Growth factor-mediated coupling between lineage size and cell fate choice underlies robustness of mammalian development

Néstor Saiz[1]*, Laura Mora-Bitria[2], Shahadat Rahman[1], Hannah George[1], Jeremy P Herder[1], Jordi García-Ojalvo[2]* and Anna-Katerina Hadjantonakis[1,3]*

[1]Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[2]Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona Biomedical Research Park 08003 Barcelona, Spain
[3]Lead contact

* Correspondence: Anna-Katerina Hadjantonakis (hadjanta@mskcc.org), Jordi García-Ojalvo (jordi.g.ojalvo@upf.edu), Néstor Saiz (saizaren@mskcc.org)

Summary

Precise control and maintenance of the size of cell populations is fundamental for organismal development and homeostasis. The three cell types that comprise the mammalian blastocyst-stage embryo are generated in precise proportions and in a short time, suggesting a size control mechanism ensures a reproducible outcome. Guided by experimental observations, we have developed a minimal mathematical model that explains how growth factor signaling is sufficient to guarantee this robustness. Using laser cell ablation to alter lineage composition in mouse blastocysts we show that cell elimination biases the specification of progenitors towards the targeted cell type, both *in vivo* and *in silico*. Conversely, introduction of supernumerary, lineage-restricted cells shifts the fate of progenitors away from that cell type. Finally, by controlling lineage specification in a growth factor-depleted context we show that FGF4 couples cell behaviors with lineage size. Our results reveal how individual cell fate decisions are coordinated in an *in vivo* self-organizing system to robustly generate tissues of appropriate size providing a basis for the regulative abilities of the mammalian embryo.

Introduction

Across metazoa, coordination between cell fate specification and population size ensures robust developmental outcomes and the maintenance of adult stem cell pools, while loss of control results in malformations and tumorigenesis. Integration of cell behaviors at the population level allow a coordinated response to injury in both embryos and adults (Chen et al., 2015; Wojcinski et al., 2017; Young et al., 2019). The preimplantation mammalian embryo is a paradigm of self-organization where patterning and morphogenesis occur without the need for maternal determinants or external cues. It thus represents a prime *in vivo* system to understand the processes that ensure precision and robustness during the development of multicellular organisms. These embryos can tolerate cell loss, exemplified by preimplantation genetic testing, and can incorporate foreign cells to generate chimeric animals (Bradley et al., 1984; Gardner, 1968; Mintz and Illmensee, 1975; Tachibana et al., 2012; Tarkowski, 1959; 1961). Remarkably, neither of these perturbations impair the developmental potential of the embryo, suggesting that local coordination of cell behavior enables adaptation. Despite recent interest in understanding and exploiting the capacity of the early mammalian embryo for autonomous development (Bedzhov and Zernicka-Goetz, 2014; Deglincerti et al., 2016; Harrison et al., 2017; Morgani et al., 2018a; Rivron et al., 2018; Shahbazi et al., 2019; Sozen et al., 2018; Warmflash et al., 2014), little is known about the underlying mechanisms that enable such resiliency.

The blastocyst-stage embryo is the hallmark of mammalian preimplantation development. It comprises three cell types – the pluripotent epiblast, which gives rise to the fetus, and the extra-embryonic trophectoderm (TE) and primitive endoderm (PrE, or hypoblast), which form supporting tissues. In the mouse, these lineages are specified during the 48 hours between embryonic day (E) 2.5 and E4.5, the time of implantation. Epiblast and PrE cells comprise the inner cell mass (ICM) of the blastocyst and both arise from a population of bipotent progenitors (Chazaud et al., 2006; Plusa et al., 2008). In the mouse, epiblast specification is driven by the transcription factors NANOG, SOX2 and OCT4 (Avilion et al., 2003; Chambers et al., 2003; Mitsui et al., 2003; Nichols et al., 1998). PrE specification is driven cell-autonomously by GATA6 (Bessonnard et al., 2014; Schrode et al., 2014), which requires activation of the MAPK cascade downstream of fibroblast growth factor (FGF) receptors 1 and 2, stimulated by FGF4 (Brewer et al., 2015; Chazaud et al., 2006; Kang et al., 2017; 2013; Krawchuk et al., 2013; Meng et al., 2018; Molotkov et al., 2017; Nichols et al., 2009; Yamanaka et al., 2010).

The three blastocyst lineages must be specified in appropriate numbers for development to proceed and in a short timespan (48h in the mouse, 72-96h in humans). In the mouse embryo, uncommitted ICM progenitors, which co-express NANOG and GATA6, adopt epiblast or PrE identity asynchronously and irreversibly over the course of blastocyst development (Nichols et al., 2009; Plusa et al., 2008; Saiz et al., 2016b; Xenopoulos et al., 2015). Epiblast and PrE are generated in precise proportions, irrespective of the absolute size of the embryo or the ICM (Saiz et al., 2016b), and loss of key regulators such as *Nanog*, *Gata6*, *Fgf4* or *Fgfr1* alter these proportions and cause peri-implantation lethality (Bessonnard et al., 2014; Brewer et al., 2015; Frankenberg et al., 2011; Kang et al., 2013; 2017; Krawchuk et al., 2013; Messerschmidt and

Kemler, 2010; Mitsui et al., 2003; Molotkov et al., 2017; Schrode et al., 2014; Silva et al., 2009). The ratio of these lineages is likely critical for development of the embryo beyond implantation and therefore, ICM composition must be precisely regulated (Saiz et al., 2016b). However, details of such a tissue size control mechanism remained an open question.

In this study, we address this fundamental question by combining manipulations of ICM composition with predictions from *in silico* simulations. We develop a minimal mathematical model in which cell fate decisions in the ICM are mediated solely by intercellular signaling. In this model, the ICM spontaneously and robustly segregates into two lineages, which scale with embryo size, as they do *in vivo*. We optimize two-photon laser excitation for ablation of single cells and demonstrate efficient elimination of specific cell types in live blastocysts. We demonstrate that reducing lineage size (i.e., cell number) using laser ablation, or increasing it by adding supernumerary cells, changes progenitor differentiation to restore lineage composition, both *in vivo* and *in silico*. The ability to recover from these perturbations is reduced as uncommitted progenitors are depleted over time. Finally, we alter the rate of PrE specification by experimentally tuning the size of the epiblast compartment in an FGF4-depleted context. Using this system, we show that FGF4 is the growth factor providing the feedback necessary to couple lineage size with cell fate decisions. Our results provide a basis for the regulative abilities of the early mammalian embryo and illustrate how self-organizing systems can develop robustly and reproducibly without the need for external inputs.

Results

Cell fate decisions in the inner cell mass of the blastocyst are made at the population level

Epiblast and PrE cells originate from a population of bipotent progenitor cells, which co-express markers for both lineages (Chazaud et al., 2006; Plusa et al., 2008; Saiz et al., 2016b), and which we refer to as double positive (DP) (Fig. S1a, b). Specified epiblast and PrE cells can in turn be identified by the exclusive expression of the markers NANOG and GATA6, respectively (Chazaud et al., 2006; Plusa et al., 2008) (Fig. S1c). We previously established a method to automatically classify blastocyst cell types and quantify population size based on protein expression (Saiz et al., 2016b; 2016a). PrE cells identified this way express later markers, such as SOX17 and GATA4, in a pattern consistent with previous observations (Fig. S1c-e) (Artus et al., 2011; Kurimoto et al., 2006; Niakan et al., 2010; Nowotschin et al., 2019; Plusa et al., 2008). At later stages, epiblast cells downregulate *Nanog* expression (Chambers et al., 2003), as they transition from a naïve to a primed state (Nichols and Smith, 2009), and thus we refer to these epiblast cells as NANOG-low (EPI-lo) (Fig. S1a-b, e).

We have previously shown that lineage size scales with embryo size to maintain a consistent ICM composition (Saiz et al., 2016b). However, tissue scaling could be the result of an active control mechanism or it could be explained with a probabilistic model, in which cell fate decisions take place independently of each other with a given chance. To determine whether these decisions are independent, we designed a biological probability test in which we mixed labelled (GFP+) with unlabeled cells (GFP-) from 8-cell stage embryos to generate series of chimeric embryos. If these lineage decisions are truly independent events, the differentiation pattern of the progeny of either cell population (GFP+ or GFP-) should be unaffected by the pattern of the other one. Alternatively, if they are not, the probability of any cell adopting a particular fate will be conditional on the fate choice of another.

We first mixed labelled and unlabeled wild type cells (Fig. 1a; Fig. S2a), which have unrestricted differentiation potential. In this case, the progeny of each population contributed proportionally to all embryo compartments (Fig. S2b-c) and to all ICM cell types, as they would in an intact embryo, irrespective of their representation in the resulting chimera (Fig. 1b-c; Fig. S2d-f). Chimeras resulting from aggregating two intact 8-cell stage embryos (2x size) showed the same behavior as those made with two half embryos (2x quartets, 1x size) (Fig. S2a-e).

Next, we fixed the probability of differentiation of the GFP- population by restricting its fate choice. To this end, we repeated the above experiment, using *Gata6*$^{-/-}$ cells instead of wild type (wt) cells as the GFP- population (Fig. 1d) and monitored the differentiation pattern of the GFP+ wt population. *Gata6*$^{-/-}$ embryos are cell-autonomously unable to specify PrE cells but can make TE and epiblast (Fig. 1e) (Bessonnard et al., 2014; Schrode et al., 2014). Accordingly, *Gata6*$^{-/-}$ cells combined with wt, GFP+ cells (Fig. 1d) give rise to morphologically normal chimeras (Fig. S2g-i) in which *Gata6*$^{-/-}$ ICM cells only make epiblast (Fig. 1e-f; Fig. S2j-k). In these chimeras, the differentiation pattern of the wt compartment (GFP+) changes: these wt

cells now become biased towards the PrE (Fig. 1f; Fig. S2k), despite having unrestricted differentiation potential. Moreover, when we stratified chimeras by the final contribution of each subpopulation (GFP+, GFP-) to the ICM, we observed that the fate of wt ICM cells depends on the amount of *Gata6*$^{-/-}$ cells present: in chimeras with 60% or more mutant cells (thus ≤40% wt, GFP+ cells), wt cells contribute almost exclusively to the PrE, whereas in chimeras with fewer mutant cells (≤40%), wt cells contribute to both the epiblast and PrE and generate ICMs with a normal composition (Fig. 1f; Fig. S2k-l).

Therefore, the chance of wt ICM cells adopting epiblast or PrE fate is conditional on the fate choice made by other ICM cells (Fig. S2m-n). These data demonstrate that lineage decisions in the ICM are not independent events, but rather are made at the population level, and suggest that intercellular communication coordinates cell behaviors to ensure appropriate ICM composition.

A minimal model of cell fate decisions solely mediated by growth factor signaling explains robust lineage specification in the ICM

Previously published models of the gene regulatory network (GRN) regulating cell fate decisions in the blastocyst (Bessonnard et al., 2014; De Mot et al., 2016; Tosenberger et al., 2017) combine two modules: an intracellular switch that takes the form of a direct mutual inhibition (Huang et al., 2007) and an intercellular signaling module that couples the levels of NANOG and GATA6 via FGF4-MAPK-ERK signaling in neighboring cells. However, the complexity of these models prevents discerning the relative contributions of each module to the decision. Cell fate choice in the ICM is non-cell autonomous (Fig. 1) and the PrE:epiblast ratio is maintained and refined during blastocyst development (Fig. 2a, b; S3a, b). Therefore, intercellular signaling on its own could be able to generate a robust cell fate decision.

To test this hypothesis, we have established a minimal model (see Box 1) in which a cell fate switch depends on the levels of a growth factor (GF), under the explicit control of a lineage-specific transcription factor (TF). In the context of the blastocyst, FGF4 and NANOG fit those categories (Fig. 2c and Supplementary Text), although this model is generalizable to any other system with similar characteristics.

When we applied this biochemical model to an agent-based model to simulate the growth of the ICM (see Supplementary Text), it recapitulated the lineage dynamics observed in embryos (Fig. 2d, e; Movie S1). Furthermore, when we altered the absolute cell number of this *in silico* ICM, it readily scaled lineage size to maintain ICM composition (Fig. 2f) in agreement with previous experimental observations (Saiz et al., 2016b). These data suggest that growth factor-mediated feedback could be sufficient to endow the embryo with robustness to perturbations.

The lineage composition of the ICM is robust to expansion of the epiblast

To probe the capacity of the system to perceive and adjust for changes in tissue size, we expanded the epiblast by introducing mouse embryonic stem cells (mESCs) into embryos. ESCs are derived from the epiblast and contribute to this lineage when re-introduced into an embryo (Beddington and Robertson, 1989; Boroviak et al., 2014; Brook and Gardner, 1997). In our embryo <-> embryo chimeras (Fig. 1d-f), where there is no net increase in embryo size, expansion of the epiblast-biased *Gata6*$^{-/-}$ (GFP-) compartment implied a reduction in the size of the wt (GFP+) compartment. This reduction in numbers limited the ability of wt cells to adjust to expansion of the epiblast (Fig. S2l). The use of ESCs, allowed us to specifically expand the epiblast without altering the number of host cells.

We introduced increasing amounts of ESCs into 8-16 cell embryos, which were allowed to develop in culture for 48h, until the late blastocyst stage (Fig. 3a). To control for variation in the contribution of ESCs to the chimeras, we stratified embryos based on the final size of the ESC compartment, relative to the average size of the epiblast in controls (1x control EPI = 5-10 cells, 2xEPI = 10-20 cells, etc) (Fig. S4a). This way we generated chimeras with increasingly larger epiblasts and asked how ICM composition was affected (Fig. 3b; S4b; Movies S2-3).

Increasing the number of ESCs in chimeras increased the overall size of the ICM (host + ESCs) (Fig. 3c). However, the contribution of the host embryo to the ICM remained unchanged, irrespective of the amount of ESCs present (Fig. 3d), suggesting that ESCs have no net effect on the survival or proliferation of host cells, nor in the first cell fate decision, between TE and ICM. By contrast, increasing the number of ESCs in chimeras reduced the contribution of host cells to the epiblast (Fig. 3e; S4c-e) and increased their contribution to the PrE (Fig. 3f; S4e). This shift resulted in maintenance of the ICM composition in chimeras comprising as many ESCs as 2-4xEPI (Fig. 3e). In chimeras hosting an equivalent of 4xEPI, the ICM composition in most cases was comparable to that of *Gata6*$^{+/-}$ or *Fgf4*$^{+/-}$ blastocysts (40% PrE, 60% EPI, Fig. 3e), which develop into fertile adults (Kang et al., 2013; Krawchuk et al., 2013; Schrode et al., 2014). Increasing epiblast size further, up to 8xEPI, progressively shifts ICM composition away from the wild type ratio, which is presumably not compatible with development to term. In these embryos, most host ICM cells adopted a PrE fate (Fig. 3e; S4e), but the population did not grow any further, indicating there is no compensatory proliferation in this context and that there is an upper limit to the regulative capacity of this system.

We next performed *in silico* simulations of these chimera experiments using our minimal model (Fig. 2). Increasing amounts of epiblast-equivalent cells were added to embryos before activation of the molecular circuit (Fig. 3g, arrows). Similar to our experimental results, increasing the amount of ESCs reduced the fraction of host cells contributing to epiblast (Fig. 3g). ESCs effectively replaced the epiblast compartment, to maintain the overall ICM composition for a wide range of epiblast sizes, as observed experimentally (Fig. 3h, e). Overall, these data underscore the robustness of the system and further demonstrate a population-level control of cell fate choice.

The lineage composition of the ICM is robust to in silico *reduction of lineage size*

As blastocysts develop, the number of uncommitted ICM progenitors (DP cells) decreases (Morgani et al., 2018b; Plusa et al., 2008; Saiz et al., 2016b; Schrode et al., 2014) (Fig. S1a-b). These cells give rise to both epiblast and PrE and thus, their rate of specification toward either lineage represents a candidate variable to control ICM composition. To test this hypothesis, we used our model to modify ICM composition at sequential time points and assessed the response of the system to the perturbation.

We first eliminated 30% of the cells of all three ICM lineages when each represents ~1/3 of the ICM (Fig. 4a, b). This perturbation does not alter the relative composition of the ICM and is equivalent to scaling down the absolute size of the ICM by 30%. As in our experiments (Fig. 2f and (Saiz et al., 2016b)), the relative ICM composition at the end of the simulation was maintained (Fig. 4b). Eliminating cells in this way had no effect on the lineage composition of the ICM, irrespective of the developmental stage at which it was carried out or the magnitude of the perturbation (Fig. 4c).

Next we eliminated PrE or epiblast biased cells only, therefore causing an acute deviation in the normal ICM ratio. When we removed 100% of either lineage at the same developmental stage as above, we observed only a partial recovery of the targeted lineage (Fig. 4d- g). However, a partial reduction (-30%) in either lineage was completely compensated for to restore normal ICM composition at any developmental stage (Fig. 4h, i). By contrast, the ability to recover from loss of 100% of a lineage was reduced over time, as progenitor cells are lost (Fig. 4h, i). These results suggest that the ability of the system to recover from cell loss and control ICM composition depends on both uncommitted progenitor cells and the magnitude of the perturbation.

Laser ablation enables alteration of lineage size with high spatiotemporal control in mouse embryos

We have altered lineage size to different degrees by introducing additional epiblast cells into morula-stage embryos (Fig. 3). However, these chimeras do not afford the precise temporal control we can achieve *in silico* (Fig. 4), even if ESCs are injected into blastocysts (not shown). We therefore sought a tool to alter lineage composition with high spatial and temporal control. Laser cell ablation is routinely used in non-mammalian systems to eliminate cells in a non-invasive way. However, while it has been previously applied to disrupt tissues in mammalian embryos (Eiraku et al., 2011; Reupke et al., 2009; Takaoka et al., 2017), its potential to target single cells at will and its effect on developmental competence has not been determined. Here, we implement and validate this non-invasive approach for the precise alteration of tissue size in mammalian blastocysts.

