Supplementary Information for

"A kinematic equation for the morphogenetic reproducibility of an animal"

Jianguo Wang[1, #], Long Xiao[2, #], Miaoling Yang[2], Zeqi Yao[1], Shanjun Deng[1], Zhuo Du[2] & Xionglei He[1]

[1] Hongkong Institute for Advanced Studies, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China

[2] State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

[#] These authors contribute equally

Correspondence should be addressed to:

X.H. (hexiongl@mail.sysu.edu.cn) or Z.D. (zdu@genetics.ac.cn)

This file contains:

Supplementary Figures 1-16

Supplementary Notes I-V

Legends for Supplementary Tables 1-5

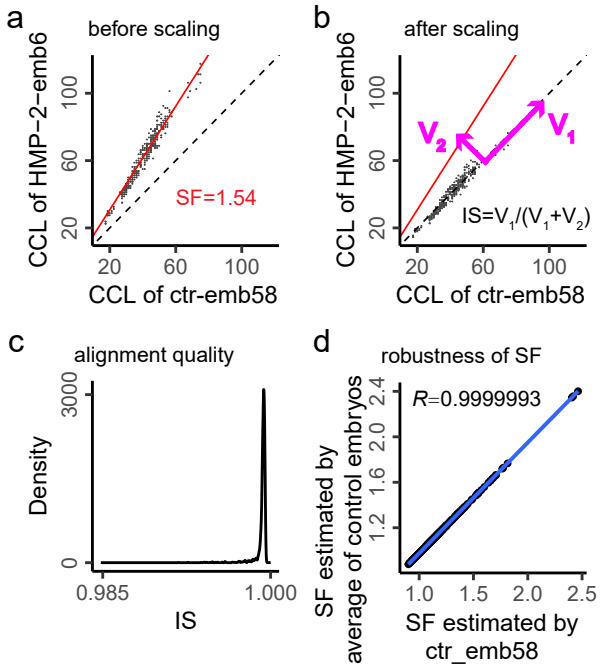

Fig. S1. Alignment of cell cycle length

(a) An example of an HMP-2 RNAi embryo (HMP-2-emb6). The Cell Cycle Lengths (CCLs) of HMP-2-emb6 are approximately 1.54 times those of the reference embryo (ctr-emb58), termed the Scaling Factor (SF). The SF is estimated by the slope of an orthogonal regression line passing through the origin, between the CCLs of HMP-2-emb6 and those of ctr-emb58. The dashed line represents the diagonal line.
(b) Following SF estimation, the scaled CCLs are derived by dividing the raw CCLs by the SF. The scaled CCLs of HMP-2-emb6 are expected to align with the CCLs of ctr-emb58, meaning that data points should distribute along the diagonal line. The variance of these points can be decomposed into two components: one along and the other orthogonal to the diagonal line (V_1 and V_2). The proportion of variance along the diagonal line is defined as the Identity Score (IS), which is used to evaluate alignment quality.
(c) The distribution of Identity Scores across all 2039 embryos.
(d) Using the average CCLs of 105 control embryos as a reference, we re-estimated the SF for each embryo. The SFs obtained based on the two references can be well-fitted by a straight line passing through the origin (blue color), indicating that scaled CCLs based on two references can be approximately transformed by a fold change. This suggests that the choice of reference does not impact the analyses in this study. Ultimately, we chose a random embryo, ctr-emb58, as the reference instead of the average of control embryos. This is because the 105 embryos still have SFs (albeit with smaller variations than the perturbed embryos), which could introduce bias given that data with different scales do not follow the same distribution and are thus unsuitable for average calculation.

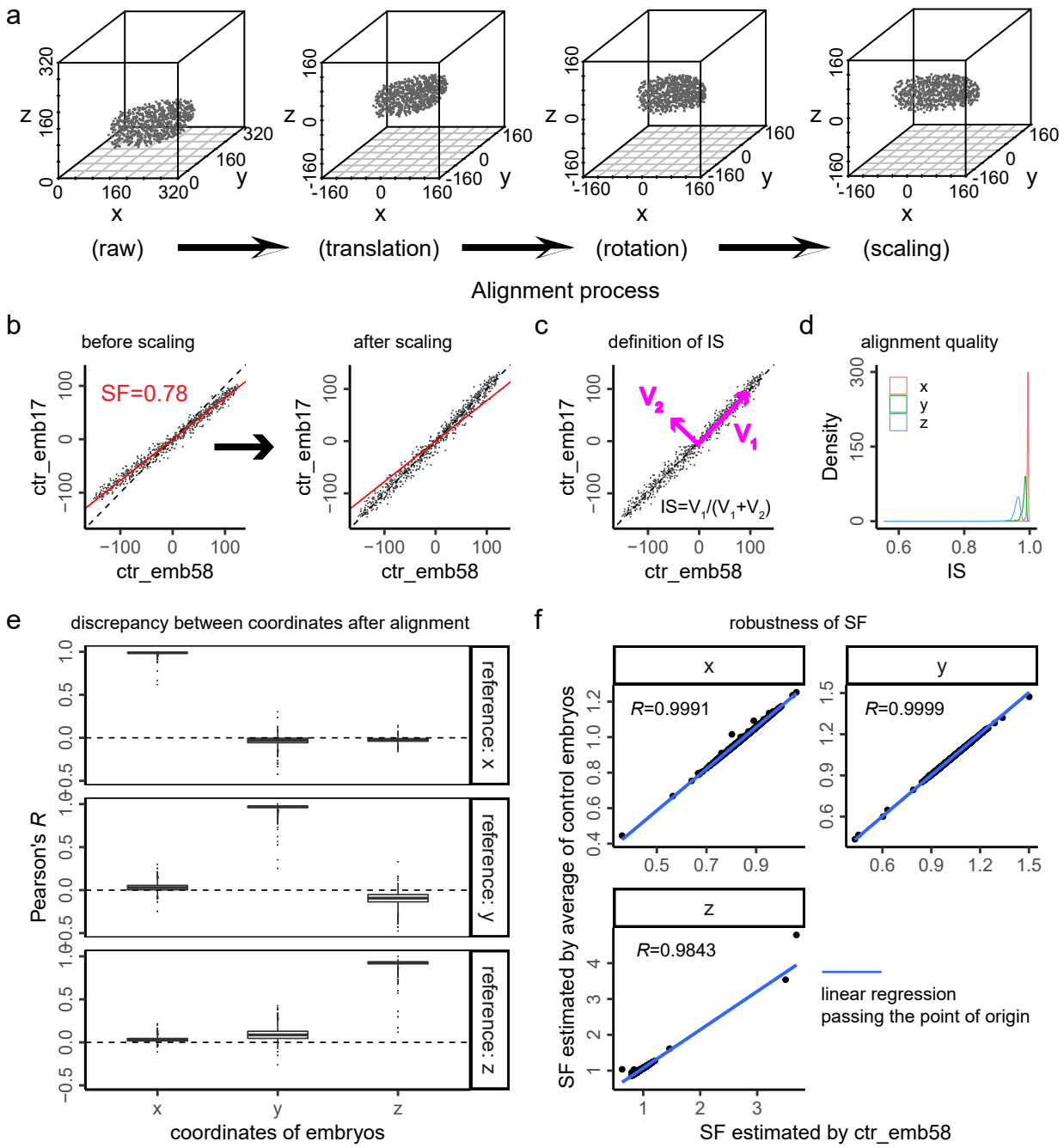

Fig. S2. Alignment of spatial coordinates

(a) The raw *x*, *y*, and *z* coordinates for each embryo undergo a sequential process of translation, rotation, and scaling to generate aligned coordinates across embryos. Each point in the three-dimensional plot represents a cell in an embryo.

(b) The same scaling process employed for CCLs is used to scale the three position features.

(c) Definition of the Scaling Factor (SF) for an embryo.

(d) The distribution of SFs for embryos for *x*, *y*, and *z* coordinates, respectively.

(e) Evaluation of coordinate misassignment in alignment. The horizontal axis represents *x*, *y*, and *z* coordinates of a focal embryo while the vertical axis represents the coordinates of the reference embryo. Pearson's *R* is used to evaluate the similarity between any coordinate of a focal embryo and that of the reference embryo.

(f) SFs remain robust when the average coordinates of control embryos are used as a reference.

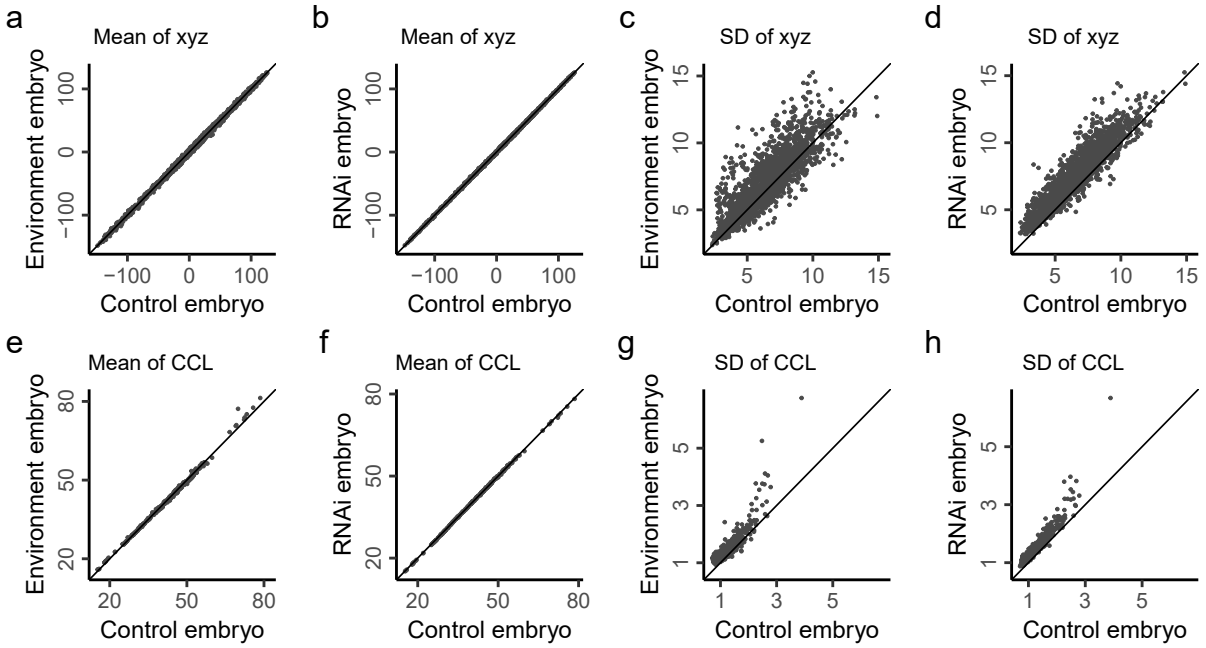

Fig. S3. Comparison of mean and standard deviation of spatio-temporal features
In each panel, each point represents the mean or standard deviation (SD) of a particular spatio-temporal feature (after alignment), estimated based on two types of embryos.
(a-d) Comparison for spatial features.
(e-h) Comparison for CCLs.
(a-b and e-f) Comparison for the mean of features.
(c-d and g-h) Comparison for the standard deviation (SD) of features.
(a, c, e, and g) Comparison between control and environment-perturbed embryos.
(b, d, f, and h) Comparison between control and RNAi-perturbed embryos.

