Global community effect: large-scale cooperation yields collective survival of differentiating embryonic stem cells

Hirad Daneshpour[1,2], Pim van den Bersselaar[1,2], and Hyun Youk[1,3,4,5*]

[1]Kavli Institute of Nanoscience, Delft, The Netherlands

[2]Department of Bionanoscience, Delft University of Technology, Delft, The Netherlands

[3]CIFAR, CIFAR Azrieli Global Scholars Program, Toronto, ON, Canada

[4]Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA.

[5]Program in Systems Biology, University of Massachusetts Medical School, Worcester, MA, USA.

[*]Corresponding author. E-mail: hyun.youk@umassmed.edu

SUMMARY

"Community effect" conventionally describes differentiation occurring only when enough cells help their local (micrometers-scale) neighbors differentiate. Although community effects are known in various settings, macroscopic-scale community effects - fates of millions of cells entangled across centimeters - remain elusive. By developing systematic approach for determining cell-cell communication distances, we discovered that differentiating mouse Embryonic Stem (ES) cells - scattered over centimeters - form one macroscopic entity via long-range communications. The macroscopic population avoids extinction only if its centimeter-scale density is above a threshold value. Single-cell-level measurements, transcriptomics, and mathematical theory revealed that this "global community effect" arises from differentiating ES-cells secreting, accumulating, and sensing survival-promoting factors, including FGF4, that diffuse over many millimeters and activate Yap1-induced survival mechanisms. Only above-threshold-density populations accumulate above-threshold-concentrations of survival factors. We thus establish that large-scale cooperation underlies ES-cell differentiation. Tuning such large-scale cooperation may engender spatially disconnected, coherently functioning, macroscopic structures for synthetic biology.

Keywords:

Cell population control; secrete-and-sense cells; quorum sensing; autocrine signaling; cooperative behaviors; differentiation; systems biology; stochastic modeling; phase diagrams; synthetic biology

INTRODUCTION

Embryonic Stem (ES) cells, which are important for synthetic biology as ex vivo cell cultures, secrete and sense myriad diffusive factors that control their own proliferation, death, or exit from pluripotency on cell-culture plates (*1*). By secreting and sensing the same molecule - that is, through an "autocrine signaling" (*2*) - an ES cell can communicate with itself (self-communicate) by capturing the molecule that it had just secreted or communicate with other ES cells (neighbor-communicate) due to its molecule diffusing to and being captured by those other cells (*3-6*). Although many autocrine-signaling molecules for ES cells are known - such as the Fibroblast Growth Factors (FGFs) that promote cell proliferation (*7-11*) - it is unclear to what extent each of these autocrine-signaling molecules are used for self- versus neighbor-communication and how each type of communication controls one cell's differentiation. Determining which cell is communicating with which cell is challenging because a molecule does not leave a visible trace of its diffusive path, from a cell that secretes it to a cell that captures it (which may be the same cell). Complicating the matter is that having two cells next to each other does not necessarily mean that they are communicating with each other. This is because a receptor may have a high binding-affinity for the autocrine-signaling molecule (e.g., EGF receptor), meaning that most copies of the molecule can be captured by the cell that secreted them and, therefore, very few remaining copies of the molecule are left for communicating with other cells (*3,12*). Conventional approaches for proving the existence of secreted factors and/or manipulating them – such as transferring medium from one cell-culture plate to another (i.e., transferring a “conditioned medium”) or washing/flowing media over cells with microfluidics – cannot determine the degree of self- versus neighbor-communication for an autocrine factor and the spatial range of a neighbor communication. This is because the conventional approaches - despite providing valuable insights and showing that secreted factors exist - involve either pooling together all molecules from everywhere on a cell-culture plate, uniformly mixing them, and then giving this mixture to cells on a new plate (e.g., in the case of transferring a conditioned medium) or accumulating all molecules with a gradient along one direction (e.g., in the case of microfluidics). Hence, these widely used methods destroy crucial, spatial information such as how far each autocrine factor travels when undisturbed, which cells secreted the factor, and which cell senses the factor. Moreover, while several effects that depend on ES-cell density are known – such cell growth enhanced by accumulation of secreted growth factors (e.g., FGF4) at high cell-densities – it is unclear, due to the ambiguities mentioned above, whether these density-dependent effects are due to a local communication between cells that are packed close to each other or due to cells that are millimeters-to-centimeters apart potentially stimulating one another through long-distance communication (the latter has not yet been established). In other words, there is an ambiguity that remains unsettled: does a cell survive because there are sufficiently many cells nearby that are all helping each other by secreting

growth factors (as can be the case in a high density cell-culture in which cells tend to be near each other and only local communication exists) or does a cell survive because distant cells that are centimeters away help one another grow and survive (as can be the case in a high density cell-culture in which distant cells can communicate but local communication is weak or absent)? Ambiguity arises here because both scenarios can result in the same abundance of growth factors in the pooled medium. Complicating the issue even further is that some ES-cell secreted factors remain unidentified and many well-known ones, such as FGF4, have new roles that are still being elucidated (*11*). Given all these reasons, we generally do not know which pairs of cells - whether they are ES cells or not - are communicating through diffusive molecules and exactly how far apart they can be before communicating becomes impossible. These questions are simple to state and are fundamental to understanding cell-cell communication. But they remain difficult to answer. Consequently, in the context of ES-cell cultures, we currently lack a coherent, quantitative picture of how all the secreted factors - both known and unknown - collectively and spatiotemporally regulate ES cells' proliferation, death, and exit from pluripotency. Such a picture would rigorously reveal to what extent ES cells form a collective entity in which they cooperate to survive and differentiate. Establishing such a quantitative, comprehensive picture may reveal that a large-scale (nonlocal) cooperation exists among differentiating ES cells. More generally, quantitatively establishing the "community size" and thereby a cell's degree of autonomy is a conceptual challenge that is relevant in many contexts, including for microbial communities (*13*) and for rigorously settling an open question of how exactly cell-cell communication may influence the efficiency of reprogramming adult cells into induced Pluripotent Stem Cells (iPSCs) (*14*). The goal of our study is to use ES cell-culture as a testbed to address this conceptual challenge - rigorously establish how autonomous/collective a cell is - by developing a systematic approach that integrates experiments and theory.

In this study, we devised a widely applicable method – that combines experimental and theoretical approaches – to determine, without having to identify all the secreted molecules involved, whether a nonlocal communication exists among adherent cells and, if so, whether and how it controls mouse ES cells' proliferation, death, and differentiation. We discovered that neither a single ES cell nor a few ES cells can survive the differentiation process. Instead, we found that ES cells, while exiting pluripotency, communicate over a macroscopic distance by secreting and sensing molecules, including FGF4, that travel at least several millimeters. Through this nonlocal communication, we found that a differentiating ES cell's survival depends on vast numbers of other cells surviving: the population's density measured over centimeters, which the cells determine through communication as they exit pluripotency, determines - in a switch-like manner - whether the entire population goes extinct or

survives during differentiation. We thus uncovered cooperation that becomes triggered at the start of pluripotency exit and causes ES cells to act as one macroscopic entity that spans many centimeters, contrary to the usual picture of ES cells being autonomous or locally interacting in aggregates or microcolonies. This phenomenon, which we call a "global community effect", extends – for ES cells – the concept of “community effect” that John Gurdon first introduced to describe his discovery of a cooperative differentiation in frog embryos (*Xenopus laevis*) (*15*) whereby a few embryonic cells cannot differentiate into a mesoderm - they die instead - and the cells can differentiate only if they are aggregated together in sufficiently high numbers (*15,16*) because only then the cells can accumulate sufficiently high concentrations of molecules that the cells must secrete and sense to survive differentiation. Since Gurdon's discovery, researchers observed such cooperative differentiations in a variety of cell types and organisms (e.g., mouse T-cells locally interacting with each other in a microwell by secreting and sensing the cytokines, IL-2 and IL-6, that induce differentiation into memory T-cells (*17*)). Past studies of community effects have thus far focused on aggregates or microcolonies of cells that differentiate by locally cooperating with each other (i.e., by communicating with each other in a microenvironment) (*17-20*). But largely unexplored is the question of whether cells can or must cooperate beyond the confines of their microenvironment - by communicating with cells that are millimeters or centimeters away - in order to differentiate. Our study proves that such a global community effect exists by rigorously determining the spatial range of communication and degree of self-versus-neighbor communication. Understanding the full spatial range of communications and the cooperative behaviors that they enable will likely open new doors for engineering synthetic multicellular structures. To that end, our work quantitatively provides a new, potential route for engineering structures in which large numbers of these structures may be engineered to cooperate over a long-range on a plate, beyond the local cooperation that may exist between cells within the same structure. Moreover, by revealing a large-scale cooperation during differentiation, our work provides a basis for and shows that it would be possible to harness and control long-range communication for building a completely synthetic, macroscopic structure - one without an in vivo analogue - consisting of spatially disconnected cells that span many centimeters and yet coherently function as one entity. Such structures may be useful for synthetic biology and help in quantitatively studying how microscopic cells use cell-cell communication to bridge vast length scales to perform a coherent, biological function as well as help in elucidating physical principles that underlie multicellular systems (*21-26*).

RESULTS

Survival is collective during differentiation but not during pluripotency

A common knowledge is that seeding too few ES cells on a cell-culture plate can cause suboptimal growths of cells in various types of cell-culture media. Although this is widely accepted and thus one usually plates enough ES cells to keep many of them "healthy", it is unclear how many is "too few", what actually constitutes "enough" numbers of ES cells, whether or not there is a sharp threshold in the number of cells that separates the "too few" from the "enough" - and if so, what the underlying molecular mechanism is - and the area over which the cell density matters (e.g., # cells / mm^2 or # cells / cm^2?). Currently lacking is a systematic exploration of these issues, in which one precisely sets and varies the initial numbers of ES cells over a wide range and then studies the resulting growth of ES cells. We posited that this phenomenon, which is well known but poorly understood, may be masking a community effect that might be mediated by a long-range communication. We systematically re-examined this phenomenon as our starting point. Specifically, we examined how mouse ES cells proliferate during self-renewal (pluripotency) and differentiation when they are sparsely scattered as near-single cells across a 10-cm diameter plate, at lower densities than is the standard for both self-renewal and differentiation conditions. We kept ES cells pluripotent (undifferentiated) by culturing them without any feeder cells and with Leukemia Inhibitory Factor (LIF) in either a serum-containing (FBS) or a serum-free (2i) medium. To initiate differentiation, we detached the pluripotent cells from plates, resuspended them in PBS, counted their numbers in suspension (Fig. S1), and then randomly scattered a fewer-than-usual number of cells across a 10-cm diameter plate to obtain a desired population density (# of cells per cm^2 of plate area, up to ~15,000 cells / cm^2). This plate contained a "differentiation medium" (N2B27 (*27*)) which, lacking LIF, triggered the cells' exit from pluripotency (Fig. 1A). With a wide-field microscope, we verified that this method of seeding led to virtually every microcolony (colony seeds) on a plate, at the start of differentiation, consisting of just ~5 ES cells on average for every population density that we examined (Fig. 1B - top; each cell's diameter is ~10 μm). Even for some of the highest population densities examined, cells collectively covered less than 1% of the total plate area (Fig. 1B - bottom). Hence, any result that we will discuss below is not due to having highly confluent microcolonies or plates. Two days after initiating differentiation, we added one of two inducers - Retinoic Acid (RA) or CHIR - to induce differentiation towards either a Neural Ectoderm (NE) lineage (with RA) or a Mesendoderm (ME) lineage (with CHIR) (Fig. 1A). An ES cell must first choose to enter either one of these two lineages before becoming any differentiated, somatic cell. We waited two days before adding the inducers because ES cells take at least two days to exit pluripotency by degrading their pluripotency factors (e.g., Oct4) (*28*). To not disturb any diffusible factors that mediate cell-cell communications, we did not shake or move the plates in any way during their days-long incubation and we discarded each plate after detaching all its cells to measure its population density.

We first examined the NE-lineage commitment with RA using a cell line, denoted 46C, that expresses GFP driven by the Sox1 promoter, which becomes expressed early in NE-lineage commitment. We found that cell populations that began with a sufficiently high density (above ~1700 cells / cm^2) grew towards the carrying capacity (Fig. 1C - top row) whereas populations that began with a sufficiently low density (below ~1700 cells / cm^2) approached extinction over six days (Fig. 1C - bottom row). A population that nearly began with a "threshold density" (~1700 cells / cm^2) neither noticeably grew nor shrank during the first six days (Fig. 1C - middle row). But, sometime after, the population either suddenly grew towards the carrying capacity or shrank towards extinction (Fig. S2). Specifically, our method of using an ensemble of many plates that all started with the same density revealed that two populations starting with the same near-threshold density could exhibit two distinct fates: one expanding and one becoming extinct (Fig. S2). Furthermore, using a flow cytometer to measure the percentage of cells expressing GFP over time revealed that higher initial population densities led to higher percentages of cells committing to the RA-induced NE-lineage after six days (Fig. 1C - right column).

We determined the "survival-versus-extinction" fate of a population for a wide range of starting densities (Fig. 1D). Importantly, we observed this phenomenon for all three widely used ES-cell lines (E14, 46C (*27*), and Brachyury-GFP cell lines (*29,30*)) (Fig. 1D and Figs. S3-S5). We also observed this phenomenon occurring for all three possible lineage inductions: NE-lineage induction with RA (Fig. 1D - green and Figs. S3-S5), ME-lineage induction with CHIR (Fig. 1D - purple and Fig. S5), and unguided differentiation (i.e., LIF removal without any inducers) (Fig. 1D - grey and Fig. S5). In all these different conditions, the threshold value that separates the initial population-densities that lead to survival from those that lead to extinction remained nearly identical. Mapping the survival-versus-extinction fate over a wide-range of initial population densities revealed that near the threshold density, the initial population density sharply determines the survival-versus-extinction fate, in a switch-like manner (Fig. S2): a mere two-fold difference in the initial density could mean extinction as opposed to the population surviving (e.g., ~1500 cells / cm^2 leads to extinction whereas ~3000 cells / cm^2 leads to population expansion towards the carrying capacity, despite both densities initially covering less than 1% of the plate area). To our knowledge, this switch-like population-level outcome that so sharply depends on the initial population density has not been reported before, despite the common, colloquial knowledge that ES cells secrete growth factors and thus having more cells should mean a "better" growth. In fact, our quantitative approach revealed an outcome that is qualitatively different from this colloquial notion: it is not that having more cells means increased/faster growth; rather, we uncovered here that the entire population either becomes extinct or expands to the carrying capacity, with the outcome set by a

relatively small difference in initial cell numbers. With a mathematical model and further quantitative experiments, we will later show how this sharp, switch-like response to the initial population density arises and show that the conventional, qualitative notion of "more cells mean enhanced growth" actually corresponds to a qualitatively different outcome than the one that we observed here (i.e., the conventional notion does not hold because if it were true, then it can allow even lowest starting population densities to expand towards the carrying capacity, which contradicts our data in Fig. 1D).

In contrast to the differentiating cells, a pluripotent (self-renewing) population in both serum-free and serum-containing medium always grew towards the carrying capacity regardless of its initial density, even when starting with as low as ~10 cells per cm^2 (Fig. 1D – orange). This is consistent with the fact that one can isolate a colony grown from a single, pluripotent ES-cell in gene-targeting experiments. In our experiments, every population started with microcolonies that were, on average, hundreds of microns apart from one another (Fig. S6) and each microcolony started with ~5 cells on average (Fig. 1B). Crucially, for each initial population density, we used time-lapse microscopy to measure the distance between every pair of colonies in a field-of-view (1.40 mm x 0.99 mm) and then, from these measurements, determined the distance between a colony and its nearest-neighboring colony at the start of the time-lapse movie (i.e., at the start of pluripotency loss). We found that whether a colony dies or not during the next four days and its growth rate were both independent of (uncorrelated with) the distance to the nearest-neighboring colony for any of the initial population densities (Fig. S7). Considering these results together, we reasoned that if cell-cell communication were responsible for above results, then it is likely a nonlocal communication (i.e., over many cell-lengths) that regulates the differentiating cells' proliferation and/or death. Since we observed the same phenomenon for all three lineage inductions and for all three different cell lines, we will focus on the NE-lineage induction with the 46C cell-line for simplicity.

Populations commit to either surviving or becoming extinct in the first 3 days by extracellularly accumulating secreted molecules

We sought to determine whether the density-dependent survival of populations is due to cells secreting and then sensing each other's survival-promoting factors and, if so, to determine the spatial range of this communication. To do so, we used two methods to investigate whether the differentiation medium of a high-density population (5172 cells / cm^2) rescues a low-density population (862 cells / cm^2) from extinction (Fig. 2A). In one method, we collected the high-density population's differentiation medium X days after triggering the differentiation and then triggered differentiation of the low-density of pluripotent

cells in this medium (Fig. 2A - labeled "1"). In this "Day X to day 0" experiment, the low-density population - which would have gone extinct - expanded by ~4-fold after spending 6 days in the medium of the high-density population that was either 2-days-old, 3-days-old, or 4-days-old (Fig. 2B - left column; X = 2, 3, 4). This suggests that the high-density population accumulated enough survival-promoting factors in its medium after two days. In contrast, the low-density population still went extinct in the medium of the high-density population that was just 1-day old (Fig. 2B - left column; X = 1). Taken together, these results suggest that the high-density population cannot secrete and accumulate enough survival-promoting factors in one day whereas in two days, they do. The results also suggest that on the third and fourth days after starting differentiation, the high-density population still has enough survival-promoting factors in its medium to prevent extinction.

Finally, we found the low-density population barely growing in the medium of a 5-days-old high-density population (Fig. 2B - left column; X = 5). Moreover, the low-density population approached extinction when incubated in the medium of a 6-days-old high-density population (Fig. 2B - left column; X = 6). These results suggest that the cells eventually stop secreting the survival-promoting factors at least 5 days after starting differentiation. They also suggest that the molecules degrade so that after 6 days, there are not enough of survival-promoting factors to prevent population extinction.

As a modified version of the "Day X to day 0" experiment above, we collected the medium of the high-density population X days after triggering its differentiation and then transplanted in it a low-density population that was already differentiating, before being transplanted, for the same number of days (X days) (Fig. 2A - labeled "2"). In this "Day X to day X" experiment, the low-density population survived and grew only when we transplanted it on the 2nd or 3rd day of differentiation (Fig. 2B - right column; X = 2, 3). Transplanting on the 2nd day (X = 2) caused the low-density population to grow by ~4-fold whereas it grew less when transplanted on the 3rd day. These results establish that four days after starting differentiation, low-density populations no longer respond to the survival-promoting factors anymore and thus they proceed to extinction.

Multiscale analysis excludes local communication as being a determinant of population's survival-versus-extinction fate

To determine how far the secreted molecules (survival-promoting factors) travel, we used a wide-field microscope to continuously track and measure the area of individual microcolonies in multiple, millimeter-scale (1.40 mm x 0.99 mm) field-of-views on each plate (Fig. 2C). We found that the area of

a colony was an accurate proxy for the number of cells in the colony, leading to the fold-change in number of cells over time (Fig. 1D) closely matching the fold-change in colony area over the same duration (Fig. S8). From the resulting time-lapse movies, we sought to answer whether any cell-cell communication within a microcolony (Fig. 2C - orange arrows) or any communication between nearby microcolonies - defined as being in the same 1-mm^2 field of view (Fig. 2C - purple arrows) - could account for the population's survival-versus-extinction fate. We tracked the area of each microcolony over 4 days (Fig. 2D and Fig. S8). Most microcolonies started with a few cells (~5 cells on average, each cell is ~100 μm^2; see Fig. 1B) and rarely started with more than ~10 cells, as indicated by the distribution of their initial areas (Fig. 2E - horizontal axis). Using the variations in the initial microcolony sizes, we found virtually no correlation between the final area of a colony after 4 days of differentiation and its initial area, for both high-density (e.g., 2857 cells / cm^2) and low-density (e.g., 857 cells / cm^2) populations (Fig. 2E; see Fig. S9 for all population densities). Hence, simple forms of intra-colony communication - each cell secreting a molecule whose intra-colony concentration is high enough to rescue the colony if the colony starts with enough cells - cannot account for the fate of the population on that plate. To be clear, this does not mean that there is absolutely no intra-colony communication. But our analysis here suggests that local (intra-colony) communication cannot be determining whether the entire population survives or not. We will later support this with a mathematical model that will also explain why, if secreted molecules are indeed responsible for the survival-versus-extinction fate, a larger colony does not have a higher probability of surviving due to a higher concentration of these molecules accumulating in the colony (i.e., analyzing reaction-diffusion equation will show that molecules that diffuse far away does not create a sharp local gradient around a secreting cell).

A microcolony's growth rate – a rate at which a microcolony's area grows in the time-lapse movies - was virtually uncorrelated with the total number of microcolonies in its 1-mm^2 field of view (Fig. 2F; see Fig. S10 for all population densities). Hence, the number of microcolonies in a 1-mm^2 region does not dictate whether or not any of the cells in that region survive and their growths. To see this, consider two separate mm^2-scale fields of view - one from a high-density population and another from a low-density population - that appear identical because both initially have the same number of microcolonies. Yet, the field from the high-density population would have microcolonies subsequently expanding while the field from the low-density population would have microcolonies subsequently dying. Hence, a mm^2-scale field of view contains insufficient information: there is no parameter - such as distance between microcolonies (Fig. S11) - that we can extract from a field of view for predicting whether the cells in the field will expand or die. Finally, as previously mentioned, we found that whether a colony dies or not during the next four days and its growth rate were both independent of (uncorrelated

with) the distance to the nearest-neighboring colony - this distance indeed decreases as the initial population density increases - for any of the initial population densities (Fig. S7). Taken together, the time-lapse movies suggest that nonlocally diffusing molecules (i.e., at least over the 1-mm^2 scale) dictate the differentiating population's survival-versus-extinction fate.

Survival-versus-extinction fate of population occurs due to millimeter-scale (nonlocal) communication

If the secreted molecules indeed diffuse over several millimeters, then changing the height of the liquid medium by millimeters should disturb the diffusion profiles (concentration profiles) formed by the secreted molecules (Fig. 2G - blue region) whereas it should not alter the diffusion profiles formed by secreted molecules that diffuse across a less-than-millimeter distance (Fig. 2G - purple region), as dictated by the standard reaction-diffusion equation that governs secreted molecules (see Supplementary text). For example, decreasing the liquid height - which is ~2-mm above the cells (plate bottom) in the experiments thus far - would lower the "ceiling" (top of liquid) that the secreted (long-range) molecules can reach and thereby concentrate more of the molecules at the bottom where the secreting cells are (Fig. 2G). Conversely, increasing the liquid height would let the long-range molecules escape further away from the cells and thereby concentrate less of these molecules in the cells' immediate vicinity. Changing the liquid height would not change the initial population density (number of cells per plate area). In accordance with this line of reasoning, a low-density population (862 cells / cm^2) expanded towards the carrying capacity in a less-than-usual volume of the differentiation medium - having a liquid height of ~0.3 mm - whereas it became extinct in the usual ~2-mm liquid height (Fig. 2H - pink points). Conversely, a high-density population (3448 cells / cm^2) progressively grew less over 6 days as the initial height of the differentiation medium increased, eventually becoming extinct in a ~7-mm high liquid (Fig. 2H - blue points; see Fig. S12 for corresponding percentages of alive cells that differentiated). Together, these results reinforce the conclusion that cells nonlocally communicate, by secreting diffusive factors that spread over at least several millimeters.

Transcriptome analysis suggests that nonlocal communication controls Yap1's activity

To identify the intracellular pathways that the secreted molecules control - which may help us in identifying the secreted molecules themselves - we performed RNA-Seq on a high-density population (5172 cells/cm^2) that expands towards the carrying capacity, a near-threshold-density population (1931 cells/cm^2) that neither appreciably expands or shrinks for 6 days, and a low-density population (862

cells/cm^2) that becomes extinct (Fig. 3A and Figs. S13-S14). We focused on the first 2 days of differentiation - the period of exiting pluripotency - since this is when a population's survival-versus-extinction fate is primarily determined (Fig. 2B - right column). The high-density population, compared to the low-density population, showed enrichment in multicellular processes, cell-cell signaling, neurological processes, and cell adhesion (Fig. S13). We found that many of these enriched processes involve genes that are regulated by Yap1 - a key component of the widely conserved Hippo signaling pathway and a master regulator of transcription for genes that control cell proliferation, apoptosis, and differentiation (*31-34*) (Fig. S13). Yap1 is primarily known as being regulated by cell-cell contact-mediated signaling (*35*), which is intriguing because we found that diffusive factors control a population's survival-versus-extinction fate. Nonetheless, we examined various known, downstream targets of Yap1 in populations of various initial densities. Compared to the low-density population, the high-density population had higher expression of genes that Yap1 directly or indirectly activates (e.g., *Cyr61* and *Amotl2*) and lower expression of genes that Yap1 directly or indirectly downregulates (e.g., *Angptl4* and *Tmem79*) (Fig. 3A and Fig. S14) (*33, 34, 36-41*). These results suggest that Yap1 becomes more active in populations that begin with higher densities.

Nonlocal communication requires active Yap1 during pluripotency exit to control population's survival-versus-extinction fate

We next focused on Yap1's role during a population's approach to extinction. Yap1 exists in either a phosphorylated or a dephosphorylated form (Fig. 3B). The phosphorylated (inactive) Yap1 cannot enter the nucleus to regulate transcription. The dephosphorylated (active) Yap1 enters the nucleus to transcriptionally regulate numerous genes, including anti-apoptotic (e.g., *Bcl2*) and cell-signaling (e.g., *Cyr61*) genes. To first determine if active Yap1 is required for a population's survival-versus-extinction fate determination, we inhibited Yap1 with a 1 μM of verteporfin (VP), which is a standard Yap1-inhibitor (*33, 42*) (Fig. 3B). Pluripotent populations, with the VP in a serum based (FBS) medium, grew normally as if no VP were present (Fig. 3C). But in the differentiation medium, a high-density population (5172 cells/cm^2) now went extinct if it continuously encountered the VP from the start of differentiation whereas it still expanded towards the carrying capacity if the VP encounter started only two days after beginning differentiation (Fig. 3D - left panel). The differentiation medium of a 2-days-old high-density population - which previously rescued the low-density population (862 cells/cm^2) - no longer rescued the low-density population from extinction if the VP was present from the start of pluripotency exit whereas it rescued the low-density population if the VP encounter started only two days after beginning differentiation - the low-density population then grew nearly as much as it did without the VP (Fig. 3D -

right panel). These results establish that the putative nonlocal communication requires Yap1 to be activatable to save a differentiating population from extinction during the first 2 days of pluripotency exit.

Nonlocal communication promotes Yap1's activation of anti-apoptotic genes that counter apoptosis during differentiation

To determine how active Yap1 is in various populations, we used ELISA that specifically detected inactive Yap1 which was phosphorylated at Ser397 - this is a primary phosphorylation site (*31, 43*). For each initial population density, we measured the average amount of inactive Yap1 per cell, 3 days after starting differentiation and relative to the amount of inactive Yap1 per cell found in pluripotent populations of the same density. Compared to the pluripotent cells, cells of the extinction-bound, low-density population (862 cells/cm^2) had only ~10% less inactive Yap1 (Fig. 3E – pink; and Fig. S15) whereas cells of the expanding, high-density populations (5172 cells/cm^2) had ~60% less inactive Yap1 (Fig. 3E – blue; and Fig. S15). Cells of the low-density population that began differentiating in the medium of a 2-days-old high-density population - and thus survived - had ~50% fewer inactive Yap1 than pluripotent cells (Fig. 3E - green; and Fig. S15), similar to the cells of the high-density population. We can compare these levels of inactive Yap1 between populations because we also found that all three differentiating populations expressed nearly the same level of *Yap1* (i.e., same total level of inactive and active Yap1 combined) (Fig. S15). These results establish that, after exiting pluripotency (day 3), cells of proliferating populations have more active (dephosphorylated) Yap1 than cells of extinction-bound populations, consistent with active Yap1 being required for populations to survive pluripotency exit (Fig. 3D).

To further elucidate Yap1's role in deciding a population's survival-versus-extinction fate, we used RT-qPCR to examine the expression of a well-known anti-apoptotic gene, *Bcl2*, that Yap1 regulates (*33*) (see Fig. S16 and Supplementary Table 1 for other Yap1-mediated genes that we examined). Over several days, we examined three differentiating populations: a low-density population (862 cells/cm^2), a high-density population (5172 cells/cm^2), and a low-density population that, after two days of differentiation, was transplanted into and thus rescued by the medium of a high-density population. On each day, we normalized a population's *Bcl2* expression level by that population's *Gapdh* (housekeeping gene) level. Afterwards, we compared each population's normalized *Bcl2* expression level for a given day to that of one-day-old low-density population (whose value is thus "1x" in Fig. 3F). The low-density population's Bcl2 level remained constant at a low value (~1x) while it approached extinction (Fig. 3F - pink). The high-density population's Bcl2 level was nearly identical to that of the low-

density population for the first two days but then it increased by ~10-fold over the next two days (Fig. 3F - blue). The low-density population, when we rescued it right after the 2nd day by giving it the medium of the high-density population, behaved similarly as the high-density population: its Bcl2 level increased by ~10-fold from the 2nd to 4th day of differentiation (Fig. 3F - green). This is consistent with both the rescued and high-density populations, due to extracellular induction, having nearly the same levels of activated Yap1 since Yap1 activates Bcl2 expression (Fig. 3E - blue and green). Together, these results suggest that low-density populations become extinct because they cannot accumulate enough secreted factors to sufficiently activate the anti-apoptotic pathways and overcome the pro-apoptotic pathways whereas higher-density populations do (Fig. 3G). It is known that pro-apoptotic pathways are activated during differentiation (*33, 44-46*) and we verified this with RT-qPCR (Fig. S16). Consistent with this picture, we observed that the protein level of active caspase-3, a well-known apoptosis executioner (*33*), was ~2-fold higher in the low-density population than in the high-density and rescued populations (Fig. S17).

