Cell-cell communication through FGF4 generates and maintains robust proportions of differentiated cell types in embryonic stem cells

Dhruv Raina[1§], Azra Bahadori[1,2§], Angel Stanoev[1], Michelle Protzek[1], Aneta Koseska[1,3], Christian Schröter[1,*]

[1] Department of Systemic Cell Biology, Max Planck Institute of Molecular Physiology, Dortmund, Germany

[2] Current address: Center for Chromosome Stability, University of Copenhagen, Copenhagen, Denmark

[3] Current address: Cellular computations and learning, research center caesar, Bonn, Germany

§ These authors contributed equally

[*] For correspondence: christian.schroeter@mpi-dortmund.mpg.de

Summary

During embryonic development and tissue homeostasis, reproducible proportions of differentiated cell types are specified from populations of multipotent precursor cells. Molecular mechanisms that enable both robust cell type proportioning despite variable initial conditions in the precursor cells, as well as the re-establishment of these proportions upon perturbations in a developing tissue remain to be characterized. Here we report that the differentiation of robust proportions of epiblast-like and primitive endoderm-like cells in mouse embryonic stem cell cultures emerges at the population level through cell-cell communication via a short-range FGF4 signal. We characterize the molecular and dynamical properties of the communication mechanism, and show how it controls both robust cell type proportioning from a wide range of experimentally controlled initial conditions, as well as the autonomous re-

establishment of these proportions following the isolation of one cell type. The generation and maintenance of reproducible proportions of discrete cell types is a new function for FGF signaling that may operate in a range of developing tissues.

Introduction

The differentiation of specialized cell types from populations of multipotent precursor cells is the basis of embryonic development and tissue homeostasis in the adult. Developing tissues generally produce and maintain a standard end result consisting of defined proportions of differentiated cell types despite biological noise and perturbations, a behavior termed canalization (Waddington, 1942). Molecular mechanisms that contribute to canalized development by controlling the proportions of differentiated cell types remain to be characterized.

Mammalian preimplantation development is a prime example for developmental canalization. The size of the three lineages, trophoectoderm (TE), epiblast (Epi), and primitive endoderm (PrE) is remarkably constant between mouse preimplantation embryos (Saiz et al., 2016). Furthermore, mammalian embryos can regulate the proportions of the three lineages following splitting, fusing, or the addition of embryonic stem cells (ESCs), such that the embryo is capable of post-implantation development (Bedzhov et al., 2014; Bradley et al., 1984; Gardner, 1968; Martinez Arias et al., 2013; Tarkowski, 1961; 1959). The differentiation of Epi and PrE cells from inner cell mass (ICM) precursor cells is controlled by transcription factors such as NANOG and GATA6, which mark and specify Epi and PrE cells, respectively. These factors are initially co-expressed in ICM cells, and become mutually exclusive as cells differentiate (Chazaud et al., 2006; Plusa et al., 2008; Simon et al., 2018). In addition, PrE differentiation requires fibroblast growth factor (FGF)/ extracellular regulated kinase (ERK) signaling (Chazaud et al., 2006; Kang et al., 2017; 2012; Krawchuk et al., 2013; Molotkov et al., 2017). The expression dynamics of lineage-specific transcription factors, together with extensive

53 analysis of genetic mutants and pharmacological interventions in the embryo have led to models
54 in which mutually repressive interactions between transcriptional regulators together with
55 heterogeneous FGF/ERK signaling allocate individual ICM cells to one of the two lineages
56 (Yamanaka et al., 2010; Bessonnard et al., 2014; Chickarmane and Peterson, 2008; De Caluwé
57 et al., 2019; De Mot et al., 2016). More recently, FGF signaling has been shown to underlie the
58 regulation of Epi and PrE-lineage sizes following the addition or ablation of cells, suggesting
59 that it orchestrates cell differentiation at the population level (Saiz et al., 2020).
60 Population-level mechanisms for cell differentiation are an attractive solution to the problem
61 of developmental canalization, as it has been shown theoretically that populations of
62 communicating cells can re-establish specific cell type proportions following perturbations
63 (Stanoev et al., 2019). This theory furthermore predicts that the differentiation outcomes are
64 robust towards initial conditions, since the heterogeneous cell identities are generated and
65 maintained on the level of the communicating population, rather than being specified
66 intrinsically in each cell. Identifying and experimentally verifying a molecular mechanism that
67 leads to such emergent phenomena however requires an experimental system in which both the
68 initial conditions of the precursor population as well as cell-cell communication can be
69 precisely controlled.
70 Here we approach this challenge using the differentiation of Epi- and PrE-like cells from ESCs
71 upon the transient expression of exogenous GATA factors as a model system. PrE-like
72 differentiation in ESCs requires above-threshold levels of GATA factor expression and ERK
73 activity (Fig. 1A, (Schröter et al., 2015)), which supports the concept that mutually repressive
74 interactions between NANOG and GATA factors together with FGF/ERK signalling control
75 differentiation in single cells.
76 We exploit the ability to experimentally control GATA expression levels in this system to
77 demonstrate that robust proportions of Epi- and PrE-like cells differentiate even when starting
78 from a wide range of initial conditions. Mutant analysis identifies recursive cell-cell

communication via local FGF4 signaling as the minimal molecular mechanism for this emergent property of cell type proportioning. These experimental observations can be described as a dynamical property of an inhomogeneous steady state (IHSS), a population-based solution that is characterized with mutually exclusive gene expression patterns (Stanoev et al., 2019). We experimentally confirm a central prediction of this theory by demonstrating that cell-cell communication allows populations to re-establish a mixture of Epi- and PrE-like cells from isolated PrE-like cells. Our work thus provides direct experimental evidence that local cell-cell communication via a secreted growth factor can contribute to robust and reproducible global proportions of differentiated cell types across a wide range of initial conditions.

Results

Differentiation of robust proportions of Epi- and PrE-like cells in ESC cultures

To study mechanisms that control the proportions of PrE-like and Epi-like cells differentiating in ESC populations, we integrated a doxycycline-inducible GATA4-mCherry construct via piggybac transgenesis to generate ESC lines that allowed sampling a wide range of GATA4-mCherry expression levels. We induced the transgene while cells were cultured in chemically defined minimal N2B27 medium supplemented with the MEK inhibitor PD0325901 (PD03), the GSK3 inhibitor CHIR99021, and the cytokine LIF (2i + LIF medium, (Ying et al., 2008)) to maintain pluripotency. Differentiation was initiated by concomitantly removing doxycycline and switching to N2B27 medium lacking inhibitors and cytokines. We first analyzed how the proportions of Epi- and PrE-like cells evolved over time in a selected clonal line, using the expression of the endogenous NANOG and GATA6 proteins as markers for the two cell types. Mutually exclusive expression of NANOG and GATA6 could already be observed as early as 16 h after the initiation of differentiation. From 40 h onwards, the initially large proportion of GATA6+; NANOG+ double positive cells was strongly reduced, and two cell populations were

105 clearly separated in the NANOG-GATA6 expression space (Figures S1A, S1B). One
106 population encompassed most of the GATA6+; NANOG- (PrE-like) cells. Their proportion
107 increased until 40 h and slightly declined thereafter. The second population encompassed both
108 GATA6-;NANOG+ (Epi-like) cells, whose proportion remained stable throughout the time
109 course, as well as double-negative cells whose proportion increased continuously (Figures S1A,
110 S1B). The clear separation between the two populations suggests that the increase in the
111 proportion of double negative cells at the expense of GATA6+; NANOG- PrE-like cells beyond
112 40 h is mostly fueled by the downregulation of NANOG expression in the GATA6-negative
113 cell population, combined with a slower proliferation of the GATA6-positive population, rather
114 than by the reversion of PrE-like into double negative cells. Together, these observations
115 therefore suggest that stable proportions of Epi- and PrE-like cells are established at 40 h after
116 the initiation of differentiation.
117 To determine how expression levels of the inducible GATA4-mCherry protein affect cell type
118 proportions, we titrated GATA4-mCherry expression by staggering the initiation of
119 doxycycline induction in time, while keeping the duration of the entire experiment constant
120 (Figure 1B). GATA4-mCherry levels differed by more than 5-fold between 1 h and 8 h of
121 doxycycline induction in 2i + LIF (Figures 1C and 1D), and were paralleled by a concentration-
122 dependent downregulation of NANOG expression (Figures 1D and S2A). After 40 h of
123 differentiation, both GATA6+;NANOG- PrE-like and GATA6-;NANOG+ Epi-like cells were
124 observed (Figures 1D and S2A). Quantitative immunofluorescence revealed that the
125 proportions of PrE-like and Epi-like cells fell within a narrow range for the different induction
126 times. While the proportion of PrE-like cells slightly increased from 43.6 ± 12.1% for 1 h to
127 59.7 ± 11.9% and 57.2 ± 8.7% for 4 h and 8 h of induction, respectively, the proportion of Epi-
128 cells was stable for different induction times (minimum 26.9 ± 6.9% for 2 h induction,
129 maximum 30.8 ± 3.7% for 8 h induction) (Figures 1E and S2B). Thus, a wide range of inducible
130 GATA4-mCherry expression levels leads to similar proportions of differentiated cell types. To

131 rule out that the robust proportioning reflected pre-specification of cell types or a limited
132 differentiation potential of a subset of cells, we added recombinant FGF4 during the
133 differentiation phase to boost ERK activity and lower the GATA4-mCherry threshold required
134 for PrE-like differentiation. The presence of FGF4 strongly increased the proportion of PrE-
135 like cells at the expense of Epi-like cells compared to differentiation in N2B27 alone, and
136 reached a maximum of 93.5 ± 2.1% for 8 h induction (Figure S2C).
137 The observation of robust cell type proportioning was corroborated in four clonal cell lines with
138 independent integrations of the GATA4-mCherry transgene, which displayed a more than 8-
139 fold difference in GATA4-mCherry expression levels following 8 h of doxycycline induction
140 (Figures S3A and S3B). In three out of the four clones, similar proportions of both Epi- and
141 PrE-like cells differentiated upon doxycycline removal and culture in N2B27. The one clone in
142 which fewer PrE-like and more Epi-like cells differentiated had the lowest GATA4-mCherry
143 induction levels (Figures S3C and S3D). In the presence of recombinant FGF4 during the
144 differentiation phase, the fraction of PrE-like cells consistently increased compared to
145 differentiation in N2B27 alone, and mirrored clonal differences in GATA4-mCherry induction
146 levels. In the clonal line with the highest GATA4-mCherry levels, the proportion of PrE-like
147 cells reached a maximum of 98.8 ± 2.0% upon addition of exogenous FGF4 (Figure S3D),
148 indicating that almost all cells have PrE-like differentiation potential following sufficiently
149 strong GATA4-mCherry expression. Taken together, these results suggest that the robust
150 proportioning of cell types is established at the population level through cell-cell signaling.
151

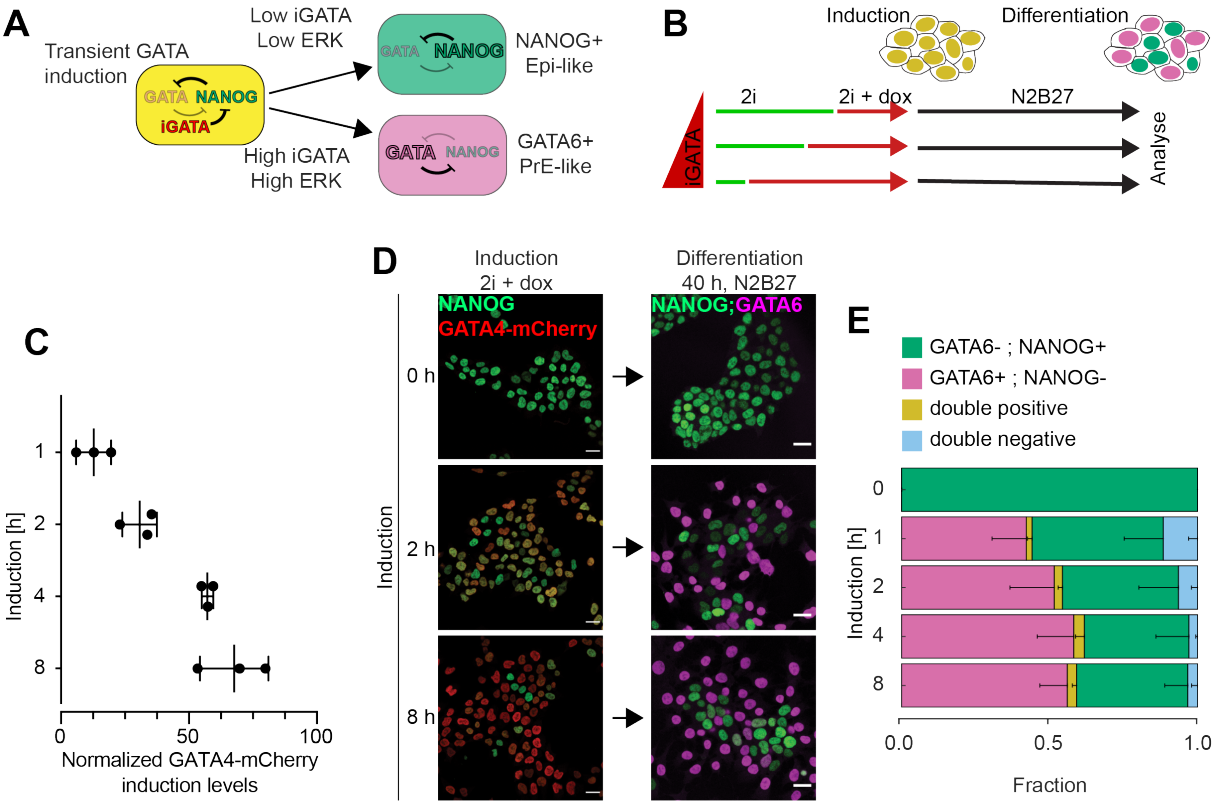


Figure 1. Proportions of differentiated cell types are independent from GATA4-mCherry induction levels

A *Schematic representation of cell differentiation following transient doxycycline-controlled expression of inducible GATA factors (iGATA). For sufficiently high inducible GATA levels and ERK activity, ESCs undergo PrE-like differentiation marked by endogenous GATA6 expression (magenta, bottom). Otherwise they adopt an Epi-like identity marked by NANOG expression (green, top).* ***B*** *Scheme of experimental protocol for titrating inducible GATA4-mCherry expression levels by doxycycline addition to individual samples at different time points. All samples were switched simultaneously to N2B27 medium, such that the total time from seeding to analysis was held constant.* ***C*** *GATA4-mCherry expression levels for different durations of doxycycline induction in 2i + LIF medium measured by flow cytometry, normalized to the non-induced control. Individual data points show mean fluorescence intensities from at least 20.000 cells in an individual experiment, bars indicate mean ± SD across three independent experiments.* ***D*** *Left: immunostaining for NANOG (green) and GATA4-mCherry (red) in inducible cell lines immediately after the end of a doxycycline pulse of the indicated durations. Right: Immunostaining for NANOG (green) and GATA6 (magenta) in cells treated with doxycycline for the*

168 *indicated durations, followed by 40 h of differentiation in N2B27 medium. Cells without doxycycline*
169 *induction have been continuously maintained in 2i medium.* ***E*** *Average cell type proportions from N =*
170 *4 independent experiments; fraction of GATA6+, NANOG- cells in magenta, GATA6-, NANOG+ cells*
171 *in green, double positive cells in yellow, and double negative cells in blue. Error bars: 95% confidence*
172 *interval (CI).*
173

174 **Differentiating ESCs communicate via a local FGF4 signal**

175 To identify candidate mechanisms for cell-cell communication that could underlie this robust
176 cell type proportioning we focused on FGF4, since FGF/ERK signaling is required for PrE
177 differentiation both in ESCs and in the embryo (Kang et al., 2012; Krawchuk et al., 2013;
178 Kunath et al., 2007; Schröter et al., 2015), and paracrine FGF4 is the main activator of ERK in
179 ESCs (Kunath et al., 2007).