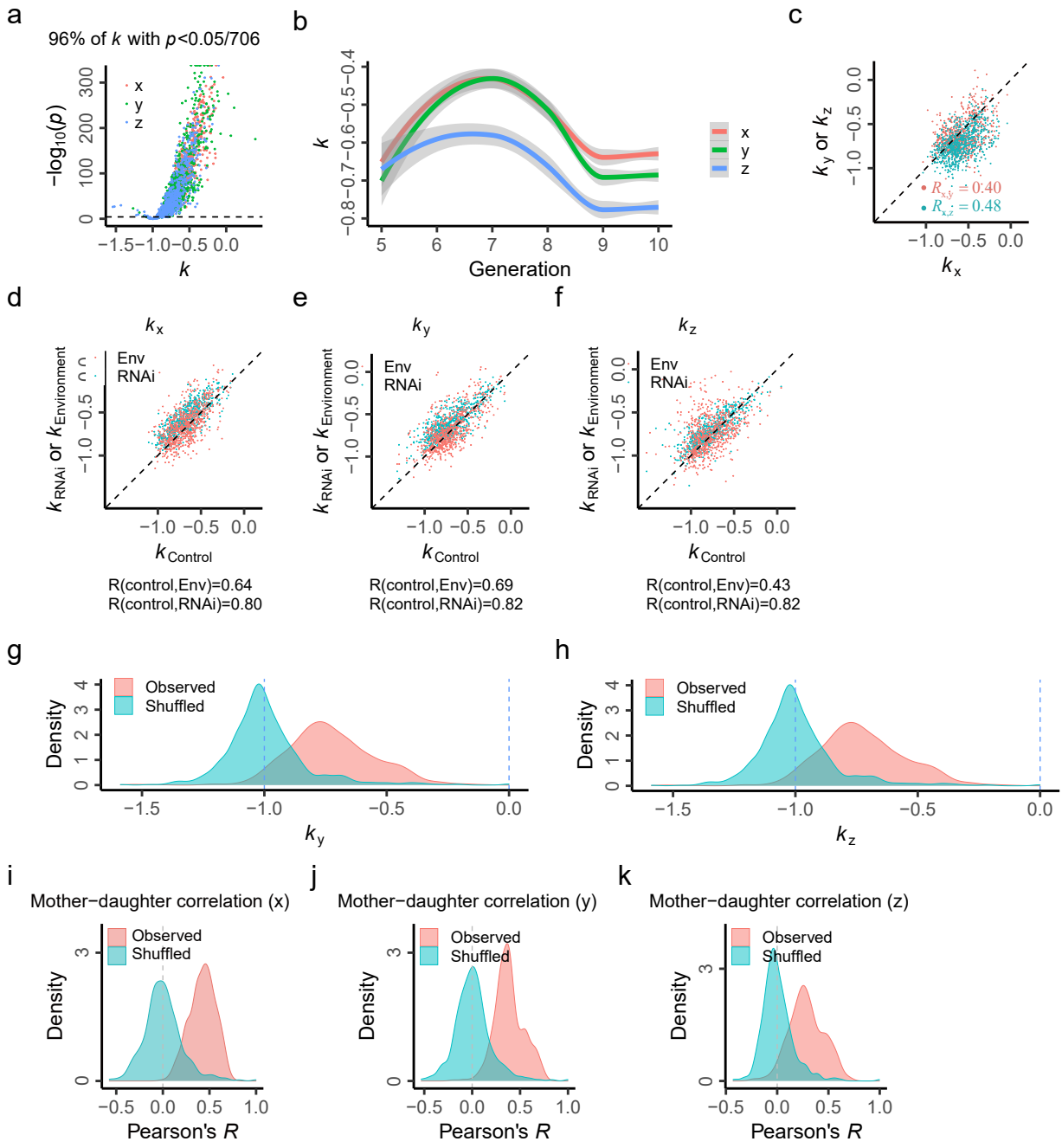

Fig. S4. The properties of the *k* values for spatial coordinates.
(a) Note that the significance of the *k* value between a mother cell and the corresponding mother-daughter difference is determined by the significance of *a* in Eq. (1), since *k*=*a*−1. Therefore, we test the significance of *k* using the p-value (t-test) for the Pearson's correlation between mother cell noise and daughter cell noise. The relationship between the *k* values and the p-values are shown for all mother-daughter cell pairs, for three coordinates respectively. The dashed line represents the threshold of *p*=0.05.
(b) The variation of *k* values across generations. The fitted lines are obtained from loess regression with 95% confidence interval shown, for three coordinates, respectively.
(c) The correlated relationships between the kx, ky and kz.
(d-f) The correlated relationships between the k values of control, Env and RNAi embryos.
(g-h) The distributions of ky and kz compared with those calculated among shuffled M-D pairs.
(i-k) The distributions of Pearson's R between the noise of mother and daughter at each coordinate compared with those calculated among shuffled M-D pairs.

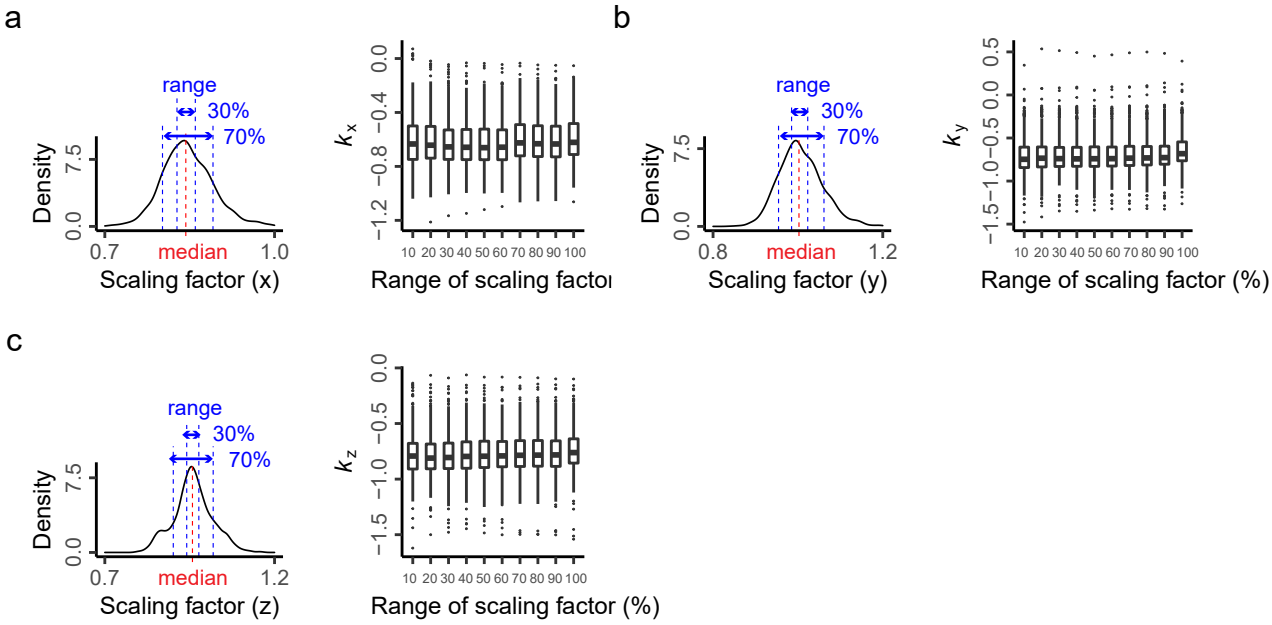

Fig. S5. Robustness of *k* against scaling factor
The left side of each panel shows the definition of the scaling range, which describes the percentage around the median of Scaling Factors (SFs). The distribution of SFs for the *x*, *y*, and *z* coordinates are shown, respectively. Outliers are not shown for clarity. The right side of each panel shows the estimation of k_x, k_y, and k_z, respectively, given a range of scaling factors. A series of scaling factor ranges (from 10% to 100%) are considered.

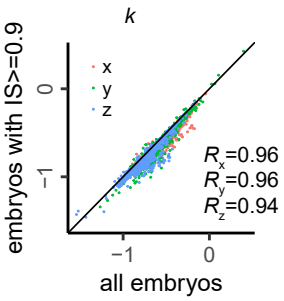

Fig. S6. Robustness of *k* against alignment quality
We define high alignment quality for an embryo if the Identity Score (IS) for all three coordinates of the embryo is greater than or equal to 0.9. This cutoff filters out 23 embryos. The re-estimated *k* values are plotted against those derived from all the embryos. Three coordinates are labeled in different colors. The diagonal line is shown for reference.

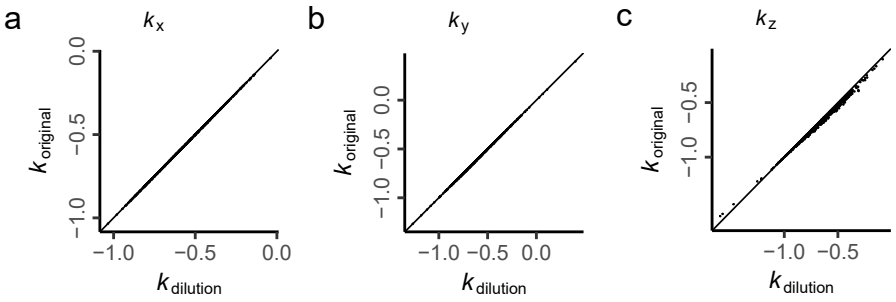

Fig. S7. Robustness of *k* agains dilution effect
Each panel in this figure compares the *k* values estimated by considering the dilution effect from measurement precision with those estimated without considering this effect. The *k* values for the three coordinates are displayed separately, and a diagonal line is shown for reference in each panel.

