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

Dhruv Raina[1], Angel Stanoev[1], Azra Bahadori[1,2], Michelle Protzek[1], Aneta Koseska[1], 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

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

Abstract

During embryonic development and tissue homeostasis, reproducible proportions of differentiated cell types need to be specified from homogeneous precursor cell populations. How this is achieved despite uncertainty in initial conditions in the precursor cells, and how proportions are re-established upon perturbations in the developing tissue is not known. Here we report the differentiation of robust proportions of epiblast- and primitive endoderm-like cells from a wide range of experimentally controlled initial conditions in mouse embryonic stem cells. We demonstrate both experimentally and theoretically that recursive cell-cell communication via FGF4 establishes a population-based mechanism that generates and maintains robust proportions of differentiated cell types. Furthermore, we show that cell-cell communication re-establishes heterogeneous cell identities following the isolation of one cell type. The generation and maintenance of robust cell fate proportions is a new function for FGF signaling that may extend to other cell fate decisions.

Introduction

The differentiation of cell types with discrete identities from an equipotent precursor population is the basis of embryonic development and tissue homeostasis in the adult. Canalization of development, the tendency of tissues to produce a standard end result despite developmental noise and to compensate for perturbations (Waddington, 1942), indicates there are mechanisms to ensure that specialized cell types are differentiated in reproducible proportions. Current frameworks to conceptualize cell differentiation emphasize the role of multistable gene regulatory networks in single cells, which establish stable gene expression states that correspond to precursor and differentiated states (Enver et al., 2009; Huang et al., 2007). Extracellular signals facilitate the switching from the precursor to a differentiated state by modulating the attractor landscape of the single cell circuits (Huang et al., 2007; Schröter et al., 2015). In this single-cell view of cell differentiation, the specific differentiated state adopted by an individual cell strongly depends on its initial conditions while in the precursor state. Accordingly, reliable embryonic development and proportioning of differentiated fates would require appropriately constraining these initial conditions (Briscoe, 2019).

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 compensate for splitting, fusing, or the addition of embryonic stem cells (ESCs), and regulate the proportions of the three lineages such that the blastocyst is capable of postimplantation development (Bedzhov et al., 2014; Martinez Arias et al., 2013). The core gene regulatory networks and signals that first segregate the TE from the inner cell mass (ICM), and then specify embryonic Epi and extraembryonic PrE identities in undifferentiated ICM cells are well known. ICM cells initially co-express transcriptional regulators for both the Epi and the PrE lineage, such as GATA6 and NANOG, before mutually exclusive expression patterns are established (Chazaud et al., 2006; Plusa et al.,

2008; Simon et al., 2018). The dynamics of gene expression and fate allocation, together with extensive analysis of genetic mutants and pharmacologic interventions in the mouse embryo have led to a model of mutually repressive interactions between Epi- and PrE-specific transcriptional regulators in individual cells (Bessonnard et al., 2014; Chickarmane and Peterson, 2008). Under the influence of fibroblast growth factor (FGF)/ extracellular regulated kinase (ERK) signaling, which inhibits Epi-specific genes such as *Nanog* and promotes the expression of PrE-specific genes, these mutually repressive interactions are thought to progressively establish the identity of single cells (Bessonnard et al., 2014; De Caluwé et al., 2019; De Mot et al., 2016). The mutual repression model and its regulation by FGF/ERK signaling has been substantiated by experiments in embryonic stem cells (ESCs), in which PrE-like differentiation from an ICM-like state generated through induced expression of GATA6 or GATA4 requires both a threshold level of induced GATA proteins as well as ERK activity (Schröter et al., 2015). A central prediction from these models for the differentiation of Epi- and PrE-identities is that the ratio of the two cell types should critically depend on initial conditions.

Here we report the differentiation of robust proportions of Epi- and PrE-like cell types from a wide range of initial conditions of an induced ICM-like state in ESCs. Through mutant analysis, we identify recursive cell-cell communication via local FGF4 signaling as a minimal molecular mechanism for this emergent property of robust fate proportioning. The experimental observations can be described by a generic dynamical mechanism, an inhomogeneous steady state (IHSS), in which robust ratios of differentiated cell fates are generated and maintained on a population level (Stanoev et al., accompanying manuscript). 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 cell identities from isolated PrE-like cells.

Results

Differentiation of robust proportions of Epi- and PrE-like cell types from a wide range of expression levels of lineage-specific transcriptional regulators.

To study mechanisms that control proportions of cell types in populations, we used an ESC system in which cells with Epi- or PrE-like identity differentiate from an ICM-like cell state that is generated via transient doxycycline-induced expression of GATA factors. We had previously shown that PrE-like differentiation from this ICM-like state occurs above a threshold level of GATA expression and ERK signaling (Schröter et al., 2015) (Fig. 1A). To further test how the ratio of PrE-like and Epi-like cells depends on the initial transcription factor expression levels in the ICM-like state, we established new inducible GATA4-mCherry transgenic ESC lines with independent integrations of the inducible transgene. We induced the transgene while cells were cultured in 2i + LIF medium (Ying et al., 2008), which prevents any differentiation through the MEK inhibitor PD0325901 (PD03), and then initiated differentiation by concomitantly removing doxycycline and switching to chemically defined N2B27 medium lacking PD03. Varying the doxycycline induction time in a selected clonal line allowed us to titrate expression levels of the inducible GATA4-mCherry protein by more than 5-fold (Supplementary Fig. 1A), and thereby to start differentiation from a wide range of NANOG-to-GATA4-mCherry expression ratios (Fig. 1B). After 40 h of differentiation, we detected both NANOG-expressing Epi-like cells and PrE-like cells expressing the endogenous GATA6 protein for all GATA4-mCherry induction levels tested (Fig. 1C). Quantitative immunofluorescence (QIF) revealed that the proportions of both PrE-like (GATA6+; NANOG) and Epi-like (GATA6-;NANOG+) cell populations fell within a narrow range for different induction times (Fig. 1C). The Epi-like population comprised between 44.3 ± 13.2% (95% confidence interval (CI), N=4 independent experiments) and 37.5 ± 7.8%

of cells for 1 h and 8 h of induction, respectively, while the proportion of PrE-like cells increased slightly from 42.1 ± 11.6% for 1 h induction to 51.6% ± 15.0% for 2 h, and then plateaued between 58.2 ± 12.4% and 56.0 ± 9.3% for 4 h and 8 h of induction, respectively (Fig. 1C). Thus, a wide range of inducible GATA4-mCherry expression levels lead to similar distributions of differentiated fates. The robust fate proportioning did not reflect pre-specified fates or a limited differentiation potential of the cells, as supplementation of the medium with 10 ng/ml FGF4 during the differentiation phase strongly increased the proportion of PrE-like cells at the expense of Epi-like cells compared to differentiation in N2B27 alone (Supplementary Fig. S1B). We confirmed the observation of robust fate proportioning in four clonal cell lines with independent integrations of the GATA4-mCherry transgene, which displayed a more than 8-fold difference in GATA4-mCherry expression levels following 8 h of doxycycline induction (Supplementary Fig. S2A, B). Despite these differences in initial conditions, similar proportions of both GATA6-positive and NANOG-positive cells appeared upon doxycycline removal and differentiation in N2B27 in all clones (Supplementary Fig. S2C). The fraction of PrE-like cells, for example, ranged from 40.0 ± 12.1% to 53.0 ± 7.9% for the lowest and highest expressing clone, respectively (Supplementary Fig. S2D, left). The two lines with intermediate GATA4-mCherry induction levels gave a slightly higher percentage of PrE-like cells of up to 65.9 ± 11.9%. Supplementation of the medium with 10 ng/ml FGF4 during the differentiation phase strongly increased the proportion of PrE-like cells also in this case, to a maximum of 98.8 ± 2.0% PrE-like cells in the clonal line with the highest GATA4-mCherry expression levels (Supplementary Fig. S2D, right). These observations indicate that all cells have PrE-like differentiation potential upon doxycycline-induced GATA4-mCherry expression, and suggests that the robust proportioning of cell fates is established at the population level through cell-cell communication.

Robust specification of cell fate proportions arises from recursive cell-cell communication via FGF4.

To identify the molecular mechanism that mediates cell fate proportioning, we focused on *Fgf4*, as it is the main paracrine activator for ERK in ESCs, and FGF/ERK signaling is required for PrE differentiation both in ESCs and in the embryo (Kang et al., 2012; Krawchuk et al., 2013; Kunath et al., 2007; Schröter et al., 2015). In an *Fgf4* mutant GATA4-mCherry inducible cell line, the differentiation of GATA6+ PrE-like cells was almost completely abrogated (Fig. 2A, B), in contrast to previous studies in ESCs (Kang et al., 2012; Wamaitha et al., 2015), but recapitulating the *Fgf4* mutant phenotype in the embryo (Feldman et al., 1995; Kang et al., 2012; Krawchuk et al., 2013). The differentiation of PrE-like cells could efficiently be rescued by supplementing the medium with recombinant FGF4 during the differentiation phase (Fig. 2A, B). The number of PrE-like cells for a given GATA4-mCherry induction level smoothly increased with FGF4 concentration (Fig. 2B), suggesting that the robustness of cell fate proportions in wild type cells could be mediated by regulated FGF4 signaling. To directly test this hypothesis, we compared differentiation outcomes in wild type cells where FGF4 signaling is triggered by endogenous ligands, with that of mutant cells treated with a constant exogenous dose of FGF4, upon titrating GATA4-mCherry levels by varying induction time. In contrast to wild type cells, cell fate proportions in *Fgf4* mutant cells continuously changed with induction times, thus reflecting the changes in initial conditions (Fig. 2C). The proportion of GATA6+;NANOG- PrE-like cells for example increased from 11.7 ± 3.8% following 1 h of doxycycline induction to 33.0 ± 6.9% for 2 h, 59.5 ± 7.6% for 4 h and 76.8 ± 6.5% for 8 h of induction in *Fgf4* mutant cells differentiating in the presence of 10 ng/ml FGF4, whereas in wild type cells, this fraction slightly increased from 33.0 ± 10.4% for 1 h induction to 47.9 ± 3.5% for 2 h and then plateaued, reaching a maximum of 51.9 ± 2.0% for 8 h induction (Fig. 2C). Together, these results indicate that cell-cell communication via endogenous FGF4 signaling is the molecular basis of robust and

reliable cell fate proportioning, irrespective of the variance in initial gene expression levels of the fate specifiers.

