

Cell Stem Cell, Volume 30

Supplemental Information

Multimodal characterization of murine gastruloid development

Simon Suppinger, Marietta Zinner, Nadim Aizarani, Ilya Lukonin, Raphael Ortiz, Chiara Azzi, Michael B. Stadler, Stefano Vianello, Giovanni Palla, Hubertus Kohler, Alexandre Mayran, Matthias P. Lutolf, and Prisca Liberali

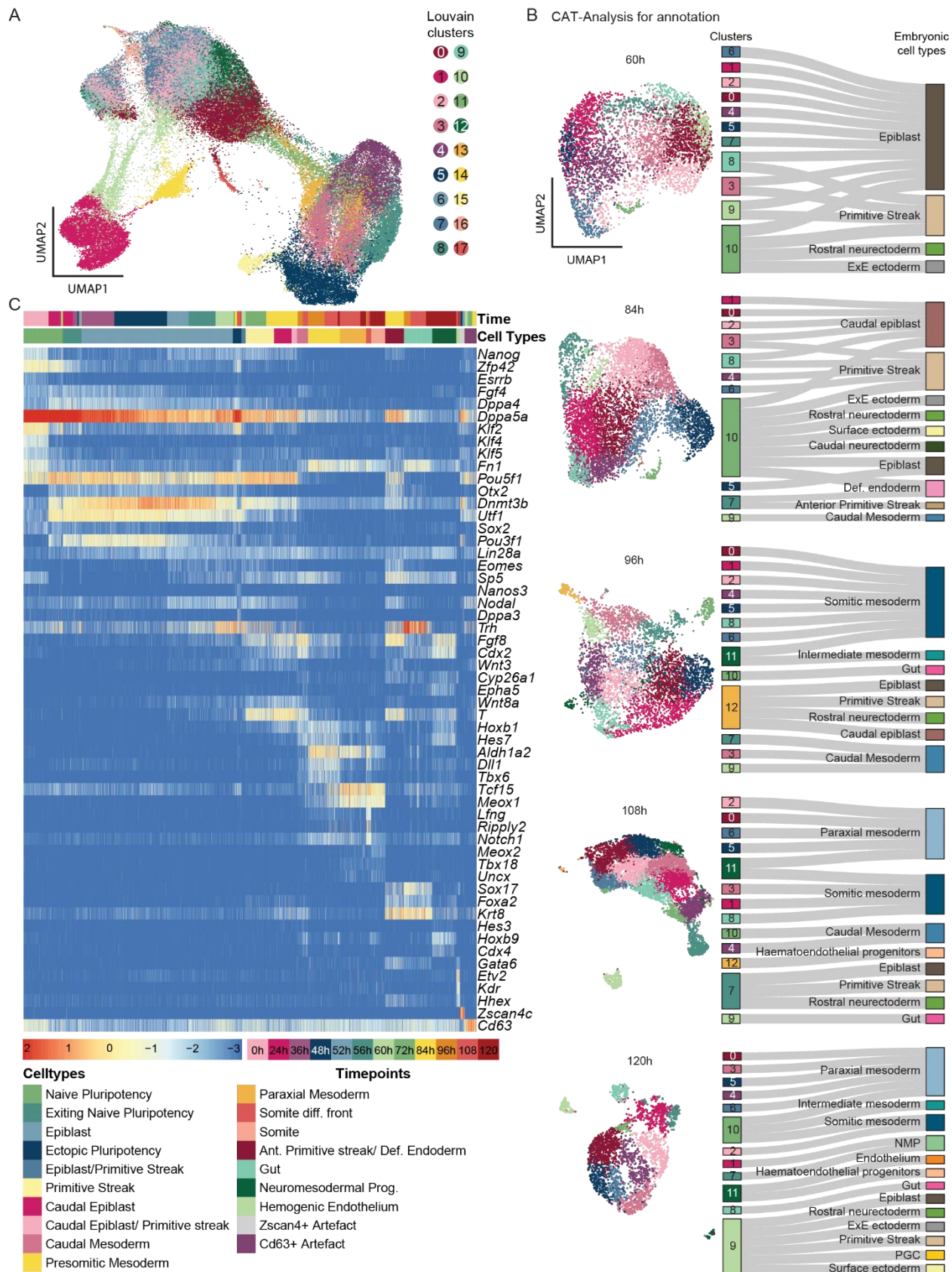


Figure S1 scRNA-seq time course of gastruloid development: Cell type annotation. *Related to Figure 1*

- (A) UMAP of single-cell transcriptomes from gastruloids highlighting Louvain clusters.
- (B) UMAPs highlighting gastruloid clusters from individual timepoints (Left). Sankey plots summarising CAT analysis at individual timepoints. Thickness of the matching bands is directly proportional to the Euclidean distance, i.e, the thicker the band, the greater the Euclidean distance.
- (C) Heatmap showing the expression of cell type markers. Scalebar, log₂-transformed normalised expression. Ant. Primitive streak/ Def. Endoderm (Anterior Primitive streak/ Definitive Endoderm).