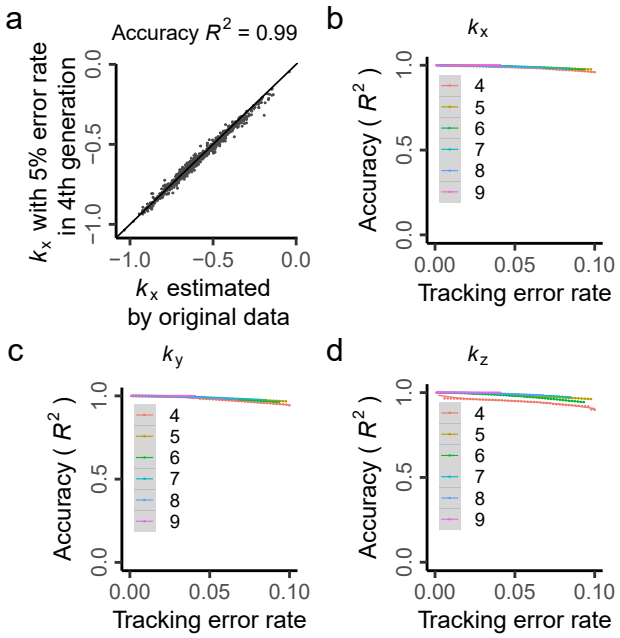

Fig. S8. The robustness of *k* against lineage tracking accuracy
Accuracy is defined as the square of Pearson's *R* between the original *k* values and the *k* values re-estimated after introducing tracking error. When tracking error is introduced into a specific generation, the error rate is calculated as the ratio of the number of misidentified cells to the total cell count.
(a) For example, we introduced a 5% error in the 4th generation for the x coordinate. Specifically, we randomly selected 5% of mother cells in the 4th generation across all embryos and interchanged the two daughter cells for each of these selected mother cells, as well as their descendants in subsequent generations. This corresponds to a 5% error rate. We then re-estimated k_x based on the data with this error. The k_x based on original data and the k_x estimated with a 5% error rate are highly correlated, resulting in an accuracy of R^2=0.995.
(b-d) Display the relationship between accuracy and error rate when tracking errors are introduced into different generations (mother generation: 4th-9th) for *x*, *y*, and *z* coordinates, respectively. The fitted lines are obtained by loess regression, and the grey bands represent the 95% confidence interval. As the previous study demonstrated, the tracking error rate in RNAi embryos is less than 1%.

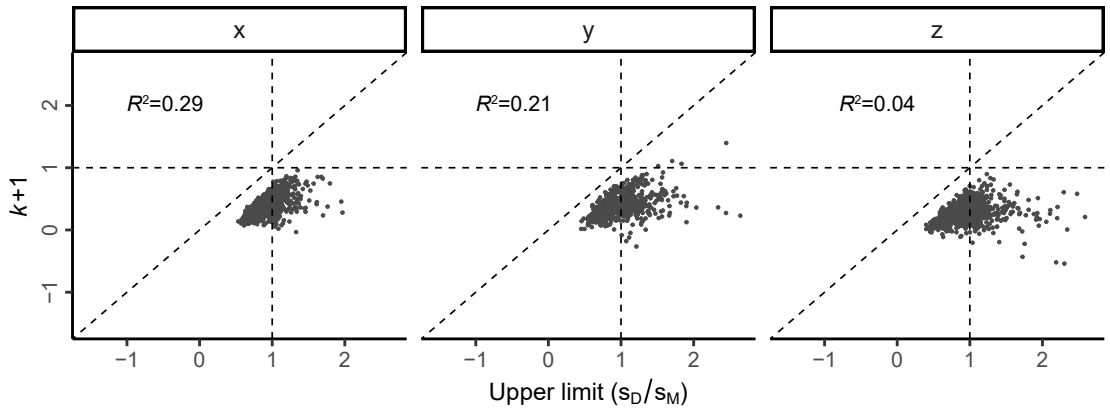

Fig. S9. The robustness of *k* agains the ratio of σ_D/σ_M
In linear regression, if the independent variable has a smaller standard deviation (SD) than the dependent variable, a slope smaller than 1 will always be obtained due to the relationship between slope and correlation, $k=\sigma_D/\sigma_M \times R$. The horizontal axis shows the σ_D/σ_M for each mother-daughter cell pair at three coordinates, respectively. The vertical dashed line signifies $\sigma_D/\sigma_M=1$ in each panel, indicating that the SD is overall comparable between mother and daughter cells. The vertical axis is $a=k+1$, and the horizontal dashed line corresponds to $k=0$ in each panel. The diagonal line is shown in each panel. Given that σ_D/σ_M is the upper limit of $a=k+1$, all the points are distributed below the diagonal line. The small R^2 in each panel suggests that $a=k+1$ cannot be simply explained by their upper limits.

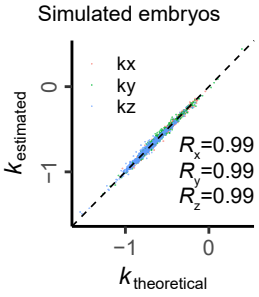

Fig. S10. The robustness of *k* against alignment procedure
This figure compares the theoretical and observed *k* values of the three coordinates in simulated embryos. The three coordinates are labeled in different colors. The diagonal line and Pearson's *R* for each coordinate are shown.

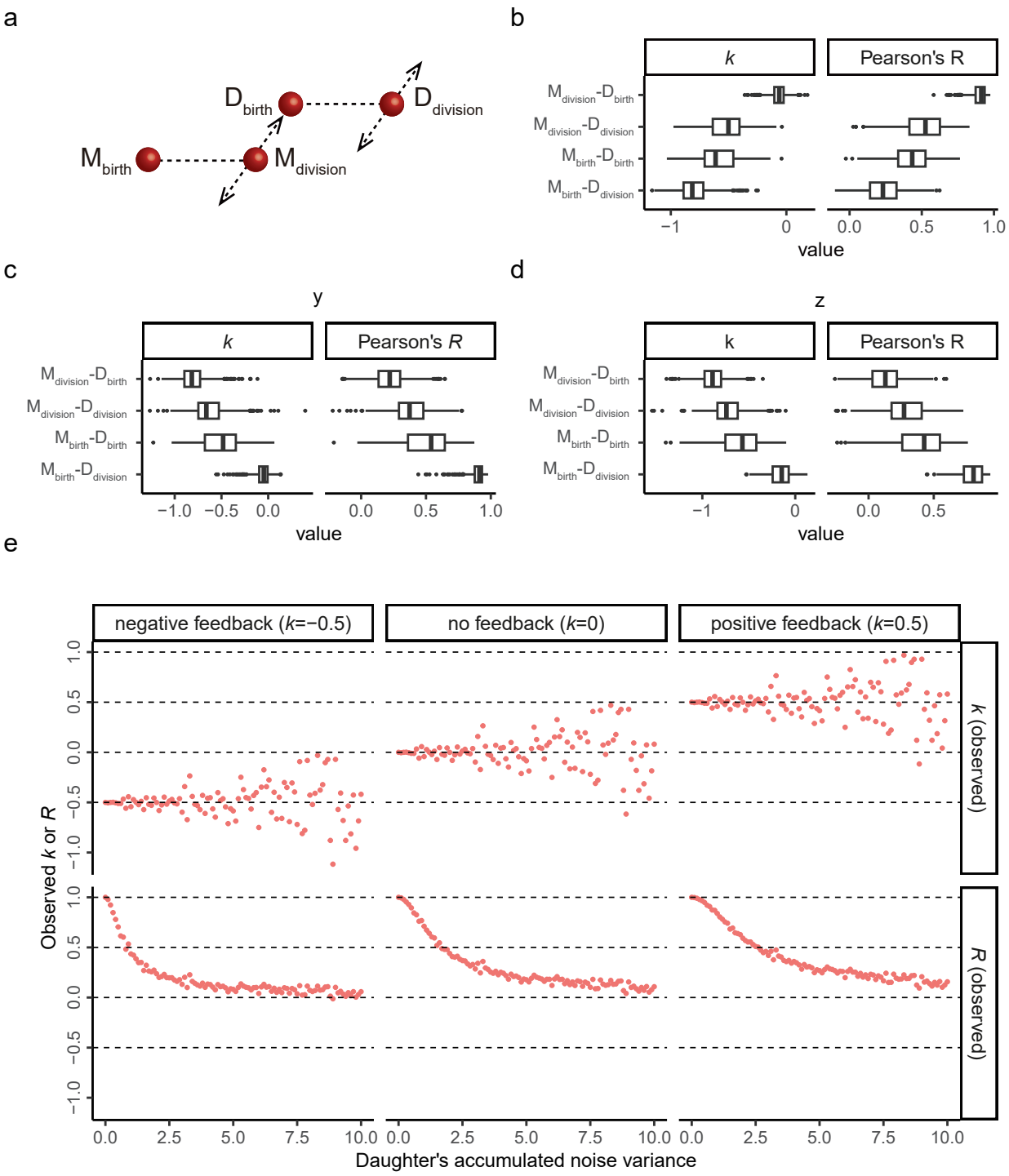
a
b
c
d
e
k
Pearson's R
value
y
z
negative feedback (k=−0.5)
no feedback (k=0)
positive feedback (k=0.5)
k (observed)
R (observed)
Observed k or R
Daughter's accumulated noise variance

Fig. S11. Negative feedback across different cell stages
(a) In addition to the birth position (the default considered in this study), the division position of each cell can also be analyzed. As such, for each M-D cell pair there are four combinations to be examined.
(b-d) The distributions of k and R values of all M-D cell pairs obtained for the four combinations described in the panel a. The position noises of x, y and z coordinates are shown, respectively.
(e) The Effect of Daughter's Noise Accumulation on *k* and *R*. This panels illustrates how the noise accumulation in daughter cells can impact the estimated *k* values and Pearson's *R*, based on mother's and daughter's noise. The process is simulated as follows: First, we generate 1,000 instances of mother's noise from a standard normal distribution (N(0,1)). Then, the accumulated noise ε is generated from N(0,1) and multiplied by a factor denoting the variance level, ranging from 0 to 10. Next, the daughter's noise is calculated using $X_D = (k+1)X_M + \varepsilon$, considering three different *k* values that correspond to negative feedback, no feedback, and positive feedback, respectively. For each scenario, we estimate the *k* values. The results show that the observed *R* between mother's and daughter's noise gradually decreases to zero as the variance of the daughter's accumulated noise increases. However, the expected values of the observed *k* remain constant, while their variation increases. Therefore, the negative *k* values we obtained cannot simply be attributed to the accumulation of noise in the daughter cells. Considering that the daughter at birth inherits the main variance of its mother, the negative *k* values demonstrate the presence of negative feedback.

