e Visualization of transcription factor (TF) activities within and between MBEN cell stages. Above each cell stages, the top five expressed TFs are indicated. Furthermore, the ve most up- and down-

regulated TFs that changed during each developmental step of the MBEN trajectory are given (see also Fig. 3d). In order to design e, the image“Blue Astrocyte” (by Andrew Hardaway, licensed under <https://creativecommons.org/licenses/by/4.0/>) from the database https://scidraw.io/ was used. Source data are provided as a Source Data le.

In order to complement these findings, we performed a cluster-specific transcription factor (TF) activity analysis using the DoRothEA

tool65,66. To this end, overall TF activity per cluster and changes in TF activity between MBEN stages that followed each other were calculated (Fig. 4e, Supplementary Fig. 7, Supplementary Tables 4, 5).

For Transcription factor (TF) activity estimation the python

version of DecoupleR v1.3.0 package was used99. The needed con-

version from Seurat objects to AnnData objects was done with the

SeuratDisk library (https://mojaveazure.github.io/seurat-disk/). TF

activity estimation was performed based on DoRothEA which is a

comprehensive prior knowledge resource containing curated TFs

and their targets66. This network was derived from the OmniPath

database100 via DecoupleR (R package: OmnipathR v3.7.0). In DoR-

othEA, each TF-target interaction includes a con dence level anno-

tation ranging from A to E based on the supporting evidence where A

is the highest con dence level and E is the lowest. For this analysis,

TF-target pairs coming from the three highest con dence levels (i.e.,

A, B and C) were used to create a predictive model for TF activity

estimation. Here, we apply a multivariate linear model to every cell in

our samples to estimate the log-transformed gene expressions using

weights assigned to the interactions between TFs and genes. After

the model is trained, the resulting t-values of the slopes serve as

scores. A positive score indicates an active pathway, while a negative

score indicates an inactive pathway. The resulting activities were

summarized per cluster by their mean with the summarize\_acts

function. The minimum standard deviation was set to zero to retrieve

all results.

A diagram of a cell cycle

AI-generated content may be incorrect.

d Transcription factor activities computed with DoRothEA for the identified cell types present at day3 of the in vitro differentiation protocol and matched cell types in the in vivo gastrulation data set23, relevant TFs are shown in bold, asterisks highlight significantly different activity vs its progenitor, Bonferroni adjusted p < 0.05

A red and blue squares

AI-generated content may be incorrect.

A group of cells with different colored squares

AI-generated content may be incorrect.

Transcription factor activity analysis. Transcriptomic changes across trajectories

and time points were studied based on transcription factor activities using DoR-

othEA and VIPER analysis30. DoRothEA v1.2.1 (https://saezlab.github.io/

dorothea) required Seurat v4.0.2 (https://satijalab.org/seurat/). Both in vitro and

in vivo data sets were subset based on connected cell types according to the

trajectory analysis. Normalised data was scaled within each subset, and TF activity

scores were computed for each cell for 271 TFs with high-confidence target-gene

annotation (A, B and C confidence levels, https://saezlab.github.io/dorothea/).

Heatmaps for in vitro vs in vivo comparison were produced by selecting the top 50

most variable TFs in each data set, and results were merged and plotted using

pheatmap (https://www.rdocumentation.org/packages/pheatmap/versions/1.0.12).

Significance analysis was performed with FindMarkers function from Seurat

(https://satijalab.org/seurat/, v4.0.2) using the LR method, p values were adjusted

using Bonferroni correction.