Article

Precise and scalable self-organization in mammalian pseudo-embryos

In the format provided by the
authors and unedited

I. Supplemental tables

N_0 / Time	50	100	150	200	250	300	350	400	450	500	550	600	700	800	900	1000	1100
24h	3	28	18	15	10	30	17	16	11	31	14	16	16	4	18	4	6
48h	11	27	17	15	12	34	15	18	12	31	17	16	16	6	18	6	6
72h	3	28	17	18	12	33	16	17	12	34	17	18	16	6	15	6	4
96h	5	29	18	16	10	32	17	17	12	34	15	18	16	6	17	5	5
120h	5	26	18	15	10	31	15	13	11	30	14	15	10	2	15	5	2

TABLE S1. Number of samples for the size scaling data over a wide N_0 range 50–1100 with FACS seeding. Used in Fig. 1b, d, and Extended Data Fig. 2d, e, g.

N_0 / Time	100	300	500	800
24h	25	25	30	27
48h	23	26	26	24
72h	26	25	28	23
96h	25	23	30	27
120h	23	22	20	14

TABLE S2. Number of samples for the size scaling data with FACS seeding. Used in Fig. 1b inset, and Extended Data Fig. 2e.

Time	Regression slope	Standard error
24h	1.002	0.016
48h	0.999	0.010
72h	1.001	0.013
96h	1.001	0.013
120h	1.003	0.033

TABLE S3. Results of linear fits for the data points of individual gastruloids with $\overline{N}_0 = 300$ in the inset of Fig. 1d. We performed a linear fit for the cell count $N(t)/\overline{N}_{300}(t)$ as a function of the initial seed cell count $N_0/300$. We find that the slopes are equal to one within error bars, i.e. they are statistically indistinguishable from one at all time points.

N_0 / Time	220	240	260	280	300	320	340	360
24h	1	1	5	7	13	19	5	2
48h	0	1	4	7	13	14	5	2
72h	0	1	4	6	14	15	4	2
96h	1	1	6	9	14	13	6	2
120h	1	0	2	4	8	13	4	2

TABLE S4. Number of samples for the size scaling around $N_0 = 300$ by bin. Used in Fig. 1d inset.

$\overline{N}_0$ / Gene	100	200	300	400	500	800
SOX2, CDX2	43	53	49	41	35	
BRA	35	38	27	25	42	
FOXC1	41		36		28	24

TABLE S5. Number of samples for the gene expression pattern scaling data. Used in Fig. 4b, c, and Extended Data Fig. 9, 10.

Exp № \ Year	1	2	3	4	5	6
2020	6	5	6	10	4	6
2021	6	6	8			
2022	24	51				
2023	44					

TABLE S6. Number of samples for the average midline length determination with FACS seeding as a function of time and for different experiments. Used in Extended Data Fig. 2f.

Exp № \ Year	1	2	3	4	5	6	7	8	9
2021	52								
2022	54	15	13	13	18	21	15	48	44

TABLE S7. Number of samples for the average midline length determination with manual seeding as a function of time and for different experiments. Used in Extended Data Fig. 2f.

Replica \ Exp №	1	2	3
I	48	43	46
II	39	36	33
III	25	32	27
IV	30	33	32

TABLE S8. Number of samples for the repeatability and reproducibility of a single experiment with three same-day replicas. Used in Extended Data Fig. 8.

Position \ Gene	25%	50%	75%
SOX2	$1.6\ 10^{-4}$	$7.4\ 10^{-12}$	0
CDX2	$4.7\ 10^{-6}$	$6.7\ 10^{-16}$	0
BRA	0	0	0
FOXC1	$4.3\ 10^{-11}$	$2.1\ 10^{-3}$	0.94

TABLE S9. p-values for the linear regression of the relative position of the 25%, 50% and 75% maximum intensity levels as a function of L for each gastruloid (Extended Data Fig. 10c).

Antibodies	Species	Reference	Provider	Dilution
SOX2	Rat	14-9811-82	eBioscience	1:200
CDX2	Rabbit	EPR2764Y	Invitrogen	1:200
BRA	Rabbit	ab209665	Abcam	1:150
FOXC1	Rabbit	ab223850	Abcam	1:500
Anti-Rat AF488	Donkey	A-21208	Invitrogen	1:500
Anti-Rabbit AF647	Donkey	A-31573	Invitrogen	1:500

TABLE S10. List of antibodies used for immunofluorescence staining (see Methods).

Initial mESC seed size	Midline length	Volume	Cell count
$\overline{N}_0 = 100$	$322 \pm 30\,\mu m$ $(9.4\%,\ n = 23)$	$4.2 \pm 0.6\ 10^6\ \mu m^3$ $(14.5\%,\ n = 23)$	$3.0 \pm 0.4\ 10^3$ cells $(12.6\%,\ n = 23)$
$\overline{N}_0 = 300$	$494 \pm 52\,\mu m$ $(10.5\%,\ n = 22)$	$1.4 \pm 0.2\ 10^7\ \mu m^3$ $(13.6\%,\ n = 22)$	$1.0 \pm 0.1\ 10^4$ cells $(9.5\%,\ n = 22)$
$\overline{N}_0 = 500$	$673 \pm 68\,\mu m$ $(10.2\%,\ n = 20)$	$2.5 \pm 0.4\ 10^7\ \mu m^3$ $(14.6\%,\ n = 20)$	$1.9 \pm 0.2\ 10^4$ cells $(9.4\%,\ n = 20)$
$\overline{N}_0 = 800$	$754 \pm 59\,\mu m$ $(7.8\%,\ n = 14)$	$3.7 \pm 0.7\ 10^7\ \mu m^3$ $(19.3\%,\ n = 14)$	$2.7 \pm 0.4\ 10^4$ cells $(13.7\%,\ n = 14)$

TABLE S11. Gastruloid midline length, volume and cell count at 120 h for different values of $\overline{N}_0$. These values were obtained from the chemical dissociation of gastruloids seeded using FACS (Extended Data Fig. 1e and 7c). All numbers are means $\pm$ standard deviations across n gastruloids. The coefficient of variation, defined by the ratio of standard deviation over mean, is reported in %.