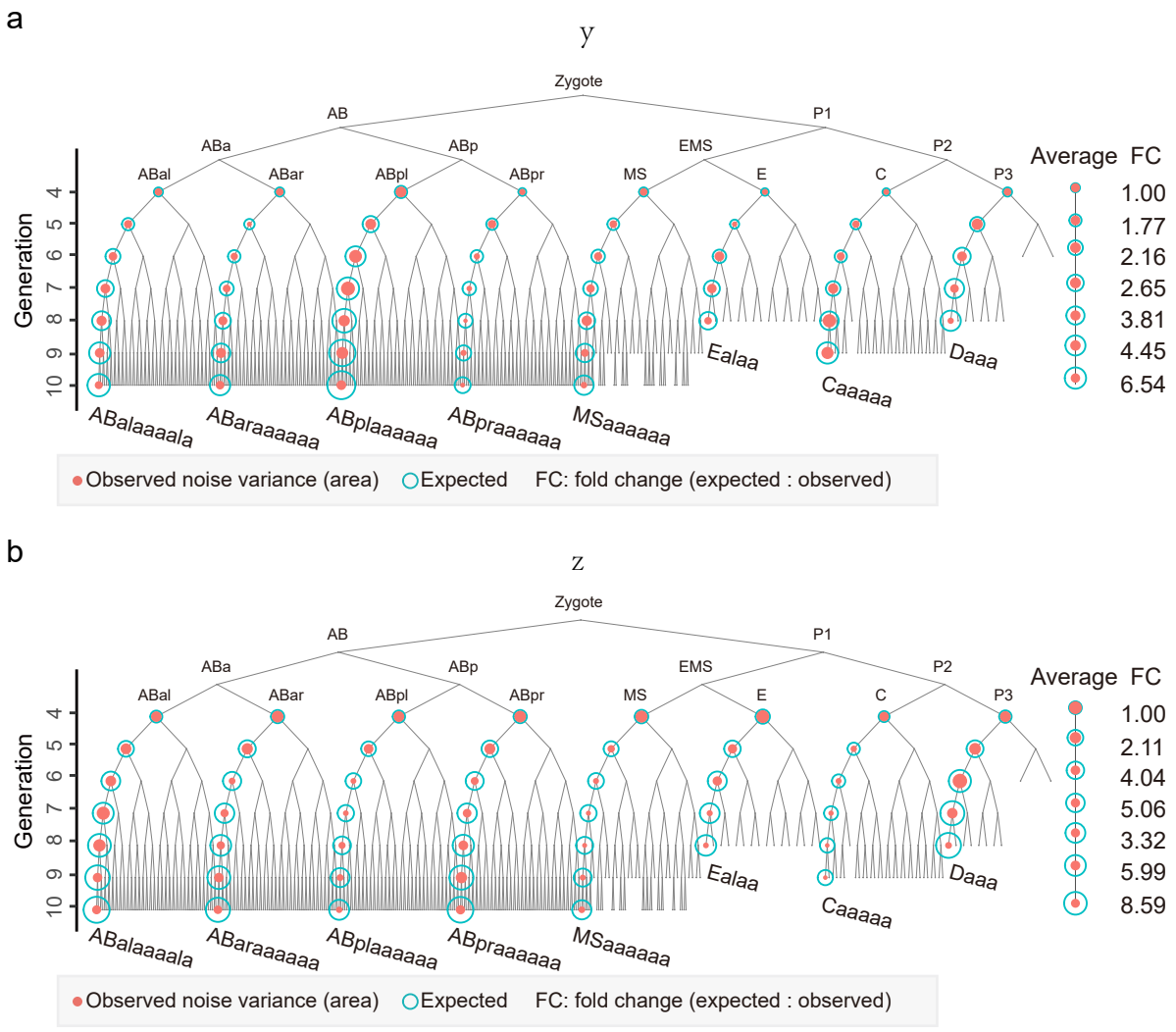

Fig. S12. Expected and observed noise variance along lineages.
This figure presents the results for the y-coordinate and z-coordinate, corresponding to Fig. 2f. The theoretical and observed noise variances along different lineages are compared for these two spatial coordinates. Each lineage's theoretical noise variance is calculated based on the noise variance of the initial cell and the estimated daughter-specific noise variance. The observed noise variance is directly measured from the data.

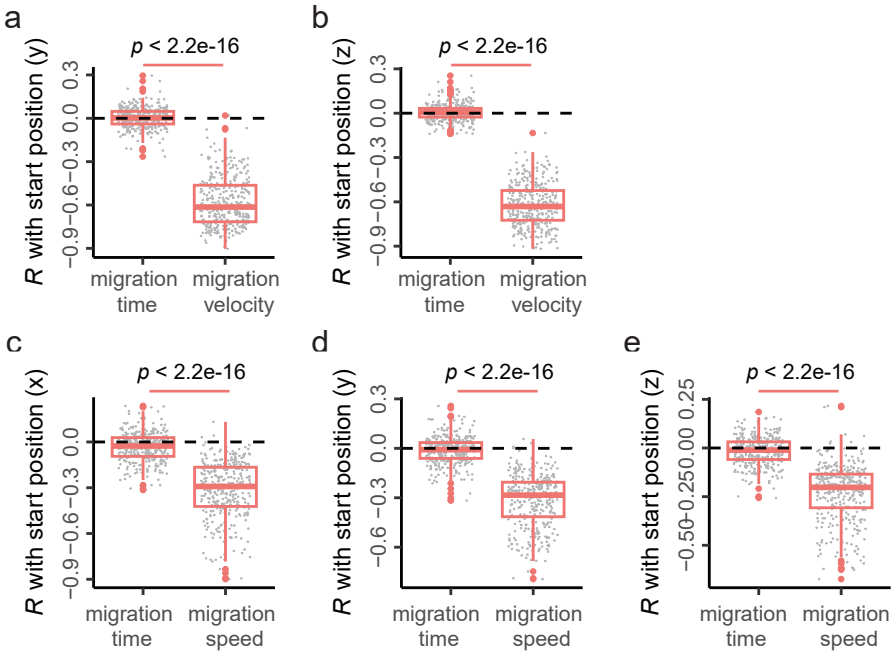

Fig. S13. Cell velocity mediates negative-feedback regulation of positional noise
(a-b) The same analysis with Fig. 4h for *y* and *z* coordinates, respectively. The *R* between start position and velocity for *x-*, *y-* and *z*-coordinates is almost always negative, significantly different from that between start position and time.
(c-e) Correlation anlaysis is conducted on start position (direction adjusted) and migration speed or migration time. Others are the same with analysis on migration velocity.

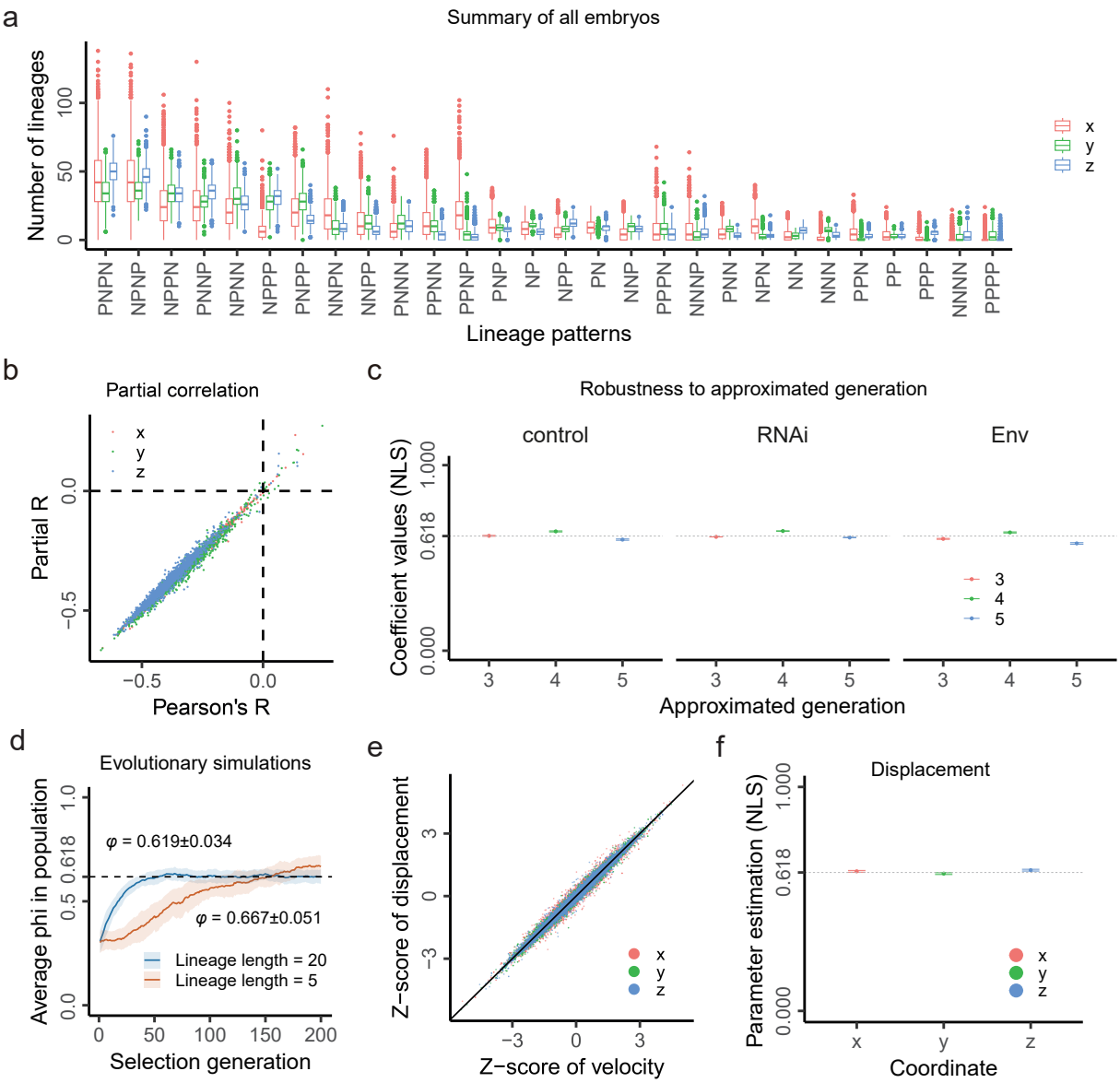

Fig. S14. Extended analysis for the self-similarity equation of cell velocity
(a) The cell velocity noise dynamics across the 5 generations can form various lineage patterns. We define a positive velocity change from one generation to the next as “P” while a negative change as “N”. The number of lineages for different lineage patterns are counted for each embryo and considered for three coordinates, respectively.
(b) Partial correlation analysis with position information as a confounding factor, show a consistent negative relationship between M-D pairs.
(c) In main text, we have used 3-generation approximation for the GDL self-similar equation. In this panel, we also conducted φ-estimation for 4-generation and 5-generation approximation. Results show robust estimation around 0.618.
(d) Evolutionary simulations for the evolution of φ with a lineage length equal to 5 or 20.
(e) Comparison between Z-scores of velocity and displacement based on 105 controls.
(f) Estimated φ for the same form self-similarity equation for displacement as velocity.