To identify PrE cells, we used a nuclear reporter for *Pdgfra* expression (*Pdgfra*$^{H2B\text{-}GFP/+}$ (Hamilton et al., 2003; Plusa et al., 2008)) (Figure S5a). To visualize epiblast cells, we generated dual reporter embryos carrying both the PrE reporter and a spectrally-distinct, ubiquitous nuclear mKate2 reporter (Susaki et al., 2014), in which epiblast cells are labelled by nuclear mKate2, but not GFP (Fig. S5a-b). PrE and epiblast cells (corresponding to high and low *Pdgfra*, respectively, Fig. S5b) were marked by the exclusive expression of GATA6 or NANOG in equivalent fixed samples (Fig. S5c-d). Moreover, prospective progenitor cells could be identified as those expressing mKate2 and intermediate levels of GFP (Fig. S5a-d), thus allowing the simultaneous identification of all three ICM cell types (Fig. S5b, d-e). Using time-lapse imaging, we confirmed that cells classified as PrE maintained or upregulated *Pdgfra* expression as they developed, whereas epiblast cells maintained low levels (Fig. S5f, see Methods for details on cell classification). On the other hand, cells classified as progenitors based on GFP levels, upregulated or downregulated *Pdgfra* expression over time, as they adopted PrE or epiblast identity, respectively (Fig. S5f).

We established conditions to target individual PrE nuclei using an 800 nm, two-photon laser, without damaging the surrounding tissue (Fig. 5a, b). The use of a two-photon laser enabled the precise targeting of cells on any z-plane within the ICM without affecting other cells in the light path. We used an equivalent experimental design to target epiblast nuclei, labelled with mKate2 but not GFP (Fig. 5c-d). Embryos expressing only mKate2 were used as either intact controls or as random controls, in which we targeted ICM cells without knowledge of their identity. We monitored cell behavior during the 20 hours following ablation using time lapse microscopy (Fig. 5e; Movie S4) to confirm the death of targeted cells, which had a half-life of ~3h (Fig. 5f). Critically, untargeted cells in experimental embryos showed identical survival to ICM cells in intact control embryos (Fig. 5f), including cells immediately adjacent to targeted cells (Fig. 5g).

Using this approach, we could eliminate any fraction of either cell type without compromising development to the late blastocyst stage (Fig. 5h). The gross morphology and size of experimental embryos was comparable to that of intact controls, even after targeting 100% of either lineage (Fig. 5h; S6a; Movies S5-6). Nonetheless, cell ablation resulted in a reduction in total ICM size (Fig. S6b), indicating that embryos have a limited ability to compensate for changes in absolute cell numbers at this stage. In summary, this method leverages existing reagents and technology to provide a simple, non-invasive way to eliminate individual cells with high spatial resolution at any point of blastocyst development. This microsurgical technique could be readily applicable to other systems amenable to optical imaging.

The cell fate choice of uncommitted ICM progenitors is dictated by lineage size

To determine the role of ICM progenitors in the regulative capacity of the embryo, we used laser ablation to specifically eliminate PrE or epiblast cells (Fig. 5a, c) at sequential stages of blastocyst

development, as progenitors are depleted (Fig. 6a) and as we did *in silico* (Fig. 4c, h, i). After a recovery period of 24h, embryos contained both PrE and epiblast cells, even if all cells of either lineage had been eliminated (Fig. 6b). This result agreed with the prediction of our model (Fig. 4f, g) and indicated that embryos could, at least partly, regenerate the targeted population.

To characterize this regenerative ability, we targeted increasing fractions (50-100%) of the PrE at sequential blastocyst stages (Fig. 6a) and assessed the composition of the ICM after they were allowed to recover in culture for 24h (Fig. 6a, b). The ability of embryos to re-establish a normal PrE:epiblast ratio was reduced the later in development the perturbations were carried out (Fig. 6c; S6c). As in our simulations (Fig. 4h), this reduction in the ratio was more pronounced the larger the fraction of PrE cells eliminated (Fig. 6c; S6c). The opposite experiment, in which we eliminated different amounts of epiblast cells, had a similar effect in the opposite direction (Fig. 4i; 6d; S6c). This finding indicates that the ability of the system to respond to changes in tissue size depends on the presence of uncommitted progenitor cells.

We therefore assessed the response of progenitor cells to changes in ICM composition. We tracked individual cells in time-lapse movies over the first 15h hours following cell ablation. Progenitor cells in intact embryos contribute to both epiblast and PrE, as assessed by *Pdgfra* expression (Fig. S5f; 6e). We found a bias of these cells towards the PrE in intact embryos (Fig.6e). This is likely due to oversampling of the PrE, given the presence of both the *Pdgfra* and the ubiquitous mKate2 reporter, as well as to the overrepresentation of PrE cells in the ICM. Progenitor cells in embryos where the PrE was targeted showed a comparable trend to control embryos, generally upregulating *Pdgfra* expression and acquiring PrE identity (Fig. 5g, purple line). Conversely, in embryos where epiblast cells were targeted, progenitors had a propensity to downregulate or maintain low levels of *Pdgfra* and become epiblast (Fig. 6e). Accordingly, at the end of the movies, progenitor cells were found more frequently in the lineage that had been targeted (Fig. 6f). When we assessed the final PrE:epiblast ratio in movies where we could track all or most of the ICM cells, we saw a comparable result to that obtained from fixed data (Fig. 6g and c-d; Fig. S6c-d).

Intriguingly, in embryos in which we targeted 100% of the PrE, progenitor cells had a propensity to downregulate *Pdgfra*, similar to embryos in which the epiblast had been targeted (Fig. 6e). We therefore assessed the cell behaviors that might account for this result. In both intact controls and in embryos where 50-75% of the PrE was targeted, we observed cell proliferation and generation of new cells from progenitors in both the PrE and epiblast compartments (Fig. S7a-b). However, in embryos where we targeted 100% of the PrE, we found a high degree of cell death throughout the movie that was not compensated for by proliferation or specification of new PrE cells (Fig. S7b). This was due to the delayed death of targeted cells in this subset of embryos, which likely phenocopied the result of cell targeting at a later developmental stage, when embryos show a decreased capacity to recover (Fig. 6c; S6c).

We then investigated the relative contribution of cell death and cell proliferation to recovery of ICM composition from cell ablation. Targeting either PrE or epiblast increased the

survival of progenitor cells but had no clear effect on neither PrE nor epiblast cells (Fig. S7f). Additionally, we found a slight increase in proliferation among progenitor cells following PrE ablation, but not epiblast ablation, which instead increased proliferation of epiblast cells (Fig. S7g). Despite a relatively low sample size, overall, these data suggest that a combination of both survival and proliferation might help balance the size of each compartment.

Reduction in tissue size biases the overall contribution of progenitor cells towards the targeted tissue, which accounts for the ability to recover from moderate shifts in ICM composition. This behavior is consistent with that found in chimeras (Fig. 3e), in which expansion of one lineage biased cells away from it. These results agree with the prediction of our mathematical model and support the growth factor-mediated coupling of tissue size to progenitor behavior.

FGF4 provides the dynamic readout of lineage size that determines cell fate specification

We have shown here that the lineage composition of the ICM is robust to changes in absolute tissue size, both *in vivo* and *in silico* and propose that growth factor-mediated feedback, ensures this robustness (Fig. 2c). Fibroblast growth factor 4 (FGF4), secreted by ICM cells, is necessary for PrE specification (Kang et al., 2013; Krawchuk et al., 2013), making it an attractive candidate for a feedback signal. To test this hypothesis, we used *Fgf4*$^{-/-}$ embryos – which lack the ligand and thus cannot make PrE (Movie S7) – and introduced increasing amounts of wild type (wt) ESCs, which act as a localized source of FGF4 and a proxy for wt epiblast cells (Fig. 7a). Increasing amounts of ESCs are equivalent to a growing epiblast in wt embryos and would allow us to recapitulate and control the process of PrE specification.

We found that wt ESCs can induce PrE specification in mutant embryos (Fig. 7b; S8a; Movie S8). Conversely, *Fgf4*$^{-/-}$ ESCs gave rise to chimeras without PrE, which phenocopied *Fgf4* null embryos (Fig. 7b; S8a; Movie S9), thus demonstrating FGF4 is the ligand responsible for the rescue. Wt ESCs contributed to chimeras with *Fgf4*$^{-/-}$ and wt hosts comparably (Fig. S8b; Fig. S4a) and, although we observed more variation in the size of the host-derived ICM compartment in wt ESC <-> *Fgf4*$^{-/-}$ embryo chimeras than in wt ESC <-> wt embryo chimeras, this was not related to the number of ESCs present (Fig. S8c; Fig. 3d). More importantly, the number of wt ESCs present in chimeras dictated the size of the PrE (Fig. 7c-d). Chimeras in which the ESCs were present in equivalent numbers to the epiblast in wt embryos (1xEPI) show a full rescue of absolute PrE size (Fig. 7c). Both the absolute and relative size of the PrE in these chimeras depended on the size of the ESC compartment (Fig. 7c-d), with chimeras frequently showing a full rescue of the wt ICM composition (Fig. 7d, dashed line). Of note, chimeras with a high number of ESCs (4x-8xEPI) showed a deviation from the wt ICM composition comparable to that observed in wt ESC <-> wt embryo chimeras (Fig. 7d and 3e; S8d-e and S4c-d). In both cases, although most of the host ICM cells adopted a PrE fate, they did not proliferate to compensate for the large excess of ESCs, and thus ICM composition is disrupted with very high numbers of

ESCs (Fig. 7d; 3e).

Lastly, we introduced wild type ESCs into *Fgf4*$^{-/-}$ embryos at the blastocyst stage to test the temporal variable of the recovery (Fig. S8g). Early *Fgf4*$^{-/-}$ blastocysts show co-expression of GATA6 and NANOG in the ICM, but gradually lose GATA6 due to the absence of FGF4 and consequently all ICM cells adopt an epiblast fate (Kang et al., 2013; Ohnishi et al., 2014). Accordingly, we found that some blastocysts at the time of injection showed DP cells, while others only had epiblast cells in the ICM (Fig. S8h). Consequently, addition of ESCs resulted in chimeras that mostly resembled uninjected controls, with only some chimeras specifying low numbers of PrE cells (Fig. S8h-i). Interestingly, ESCs in these chimeras failed to mix with host ICM cells and tended to remain at the ICM surface, suggesting that differences in adhesion between cell types are necessary for cell mixing within the ICM (Movie S10-11).

Taken together, these data demonstrate that FGF4 provides the feedback that couples lineage size with the lineage specification of progenitor cells, effectively acting as the cell-counting mechanism. Such feedback control enables the system to dynamically adjust the differentiation rate of progenitors in response to perturbations to ensure a robust and consistent developmental outcome.

Discussion

In multicellular organisms, both local and systemic control are necessary to guide embryonic development and maintain homeostasis (Kicheva et al., 2014; Roselló-Díez et al., 2018). Local control of the patterning and morphogenesis of developmental modules underlies their ability for autonomous development, demonstrated by the growth of organoids *in vitro* (Boj et al., 2015; Eiraku et al., 2011; Sato et al., 2009). The early mammalian embryo is one such self-organizing system in which patterning and morphogenesis can take place in the absence of maternal control (Bedzhov and Zernicka-Goetz, 2014; Deglincerti et al., 2016; Shahbazi et al., 2016). Despite insights provided by stem cell-based systems (Beccari et al., 2018; Harrison et al., 2017; Rivron et al., 2018; Sozen et al., 2018), the local control mechanisms that enable tissue scaling and reproducible patterning remain poorly understood. In this study we have established that patterning in the ICM of the mouse blastocyst is a phenomenon controlled at the population level. We demonstrate both theoretically and empirically that growth factor-mediated feedback couples progenitor fate decisions and lineage size, thus enabling the system to scale and compensate for experimental changes in lineage size. We show that the probability of epiblast or PrE specification is dynamic and determined by the size of each compartment, reflected in the amount of growth factor available in the immediate neighborhood of uncommitted progenitors. Our results uncover a mechanism for local control of patterning through FGF4-mediated coupling of lineage size and fate specification.

Previous models of cell fate specification in the blastocyst incorporate direct mutual inhibition between NANOG and GATA6 (Bessonnard et al., 2014; Nissen et al., 2017; Schröter et al., 2015; Tosenberger et al., 2017). This configuration was inspired by pioneering studies modelling cell-autonomous fate decisions (Huang et al., 2007). However, in this non cell-autonomous process, such configuration becomes unnecessary. In our minimal model, indirect mutual inhibition via growth factor-mediated signaling is sufficient to robustly partition the ICM into two cell types, explain the scaling of tissue size with embryo size and to predict the outcome of expanding or reducing ICM lineages. Our model requires fitting fewer parameters, involves fewer assumptions and has a more predictable behavior than previous models. In this model, cell mixing and short-range communication between cells are sufficient to explain the observed salt-and-pepper distribution of cell types in the ICM – a behavior resembling Notch-mediated lateral inhibition. Given that lineage specification in mouse and rat embryos is entirely dependent on FGF-RTK signaling activity, we cannot empirically disprove a direct interaction between NANOG and GATA6 *in vivo*. However, our results suggest that if occurring, it must have merely an accessory or redundant role in rodents. The constrains of smaller size and faster development in rodents compared to larger mammals could have driven the implementation of feedback as a failsafe mechanism during evolution.

Preimplantation embryos face two main challenges: to make enough cells of each type and to do so before implantation into the uterus. Blastocyst lineages scale with embryo size (Papaioannou and Ebert, 1995; Saiz et al., 2016b), and perturbations in absolute cell numbers are compatible with development to term (Papaioannou et al., 1989; Tarkowski, 1959; 1961),

suggesting the variable under control at preimplantation stages is relative, not absolute lineage size. Feedback is widely used in complex systems (such as multicellular organisms) to control noise and ensure robust behavior (Lander, 2011). Therefore, in the mammalian blastocyst, feedback control enables the coupling of cell fate decisions to changes in lineage size over time. This quorum sensing-like behavior would ensure the appropriate lineage composition for development beyond implantation.

To test the robustness of this system, we devised experiments to modify ICM composition at will. We established a protocol for direct elimination of ICM cells using two-photon laser excitation and we systematically validated its application to precisely eliminate individual cells of either ICM lineage with high spatiotemporal control and without damaging neighboring cells. We combined this technology with ESC chimeras and *in silico* simulations to modify ICM composition in embryos and determine the role of uncommitted progenitors in the adaptation of the system. We established that the amount of progenitors present dictates the ability of the population to recover from changes in lineage composition. Using time-lapse imaging we determined that their fate choice counters the experimental perturbation, so as to maintain ICM composition. These data also suggest that cell cycle changes play a minor role, if any, in this process. We did observe an increase in progenitor survival compared to controls as a result of cell ablation, perhaps indicating progenitor cells in intact controls are present in excess and subject to selection. However, we failed to observe any clear change in proliferation to compensate for cell loss in either lineage.

Given the consistency between the predictions of our minimal model and our experimental results, direct mutual inhibition between NANOG and GATA6 is unlikely to drive cell fate specification in the mouse blastocyst. Instead, we assume that FGF4 is both necessary and sufficient for this binary decision. Epiblast cells (and to a lower extent progenitor cells) are the main source of FGF4 (Frankenberg et al., 2011; Guo et al., 2010; Nowotschin et al., 2019; Ohnishi et al., 2014). Activation of the RKT-MAPK axis induces PrE fate, while low or no FGF signaling promotes epiblast fate (Chazaud et al., 2006; Kang et al., 2013; Krawchuk et al., 2013; Nichols et al., 2009; Ohnishi et al., 2014). In this small and well mixed system, no signaling gradient forms, but local differences in exposure to FGF4 suffice to drive the cell fate decision, as previously proposed (Bessonnard et al., 2014; Tosenberger et al., 2017).

To directly test whether changes in FGF4 concentration reflect changes in lineage size, we established an experimental paradigm to control PrE formation in the embryo by controlling the size of the population secreting FGF4. *Fgf4*$^{-/-}$ embryos have progenitor cells in the ICM, which default to epiblast in the absence of the ligand (Kang et al., 2013; Krawchuk et al., 2013). Treating these embryos from the morula stage with saturating amounts of exogenous FGF4 converts all ICM cells to PrE (Kang et al., 2013; Krawchuk et al., 2013), as it does in wild type embryos (Saiz et al., 2016b; Yamanaka et al., 2010). In contrast, we titrated the size of the source of FGF4 (the ESC population), to recapitulate local heterogeneities in signal availability. As a result, we could modulate the number of PrE cells specified to ultimately phenocopy equivalent chimeras made with wild type host embryos (compare Fig. 3e and 7d). Therefore,

FGF4 availability both reflects each lineage's relative size and couples it to the chance of progenitor differentiation towards either cell type. These results demonstrate that FGF4 is not only a necessary factor for PrE specification, but that it provides the dynamic readout of lineage size necessary to coordinate cell behaviors across the population and control ICM composition.

Our study uncovers how cell fate choice and lineage size are coupled via growth factor signaling to ensure robust patterning and morphogenesis during early mammalian development. These findings provide a framework for our current understanding of signaling and cell fate decisions in the early mammalian embryo and may be generalizable to the formation of other self-organizing, autonomous developmental units.

Materials and Methods

Mouse strains and husbandry

All animal work was approved by Memorial Sloan Kettering Cancer Center's Institutional Animal Care and Use Committee. Animals were housed in a specific pathogen-free facility under a 12h light cycle. Embryos were obtained from natural matings of 4-12 week old females to stud males. The alleles used and their original source are summarized in Table 1. Specific combinations of alleles used are indicated in the text and figures. All alleles were outbred onto CD1 for one or more generations and maintained on a mixed background.

Embryo staging

Embryonic days (E) are determined by considering noon of the day of detection of the copulation vaginal plug as E0.5, and thus, the preceding midnight as the consensus time of fertilization. It is worth noting, however, that the exact time of fertilization will inevitably vary between litters as it may happen at any point throughout the 12h dark cycle. It is for this reason that we and others use the total cell number as a more reliable way to stage embryos, assuming invariance in the average length of each cell cycle from embryo to embryo. This can be achieved either directly, by counting all cells in embryos expressing a ubiquitous fluorescent reporter, or retrospectively, after fixing the sample and staining all nuclei. To determine the initial developmental stage in experiments involving in vitro culture of blastocysts without a ubiquitous nuclear reporter, we fixed a subset of embryos from each litter (2-4 embryos) at the time of collection, as a reference for the developmental stage of the entire litter (referred to as Littermates in data tables). Therefore, we use embryonic days at the bench, as an approximate guide for the time to collect embryos, but ultimately classify them retrospectively based on their cell number. Importantly, in all experiments, individual litters were treated as experimental units to reduce variability due to stage differences between litters and to control for batch effects. Finally, it is well established that embryos develop slower *in vitro* than they do *in vivo* (Smith and McLaren, 1977). Therefore, we normally allow cultured embryos to develop as late as possible on the final day of culture, so that they reach the late blastocyst stage and use landmarks such as resolution of the PrE layer (Plusa et al., 2008) and loss of cells co-expressing PrE and epiblast markers, in addition to total cell count, as indicators of developmental stage.