The independence of fate proportions from initial conditions in a communicating cell population is a central property of an IHSS, a dynamical solution of a network of mutual repression switches (Stanoev et al., accompanying manuscript). To test whether the IHSS solution represents a dynamical basis for our experimental observations, we compared a single-cell model where FGF4 acts as a unidirectional input to each cell operating an intracellular *Nanog-Gata* toggle switch (Schröter et al., 2015), with a population-based model in which cells were recursively communicating via FGF4 (Fig. 2D). Motivated by our observation that differentiation of PrE-like cells is almost completely abrogated in the *Fgf4* mutants (Fig. 2A), as well as previous work on regulation of *Nanog* by FGF signaling (Hamilton and Brickman, 2014; Schröter et al., 2015), we assumed that both secreted as well as recombinant FGF4 ligands quantitatively reduce *Nanog* expression. To establish coupling between the cells in a population, FGF4 production must be regulated by the intracellular circuit. While different coupling topologies can establish an IHSS (Stanoev et al., accompanying manuscript), for simplicity we focus here on inhibition of *Fgf4* transcription by GATA factors, and provide further evidence for this topology below. In the mouse embryo, FGF ligands have a restricted signaling range (Shimokawa et al., 2011). We therefore implemented short-range communication via FGF4 in our model (methods).

For both the single cell as well as the coupled model, we quantified the proportion of (NANOG+, GATA-) and (NANOG-, GATA+) cells when starting from a broad range of initial NANOG/GATA expressions. When FGF is considered as a unidirectional input to the cells, the numerical simulations demonstrated that the fraction of (NANOG-, GATA+) cells was highly biased by the distribution of initial gene expressions (Fig. 2D, top and middle rows). In contrast, for the same initial expression distributions, robust proportions between the

two cell types were obtained from a model of a cell population (N=10000), coupled by short-range communication on a 100×100 grid. (Fig. 2D, bottom row). Bifurcation analysis of a minimal system of N=2 coupled cells with same parameters demonstrated that this system indeed exhibits an IHSS solution (Supplementary Fig. S3). This parallel between the theoretical and experimental results thus substantiates the conceptual differences between a single-cell and a population-based formulation of differentiation in terms of robust Epi/PrE-like cell fate proportioning, indicating that recursive communication via FGF signaling uniquely underlies this dynamical feature.

FGF4 expression is repressed by GATA factors and acts locally

To further characterize the FGF4-mediated communication mechanism underlying the robust cell type proportioning in populations, we first investigated how *Fgf4* transcription is regulated. In situ mRNA staining (Choi et al., 2018) for Fgf4, Nanog and Gata6 transcripts along the differentiation time-course showed that Nanog and Fgf4 mRNA were co-expressed in most cells before induction in 2i medium (Fig. 3A), but became markedly downregulated after 8 h of GATA4-mCherry induction in 2i, especially in cells that expressed the GATA4-mCherry transgene most strongly (Fig. 3A). Fgf4 transcripts could be detected again following 40 h of differentiation in N2B27, but now their expression was mutually exclusive with Gata6 mRNA (Fig. 3A). Cells transferred from 2i to N2B27 for 40 h without doxycycline-induction stained positive for Fgf4 mRNA, but Nanog mRNA could barely be detected (Fig. 3A). These *Fgf4* expression dynamics suggest that GATA factors negatively regulate *Fgf4* transcription, although they do not rule out a positive influence of NANOG on *Fgf4* expression that has been identified in the embryo (Frankenberg et al., 2011; Messerschmidt and Kemler, 2010).

If cell fate proportioning in wild type cells was based on regulation of *Fgf4* expression by GATA factors, one would expect a gradual reduction in FGF4 signaling activity with

increasing GATA induction levels. To directly test this prediction, we integrated a *Sprouty4* (*Spry4*) transcriptional reporter construct that we have previously established as a quantitative readout for long-term FGF4 signaling in ESCs (Morgani et al., 2018) in the inducible cell lines. We triggered different GATA4-mCherry expression levels by varying doxycycline induction time, and measured mean reporter expression levels after 24 h of differentiation in N2B27 medium. At this timepoint, reporter expression is expected to reflect the integrated FGF signal of the period during which cells transition from the ICM-like state to differentiated identities. Consistent with our expectation, we found that higher GATA4-mCherry expression levels induced by longer doxycycline induction times resulted in reduced mean fluorescence levels of the *Spry4* reporter at the end of the experiment (Fig. 3B). For the longest induction time, reporter expression levels were reduced to 57.8 ± 5.1% compared to the uninduced control (Fig. 3 B bottom, N = 4 independent experiments). These data support the idea that FGF4 signaling levels in the cell population are inversely related with GATA induction levels, as expected for a molecular mechanism that buffers initial conditions to lead to robust cell fate proportioning.

To test coupling range mediated by FGF4, we next sought to identify the spatial extent of FGF4 signaling in our system. We first tested the role of global communication through FGF4 ligands by comparing differentiation outcomes at different medium-to-cell ratios during the differentiation step (Fig. 3C). We reasoned that if FGF4 ligands equilibrated in the medium, larger volumes would effectively reduce FGF4 concentration and thereby reduce the proportion of FGF-dependent PrE-like cells. In contrast to this expectation, the proportions of GATA6+; NANOG- PrE-like cells detected by flow cytometry 40 h after an 8 h doxycycline pulse slightly increased with media volume, while the proportions of the GATA6-; NANOG+ Epi-like cells slightly decreased (Fig. 3C). These results indicate that dilution of FGF4 ligands in the medium does not strongly affect cell fate proportioning, and suggested that local effects might be more relevant. To test this, we disrupted cell-cell contacts by trypsinizing and re-

seeding cells at different densities immediately after doxycycline induction (Fig. 3D). This treatment strongly reduced the proportion of PrE-like cells compared to the non-trypsinized control (Fig. 3D). Furthermore, the proportion of PrE-like cells systematically increased with cell density (Fig. 3D). Together, these data suggest that cell-cell communication via FGF4 occurs locally and is positively influenced by cell-cell contacts, in line with the assumptions of our model.

To directly measure the spatial range of FGF4 signaling in ESC colonies, we seeded isolated *Fgf4* wild type cells that were labelled by a constitutively expressed dsRed marker onto a layer of *Fgf4* mutant cells bearing a *Spry4*$^{H2B\text{-}Venus}$ transcriptional reporter allele for FGF signaling (Morgani et al., 2018). After 12 h, H2B-Venus was strongly expressed in a halo of reporter cells immediately surrounding the *Fgf4* wild type cells, but reporter expression dropped precipitously in cells further away from the signal-emitting cells (Fig. 3E). The spatial decay of the H2B-Venus signal was well-approximated by an exponential fit with a decay length of ~ 11 µm (9.2 - 12.4 µm, 95% confidence interval, Fig. 3F). This is likely an overestimate of the immediate effective range of paracrine FGF4 signaling, as the transcriptional reporter integrates signaling activity over long timescales, during which cell divisions and movement will increase the distance between signal-sending and –receiving cells. Thus, the range of repressive coupling via FGF4 that drives robust Epi/PrE-like cell fate proportioning in ESCs is spatially restricted and acts most efficiently at cell-cell contacts.

A distribution of cell fates is maintained by intercellular communication

A key property of a population-based mechanism for cell differentiation, such as the IHSS, is the interdependence of cells with different fates. In numerical simulations, this property of the IHSS solution manifests in the regeneration of the heterogeneous populations following the isolation of a single cell type (Stanoev et al., accompanying manuscript). To test whether different cell identities were likewise re-established through intercellular communication in

ESCs, we isolated cells expressing a live reporter for the PrE-like fate, and followed reporter expression in a cell population in defined medium over time (Fig. 4A).

To track PrE-like identity, we used a transcriptional reporter construct for *Gata6* expression in the background of a GATA4-mCherry inducible cell line (Freyer et al., 2015; Schröter et al., 2015). To be able to follow both the up- and the downregulation of reporter expression, we used VNP as a short-lived reporter protein (Abranches et al., 2013; Nagoshi et al., 2004). Gata6:VNP-positive cells sorted 16 h after the end of a doxycycline pulse regenerated a mixture of VNP-positive and –negative cells within 10 h of culture in N2B27, similar to the distribution in cell colonies that had not been disrupted and sorted (Fig. 4B, C). In contrast, VNP expression was maintained in most sorted cells upon supplementation of the medium with FGF4, while inhibition of FGF/ERK signaling with the MEK inhibitor PD03 completely abrogated Gata6:VNP expression following sorting (Fig. 4B, C). A similar loss of Gata6:VNP expression was observed when *Fgf4* mutant Gata6:VNP-positive cells that had been differentiated in the presence of recombinant FGF4 for 16 h were sorted and seeded in N2B27 (Supplementary Fig. S4). This indicates that ongoing cell-cell communication via the FGF4/ERK axis organizes the re-establishment of a mixture of cell identities in a population. Following the dynamics of reporter expression over time revealed that in N2B27, Gata6:VNP levels were transiently downregulated before the heterogeneous expression pattern was re-established (Fig. 4D, Supplementary Movie S1). This was in contrast to reporter expression dynamics in the presence of PD03, where all cells rapidly downregulated VNP, and also differed from expression dynamics both in unperturbed colonies as well as in Gata6:VNP-positive cells cultured in the presence of FGF4, where reporter expression remained constant over time in the majority of cells (Fig. 4D, Supplementary movies S2 – S4). The transient dynamics of Gata6 reporter expression in N2B27 were paralleled by characteristic changes in *Fgf4* expression: While Fgf4 transcripts could hardly be detected immediately after sorting, they re-appeared in a fraction of VNP-low cells 6 h later, at around the same time that reporter