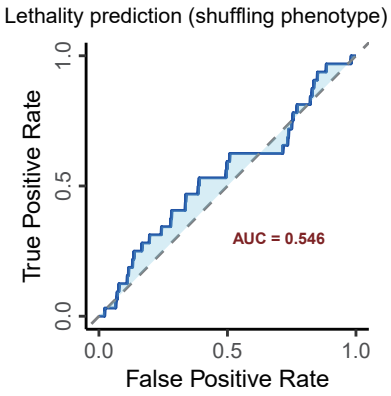

Fig. S15. Shuffling analysis for predicting hatching lethality phenotype
Logistic regression predicting shuffled hatching outcome using MAEs from the three coordinates as predictors. Seventy percent of embryos are used for training and the remainder for testing to calculate the AUC.

$$\begin{cases} V_{\mathrm{D1}} = -\varphi^2 V_{\mathrm{M}} - \varphi \sum_{i=2}^{t-1} \varphi^i V_{t-i} + \varepsilon_{\mathrm{D1}} \\ V_{\mathrm{D2}} = -\varphi^2 V_{\mathrm{M}} - \varphi \sum_{i=2}^{t-1} \varphi^i V_{t-i} + \varepsilon_{\mathrm{D2}} \end{cases}$$

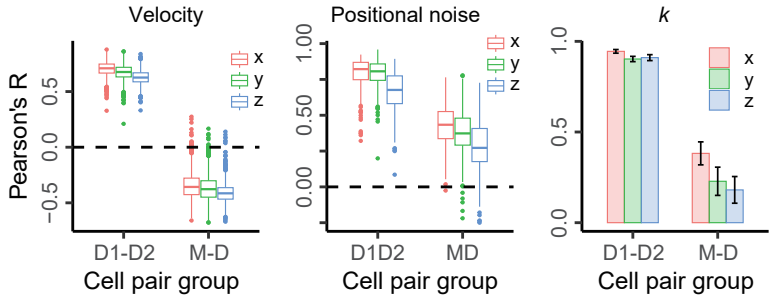

Fig. S16. Self-similarity equation explains mother-daughter similarity paradox
The mother–sister similarity paradox for cell velocity, positional noise, and negative-feedback strength *k*. Pearson's correlations are compared between sister pairs (D1–D2) and mother–daughter pairs (M–D), respectively.

Supplementary Note I

The convergence of cell noise level along lineage under negative feedback

1. Introduction

In this study, we utilized a Cartesian coordinate system to map the positions of individual cells within each embryo. We examined embryos subjected to various disturbances and discovered that the cells with the same identity within these embryos exhibit consistent expected positions after appropriate alignment. This allowed us to define the cell noise for each individual cell in each embryo, relative to the expected position. By analyzing the variation of cell noise across embryos, we quantitatively modeled the relationship between the noise of mother cells and that of their daughter cells.

As a result, we derived a metric, denoted as k, which represents the slope between mother cell noise and the noise difference between mother and daughter cells. This metric helps describe the sign and strength of feedback in the transmission of noise from mother to daughter cells. Almost all k values were negative with four exceptions at the y-coordinate, indicating a ubiquitous negative feedback mechanism in mother-daughter cell noise transmission. In this supplementary note, we will provide a detailed mathematical derivation of the significant property associated with the widespread negative feedbacks: the convergence of cell noise levels along the lineage. Notably, the convergence in this study denotes that cell noise levels along lineage will not continuously increase with generation but instead remain no more than an upper limit.

2. Model and evaluation

Consider an iterative model represented by the equation:

$$X_t = a_t X_{t-1} + \varepsilon_t \tag{1}$$

where X_t denotes the cell noise at the t-th generation of a cell lineage; $a = k+1$ represents

the expected proportion of cell noise inherited by a daughter cell from its mother; and ε_t is the cell-specific noise of the cell at the t-th generation along a lineage, assumed to be independent of the mother cell noise X_{t-1}. For the initial mother-daughter cell pair, we have: $X_1=a_1X_0+\varepsilon_1$. When we express formula (1) iteratively, we can derive the following formula:

$$\begin{aligned} X_t &= a_t X_{t-1} + \varepsilon_t \\ &= a_t(a_{t-1}X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= a_t a_{t-1} X_{t-2} + a_t\varepsilon_{t-1} + \varepsilon_t \\ &= \ldots \\ &= (\prod_{i=1}^{i=t} a_i)X_0 + \sum_{j=1}^{j=t-1}(\prod_{i=j+1}^{i=t} a_i)\varepsilon_j + \varepsilon_t \end{aligned} \tag{2}$$

Considering the random properties of cell noise, we use the expected variance, $\mathrm{Var}(X_t)$, to evaluate the cell noise level accumulated along cell lineage. For convenience, we set $\mathrm{Var}(X_0) = S_{X_0}^2$ and $\mathrm{Var}(\varepsilon_i) = S_{\varepsilon_i}^2$. Assuming that X_0 and ε_i are independent. We can estimate the variance of X_t as follows:

$$\begin{aligned} \mathrm{Var}(X_t) &= (\prod_{i=1}^{i=t} a_i)^2\,\mathrm{Var}(X_0) + \sum_{j=1}^{j=t-1}(\prod_{i=j+1}^{i=t} a_i)^2\,\mathrm{Var}(\varepsilon_j) + \mathrm{Var}(\varepsilon_t) \\ &= (\prod_{i=1}^{i=t} a_i)^2 S_{X_0}^2 + \sum_{j=1}^{j=t-1}(\prod_{i=j+1}^{i=t} a_i)^2 S_{\varepsilon_j}^2 + S_{\varepsilon_t}^2 \end{aligned} \tag{3}$$

3. Theoretical analysis for a constant *k*

3.1. *k* is a number satisfying -1 < *k* < 0

In this situation, $a_i = a = k+1$ is a constant value between 0 and 1. We can rewrite formula (3) as:

$$\mathrm{Var}(X_t) = a^{2t} S_{X_0}^2 + \sum_{j=1}^{j=t} a^{2(t-j)} S_{\varepsilon_j}^2 \tag{4}$$

Assuming that there is an upper limit and an lower limit for $\{S_{X_0}^2, S_{\varepsilon_i}^2\}$, denoted as S_u^2 and S_l^2, respectively, we have:

$$\begin{aligned}\mathrm{Var}(X_t)_u &= a^{2t}S_u^2 + \sum_{j=1}^{j=t} a^{2(t-j)}S_u^2 = a^{2t}S_u^2 + \frac{1-a^{2t}}{1-a^2}S_u^2 = \frac{1}{1-a^2}S_u^2 - (\frac{a^{2(t+1)}}{1-a^2})S_u^2 \\ \mathrm{Var}(X_t)_l &= a^{2t}S_l^2 + \sum_{j=1}^{j=t} a^{2(t-j)}S_l^2 = a^{2t}S_l^2 + \frac{1-a^{2t}}{1-a^2}S_l^2 = \frac{1}{1-a^2}S_l^2 - (\frac{a^{2(t+1)}}{1-a^2})S_l^2\end{aligned} \quad (5)$$

From this formula, we can obtain that both of the upper limit ($\mathrm{Var}(X_t)_u$) and lower limit ($\mathrm{Var}(X_t)_l$) of noise level along lineage are converged with generation. The limit forms are given as:

$$\begin{aligned}\mathrm{Var}(X_t)_u &= \frac{1}{1-a^2}S_u^2 = \frac{1}{1-(k+1)^2}S_u^2 \\ \mathrm{Var}(X_t)_l &= \frac{1}{1-a^2}S_l^2 = \frac{1}{1-(k+1)^2}S_l^2\end{aligned} \quad (6)$$

Therefore, noise level along lineage is convergent at the situation of $-1 < k < 0$.

3.2. *k* is a number satisfying -2 < *k* < -1

In this situation, $a_i = a = k+1$ is a constant value between -1 and 0. We can still obtain the same convergent limits as formula (6). Therefore, noise level along lineage is convergent at the situation of $-2 < k < -1$.

3.3. *k* = -1

In this situation, $a_i = a = k+1=0$. We will obtain a trivial form from formula (3) as:

$$\mathrm{Var}(X_t) = S_{\varepsilon_t}^2 \quad (7)$$

Therefore, the noise level will be no larger than S_u^2. Therefore, noise level along lineage is convergent at the situation of $k = -1$.

3.4. *k* is a number satisfying *k* > 0 or *k* < -2

At these situations, $a_i = a = k+1$ has an absolute value larger than 1. Then, the lower limit $\mathrm{Var}(X_t)_l$ in formula (5) will keep increasing with t increasing. Therefore, we can't deduce the convergent limits. Therefore, noise level along lineage is divergent at the situation of $k > 0$ or $k < -2$.

3.5. $k = 0$ or $k = -2$

At these situations, $a_i = a = k+1$ has an absolute of 1. We can obtain the trivial forms from formula (4) as:

$$\begin{aligned} \mathrm{Var}(X_t)_u &= (t+1)S_u^2 \\ \mathrm{Var}(X_t)_l &= (t+1)S_l^2 \end{aligned} \tag{8}$$

Since $\mathrm{Var}(X_t)_l$ increases with t increasing, there is not limit forms. Therefore, noise level along lineage is divergent at the situation of $k = 0$ or $k = -2$.

3.6. Special case where ε_t follows the standard normal distribution

This part corresponds to the theoretical analysis for the convergence of noise accumulation along lineage through a continuous negative-feedback canal. At this special case, the accumulated variance is simplified as: Var $(X_t) = (1-a^{2t})/(1-a^2)$. When $a^2 = 0.5$ (corresponding to $k = a - 1 = -0.29$), the limit of Var (X_t) equals to 2. When $a^2 = 1.5$ ($k = a - 1 = 0.22$) or $a = 1$, Var (X_t) diverges with generation.

4. Theoretical analysis for variational k

The above situations are simplified by analyzing a constant k value. In practice, we observed that the k values vary with cell pairs. Assuming that there are an upper limit and a lower limit for the absolute of a_i, named as a_u and a_l, respectively. We can simply update the above results.

