Precise and scalable self-organization in mammalian pseudo-embryos

Melody Merle, Leah Friedman, Corinne Chureau, Thomas Gregor

Precise and scalable self-organization in mammalian pseudo-embryos

Mélody Merle,[1,*] Leah Friedman,[1,*] Corinne Chureau,[1,*] and Thomas Gregor[1,2,†]

[1] *Department of Developmental and Stem Cell Biology, CNRS UMR3738 Paris Cité, Institut Pasteur, Paris, France*

[2] *Joseph Henry Laboratories of Physics & Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA*

(Dated: March 31, 2023)

During multi-cellular development, highly reproducible gene expression patterns determine cellular fates precisely in time and space. These processes are crucial during the earliest stages when the body plan and the future asymmetric body axes emerge at gastrulation. In some species, such as flies and worms, these early processes achieve near-single-cell spatial precision, even for macroscopic patterns. However, we know little about such accuracy in mammalian development, where quantitative approaches are limited. Using an *in vitro* model for mammalian development, i.e., gastruloids, we demonstrate that gene expression patterns are reproducible to within 20% in protein concentration variability, which translates to a positional error close to a single cell diameter at the tissue scale. In addition, 2–3 fold system size changes lead to scaled gene expression patterns again on the order of an individual cell diameter. Our results reveal developmental precision, reproducibility, and size scaling for mammalian systems. All three properties spontaneously arise in self-organizing cell aggregates and could thus be fundamental features of multicellularity.

Identical body plans across individuals of multi-cellular organisms result from precise and reproducible gene expression patterning that is molecularly controlled by intrinsically noisy processes [1–3]. High reproducibility and precision have been demonstrated in animal models ranging from worms to vertebrates [3–8]. In the early fly embryo, the precision of the macroscopic features of the body plan could be traced back to the reproducibility of maternal signaling gradients [9, 10], and this reproducibility translates to a spatial accuracy that is proportional to system size and sufficient to distinguish individual cells from their neighbors [11]. For decades development was viewed as a succession of error-reducing steps coping with molecular noise [2, 12]. These steps have been optimized over evolutionary time scales within the specific boundaries of each organism [13]. Here we test these ideas in the context of mammalian development.

Developmental reproducibility and precision have rarely been assessed in mammals where quantitative approaches are limited. Recent progress with *in vitro* models offers new opportunities to overcome these difficulties [14, 15]. Gastruloids are small aggregates derived from mouse embryonic stem cells (mESCs), forming three-dimensional pseudo-embryos that mimic key events of mammalian gastrulation through self-organized patterning (Fig. S1A). In particular, they break symmetry and elongate along an axis that resembles the most posterior part of the mouse embryo's anterior-posterior (AP) axis [16]. They can be efficiently grown in large quantities, opening new avenues for quantitative approaches. Yet, the reproducibility of these systems has been questioned [17].

Here we ask how precisely a self-organizing multicellular system is regulated. We show that gastruloids can indeed be grown in a highly reproducible manner and lend themselves as a quantitative model for mammalian development. We demonstrate the intrinsic reproducibility of the self-organization process, both for growth dynamics and gene expression patterns. Expression levels are tightly controlled along the AP axis, and pattern boundaries are positioned with single-cell precision. Gastruloid growth scales with the initial number of seed cells, and AP gene expression patterns scale precisely with the length of the gastruloids' midline.

RESULTS

Reproducible gastruloid growth and size scaling. The general problem of mammalian systems for quantitative analysis stems in part from the difficulty of repeating experiments under the exact same conditions. This is usually exacerbated by relatively slow and therefore long processes and the often inaccessible nature of the specimen, e.g. for *in vivo* studies. In principle, the gastruloid model can overcome this hurdle as it can quickly be produced to run hundreds of identical *in vitro* experiments in parallel. However, before we can address the intrinsic precision and reproducibility of these self-organized systems, we need to assess how strictly experimental conditions can be controlled to achieve the highest possible degree of experimental reproducibility in gastruloids. Therefore, we start by examining general physical properties, such as reproducibility of growth and dependence

* These authors contributed equally.
† Correspondence: thomas.gregor@pasteur.fr

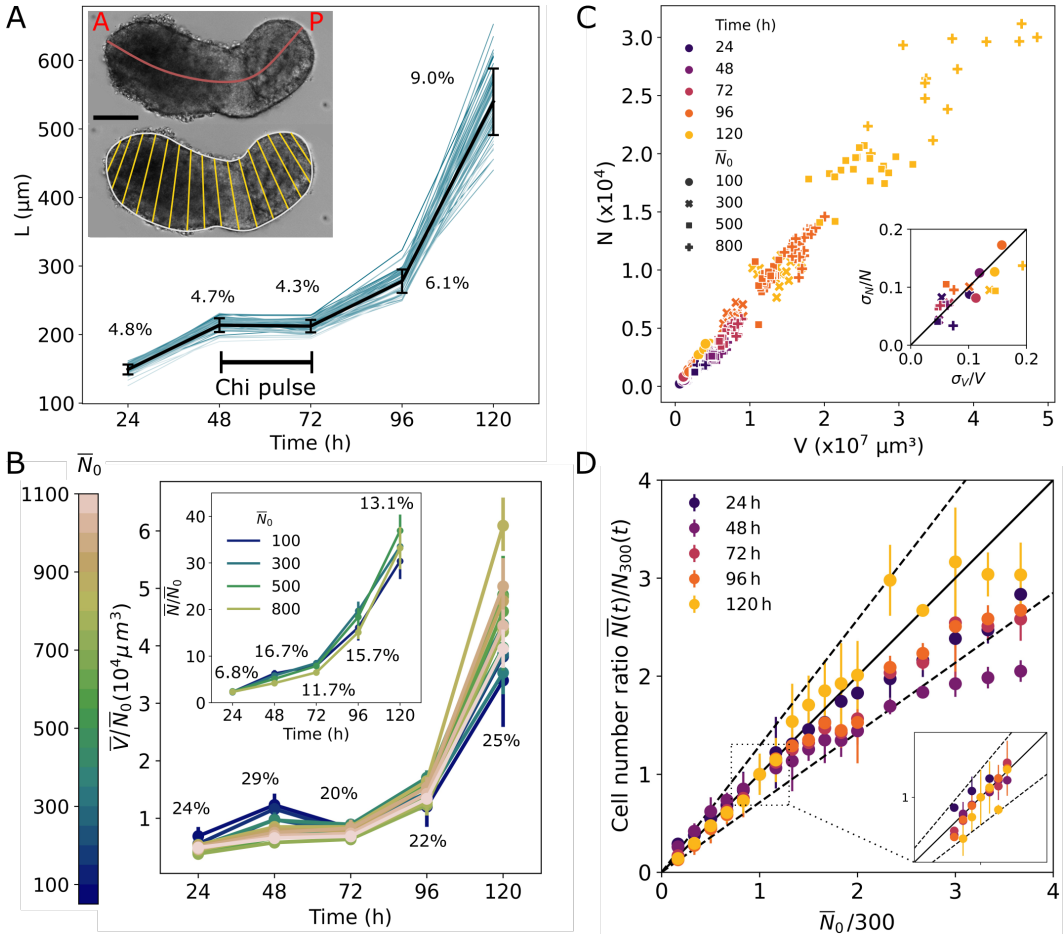

FIG. 1. Reproducible gastruloid growth, scaled to system size. **A**: Gastruloid midline length variation as a function of time. Curves shown for 57 gastruloids followed individually over time (blue) and mean (black). Percent variation around the mean is reported for each time point. Spread of initial number of seeded mESCs is $N_0 = 305 \pm 28$ cells (Fig. S1A). (For an equivalent relationship for volume see Fig. S2A.) Inset shows a bright-field image of a gastruloid at 120 h, overlaid with its midline ranging from anterior (A) to posterior (P) pole (red, top) and sliced evenly for volume reconstruction (yellow, bottom); scalebar is 100 μm; also see Fig. S1B-D. **B**: Gastruloid volume and total cell count (inset and Fig. S1E) as a function of time. Volumes are normalized by the average number of initial seed cells $\overline{N}_0$ at time zero (color code). Each line represents the mean of on average 15 gastruloids with the same N_0. Percentages correspond to residual variations within which normalized volumes collapse for 15 different values of $\overline{N}_0$. Similar collapse for normalized gastruloid cell counts for four values of $\overline{N}_0$ (inset). **C**: Scatter plot of total cell count versus the measured volume for 492 individual gastruloids at different time points (color code) and with varying $\overline{N}_0$ (symbol); Pearson correlation coefficient is $r = 0.99$. Inset shows correlation ($r = 0.78$) of variability for N and V for sets of gastruloids with identical age and N_0. This is evidence that the independent methods for measuring N and V are accurate estimates of gastruloid growth. **D**: Cell count ratio $\overline{N}(t)/N_{300}(t)$ as a function of initial seed cell count ratio $\overline{N}_0/300$. In both the cell count of gastruloids seeded from $\overline{N}_0 = 300$ is chosen as a reference. Time is encoded by color (see legend). Black diagonal (slope = 1) represents perfect scaling of gastruloid size at time t upon changes in $\overline{N}_0$ ranging over $50 \leq N_0 \leq 1100$. Dashed line estimates expected deviations from perfect scaling at 120 h due to fluctuations in $\overline{N}_0/300$ and the doubling time t_D given a simple exponential growth model (Fig. S2F). Inset shows the same relationship centered around $\overline{N}_0 = 300$ where we observe close to perfect scaling at all time points.

on initial conditions. We measure growth by monitoring for five days the length of the midline, the total volume, and the total cell count of individual gastruloids (Fig. S1, Methods).

The growth curves of individual gastruloids collapse onto a tight relationship, both in terms of length and volume measurements, indicating high reproducibility of the growth dynamics at all time points (Fig. 1A and S2A). The residual spread in these curves can be partially traced back to the variability of the initial number of seeded cells N_0 and the fluctuations in effective doubling time, which we measure to be on average 25 ± 1.5 h for $\overline{N}_0 = 300$ (Fig. S2B). Gastruloid volumes correlate with N_0 at all time points (Fig. S2C), and when we scan a large range of average $\overline{N}_0$ (up to 20-fold changes) the correlations increase significantly (Fig. S2D). Growth