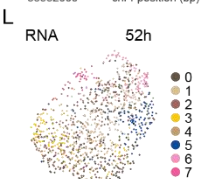
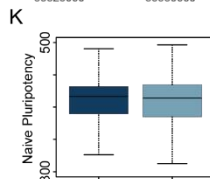
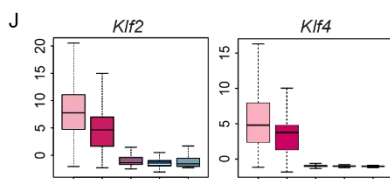
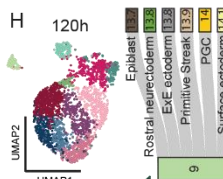
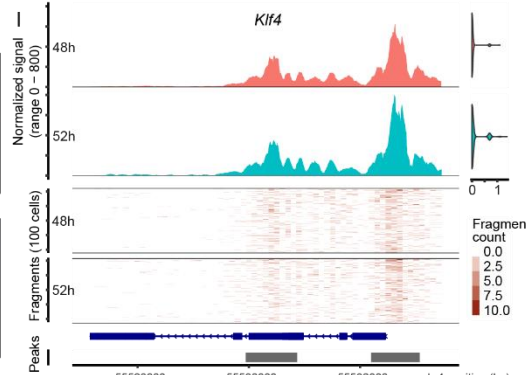
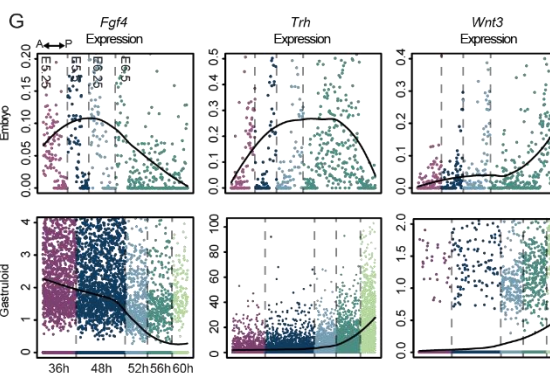
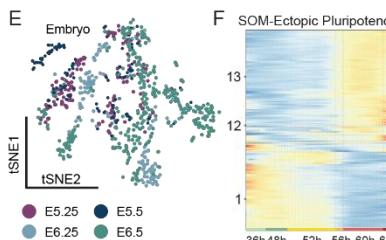
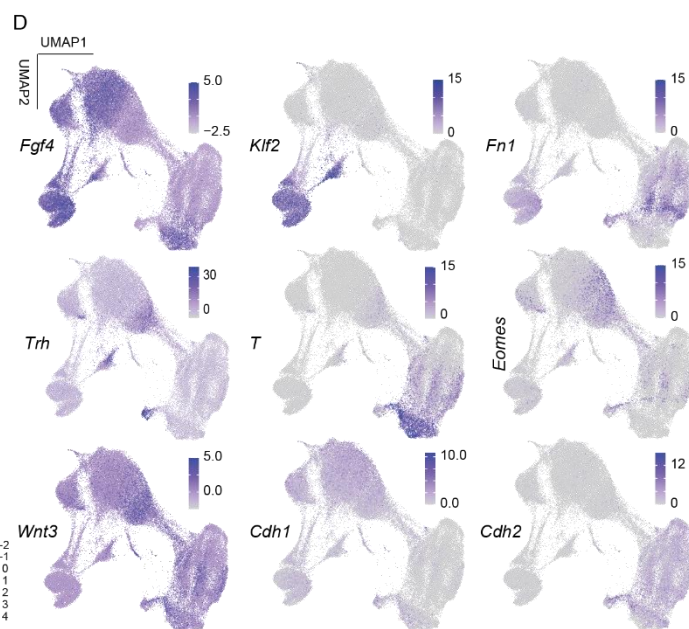
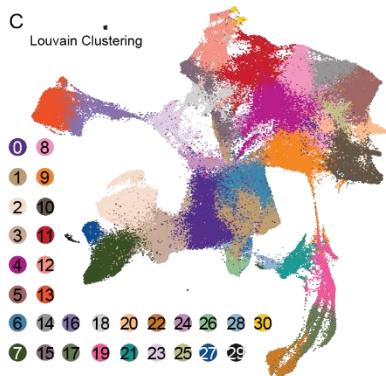
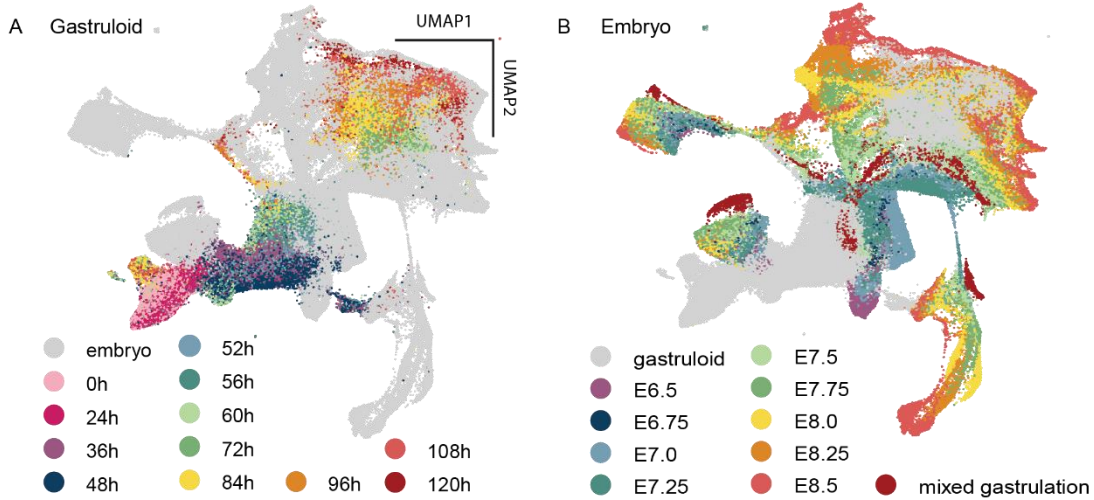
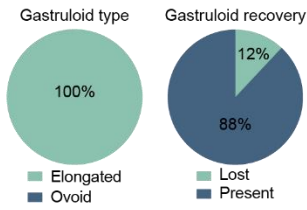


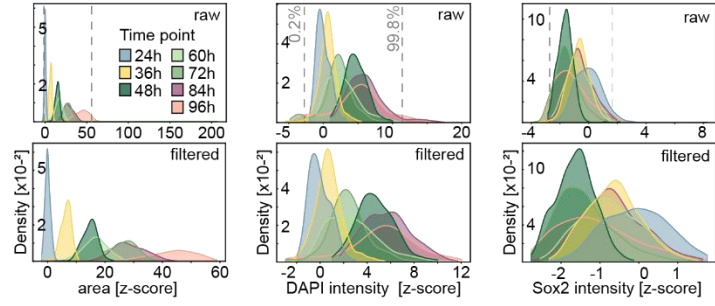
Figure S2 *In vivo* comparison and characterization of different epiblast and pluripotency states.
Related to Figure 2