180 To investigate how GATA4-mCherry induction levels affect FGF4 signaling in the cell
181 population, a *Sprouty4* $^{H2B\text{-}Venus}$ (*Spry4* $^{H2B\text{-}Venus}$) transcriptional reporter construct that we have
182 previously established as a quantitative readout for long-term FGF4 signaling (Morgani et al.,
183 2018) was integrated in the inducible cell lines. Longer doxycycline induction times
184 corresponding to higher GATA4-mCherry expression levels resulted in reduced mean
185 fluorescence of the *Spry4* reporter after 24 h of differentiation (Figure 2A). This indicates that
186 GATA4-mCherry expression negatively regulates FGF4 signaling during cell type
187 specification. *In situ* mRNA staining (Choi et al., 2018) revealed that this regulation occurred
188 at the transcriptional level. While *Fgf4* mRNA was strongly expressed in most cells before
189 induction in 2i medium (Figure 2B, left), *Fgf4* staining was strongly reduced after 8 h of
190 GATA4-mCherry induction, and confined to cells with low GATA4-mCherry expression levels
191 (arrowheads in Figure 2B, middle). After 40 h of differentiation in N2B27, *Fgf4* expression
192 was mutually exclusive with *Gata6* mRNA expression (Figure 2B, right). This demonstrates

193 that *Fgf4* mRNA expression is negatively correlated with GATA factor expression and PrE-
194 like differentiation in ES cells.
195 We next tested how FGF4 dose affected cell type proportions in ESCs. To be able to control
196 FGF4 levels, we mutated the *Fgf4* gene in the GATA4-mCherry inducible cells. In contrast to
197 the wild type control, *Fgf4* mutant cells continued to express high levels of NANOG, and
198 showed almost no signs of differentiation after an 8 h doxycycline pulse followed by 40 h of
199 culture in N2B27 medium alone (Figures 2C and 2D). Supplementing recombinant FGF4 after
200 the doxycycline pulse rescued the differentiation of two discrete cell types that had similar
201 NANOG- and GATA6-expression profiles as differentiating wild type cells (Figure S4). The
202 proportions of these cell types depended on FGF4 concentration (Figures 2D and S4). Thus,
203 FGF4 signaling and the cell-intrinsic transcriptional circuits underlying cell differentiation
204 mutually regulate each other in a dose-dependent fashion, establishing communication via
205 FGF4 as a potential mechanism for coordinating cell differentiation in the population.
206 We next sought to determine the spatial range of signaling via FGF4 ligands in ESCs. We first
207 tested the role of global communication through FGF4 by comparing differentiation outcomes
208 at different medium-to-cell ratios (Figure 2E). We reasoned that if FGF4 acted globally, ligand
209 concentration would equilibrate in the medium, such that larger volumes would effectively
210 reduce FGF4 concentration, and decrease the proportion of PrE-like in favor of Epi-like cells.
211 In contrast to this expectation, cell type proportions changed negligibly with media volume,
212 (Figure 2E), indicating that dilution of FGF4 ligands in the medium does not strongly affect
213 cell type proportioning.
214 To test whether in contrast the communication is governed by local FGF4 signaling, we
215 disrupted cell-cell contacts by trypsinizing and re-seeding cells at different densities
216 immediately after doxycycline induction (Figure 2F). This treatment strongly reduced the
217 proportion of PrE-like cells compared to the non-trypsinized control (compare second to three
218 rightmost columns in Figure 2F). Furthermore, the proportion of PrE-like cells systematically

219 increased with cell density (three rightmost columns in Figure 2F, $p < 0.05$). Together, these
220 data suggest that cell-cell communication via FGF4 occurs locally and is positively influenced
221 by cell-cell contacts.
222 To directly measure the spatial range of FGF4 signaling in ESC colonies, we seeded single wild
223 type cells labelled with a constitutively expressed dsRed marker onto a layer of *Fgf4* mutant
224 *Spry4*$^{H2B\text{-}Venus}$ reporter cells (Morgani et al., 2018). 12 h after the addition of wild type cells,
225 H2B-Venus was strongly expressed in a halo of reporter cells immediately surrounding the
226 signal-emitting cells, but reporter expression dropped precipitously further away from them
227 (Figure 2G). The spatial decay of the H2B-Venus signal was well approximated by a plateau
228 of ~ 14.4 µm (95% confidence interval 12.8 – 17.0 µm), followed by an exponential decay with
229 a decay length of ~ 11 µm (95% confidence interval 9.2 - 12.4 µm, Figure 2H). This is likely
230 an overestimate of the immediate effective range of paracrine FGF4 signaling, as the
231 transcriptional reporter integrates signaling activity over the entire duration of the experiment,
232 during which cell divisions and movement will increase the distance between signal-sending
233 and –receiving cells. Delaunay triangulation revealed that 16 h after the initiation of
234 differentiation, the mean distance between a cell and its nearest and second-nearest neighbors
235 was 14.0 ± 3.2 µm and 25.5 ± 5.3 µm, respectively (Methods, Figure S5). Thus, the range of
236 cell-cell communication via FGF4 is spatially restricted and mainly couples nearest and second-
237 nearest neighbors.
238

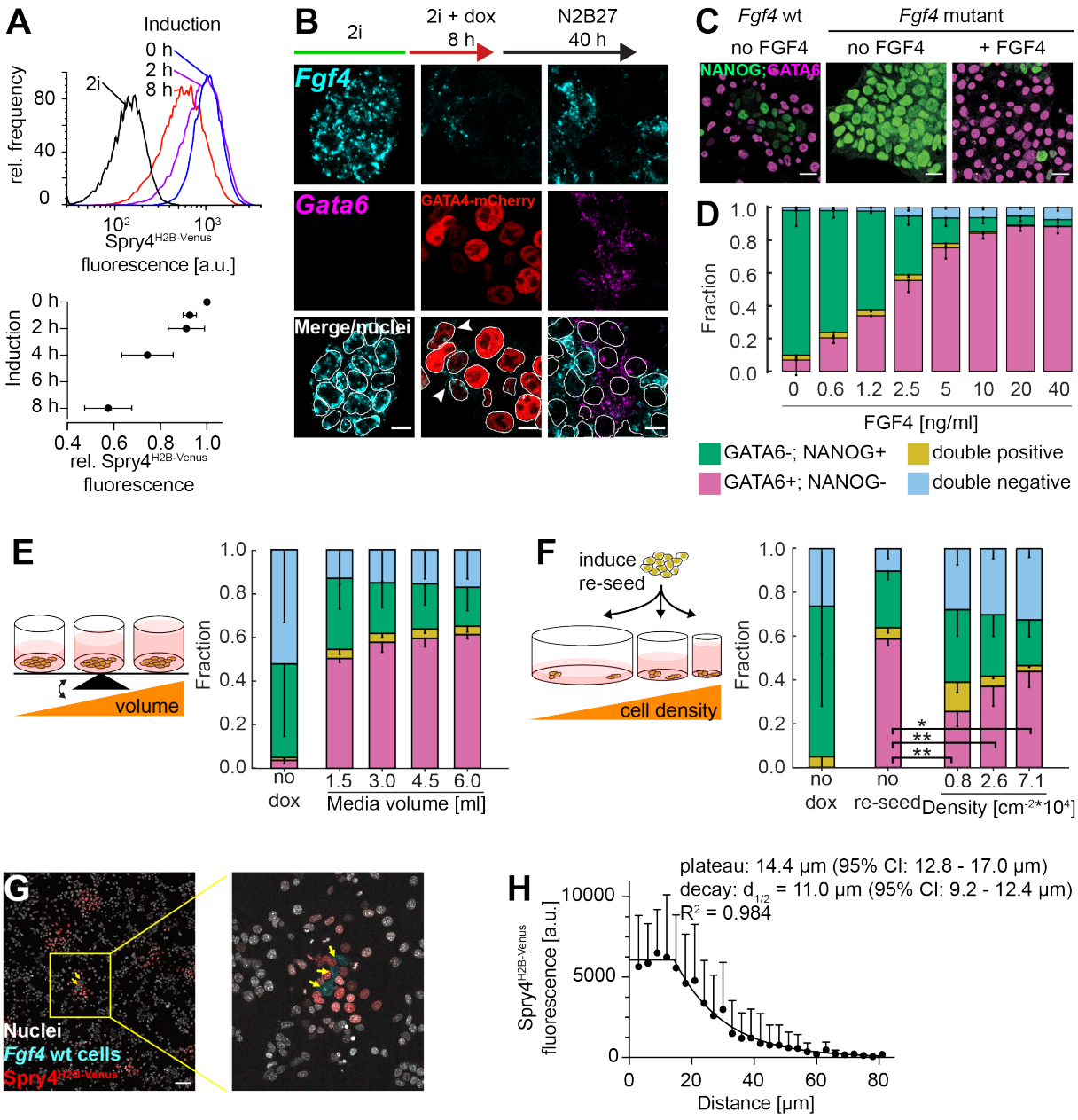


Figure 2. Differentiating ESCs communicate via a local FGF4 signal

A *Top: Flow cytometry histograms depicting expression of a Spry4$^{H2B-Venus}$ reporter integrated into GATA4-mCherry inducible cells, quantified after 24 h of differentiation in N2B27 medium following the indicated durations of doxycycline induction. Black line indicates reporter expression in cells maintained in 2i medium. Bottom: Mean ± SD of reporter expression from N = 4 independent experiments, normalized to fluorescence levels of cells transferred to N2B27 without doxycycline induction.* ***B*** *Fgf4 (top, cyan) and Gata6 (middle, magenta) mRNA staining by in situ hybridization chain reaction at indicated times during the differentiation time-course. At the end of the induction period,*

249 *GATA4-mCherry protein (red) instead of Gata6 mRNA is detected (middle row, middle panel). Bottom*
250 *row: Merged images with outlines of nuclei indicated by thin white lines. Scale bar, 10 µm.* ***C***
251 *Immunostaining for GATA6 (magenta) and NANOG (green) in wild type (left) and Fgf4 mutant cells*
252 *differentiated for 40 h in N2B27 without or with (right) 10 ng/ml FGF4 after an 8 h doxycycline pulse.*
253 *Scale bars, 20 µm.* ***D*** *Average proportions of cell types in Fgf4 mutant cells induced with doxycycline*
254 *for 8 h, followed by differentiation in N2B27 in the presence of the indicated concentrations of FGF4.*
255 *GATA6 and NANOG expression were detected by immunostaining and measured by flow cytometry (see*
256 *Figure S4). N = 4; fraction of GATA6+; NANOG- cells in magenta, GATA6-; NANOG+ cells - green,*
257 *double positive cells - yellow, and double negative cells - blue. Error bars: 95% CI.* ***E*** *Left: Schematic*
258 *representation of experimental setup to test effects of media volume on cell type proportions. Right:*
259 *Average cell type proportions determined by flow cytometry, N = 3.* ***F*** *Left: Schematic representation*
260 *of experimental setup to test effects of cell density on cell type proportions. Right: Average cell type*
261 *proportions determined by flow cytometry, N = 4. ** and * indicate p < 0.005 and 0.05, respectively*
262 *(one-way ANOVA). Color code and error bars in* ***E***, ***F*** *same as in* ***D***. ***G*** *Single labeled Fgf4 wild-type*
263 *cells (cyan, yellow arrowheads) seeded on a layer of Fgf4 mutant Spry4$^{H2B\text{-}Venus}$ transcriptional reporter*
264 *cells. Nuclei are labeled by siR-Hoechst (white), H2B-Venus fluorescence in red. Scale bar, 100 µm.* ***H***
265 *Quantitative analysis of FGF4 signaling range. Data points show mean ± SD of background-subtracted*
266 *H2B-Venus fluorescence intensities in nuclei of Spry4$^{H2B\text{-}Venus}$ reporter cells quantified in distance bins*
267 *of 3 µm width relative to the center of mass of Fgf4 wild type cells. Data from N = 9 independent*
268 *signaling centers. The fluorescence decay length was estimated by fitting a plateau followed by one-*
269 *phase exponential decay to the data (black line).*
270

271 **Cell-cell communication via FGF4 underlies robustness of cell type proportions**

272 We next asked if communication via FGF4 was the molecular mechanism underlying cell type
273 proportioning. To this end, we titrated GATA4-mCherry induction levels, and compared the
274 robustness of cell type proportions between wild type cells that can sense and secrete FGF4,
275 and communication-deficient *Fgf4* mutant cells rescued with a fixed dose of recombinant FGF4
276 (Figures 3A - C). While cell type proportions remained relatively constant in the wild type

277 (Figures 3 A - C, left column), they were strongly dependent on induction times - and hence
278 initial GATA4-mCherry levels - in the rescued *Fgf4* mutant (right column). These broad
279 deviations of cell type proportions in the mutant demonstrate that recursive communication via
280 FGF4 underlies a population-level phenotype that is characterized by the differentiation of
281 robust proportions of cell types from a wide range of starting conditions.
282 To further explore how this population-level behavior of robust proportioning arises from cell-
283 cell communication, we used numerical simulations to determine the dynamical properties of a
284 communicating cell population. We developed a mathematical description by considering a
285 previously characterized circuit consisting of mutually repressive interactions between GATA
286 factors and NANOG (Schröter et al., 2015) that communicates among single cells through FGF
287 signaling (Figure 3D, left). Specifically, we pose that GATA factors repress FGF4 expression
288 (Figure 2), and that FGF signaling represses Nanog expression (Hamilton and Brickman, 2014;
289 Schröter et al., 2015). Cell-cell communication was set between nearest and second-nearest
290 neighbors in a population of N = 10000 cells on a 100 x 100 square grid (Figure S6A).
291 In the simulations, we considered several different scenarios of initial conditions, ranging from
292 all cells being GATA-positive initially to all cells being NANOG-positive initially, to mimic
293 the experimental settings (Figure 3E, left column). Within the coupled population, a stable
294 proportion of two distinct gene expression patterns - (NANOG+, GATA-) or (NANOG-,
295 GATA+) - was established (Figure 3E - G, middle column) from all starting conditions. In
296 contrast, removing the communication in the model such that the dynamics of the mutually
297 repressive NANOG-GATA circuit in single cells is only affected by a constant exogenous FGF
298 input, the cell type proportions strongly depended on the initial conditions distributions (Figure
299 3D - G, right column). This congruence between the theoretical and experimental results
300 indicates that recursive cell-cell communication via FGF signaling is sufficient to recapitulate
301 our observation of robust cell type proportioning.

302 Bifurcation analysis of a minimal system of N = 2 communicating cells indicated that the robust
303 proportioning of the two differentiated cell types in the wild type is a consequence of a new
304 joint dynamical state of the system, characterized with an inhomogeneous steady state (IHSS,
305 Figure S6B). An IHSS is generated at the population level, it is characterized by mutually
306 exclusive gene expression patterns in the population, and has recently been proposed as a
307 generic mechanism for robust cell type proportioning (Stanoev et al., 2019). This suggests that
308 the emergence of an IHSS as a new dynamical solution in the population of FGF-coupled
309 NANOG-GATA mutually repressive circuits is a possible dynamical basis for robust cell type
310 proportioning in differentiating ESCs.