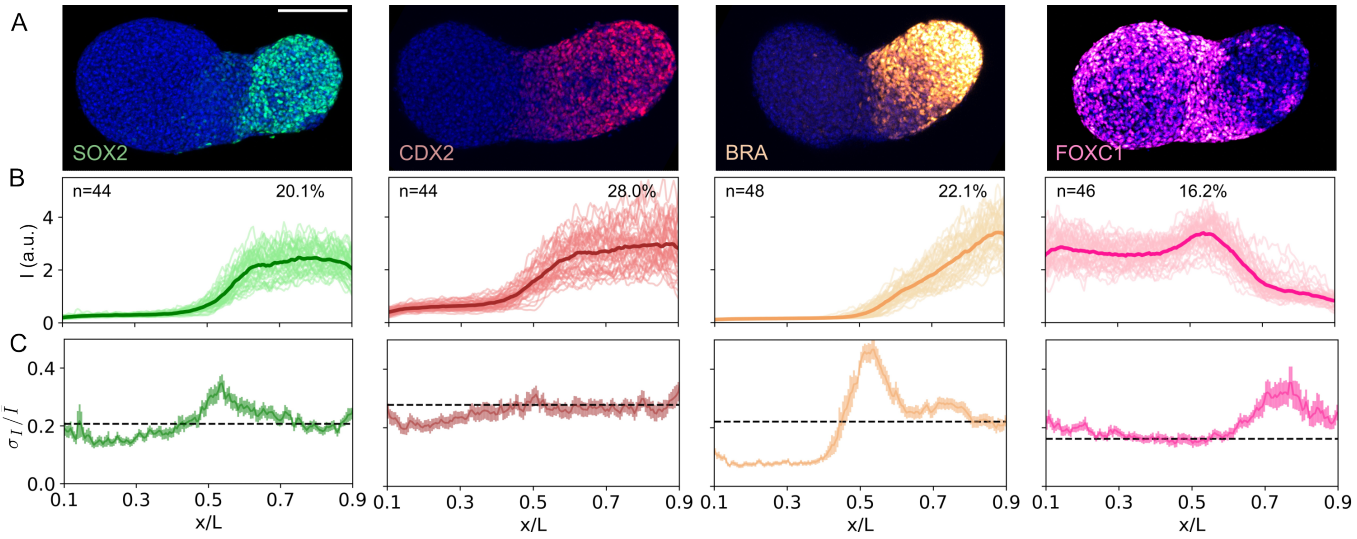

FIG. 2. Reproducibility of gene expression patterns in gastruloids. **A**: Maximum projections of four confocal image stacks of 120 h old gastruloids stained by immunofluorescence for SOX2, CDX2, BRA, FOXC1. AP-axis is in a left–right orientation. Scalebar is 100 μm. **B**: n individual gene expression profiles (light color) for the four markers in A and the corresponding average profile (dark bold) projected on the midline and reported relative to gastruloid length L. **C**: Variability ($\sigma_I/\bar{I}$) of the respective gene expression patterns from B. Error bars are obtained by bootstrapping. Dashed lines represent the average variability in the region where genes are most highly expressed (values in B).

curves of gastruloid volumes normalized by $\overline{N}_0$ all collapse (Fig. 1B), demonstrating that growth is reproducible across all measured time points and can be traced back to the initial number of seeded mESCs.

A similar relationship is seen for the growth curves of gastruloid's total number of cells, N (obtained from chemically dissociated gastruloids, Fig. S1E). We see again for different initial $\overline{N}_0$ that all curves collapse (Fig. 1B, inset) and that the effective doubling time is conserved (Fig. S2E). The residual spread in the collapsing growth curves can be related to other factors, i.e. the variability between experiments (Fig. S2F). Overall, these collapses seem to indicate that size control is an emergent property, and a refinement to achieve a target size with error reduction seems unnecessary.

The observed collapse of the growth curves indicates that gastruloid size scales with the initial seed number. Contrary to mouse embryos [18–20], there seems to be no control over the absolute size of the system. To verify this assertion more directly, we determine how well gastruloid volume is a predictor of the total cell count. Measurements of cell count and volume on the same gastruloid demonstrate a tight linear relationship across multiple time points and for various $\overline{N}_0$ (Fig. 1C). This strong correlation suggests that the underlying dispersion in single cell size is highly conserved across gastruloids and across external conditions (Fig. S7G). Using this relationship, we demonstrate quasi-perfect cell count scaling with $\overline{N}_0$ for all but the earliest stages (pre-Chi pulse) of gastruloid growth (Fig. 1D). Together, these results demonstrate that growth proceeds with high fidelity. Under well-controlled experimental conditions, gastruloids self-organize with very close control over variability in growth rate (Fig. S2E) and other noisy processes of size control.

Reproducible gene expression patterning. Coordinated gastruloid growth and especially axis elongation are coupled to gene expression patterns along the body axes. Thus we ask next how the physical properties of reproducibility and scaling are reflected in the anterior-posterior patterning of gene expression. To this end, we measure gene expression profiles five days after seeding when the pseudo-AP-axis is morphologically well established. We performed immunofluorescence staining of four germ-layer markers (SOX2, CDX2, BRA, and FOXC1) that are part of the patterning process (Fig. 2A). From 2D maximum projections of a confocal image stack, we extracted 1D intensity profiles projected isometrically on each gastruloid's midline (Fig. 2B, S3B, Methods).

Perhaps unexpectedly for mammalian systems, for all four genes, individual profiles are very similar, clustering around the average profile tightly. The overall variability is small (Fig. 2C): it peaks in regions where gene expression levels change significantly over short distance intervals, such as pattern boundaries. Such a region is observed, e.g., at 0.5 x/L in the BRA profile, showing that even the maximally observed variation in those regions rarely surpasses 40%. The variability in these gene expression profiles from one gastruloid to the next follows a trend similar to what was observed in other organisms [9, 21–23]. Pioneering experiments in engineered systems have shown that even when expression inducer

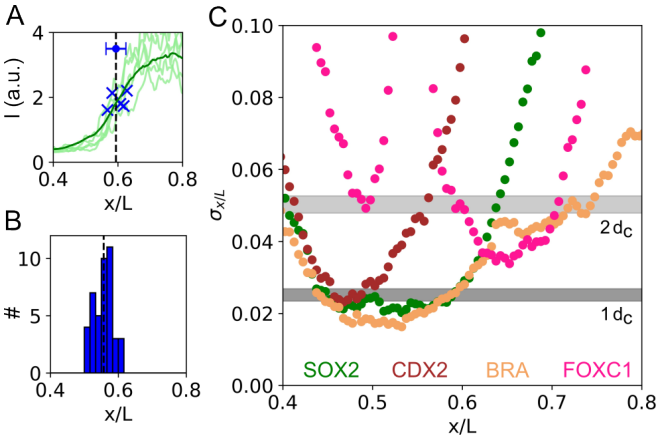

FIG. 3. Single-cell pattern boundary precision in gastruloids. **A**: Close-up of Fig. 2B for four SOX2 expression profiles as a function of position along the midline (green), with pattern boundary positions of four individual profiles marked at the half-maximal expression value (EC50, blue crosses). Mean profile of gastruloid midlines in dark green (n=44). **B**: Distribution of SOX2 pattern boundary positions from Fig. 2B. The mean defines the pattern boundary position x_B; the standard deviation of the distribution (blue bar in A) of 2.4% defines the positional error for pattern boundary establishment. **C**: Positional error directly calculated from the standard deviation of intensity values across the individual expression profiles in A, $\sigma_I(x/L)$. For each position x/L, this expression error is propagated into an error in position, $\sigma_{x/L}$ (see Methods). Color code as in Fig. 2; gray areas correspond to one and two effective cell diameters d_c, respectively, including measurement errors (Fig. S6 and S7).

concentrations are high, resulting levels of gene expression fluctuate [21, 22]. Here we see something similar in the regions where genes are expressed at their maximal levels of expression (Fig. S5B), where the variability hovers around 20% (dashed lines in Fig. 2C). A similar variability near maximum expression is also seen in other organisms [9, 23].

Note that much of the noise we observe could in principle result from measurement errors. In separate experiments (see Methods and Fig. S3, S4, S6, S8) we estimate the component of measurement noise which arises in the experimental process. Some sources of experimental noise are inherent to immunofluorescence staining and imaging [10], and others are due to the arbitrary choices made during the image analysis routines, such as axis definition and projection method to measure the gene expression patterns (Fig. S3). Overall, we estimated that all sources of measurement error combined correspond to less than 10% of the total variance (Fig. S4, Methods). The measured variability thus represents an upper bound for the biological variability of the system and the true value is even lower.

Single-cell precision of pattern boundaries. During development, cells implement different actions along the body axis according to patterning signals executed by, e.g., the four genes we analyzed above [24, 25]. System intrinsic variability from one gastruloid to the next (Fig. 2B) limits the positional precision with which such actions can be implemented. In the case of the four analyzed genes, we estimate the positional precision by determining the positions x/L at which the half-maximal expression level in the boundary regions for each pattern is reached (Fig. 3A and S5A). In the case of SOX2, we obtain a narrow distribution of these positions (Fig. 3B) with a standard deviation of 2.4%, and similar values for the other genes (Fig. S5C).

Instead of examining a singular boundary point, a more general way to estimate the positional error is to consider the full extent of the pattern and to propagate the fluctuations in expression levels (Fig. 2C) into an error in position (Fig. 3C, Methods) [26]. The generalization shows that a precision of 2–3% is obtained in domains that span between 5 to 10% of the gastruloid length. These domains correspond to the respective intermediate levels of expression for each gene (Fig. S5D). Moreover, the values obtained with both methods match at the mean pattern boundary positions (Fig. S5E). These results indicate that in principle, cells could use these genes' expression levels to determine their position along the pattern boundary precisely.

To understand how 2–3% spatial precision along the midline translates into a relevant length for the system in absolute units, we measured the average size of an individual cell within gastruloids. To this end, we return to our simultaneous measurement of cell count N and volume V for several hundred gastruloids (Fig. 1C). The strong linear dependence N=s*V has a slope s, which is the inverse of the mean cell volume. We can thus extract the effective diameter of cells in developing gastruloids. This measurement is recapitulated within error bars by an alternative method that is based on a high-resolution 3D reconstruction of individual gastruloids with fluorescently labeled cell membranes. The consensus effective cell diameter at 120 h is $d_c = 13.5 \pm 0.8\,\mu m$ and represents the relevant linear size of the system (Fig. S7). With this system-intrinsic length scale measurement, we determine that the achieved patterning precision corresponds to 1–2 cell diameters along the midline of the gastruloid (Fig. 3C). This result shows that mammalian gastruloids exhibit a patterning precision comparable to patterning systems in fly embryos [9] and worms [27].