Molecular weights of secreted, survival-promoting factors must be between 50 kDa to 300 kDa

To identify the secreted molecule(s) that activate Yap1, we first determined their approximate molecular weight(s). We used a commercial, membrane-based filter that captures all molecules in a liquid medium that are larger (in kDa) than the membrane's "filter size" (in kDa) while letting through all other molecules in the medium, which are smaller than the filter size. We used filter sizes that ranged from 3 kDa to 300 kDa, all of which are too large for capturing standard media ingredients that are essential for ES-cell growth (Fig. S18 and Supplementary Table 2) (*47*). After two days of culturing a high-density population (5172 cells / cm^2), we flowed its differentiation medium through various filters (Fig. 4A). We then transplanted an extinction-bound, 2-days-old low-density population (862 cells / cm^2) into two types of media. One was the filtered medium from the high-density population which contained all molecules that are smaller than the filter size (Fig. 4A - orange). The second was a fresh differentiation medium into which we dissolved all the molecules from the high-density population's medium that are larger than the filter size (Fig. 4A - purple). A filtered medium with all molecules that passed through the 100-kDa filter, but not through the 50 kDa filter, rescued the low-density population and caused it to expand up to ~3-fold in four days (Fig. 4B - top). Hence some rescuing molecule(s) weigh between 50 kDa and 100 kDa. Moreover, a fresh medium containing all molecules larger than 100 kDa, but not larger than 300 kDa, rescued the low-density population from extinction by causing it to expand up to ~3-fold in 4 days (Fig. 4B - bottom). Hence some rescuing molecule(s) must weigh between 100 kDa and 300 kDa. Together, these results establish that there are at least two species of secreted molecules that determine a

population's survival, whose apparent weights are between 50 kDa and 300 kDa. More precisely, each filter size has, conservatively, an error of +/-50% according to the manufacturer (see Methods). Hence, the secreted molecules of interest can be as low as 25 kDa. Using the range of molecular weights determined here and our RNA-Seq dataset (Fig. 3A and Figs. S13-S14), we next sought to identify and screen potentially secreted factors that fall within the identified molecular-weight range (as we will show, this led us to identifying 11 such potential factors and testing the effect of each one and their combinations on population survival).

FGF signaling is necessary to avoid population extinction

We first focused on two of the smallest candidate secreted factors. Fibroblast Growth Factors (FGFs) - notably FGF4 (~22-kDa) and FGF5 (~29-kDa) - have weights that are close to the ~25-kDa lower limit that we established in our filtration experiments. Our RNA-Seq dataset also showed that cells highly expressed FGF4, FGF5, and several FGF receptors (FGFRs) during the crucial, first two days of differentiation - the time window in which rescuing a population is still possible (Fig. S19). To test whether FGFs are important for determining a population's survival-versus-extinction fate, we blocked FGFRs by incubating various differentiating populations with PD173074, which is a standard FGFR-inhibitor (*48, 49*) (Fig. 4C). Inhibiting FGFRs continuously for 6 days drove every population towards extinction, including those with densities higher than the threshold value (~1700 cells / cm^2) (Fig. 4C - bottom graph). Interestingly, with the FGFRs blocked, having a higher initial density enabled staving off extinction for a longer time: fold-change in population density was closer to but still below one after 6 days for populations with higher initial densities (Fig. 4C - bottom graph). This suggests that even if FGFs turn out to be one of the major autocrine-signaling factors that determine a population's survival, there are other factors (non-FGFs) that promote a population survival (i.e., higher-density populations would accumulate more of these factors than lower-density populations). This is consistent with our finding from the membrane-filter experiments that there are at least two survival-promoting factors (Fig. 4B). Yet, it is important to note that none of the high-density populations expanded - their fold-changes were either far or just below one after 6 days of continuous blockage of FGFRs. This means that FGF-signaling is crucial for population expansion. The FGFR-blockage also drastically decreased the percentage of cells that entered the NE-lineage (expressed GFP) for all populations (Fig. 4C - top graph). These results indicate that FGF signaling, likely involving the highly expressed FGF4 or FGF5, is crucial for determining a population's survival-versus-extinction fate. Of the two FGFs that we found to be highly expressed - FGF4 and FGF5 - we will focus on FGF4, which is one of the most well-studied versions of FGFs.

Adding sufficiently high concentrations of FGF4 prevents population extinction

Having established that blocking FGF receptors drives populations of any density towards extinction, we next investigated whether adding a recombinant mouse FGF4 into the differentiation medium rescues a low-density population (862 cells / cm^2) from extinction (Fig. 4D - line graph). We added various concentrations of FGF4 at the start of the differentiation and then, 6 days later, we measured the population density. The low-density population approached extinction if the FGF4 concentration was below a threshold value (~1 ng/mL = 0.046 nM) whereas above the threshold value, it survived while barely expanding during the 6 days (Fig. 4D - line graph). Despite the minimal growths, the highest added concentration of FGF4 (200 ng/mL) caused the surviving population to expand by ~10-fold after ten days (Fig. 4D - bar graph). These results show that extracellular FGF4 alone is sufficient for rescuing a population from extinction.

High-density populations accumulate detectable amounts of extracellular FGF4 in two days

Adding FGF4, by hand, into the medium rescues low-density populations. The remaining question is whether differentiating cells secrete and accumulate FGF4 during the first two days in which a population's survival-versus-extinction fate is sealed. Consistent with our RNA-Seq dataset, we found with RT-qPCR that pluripotent cells and differentiating cells, during the first 2 days of differentiation, indeed express FGF4 and FGF receptors (FGFR1-4) (Fig. S20). Studies typically assume that cells that express FGF4 also secrete it. But for our purpose, we need to directly detect extracellular FGF4. Specifically, rather than the absolute concentration of extracellular FGF4, we are interested in seeing whether and how fast the concentration of extracellular FGF4 increases over time during differentiation. Hence, we decided to use ELISA to measure the fold change in the concentration of extracellular FGF4 during the first 2 days of differentiation. To do so, we decided to measure the concentration of extracellular FGF4 relative to that of a highly confluent population of pluripotent cells for comparison. The latter is expected to have a high concentration of extracellular FGF4, based on previous studies' finding that pluripotent ES cells highly express FGF4 (*49, 50*). Normalizing all our ELISA measurements of FGF4 concentration by that of the pluripotent population also makes our result interpretable in the event that ELISA does not detect 100% of all FGF4s that are secreted (e.g., due to antibodies not binding to all their targets) (see Fig. S21 for validations of these justifications). With these justifications in mind, we used ELISA to measure the concentration of extracellular FGF4 in the medium of a high-density population ("5x" = 8620 cells / cm^2) during the first two days of differentiation (Fig. 4E and Fig.

S21). We also used ELISA to measure the concentration of extracellular FGF4 in the medium of a highly confluent population of pluripotent cells (population density of ~80x) (Fig. 4E - yellow line). The differentiation medium did not initially have any detectable amounts of FGF4 (Fig. 4E - "day 0"). But during the next two days, the high-density population secreted and gradually accumulated extracellular FGF4, eventually reaching ~10% of the FGF4 concentration that we found in the medium of the pluripotent population (Fig. 4E and Fig. S21).

FGF4 is sufficient and FGF signaling is necessary for population's survival-versus-extinction fate-decision

Together, the results above (Figs. 4C-E) establish that FGF signaling is necessary and that the secreted FGF4 alone is sufficient for determining whether a population survives or becomes extinct (Fig. 4F). To further support this claim, we rescued a low-density population by adding a high concentration of FGF4 into its medium, and then with ELISA found that the rescued cells had nearly identical, high level of activated (dephosphorylated) Yap1 as a high-density population and a low-density population that was rescued by the medium of a high-density population (Fig. S22). Moreover, these FGF4-rescued cells of the low-density population increased their expressions of anti-apoptotic (*Bcl2*) and cell-signaling (*Cyr61*) genes that Yap1 regulates (Fig. S23), just like the cells that were rescued by the high-density population's medium. Finally, we found that FGF4 can indeed diffuse over several millimeters. To determine this, we first determined that extracellular FGF4 in medium without any cells does not appreciably degrade for three days at 37°C (incubation temperature for our cell cultures) (Fig. S24). As an additional support, we found that taking the medium of a high-density population and then incubating it at 37°C without any cells for 4 days only minorly reduced the medium's ability to rescue low-density population (i.e., observed similar fold-changes in the population density of rescued populations) (Fig. S25). This is consistent with the extracellular FGF4, which is contained in the medium taken from the high-density population, being stable for at least 3 days. By using the relatively slow degradation of FGF4 and the relatively small weight of FGF4 (~25 kDa) in the Stokes-Einstein equation that governs diffusive molecules, we mathematically established that FGF4 diffuses over several millimeters (see Supplementary text). To additionally support the idea that diffusion is the primary means of spreading secreted factors, we used a microscope-based time-lapse movie to determine how fast a droplet of small dye molecules spreads in the differentiation medium at 37°C without any cells (Fig. S26). We found that diffusion alone, without any other mechanism such as liquid convection, can quantitatively explain a fast spreading of the dye molecules that we observed, which is consistent with diffusion driving FGF4's long-range spreading.

While FGF4 is sufficient and FGF signaling is necessary for a population to determine its survival-versus-extinction fate, our experiments also show that additional, secreted molecules likely also contribute to determining a population's survival-versus-extinction fate. As an illustration of this, note that a low-density population can expand by ~3-fold in a filtered medium taken from a high-density population (Fig. 4B) whereas it barely expands even after 6 days in a high concentration of added FGF4 (but it does expand by ~10-fold after 10 days (Fig. 4D)). Hence, the filtered medium contains survival-promoting factors other than FGF4. Aside from FGF4 and FGF5, our RNA-Seq dataset revealed nine other autocrine-signaling molecules that are known to regulate cell proliferation and/or death such as VEGFB and PDGFA (Figs. S27). But adding recombinant versions of any of these nine molecules, either one-by-one or in combinations (including in combination with FGF4), did not result in rescuing of a low-density population (Fig. S28). These results establish that being an autocrine-signaling molecule that controls cell proliferation and/or death - a knowledge typically derived without examining the molecule's diffusion length and the extent of self- and neighbor-communication - does not mean that it participates in determining a population's survival-versus-extinction fate.

Mathematical model with one fit parameter recapitulates main experimental observations

To see how the various experimental findings merge into a coherent picture, we built a simple mathematical model that takes some of the experimental findings as its basic ingredients and aims to recapitulate all other experimental findings (Fig. 5A-B and Supplementary text). Our model has just one free parameter and is a modification of another model that we previously built to explain cooperative replications of yeasts at high temperatures (*51*). In our stochastic model, a cell randomly chooses to either divide, die, or remain alive without dividing (Fig. 5A). Each of these choices occur with a certain probability which is determined by, for simplicity, just one molecule (FGF4) that every alive cell secretes at a constant rate. The probability of a cell replicating nonlinearly increases as the extracellular concentration of FGF4 increases (Fig 5B - blue). The probability of a cell dying is constant - it is independent of the FGF4 concentration (Fig. 5B - red) - which we experimentally verified in microscope-based time-lapse movies of cell growths and deaths (Figs. S29-S31). Given that our experiments (Fig. 2H) and the Stokes-Einstein equation both show that FGF4 diffuses across a macroscopic distance (i.e., at least several millimeters), our model assumes, for simplicity, that FGF4 becomes well-mixed (uniformly mixed) throughout the liquid medium immediately after cells secrete it. With this assumption and the probabilities of replicating and dying as described, our model recapitulated all the main experimental results that we have thus far described (Figs. 5C-D). It also predicted new phenomena

which we confirmed with new experiments (Figs. 5E-F) as we will explain below. Crucially, the model agreed with all the data for just a single set of fixed parameter values that were constrained by our experimental data (see Supplementary text).

The model recapitulates the population dynamics as follows. As a population begins to differentiate, more cells die than replicate. This is because FGF4 is initially absent and thus the probability of replicating is initially zero while the probability of dying is non-zero. As the deaths occur, alive cells secrete and extracellularly accumulate FGF4 which, in turn, gradually increases the probability of replicating. If the population's initial density is below a threshold value, all cells would have died before the population had enough time to accumulate sufficient amount of FGF4 to save itself from extinction (Fig. 5C - red). If the population's initial density is above the threshold value, enough FGF4 accumulates and thus the probability of replicating eventually matches and then increases above the probability of dying (Fig. 5C - blue). The FGF4 concentration at which the probabilities of replicating and dying match is the model's sole fit parameter. If the population's initial density is near the threshold value, it either expands or becomes extinct (Fig. 5C - green) - which one of the two occurs is randomly determined - because the probabilities of replication and death nearly match for a prolonged time. This leads to the days-long, nearly static population density seen in our experiments (Fig. S2), until a few more cells stochastically replicating (dying) pushes the population towards expansion (extinction). The model correctly recapitulates the threshold population density at which this stochastic population-level fate occurs (Fig. 5D). The model and analysis of the reaction-diffusion equation for secreted factor also explains why we see global effects but not local effects (i.e., larger colony is not more likely to survive because the far-away diffusing molecule does not locally linger around the secreting cells to create sharp concentration gradients around them; moreover, the factor concentration at which the probability of dying and probability of surviving match (Fig. 5B) contributes to why no local effect is seen) (see Supplemental text).

Aside from recapitulating the existing data, the model quantitatively predicted how populations would behave - whether it survives or becomes extinct - when one simultaneously changes the volume of liquid medium and the initial population density over wide ranges. We confirmed these predictions with additional experiments (Figs. 5E-F and Fig. S32). In particular, the model predicted the precise shape of a phase boundary - precise combinations of volume (height) of liquid medium and initial population density for which a population's survival-versus-extinction fate is random (Fig. 5F - grey curve) - which we have not yet investigated (we investigated the liquid volume for only two initial population densities thus far (Fig. 2H)). We performed additional experiments in which we

simultaneously varied the initial population density and the liquid volume (height) over wide ranges. The results of these experiments closely matched the model's quantitative prediction of the phase boundary shape (Fig. 5F - circles; and Fig. S32). By accurately and comprehensively summarizing the experimental findings into a coherent picture, our simple model supports the conclusions that we have drawn from the experiments.

DISCUSSION

Here we used ES-cell cultures as testbeds to develop a generally applicable, systematic approach for quantitatively addressing a basic question about cell-cell communication that is easy to state and yet is difficult to address: which cells in a population are communicating with one another through diffusive molecules? Addressing this fundamental question has been difficult because conventional methods - such as transferring of conditioned media and microfluidics - destroy the necessary, spatial information such as how far each autocrine factor travels when undisturbed.

A common knowledge states that ES cells secrete growth factors - including FGF4 - and thus having more cells should mean a "better" growth. But our quantitative approach revealed a phenomenon that is qualitatively different from this commonly accepted view: it is not that having more cells simply means increased/faster growth but that, as we uncovered here, the entire population either becomes extinct or expands to the carrying capacity - a population-level fate is decided in a switch-like manner - with this switch being flipped by a relatively small difference - less than a two-fold difference - in the initial number of cells that are scattered across multiple centimeters. This sharp, switch-like fate of a population is qualitatively different from the conventional notion of higher cell densities leading of faster growth because in the latter scenario, population extinction does not need to occur (i.e., contrary to our observation, the conventional view would potentially allow for even a lowest starting population density to expand towards the carrying capacity, albeit more slowly than a high-density population). Moreover, one conventionally thinks of a high cell density to mean a locally high cell density (i.e., cells tightly packed in a cluster or colony) which is distinct from the phenomenon that we uncovered: a population's survival depends on the number of cells spread out over a macroscopic length-scale due to a global communication, not local communication, determining the population's survival-versus-extinction fate.

Even when one knows that secreted growth factors can accumulate at higher cell densities, there is an ambiguity: does a cell survive because there are sufficiently many cells nearby that are all helping each other by secreting growth factors (as can be the case in a high density population in which the

distances between cells may be small and only local communication exists) or does a cell survive because distant cells that are centimeters away help one another grow and survive (as can be the case in high density population in which a global communication exists but insufficient local communication exists due to fast, far-away diffusing molecule that does not linger locally)? Ambiguity arises here because both scenarios can result in the same abundance of growth factors in the pooled medium. In developing an approach to address this ambiguity, we discovered that when mouse ES cells begin to differentiate, they also begin to cooperate on a macroscopic (centimeters) scale to help each other survive the differentiation process. We found that an insufficient amount of cooperation drives a population to extinction and that the "amount" of cooperation is proxied by the population density that is measured over a centimeter-scale. A nonlocal communication, tunable up to at least ~1 cm (Fig. 2H), mediates the large-scale cooperation and involves multiple secreted molecules, a crucial one of which is FGF4. Consequently, ES cells form patches of communities that are several millimeters in size. The same long-range communication then connects these patches of communities and they, in turn, form a global community that spans the entire cell-culture plate. Hence, differentiating ES cells form one macroscopic entity that either lives or dies as one body during pluripotency exit and whose constituents (cells) are distributed across many centimeters yet whose livelihoods depend on one another (Fig. 6 - bottom). We found that this phenomenon occurs in the first ~2 days after removing LIF (i.e., during pluripotency loss and before adding any inducers such as RA), meaning that the phenomenon that we observed is for pluripotency loss - which is relevant for differentiation into any lineage - rather than for differentiation into a specific lineage (e.g., NE-differentiation) (Fig. 1D, Fig. S5). Moreover, we found that the collective survival-versus-extinction behavior is absent during pluripotency (Fig. 1D). We can cast this result as an extension of the classic, community effect that John Gurdon discovered (*15,16*). To date, studies of community effects have focused on how a size of cell aggregates or colonies determines cell fates as recently demonstrated, for example, in T-cells that help each other differentiate into memory T-cells in microwells (*17-20*) (Fig. 6 - top). Our work extends this classic framework for cooperative differentiation by showing that ES cells exhibit a "global community effect", in which isolated single cells cooperate across macroscopic distances to determine their fate - dying or living and differentiating - through a long-range communication (Fig. 6 - bottom).

The global community effect for ES cells is reminiscent of quorum sensing (*52*). Indeed, autocrine signaling with negligible self-communication - meaning that each cell captures very few copies of its own secreted molecule - can lead to quorum sensing in some contexts (*3-6*). But an autocrine-signaling molecule, in general, does not ensure a quorum-sensing-type behavior because it enables both a self-communication (i.e., cell captures its own molecule) and a neighbor-communication (i.e., cell

sends the molecule to another cell). Despite the fact that we know of many autocrine-signaling molecules that promote growth of ES cells, including FGF4 (*7*-*11*), it has generally remained unclear to what degree a cell stimulates its own growth (self-communication) versus being stimulated by the other cells (neighbor-communication) and, if neighbor-communication occurs, then which pairs of cells stimulate one another (either nearby or distant cells) and by how much (and how far apart two cells can be to stimulate one another through the autocrine growth factors). In other words, every cell can be secreting the same autocrine-signaling molecule that promotes the growth a cell that senses it, but that does not necessarily mean that every cell is helping each other grow. For example, EGF receptors have such a high binding affinity for EGF that an EGF-secreting cell can capture most of its own copies of secreted EGF (*12*), which would result in negligible or highly reduced neighbor-communication (*3*-*6*). Resolving the issues mentioned above regarding self- versus neighbor-communication requires systematically and quantitatively decoupling communication at different length-scales, which we have done here. Consequently, we found that a long-range communication (neighbor-communication) occurs for differentiating ES cells through at least two secreted-and-sensed molecules, a crucial one of which is FGF4. Interestingly, researchers recently showed that cells in the inner cell mass collectively control the relative abundances of the three cell types that form in a mouse blastocyst, by communicating through FGF4 at length scales of up to ~100 microns (*11*). Our work extends the range of FGF4 to tens of millimeters. According to our calculation of typical diffusion lengths (see Supplementary text), we expect that other long-range communications among cells and colonies are likely and thus other global community effects likely remain to be discovered for ES cells. One can follow the systematic approach that we demonstrated here: start by sparsely seeding the cells (Fig. 1B) and performing the multiscale analysis as we have (Fig. 2). Our approach is simple. Thus, we expect it to be easily adaptable to a variety of cell cultures.

Cell cultures, despite their limitations of not faithfully capturing numerous in vivo processes, are still important testbeds in synthetic biology for revealing principles of cell-cell communication and for engineering synthetic, multicellular systems. On this note, we note that while ES cells are derived from the inner cell mass and the global community effect occurs just after removing LIF - a stage that mimics early development whereby blastocysts become epiblasts that arise prior to the gastrulation that precedes the generation of the three germ layers - one must be careful in translating any results from ES cell cultures to in vivo embryos because ex vivo growth conditions clearly differ from the in vivo conditions (i.e., inside the blastocyst). But our discovery of the long-range communication still reveals an important design principle that is relevant for tightly packed, differentiating ES cells inside ex vivo colonies or an in vivo blastocyst. Namely, the extremely long-distance nature of autocrine-signaling

molecules such as FGF4 tells us the following scenario. Suppose that a cell must detect some minimum concentration of a molecule to live and that its nearby cell must supply this molecule by secretion. If the molecule has a long diffusion length, then the nearby cell would need to secrete it at a higher rate than if the molecule had a shorter diffusion length because the more diffusive molecule spreads out more and thus concentrates less in a region than the less diffusive molecule. Alternatively, there would need to be more of the molecule-secreting cells surrounding the molecule-receiving cell (receiver) in order to elicit the receiver's response - to create more than the threshold concentration required for response - which creates an ideal setting for engineering or naturally developing a switch-like response to a local population density (i.e., create locally collective behaviors). Trade-offs seen in these two examples, derived from studying macroscopic distances, are important to consider to better understand design principles of local communication among cells in compact aggregates. In addition to revealing such design principles, we expect that using our approach to study cell-cell communication in cell cultures will aid in rigorously investigating how communication among cells - over distances not yet specified - may influence the success rate of reprogramming adult cells into induced Pluripotent Stem Cells (iPSCs) on a plate, which has been hypothesized but difficult to rigorously evaluate (*14*).

Given that ES cells are important, ex vivo tools for synthetic biology, it would be useful to examine whether global community effects also exist among ex vivo, multicellular structures on plates such as synthetic blastocysts (*53-57*). Indeed, there are long-standing hints, but not yet quantitative proofs, of community effects occurring among ex vivo structures - these structures may be sharing their secreted mitogens to help each other develop (*58,59*). This, combined with our study, suggests that one may find global community effects among ex vivo multicellular structures, in which large numbers of multicellular structures communicate amongst them on a plate. Using our type of systematic approach, one may uncover such global community effects which may involve both secreted molecules (*17,60-65*) and cell-cell contact-mediated communication (*66-69*). The latter did not enter our analysis as we sparsely seeded near-single cells. If there are global community effects among synthetic multicellular structures, then one may engineer cooperative decision-making circuits by dynamically (temporally) tuning the secretion and sensing of autocrine-signaling molecules and controlling cells' temporally changing response to the molecules. Indeed, we found that the large-scale cooperation among the differentiating ES cells involves dynamically tuned secretion of and response to extracellular factors (e.g., Fig. 2B). Our work shows that manipulating the macroscopic environment - changing the liquid medium as in our study and, equivalently, changing the macroscopic shape of culture plates - which may be more accessible than genetic manipulations, can tune the global community effects. Such environmental manipulations may be promising yet underexplored routes for building ex vivo synthetic

structures (*21,70*). Moreover, by revealing a large-scale cooperation during differentiation, our work provides a basis for and shows that it would be possible to tune long-range communication and cooperation (*71*) to build a completely synthetic, macroscopic structure - one without an in vivo analogue - that consists of spatially disconnected cells spanning many centimeters and coherently functioning as a single entity. Such structures, together with our systematic approach for determining spatial range of cell-cell communication, may provide new testing grounds in synthetic biology for quantitatively understanding how microscopic cells can bridge vast length-scales to perform a coherent, biological function as well as help in elucidating physical principles that underlie multicellular systems (*21,72-78*).

METHODS

Cell lines and cell-culture media. We used three murine ES cell lines: E14Tg2A (129/Ola), 46C and Brachyury-eGFP. The 46C cell line (Sox1 promoter driving GFP) was previously described by Ying et al. (*27*) and was a kind gift from Austin Smith whose lab constructed it by targeting GFP to the endogenous Sox1 locus. Thus, the 46C cells had the Sox1 promoter controlling a GFP expression. The Brachyury-eGFP cell line was previously described by Fehling et al. (*29*) and Pearson et al. (*30*) and was a kind gift from Valery Kouskoff whose lab constructed it by targeting eGFP to the endogenous Brachyury (T) locus. Thus, this cell line had the Brachyury (T) promoter controlling an eGFP expression. To keep ES cells pluripotent, we routinely (every 2 days, with 1/10 dilution) passaged them in either a serum-based (FBS) or a serum-free (2i) pluripotency medium. The serum-based medium (denoted serum+LIF) consisted of high-glucose DMEM (Gibco) supplemented with 15% fetal bovine serum (Gibco, ES qualified), 1X MEM non-essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), 1X glutaMAX (Gibco), 0.1 mM 2-mercaptoethanol (Gibco), 1000 U/mL penicillin-streptomycin (Gibco) and 1000 U/mL Leukemia Inhibitory Factor (LIF, Polygene). The serum-free medium (denoted 2i+LIF) consisted of approximately half-and-half mixture of Neurobasal (Gibco) and DMEM/F12 (Gibco) abd was supplemented with 1X MEM non-essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), 1X glutaMAX (Gibco), 1X N-2 (Gibco), 1X B-27 minus vitamin A (Gibco), 0.1 mM 2-mercaptoethanol (Gibco), 50 μg/mL BSA (Sigma, fraction V), 1000 U/mL penicillin-streptomycin (Gibco), 1000 U/mL Leukemia Inhibitory Factor (LIF, Polygene), 3 μM CHIR99021 and 1 μM PD0325901. Cells were maintained in the pluripotency medium on 10-cm diameter tissue-culture plates (Sarstedt, TC Dish 100 standard) that were coated with 0.1% gelatin in water (Sigma, from bovine skin Type B) at 37°C for at least 20 minutes prior to seeding cells.

Differentiation experiments. To initiate differentiation, we detached and collected ES cells from plates by incubating them with 1 mL of accutase (Gibco, StemPro Accutase Cell Dissociation Reagent) in 37°C for 5 minutes. After collecting the cells, we washed them twice with PBS and then centrifuged them to remove any remaining accutase from the resulting cell pellet. We then resuspended the cell pellet in N2B27 and then counted the number of cells per mL in this resuspension, as described in the "Cell counting" section below. Afterwards, we calculated the volume of this resuspension required to achieve a desired number of cells per cm^2 on a plate and then pipetted this volume into a tube containing 10 mL of N2B27 that was pre-warmed N2B27 to 37 ^{O}C. We then transferred this onto a 10-cm diameter plate whose bottom was coated with a 0.1%-gelatin. We distributed the cells across the area of the plate by gently shaking the plate and then incubated the cells at 37°C with 5% CO_2. Importantly, we did not disturb the plate for at least 6 hours after the plating to allow the cells to sediment and attach to the gelatinized plate bottom. We defined this moment to be the start of differentiation time-course. Cells were left for 2 days in N2B27 and then the spent medium was replaced with either fresh, pre-warmed N2B27 (for unguided differentiation), or N2B27 supplemented with 500 nM of Retinoic Acid (for Neural Ectodermal differentiation), or N2B27 supplemented with 3 μM of CHIR99021 (for Mesendodermal differentiation). We then left the plate for further incubation at 37°C with 5% CO_2. Subsequently, we collected the cells from plates for counting (see "Cell counting" section below) and flow cytometry (see "Flow cytometry" section below).

Flow cytometry. We collected cells using accutase, washed them with PBS, resuspended them in PBS + 4% FBS and kept them on ice before measurements with a flow cytometer. We used a BD FACSCelesta system with a High-Throughput Sampler (HTS) and lasers with the following wavelengths: 405 nm (violet), 488 nm (blue) and 561 nm (yellow/green). We calibrated the FSC and SSC gates to detect only mouse ES cells (FSC-PMT = 231 V, SSC-PMT = 225 V, GFP-PMT = 476 V; as a control,

flowing plain PBS yielded no detected events). We measured the GFP fluorescence using the FIT-C channel. We analysed the flow-cytometry data using FlowJo and custom MATLAB script (MathWorks).

Time-lapse microscopy. We used a wide-field microscope (Nikon SMZ25) to track the growth of microcolonies as a function of the initial population density (E14Tg2A cell line). Cells were cultured on a 6-cm diameter tissue-culture plates (Sarstedt, TC Dish 60, Standard) coated with 0.1% gelatin (Sigma, from bovine skin Type B) in water, fed with 5-mL differentiation medium (N2B27) and incubated inside a temperature-, CO2- and humidity-controlled microscope chamber (Okolab) with steady conditions of 37°C with 5% CO_2. We imaged microcolonies with the following initial population densities (in # cells / cm^2): 476, 857, 1285, 1714, 2142, 2857, 3571 and 4285 cells per cm^2 plate area. Previously, we kept cells in pluripotency medium on 10-cm diameter dishes and routinely passaged (every 2 days). Importantly, cells were allowed for approximately 6 hours to settle and attach to the gelatinized bottom of the plate before the image acquisition. The microscope's mono acquisition settings included a 1X microscope objective, 90.0X magnification, 28.5 (arbitrary units) of DIA intensity, 300-ms exposure time and 2.2X analog gain. Before image acquisition we picked a total of 17 fields-of-view that were evenly spread across the entire 6-cm diameter plate with each field-of-view being 1399.16 μm x 994.95 μm. Images were acquired with 1-hour intervals in a total of 96-hour duration (i.e., 4 days of imaging) during which cells were maintained in N2B27 without refreshing and thus disturbing cells. We analyzed the microscope images using custom MATLAB script (MathWorks). We found at most ~10 colonies per field-of-view for the lowest population densities and at most ~50 colonies per field-of-view for the highest population densities. For each population density, we analyzed three separate plates, each in a different week (three biological replicates).

Cell counting. We detached cells from a cell-culture plate with accutase and then washing them twice with PBS. We used a brightfield microscope (Motic AE31, 100X total magnification) and a hemacytometer (Marienfeld Buerker, #631-0921) to count alive cells by excluding dead cells with a Trypan Blue staining (i.e., dead cells appear blue). We counted the total number *N* of alive, non-stained cells in all 9 large squares – consistently excluding alive cells on two out of the four edges of each square. This way of counting enabled us to determine the total number of harvested, alive cells per unit of volume (mL) by back calculating, with the following formula: *N* x [dilution factor] x 10,000. In case we did not visibly detect any alive cells with this hemacytometer-based method of counting, we used a flow cytometer to estimate the cell counts (see "Flow cytometry" section and Fig. S1).

Medium-transfer experiments. We collected liquid medium of a high-density population (5172 cells/cm^2), centrifuged it at 200g for 5 minutes to pellet and eliminate any remaining cells or other debris from the medium, and then transferred the medium to a low-density population (862 cells/cm^2) after first removing the liquid medium of the low-density population. Specifically, we did this medium transfer in two ways. In one scenario (Fig. 2A – labelled as "1"), we collected the medium of a high-density population as described above on day X – the X here means X days after we initiated differentiation– and then incubated a low-density population in this medium to initiate its differentiation (i.e., the low-density population was in a self-renewing state before this medium transfer). In the second scenario (Fig. 2A – labelled as "2"), we collected the medium of a high-density population as described above on day X, and then incubated in it a low-density population that was differentiating for X days in its own medium. In this method, we measured the population density of the low-density population four days after the medium transfer (i.e., X+4 days) rather than on the same day for all values of X. This ensured that we could fairly compare the different low-density populations (i.e., same number of days (4) spent in the medium of the high-density population).