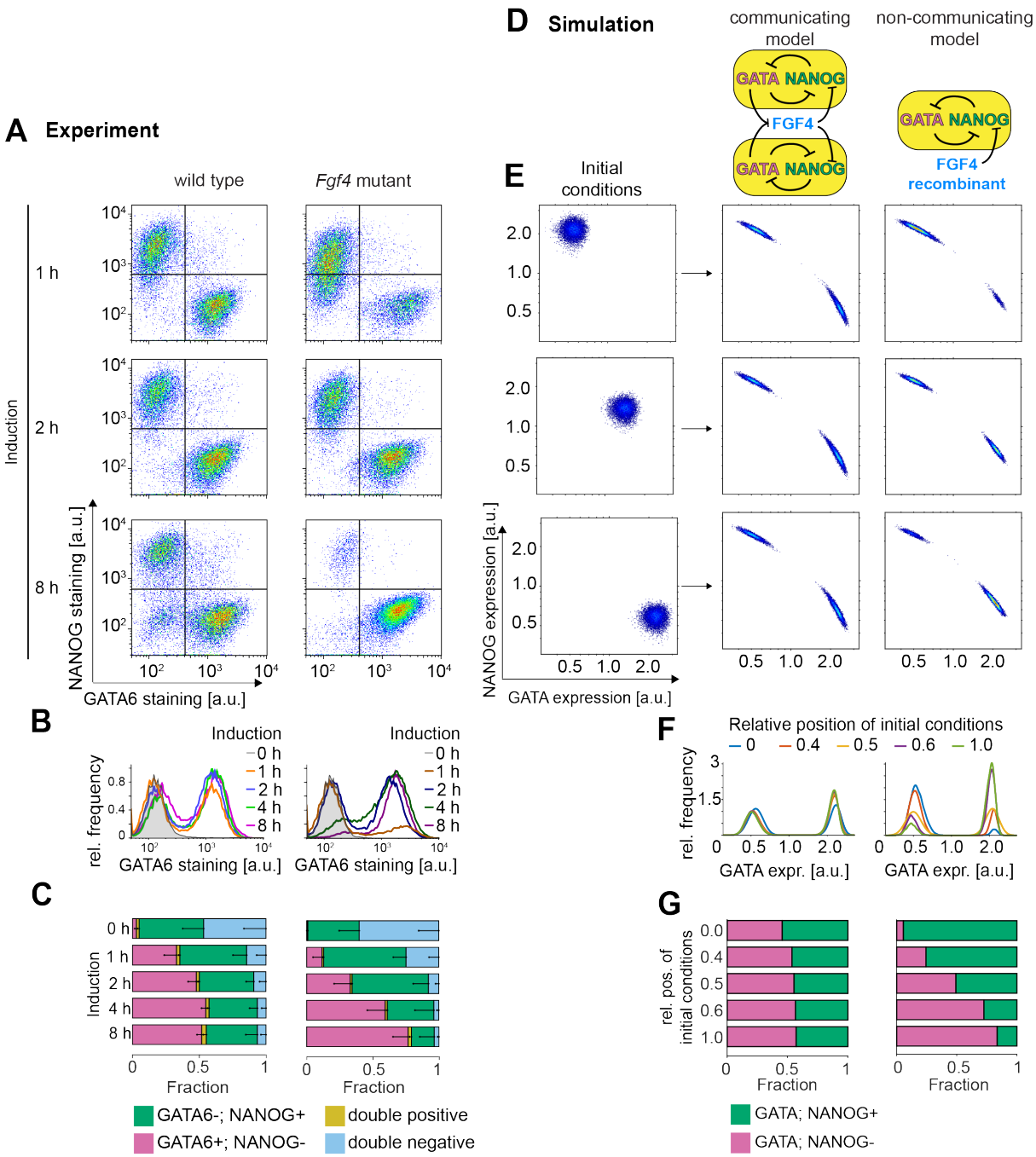


Figure 3. Cell-cell communication via FGF4 mediates cell type proportioning

A – C *Cell differentiation in wild type (left) and Fgf4 mutant cells (right) after doxycycline pulses of variable length, followed by 40 h of differentiation in N2B27 alone (wild type) or in N2B27 supplemented with 10 ng/ml FGF4 (Fgf4 mutant).* ***A*** *Flow cytometry profiles depicting NANOG and GATA6 expression. Lines indicate gates to assign cell types.* ***B*** *Marginal distributions of GATA6 staining.* ***C*** *Quantification of average cell type proportions. GATA6+; NANOG- cells – magenta, GATA6-; NANOG+ cells - green, double positive cells - yellow, and double negative cells - blue. N =*

4, error bars: 95% CI. ***D*** *Schematic representation of the model with (left) and without cell-cell communication (right).* ***E - G*** *Influence of initial conditions (left column in* ***E****) on (GATA6+; NANOG-) and (GATA6-; NANOG+) cell type proportions in the model with (middle column) or without cell-cell communication (right column).* ***E*** *Distributions of initial conditions, shifting from NANOG+ to GATA+ (left column), and corresponding results of numerical simulations with (middle column) or without cell-cell communication (right column).* ***F*** *Respective GATA6 projections, equivalent to* ***B****.* ***G*** *Quantification of the cell type proportions obtained from the numerical simulations in* ***E.*** *See Figure S6 and Methods for model details and parameters.*

Spatial organization of cell types indicates a shift from FGF4-dependent to FGF4-independent patterning mechanisms

The short spatial range of FGF4 signals in ESC cultures suggests that communication via FGF4 not only leads to robust global cell type proportions, but that the differentiated cell types should be also arranged in spatial patterns with local structure. We therefore analyzed the spatial arrangement of cell types at different time points after the initiation of differentiation. Staining for GATA6 and NANOG indicated that cell types were well mixed after 16 h and 24 h of differentiation, but clustering, particularly of GATA6-; NANOG+ cells, was observed after 40 h (Figure 4A). Analysis of the cell type composition of the immediate neighborhood of GATA6+; NANOG- (G+) and GATA6-; NANOG+ (N+) cells, identified with a 2D Gaussian mixture model (Figure S7, methods), corroborated a transition from well-mixed to clustered patterns (Figure S8A).

Next, we set out to quantify cluster sizes during the differentiation time course. We first computed the scaled fraction of N+ cells in neighborhoods of increasing radius around all N+ cells in a field of view (Figure 4B, Methods). The value of the scaled fraction is 1 as long as all cells in the neighborhood are N+, and approaches zero when the composition of the local neighborhood equals the global composition of cell types. We defined the cluster radius as the distance around N+ cells at which this scaled fraction drops to 0.5 (dashed lines in Figure 4B).

347 The measure of cluster radius allows quantifying the length scales of spatially irregular features,
348 however, its absolute value will be smaller than the physical size of the spatial features. In wild
349 type cells, the median cluster radius was 13.0 µm and 14.7 µm at 16 h and 24 h of
350 differentiation, respectively, and increased to 29.0 µm at 40 h of differentiation, corresponding
351 to approximately 0.9, 1.0, and 2.1 cell diameters (Figure 4C, orange).
352 To investigate whether these experimentally determined wild type cluster patterns are
353 consistent with short range signaling, we quantified the cell type spatial organization in model
354 simulations. To be able to capture cluster formation due to cell division events, we considered
355 a dividing cell population, where the grid size was doubled at each cell division event, and the
356 daughter cells inherited gene expression states from the mother cell (Figure S9A). In the wild
357 type case, the cluster radius increased only slightly, from 0.6 to 0.8 cell diameters over five
358 divisions (Figures 4D, S9B, grey). This constant cluster size despite the state propagations
359 during the cell divisions is maintained by the short-range communication that induces cell type
360 transitions following each division in the simulation. The cluster radius in simulations is thus
361 broadly consistent with the experimentally measured values at early, but not at late stages of
362 differentiation.
363 Cluster formation at later stages of differentiation could be driven by long-range
364 communication via FGF4, or by FGF4-independent mechanisms. To distinguish between these
365 possibilities, we analyzed the spatial arrangement of cell types in rescued *Fgf4* mutant cells,
366 using 4 h or 8 h of GATA4-mCherry induction together with differentiation in the presence of
367 10 ng/ml or 2.5 ng/ml FGF4, respectively, to obtain similar cell type proportions as in the wild
368 type (Figures 2B and 3C). In both *Fgf4* mutant conditions, cell types were initially well mixed,
369 and clustered at later stages of differentiation (Figures 4A and S8B – D). The median cluster
370 radius in *Fgf4* mutants was similar to the wild type at 16 h, but then increased more rapidly
371 until 24 h, to reach similar sizes as in the wild type at 40 h (Figures 4B, 4C, blue, and S10). The
372 continuous increase in cluster size was recapitulated in simulations of the mutant case, where

373 differentiation in the individual cells was initialized from a distribution of initial conditions. In
374 these simulations, cluster sizes were initially comparable to the wild type case, but increased
375 rapidly as cells divided, since in the absence of coupling-dependent cell type transitions gene
376 expression states and thus cell types propagated locally (Figures 4D and S9, blue).
377 Consequently, the transition from a well-mixed to a clustered cell type arrangement observed
378 experimentally in the wild type could be recapitulated if following the settling to committed
379 cell types, the communication was removed from the system (Figures 4D and S9, orange).
380 Taken together, these results suggest that at early stages, the observed spatial organization of
381 cell types is consistent with regulation by a local cell-cell communication mechanism. Later on
382 however, additional FGF4-independent mechanisms such as cell division and active cell sorting
383 dominate the spatial organization.

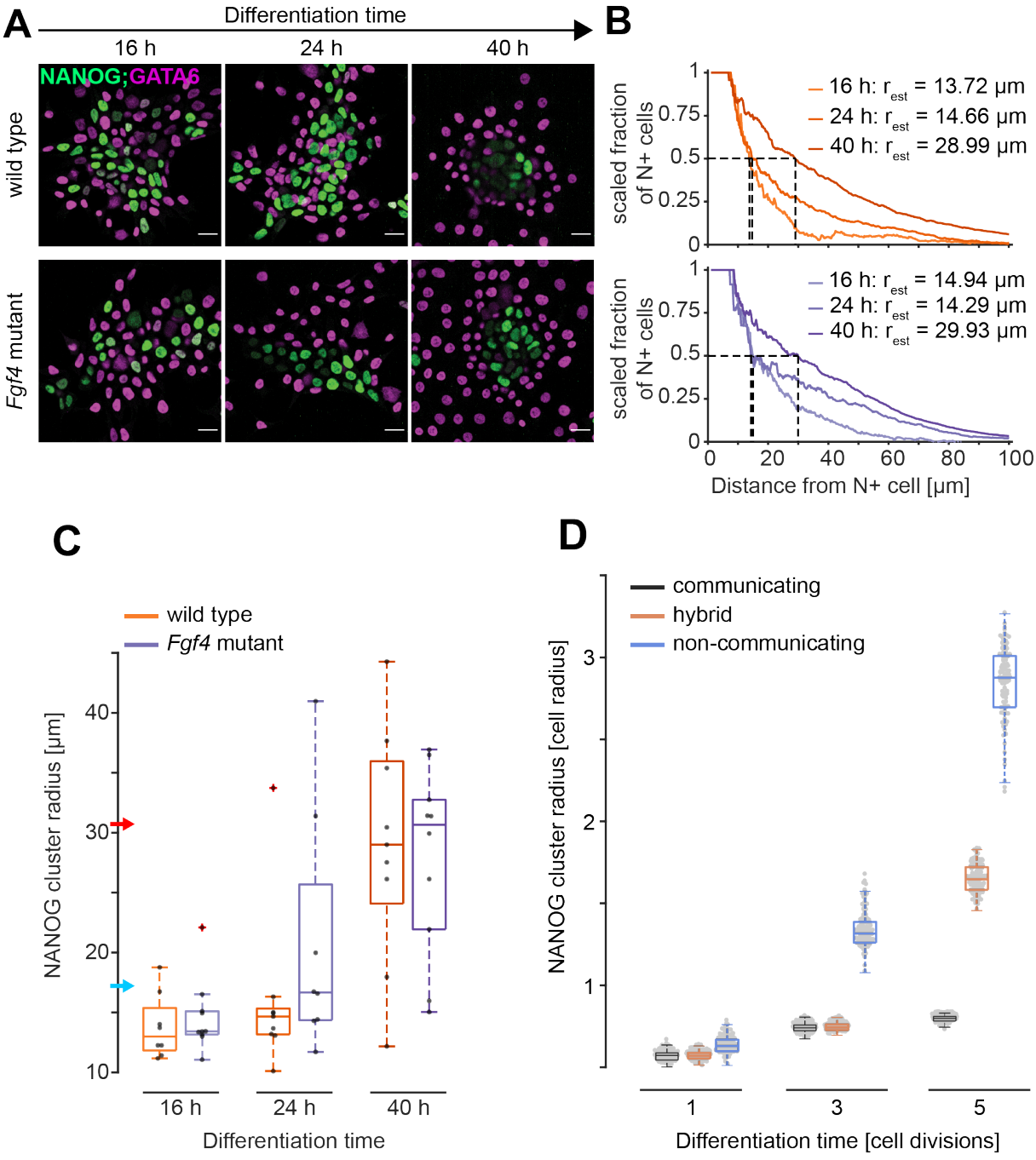


Figure 4. Spatial arrangement of cell types in wild type and Fgf4 mutant cells.

A *Immunostaining of wild type and Fgf4 mutant cells for NANOG (green) and GATA6 (magenta) at different time points after the initiation of differentiation. Wild type cells were induced with doxycycline for 8 h and differentiated in N2B27, Fgf4 mutant cells were induced for 4 h and differentiated in N2B27 supplemented with 10 ng/ml FGF4.* ***B*** *Estimation of NANOG cluster radius in a single field of view for wild type (orange, top) and Fgf4 mutant cells (blue, bottom) differentiated as in* ***A****. Graphs depict the scaled fraction of N+ cells within a specific radius around seed cells. Dashed lines indicate determination of cluster radius, the distance around NANOG+ cells from which the scaled fraction is*

equal to 0.5. ***C*** *Summary statistics of quantified cluster radii for wild type (orange) and Fgf4 mutant cells (blue) differentiated as in* ***A****. Dots indicate values from individual fields of view (n ≥ 8), box plots show median, interquartile ranges and outliers (red cross). Blue and red arrowheads indicate mean distance between nearest and second nearest neighbors, respectively (Figure S5).* ***D*** *NANOG+ cluster radii quantified from 100 independent numerical realizations of the model with continuous communication (grey), without communication (blue), and of a hybrid model where communication is switched off after the third division (orange). Cluster radii were obtained after each cell cycle of a lineage tree simulation, before cell divisions take place.*

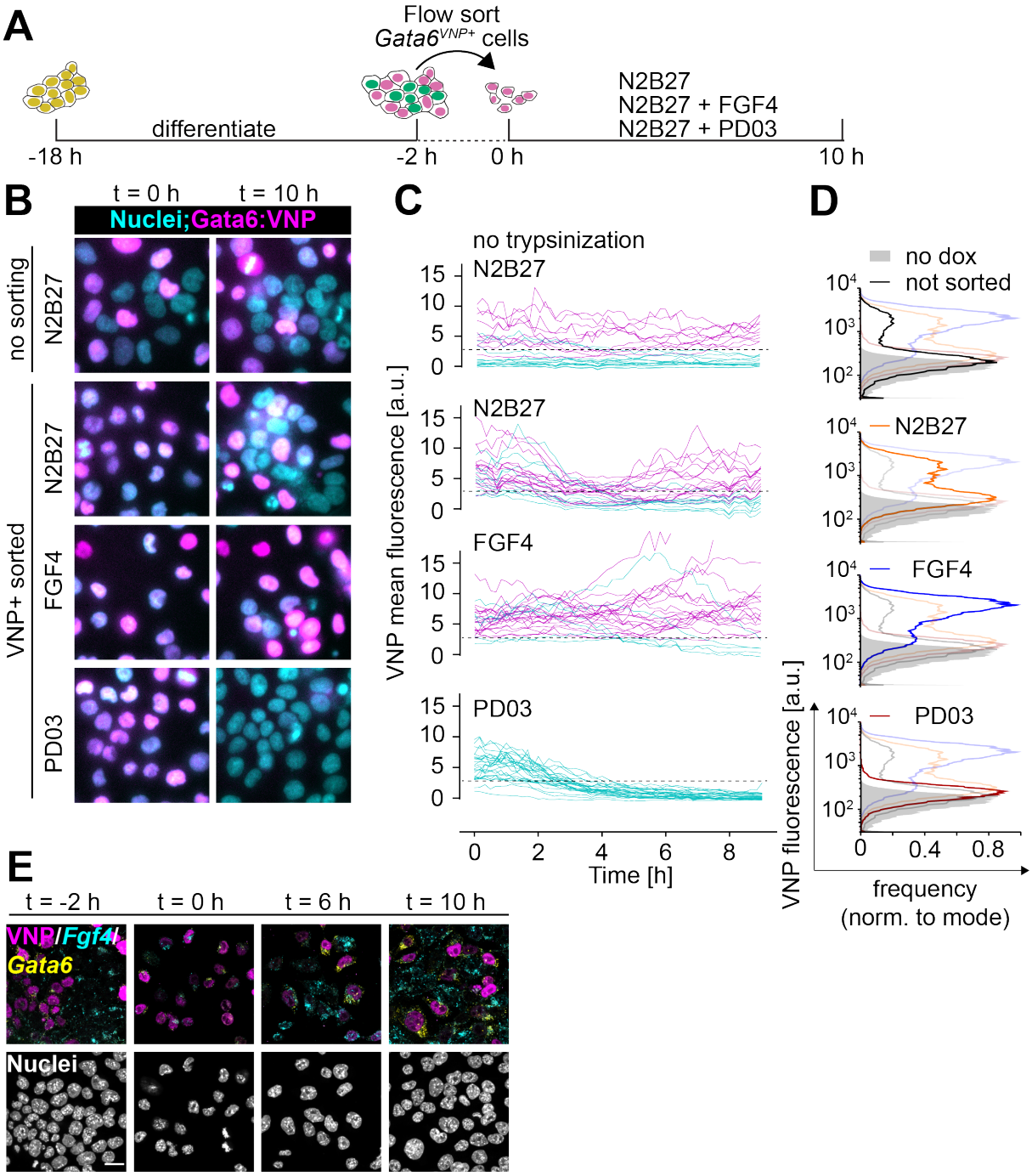
Heterogeneous differentiated cell types are maintained by intercellular communication