- (A) UMAP of the integration and co-embedding of gastruloid and embryonic cells highlighting gastruloid timepoints.
- (B) UMAP of the integration and co-embedding of gastruloid and embryonic cells highlighting embryonic timepoints.
- (C) UMAP of the integration and co-embedding of gastruloid and embryonic cells highlighting Louvain clusters.
- (D) Expression UMAPs of *Fgf4*, *Trh*, *Wnt3*, *Klf2*, *T*, *Cdh1*, *Fn1*, *Eomes*, and *Cdh2*.
- (E) t-SNE map of single-cell transcriptomes from Cheng *et al.* 2019^[1] dataset highlighting sampling timepoints.
- (F) SOM of temporal profiles for the ectopic pluripotency population. Bottom: Colour bar shows timepoints.
- (G) Temporal gene expression maps of *Fgf4*, *Trh*, and *Wnt3*, for embryonic and gastruloid epiblast cells. The y-axis: normalised expression.
- (H) UMAP of single-cell transcriptomes from 120h highlighting clusters. Sankey plot highlighting results of the CAT analysis for EP cluster at 120h (cluster 9) in relation to the embryonic cell types. The thickness of the matching bands is directly proportional to the Euclidean distance. Greater thickness corresponds to greater Euclidean distance.
- (I) Coverage plot showing the chromatin accessibility for *Klf4* and 1000 bp upstream region of transcription start site (TSS) containing the promoter for the 48h and 52h timepoints. Upper right panel: multiome RNA expression of *Klf4* in the same cells. Top panel: averaged frequency of sequenced DNA fragments within the genomic region for 48h and 52h timepoints. Middle panel shows frequency of sequenced fragments within the genomic region for single cells from 48h and 52h timepoints. The panel underneath shows gene annotation; arrows indicate the direction of transcription. Bottom panel shows peak coordinates.
- (J) Boxplots showing scRNA-seq derived expression profiles of *Klf2* and *Klf4* at 0h, 24h, 36h, 48h, and 52h.
- (K) Box plots showing the aggregated gene activity scores for the naive pluripotency signature in single cells.
- (L) UMAP of single nucleus RNA-seq data/modality from the multiome highlighting clusters. The 52h timepoint is shown.

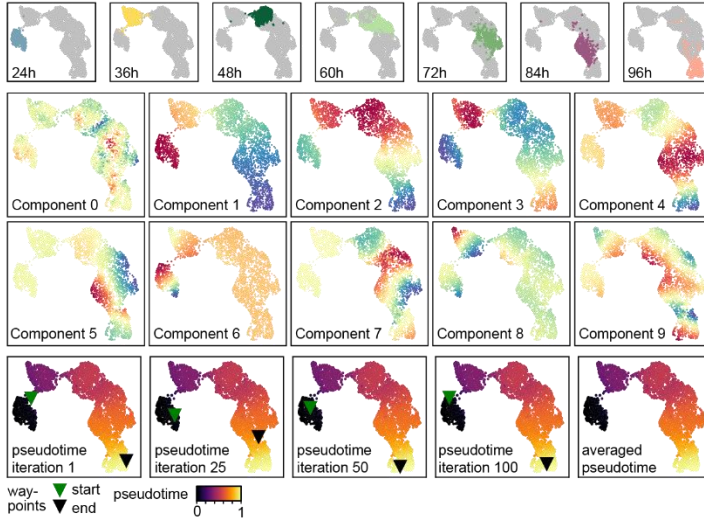
A



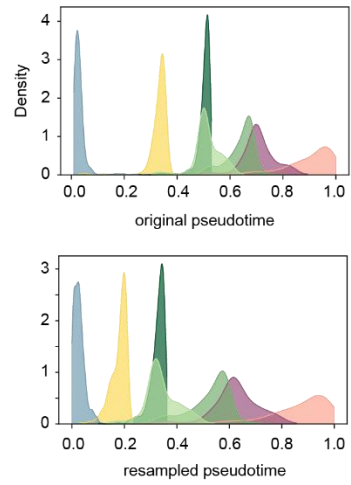
B



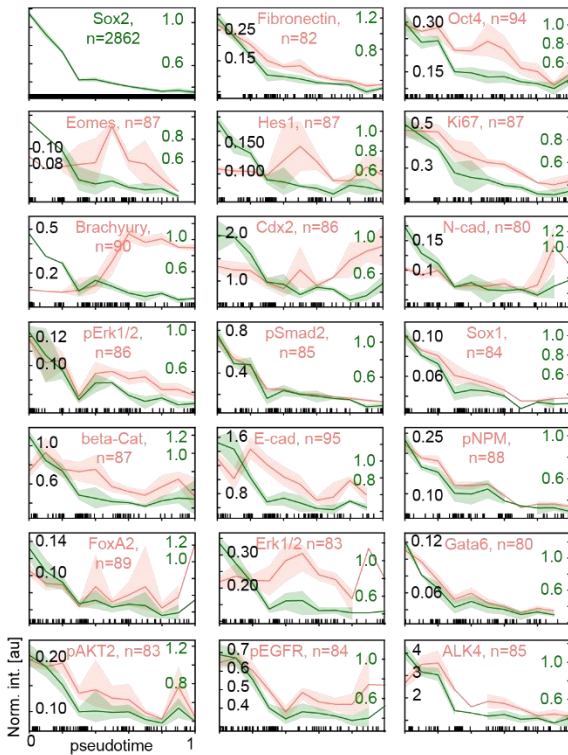
C



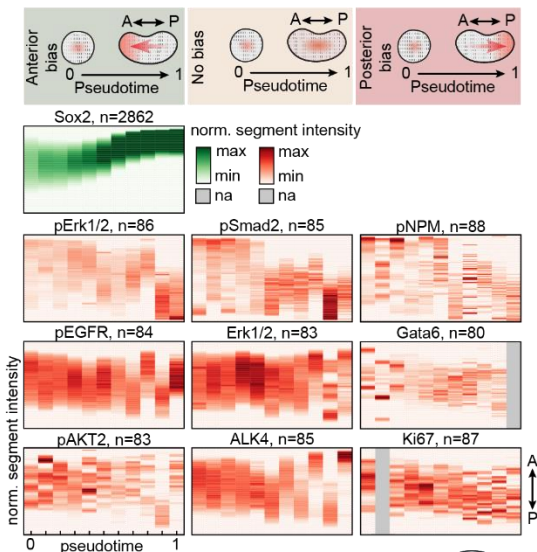
D



E



F



G

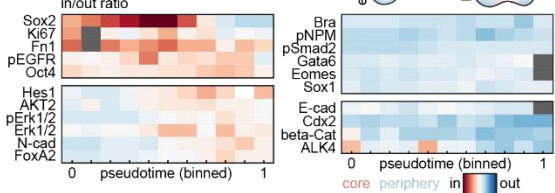


Figure S3. Trajectory inference of gastruloid imaging time course. Related to Figure 3.