4.1. $-2 < k < 0$

In this situation, $a_i = k_i+1$ satisfies $-1 < a_i < 1$ or $|a_i| < 1$. First, consider $k_i \neq -1$ or $a_i \neq 0$. We can simply update the results of formula (6) as:

$$\begin{aligned} \mathrm{Var}(X_t)_u &= \frac{1}{1-a_u^2} S_u^2 \\ \mathrm{Var}(X_t)_l &= \frac{1}{1-a_l^2} S_l^2 \end{aligned} \tag{9}$$

Then, consider some of k_i satisfy $k_i = -1$, equivalent with $a_i = 0$, which can be formulated

as $X_t = a_iX_{t-1} + \varepsilon_t = \varepsilon_t$. Therefore, for these cell pairs, actually no noise is inherited from mother to daughter. Assuming that the number of cell pairs with $k_i = -1$ equals to n, the actual iteration times turns to be t-n. We can update formula (5) as:

$$\begin{aligned} \mathrm{Var}(X_t)_u &= \frac{1}{1-a^2}S_u^2 - (\frac{a^{2(t-n+1)}}{1-a^2})S_u^2 \\ \mathrm{Var}(X_t)_l &= \frac{1}{1-a^2}S_l^2 - (\frac{a^{2(t-n+1)}}{1-a^2})S_l^2 \end{aligned} \tag{10}$$

When n accounts for a small proportion, denoted as $r = n / t$. Then, we have t-$n = (1-r)$ t. Therefore, with t increasing, we have the same limit forms as formula (9). Therefore, noise level along lineage is convergent when k varies, satisfying $-2 < k < 0$.

4.2. $k \geq 0$ or $k \leq -2$

In this situation, $|a_i| \geq 1$. For $|a_i| > 1$, we can update the lower limit in formula (5) as:

$$\mathrm{Var}(X_t)_l = \frac{1}{1-a_l^2}S_l^2 - (\frac{a_l^{2(t+1)}}{1-a_l^2})S_l^2 \tag{11}$$

For $|a_i| = 1$, the formula (8) remains. Therefore, the noise level will keep divergent at this situation. Therefore, noise level along lineage is divergent when k varies, satisfying $k \geq 0$ or $k \leq -2$.

4.3. Mixed situation

Consider the mixed situation where both $-2 < k < 0$ and $k \geq 0$. This situation means the cell noise level will diverge after transmission from mother cell to daughter cell for some cell pairs (divergent cell pairs) but convergent for the other cell pairs (convergent cell pairs). If all the k values satisfy $-2 < k < 0$ after certain generation, the cell noise level will continue to be convergent from this generation. However, previous noise accumulation in the divergent cell pairs could cause noise level along lineage to surpass certain noise tolerance level of specific cells, leading to developmental failure. Therefore, in a typical multicellular organism, it's expected that negative-feedback regulation on cell noise should be a general mechanism used by almost all cell pairs, which is consistent with the results in this study.

5. Conclusion and discussion

First, noise level increases along lineage with the absolute of a_i, the noise level of initial cell and those of cell-specific noises. Without considering other restrictions like spatial restrictions, cell noise level increases with generation except that all cell pairs satisfy $k = -1$ (no inheritance for all cell pairs). However, this is distinct from the real observation. Additionally, the upper limit and the lower limit are used to determine whether cell noise along lineage is convergent or diverge, whereas they might not be the exact attainable values.

Second, in a typical multicellular organism like C. elegans with multiple generations in each lineage, $k < 0$ (negative feedback) should be satisfied for almost all cell pairs to ensure the convergence of cell noise level, while $k \geq 0$ (positive feedback) in too many cell pairs could cause accumulated cell noise to surpass the noise tolerance level of specific cells. In much simpler multicellular organisms, a larger proportion of $k \geq 0$ could be permitted, which is an interesting hypothesis to be tested in future studies.

Third, in this study, we observed that most of cell pairs satisfy $-1 < k < 0$ and a minority of cell pairs satisfy $-2 < k \leq -1$. Here, $-2 < k < -1$ can be regarded as an over-strong negative-feedback regulation, which means that mother's noise has an extra penalty on daughter's noise, reflecting a trade-off between mother and daughter's cell noise. We didn't find $k \leq -2$ for any cell pairs, which could mean a mathematical but not practical situation. Therefore, we have simplified the expression in this study as: positive feedback denoted by $k > 0$; no feedback denoted by $k = 0$; and negative feedback denoted by $k < 0$ (actually, $-2 < k < 0$).

Supplementary Note II

Corr (D, V) ≈ 1 when CV(T) ≪ CV(D) or CV(T) ≪ CV(V)

1. Introduction

The objective is to determine whether the extremely high correlation between displacement (D) and velocity (V) can be explained by the fact that the coefficient of variation (CV) of time (T) is much smaller than that of displacement or velocity. Specifically, we provide mathematical proofs for two scenarios: Corr (D, V) ≈ 1 when CV(T) ≪ CV(D), and Corr (D, V) ≈ 1 when CV(T) ≪ CV(V). In both cases, the derivations yield the same mathematical form, demonstrating that the near-perfect correlation between displacement and velocity is an expected consequence of the much smaller variability in time.

2. Corr (D, V) ≈ 1 when CV(T) ≪ CV(D)

2.1. Problem Setting

Let D be a random variable representing a component of displacement (can be positive or negative) and $T > 0$ be time. The ratio $V = \frac{D}{T}$ represents the corresponding component of velocity. Assume D and T have finite second moments, with means μ_D (assumed nonzero) and $\mu_T > 0$, standard deviations σ_D and σ_T, and correlation $\rho = \mathrm{Corr}(D, T)$.

Define the **coefficients of variation**:

$$CV(D) = \frac{\sigma_D}{|\,\mu_D\,|}, CV(T) = \frac{\sigma_T}{\mu_T}. \quad (1)$$

(For simplicity, we assume $\mu_D > 0$; if $\mu_D < 0$, the analysis follows analogously by considering $|\,\mu_D\,|$ in the expansion.)

Given: $CV(D) \gg CV(T)$.

Example: $CV(D) \approx 26,\ CV(T) \approx 0.26$, so $\frac{CV(T)}{CV(D)} \approx 0.01$.

Goal: Show that under these conditions, $\text{Corr}\left(D, \frac{D}{T}\right) \approx 1$.

2.2. First-Order Taylor Expansion (Delta Method)

Let $V = \frac{D}{T}$. Expand V around (μ_D, μ_T):

$$V \approx \frac{\mu_D}{\mu_T} + \frac{1}{\mu_T}(D - \mu_D) - \frac{\mu_D}{\mu_T^2}(T - \mu_T). \tag{2}$$

2.3. Covariance and Variance Approximations

Using the linear approximation:

- **Covariance:**

$$\text{Cov}(D, V) \approx \frac{1}{\mu_T}\sigma_D^2 - \frac{\mu_D}{\mu_T^2}\rho\sigma_D\sigma_T. \tag{3}$$

- **Variance of V:**

$$\text{Var}(V) \approx \frac{1}{\mu_T^2}\sigma_D^2 + \frac{\mu_D^2}{\mu_T^4}\sigma_T^2 - 2\frac{\mu_D}{\mu_T^3}\rho\sigma_D\sigma_T. \tag{4}$$

2.4. Expression in Terms of Coefficients of Variation

Substitute $\sigma_D = CV(D)\mu_D$ and $\sigma_T = CV(T)\mu_T$:

$$\begin{aligned} \text{Cov}(D, V) &\approx \frac{\mu_D^2}{\mu_T}(CV(D)^2 - \rho\, CV(D)\, CV(T)), \\ \text{Var}(V) &\approx \frac{\mu_D^2}{\mu_T^2}(CV(D)^2 + CV(T)^2 - 2\rho\, CV(D)\, CV(T)), \\ \text{Var}(D) &= CV(D)^2\mu_D^2. \end{aligned} \tag{5}$$

2.5. Correlation Formula (Within Linear Approximation)

$$\begin{aligned} \text{Corr}(D, V) &\approx \frac{\text{Cov}(D, V)}{\sqrt{\text{Var}(D)\,\text{Var}(V)}} \\ &= \frac{CV(D)^2 - \rho\, CV(D)\, CV(T)}{CV(D)\sqrt{CV(D)^2 + CV(T)^2 - 2\rho\, CV(D)\, CV(T)}} \\ &= \frac{CV(D) - \rho\, CV(T)}{\sqrt{CV(D)^2 + CV(T)^2 - 2\rho\, CV(D)\, CV(T)}}. \end{aligned} \tag{6}$$

2.6. Asymptotic Expansion for $CV(D) \gg CV(T)$

Let $\varepsilon = \frac{CV(T)}{CV(D)} \ll 1$. Then:

$$\text{Corr}(D, V) = \frac{1 - \rho\varepsilon}{\sqrt{1 + \varepsilon^2 - 2\rho\varepsilon}}. \tag{7}$$

2.7. Taylor Expansion to Second Order in ε

Expand the denominator using $\sqrt{1+u} = 1 + \frac{1}{2}u - \frac{1}{8}u^2 + \cdots$ with $u = \varepsilon^2 - 2\rho\varepsilon$:

$$\sqrt{1 + \varepsilon^2 - 2\rho\varepsilon} \quad = 1 - \rho\varepsilon + \frac{1}{2}(1 - \rho^2)\varepsilon^2 + O(\varepsilon^3). \tag{8}$$

Then, using $(1+v)^{-1} = 1 - v + v^2 - \cdots$ with $v = -\rho\varepsilon + \frac{1}{2}(1-\rho^2)\varepsilon^2 + O(\varepsilon^3)$:

$$\frac{1}{\sqrt{1 + \varepsilon^2 - 2\rho\varepsilon}} \quad = 1 + \rho\varepsilon + \left(\frac{3}{2}\rho^2 - \frac{1}{2}\right)\varepsilon^2 + O(\varepsilon^3). \tag{9}$$

Multiplying by the numerator $(1-\rho\varepsilon)$:

$$\begin{aligned} \text{Corr}(D, V) \quad &= (1 - \rho\varepsilon)\left[1 + \rho\varepsilon + \left(\frac{3}{2}\rho^2 - \frac{1}{2}\right)\varepsilon^2 + O(\varepsilon^3)\right] \\ &= 1 - \frac{1}{2}(1 - \rho^2)\varepsilon^2 + O(\varepsilon^3). \end{aligned} \tag{10}$$

Substituting $\varepsilon = \frac{CV(T)}{CV(D)}$ yields the final result:

$$\boxed{\text{Corr}\left(D, \frac{D}{T}\right) = 1 - \frac{1}{2}(1 - \rho^2)\left(\frac{CV(T)}{CV(D)}\right)^2 + O\left(\left(\frac{CV(T)}{CV(D)}\right)^3\right)}. \tag{11}$$

2.8. Interpretation and Numerical Example

- The correlation deviates from 1 only by a term of order $\left(\frac{CV(T)}{CV(D)}\right)^2$, which is negligible when $CV(D) \gg CV(T)$.
- The result is independent of the sign of μ_D (displacement direction) and holds for any correlation ρ between D and T.
- Given $CV(D) \approx 26$ and $CV(T) \approx 0.26$, we have $\frac{CV(T)}{CV(D)} \approx 0.01$, so $\left(\frac{CV(T)}{CV(D)}\right)^2 \approx 10^{-4}$. The maximum deviation from 1 occurs when $\rho = 0$,

yielding $\text{Corr}(D,V) \approx 1 - 0.5 \times 10^{-4} = 0.99995$.