A central characteristic of a population-based mechanism for cell differentiation, such as the IHSS, is the interdependence of different cell types (Koseska and Bastiaens, 2017). It has been theoretically demonstrated that this property of the IHSS solution manifests in the regeneration of heterogeneous populations following separation of cell types after differentiation (Stanoev et al., 2019). We asked whether a similar behavior could be observed when PrE-like cells are isolated after both cell types have been established, using a live reporter for *Gata6* expression as a proxy to isolate PrE-like cells, and to subsequently monitor their differentiation state in different media (Figure 5A). Detection of both the up- and the downregulation of expression from the *Gata6* locus requires a short-lived transcriptional reporter protein. We therefore modified an established transcriptional reporter design (Freyer et al., 2015; Schröter et al., 2015) and knocked the coding sequence for the Venus-NLS-PEST protein (Abranches et al., 2013; Nagoshi et al., 2004) into the *Gata6* locus of a GATA4-mCherry inducible cell line.

Putative PrE-like cells were isolated by flow sorting for VNP expression 16 h after the end of a doxycycline pulse, a time-point when their spatial arrangement suggested that differentiation into heterogenous cell types is established by FGF4 communication. When cultured in N2B27 medium, sorted VNP-positive cells regenerated a mixture of VNP-positive and –negative cells

419 within 10 h, resembling cell colonies that had not been disrupted and sorted (Figures 5B – D,
420 first and second row).
421 Flow cytometry indicated a bimodal distribution of VNP expression both in unperturbed
422 cultures and cultures that had been sorted for VNP-expression after 10 h of culture in N2B27
423 (Figure 5D). The smaller proportion of VNP-positive cells in the unperturbed cultures is likely
424 due to insufficient induction of the GATA4-mCherry transgene in this cell line. The dynamics
425 of VNP expression differed markedly between unperturbed cultures and sorted VNP-positive
426 cells growing in N2B27: In unperturbed colonies, VNP expression was stable over time in most
427 cells, while in sorted populations, the reporter was first globally downregulated before the
428 heterogeneous expression patterns emerged (Figure 5C, first and second panel, Movies S1, S2).
429 Similar transitions have been predicted *in silico* as a generic feature of the IHSS solution
430 (Stanoev et al., 2019). The transient downregulation of VNP expression could be prevented by
431 supplementing the culture medium with FGF4, which led to the maintenance of reporter
432 expression in the majority of cells (Figures 5B – D, third row, Movie S3). Inhibition of
433 FGF/ERK signaling with the MEK inhibitor PD03 in contrast resulted in the rapid
434 downregulation of VNP expression following sorting in all cells (Figures 5B – D, fourth row,
435 Movie S4). Reporter expression was likewise downregulated in sorted VNP-positive *Fgf4*
436 mutant cells upon culture in N2B27 alone (Figure S11). These data indicate that cell-cell
437 communication via FGF4/ERK regulates the re-establishment of a mixture of heterogeneous
438 cell types in a population. The dynamics of *Fgf4* mRNA expression after sorting of VNP-
439 positive cells and culture in N2B27 confirmed this (Figure 5E): Immediately after sorting, *Fgf4*
440 transcripts could hardly be detected, as would be expected from the repression of *Fgf4* by
441 endogenous GATA6 in VNP-positive cells. 6 h later, when the VNP reporter and hence
442 endogenous GATA6 expression had dropped and repression was relieved, a subset of cells
443 started to re-express *Fgf4* transcripts. 10 h after sorting, another subset of cells re-expressed
444 *Gata6* mRNA and VNP, such that expression of *Fgf4* transcripts, *Gata6* mRNA and VNP in

445 the population was mutually exclusive, similar to the situation before sorting. Taken together,
446 these results indicate FGF/ERK signaling re-establishes populations with different cell types
447 following the isolation of PrE-like cells. In unperturbed cell colonies, intercellular
448 communication via FGF/ERK therefore not only generates, but also actively maintains
449 balanced proportions of differentiated cell types.
450



451
452

453 ***Figure 5. Heterogeneous cell identities are re-established by cell-cell communication***

454 ***A*** *Schematic representation of the cell-sorting experimental protocol.* ***B*** *Representative images of*
455 *Gata6^{VNP} reporter expression in live cells of a non-trypsinized control (upper row) and in cells sorted*
456 *for VNP expression. Left column is immediately after sorting, right column is after 10 h of culturing in*
457 *N2B27 medium with the indicated supplements.* ***C*** *Live-cell traces of VNP expression in individual cells*
458 *from non-trypsinized colonies (upper panel), or cells sorted for VNP expression upon culture in the*
459 *indicated media. Traces are color coded according to expression levels at the end of the experiment.*
460 *Dashed line indicates the threshold to separate putative VNP-positive cells (magenta) from VNP-*
461 *negative cells (cyan).* ***D*** *Flow cytometry histograms of VNP expression of cells that had not been*
462 *trypsinized and sorted (black, top), and of cells that had been sorted for VNP expression followed by 10*
463 *h of culture in the indicated media. Each panel shows the histogram of the relevant condition as a dark*
464 *line, and distributions of all other conditions shaded for comparison. Non-induced control is in grey in*
465 *all panels.* ***E*** *Staining for Fgf4 (cyan) and Gata6 (yellow) mRNA in Gata6^{VNP} reporter cells before*
466 *sorting (left) and at 2 h, 6 h and 10 h after flow sorting of VNP-positive cells. VNP fluorescence in*
467 *magenta. Scale bar, 20 µm.*

468

469 **Discussion**

470 Here we report emergent population-level behavior during the differentiation of Epi- and PrE-
471 like cells from ESCs expressing inducible GATA factors: Robust proportions of the two cell
472 types are specified from a wide range of GATA induction levels, and re-established from
473 isolated PrE-like cells. This emergent global behavior relies on local cell-cell communication
474 via FGF4. The observed differentiation characteristics recapitulate the properties of a
475 population-based dynamical solution, an inhomogenous steady state, recently proposed as a
476 generic mechanism underlying robust differentiation (Stanoev et al., 2019). Our results suggest
477 a new function for FGF signaling, which is to generate and maintain robust proportions of
478 differentiated cell types.

479

480 The specification of Epi- and PrE-like cells in ESCs shows both molecular and functional
481 parallels to the patterning of the ICM of the mouse preimplantation embryo. In contrast to
482 previous studies that reported PrE-like differentiation in *Fgf4* mutant ESCs upon permanent
483 high-level expression of exogenous GATA factors (Kang et al., 2012; Wamaitha et al., 2015),
484 we find that PrE-like cells do not differentiate from *Fgf4* mutant ESCs upon transient GATA
485 induction. This recapitulates the *Fgf4* mutant phenotype in the embryo (Feldman et al., 1995;
486 Kang et al., 2012; Krawchuk et al., 2013). In both *Fgf4* mutant ESCs and embryos, cell type
487 proportions can be controlled by recombinant FGF4 in a dose-dependent manner (Yamanaka
488 et al., 2010; Krawchuk et al., 2013). Furthermore, the differentiation of Epi- and PrE-like cells
489 *in vitro* recapitulates the remarkably constant proportions of cell types seen in the developing
490 embryo (Saiz et al., 2016). Lastly, cell identities in ESC populations are plastic and can be re-
491 specified upon changing a cell's environment, again similar to observations in the embryo
492 (Grabarek et al., 2012; Martinez Arias et al., 2013). A recent study using chimaeras and targeted
493 ablation of specific cell types concluded that an FGF4-based population-level mechanism
494 balances the size of the Epi and the PrE lineage in the mouse embryo (Saiz et al., 2020). We
495 used doxycycline-inducible transgenes in wild type and *Fgf4* mutant ESCs as an orthogonal
496 experimental system to demonstrate that a similar molecular and dynamical mechanism
497 operates during cell differentiation *in vitro*. In addition to the previous observations, the ESC
498 system also allowed us to show that cell type proportions can be robustly established from a
499 broad range of initial conditions. The first direct measurements of the spatial range of cell-cell
500 communication via FGF4 in differentiating ESCs finally provides indications how the spatial
501 arrangement of cell types is established during development.
502
503 The short signaling range of FGF4 during the differentiation of Epi- and PrE-like cells *in vitro*,
504 together with the progressive FGF-independent spatial clustering once cell types have been
505 specified, suggest an evolutionary explanation for the spatially random differentiation of Epi

506 and PrE cells in the ICM (Chazaud et al., 2006; Plusa et al., 2008; Rossant et al., 2003). Rather
507 than settling on a mechanism that produces the end result of spatially separated tissues in one
508 step, evolution might have taken a "tinkering" approach (Jacob, 1977), combining
509 communication via FGF4 as a mechanism that allows specifying robust proportions of cell
510 types, with a subsequent sorting step, a sequence that is conserved *in vitro.*
511
512 FGF4 signaling during the differentiation of Epi- and PrE-like cells shows several parallels to
513 general properties of the Delta-Notch signaling system. The signaling range of FGF4 is spatially
514 restricted, akin to the juxtacrine signaling mode of Delta-Notch. Communication via FGF4 in
515 differentiating ESCs establishes and maintains discrete cell types through repressive coupling
516 - signals sent by one cell type repress the identity of that same type in neighboring cells – much
517 like Delta-Notch signaling during lateral inhibition (Ferrell, 2012; Henrique and Schweisguth,
518 2019; Hori et al., 2013; Simpson, 1990). The short spatial range of the FGF4 signal requires
519 cell types to be well mixed immediately after specification, similar to the spatial patterning of
520 cell types arising from Delta-Notch-mediated lateral inhibition (Collier et al., 1996).
521 Consequently, several hallmarks of the emergent population-level behavior established by
522 FGF4 in differentiating ESCs have been observed in an engineered cell system where cells
523 communicate via Delta-Notch, such as the differentiation of discrete cell types in reproducible
524 proportions, the re-establishment of those proportions upon removal of one cell type, and the
525 dependence of cell type proportions on cell density or contact (Matsuda et al., 2015). Thus,
526 when embedded in appropriate intracellular regulatory circuits, molecularly diverse
527 intercellular communication systems can yield similar functional outputs.
528
529 In rescued *Fgf4* mutant cells that do not communicate, differentiation outcomes in the cell
530 population strongly depend on the distribution of GATA4-mCherry induction levels. This is in
531 line with predictions from single cell models for cell differentiation, in which initial conditions

532 in individual cells strongly influence their differentiation path (Huang et al., 2007). When cells
533 communicate via FGF4 in contrast, the collective differentiation outcome is robust and
534 becomes independent from the distribution of GATA4-mCherry induction levels. Thus, the
535 behavior of the communicating cell population cannot directly be extrapolated from the
536 behavior of single isolated cells.
537 Theoretically, the conceptual differences between single-cell and population-based modes of
538 differentiation manifest in the emergence of a new type of solution that jointly describes the
539 heterogeneous cell identities, an IHSS, in the communicating cell population. The robust
540 generation of cell type proportions irrespective of initial conditions, and their active
541 maintenance through intercellular communication that we observe experimentally, are two key
542 properties of the IHSS (Stanoev et al., 2019). This suggests that the IHSS is a likely dynamical
543 mechanism underlying the differentiation of cells with discrete identities during mammalian
544 preimplantation development. Given the pervasiveness of robust cell type proportioning during
545 development and homeostasis (Viader-Llargués et al., 2018), it is likely that similar population-
546 based mechanisms underlie canalized development in diverse systems in which multipotent
547 progenitor cells give rise to several differentiated cell types.
548

549 **Methods**

550

551 **Cell lines**

552 Cell lines used in this study were E14tg2a (Hooper et al., 1987) and an *Fgf4* mutant *Spry4*$^{H2B-}$
553 $^{Venus/+}$ line that we have previously described (Morgani et al., 2018). dsRed-labelled cells were
554 from an E14tg2a-background and kindly supplied by J Nichols. The Gata6:VNP reporter was
555 established in the background of a line carrying a doxycycline-inducible GATA4-mCherry
556 transgene in the Col1a1 locus as well as a randomly integrated H2B-Cerulean nuclear marker
557 driven by a CAGS promoter described in (Schröter et al., 2015).

558 E14tg2a-based inducible cell lines were maintained on fibronectin-coated tissue culture plastic
559 in 2i + LIF medium, which consists of a N2B27 basal medium supplemented with 3 μM
560 CHIR99021 (Tocris), 1μM PD0325901 (SelleckChem) and 10 ng/ml LIF (protein expression
561 facility, MPI Dortmund). For maintenance of *Fgf4* mutant subclones, we supplemented the 2i
562 + LIF medium with 10% fetal bovine serum (FBS), as *Fgf4* mutant lines showed severely
563 decreased proliferation upon long-term culture in 2i + LIF alone. FBS was removed at least one
564 day before the experiment.

565 *Spry4*-reporter cell lines to measure signaling range, as well as *Gata6*-reporter cell lines were
566 maintained on gelatin coated dishes in GMEM-based medium supplemented with 10% FBS,
567 sodium pyruvate, 50 μM β-mercaptoethanol, glutamax, non-essential amino acids and 10 ng/ml
568 LIF. 1 μM PD0325901 was added to the cultures of *Spry4*- and *Gata6*-reporters three days
569 before the experiment, to downregulate *Spry4* reporter expression, or to capacitate cells for PrE-
570 like differentiation (Schröter et al., 2015).

571 FGF4 was from Peprotech and supplied in the indicated concentrations, together with 1 μg/ml
572 heparin (Sigma).

573

574

575 **Genetic engineering of ESC lines**

576 Doxycycline-inducible *GATA4*-mCherry inducible ES cells were generated by electroporation
577 of 50.000 E14tg2a ES cells with 4 µg of pPB-TET-GATA4-mCherry, 4 µg pCAG-rtTA-Neo,
578 and 4 µg pCAG-PBase (Wang et al., 2008) followed by G418 selection (400 µg/ml) one day
579 after transfection. We established more than 10 independent clonal lines and assayed induction
580 levels and homogeneity by flow cytometry 2 – 8 h after induction of transgene expression by
581 adding 500 ng/ml doxycycline to the culture medium. Four clones with homogeneous induction
582 levels were chosen and maintained under G418 selection, to circumvent silencing of the
583 inducible transgene.