These levels of reproducibility and precision are conserved for gastruloids grown in parallel from the same identical population of cells. Minimizing sources of variability of the whole chain of experimental protocols, from gastruloid seeding to imaging, we obtained profiles of very similar average absolute concentration levels, variability, and positional error in several experiments and for several genes (Fig. S8). Note that we are reporting the total variance, which includes measurement errors. Thus, the actual values for the reproducibility and pre-

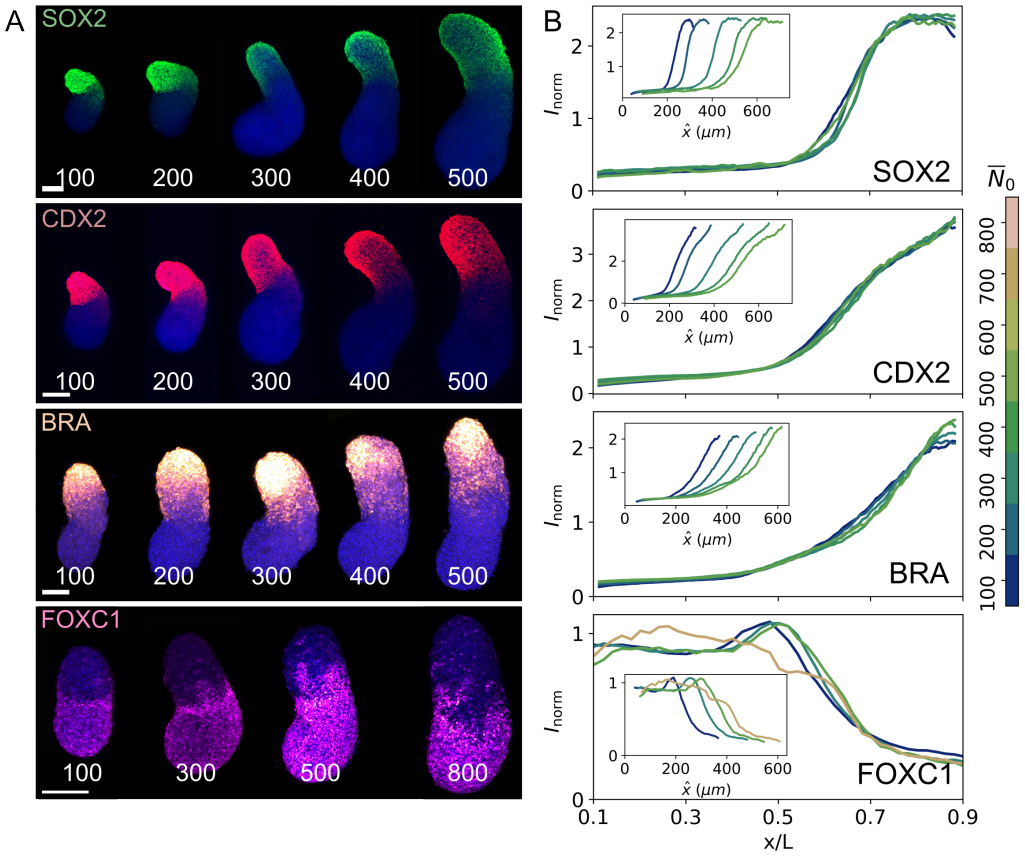

FIG. 4. Scaling of AP gene expression patterns in gastruloids. **A**: Confocal images of gastruloids immunofluorescently stained for four different genes each representing a different initial seed number $\overline{N}_0$ (in white, also Fig. S9). AP-axis from bottom to top. Scalebars are 100 μm. **B**: Normalized mean expression profiles for sets of gastruloids with the same $\overline{N}_0$ (color code) as a function of the relative position x/L along the average midline of the respective set (n=15–50 gastruloids per $\overline{N}_0$). AP-axis is in a left–right orientation. Inset shows normalized mean expression profiles as in B as a function of average position in absolute units of the respective set.

cision of the system are even higher than what we report here. These results provide an absolute scale for the reproducibility of the self-organized patterning process; cells at each point along the pattern produce the same amount of gene product in absolute units. These units translate along the pattern axis into a spatial precision of the linear dimension of a single cell, arguably the physical limit that a biological system can achieve.

Gene expression pattern scaling with size. While gastruloid lengths vary by 7–11% five days after seeding (Fig. 1A and S1F), the relative positional error is less than 3%, indicating that the mechanisms underlying pattern formation in gastruloid are able to measure the system size and correct for it [28]. To assess this conjecture directly, we generate gastruloids of different sizes by varying the initial number of seeded cells in the range with the most robust elongation results achieving a 2.3-fold change in gastruloid lengths (Fig. 4A and S9).

For each of the four genes examined above, we generate sets of immunofluorescently labeled gastruloids with different $\overline{N}_0$. When plotting the average profiles for each set of $\overline{N}_0$ and each gene as a function of absolute length, they are spread out along the x-axis (Fig. 4B, inset). However, when rescaled by the mean gastruloid length of a given set, the average profiles of all sizes collapse (Fig. 4B). The remaining positional error of the rescaled profiles is within 1–2 cells (Fig. S9C), similar to the intrinsic precision of the pattern boundaries. These results demonstrate that the expression patterns of the four genes have information locally on the global length of the system. At each position along the gastruloid's midline, a cell produces and maintains an exact absolute amount of protein with an accuracy of a few tens of percent of its mean value.

DISCUSSION

Our results argue for an intrinsic potential for reproducible, precise, and scalable self-organization in gastruloids. Hence these properties can be achieved in mam-

malian systems and are not specific to fly or worm development. In addition, it suggests that reproducibility and scaling, identified both in developing embryos and in these synthetic structures, can be context-independent properties. Our findings point to underlying principles that govern the self-organization processes in multicellular systems, reaching across half a billion years of evolutionary change. The fact that synthetic systems have similar intrinsic reproducibility and precision to *in vivo* systems expands the possibilities for advanced engineering applications in the field of organoids and more generally cell aggregates.

The perhaps surprising properties uncovered here about mammalian cell aggregates cannot apply to all features of these systems. For example, gastruloid shape is clearly not reproducible (see e.g., Fig. 2A and 4A). In fact, gastruloid length, which is a feature of shape, varies more at 120 h than a pattern boundary position. It suggests that only truly self-organized processes, such as the emergence of a gene expression pattern in these aggregates, are precisely regulated, while the shape is largely influenced by external and environmental conditions and thus not subject to stringent control [29].

In mammals, these findings can currently not be obtained *in vivo*; only *in vitro* systems provide the necessary experimental accessibility and workability. Not only does the use of an *in vitro* system allow for the necessary experimental accessibility, but it also enables easy control of parameters such as the system size and allows for perturbing the system continuously and outside its natural limits. Our approach and findings suggest that the gastruloid model, and maybe more generally also other stem-cell-derived aggregates, have the potential to serve as a powerful tool to study mammalian development and other processes, quantitatively.

ACKNOWLEDGMENTS

We thank Isma Bennabi, David Brückner, Michele Cerminara, Michel Cohen-Tannoudji, Pauline Hansen, Miloš Nikolić, Camil Mirdas, Judith Pineau, Armin Shoushtarizadeh, Jerome Wong-Ng, Benjamin Zoller, and the late Roel Neijts. This work was supported by Institut Pasteur, Centre National de la Recherche Scientifique, CFM Foundation for Research, and the French National Research Agency (ANR-10-LABX-73 'Revive', ANR-19-CE45-0016 'Polychrome', and ANR-20-CE12-0028 'ChroDynE').

[1] P. A. Lawrence, *The making of a fly: The genetics of animal design* (Blackwell Scientific Publications, Oxford, U.K., 1992).

[2] A. M. Arias and P. Hayward, Filtering transcriptional noise during development: concepts and mechanisms, Nature Reviews Genetics **7**, 34 (2006).

[3] J. Briscoe and S. Small, Morphogen rules: design principles of gradient-mediated embryo patterning., Development (Cambridge, England) **142**, 3996 (2015).

[4] J. E. Sulston, E. Schierenberg, J. G. White, and J. N. Thomson, The embryonic cell lineage of the nematode Caenorhabditis elegans, Dev. Biol. **100**, 64 (1983).

[5] T. Bollenbach, P. Pantazis, A. Kicheva, C. Bökel, M. González-Gaitán, and F. Jülicher, Precision of the Dpp gradient, Development **135**, 1137 (2008).

[6] E. Bier and E. M. De Robertis, Embryo Development. BMP gradients: A paradigm for morphogen-mediated developmental patterning., Science (New York, N.Y.) **348**, aaa5838 (2015).

[7] L. Bentovim, T. T. Harden, and A. H. DePace, Transcriptional precision and accuracy in development: from measurements to models and mechanisms, Development **144**, 3855 (2017).

[8] M. Zagorski, Y. Tabata, N. Brandenberg, M. P. Lutolf, G. Tkačik, T. Bollenbach, J. Briscoe, and A. Kicheva, Decoding of position in the developing neural tube from antiparallel morphogen gradients, Science (New York, N.Y.) **356**, 1379 (2017).

[9] T. Gregor, D. Tank, E. Wieschaus, and W. Bialek, Probing the Limits to Positional Information, Cell **130**, 153 (2007).

[10] J. Dubuis, R. Samanta, and T. Gregor, Accurate measurements of dynamics and reproducibility in small genetic networks, Molecular Systems Biology **9**, 10.1038/msb.2012.72 (2013).

[11] G. T. Petkova MD, Tkacik G, Bialek W, Wieschaus EF, M. D. Petkova, G. Tkačik, W. Bialek, E. F. Wieschaus, and T. Gregor, Optimal decoding of cellular identities in a genetic network, Cell **176**, 844 (2019), arXiv:1612.08084.

[12] C. H. Waddington, Canalization of development and the inheritance of acquired characters, Nature 1942 150:3811 **150**, 563 (1942).

[13] T. C. Lacalli, Patterning, From Conifers to Consciousness: Turing's Theory and Order From Fluctuations, Frontiers in Cell and Developmental Biology **10**, 10.3389/fcell.2022.871950 (2022).

[14] N. Gritti, D. Oriola, and V. Trivedi, Rethinking embryology in vitro: A synergy between engineering, data science and theory, Developmental Biology **474**, 48 (2021).

[15] E. A. Rosado-Olivieri and A. H. Brivanlou, Synthetic by design: Exploiting tissue self-organization to explore early human embryology, Developmental Biology **474**, 16 (2021).

[16] L. Beccari, N. Moris, M. Girgin, D. A. Turner, P. Baillie-Johnson, A.-C. Cossy, M. P. Lutolf, D. Duboule, and A. M. Arias, Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids, Nature **562**, 272 (2018).

[17] J. Fu, A. Warmflash, and M. P. Lutolf, Stem-cell-based embryo models for fundamental research and translation, Nature Materials **20**, 132 (2021).

[18] M. H. Snow and P. P. Tam, Is compensatory growth a complicating factor in mouse teratology?, Nature **279**, 555 (1979).
[19] N. E. Lewis and J. Rossant, Mechanism of size regulation in mouse embryo aggregates., Journal of embryology and experimental morphology **72**, 169 (1982).
[20] G. F. Rands, Size regulation in the mouse embryo. II. The development of half embryos., Journal of embryology and experimental morphology **98**, 209 (1986).
[21] M. Elowitz, A. Levine, E. Siggia, and P. Swain, Stochastic gene expression in a single cell, Science **297**, 1183 (2002).
[22] J. Raser and E. O'Shea, Control of stochasticity in eukaryotic gene expression, Science **304**, 1811 (2004).
[23] T. Carolina de Souza-Guerreiro, X. Meng, E. Dacheux, H. Firczuk, and J. McCarthy, Translational control of gene expression noise and its relationship to ageing in yeast., The FEBS journal **288**, 2278 (2021).
[24] R. Neijts, S. Simmini, F. Giuliani, C. van Rooijen, and J. Deschamps, Region-specific regulation of posterior axial elongation during vertebrate embryogenesis, Developmental Dynamics **243**, 88 (2014).
[25] S. Amin, R. Neijts, S. Simmini, C. van Rooijen, S. C. Tan, L. Kester, A. van Oudenaarden, M. P. Creyghton, and J. Deschamps, Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche., Cell reports **17**, 3165 (2016).
[26] J. Dubuis, G. Tkacik, E. Wieschaus, T. Gregor, and W. Bialek, Positional information, in bits, Proceedings of the National Academy of Sciences of the United States of America **110**, 16301 (2013).
[27] J. Cao, G. Guan, V. W. S. Ho, M. K. Wong, L. Y. Chan, C. Tang, Z. Zhao, and H. Yan, Establishment of a morphological atlas of the Caenorhabditis elegans embryo using deep-learning-based 4D segmentation, Nature Communications 2020 11:1 **11**, 1 (2020).
[28] S. Werner, T. Stückemann, M. Beirán Amigo, J. C. Rink, F. Jülicher, and B. M. Friedrich, Scaling and Regeneration of Self-Organized Patterns, Physical Review Letters **114**, 138101 (2015), arXiv:1411.2359.
[29] J. V. Veenvliet, P.-F. Lenne, D. A. Turner, I. Nachman, and V. Trivedi, Sculpting with stem cells: how models of embryo development take shape, Development **148**, 10.1242/dev.192914 (2021).
[30] L. Beccari, M. Girgin, D. A. Turner, P. Baillie-Johnson, A.-C. Cossy, L. Beccari, N. Moris, M. Lutolf, D. Duboule, and A. Martinez Arias, Generating Gastruloids from Mouse Embryonic Stem Cells, Protocol Exchange 10.1038/protex.2018.094 (2018).
[31] S. C. van den Brink, A. Alemany, V. van Batenburg, N. Moris, M. Blotenburg, J. Vivié, P. Baillie-Johnson, J. Nichols, K. F. Sonnen, A. Martinez Arias, and A. van Oudenaarden, Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids., Nature 10.1038/s41586-020-2024-3 (2020).
[32] M. Mansoury, M. Hamed, R. Karmustaji, F. Al Hannan, and S. T. Safrany, The edge effect: A global problem. The trouble with culturing cells in 96-well plates., Biochemistry and biophysics reports **26**, 100987 (2021).
[33] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, Cellpose: a generalist algorithm for cellular segmentation., Nature methods **18**, 100 (2021).
[34] G. Tkačik, J. Dubuis, M. Petkova, and T. Gregor, Positional information, Positional error, and readout precision in morphogenesis: A mathematical framework, Genetics **199**, 39 (2015).