286 expression started to increase in a subset of cells (Fig. 4E). Taken together, these results
287 indicate FGF/ERK signaling re-establishes populations with heterogeneous cell identities
288 following the isolation of a PrE-like cells. In unperturbed cell colonies, intercellular
289 communication via FGF/ERK therefore not only generates, but also actively maintains
290 balanced proportions of cells with heterogeneous identities.
291
292 **Discussion**
293 Here we report emergent population-level behaviors in the differentiation of Epi- and PrE-like
294 cells from an induced ICM-like state in ESCs: Robust proportions of cells with the two
295 identities are specified from a wide range of initial conditions, and re-established from
296 isolated PrE-like cells. We demonstrate that these emergent behaviors depend on a short-
297 range FGF4-signal that couples cell identities across the population. The experimental system
298 recapitulates the behavior of a general class of dynamical systems that break symmetry,
299 establish and maintain robust proportions of differentiated cell fates through an IHSS as a
300 population-based solution (Stanoev et al., accompanying manuscript). Our results suggest a
301 new function for FGF signaling, which is to generate and maintain cell type diversity in
302 defined proportions.
303
304 The specification of Epi- and PrE-like identities from an induced ICM-like state in ESCs
305 displays both molecular and functional parallels to the patterning of the ICM of the mouse
306 preimplantation embryo. Similar to the situation in the embryo, we find that PrE-like
307 differentiation is lost in *Fgf4* mutant cells, a phenotype that previous *in vitro* studies missed,
308 possibly due to permanent high-level expression of exogenous GATA factors (Kang et al.,
309 2012; Wamaitha et al., 2015). Furthermore, we recapitulate *in vitro* the remarkably constant
310 proportions of cell types that are also seen in the developing embryo (Saiz et al., 2016).
311 Similar to the situation in the embryo, cell fates in ESC populations are plastic and can be re-

specified upon changing a cell's environment (Grabarek et al., 2012; Martinez Arias et al., 2013). Thus, the functional behavior of the ESC system mirrors the robust and regulative development of the mouse preimplantation embryo.

Molecularly, recursive communication in ESCs is realized through the regulation of *Fgf4* expression by transcriptional regulators of cellular identity such as GATA factors and NANOG, as well as through the dose-dependent response of cells to FGF4 ligand levels. In the embryo, *Fgf4* expression in single cells correlates with markers of cellular identity from an early timepoint onwards (Guo et al., 2010; Ohnishi et al., 2014), and depends on the epiblast-specific transcriptional regulators *Oct4* and *Nanog* (Frankenberg et al., 2011; Messerschmidt and Kemler, 2010; Nichols et al., 1998). Furthermore, a dose-dependent increase in the number of PrE-cells has been reported for both *Fgf4* mutant and wild type embryos treated with increasing concentrations of recombinant FGF4 (Krawchuk et al., 2013). These parallels between ESCs and the embryo suggest that similar molecular and dynamical mechanisms underlie cell fate patterning in the two systems. Recent work employing targeted manipulations to lineage sizes in *Fgf4* wild type and mutant embryos has likewise led to the conclusion that an FGF4-based population-level mechanism for cell fate decisions balances lineage size in the embryo (Saiz et al., 2019). In the future, it will be interesting to investigate whether robust cell fate proportioning through recursive communication via FGF extends to cell fate decisions beyond those of preimplantation development.

In the differentiation paradigm studied here, repressive coupling via FGF4 establishes discrete cell identities in a population, a function that has classically been associated with Delta-Notch signaling (Henrique and Schweisguth, 2019; Hori et al., 2013). In an engineered cell system, communication via Delta-Notch generates population-level behaviors similar to those caused by FGF4 in differentiating ESCs, such as the spontaneous emergence of discrete fates in

stable ratios, the re-establishment of those ratios upon removal of one cell type, and the dependence of cell fate ratios on cell density or contact (Matsuda et al., 2015). In the case of the engineered system, repressive coupling is realized through a Notch-responsive artificial repressor of Delta expression (Matsuda et al., 2015), while in differentiating ESCs, FGF4 signaling likely mediates repressive coupling through its connection to the mutually repressive transcriptional regulatory programs that specify fates. Thus, when embedded in appropriate intracellular regulatory circuits, molecularly diverse intercellular communication systems can yield similar functional outputs.

A central effect of intercellular communication during the differentiation of Epi- and PrE-like cells is to establish population level robustness, making the collective differentiation outcome independent from the distribution of initial conditions in single cells. This is because the behavior of individual cells in a communicating system is different from that of the cells under non-communicating conditions. As the population-based behavior cannot directly be extrapolated from that of single cells, the generation of heterogeneous cell identities within a population requires a different theoretical treatment. In an accompanying manuscript, a new generic dynamical mechanism has been identified, an IHSS, that describes how heterogeneous entities can emerge from a homogenous population in presence of cell-cell communication (Stanoev et al., accompanying manuscript). The robust generation of specific cell fate proportions irrespective of initial conditions, and their active maintenance through intercellular communication that we observe experimentally are two key properties of the IHSS, suggesting that the IHSS is a likely dynamical mechanism underlying the differentiation of cells with discrete identities during mammalian preimplantation development. Current models for the specification of Epi- and PrE-identities in the ICM likewise consist of cell-intrinsic multistable regulatory networks coupled through cell-cell communication (Bessonnard et al., 2014; De Caluwé et al., 2019; De Mot et al., 2016).

364 Although it is possible that these models also show emergent population-level behavior, this
365 has not been explored, neither theoretically nor experimentally.
366 The population-level mechanism for cell differentiation that we describe here both generates
367 and evenly populates discrete cell states. It thus embodies both the tendency of single cells to
368 differentiate towards well-defined identities, as well as the tendency of tissues, organs and
369 organisms to develop towards a reproducible and stereotyped appearance, thereby reuniting
370 two connotations of the term developmental canalization (Ferrell, 2012; Guignard et al.;
371 Waddington, 1942).
372
373

Methods

Cell lines

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

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

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

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

400 **Genetic engineering of ESC lines**

401 Doxycycline-inducible *GATA4*-mCherry inducible ES cells were generated by electroporation
402 of 50.000 E14tg2a ES cells with 4 µg of pPB-TET-GATA4-mCherry, 4 µg pCAG-rtTA-Neo,
403 and 4 µg pCAG-PBase (Wang et al., 2008) followed by G418 selection (400 µg/ml) one day
404 after transfection. We established more than 10 independent clonal lines and assayed
405 induction levels and homogeneity by flow cytometry 2 – 8 h after induction of transgene
406 expression by adding 500 ng/ml doxycycline to the culture medium. Four clones with
407 homogeneous induction levels were chosen and maintained under G418 selection, to
408 circumvent silencing of the inducible transgene.
409 Mutagenesis of the *Fgf4* locus was performed as previously described (Morgani et al., 2018).
410 *Fgf4* loss of function clones were identified by PCR-amplification, cloning and sequencing of
411 a sequence around the *Fgf4* start codon. We either selected clones with a targeted mutation
412 delivered by a single-stranded DNA repair template that we have previously shown to disrupt
413 *Fgf4* function (Morgani et al., 2018), or selected at least two independent clones carrying
414 indels around the start codon that introduced frameshift as well as nonsense mutations. All
415 independent clones with random indels showed indistinguishable behavior in the
416 differentiation assays.
417
418 The *Gata6* reporter cell line was generated using previously described knock-out first
419 targeting arms of the EUCOMM project (Skarnes et al., 2011), combined with a VNP reporter
420 cassette (Nagoshi et al., 2004) and a neomycin resistance gene driven from a human β-actin
421 promoter. This construct was integrated by homologous recombination into a line carrying a
422 doxycycline-inducible GATA4-mCherry transgene in the *Col1a1* locus as well as a randomly
423 integrated H2B-Cerulean nuclear marker driven by a CAGS promoter described in (Schröter
424 et al., 2015). Clones were screened for correct integration of the reporter construct by long
425 range PCR spanning the targeting arms.

The targeting construct to generate the *Spry4*$^{H2B\text{-}Venus}$ allele in GATA4-mCherry inducible cell lines was based on the one used in (Morgani et al., 2018), except that the puromycin selectable marker was exchanged for a neomycin cassette. The construct was integrated into ESCs by homologous recombination, and neomycin-resistant clones were expanded and screened for correct integration of the reporter construct by long range PCR spanning the targeting arms. All genetically modified lines were karyotyped using standard procedures (Nagy et al., 2008), and only lines with a median chromosome count of n = 40 were used for experiments.

Immunostaining and QIF

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

In situ HCR

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

minutes. Transcript-specific probes were used at a concentration of 4 nM and hybridized overnight. Excess probe was removed through several washes with probe wash buffer and 5x SSCT, and cells were equilibrated in amplification buffer for at least 30 minutes. Fluorescently labeled amplifiers were used at a concentration of 60 nM. Amplification was allowed to proceed for 16 – 24 hours at room temperature. Excess amplifier was removed by several washes with 5x SSCT, followed by counterstaining with Hoechst 33342 and mounting in glycerol-based medium. Imaging was performed on an SP8 confocal microscope with a 63x (NA1.4) lens.

Flow cytometry

Staining for flow cytometric analysis of intracellular antigens was performed as previously described (Schröter et al., 2015). Primary and secondary antibodies were the same as used for immunostaining. mCherry fluorescence measurements and cell sorting were performed on a BD FACS Aria. All other flow cytometric analysis was carried out using a BD LSR II. Single cell events were gated based on forward and side scatter properties. Gates to separate marker-positive from marker-negative cells were determined visually as the threshold that best bisected the bimodal distribution of marker expression across all samples within one experiment.

