nature structural & molecular biology

Article

https://doi.org/10.1038/s41594-024-01251-4

Precise and scalable self-organization in mammalian pseudo-embryos

Received: 11 June 2023

Accepted: 8 February 2024

Published online: 15 March 2024

Check for updates

Mélody Merle[1,3], Leah Friedman[1,3], Corinne Chureau[1,3], Armin Shoushtarizadeh[1] & Thomas Gregor[1,2]

Gene expression is inherently noisy, posing a challenge to understanding how precise and reproducible patterns of gene expression emerge in mammals. Here we investigate this phenomenon using gastruloids, a three-dimensional in vitro model for early mammalian development. Our study reveals intrinsic reproducibility in the self-organization of gastruloids, encompassing growth dynamics and gene expression patterns. We observe a remarkable degree of control over gene expression along the main body axis, with pattern boundaries positioned with single-cell precision. Furthermore, as gastruloids grow, both their physical proportions and gene expression patterns scale proportionally with system size. Notably, these properties emerge spontaneously in self-organizing cell aggregates, distinct from many in vivo systems constrained by fixed boundary conditions. Our findings shed light on the intricacies of developmental precision, reproducibility and size scaling within a mammalian system, suggesting that these phenomena might constitute fundamental features of multicellularity.

Multicellular development entails the meticulous organization of cellular identities and body proportions in both spatial and temporal dimensions[1–3]. Gastrulation is a key event in this process, during which the body plan and the subsequent establishment of asymmetric body axes occur. Coordinated gene expression during this stage leads to reproducible patterns between individuals despite the noisiness of the underlying molecular events of gene regulation[4,5].

The challenge of translating transcriptional variability into precise and reproducible gene expression patterns has captivated research across a spectrum of animal models, from nematodes to vertebrates[6–11]. Developmental processes have been conceptualized as a sequence of steps aimed at mitigating and correcting errors in the face of molecular noise[4,12]. In vertebrates, mechanisms such as differential specification rates and cell sorting have been described as error-correction strategies[13,14].

However, in the context of the early fly embryo, the precision of macroscopic body plan features can be traced back to the precision of maternal inputs[15]. Exemplified by the morphogen gradient of Bicoid[16], this precision is transmitted at the single-cell level along the major body axis to zygotic genes before gastrulation[17]. Such precision in flies challenges the limits of molecular noise[18,19], suggesting that successive developmental stages may have evolved to minimize noise transmission at each step, both across evolutionary time scales and within the spatiotemporal boundaries of individual organisms[20].

An intriguing consequence of this precision is the scaling of gene expression patterns relative to system size[3,21]. Scaling, observed in both invertebrates and vertebrates[22–25], entails the preservation of body plan proportions among different individuals. During development, scaling manifests at various levels, encompassing morphogenetic movements, gene expression domains and other phenomena, reflecting the intricate interplay of regulatory mechanisms[24,26–31].

In contrast to organisms with well-defined developmental boundaries, such as flies, frogs or worms, mammalian development relies on self-organization and continuous growth. Quantitative assessments of reproducibility, precision and scaling in mammalian systems have been limited, prompting inquiries into whether the precision observed in flies is even necessary, as mammals rely on different developmental mechanisms. These properties have been found in other

[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. [3]These authors contributed equally: Mélody Merle, Leah Friedman, Corinne Chureau. e-mail: thomas.gregor@pasteur.fr

types of self-organizing systems, such as scaling in flatworms during regeneration[32], suggesting that scaling could also be achieved during self-organization-driven development.

Recent progress with in vitro models[33,34], including gastruloids[35], derived from mouse embryonic stem (mES) cells, presents promising avenues for investigation. These three-dimensional pseudo-embryos mimic critical events of mammalian gastrulation through self-organized patterning. They break symmetry and elongate along an axis that resembles the most posterior part of the mouse embryo's anterior–posterior (AP) axis[36], and they can be cultivated in substantial quantities, rendering them conducive to quantitative approaches[37,38]. Yet, concerns regarding the reproducibility of these systems have been raised[39].

In this article, we investigate the regulatory precision of self-organizing multicellular systems using gastruloids as a model for mammalian development. Our findings reveal highly reproducible gene expression patterns and growth dynamics in gastruloids, showcasing their potential as a quantitative model. We observe precise control of gene expression along the anterior–posterior axis, with single-cell precision in pattern boundary positioning. Additionally, gastruloid growth scales precisely with the initial number of seed cells. These results highlight the intrinsic reproducibility and precision of gastruloid self-organization, underscoring their utility for studying fundamental aspects of mammalian development.

Results

Reproducible gastruloid growth and size scaling

Quantitative analysis in mammalian development faces a substantial hurdle due to the inherent difficulty of achieving experiment reproducibility, particularly when replicating experiments under identical conditions is challenging. The protracted and often inaccessible nature of embryos further complicates this issue. However, the gastruloid model presents a promising solution to these challenges, offering the capability to culture hundreds of specimens simultaneously and under equivalent conditions.

Under tightly controlled experimental conditions, we have maximized the degree of reproducibility in gastruloid cultures. Within the confines of the original protocol[40], we achieved a 97% success rate in inducing the elongation of gastruloids along a single axis (Extended Data Fig. 1). To further assess the intrinsic reproducibility of the self-organization processes within these systems, we examined general physical properties, including the uniformity of growth and the influence of initial conditions. We quantified growth dynamics by monitoring the length of the midline, the total volume and the total cell count of individual gastruloids over 5 days (Extended Data Fig. 1 and Methods).

The growth curves of individual gastruloids exhibit a remarkable degree of consistency, as evidenced by the overlapping measurements for both length and volume over time (Fig. 1a and Extended Data Fig. 2a). This observation underscores the high reproducibility of growth dynamics at various time points. The residual variability observed in these growth curves can, in part, be attributed to two primary factors. First, it is influenced by the variability in the initial number of seeded cells N_0 (Extended Data Fig. 1a). Additionally, fluctuations in the effective doubling time, which we have measured to average at 26.4 ± 1.7 h for $\overline{N}_0 = 300$ (Extended Data Fig. 2b), contribute to this variability.

The volumes of gastruloids display a significant correlation with the initial number of seeded cells (N_0) at all measured time points (Extended Data Fig. 2c). Moreover, when we extend our investigation to encompass a wide range of average $\overline{N}_0$ values (up to 22-fold changes), these correlations become far more pronounced (Extended Data Fig. 2d). It is worth noting that, as the average $\overline{N}_0$ increases, the percentage of gastruloids exhibiting multipolarity also increases (Extended Data Fig. 1d). Yet, our volume measurement algorithm is not applicable to multipolar gastruloids, introducing a bias in the data, which becomes more apparent as $\overline{N}_0$ increases. Nevertheless, a critical finding emerges from our data: growth curves of gastruloid volumes, when normalized by $\overline{N}_0$, consistently collapse (Fig. 1b). This convergence serves as compelling evidence that growth is not only reproducible across all observed time points but also scales in relation to the initial number of seeded mES cells.

A similar relationship becomes apparent when examining the growth curves of the total number of cells, N (obtained from chemically dissociated gastruloids, Extended Data Fig. 1e). For varying initial $\overline{N}_0$ values, the growth curves collapse, as illustrated in Fig. 1b (inset), while the effective doubling time remains consistent (Extended Data Fig. 2e). The residual spread in these overlapping growth curves can be attributed to other factors, such as variability between experiments (Extended Data Fig. 2e).

The observed common relationship between growth and developmental time hints at the emergence of size control as an intrinsic property within gastruloids. Interestingly, it suggests that a refinement mechanism aimed at achieving specific sizes with error reduction may not be a necessary component of this system. Instead, the scaling of gastruloid size with the initial seed number implies a departure from the principles governing mouse embryos[41–43]. Unlike the self-regulation mechanisms observed in mouse embryos, our findings suggest that gastruloid growth dynamics operate independently of system size, with absolute size (both in terms of cell number and volume) being directly linked to the initial seeding number.

To substantiate this assertion, we conducted a direct assessment of the extent to which gastruloid volume can serve as a predictor of the total cell count. Measurements of cell count and volume within the same gastruloid revealed a strong, linear relationship that remained consistent across various time points and different average $\overline{N}_0$ values (Fig. 1c). This strong correlation suggests that the inherent dispersion in cell size is remarkably conserved across diverse gastruloids and under varying external conditions (Extended Data Fig. 8g).

To further validate our observations, we compared our results (Fig. 1b) with the hypothesis of perfect scaling. Perfect scaling, in this context, denotes a linear relationship between $N(t)$ and N_0. When we represent these values with respect to the reference seeding number $\overline{N}_0 = 300$ (Fig. 1d), perfect scaling is achieved by a slope of 1 (black line). This signifies that starting with twice as many cells (in units of $\overline{N}_0 = 300$) results in precisely twice as many cells at any given time point. We utilized the measured errors associated with both the initial seeding number and the doubling time to estimate the expected error at 120 h.

For nearly all stages of gastruloid growth, the data points fall within the boundaries of expected deviations. Notably, at earlier time points (24 h and 48 h), growth scaling is observed in a range of $\overline{N}_0$ values spanning between 100 and 600 (Extended Data Fig. 2g). In the case $\overline{N}_0 = 300$ (Fig. 1d, inset), the slope is statistically indistinguishable from one at all time points. These results collectively underscore the high fidelity of the monitored 5-day growth process. Under carefully controlled experimental conditions, gastruloids exhibit a remarkable capacity for self-organization with meticulous control over the variability in growth rate (Extended Data Fig. 2e) and other noisy processes of size regulation.

Reproducible gene expression patterning

Coordinated growth and, in particular, axis elongation in gastruloids is closely associated with the expression patterns of genes along the body axes. Consequently, we ask how the observed physical properties of reproducibility and scaling manifest in the AP patterning of gene expression. To this end, we conducted gene expression profile measurements 5 days after seeding, a time point when the pseudo-AP axis is morphologically well established. Our focus was on the AP patterning of four germ-layer markers, namely *Bra*, *Cdx2*, *FoxC1* and *Sox2*. These markers

https://doi.org/10.1038/s41594-024-01251-4

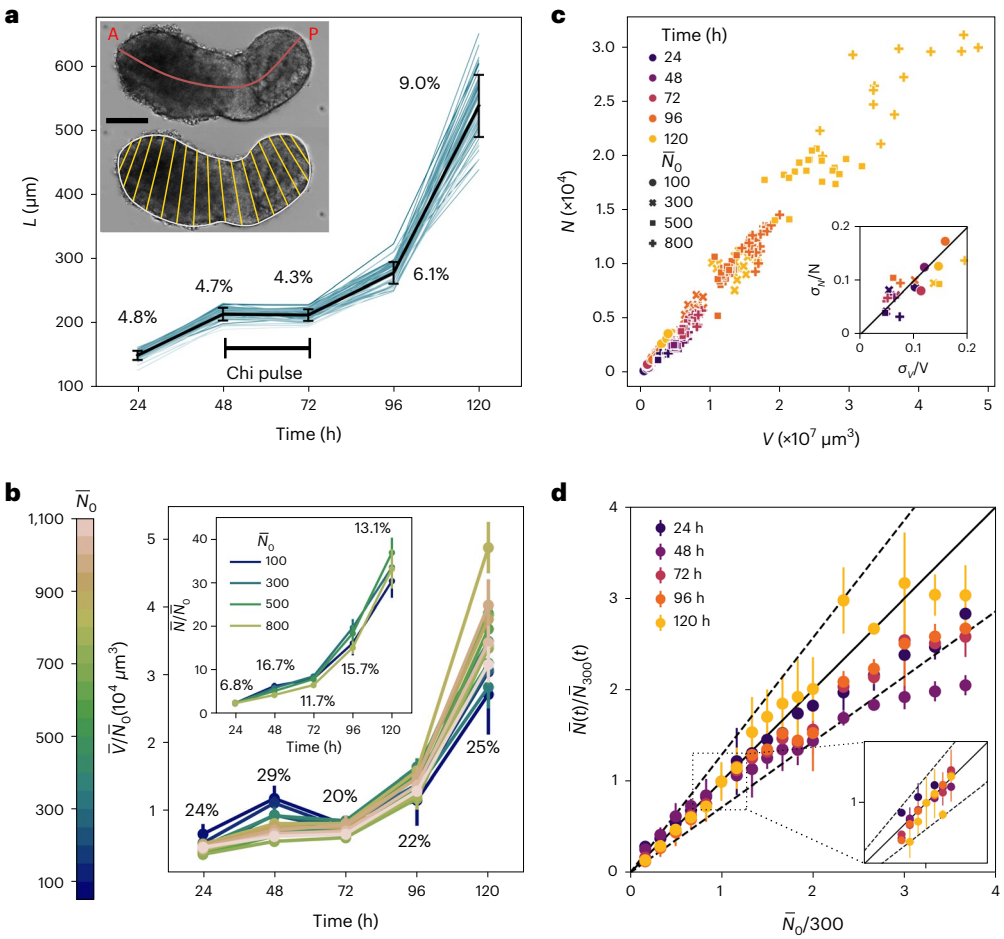


Fig. 1 | Reproducible gastruloid growth, scaled to system size. a, Gastruloid midline length variation as a function of time. Curves shown for 57 gastruloids followed individually over time (blue) and mean (black). Percent variation around the mean is reported for each time point. Spread of initial number of seeded mES cells is $N_0 = 305 \pm 28$ cells (Extended Data Fig. 1a; for an equivalent relationship for volume, see Extended Data Fig. 2a). Inset shows a brightfield 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); scale bar, 100 μm; also see Extended Data Fig. 1b–d. **b**, Gastruloid volume and total cell count (inset and Extended Data Fig. 1e) as a function of time. Volumes are normalized by the average number of initial seed cells $\bar{N}_0$ at time zero (color code). Each line represents the mean of on average 15 gastruloids with the same $\bar{N}_0$. Percentages correspond to residual variations within which normalized volumes collapse for 17 different values of $\bar{N}_0$. Similar collapse for normalized gastruloid cell counts for four values of $\bar{N}_0$ (inset). See Supplementary Tables 1 and 2 for sample numbers. **c**, Scatter plot of total cell count versus the measured volume for 492 individual gastruloids at different time points (color code) and with varying $\bar{N}_0$ (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 $\bar{N}_0$. This is evidence that the independent methods for measuring N and V are accurate estimates of gastruloid growth. **d**, Cell count $\bar{N}(t)/\bar{N}_{300}(t)$ as a function of the initial seed cell count $\bar{N}_0/300$ in units of the reference seed at $\bar{N}_0 = 300$. Time is encoded by color (see legend). Black diagonal (slope of 1) represents perfect scaling of gastruloid size at time t upon changes in $\bar{N}_0$ ranging over $50 \leq \bar{N}_0 \leq 1100$. Dashed line estimates expected deviations from perfect scaling at 120 h due to fluctuations in $\bar{N}_0/300$ and in the doubling time t_D given a simple exponential growth model (Extended Data Fig. 2e and Methods). Detailed representations for individual time points can be found in Extended Data Fig. 2g. Inset shows the same relationship centered around $\bar{N}_0 = 300$ where the regression slope is statistically indistinguishable from one at all time points (Supplementary Table 3). See Supplementary Tables 1 and 4 for sample numbers. All error bars are standard deviations.

are well documented for their pivotal roles in the differentiation of tissue progenitors and the establishment of the AP axis during gastrulation[44–47].