- (A) Left: Assessment of elongation efficiency (elongated or ovoid) with high-throughput culture. n=329 gastruloids. Right: quantification of gastruloid-loss with high-throughput culture. n=2 plates, n=717 gastruloids.
- (B) Kernel density estimation plots of selected features in gastruloids at indicated time points. Top: dashed lines show the 0.2 and 99.8 percentiles of the data distributions on the raw data. Bottom: data after clipping by percentiles. Z-score normalisation. n=3165 and 2862 individual gastruloids for raw and filtered data, respectively.
- (C) UMAP plots for n=2862 individual gastruloids colour-coded by time point (top), diffusion components 0-9 (middle), or pseudotime (bottom). In bottom row, colour coding by pseudotime from individual iterations 1, 25, 50, and 100 or averaged pseudotime, respectively. For individual iterations, starting and terminal gastruloids are highlighted.
- (D) Kernel density estimation plots of gastruloids from indicated time points over the original (top) and resampled (bottom) averaged pseudotime.
- (E) Line plots of mean staining intensity for indicated markers and Sox2 from same gastruloids along pseudotime. n, number of gastruloids, see panel (F). Opaque interval shows standard deviation. Intensity of the marker labelled in red plotted on the main axis (left), intensity of Sox2 from the same gastruloids plotted on the secondary axes (right).
- (F) Heatmaps depicting distribution of indicated stainings from anterior (top) to posterior (bottom) pole of gastruloids for indicated pseudotime intervals. n, number of gastruloids stained for indicated markers.
- In E and F, rug plots depict positions of individual data points. A, anterior; a.u., arbitrary units; na, missing data point; Norm. int., normalised intensity; P, posterior; pt, pseudotime; see Table 2.
- (G) Heatmaps depicting distribution of indicated stainings between inner and outer region of gastruloid middle plane (in/out ratio) at indicated pseudotime intervals.

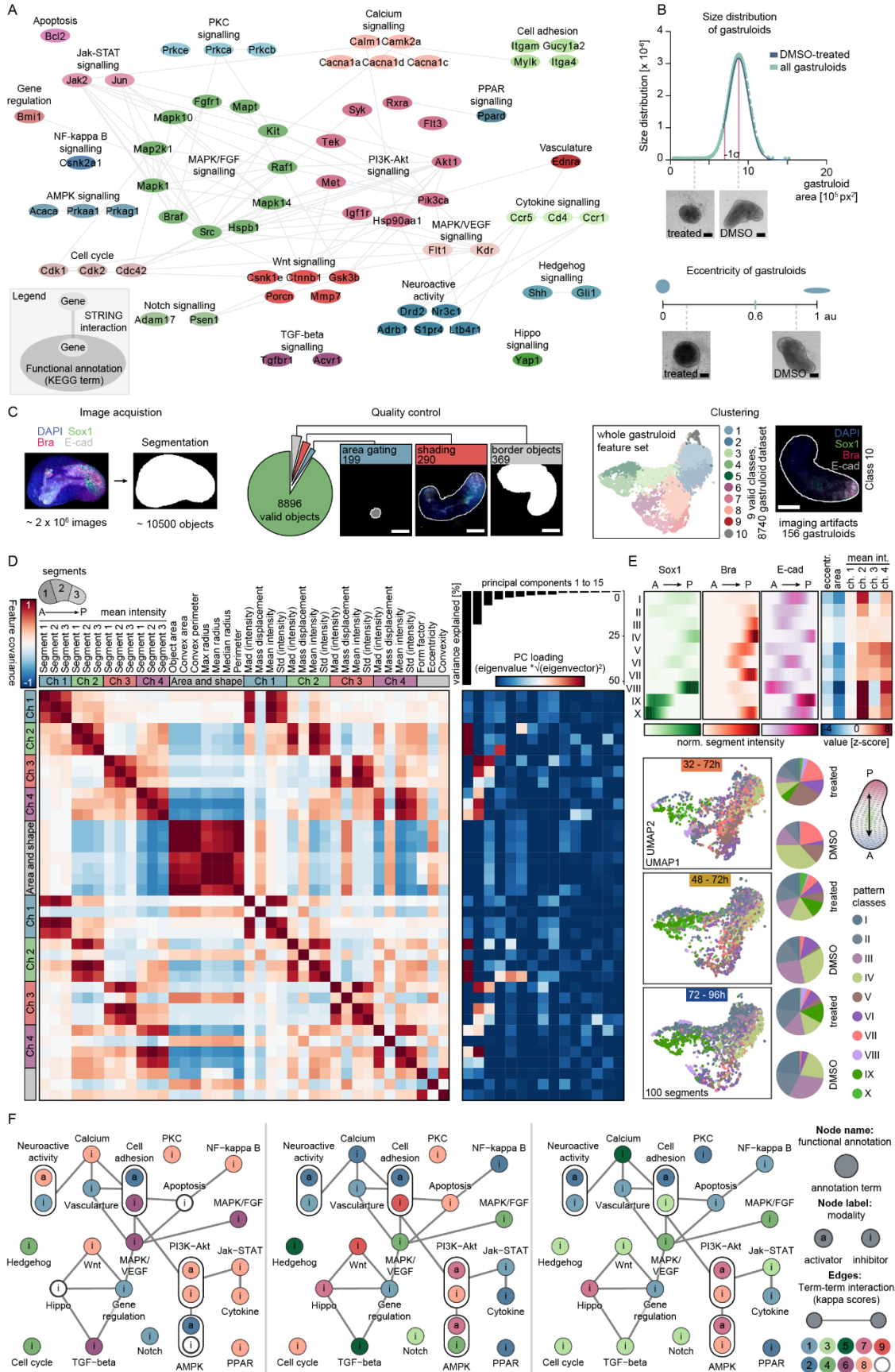


Figure S4. Library and phenotypic classes of gastruloid screen. Related to Figure 4.