Embryo recovery and handling

Embryos were recovered from dissected oviducts (up to the 8-16 cell stage) or uterine horns (blastocyst stages) by forcing Flushing and Holding Media (FHM, Millipore) through them, as previously described (Behringer et al., 2014). Live embryos were manipulated in FHM. All solutions used to handle or recover live embryos were pre-warmed at 37ºC. Whenever not being handled under the microscope, embryos were kept at 37ºC in a bench top, dry incubator box. When necessary, zona pellucidae were removed from morulae or blastocysts by brief washes in

Acid Tyrode's solution (Millipore) (Behringer et al., 2014) and returned to FHM as soon as the zona dissolved. Blastocysts were fixed in a solution of 4% paraformaldehyde (PFA, Bio-Rad) in phosphate buffered saline (PBS) for 10 minutes at room temperature and preserved in PBS at 4ºC.

Embryo genotyping

Individual embryos from mutant backgrounds (chimeric or else) were lysed for genotyping in 10μl of lysis buffer (10mM Tris, pH 8.5, 50mM KCl, 0.01% gelatin and 300μg/ml of Proteinase K) for 50 minutes at 50ºC, followed by 10 minutes at 90ºC to inactivate the Proteinase K (Artus et al., 2005). 2μl of the lysate were then used for each genotyping reaction using the primers indicated in Table 2 and the Clontech Advantage 2 Taq Polymerase kit.

Embryo culture and live imaging

Embryos at all stages were cultured on 35mm Petri dishes (Falcon) within microdrops of 10-15μl of Potassium Simplex Optimized Media with amino acids (KSOM-AA, Millipore), at 37ºC, in a humidified 5% CO2 atmosphere. KSOM-AA drops were always covered with a layer of mineral oil (Sigma) to prevent desiccation. Prior to culture, embryos were always rinsed 3x in drops of KSOM-AA. For live imaging, embryos were cultured on 35mm glass-bottom dishes (MatTek) within 5-10μl drops of KSOM-AA.

Images of live embryos were acquired using Zeiss LSM880 laser-scanning confocal microscopes. In ablation experiments, images were acquired using a water-immersion Zeiss C-Apochromat 40x/NA1.1 objective with a 0.21mm working distance. For all other experiments, images were acquired through an oil-immersion Zeiss EC Plan-Neofluar 40x/NA1.3 with a 0.17mm working distance. GFP was excited using a 488nm Argon laser at 20μW of power. mKate2 was excited using a 543 HeNe laser at 90μW. Laser power was measured through a Zeiss Plan-Neofluar 10x/NA0.3 objective prior to each imaging session with a light meter (Coherent) and the laser output adjusted to match laser power across experiments. Unless indicated otherwise, 80μm stacks were acquired through embryos, with a 2μm step, every 15 min. Although these stacks do not capture an entire blastocyst, they can typically capture the whole ICM while limiting laser exposure to about 30-40s per time point and embryo. Time lapse movies had a length of 16 to 20h.

Cell culture

Embryonic Stem Cells (ESCs) used in this study were *CAG:H2B-EGFP*$^{Tg/+}$ (Hadjantonakis and Papaioannou, 2004) and *Hex*$^{Venus/+}$; *CAG:H2B-tdTomato*$^{Tg/+}$ (Morgani et al., 2013), which, for simplicity, we refer to as CAG:H2B-tdTomato throughout the text. ESCs were cultured without feeder cells, in 0.1% gelatin-coated tissue culture grade dishes (Falcon). For long term culture we

used 1x DMEM supplemented with 2mM L-glutamine, 0.1mM MEM non-essential amino acids (NEAA), 1mM sodium pyruvate, 100 U/ml penicillin, 100μg/ml streptomycin (all from Life Technologies), 0.1mM 2-mercaptoethanol (Sigma), 10% Fetal Bovine Serum (Hyclone) and 1000 U/ml of recombinant leukemia inhibitory factor (LIF). To passage cells, they were washed with PBS and detached by brief incubation at 37° C in Trypsin-EDTA (0.25% or 0.05%, Life Technologies), which was neutralized with Serum/LIF media. Cells were precipitated by centrifugation and resuspended in the desired culture media before replating.

Embryo-embryo aggregation chimeras

Embryos were isolated in the morning of the third day of development (E2.5) and those with 4 or less cells were discarded. Morulae (8-16 cells) from CD1 x *CAG:H2B-EGFP*$^{Tg/+}$ (Hadjantonakis and Papaioannou, 2004) crosses were sorted based on the presence of GFP. *CAG:H2B-EGFP*$^{Tg/+}$ embryos were used in the generation of both wt;GFP+ and *Gata6*$^{-/-}$;GFP+ chimeras (Figure 1b), whereas wild type morulae were either used for wt;GFP+ chimeras or cultured as un-manipulated controls. For *Gata6*$^{-/-}$;GFP+ chimeras, the *Gata6*$^{-/-}$ component was generated as indicated in Figure 1b (see Table 1 for references on the alleles used). The use of the *Gdf9*^{iCre} (Lan et al., 2004) allele and the floxed *Gata6* allele (Sodhi et al., 2006) gave greater numbers (>25%) of *Gata6*$^{-/-}$ embryos than would normally result from standard heterozygous crosses.

After removal of the zona pellucida, uncompacted, 8-cell stage embryos, were dissociated individually by gently aspirating them through a mouth-controlled, finely drawn glass capillary (Sutter Instruments), which had been manually broken to the appropriate opening diameter (Tarkowski and Wróblewska, 1967). Whenever possible, embryos were dissociated into doublets or quartets to minimize cell lysis and facilitate re-agreggation. While uncompacted embryos readily fall apart upon pipetting, compacted 8-cell or 8-to-16-cell stage morulae were incubated for 10 to 15 minutes in Ca^{2+}/Mg^{2+}-free KSOM-AA to induce de-compaction (home-made as per Behringer et al., 2014), then dissociated as described above.

Dissociated blastomeres were placed in the desired proportions (Figure 1b) in close proximity within indentations made on the bottom of 35-mm Petri dishes (Falcon) using an aggregation needle (BLS, Hungary), under a drop of KSOM-AA (see Behringer et al., 2014; Eakin and Hadjantonakis, 2006 for a more detailed description of the aggregation chamber). These concave indentations, or depressions, keep blastomeres in contact to allow re-aggregation, while keeping each chimera apart from one another - thus, multiple chimeras can be cultured in a single KSOM microdrop, one in each indentation. To generate double embryo chimeras, morulae were not dissociated and instead two whole denuded morulae of the desired genotypes were placed next to each other within a depression.

mz*Gata6*$^{-/-}$ embryos generated as described in Fig. 1b were found not to be always null, and in occasion can be mosaic, which cannot be determined through genotyping. Therefore, in all *Gata6*$^{-/-}$;GFP+ chimeras the *Gata6*$^{-/-}$ complement was clonal. All *Gata6*$^{-/-}$;GFP+ chimeras were

genotyped retrospectively for the presence of both the conditional and null *Gata6* alleles, and phenotyped based on the presence/absence of GATA6 in immunofluorescence images. Unclear genotypes were excluded from the analysis. On the other hand, while sister GFP+ blastomeres were used when possible, the GFP+ complement of these chimeras was not necessarily clonal, so as to maximize the number of chimeras generated per experiment.

ESC-embryo chimeras

ESCs were cultured as described above and prepared for chimeras exactly as described in (Behringer et al., 2014). To make chimeras of ESCs with morula-stage embryos, aggregation dishes were prepared as described above for embryo-embryo chimeras. Likewise, embryos were collected at E2.5 and individual denuded morulae of the desired genotype were placed within the indentations on the culture dish. No de-compaction or dissociation is needed, although experiments were carried out in the morning to obtain embryos at the 8- or 8-16 cell stage and ensure efficient aggregation. Fluorescently-labelled ESCs were rinsed twice in KSOM-AA to dilute the FBS and LIF present in the ESC media and, subsequently, clumps of the desired number of ESCs were placed in contact with each embryo (the number of ESCs per embryo added in each experiment is indicated in the main text and figures). Aggregation chimeras thus generated (Fig. 3a) were allowed to develop in vitro, as described above, for 48-56h, until the late blastocyst stage, when they were fixed in 4% PFA. Chimeras transferred to surrogate mothers were allowed to develop in vitro for 30h before uterine transfer to pseudopregnant females at day E2.5 post mating.

When making chimeras at the blastocyst stage, embryos were collected in the morning of the fourth day of development (E3.5) and placed in FHM drops on a stage at 37° C, under an opaque lid. Chimeras were then generated using standard procedures (described elsewhere and compiled in Behringer et al., 2014). A suspension of dissociated ESCs was placed in a drop of FHM on a Petri dish, while one litter of blastocysts (5-10 embryos) was placed in an adjacent drop, all covered with mineral oil. A PrimeTech Piezo drive (Sutter Instruments) attached to an Eppendorf CellTram microinjectiong system was used to assist with injections. Individual embryos were held with a standard holding pipette (VacuTip, Eppendorf) by the ICM end, while a Piezo Drill flat tip needle (Eppendorf) loaded with ESCs was pressed against the zona pellucida at a 180° angle (Fig. S4e). The needle was pushed through both the zona and the TE while triggering a sustained pulse of the Piezo drive and 8-12 ESCs were injected into the blastocyst cavity. After injection, each litter was transferred to drops of KSOM-AA and allowed to develop for 28-31h in an incubator, as described above and in Fig. S4e. 3-4h after injection, injected embryos were denuded to allow for cavity expansion while maintaining blastocyst morphology. Reference littermates were fixed in 4% PFA right after injection of ESCs into the rest of the embryos.

Laser cell ablation

The steps of the ablation experiment and timeline are summarized in Figure 4a and c. For ablation of PrE cells, most embryos used were obtained from CD1 females intercrossed with *Pdgfra*$^{H2B\text{-}GFP/+}$ stud males (Fig. 5a). For ablation of epiblast cells (and some PrE ablation experiments) all embryos used were from intercrosses of *ROSA26:CAG:3x-nls-mKate2*$^{Tg/Tg}$ females and *Pdgfra*$^{H2B\text{-}GFP/+}$; *ROSA26:CAG:3x-nls-mKate*$^{Tg/Tg}$ males (henceforth referred to as *R26:mKate*$^{Tg/Tg}$) (Fig. 5c).

Embryos were collected at different time points between noon and late evening of the fourth day of development (~E3.5 to E4.0) to cover the entire period of PrE and epiblast specification (~30 to 110 cells). To ascertain the number of PrE or epiblast cells in each embryo, a z-stack through the whole ICM was acquired for each embryo prior to cell ablation on a Zeiss LSM880 microscope equipped with a water-immersion Zeiss C-Apochromat 40x/NA1.1 objective. PrE nuclei were visually identified as those where the GFP signal was markedly higher than their neighbors on each optical plane (Fig. 5b; Fig. S5a). Nuclei with intermediate levels of GFP were scored as putative progenitor cells, whereas nuclei with no GPF signal were deemed to belong to epiblast cells (Fig. S5b, see Fig. S6a for estimates of each population size in live embryos compared to fixed samples stained for molecular markers). In embryos expressing mKate2, stacks were acquired through the whole embryo prior to cell ablation to calculate the initial cell count of each embryo.

After this imaging step, cells were eliminated by repeatedly focusing the beams of an 800-nm Ti:Sapphire femtosecond laser at 10-12% output (Coherent), onto the central region of each nucleus to be ablated (approximately 30-50% of the nuclear area in the section). The number of pulses to use was determined empirically on a test embryo for each experiment, so that a clear wound was observed on the fluorescent nucleus, but no obvious damage was inflicted to the neighboring cells (typically between 100-250x iterations) (Fig. 5b, d). Although target nuclei were selected at random throughout the ICM, when possible, nuclei located on the same z plane were targeted, in order to minimize the length of the procedure. Intact control embryos were only subject to the initial imaging step to estimate the size of each ICM population. Random control embryos were *R26:mKate*$^{Tg/Tg}$ embryos (not carrying the *Pdgfra*$^{H2B\text{-}GFP}$ allele) where a number of ICM cells equivalent to the total number of PrE or epiblast cells found in GFP+ littermates was randomly selected and ablated using the same settings.

After ablation of the desired number of cells in all experimental embryos, all embryos, intact controls and ablated, were live imaged for 16-20h, as described above. Laser ablation generated a visible wound in nuclei expressing H2B-GFP, which allowed their tracking over the course of the movie and assessing their fragmentation as a mark of cell death. However, mKate2 was extinguished as a result of the repeated laser illumination – presumably through a combination of bleaching and cytoplasmic diffusion of the fluorophore – and targeted epiblast cells could not be tracked over time. After time-lapse imaging, embryos were allowed to develop further in an incubator until a total of 24h after the time of collection and were fixed individually in 4% PFA so that end-point immunofluorescence images could be related to time-lapse data.

Immunofluorescence

The protocol for whole-mount embryo immunofluorescence was the same described in (Saiz et al., 2016a; 2016b). Fixed embryos were washed for 5 minutes in 0.1% Triton X-100 (Sigma) in PBS (PBX), permeabilized for 5 minutes in a solution of 0.5% Triton X-100, 100μM Glycine (Sigma) in PBS, rinsed again for 5 minutes in PBX and blocked for 30 minutes to 1h in 2% horse serum (Sigma) in PBS. All these steps were carried out at room temperature. After blocking, embryos were incubated in primary antibodies diluted in blocking solution overnight at 4ºC. All primary antibodies and the dilutions used are provided in Table 3. After incubation in primary antibody, embryos were washed 3x in PBX for 5 minutes each and blocked again for 30 minutes to 1h, all at room temperature. Embryos were then incubated in secondary antibody solution for 60-75min at 4ºC, then washed 2x in PBX for 5 minutes each and placed in a 5μg/ml solution of Hoechst 33342 (Life Technologies) in PBS. All secondary antibodies were diluted 1:500 in blocking solution. All secondary antibodies were from Life Technologies, except the AlexaFluor®488 anti-chicken, which was from Jackson ImmunoResearch.

Image acquisition of fixed samples

Immunolabeled embryos were mounted on 35mm glass-bottom dishes (MatTek), within micro drops of a 5μg/ml solution of Hoechst 33342 in PBS and imaged using a Zeiss LSM880 laser-scanning confocal microscope, equipped with an oil-immersion Zeiss EC Plan-Neofluar 40x/NA1.3 with a 0.17mm free working distance. Z-stacks were acquired through whole embryos with 1μm step between optical slices. Considerations for image acquisition are discussed in detail in (Saiz et al., 2016a). Parameters were set up initially for each primary-secondary antibody combination so as to cover the full dynamic range of 12-bit images, while avoiding excessive saturation. Laser power was measured for each laser line prior to each imaging session and parameters adjusted so as to keep laser power consistent for each primary-secondary antibody combination across experiments over time. For the 405 laser line only, parameters were adjusted so as to capture bright, sharp nuclei in each experiment, regardless of laser power. The 405 channel was always used to excite the nuclear label (Hoechst 33342) and this channel was only used for image segmentation, not used in the quantification process.

Image processing

Nuclear image segmentation of still images was performed using the MINS software (Lou et al., 2014) as previously described (Saiz et al., 2016a). MINS is freely available at https://github.com/therealkatlab/MINS (requires MATLAB license). All segmentation data tables were manually curated to correct for over- and under-segmentation as described in (Morgani et al., 2018b; Saiz et al., 2016a). Missing nuclei, or multiple nuclei segmented as one (under-segmentation) were measured manually using ImageJ (Rasband, W.S., ImageJ, U. S. National

Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, 1997-2018) and the measured values for each channel, as well as XYZ coordinates, added to the data table. For nuclei segmented as multiple masks (over-segmentation), only quantification values for the object of largest volume were preserved.

Time lapse images were processed using Imaris (Oxford Instruments). Individual ICM cells were tracked manually over the course of the movie using the spots function. Cells were identified based on the presence of H2B-GFP and/or nuclear mKate2. Cellular events like cell death or mitosis were labelled as such for each individual cell (Fig. S5f). GFP levels served as a proxy for *Pdgfra* expression and allowed the classification of cell types over time (Fig. S5f). Cell identity at the start of the experiment was assigned visually based on GFP levels on z-stacks acquired before ablation (Fig. S5a, e). Cell identity was assigned to targeted cells at the time of the experiment and verified (and reclassified when necessary) after the completion of the experiment and quantification of GFP levels. The cells with highest levels of GFP relative to their neighbors were considered PrE, those with intermediate levels were deemed DP, while cells negative for GFP were considered epiblast (see Fig. S5a-b, d).

Data processing

Fluorescence data obtained after segmentation with MINS and manual curation was processed essentially as described in (Morgani et al., 2018b; Saiz et al., 2016b), although certain modifications were made for subsets of the data. For samples where GFP fluorescence was detected indirectly with an anti-GFP antibody (Fig. 1b-f; Fig. S2), anti-GFP::AF®488 fluorescence was log transformed and corrected for fluorescence decay along the Z axis by fitting a linear model, as described in (Saiz et al., 2016a). For samples stained with antibodies against transcription factors or for samples where GFP/tdTomato fluorescence was detected directly (either live or fixed samples), an Empirical Bayesian slope correction step was applied to log-transformed, Z-corrected data, as in (Saiz et al., 2016b). Detailed descriptions of the correction steps can be found in https://github.com/nestorsaiz/saiz-et-al_2020/notebooks

Two different antibodies were used to detect NANOG and GATA6 expression (Table 3). Rabbit anti-NANOG (NANOG(rb)) and goat anti-GATA6 (GATA6(gt)) were used for most samples, as well as in our previous studies (Kang et al., 2017; Morgani et al., 2018b; Saiz et al., 2016b; Schrode et al., 2014). For certain samples a rat anti-NANOG (NANOG(rat)) and a rabbit anti-GATA6 (GATA6(rb)) were used (Table3). Fluorescence values obtained with these antibodies were transformed to NANOG(rb)- and GATA6(gt)-equivalents, respectively. To do this, we generated a dataset of samples where each epitope was labelled with two antibodies (NANOG(rb) and NANOG(rat) or GATA6(gt) and GATA6(rb)) and fitted linear regression models to calculate a transforming factor.