We performed immunofluorescence staining (Fig. 2a), and from two-dimensional (2D) maximum projections of a confocal image stack, we extracted one-dimensional (1D) intensity profiles projected isometrically along each gastruloid's midline (Fig. 2b, Extended Data Fig. 3b and Methods). Unexpectedly, individual profiles for all four genes in mammalian systems are closely clustered around the average profile. The overall variability in these profiles is notably small (Fig. 2c), peaking in regions where gene expression levels change sharply over short distance intervals, such as boundaries between high and low expression domains.

In pioneering experiments with engineered systems, it has been shown that, even when gene expression is highly induced, the resulting expression levels fluctuate[48,49]. We observe something similar in the regions where our examined genes are expressed at their maximal levels (Extended Data Fig. 5b), with the variability hovering around 20% (Fig. 2c). This variability near the maximum expression level has also been observed in other organisms[18,50].

The observed noise may, in principle, result from measurement errors. In separate experiments (Methods and Extended Data Figs. 3, 4, 6 and 8), we estimate the component of measurement noise arising in the experimental process. Some sources of experimental noise are inherent to the immunofluorescence staining and imaging processes[51] (Extended Data Fig. 4), while others result from arbitrary choices made during image analysis routines, such as axis definition and the projection method for measuring gene expression patterns (Extended Data Fig. 3). Overall, we estimated that all sources of measurement error

https://doi.org/10.1038/s41594-024-01251-4

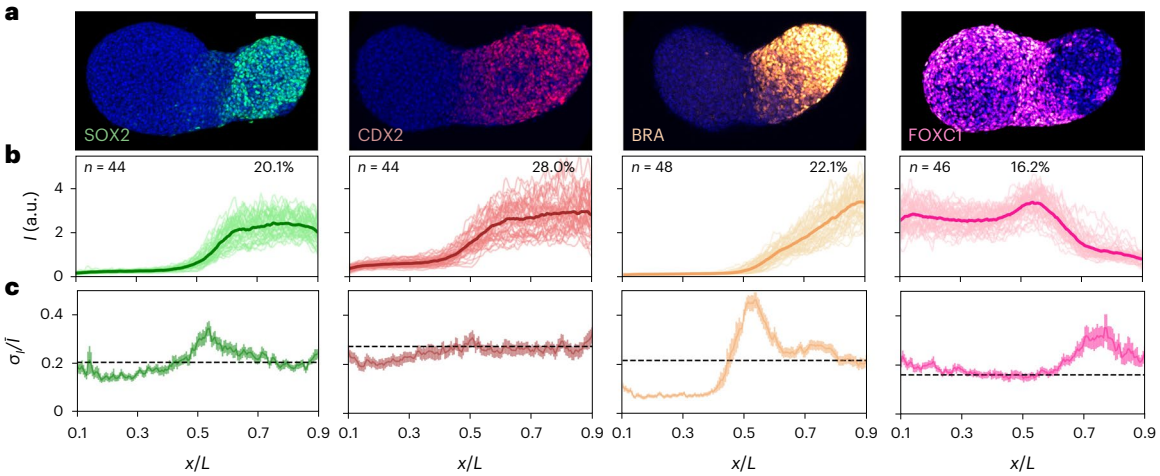


Fig. 2 | Reproducibility of gene expression patterns in gastruloids. **a**, Maximum projections of four confocal image stacks of 120 h old gastruloids stained by immunofluorescence for SOX2, CDX2, BRA and FOXC1. AP axis is in a left–right orientation. Scale bar, 100 µm. **b**, *n* individual raw gene 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** as a function of relative position along the midline x/L. Error bars are standard errors of the mean obtained from bootstrapping. Dashed lines represent the average variability in the region where genes are most highly expressed (values in **b**; Extended Data Fig. 5b).

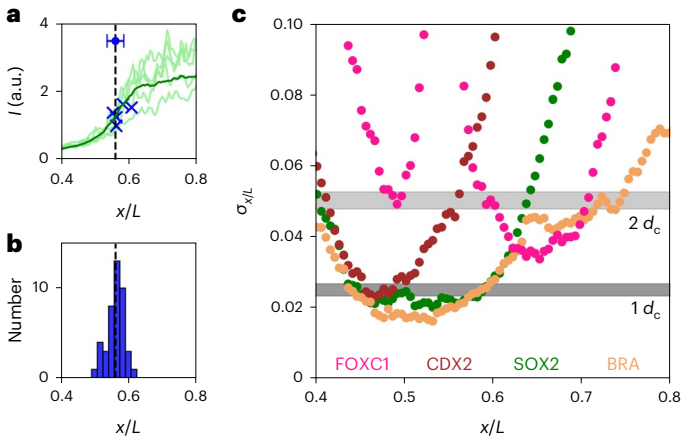


Fig. 3 | Single-cell pattern boundary precision in gastruloids. **a**, Close-up of Fig. 2b for five SOX2 expression profiles as a function of position along the midline (green), with pattern boundary positions of individual profiles marked at the half-maximal expression value (EC_{50}, blue crosses). Mean profile of gastruloid midlines in dark green ($n = 44$). Dashed line is at mean position for these five profiles. Blue point is from distribution in **b** with standard deviation. **b**, Distribution of SOX2 pattern boundary positions from Fig. 2b. The mean defines the pattern boundary position x_B ($n = 44$); the standard deviation of this 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, $\sigma_{x/L}$ (Methods). Color code as in Fig. 2; gray areas correspond to one and two effective cell diameters d_c, respectively, including measurement errors (Extended Data Figs. 6 and 7).

combined correspond to less than 10% of the total variance (Extended Data Fig. 4 and Methods). The values reported here thus represent an upper bound for the biological variability of the system, and the true value is even lower (Extended Data Fig. 5c).

Single-cell precision of pattern boundaries

During development, cells rely on patterning signals executed by genes like those analyzed above[45,46]. However, inherent variability between individual gastruloids (Fig. 2b) limits the precision with which cells can execute their functions and fates at specific positions. We estimate the positional precision for the four analyzed genes by determining the positions along the midline where the half-maximal expression level is reached within the boundary regions for each patterning gene (Fig. 3a and Extended Data Fig. 5a). For instance, in the case of SOX2, we observe a narrow distribution of these positions (Fig. 3b), with a standard deviation of only 2.4%. The other genes exhibit a similar level of boundary precision (Extended Data Fig. 5d).

Instead of focusing solely on a single boundary point, a more comprehensive approach involves considering the entire extent of the pattern and translating the fluctuations in expression levels (Fig. 2c) into positional errors (Fig. 3c and Methods)[19]. This broader analysis reveals that a positional precision of 2–4% is achieved within domains spanning approximately 5–10% of the gastruloid length. These domains align with the respective boundary regions for each gene (Extended Data Fig. 5e). The values obtained through both methods are consistent at the mean pattern boundary positions (Extended Data Fig. 5f). In principle, this precision allows cells to use the expression levels of these genes to precisely determine their positions along the pattern boundary.

To gain insight into the practical significance of achieving 2–3% spatial precision along the midline, we measured the average size of individual cells within gastruloids. We revisited our simultaneous measurement of cell count *N* and volume *V* for several hundred gastruloids (Fig. 1c). The strong linear relationship $N = sV$, with a slope *s* representing the inverse of the mean cell volume, allowed us to determine the effective diameter of cells in developing gastruloids. To validate this measurement, we employed high-resolution 3D reconstructions of individual gastruloids[52], in which we fluorescently labeled all cell membranes (Extended Data Fig. 7). The consensus between both methods yielded an effective cell diameter of $d_c = 13.5 \pm 0.8\,\mu m$ (after 72 h of development). This value serves as the relevant linear size unit for the system.

With this system-intrinsic length scale measurement, we determined that the achieved patterning precision corresponds to one to two cell diameters along the midline of the gastruloid (Fig. 3c). This finding demonstrates that mammalian gastruloids exhibit patterning precision on par with patterning systems in fly embryos[18], worms[53] and ascidians[11]. In all these systems, the positional error resulting from gene expression fluctuations allows for distinguishing neighboring cells.

https://doi.org/10.1038/s41594-024-01251-4

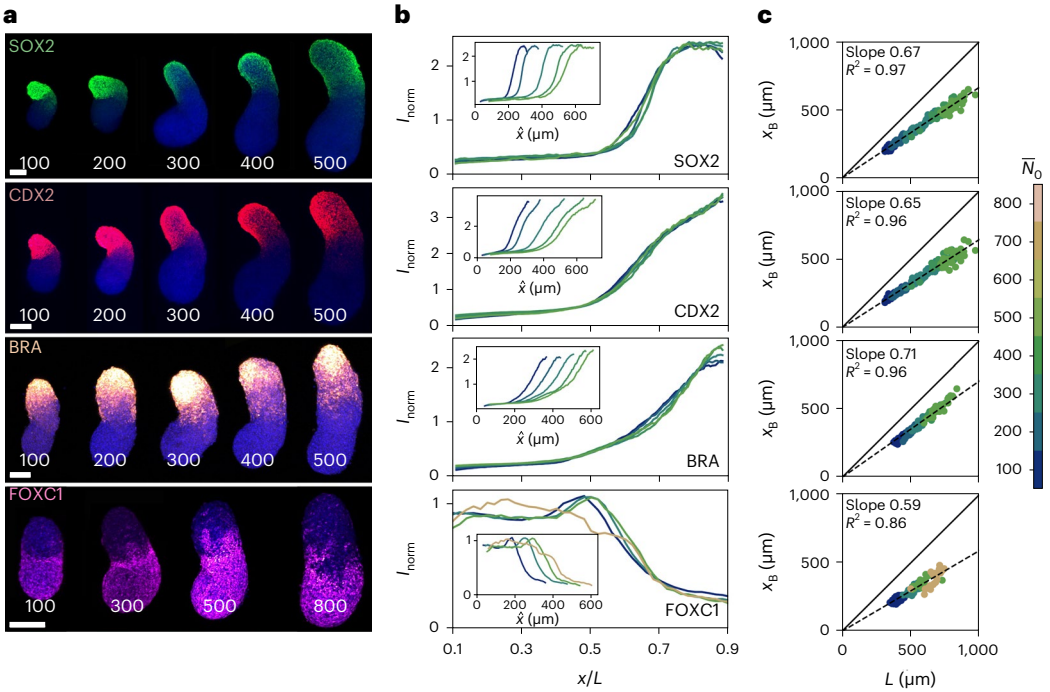


Fig. 4 | Scaling of AP gene expression patterns in gastruloids. a, Confocal images of gastruloids immunofluorescently stained for four different genes each representing a different initial seed number $\overline{N}_0$ (in white, also Extended Data Fig. 9). AP axis from bottom to top. Scale bars, 100 µm. **b**, Normalized mean expression profiles for sets of gastruloids with the same $\overline{N}_0$ (color code as in **c**) 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 $\hat{x}$ of the respective set. **c**, Boundary position x_B in absolute units of individual gastruloids seeded with varying $\overline{N}_0$ (same color code as in **b**) as a function of absolute individual gastruloid length L. Bold diagonal line indicates gastruloid length ($x_B = L$). Dashed line shows linear fit with intercept at zero. Perfect scaling would imply $R^2 = 1$, meaning that 100% of the observed boundary position variance is related to gastruloid length. Here, for the genes *Sox2*, *Cdx2* and *Bra*, the scaling relationship with respect to gastruloid length explains 96–97% of the boundary position variance. See Supplementary Table 5 for sample numbers.

These levels of reproducibility and precision remain consistent for gastruloids grown in parallel from the same population of cells (Extended Data Fig. 8). By minimizing sources of variability throughout the entire chain of experimental protocols, from gastruloid seeding to imaging, we obtained profiles with very similar average absolute concentration levels, variability and positional error in multiple experiments and for various genes. Note that we are reporting the total variance, which includes measurement errors. Therefore, the actual values for the reproducibility and precision 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 equivalent to the linear dimension of a single cell, arguably the maximum precision that is functionally beneficial for a multicellular system.