Appendix A: Methods

Disclaimer: A full detailed account of the methods will be added in a subsequent version.

Appendix B: Supplemental Figures

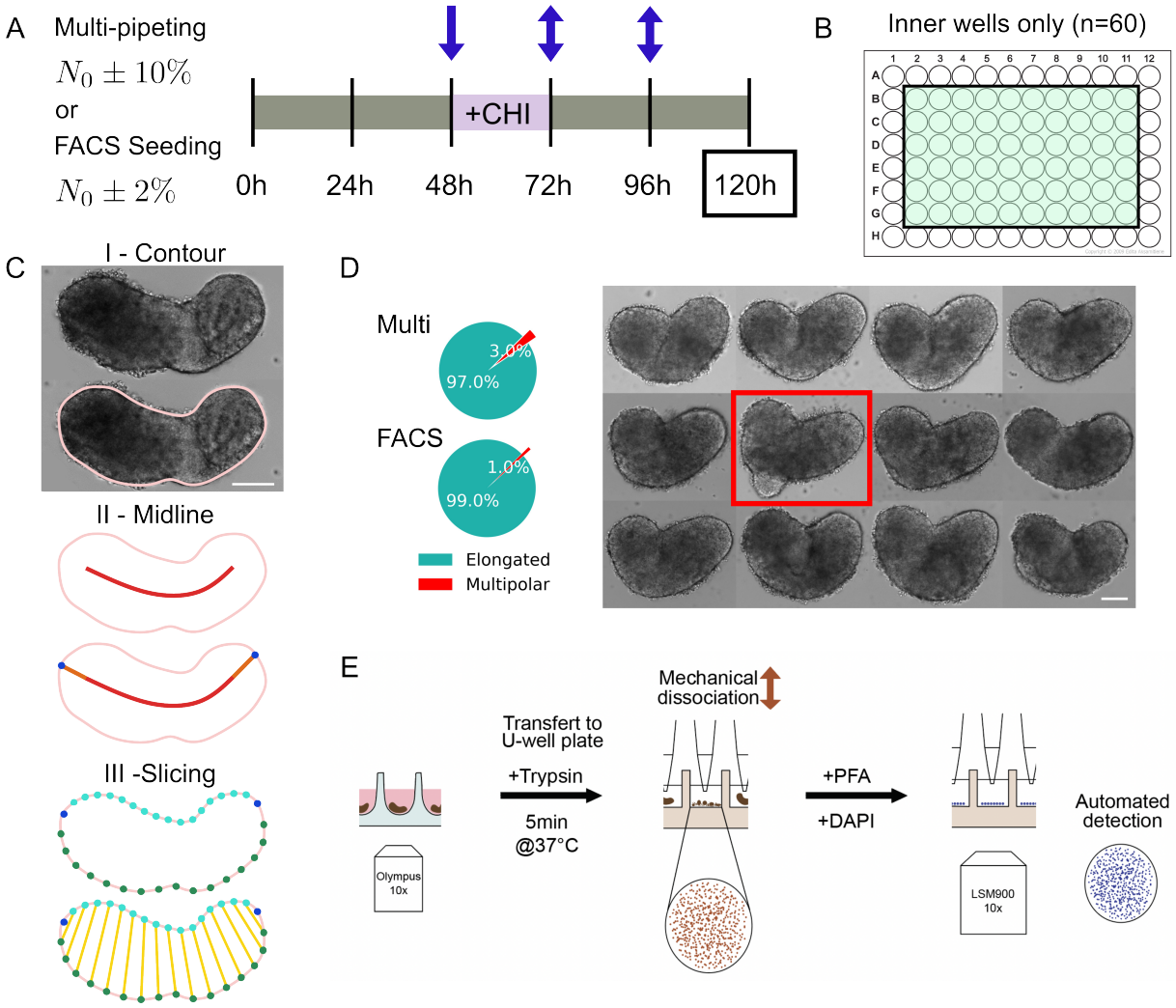

FIG. S1. Experimental detail, protocols, and analysis. **A**: Gastruloid protocol as described before with a Chi pulse on day three [30]. Initial seeding either done by manual multi-pipetting or using Fluorescence-activated Cell Sorting (FACS) [31], implying a different variability in the initial number of seeded cells N_0; 10% vs. 2%, respectively. Blue arrows indicate addition of Chiron and change of medium. **B**: Discarding all gastruloids grown in outer wells for increasing reproducibility. Empirical observation determined largely from different behaviors for gastruloids grown in inner versus outer wells [32]. **C**: Image analysis steps include the definition of a smooth contour (I), drawing the midline (II), and slicing along this midline using an equidistant positioning of two sets of equal-number points on each side of the contour (III). For III, the points in left half (light blue) and in right half (dark blue) are equidistant along the contour, respectively. Gastruloid volume is reconstructed by assuming each slice is rotationally symmetric (i.e., a truncated cone). Scalebar is 100 μm. **D**: Gastruloids imaged with bright-field microscopy. Gastruloid elongation efficiency is 97% for multi-pipetting and 99% for FACS seeding. The remaining gastruloids have multiple poles (e.g., red framed image). Scalebar is 100 μm. **E**: Schematic of the protocol to measure the volume and cell count of individual gastruloids. Bright-field images of gastruloids are acquired before chemical dissociation, left; fluorescent images of *all* individual cells composing the gastruloid are acquired after dissociation using confocal microscopy (see Methods).

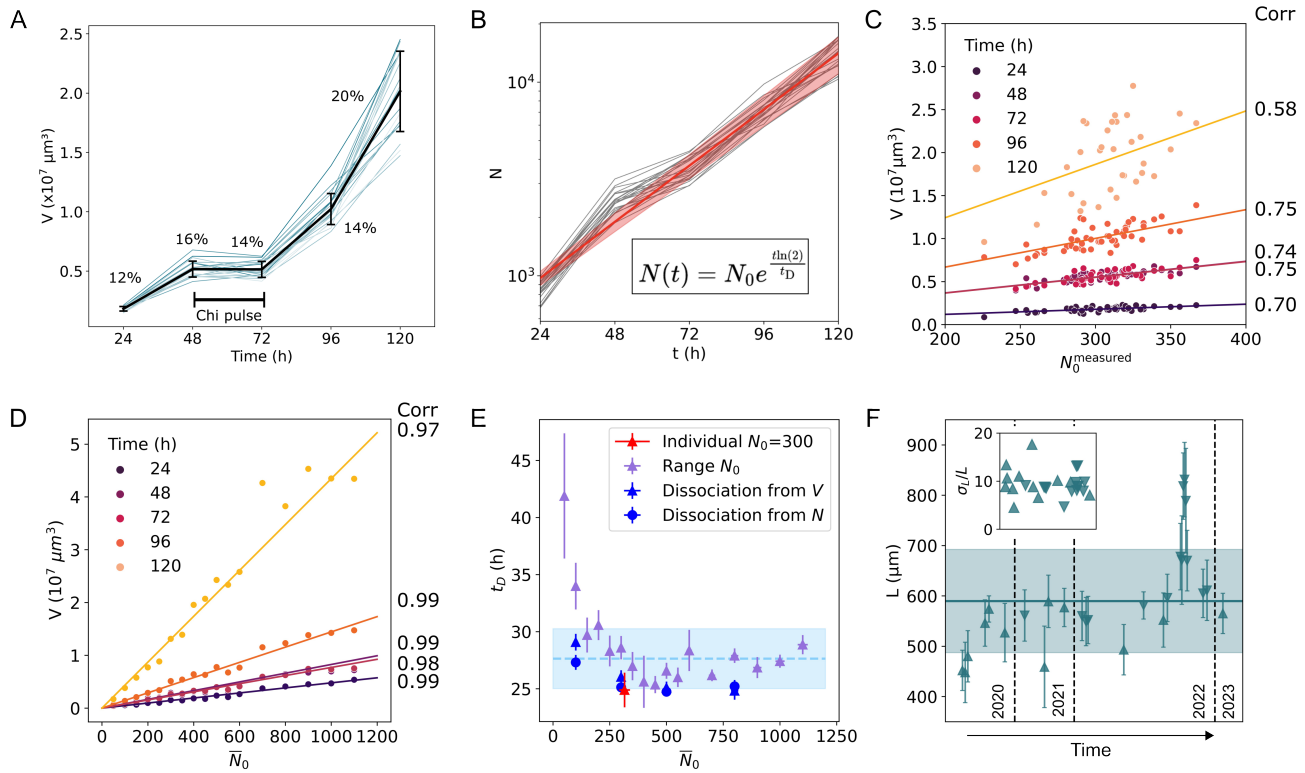

FIG. S2. Growth reproducibility and scaling. **A**: Gastruloid volume as a function of time. Volumes are obtained from 2D reconstruction in Fig. S1C. Curves shown for 23 gastruloids (subset of Fig. 1A) followed over time individually (blue) and mean (black). Percent variation around the mean is reported for each time point. **B**: Exponential growth of the total number of cells in individual gastruloids (same as in A). The total cell count N shown in log scale as a function of time t is obtained from the proportionality between V and N (Fig. 1C and S7). Exponential growth (framed formula) is assumed for each individual growth curve (in grey) to extract the effective doubling time t_D for each gastruloid (via linear fitting). Red line corresponds to exponential growth with mean effective doubling time $t_D = 25 \pm 1.5$ h. Red shaded area was computed from error propagation. **C**: Gastruloid volumes correlate with N_0 at all time points. Scatter plot of individual gastruloid volumes from A at different time points versus N_0, measured just after seeding, overlaid by a linear regression fit. The correlation coefficient for each fit is reported on the right y-axis. **D**: Scatter plot of mean gastruloid volume at different time points versus $\overline{N}_0$, measured just after seeding, overlaid by a linear regression fit. These are the same gastruloids shown in Fig. 2B. Right y-axis shows the correlation between volume and $\overline{N}_0$ for different time points (color). When scanning a large range of average $\overline{N}_0$ ($50 \leq \overline{N}_0 \leq 1100$), the correlations increase significantly. **E**: Effective doubling time t_D as a function of $\overline{N}_0$. The effective doubling time is obtained by fitting growth curves of the number of cells by an exponential growth model (see Methods). For round markers, t_D is extracted from cell counts measured directly by chemical dissociation. For triangle markers, cell counts are obtained from volume measurements using the relationship in Fig. 1C. Red markers correspond to the individual gastruloids in Fig. 1A; purple markers correspond to averaged data in Fig. 1B; blue markers to the inset in Fig. 1B. Average effective doubling time for gastruloids seeded with $150 \leq \overline{N}_0 \leq 1100$ is $\bar{t}_D = 28 \pm 3$ h (mean as blue dashed line; light blue area standard deviation). **F:** Evolution of mean midline length per experiment over three years (2020–2023) for gastruloids with $\overline{N}_0 = 300$ at 120 h. Upward triangles are experiments seeded by multi-pipetting; downward triangles are experiments seeded using FACS. Each experiment had between 30–40 samples. Error bars are standard deviations across individual samples. The blue line represents the overall average across all experiments with blue shaded area as the standard deviation: $\overline{L} = 590 \pm 102\,\mu$m (17%, $n = 30$). Inset shows the corresponding evolution of the variability of the mean gastruloid midline length per experiment. Intra-experiment variability in length is on average $\langle \sigma_L/L \rangle = 9.4 \pm 2.7\%$ ($n = 30$). Over three years, both the gastruloid midline length and its variability are highly consistent.