Finally, to account for variability in fluorescence values between litters and over time when classifying ICM cells, we also re-scaled values for each channel in each litter by making

them relative to the maxima of that channel in the litter. These data were only used for ICM cell type classification, as it eliminates temporal changes in protein levels (associated with developmental stage).

Cell identity was assigned to ICM cells based on relative NANOG and GATA6 expression. When separated based on NANOG and GATA6 levels, ICM cells typically form four clusters, corresponding to the epiblast (NANOG+), PrE (GATA6+), double positive (DP, progenitor cells) and NANOG-low or NANOG- epiblast (GATA6-, NANOG-). We have previously applied k-means clustering to automatically identify cell populations in the ICM (Morgani et al., 2018b; Saiz et al., 2016b). We found that while k-means generally performs well, it tends to overestimate the size of the DP population, as well as the NANOG-low population at late blastocyst stages. To overcome this limitation, we implemented hierarchical clustering to identify cell types in the ICM. We determined empirically that the agglomerative UPGMA (unweighted pair method with arithmetic mean) captured the PrE and epiblast clusters comparably to k-means clustering across all blastocyst stages. Importantly, this method outperformed k-means when classifying DP and NANOG-low cells. Although we classified cells using both k-means and hierarchical clustering, we use the results of the hierarchical clustering classifier throughout the paper. More detailed descriptions of the transformation and classification steps can be found in https://github.com/nestorsaiz/saiz-et-al_2020/notebooks

Corrected fluorescence values of H2B-GFP (*Pdgfra* expression) for ICM cells were used to determine changes in cell identity over time in time-lapse movies. To reduce noise in the data, a moving average was calculated for each cell, with a window of 4 timeframes (1h). A simple classifier was thus devised to assign cell identity automatically to individual cells based on these GFP levels. Thresholds for fluorescence were manually determined for each litter analyzed, both at the start and the end of the movie. These thresholds were based on visual inspection of temporal GFP dynamics for each cell and informed by the cell identities assigned manually, as described above. These initial and final thresholds were used to determine a threshold slope for PrE and epiblast identities. The classifier followed two rules: (1) cells classified as PrE or epiblast maintain that identity for the remainder of the movie – based on (Saiz et al., 2016b; Xenopoulos et al., 2015) and the lack of oscillation on GFP levels or obvious, systematic shifts from high to low levels, or vice-versa (Fig. S5e, f) – (2) cells classified as DP at t=0 become PrE or epiblast once their level of GFP remains above or below the respective thresholds for at least 2h – to account for fluctuations and noise in data (Fig. S5f).

Data and code availability

All data processing was done in R version 3.4.2, using RStudio as an interactive development environment. All processed data as well as the code used to transform data and classify cells is publicly available at https://github.com/nestorsaiz/saiz-et-al_2020 and upon request. All raw confocal images and data tables are freely available on Figshare with DOI 10.6084/m9.figshare.c.4736507.

Box 1. A minimal model of growth-factor-mediated cell-fate decision

We use the following model to represent the dynamics of a population of cells in which the transcription factors NANOG and GATA6 mutually repress each other through extracellular growth factor signaling:

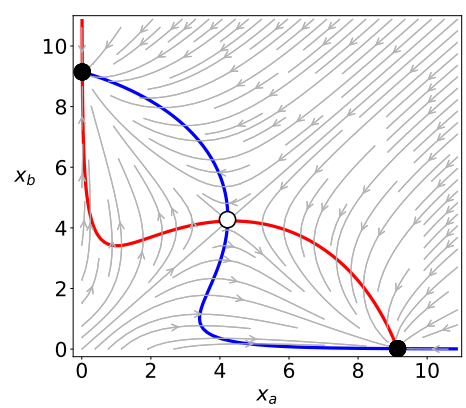
$$\frac{dx_i}{dt} = \frac{\alpha(1+x_i^n)^m}{\left(1+x_i^n\right)^m + \left(\frac{\langle x\rangle_i}{K}\right)^{2m}} - x_i, \qquad i = 1,2\ldots,N$$

Here x_i represents the (dimensionless) concentration of NANOG in cell i, and $\langle x\rangle_i$ denotes the average value of x over the immediate neighborhood of cell i, including $\langle x\rangle_i$ itself. This model can be derived from a specific molecular-level circuit explicitly involving NANOG, GATA6 and the growth factor FGF4, and implicitly ERK downstream of FGF receptors, as shown in the Supplementary Text and Fig. 2c. However, due to its minimal character, it is not unique to the specifics of the molecular interactions assumed to be involved in this cell-fate decision; other molecular circuits could be reduced to it.

In the model above, the level of NANOG is inhibited by its neighbors, in a manner that resembles lateral inhibition-mediated signaling. We ask whether such a model can sustain solutions in which cells cluster into two distinct cell types, expressing NANOG at two different levels (which we could interpret as epiblast and PrE cells). Epiblast and PrE cells display a salt-and-pepper distribution in the ICM at early blastocyst stages (Chazaud et al., 2006). Therefore, we assume perfect mixing among the cells within the population, in which case such a two-cluster state would be described by the following two-dimensional dynamical system:

$$\frac{dx_a}{dt} = \frac{\alpha(1+x_a^n)^m}{(1+x_a^n)^m + \left(\frac{x_a+x_b}{2K}\right)^{2m}} - x_a$$

$$\frac{dx_b}{dt} = \frac{\alpha(1+x_b^n)^m}{(1+x_b^n)^m + \left(\frac{x_a+x_b}{2K}\right)^{2m}} - x_b$$



The dynamics of this potential two-cluster solution can be examined via the phase-plane portrait shown in the figure, which displays the nullclines of the system in which each of the derivatives is zero (in red and blue), the typical flow exhibited by the system from different initial conditions (in grey) and the equilibria of the system, two of them stable (black circles) and the third one unstable (white circle). Note that the two stable equilibria correspond to the same behavior of the system, since the labels a and b of the two clusters are interchangeable.

The existence of the two symmetric stable equilibria ensures that the two-cluster state is a solution of the system, and that the population splits spontaneously in two distinct fates, as we

show by means of agent-based simulations (described in the Supplementary Text) throughout the text.

Acknowledgements

We would like to thank Drs Frederic Geissmann, Pierre-Louis Loyher, Alison North and Christina Pyrgaki for access to their two-photon microscope systems and for invaluable help in setting up the experimental conditions for cell ablation in embryos, Dr Pavak Shah for expert advice on laser ablation in different contexts, Drs Jennifer Nichols and Stanley Strawbridge for open discussion of complementary projects, Dr Alfonso Martinez Arias and Dr Christian Schröter for insightful discussions, and last but not least, we thank Drs Pedro Rocha, Joana Vidigal, and members of the Hadjantonakis lab for critical feedback and comments on the text. The authors would like to dedicate this work to the memory of Profs Andrzej K Tarkowski and Yoshiki Sasai, whose pioneering work has long been a source of inspiration.

NS was supported by a fellowship from the Tri-Institutional Stem Cell Initiative, funded by the Starr Foundation; SR was funded by the City College of New York-MSKCC Partnership Undergraduate Research Training (PURT) (NIH/NCI U54-CA137788/U54-CA132378); JPH was funded by the Louis V Gerstner-Sloan Kettering (GSK) Summer Undergraduate Research Program (SURP). Work in the Hadjantonakis lab is supported by the National Institutes of Health (R01-HD094868, R01-DK084391 and P30-CA008748) and by the MSKCC Functional Genomics Initiative. Work in the García-Ojalvo lab is funded by the Spanish Ministry of Science, Innovation and Universities and FEDER (projects PGC2018-101251-B-I00 and MDM-2014-0370) and by the ICREA Academia Program.

Author contributions

NS and AKH conceived and planned the experiments. NS performed all embryo experiments. LMB and JGO built the mathematical model and performed *in silico* experiments. NS, SR, HG and JPH manually curated immunofluorescence data. NS and SR processed time-lapse data. SR and JPH assisted with sample genotyping. NS designed and wrote data processing pipelines and analyzed experimental data with assistance from SR and HG. NS and JGO generated graphs and figures. NS, JGO and AKH wrote the manuscript with input and approval from all authors.

References

Artus, J., Piliszek, A., and Hadjantonakis, A.-K. (2011). The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. Dev Biol *350*, 393–404.

Artus, J., Vandormael-Pournin, S., Frödin, M., Nacerddine, K., Babinet, C., and Cohen-Tannoudji, M. (2005). Impaired mitotic progression and preimplantation lethality in mice lacking OMCG1, a new evolutionarily conserved nuclear protein. Mol Cell Biol *25*, 6289–6302.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes & Development *17*, 126–140.

Beccari, L., Moris, N., Girgin, M., Turner, D.A., Baillie-Johnson, P., Cossy, A.-C., Lutolf, M.P., Duboule, D., and Arias, A.M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. Nature *562*, 272–276.

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733–737.

Bedzhov, I., and Zernicka-Goetz, M. (2014). Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. Cell *156*, 1032–1044.

Behringer, R.R., Gertsenstein, M., Vintersten Nagy, K., and Nagy, A. (2014). Manipulating the mouse embryo: a laboratory manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

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.

Boj, S.F., Hwang, C.-I., Baker, L.A., Chio, I.I.C., Engle, D.D., Corbo, V., Jager, M., Ponz-Sarvise, M., Tiriac, H., Spector, M.S., et al. (2015). Organoid Models of Human and Mouse Ductal Pancreatic Cancer. Cell *160*, 324–338.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat Cell Biol *16*, 516–528.

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

Brewer, J.R., Molotkov, A., Mazot, P., Hoch, R.V., and Soriano, P. (2015). Fgfr1 regulates development through the combinatorial use of signaling proteins. Genes & Development *29*, 1863–1874.

Brook, F.A., and Gardner, R.L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. Proc Natl Acad Sci USA *94*, 5709–5712.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643–655.

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.

Chen, C.-C., Wang, L., Plikus, M.V., Jiang, T.X., Murray, P.J., Ramos, R., Guerrero-Juarez, C.F., Hughes, M.W., Lee, O.K., Shi, S., et al. (2015). Organ-Level Quorum Sensing Directs Regeneration in Hair Stem Cell Populations. Cell *161*, 277–290.

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. Biophysical Journal *110*, 710–722.

Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. Nature *533*, 251–254.

Eakin, G.S., and Hadjantonakis, A.-K. (2006). Production of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. Nature Protocols *1*, 1145–1153.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature *472*, 51–56.

Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., and Chazaud, C. (2011). Primitive Endoderm Differentiates via a Three-Step Mechanism Involving Nanog and RTK Signaling. Dev Cell *21*, 1005–1013.

Gardner, R.L. (1968). Mouse chimeras obtained by the injection of cells into the blastocyst. Nature *220*, 596–597.

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

Hadjantonakis, A.-K., and Papaioannou, V.E. (2004). Dynamic in vivo imaging and cell tracking using a histone fluorescent protein fusion in mice. BMC Biotechnol. *4*, 33.

Hamilton, T.G., Klinghoffer, R.A., Corrin, P.D., and Soriano, P. (2003). Evolutionary divergence of platelet-derived growth factor alpha receptor signaling mechanisms. Mol Cell Biol *23*, 4013–4025.

Harrison, S.E., Sozen, B., Christodoulou, N., Kyprianou, C., and Zernicka-Goetz, M. (2017). Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. Science *356*, eaal1810.

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

Kang, M., Garg, V., and Hadjantonakis, A.-K. (2017). Lineage Establishment and Progression within the Inner Cell Mass of the Mouse Blastocyst Requires FGFR1 and FGFR2. Dev Cell *41*, 496–510.e5.

Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.-K. (2013). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development *140*, 267–279.

Kicheva, A., Bollenbach, T., Ribeiro, A., Valle, H.P., Lovell-Badge, R., Episkopou, V., and Briscoe, J. (2014). Coordination of progenitor specification and growth in mouse and chick spinal cord. Science *345*, 1254927–1254927.

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

Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K.D., Yamada, R.G., Ueda, H.R., and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res *34*, e42.

Lan, Z.-J., Xu, X., and Cooney, A.J. (2004). Differential oocyte-specific expression of Cre recombinase activity in GDF-9-iCre, Zp3cre, and Msx2Cre transgenic mice. Biol Reprod *71*, 1469–1474.

Lander, A.D. (2011). Pattern, Growth, and Control. Cell *144*, 955–969.

Lou, X., Kang, M., Xenopoulos, P., Muñoz-Descalzo, S., and Hadjantonakis, A.-K. (2014). A Rapid and Efficient 2D/3D Nuclear Segmentation Method for Analysis of Early Mouse Embryo and Stem Cell Image Data. Stem Cell Reports *2*, 382–397.

Meng, Y., Moore, R., Tao, W., Smith, E.R., Tse, J.D., Caslini, C., and Xu, X.-X. (2018). GATA6 phosphorylation by Erk1/2 propels exit from pluripotency and commitment to primitive endoderm. Dev Biol *436*, 55–65.

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.

Mintz, B., and Illmensee, K. (1975). Normal genetically mosaic mice produced from malignant teratocarcinoma cells. Proceedings of the National Academy of Sciences *72*, 3585–3589.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Molotkov, A., Mazot, P., Brewer, J.R., Cinalli, R.M., and Soriano, P. (2017). Distinct Requirements for FGFR1 and FGFR2 in Primitive Endoderm Development and Exit from Pluripotency. Dev Cell *41*, 511–526.e514.

Morgani, S.M., Canham, M.A., Nichols, J., Sharov, A.A., Migueles, R.P., Ko, M.S.H., and Brickman, J.M. (2013). Totipotent Embryonic Stem Cells Arise in Ground-State Culture Conditions. Cell Reports.

Morgani, S.M., Metzger, J.J., Nichols, J., Siggia, E.D., and Hadjantonakis, A.-K. (2018a). Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. Elife *7*, 1040.

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

Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M., Dimos, J.T., Chen, A.E., Melton, D.A., et al. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Genes & Development *24*, 312–326.

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.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215–3222.

Nissen, S.B., Perera, M., Gonzalez, J.M., Morgani, S.M., Jensen, M.H., Sneppen, K., Brickman, J.M., and Trusina, A. (2017). Four simple rules that are sufficient to generate the mammalian blastocyst. PLoS Biol *15*, e2000737.

Nowotschin, S., Setty, M., Kuo, Y.-Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature *569*, 361–367.

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. Nat Cell Biol *16*, 27–37.

Papaioannou, V.E., and Ebert, K.M. (1995). Mouse half embryos: viability and allocation of cells in the blastocyst. Dev Dyn *203*, 393–398.

Papaioannou, V.E., Mkandawire, J., and Biggers, J.D. (1989). Development and phenotypic variability of genetically identical half mouse embryos. Development *106*, 817–827.

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.

Reupke, T., Püschel, B., and Viebahn, C. (2009). Tracing and ablation of single cells in the mammalian blastocyst using fluorescent DNA staining and multi-photon laser microscopy. Histochem. Cell Biol. *131*, 521–530.

Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.-C., Korving, J., Vivié, J., Truckenmüller, R.K., Oudenaarden, A., Blitterswijk, C.A., and Geijsen, N. (2018). Blastocyst-like structures generated solely from stem cells. Nature *557*, 106–111.

Roselló-Díez, A., Madisen, L., Bastide, S., Zeng, H., and Joyner, A.L. (2018). Cell-nonautonomous local and systemic responses to cell arrest enable long-bone catch-up growth in developing mice. PLoS Biol *16*, e2005086.

Saiz, N., Kang, M., Schrode, N., Lou, X., and Hadjantonakis, A.-K. (2016a). Quantitative Analysis of Protein Expression to Study Lineage Specification in Mouse Preimplantation Embryos. J Vis Exp 53654.

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

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature *459*, 262–265.

Schrode, N., Saiz, N., Di Talia, S., and Hadjantonakis, A.-K. (2014). GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. Dev Cell *29*, 454–467.

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

Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.M.E., Campbell, A., Devito, L.G., Ilic, D., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. Nat Cell Biol *18*, 700–708.

Shahbazi, M.N., Siggia, E.D., and Zernicka-Goetz, M. (2019). Self-organization of stem cells into embryos: A window on early mammalian development. Science *364*, 948–951.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722–737.

Smith, R., and McLaren, A. (1977). Factors affecting the time of formation of the mouse blastocoele. Journal of Embryology and Experimental Morphology *41*, 79–92.

Sodhi, C.P., Li, J., and Duncan, S.A. (2006). Generation of mice harbouring a conditional loss-of-function allele of Gata6. BMC Dev Biol *6*, 19.

Sozen, B., Amadei, G., Cox, A., Wang, R., Na, E., Czukiewska, S., Chappell, L., Voet, T., Michel, G., Jing, N., et al. (2018). Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. Nat Cell Biol *20*, 979–989.

Susaki, E.A., Tainaka, K., Perrin, D., Kishino, F., Tawara, T., Watanabe, T.M., Yokoyama, C., Onoe, H., Eguchi, M., Yamaguchi, S., et al. (2014). Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. Cell *157*, 726–739.

Tachibana, M., Sparman, M., Ramsey, C., Ma, H., Lee, H.-S., Penedo, M.C.T., and Mitalipov, S. (2012). Generation of Chimeric Rhesus Monkeys. Cell 1–11.

Takaoka, K., Nishimura, H., and Hamada, H. (2017). Both Nodal signalling and stochasticity select for prospective distal visceral endoderm in mouse embryos. Nature Communications *8*, 1492.

Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. Nature *184*, 1286–1287.

Tarkowski, A.K. (1961). Mouse chimaeras developed from fused eggs. Nature *190*, 857–860.

Tarkowski, A.K., and Wróblewska, J. (1967). Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. Journal of Embryology and Experimental Morphology *18*, 155–180.