584 Mutagenesis of the *Fgf4* locus was performed as previously described (Morgani et al., 2018).
585 *Fgf4* loss of function clones were identified by PCR-amplification, cloning and sequencing of
586 a sequence around the *Fgf4* start codon. We either selected clones with a targeted mutation
587 delivered by a single-stranded DNA repair template that we have previously shown to disrupt
588 *Fgf4* function (Morgani et al., 2018), or selected at least two independent clones carrying indels
589 around the start codon that introduced frameshift as well as nonsense mutations. All
590 independent clones with random indels showed indistinguishable behavior in the differentiation
591 assays.

592 The *Gata6* reporter cell line was generated using previously described knock-out first targeting
593 arms of the EUCOMM project (Skarnes et al., 2011), combined with a VNP reporter cassette
594 (Nagoshi et al., 2004) and a neomycin resistance gene driven from a human β-actin promoter.
595 This construct was integrated by homologous recombination into a line carrying a doxycycline-
596 inducible GATA4-mCherry transgene in the *Col1a1* locus as well as a randomly integrated
597 H2B-Cerulean nuclear marker driven by a CAGS promoter described in (Schröter et al., 2015).
598 Clones were screened for correct integration of the reporter construct by long range PCR
599 spanning the targeting arms.

600 The targeting construct to generate the *Spry4*$^{H2B\text{-}Venus}$ allele in GATA4-mCherry inducible cell

601 lines was based on the one used in (Morgani et al., 2018), except that the puromycin selectable
602 marker was exchanged for a neomycin cassette. The construct was integrated into ESCs by
603 homologous recombination. To increase targeting efficiency, cells were co-transfected with a
604 plasmid expressing Cas9 and a sgRNA that targets a sequence at the 5' end of the 5' targeting
605 arm which is present in the endogenous *Spry4* locus, but not the targeting construct. Neomycin-
606 resistant clones were expanded and screened for correct integration of the reporter construct by
607 long range PCR spanning the targeting arms. All genetically modified lines were karyotyped
608 using standard procedures (Nagy et al., 2008), and only lines with a median chromosome count
609 of n = 40 were used for experiments.
610

611 **Immunostaining and image analysis**

612 Immunostaining of adherent cells was performed as previously described (Schröter et al., 2015).
613 Antibodies used were anti-NANOG (e-bioscience, eBioMLC-51, 14-5761-80, final
614 concentration 2.5 µg/ml), anti-GATA6 (R&D AF1700, final concentration 1 µg/ml), and anti-
615 FLAG (Sigma-Aldrich F1804-200, final concentration 1 µg/ml). Secondary antibodies were
616 from Invitrogen/LifeTech. Images were acquired using a 63x 1.4 N.A. oil-immersion objective
617 on a confocal Leica SP8 microscope, with all settings held constant between replicates. Images
618 were quantified using custom scripts written for ImageJ (NIH) and in Matlab (The Mathworks).
619

620 **In situ HCR**

621 Probe sets for *Nanog*, *Gata6* and *Fgf4* and corresponding Alexafluor-labelled amplifiers for
622 staining of mRNA molecules via third generation in situ HCR (Choi et al., 2018) were sourced
623 from Molecular Instruments. Staining was performed according to manufacturer's instructions.
624 Briefly, adherent cells were fixed for 15 minutes with 4% paraformaldehyde, washed with PBS
625 and permeabilized for several hours in 70% ethanol at -20ºC. Cells were then washed twice
626 with 2x SSC and equilibrated in probe hybridization buffer for at least 30 minutes. Transcript-

627 specific probes were used at a concentration of 4 nM and hybridized overnight. Excess probe
628 was removed through several washes with probe wash buffer and 5x SSCT, and cells were
629 equilibrated in amplification buffer for at least 30 minutes. Fluorescently labeled amplifiers
630 were used at a concentration of 60 nM. Amplification was allowed to proceed for 16 – 24 hours
631 at room temperature. Excess amplifier was removed by several washes with 5x SSCT, followed
632 by counterstaining with Hoechst 33342 and mounting in glycerol-based medium. Imaging was
633 performed on an SP8 confocal microscope with a 63x (NA1.4) lens.
634
635 **Flow cytometry**
636 Staining for flow cytometric analysis of intracellular antigens was performed as previously
637 described (Schröter et al., 2015). Primary and secondary antibodies were the same as used for
638 immunostaining. mCherry fluorescence measurements and cell sorting were performed on a
639 BD FACS Aria. All other flow cytometric analysis was carried out using a BD LSR II. Single
640 cell events were gated based on forward and side scatter properties. GATA4-mCherry
641 expression measurements were normalized to the respective uninduced control. Gates to
642 separate marker-positive from marker-negative cells were determined visually as the threshold
643 that best bisected the bimodal distribution of marker expression across all samples within one
644 experiment.
645
646 **Decay length measurements**
647 *Fgf4* mutant Spry4$^{H2B\text{-}Venus}$ reporter cells (Morgani et al., 2018) were seeded at a density of
648 $5*10^4$ cells/cm^2 in N2B27. 2 hours later, dsRed-expressing cells were added at a density of 500
649 cells/cm^2. For the first 3 hours of co-culture, the medium was supplemented with 250 to 500 nM
650 siR-Hoechst (Lukinavičius et al., 2015) to label nuclei. 12 h later, live cells were imaged on a
651 Leica SP8 confocal system. Nuclei were segmented in FIJI (Schindelin et al., 2012), and for
652 each Spry4$^{H2B\text{-}Venus}$ reporter cell in the vicinity of a ds-Red expressing cell, the background-

subtracted Venus fluorescence intensity as well as the distance to the center of mass of the dsRed expressing cells was determined. Cells were grouped according to their distance from dsRed expressing cells in 3 μm bins, and mean fluorescence intensities for each bin plotted versus their distance. Decay length was estimated in GraphPad Prism by fitting a plateau followed by a one-phase decay function.

Determination of cell-cell distances

Neighborhood graphs were constructed for each field of view by using the cell positions to generate a Delauney triangulation and the corresponding Voronoi diagram. Spurious links between non-adjacent cells were trimmed by excluding links in the Delauney graph that do not directly pass through the shared Voronoi edge between the two respective Voronoi cells. Links between adjacent cells (purple links in Fig S5A) were pooled to generate a distance distribution between nearest neighbors (purple histogram, Fig S5B). The distances of links between unconnected cells that share a neighbor (yellow in Fig S5A), were likewise pooled to give the distribution of second-nearest neighbor distances (yellow histogram, Fig S5B). Both distributions were independently fit with a 2-component Gaussian mixture model, to separate the true distributions of nearest and second-nearest neighbor distances from higher-order distributions arising from erroneously assigned links (purple and yellow dashed lines, Fig S5B). The mean and SD of nearest- and second-nearest neighbor distances was estimated from the first component of these Gaussian fits.

Assignment of cell types with Gaussian mixture model

For spatial clustering analysis, we focused on GATA6+; NANOG- (G+) and GATA6-; NANOG+ (N+) cells only. To identify these cell types specifically, we applied a two-component Gaussian mixture model (GMM) fit to single cell distributions in the GATA6; NANOG expression space, using the MATLAB function `fitgmdist()`. The fitted GMM

679 assigns each cell a posterior probability of its association with one of the two component
680 distributions. Cells with a posterior probability of >0.9 for one of the two components were
681 classified to the corresponding cell type.
682
683 **Analysis of neighborhood composition**
684 To determine local neighborhood composition, we first generated a matrix of Euclidean
685 distances between every cell classified as either G+ or N+ for every individual field of view.
686 For every cell, we then computed the fraction of G+ and N+ cells within a Euclidean distance
687 of 31.7 μm, encompassing most of the nearest and second nearest neighbors. Random
688 distributions (blue lines in Figure S8) were calculated by populating the local neighborhood of
689 each cell randomly from the global distribution in each treatment.
690
691 **Measurement of cluster radius**
692 The average cluster radius of N+ cells was estimated with the same method in experimental
693 data where cell types had been categorized with a Gaussian mixture model, and in simulations
694 of the model. We calculated for all N+ cells in a field of view or simulation run the decrease in
695 the fraction of N+ cells in neighborhoods of increasing radius, settling to the overall fraction of
696 N+ cells. For scaling, we first subtracted the overall fraction of N+ cells in the respective field
697 of view or simulation run, and then normalized values such that the scaled fraction at zero
698 distance was set to one again. The cluster radius is the distance at which the scaled fraction of
699 N+ cells is equal to 0.5.
700
701 **Statistical analysis**
702 GraphPad Prism was used to perform significance testing (ANOVA).
703
704 **Live cell imaging and tracking**

705 To track *Gata6* reporter expression in live cells, PrE-like differentiation was induced by a 6 h
706 pulse of doxycycline-treatment in serum-containing medium as described in (Schröter et al.,
707 2015). 16 hours after doxycycline-removal, cells were either switched directly to N2B27
708 medium lacking phenol red, or trypsinized, sorted for reporter expression, and seeded on
709 fibronectin-coated imaging dishes (ibidi μ-slides). Time-lapse imaging was started within 2 h
710 after sorting on an Olympus IX81 widefield microscope equipped with LED illumination
711 (pE4000, CoolLED) and a Hamamatsu c9100-13 EMCCD camera. Hardware was controlled
712 by MicroManager software (Edelstein et al., 2001). Time-lapse movies were acquired using a
713 40x oil immersion lens (NA 1.2), with 10-minute time intervals.
714 Cell tracking was carried out with TrackMate (Tinevez et al., 2017) based on the constitutively
715 expressed H2B-Cerulean nuclear marker. Fluorescence intensity was measured in a circular
716 region of interest in the center of the nucleus, and background-subtracted fluorescence
717 intensities plotted in Python. Trace color in Figure 4D was assigned according to fluorescence
718 intensity in the last frame of the movie, with respect to the estimated intensity threshold used
719 for flow sorting (dashed line).
720
721 **Computational model for cell type proportioning**
722 The model of the intercellular communication system (Figure 3D) is adapted from Stanoev et
723 al. (2019), and is described with the following set of equations:

724 $$\frac{1}{\lambda}\frac{dN_i}{dt} = \alpha_N \frac{1}{1+G_i^{\beta}} + \alpha_{N,F}\frac{1}{1+F_{ext,i}^{\eta}} - N_i$$

725 $$\frac{1}{\lambda}\frac{dG_i}{dt} = \alpha_G \frac{1}{1+N_i^{\gamma}} - G_i$$

726 $$\frac{1}{\lambda}\frac{dF_i}{dt} = \alpha_F \frac{1}{1+G_i^{\delta}} - F_i$$

727 N_i and G_i describe NANOG and GATA6 protein expression levels in cell i, regulated by mutual
728 inhibition, while F_i is the secreted FGF4 whose production is downregulated by GATA6.

729 $F_{ext,i} = \frac{1}{|N(i)|+1} \sum_{j \in (N(i) \cup i)} F_j$ is the extracellular FGF4 concentration that is sensed by cell i

730 from its neighborhood $N(i)$, resulting in downregulation of NANOG production in the cell.

731 $\alpha_N = 2.5, \alpha_{N,F} = 0.5, \alpha_G = 3$ and $\alpha_F = 3$ denote production rate constants, $\beta = \eta = \gamma = \delta =$

732 2 are the Hill coefficients, degradation rates were set to 1 as $\lambda = 50$ was used as a scaling

733 kinetic parameter. 10000 cells were deployed on a regular 100x100 two-dimensional lattice

734 with no-flux boundary conditions. Cell-cell communication was modeled to be short-range,

735 reflecting the experimental wild-type case, i.e. communication between direct neighbors and

736 cells two hops away on the lattice (Figure S6A). When mimicking the *Fgf4* mutant case,

737 communication between cells was excluded, and an external input was modeled with $F_{ext} =$

738 1.2.

739 The cell populations were initiated analogously to the experimental case, by varying the initial

740 conditions of all cells from being NANOG-expressing, through intermediate NANOG and

741 GATA6 expression, to being GATA6-expressing (Figure 3E, left column). More specifically,

742 the variables were sampled independently from unimodal Gaussian distributions

743 $\mathcal{N}(\mu_{ics}(p), \sigma_{ics} = 0.1 * \mu_{ics}(p))$, with the mean $\mu_{ics}(p) = (1-p) * \mu_{G-;N+} + p * \mu_{G+;N-}$

744 placed on the line segment connecting the GATA6-; NANOG+ state $\mu_{G-;N+}$ and the GATA6+;

745 NANOG- state $\mu_{G+;N-}$, partitioning it in proportion p. $p \in \{0, 0.4, 0.5, 0.6, 1\}$ was used for the

746 quantifications in Figure 3G. Samples from around the endpoints and the midpoint ($p \in$

747 $\{0, 0.5, 1\} \Rightarrow \mu_{ics} \in \{\mu_{G-;N+}, \frac{1}{2}(\mu_{G-;N+} + \mu_{G+;N-}), \mu_{G+;N-}\}$) are shown in Figure 3E, left

748 column.

749 Cell heterogeneity was introduced by varying all of the parameters independently with standard

750 deviation of 0.02 from the respective values for each cell. Stochastic differential equation model

751 was constructed from the deterministic equations by adding a multiplicative noise term $\sigma X dW_t$,

752 where dW_t is the Brownian motion term, X is the variable state and $\sigma = 0.1$ is the noise term.

753 The model was solved with $\Delta t = 0.01$ using the Milshtein method (Milshtein, 1974). Following

integration, cell identities were estimated by comparing the NANOG and GATA6 values from the final states of the cells, and the ratios were computed.

For comparing the spatial organizations between communicating and non-communicating cells at different time points, periodic synchronous cell divisions in the population were included in the model as in (Stanoev et al., 2019), spanning 5 cell cycles. Cell divisions occur along the horizontal and vertical axes on the grid alternately, sequentially yielding lattices of 10x10, 10x20, 20x20, 20x40 and 40x40 (Figure S9A). At every cell division, the final state of the mother cell is passed on to the daughter cells' initial conditions. The mother cell's parameter set is also inherited. Spatial organizations were analyzed at the end of each cell cycle, after the collective state is allowed to reach a steady state in a deterministic fashion, by estimating the N+ cluster radius as described above.

For the hybrid model, it was assumed that after the third cell cycle the cells commit to their current fates and the communication becomes inconsequential, effectively bringing about a switch to a non-communicating grid. For all conditions, cells' states were initialized with Gaussians with $\mu_{ics}(0.5)$, as described above.

Acknowledgements

We thank J. Nichols for sharing dsRed-labelled ESCs, L. Süther for help with generating inducible cell lines, M. Schulz and S. Müller for assistance with microscopy and flow cytometry, and P. Bastiaens for stimulating discussion and conceptual input on the project. We thank P. Bieling, J. de Navascues, G. Vader, S. Fischer and the members of the Schröter and Koseska groups for discussions and comments on the manuscript. All authors are supported by the Max Planck Society.

Author contributions

Conceptualization, A.K. and C.S.; Methodology, A.S. and C.S.; Investigation, D.R., A.B., A.S., M.P., and C.S.; Validation, D.R., A.B., and M.P.; Visualization, D.R. and A.B.; Formal analysis, D.R. and A.B.; Writing – original draft, C.S.; Writing – review and editing, all authors; Supervision, A.K. and C.S..

References

Abranches, E., Bekman, E., Henrique, D., 2013. Generation and Characterization of a Novel Mouse Embryonic Stem Cell Line with a Dynamic Reporter of Nanog Expression. PLoS ONE 8, e59928. doi:10.1371/journal.pone.0059928.s007

Bedzhov, I., Graham, S.J.L., Leung, C.Y., Zernicka-Goetz, M., 2014. Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. Philos Trans R Soc Lond, B, Biol Sci 369, 20130538. doi:10.1098/rstb.2013.0538

Bessonnard, S., De Mot, L., Gonze, D., Barriol, M., Dennis, C., Goldbeter, A., Dupont, G., Chazaud, C., 2014. Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. Development 141, 3637–3648. doi:10.1242/dev.109678

Bradley, A., Evans, M., Kaufman, M.H., Robertson, E., 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature 309, 255–256.