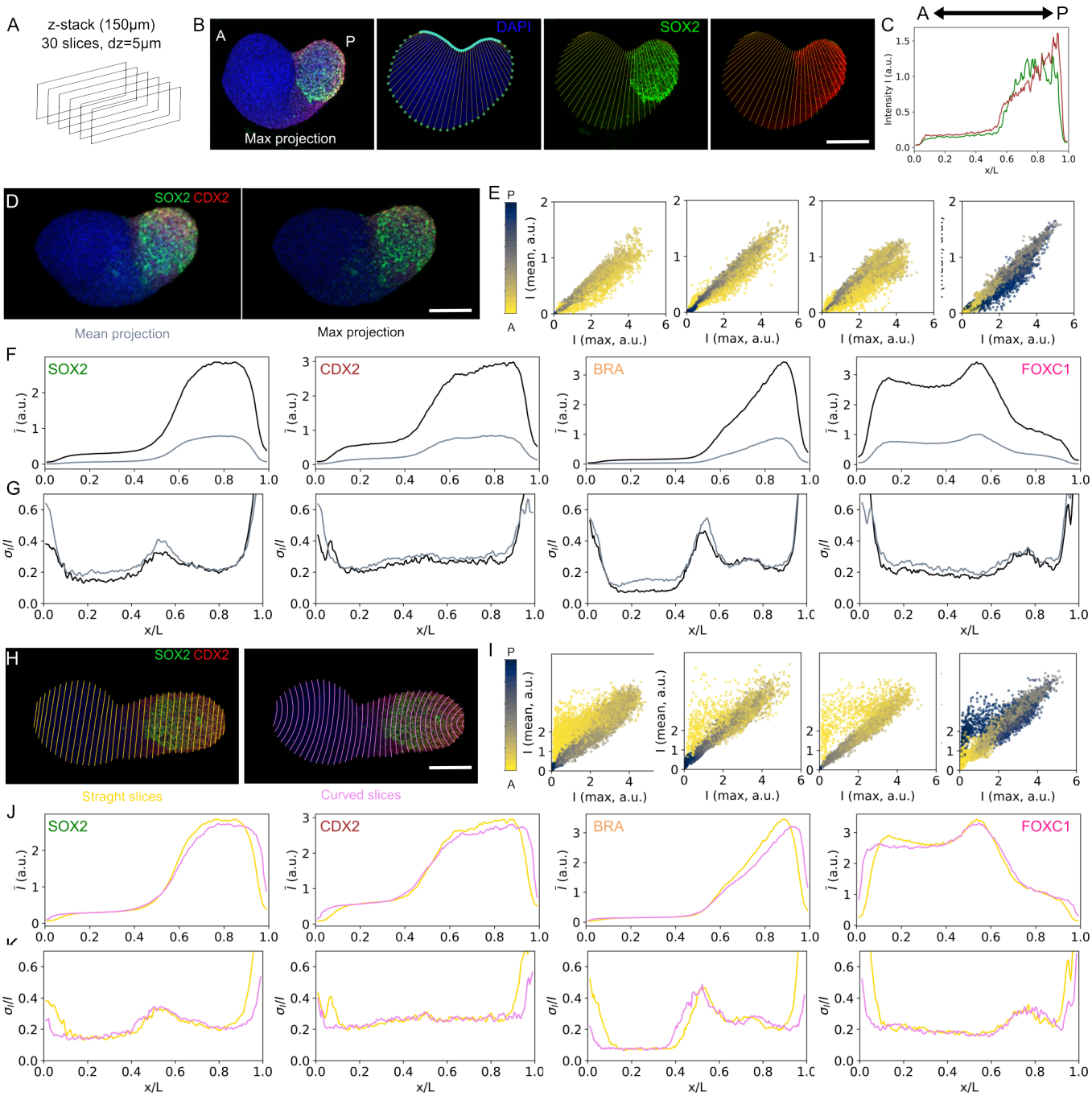

FIG. S3. Immunofluorescence image analysis. [Caption see next page.]

FIG. S3. Immunofluorescence image analysis. **A**: Fixed gastruloids are imaged by confocal microscopy in z-stacks of 150 μm (30 slices, dz = 5μm). **B**: Analysis pipeline of Fig. S1C is applied to the DAPI channel for each gastruloid to extract midline, contour, and equidistant slices. Fluorescence intensities of the other channels are max projected (here illustrated with SOX2 (green) and CDX2 (red)) and intensities of individual slices are integrated to obtain a single value per slice and to construct one-dimensional expression profiles as a function of slice position along the midline. Scalebar is 100 μm. **C**: One-dimensional profiles of SOX2 (green) and CDX2 (red) along the midline obtained for the gastruloid in B. **D**: Visual comparison of maximum (top) versus mean (bottom) projection of a gastruloid stained for SOX2 (green) and CDX2 (red). Scalebar is 100 μm. **E**: Quantitative comparison of maximum (top) versus mean (bottom) projection of intensities for the four examined genes in individual gastruloids from Fig. 2 (n={44, 44, 48, 46} respectively for SOX2, CDX2, BRA and FOXC1). Color code corresponds to the position of each slice along the midline (yellow towards the anterior pole, gray-blue towards the posterior pole). **F**: Mean profiles of expression of the four genes as a function of relative position x/L using either maximum (black) or mean (gray) projection. **G**: Variability as a function of the relative position x/L along the midline of each set of gastruloids for the four genes. Gray and black lines correspond to the variability computed respectively from either mean or maximum projections. Measured variability is lower when using maximum projection. **H**: Visual comparison of gastruloid slicing methods, straight lines (top, yellow) versus curved lines (bottom, pink); immunostained gastruloid stained for SOX2 (green) and CDX (red). Straight lines are line segments calculated between the equidistant points along both sides of the contour as in Fig. S1C. Curved lines are obtained using both equidistant points along the contour and along the midline. From this combination of points, a parabolic equation is calculated using a polynomial equation of degree 2. This procedure is meant to recapitulate the overall curvature of the gastruloid. **I**: Quantitative comparison of maximum (top) versus mean (bottom) projection of intensities for the four examined genes in individual gastruloids from Fig. 2 (n={44, 44, 48, 46} for SOX2, CDX2, BRA and FOXC1, respectively). Color code corresponds to the position of each slice along the midline (yellow towards the anterior pole, gray-blue towards the posterior pole). **J** :Comparison of mean profiles of the four stained sets of gastruloids from Fig. 2 as a function of relative position x/L using either straight (yellow) or curved (pink) line slicing. **J**: Variability as a function of the relative position x/L along the midline of each set of gastruloids for the four genes. Yellow and purple lines correspond to straight and curved line slicing, respectively. Using the curved lines method diminishes border effects on profiles of the four genes (mean and variability). No significant change is observed for the most part of the gastruloid midline, making both methods essentially equivalent. For computational simplicity, we employ the straight lines method. All profiles are represented between 0.1 and 0.9 x/L in the rest of the paper.