Tosenberger, A., Gonze, D., Bessonnard, S., Cohen-Tannoudji, M., Chazaud, C., and Dupont, G. (2017). A multiscale model of early cell lineage specification including cell division. Npj Syst Biol Appl *3*, 32.

Warmflash, A., Sorre, B., Etoc, F., Siggia, E.D., and Brivanlou, A.H. (2014). A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. Nature Methods *11*, 847–854.

Wojcinski, A., Lawton, A.K., Bayin, N.S., Lao, Z., Stephen, D.N., and Joyner, A.L. (2017). Cerebellar granule cell replenishment postinjury by adaptive reprogramming of Nestin+ progenitors. Nat Neurosci *20*, 1361–1370.

Xenopoulos, P., Kang, M., Puliafito, A., Di Talia, S., and Hadjantonakis, A.-K. (2015). Heterogeneities in Nanog Expression Drive Stable Commitment to Pluripotency in the Mouse Blastocyst. Cell Reports *10*, 1508–1520.

Yamanaka, Y., Lanner, F., and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development *137*, 715–724.

Young, R.M., Hawkins, T.A., Cavodeassi, F., Stickney, H.L., Schwarz, Q., Lawrence, L.M., Wierzbicki, C., Cheng, B.Y., Luo, J., Ambrosio, E.M., et al. (2019). Compensatory growth renders Tcf7l1a dispensable for eye formation despite its requirement in eye field specification. Elife *8*, 12792.

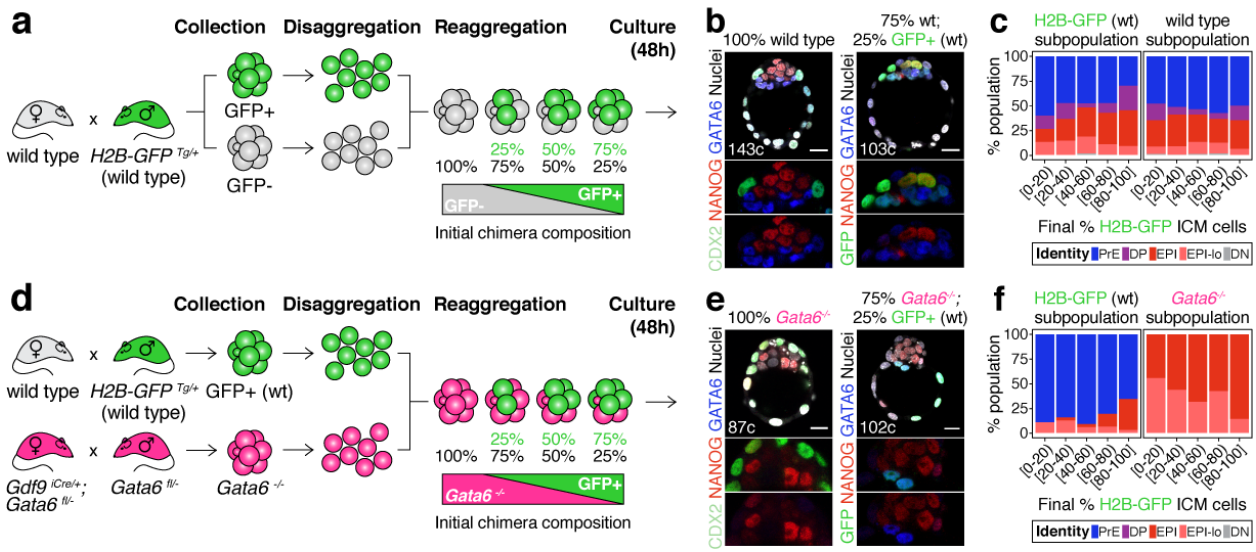
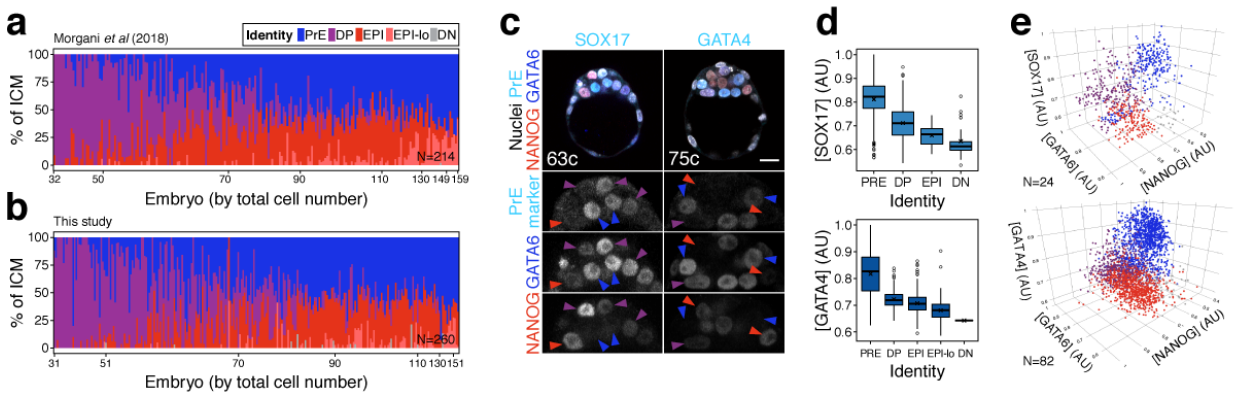
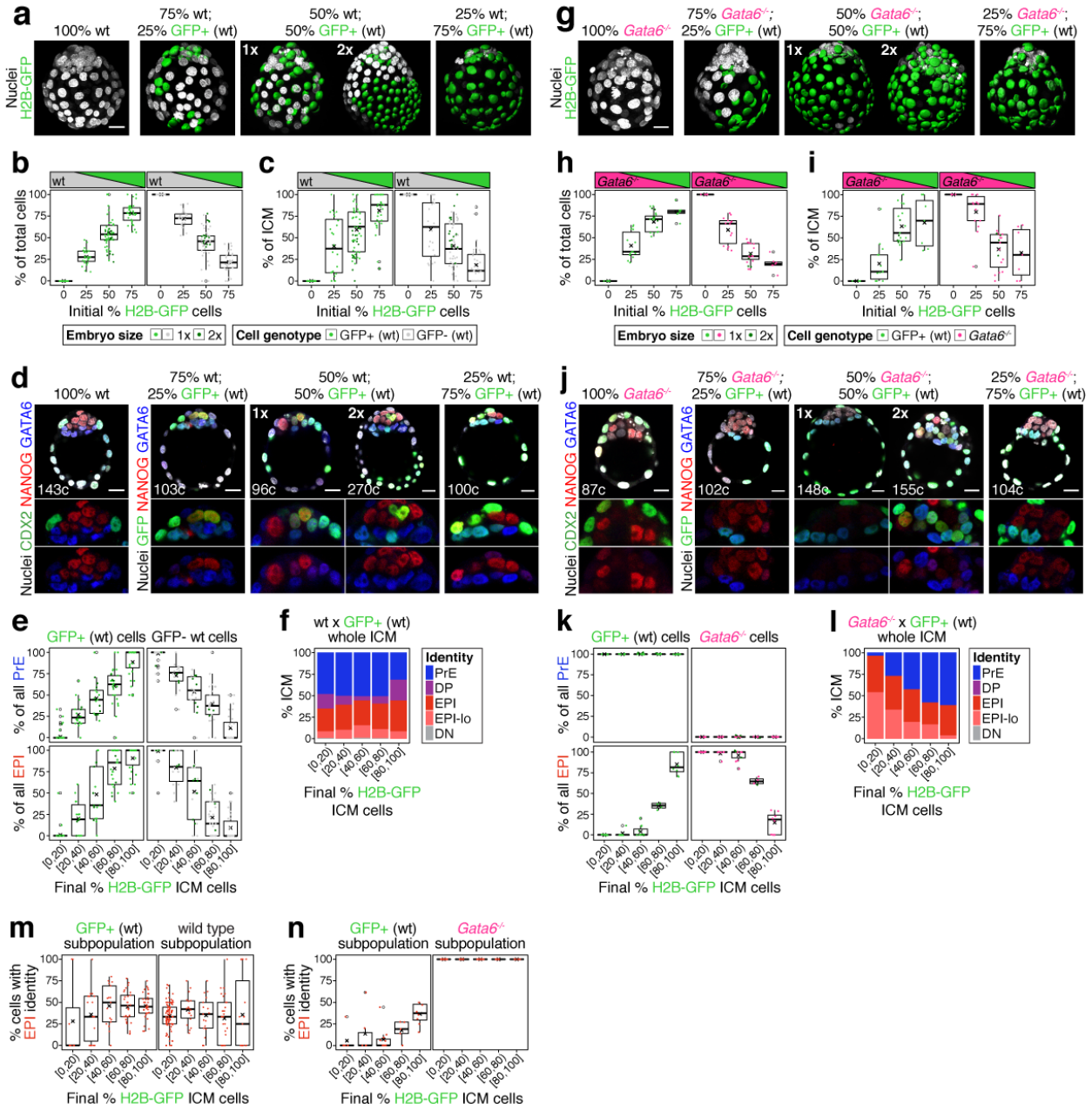


Figure 1. Cell fate decisions in the ICM of the blastocyst are made at the population level. **(a)** Experimental design to test independence of cell fate decisions in the ICM. Wild type embryos (H2B-GFP+ or GFP-) were collected at the 8-cell stage, disaggregated into single cells or clumps, re-aggregated in the combinations indicated and allowed to develop in culture for 48h, until the late blastocyst stage (equivalent to ~4.5 days post-fertilization). For further experimental details and information on alleles, see Methods. **(b)** Optical cross sections through representative immunofluorescence images of a non-chimeric, wild type control (100% wt) and a chimera made with 75% GFP-; 25% GFP+ cells (all wild type). **(c)** Stacked bar plots showing the lineage distribution of GFP+ (left) and GFP- (right) wild type ICM cells in embryos, stratified by the final % of H2B-GFP+ ICM cells, as indicated on the x-axis. **(d)** Experimental design to test independence of cell fate decisions in the ICM. Wild type embryos (H2B-GFP+) and *Gata6*$^{-/-}$ embryos were collected at the 8-cell stage, disaggregated into single cells or clumps, re-aggregated in the combinations indicated and allowed to develop in culture for 48h, until the late blastocyst stage (equivalent to ~4.5 days post-fertilization). For further experimental details and information on alleles, see Methods. **(e)** Optical cross sections through representative immunofluorescence images of a non-chimeric, *Gata6*$^{-/-}$ control (100% *Gata6*$^{-/-}$) and a chimera made with 75% *Gata6*$^{-/-}$; 25% GFP+ (wt) cells. **(f)** Stacked bar plots showing the lineage distribution of GFP+ (wt, left) and GFP- *Gata6*$^{-/-}$ (right) ICM cells in embryos, stratified by the final % of H2B-GFP+ ICM cells, as indicated on the x-axis. Color coding is indicated. All embryos labeled for NANOG (red), GATA6 (blue) and either CDX2 (controls) or GFP (chimeras) (green). All optical cross sections are 5μm maximum intensity projections. Total cell counts are indicated for each embryo within the merged images. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, DN: Double Negative (for NANOG and GATA6). Scale bars = 20μm.



Supplementary Figure 1 (Related to Figure 1). (a) Stacked bar plots showing progression of lineage composition in the ICM of wild type blastocysts fixed upon collection analyzed in (Morgani et al., 2018b). ICM cell types were re-assigned using hierarchical clustering, as described in the Methods section. **(b)** Stacked bar plots showing progression of lineage composition in the ICM of wild type blastocysts fixed upon collection generated in this study. Color coding for cell types is indicated. Each bar represents the ICM of a single embryo. Embryos are arranged by ascending total cell number (from 32 up to >150 cells, comprising all the stages of blastocyst development). N indicates number of embryos in the plot. **(c)** Optical cross sections through representative immunofluorescence images of blastocysts stained for NANOG (red), GATA6 (blue) and either SOX17 or GATA4 (cyan), as indicated. Arrowheads point at cells classified as PrE (blue), epiblast (red) or DP (purple). **(d)** Average SOX17 and GATA4 levels in each ICM cell type in a cohort of 24 embryos (SOX17) and 82 embryos (GATA4). Values are natural log re-scaled to 0-1 for each dataset. **(e)** 3D scatterplots for the three markers shown in (c, d). Each spot represents one cell. Spots are color-coded for cell identities as indicated in (a). Cells were classified using only relative NANOG and GATA6 levels, as described in the Methods. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR) and cross represents the arithmetic mean. All optical cross sections are 5μm maximum intensity projections. Total cell counts are indicated for each embryo within the merged images. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, DN: Double Negative (for NANOG and GATA6). Scale bars = 20μm.



Supplementary Figure 2 (Related to Figure 1). (a) Maximum intensity projections of all nuclei of a series of representative wild type control and wt;GFP+ chimeras made with increasing % of wild type GFP+ cells, as described in Fig. 1a. 1x indicates normal size embryo, resulting from aggregating 8 cells, 2x indicates double-sized embryo, resulting from aggregating 2x intact 8-cell stage embryos. **(b)** Box plots showing the relative contribution of GFP+ (green) and GFP-, wild type (gray) cells to the embryo for each of the initial % of GFP+ cells, as shown in (a) Dark green spots represent double embryos (2x). **(c)** Box plots showing the relative contribution of GFP+ (green) and GFP-, wild type cells (gray) to the ICM for each of the initial % of GFP+ cells, as shown in (a). **(d)** Optical cross sections through embryos shown in (a) and immunostained for

the markers indicated. **(e)** Box plots showing the relative contribution of GFP+ (green) and GFP-, wild type cells (gray) to each the PrE and the epiblast. Embryos are grouped by the final fraction of GFP+ ICM cells, as indicated on the x-axis **(f)** Stacked bar plots showing the ICM composition in wt;GFP+ chimeras, grouped by the final fraction of GFP+ ICM cells, as in (e) and as indicated on the x-axis. Color coding is indicated. **(g)** Maximum intensity projections of all nuclei of a series of representative *Gata6*$^{-/-}$ control and *Gata6*$^{-/-}$;GFP+ chimeras made with increasing % of GFP+ (wt) cells, as described in Fig. 1d. 1x indicates normal size embryo, resulting from aggregating 8 cells, 2x indicates double-sized embryo, resulting from aggregating 2x intact 8-cell stage embryos. **(h)** Box plots showing the relative contribution of GFP+ (green, wt) and GFP-, *Gata6*$^{-/-}$ cells (magenta) to the embryo for each of the initial % of GFP+ cells, as shown in (g) Dark green spots represent double embryos (2x). **(i)** Box plots showing the relative contribution of GFP+ (green, wt) and GFP-, *Gata6*$^{-/-}$ cells (magenta) to the ICM for each of the initial % of GFP+ cells, as shown in (g) **(j)** Optical cross sections through embryos shown in (g) and immunostained for the markers indicated. **(k)** Box plots showing the relative contribution of GFP+ (green, wt) and GFP-, *Gata6*$^{-/-}$ (magenta) cells to each the PrE and the epiblast for embryos grouped by the final fraction of GFP+ ICM cells, as in (e) and as indicated on the x-axis. **(l)** Stacked bar plots showing the ICM composition in *Gata6*$^{-/-}$;GFP+ chimeras, grouped by the final fraction of GFP+ ICM cells, as in (k) and as indicated on the x-axis. Color coding is indicated. **(m)** Box plots showing the % of GFP+ or GFP- cells (both wild type) that adopt EPI/DN identity in wt;GFP+ chimeras, grouped by the final fraction of GFP+ ICM cells, as in (e,k) and as indicated on the x-axis. **(n)** Box plots showing the % of GFP+ (wt) or GFP-, *Gata6*$^{-/-}$ cells that adopt EPI/DN identity in *Gata6*$^{-/-}$;GFP+ chimeras, grouped by the final fraction of GFP+ ICM cells, as in (m). In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. All embryos labeled for NANOG (red), GATA6 (blue) and either CDX2 (controls) or GFP (chimeras) (green). All optical cross sections are 5μm maximum intensity projections. Total cell counts are indicated for each embryo within the merged images. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, DN: Double Negative (for NANOG and GATA6). Scale bars = 20μm.

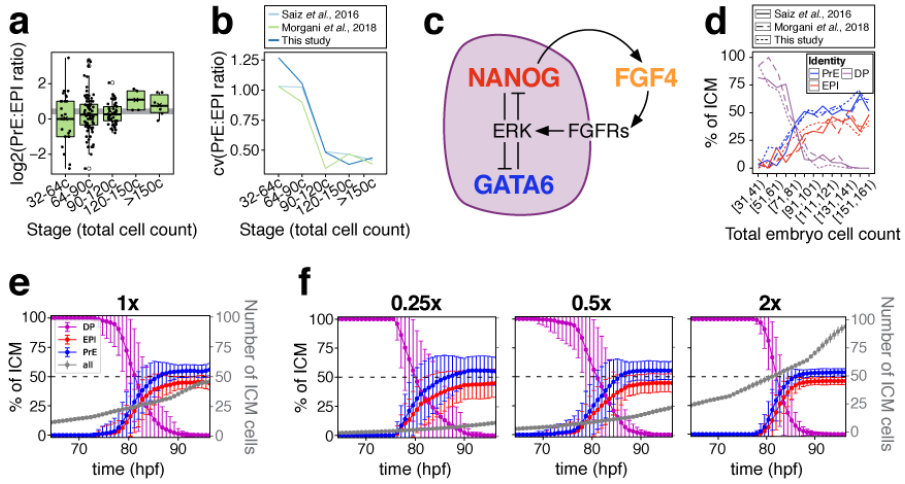
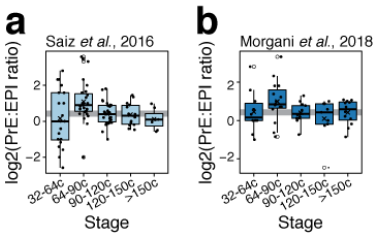


Figure 2. A minimal model of cell fate decisions solely mediated by growth factor signaling explains robust lineage specification in the ICM. (a) Box plot showing the ratio of PrE to epiblast cells in embryos at sequential stages of blastocyst development shown in Fig. S1b. **(b)** Coefficient of variation of the PrE:epiblast ratio over developmental time in three different datasets – this study (shown in (e)), data from (Saiz et al., 2016b) (Fig. S3b) and data from (Morgani et al., 2018b) (Fig. S3c), as indicated. **(c)** Diagram of our proposed model of molecular control of cell fate in the ICM. **(d)** Lineage dynamics in *in silico* simulations of ICM development using our proposed model. **(e)** *In vivo* ICM lineage dynamics from three experimental datasets, as indicated and as in (a). **(f)** Lineage dynamics in *in silico* simulations of scaling experiments, as published in (Saiz et al., 2016b). Absolute ICM size was modified to 1x, 0.25x, 0.5 or 2x the normal size, as indicated.



