the previous coarse-grained models by introducing a short-range negative-feedback force and an elongated range of attractive force. With the consistent cell-cell connection atlas obtained empirically, we further reconstructed the multicellular structures at ~50-cell stage, beyond the initiation of gastrulation. This work improved and verified a general applicable modeling framework for mechanical reconstruction, which could be used in embryogenesis and morphogensis of *C. elegans* as well as other species. The simulations in this work were performed by Matlab and all the details are available upon request.

Acknowledgments
This work was supported by the Ministry of Science and Technology of China (2015CB910300) and the National Natural Science Foundation of China (91430217). Part of the analysis was performed on the High-Performance Computing Platform of the Center for Life Sciences at Peking University.

References
[1] Sulston J E, Schierenberg E et al. 1983 *Dev. Biol.* **100** 64-119
[2] Corsi A K, Wightman B and Chalfie M 2015 *Genetics* **200** 387-407
[3] Guan G, Wong M K et al. 2019 System-level quantification and phenotyping of early embryonic morphogenesis of *Caenorhabditis elegans*. *Preprint* bioRxiv: 776062
[4] Thorpe C J, Schlesinger A et al. 1997 *Cell* **90** 695-705
[5] Chen L, Ho V W S et al. 2018 *Genetics* **209** 37-49
[6] Walston T, Tuskey C et al. 2004 *Dev. Cell* **7** 831-41
[7] Langenhan T, Promel S et al. 2009 *Dev. Cell* **17** 494-504
[8] Lee J Y, Marston D J et al. 2006 *Curr. Biol.* **16** 1986-97
[9] Raich W B, Agbunag C and Hardin J 1999 *Curr. Biol.* 1999, **9** 1139-46
[10] Kajita A, Yamamura M and Kohara Y 2002 *Genome Inform.* **13** 224-32
[11] Kajita A, Yamamura M and Kohara Y 2003 *Bioinformatics* **19** 704-16
[12] Fickentscher R, Struntz P and Weiss M 2013 *Biophys. J.* **105** 1805-11
[13] Fickentscher R, Struntz P and Weiss M 2016 *Phys. Rev. Lett.* **117** 188101
[14] Yamamoto K and Kimura A 2017 *Development* **144** 4437-49
[15] Cao J, Guan G et al. 2019 Establishment of morphological atlas of *Caenorhabditis elegans* embryo with cellular resolution using deep-learning-based 4D segmentation. *Preprint* bioRxiv: 797688
[16] Jelier R, Kruger A et al. 2016 *Cell Syst.* **3** 160-71
[17] Naganathan S R, Fürthauer S et al. 2014 *eLife* **3** e04165
[18] Kenji S and Bruce B 2018 *Dev. Cell* **46** 257-70
[19] Wong M K, Guan D et al. 2016 *J. Biol. Chem.* **291** 12501-13
[20] Mello C C, Draper B W and Priess J R 1994 *Cell* **77** 95-106
[21] Mango S E, Thorpe C J et al. 1994 *Development* **120** 2305-15
[22] Park F D and Priess J R 2003 *Development* **130** 3547-56
[23] Simonsen K T, Moerman D G and Naus C C 2014 *Front. Physiol.* **5** 40
[24] Nance J, Munro E M and Priess J R 2003 *Development* **130** 5339-50
[25] Lee J Y, Marston D J et al. 2006 *Curr. Biol.* **16** 1986-97
[26] Nance J, Lee J Y and Goldstein B 2005 Gastrulation in *C. elegans*, WormBook, ed The *C. elegans* Research Community pp 1-13
[27] Priess J 2005 Notch signaling in the *C. elegans* embryo, WormBook, ed The *C. elegans* Research Community pp 1-16
[28] Schierenberg E and Junkersdorf B 1992 *Roux. Arch. Dev. Biol.* **202** 10-16