- (A) STRING network of annotated genes of the screening library. Nodes are colour-coded by KEGG terms. Nodes show genes and edges show STRING interactions.
- (B) Top: Size distribution of gastruloids treated with pre-screening library and representative images. Bottom: Spectrum of eccentricity and representative images. Compounds causing gastruloids to be one standard deviation (s) smaller than average and below eccentricity 0.6 were chosen for final screening library.
- (C) Overview of image processing, quality control, and quantification. left: image of a representative gastruloid and corresponding segmentation mask. Middle: pie chart showing distribution of objects discarded by indicated quality control gating. right: UMAP plot colour-coded by whole gastruloid class, representative image of a gastruloid from discarded class 10.
- (D) Left: covariance matrix of extracted features, grouped by feature type. Right: heat map showing principal component loading of the extracted features; bar chart above, variance explained by the first 15 principal components.
- (E) Top: distribution of indicated stainings from anterior (left) to posterior (right) in indicated pattern classes, heatmap view of the mean feature values for the pattern classes. Below: UMAP plot colour-coded by pattern class for gastruloids from the indicated treatment regimens (left to right: 32h-72h, 48h-72h, 72h-96h). Pie charts display abundance of each class in the respective subset for DMSO controls and other conditions separately.
- (F) Networks of functional interactions of all annotated terms colour-coded by most frequent whole gastruloid class for each annotation term in indicated treatment regime (left to right: 32h-72h, 48h-72h, 72h-96h). Node name, annotation term. Node label, modality (a, activator, i, inhibitor). Node colour, most frequent whole gastruloid phenotype. White node indicates no detected phenotype. Edges, term-term interaction.

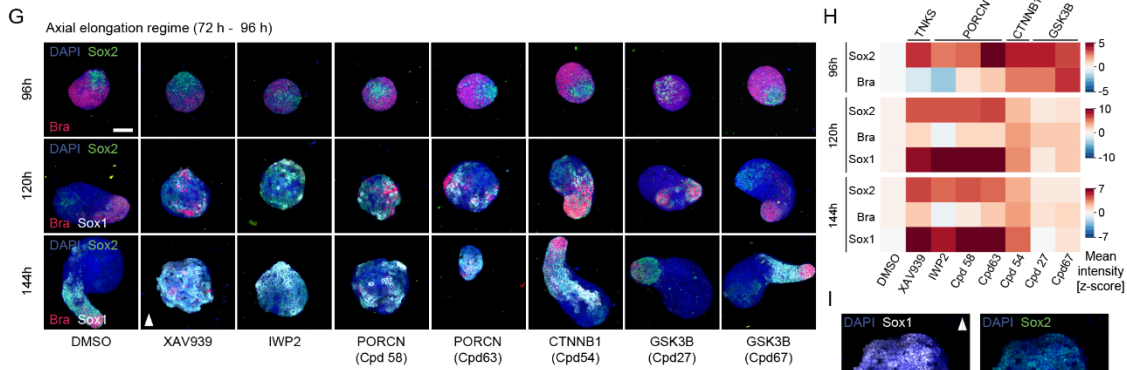
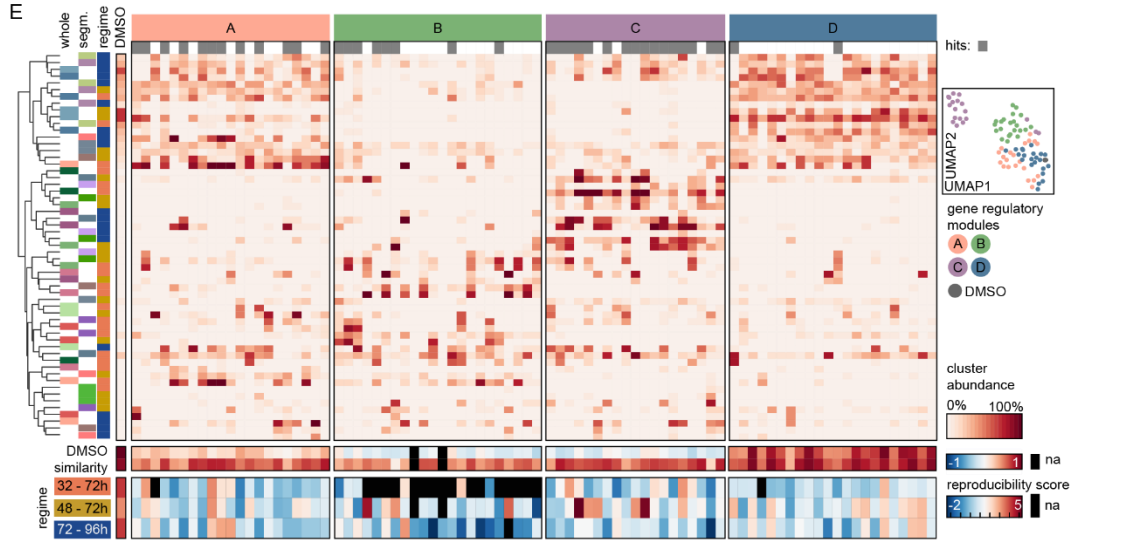
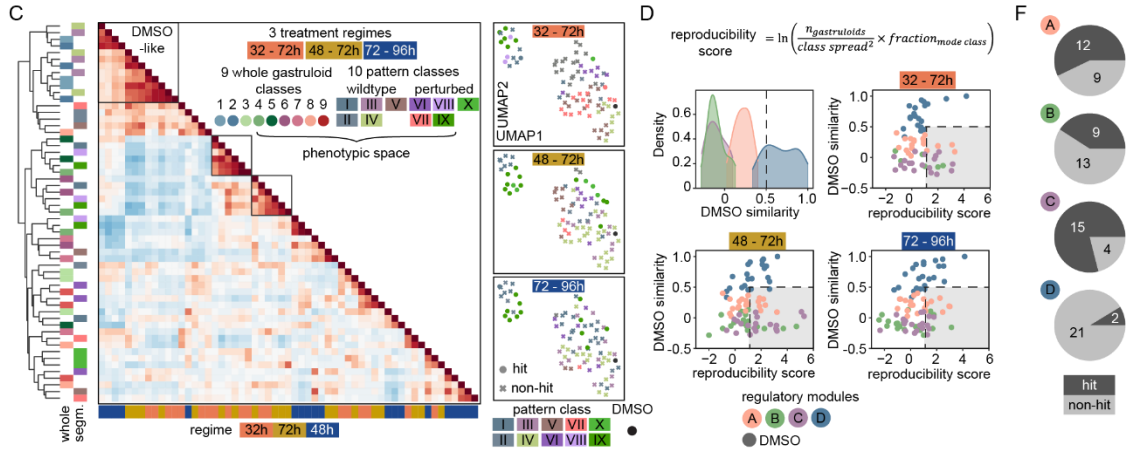
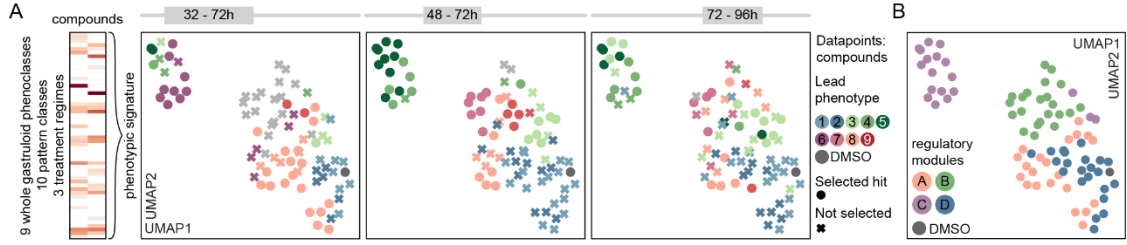
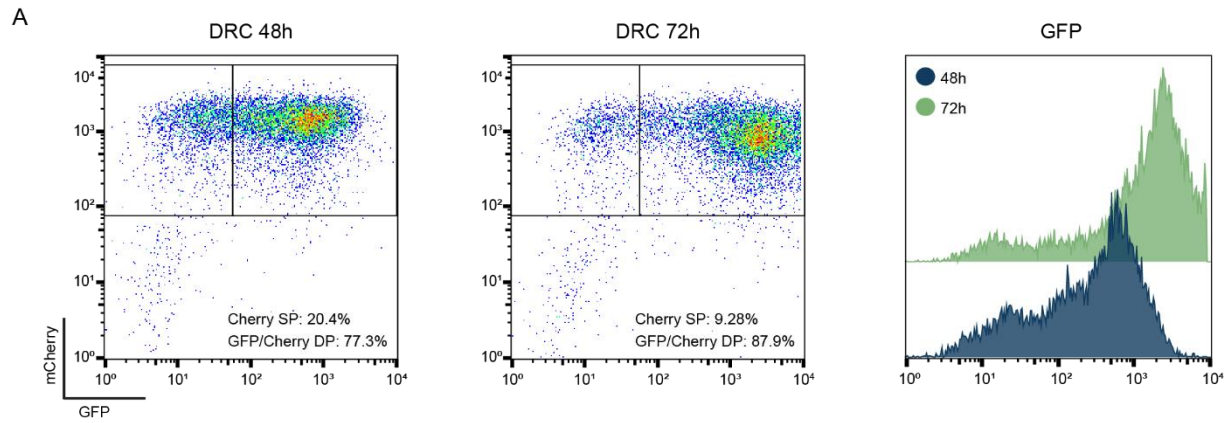