Chazaud, C., Yamanaka, Y., Pawson, T., Rossant, J., 2006. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell 10, 615–624. doi:10.1016/j.devcel.2006.02.020

Chickarmane, V., Peterson, C., 2008. A Computational Model for Understanding Stem Cell, Trophectoderm and Endoderm Lineage Determination. PLoS ONE 3, e3478–8. doi:10.1371/journal.pone.0003478

Choi, H.M.T., Schwarzkopf, M., Fornace, M.E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., Pierce, N.A., 2018. Third-generation in situhybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. Development 145, dev165753–122. doi:10.1242/dev.165753

Collier, J.R., Monk, N.A., Maini, P.K., Lewis, J.H., 1996. Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling. J Theor Biol 183, 429–446. doi:10.1006/jtbi.1996.0233

De Caluwé, J., Tosenberger, A., Gonze, D., Dupont, G., 2019. Signalling-modulated gene regulatory networks in early mammalian development. J Theor Biol 463, 56–66. doi:10.1016/j.jtbi.2018.12.008

De Mot, L., Gonze, D., Bessonnard, S., Chazaud, C., Goldbeter, A., Dupont, G., 2016. Cell Fate Specification Based on Tristability in the Inner Cell Mass of Mouse Blastocysts. Biophysj 110, 710–722. doi:10.1016/j.bpj.2015.12.020

Edelstein, A., Amodaj, N., Hoover, K., Vale, R., Stuurman, N., 2001. Computer Control of Microscopes Using µManager. John Wiley & Sons, Inc., Hoboken, NJ, USA. doi:10.1002/0471142727.mb1420s92

Feldman, B., Poueymirou, W., Papaioannou, V.E., DeChiara, T.M., Goldfarb, M., 1995. Requirement of FGF-4 for postimplantation mouse development. Science 267, 246–249. doi:10.2307/2885904

824 Ferrell, J.E., Jr, 2012. Bistability, Bifurcations, and Waddington's Epigenetic Landscape.
825 Current Biology 22, R458–R466. doi:10.1016/j.cub.2012.03.045
826 Freyer, L., Schröter, C., Saiz, N., Schrode, N., Nowotschin, S., Martinez Arias, A.,
827 Hadjantonakis, A.-K., 2015. A loss-of-function and H2B-Venus transcriptional reporter
828 allele for Gata6 in mice. BMC Dev Biol 15, 38. doi:10.1186/s12861-015-0086-5
829 Gardner, R.L., 1968. Mouse chimeras obtained by the injection of cells into the blastocyst.
830 Nature 220, 596–597. doi:10.1038/220596a0
831 Grabarek, J.B., Zyzynska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J.,
832 Hadjantonakis, A.-K., Plusa, B., 2012. Differential plasticity of epiblast and primitive
833 endoderm precursors within the ICM of the early mouse embryo. Development 139, 129–
834 139. doi:10.1242/dev.067702
835 Hamilton, W.B., Brickman, J.M., 2014. Erk Signaling Suppresses Embryonic Stem Cell Self-
836 Renewal to Specify Endoderm. CellReports 1–16. doi:10.1016/j.celrep.2014.11.032
837 Henrique, D., Schweisguth, F., 2019. Mechanisms of Notch signaling: a simple logic
838 deployed in time and space. Development 146, dev172148–11. doi:10.1242/dev.172148
839 Hooper, M., Hardy, K., Handyside, A., Hunter, S., Monk, M., 1987. HPRT-deficient (Lesch-
840 Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature 326,
841 292–295. doi:10.1038/326292a0
842 Hori, K., Sen, A., Artavanis-Tsakonas, S., 2013. Notch signaling at a glance. Journal of Cell
843 Science 126, 2135–2140. doi:10.1242/jcs.127308
844 Huang, S., Guo, Y.-P., May, G., Enver, T., 2007. Bifurcation dynamics in lineage-
845 commitment in bipotent progenitor cells. Dev Biol 305, 695–713.
846 doi:10.1016/j.ydbio.2007.02.036
847 Jacob, F., 1977. Evolution and tinkering. Science 196, 1161–1166.
848 doi:10.1126/science.860134
849 Kang, M., Garg, V., Hadjantonakis, A.-K., 2017. Lineage Establishment and Progression
850 within the Inner Cell Mass of the Mouse Blastocyst Requires FGFR1 and FGFR2. Dev
851 Cell. doi:10.1016/j.devcel.2017.05.003
852 Kang, M., Piliszek, A., Artus, J., Hadjantonakis, A.K., 2012. FGF4 is required for lineage
853 restriction and salt-and-pepper distribution of primitive endoderm factors but not their
854 initial expression in the mouse. Development 140, 267–279. doi:10.1242/dev.084996
855 Koseska, A., Bastiaens, P.I., 2017. Cell signaling as a cognitive process. EMBO J 36, 568–
856 582. doi:10.15252/embj.201695383
857 Krawchuk, D., Honma-Yamanaka, N., Anani, S., Yamanaka, Y., 2013. FGF4 is a limiting
858 factor controlling the proportions of primitive endoderm and epiblast in the ICM of the
859 mouse blastocyst. Dev Biol 384, 65–71. doi:10.1016/j.ydbio.2013.09.023
860 Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., Smith, A., 2007.
861 FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent
862 embryonic stem cells from self-renewal to lineage commitment. Development 134, 2895–
863 2902. doi:10.1242/dev.02880
864 Lukinavičius, G., Blaukopf, C., Pershagen, E., Schena, A., Reymond, L., Derivery, E.,
865 González-Gaitán, M., D'Este, E., Hell, S.W., Gerlich, D.W., Johnsson, K., 2015. SiR-
866 Hoechst is a far-red DNA stain for live-cell nanoscopy. Nature Communications 6, 8497.
867 doi:10.1038/ncomms9497
868 Martinez Arias, A., Nichols, J., Schröter, C., 2013. A molecular basis for developmental
869 plasticity in early mammalian embryos. Development 140, 3499–3510.
870 doi:10.1242/dev.091959
871 Matsuda, M., Koga, M., Woltjen, K., Nishida, E., Ebisuya, M., 2015. Synthetic lateral
872 inhibition governs cell-type bifurcation with robust ratios. Nature Communications 6,
873 6195. doi:10.1038/ncomms7195
874 Molotkov, A., Mazot, P., Brewer, J.R., Cinalli, R.M., Soriano, P., 2017. Distinct

875 Requirements for FGFR1 and FGFR2 in Primitive Endoderm Development and Exit from
876 Pluripotency. Dev Cell 1–35. doi:10.1016/j.devcel.2017.05.004
877 Morgani, S.M., Saiz, N., Garg, V., Raina, D., Simon, C.S., Kang, M., Arias, A.M., Nichols,
878 J.N., Schröter, C., Hadjantonakis, A.-K., 2018. A Sprouty4 reporter to monitor FGF/ERK
879 signaling activity in ESCs and mice. Dev Biol 441, 104–126.
880 doi:10.1016/j.ydbio.2018.06.017
881 Nagoshi, E., Saini, C., Bauer, C., Laroche, T., Naef, F., Schibler, U., 2004. Circadian gene
882 expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass
883 time to daughter cells. Cell 119, 693–705. doi:10.1016/j.cell.2004.11.015
884 Nagy, A., Gertsenstein, M., Vintersten, K., Behringer, R., 2008. Karyotyping mouse cells.
885 CSH Protoc 2008, pdb.prot4706–pdb.prot4706. doi:10.1101/pdb.prot4706
886 Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., Hadjantonakis, A.-K., 2008. Distinct
887 sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst.
888 Development 135, 3081–3091. doi:10.1242/dev.021519
889 Rossant, J., Chazaud, C., Yamanaka, Y., 2003. Lineage allocation and asymmetries in the
890 early mouse embryo. Philos Trans R Soc Lond, B, Biol Sci 358, 1341–8; discussion 1349.
891 doi:10.1098/rstb.2003.1329
892 Saiz, N., Mora-Bitria, L., Rahman, S., George, H., Herder, J.P., Garcia-Ojalvo, J.,
893 Hadjantonakis, A.-K., 2020. Growth-factor-mediated coupling between lineage size and
894 cell fate choice underlies robustness of mammalian development. eLife 9, 6289.
895 doi:10.7554/eLife.56079
896 Saiz, N., Williams, K.M., Seshan, V.E., Hadjantonakis, A.-K., 2016. Asynchronous fate
897 decisions by single cells collectively ensure consistent lineage composition in the mouse
898 blastocyst. Nature Communications 7, 13463. doi:10.1038/ncomms13463
899 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
900 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J.,
901 Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source
902 platform for biological-image analysis. Nat Methods 9, 676–682.
903 doi:10.1038/nmeth.2019
904 Schröter, C., Rué, P., Mackenzie, J.P., Martinez Arias, A., 2015. FGF/MAPK signaling sets
905 the switching threshold of a bistable circuit controlling cell fate decisions in embryonic
906 stem cells. Development 142, 4205–4216. doi:10.1242/dev.127530
907 Simon, C.S., Hadjantonakis, A.-K., Schröter, C., 2018. Making lineage decisions with
908 biological noise: Lessons from the early mouse embryo. WIREs Dev Biol 61, e319–16.
909 doi:10.1002/wdev.319
910 Simpson, P., 1990. Lateral inhibition and the development of the sensory bristles of the adult
911 peripheral nervous system of Drosophila. Development 109, 509–519.
912 Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O.,
913 Thomas, M., Harrow, J., Cox, T., Jackson, D., Severin, J., Biggs, P., Fu, J., Nefedov, M.,
914 de Jong, P.J., Stewart, A.F., Bradley, A., 2011. A conditional knockout resource for the
915 genome-wide study of mouse gene function. Nature 474, 337–342.
916 doi:10.1038/nature10163
917 Stanoev, A., Schröter, C., Koseska, A., 2019. Robustness and timing of cellular
918 differentiation through population-based symmetry breaking. bioRxiv 14, 578898.
919 doi:10.1101/578898
920 Tarkowski, A.K., 1961. Mouse chimaeras developed from fused eggs. Nature 190, 857–860.
921 doi:10.1038/190857a0
922 Tarkowski, A.K., 1959. Experiments on the development of isolated blastomers of mouse
923 eggs. Nature 184, 1286–1287. doi:10.1038/1841286a0
924 Tinevez, J.-Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E.,
925 Bednarek, S.Y., Shorte, S.L., Eliceiri, K.W., 2017. TrackMate: An open and extensible

platform for single-particle tracking. Methods 115, 80–90. doi:10.1016/j.ymeth.2016.09.016

Viader-Llargués, O., Lupperger, V., Pola-Morell, L., Marr, C., López-Schier, H., 2018. Live cell-lineage tracing and machine learning reveal patterns of organ regeneration. eLife 7, e08201. doi:10.7554/eLife.30823

Waddington, C.H., 1942. Canalization of development and the inheritance of acquired characters. Nature 150, 563–565.

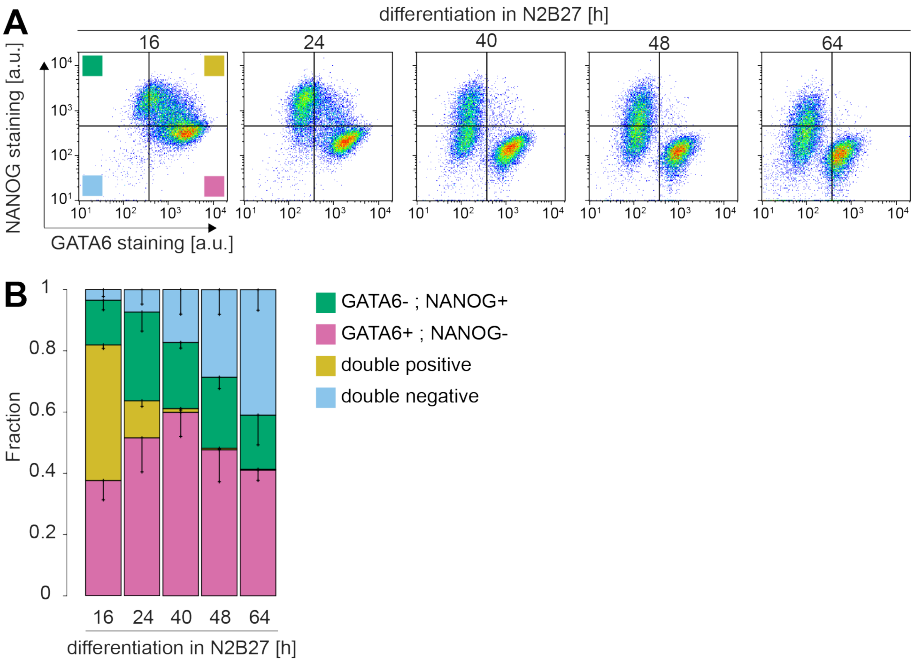
Wamaitha, S.E., del Valle, I., Cho, L.T.Y., Wei, Y., Fogarty, N.M.E., Blakeley, P., Sherwood, R.I., Ji, H., Niakan, K.K., 2015. Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. Genes Dev 29, 1239–1255. doi:10.1101/gad.257071.114

Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., Liu, P., 2008. Chromosomal transposition of PiggyBac in mouse embryonic stem cells. Proceedings of the National Academy of Sciences 105, 9290–9295. doi:10.1073/pnas.0801017105

Yamanaka, Y., Lanner F., Rossant, J., 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development 137, 715 - 724. doi:10.1242/dev.043471

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., Smith, A., 2008. The ground state of embryonic stem cell self-renewal. Nature 453, 519–523. doi:10.1038/nature06968

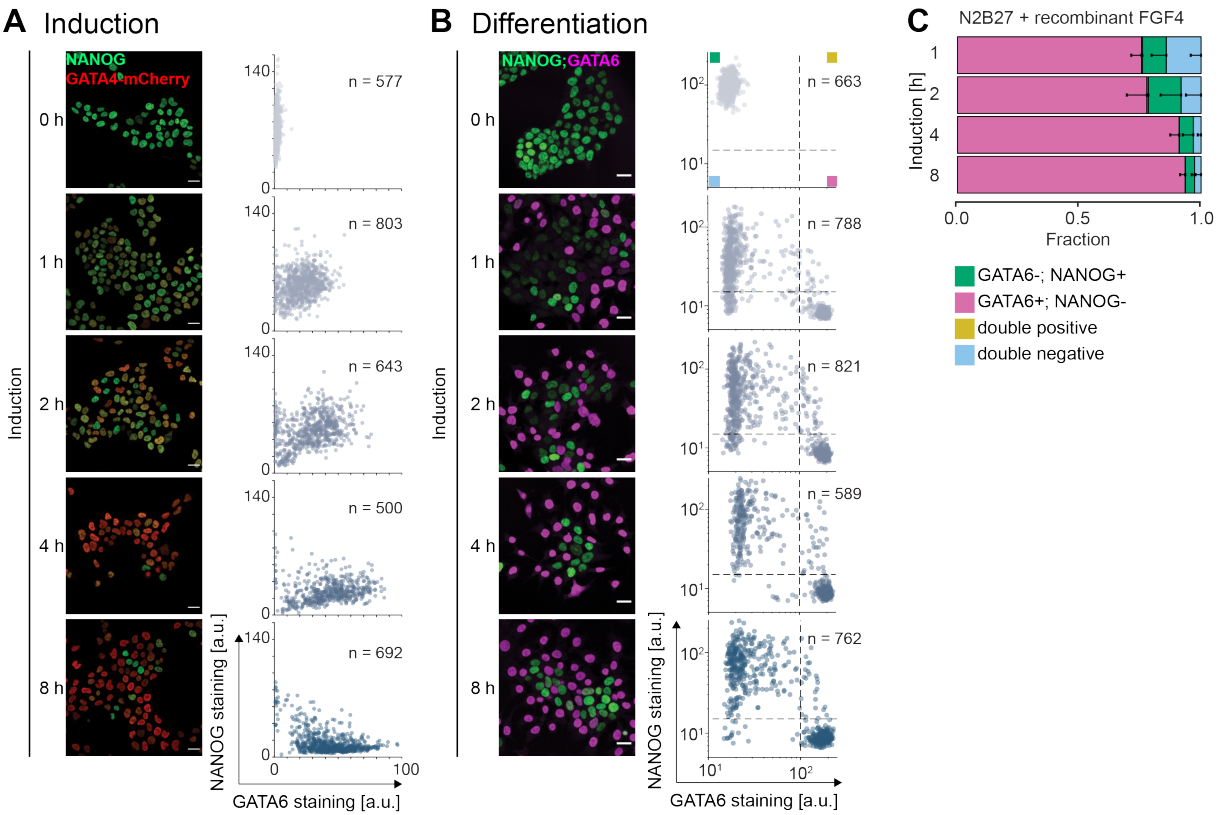
949 **Supplementary Figures**



950

951 ***Figure S1. Time course of cell type proportions following pulsed GATA4-mCherry expression***