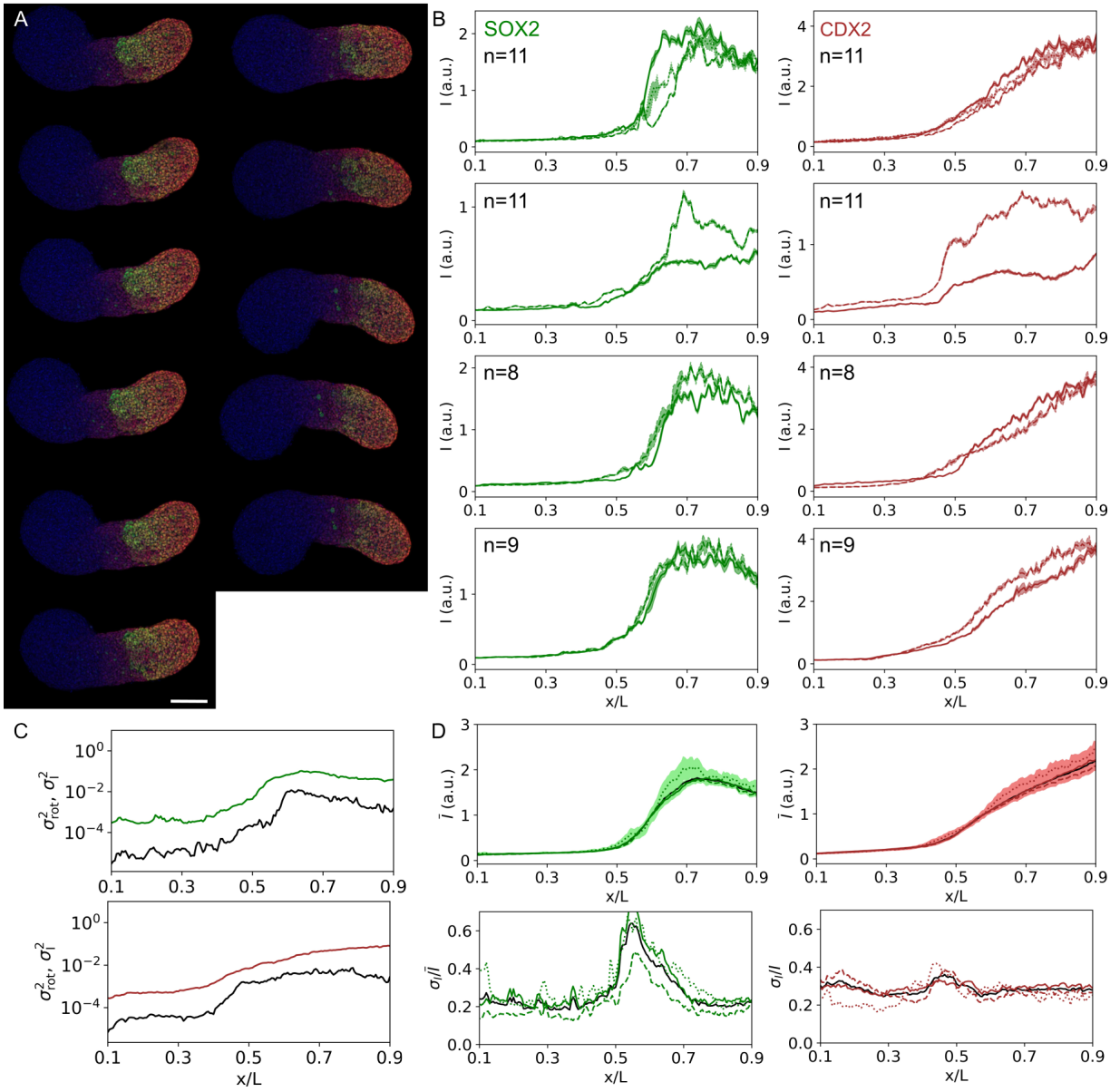

FIG. S4. Immunofluorescence measurement error estimation: specimen rotation. **A**: 11 images of a single gastruloid immunofluorescently stained for SOX2 and CDX2. Images are taken from different view angles (gastruloid is mounted in PBS medium and rotated manually via flushing for each exposure). Images are categorized for three different orientations of the gastruloid view angle: a "side view" (n=6, left column), a "top view" (n=2, top of the right column), and a "backside view" (n=3, bottom of the right column). The preferential orientation is determined by the gastruloid shape and is different from gastruloid to gastruloid. Scalebar is 100 μm. **B**: Mean profiles of SOX2 (green) and CDX2 (red) expression for each of the orientation categories observed in A: side view (full line), top view (dotted line), and backside view (dashed line). Panels in each row correspond to four experiments with a different individual gastruloid (n={11, 11, 8, 9} images, respectively). Shaded areas are standard errors in all graphs. **C**: Variance of mean profile in the four gastruloids due to specimen rotation for SOX2 (black, top) and CDX2 (black, bottom) calculated by bootstrapping data in B. This variance is compared to the total variance (SOX2: green, top; CDX2: red, bottom) of n=88 gastruloids from which the four gastruloids in B were extracted. The rotation-induced variance represents less than 10% of the total variance. **D**: Mean profiles of expression (top) and variability (bottom) for SOX2 (green) and CDX2 (red) of the n=88 gastruloids from C classified according to their orientation (line style as in B). Black lines are the mean profile and standard deviation of gene expression in the total population. This classification based on specimen rotation has no effect on the values of mean expression or on the variability.

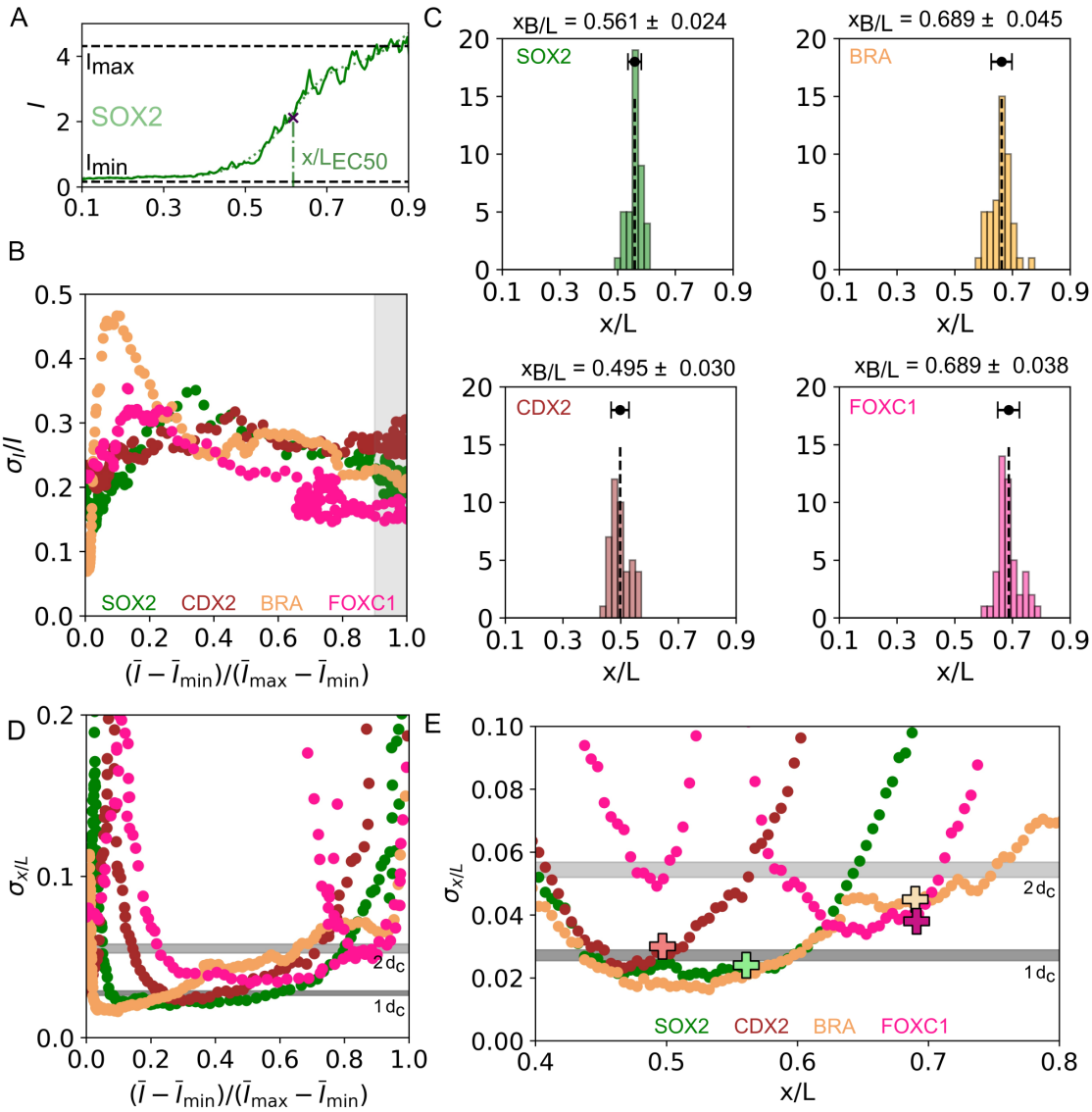

FIG. S5. Reproducibility and precision of gene expression profiles. **A**: EC50 determination: $I_{\max}$ and $I_{\min}$ for individual one-dimensional gene expression profiles are defined as the average value of the 10% largest and lowest expressing bins, respectively. The raw profile (green, plain curve) is spline fitted (green, dotted curve), and the position where the fit is equal to $(I_{\max}+I_{\min})/2$ defines x/L_{EC50}. **B**: Variability ($\sigma_I/\bar{I}$) as a function of normalized intensity $\bar{I}$. $\bar{I}_{\max}$ and $\bar{I}_{\min}$ are determined as in A. Error bars are from bootstrapping. The average value in the gray region (defined by the gene being expressed at more than 90% of its max level) is used as a measure of gene expression reproducibility for the fully *induced* gene. **C**: Distribution of x/L_{EC50} for each of the four markers. The average value is the gene boundary position x_B/L and the standard deviation around this value is a measure of the positional error of the boundary position. **D**: Generalized positional error as a function of the normalized intensity for each marker (color code as in B). The zones of highest precision (i.e., $\sigma_{x/L} \leq 5\%$) correspond to the transition regions between low- and high-expression domains. **E**: Positional error $\sigma_{x/L}$ calculated for four genes as in Fig. 3C. The positional errors at the boundaries are shown here at the mean boundary position x_B/L extracted in C (big crosses, bootstrapped errors are within marker size). The values from both methods are consistent and for all genes, the positional errors at the boundaries correspond to a linear dimension of 1–2 cell diameter (gray bands).