Figure S5. Regulatory modules of gastruloid screen. Related to Figure 5.

- (A) Left: depiction of combined phenotypic space (see S5D). Right: UMAP plots colour-coded by most frequent whole gastruloid phenotype in indicated treatment regime. Marker type specifies compounds selected as hits in indicated regimes.
- (B) UMAP plot colour-coded by regulatory modules, DMSO condition depicted separately. In UMAP plots, data points are individual compounds, n=85.
- (C) Left: heatmap representation of hierarchically clustered covariance matrix between whole gastruloid and pattern classes, heatmap rows and columns are abundances of indicated classes in indicated treatment regimes. Right: UMAP plot colour-coded by most frequent pattern class in compound treatment in indicated treatment regime (top to bottom: 32h-72h, 48h-72h, 72h-96h). Marker type indicates compounds selected as hits in indicated regimes. DMSO controls highlighted separately.
- (D) Definition of reproducibility score. Top left: kernel density estimation plot of phenotypic similarity to DMSO for compounds from indicated similarity clusters (see Figure 5). Top right and bottom row: scatter plots of phenotypic similarity to DMSO and reproducibility score for the compound library colour-coded by compound-level similarity clusters. Highlighted areas correspond to hit selection criteria, thresholds for DMSO similarity, and reproducibility score shown as dashed lines.
- (E) Top: heatmap representation of the 57-dimensional combined phenotypic space for all compounds, grouped by compound-level similarity cluster. Heatmap rows are grouped by covariance, see panel (A). Colour coding by whole gastruloid and pattern classes and treatment regime as in panel (A). Middle: correlation of phenotypic signature to that of DMSO controls. Bottom: reproducibility score of individual compounds in indicated regimes (top to bottom: 32h-72h, 48h-72h, 72h-96h). Compounds selected as hits indicated by grey highlight above the heatmap.
- (F) Pie charts depicting abundance of selected hits in regulatory modules.
In UMAP and scatter plots, data points are individual compounds, n=85. na, missing data point; segm., pattern class; whole, whole gastruloid class.
- (G) Representative images of gastruloids treated with Wnt pathway compounds in the “Axial elongation regime” (48h-72h) corresponding to H. MIPs of confocal z-stacks, DAPI and antibody stainings. Compounds indicated below were used at 5 μ M. White arrowhead depicts gastruloid shown in I. Scale bar, 150 μ m.
- (H) Heatmaps of z-scored mean intensity of Sox2, Sox1 and Bra of gastruloids treated with indicated small compounds at 72h and fixed at indicated time points. n = minimum of 23 per timepoint and condition.
- (I) Enlarged images of XAV939 treated gastruloids fixed at 144h shown in G. DAPI, Sox1 and Sox2 are shown. Scalebar, 200 μ m.