Supplementary Figure 3 (related to Figure 2). (a) Box plot showing the ratio of PrE to epiblast cells in embryos at sequential stages of blastocyst development analyzed in (Saiz et al., 2016b). **(b)** Box plot showing the ratio of PrE to epiblast cells in embryos at sequential stages of blastocyst development analyzed in (Morgani et al., 2018b) and shown in Fig. S1a. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Crosses indicate the arithmetic mean and each dot represents one embryo.

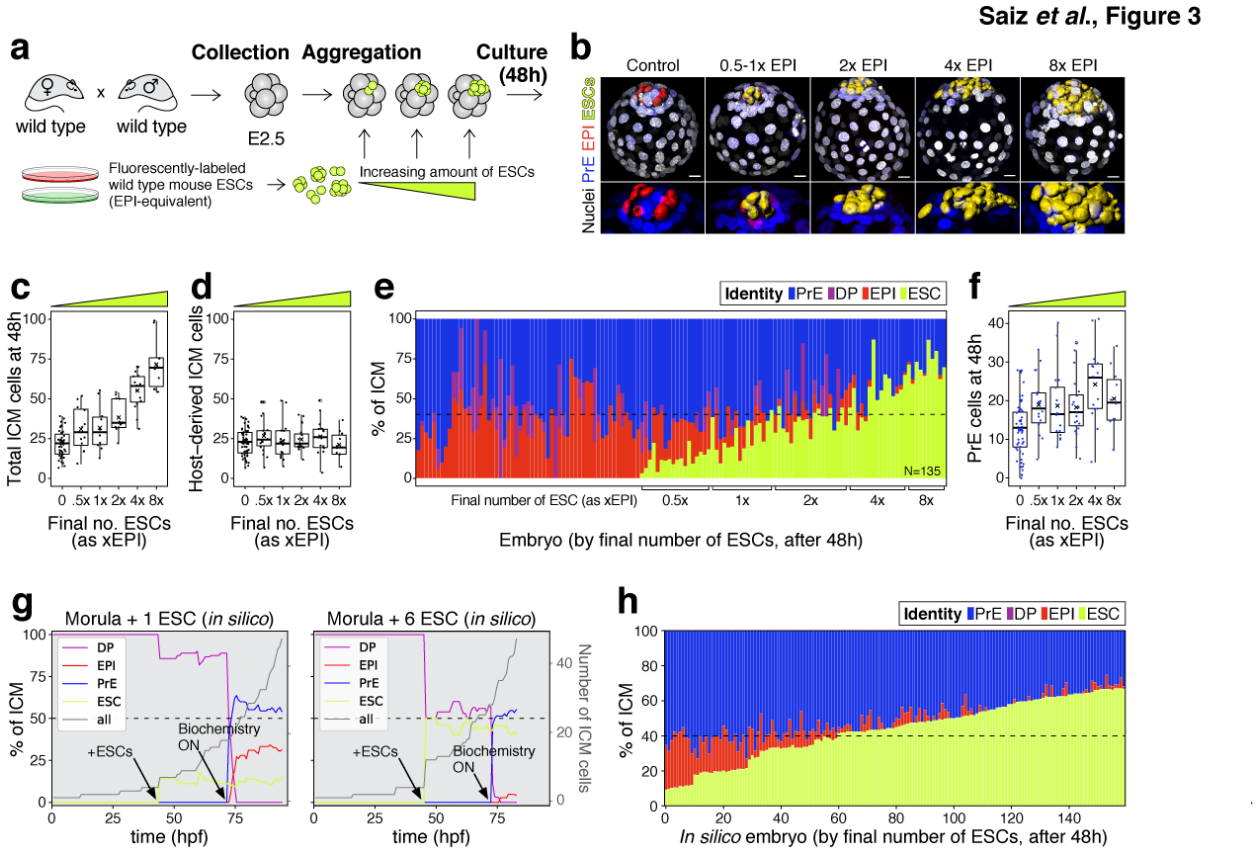
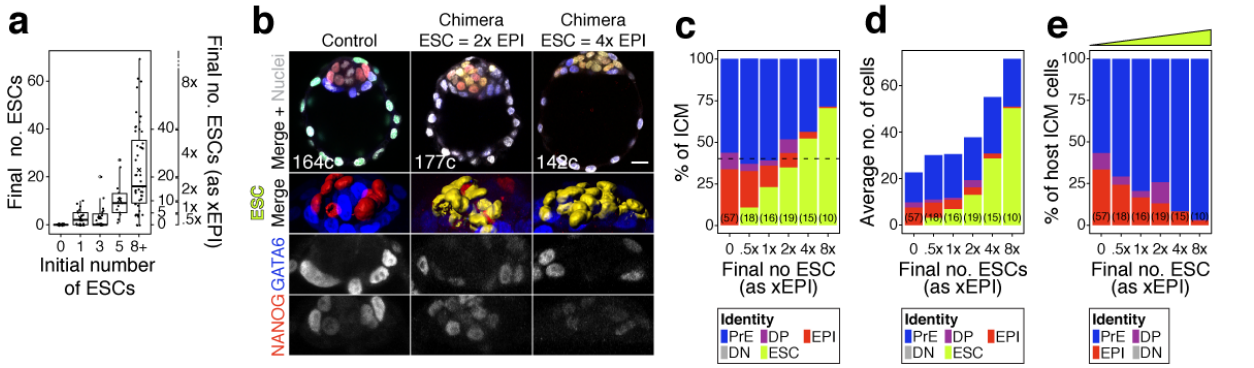


Figure 3. The lineage composition of the ICM is robust to expansion of the epiblast. (a) Experimental design. 8- or 8-16 cell stage embryos (2.5 days post-fertilization) were recovered from CD1 (wild type) females crossed with CD1 males. Embryos were denuded, aggregated with clumps of fluorescently labelled ESCs and cultured for 48-56h, until the late blastocyst stage (equivalent to ~4.5 days post-fertilization). **(b)** 3D renders of a series of control (no ESCs) and chimeras, generated as indicated in (a), carrying increasing amounts of ESCs (indicated as EPI-equivalent size (x EPI)). Control epiblast and ESCs in chimeras are highlighted as computer-rendered volumes and color coded as indicated. GATA6+ PrE is shown in blue, NANOG+ host epiblast is shown in red where applicable. **(c)** Box plot showing the total size of the ICM (host-derived + ESCs) at the end of the experiment in each group of embryos, as defined by the size of the ESC compartment. **(d)** Box plot showing the size of the host-derived ICM component at the end of the experiment in each group of embryos, as in (c). **(e)** Stacked bar plot showing the relative ICM composition at the end of the experiment for all embryos analyzed. Each bar represents the ICM of one embryo, ordered by increasing absolute number of ESCs at the end of the experiment. Dashed line indicates the normal ratio of 60% PrE:40% epiblast found in intact wild type embryos. Number of embryos analyzed is indicated (N). Brackets on x-axis indicate the number of ESCs in those embryos, relative to the size of the average wt control epiblast (xEPI). **(f)** Box plot showing the size of the PrE at the end of the experiment in each group of embryos. **(g)** Growth curves showing lineage dynamics in *in silico* simulations of the aggregation experiments shown in (a). Left Y-axis and curves for each lineage indicate relative size (as % of

ICM). Right Y-axis and gray curves indicate total number of ICM cells (including ESCs). **(h)** Stacked bar plot showing the relative ICM composition at the end of the experiment in *in silico* simulations of the experiments shown in (a). Each bar represents the ICM of one simulated embryo (i.e., a single iteration), and bars are arranged by increasing absolute number of ESCs at the end of the simulation, as in (e). Dashed line indicates the normal ratio of 60% PrE:40% epiblast found in intact wild type embryos. Color coding is indicated for (e, g, h). In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Crosses indicate the arithmetic mean and each dot represents one embryo. Yellow wedges represent the increasing amount of ESCs in each group. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, ESC: embryonic stem cell. Scale bars = 20μm.



Supplementary Figure 4 (related to Figure 3). (a) Box plots indicating ESC contribution to chimeras. Number of ESCs aggregated with morulae is shown against the final number of ESC in the chimeric embryo after 48h in culture. Right y axis shows epiblast-equivalent size bins used to categorize chimeric embryos (xEPI). **(b)** Optical cross sections through representative immunofluorescence images of a control and chimeras with 2xEPI- and 4xEPI-equivalent ESC compartments. Embryos are labelled for NANOG (red) and GATA6 (blue) to identify all ICM cell types. ESCs are shown in yellow. Lower panels show magnifications of the ICM, with all markers overlaid, for each individual marker as grayscale and for ESCs as 3D surface renders. Total cell count for each embryo is shown within the merged panel. **(c)** Stacked bar plot showing the relative ICM composition of each group of chimeras shown in Fig. 3e. Number of embryos in each group is indicated in parenthesis. **(d)** Stacked bar plot showing the average number of each ICM cell type for each group of chimeras shown in Fig. 3e and in (c). Number of embryos in each group is indicated in parenthesis. **(e)** Stacked bar plot showing the relative contribution of host cells to each of the ICM lineages in each group of embryos. Yellow wedge represents the increasing amount of ESCs in each group. Number of embryos in each group is indicated in parenthesis. Color coding is indicated. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. All optical cross sections are 5μm maximum intensity projections. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, ESC: embryonic stem cell. Scale bars = 20μm.

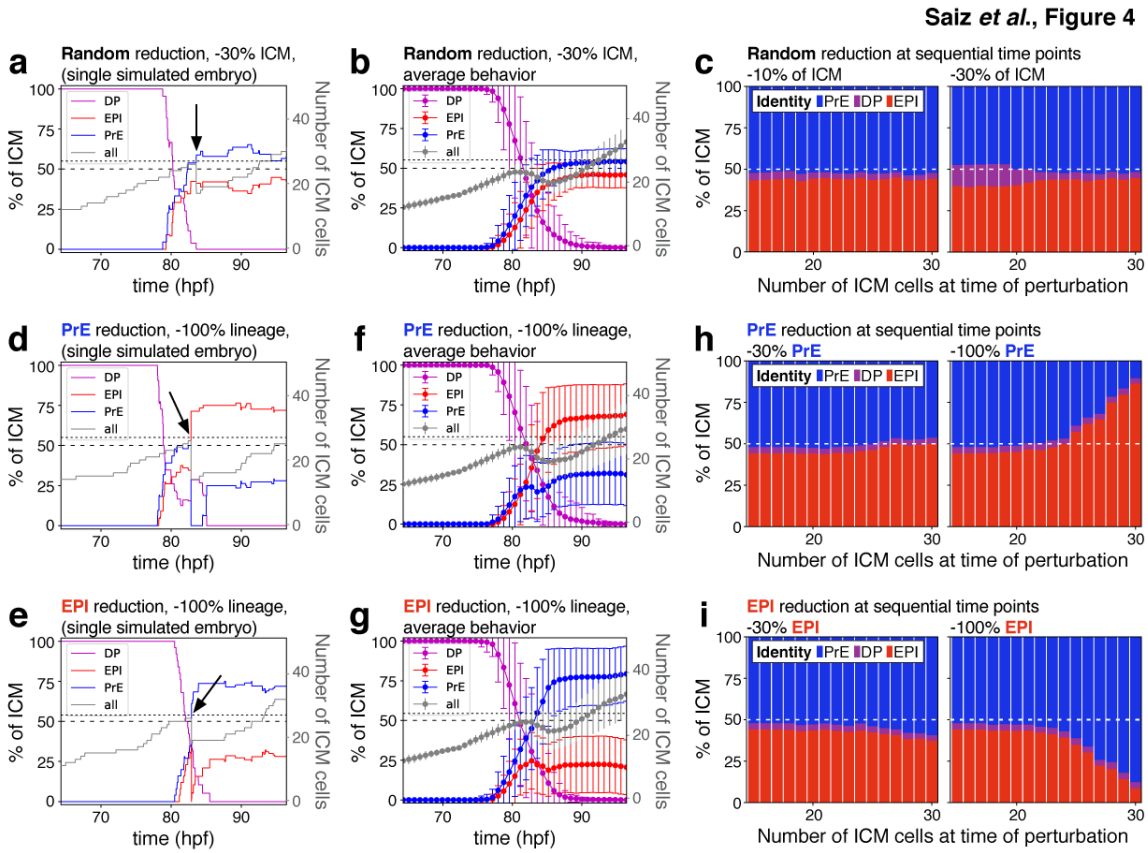


Figure 4. The lineage composition of the ICM is robust to *in silico* reduction of lineage size. **(a)** Growth curves for each ICM lineage after simulation of a 30% reduction in ICM size (3 PrE, 3 DP and 3 epiblast cells removed from a 27-cell ICM) using our model described in Fig. 2. Arrow indicates the time of cell elimination. Lines are color-coded for each lineage, as indicated and represent relative lineage size (scale on the left Y-axis). Grey line indicates the absolute size of the ICM, as shown on the right Y-axis. Dotted line indicates 27 ICM cells. Dashed line indicates 50% of the ICM, for reference. **(b)** Growth curves as those in (a) showing the average behavior for 100 simulations. Error bars indicate the standard deviation. **(c)** Stacked bar plots showing the final ICM composition after simulating the elimination of 10% (left) or 30% (right) of ICM cells at sequential points in embryo development. Developmental stage at the time of cell elimination is indicated on the x-axis as number of ICM cells (15-30 ICM cells, equivalent to ~50-100 total cells). **(d, e)** Growth curves for each ICM lineage after simulation of a 100% reduction in PrE (d) or epiblast (e), when the ICM reaches 27 cells, as shown in (a) and indicated by the arrow. Lines are color-coded for each lineage, as indicated and represent relative lineage size (scale on the left Y-axis). Grey line indicates the absolute size of the ICM, as shown on the right Y-axis. Dotted line indicates 27 ICM cells. Dashed line indicates 50% of the ICM, for reference. **(f, g)** Growth curves as those in (d, e) showing the average behavior for 100 simulations of PrE (f) and epiblast (g) reduction. Error bars indicate the standard deviation. **(h, i)** Stacked bar plots showing the final ICM composition after simulating the elimination of 30% (left) or 100% (right) of the PrE (h) or the epiblast (i) at sequential points in embryo development. Developmental stage at the time of

cell elimination is indicated on the x-axis as number of ICM cells (15-30 ICM cells, equivalent to ~50-100 total cells). Color coding is indicated. hpf: hours post-fertilization, PrE: Primitive Endoderm, EPI: epiblast, DP: double positive.