Decay length measurements

Fgf4 mutant Spry4:H2B-reporter cells (Morgani et al., 2018) were seeded at a density of $5*10_4$ cells/cm_2 in N2B27. 2 hours later, dsRed-expressing cells were added at a density of 500 cells/cm_2. For the first 3 hours of co-culture, the medium was supplemented with 250 to 500 nM siR-Hoechst (Lukinavičius et al., 2015) to label nuclei. 12 h later, live cells were imaged on a Leica SP8 confocal system. Nuclei were segmented in FIJI (Schindelin et al., 2012), and for each *Spry4*$_{H2B\text{-}Venus}$ 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.

Live cell imaging and tracking

To track *Gata6* reporter expression in live cells, PrE-like differentiation was induced by a 6 h pulse of doxycycline-treatment in serum-containing medium as described in (Schröter et al., 2015). 16 hours after doxycycline-removal, cells were either switched directly to N2B27 medium lacking phenol red, or trypsinized, sorted for reporter expression, and seeded on fibronectin-coated imaging dishes (ibidi µ-slides). Time-lapse imaging was started within 2 h after sorting on an Olympus IX81 widefield microscope equipped with LED illumination (pE4000, CoolLED) and a Hamamatsu c9100-13 EMCCD camera. Hardware was controlled by MicroManager software (Edelstein et al., 2001). Time-lapse movies were acquired using a 40x oil immersion lens (NA 1.2), with 10-minute time intervals.

Cell tracking was carried out with TrackMate (Tinevez et al., 2017) based on the constitutively expressed H2B-Cerulean nuclear marker. Fluorescence intensity was measured in a circular region of interest in the center of the nucleus, and background-subtracted fluorescence intensities plotted in Python. Trace color in Fig. 4D was assigned according to fluorescence intensity in the last frame of the movie, with respect to the approximated intensity threshold used for flow sorting (dashed line).

Computational model for cell fate proportioning

The model of the intercellular communication system (Fig. 2D) is adapted from an accompanying manuscript by Stanoev et al., and is described with the following set of equations:

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

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

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

N_i and G_i describe NANOG and GATA6 protein expression levels in cell i, regulated by mutual inhibition, while F_i is the secreted FGF4 whose production is downregulated by GATA6. $F_{ext,i} = \frac{1}{|N(i)|+1}\sum_{j\in(N(i)\cup i)} F_j$ is the extracellular FGF4 concentration that is sensed by cell i from its neighborhood $N(i)$, resulting in downregulation of NANOG production in the cell. $\alpha_N = 2.5, \alpha_{N,F} = 0.5, \alpha_G = 3$ and $\alpha_F = 3$ denote production rate constants, $\beta = \eta = \gamma = \delta = 2$ are the Hill coefficients, degradation rates were set to 1 as $\lambda = 50$ was used as a scaling kinetic parameter. 10000 cells were deployed on a regular 100x100 two-dimensional lattice with no-flux boundary conditions. Cell-cell communication was modeled to be short-range, reflecting the experimental wild-type case, i.e. communication between direct neighbors and cells on two hops away on the lattice, as described in (Stanoev et al., accompanying manuscript). When mimicking the *Fgf4* mutant case, communication between cells was excluded, and an external input was modeled with $F_{ext} = 1.2$.

The cell populations were initiated analogous to the experimental case (Fig. 2C), by varying the initial conditions of all cells from being NANOG-expressing, through intermediate NANOG and GATA6 expression, to being GATA6-expressing. More specifically, the variables were sampled independently from unimodal Gaussian distributions

$\mathcal{N}(\mu_{ics}(p), \sigma_{ics} = 0.1 * \mu_{ics}(p))$, with the mean $\mu_{ics}(p) = (1-p) * \mu_{G-;N+} + p * \mu_{G+;N-}$ placed on the line segment connecting the GATA6-; NANOG+ state $\mu_{G-;N+}$ and the GATA6+; NANOG- state $\mu_{G+;N-}$, partitioning it in proportion p. $p \in \{0, 0.4, 0.5, 0.6, 1\}$ was used for the quantifications in Fig. 2D, right. Samples from around the endpoints and the midpoint ($p \in \{0, 0.5, 1\} \Rightarrow \mu_{ics} \in \{\mu_{G-;N+}, \frac{1}{2}(\mu_{G-;N+} + \mu_{G+;N-}), \mu_{G+;N-}\}$) are shown in Fig. 2D, top row.

Cell heterogeneity was introduced by varying all of the parameters independently with standard deviation of 0.02 from the respective values for each cell. Stochastic differential equation model was constructed from the deterministic equations by adding a multiplicative noise term $\sigma X dW_t$, where dW_t is the Brownian motion term, X is the variable state and $\sigma = 0.1$ is the noise term. The model was solved with $\Delta t = 0.01$ using the Milstein 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.

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 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

Dhruv Raina, Investigation, Validation, Visualization, Formal analysis, Writing – review & editing

Angel Stanoev, Methodology, Writing – review & editing;

Azra Bahadori, Investigation, Visualization, Writing – review & editing;

Michelle Protzek, Investigation, Validation; Writing – review & editing

Aneta Koseska, Conceptualization, Supervision, Writing – review & editing

Christian Schröter, Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Supervision, Project administration;

References

Abranches, E., Bekman, E., and 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.

Bedzhov, I., Graham, S.J.L., Leung, C.Y., and 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.

Bessonnard, S., De Mot, L., Gonze, D., Barriol, M., Dennis, C., Goldbeter, A., Dupont, G., and 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.

Briscoe, J. (2019). Understanding Pattern Formation in Embryos: Experiment, Theory, and Simulation. Journal of Computational Biology *26*, 696–702.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell *10*, 615–624.

Chickarmane, V., and Peterson, C. (2008). A Computational Model for Understanding Stem Cell, Trophectoderm and Endoderm Lineage Determination. PLoS ONE *3*, e3478–e3478.

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

De Caluwé, J., Tosenberger, A., Gonze, D., and Dupont, G. (2019). Signalling-modulated gene regulatory networks in early mammalian development. J Theor Biol *463*, 56–66.

De Mot, L., Gonze, D., Bessonnard, S., Chazaud, C., Goldbeter, A., and Dupont, G. (2016). Cell Fate Specification Based on Tristability in the Inner Cell Mass of Mouse Blastocysts. Biophysj *110*, 710–722.

Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2001). Computer Control of Microscopes Using μManager (Hoboken, NJ, USA: John Wiley & Sons, Inc.).

Enver, T., Pera, M., Peterson, C., and Andrews, P.W. (2009). Stem cell states, fates, and the

587 rules of attraction. Cell Stem Cell *4*, 387–397.

588 Feldman, B., Poueymirou, W., Papaioannou, V.E., DeChiara, T.M., and Goldfarb, M. (1995).
589 Requirement of FGF-4 for postimplantation mouse development. Science *267*, 246–249.

590 Ferrell, J.E., Jr (2012). Bistability, Bifurcations, and Waddington's Epigenetic Landscape.
591 Current Biology *22*, R458–R466.

592 Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., and
593 Chazaud, C. (2011). Primitive endoderm differentiates via a three-step mechanism involving
594 Nanog and RTK signaling. Dev Cell *21*, 1005–1013.

595 Freyer, L., Schröter, C., Saiz, N., Schrode, N., Nowotschin, S., Martinez Arias, A., and
596 Hadjantonakis, A.-K. (2015). A loss-of-function and H2B-Venus transcriptional reporter
597 allele for Gata6 in mice. BMC Dev Biol *15*, 38.

598 Grabarek, J.B., Zyzynska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J.,
599 Hadjantonakis, A.-K., and Plusa, B. (2012). Differential plasticity of epiblast and primitive
600 endoderm precursors within the ICM of the early mouse embryo. Development *139*, 129–139.

601 Guignard, L., Fiuza, U.-M., Leggio, B., Faure, E., Laussu, J., Hufnagel, L., Malandain, G.,
602 Godin, C., and Lemaire, P. (2017). Contact-dependent cell-cell communications drive
603 morphological invariance during ascidian embryogenesis. bioRxiv, doi: 10.1101/238741.

604 Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010).
605 Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote
606 to blastocyst. Dev Cell *18*, 675–685.

607 Hamilton, W.B., and Brickman, J.M. (2014). Erk Signaling Suppresses Embryonic Stem Cell
608 Self- Renewal to Specify Endoderm. CellReports 1–16.

609 Henrique, D., and Schweisguth, F. (2019). Mechanisms of Notch signaling: a simple logic
610 deployed in time and space. Development *146*, dev172148–11.

611 Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient
612 (Lesch|[ndash]|Nyhan) mouse embryos derived from germline colonization by cultured cells.
613 Nature *326*, 292–295.

614 Hori, K., Sen, A., and Artavanis-Tsakonas, S. (2013). Notch signaling at a glance. Journal of
615 Cell Science *126*, 2135–2140.

616 Huang, S., Guo, Y.-P., May, G., and Enver, T. (2007). Bifurcation dynamics in lineage-
617 commitment in bipotent progenitor cells. Dev Biol *305*, 695–713.

618 Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.K. (2012). FGF4 is required for
619 lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their
620 initial expression in the mouse. Development *140*, 267–279.

621 Krawchuk, D., Honma-Yamanaka, N., Anani, S., and Yamanaka, Y. (2013). FGF4 is a
622 limiting factor controlling the proportions of primitive endoderm and epiblast in the ICM of
623 the mouse blastocyst. Dev Biol *384*, 65–71.

624 Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A.

(2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

Lukinavičius, G., Blaukopf, C., Pershagen, E., Schena, A., Reymond, L., Derivery, E., González-Gaitán, M., D'Este, E., Hell, S.W., Gerlich, D.W., et al. (2015). SiR-Hoechst is a far-red DNA stain for live-cell nanoscopy. Nature Communications *6*, 8497.