Gene expression pattern scaling with size

To comprehend the relationship between the growth dynamics of these self-organizing structures and their gene expression profiles, we examine how gene expression patterns scale with gastruloid size. This will shed light on the system's capacity to sustain proportional patterns as it undergoes growth. Despite variations in gastruloid lengths by 7–11% 5 days after seeding (Fig. 1a and Extended Data Fig. 2f), the relative positional error in gene expression boundaries remains below 3%. This suggests that the mechanisms underlying pattern formation in gastruloids can adapt to system size[54]. To directly test this hypothesis, we deliberately manipulated the sizes of gastruloids by adjusting the initial number of seeded cells. Within the range where elongation results are most robust, we achieve up to a 2.3-fold change in gastruloid lengths (Fig. 4a and Extended Data Fig. 9).

For each of the four genes studied earlier, we create sets of immunofluorescently labeled gastruloids with different $\overline{N}_0$. Upon plotting the average gene expression profiles for each $\overline{N}_0$ set against absolute units, they exhibit dispersion along the x axis in direct proportion to the corresponding average gastruloid length (Fig. 4b, inset). However, upon normalization by the mean gastruloid length within a given set, the average profiles collapse (Fig. 4b), which can also be seen in individual profiles (Extended Data Fig. 9c,d). These findings indicate a linear relationship with a zero intercept between the absolute boundary position (as defined in Fig. 3a) and the length in absolute units (Fig. 4c). Our ability to consistently cultivate gastruloids with varying initial numbers of seeded cells N_0 enables us to evaluate this scaling relationship across a broad range of gastruloid lengths spanning approximately 600 µm.

We can further quantify the scaling effect by examining the position of other key points along the gene expression profiles, such as positions where the intensity equals 25% and 75% of the maximum expression (Extended Data Fig. 10). We observe that our smallest (~300 µm) and largest (~900 µm) gastruloids display boundary shifts of one cell diameter compared to the $\overline{N}_0 = 300$ reference case, with the exception of BRA, where the shift amounts to four cell diameters. Therefore, at each relative position along the gastruloid's midline, a cell consistently produces and maintains an absolute amount of protein with an accuracy within a few tens of percent of its mean value. These results indicate that the expression patterns of these four genes contain information locally about the overall length of the entire system. Furthermore, the residual positional error after rescaling to relative coordinates is consistently within one to two cell diameters (Extended Data Fig. 10d), for gastruloids with $100 \leq \overline{N}_0 \leq 500$. Note that this falls within the range for which we observed growth scaling. This level of

precision matches the inherent precision of the pattern boundaries, showcasing a remarkable sensitivity to global parameters.

Discussion

Our results reveal intriguing properties of mammalian cell aggregates and their potential implications for developmental biology and regenerative medicine. They underline the intrinsic potential for reproducible, precise and scalable self-organization in gastruloids. These properties seem to go beyond the development of organisms like worms, flies or frogs and are also achieved by mammalian systems, shedding light on the existence of a general principle of pattern formation that acts at the multicellular scale independently of boundary conditions[55].

The remarkable reproducibility, precision and scalability observed in gastruloids carry profound biological implications. They challenge our conventional understanding of mammalian development, hinting at the precise regulation of developmental features, such as gene expression patterns, during self-organized processes. Additionally, these findings suggest that reproducibility and scaling, evident in both developing embryos and synthetic structures such as gastruloids, may represent context-independent properties[56,57]. Collectively, these insights point toward fundamental principles that govern self-organization processes in multicellular systems.

These properties underscore the importance of exploring how these principles relate to the complex and dynamic processes of in vivo development in mammals. Our findings raise questions about the extent to which these principles govern embryonic development and tissue formation. The fact that synthetic systems exhibit intrinsic reproducibility and precision similar to in vivo systems expands the possibilities for advanced engineering applications in the field of organoids and, more broadly, cell aggregates[58–61].

In mammals, achieving these findings is currently not feasible in vivo; only in vitro systems offer the necessary experimental accessibility and manipulability. Utilizing an in vitro system not only provides essential experimental access but also enables precise control over parameters, including system size, and facilitates the ability to continuously perturb the system beyond its natural limits. Our approach and findings suggest that the gastruloid model, and possibly other stem-cell-derived aggregates, have the potential to serve as a powerful tool for quantitative studies of mammalian development and various other biological processes.

It is important to note that, while our findings highlight the surprising properties of mammalian cell aggregates, they do not necessarily apply to all features of these systems. For instance, gastruloid shape is not consistently reproducible (as evident in Figs. 2a and 4a). Gastruloid length, which is a component of shape, exhibits more variation at 120 h compared to pattern boundary positions. This distinction separates self-organized processes, such as the emergence of gene expression patterns in these cell aggregates, from those that are predominantly influenced by external and environmental conditions and, therefore, not subject to stringent control[13,62]. Only the former appear to be subject to precise regulation.

While the study highlights the potential of gastruloids as a model system, we acknowledge that it has limitations. Gastruloids, while powerful tools, are not identical to in vivo embryos, and thus there are aspects of development they cannot fully replicate. Additionally, translating these findings to in vivo contexts remains a challenge[63,64]. Nonetheless, the study underscores the value of using in vitro systems as accessible and controllable models for studying development.

In conclusion, our findings provide a fresh perspective on the properties of self-organization in mammalian cell aggregates. These properties are not confined to invertebrate model systems such as flies or worms but could be context independent, spanning millions of years of evolutionary change. Understanding these principles can transform our approach to developmental biology, advancing our knowledge of embryonic development in mammals. Gastruloids and similar in vitro systems, with their accessibility and controllability, are poised to play a pivotal role in future research in this field, offering opportunities for quantitative exploration and applications.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-024-01251-4.

References

1. Conklin, E. Organ-forming substances in the eggs of ascidians. *Biol. Bull.* **8**, 205–230 (1905).
2. Kirschner, M. & Gerhart, J. *Cells, Embryos and Evolution* (Blackwell Science, 1997).
3. Houchmandzadeh, B., Wieschaus, E. & Leibler, S. Establishment of developmental precision and proportions in the early *Drosophila* embryo. *Nature* **415**, 798–802 (2002).
4. Arias, A. M. & Hayward, P. Filtering transcriptional noise during development: concepts and mechanisms. *Nat. Rev. Genet.* **7**, 34–44 (2006).
5. Briscoe, J. & Small, S. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996–4009 (2015).
6. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
7. Bollenbach, T. et al. Precision of the Dpp gradient. *Development* **135**, 1137–1146 (2008).
8. Bier, E. & De Robertis, E. M. Embryo development. BMP gradients: a paradigm for morphogen-mediated developmental patterning. *Science* **348**, aaa5838 (2015).
9. Bentovim, L., Harden, T. T. & DePace, A. H. Transcriptional precision and accuracy in development: from measurements to models and mechanisms. *Development* https://doi.org/10.1242/dev.146563 (2017).
10. Zagorski, M. et al. Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* **356**, 1379–1383 (2017).
11. Guignard, L. et al. Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis. *Science* **369**, eaar5663 (2020).
12. Waddington, C. H. Canalization of development and the inheritance of acquired characters. *Nature* **150**, 563–565 (1942).
13. Kicheva, A. et al. Coordination of progenitor specification and growth in mouse and chick spinal cord. *Science* **345**, 1254927 (2014).
14. Tsai, T. Y.-C. et al. An adhesion code ensures robust pattern formation during tissue morphogenesis. *Science* **370**, 113–116 (2020).
15. Petkova, M. D., Little, S. C., Liu, F. & Gregor, T. Maternal origins of developmental reproducibility. *Curr. Biol.* **24**, 1283–1288 (2014).
16. Driever, W. & Nüsslein-Volhard, C. The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell* **54**, 138–143 (1988).
17. Petkova, M. D., Tkacik, G., Bialek, W., Wieschaus, E. F. & Gregor, T. Optimal decoding of cellular identities in a genetic network. *Cell* **176**, 844–855.e15 (2019).
18. Gregor, T., Tank, D. W., Wieschaus, E. F. & Bialek, W. Probing the limits to positional information. *Cell* **130**, 153–164 (2007).
19. Dubuis, J. O., Tkacik, G., Wieschaus, E. F., Gregor, T. & Bialek, W. Positional information, in bits. *Proc. Natl Acad. Sci. USA* **110**, 16301–16308 (2013).

20. Lacalli, T. C. Patterning, from conifers to consciousness: Turing's theory and order from fluctuations. *Front. Cell Dev. Biol.* **10**, 871950 (2022).
21. Nikolić, M. et al. Scale invariance in early embryonic development. Preprint at https://doi.org/10.48550/arXiv.2312.17684 (2023).
22. Ishimatsu, K. et al. Size-reduced embryos reveal a gradient scaling based mechanism for zebrafish somite formation. *Development* https://doi.org/10.1242/dev.161257 (2018).
23. Uygur, A. et al. Scaling pattern to variations in size during development of the vertebrate neural tube. *Dev. Cell* **37**, 127–135 (2016).
24. Almuedo-Castillo, M. et al. Scale-invariant patterning by size-dependent inhibition of nodal signalling. *Nat. Cell Biol.* **20**, 1032–1042 (2018).
25. Leibovich, A., Edri, T., Klein, S. L., Moody, S. A. & Fainsod, A. Natural size variation among embryos leads to the corresponding scaling in gene expression. *Dev. Biol.* **462**, 165–179 (2020).
26. Al Asafen, H. et al. Robustness of the dorsal morphogen gradient with respect to morphogen dosage. *PLoS Comput. Biol.* **16**, e1007750 (2020).
27. Cheung, D., Miles, C., Kreitman, M. & Ma, J. Scaling of the Bicoid morphogen gradient by a volume-dependent production rate. *Development* **138**, 2741–2749 (2011).
28. Ben-Zvi, D., Shilo, B.-Z. & Barkai, N. Scaling of morphogen gradients. *Curr. Opin. Genet. Dev.* **21**, 704–710 (2011).
29. Huang, A., Rupprecht, J.-F. & Saunders, T. E. Embryonic geometry underlies phenotypic variation in decanalized conditions. *eLife* **9**, e47380 (2020).
30. Romanova-Michaelides, M. et al. Morphogen gradient scaling by recycling of intracellular Dpp. *Nature* **602**, 287–293 (2022).
31. Saiz, N. & Hadjantonakis, A.-K. Coordination between patterning and morphogenesis ensures robustness during mouse development. *Philos. Trans. R. Soc. B* **375**, 20190562 (2020).
32. Stückemann, T. et al. Antagonistic self-organizing patterning systems control maintenance and regeneration of the anteroposterior axis in planarians. *Dev. Cell* **40**, 248–263.e4 (2017).
33. Gritti, N., Oriola, D. & Trivedi, V. Rethinking embryology in vitro: a synergy between engineering, data science and theory. *Dev. Biol.* **474**, 48–61 (2021).
34. Rosado-Olivieri, E. A. & Brivanlou, A. H. Synthetic by design: exploiting tissue self-organization to explore early human embryology. *Dev. Biol.* **474**, 16–21 (2021).
35. van den Brink, S. C. et al. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* https://doi.org/10.1242/dev.113001 (2014).
36. Beccari, L. et al. Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**, 272–276 (2018).
37. Hashmi, A. et al. Cell-state transitions and collective cell movement generate an endoderm-like region in gastruloids. *eLife* **11**, e59371 (2022).
38. Underhill, E. J. & Toettcher, J. E. Control of gastruloid patterning and morphogenesis by the Erk and Akt signaling pathways. *Development* **150**, dev201663 (2023).
39. Fu, J., Warmflash, A. & Lutolf, M. P. Stem-cell-based embryo models for fundamental research and translation. *Nat. Mater.* **20**, 132–144 (2021).
40. Beccari, L. et al. Generating gastruloids from mouse embryonic stem cells. *Protoc. Exch.* https://doi.org/10.1038/protex.2018.094 (2018).
41. Snow, M. H. & Tam, P. P. Is compensatory growth a complicating factor in mouse teratology? *Nature* **279**, 555–557 (1979).
42. Lewis, N. E. & Rossant, J. Mechanism of size regulation in mouse embryo aggregates. *J. Embryol. Exp. Morphol.* **72**, 169–181 (1982).
43. Rands, G. F. Size regulation in the mouse embryo. II. The development of half embryos. *J. Embryol. Exp. Morphol.* **98**, 209–217 (1986).
44. Mittnenzweig, M. et al. A single-embryo, single-cell time-resolved model for mouse gastrulation. *Cell* **184**, 2825–2842.e22 (2021).
45. Neijts, R., Simmini, S., Giuliani, F., van Rooijen, C. & Deschamps, J. Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Dev. Dyn.* **243**, 88–98 (2014).
46. Amin, S. et al. Cdx and T brachyury co-activate growth signaling in the embryonic axial progenitor niche. *Cell Rep.* **17**, 3165–3177 (2016).
47. Blassberg, R. et al. Sox2 levels regulate the chromatin occupancy of Wnt mediators in epiblast progenitors responsible for vertebrate body formation. *Nat. Cell Biol.* **24**, 633–644 (2022).
48. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
49. Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).
50. Carolina de Souza-Guerreiro, T., Meng, X., Dacheux, E., Firczuk, H. & McCarthy, J. Translational control of gene expression noise and its relationship to ageing in yeast. *FEBS J.* **288**, 2278–2293 (2021).
51. Dubuis, J. O., Samanta, R. & Gregor, T. Accurate measurements of dynamics and reproducibility in small genetic networks. *Mol. Syst. Biol.* **9**, 639 (2013).
52. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
53. Moore, J. L., Du, Z. & Bao, Z. Systematic quantification of developmental phenotypes at single-cell resolution during embryogenesis. *Development* **140**, 3266–3274 (2013).
54. Werner, S. et al. Scaling and regeneration of self-organized patterns. *Phys. Rev. Lett.* **114**, 138101 (2015).
55. Turing, A. M. The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B* **237**, 37–71 (1952).
56. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
57. Stanton, B. C. et al. Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2014).
58. Clevers, H. Modeling development and disease with organoids. *Cell* **165**, 1586–1597 (2016).
59. Fatehullah, A., Tan, S. H. & Barker, N. Organoids as an in vitro model of human development and disease. *Nat. Cell Biol.* **18**, 246–254 (2016).
60. Lancaster, M. A. & Knoblich, J. A. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247125 (2014).
61. Shariati, L., Esmaeili, Y., Haghjooy Javanmard, S., Bidram, E. & Amini, A. Organoid technology: current standing and future perspectives. *Stem Cells* **39**, 1625–1649 (2021).
62. Veenvliet, J. V., Lenne, P.-F., Turner, D. A., Nachman, I. & Trivedi, V. Sculpting with stem cells: how models of embryo development take shape. *Development* **148**, dev192914 (2021).
63. Rossi, G., Manfrin, A. & Lutolf, M. P. Progress and potential in organoid research. *Nat. Rev. Genet.* **19**, 671–687 (2018).
64. Jensen, K. B. & Little, M. H. Organoids are not organs: sources of variation and misinformation in organoid biology. *Stem Cell Rep.* **18**, 1255–1270 (2023).