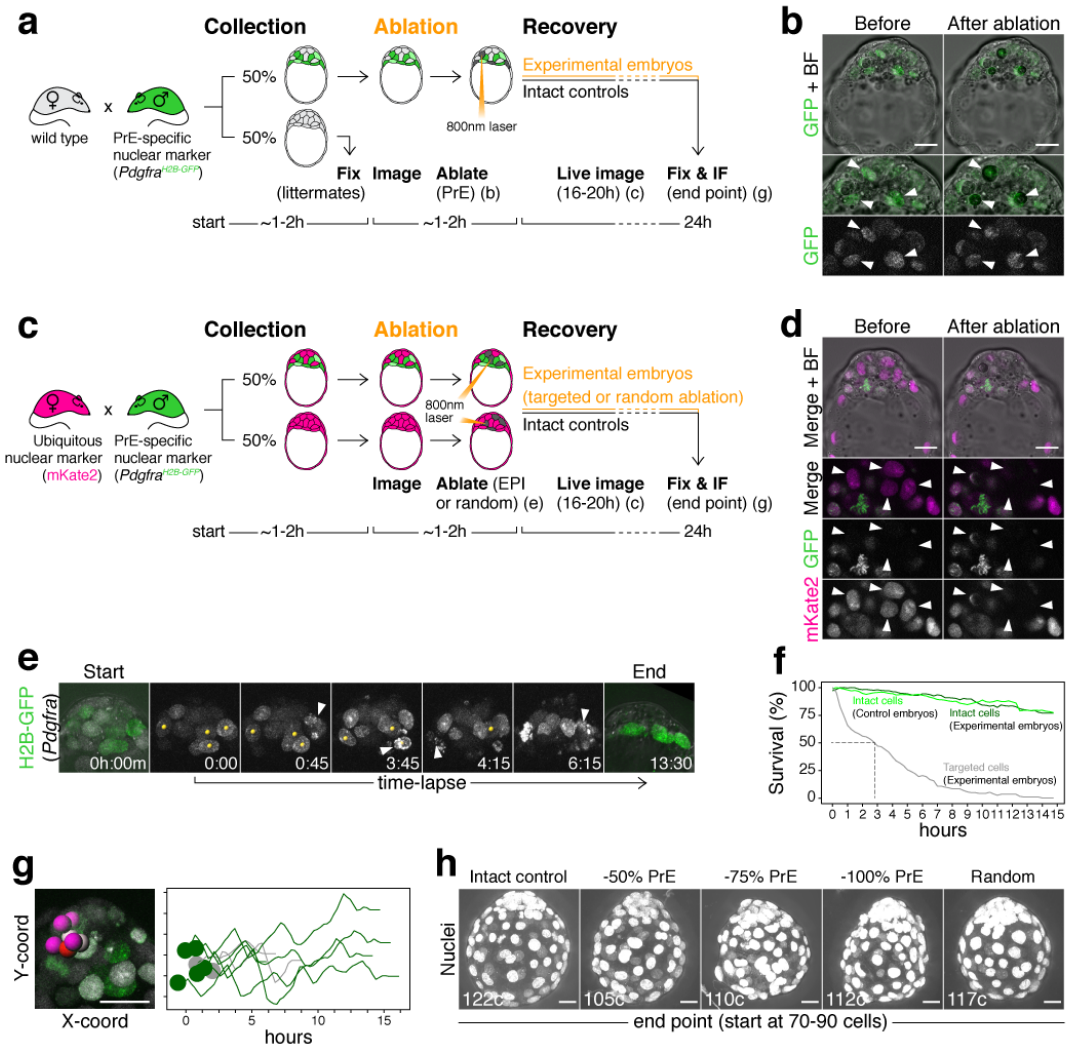
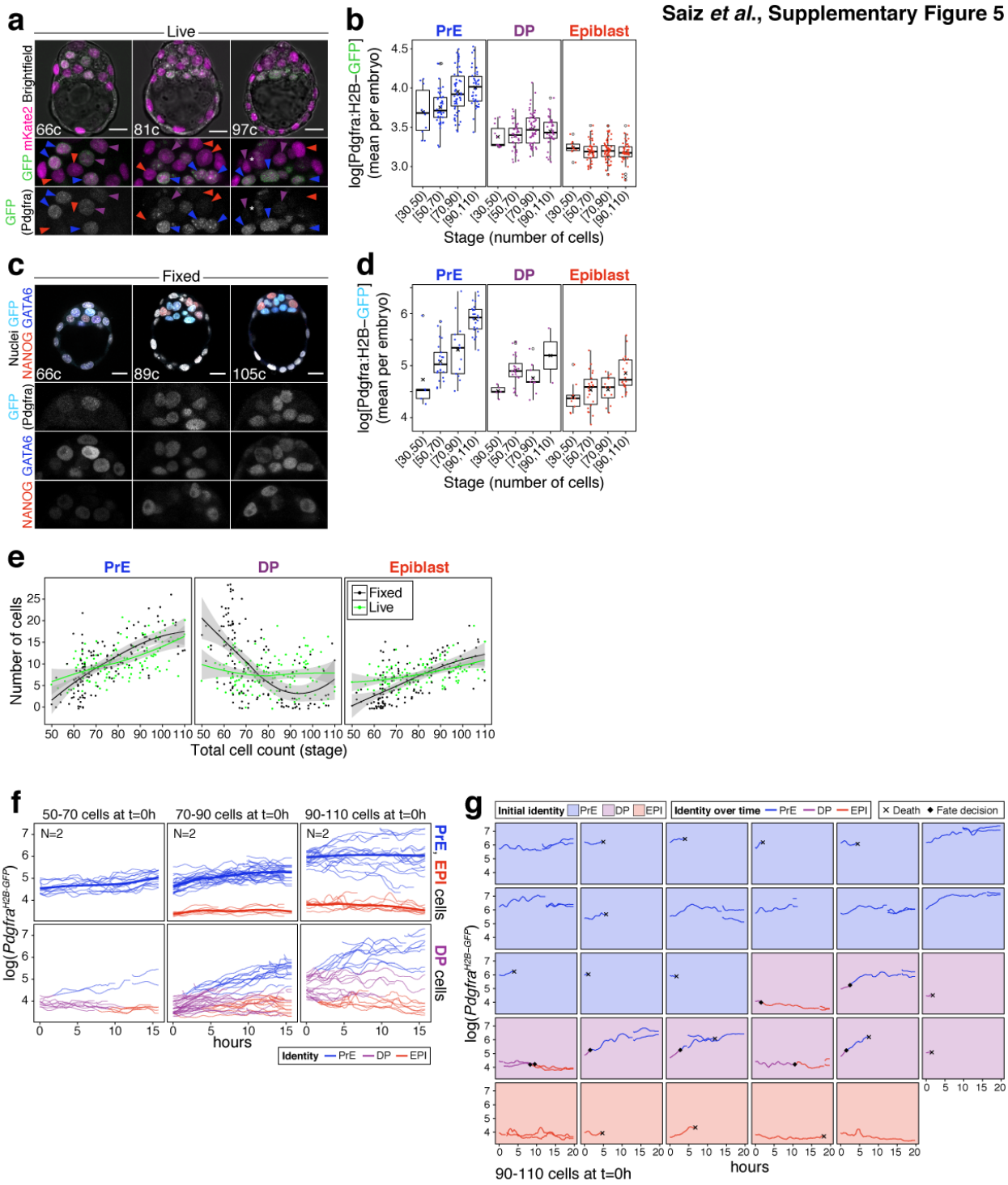


Figure 5. Laser ablation enables alteration of lineage size with high spatiotemporal control in mouse embryos. (a) Experimental design for PrE ablation. Blastocysts recovered from CD1 (wild type) females crossed with *Pdgfra*$^{H2B\text{-}GFP/+}$ or *Pdgfra*$^{H2B\text{-}GFP/+}$*; R26:CAG:3x-nls-mKate2*$^{Tg/Tg}$ males were sorted for GFP. GFP- embryos were fixed as reference littermates to estimate the developmental stage of the entire litter. GFP+ embryos were used for the experiment and subject to ablation of different amounts of PrE cells (GFP+) followed by 16-20h of live imaging to visualize the response to cell ablation, followed by 8-4h of *in vitro* culture (for a total of 24h). **(b)** Representative images of live GFP+ embryos before and after ablation of PrE cells. Bottom panels show magnifications of the ICM with GFP alone on grayscale, as indicated. Arrowheads point at targeted PrE cells. **(c)** Experimental design for epiblast ablation. Blastocysts were recovered from *R26:CAG:3x-nls-mKate2*$^{Tg/Tg}$ females crossed with *Pdgfra*$^{H2B\text{-}GFP/+}$*; R26:CAG:3x-nls-mKate2*$^{Tg/Tg}$ males. GFP-, mKate2+ embryos were used as either intact or random ablation controls. GFP+, mKate2+ embryos were used for the experiment and subject to ablation of different amounts of epiblast cells (GFP-) or PrE cells (GFP+), followed by 16-20h of live

imaging to visualize the response to cell ablation, followed by 8-4h of *in vitro* culture (for a total of 24h). **(d)** Representative images of live GFP+ embryos before and after ablation of epiblast cells. Bottom panels show magnifications of the ICM, for both markers together and each of them on grayscale, as indicated. Arrowheads point at targeted epiblast cells. **(e)** Still images of a representative embryo in the hours after ablation. See also Movie S4. Yellow spots on grayscale images mark targeted cells. Arrowheads point at cell death of each targeted cell. All images are timestamped as h:mm. **(f)** Survival curves for targeted (gray) and intact cells in both experimental (dark green) and intact control embryos (light green). Dashed line marks half-life of targeted cells (~3h) **(g)** Survival of intact cells neighboring targeted cells. Image shows selected ICM cells (intact DP and epiblast cells, color coded, and targeted PrE cells, gray). Graph shows X-Y coordinates of cells shown in picture and survival (hours) for each cell. Time scale for each cell is shifted based on their initial X position, for visualization purposes. **(h)** Maximum intensity projections of representative embryos fixed after 24h in culture, showing all nuclei over bright field image. Treatment was done at the 70-90 cell stage and is indicated above. Total cell count for each embryo is shown within each image. PrE: Primitive Endoderm, EPI: Epiblast. Scale bars = 20μm.



Supplementary Figure 5 (related to Figure 5). (a) Optical cross sections through representative live blastocysts at sequential developmental stages (total cell count is indicated within each image). Lower panels show magnifications of the ICM for both markers, as indicated, and for GFP alone in grayscale. Colored arrowheads point at nuclei considered to be PrE (blue arrowheads), DP (purple arrowheads) or epiblast (red arrowheads) based on GFP level. **(b)** Box

plots showing the average level of GFP (*Pdgfra*) per embryo, in each ICM cell type, for each developmental stage considered. **(c)** Optical cross sections through representative immunofluorescence images of reference *Pdgfra*$^{H2B\text{-}GFP/+}$ littermates fixed upon collection and labelled for NANOG (red) and GATA6 (blue) to identify ICM cell types. Lower panels show ICM magnifications for each marker in grayscale, as indicated. **(d)** Box plots showing the average level of GFP (*Pdgfra*) per embryo, in each ICM cell type (as determined by NANOG and GATA6 expression), for each developmental stage considered – for in embryos like those in (c). **(e)** Growth curves for each ICM population over time. Black lines and dots show numbers corresponding to fixed samples, where cell identities were assigned automatically based on relative NANOG and GATA6 levels, as described in the Methods. Green lines and dots show numbers corresponding to live samples, where cell identities were assigned manually based on GFP (*Pdgfra*) levels alone, as described in the Methods. Curves are local regression lines for each subset of data, fitted using the LOESS method. **(f)** Temporal dynamics of cell fate specification in live embryos. Expression levels of the PrE reporter allele *Pdgfra*$^{H2B\text{-}GFP}$ (Hamilton et al., 2003) are shown over time for individual ICM cells in intact embryos at sequential stages of development. Reporter expression was assessed using time lapse imaging over a 15h time window. Top panels show dynamics of reporter expression in cells classified as PrE or epiblast at the beginning of the movie. Smoothing curves for each PrE and epiblast are shown as thicker lines and color coded. Bottom panels show dynamics of reporter expression in cells classified as DP (progenitors) at the beginning of the movie. DP cells become PrE or epiblast over the course of the movie, as determined by *Pdgfra*$^{H2B\text{-}GFP}$ expression. **(g)** *Pdgfra* expression dynamics for individual cells in one embryo from (f) imaged from the 90-110 cell stage onwards. Panels are color coded for the initial identity of the cell. Traces are color coded for identity over time. Acquisition of PrE or epiblast identity is denoted by a black diamond. Black cross indicates cell death. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast. Scale bars = 20μm.

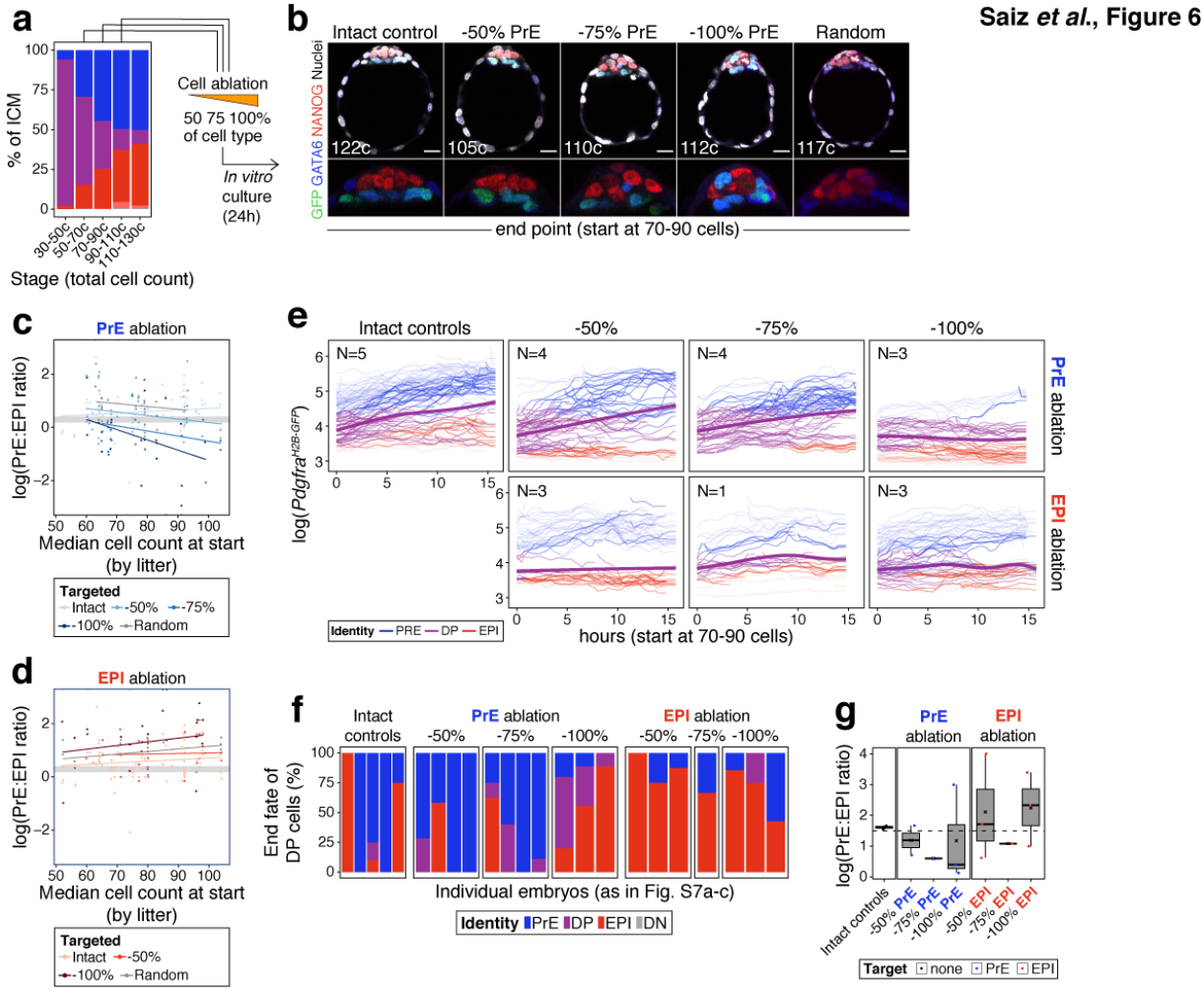
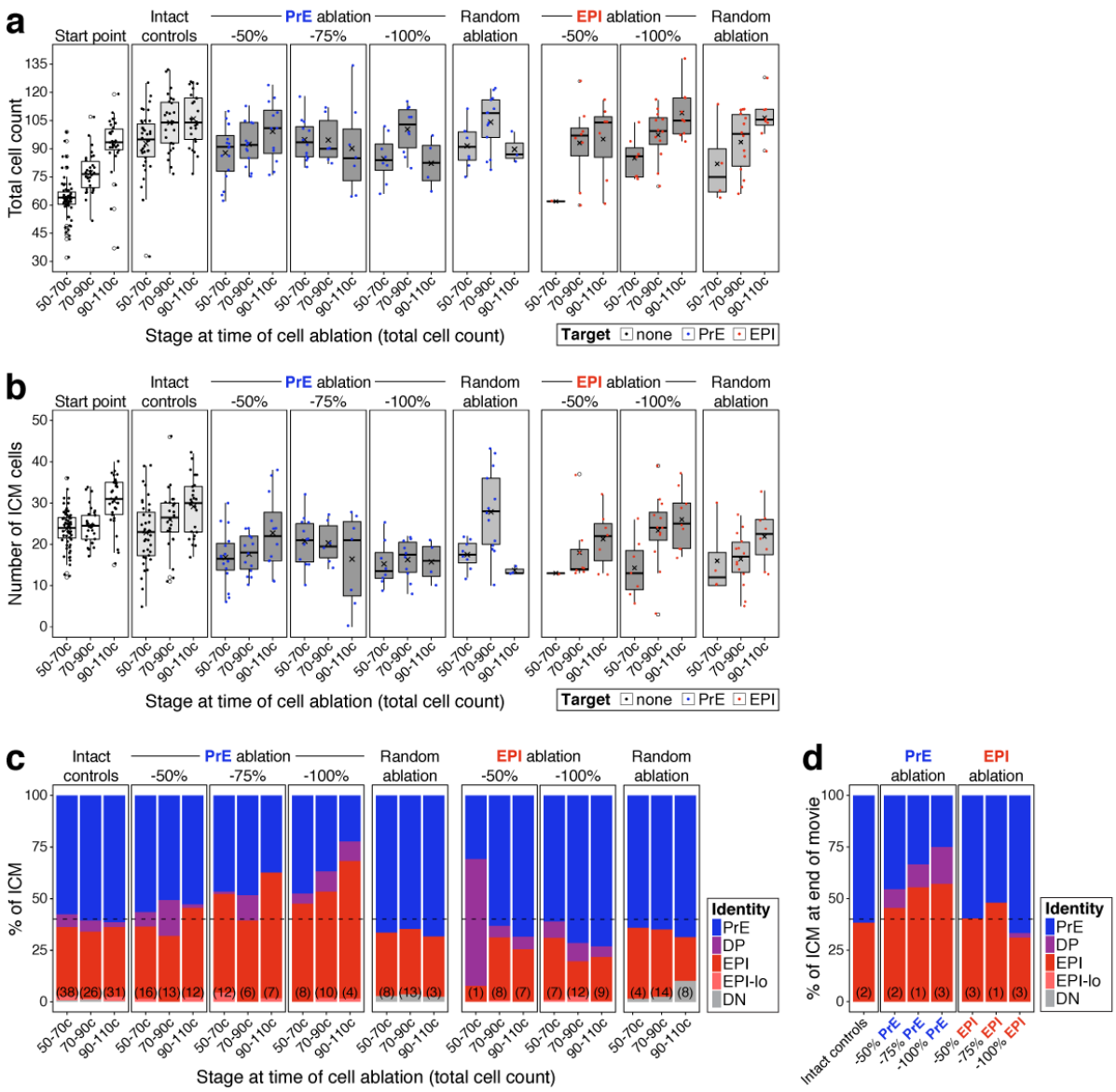