Medium-filtration experiments. We collected the liquid medium of a high-density population (5172 cells/cm^2) 2 days after we initiated its differentiation, centrifuged the medium at 200g for 5 minutes to eliminate any cells and other debris from the medium, and then transferred the medium to a second centrifugal tube for ultrafiltration. The filter unit consisted of two compartments that were physically separated by a regenerated cellulose membrane which separates soluble molecules, depending on their molecular size and shape. Specifically, the membrane has pores that either pass or hold soluble molecules based on their molecular weight (in kDa) during a high-speed centrifugation. We used filter sizes of 3 kDa (Merck, Amicon Ultra-15 Centrifugal Filter Unit, UFC900324), 30 kDa (idem, UFC903024), 50 kDa (idem, UFC905024), 100 kDa (idem, UFC910024) and 300 kDa (Merck, Vivaspin 20 centrifugal concentrator, Z629472). We centrifuged the medium of the high-density population with spin times specified by the manufacturer. After the filtration, the centrifugal tube with the membrane filter contained two supernatants, each in separate compartments: one that contained all molecules that were larger than the filter size - this portion was much less than 1 mL and stayed on top of the membrane filter - and one that contained all molecules that were smaller than the filter size. We added the supernatant containing larger-than-filter-size molecules to a 10-mL N2B27 with 500 nM of retinoic acid (for neuroectoderm differentiation). In this mixed medium, we incubated a low-density population that had been differentiating for 2 days in its own N2B27. The results of this experiment are in the bottom graph of Fig. 4B. For a second experiment, we added 500 nM of retinoic acid to the ~9 mL of the supernatant that contained all the molecules that were smaller than the filter size. We then incubated in it a low-density population that had been differentiating for 2 days in its own N2B27. The results of this experiment are in the top graph of Fig. 4B. According to the manufacturer, in order to ensure that one captures proteins of a desired molecular weight, one needs to use a filter size that is at least two times smaller than the desired molecular weight. This sets a conservative safety/error margin that we took into account in the conclusions that we drew from the results in Fig. 4B, as explained in the main text. Finally, the few large molecules (> 3 kDa) that are ingredients of N2B27 were previously shown to have either no effect or a small growth-promoting effect on ES cells (i.e., they do not inhibit ES cell growth) (*47*).

RNA-Seq. We performed RNA-Seq on 46C populations of three different initial densities which were (in # cells/cm^2): 862, 1724 and 5172. These populations were undergoing an unguided differentiation (i.e., in N2B27 without any inducers) and we examined their transcriptome 1 day after and 2 days after initiating their differentiations (Fig. 3A). We also performed RNA-Seq on a pluripotent 46C population, which would show the initial transcriptome of the three differentiating populations (Fig. 3A - first column). To perform RNA-Seq, we collected cells from each population and then centrifuged them using a pre-cooled centrifuge. We then extracted RNA from each cell pellet using the PureLink RNA Mini Kit (Ambion, Life Technologies) according to the manufacturer's protocol. We next prepared the cDNA library with the 3′ mRNASeq library preparation kit (Quant-Seq, Lexogen) according to the manufacturer's protocol. We then loaded the cDNA library onto an Illumina MiSeq system using the MiSeq Reagent Kit v3 (Illumina). We analyzed the resulting RNA-seq data as previously described in Trapnell et al., *Nature Protocols* (2012). We performed the read alignment with TopHat, read assembly with Cufflinks, and analyses of differential gene expression with Cuffdiff. As a reference genome, we used the genome sequence of *Mus musculus* from UCSC (mm10). We performed enrichment analysis of genes based on their FPKM values (i.e., more than 2-fold expressed when two initial population densities are compared) by using GO-terms from PANTHER and a custom MATLAB script (MathWorks). We visualized the results of pre-sorted, Yap1-related genes (*33,34,36-41*) as heat maps which displayed the normalized expression value (row Z-score) for each gene and each condition.

RT-qPCR. We performed RT-qPCR on 46C populations of two different initial densities which were (in # cells/cm^2): 862 and 5172. We observed them on each of four days of differentiation in N2B27 that was

supplemented with 500 nM of Retinoic Acid (for Neural Ectodermal differentiation). We collected the cells and extracted their RNA with PureLink RNA Mini Kit (Ambion, Life Technologies) according to the manufacturer's protocol. Then, we reverse transcribed RNA into cDNA using iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad). Next, we performed qPCR in 10-μL reactions with iTaq Universal SYBR Green Supermix (Bio-Rad) and 100 nM of forward and reverse primers. We verified the primer specificity and primer-dimer formation by using the melt curve analysis which showed one peak. See the list of primers that we used in Supplementary Table 1. We normalized all expression levels for a population by that population's *Gapdh* expression level. We then further normalized the resulting value for a gene by dividing it by that gene's normalized (by Gapdh) expression level in one-day-old low-density population. We performed all reaction in triplicates on a QuantStudio 5 Real-Time PCR System (Thermo Fisher).

Attempting to rescue low-density population by supplementing its medium with candidate molecules. We examined whether we could rescue a low-density population from extinction by adding one or a combination of 11 different autocrine-signaling molecules (all recombinant versions from mouse/human). We considered 46C cells in a low-density population (862 cells/cm^2 initially) After 2 days of culturing in N2B27, we added one or combinations of the following recombinant proteins to the medium: 200 ng/mL of recombinant mouse FGF4 (R&D Systems, #7486-F4), 200 ng/mL of recombinant human FGF5 (R&D Systems, #237-F5), 100 ng/mL of recombinant mouse PDGFA (Novus, NBP1-43148), 100 ng/mL of recombinant mouse VEGFB 186 (Novus, #767-VE), 100 ng/mL of recombinant mouse VEGFA (Novus, #493-MV), 500 ng/mL of recombinant human CYR61/CCN1 (Novus, #4055-CR), 500 ng/mL of recombinant human CTGF/CCN2 (Novus, #9190-CC), 200 ng/mL of recombinant mouse CLU (Novus, #2747-HS), 500 ng/mL of recombinant human HSPA8/HSC70 (Novus, #NBP1-30278), 1000 ng/mL of recombinant human PPIA (Novus, #NBC1-18425), or 2000 ng/mL of mouse recombinant SCF (STEMCELL, #78064). After incubating in a medium containing one or a combination of these molecules for four days, we collected the cells for counting (see “Cell counting” section) and flow cytometry (see “Flow cytometry” section) to determine whether the population survived or not and its differentiation efficiency. Results of these experiments are in Fig. S28.

FGF4 ELISA. We measured concentrations of extracellular FGF4 in 10-mL liquid media (N2B27) as follows. We used Mouse FGF4 ELISA Kit (ELISAGenie / Westburg, MOES00755) and followed the manufacturer's protocol. The assayed involved measuring the absorbance at 450 nm for various samples as a direct measure of the FGF4 concentration in the sample. We verified the absorbance signals are real and sufficiently high relative to the lower detection-limit of the ELISA kit by constructing a standard curve (Fig. S21). We measured the absorbances on a Synergy HTX Multi-Mode Reader (BioTek). We measured the concentration of extracellular FGF4 relative to that of a highly confluent population of pluripotent cells for comparison. It makes sense to normalize all our ELISA measurements of FGF4 concentration by that of the pluripotent population because the antibodies used in the ELISA kit may not detect 100% of all FGF4s that are secreted (e.g., due to antibodies being designed for specific, recombinant forms of the mouse FGF4). By using three different forms of FGF4 - one from the ELISA kit, FGF4 secreted by the cells in our experiments, and a recombinant form of FGF4 from a different company - we found that indeed the ELISA kit does not detect all forms of mouse FGF4 but that it does detect the form secreted by the cells in our experiments, though less efficiently than it detects the recombinant FGF4 that accompanied the ELISA kit (Fig. S21).

Phospho-Yap1 ELISA. We examined the endogenous levels of phosphorylated Yap1 protein in four different conditions (Figs. 3C-E and Fig. S15). For each measurement, we collected cells in 10-mL tubes, counted the total number of collected cells with the counting method described in “Cell counting”

section, and then centrifuged them to form a pellet. We then lysed the cells with a lysis buffer (Cell Signaling Technology, #9803) and 1 mM of PMSF (Sigma-Alderich, P7626). We incubated the cell lysates with Phospho-YAP (Ser397) rabbit antibody and performed a sandwich-ELISA assay by using PathScan Phospho-YAP (Ser397) Sandwich ELISA Kit (Cell Signaling Technology, #57046). We then used a Synergy HTX Multi-Mode Reader (BioTek) to measure each sample's absorbance at 450 nm. The absorbance is a direct measure of the abundance of phosphorylated Yap1. To compare the different absorbances, we constructed a standard curve by serially diluting a lysate of pluripotent cells (Fig. S15). We used the standard curve to report the levels of phosphorylated Yap1 in all differentiating populations.

Caspase-3 assay. We measured the protein-level activity of Caspase-3 (well-known apoptosis executioner inside cells) in E14 cells that were maintained in a pluripotency medium (serum+LIF) or were differentiating in N2B27 with 500 nM of Retinoic Acid. We examined three differentiating populations: high-density population (6896 cells/cm^2 initially), a low-density population (517 cells/cm^2 initially), and a low-density population that was rescued by the medium of the high-density population after two days of differentiating. We collected the cells of each of these populations and then performed a membrane-permeable DNA-dye-based assay that measures the amounts of active Caspase 3/7 in intact, alive cells (NucView 488 Caspase-3 Assay Kit for Live Cells). We followed all steps according to the manufacturer's protocol. We used a flow cytometer to measure the amounts of active Caspase-3 in single cells. We normalized the Caspase-3 level per cell by the average Caspase-3 level of an ES cell (i.e., mean level per cell of the pluripotent population.) Results of these experiments are in Fig. S17.

Inhibiting FGF receptors. We examined the fold changes in population densities after several days of inhibiting FGF receptors (FGFRs) with a small-molecule inhibitor, PD173074 (PD17, Tocris, #3044) (previous studies characterized this inhibitor: see (*48,49*)). We used 46C cells differentiating into the NE lineage. We examined the following initial population densities (in # cells / cm^2): 172, 431, 862, 1931, 5172, 8620 and 15517. To inhibit the FGFRs, we added 2 μL of 10-mM PD173074 (PD17) to a 10-mL N2B27 medium. We dissolved the stock of PD17 in DMSO to a final concentration of 2 μM (1056 ng/mL). After 6 days, we measured each population's density (see “Cell counting” section) and differentiation efficiency (see “Flow cytometry” section). As a control, we examined the effect of adding 2 μL of DMSO to cell-culture medium without any PD17. This ensured that our results were not due to any side effects of having DMSO that was carried over with the PD17 that we added to each medium. Results of these experiments are in Fig. 4C.

Verteporfin experiments. We examined the fold changes in population densities after several days of incubation verteporfin (R&D Systems, #5305), which prevents active Yap1 from entering the nucleus to control expression of its myriad target genes. We used 46C cells that were differentiating into the Neuroectoderm lineage. We supplemented the N2B27 with 1 μM of verteporfin that was dissolved in DMSO (based on LeBlanc et al., (*33*)). After 6 days, we measured each population's density (see “Cell counting” section) and differentiation efficiency (see “Flow cytometry” section). As a control, we examined the effect of adding only DMSO to cell-culture medium without any verteporfin. This ensured that our results were not due to any side effects of having DMSO that was carried over with the verteporfin that we added to each medium. Results of these experiments are in Figs. 3C-D.

Mathematical model. Section 2 of the Supplementary text describes the mathematical model in detail and states the parameter values that we used for the stochastic simulations in Figs. 5C-F.

SUPPORTING INFORMATION

Supplementary Figures 1-32

Supplementary Tables 1-2

Supplementary Text

ACKNOWLEDGEMENTS

We thank Thomas Fazzio, Sally Lowell, Amir Mitchell, and Erwin Frey for insightful discussions and/or comments on our manuscript; Marloes Arts for help with experiments; Brian Analikwu and William Verstraeten for help with analyzing the RNA-Seq data; and members of the Youk laboratory for discussions. We thank Austin Smith for sharing his 46C cell line, Valerie Kouskoff for sharing her Brachyury-GFP cell line, and Angie Rizzino. H.Y. was supported by the European Research Council (ERC) Starting Grant (MultiCellSysBio, #677972), Netherlands Organisation for Scientific Research (NWO) Vidi Award (#680-47-544), CIFAR Azrieli Global Scholars Program, and EMBO Young Investigator Award.

AUTHOR CONTRIBUTIONS

H.D. and H.Y. conceived the project and designed the experiments. H.D. performed all the experiments and data analyses. P.vd.B. assisted with some experiments and performed the mathematical modeling with help from H.D. H.Y. provided overall guidance. H.D. and H.Y. wrote the manuscript with inputs from P.vd.B.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY

All data supporting the conclusions of this study are available in this paper. Data that support the findings of this study are also available from the corresponding author upon reasonable request. RNA-Seq data in this work is available on NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE157642.

REFERENCES:

1. Ngangan AV, Waring JC, Cooke MT, Mandrycky CJ, McDevitt TC: Soluble factors secreted by differentiating embryonic stem cells stimulate exogenous cell proliferation and migration. *Stem Cell Res Ther* **5**:26 (2014).

2. Sporn MB, Todaro, GJ: Autocrine secretion and malignant transformation of cells. *N Engl J Med* **303**:878-880 (1980).

3. Youk H, Lim WA. Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science* **343**:1242782 (2020).

4. Coppey M, Berezhkovskii AM, Sealfon SC, Shvartsman SY: Time and length scales of autocrine signals in three dimensions. *Biophys J* **93**:1917-1922 (2007).

5. Doganer BA, Yan LKQ, Youk H. Autocrine signaling and quorum sensing: Extreme ends of a common spectrum. *Trends in Cell Biology* **26**:262-271 (2016).

6. Maire T, Youk H. Molecular-level tuning of cellular autonomy controls the collective behaviors of cell populations. *Cell Systems* **1**:349-360 (2015).

7. Wilder PJ, Kelly D, Brigman K, Peterson CL, Nowling T, Gao QS, McComb RD, Capecchi MR, Rizzino A: Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Dev Biol* **192**:614-629 (1997).

8. Kook SH, Jeon YM, Lim SS, Jang MJ, Cho ES, Lee SY, Choi KC, Kim JG, Lee JC: Fibroblast Growth Factor-4 enhances proliferation of mouse embryonic stem cells via activation of c-Jun signaling. *PLOS One* **8**:e71641 (2013).

9. Raina D, Stanoev A, Bahadori A, Protzek M, Koseska A, Schröter C: Cell-cell communication through FGF4 generates and maintains robust proportions of differentiated cell fates in embryonic stem cells. *BioRxiv* doi:10.1101/2020.02.14.949701 (2020).

10. Mistri TK, Arindrarto W, Ng WP, Wang C, Lim LH, Sun L, Chambers I, Wohland T, Robson P: Dynamic changes in Sox2 spatio-temporal expression direct the second cell fate decision through Fgf4/Fgfr2 signaling in preimplantation mouse embryos. *Biochem J* **475**:1075-1089 (2020).

11. Saiz N, Mora-Bitria L, Rahman S, George H, Herder JP, Garcia-Ojalvo J, Hadjantonakis AK: Growth-factor-mediated coupling between lineage size and cell fate choice underlies robustness of mammalian development. *eLife* **9**:e56079 (2020).

12. DeWitt AE, Dong JY, Wiley HS, Lauffenburger DA. Quantitative analysis of the EGF receptor autocrine system reveals cryptic regulation of cell response by ligand capture. *J Cell Sci* **114**:2301-2313 (2001).

13. Dal Co A, Van Vliet S, Kiviet DJ, Schlegel S, Ackermann M. Short-range interactions govern the dynamics and function of microbial communities. *Nat Ecol Evol* **4**:366-375 (2020).

14. Liu G, David BT, Trawczynski M, Fessler RG. Advances in pluripotent stem cells: history, mechanisms, technologies, and applications. *Stem Cell Reviews and Reports* **16**:3-32 (2020).

15. Gurdon JB: A community effect in animal development. *Nature* **336**:772-774 (1988).

16. Gurdon JB: Community effects and related phenomena in development. *Cell* **75**:831-834 (1993).

17. Polonsky M, Rimer J, Kern-Perets A, Zaretsky I, Miller S, Bornstein C, David E, Kopelman NM, Stelzer G, Porat Z, *et al.*: Induction of CD4 T cell memory by local cellular collectivity. *Science* **360**:1-11 (2018).

18. Alladin A, Chaible L, Garcia del Valle L, Sabine R, Loeschinger M, Wachsmuth M, Hériché JK, Tischer C, Jechlinger M: Tracking cells in epithelial acini by light sheet microscopy reveals proximity effects in breast cancer initiation. *eLife* **9**:e54066 (2020).

19. Bauwens CL, Peerani R, Niebruegge S, Woodhouse KA, Kumacheva E, Husain M, Zandstra PW: Control of human embryonic stem cell colony and aggregate size heterogeneity influences differentiation trajectories. *Stem Cells* **26**:2300-2310 (2008).

20. Paria BC & Dey SK: Preimplantation embryo development in vitro: cooperative interactions among embryos and role of growth factors. *Proc Natl Acad Sci USA* **87**:4756-4760 (1990).

21. Elowitz MB, Lim WA. Build life to understand it. *Nature* **468**:889-890 (2010).

22. Bauer G, Fakhri N, Kicheva A, Kondev J, Kruse K, Noji H, Riveline D, Saunders TE, Thattai M, Wieschaus E. The science of living matter for tomorrow. *Cell Systems* **6**:400-402 (2018).

23. Hannezo E, Heisenberg CP. Mechanochemical feedback loops in development and disease. *Cell* **178**:12-25 (2019).

24. Huang S. Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. *Phil Trans R Soc B* **366**:2247-2259 (2011).

25. Garcia-Ojalvo J. Physical approaches to the dynamics of genetic circuits: a tutorial. *Contemporary Physics* **52**, 439-464 (2011).

26. Estrada J, Wong F, DePace A, Gunawardena J. Information integration and energy expenditure in gene regulation. *Cell* **166**:234-244 (2016).

27. Ying QL & Smith AG: Defined conditions for neural commitment and differentiation. *Methods in Enzymology* 365:327-341 (2003).

28. Thomson M, Liu SJ, Zhou LN, Smith Z, Meissner A, Ramanathan S: Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**:875-889 (2011).

29. Fehling HJ, Lacaud G, Kubo A, Kennedy M, Robertson S, Keller G, Kouskoff V: Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Dev Biol* **130**:4217-4227 (2003).

30. Pearson S, Cuvertino S, Fleury M, Lacaud G, Kouskoff V: In vivo repopulating activity emerges at the onset of hematopoietic specification during embryonic stem cell differentiation. *Stem Cell Rep* **4**:431-444 (2015).

31. Varelas X: The Hippo pathway effectors TAZ and YAP in development, homeostasis and disease. *Dev* **141**:1614-1626 (2014).

32. Chung H, Lee BK, Uprety N, Shen W, Lee J, Kim J: Yap1 is dispensable for self-renewal but required for proper differentiation of mouse embryonic stem (ES) cells. *EMBO Rep* **17**:519-529 (2016).

33. LeBlanc L, Lee B-K, Yu AC, Kim M, Kambhampati AV, Dupont SM, Seruggia D, Ryu BU, Orkin SH, Kim J: Yap1 safeguards mouse embryonic stem cells from excessive apoptosis during differentiation. *eLife* **7**:e40167 (2018).

34. Mugahid D, Kalocsay M, Liu X, Gruver JS, Peshkin L, Kirschner MW: YAP regulates cell size and growth dynamics via non-cell autonomous mediators. *eLife* **9**:e53404 (2020).

35. Walko G. et al: A genome-wide screen identifies YAP/WBP2 interplay conferring growth advantage on human epidermal stem cells. *Nature Communications* **8**:14744 (2017).

36. Vigneron AM & Vousden KH: An indirect role for ASPP1 in limiting p53-dependent p21 expression and cellular senescence. *EMBO J* **31**:471-480 (2012).

37. Kim M, Kim T, Johnson RK, Lim DS: Transcriptional co-repressor function of the Hippo pathway transducers YAP and TAZ. *Cell Reports* **11**:270-282 (2015).

38. Zhou Y, Huang T, Cheng ASL, Yu J, Kang W, To KF: The TEAD Family and Its Oncogenic Role in Promoting Tumorigenesis. *Int J Mol Sci* **17**:1-15 (2016).

39. Yu OM, *et al.*: YAP and MRTF-A, transcriptional co-activators of RhoA-mediated gene expression, are critical for glioblastoma tumorigenicity. *Oncogene* **37**:5492-5507 (2018).

40. Zhu Q, *et al.*: The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. *eLife* **3**:e02224 (2014).

41. Huh HD, Kim DH, Jeong HS, Park WE: Regulation of TEAD Transcription Factors in Cancer Biology. *Cells* **8**:600 (2019).

42. Brodowksa K, Moujahed A, Marmalidou A, Meyer zu Horste M, Cichy J, Miller JW, Gragoudas E, Vavvas DG: The clinically used photosensitizer Verteporfin (VP) inhibits YAP-TEAD and human retinoblastoma cell growth in vitro without light activation. *Exp Eye Res* **124**:67-73 (2015).

43. Moon S, *et al.*: Phosphorylation by NLK inhibits YAP-14-3-3-interactions and induces its nuclear localization. *EMBO Rep* **18**:61-71 (2017).

44. Duval D, Trouillas M, Thibault C, Dembele D, Diemunsch F, Reinhardt B, Mertz AL, Dierich A, Boeuf H: Apoptosis and differentiation commitment: novel insights revealed by gene profiling studies in mouse embryonic stem cells. *Cell Death & Differentiation* **13**:564-575 (2006).

45. Fujita J, Crane AM, Souza MK, Dejosez M, Kyba M, Flavell RA, Thomson JA, Zwaka TP: Caspase activity mediates the differentiation of embryonic stem cells. *Cell Stem Cell* **2**:595-601 (2008).

46. Wang ES, Reyes NA, Melton C, Huskey NE, Momcilovic O, Goga A, Blelloch R, Oakes SA: Fas-activated mitochondrial apoptosis culls stalled embryonic stem cells to promote differentiation. *Current Biology* **25**:3110-3118 (2015).

47. Mittal N, Voldman J: Nonmitogenic survival-enhancing autocrine factors including cyclophilin A contribute to density-dependent mouse embryonic stem cell growth. *Stem Cell Res* **6**:168-176 (2011).

48. Mohammadi M, Froum S, Hamby JM, Schroeder MC, Panek RL, Lu GH, Eliseenkova AV, Green D, Schlessinger J, Hubbard SR: Crystal structure of an angiogenesis inhibitor bound to the FGF receptor tyrosine kinase domain. *EMBO J* 1998, 17:5896-5904.

49. Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith AG: FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**:2895-2902 (2007).

50. Mossahebi-Mohammadi M, Quan M, Zhang JS, Li X: FGF Signaling Pathway: A Key Regulator of Stem Cell Pluripotency. *Front Cell Dev Biol* **8**:1-10 (2020).

51. Laman Trip DS, Youk H: Yeasts collectively extend the limits of habitable temperatures by secreting glutathione. *Nature Microbiology* **5**:943-954 (2020).

52. Miller MB & Bassler BL: Quorum sensing in bacteria. *Annu Rev Microbiol* **55**:165-199 (2001).

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

54. Deglincerti A, Croft GF, Pietila LN, Zernicka-Goetz M, Siggia ED, Brivanlou AH. Self-organization of the in vitro attached human embryo. *Nature* **533**:251-254 (2016).

55. Rivron NC, Frias-Aldeguer J, Vrij EJ, Boisset JC, Korving J, Vivié J, Truckenmüller RK, Van Oudenaarden A, Van Blitterswijk CA, Geijsen N: Blastocyst-like structures generated solely from stem cells. *Nature* **557**:106-111 (2018).

56. Sozen B, Cox AL, De Jonghe J, Bao M, Hollfelder F, Glover DM, Zernicka-Goetz M: Self-organization of mouse stem cells into an extended potential blastoid. *Dev Biol* **51**:698-712 (2019).

57. Moris N, Anlas K, Van den Brink SC, Alemany A, Schröder J, Ghimire S, Balayo T, Van Oudenaarden A, Martinez-Arias A: An in vitro model of early anteroposterior organization during human development. *Nature* **582**:410-415 (2020).

58. O'Neill C, Li Y, Jin XL: Survival Signalling in the Preimplantation Embryo. *Cell Sign During Mammalian Early Embryo Dev.* pp.278-288 (Springer, 2015).

59. Wong AP, Shojaie S, Liang Q, Xia S, Di Paola M, Ahmadi S, Bilodeau C, Garner J, Post M, Duchesneau P, Waddell TK, Bear CE, Nagy A, Rossant J: Conversion of human and mouse

fibroblasts into lung-like epithelial cells. *Sci Rep* **9**:1-15 (2019).

60. Gunne-Braden A et al. GATA3 mediates a fast, irreversible commitment to BMP4-driven differentiation in human embryonic stem cells. *Cell Stem Cell* **26**:693-706 (2020).

61. Oyler-Yaniv A, Oyler-Yaniv J, Whitlock BM, Liu Z, Germain RN, Huse M, Altan-Bonnet G, Krichevsky O: A tunable diffusion-consumption mechanism of cytokine propagation enables plasticity in cell-to-cell communication in the immune system. *Immunity* **46**:609-620 (2017).

62. Hart Y, Reich-Zeliger S, Antebi YE, Zaretsky I, Mayo AE, Alon U, Friedman N: Paradoxical signaling by a secreted molecule leads to homeostasis of cell levels. *Cell* **158**:1022-1032 (2014).

63. Shvartsman SY, Wiley HS, Deen WM, Lauffenburger DA: Spatial range of autocrine signaling: Modeling and computational analysis. *Biophys J* **81**:1854-1867 (2001).

64. Augustin I, Dewi DL, Hundshammer J, Erdmann G, Kerr G, Boutros M: Autocrine Wnt regulates the survival and genomic stability of embryonic stem cells. *Science Signaling* **10**:eaah6829 (2017).

65. Chen CC, Wang L, Plikus MV, Jiang TX, Murray PJ, Ramos R, Guerrero-Juarez CF, Hughes MW, Lee OK, Shi S, *et al.*: Organ-Level Quorum Sensing Directs Regeneration in Hair Stem Cell Populations. *Cell* **161**:277-290 (2015).

66. Chaigne A, Labouesse C, White IJ, Agnew M, Hannezo E, Chalut KJ, Paluch EK: Abscission couples cell division to embryonic stem cell fate. *Dev Cell* **55**:1-14 (2020).

67. Malaguti M, Migueles RP, Blin G, Lin CY, Lowell S: Id1 stabilizes epiblast identity by sensing delays in nodal activation and adjusting the timing of differentiation. *Dev Cell* **50**:462-477 (2019).

68. Segel M, Neumann B, Hill MFE, Weber IS, Viscomi C, Zhao C, Young A, Agley CC, Thompson AJ, Gonzalez GA, Sharma A, Holmqvist S, Rowitch DH, Franze K, Franklin RJM, Chalut KJ: Niche stiffness underlies the ageing of central nervous system progenitor cells. *Nature* **573**:130-134 (2019).

69. Nandagopal N, Santat LA, LeBon L, Sprinzak D, Bronner ME, Elowitz MB. Dynamic ligand discrimination in the notch signaling pathway. *Cell* **172**:1-12 (2018).

70. Daneshmandi L, *et al.*: Emergence of the stem cell secretome in regenerative engineering. *Trends in Biotech* 38:1373-1384 (2020).

71. Liu J, Martinez-Corral R, Prindle A, Lee DYD, Larkin J, Gabalda-Sagarra M, Garcia-Ojalvo J, Süel GM: Coupling between distant biofilms and emergence of nutrient time-sharing. *Science* **356**:638-642 (2017).

72. Wrenn ED, Yamamoto A, Moore BM, Huang Y, McBirney M, Thomas AJ, Greenwood E, Rabena YF, Rahbar H, Partridge SC, Cheung KJ. Regulation of collective metastasis by nanolumenal signaling. *Cell* **183**:395-410 (2020).

73. Mojtahedi M, Skupin A, Zhou J, Castano IG, Leong-Quong RYY, Chang H, Trachana K, Giuliani A, Huang S. Cell fate decision as high-dimensional critical state transition. *PLoS Biol* **14**:e2000640 (2016).

74. Grosse-Wilde A, Kuestner RE, Skelton SM, MacIntosh E, d'Hérouel AF, Ertaylan G, Del Sol A, Skupin A, Huang S. Loss of inter-cellular cooperation by complete epithelial-mesenchymal transition supports favorable outcomes in basal breast cancer patients. *Oncotarget* **9:**20018-20033 (2018).

75. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**:544-547 (2008).

76. Petridou N, Corominas-Murtra B, Heisenberg CP, Hannezo E. Rigidity percolation uncovers a structural basis for embryonic tissue phase transitions. *Cell* (2021).

77. Corominas-Murtra B, Scheele C, Kishi K, Ellenbroek SIJ, Simons BD, Van Rheenen J, Hannezo E. Stem cell lineage survival as noisy competition for niche access. *Proc Natl Acad Sci USA* **17**:16969-16975 (2020).