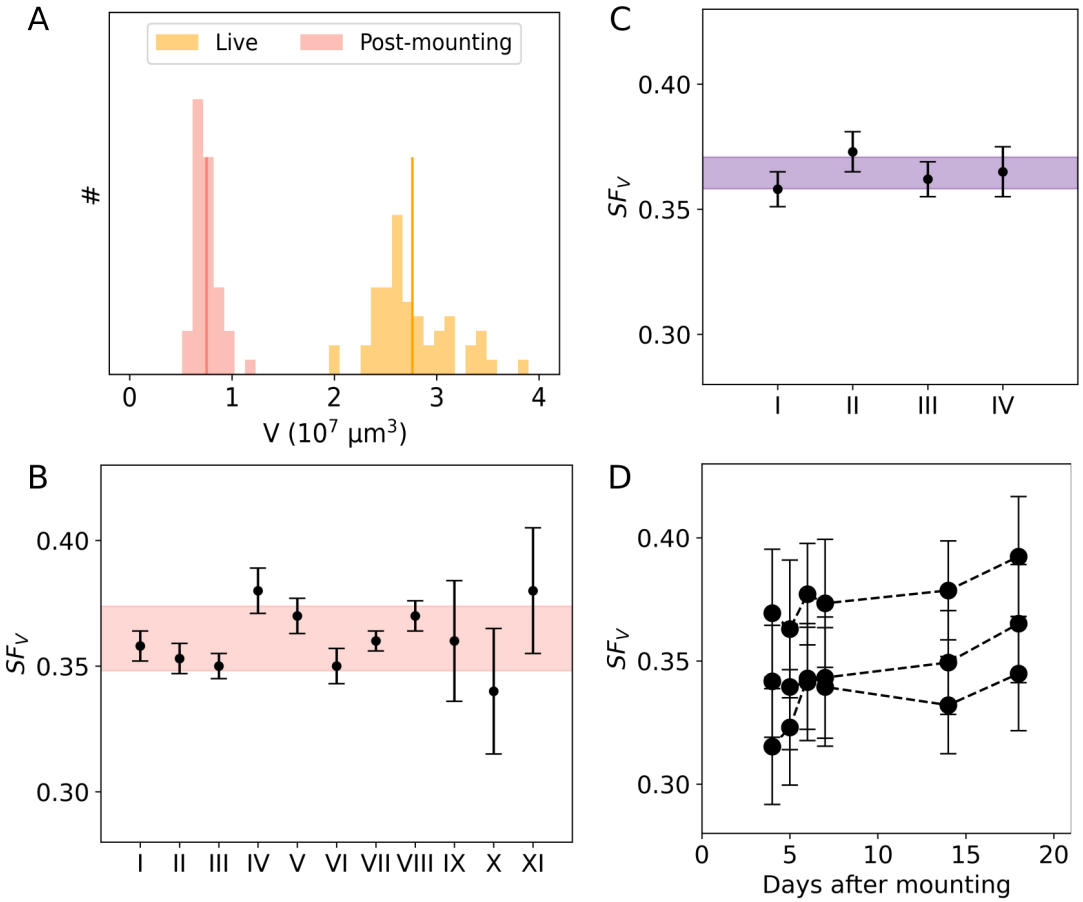

FIG. S6. Shrinkage due to fixation and sample mounting. **A**: Distribution of gastruloid volumes V_{BF} (gastruloids seeded with $\overline{N}_0 = 300$ cells) at 120 h; 2D volume reconstruction from either bright-field images or maximum projection of confocal images on the DAPI channel. Gastruloid volumes after fixation and mounting (red, n=47) are ~ 3 times smaller than the same set of gastruloids imaged live before fixation (yellow, n=52). The number of gastruloids after fixation and mounting is always smaller than during live imaging as gastruloids are lost during the protocol. **B**: A one-dimensional shrinkage factor is defined by the ratio of the average values in A: $SF_V = 1 - (V_{IF}/V_{BF})^{1/3}$. This factor quantifies by how much gastruloid size is reduced during the staining protocol. It is applied to all measured lengths of midlines from stained gastruloids. Gastruloids are mounted in 50% PBS and 50% aqueous mounting medium (Aqua-Poly/Mount, Polysciences). I–XI are 11 independent experiments where SF_V was calculated on gastruloids initially seeded with $\overline{N}_0 = 300$ cells and imaged at 120 h after seeding. Error bars are from bootstrapping with on average n=51 for live images and n=42 gastruloids after fixation and mounting (experiments I-VIII,) or n=20 for live images and n=10 after fixation and mounting (IX-XI). The shrinkage factor in these experimental conditions is of $SF_V = 0.35 \pm 0.03$. **C**: Same as B for a glycerol-based SlowFade$^{\text{TM}}$ Glass Antifade mounting medium (Invitrogen) used in the phalloidin-staining protocol. Each data point corresponds to an average gastruloid pool of n=49 for live and n=27 after fixation and mounting. Error bars from bootstrapping. Experiment I corresponds to $\overline{N}_0 = 100$ cells at at 120 h, experiments II-IV correspond to $\overline{N}_0 = 300$ at 72 h, 96 h and 120 h, respectively. The shrinkage factor in this mounting medium is $SF_V = 0.36 \pm 0.1$. Note that gastruloids are fixed for 1 h in the phalloidin-staining protocol while they are fixed for 2 h in the immunostaining protocol. **D**: Shrinkage factor stability over time for three different mounting techniques in 50% PBS and 50% aqueous mounting medium (Aqua-Poly/Mount, Polysciences): on a slide with a 250µm spacer or in a glass bottom dish w/ or w/o coverslip. Shrinkage factor measured repeatedly in the same set of gastruloids from IX-XI of B over three weeks. Error bars from bootstrapping.

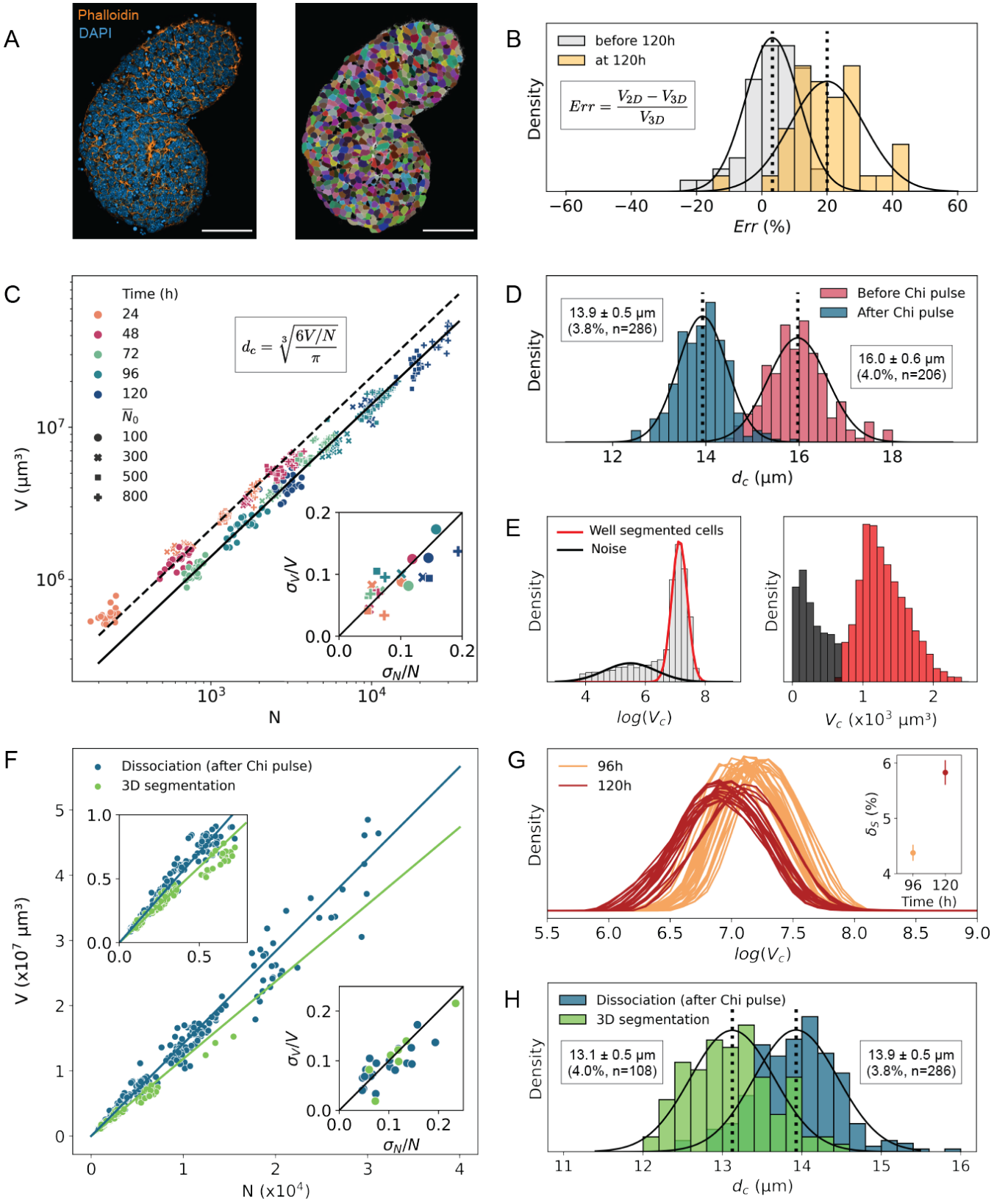

FIG. S7. Determination of total cell count and effective cell diameter. [Caption see next page.]

FIG. S7. Determination of total cell count and effective cell diameter. **A**: Visualisation of the cell masks obtained by 3D segmentation [33]. (Left:) Slice of a confocal image stack of a 120 h old gastruloid, seeded from $\overline{N}_0 = 100$ cells, stained for Phalloidin (orange) and DAPI (blue). (Right:) Phalloidin channel from left in grayscale overlaid with cell masks obtained by 3D segmentation. Scalebar is 50 μm. **B**: Estimation of the discrepancy between 3D and 2D volume reconstruction. The pipeline presented in Fig. S1C overestimates gastruloid volumes; we estimate by how much using the volume determined by 3D segmentation as a ground truth. Distribution of the error Err on the volume determined by 2D volume reconstruction V_{2D}, before 120 h and at 120 h, overlaid by a Gaussian distribution fit for each distribution. Vertical dashed lines correspond to the mean of each distribution. The ground truth 3D volume V_{3D} was obtained from the 3D segmentation. Before 120 h, $Err = 3.2 \pm 8.2\%$ ($n = 56$). After 120 h, $Err = 20.0 \pm 11.2\%$ ($n = 40$). The volume was overestimated in both time classes but more so when the gastruloid elongated. Note that this evaluation of the discrepancy between 3D and 2D volume reconstruction is independent of the shrinkage factor (Fig. S6) because 3D and 2D volume reconstructions are applied to the same shrunken gastruloid mounted with the phalloidin-staining protocol. **C**: Scatter plot of the measured volume from 2D reconstruction V (corrected for the error determined in B) versus the total cell count N obtained by chemical dissociation (with the protocol in Fig. S1E), for 492 individual gastruloids at different time points (color code) and with varying $\overline{N}_0$ (symbol). From V and N for each individually dissociated gastruloid an effective cell volume $V_c = V/N$ was computed, and from there we obtain the slope (black lines). The mean $\overline{V}_c$ for gastruloids aged from 24 to 48 h (before Chi pulse) and the mean $\overline{V}_c$ for gastruloids aged from 72 to 120 h (after Chi pulse) correspond to dashed and full lines, respectively. Inset shows correlation ($r = 0.78$) of variability for V and N for sets of gastruloids with identical age and $\overline{N}_0$. The framed formula shows how the effective cell diameter d_c is computed from the volume and cell count. It can be obtained from the distribution of V_c, or directly from the slopes (see D). **D**: Distribution of the effective cell diameters d_c per dissociated gastruloid, calculated from each effective single cell volume (V/N), before (red) and after (blue) Chi pulse. Black lines are a Gaussian fit for each distribution. Vertical dashed lines correspond to the mean of each distribution. Before Chi pulse, $d_c = 16.0 \pm 0.6\,\mu$m ($4.0\%, n = 206$); after Chi pulse, $d_c = 13.9 \pm 0.5\,\mu$m ($3.8\%, n = 286$). This is evidence of a Chi pulse-induced reduction in gastruloid's effective cell size by 13% (linear dimension). **E**: Single cell volume distributions serve to reject noisy masks from 3D segmentation results. After an initial rejection of any 3D masks smaller than 10^4 voxels, a bimodal distribution of the logarithm of single cell volumes V_c (obtained by 3D segmentation of a 120 h old gastruloid with $\overline{N}_0 = 100$) is fit by a two-component Gaussian mixture model (left). The mode in black corresponds to the distribution of small noisy masks, the mode in red corresponds to the distribution of well-segmented cells. Morphological closing is performed on the latter and the corresponding distribution of single cell volumes V_c is shown in right panel, with noise masks (black) and well-segmented masks (red). **F**: Scatter plot of gastruloid volume versus total cell count obtained by two independent methods. Blue: chemical dissociation and 2D volume reconstruction (for gastruloids dissociated after Chi pulse only). Green: 3D segmentation for volume and cell count measurement (well-segmented cells only, see E). Blue and green lines correspond to the mean V_c for chemically dissociated and 3D segmented gastruloids, respectively. Upper left inset shows a close-up for small V and N. Lower right inset shows correlation of variability for V and N for both methods. Note that the main error attached to the 3D segmentation volume is due to the estimation of the shrinkage factor of the mounting medium used in the phalloidin-staining protocol (see Fig. S5C). 2D volume reconstruction from dissociated gastruloids is applied to images of live gastruloids (i.e., they are not shrunken). **G**: Distribution of the logarithm of single cell volumes V_c obtained by 3D segmentation after filtering and reconstruction for 96 h (n=28) and 120 h (n=20) old gastruloids with $\overline{N}_0 = 100$. Inset shows dispersion self-similarity δ_S, defined as $\langle \sigma_{\log(\mathrm{V_C})} / \log(\mathrm{V_C}) \rangle$ for each set of distributions. It demonstrates the reproducibility of the dispersion in cell size in individual gastruloids and a further reduction in gastruloid cell size during the elongation process. The low variability indicates that the dispersion is highly conserved across gastruloids. **H**: Distribution of the effective cell diameter per gastruloid, obtained by chemical dissociation (only data from gastruloids dissociated after Chi pulse) and 3D segmentation, overlaid by a Gaussian fit for each distribution. Vertical dotted lines correspond to the mean of each distribution. With the dissociation protocol, $d_c = 13.9 \pm 0.5\ \mu m$ ($3.8\%, n = 286$). With the 3D segmentation method, $d_c = 13.1 \pm 0.5\ \mu m$ ($4.0\%, n = 108$). Taking into account the different sources of error and our two independent methods of determination of the effective cell diameter, the relevant linear size of the system at 120 h is $d_c = 13.5 \pm 0.8\,\mu m$.