Martinez Arias, A., Nichols, J., and Schröter, C. (2013). A molecular basis for developmental plasticity in early mammalian embryos. Development *140*, 3499–3510.

Matsuda, M., Koga, M., Woltjen, K., Nishida, E., and Ebisuya, M. (2015). Synthetic lateral inhibition governs cell-type bifurcation with robust ratios. Nature Communications *6*, 6195.

Messerschmidt, D.M., and Kemler, R. (2010). Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. Dev Biol *344*, 129–137.

Morgani, S.M., Saiz, N., Garg, V., Raina, D., Simon, C.S., Kang, M., Arias, A.M., Nichols, J.N., Schröter, C., and Hadjantonakis, A.-K. (2018). A Sprouty4 reporter to monitor FGF/ERK signaling activity in ESCs and mice. Dev Biol *441*, 104–126.

Nagoshi, E., Saini, C., Bauer, C., Laroche, T., Naef, F., and Schibler, U. (2004). Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. Cell *119*, 693–705.

Nagy, A., Gertsenstein, M., Vintersten, K., and Behringer, R. (2008). Karyotyping mouse cells. CSH Protoc *2008*, pdb.prot4706–pdb.prot4706.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379–391.

Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oleś, A.K., Araúzo-Bravo, M.J., Saitou, M., Hadjantonakis, A.-K., et al. (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. Nature Cell Biology *16*, 27–37.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.-K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081–3091.

Saiz, N., Mora-Bitria, L., Rahman, S., George, H., Herder, J.P., Garcia-Ojalvo, J., and Hadjantonakis, A.-K. (2019). Growth factor-mediated coupling between lineage size and cell fate choice underlies robustness of mammalian development. bioRxiv. doi:10.1101/2019.12.27.889006

Saiz, N., Williams, K.M., Seshan, V.E., and Hadjantonakis, A.-K. (2016). Asynchronous fate decisions by single cells collectively ensure consistent lineage composition in the mouse blastocyst. Nature Communications *7*, 13463.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676–682.

665 Schröter, C., Rué, P., Mackenzie, J.P., and Martinez Arias, A. (2015). FGF/MAPK signaling
666 sets the switching threshold of a bistable circuit controlling cell fate decisions in embryonic
667 stem cells. Development *142*, 4205–4216.

668 Shimokawa, K., Kimura-Yoshida, C., Nagai, N., Mukai, K., Matsubara, K., Watanabe, H.,
669 Matsuda, Y., Mochida, K., and Matsuo, I. (2011). Cell surface heparan sulfate chains regulate
670 local reception of FGF signaling in the mouse embryo. Dev Cell *21*, 257–272.

671 Simon, C.S., Hadjantonakis, A.-K., and Schröter, C. (2018). Making lineage decisions with
672 biological noise: Lessons from the early mouse embryo. WIREs Dev Biol *61*, e319–16.

673 Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O.,
674 Thomas, M., Harrow, J., Cox, T., et al. (2011). A conditional knockout resource for the
675 genome-wide study of mouse gene function. Nature *474*, 337–342.

676 Tinevez, J.-Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E.,
677 Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: An open and extensible
678 platform for single-particle tracking. Methods *115*, 80–90.

679 Waddington, C.H. (1942). Canalization of development and the inheritance of acquired
680 characters. Nature *150*, 563–565.

681 Wamaitha, S.E., del Valle, I., Cho, L.T.Y., Wei, Y., Fogarty, N.M.E., Blakeley, P., Sherwood,
682 R.I., Ji, H., and Niakan, K.K. (2015). Gata6 potently initiates reprograming of pluripotent and
683 differentiated cells to extraembryonic endoderm stem cells. Genes Dev *29*, 1239–1255.

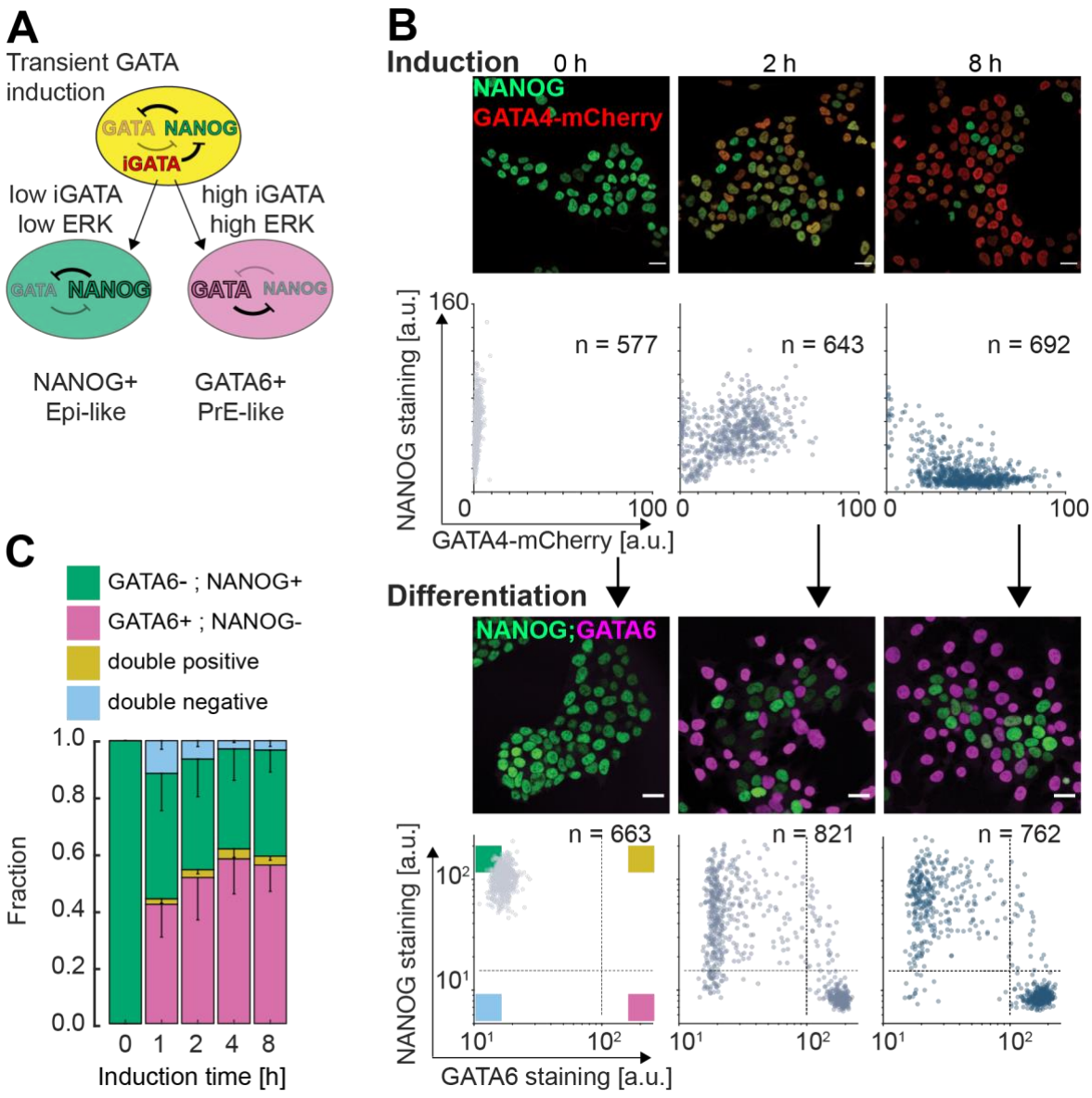
684 Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., and Liu, P.
685 (2008). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. Proceedings
686 of the National Academy of Sciences *105*, 9290–9295.

687 Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and
688 Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–
689 523.

690

691

692 **Figures**



693

694 **Fig. 1: Proportions of differentiated cell types are independent from GATA4-mCherry**

695 **induction levels**

696 **A** Schematic representation of differentiation paradigm. From an ICM-like state (yellow)

697 generated by transient doxycycline-mediated expression of GATA factors, cells return to an

698 Epi-like state (green, left) or adopt a PrE-like identity (magenta, right) depending on GATA

699 levels and ERK activity. **B** Top rows: Immunostaining and quantitative analysis of

700 fluorescence in individual nuclei of NANOG (green) and GATA4-mCherry (red) expression

701 in inducible cell lines after indicated durations of doxycycline stimulation. Bottom rows:

Immunostaining and quantification of GATA6 and NANOG expression in cells treated with doxycycline for indicated periods of time and differentiated in N2B27 for 40 h. Cells without doxycycline induction have been continuously maintained in 2i medium. Dashed lines in scatter plots indicate thresholds to determine cell identities: Upper left quadrant: GATA6-; NANOG+ (G-,N+); lower right quadrant GATA6+; NANOG- (G+,N-); upper right quadrant double positive (DP); lower left quadrant double negative (DN). Scale bars, 20 µm. **C** Summary of mean proportions from N = 4 independent experiments; fraction of G+,N- cells in magenta, G-,N+ in green, DP cells in yellow, and DN cells in blue. Error bars indicate 95% confidence intervals.