78. Gunawardena J. Models in biology: 'accurate descriptions of our pathetic thinking'. *BMC Biol* **12**:29 (2014).

Fig. 1. Survival and differentiation are collective phenomena for differentiating embryonic stem cells

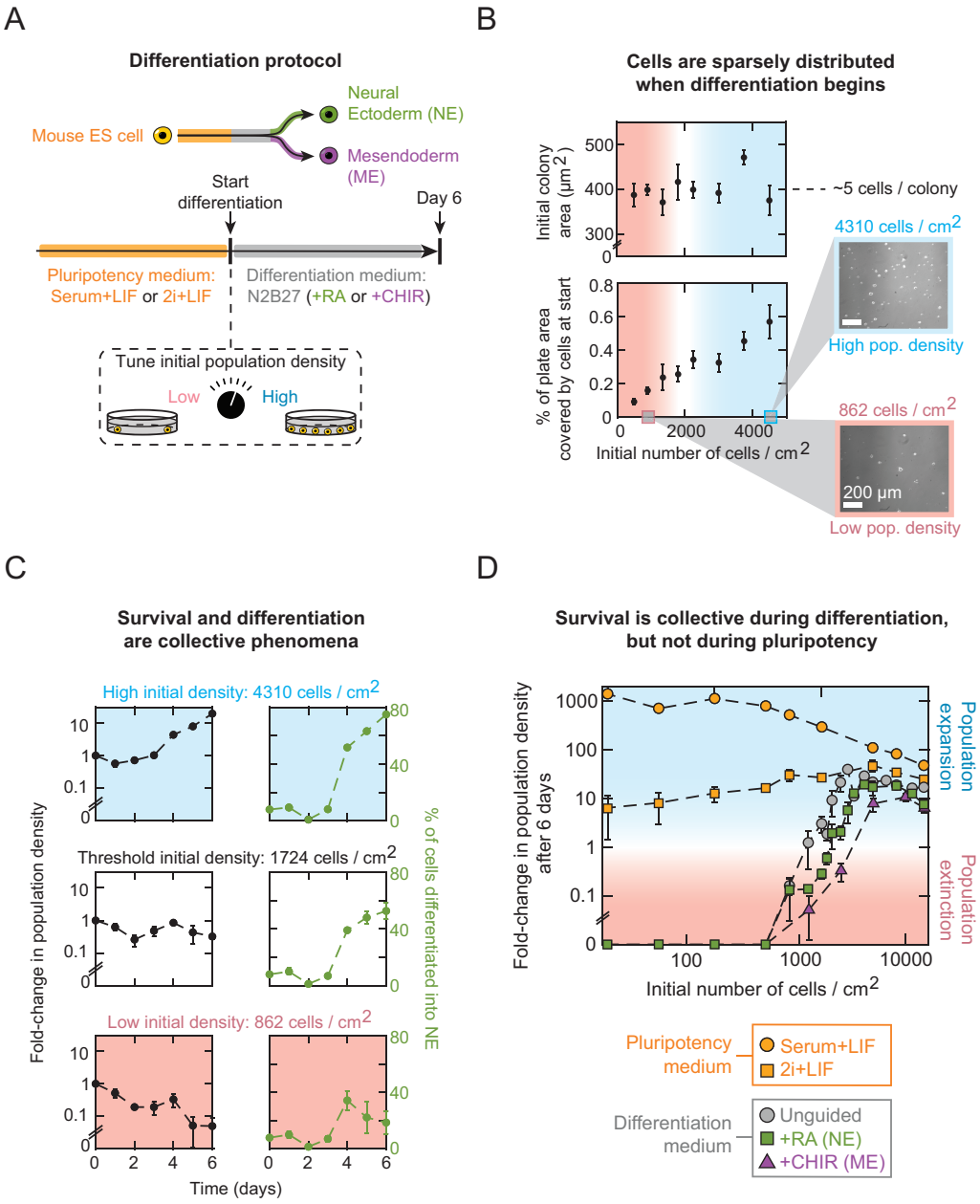


Figure 1 – Survival and differentiation are collective phenomena for differentiating embryonic stem cells.
(A) Differentiation protocol used in this study. See Methods.
(B) As a function of initial number of cells seeded on a 10-cm dish: initial area of a microcolony on average (top graph) and percentage of the plate area initially covered by microcolonies (bottom graph). Microscope used to measure both. Blue shade indicates population expansion. Red shade indicates population extinction. Two example microscope images shown. Scale bar = 200 μm. *n* = 3; Error bars are s.e.m.
(C) Fold-change in population density (black) and percentage of cells entering Neuroectoderm (NE) lineage (i.e., percentage of cells expressing Sox1-GFP) (green) as a function of time after RA-induced differentiation starts. Indicated above each row is the initial population density for the population shown in that row (See also Supplementary Figs. S2-S4). *n* is at least 3. Error bars are s.e.m.
(D) As a function of initial population density, fold change in population density after 6 days in pluripotency medium (orange) or one of three differentiation media (grey - N2B27 without any inducers; green - N2B27 with Retinoic Acid (RA) that induces NE-lineage differentiation; purple - N2B27 with CHIR that induces Mesendoderm (ME) lineage differentiation). Cell lines used: E14 (orange and grey), 46C (green), and Brachyury-GFP (purple). See Supplementary Fig. S5 for all cell lines and media types. Blue shade indicates population expansion. Red shade indicates population extinction. *n* is at least 3; Error bars are s.e.m.

Fig. 2. Nonlocal communication determines survival of differentiating population

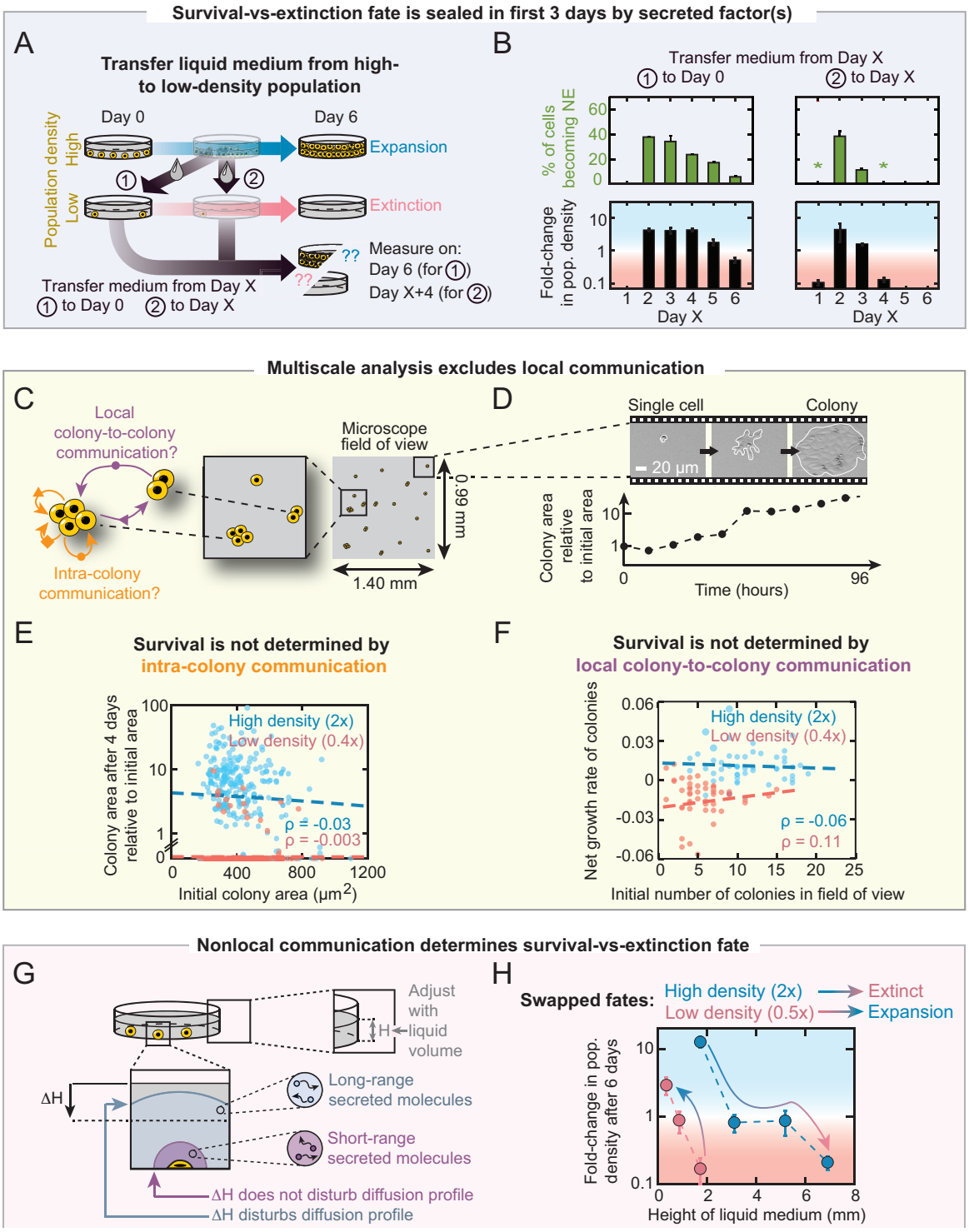


Figure 2 – Nonlocal communication determines survival of differentiating population.
(A) Over 6 days, High-density population (5172 cells / cm^2) expands (top: blue arrow) and low-density population (862 cells / cm^2) approaches extinction (middle: pink arrow). Two methods shown for replacing medium of the low-density population by that of the high-density population. Method (labeled "1"): replace initial medium of low-density population with the medium from X-days-old high-density population. Method 2 (labeled "2"): replace medium of X-days-old low-density population with that of X-days-old high-density population.

(B) Results of two experiments described in (A). Left column for method 1 and right column for method 2. Data for 46C cells induced to enter Neuroectoderm (NE) lineage. As a function of Day X, green bars show percentage of cells in population becoming NE (expressing Sox1-GFP) and black bars show fold change in population density. Blue shade indicates population expansion and red shade indicates population extinction. Asterisks indicate too few cells for reliable measurement on flow cytometer (less than 0.1-fold change in population density). Error bars are s.e.m.; $n = 3$.
(C) Two types of communication to exclude as possible mechanisms that control population's survival-versus-extinction fate by using time-lapse microscopy. Each field of view in microscope is 1.40 mm x 0.99 mm.
(D) Example of 96-hour time-lapse movie that tracks growth of a microcolony (subregion of a field-of-view shown). Scale bar = 20 μm. See also Supplementary Fig. S8.
(E) Each blue dot ($n = 224$) represents a microcolony from high-density population (initial density = 2857 cells / cm^2) and each red dot ($n = 94$) represents a microcolony from low-density population (initial density = 857 cells / cm^2). Data from multiple fields of view are shown for E14 cells in unguided differentiation. Area of each microcolony after 4 days of differentiation normalized by its initial area (vertical axis) versus its initial area (horizontal axis). Blue line: linear regression with $\rho = -0.03$. Red line: linear regression with $\rho = -0.003$. See Fig. S9 for all population densities.
(F) Blue and red data points represent the exact same microcolonies as in (E) but now showing microcolony's net growth rate (vertical axis) versus the initial number of microcolonies in the ~1-mm^2 field-of-view that contains the colony (horizontal axis). Blue line: linear regression with $\rho = -0.06$. Red line: linear regression with $\rho = 0.11$. See Figs. S10-S11 for all population densities. Also see Fig. S7 for experiments that show that the distance between one colony and its nearest-neighboring colony does not determine the colony's survival and its growth rate.
(G) Diffusion profile (volume filled by diffusing molecules) for molecule that diffuses far (blue region) and molecule that diffuses short range (purple region). Cell, in yellow, that secretes both molecules is adhered to the plate bottom. ΔH is change in liquid height. H is the total height of liquid medium that we tune in (H).
(H) Results of experiment described in (G). Fold-change in population density of 46C cells after 6 days of differentiation into NE lineage, as a function of the height of the liquid medium ("H" in (G)). Red points are for low-density population (initially 862 cells / cm^2) and blue points are for high-density population (initially 3448 cells / cm^2). Error bars are s.e.m.; $n = 3$.

Fig. 3. Nonlocal communication uses YAP1 to determine population survival within first two days of lineage commitment

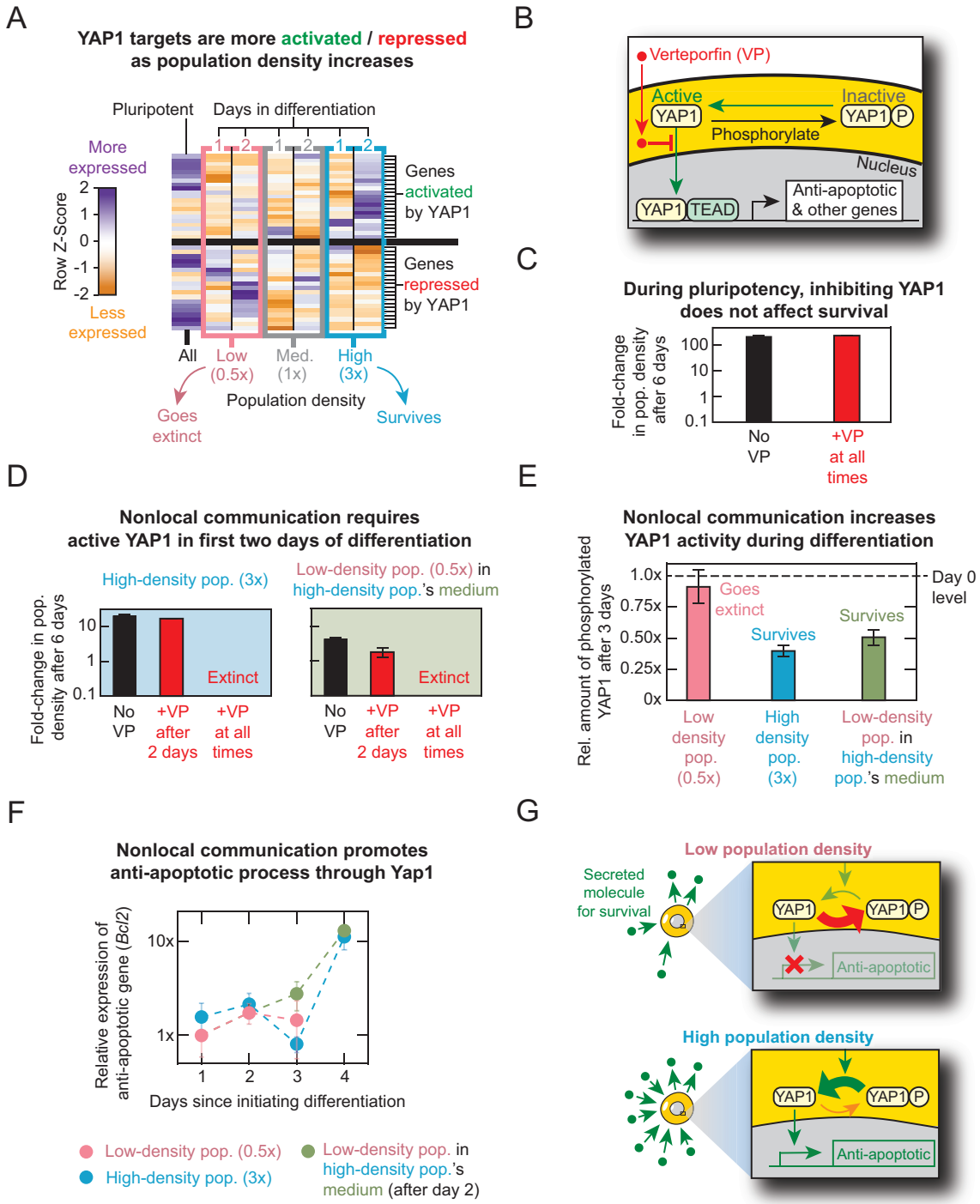


Figure 3 – Nonlocal communication uses Yap1 to determine population survival within first two days of lineage commitment.
(A) Heat map showing transcriptome-wide changes in unguided differentiation of 46C cells of low-density population (862 cells / cm^2 ; enclosed in pink box), near-threshold (medium-density) population (1931 cells / cm^2; enclosed in grey box), and high-density population (5172 cells / cm^2; enclosed in blue box). Leftmost column shows data for self-renewal (pluripotent) population before differentiation begins (labeled "All" since every population starts as this population before differentiation). Each column of differentiating population shows data for 1 day after (labeled "1") or 2 days after (labeled "2") starting differentiation. Each row shows a different gene, each of which are either activated (21 genes) or repressed (19 genes) by Yap1, either directly or indirectly. Fig. S14 lists all genes. Color represents row Z-score: a measure of by how much

a gene's expression level for a given condition deviates from that gene's expression level averaged across all different conditions (i.e., different populations and days). Purple represents a positive row Z-score (more expressed than average). Orange represents a negative row Z-score (less expressed than average). Data based on 3 biological replicates.

(B) Yap1 exists as either phosphorylated (labeled "P") or dephosphorylated. Verteporfin (VP) inhibits active (dephosphorylated) Yap1 from entering the nucleus and regulating target gene expression.

(C) Fold-change in population density (46C cells) after 6 days in serum-based, pluripotency medium. Red bar is for populations continuously incubated with 1-μM verteporfin (VP) in DMSO for six days. Black bar is for populations without VP but with the same amount of DMSO as control. All populations started at 862 cells / cm^2. Error bars are s.e.m.; $n = 3$.

(D) Fold-change in population density for high-density population (5172 cells / cm^2 initially, in blue box) and low-density population that was rescued with medium of 2-days-old high-density population (862 cells / cm^2 initially, in green box) after 6 days of differentiation towards NE-lineage. Data for 46C cells. Black bar: Verteporfin (VP) was always absent. Red bar in middle of each box: VP was added to medium after the first two days. Third column of each box shows absence of cells (extinction) when VP was present from the start of differentiation. Error bars are s.e.m.; $n = 3$.

(E) ELISA measurements showing amounts of Yap1 protein phosphorylated at Ser397 (inactive Yap1) (see Methods and also Fig. S15). Vertical axis shows the relative amount of inactive Yap1: the level of inactive Yap1 for a differentiating population divided by the amount of inactive Yap1 for a pluripotent population that has the same cell numbers as the differentiating population at the time of lysing the cells for ELISA. Values are for 46C cells, 3 days after starting differentiation towards NE lineage. Pink: low-density population (862 cells / cm^2 initially). Blue: high-density population (5172 cells / cm^2 initially). Green: low-density population rescued after two days by medium from a 2-day-old high-density population. Error bars are s.e.m.; $n = 3$.

(F) *Bcl2* (anti-apoptotic gene) expression level over time after initiating differentiation towards NE lineage. Data obtained from 46C cells with RT-qPCR. Colors represent same initial population densities as in (E). Also see Fig. S16. On each day, we first normalized a population's *Bcl2* expression level by that population's *Gapdh* level. Afterwards, plotted on the vertical axis, we divided each population's *Gapdh*-normalized *Bcl2* expression level on a given day by the *Gapdh*-normalized value for one-day-old low-density population (whose value is thus "1x" here). Error bars are s.e.m.; $n = 3$.

(G) Conclusion from the experiments in (C-F). Diagrams show how secreted molecules activate Yap1, which in turn controls anti-apoptotic processes.

Fig. 4. Secreted FGF4 determines population's survival-vs-extinction fate during lineage commitment

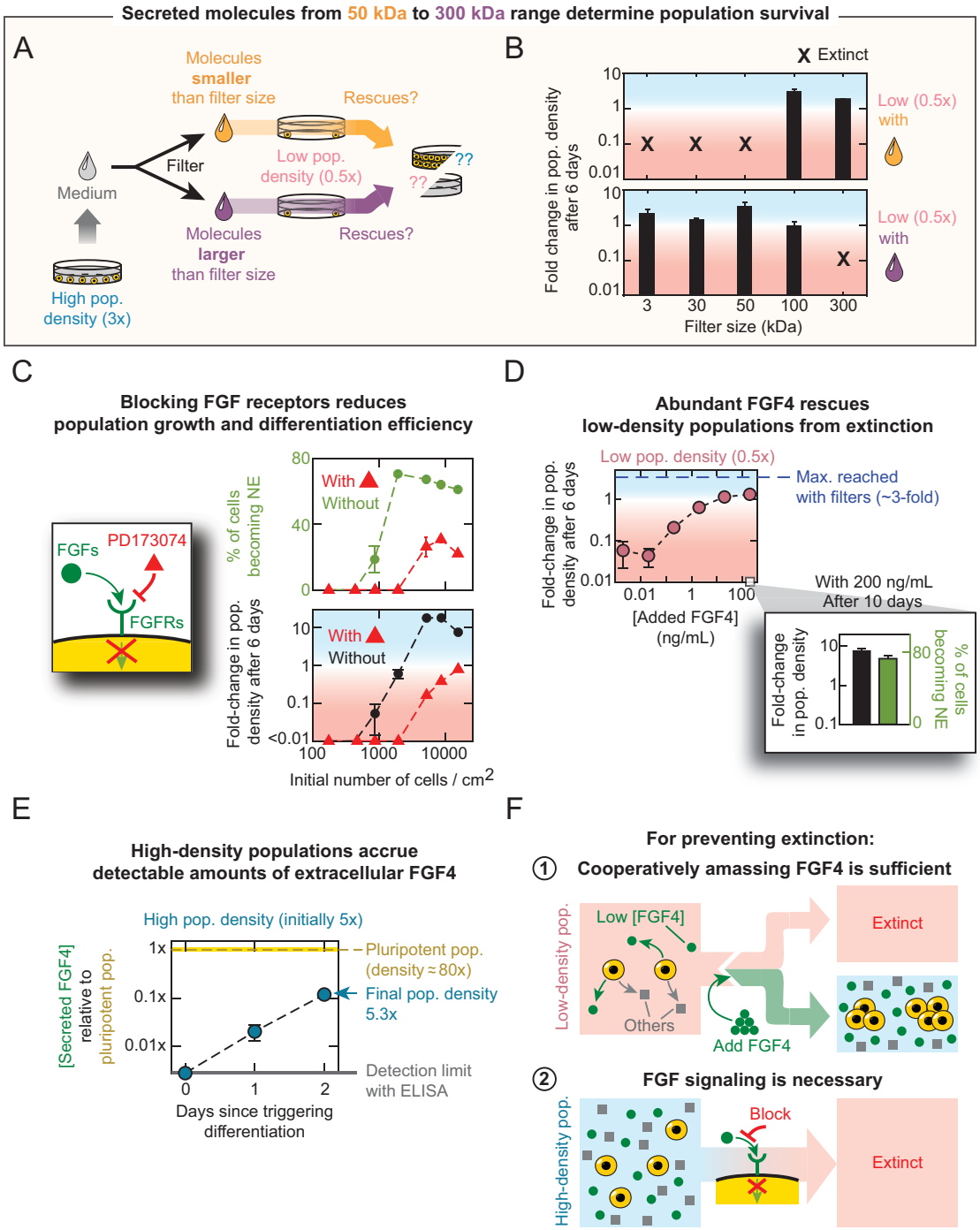


Figure 4 – Secreted FGF4 determines population's survival-vs-extinction fate during lineage commitment.
(A) Protocol for separating secreted molecules based on their molecular weights, through membrane-based filters (see Methods). We used 46C cells for all experiments. Grey arrow: medium is first taken from a High-density population (5172 cells / cm^2 initially). We then run the medium through a membrane filter of a certain filter (pore) size, resulting in splitting of medium in two parts (splitting black arrows): one part (orange) contains all molecules that are smaller than the filter size (in kDa) and a low-density population (862 cells / cm^2 initially) is then incubated in this medium to see if it expands or becomes extinct (orange arrow). The other part of the filtered medium (purple) contains all molecules that are larger than the filter size (in kDa). Same procedure is carried out with this medium (purple arrow) as with the medium containing all the lighter molecules. Full detains in Methods.

(B) Results of experiment described in (A). Top graph: fold change in density of the low-density population (vertical axis) that was incubated in the medium which contained either molecules smaller than the filter size (top graph) or larger than the filter size (bottom graph). Filter size is indicated on the horizontal axis. Cross ("X") indicates that the population went extinct (i.e., fold change below 0.01). Blue shade indicates population expansion and red shade indicates population extinction. Error bars are s.e.m.; $n = 3$.
(C) Cartoon shows PD173074, a standard small-molecule inhibitor of FGF receptors (*48,49*). Fold change in population density (bottom graph) and percentage of populations that enter NE lineage (top graph), both measured 6 days after differentiation began and as a function of initial population density. Data for 46C cells induced to differentiate into NE lineage. Red data points in both graphs are for populations that were incubated with 2 μM (1056 ng/mL) of PD173074 from the start of differentiation. PD173074 was dissolved in DMSO. Thus, as a control, black and green points are for populations without PD173074 but with the same amount of DMSO (volume per volume) as the populations represented by red data points. Blue shade indicates population expansion and red shade indicates population extinction. Error bars are s.e.m.; $n = 3$.
(D) Top graph: fold change in density of a differentiating population, after 6 days in medium containing FGF4 that we added. Horizontal axis shows the concentration of mouse recombinant FGF4 that we added at the start of differentiation. All data are for initially low-density population (862 cells / cm^2) and 46C cells induced towards NE-lineage. Horizontal dashed line shows the maximum fold-change in density obtained in our study, which occurs when the same low-density population is rescued by a filtered medium (in (B)). Bottom graph: fold change in population density and percentage of cells entering NE lineage (Sox1-GFP expressing cells), both measured 10 days after differentiation begins in presence of 200-ng/mL FGF4 that we added at the start of differentiation. Error bars are s.e.m.; $n = 3$.
(E) ELISA measurements of concentrations of extracellular FGF4 in the medium of a high-density population (8620 cells / cm^2 initially) during unguided differentiation (blue points). Vertical axis shows FGF4 concentration relative to that of a ~80% confluent pluripotent population (denoted "1x" and marked with yellow horizontal line). 80% confluency equals i~8 x 10^6 cells in 10-cm dish. Lower detection limit of the ELISA assay is indicated (in grey). See Fig. S21 for ELISA standard curves. Error bars are s.e.m.; $n = 3$.
(F) Conclusions from experiments in (C-E). Top row: gathering enough extracellular FGF4 is sufficient to prevent population extinction and this occurs for populations that start with sufficiently high densities. Equivalently, adding enough FGF4 to medium of low-density population alone is sufficient for preventing extinction. Bottom row: FGF signaling is necessary for recuing a low-density population from extinction.

Fig. 5. Mathematical model recapitulates experimental data

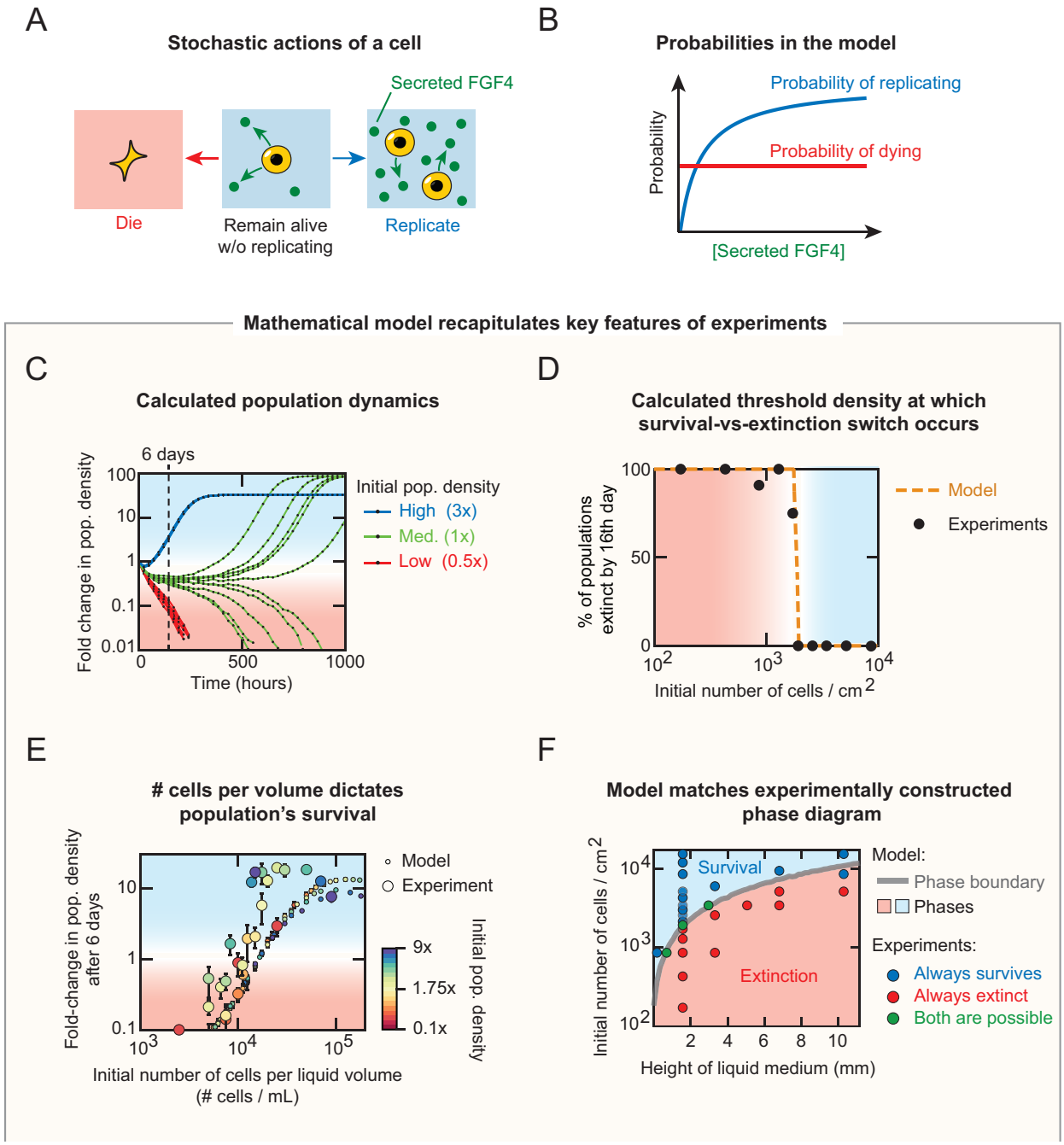


Figure 5 – Mathematical model recapitulates experimental data.
(A) Mathematical model with three possible, stochastic actions of a cell. Full details in Supplementary Text.
(B) Probabilities used in the model: probability of a cell dying (red) and probability of cell replicating (blue), both as functions of extracellular concentration of FGF4. Probability of dying is constant (see Fig. S30 for experimental validation). The probability of replicating nonlinearly increases with the concentration of extracellular FGF4 and is based on growth rates of various populations that we measured (see Fig. S31 and Supplementary Text).
(C) Fold change in population density, simulated by the model. Red: low-density population (862 cells/cm^2 initially). Green: near-threshold (medium) density population (1931 cells/cm^2 initially). Blue: high-density population (5172 cells/cm^2 initially). Vertical, dashed line marks 6 days after differentiation began (for comparison with Figs. 1C-D). Blue shade indicates population expansion and red shade indicates population extinction.

(D) Percentage of populations, in an ensemble of populations that start with the same, near-threshold density (1931 cells/cm^2), that becomes extinct either before or on the 16th day of differentiation. Black data points are from measurements and orange curve is from running the stochastic model multiple times (one for each population in an ensemble). Subset of n = 86 populations shown as black points (see Fig. S2 for full data). Population was considered to have reached extinction if its fold change decreased to below 0.1, relative to initial density. Blue shade indicates population expansion and red shade indicates population extinction.
(E) Data from new experiment performed after building the model (large circles) and model's prediction (small circles). Plotted here is the fold change in population density after 6 days as a function of the initial population density which is now measured as # cells per mL of liquid volume (instead of # cells / cm^2 that we used up to this point). Data from experiments and model predictions are from combining multiple volumes of liquid medium (in mL: 2, 5, 10, 18, 20, 30, 40 and 60) with initial # cells / cm^2 (172, 431, 862, 1293, 1724, 1931, 2155, 2586, 3017, 3448, 4310, 5172, 6034, 8621, 9483, 12069 and 15517 – indicated by the color bar). To obtain # cells / mL, we multiplied the # cells / area by the total area of the cell-culture plate (same for all conditions) and then divided the resulting value by the liquid volume. Data for 46C cells differentiating towards NE lineage. Blue shade indicates population expansion and red shade indicates population extinction. See Fig. S32 for full data set. Error bars are s.e.m.; n = 3.
(F) Model-produced phase diagram (grey curve and blue-red shadings) and experimental confirmation of the phase diagram (circles are from new experiments that were not used to build the model). Same experimental data as in Fig. 5E and Fig. S32 (n = 3; error bars are s.e.m.). Grey curve was constructed from the model by calculating, for each liquid medium height (volume), the threshold population density at which a population can either expand or become extinct (e.g., green curves in Fig. 5C). This, in turn, determines where the blue and red shades are in this plot. Circles are from experiments. Blue circles ("always survives") are for initial conditions - defined by liquid medium's height and initial number of cells / cm^2 - for which all replicate populations expanded. Red circles ("always extinct") are for initial conditions for which all replicate populations decreased in their density after 6 days (average of their fold change was below 0.6). Green circles ("both are possible") are for initial conditions for which some replicate populations expanded after 6 days while some did not (average of their fold change was between 0.6 and 1).