B Spearman Correlation – Ward clustering

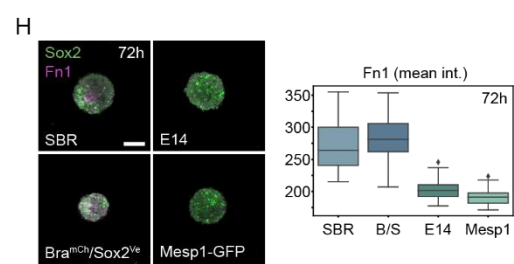
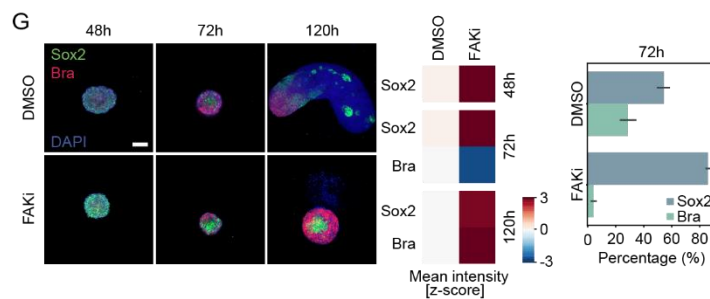
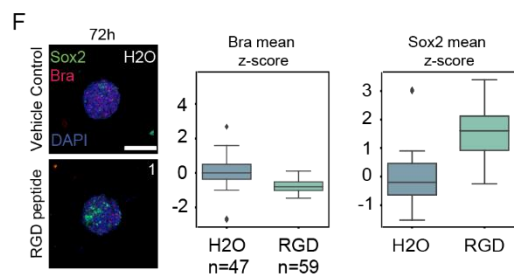
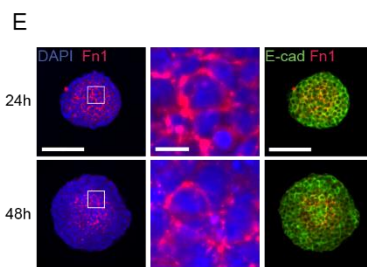
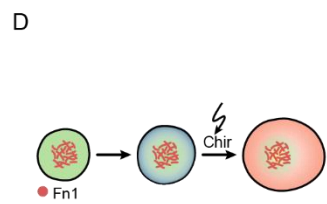
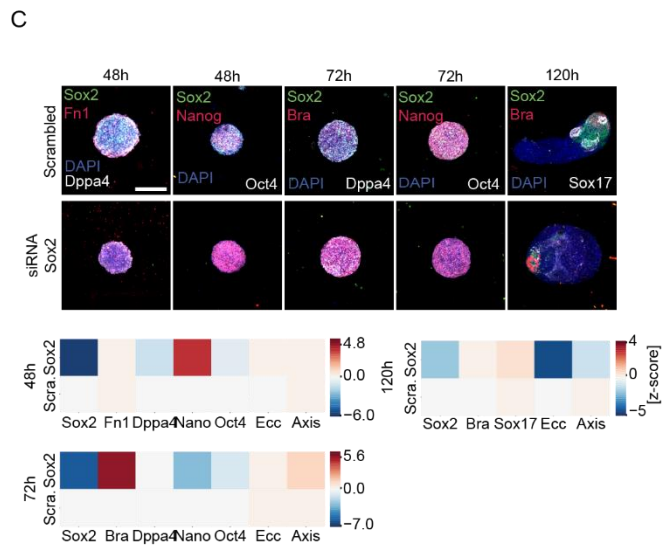
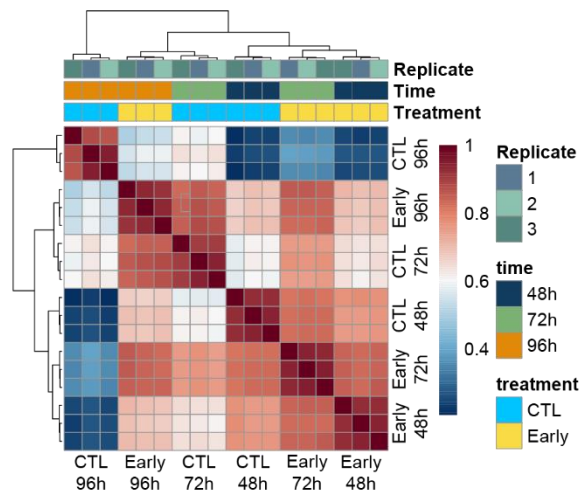


Figure S6. Core characterization and perturbation. Related to Figure 6.

- (A) Flow cytometry results of 48h and 72h DRC gastruloids. Left and middle panels depict scatterplots of single cells. x-axis depicts GFP intensities, y axis depicts mCherry intensities. The gatings chosen to differentiate mCherry single positive (SP) and GFP/mCherry double positive (DP) as well as relative cell proportions are shown. Right panel compares GFP histograms of the 48h and 72h timepoints.
- (B) Spearman Correlation Matrix showing similarities between bulk RNA-seq samples using the 2000 most variable genes across all samples. Respective replicates, timepoints and treatments are indicated (right).
- (C) siRNA KD experiment: representative MIP images of gastruloids at 48h, 72h and 120h. MIP of confocal z-stacks showing antibody stainings for Sox2, Fn1, Dppa4, Oct4, Sox17, Bra and Nanog. Scale bar, 200 μ m. Bottom: Heatmaps of z-scored mean intensity of Sox2, Fn1, Dppa4, Nanog (Nano) and Oct4 as well as z-scored eccentricity (Ecc.) and major axis length (Axis) of gastruloids treated either with siRNA targeting Sox2 (Sox2) or scrambled control (Scra.) fixed at indicated time points. For siRNA KD 129 (120h), 138 (72h) and 141 (48h) gastruloids were analysed. For the scrambled control 99 (120h), 116 (72h) and 138 (48h) gastruloids were analysed.
- (D) Scheme depicting Fn1 expression pattern.
- (E) Representative images of gastruloids at indicated time points. Images are middle z-plane showing DAPI and antibody stainings for Fn1 and E-cadherin. Images in second column show ROI with Fn1 localisation. Scale bars, 100 μ m (left and right) or 25 μ m (middle).
- (F) Left: Representative MIP images of gastruloids treated either with RGD peptide or vehicle control, fixed at 72h and stained for Bra, Sox2 and DAPI. Scale bar, 200 μ m. Right: Boxplots showing z-scored mean intensities of Bra and Sox2 stainings of gastruloids treated with 1mg/ml RGD peptide or control. Number of gastruloids (n) are indicated.
- (G) Left: Representative images of gastruloids treated either with FAK inhibitor or DMSO, fixed at 48h, 72h and 120h and stained for Bra, Sox2 and DAPI. Scale bar, 100 μ m. Middle: Heatmaps of z-scored mean intensity of Sox2 and Bra of gastruloids treated with FAK inhibitor or DMSO and fixed at indicated time points. Right: Percentage of super pixels positive for indicated markers in indicated conditions at 72h. For FAKi 149 (120h), 151 (72h) and 165 (48h) gastruloids were analysed. For the DMSO control 140 (120h), 150 (72h) and 152 (48h) gastruloids were analysed.
- (H) Left: Representative middle z-stack images of 72h gastruloids generated from SBR, E14, BramCh/Sox2Ve and Mesp1-GFP cell lines. Gastruloids were stained for Fn1, Sox2 and DAPI. Scale bar, 150 μ m. Right: Boxplots showing mean intensity of Fn1 stainings of gastruloids generated from indicated cell lines.