Methods

Cell culture

129/SvEv (EmbryoMax) mES cells were cultured on gelatinized tissue-culture six-well plates in a humidified incubator (5% CO_2, 37 °C). They were cultured in Dulbecco's modified Eagle medium (DMEM 1× + GlutaMAX, Fisher 11584516) supplemented with 10% fetal bovine serum (FBS, Pansera), 1× non-essential amino acids (Gibco 11140-035), 1 mM sodium pyruvate (Gibco 11360-039), 1× penicillin–streptomycin (Sigma-Aldrich P4333), 100 µM 2-mercaptoethanol (Gibco 31350-010), 10 ng ml^{-1} leukemia inhibitory factor (Miltenyi Biotec 130-099-895), 3 µM GSK3 inhibitor CHIR 99021 (Sigma-Aldrich SML1046) and 1 µM MEK inhibitor PDO35901 (Sigma-Aldrich PZ0162). Cells were passaged every other day and seeded at 5×10^4 cells ml^{-1} using an automatic cell counter (Logos Biosystems LUNA-II). Medium was half-replaced when cells were not passaged. Cells were tested regularly for mycoplasm (Eurofins Mycoplasmacheck).

Gastruloid culture

A complete description of the protocol to generate gastruloids is described in ref. 40. N2B27 medium was prepared in-house, at least every 3 weeks. Initial cell seeding was performed by manual multi-pipetting using an automatic cell counter (Logos Biosystems LUNA-II) or by flux cytometry (BD FACSAriaIII) in single-cell mode. When prepared for fluorescence-activated cell sorting (FACS) seeding, cells were rinsed twice with phosphate-buffered saline (PBS) as usual, then resuspended in N2B27 at a concentration of $5–10 \times 10^6$ cells ml^{-1} and strained using cell strainers (Falcon 352235). Exposure to the GSK3 inhibitor CHIR 99021 (a Wnt agonist called Chi throughout the text) during 48 and 72 h starts the elongation process. The error in initial seeding number for both methods (Extended Data Fig. 1a) was measured by manually counting the number of cells in individual wells between 30 min up to 1 h after seeding. This way, cells had sedimented but had not started to aggregate.

Immunofluorescence staining

Gastruloids were collected in a 15 ml Falcon tube precoated with PBSF (10% FBS in PBS with Mg^{2+} and Ca^{2+}) with a cut P1000 tip and washed once with PBS. Gastruloids were fixed in 4% paraformaldehyde for 2 h, washed once with PBSF and twice with PBS, then resuspended in 1 ml PBS. After this step, they are stored at 4 °C (up to several months). For immunofluorescence staining procedure, gastruloids were permeabilized in 10 ml of PBSFT (10% FBS + 0.03% Triton in PBS with Mg^{2+} and Ca^{2+}) for 1 h at room temperature (RT); incubated overnight at 4 °C in 0.5 ml PBSFT with 1 µl of 1 mg ml^{-1} 4′,6-diamidino-2-phenylindole (DAPI) and with the diluted primary antibody added (see Supplementary Table 10 for primary antibody information). Gastruloids were washed three times with 10 ml of PBSFT, each wash for 30 min at RT. They were incubated overnight at 4 °C in 0.5 ml PBSFT with 1 µl DAPI and the secondary antibody diluted 1:500 for anti-rat IgG (Invitrogen A21208) and anti-rabbit IgG (Invitrogen A3157). They were subsequently washed at RT twice in PBSFT each time for 30 min and once in PBS for 30 min before mounting. All incubations and washes were done on a shaker and all centrifugations were at 10*g* for 2 min to pellet gastruloids at the bottom of the 15 ml tube. For mounting, gastruloids were transferred for cleaning with a cut P1000 tip in a six-well dish with 3 ml PBS then transferred to a 1.5 ml low-binding tube to remove all impurities. After removing all PBS, 150 µl of mounting medium (50% Aqua-Poly/Mount (Polysciences 18606-20)/50% PBS with Mg^{2+} and Ca^{2+}) were added and gastruloids were transferred with a cut P200 tip to a glass bottom dish (Cellvis D35-10-1.5-N). Finally, after rearranging the gastruloids to avoid any contact, they were covered by a cover glass.

Phalloidin staining

For phalloidin staining, gastruloids were collected, fixed and stored similarly to the immunofluorescence staining protocol, except the fixation step was for 1 h. Fixed gastruloids stored in PBS were transferred to a staining solution containing 1 ml of PBST (0.1% Triton in PBS with Mg^{2+} and Ca^{2+}) with 1 µl of 1 mg ml^{-1} DAPI and 2.5 µl 400× phalloidin. The tube was placed at 4 °C on a shaker overnight. Gastruloids were subsequently washed three times with 10 ml of PBST. During each wash, the tube was placed under agitation at 4 °C for 20 min, centrifuged for 2 min at 10*g*, and the supernatant was aspirated. Finally, the samples were mounted on microscopy coverslips with SlowFade Glass Antifade mounting medium (Invitrogen S36917) using homemade spacers.

Confocal imaging

Gastruloid fluorescence imaging was performed on an LSM900 Airyscan 2 microscope equipped with an Airyscan detector with GaAsP-PMT (Zeiss). We used Airyscan mode to speed up acquisitions and checked that gene expression profile variability and precision were identical in both Airyscan and confocal modes. Before image analysis, raw acquisitions are subjected to Airyscan processing by Zen 3.3 software (Zeiss). Gastruloids were imaged using a 10×, 0.45 numerical aperture (NA) air objective (Zeiss) with a zoom setting between 1 and 1.7, and 150-µm-thick *z*-stacks of 30 slices with a voxel size of $220 \times 220 \times 5{,}000$ nm^3 were acquired, encompassing the lower half of the gastruloid (set on an inverted microscope objective). Laser lines of 405 nm, 488 nm, and 639 nm were used to image DAPI, Alexafluor-488 (for SOX2) and Alexafluor-647 (for CDX2/BRA/FOXC1), respectively. Gastruloids with phalloidin staining were imaged using a 40×, 1.43 NA oil immersion objective (Zeiss) with a voxel size of $76 \times 76 \times 190$ nm^3. The size of the *z*-stack was adjusted for each gastruloid to cover the full width of the specimen. The 405 nm and 561 nm lasers were used to excite the DAPI and phalloidin fluorophores, respectively.

Image analysis

Morphological analysis: midline length determination. The main body axis of each gastruloid (that is, pseudo-AP axis) was defined by computing the medial axis and extending it with straight lines, tangent to each end of the medial axis (Extended Data Fig. 1c). The intersection between this extension and the contour defines the anterior and posterior gastruloid tips. The length of the midline is defined as the length of the curved segment between these two tips.

2D volume reconstruction. An equal number of equidistant points were placed along both sides of the 2D contour (left and right contour sides were defined as segments between anterior and posterior tips). Gastruloids were segmented into bins by pairwise joining equivalent points on the left and right sides. From this midsection plane, an approximate volume was reconstructed assuming an approximation of radial symmetry. The volume of the most extreme bins at the tips was computed using the equation for a sphere cap volume. The volume of all other bins was computed assuming a truncated cone volume, with the following formula (n_B is the total number of bins including caps, h_i the width of bin i along the medial axis and R_i half of its length):

$$V = \frac{1}{3}\pi\left[h_0^2\left(3R_0 + h_0\right) + \sum_{i=1}^{n_B} h_i\left(R_i^2 + R_i R_{i+1}\right) + h_{n_B}^2\left(3R_{n_B} + h_{n_B}\right)\right] \quad (1)$$

This pipeline was applied either to images taken in brightfield at the focal plane for live gastruloids (plane of maximal area) or on 2D maximal projections of the DAPI channel for fixed gastruloids. Using 3D segmentation (see below), which provides a more accurate measurement of the volume but also a more elaborate imaging protocol, we estimated the measurement error of this 2D-based volume reconstruction to be 3–20% (that is, an overestimation of the actual gastruloid volume depending on the state of elongation; Extended Data Fig. 7b). All volumes obtained by 2D volume reconstruction were corrected for this systematic error.

3D cell segmentation. 3D segmentation on image stacks of small gastruloids (initial number of seeded cells smaller or equal to 300) was carried out using Cellpose[52], a state-of-the-art neural network-based segmentation framework. A Cellpose model was trained and fine-tuned on manually segmented 2D images extracted from *z*-stacks of fixed and stained gastruloids with dual labels for actin and cell nuclei (Extended Data Fig. 7a, Methods). After 3D segmentation, cell masks were analyzed to filter out poorly segmented cells and noisy pixels erroneously identified as cells. To reject noisy masks, a bimodal Gaussian mixture model was fitted to the distribution of the logarithm of single-cell volumes (Extended Data Fig. 7e). Only cells belonging to the main mode with the highest mean were kept. The resulting filtered segmentation was checked to be in accordance with the imaging data. Finally, morphological closing was performed on the individual cell masks to avoid holes and cell-in-cell detections.

1D gene expression profile extraction. For fixed gastruloids, a 2D maximum projection was calculated for each channel across a 150-µm-thick stack. The morphological slicing used for 2D volume reconstruction was performed on the DAPI channel with $n_B = 200$ bins (for $\bar{N}_0 = 300$ at 120 h, each bin is ~3 µm wide along the gastruloid midline). The maximum projected immunofluorescence intensity was averaged over each bin (averaged over bin size, that is, ~5,500 pixels) for each channel to obtain the raw profiles of intensity as a function of the fractional position x/L. Any other analysis on the profiles was performed for $0.1 \leq x/L \leq 0.9$ (Extended Data Fig. 3h–k).

As an alternative, for Extended Data Fig. 3h only, the slicing was done using three points: equidistant along both sides of the contour and additionally equidistant points along the midline. Each bin limit was computed from a triplet of corresponding points using a second-order polynomial fit (after a coordinate rotation).

Growth analysis

To assess the reproducibility and scaling of gastruloid growth, midline length and volume were measured every 24 h. Gastruloids are either followed individually (Fig. 1a and Extended Data Fig. 2a) or by averaging the midline length and volume of multiple gastruloids (Fig. 1b). The total cell count N is calculated using the proportionality between V and N (Fig. 1c and Extended Data Fig. 7) or obtained by direct measurement via chemical dissociation (Fig. 1b inset, Methods). Exponential growth is assumed for the cell count N for individual (Extended Data Fig. 2b) and average growth curves. It can be expressed as follows with N_0 the initial number of seeded cells and t_D the effective doubling time:

$$N(t) = N_0 e^{\frac{t\ln(2)}{t_D}} \tag{2}$$

The effective doubling time t_D was extracted from growth curves via linear fitting of $\ln(N(t))$ (Extended Data Fig. 2b,e). We used error propagation to compute the expected error on the cell number ΔN due to the fluctuations in the initial number of seeded cells ΔN_0 and effective doubling time Δt_D:

$$\left(\frac{\Delta N}{N}\right)^2 = \left(\frac{\Delta N_0}{N_0}\right)^2 + \left(\frac{t\ln(2)\Delta t_D}{t_0^2}\right)^2 \tag{3}$$

Positional error analysis

Determination of boundary position. The maximum and minimum intensity of each profile, respectively I_{max} and I_{min}, were determined by calculating the average intensity in the 10% of bins with respectively the highest and the lowest values (Extended Data Fig. 5a)[51]. The profile is spline-fitted for $0.1 \leq x/L \leq 0.9$ and the position at which the intensity I is equal to $(I_{max} + I_{min})/2$ defines the individual profile's boundary position x/L_{EC50} (Extended Data Fig. 5a). For better automation, the boundary position is searched in a gene-specific region, for example, with $0.4 \leq x/L \leq 0.9$ for SOX2.