Fig. 6. Global community effect.

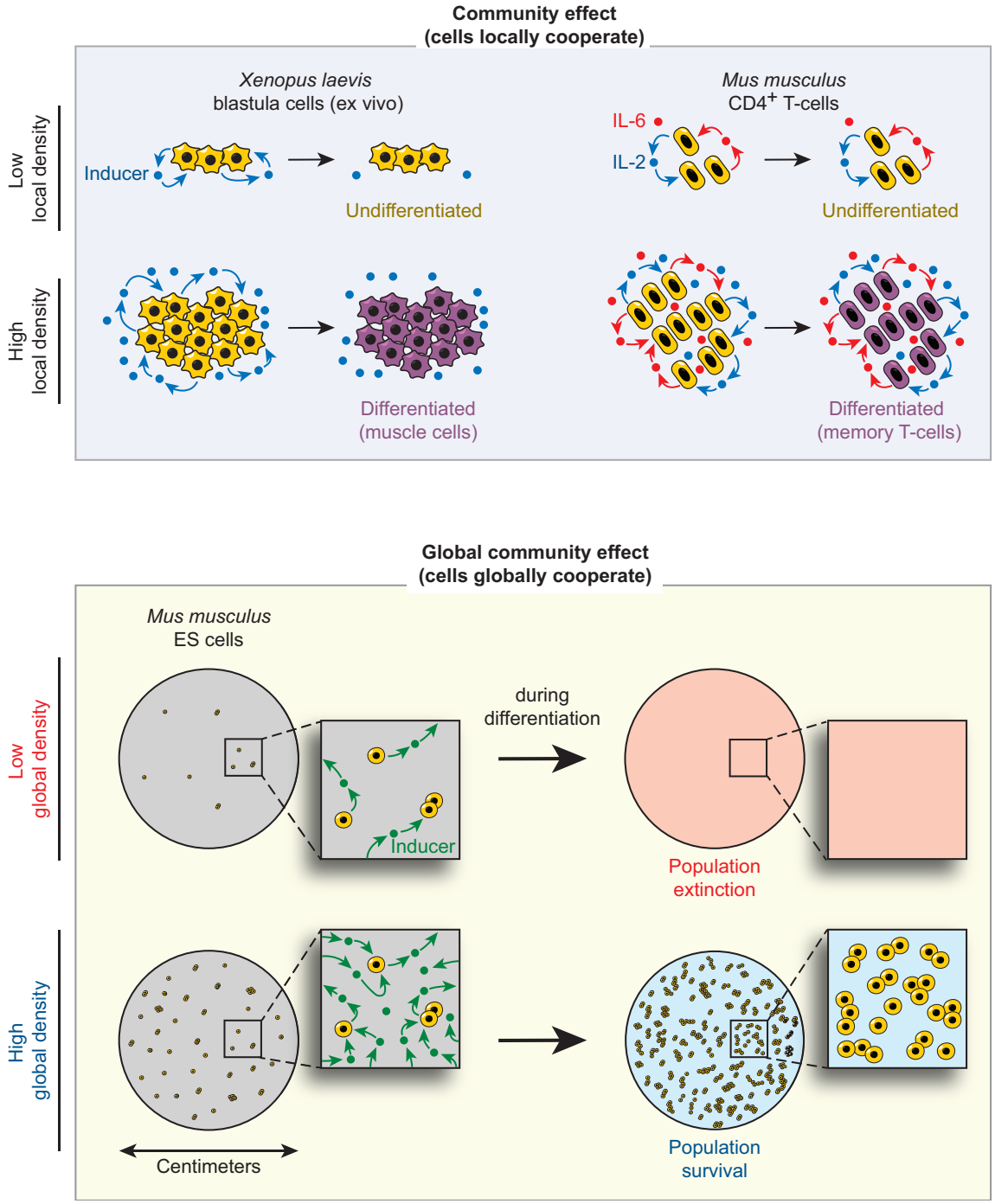


Figure 6 – Global community effect.
Upper panel: conventional community effect, as originally described by John Gurdon in frog embryonic cells (*15,16*) (left column) and recently also discovered for microwells of mouse T-cells that cooperatively differentiate into memory T-cells through local cooperation (within microwell) (*17*) (right column). In both examples, upper row shows differentiation failing to occur when the local population density (# of cells in a microwell or cell aggregate) is too low while bottom row shows differentiation occurring when the local density is sufficiently high (due to local cell-cell cooperation, mediated by secreting and sensing autocrine inducer). Lower panel: summarizes our work. Upper row shows differentiation failing (population becoming extinct) when the global density (# of cells measured over a macroscopic scale (over several centimeters)) is below a threshold value while the bottom row shows a successful differentiation occurring when the global density is higher than the threshold value (due to global cell-cell cooperation, mediated by secreting and sensing a long-range, autocrine inducer (green circle)).

Supplementary Information

Global community effect: large-scale cooperation yields collective survival of differentiating embryonic stem cells

Hirad Daneshpour[1,2], Pim van den Bersselaar[1,2], and Hyun Youk[1,3,4,5*]

[1]Kavli Institute of Nanoscience, Delft, The Netherlands

[2]Department of Bionanoscience, Delft University of Technology, Delft, The Netherlands

[3]CIFAR, CIFAR Azrieli Global Scholars Program, Toronto, ON, Canada

[4]Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA.

[5]Program in Systems Biology, University of Massachusetts Medical School, Worcester, MA, USA.

[*]Corresponding author. E-mail: hyun.youk@umassmed.edu

Table of contents:

- Figures S1-S32 ..... **Pgs. 2-50**

- Tables S1-S2 ..... **Pgs. 51-53**

- Supplementary Text (details of mathematical model) ..... **Pgs. 54-59**

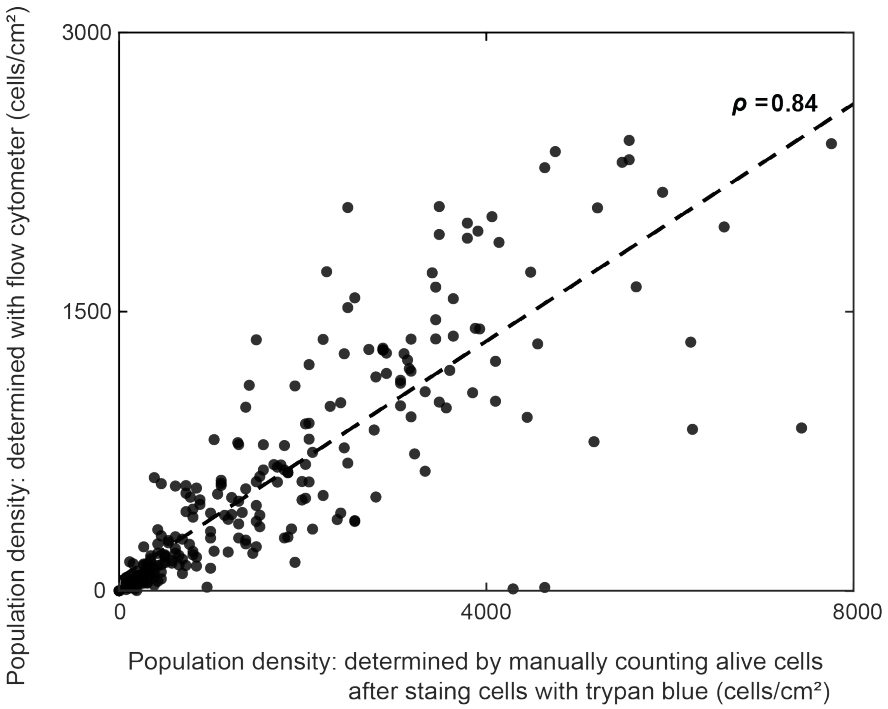


Figure S1 – Two different methods of counting cells yield population densities that are directly proportional to each other (i.e., offset by a constant factor of order one) and in the same order of magnitude, thus affirming both methods for determining population densities. Each data point represents a single population whose density (cells/cm^2) was determined by two independent methods. Each axis represents a different method. As one method, we used a standard hemocytometer to count individual cells after subjecting the cells to the dye, trypan blue. Trypan blue penetrated only dead cells within the population. We counted the unstained (non-blue) cells with a hemocytometer to determine the resulting population density (cells / cm^2) of alive cells on a cell-culture plate. As another method to determine the population density, we used a flow cytometer to count the number of cells (events) that belonged to a specified FSC-SSC gate. We set the FSC-SSC gate so that it captured alive cells while excluding dead cells. These two methods yielded cell counts - and thus the corresponding population densities - that were directly proportional to one another, as indicated by a high Pearson linear correlation (black line; ρ = 0.84). The proportionality factor is on the order of one, meaning that, for the same population, the two methods yield numbers that are in the same order of magnitude as the population density. Given this result, throughout our study, we primarily counted cells manually (i.e., with Trypan blue) and used the flow cytometer when this was not possible. Specifically, we used the flow cytometer to count populations that had very few cells such as those near extinction (i.e., populations whose fold-change in density was near or below 0.1 after some days).

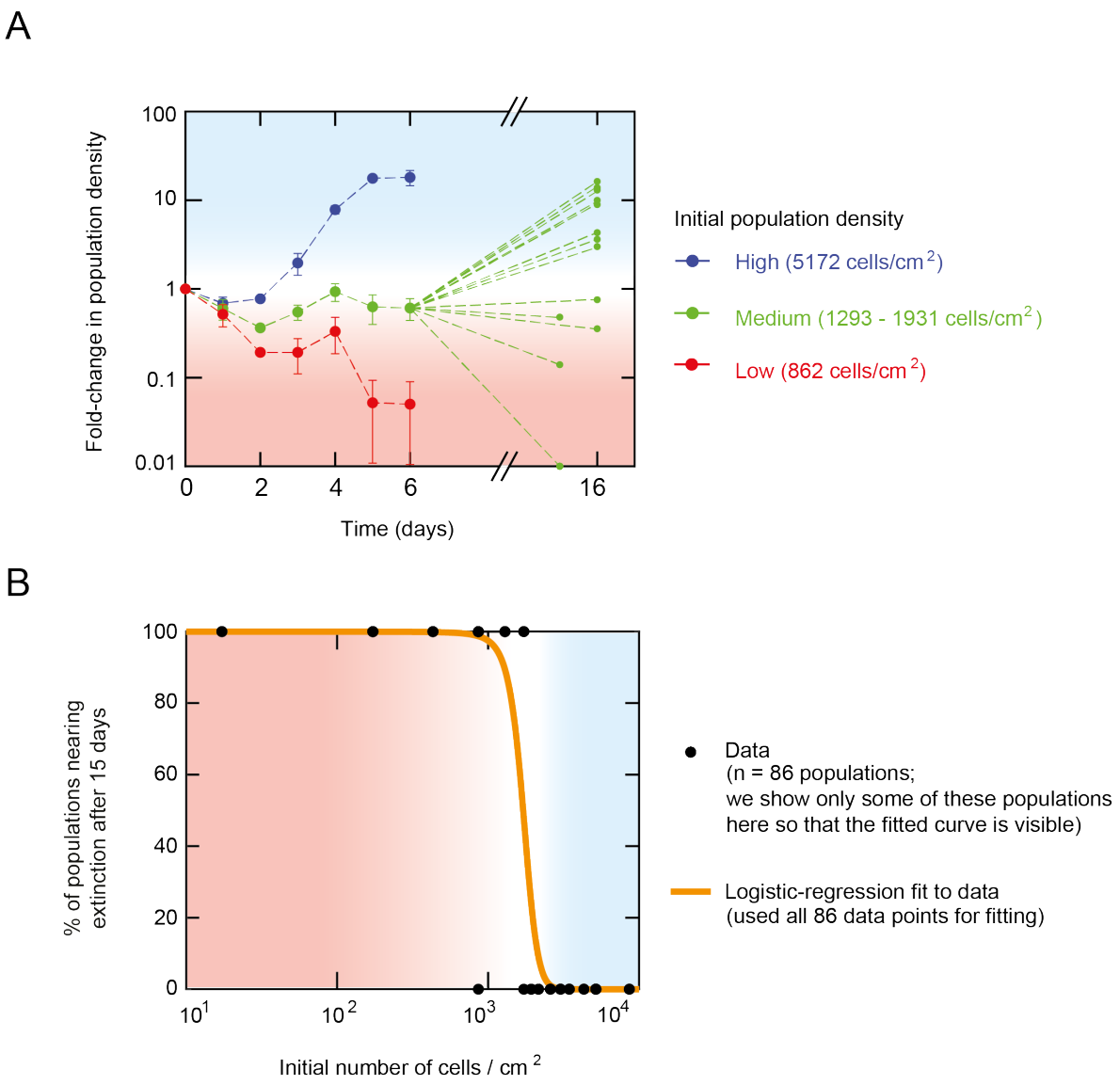


Figure S2 – Differentiating populations that start with a (near) "threshold density" (~1700 cells/cm^2) neither expand nor shrink during the first 6 days. But, after more days, these populations either expand towards the carrying capacity or shrink towards extinction in a stochastic manner (i.e., two populations of the same starting density can have two different fates (one survives and one becomes extinct)). Here, we have used Retinoic Acid (RA) to induce the NE-lineage commitment in the 46C cells (i.e., cells that have Sox1 promoter controlling GFP expression) (see Methods). Note that Sox1 is a marker of NE lineage. **(A)** Blue: populations that started with a sufficiently high density (5172 cells/cm^2) all grew towards the carrying capacity. Red: populations that started with a sufficiently low density (862 cells/cm^2) became extinct within the first 6 days. Green: populations that started near a "threshold density" (between 1293 and 1931 cells/cm^2) neither grew or shrank in the first 6

days. However, by 15-16 days after removing LIF (i.e., beginning differentiation), some of these populations reached the carrying capacity (i.e., population density increased by ~10-folds) while some others became extinct (i.e., population density became ~0.1-fold or less than its starting value). Still, some populations maintained nearly the same density for these 15-16 days (i.e., green curves with fold change of nearly one after 15-16 days). Thus, populations having the same initial density can have distinct fates: some would become extinct and some would survive (i.e., the fate is stochastically determined). For all the data shown for the first six days, n = 3 and error bars are s.e.m. Each green data point, taken 15-16 days after triggering differentiation, represents a single population (to show the stochasticity) rather than being averaged over multiple populations. They thus do not show error bars. **(B)** To quantify the stochastic nature of the survival-versus-extinction fate for populations that start with a near-threshold density (green data in (A)), we measured the fold-change in population density for 86 populations that collectively spanned a wide range of initial densities (black points). To each population, we either assigned a value of zero if it eventually grew towards the carrying capacity (i.e., fold-change of larger than 1) or a value of one if it eventually approached extinction within the first 16 days of differentiation. Then, we performed a logistic regression on these black data points by fitting a logistic function, $p(x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$ (orange curve). Consequently, $p(x)$ represents the probability that a population approaches extinction by 15-16 days after differentiation began. By fitting, we found β_1 = -8.7856 ± 0.8780 with a p-value of 3.76 x 10^{-25} according to the Wald test. This logistic regression is the simplest model (null model) that we can have for describing the probability of becoming extinct without any information about the mechanisms that determine the survival-versus-extinction fate. Later, we will introduce a mechanistic model that replaces this logistic regression fit (in Fig. 5).

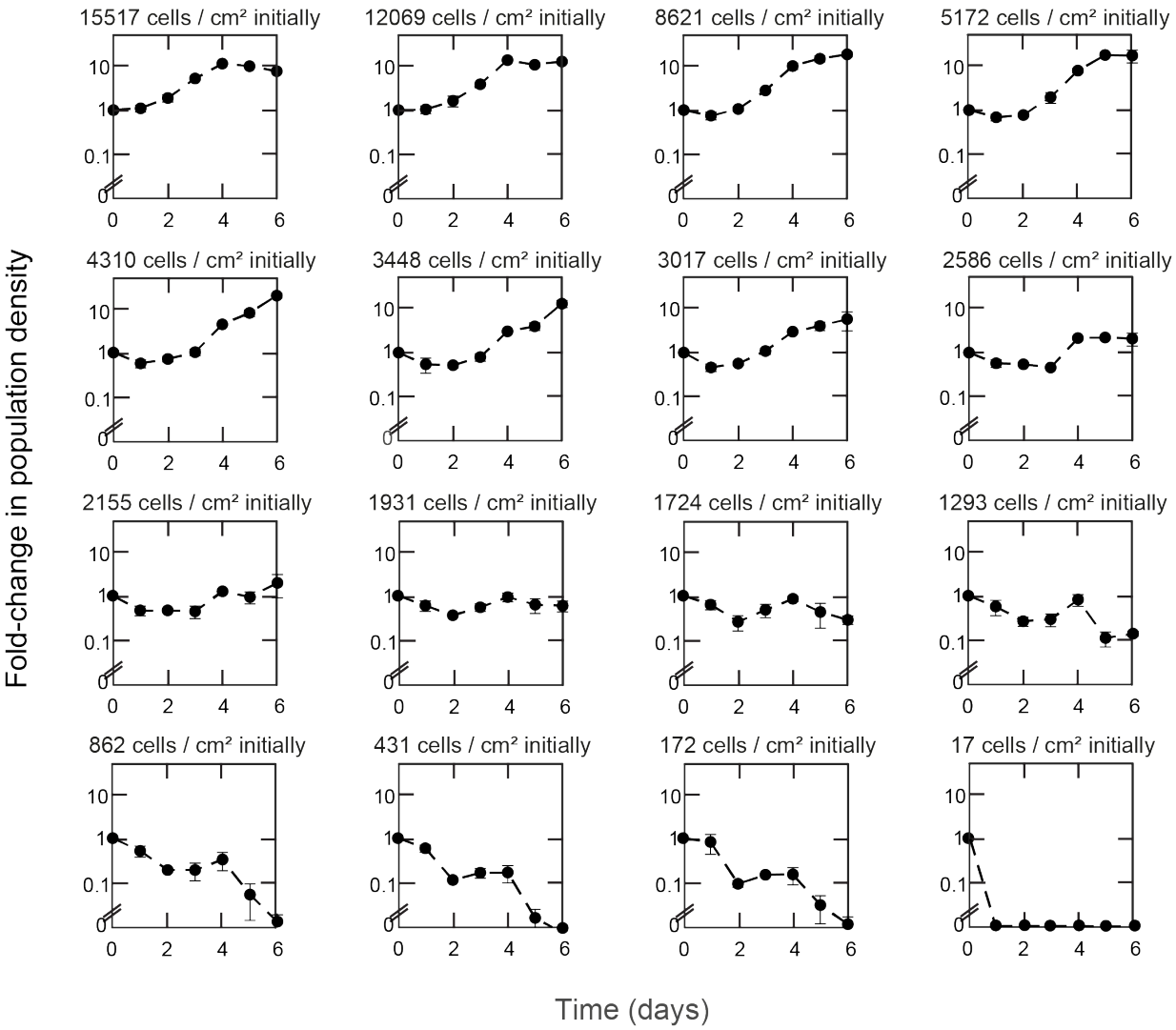


Figure S3 – Population growth during differentiation towards the NE-lineage for a wide range of starting population densities. Data shown for populations of 46C cells which have Sox1 promoter driving GFP expression. 46C cells are undergoing Retinoic Acid (RA) induced differentiation towards the NE-lineage (see Methods). We triggered pluripotency exit to begin each time course shown here (protocol in Fig. 1A). Each box shows the population dynamics for a different starting population-density (indicated above each box). Each box shows the fold-change in population density (vertical axis) as a function of the time passed since triggering differentiation (horizontal axis). $n \geq 3$ for each data point. Error bars are s.e.m.

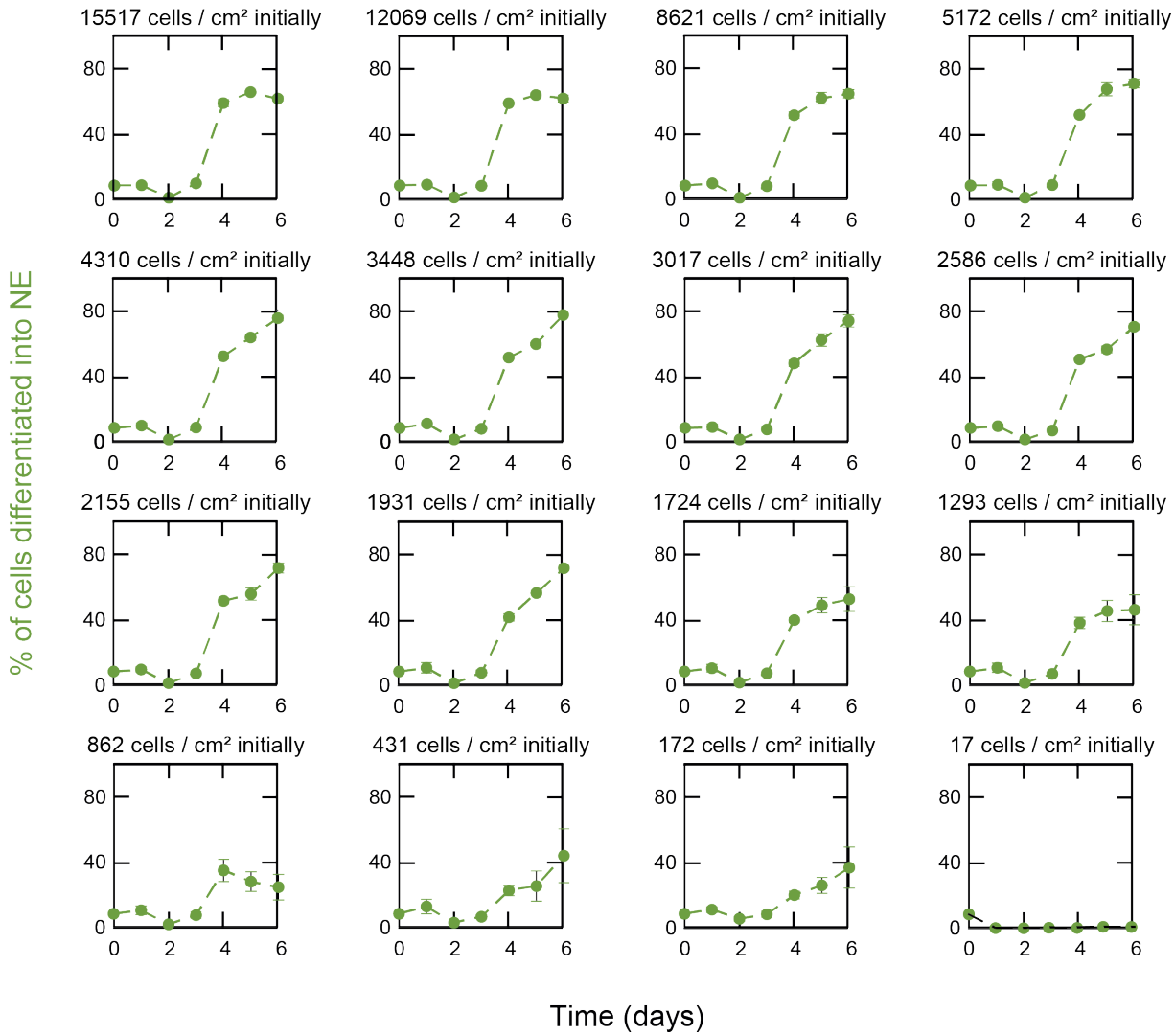


Figure S4 – Populations with higher initial densities achieve a higher differentiation efficiency (% of cells successfully entering NE-lineage). Retinoic Acid (RA) induced differentiation towards the NE-lineage by 46C cell-line populations. The 46C cell line has Sox1 promoter driving GFP expression. Sox1 - and thus GFP - is expressed only when the cell enters the NE-lineage (see Methods). We triggered pluripotency exit to begin each time course (protocol in Fig. 1A). Each box shows the differentiation efficiency for a different starting population-density (indicated above each box). On each day, we collected all cells from a cell-culture plate and then flowed them into a flow cytometer to measure the percentage of alive cells in the population that expressed GFP (i.e., percentage of cells that expressed Sox1 - a marker of NE-lineage commitment) (see Methods). Differentiation efficiency reached a maximum ~80% for populations that started above the threshold density of ~1700 cells/cm^2. Differentiation efficiency was below ~50% for populations that began with a below-threshold density. $n \geq 3$; error bars are s.e.m.

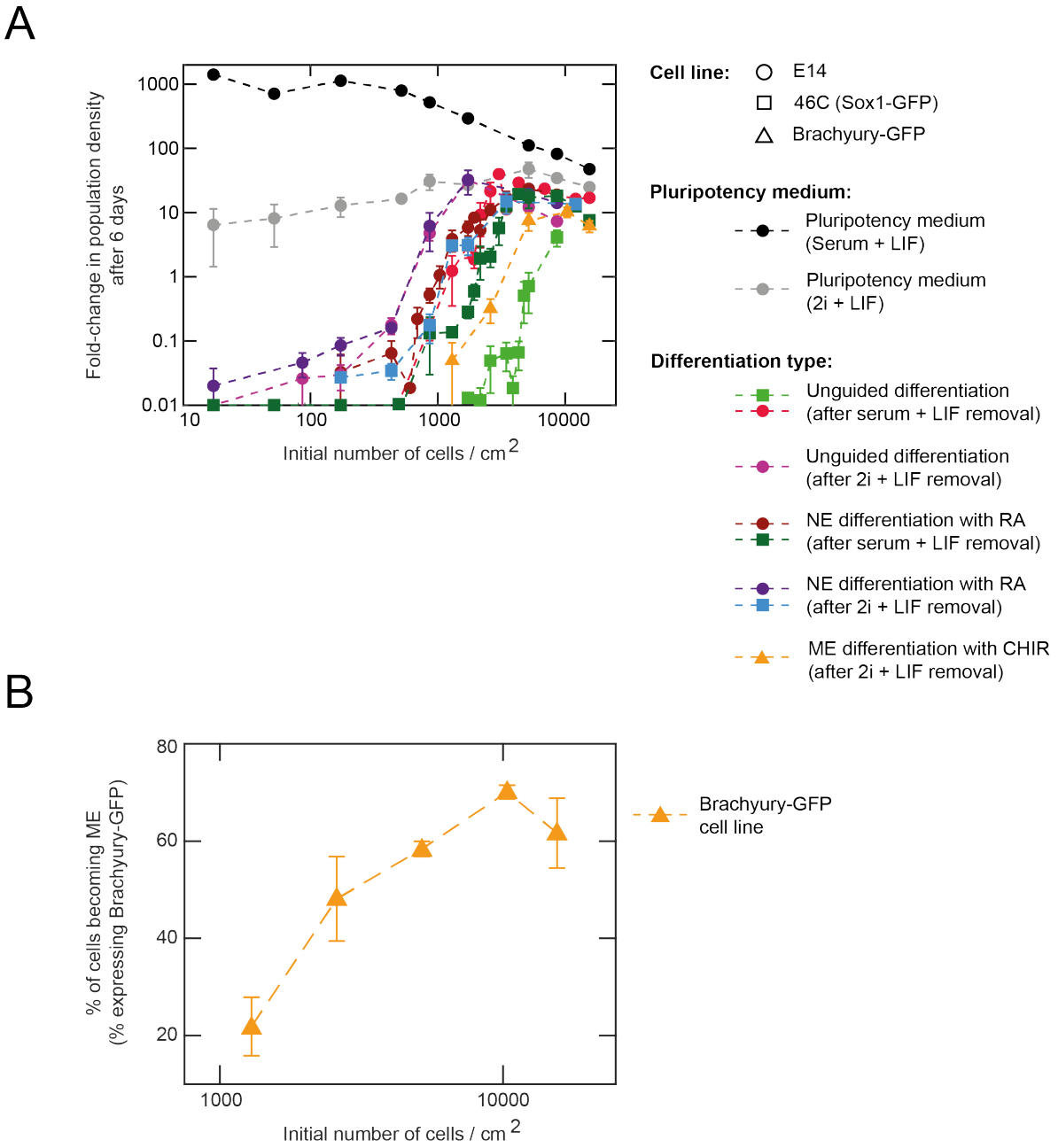


Figure S5 – Different cell lines and multiple types of differentiation all exhibit the same phenomenon: differentiating population's initial density determines its survival-versus-extinction fate. (A) Data shown for three different cell lines: E14 (circles), 46C (squares), and Brachyury-GFP (triangles). The 46C cells have a Sox1 promoter controlling their GFP expression (Sox1 is a marker of the neuro-ectoderm (NE) lineage). The Brachyury-GFP cells have a Brachyury promoter controlling their GFP expression (gift from V. Kouskoff and described in (*29,30*). Brachyury is a marker of mesendoderm (ME) lineage. Different colors

represent different types of cell-culture media as indicated in the legend. For each differentiation, we took one of the three cell lines that were kept pluripotent with LIF in either a serum-based medium or a serum-free (2i) medium. The three types of differentiations are: unguided differentiation in which no inducer was added after triggering pluripotency loss, NE differentiation in which we added retinoic acid (RA), and ME differentiation in which we added the small molecule, CHIR (see Methods). As seen here, regardless of the cell line and differentiation type, a differentiating population's initial density determined its survival-versus-extinction fate. **(B)** After four days of CHIR-induced differentiation, we used a flow cytometer to measure the percentage of the Brachyury-GFP cells that expressed GFP (i.e., percentage of cells in a population that entered the ME lineage (*28*). As with the 46C cells that differentiated towards the NE lineage (Fig. 1C), populations of Brachyury-GFP cells that start with higher densities have higher differentiation efficiencies (larger percentages of cells entering the ME lineage). For both (A) and (B): $n = 3$; error bars are s.e.m.