Thus, under the given conditions, displacement D and velocity $V = \frac{D}{T}$ are essentially perfectly correlated.

3. Corr (D, V) ≈ 1 when CV(T) ≪ CV(V)

3.1. Problem Setting

Let V be a random variable representing a component of velocity (can be positive or negative) and $T > 0$ be time. The product $D = V \cdot T$ represents the corresponding component of displacement. Assume V and T have finite second moments, with means μ_V (assumed nonzero) and $\mu_T > 0$, standard deviations σ_V and σ_T, and correlation $\rho = \text{Corr}(V,T)$.

Define the **coefficients of variation**:

$$CV(V) = \frac{\sigma_V}{|\,\mu_V\,|}, CV(T) = \frac{\sigma_T}{\mu_T}. \tag{12}$$

(For simplicity, assume $\mu_V > 0$; the analysis is analogous if $\mu_V < 0$.)

Given: $CV(V) \gg CV(T)$.

Example: $CV(V) \approx 26$, $CV(T) \approx 0.26$, so $\frac{CV(T)}{CV(V)} \approx 0.01$.

Goal: Show that under these conditions, $\text{Corr}(V, V \cdot T) \approx 1$.

3.2. First-Order Taylor Expansion (Delta Method)

Let $D = V \cdot T$. Expand D around (μ_V, μ_T):

$$D \approx \mu_V\mu_T + \mu_T(V - \mu_V) + \mu_V(T - \mu_T). \tag{13}$$

3.3. Covariance and Variance Approximations

Using the linear approximation:

- **Covariance:**

$$\text{Cov}(V,D) \approx \mu_T\text{Cov}(V,V) + \mu_V\text{Cov}(V,T) = \mu_T\sigma_V^2 + \mu_V\rho\sigma_V\sigma_T. \tag{14}$$

- **Variance of D:**

$$\mathrm{Var}(D) \approx \mu_T^2\sigma_V^2 + \mu_V^2\sigma_T^2 + 2\mu_V\mu_T\rho\sigma_V\sigma_T. \tag{15}$$

3.4. Expression in Terms of Coefficients of Variation

Substitute $\sigma_V = CV(V)\mu_V$ and $\sigma_T = CV(T)\mu_T$ (with $\mu_V > 0$):

$$\begin{aligned}
\mathrm{Cov}(V,D) &\approx \mu_T(CV(V)^2\mu_V^2) + \mu_V\rho(CV(V)\mu_V)(CV(T)\mu_T) \\
&= \mu_V^2\mu_T(CV(V)^2 + \rho\, CV(V)\, CV(T)), \\
\mathrm{Var}(D) &\approx \mu_T^2(CV(V)^2\mu_V^2) + \mu_V^2(CV(T)^2\mu_T^2) + 2\mu_V\mu_T\rho(CV(V)\mu_V)(CV(T)\mu_T) \\
&= \mu_V^2\mu_T^2(CV(V)^2 + CV(T)^2 + 2\rho\, CV(V)\, CV(T)), \\
\mathrm{Var}(V) &= CV(V)^2\mu_V^2.
\end{aligned} \tag{16}$$

3.5. Correlation Formula (Within Linear Approximation)

$$\begin{aligned}
\mathrm{Corr}(V,D) &\approx \frac{\mathrm{Cov}(V,D)}{\sqrt{\mathrm{Var}(V)\,\mathrm{Var}(D)}} \\
&= \frac{\mu_V^2\mu_T(CV(V)^2 + \rho\, CV(V)\, CV(T))}{\sqrt{CV(V)^2\mu_V^2 \cdot \mu_V^2\mu_T^2(CV(V)^2 + CV(T)^2 + 2\rho\, CV(V)\, CV(T))}} \\
&= \frac{CV(V)^2 + \rho\, CV(V)\, CV(T)}{CV(V)\sqrt{CV(V)^2 + CV(T)^2 + 2\rho\, CV(V)\, CV(T)}} \\
&= \frac{CV(V) + \rho\, CV(T)}{\sqrt{CV(V)^2 + CV(T)^2 + 2\rho\, CV(V)\, CV(T)}}.
\end{aligned} \tag{17}$$

3.6. Asymptotic Expansion for $\boldsymbol{CV(V) \gg CV(T)}$

Let $\varepsilon = \frac{CV(T)}{CV(V)} \ll 1$. Then:

$$\mathrm{Corr}(V,D) = \frac{1 + \rho\varepsilon}{\sqrt{1 + \varepsilon^2 + 2\rho\varepsilon}}. \tag{18}$$

3.7. Taylor Expansion to Second Order in $\boldsymbol{\varepsilon}$

Expand the denominator using $\sqrt{1+u} = 1 + \frac{1}{2}u - \frac{1}{8}u^2 + \cdots$ with $u = \varepsilon^2 + 2\rho\varepsilon$:

$$\sqrt{1 + \varepsilon^2 + 2\rho\varepsilon} \quad = 1 + \rho\varepsilon + \frac{1}{2}(1-\rho^2)\varepsilon^2 + O(\varepsilon^3). \tag{19}$$

Then, using $(1+v)^{-1} = 1 - v + v^2 - \cdots$ with $v = \rho\varepsilon + \frac{1}{2}(1-\rho^2)\varepsilon^2 + O(\varepsilon^3)$:

$$\frac{1}{\sqrt{1 + \varepsilon^2 + 2\rho\varepsilon}} \quad = 1 - \rho\varepsilon + \left(\frac{3}{2}\rho^2 - \frac{1}{2}\right)\varepsilon^2 + O(\varepsilon^3). \tag{20}$$

Multiplying by the numerator $(1+\rho\varepsilon)$:

$$\begin{aligned}\text{Corr}(V,D) &= (1+\rho\varepsilon)\left[1-\rho\varepsilon+\left(\frac{3}{2}\rho^2-\frac{1}{2}\right)\varepsilon^2+O(\varepsilon^3)\right] \\ &= 1-\frac{1}{2}(1-\rho^2)\varepsilon^2+O(\varepsilon^3).\end{aligned} \tag{21}$$

Substituting $\varepsilon = \frac{CV(T)}{CV(V)}$ yields the final result:

$$\boxed{\text{Corr}(V,V\cdot T) = 1-\frac{1}{2}(1-\rho^2)\left(\frac{CV(T)}{CV(V)}\right)^2+O\left(\left(\frac{CV(T)}{CV(V)}\right)^3\right)}. \tag{22}$$

3.8. Interpretation and Numerical Example

- The correlation deviates from 1 only by a term of order $\left(\frac{CV(T)}{CV(V)}\right)^2$, which is negligible when $CV(V) \gg CV(T)$.
- The result holds for any correlation ρ between V and T.
- Given $CV(V) \approx 26$ and $CV(T) \approx 0.26$, we have $\frac{CV(T)}{CV(V)} \approx 0.01$, so $\left(\frac{CV(T)}{CV(V)}\right)^2 \approx 10^{-4}$. The maximum deviation from 1 occurs when $\rho = 0$, yielding $\text{Corr}(V,D) \approx 1 - 0.5\times 10^{-4} = 0.99995$.

Thus, under the given conditions, velocity V and displacement $D = V\cdot T$ are essentially perfectly correlated.

4. Conclusion and discussion

Whether we consider the pair (D, T) or (V, T), the correlation between the variable (displacement or velocity) and the derived quantity (velocity or displacement) is approximately 1 when the coefficient of variation of the variable is much larger than that of time. The small deviation is quadratic in the ratio of the coefficients of variation.

Supplementary Note III

Theoretical and numerical analysis of the optimal φ under finite lineage length

1. Introduction

Under finite lineage length, later-generation cells have fewer downstream descendants to implement delayed negative feedback. When stochastic velocity noise is generated at all generations, this asymmetry could, in principle, favor larger values of φ, because stronger immediate feedback partially compensates for insufficient long-range correction. Below, we show rigorously that although the optimal φ indeed depends on lineage length n, this dependence is highly structured: the variance function contains both **linearly growing** and **geometrically saturating** components, whose interplay causes the optimal φ to converge rapidly toward the golden ratio as lineage length increases.

2. Exact finite-lineage variance

2.1. Residual kernel and lineage distance

We start strictly from the proven residual kernel (Eq. 1):

$$R_{n,k} = \varepsilon_k(1-\sum_{i=2}^{n-k+1}\varphi^i), \varepsilon_k \sim N(0,\sigma^2) \tag{1}$$

Define the generation distance $\Delta = n - k \geq 0$. Then

$$R_{n,k} = \varepsilon_k(1-\sum_{i=2}^{\Delta+1}\varphi^i) \tag{2}$$

2.2. Exact simplification of the geometric sum

The geometric sum is evaluated exactly as

$$\sum_{i=2}^{\Delta+1}\varphi^i = \varphi^2\frac{1-\varphi^{\Delta}}{1-\varphi} \tag{3}$$

Substituting into Eq. (1) gives

$$R_{n,k} = \varepsilon_k \left[1 - \frac{\varphi^2}{1-\varphi} + \frac{\varphi^{\Delta+2}}{1-\varphi} \right] \tag{4}$$

For algebraic clarity only, define

$$A(\varphi) \equiv 1 - \frac{\varphi^2}{1-\varphi} \tag{5}$$

Then

$$R_{n,k} = \varepsilon_k \left(A(\varphi) + \frac{\varphi^{\Delta+2}}{1-\varphi} \right) \tag{6}$$

This step involves no approximation.