Determination of $\sigma_{x/L}$. For each gene, a generalized version of the positional error is calculated using error propagation[19]:

$$\sigma_{x/L} = \sigma_I \left| \frac{d\bar{I}}{d(x/L)} \right|^{-1} \tag{4}$$

The error σ_I is measured as the standard deviation of the raw intensity profiles. The derivative is calculated at each position $\sigma_{x/L}$ as the slope of a fitted order-three polynomial on 15 consecutive data points, that is, bins (~45 µm or ~3 cells).

This value was converted to cell diameter (d_c) units using $\varepsilon = \sigma_{x/L} L_{fixed}/d_c SF$. The measured length L in fixed gastruloids is impacted by the shrinking of gastruloids due to fixation and mounting media, and we can write $\varepsilon = \sigma_{x/L} L_{fixed}/d_c SF$, where SF is the shrinkage factor (Extended Data Fig. 6, Methods). Both the measurements of d_c and SF come with an error (see below) and by error propagation, we obtain an error on ε that is represented as the width of a gray band in Fig. 3c and Extended Data Figs. 5 and 8d:

$$\Delta\varepsilon^2 = \left(\frac{1 - SF}{L_{fixed}}\right)^2 + \left(\frac{d_c}{L_{fixed}}\Delta SF\right)^2 \tag{5}$$

Variance-based analysis of profiles

χ^2-Minimization. We minimize the total deviation χ^2 of the intensity profiles from the mean across n gastruloids. We assume that for each profile $I^{(\mu)}$ ($\mu = 1, \ldots, n$) we can quantify this deviation for the (unknown) true intensity profile $i^{(\mu)}$ by an additive constant α_μ and a scale factor β_μ, such that $I^{(\mu)} = \alpha_\mu + {}_\mu i^{(\mu)}$. The total deviation χ^2 can be written as

$$\chi^2(\{\alpha_\mu, \beta_\mu\}) = \sum_{\mu=1}^{n} \int_0^1 dx \left(\left[\frac{I^{(\mu)} - \alpha_\mu}{\beta_\mu} \right] - \bar{i} \right)^2 \tag{6}$$

We minimized χ^2 to learn α_μ and β_μ for each intensity profile (either individual in Extended Data Fig. 5c or mean in Fig. 4b, Extended Data Fig. 8e,f) under the constraints $\sum_{\mu=1}^{n} \alpha_\mu = 0$ and $\prod_{\mu=1}^{n} \beta_\mu = 1$.

Variance decomposition. In the case of the comparison of mean profiles for different plates, we used variance decomposition (Extended Data Fig. 8e) to estimate which part of the variance was due to intrinsic versus inter-plate variance:

$$\sigma_{tot^2} = \frac{1}{\sum n_i} \sum_{i=1}^{n_p} n_i (\bar{I}_i - \langle \bar{I} \rangle)^2 + \frac{1}{\sum n_i} \sum_{i=1}^{n_p} n_i \sigma_{I_i}^2 \tag{7}$$

where n_p is the number of plates and n_i the number of gastruloids in plate i. The first term represents the inter-plate (variance of the means) and the second term represents the intra-plate variance (mean of variances).

The results of this decomposition depend on the position x/L along the midline of the gastruloid. To extract a single inter-plate component of the variance per gene and experiment (Extended Data Fig. 8f), we calculated a weighted average of σ_{inter}:

$$\bar{\sigma}_{inter}^2 = \frac{\sum_x \sigma_{inter}^2(x) I(x)}{\sum_x I(x)} \tag{8}$$

Determination of the shrinkage factor

The various steps of the immunofluorescence and phalloidin staining protocols affect the geometry of the gastruloid isotropically. In particular, the fixation and the mounting media tend to shrink gastruloids.

We determined the shrinkage factor by comparing the 2D reconstructed volume of mounted gastruloids from images obtained by confocal microscopy to the volume of the same gastruloids imaged in brightfield just before collection. A one-dimensional shrinkage factor was defined by the ratio: $SF = 1 - (V_{IF}/V_{BF})^{1/3}$. This factor quantifies by how much gastruloid size is reduced during the staining protocol. As the staining protocols are done in batches, we cannot calculate a shrinkage factor per individual gastruloid. We calculated SF on the average volume before and after protocol for a given plate (Extended Data Fig. 6a). For 50% PBS and 50% aqueous mounting medium (Aqua-Poly/Mount, Polysciences 18606-20) used in the immunofluorescence staining protocol, the shrinkage factor is $SF = 0.35 \pm 0.03$ (Extended Data Fig. 6b). For SlowFade Glass Antifade mounting medium (Invitrogen S36917) used in the phalloidin staining protocol, the shrinkage factor is $SF = 0.36 \pm 0.1$ (Extended Data Fig. 6c). In addition, for one dataset, we monitored the evolution of the estimation of the shrinkage factor from 3 days to 3 weeks after mounting and found that it remained constant within this window of time (Extended Data Fig. 6d). The shrinkage factor determined for both these mounting media was applied to all measured lengths and volumes from stained gastruloids.

Gastruloid cell counting by chemical dissociation

Individual gastruloid chemical dissociation was carried out on gastruloids seeded using FACS. First, 140 µl of medium per well was removed from the 96-well plates where the gastruloids were grown, and gastruloids were washed with 200 µl of PBS with Mg^{2+} and Ca^{2+}. Brightfield images of the gastruloids were taken with an Olympus IX83 microscope with a 10× objective to later determine individual gastruloid volumes via 2D volume reconstruction (Extended Data Fig. 1c, Methods). Gastruloids were then washed with 200 µl of PBS without Mg^{2+} and Ca^{2+} and transferred to a flat-bottom 96-well plate (TPP 92096) with a cut P100 tip coated with FBS to prevent gastruloid deformation and sticking. Then 160 µl of medium was removed and 80 µl of 10× trypsin was added per well. The plate was incubated at 37 °C for 5 min. Next, the solution in each well was pipetted up and down to dissociate each gastruloid into a single-cell suspension. The plate was again incubated at 37 °C for 5 min. Then 130 µl of FBS was added per well to neutralize the trypsin. To continue dissociating the gastruloids, the solution in each well was again pipetted up and down with the same pipette tips to avoid cell loss. The cells in each well were then fixed in 1% paraformaldehyde for 15 min and stained with a solution of 0.5% Triton and 0.2% DAPI. The wells were imaged using the Zeiss LSM900 with a 10×, 0.45 NA air objective with z-stacks of 60 µm and a voxel size of $829 \times 829 \times 7{,}500\ nm^3$. Finally, the cells were automatically detected by watershed segmentation. Individual gastruloid dissociation results agreed with mean cell count data obtained by pooling 48 gastruloids and carrying out bulk dissociation.

Equivalent cell diameter determination

We developed two independent methods to determine gastruloid volume and cell count. The first method involves chemical dissociation-based cell counts of individual gastruloids (Methods) and a 2D volume reconstrution from images taken of the gastruloids before dissociation. We applied this protocol to gastruloids with a range of initial number of seeded cells ($\overline{N}_0 = 100, 300, 500$ and 800) and at five daily time points (ranging from 24 h to 120 h). For each gastruloid, we first reconstructed their volume V from a 2D cross-section image at the center of the gastruloid using brightfield microscopy (Methods). Subsequently the imaged gastrluoid was dissociated chemically to obtain its cell count N (Fig. 1c and Extended Data Fig. 7c). Gastruloid length, volume and cell count at 120 h are reported in Supplementary Table 11. From V and N for each individually dissociated gastruloid, we computed their effective cell volume $V_c = N/V$ and effective cell diameter $d_c = \sqrt[3]{\frac{6V_c}{\pi}}$. The effective cell diameter corresponds to the mean of the distribution of d_c, its error to the standard deviation. Before Chi pulse, $d_c = 16.0 \pm 0.6$ µm (4.0%, $n = 206$); after Chi pulse, $d_c = 13.9 \pm 0.5$ µm (3.8%, $n = 286$) (Extended Data Fig. 7d). This is evidence of a Chi-pulse-induced reduction in gastruloids' effective cell size by ~13% (linear dimension).

The second method is based on 3D cell segmentation (Methods). We used single-cell data to extract the volume (sum of single-cell volumes) and cell count (number of 3D segmented cells) from gastruloids aged at least 72 h with $\overline{N}_0$ ranging from 50 to 300 cells. A comparison of V and N from these two methods is presented in Extended Data Fig. 7f,h (with extensive additional information in the caption). With the 3D segmentation method, $d_c = 13.1 \pm 0.5$ µm (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 $d_c = 13.5 \pm 0.8$ µm.

Note, in our 3D segmentation analyses, we observe small amounts of extracellular space that accumulates during gastruloid formation. These lumen-like structures occur in some gastruloids, but their size is negligible compared to the overall size of the gastruloid. We estimate the amount of extracellular space to be less than 1% of the total gastruloid volume.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Processed immunofluorescence staining data is available as maximum projection images for individual gastruloids, organized by figure number. All images have been deposited on the Zenodo repository under https://doi.org/10.5281/zenodo.8108188. Raw images are available upon request.

Code availability

Custom Python-based analysis code for data processing is available at the GitLab repository (https://gitlab.pasteur.fr/tglab/gastruloids_precisionandscaling).

References

65. van den Brink, S. C. et al. Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* https://doi.org/10.1038/s41586-020-2024-3 (2020).
66. Mansoury, M., Hamed, M., Karmustaji, R., Al Hannan, F. & Safrany, S. T. The edge effect: a global problem. The trouble with culturing cells in 96-well plates. *Biochem. Biophys. Rep.* **26**, 100987 (2021).
67. Tkačik, G., Dubuis, J., Petkova, M. & Gregor, T. Positional information, positional error, and readout precision in mor- phogenesis: a mathematical framework. *Genetics* **199**, 39 (2015).

Acknowledgements

We thank I. Bennabi, D. Brückner, M. Cerminara, M. Cohen-Tannoudji, P. Hansen, M. Nikolić, C. Mirdas, J. Pineau, J. Wong-Ng, B. Zoller and the late R. Neijts. This work was supported by Institut Pasteur (particularly the cytometry platform), Centre National de la Recherche Scientifique, CFM Foundation for Research and the French National Research Agency (ANR-10-LABX-73'Revive', ANR-16-CONV-0005 'Inception', ANR-20-CE12-0028'ChroDynE' and ANR-23-CE13-0021'GastruCyp').

Author contributions

M.M., L.F., C.C., A.S. and T.G. designed experiments. M.M., L.F., C.C. and A.S. developed experimental protocols. M.M., L.F. and C.C. performed experiments. M.M. and L.F. performed computational

image analysis. M.M., L.F. and T.G. wrote the manuscript. T.G. secured funding and supervised the work.

Competing interests
The authors declare no competing interests.

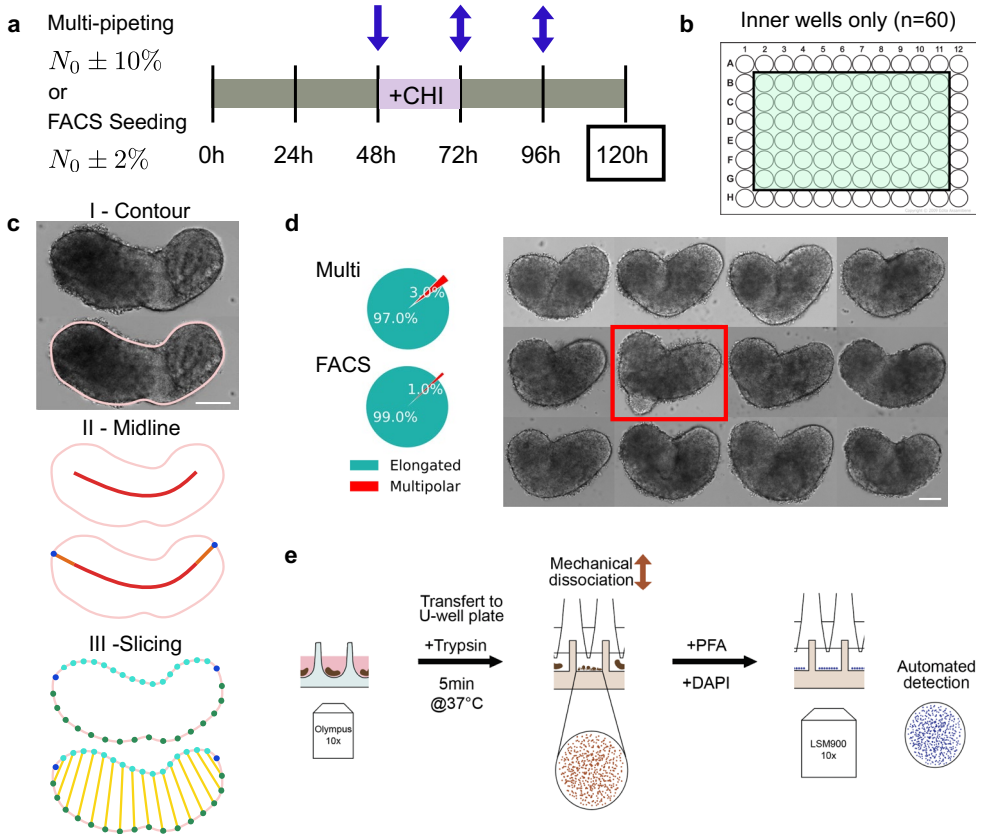
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41594-024-01251-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-024-01251-4.

Correspondence and requests for materials should be addressed to Thomas Gregor.