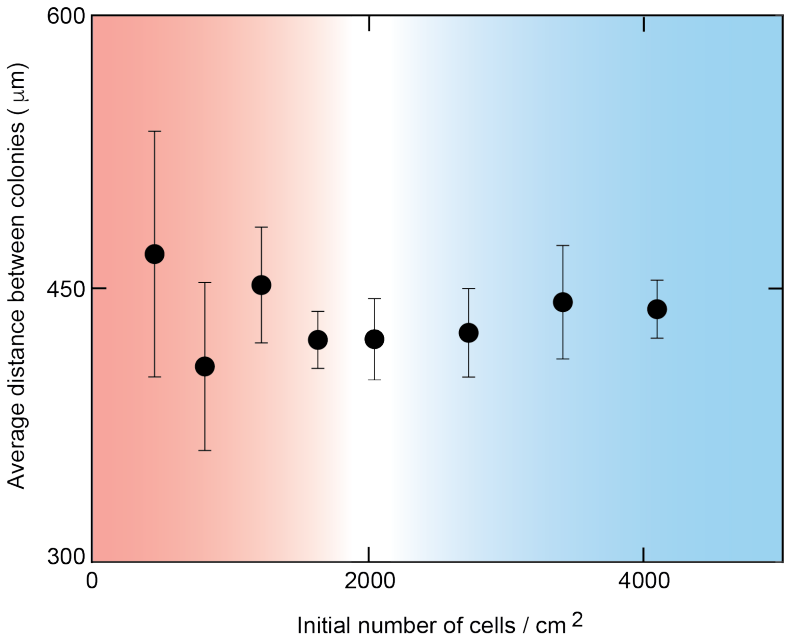


Figure S6 – For every population density, microcolonies are hundreds of microns apart from each other when differentiation begins. Data shown for E14 cells in unguided differentiation. We scattered a relatively a desired number of cells across a 6-cm diameter plate containing N2B27 medium to trigger pluripotency loss. Then, we used a wide-field microscope to locate and image the microcolonies in seventeen fields of view. Each field of view has a dimension of 1.40 mm x 0.99 mm. From these images, we determined the distance between every pair of colonies that resided in the same field of view. Then, we averaged these distances (averaging over all pairs of colonies from all seventeen fields of view per plate). The resulting, average distance between colonies is plotted here as a function of the initial population density. As shown here, for a wide range of initial population densities, the average distance between microcolonies were virtually identical (~450 μm). Each cell has a diameter of ~10 μm, meaning that microcolonies, if they resided in the same field of view, were hundreds of microns apart (~45 cell-diameters apart). n = 3 plates for each initial population density; error bars are s.e.m.

A way to understand why the average colony-colony distance is nearly independent of the initial population density is that the plate area is much larger than can be covered by the cells, even for the high-density populations (at most ~1% of plate area is covered by the cells, as one can also check by a back-of-envelope calculation). In short, the vast difference in length-scale between individual cells and the plate area leads to above result. Given this, more informative metric is actually the distance between a colony and its nearest-neighboring colony (see Fig. S7).

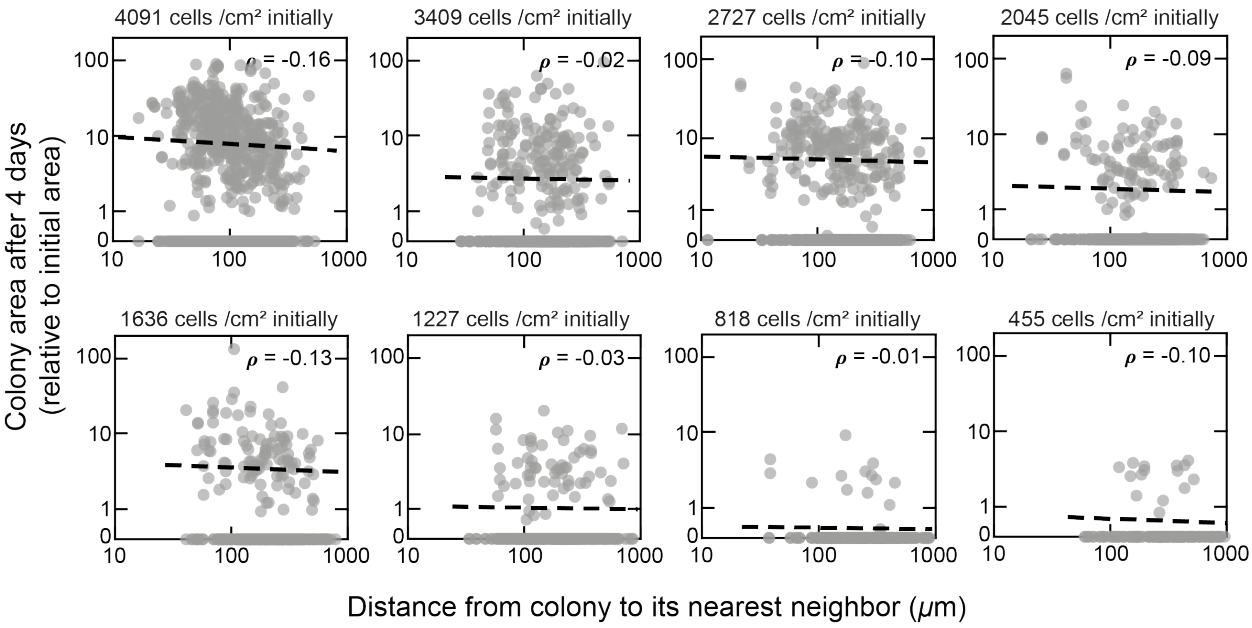


Figure S7 – Growth and survival of differentiating cells do not depend on the distance between nearest-neighboring colonies for any initial population density. Data for E14 cells undergoing unguided differentiation (i.e., no inducers such as RA). We scattered a desired number of cells across a 6-cm diameter plate containing N2B27 medium to initiate exit from pluripotency. Then, we used a wide-field microscope to locate and image the microcolonies in seventeen fields of view (see Methods). Each field of view has a dimension of 1.40 mm x 0.99 mm. From these images, we determined the distance between every pair of colonies that resided in the same field of view. From these measurements, we found the smallest distance (i.e., distance from one colony to its nearest-neighboring colony). We did this for each colony in a field of view. The resulting, nearest-neighboring distance for each colony is plotted here as a function of the initial population density. Each grey point represents the final area (after 96 hours) of a colony relative to its initial area as a function its distance to its nearest neighbor (at the start of the time-lapse movie). If a colony died during the time-lapse, then we assigned it a value of zero as the fold-change in its area. Each box shows the fold-change in the colony area (vertical axis) as a function of the distance from colony to its nearest neighbor (horizontal axis) for 17 fields of view and a specific population density. Dashed curves in each box denote the Pearson correlation with the correlation coefficient ρ in each box. There is virtually no correlation between a colony's final area (and whether it dies or not) and its initial, nearest-neighbor distance, for any starting population density. These results strongly indicate that the presence of a nearby colony does not predict whether a colony survives or not and also does not predict how fast each colony grows during differentiation.

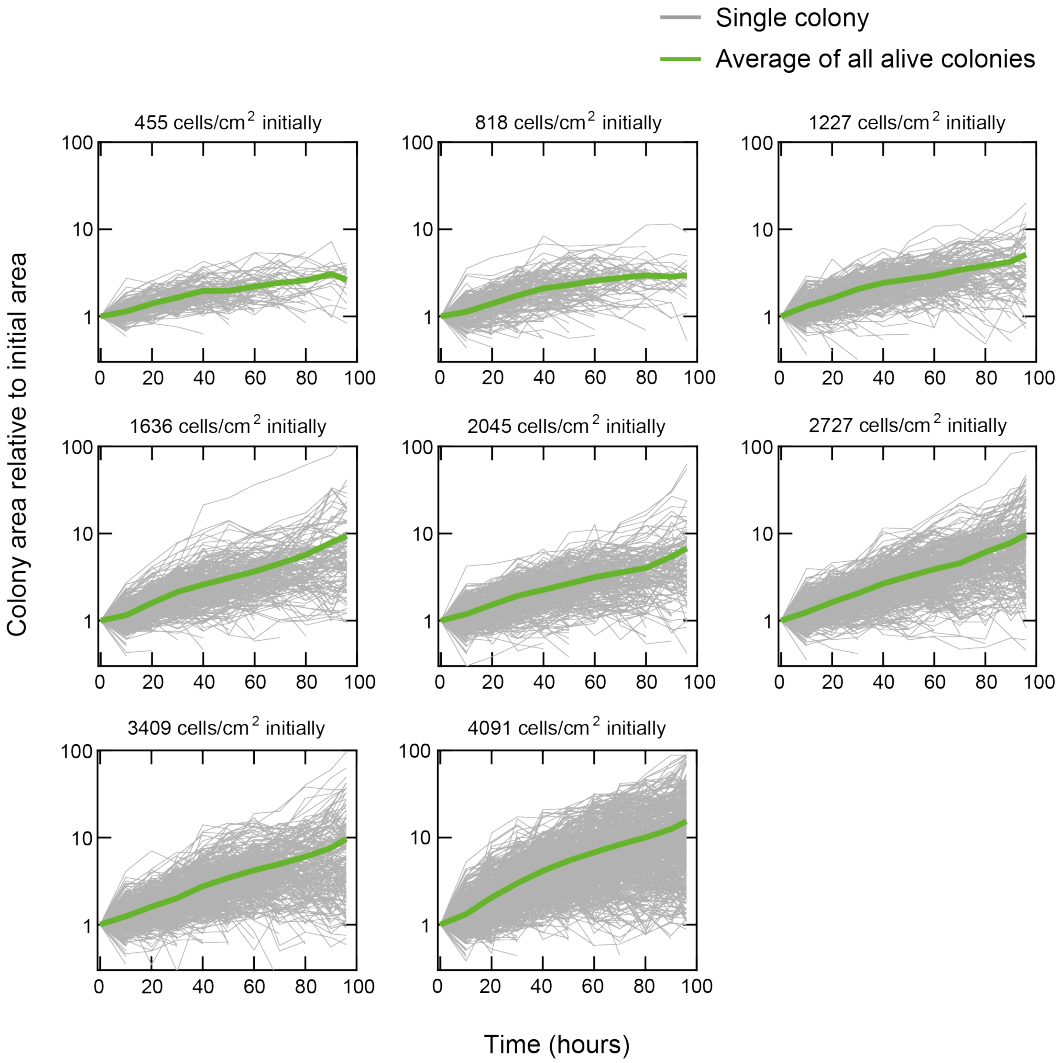


Figure S8 – Time-lapse microscopy over four days reveals the growth rate of each microcolony for a wide range of initial population densities. Data shown for E14 cells in unguided differentiation. that were triggered to exit pluripotency at the start of each time course shown here. We used a wide-field microscope to image and measure the area of each microcolony over four days (each grey curve). During the four days, we measured the area of a given microcolony every 10 hours (see Methods). Each box shows the fold-change in the colony area (vertical axis) as a function of the time passed since triggering differentiation (horizontal axis) for different starting population densities. In these movies, we observed both growing and dying colonies. For the colonies that died, the grey curves abruptly end, at the last time frame in which they were alive and before the end of the four-day period. The dying colonies visibly stood out as they typically displayed apoptotic bodies or lifted off the plate and thus disappeared from the focal plane. At each time frame, we computed the average area of all living colonies (green curve) from which we can extract the average growth rate of a colony for each population density (used for a model that we will introduce later in Fig. 5).

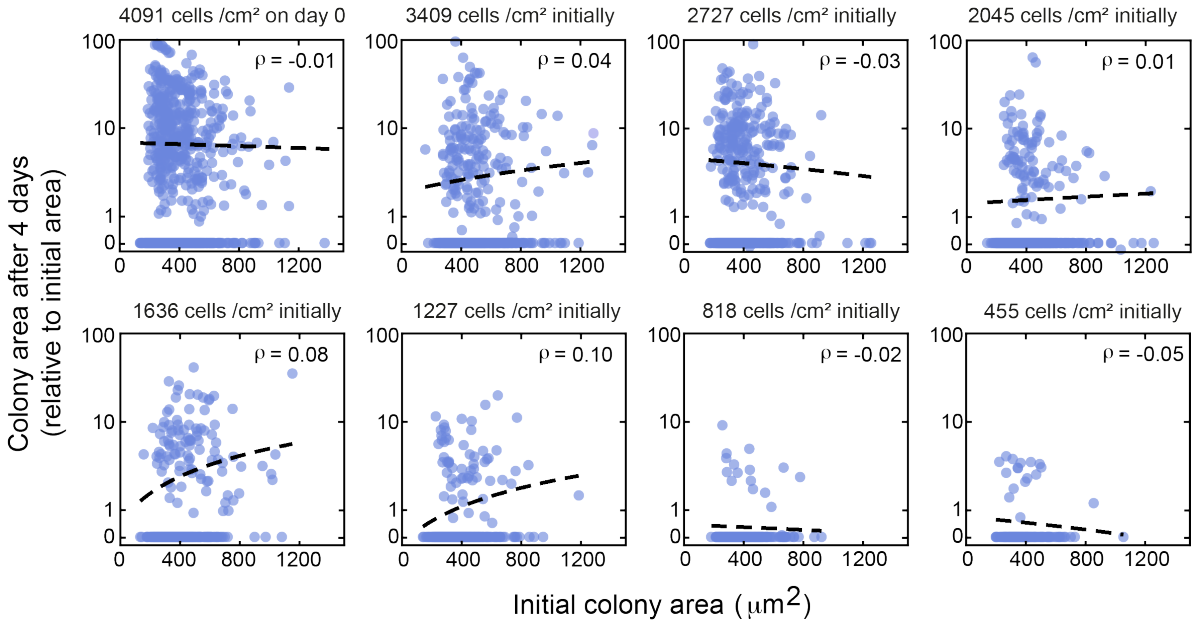


Figure S9 – Growth and survival of differentiating cells within a colony do not depend on how many cells are initially in the colony (i.e., colony area) for any population density. Data for E14 cells in unguided differentiation. Same protocol as in Fig. S8. We used a time-lapse microscope to measure the colony area during four days of differentiation. Each blue point represents the final area (after 96 hours) of a colony relative to its initial area (i.e., fold-change in colony area compared to the initial area). If a colony died during the time-lapse (spotted as explained in caption for Fig. S8), then we assigned it a value of zero as the fold-change in its area. Each box shows the fold-change in the colony area (vertical axis) as a function of the initial colony area (horizontal axis) for 17 fields of view and a specific population density. Dashed curves in each box denote the Pearson correlation with the correlation coefficient ρ in each box. There is virtually no correlation between a colony's final area (and whether it dies or not) and its initial area, for any starting population density. These results suggest that the initial area of a microcolony – which in turn is set by the number of cells in a microcolony – does not predict whether any cells in the colony survive or not and also does not predict how fast each cell grows during differentiation.

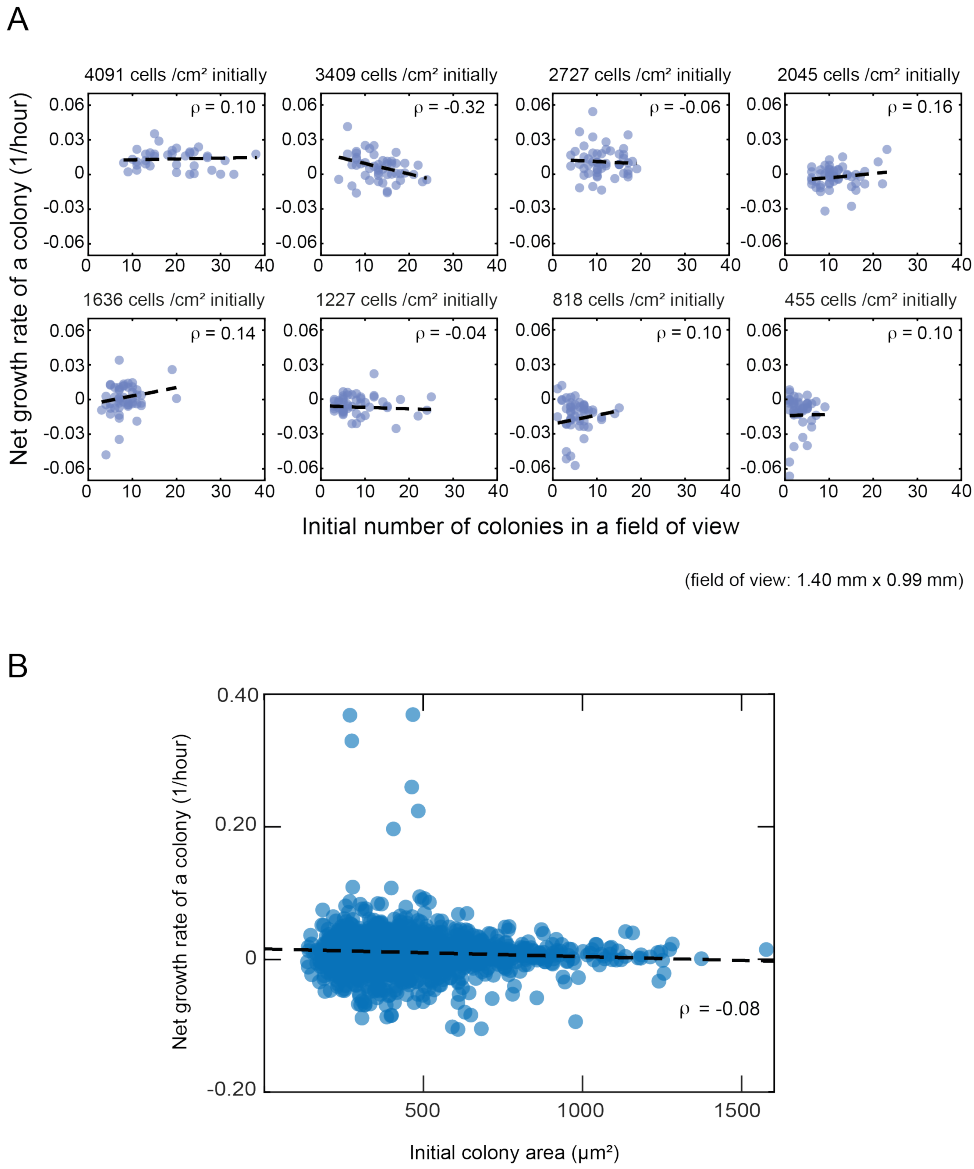


Figure S10 – Net growth rate of a colony does not depend on how many cells are in the colony (i.e., colony area) for any population density. Data shown for E14 cells in unguided differentiation. Same protocol as in Fig. S8. **(A)** We fitted an exponential function, Area(t) = $A_0\exp(\mu t)$, to each grey trace shown in Fig. S8 (i.e., we estimated colony area as exponentially growing as a function of time). From this fit, we determined net growth rate μ for each colony (i.e., for every grey trace shown in Fig. S8). For a given colony, we determined its net growth rate and the initial number of colonies that resided in the same field of view. Each blue point shows the data for a single colony. The net growth rate is positive ($\mu > 0$) if the colony grew over the four days of differentiation whereas it is negative ($\mu < 0$) or zero if the colony died (i.e., shrank or did not grow after which it often detached from the plate and thus disappeared from the field of view). Each box shows a colony's net growth rate (vertical axis) as a function

of the initial number of colonies that resided in the same field of view (horizontal axis) for a specific population density. We analyzed 17 fields of view for each starting population density. Dashed lines in each box shows the Pearson correlation with the correlation coefficient ρ denoted in each box. Here we see that how fast a colony grows - and whether it survives or dies - is virtually uncorrelated with how many other colonies there are in its ~1 mm x 1 mm neighborhood. **(B)** There is virtually no correlation between a colony's net growth rate and its initial area, as shown here. Data points come from many different, initial population densities and pooled together here. Dashed line shows the Pearson correlation with the correlation coefficient $\rho = -0.08$.

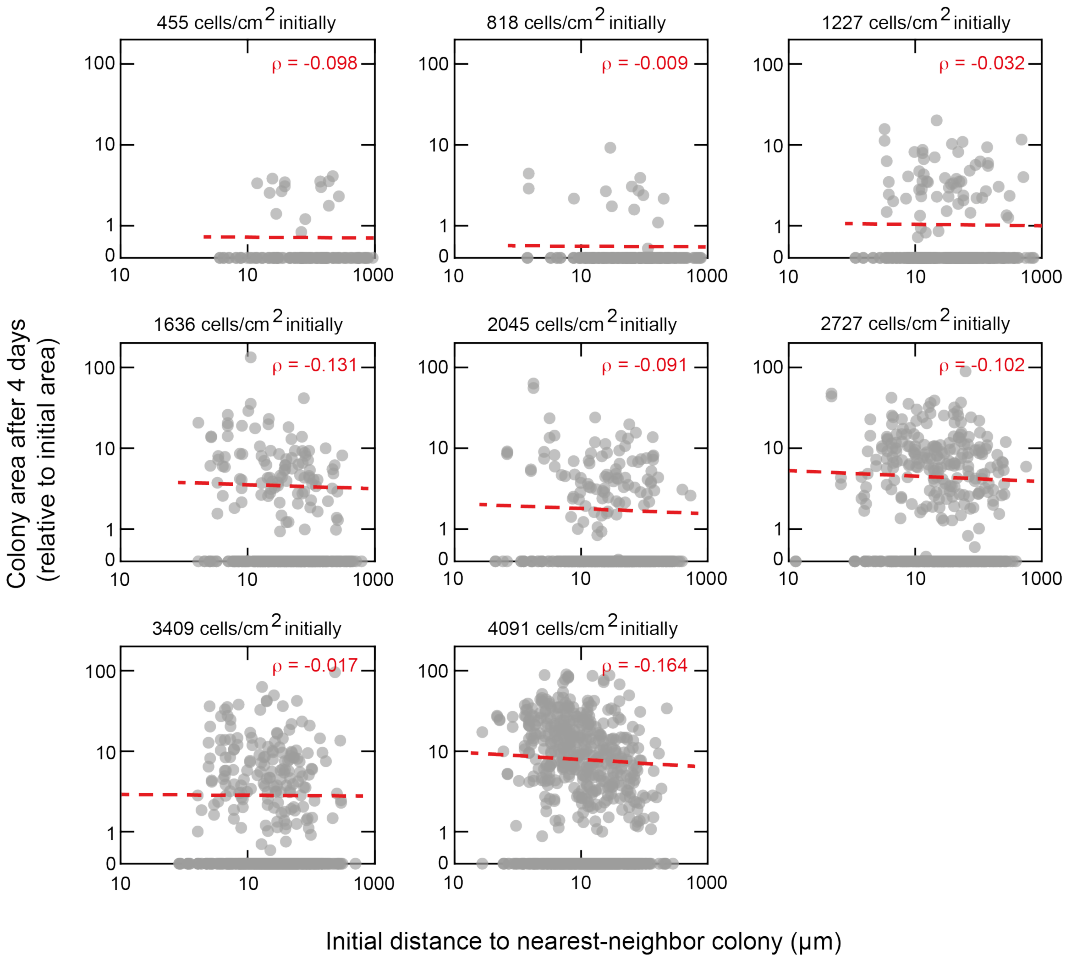


Figure S11 – Growth and survival of differentiating cells within a colony does not depend on having another colony nearby. Data shown for E14 cells in unguided differentiation. Same protocol as in Fig. S8. Each grey data point represents a single colony. For each colony, we measured the distance from its center to the center of every other colony that resided in the same field of view. We then determined the smallest of these distances: the colony's distance to its nearest neighbor. Each colony typically started with 1~10 cells (Fig. 1B). We then measured the colony's final area (after four days of differentiation). The grey data points show the final area of each colony as a function of the distance to its nearest-neighbor colony. We assigned a value of zero to the colony's final area if the colony died during the four-day period. Each box shows a different, initial population density. We analyzed 17 fields of view for each starting population density. In each box, a red line shows the Pearson correlation whose correlation coefficient ρ is indicated. For all population densities, a colony's growth rate and its chance of survival are virtually uncorrelated with how close it was to another colony when the differentiation began.

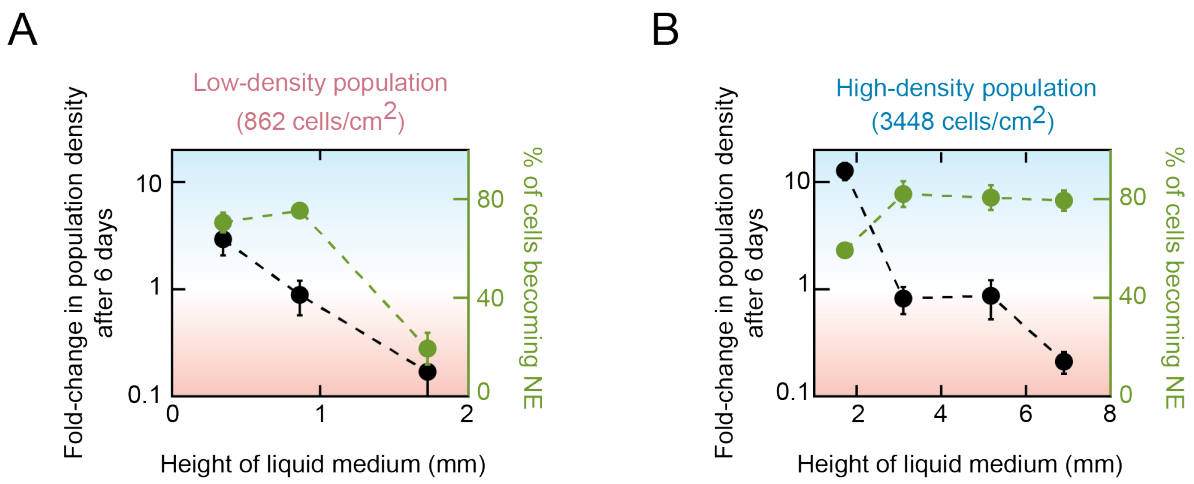
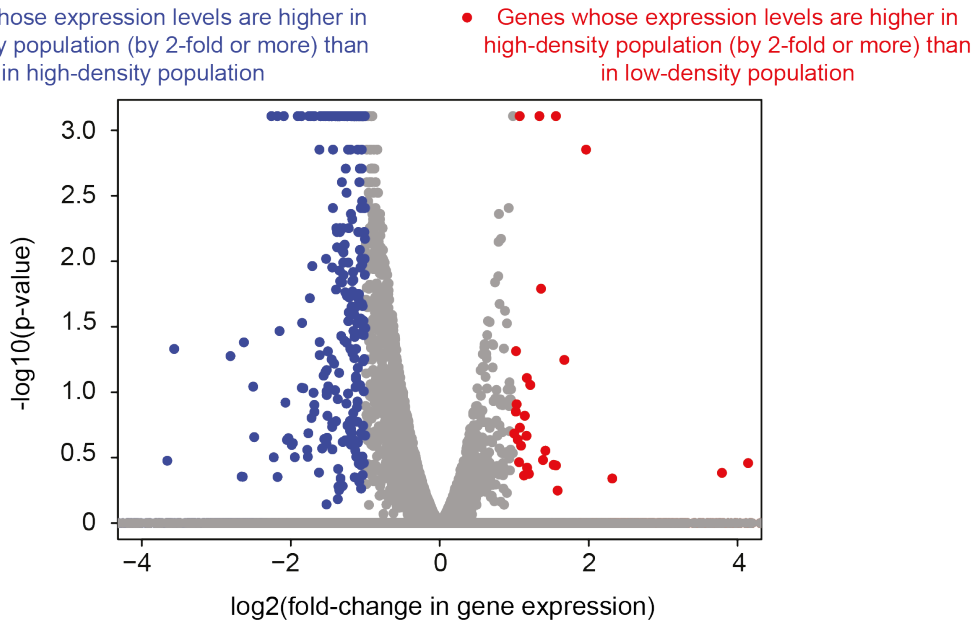
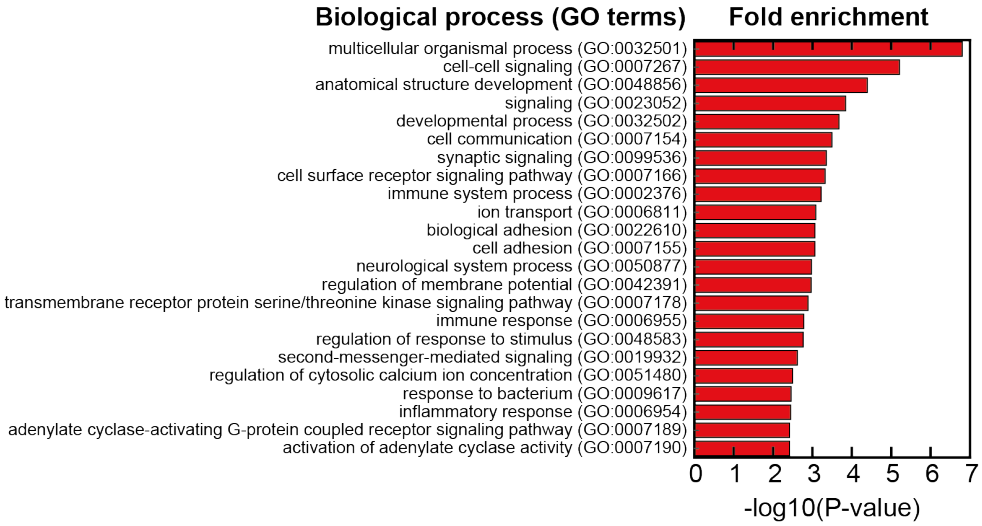


Figure S12 – Changing the height of the cell-culture medium over several millimeters alters the survival-versus-extinction fate of the population on the bottom of the cell-culture plate. Data shown for 46C (Sox1-GFP) cells undergoing Retinoic Acid (RA) induced differentiation towards the Neuroectoderm (NE) lineage. Black data points are duplicates of the data shown in Fig. 2H and they indicate the fold change in the population density (relative to the initial population density) as a function of the liquid-medium height above the cells. Green data points indicate the percentage of the cells that expressed GFP (i.e., Sox1 - a marker for NE lineage), which we measured with a flow cytometer, as a function of the liquid height above the cells. In our study, we used a 10-mL liquid medium (e.g., in Fig. 1) unless we explicitly state that we used a different volume (e.g., in Fig. 2H and here). A 10-mL liquid has a height that is just below 2 mm in a 10-cm diameter plate. **(A)** Data for a differentiating population that starts with a low density (862 cells / cm^2). In a 10-mL liquid, this population becomes extinct (last data point, at ~2-mm liquid height). From the smallest to the largest liquid height, the data correspond to 2 mL, 5 mL, and 10 mL liquid media. **(B)** Data for a differentiating population that starts with a high density (3448 cells / cm^2). In a 10-mL liquid, this population survives and grows towards the carrying capacity (first data point, at ~2-mm height). From the smallest to the largest liquid height, the data correspond to 10 mL, 20 mL, 30 mL, and 40 mL liquid media. (A) shows that we can rescue a low-density population - it survives - if we decrease the liquid height by 50% or more from the usual, ~2-mm. (B) shows that we can drive a high-density population to barely survive or become extinct if we increase the liquid height by a 2-fold or more. Altogether, these results exclude local communication. They suggest that the key secreted molecules diffuse over at least several millimeters to control a population's survival-vs-extinction fate.