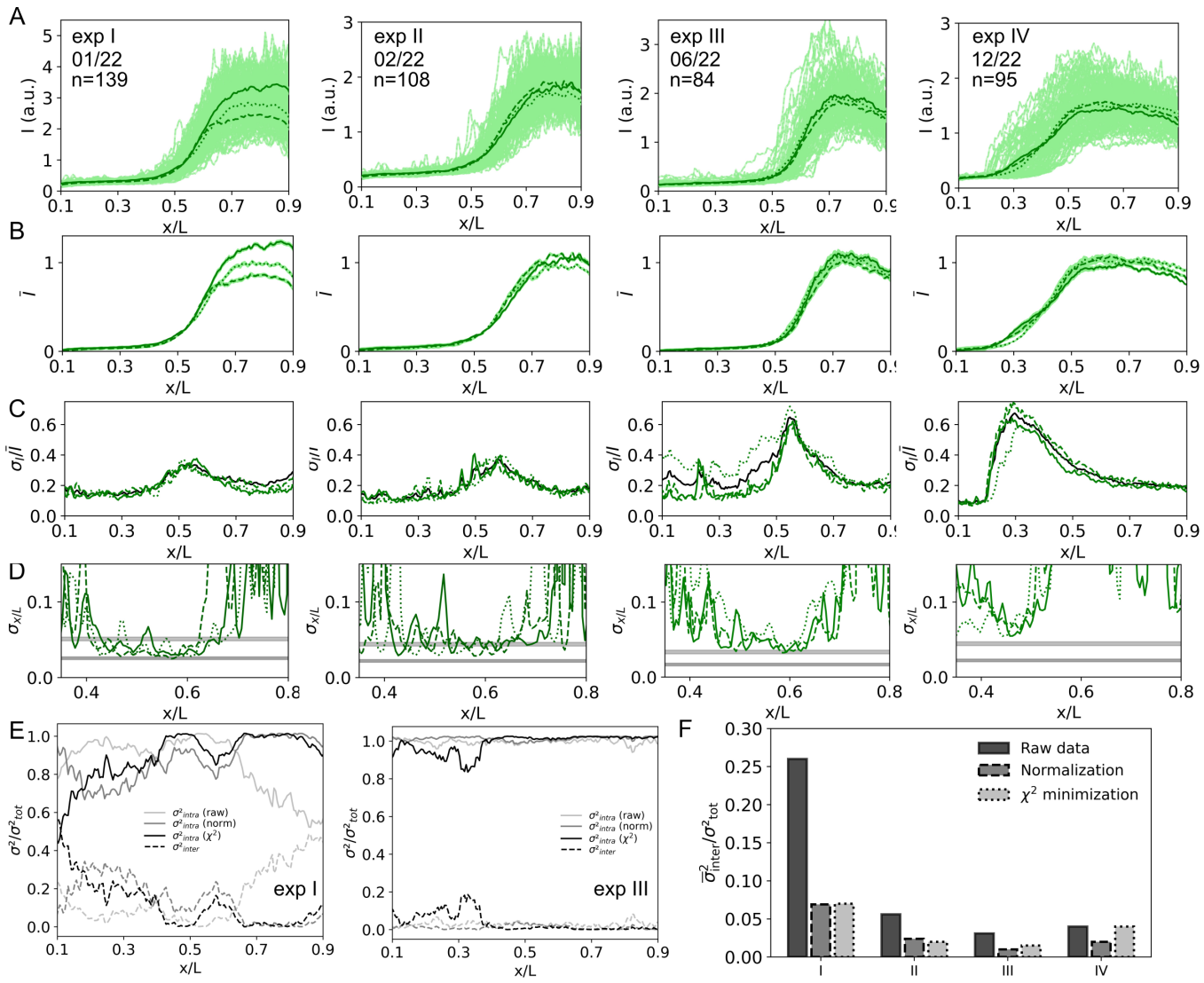

FIG. S8. Repeatability and reproducibility of a single experiment. [Caption see next page.]

FIG. S8. Repeatability and reproducibility of a single experiment. **A**: Twelve repetitions of the same experiment on different dates (exp I–exp IV, month/year, with n=139, 105, 84 and 95 gastruloids). Each panel shows raw individual gastruloids profiles (light green, no y-axis normalization) and mean profiles (dark green) of three same-day replicas of SOX2 expression in immunostained gastruloids seeded, cultured, fixed, stained, and imaged in parallel on three separate plates (i.e., in each each panel three same-day-replicas shown by full, dashed, and dotted lines). Each individual experiment (12 total) is composed of 25–50 gastruloids. Conditions are identical for all experiments except for experiment III in which gastruloids were mounted in PBS instead of Aqua-Poly/Mount. Note that same-day replicas are significantly more reproducible (i.e., self-similar) than experiments across different days (i.e., the mean expression pattern differs more across days than across same-day replicas, something *not* seen in developing embryos [34]). **B**: Mean profiles as a function of relative position x/L for each replica. Shaded areas are standard errors. Normalization was performed on the entire data set across all n gastruloids for a global maximum and minimum average intensity (i.e., a single max and a single mean for experiment day). Same-day replica can have absolute reproducibility (exp II–IV), where profile distributions collapse without y-axis normalization. **C**: Profile variability $\sigma_I/\overline{I}$ as a function of relative position x/L along the midline for each replica (green, line style as in A), or for the entire data set across same-day replicas (black). Panels run across four experiments as in A. Again, same-day replicas are highly reproducible while variability profiles differ significantly across different days. **D**: Positional error $\sigma_{x/L}$ calculated by error propagation from A and B for each replica. Gray lines correspond to one and two effective cell diameters d_c, respectively. The corresponding values in $\sigma_{x/L}$ are different between different experiments because of experiment-to-experiment variability in length (Fig. S2E). Boundary precision is maintained near 1–2 cell diameters across all replicas (i.e., same-day and across days). **E**: Variance decomposition for the SOX2 profile in experiments I and III (Methods). Plain lines correspond to the inter-plate part of the variance (for three same-day replicas) and the dashed lines to the intra-plate part of the variance. The inter-plate and intra-plate variance are represented as a fraction of the total variance of the whole population of same-day gastruloids (black lines in C). The decomposition is done in three ways: 1) on the raw profiles (black lines), 2) on normalized profiles (all profiles of individual replica are normalized by the same values, such as minimum/maximum expression levels of each replica's mean profile are set to 0/1, respectively; gray lines), and 3) on χ^2-minimized profiles (all profiles of individual replicas are normalized by the same values, obtained by χ^2-minimization of the mean profiles; light gray lines). Experiment I is an example of relative but not absolute reproducibility; experiment III is reproducible in absolute units, demonstrating that in principle the system is capable to generate absolute molarities of a gene product at well-defined positions along the gastruloid midline. **F**: Weighted average of inter-plate part of the variance, in the four experiments for either raw data, min/max normalized data, or data normalized using χ^2 minimization. Internal replicas regularly achieve absolute reproducibility (i.e., no normalization, raw data comparison) better than 5% of the total variance in the data.

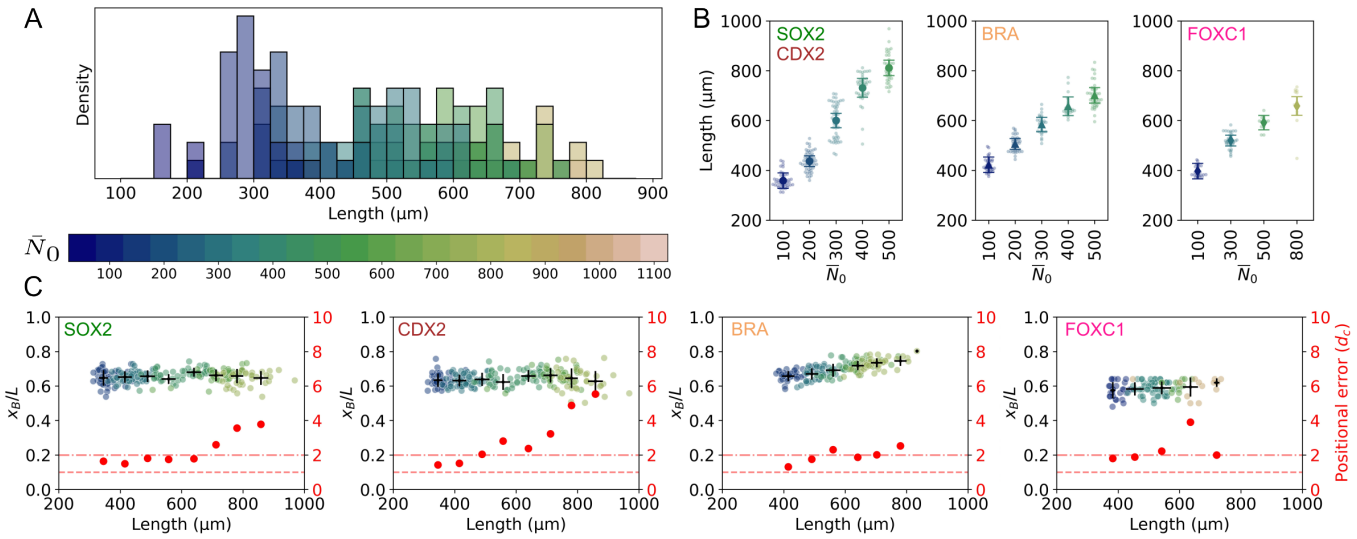

FIG. S9. Limits of precision in scaled gene expression profiles. **A**: Midline length distribution for individual gastruloids at 120 h seeded with $\overline{N}_0$ ranging from 50 to 1100 (from Fig. 1B). A 22-fold range in $\overline{N}_0$ results in gastruloids with a 3.8-fold range in average length $\overline{L}$. The full range of individual gastruloid lengths L spans 5.3-fold. **B**: Length distributions of gastruloid sets in Fig. 4 as a function of $\overline{N}_0$ (light points are individual gastruloids; dark points are average length and standard deviation per set and per gene; color code as in A). The span in length differs between experiments. For the data corresponding to SOX2 and CDX2, the 5-fold in $\overline{N}_0$ achieves a 2.3-fold in length. For the data corresponding to BRA and FOXC1, a 5-fold and 8-fold range in $\overline{N}_0$ achieves a 1.7-fold range in length, respectively. **C**: Boundary position in units of fractional midline length (x/L) for individual gastruloids as a function of length for the four genes SOX2, CDX2, BRA, and FOXC1 (seeding number color code as in A). The range of lengths is binned and the average position is plotted as a guide for the eye (black crosses indicating standard deviations on both axes). This position is independent of size for each gene, demonstrating the scaling with length of the protein profiles (a slight residual dependence after rescaling is observed in the BRA profile). The standard deviation of the boundary position in each length bin is represented in the y-direction and defines a positional error. This error is converted in cell diameter units (d_c, red y-axis), showing that the positional error remains between 1–2 cells up to a certain range of gastruloid lengths.