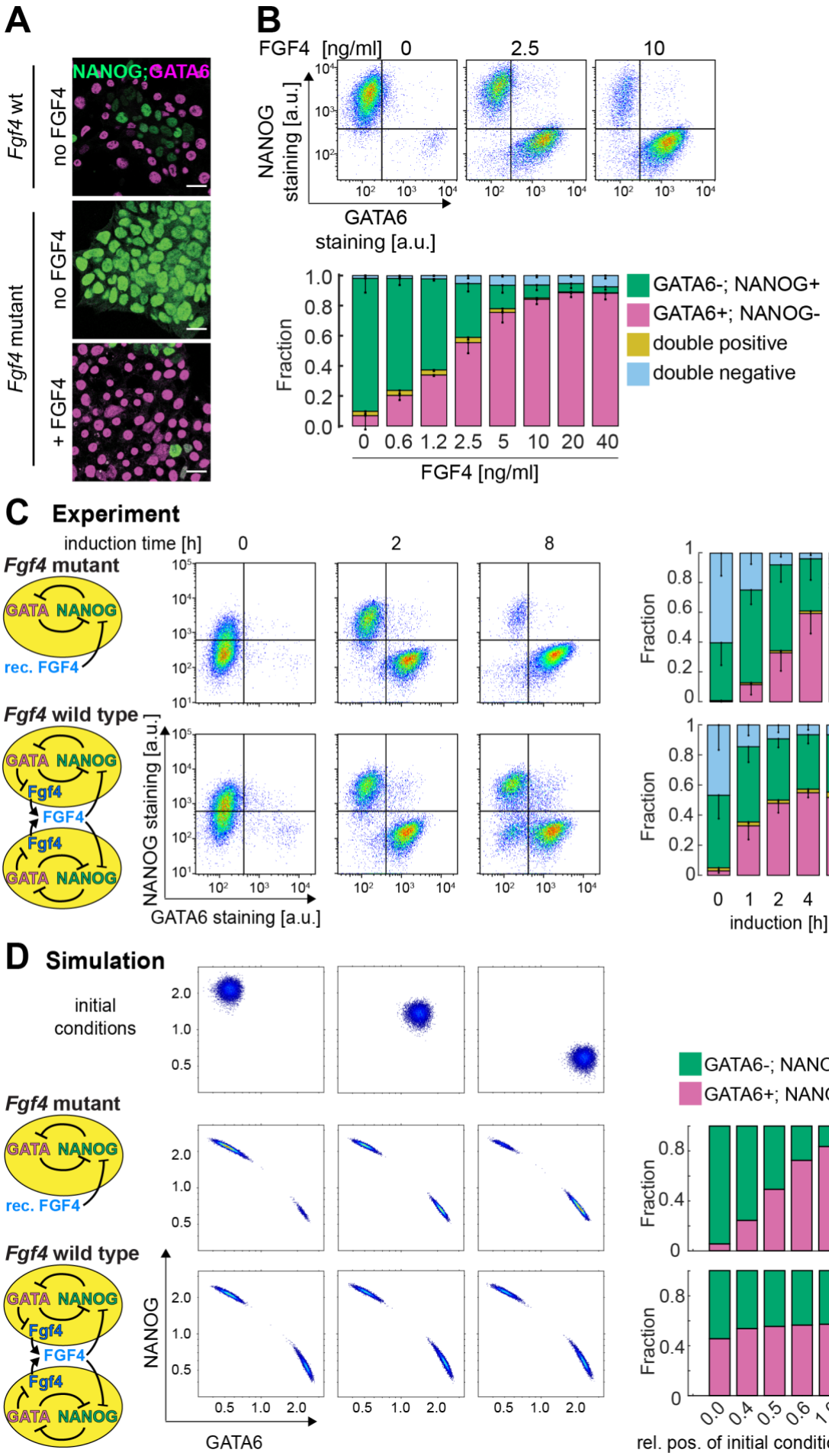
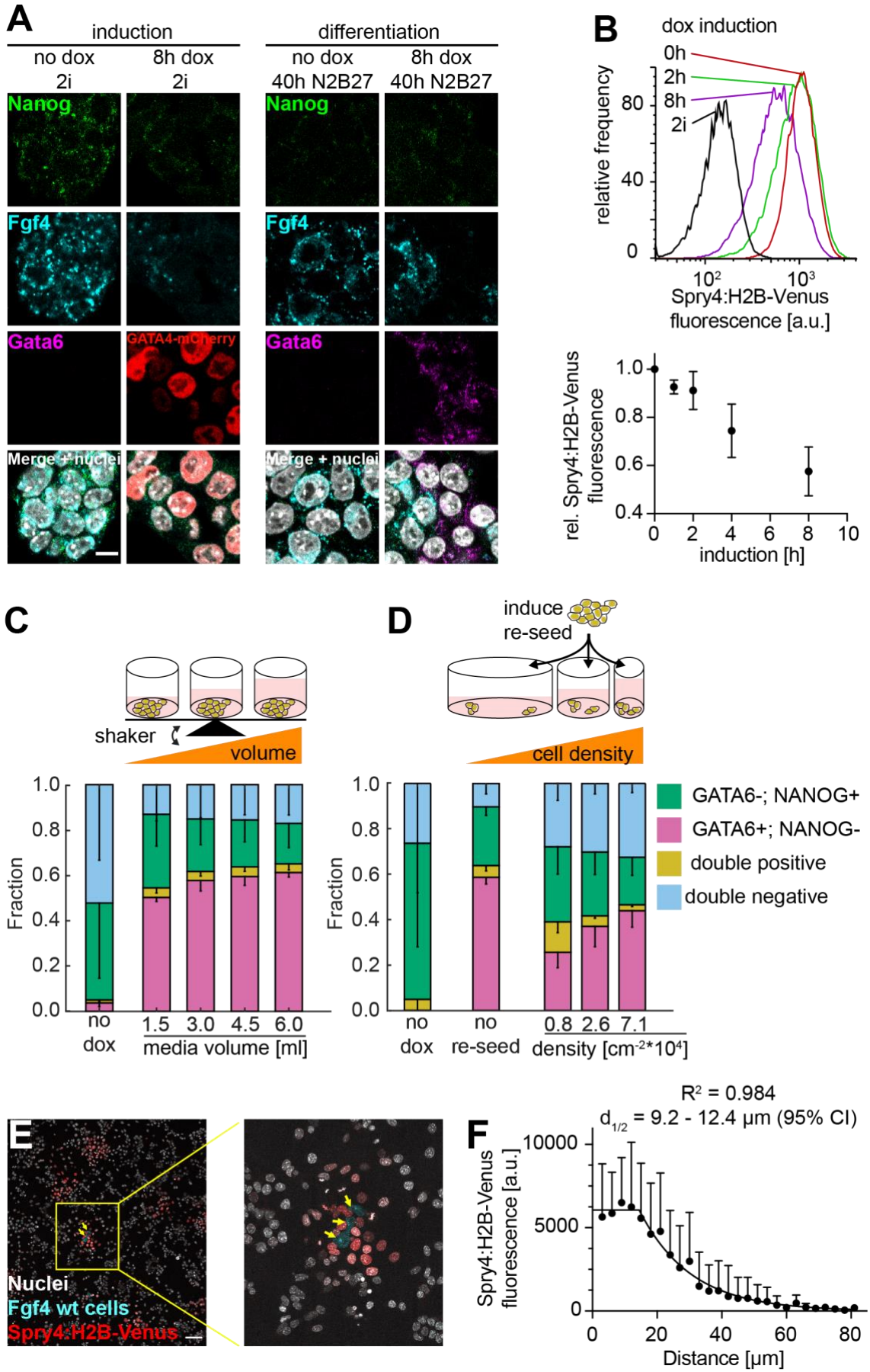


Fig. 2: Cell-cell communication via FGF4 mediates cell fate proportioning

A Immunostaining of *Fgf4* wild type (top) and *Fgf4* mutant (bottom) cells stained for GATA6 (magenta) and NANOG (green) after 8 h of doxycycline induction followed by 40 h of differentiation in N2B27 medium only, or upon supplementation with 10 ng/ml FGF4 as indicated. Scale bars, 20 µm. **B** Top: Flow cytometry of *Fgf4* mutant cells stained as in **A** following 8 h of doxycycline induction and differentiation in N2B27 supplemented with the indicated concentrations of FGF4. Lines indicate gates to assign cell identities. Bottom: Bar chart summarizing mean proportions aggregated from four independent experiments; fraction of GATA6+; NANOG- cells in magenta, GATA6-; NANOG+ in green, double positive cells in yellow, and double negative cells in blue. Error bars indicate 95% CI. **C** Left: Flow cytometry of *Fgf4* mutant (upper panels) or wild type cells (lower panels) stimulated with doxycycline for the indicated times and differentiated in N2B27 alone (wild type), or in N2B27 supplemented with 10 ng/ml FGF4 (mutant). GATA6 and NANOG expression were visualized by immunostaining. Lines indicate gates to assign cell identities. Right: Bar chart summarizing mean proportions aggregated from four independent experiments; Color code as in **B**, error bars indicate 95% CI. **D** Numerical comparison of single-cell- and population-based cell type proportioning mechanisms. Top row: Initial condition distributions used in the simulations for the two distinct models (rows below). Relative positions of the initial conditions are 0.0 (left), 0.5 (middle) and 1.0 (right). Middle row: Exemplary cell type proportions and respective quantifications (bar plots, right) for the single-cell case, supplemented with constant F_{ext}. Bottom row: Same as above, only for the population-based case. Model and parameters: Methods.



735

736

Fig. 3: Fgf4 expression is repressed by GATA factors and acts locally.