A


Genes whose expression levels are higher in low-density population (by 2-fold or more) than in high-density population
Genes whose expression levels are higher in high-density population (by 2-fold or more) than in low-density population
-log10(p-value)
3.0
2.5
2.0
1.5
1.0
0.5
0
−4
−2
0
2
4
log2(fold-change in gene expression)

B


Biological process (GO terms)
Fold enrichment
multicellular organismal process (GO:0032501)
cell-cell signaling (GO:0007267)
anatomical structure development (GO:0048856)
signaling (GO:0023052)
developmental process (GO:0032502)
cell communication (GO:0007154)
synaptic signaling (GO:0099536)
cell surface receptor signaling pathway (GO:0007166)
immune system process (GO:0002376)
ion transport (GO:0006811)
biological adhesion (GO:0022610)
cell adhesion (GO:0007155)
neurological system process (GO:0050877)
regulation of membrane potential (GO:0042391)
transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0007178)
immune response (GO:0006955)
regulation of response to stimulus (GO:0048583)
second-messenger-mediated signaling (GO:0019932)
regulation of cytosolic calcium ion concentration (GO:0051480)
response to bacterium (GO:0009617)
inflammatory response (GO:0006954)
adenylate cyclase-activating G-protein coupled receptor signaling pathway (GO:0007189)
activation of adenylate cyclase activity (GO:0007190)
0 1 2 3 4 5 6 7
-log10(P-value)

C

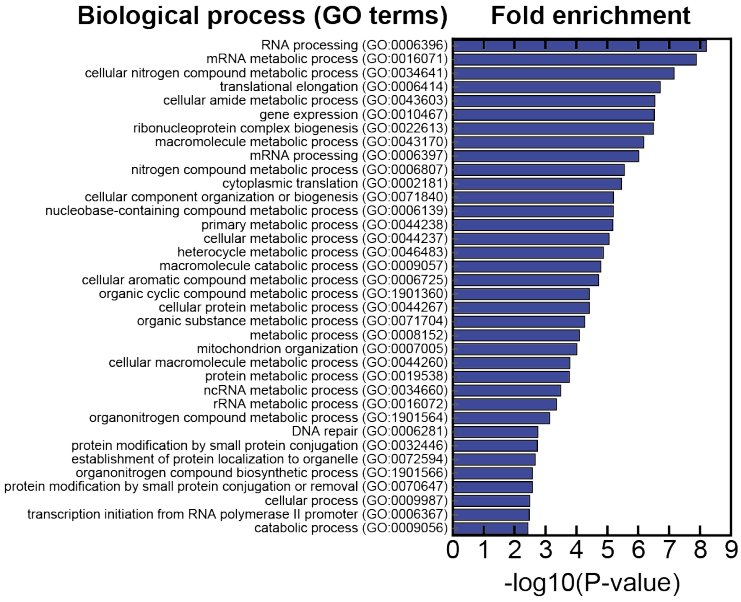
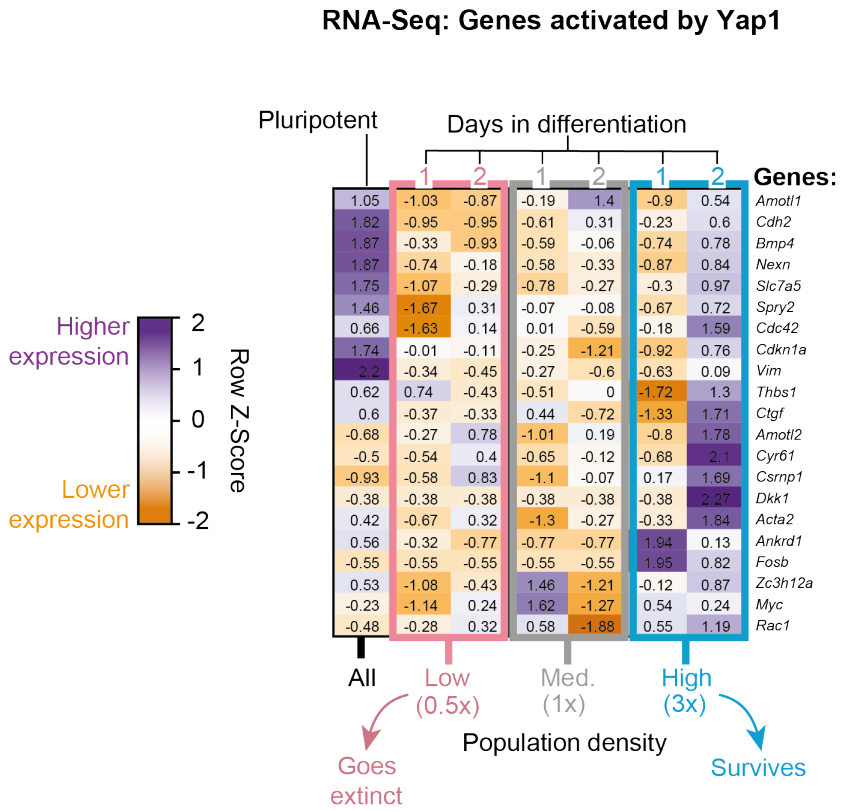

Biological process (GO terms)
Fold enrichment
RNA processing (GO:0006396)
mRNA metabolic process (GO:0016071)
cellular nitrogen compound metabolic process (GO:0034641)
translational elongation (GO:0006414)
cellular amide metabolic process (GO:0043603)
gene expression (GO:0010467)
ribonucleoprotein complex biogenesis (GO:0022613)
macromolecule metabolic process (GO:0043170)
mRNA processing (GO:0006397)
nitrogen compound metabolic process (GO:0006807)
cytoplasmic translation (GO:0002181)
cellular component organization or biogenesis (GO:0071840)
nucleobase-containing compound metabolic process (GO:0006139)
primary metabolic process (GO:0044238)
cellular metabolic process (GO:0044237)
heterocycle metabolic process (GO:0046483)
macromolecule catabolic process (GO:0009057)
cellular aromatic compound metabolic process (GO:0006725)
organic cyclic compound metabolic process (GO:1901360)
cellular protein metabolic process (GO:0044267)
organic substance metabolic process (GO:0071704)
metabolic process (GO:0008152)
mitochondrion organization (GO:0007005)
cellular macromolecule metabolic process (GO:0044260)
protein metabolic process (GO:0019538)
ncRNA metabolic process (GO:0034660)
rRNA metabolic process (GO:0016072)
organonitrogen compound metabolic process (GO:1901564)
DNA repair (GO:0006281)
protein modification by small protein conjugation (GO:0032446)
establishment of protein localization to organelle (GO:0072594)
organonitrogen compound biosynthetic process (GO:1901566)
protein modification by small protein conjugation or removal (GO:0070647)
cellular process (GO:0009987)
transcription initiation from RNA polymerase II promoter (GO:0006367)
catabolic process (GO:0009056)
0 1 2 3 4 5 6 7 8 9
-log10(P-value)

Figure S13 – Enrichment analysis of RNA-Seq data reveals that high-density populations, compared to low-density populations, have higher levels of processes (GO terms) such as multicellular organismal processes, cell-cell signaling, neurological system processes, and cell adhesion. We identified possible, intracellular pathways that the secreted molecules control by performing a transcriptome-level profiling (RNA-Seq) to examine differentially expressed genes. For this, we used Retinoic Acid (RA) to cause 46C cells to differentiate towards the Neuroectoderm (NE) lineage. We performed RNA-Seq on populations of three different starting densities: (1) a low-density (863 cells/cm^2) population that becomes extinct; (2) a high-density (5172 cells/cm^2) population that grows towards the carrying capacity; and (3) a medium-density (1724 cells/cm^2) population that is near the threshold density. For RNA-Seq, we collected all cells from these populations on the first and second days after triggering differentiation. We also collected cells that were kept pluripotent in a serum-based pluripotency medium (FBS with LIF) as a comparison. We analyzed the resulting transcriptome expression levels (FPKMs) for each gene by focusing on gene-expression levels that differed by more than 2-folds between the high- and low-density populations. **(A)** Volcano plot. Each grey dot represents a single gene. The horizontal axis shows the relative expression level: the expression level of the high-density population divided by the expression level of the low-density population. The vertical axis shows the p-value of the statistical test performed in Cufflinks (see Methods). Genes that are more highly expressed by the high-density population than the low-density population, by 2-folds or more, are shown as red points. Genes that are more highly expressed by the low-density population than the high-density population, by 2-folds or more, are shown as blue points. **(B)** Enrichment analysis for genes that are more highly expressed by the high-density population than the low-density population by 2-folds or more (red data points in (A)). We used PANTHER and a custom MATLAB script to examine which GO terms (biological processes) are enriched and determine the corresponding significance (p-value) of the enrichment. We list here the enriched GO terms with their GO-term numbers. This plot shows that GO terms such as “multicellular organismal processes”, “cell-cell signaling”, “neurological system processes” and “cell adhesion” have some of the highest fold enrichments. **(C)** Enrichment analysis for genes that are more highly expressed by the low-density population than the high-density population by 2-folds or more (blue data points in (A)). We list here the enriched GO terms with their GO-term numbers. This plot shows that GO terms such as “RNA processing” and “macromolecule catabolic processes” have some of the highest fold enrichments. See also Fig. S14 for further analyses of the RNA-Seq dataset.

A



B

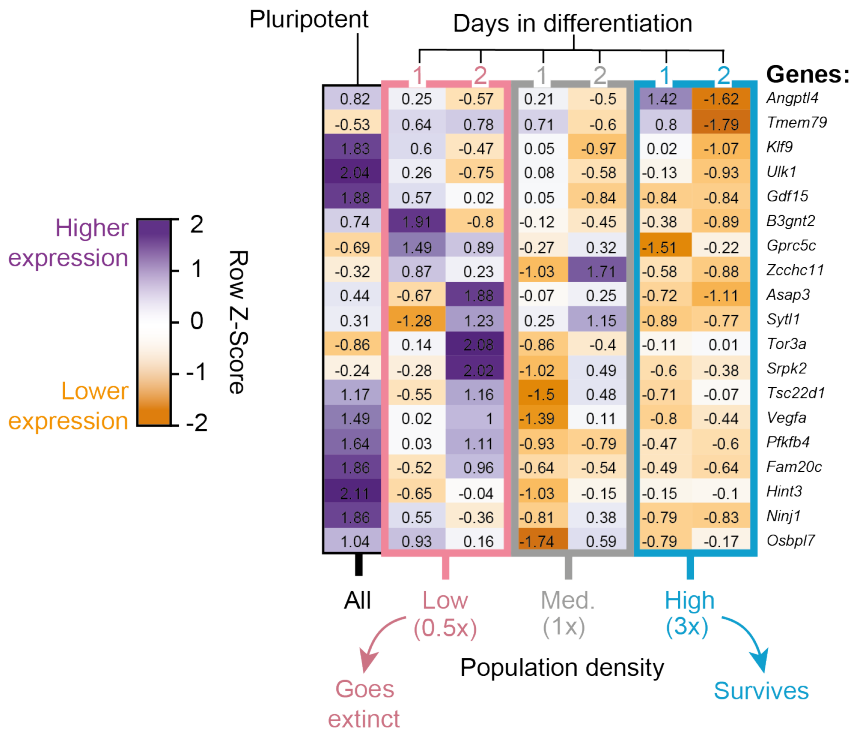
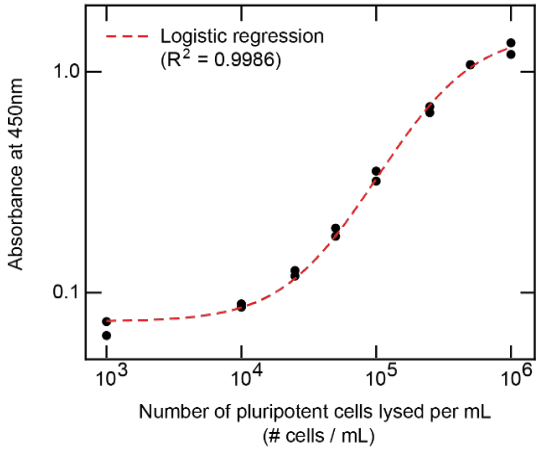


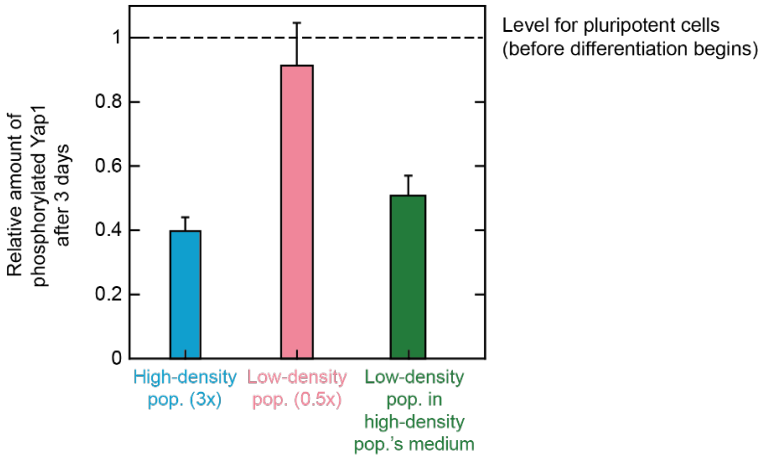
Figure S14 – RNA-Seq analysis reveals that Yap1 becomes more active in populations that start with higher densities. Same RNA-Seq dataset as in Fig. S13. Yap1 is a key

component of the Hippo signaling pathway that is important for cell proliferation and apoptosis. We zoomed into Yap1-related genes in our RNA-Seq dataset because the analysis of enriched GO terms (Fig. S13) revealed several Yap1-related genes prominently participating in the "cell adhesion" processes (one of the top enriched GO terms). Details of the RNA-Seq are in the caption for Fig. S13. In brief, we performed RNA-Seq on populations of three different starting densities: (1) a low-density (863 cells/cm^2) population that becomes extinct; (2) a high-density (5172 cells/cm^2) population that grows towards the carrying capacity; and (3) a medium-density (1724 cells/cm^2) population that is near the threshold density. For RNA-Seq, we collected all cells from these populations on the first and second days after triggering differentiation. We also collected cells that were kept pluripotent in a serum-based pluripotency medium (FBS with LIF) as a comparison. We analyzed the expression levels (FPKMs) by classifying genes into two groups: (1) genes that are activated by Yap1; and (2) genes that are repressed by Yap1 (*33,34,36-41*). For each gene, we compute its mean expression level μ by averaging the its expression level across all experimental conditions (i.e., across all densities and days). Afterwards, we determined the row Z-score for each gene and experimental condition, which is a measure of by how much a gene's expression level in a given experimental condition deviates from the mean expression level (μ) for that gene. A gene that is more highly expressed has a high row Z-score (close to ~ 2) and is indicated as a shade of purple in the heatmaps here. A gene that is more lowly expressed has a low row Z-score (close to -2) and is indicated as a shade of orange in the heatmaps here. **(A)** Heat map that shows the row Z-score for each gene (each row) and each experimental condition (each column). These genes are known to be either directly or indirectly activated by Yap. We observed that these genes, including *Cyr61* and *Amotl2,* were more highly expressed (purple color) by the higher-density populations than the lower-density populations, suggesting that Yap1 is more active in higher-density populations. **(B)** Heat map that shows the row Z-scores for each gene (each row) and each experimental condition (each column). These genes are known to be either directly or indirectly repressed by Yap1. We observed that these genes, including *Angptl4* and *Tmem79*, were more highly expressed (purple color) by lower-density populations than the higher-density populations. Taken together, the results here (A-B) are consistent with Yap1 becoming more active in higher-density populations.

A



B



C

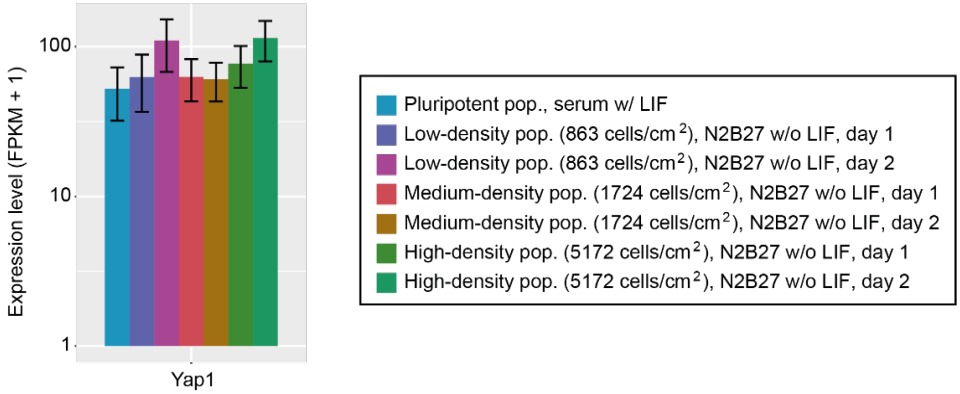


Figure S15 – Populations that survive differentiation have more active (dephosphorylated) Yap1 compared to populations that become extinct during differentiation. To determine how the Yap1 activity may be determined by the population density, we performed ELISA that specifically detects inactive (phosphorylated) Yap1 - Yap1 phosphorylated at Ser397, which is a primary phosphorylation site (*31,32*) (see Methods). **(A)**

Standard curve for the ELISA. Here we lysed pluripotent populations of various densities and then measured the amount of phosphorylated Yap1 in each lysate - this yields an optical absorbance value at 450 nm (black points). Duplicates for each lysate are shown. Red curve is the logistic fit function: $f(x) = D + \frac{A-D}{1+(\frac{x}{C})^B}$, with *A, B, C* and *D* are fit constants and x is the number of cells lysed per mL. Results: A = 0.074, B = 1.418, C = 2.95 x 10^5 and D = 1.502 with an R^2 = 0.9986. **(B)** We performed ELISA on the 46C cells that were induced to differentiate towards the Neuroectoderm (NE) lineage by Retinoic Acid (RA). We examined three populations: (1) high-density population (5172 cells/cm^2); (2) low-density population (863 cells/cm^2); and (3) low-density population that we rescued from extinction by transplanting it, after two days, into the high-density population's medium. For each population, we measured its level of phosphorylated (inactive) Yap1 three days after starting differentiation and then normalized this level to the level present in pluripotent cells of the same density (shown in (A)). This yielded a "relative abundance" for each of the three populations. Compared to the pluripotent cells of equivalent density, cells of the low-density population (pink bar) had ~10% fewer inactive Yap1 whereas cells of the high-density population (blue bar) had ~60% fewer inactive YAP1 than the pluripotent population of the same density. Cells of the rescued low-density population (green bar) had ~50% (green bar) less inactive Yap1 than pluripotent populations of the same density. Together, these results establish that, after exiting pluripotency, cells of surviving populations have more active (dephosphorylated) Yap1 than cells that head towards extinction. **(C)** Expression level of Yap1 from RNA-Seq dataset. Legend shows different conditions. Note that on each day, the low-density (863 cells/cm^2 initially) and the high-density (5172 cells/cm^2 initially) populations have virtually the same Yap1 expression level. Thus, we can compare the amounts of inactive (phosphorylated) Yap1 between the low- and high-density populations in Fig. 3E (i.e., since both populations have (nearly) the same total level of Yap1, we would be subtracting the amount of inactive Yap1 from the same value for both populations to get the amount of active Yap1). Medium-density (1724 cells/cm^2 initially) population starts with the near-threshold density. *n* = 3 for all plots; Error bars are s.e.m.

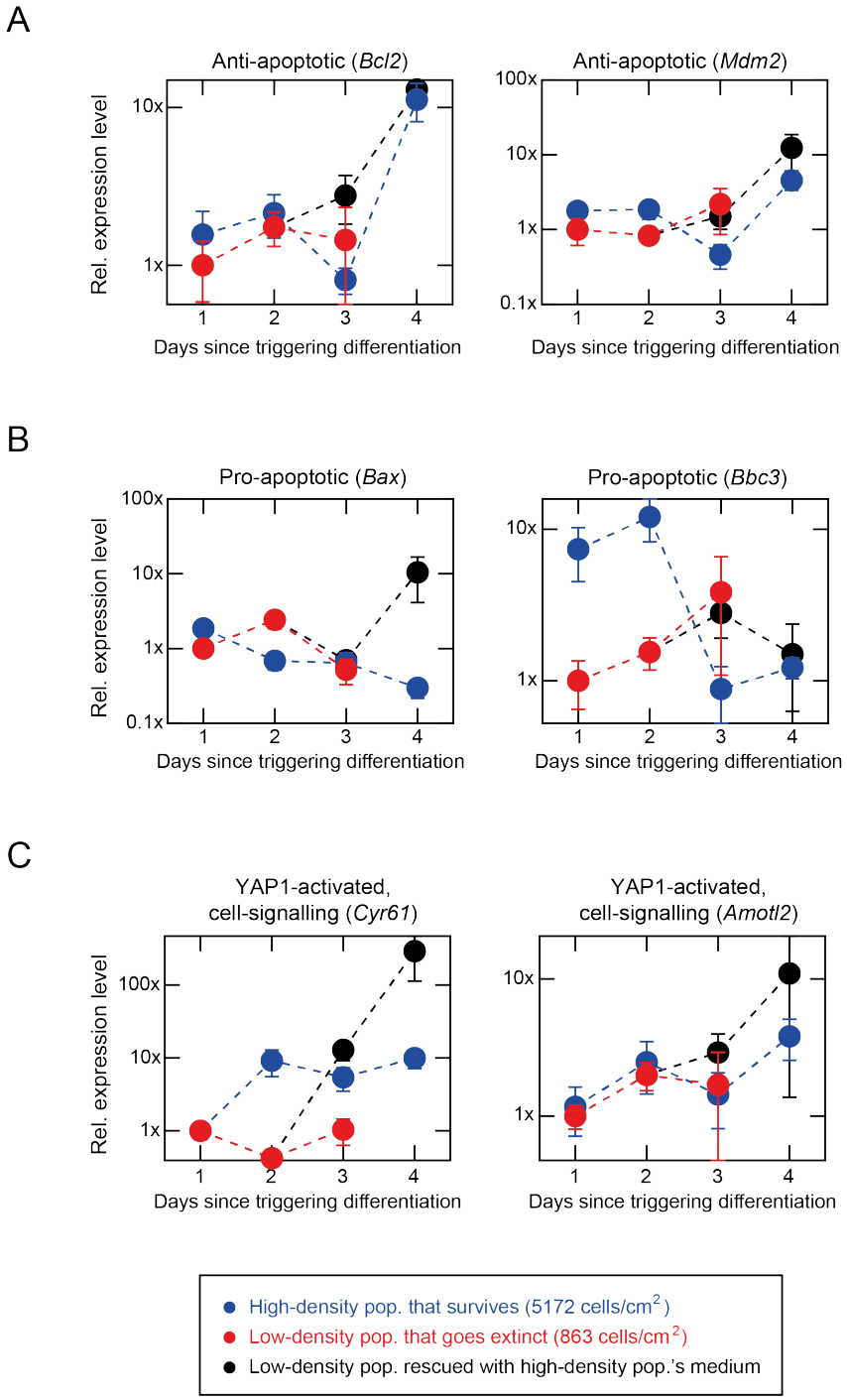


Figure S16 – High-density populations and rescued low-density populations (rescued by high-density population's medium) eventually have elevated anti-apoptotic pathway activities (e.g., expressions of *Bcl2* and *Mdm2* genes, which are controlled by Yap1), and elevated Yap-1-controlled cell-signaling activities (e.g., expressions of *Cyr61* and *Amotl2*). With real-time quantitative PCR (primers in Table S1), we measured anti-apoptotic, pro-apoptotic, and Yap1-mediated cell-signaling genes over the course of differentiation. Normalization of expression values: for each gene *g*, we first divided its expression level by

the expression level of *Gapdh*, resulting in a value N_g. For each population, we divided its N_g by the low-density population's N_g on day 1 to get the final, normalized expression level μ which is plotted in here in all graphs. Thus, "1x" is the expression level of the low-density population on the first day after starting differentiation. (**A)** Expression levels of two anti-apoptotic genes, *Bcl2* (left graph) and *Mdm2* (right graph). Data for *Bcl2* is a replicate of the data shown in Fig. 3F which we show here for comparison with *Mdm2*. Both Bcl2 and Mdm2 show increased expressions (more anti-apoptotic) for high-density population (blue) and low-density population that was rescued by the medium of the high-density population after the 2nd day (black). Low-density population that goes extinct (red) shows nearly constant, low expression level of both genes. No data for 4th day is shown for the low-density population because it becomes extinct after the 3rd day (there were already barely any cells left for the 3rd day data shown here). **(B)** Expression levels of two pro-apoptotic genes - *Bax* (left graph) and *Bbc3* (right graph). Color scheme is the same as in (A). The high-density population initially has a higher Bbc3 expression than the low-density population but eventually down-regulates and has lower Bbc3 expression than the low-density population. The low-density population, in turn, gradually increases its Bbc3 expression over time, up to the moment of extinction (~ day 3). Note that the rescued low-density population keeps its Bbc3 expression level low, past day 2 (which is when it receives the medium from a high-density population) and has nearly same low Bbc3 expression as the high-density population after being rescued. Note that differentiation is known to increase expression of apoptotic genes. **(C)** Expression levels of two cell-signaling genes that are upregulated by Yap1, *Cyr61* (left graph) and *Amotl2* (right graph). Only the high-density and the rescued low-density populations gradually increase the expression levels of both genes whereas the low-density population that heads towards extinction (red) maintains a nearly constant, low expression of both genes (consistent with our findings in Fig. 3 that secreted factors that are abundant for high-density populations increase Yap1 activity (and thus upregulate expression of *Cyr61* and *Amotl2*). In all the plots, n = 3; Error bars are s.e.m.

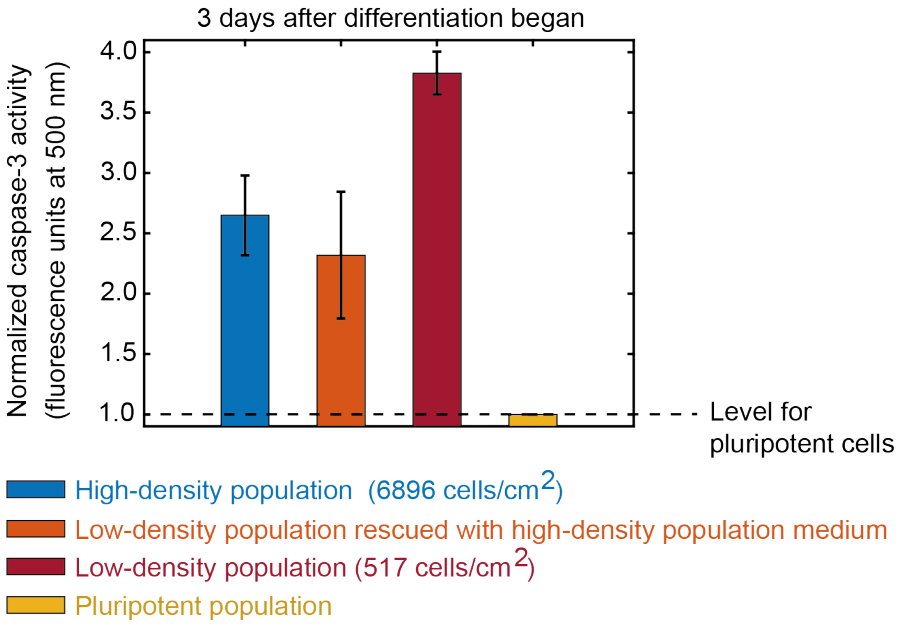


Figure S17 – Pro-apoptotic caspase-3 activity is higher in low-density populations than in high-density and rescued low-density populations. We sought to examine the activity of a well-known, pro-apoptotic marker (caspase 3) in populations that are either extinction-bound or surviving. We used a membrane-permeable, DNA-dye-based assay (NucView 488 Caspase-3 Assay Kit for Live Cells). This assay measured the amounts of active caspase 3/7 inside cells. Data shown for E14 cells that were induced to differentiation towards the neuroectoderm lineage by retinoic acid. We examined casp-3 levels in four different populations: (1) high-density population (6896 cells/cm^2 –blue bar); low-density population (517 cells/cm^2 –red bar); (3) low-density population that was rescued from extinction by transplanting it, after two days, into the high-density population's medium (orange bar); and (4) pluripotent E14 cells before we induced the differentiation (yellow bar). After 3 days of differentiation, we collected the cells from each of these populations, mixed them with the DNA-dye according to the manufacturer's protocol, and then measured the resulting fluorescence at 500 nm in single cells with a flow cytometer. Plotted here are the geometric means of the fluorescence for each population. Higher fluorescence means more Casp-3 activity. We normalized the values to the fluorescence level of the pluripotent population (yellow bar), as indicated by the dashed line. n =3; Error bars are s.e.m. This plot shows that all differentiating populations upregulated levels of Casp-3 activity relative to the pluripotent populations. Thus, the nonlocal communication causes populations to have less Casp-3 activities than extinction-bound populations.