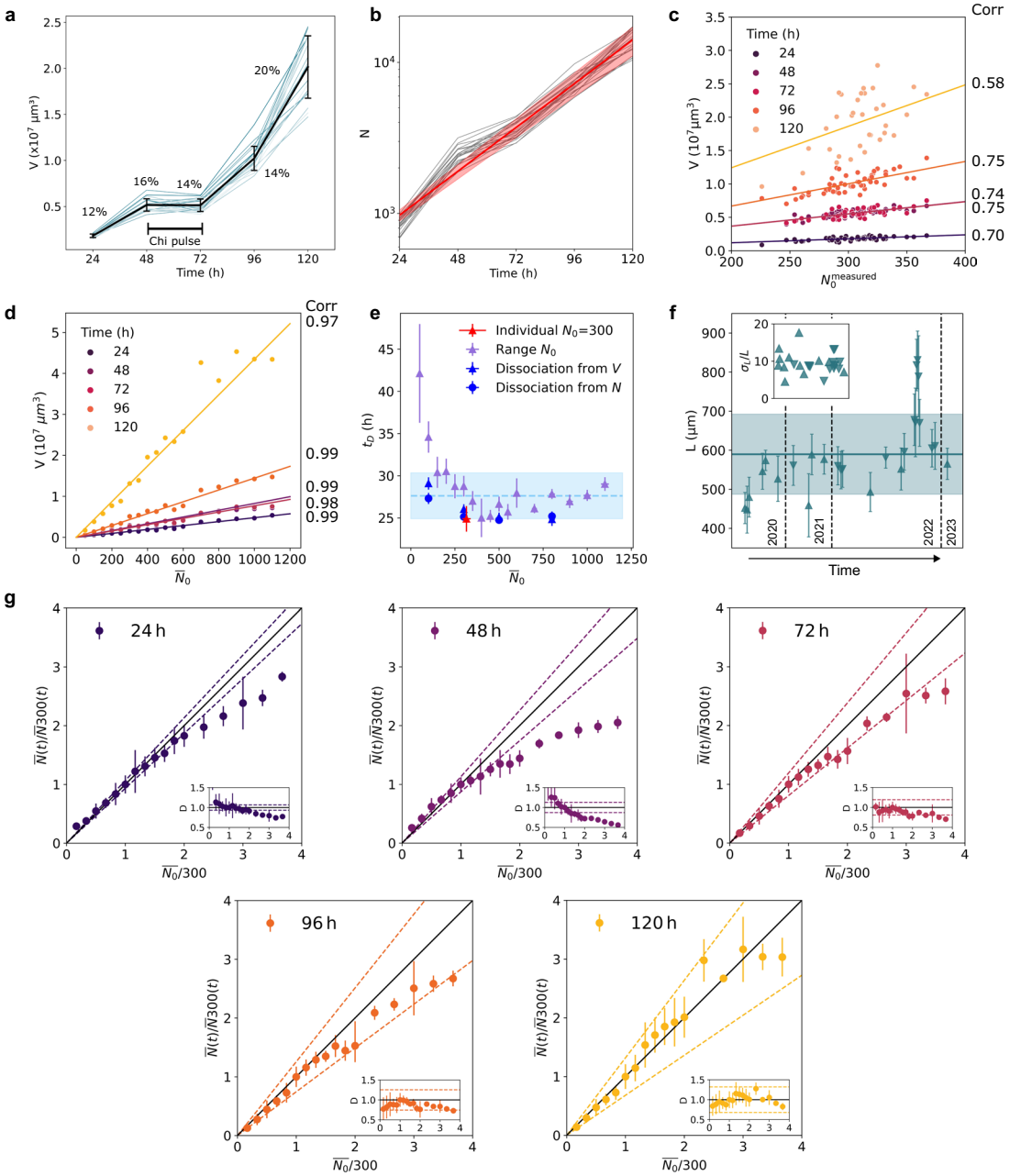
Peer review information *Nature Structural & Molecular Biology* thanks Timothy Saunders and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Carolina Perdigoto, in collaboration with the *Nature Structural & Molecular Biology* team. Peer reviewer reports are available.

https://doi.org/10.1038/s41594-024-01251-4



Extended Data Fig. 1 | Experimental detail, protocols, and image analysis. **a**: Gastruloid protocol as described before with a Chi-pulse on day three[40]. Initial seeding either done by manual multi-pipetting or using Fluorescence-activated Cell Sorting (FACS)[65], 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[66]. **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 (that is, a truncated cone). Scalebar is 100 µm. **d**: Gastruloids imaged with brightfield microscopy. Gastruloid elongation efficiency is 97% for multi-pipetting and 99% for FACS seeding, for $\bar{N}_0 = 300$. The remaining gastruloids have multiple poles (for example, red framed image). Scalebar is 100 µm. **e**: Schematic of the protocol to measure the volume and cell count of individual gastruloids. Brightfield 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).

https://doi.org/10.1038/s41594-024-01251-4

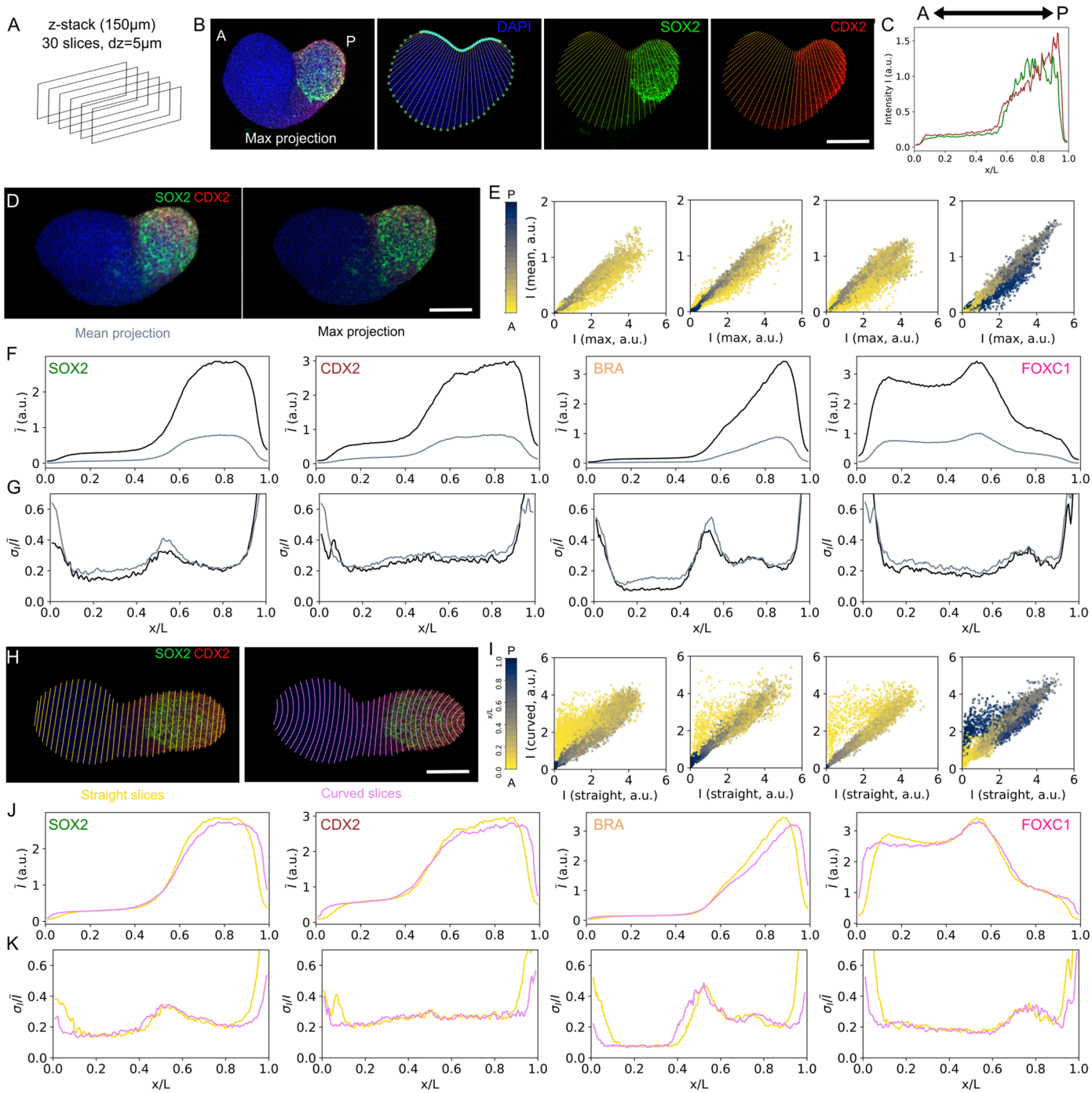


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41594-024-01251-4

Extended Data Fig. 2 | Growth reproducibility and size scaling. a: Gastruloid volume as a function of time. Volumes are obtained from 2D reconstruction in Extended Data Fig. 1c. 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 Extended Data Fig. 7). Exponential growth (see Methods) 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 $t_D = 26.4 \pm 1.7$ h. Red shaded area was computed from error propagation (Methods). **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. 1b. 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 $\bar{t}_D = 27.6 \pm 2.6$ h (mean as blue dashed line; light blue area standard deviation). **f**: Evolution of average midline length per experiment over three years (2020–2023) for gastruloids with $\overline{N}_0 = 300$ at 120 h. Downward triangles are average midline lengths of experiments seeded by multi-pipetting; upward triangles are average midline lengths of experiments seeded using FACS. 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: $\bar{L} = 590 \pm 102$ µm (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. See Tables S6 and S7 for sample numbers. **g**: Average cell count $\overline{N}(t)/\overline{N}_{300}(t)$ as a function of the initial average seed cell count $\overline{N}_0/300$ in units of the average reference seed cell count $\overline{N}_0 = 300$. Five panels correspond to gastruloid ages at 1 through 5 days (also encoded by color). Black diagonal (slope = 1) represents perfect scaling (see main text) of gastruloid size at time t upon changes in $\overline{N}_0$ ranging over $50 \leq \overline{N}_0 \leq 1100$. For each time point, using a simple exponential growth model, the dashed lines estimate the bounds on the expected deviations from perfect scaling due to fluctuations in both $\overline{N}_0/300$ and in the doubling time t_D (**e** and Methods). Insets show deviations D from perfect scaling: $D = \frac{\overline{N}(t)/\overline{N}_{300}(t)}{\overline{N}_0/300}$, as a function of the initial average seed cell count $\overline{N}_0/300$ in units of the average reference seed cell count $\overline{N}_0 = 300$. Black horizontal line represents perfect scaling and the dashed lines show expected deviations from error propagation. Statistics as in Fig. 1d. All error bars are standard deviations.

https://doi.org/10.1038/s41594-024-01251-4

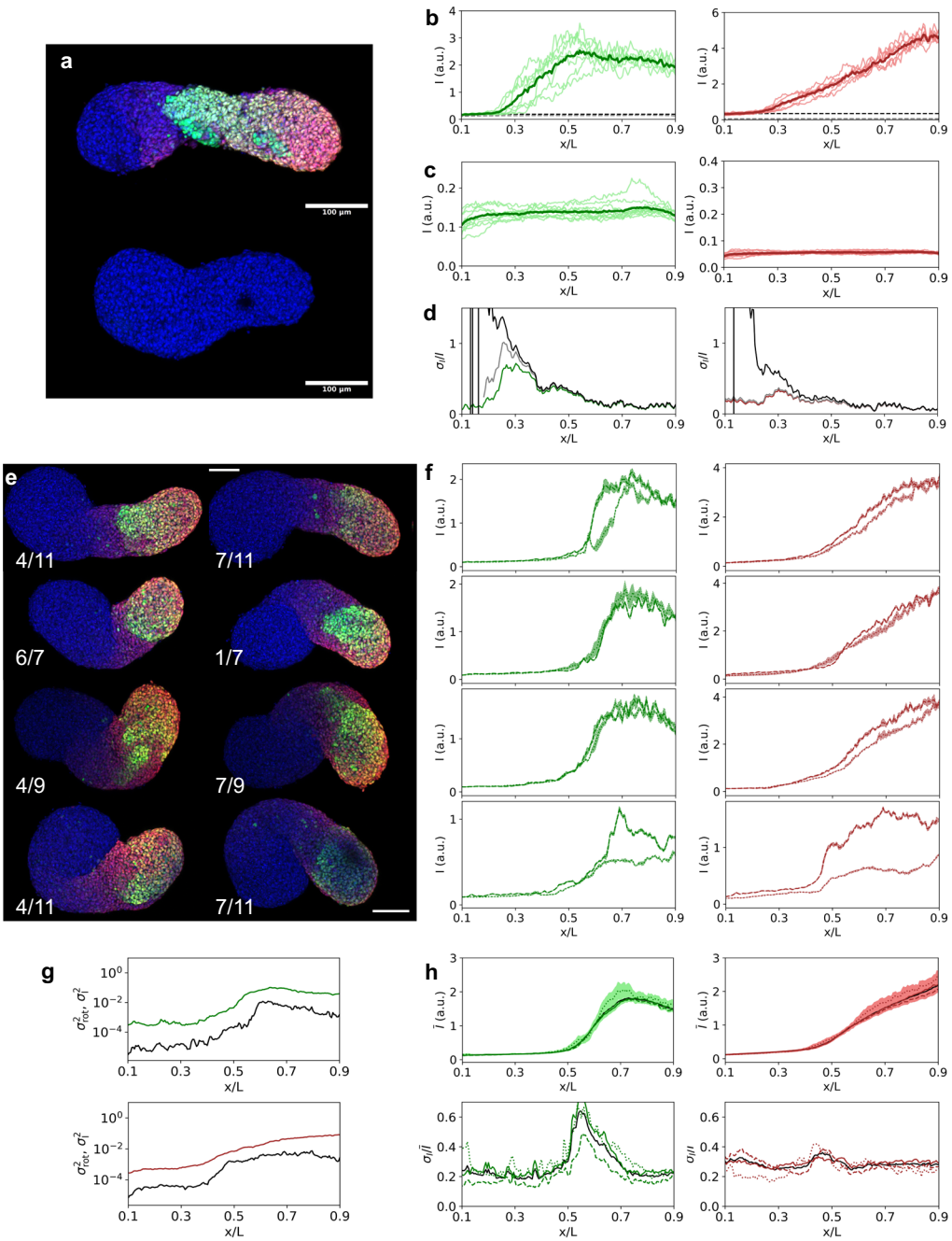


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41594-024-01251-4

Extended Data Fig. 3 | Immunofluorescence image analysis. a: Fixed gastruloids are imaged by confocal microscopy in z-stacks of 150 µm (30 slices, dz = 5 µm). **b**: Analysis pipeline of Extended Data Fig. 1c 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 mean (left) versus maximum (right) projection of a gastruloid stained for SOX2 (green) and CDX2 (red). Scalebar is 100 µm. **e**: Quantitative comparison of maximum (x-axis) versus mean (y-axis) 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 (left, yellow) versus curved lines (right, pink); immunostained gastruloid stained for SOX2 (green) and CDX2 (red). Straight lines are line segments calculated between the equidistant points along both sides of the contour as in Extended Data Fig. 1c. 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 second-order polynomial fit. This procedure is meant to recapitulate the overall curvature of the gastruloid. **i**: Quantitative comparison of intensities using straight (x-axis) versus curved (y-axis) line slicing 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**: 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. **k**: 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 for $0.1 \leq x/L \leq 0.9$ in the rest of the paper.