A Staining for Nanog (green), Fgf4 (cyan) and Gata6 (magenta) mRNA by in situ hybridization chain reaction before doxycyline induction (first column on the left), after 8 h of doxycycline induction (second column), and after 40 h in N2B27 without (third) or with previous doxycycline induction (fourth column). At the end of the induction period, GATA4-mCherry protein (red) instead of Gata6 mRNA is detected. Scale bar, 10 µm. **B** Top: Flow cytometry histograms for expression of a *Spry4*$_{H2B\text{-}Venus}$ reporter integrated into GATA4-mCherry inducible lines 24 h after indicated durations of doxycycline induction. Black line indicates reporter expression in cells maintained in 2i medium. Bottom: Mean ± SD of reporter expression from 4 independent experiments, normalized to fluorescence levels of cells transferred to N2B27 without doxycycline induction. **C** Top: Schematic representation of experimental setup to test effects of media volume on cell fate proportioning. Bottom: Quantitative analysis of cell identity analyzed by flow cytometry across N = 3 independent experiments, error bars indicate 95% CI. Fraction of GATA6+; NANOG- cells (G+,N-) in magenta, GATA6-; NANOG+ (G-,N+) in green, double positive cells (DP) in yellow, double negative cells (DN) in blue. **D** Top: Schematic representation of experimental setup to test effects of cell density on fate proportioning. Bottom: Quantitative analysis of cell identity analyzed by flow cytometry across N = 4 independent experiments, error bars indicate 95% CI. Color code as in **C**. **E** Single labeled *Fgf4* wild-type cells (cyan, yellow arrowheads) seeded on a layer of *Fgf4*-mutant *Spry4*$_{H2B\text{-}Venus}$ transcriptional reporter cells. Nuclei are labelled by siR-Hoechst (white), H2B-Venus in red. Scale bar, 100 µm. **F** Quantitative analysis of FGF4 signaling range. Relative H2B-Venus fluorescence intensities and distance from the center of *Fgf4* wild type cells were measured for individual nuclei of *Spry4*$_{H2B\text{-}Venus}$ reporter cells from nine independent signaling centers. Shown are mean ± SD of fluorescence intensities in distance bins of 3 µm width. The fluorescence decay length was estimated by fitting a plateau followed by one-phase exponential decay to the data (black line).

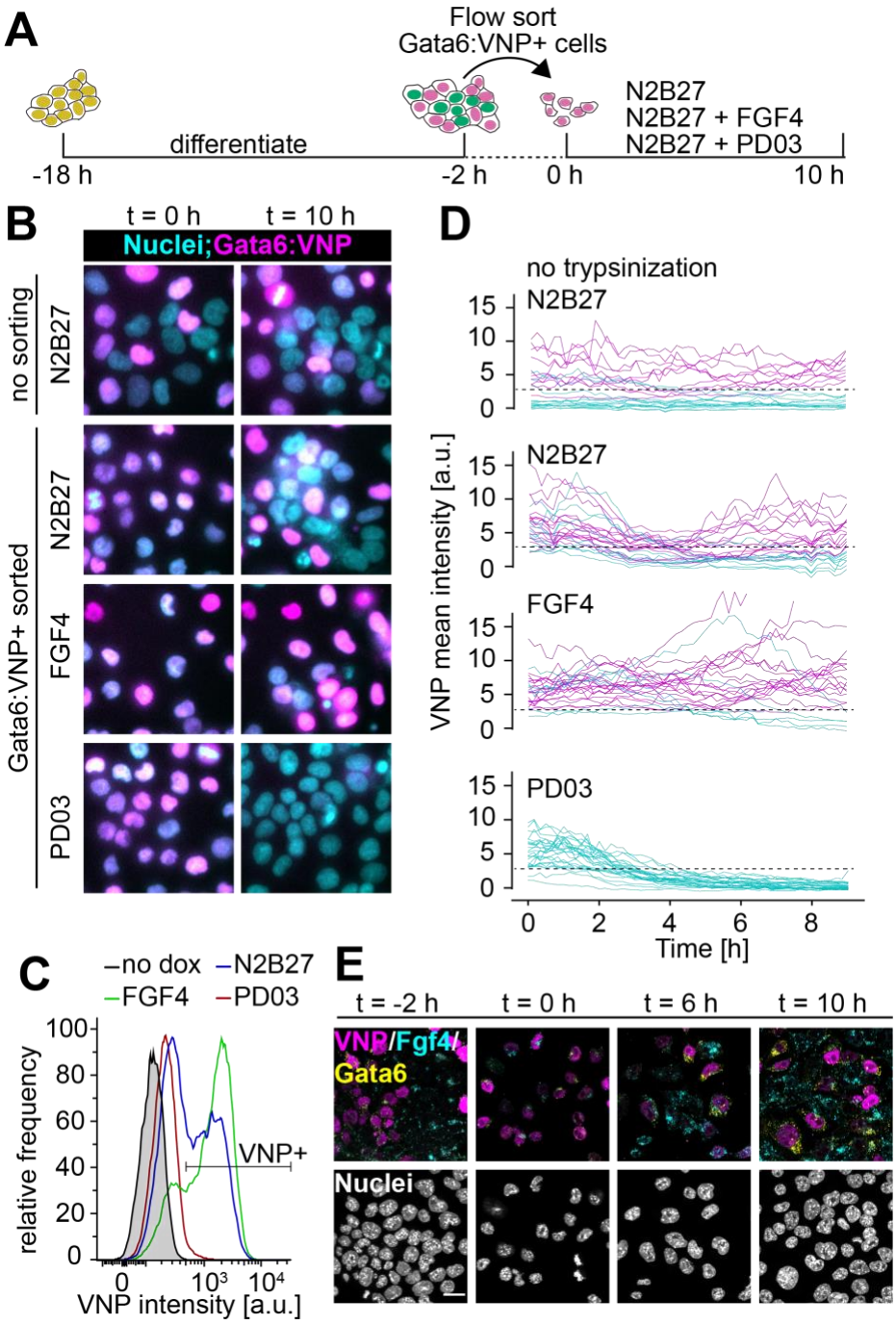
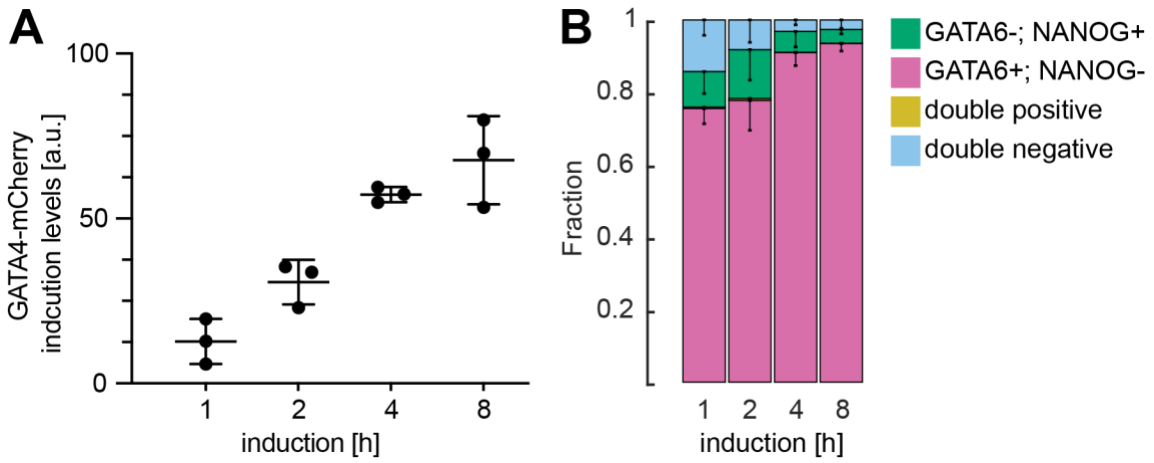


Fig. 4: Heterogeneous cell identities are re-established by cell-cell communication

A Schematic representation of experimental setup. **B** Representative images of non-trypsinized control (upper row) and cells immediately after sorting for Gata6:VNP expression (left), and after 10 h of culturing in N2B27 medium with the indicated supplements (right). **C** Flow cytometry to detect VNP expression in cells that had been sorted for VNP-expression, followed by 10 h of culture in the indicated media. **D** Live-cell traces of VNP expression in individual cells from a non-trypsinized colony (upper panel), or cells sorted for VNP-

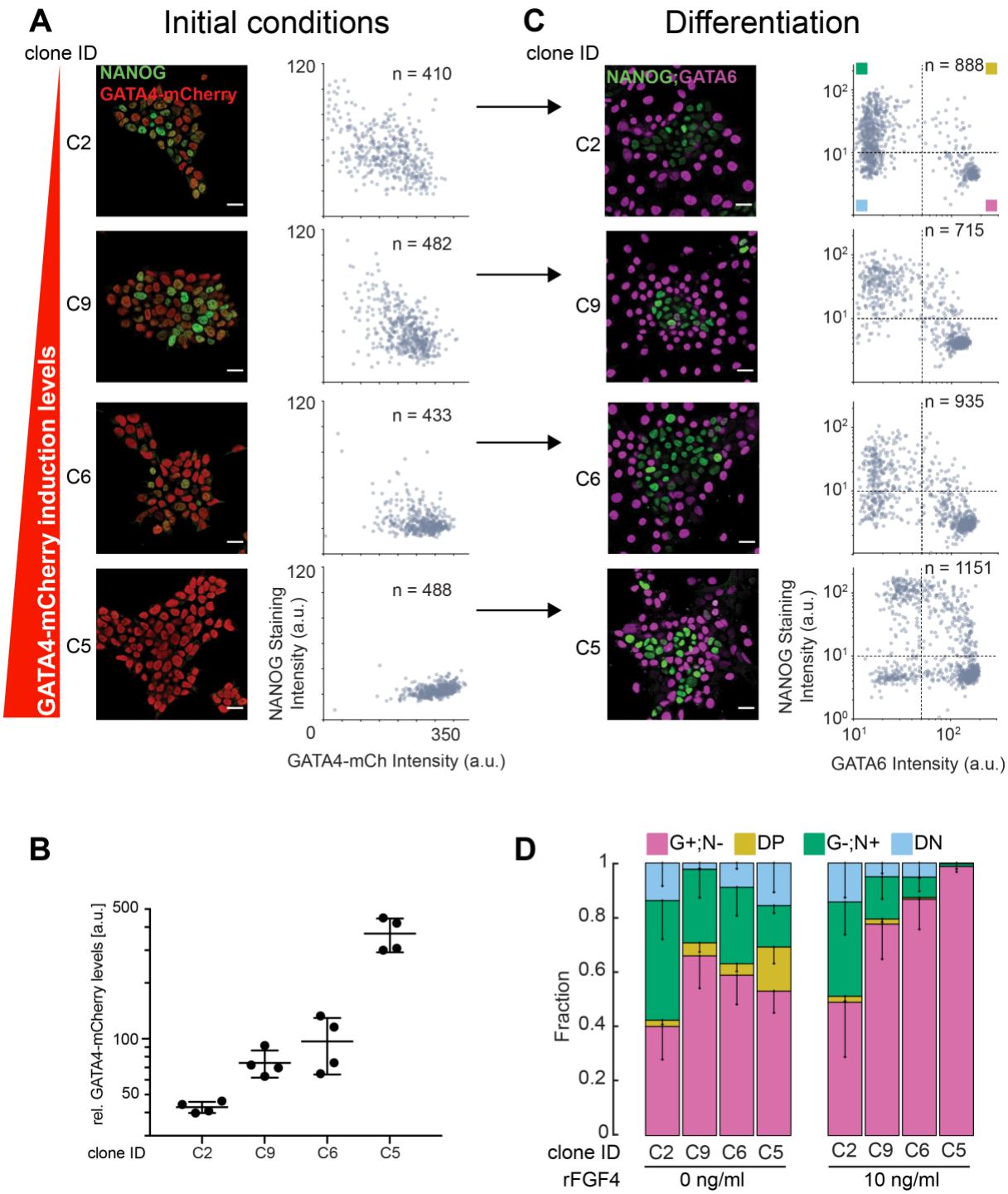
771 expression upon culture in the indicated media. Traces are color coded according to
772 expression levels at the end of the experiment. Dashed line indicates the threshold to separate
773 putative VNP-positive cells (magenta) from VNP-negative cells (cyan). **E** Staining for Fgf4
774 (cyan) and Gata6 (yellow) mRNA in Gata6:VNP-reporter cells before sorting (left) and at 2 h,
775 6 h and 10 h after flow sorting of VNP-positive cells. VNP fluorescence in magenta. Scale
776 bar, 20 µm.
777