952 ***A*** *Flow cytometry profiles of cells stained for NANOG and GATA6 after an 8 h doxycycline pulse to*
953 *induce GATA4-mCherry expression, followed by differentiation in N2B27 medium for the indicated*
954 *durations. Two clusters corresponding to NANOG-high; GATA6-low Epi-like cells and NANOG-low,*
955 *GATA6-high PrE-like cells can already be distinguished after 16 h of differentiation, but are only fully*
956 *separated at 40 h and beyond. Solid lines and color code in leftmost panel indicate gating strategy to*
957 *assign cell types.* ***B*** *Average cell type proportions from N = 3 independent experiments. Fraction of*
958 *GATA6+, NANOG- cells in magenta, GATA6-, NANOG+ cells in green, double positive cells in yellow,*
959 *and double negative cells in blue. Error bars: 95% CI.*



960

961 ***Figure S2. Quantitative immunofluorescence analysis of NANOG and GATA expression***

962 ***A*** *Immunostaining (left) and single-cell quantification (right) of NANOG (green) and GATA4-mCherry*
963 *(red) expression in inducible cells after indicated durations of doxycycline treatment in 2i + LIF*
964 *medium.* ***B*** *Immunostaining (left) and single-cell quantification (right) of NANOG (green) and GATA6*
965 *(magenta) expression in inducible cells after indicated durations of doxycycline treatment, followed by*
966 *40 h differentiation in N2B27 medium. Immunofluorescence micrographs in* ***A*** *and* ***B*** *for 0, 2, and 8 h*
967 *of doxycycline induction are reproduced from Fig. 1 for comparison. Dashed lines in scatter plots in* ***B****:*
968 *Thresholds to assign cell types; upper left quadrant: GATA6-; NANOG+; lower right quadrant*
969 *GATA6+; NANOG-; upper right quadrant double positive; lower left quadrant double negative. Scale*
970 *bars, 20 µm.* ***C*** *Proportions of cell types upon indicated durations of doxycycline induction followed by*
971 *40 h of differentiation in N2B27 medium supplemented with 10 ng/ml FGF4. Cell identities were*
972 *determined by immunostaining and quantitative immunofluorescence. Fraction of GATA6+, NANOG-*
973 *cells in magenta, GATA6-, NANOG+ cells in green, double positive cells in yellow, and double negative*
974 *cells in blue. N = 4, error bars: 95% CI.*

975

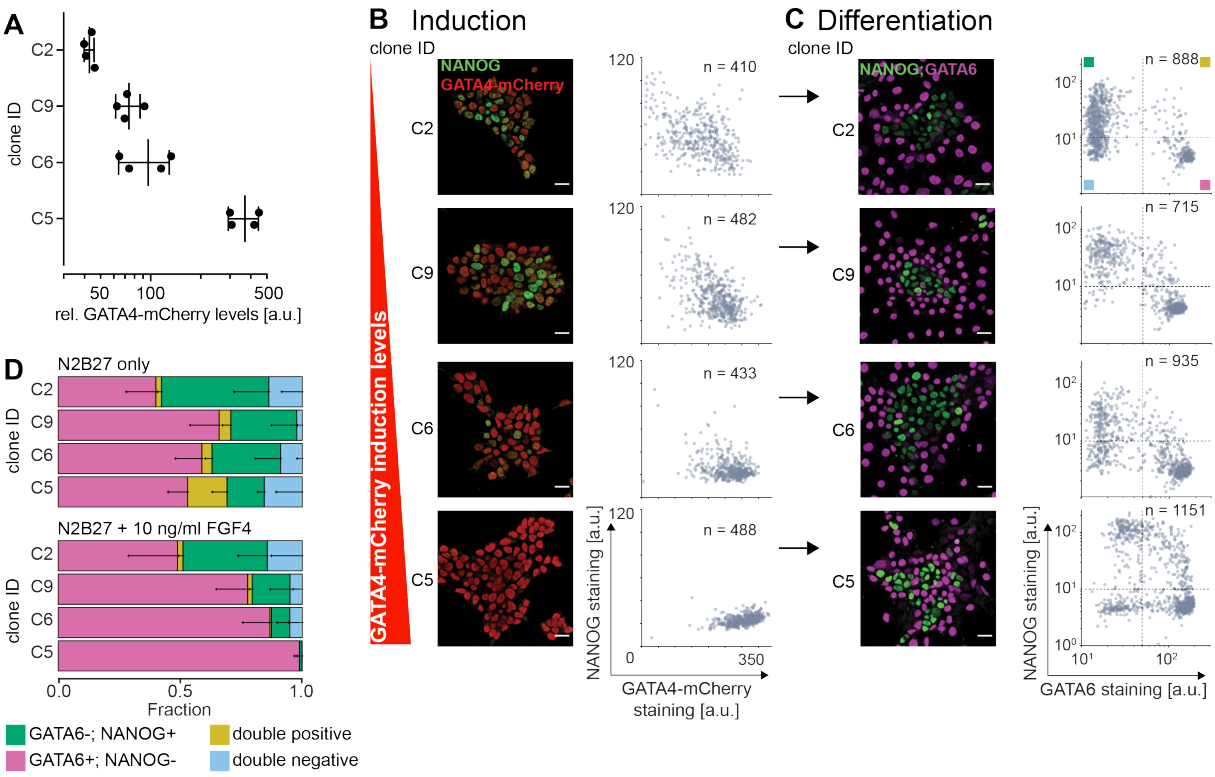
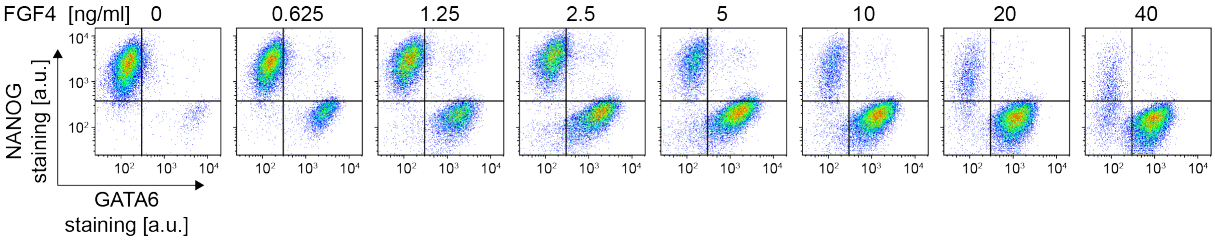


Figure S3. Robust cell type proportioning in independent clonal cell lines

A *GATA4-mCherry induction levels in four independent clonal inducible cell lines after 8 h of doxycycline treatment in 2i + LIF medium, measured by flow cytometry. Clones are ordered by GATA4-mCherry expression strength, fluorescence values were normalized to non-induced control cells. Plot shows individual data points and mean ± SD from N = 4 independent experiments.* ***B*** *Immunostaining (left) and single-cell quantification(right) of NANOG (green) and GATA4-mCherry (red) expression in the same clonal lines analyzed in* ***A*** *after 8 h of doxycycline stimulation in 2i + LIF medium.* ***C*** *Immunostaining (left) and single cell quantification (right) of NANOG (green) and GATA6 (magenta) expression in cells from independent clonal lines treated with doxycycline for indicated periods of time and differentiated in N2B27 for 40 h. Dashed lines: Thresholds to determine cell types; upper left quadrant: GATA6-; NANOG+; lower right quadrant GATA6+; NANOG-; upper right quadrant double positive; lower left quadrant double negative. Clones are ordered by GATA4-mCherry induction strength as in* ***A****. Scale bars in* ***A****,* ***C****, 20 µm.* ***D*** *Average cell type proportions in clonal lines differentiated as in* ***C*** *in N2B27 only (top) or in N2B27 supplemented with 10 ng/ml FGF4 (bottom).Fraction of GATA6+, NANOG- cells in magenta, GATA6-, NANOG+ cells in green, double positive cells in yellow, and double negative cells in blue. N = 4, error bars: 95% CI.*

993



994

995 ***Figure S4. FGF4 dose regulates cell type proportions in Fgf4 mutant cells***

996 *Flow cytometry profiles of Fgf4 mutant cells stained for NANOG and GATA6 after 8 h of doxycycline*

997 *induction, followed by 40 h of differentiation in N2B27 medium supplemented with 1 µg/ml heparin and*

998 *the indicated concentrations of recombinant FGF4. Lines indicate gates to assign cell types.*

999

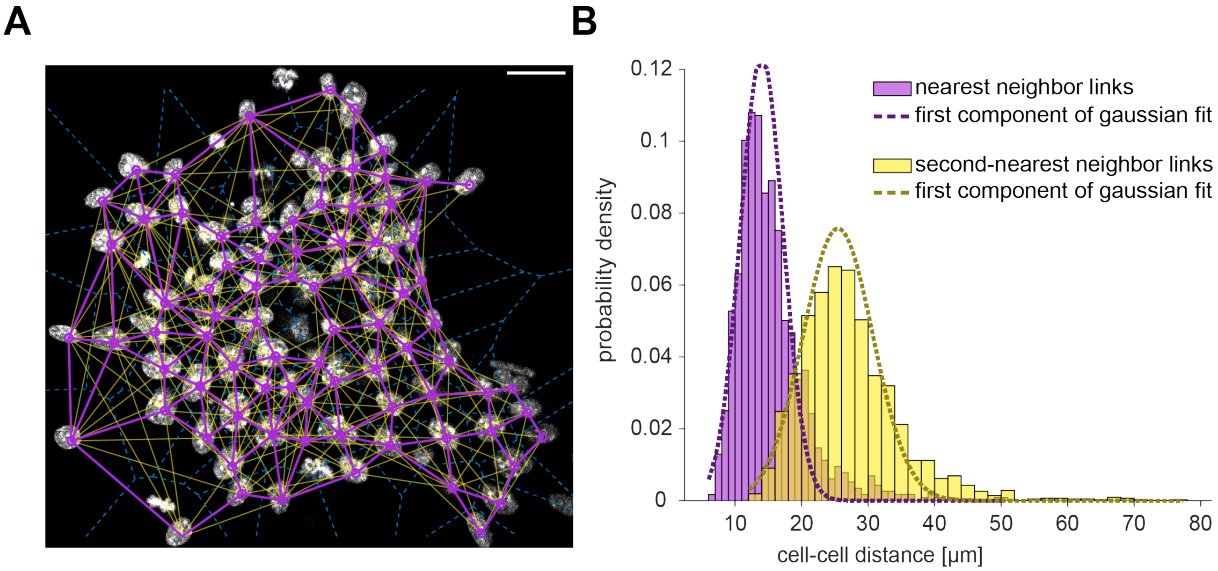


Figure S5. Estimation of cell-cell distances by Delauney triangulation

A *Nuclear staining (white) of a colony 16 h after initiation of differentiation, overlaid with centers of mass of individual nuclei (purple circles), Voronoi edges (dashed blue lines), and links to nearest (purple) and second nearest neighbors (yellow, see Methods). Scale bar, 20 µm.* ***B*** *Histogram of distance distributions between nearest neighbors (purple) and second nearest neighbors (yellow), determined as in* ***A****. To separate true nearest and second-nearest neighbor distances from spurious longer-distance links, we applied two-component Gaussian mixture model fits to each of the distributions. Dashed lines indicate the first components of each fit. Distances determined by these first components were 14.0 ± 3.2 (mean ± SD) for nearest neighbors, and 25.5 ± 5.3 µm for second nearest neighbors. n = 8 independent fields of view.*

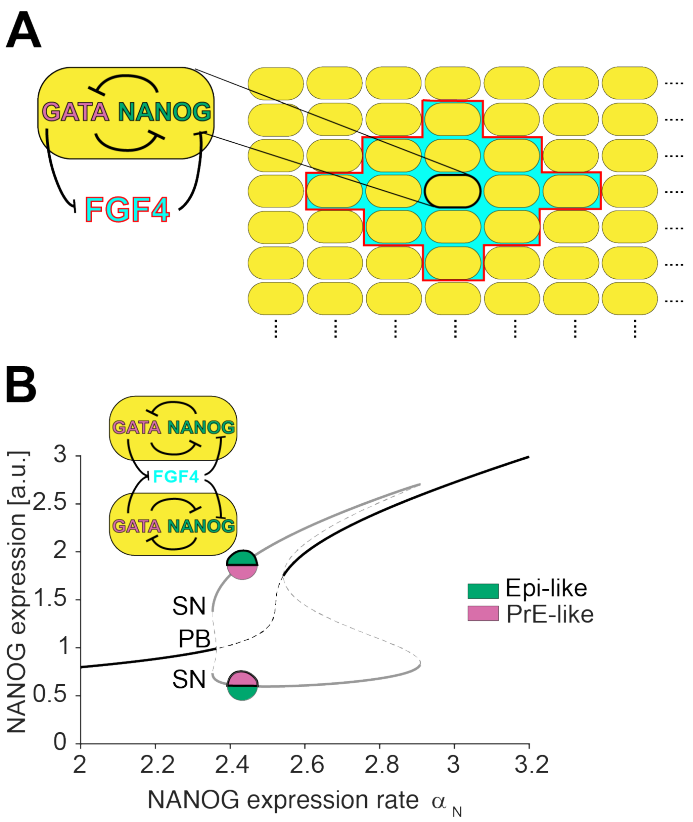


Figure S6. FGF4 signaling range and bifurcation analysis of the model

A *Scheme illustrating the spatial range of cell-cell communication implemented in the model. FGF4 (cyan with red box) secreted by an individual cell can only reach nearest and second nearest neighbors, and an individual cell will only respond to FGF4 secreted by nearest and second nearest neighbors.* ***B*** *Bifurcation diagram showing the principle of emergence of a symmetry-broken inhomogeneous steady state (IHSS) for a two-cell coupled system (inset) in dependence of the production rate constant of NANOG (αN), using NANOG expression as a representative variable. The IHSS (grey lines) is generated via a pitchfork (PB) and stabilized via saddle-node (SN) bifurcations. The two branches represent the heterogeneous attractors: (GATA6-/NANOG+) for cell 1 and (GATA6+/NANOG-) for cell 2, or (GATA6+/NANOG-) for cell 1 and (GATA6-/NANOG+) for cell 2 (representative circles). Solid/dashed lines: stable/unstable solutions. Black: homogeneous steady state; grey: IHSS solution.*