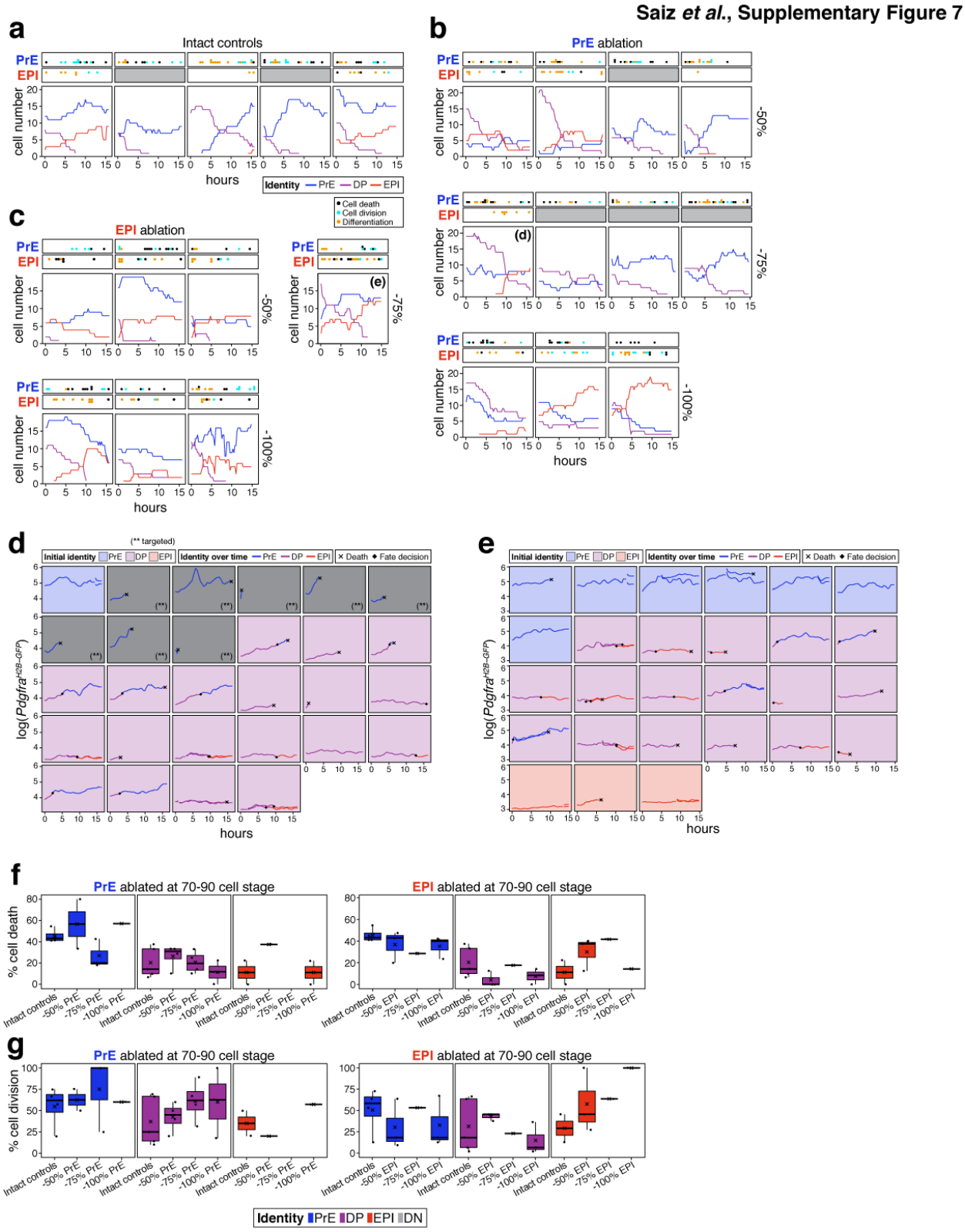
Figure 6. The cell fate choice of uncommitted ICM progenitors is dictated by lineage size. **(a)** Stacked bar plot showing ICM composition at sequential stages of blastocyst development. Embryos at each of these stages were subject to laser ablation of different fractions of either the PrE or the epiblast and allowed to recover *in vitro* for 24h (see Fig. 5a,c). **(b)** Optical cross sections through representative immunofluorescence images of embryos subject to ablation at the 70-90 cell stage and fixed at the end of the experiment (24h later) (same embryos as in Fig. 5g). Embryos are labelled for NANOG (red) and GATA6 (blue) to identify all ICM cell types. H2B-GFP, indicating *Pdgfra* expression, is shown in green where applicable. Lower panels show magnifications of the ICM Treatment for each embryo is indicated over each image. **(c)** PrE:epiblast ratio (shown as natural logarithm) at the end of the experiment (24h) in embryos where fractions of the PrE were eliminated at sequential stages of blastocyst development, as indicated on the x-axis. Shades of blue indicate the magnitude of the reduction in the PrE. Intact controls are embryos in which no cells were killed, Random controls are embryos in which randomly chosen ICM cells were killed without knowing their identity, in equivalent numbers to the -100% group. **(d)** PrE:epiblast ratio (shown as logarithm) at the end of the experiment (24h) in embryos where fractions of the epiblast were eliminated at sequential stages of blastocyst

development, as indicated on the x-axis. Shades of red indicate the magnitude of the reduction in the epiblast. **(e-f)** *In silico* simulation of the same experiments shown in (c-d), using our model, described in Fig. 2. **(g)** Dynamics of *Pdgfra*$^{H2B\text{-}GFP}$ expression in progenitor cells (DP) of experimental embryos targeted at the 70-90 cell stage. Each line represents one cell. Color coding indicates cell identity, as inferred from reporter expression (see Methods). Number of embryos per plot is indicated. Cells classified as PrE or epiblast at the beginning of the experiment are shown as color-coded semi-transparent lines behind progenitor cells, for reference. Smoothing curves for *Pdgfra* expression in progenitor cells are shown as thick purple lines. Fraction of the PrE or epiblast eliminated is indicated above each panel, lineage targeted is indicated to the right of each panel. **(h)** Stacked bar plots showing the final identity adopted by progenitor (DP) cells in each of the embryos plotted in (g). **(i)** Box plots showing the PrE:epiblast ratio (shown as logarithm) at the end of the movie, in embryos where all or most of the ICM cells could be tracked after cell ablation at the 70-90 cell stage (subset of embryos shown in (g) and (h)). Compare to panels (c) and (d). Treatment is indicated on the x-axis. In box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast. Scale bars = 20μm.



Supplementary Figure 6 (related to Figure 6). (a) Box plots showing the total cell count of embryos in each of the experimental groups shown in Fig. 6, at the end of the experiment. Embryos are binned by developmental stage at the start of the experiment, as shown on the x-axis. Start point comprises reference littermates fixed at the beginning of the experiment. Intact controls are embryos in which no cell was targeted. Random controls are embryos in which randomly chosen ICM cells were targeted, irrespective of their identity, in equivalent numbers to the -100% group for each lineage targeted (see Methods). Embryos in which the PrE or epiblast were targeted are split by the fraction of the lineage eliminated (50-100%, as indicated). **(b)** Box plots showing the total number of ICM cells in each of the experimental groups shown in (a). **(c)** Stacked bar plots showing the relative ICM composition in each of the experimental groups shown in Fig. 6c-d and in (a-b) above. The number of embryos in each group is indicated in

parenthesis. Cell types are color-coded as indicated. **(d)** Stacked bar plots showing the relative ICM composition at the end of the movie in embryos shown in Fig. 6e, for each of the treatments indicated on the x-axis. Data in (d) corresponds to embryos in which ablation was performed at the 70-90 cell stage and in which all or most of the ICM cells could be tracked throughout the movie. Cell types are color-coded as indicated. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast.



Supplementary Figure 7 (related to Figure 6). (a) Growth curves for each of the ICM populations in each of the intact embryos shown in Fig. 6e. Cell types are color-coded. Top panels indicate

cellular events in the PrE or epiblast populations: black dots represent cell death, cyan dots represent cell divisions, and orange dots represent a progenitor cell adopting either PrE or epiblast identity, respectively. Gray panels indicate no data is available **(b)** Growth curves for each of the ICM populations in each of the embryos shown in Fig. 6e in which the PrE was targeted, in the fractions indicated. **(c)** Growth curves for each of the ICM populations in each of the embryos in which the epiblast was targeted, shown in Fig. 6e. All embryos in (a-c) were manipulated at the 70-90 cell stage and live imaged for the first 15h of the 24h culture after ablation. **(d)** Temporal changes in *Pdgfra* expression and cell identity for all ICM cells tracked in an embryo in which 75% of the PrE was eliminated when it had 70-90 cells – labelled as "(d)" in (b). **(e)** Temporal changes in *Pdgfra* expression and cell identity for all ICM cells tracked in an embryo in which 75% of the epiblast was eliminated when it had 70-90 cells – labelled as "(e)" in (c). Each panel displays one cell. Panels are color coded for the initial identity of the cell, as indicated. Traces are color coded for identity over time. Acquisition of PrE or epiblast identity is denoted by a black diamond. Black cross indicates cell death. **(f)** Box plots showing frequency of cell death among intact cells of each ICM cell type, as color-coded. Each box represents one experimental group (Intact controls, or embryos in which increasing fractions of the corresponding lineage were eliminated, as indicated on the x-axis). Left plots correspond to PrE ablation, right plots to epiblast ablation, as indicated. **(g)** Box plots showing the frequency of cell division among intact cells of each ICM cell type, as color coded. Each box represents one experimental group (Intact controls, or embryos in which increasing fractions of the corresponding lineage were eliminated, as indicated on the x-axis). Left plots correspond to PrE ablation, right plots to epiblast ablation, as indicated. Color coding is indicated. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast.

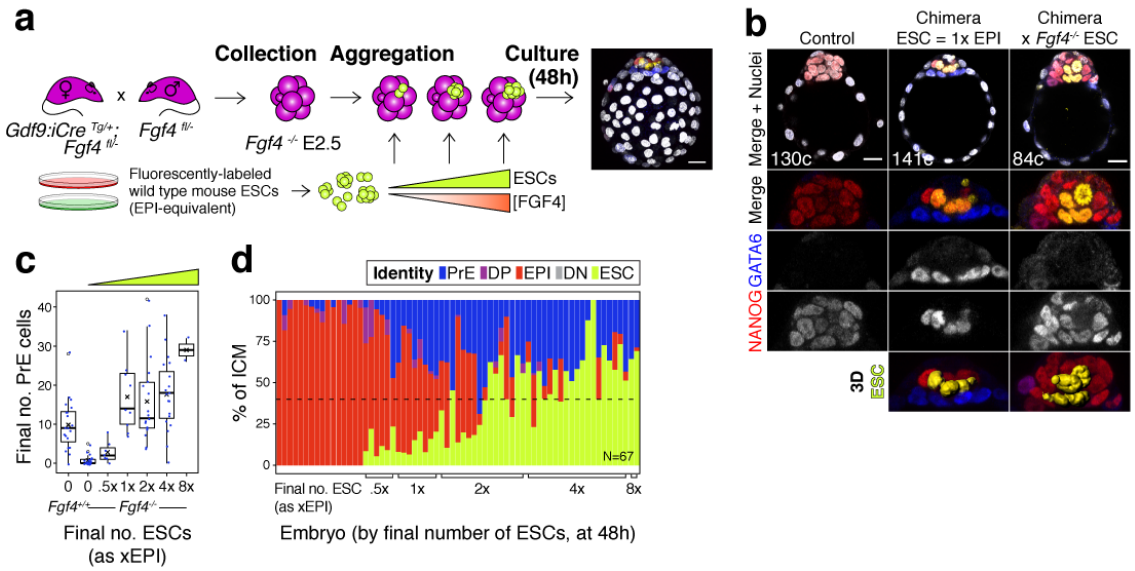
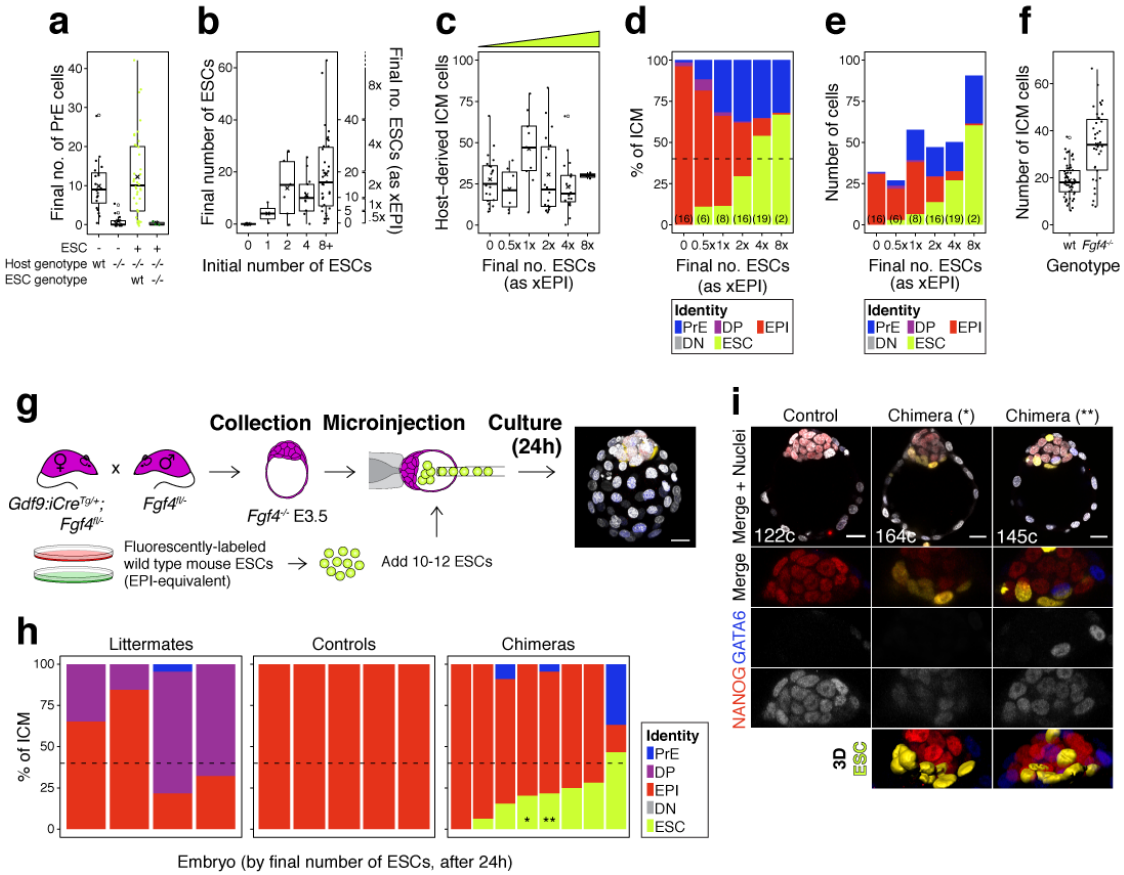


Figure 7. FGF4 provides the dynamic readout of lineage size that determines cell fate specification. (a) Experimental design. Maternal zygotic (mz) *Fgf4*$^{-/-}$ Embryos were recovered from *Gdf9*$^{iCre/+}$; *Fgf4*$^{fl/-}$ females crossed with *Fgf4*$^{fl/-}$ males at the 8-cell stage (2.5 days post-fertilization). Embryos were aggregated with clumps of fluorescently labelled wild type ESCs and cultured for 48-56h, until the late blastocyst stage (equivalent to ~4.5 days post-fertilization). Control embryos were allowed to develop without adding ESCs. Both chimeric and non-chimeric control embryos were fixed at the end of the culture period and labelled with lineage markers to assess ICM composition. **(b)** Optical cross sections through representative chimeras carrying either wild type or *Fgf4*$^{-/-}$ fluorescently labelled ESCs (as indicated) and non-chimeric control embryos labelled for NANOG and GATA6 to identify all ICM cell types. The progeny of the introduced ESCs is labelled in yellow. Total cell count is indicated for each embryo. Lower panels show magnifications of the ICM, with all markers overlaid and for each individual marker in grayscale. Surface renders of ESC compartment within the ICM are shown below. **(c)** Box plots showing absolute number of PrE cells after 48h in wild type, control embryos (no ESCs), *Fgf4*$^{-/-}$ control embryos (no ESCs) and *Fgf4*$^{-/-}$ chimeric embryos, grouped by the size of the ESC compartment, as in Figure 3. **(d)** Stacked bar plot showing the relative ICM composition in individual embryos (controls or chimeras). Each bar represents the ICM of one embryo and bars are arranged by absolute number of ESCs present in the embryo. Brackets on x-axis indicate the number of ESCs in those embryos, relative to the size of the average wt control epiblast (xEPI), as in (c). Color coding is indicated. All optical cross sections are 5μm maximum intensity projections. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. Yellow wedges represent the increasing amount of ESCs in each group. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, DN: Double Negative (for NANOG and GATA6), ESC: embryonic stem cell. Scale bars = 20μm.



Supplementary Figure 8 (related to Figure 7). (a) Box plot indicating the number of PrE cells in wild type control embryos, *Fgf4*$^{-/-}$ controls, *Fgf4*$^{-/-}$ chimeras carrying wild type ESCs and *Fgf4*$^{-/-}$ chimeras carrying GFP-tagged, *Fgf4*$^{-/-}$ ESCs, as indicated. **(b)** Box plots indicating ESC contribution to chimeras. Number of ESCs aggregated with morulae is shown against the final number of ESC in the chimeric embryo after 48h in culture. Right y-axis shows epiblast-equivalent size bins used to categorize chimeric embryos. **(c)** Box pot showing the size of the host-derived ICM component at the end of the experiment in each group of embryos. **(d)** Stacked bar plots showing the average relative ICM composition of *Fgf4*$^{-/-}$ embryos carrying wild type ESCs (shown in Fig. 7d), binned by the final size of the ESC compartment. **(e)** Stacked bar plots showing the average number of each ICM cell type for each group of chimeras shown in (d) and Fig. 7d. **(f)** Box plot showing the ICM size of wild type control embryos and *Fgf4*$^{-/-}$ control embryos. **(g)** Experimental design for blastocyst injection. Embryos were recovered from crosses equivalent to those in Fig. 7a, at the mid-blastocyst stage (~60-80 cells) and 10-12 ESCs injected into the blastocyst cavity before allowing the embryos to develop for 24-30h in culture, until a stage equivalent to ~E4.5 days post-fertilization. **(h)** Stacked bar plots showing ICM composition for individual embryos like those shown in (i), as indicated. Each bar represents the ICM of one embryo and bars are arranged by absolute number of ESCs present. Stars (*, **) denote the bars corresponding to the chimeras shown in (i). Color coding is indicated. **(i)** Optical cross sections

through representative chimeras and control embryos cultured from the blastocyst stage and labelled for NANOG and GATA6 to identify all ICM cell types. The progeny of the introduced ESCs is labelled in yellow. Lower panels show magnifications of the ICM, with all markers overlaid, for each individual marker in grayscale and for ESCs as 3D surface renders. Total cell count is indicated for each embryo. All optical cross sections are 5μm maximum intensity projections. In all box plots whiskers span 1.5x the inter quartile range (IQR) and open circles represent outliers (values beyond 1.5x IQR). Cross indicates the arithmetic mean and each dot represents one embryo. Yellow wedges represent the increasing amount of ESCs in each group. PrE: Primitive Endoderm, DP: Double Positive (for NANOG and GATA6), EPI: Epiblast, DN: Double Negative (for NANOG and GATA6), ESC: embryonic stem cell. Scale bars = 20μm.