2.3. Total residual at generation n

The total residual at generation n is the sum of all ancestral contributions:

$$R_n = \sum_{k=0}^{n} R_{n,k} = \sum_{k=0}^{n} \varepsilon_k \left(A(\varphi) + \frac{\varphi^{n-k+2}}{1-\varphi} \right) \tag{7}$$

2.4. Variance of the final residual

Because the noise terms ε_k are independent and mean zero,

$$\mathrm{Var}(R_n) = \sigma^2 \sum_{k=0}^{n} \left(A(\varphi) + \frac{\varphi^{n-k+2}}{1-\varphi} \right)^2 \tag{8}$$

Letting $j = n - k (j = 0, \ldots, n)$ yields the exact general form

$$V_n(\varphi) = \sigma^2 \sum_{j=0}^{n} \left(A(\varphi) + \frac{\varphi^{j+2}}{1-\varphi} \right)^2 \tag{9}$$

2.5. Exact closed-form expression for finite lineage depth

Expanding the square,

$$V_n(\varphi) = \sigma^2 \left[(n+1)A^2 + \frac{2A}{1-\varphi} \sum_{j=0}^{n} \varphi^{j+2} + \frac{1}{(1-\varphi)^2} \sum_{j=0}^{n} \varphi^{2j+4} \right] \tag{10}$$

Both sums are evaluated exactly:

$$\sum_{j=0}^{n} \varphi^{j+2} = \varphi^2 \frac{1-\varphi^{n+1}}{1-\varphi}, \qquad \sum_{j=0}^{n} \varphi^{2j+4} = \varphi^4 \frac{1-\varphi^{2(n+1)}}{1-\varphi^2}. \tag{11}$$

Substituting back gives the fully exact closed form:

$$V_n(\varphi) = \sigma^2 \left[(n+1)A^2 + \frac{2A\varphi^2}{(1-\varphi)^2}(1-\varphi^{n+1}) + \frac{\varphi^4}{(1-\varphi)^2(1-\varphi^2)}(1-\varphi^{2(n+1)}) \right] \tag{12}$$

Equation (7) defines the variance of the accumulative velocity deviation at the terminal cell for any finite lineage depth n, without invoking limits or approximations.

3. Structured dependence of $V_n(\varphi)$ on lineage length

Importantly, the dependence of $V_n(\varphi)$ on lineage length arises from **two qualitatively distinct contributions**:

(1) **A linear accumulation term**, $(n+1)A(\varphi)^2$, which reflects the cumulative contribution of residual ancestral noise that is not fully cancelled by feedback. This term grows linearly with lineage length and penalizes deviations of $A(\varphi)$ from zero.

(2) **Two geometric correction terms**, proportional to $1-\varphi^{n+1}$ and $1-\varphi^{2(n+1)}$, which encode the progressive effectiveness of multi-generation negative feedback. These terms increase rapidly at small n but saturate geometrically as lineage length grows.

Thus, lineage length affects $V_n(\varphi)$ through both **extensive growth** (the linear term) and **intensive saturation** (the geometric terms). The optimal φ is determined by the balance between these effects.

4. Why the optimal φ converges rapidly

For short lineages, the geometric terms have not yet saturated, and delayed negative feedback is incomplete. In this regime, slightly larger values of φ reduce residual

variance by strengthening early feedback, shifting the optimum above the golden ratio.

As lineage length increases, the geometric correction terms approach their maximal contributions, while the linear term $(n+1)A(\boldsymbol{\varphi})^2$ increasingly dominates the sensitivity of $V_n(\boldsymbol{\varphi})$ to $\boldsymbol{\varphi}$. Because $A(\boldsymbol{\varphi}) = 0$ is achieved precisely at $\boldsymbol{\varphi} = \frac{\sqrt{5}-1}{2}$, the golden ratio uniquely minimizes the coefficient of the linear growth term. Consequently, once lineage length is sufficiently large, any deviation of $\boldsymbol{\varphi}$ from the golden ratio produces a linearly accumulating penalty that cannot be compensated by further geometric correction.

This structured variance decomposition explains why the optimal $\boldsymbol{\varphi}$ depends strongly on lineage length at small n, yet converges rapidly toward the golden ratio as lineage length increases, without invoking any infinite-lineage limit.

5. Numerical estimation of the optimal $\boldsymbol{\varphi}$ for finite lineage length

To quantify this dependence, the optimal $\boldsymbol{\varphi}$ was determined numerically for each finite lineage length $n$by directly minimizing the exact variance function $V_n(\boldsymbol{\varphi})$given in Eq. (7).

For a given lineage length, $\boldsymbol{\varphi}$ was constrained to a biologically meaningful interval $[0.2, 0.85]$. The variance $V_n(\boldsymbol{\varphi})$ was evaluated exactly for candidate values of $\boldsymbol{\varphi}$, and the value minimizing the terminal variance was identified using one-dimensional bounded optimization. This procedure yields both the optimal $\boldsymbol{\varphi}$ and the corresponding minimum variance for each lineage length.

The analysis was repeated across lineage lengths ranging from early embryonic depths to values exceeding those observed in *C. elegans*. The resulting optimal $\boldsymbol{\varphi}$ values were then compared to the golden ratio.

6. Conclusion and discussion

This analysis shows that lineage lengths of approximately 10–13 generations—matching the longest developmental lineages in *C. elegans*—are already sufficient to place the system near the global optimum defined by the golden ratio. At these lineage

depths, geometric feedback saturation is nearly complete, and further increases in lineage length would yield only marginal gains. Thus, the golden ratio emerges not as an asymptotic artifact, but as a **finite-lineage optimum** dictated by the exact variance structure of the self-similar feedback system.

Supplementary Note IV

Micro-Macro transformation between GDL form and iterative form

1. Introduction

To clarify how negative feedback across multiple generations is realized, we perform a micro–macro transformation of the self-similarity equation. Starting from the **geometric distributed-lag (GDL) formulation**:

$$V_t = -\alpha \sum_{i=1}^{t-1} \beta^i V_{t-i} + \varepsilon_t , \tag{1}$$

where V_t is the velocity deviation at generation t, ε_t is stochastic noise, α is the negative-feedback strength, and β is the decaying coefficient capturing ancestral influence. The geometric decay of β^i ensures that distant progenitors contribute progressively less to the negative-feedback correction.

Although Eq. (1) captures multi-generation feedback, it requires explicit summation over all ancestral generations, which is biologically implausible. To map this to a mechanistic, adjacent-generation process, we define:

$$V_t \equiv V_t^{\text{new}} , \tag{2}$$

i.e., the velocity deviation of the daughter cell as computed in the GDL form. We then decompose the daughter's total velocity-determining factors into:

(1) $V_t^{\text{inherited}}$: fraction inherited from the mother cell;

(2) V_t^{new}: fraction newly generated and actively contributing to the daughter's velocity.

The **iterative, two-generation formulation** becomes:

$$\begin{cases} V_t^{\text{new}} = -\alpha V_t^{\text{inherited}} + \varepsilon_t \\ V_t^{\text{inherited}} = \beta (V_{t-1}^{\text{new}} + V_{t-1}^{\text{inherited}}) \end{cases}, \tag{3}$$

where α now represents both the **negative-feedback strength** in the classical sense and the **generation rate** in the mass-conservation perspective, and βrepresents both the **decay of ancestral influence** and the **degradation rate** of inherited factors.

2. Equivalence with the GDL form

By recursive substitution of $V_t^{\text{inherited}}$into the first equation, we obtain:

$$V_t^{\text{new}} = -\alpha \sum_{i=1}^{t-1} \beta^i V_{t-i}^{\text{new}} + \varepsilon_t, \tag{4}$$

which is identical to the original GDL form (Eq. 1). Therefore, the iterative form is mathematically equivalent, while relying only on local two-generation interactions.

3. Mass-conservation interpretation

The iterative form naturally reveals a **mass-conservation law** for velocity-determining factors. Let αdenote the fraction newly generated per generation and $\beta$the fraction inherited. The total velocity-determining content is conserved across generations when $\beta(1+\alpha) = 1$ or equivalently $1-\beta = \alpha\beta$. This expresses a constant turnover rate: newly generated factors replace precisely the fraction lost by inheritance. Biologically, $\alpha = \beta = \varphi$ corresponds to the **golden ratio**, representing the unique value that simultaneously optimizes feedback correction and preserves mass of velocity-determining factors.

4. Biological significance

(1) Multi-generation negative feedback emerges naturally through repeated two-generation iterations.

(2) The iterative form maps directly onto measurable quantities: inherited and newly synthesized velocity-determining components.

(3) The golden ratio φ provides both the **optimal negative-feedback coefficient** and the **optimal mass-conservation parameter**, ensuring robust suppression of stochastic deviations and consistent velocity propagation across lineages.

5. Conclusion and discussion

This transformation demonstrates a **biologically realizable mechanism** for long-range negative feedback encoded in the GDL form. It provides a direct connection between the abstract self-similarity equation and concrete molecular–cellular processes. The iterative, two-generation loop constitutes a bionic design principle capable of achieving **zero-net-noise dynamics**, highlighting a deep analogy between control theory, econometrics, and developmental biology.

Supplementary Note V

Mathematical explanations for the optimization of $\alpha = \beta$

Summary

As demonstrated in Supplementary Note 2, the iterative form means the self-similarity process is formed by the iterations of two sub-processes, represented by two parameters, α and β. In the main text, we have suggested $\alpha = \beta$ based on modelling of real data. In this supplementary Note, we will provide several mathematical explanations for the optimization of $\alpha = \beta$, taking the iterative form as two consecutive sub-processes, listed in the following Table 1.

Table 1. Summary of biologically grounded cost / optimality principles for α–β equivalence

Objective name	Objective function	Optimality condition	Optimal point	Biological / biophysical interpretation
Shannon entropy (allocation evenness)	$H(r) = -r\ln r - (1-r)\ln(1-r)$, $r = \frac{\alpha}{\alpha+\beta}$	**Maximum**	$\alpha = \beta$	Maximizes even allocation of shared molecular resources between transmission and generation, preventing pathway dominance and enhancing developmental stability
Quadratic form (energy cost from enzymes, ATP)	$E = \alpha^2 + \beta^2$ with $(\alpha+\beta)x = C$	**Minimum**	$\alpha = \beta$	Minimizes ATP dissipation, enzyme saturation, and molecular crowding under fixed total concentration
Geometric mean (turnover fraction)	$T = \alpha\beta$ with $(\alpha+\beta)x = C$	**Maximum**	$\alpha = \beta$	Maximizes effective turnover between inherited and newly generated velocity-determining factors
Maximum (bottleneck-dominated cost)	$B = \max(\alpha, \beta)$ with $(\alpha+\beta)x = C$	**Minimum**	$\alpha = \beta$	Eliminates dominant bottlenecks by balancing transmission and generation capacities
Harmonic mean (propagation efficiency)	$H_m = \frac{2}{\frac{1}{\alpha}+\frac{1}{\beta}}$ with $(\alpha+\beta)x = C$	**Maximum**	$\alpha = \beta$	Maximizes efficiency of coupled transmission–generation ratios under fixed total molecular capacity
Reciprocal ratio (energy per turnover)	$R = \frac{\alpha}{\beta} + \frac{\beta}{\alpha} = \frac{\alpha^2+\beta^2}{\alpha\beta}$	**Minimum**	$\alpha = \beta$	Minimizes energetic cost per unit turnover; penalizes imbalance between transmission and generation

Legends for Supplementary Tables

Table S1: Raw cell cycle length

Table S2: Raw positions at cell birth

Table S3: Raw positions at cell division

Table S4: Calculated *k* of all M-D pairs

Tables S5: Hatching phenotype