https://doi.org/10.1038/s41594-024-01251-4

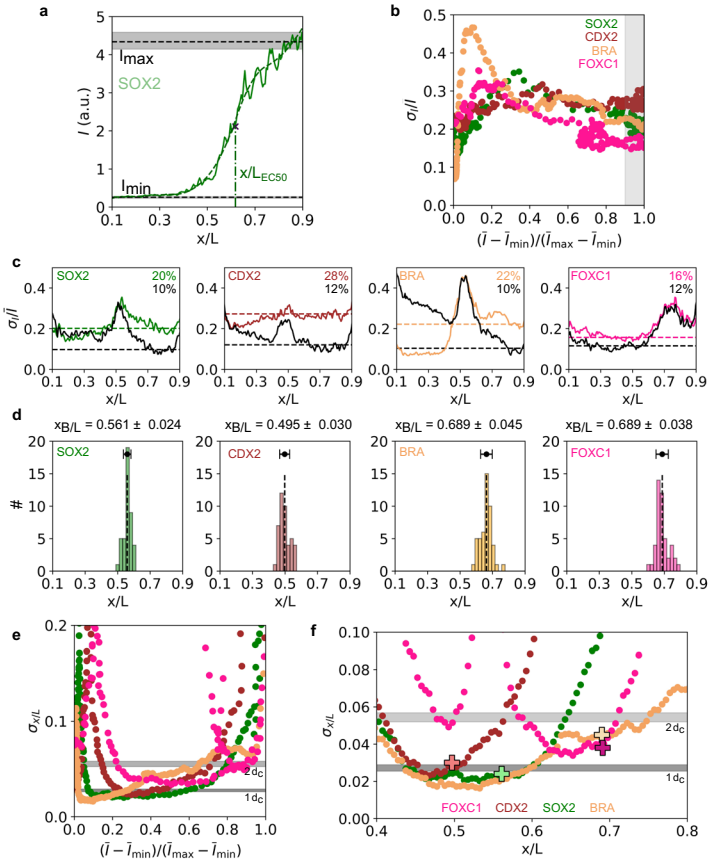


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41594-024-01251-4

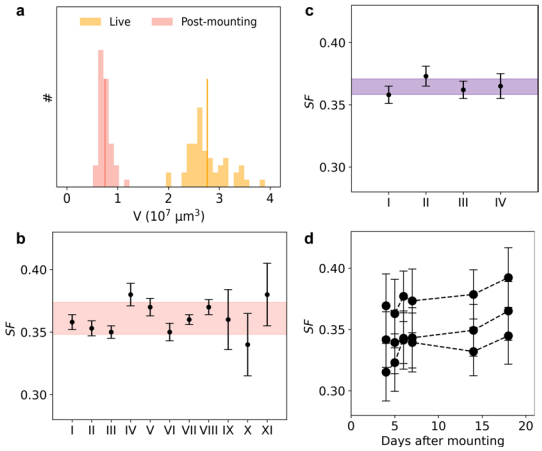
Extended Data Fig. 4 | Immunofluorescence background and measurement error estimation. a: Gastruloids dual-labeled immunofluorescently for SOX2 and CDX2 expression using the regular protocol, as in Fig. 2; bottom gastruloid is missing the primary antibodies to determine the background noise due to non-specific interactions of the secondary antibodies which are estimated the dominant source of background noise in the staining and imaging procedures[51]. **b**: Individual ($n = 10$, light color) and mean profiles (bold) for SOX2 (left, green) and CDX2 (right, red) labeled including primary antibodies. Gray dashed line is the background estimation from **c**; black dashed line is the background calculated from the raw profiles as the mean intensity level in the 10% region of lowest expression ($I_{\min}$). These two dashed lines are confounded in the case of the SOX2 profiles, confirming that the control experiment is a good estimate of the background. **c**: Control experiment without primary antibodies; individual ($n = 10$, light color) and mean profiles (bold) for SOX2 (left, green) and CDX2 (right, red). **d**: Comparison of the variability ($\sigma_I/\bar{I}$) using either the raw mean profile (bold color), the control-corrected profile (bold grey) or the $I_{\min}$ -corrected profile (bold black). **e**: 4 single gastruloids immunofluorescently stained for SOX2 and CDX2. Gastruloids are mounted in PBS medium and rotated manually via flushing for each exposure n =7–11 times, and taken from different view angles. Images are categorized for two different orientations of the gastruloid view angle: a'side view' (left column) and a'backside view' (right column). The preferential orientation is determined by the gastruloid shape and is different from gastruloid to gastruloid. Scalebar is 100 µm. **f**: Mean profiles of SOX2 (green) and CDX2 (red) expression for the four gastruloids in **a** gathered by category: side view (dashed line) and backside view (dotted line). Panels in each row correspond to four experiments with a different individual gastruloid ($n =$ {11, 9, 7, 11} images, respectively). Shaded areas are standard errors in all graphs. **g**: Variance of mean profiles in the four gastruloids due to specimen rotation for SOX2 (black, top) and CDX2 (black, bottom) calculated by bootstrapping the data in **f**. This variance is compared to the total variance (SOX2: green, top; CDX2: red, bottom) of $n = 88$ gastruloids. The rotation-induced variance represents less than 10% of the total variance. **h**: Mean expression profiles (top) and variability (bottom) for SOX2 (green) and CDX2 (red) of the $n = 88$ gastruloids from **g** classified according to their orientation (that is, determine AP orientation using SOX2 expression, determine straight versus crescent orientation, determine L/R orientation for crescent shapes; line style as in **f**). Black lines are the mean profile and standard deviation of gene expression in the total population. This classification based on specimen rotation has minimal effect on the values of mean expression or variability.

https://doi.org/10.1038/s41594-024-01251-4



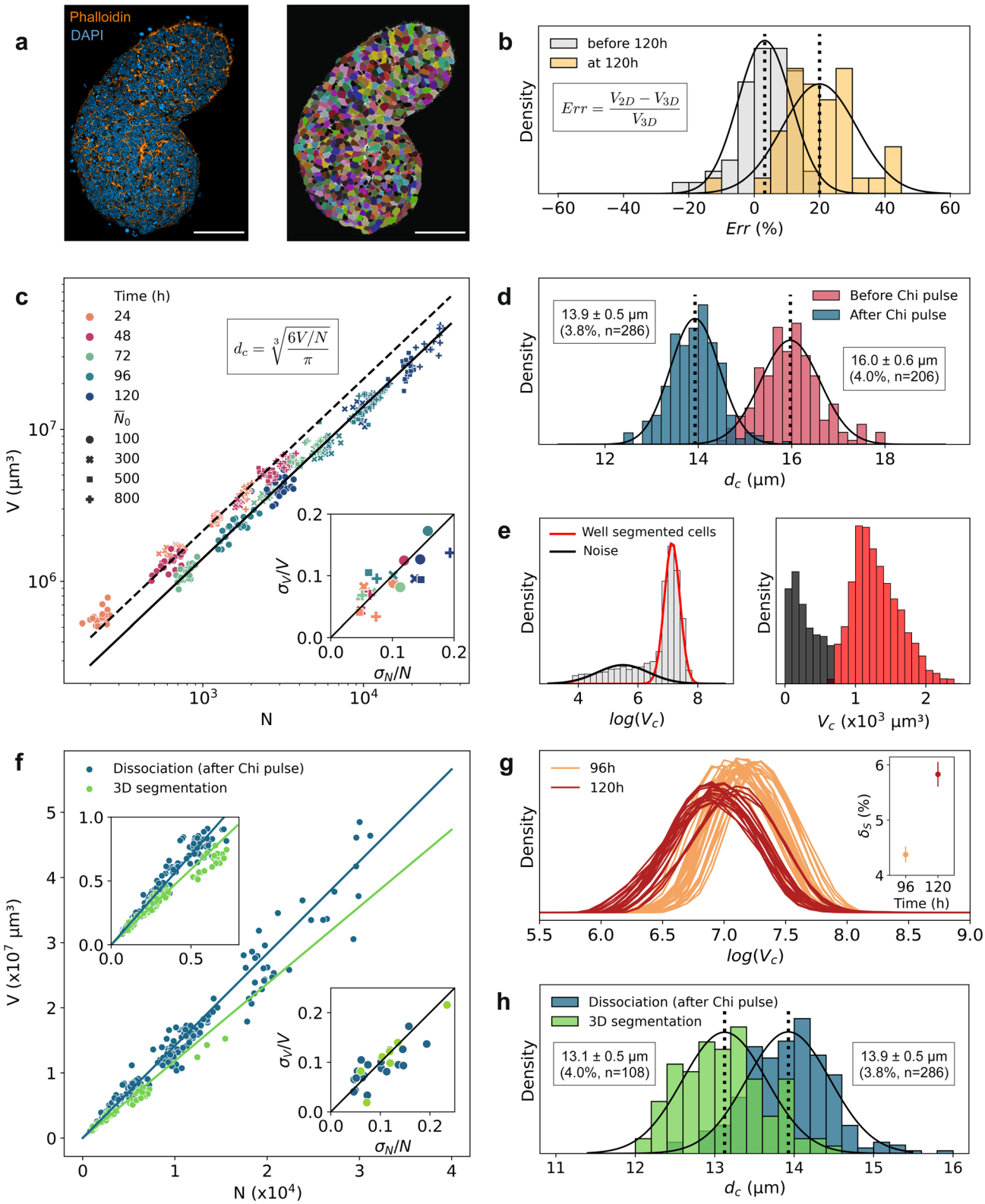
Extended Data Fig. 5 | Reproducibility and precision of gene expression profiles. a: EC50 determination: I_{max} and I_{min} for individual one-dimensional gene expression profiles are defined as the average value of the 10% largest and lowest expressing bins, respectively. The raw profile (green, plain 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 I. I_{max} and I_{min} are determined as in **a** for SOX2, CDX2, BRA, and FOXC1 (n ={44, 44, 48, 46} respectively for SOX2, CDX2, BRA and FOXC1). Same data set as in Fig. 2. 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**: Comparison between raw (colored lines) and ^{2}-minimized (black lines) variability as a function of position x/L for SOX2, CDX2, BRA, and FOXC1 for data set in **b**. Dashed lines represent the average variability in the region where genes are most highly expressed (see **b**). These values decrease from ~20% to ~10% after χ^2-minimization, showing the potential for reproducibility after systematic error reduction, similar to what is seen in the fly embryo[51]. **d**: Distribution of x/L_{EC50} for each of the four markers for data set in **b**, average and standard deviation are reported on top of each distribution. 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. **e**: Generalized positional error as a function of the normalized intensity for each marker (color code as in **b**). The zones of highest precision (that is, $\sigma_{x/L} \leq 5\%$)correspond to the transition regions between low- and high-expression domains. **f**: 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 **d** (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 diameters (gray bands).

https://doi.org/10.1038/s41594-024-01251-4



Extended Data Fig. 6 | Shrinkage factor 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 brightfield 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: $SF = 1 - \left(V_{IF}/V_{BF}\right)^{1/3}$. 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 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 $SF = 0.35 \pm 0.03$ (error is standard deviation). **c**: Same as **b** for a glycerol-based SlowFade™ 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 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 $SF = 0.36 \pm 0.1$ (error is standard deviation). 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 250 µm 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** between three days and three weeks. Error bars are standard errors of the mean obtained from bootstrapping with on average $n = 20$ for live images and $n = 10$ after fixation and mounting (IX-XI).