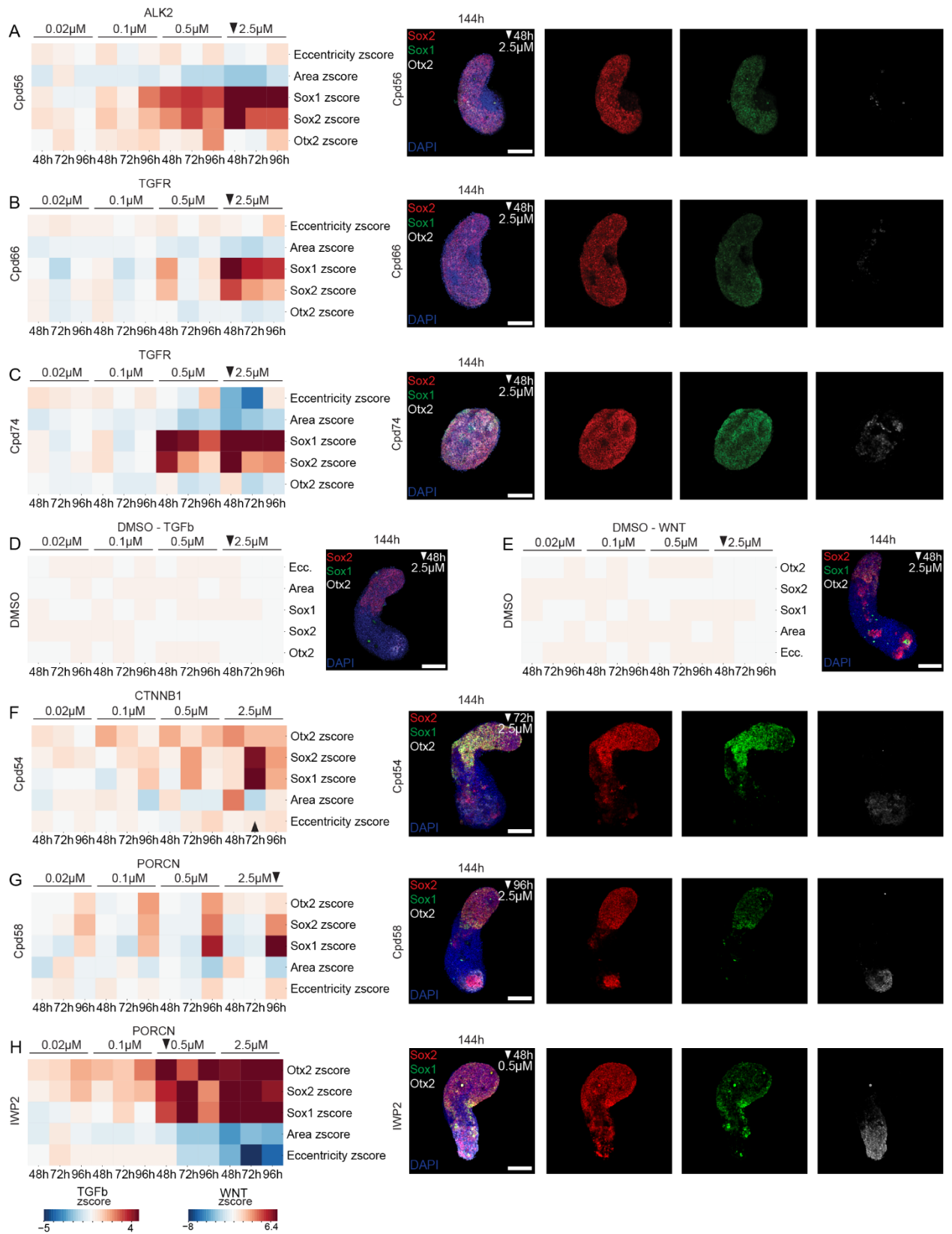


Figure S7. Limiting posterior gradients. Related to Figure 7.

(A-H) Left: Heatmaps of z-scored mean intensity of Sox2, Sox1 and Bra as well as eccentricity (Ecc.) and area of treated gastruloids or DMSO controls at indicated timepoints and fixed at 144h. Legends for both Wnt and TGFb superfamily related heatmaps in the left corner of Figure S7. Black arrow indicates condition of which a representative gastruloid is shown. Right: MIP (144h) of z-stack, showing DAPI and antibody stainings for indicated markers. Gastruloids were fixed at 144h. Respective condition is depicted in the upper right corner of the RGB image. Per condition and timepoint 24 gastruloids were used. Scale bar, 200 μ m.

References

1. Cheng, S., Pei, Y., He, L., Peng, G., Reinius, B., Tam, P.P.L., Jing, N., and Deng, Q. (2019). Single-Cell RNA-Seq Reveals Cellular Heterogeneity of Pluripotency Transition and X Chromosome Dynamics during Early Mouse Development. *CellReports* 26, 2593-2607.e2593. 10.1016/j.celrep.2019.02.031.