Supplementary Figures



Supplementary Fig. S1

A GATA4-mCherry expression levels for different durations of doxycycline induction measured by flow cytometry. Individual data points show normalized mean fluorescence intensities from at least 20.000 cells in an individual experiment, bars indicate mean ± SD across independent experiments. mCherry fluorescence in uninduced cells cultured in 2i + LIF medium was set to one. **B** Proportions of cell fates upon indicated durations of doxycycline induction followed by 40 h of differentiation in N2B27 medium supplemented with 10 ng/ml FGF4. Cell identities were determined by immunostaining and quantitative immunofluorescence, data is pooled from four independent experiments, error bars indicate 95% CI.

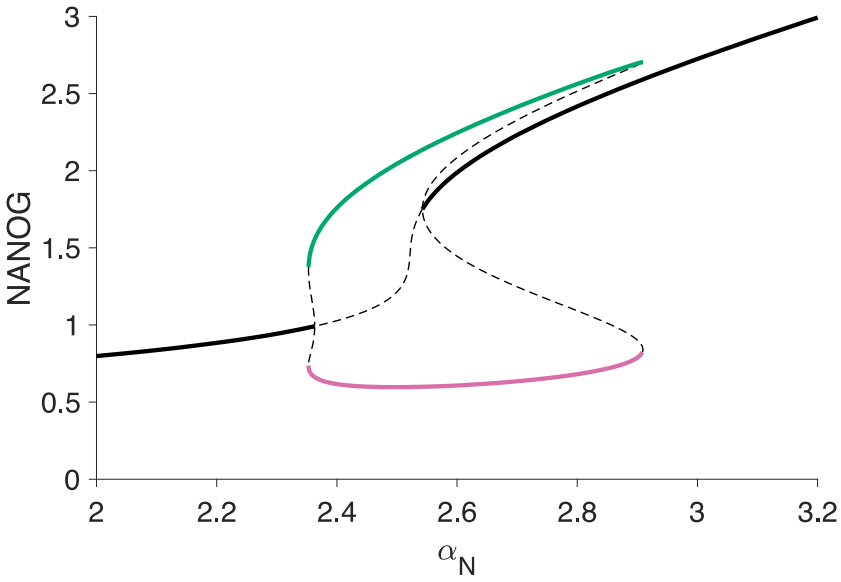


792

793 **Supplementary Fig. S2**

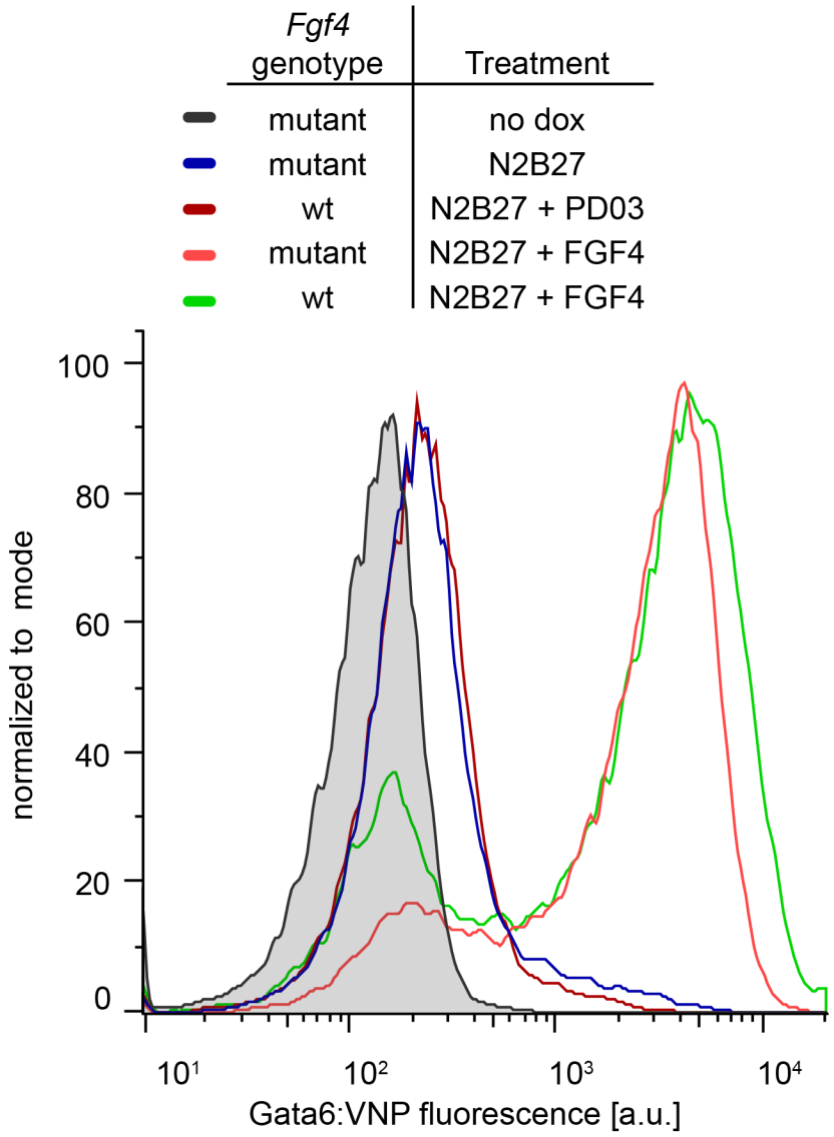
794 **A** Immunostaining (left) and quantitative analysis of fluorescence in individual nuclei (right)
795 of NANOG (green) and GATA4-mCherry (red) expression in four independent clonal
796 inducible cell lines after 8 h of doxycycline stimulation. Clones are ordered by GATA4-
797 mCherry expression strength; one out of 4 independent experiments shown. **B** GATA4-
798 mCherry induction levels in the clones shown in **A** after 8 hours of doxycycline treatment

799 measured by flow cytometry. Mean mCherry mean fluorescence intensity for every clone was
800 normalized to that of the non-induced control in each experiment to allow for comparison.
801 Plot shows individual data points and mean ± SD from four independent experiments.
802 **C** Immunostaining (left) and single cell quantification of GATA6 and NANOG expression in
803 cells from independent clonal lines treated with doxycycline for indicated periods of time and
804 differentiated in N2B27 for 40 h. Dashed lines indicate thresholds to determine cell identities:
805 Upper left quadrant: GATA6-; NANOG+ (G-,N+); lower right quadrant GATA6+; NANOG-
806 (G+,N-); upper right quadrant double positive (DP); lower left quadrant double negative
807 (DN). Clones are ordered by GATA4-mCherry induction strength as in **A.** Scale bars in **A**, **C**,
808 20 µm. **D** Quantification of results from **C** across four independent experiments for
809 differentiation in N2B27 only (left) or in N2B27 supplemented with 10 ng/ml FGF4 (right).
810 Fraction of G+,N- in magenta, G-,N+ in green, DP cells in yellow, and DN cells in blue. Error
811 bars indicate 95% CI.
812



Supplementary Fig. S3

Bifurcation diagram showing the stable IHSS solution for a two-cell coupled system in dependence of the production rate constant of NANOG. Solid lines: stable branches of the IHSS solution; dashed lines: unstable solutions. Black: HSS; green and magenta: GATA6-/NANOG+ and GATA6+/NANOG- conjugate branches of the IHSS solution.



Supplementary Fig. S4

Flow cytometry to detect Gata6:VNP expression in *Fgf4* wild type and mutant cells 10 h after sorting of VNP-positive cells and culture in the indicated media. Reporter expression levels are similar between *Fgf4* mutant cells cultured in N2B27 (dark blue) and *Fgf4* wild type cells cultured in N2B27 supplemented with the MEK inhibitor PD03 (dark red).

Supplementary Movie legends

Supplementary Movie S1 – S3

Time-lapse imaging of Gata6:VNP reporter cells flow sorted for VNP-expression 16 h after the end of a doxycycline pulse and cultured in defined N2B27 medium alone (Movie S1), or cultured in N2B27 supplemented with 10 ng/ml FGF4 (Movie S2), or cultured in N2B27 supplemented with 1 µM of the MEK inhibitor PD03 (Movie S3).

Supplementary Movie S4

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