https://doi.org/10.1038/s41594-024-01251-4

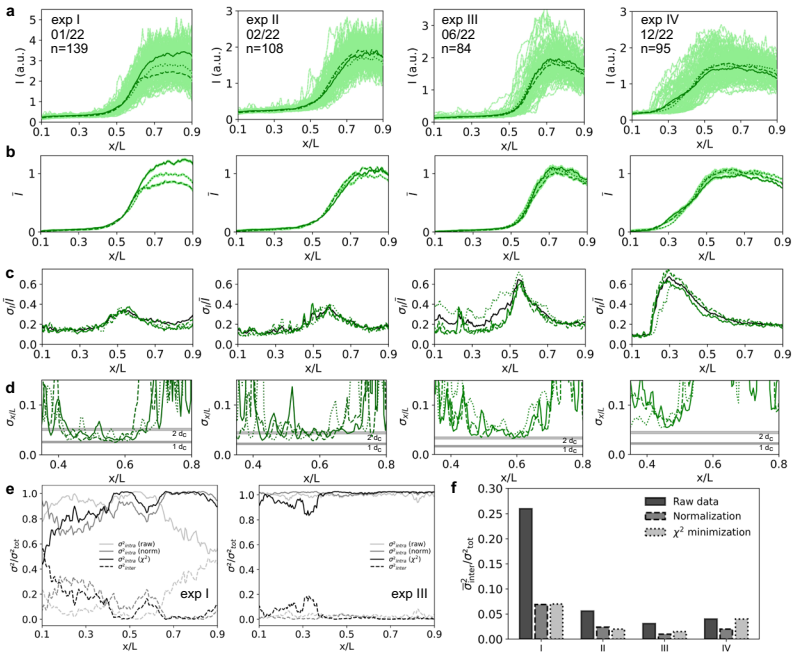


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41594-024-01251-4

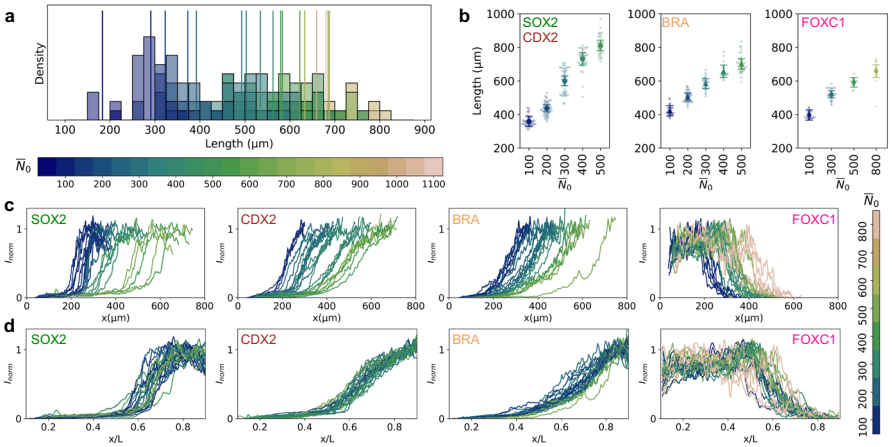
Extended Data Fig. 7 | Determination of total cell count and effective cell diameter. a: Visualisation of the cell masks obtained by 3D segmentation[52]. (Left:) Slice of a confocal image z-stack of a 120 h old gastruloid, seeded from $\overline{N}_0 = 100$ cells, stained for phalloidin (orange) and DAPI (blue). (Right:) Phalloidin channel from left in grayscale overlaid with cell masks obtained by 3D segmentation (see Methods). Scalebar is 50 µm. **b**: Estimation of the discrepancy between 3D and 2D volume reconstruction. The pipeline presented in Extended Data Fig. 1c 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 (Extended Data Fig. 6) 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 Extended Data Fig. 1e), 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 $\bar{V}_c$ for gastruloids aged from 24 to 48 h (before Chi-pulse) and the mean $\bar{V}_c$ 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 effective cell diameter d_c can be obtained from the distribution of V_c, or directly from the slopes (see Methods and **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, $d_c = 16.0 \pm 0.6$ µm (4.0%, $n = 206$); after Chi-pulse, $d_c = 13.9 \pm 0.5$ µm (3.8%, $n = 286$). This is evidence of a Chi-pulse-induced reduction in gastruloids' 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 noisy 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**). Slope of 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 (Extended Data Fig. 6C). 2D volume reconstruction from dissociated gastruloids is applied to images of live gastruloids (that is, 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 $\left\langle \sigma_{\log(V_C)} / \overline{\log(V_C)} \right\rangle$ 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, $d_c = 13.9 \pm 0.5$ µm (3.8%, $n = 286$). With the 3D segmentation method, $d_c = 13.1 \pm 0.5$ µm (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 $d_c = 13.5 \pm 0.8$ µm. All error bars are standard deviations.

https://doi.org/10.1038/s41594-024-01251-4



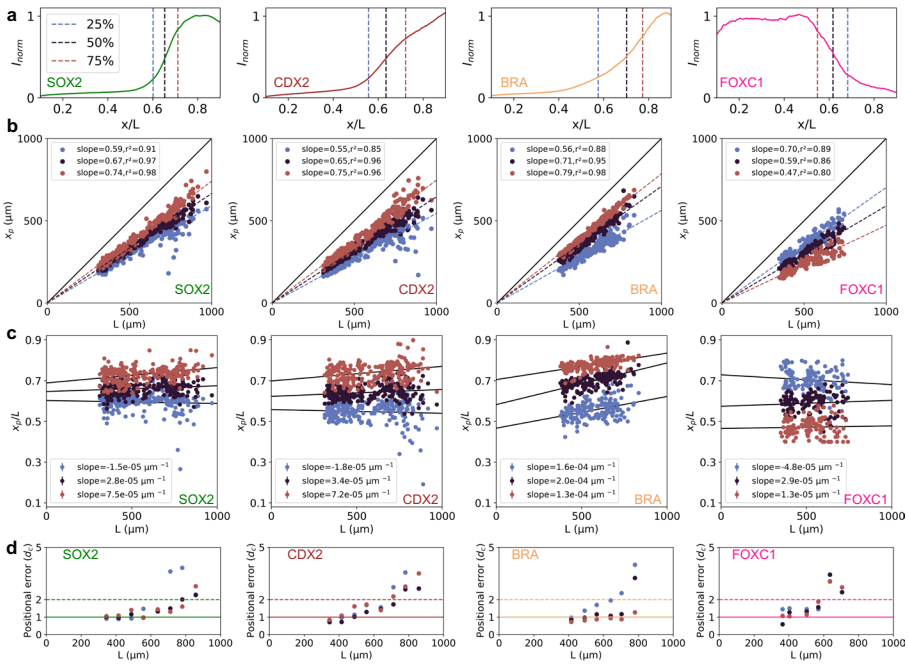
Extended Data Fig. 8 | Repeatability and reproducibility of a single experiment. a: Twelve repetitions of the same experiment on different dates (exp I–exp IV, month/year, with n =139, 105, 84 and 95 gastruloids). Each panel shows raw individual gastruloid profiles (light green, no y-axis normalization) 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 (that is, in 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 (that is, self-similar) than experiments across different days (that is, the mean expression pattern differs more across days than across same-day replicas, something *not* seen in developing embryos[67]). **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 (that is, 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 $\sigma_I/\bar{I}$ 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 (Extended Data Fig. 2f). Boundary precision is maintained near 1–2 cell diameters across all replicas (that is, 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 (that is, no normalization, raw data comparison) better than 5% of the total variance in the data. See Table S8 for sample numbers.

https://doi.org/10.1038/s41594-024-01251-4



Extended Data Fig. 9 | Scaling of gene expression in gastruloids. a: Midline length distribution for gastruloids at 120 h seeded with $\overline{N}_0$ ranging from 50 to 1100 (from Fig. 1b, on average 15 gastruloids per $\overline{N}_0$) with a 5.3-fold total length range. A 22-fold range in $\overline{N}_0$ results in gastruloids with a 3.8-fold range in average length L (bold vertical lines). **b**: Length distributions of gastruloid sets in Fig. 4 as a function of $\overline{N}_0$ (light points are individual gastruloids, $n = 24 - 53$ gastruloids per $\overline{N}_0$, for a total of $n = 517$; 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 range in $\overline{N}_0$ achieves a 2.3-fold range in gastruloid length at 120 h. For the data corresponding to BRA and FOXC1, a 5-fold and 8-fold range in $\overline{N}_0$ achieve a 1.7-fold range in length, respectively. **c**: Individual gene expression profiles (normalized between 0 and 1 for each gastruloid individually using I_{min} and I_{max} as in Extended Data Fig. 5a) for each $\overline{N}_0$ (color code on right) and each gene as a function of absolute position along each gastruloid's midline. **d**: Individual gene expression profiles (normalized as in **c**) for each $\overline{N}_0$ (color code on right) and each gene as a function of relative position (x/L) along the midline. See Table S5 for sample numbers.

https://doi.org/10.1038/s41594-024-01251-4



Extended Data Fig. 10 | Limits of precision in scaled gene expression profiles. **a**: Normalized mean expression profiles for each gene (SOX2, CDX2, BRA and FOXC1) of all gastruloids of different $\overline{N}_0$. For each gene, positional markers are defined at three positions corresponding to the 25% (x_{25}, blue), 50% (x_{50}, black), and 75% (x_{75}, red) of maximum profile intensity levels (vertical dashed lines), respectively. **b**: Absolute positions of the 25%, 50% and 75% maximum intensity levels for each gastruloid (same as in Fig. 4c) as a function of gastruloid length (same color code as above). Perfect scaling would imply $R^2 = 1$, meaning that 100% of the observed boundary position variance is related to gastruloid length. Slope values correspond to the average position of the three positional markers in relative units x_p/L. **c**: Relative position of the 25%, 50% and 75% maximum intensity levels as a function of L for each gastruloid (same color code as above). Perfect scaling predicts statistical independence of the relative boundary position (50% maximum intensity position) and the absolute gastruloid length. We performed a linear regression and found that the slopes are statistically different from zero (see Table S9 for p-values), with a 99% confidence interval; see slopes in legend. A slope of 10^{-5} μm^{-1} means that a decrease or an increase of 300 μm around the case $\overline{N}_0 = 300$ leads to a shift of the positional marker of ~1% along the AP midline, that is ~6 μm ($\leq 1d_c$). A slope of 10^{-4} μm^{-1} (as is the case for BRA) means that a decrease or an increase of 300 μm leads to a shift of the positional marker of ~10% along the AP midline, that is ~60 μm (~$4d_c$). For x_{50}/L, the slopes for the four genes SOX2, CDX2, BRA, and FOXC1 are 2.8 ± 110^{-5}, 3.4 ± 110^{-5}, 2.0 ± 210^{-5} and 2.9 ± 310^{-5} μm^{-1}, respectively. **d**: Positional error for the three markers (same color code as above) converted in cell diameter units (d_c) as a function of average gastruloid length for the four genes SOX2, CDX2, BRA, and FOXC1. The range of gastruloid lengths is binned; each data point corresponds to the bin average. The positional error remains between 1–2 cells for all genes and all markers within a certain length range (up to 600 μm for FOXC1, up to 800 μm for the other genes). This range corresponds to the mean length of gastruloids in a range $100 \leq \overline{N}_0 \leq 500$ for each experiment (Extended Data Fig. 9b). See Table S5 for sample numbers.

nature portfolio

Corresponding author(s): Thomas Gregor

Last updated by author(s): Nov 10, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Zen Blue 3.3 (Zeiss) was used to acquire data on both custom designed microscopes and a Zeiss LSM 980 confocal microscope.
Data analysis	Python (versions 3.6.13 and 3.9) codes were used to analyze all of our data; they can be obtained upon request. Some parts using cell identification use Cellpose 2.0 software. Code can be found here: https://gitlab.pasteur.fr/tglab/gastruloids_precisionandscaling

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Processed immunofluorescence staining data is available as maximum projection images for individual gastruloids, organized by figure number. All images have been deposited on the Zenodo repository under doi: 10.5281/zenodo.8108188. Raw images are available upon request.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N.A.
Reporting on race, ethnicity, or other socially relevant groupings	N.A.
Population characteristics	N.A.
Recruitment	N.A.
Ethics oversight	N.A.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All of our statistical estimates are largely within the 95% confidence interval.
Data exclusions	no outliers removed unless clear experimental flaws were identified
Replication	Experiments were repeated by two experimentalists and analyzed independently leading to the same conclusions.
Randomization	Randomization happens automatically upon seeding of cell aggregates, as a random subset of cells is chosen from a vastly larger cell population.
Blinding	Data reproducibility across three different experimenters was checked.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies against murine proteins SOX2(rat) CDX2(rabbit) BRA(rabbit) FOXC1(rabbit) Provider: eBioscience(14-9811-80), Invitrogen(EPR2764Y), Abcam(ab209665), Abcam(ab223850), respectively Secondary antibodies: Anti-Rat AF488(Intitrogen A21208) and Anti-Rabbit AF647(Invitrogen A31573)

Validation	all antibodies were commercially acquired and used in multiple studies before; their linearity was assessed in our control experiments. Antibody, reference, provider, concentration: SOX2, 14-9811-82, eBioscience, 1/200 CDX2, EPR2764Y, Invitrogen, 1/200 T/BRACHYURY, ab209665, abcam, 1/150 FOXC1, ab223850, abcam, 1/500 Anti-Rat IgG-488nm, A-21208, Invitrogen, 1/500 Anti-Rabbit IgG-647nm, A-31573, Invitrogen, 1/500 All the antibodies, were validated by the companies, they are stated to react to mouse and were validated by western blot. They are routinely used (data from CiteAb): 14-9811-82: 91 citations EPR2764Y: 9 citations (+151 citation when purchased from Abcam) ab209665: 43 citations ab223850: 6 citations

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	mouse embryonic stem cells mESCs (129/svev, EmbryoMax) are commercially available from EmbryoMax
Authentication	cell lines were not authenticated
Mycoplasma contamination	all cell lines tested negative for mycoplasma
Commonly misidentified lines (See ICLAC register)	no misidentified cell lines used