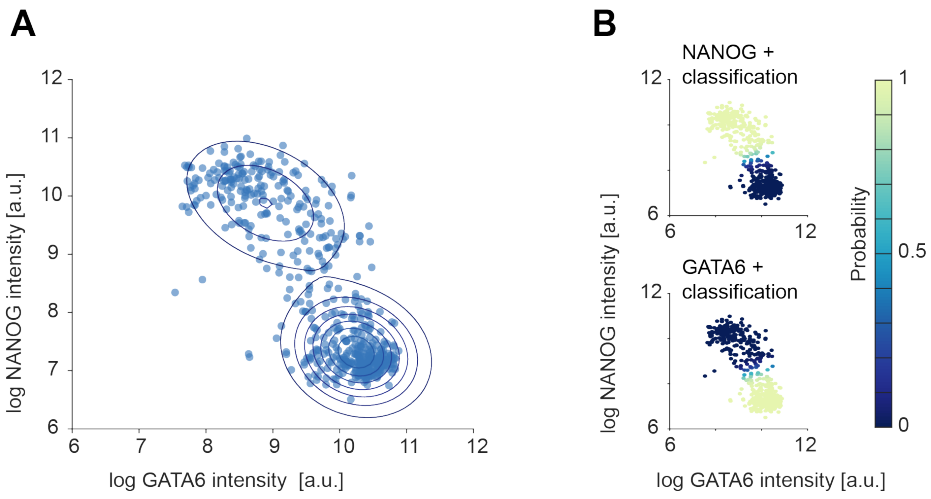
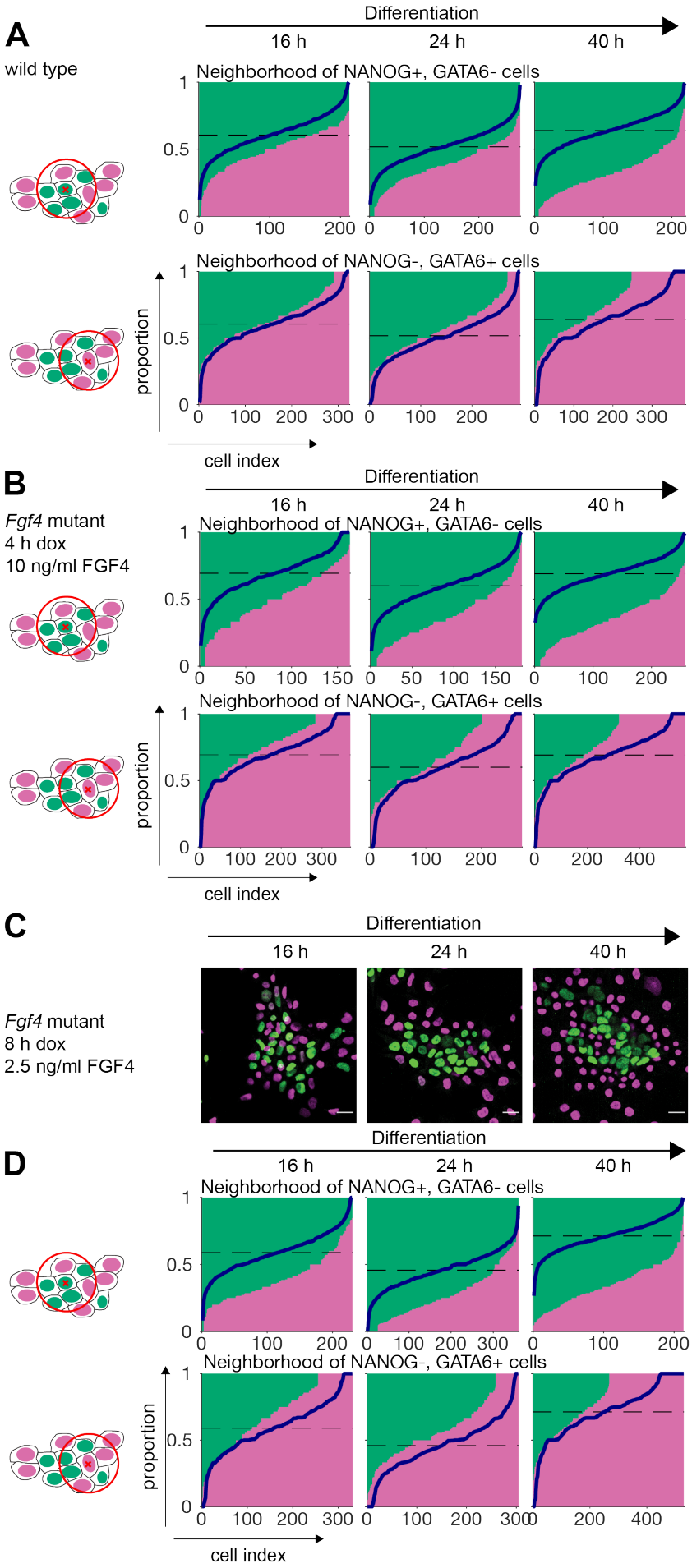


Figure S7. Cell type assignment using a two-component Gaussian mixture model

A *Blue dots indicate GATA6 and NANOG expression determined by single-cell analysis of immunostaining in cells 16 h after the initiation of differentiation (n=560 cells). Contour lines indicate levels of equal probability density determined from a two-component gaussian mixture model fit.* ***B*** *Color-coded posterior probabilities of association with one of the two component distributions of the model fit.*


A
wild type
Differentiation
16 h
24 h
40 h
Neighborhood of NANOG+, GATA6- cells
Neighborhood of NANOG-, GATA6+ cells
proportion
cell index
B
Fgf4 mutant
4 h dox
10 ng/ml FGF4
C
Fgf4 mutant
8 h dox
2.5 ng/ml FGF4
D

1033

1034 ***Figure S8. Cell type composition of local cell neighborhoods***

1035 ***A*** *Cell type proportions in a local neighborhood of 31.7 µm diameter in wild type cells stimulated for*
1036 *8 h with doxycycline and differentiated for the indicated times. Each column in a panel corresponds to*
1037 *one cell, and the size of its associated green and magenta bars correspond to the fraction of NANOG+,*
1038 *GATA6- (green) and NANOG-, GATA6+ (magenta) cells in its local neighborhood. Blue lines indicate*
1039 *the corresponding random distribution, calculated by assigning the global distribution of cell types of*
1040 *the time point randomly to experimentally determined cell positions. Clustering of cell types results in*
1041 *the deviation of the true proportions from this random distribution. The local neighborhood of*
1042 *NANOG+, GATA6- cells (top row) indicates clustering of this cell type already at 16 h, which becomes*
1043 *more pronounced over time. Strong clustering of NANOG-, GATA6+ cells (lower row) can only be*
1044 *observed at 40 h.* ***B*** *Same analysis as in* ***A****, but for Fgf4 mutant cells induced for 4 h and differentiated*
1045 *in N2B27 supplemented with 10 ng/ml FGF4 for the indicated periods of time. Deviation from random*
1046 *arrangement (blue lines) indicates clustering of NANOG+, GATA6 cells at 16 h, which becomes more*
1047 *pronounced over time. Some clustering of NANOG-, GATA6+ cells (lower row) can already be detected*
1048 *at 16 h, and becomes more pronounced until 40 h.* ***C*** *Immunostaining for NANOG (green) and GATA6*
1049 *(magenta) in Fgf4 mutant cells induced with doxycycline for 8 h and differentiated in N2B27*
1050 *supplemented with 2.5 ng/ml FGF4 for the indicated periods of time.* ***D*** *Same analysis as in* ***A, B****, for*
1051 *Fgf4 mutant cells differentiated as described in* ***C****. Deviations from random arrangement (blue lines)*
1052 *are similar to those seen in* ***B****, indicating similar clustering dynamics.*

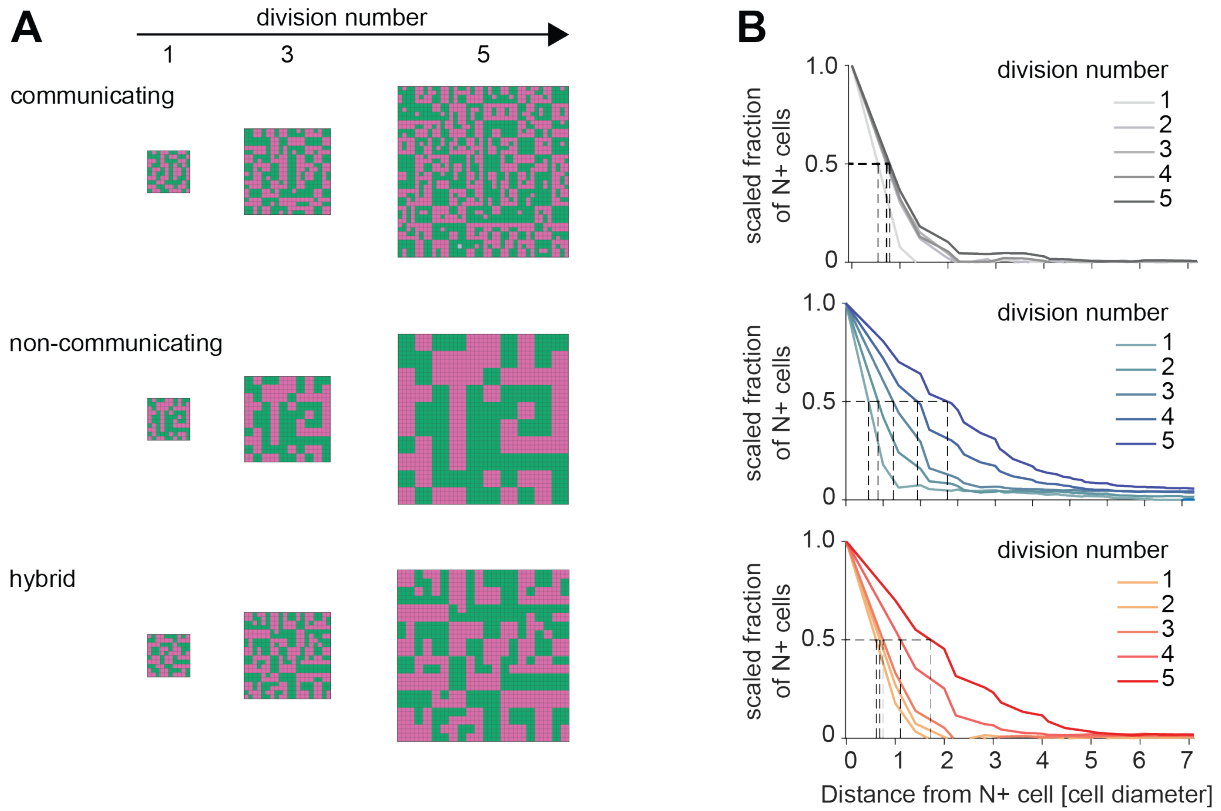
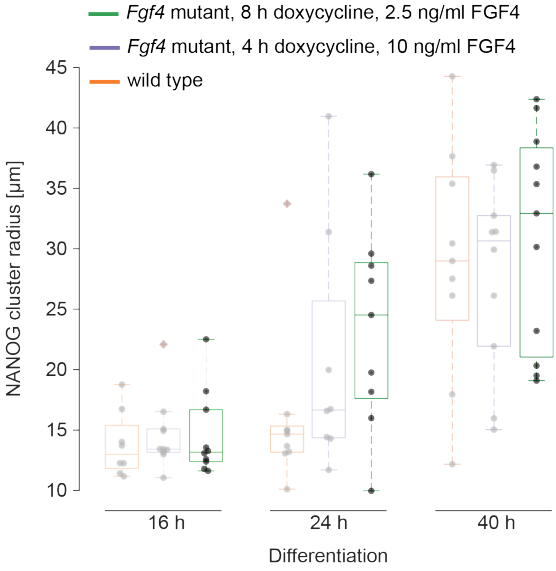


Figure S9. Spatial patterns of cell types in model simulations

A *Spatial configurations of cell types in single numerical realizations of the model with cell-cell communication (top row), without cell-cell communication (middle row) and of a hybrid model (bottom row) where communication is switched off after the third division. First column: Cell type arrangements at the end of the first cell cycle (grid size 10 x 10), second column - end of the third cell cycle (grid size 20 x 20), third column - end of the fifth cell cycle (grid size 40 x 40).* ***B*** *Corresponding estimation of NANOG cluster radius for single numerical realizations of the model variants shown in* ***A****. Graphs depict the scaled fraction of N+ cells within a specific radius around seed cells. Dashed lines indicate determination of cluster radius, the distance around NANOG+ cells at which the scaled fraction is equal to 0.5. Cluster radii were obtained after each cell cycle of the lineage tree simulation.*

1065



1066

1067

1068 ***Figure S10. Cluster radius dynamics in Fgf4 mutant cells are independent from stimulation and***

1069 ***differentiation regime***

1070 *Dark: Radius of NANOG clusters in Fgf4 mutant cells induced with doxycycline for 8 h, followed by*

1071 *differentiation in N2B27 medium supplemented with 2.5 ng/ml FGF4 for the indicated times. Dots*

1072 *indicate values from individual fields of view (n ≥ 9), box plots show median, interquartile ranges and*

1073 *outliers (red cross). Shaded data points are reproduced from Figure 4C for comparison, and show the*

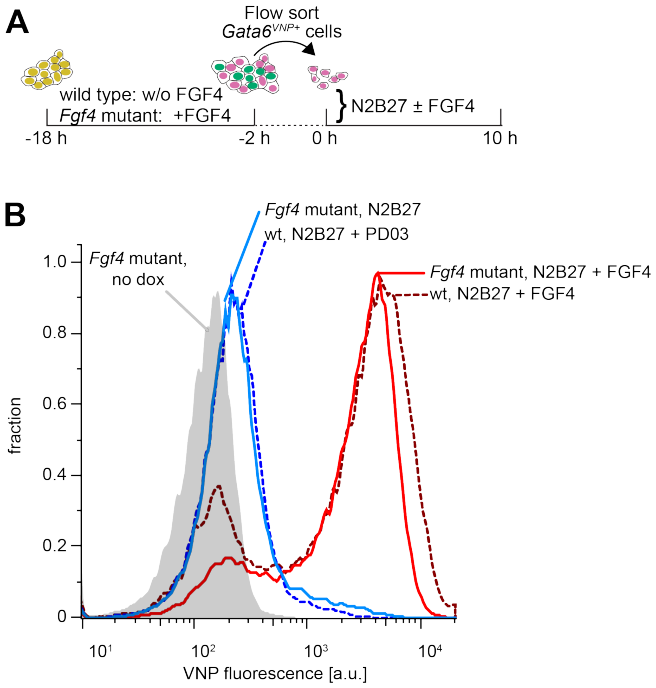
1074 *data for Fgf4 wild type cells differentiated in N2B27 only (orange) and for Fgf4 mutant cells*

1075 *differentiated in N2B27 supplemented with 10 ng/ml FGF4 following 4 h of induction with doxycycline.*

1076

1077

1078



1079

1080 ***Figure S11. Gata6^{VNP} reporter expression in sorted Fgf4 mutant cells***

1081 ***A*** *Schematic representation of the cell-sorting experimental protocol for Fgf4 mutant cells. Following*

1082 *a doxycycline pulse, Fgf4 mutant cells were initially differentiated in medium supplemented with FGF4,*

1083 *while wild type cells were differentiated without recombinant FGF4. After flow sorting of VNP positive*

1084 *cells at 16 h, wild type and Fgf4 mutant cells were cultured in N2B27 medium with or without FGF4,*

1085 *followed by analysis of VNP expression.* ***B*** *Flow cytometry histograms of VNP expression in wild type*

1086 *(dashed lines) and Fgf4 mutant (solid lines) inducible Gata6^{VNP} cells, 10 h after sorting of VNP-positive*

1087 *cells and culture in the indicated media.*

1088

1089 **Supplementary Movie legends**

1090 **Movie S1. Dynamics of VNP expression in unperturbed colonies**

1091 Time-lapse imaging of a colony of Gata6^{VNP} reporter cells starting 16 h after the end of a
1092 doxycycline pulse. Medium has been switched to N2B27 at the beginning of the recording.

1093

1094 **Movie S2 – S4. Dynamics of VNP expression in sorted cell populations**

1095 Time-lapse imaging of Gata6^{VNP} reporter cells flow sorted for VNP expression 16 h after the
1096 end of a doxycycline pulse and cultured in defined N2B27 medium alone (Movie S2), or
1097 cultured in N2B27 supplemented with 10 ng/ml FGF4 (Movie S3), or cultured in N2B27
1098 supplemented with 1 μM of the MEK inhibitor PD03 (Movie S4).