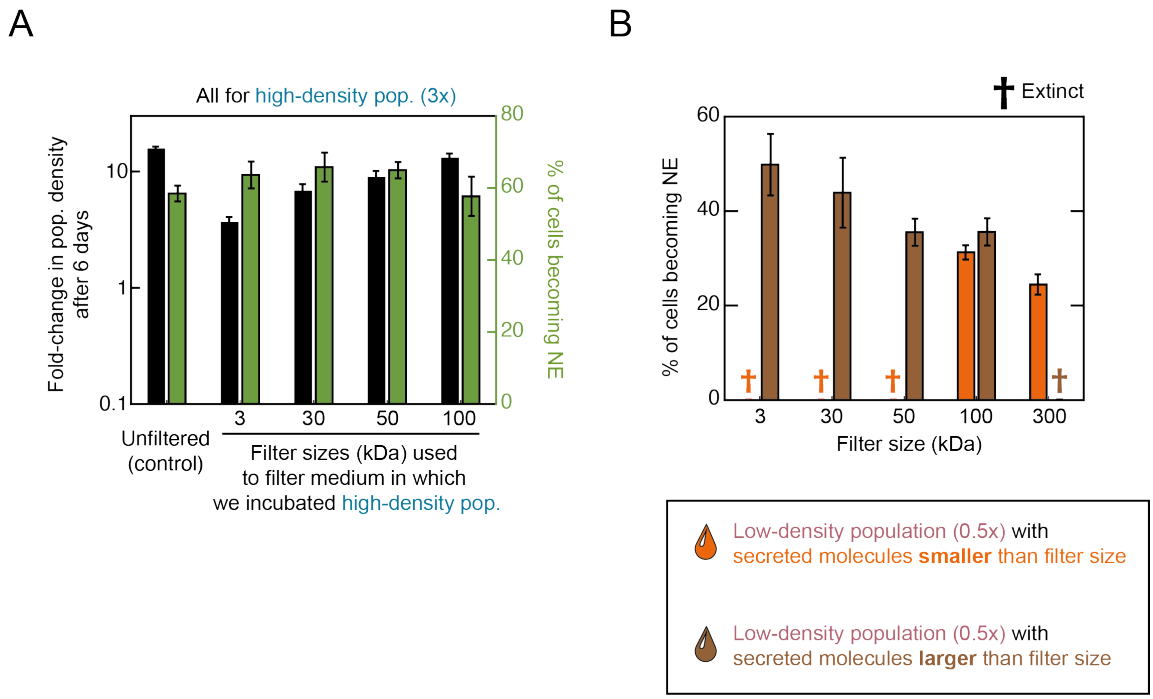


Figure S18 – Filtering culture medium does not eliminate media components that are essential for cell growth. We used commercial, membrane-based filters of various sizes (Fig. 4A-B, see Methods). Data shown for 46C cells induced to differentiate towards the neuroectoderm (NE) linage with Retinoic Acid (RA). **(A)** To check that the filters do not remove any essential media components for cell growth, we filtered the medium of a high-density population with filters of various sizes, after two days of differentiation. We then took the filtered medium (medium that passed through the filter and thus containing all molecules that are smaller than the filter size), and gave it back to the same high-density population. We incubated the population for four days in this medium. On the last day, we measured the fold-change in population density (black bars) and the percentage of cells becoming NE (green bars). As a control, we also measured the fold-change in population density after six days of growth in unfiltered medium (first black and green bars). Since all black bars have nearly the same height as do all the green bars, we can conclude that none of the filters catch any ingredients in the cell-culture medium that are essential for cell growth (e.g., vitamins and other components which are already present in the medium from the beginning of cell culture, rather than secreted by cells). **(B)** We cultured a low-density population (863 cells/cm^2) in a medium from a high-density population (5172 cells/cm^2) that went through the filter (orange bars). This medium has all molecules that are smaller than the filter size. We also cultured the low-density population in a medium that contained all the molecules that are larger (heavier) than the filter size (brown bars), which we captured by flowing the high-density population's medium through the filter. We used a wide range of filter sizes (horizontal axis). The bars show

the percentages of cells entering the NE after being cultured in the filtered medium. Crosses indicate that the population became extinct in the filtered medium. The corresponding fold-changes in population density are shown in Fig. 4B.

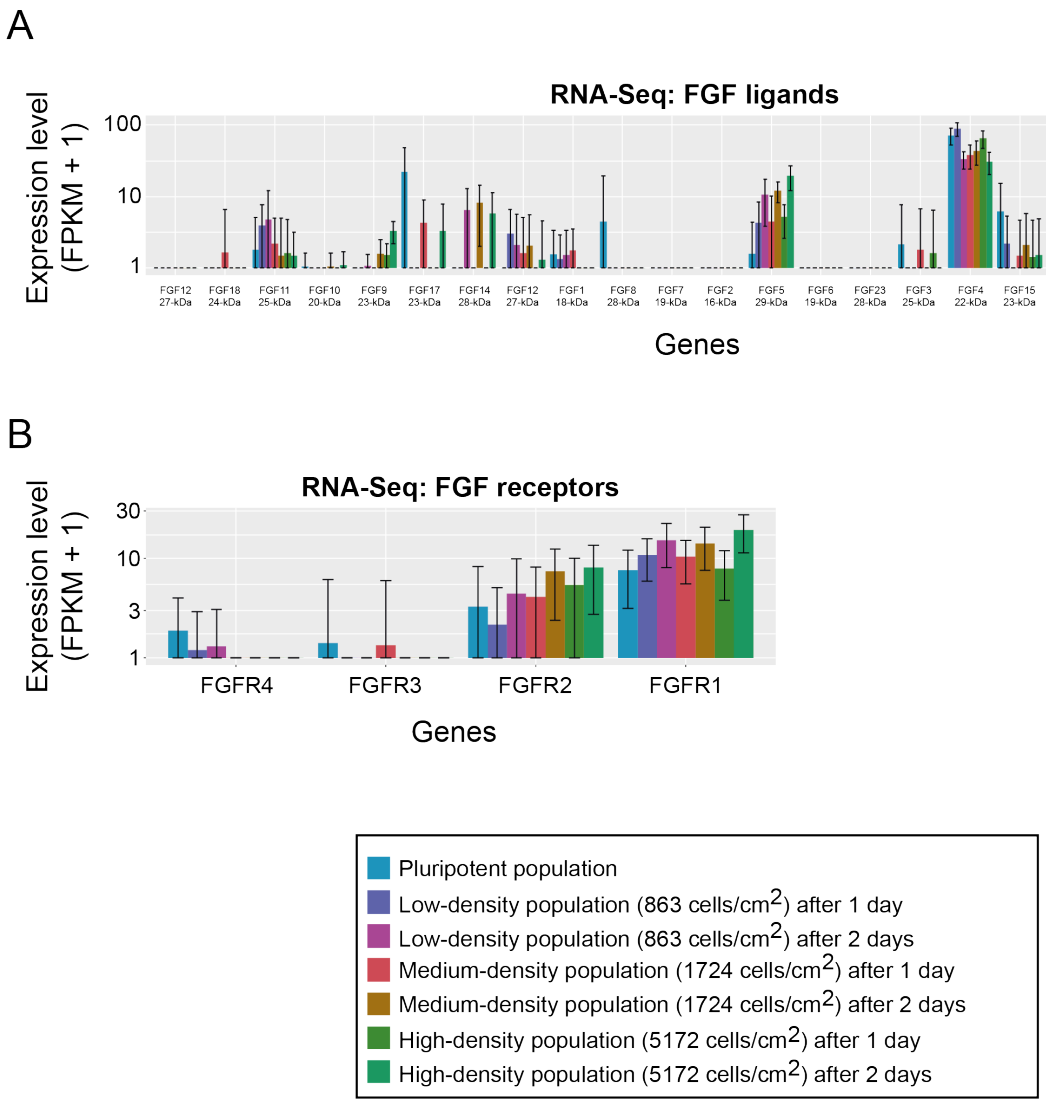


Figure S19 – RNA-Seq: Expression levels of all 22 FGFs and their receptors. We performed RNA-Seq to measure the expression levels of all 22 FGF ligands (A) and their receptors (FGFRs) (B). Data for 46C cells that were induced to differentiate towards the neuroectoderm (NE) lineage with Retinoic Acid (RA). We performed RNA-Seq on four populations: (1) pluripotent population prior to differentiation; (2) low-density (863 cells/cm^2) population; (3) high-density (5172 cells/cm^2) population; and (4) medium-density (1724 cells/cm^2) population that is near the threshold density. For the three differentiating populations, we collected their cells on the first and second day after triggering differentiation. **(A)** Expression levels of all FGF ligands. Shown are the following genes: *Fgf1-8*, *Fgf20-21* and *Fgf23*. Below each gene name is the corresponding molecular weight in kDa. Note that FGF4 expression prominently stands out among all the FGFs. **(B)** Expression levels of all FGF receptors (FGFRs). Shown are the following genes: *Fgfr1-4*. Below each gene name is the

corresponding molecular weight in kDa, according to two online resources: Uniprot and ExPASy. *n* = 3 for all plots; Error bars are s.e.m.

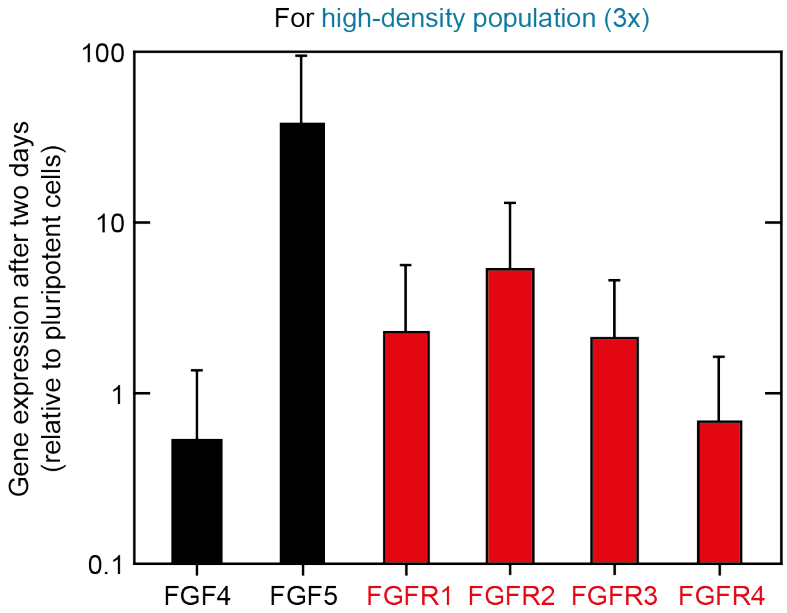


Figure S20 – Differentiating populations express *Fgf4*, *Fgf5* and all four FGF receptor genes during the first 2 days of differentiation. We used real-time quantitative PCR (RT-qPCR) to measure the expression levels of all four receptors (*Fgfr1-4*) of Fibroblast Growth Factors (FGFs) and the expression levels of the two FGFs, *Fgf4* and *Fgf5* (primers in Table S1). Data for 46 cells induced to differentiation towards the neuroectoderm (NE) lineage with Retinoic Acid (RA). We examined a high-density population (5172 cells/cm^2) after two days of differentiation. We normalized the resulting expressions of a gene relative to that of the housekeeping gene, *Gapdh* of the same population, and then further normalized the resulting value to the pluripotent population's normalized expression level (similar to the procedure described in the caption for Fig. S16). Thus, a given gene's expression level is compared to the pluripotent population's expression level for that gene. Normalized expression levels of *Fgf4* and *Fgf5* (in black) and *Fgfr1*-2 (in red). n = 3; Error bars are s.e.m. Altogether, these results show that *Fgf4*, *Fgf5* and *Fgfr1-2* are expressed – and some more so than the pluripotent population (i.e., expression value greater than 1) – during the first 2 days in which ES cells exit pluripotency.

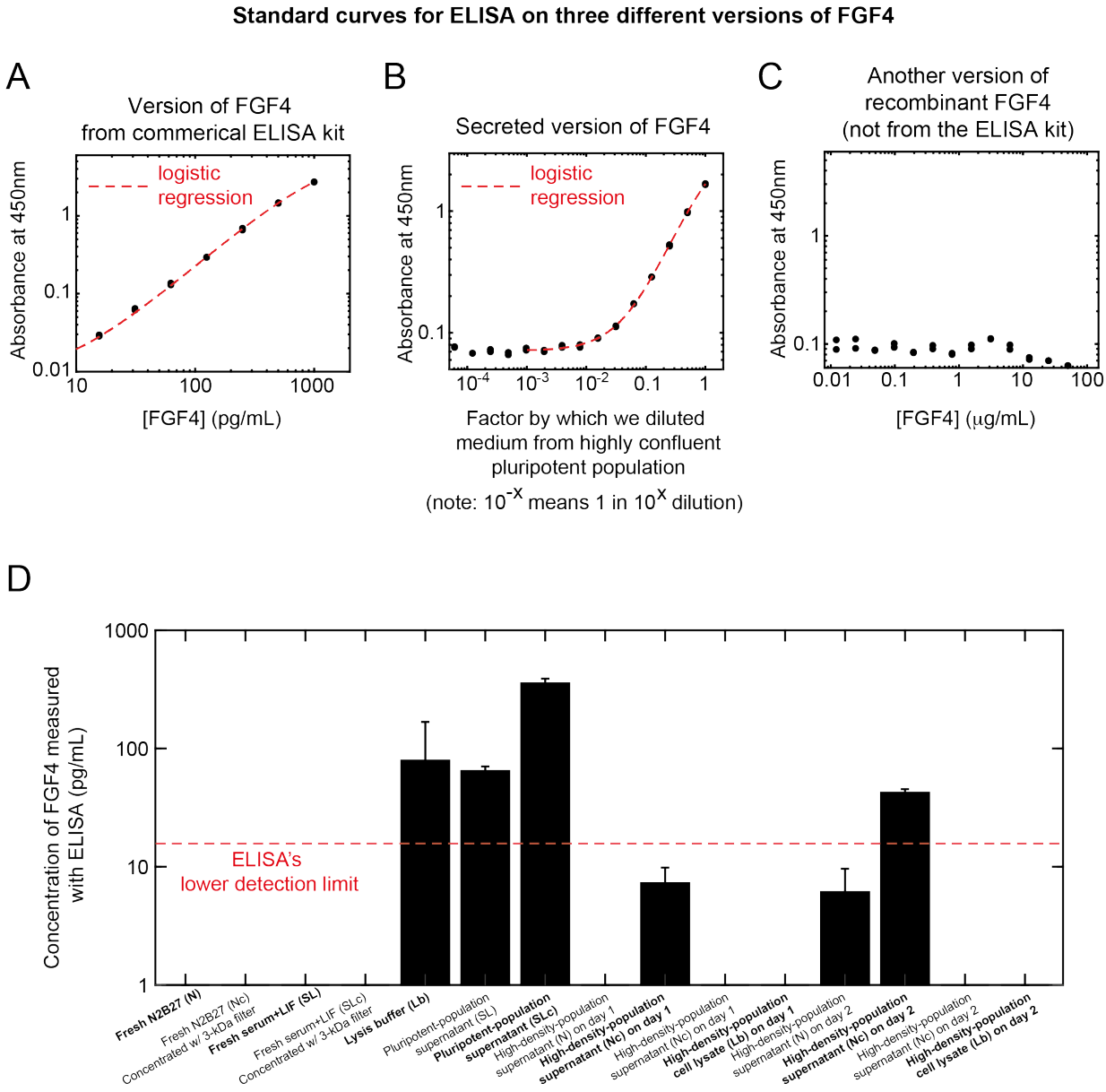


Figure S21 - Cells secrete and extracellularly accumulate appreciable amounts of FGF4 during the first 2 days of differentiation. We performed ELISA that detects mouse FGF4. **(A)** Standard curve based on a recombinant mouse FGF4 that came with the commercial ELISA kit. Note that this FGF4 is not necessarily the same version as the FGF4 that our cells secrete. Each measurement (absorbance at 450 nm) was done in duplicate (black data points). Then, we performed a logistic regression on the data by fitting a 4-parameter logistic function (red curve): $f(x) = D + \frac{A-D}{1+(\frac{x}{C})^B}$, where *A, B, C* and *D* are constant coefficients and x is the known concentration of the recombinant mouse FGF4 that we added. We found: A = 0.0084, B = 1.313, C = 1312 and D = 6.59. **(B)** Standard curve based on the version of FGF4

that pluripotent cells secrete into their medium. We first concentrated the medium taken from a highly confluent (~80% confluent) pluripotent population with a 3-kDa filter and then performed ELISA on serially diluted fractions of this concentrated medium. Each measurement (absorbance at 450 nm) was done in duplicate (black points). Then, we performed a logistic regression on the black data points by fitting a 4-parameter logistic function (red curve): $f(x) = D + \frac{A-D}{1+(\frac{x}{C})^B}$, where *A, B, C* and *D* as constant coefficients and x as amount of lysed, pluripotent cells (day 0). We found: A = 0.07123, B = 1.184, C = 1.407 and D = 4.069. The standard curve shows that pluripotent ES cells secrete a version of FGF4 that our ELISA can detect. Moreover, it also shows a limitation of our ELISA: the assay can only detect sufficiently high concentration of FGF4 as seen by the fact that it cannot detect any FGF4 in a 1:100 dilution of a concentrated medium taken from a highly confluent ES cells. **(C)** Standard curve based on recombinant mouse FGF4 from a different manufacturer (not from the ELISA kit) that we could add to the cell-culture medium to rescue low-density populations. Each measurement (absorbance at 450 nm) was done in duplicates (black points). As seen here, the ELISA cannot detect any amounts of this version of FGF4, even when its concentration is 100-folds higher than the highest concentration - of the version supplied by the ELISA kit - that we used in (A). The three standard curves (A-C) show that ELISA is highly sensitive to the form of FGF4 - we used three different forms in each of (A-C). The two versions of FGF4 that are not supplied by the ELISA kit (B-C) are detected with lower efficiency than the version supplied by the kit (A). **(D)** ELISA measurements of secreted FGF4 (in pg/mL) in various conditions (indicated with labels on the horizontal axis). We detected abundant FGF4 in the pluripotency medium (~500 pg/mL). We also performed ELISA on 46C cells that were induced to differentiate towards the neuroectoderm (NE) lineage with Retinoic Acid (RA) (see Methods). In the medium of the high-density population (8620 cells/cm^2) after two days of differentiation, we detected ~50 pg/mL of FGF4. After 1 day of differentiation, the medium of the high-density population did not contain any detectable amounts of FGF4. Hence, high-density populations take 2 days to accumulate appreciable (detectable) amounts of FGF4.

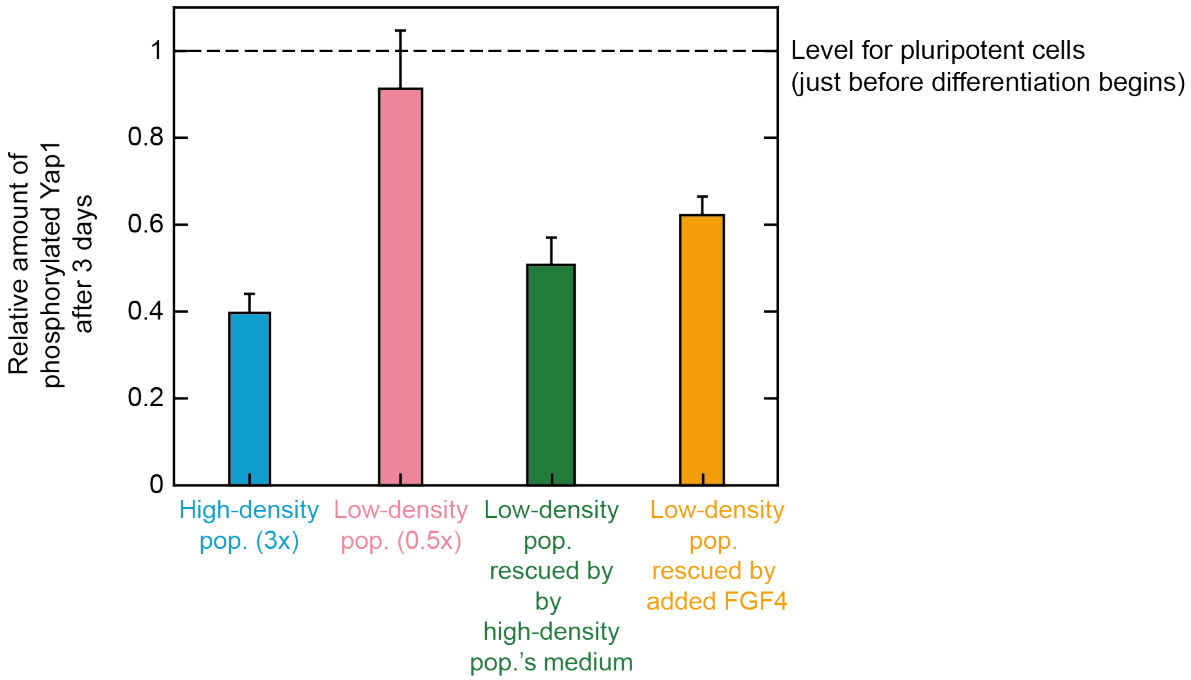


Figure S22 – Low-density population rescued by added FGF4 in the medium has similar, high level of active (dephosphorylated) Yap1 level as high-density population and a low-density population that was rescued by the medium of the high-density population. We performed ELISA that measured the phosphorylated Yap1 levels in 46C cells that were induced to differentiate towards the Neuroectoderm (NE) lineage by Retinoic Acid (RA). The first three bars (blue, pink, and green) are the same data shown in Fig. S15 and are shown here for comparison with the orange bar. Orange bar shows the level of inactive (phosphorylated) Yap1 level for a low-density population that was rescued by a high concentration of FGF4 that we added into its medium. n = 3; Error bars are s.e.m.

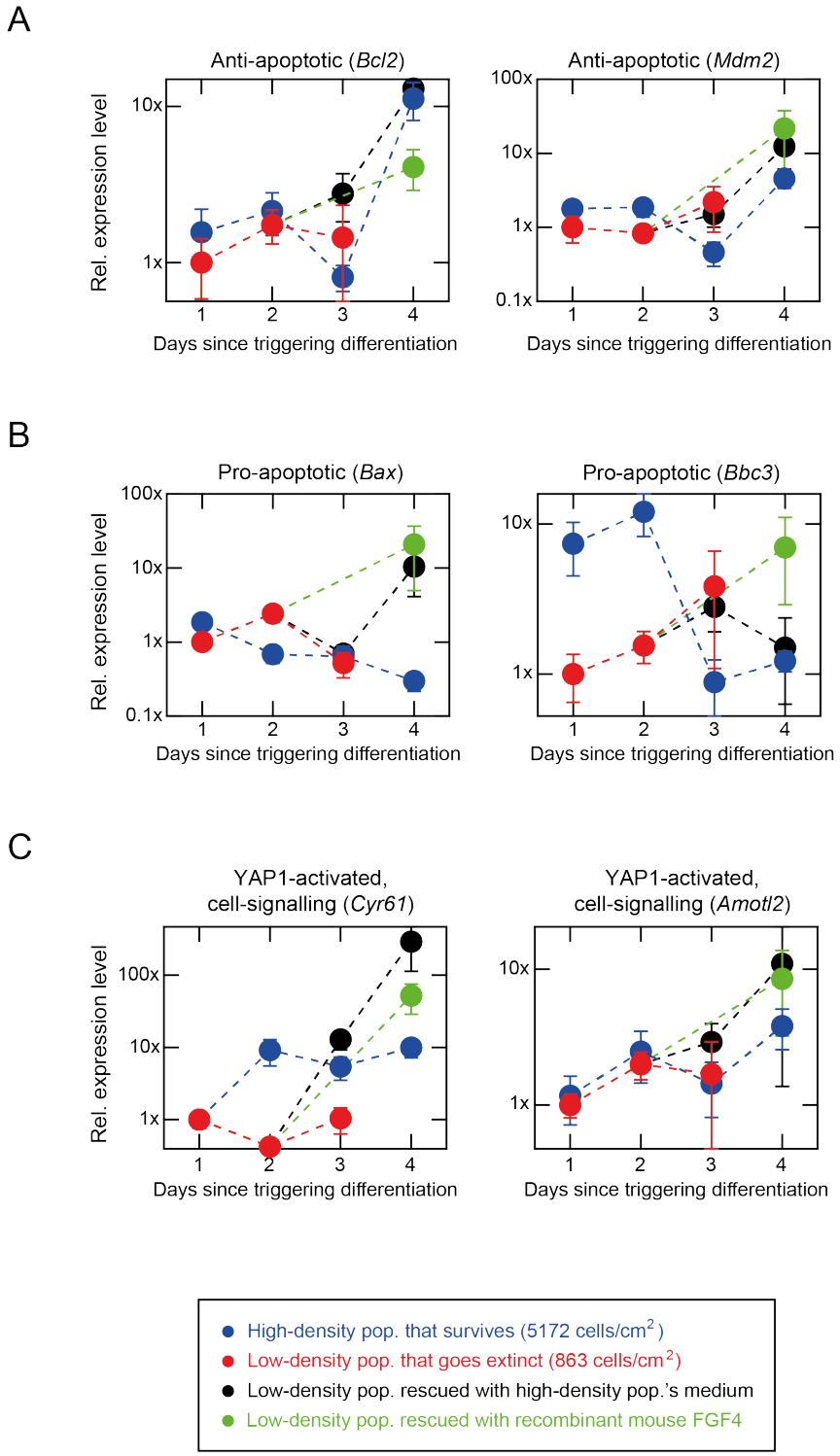


Figure S23 – Low-density populations that are rescued from extinction with recombinant mouse FGF4 activate anti-apoptotic pathways (e.g., *Bcl2* and *Mdm2* genes) and YAP1-controlled targets (e.g., *Cyr61* and *Amotl2* genes). The RT-qPCR data shown in blue, red, and black are the same as the data shown in Fig. S15. We show them here again for comparison with the green data points. Green data points are for a low-density

population that we rescued by adding sufficient amount of FGF4 to its medium, just after two days into differentiation (hence the data points for "day 1" and "day 2" are the same as the red data points). Expression levels of **(A)** anti-apoptotic genes, **(B)** pro-apoptotic genes, and **(C)** Yap1-activated cell-signaling genes measured with RT-qPCR. Primers are listed in Table S1. In all the plots, n = 3; Error bars are s.e.m.

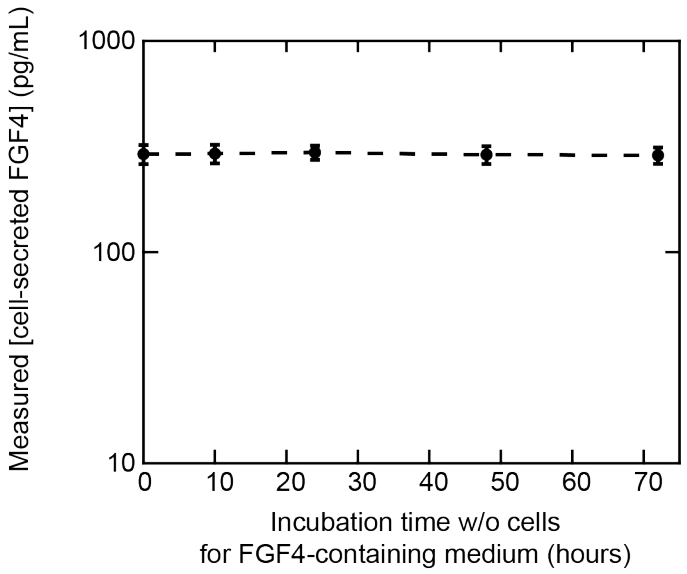
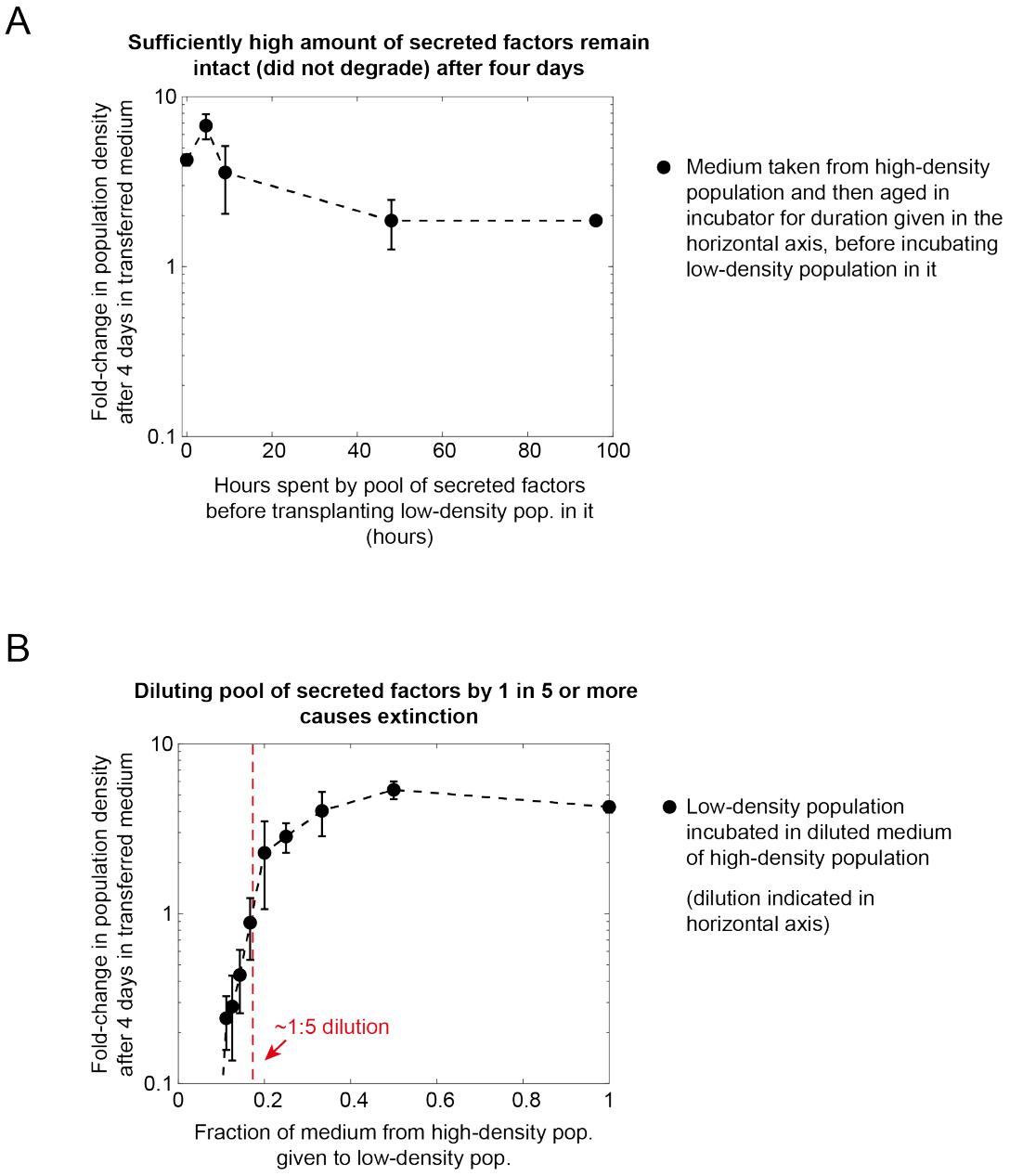


Figure S24 – Secreted FGF4 shows no appreciable degradation for 3 days at 37°C in liquid medium. Determining the degradation rate of FGF4 that cells secrete tells us whether FGF4 can diffuse by millimeters or not, through the Stokes-Einstein equation (see Supplementary text). We performed ELISA that targets FGF4 (see Methods) to determine the concentrations of secreted FGF4 when incubated without cells in liquid medium. For this, we used a 3-kDa filter to concentration the pluripotency medium taken from a confluent population of 46C cells. Then we incubated the medium without any cells in a 37°C incubator for the hours indicated on the horizontal axis. We then took it out of the incubator and performed ELISA on it to measure the remaining [FGF4]. We observed that after incubating for 72 hours (3 days), the initial concentration of the secreted FGF4 was not appreciably degraded. This result suggests that the concentration of secreted FGF4, by itself and in the absence of cells, is stable over at least 3 days and thus – according to the Stokes-Einstein equation (see Supplementary Text) – can diffuse over millimeters (see Supplementary text).



(Note: 0.1 means 1 in 10 dilution of the high-density pop.'s medium into fresh medium)

Figure S25 – Secreted factors that control differentiating population's survival-vs-extinction fate have a combined, effective half-life of at least 2 days. According to the reaction-diffusion equation, molecules of ~100 kDa in an aqueous environment need to have a half-life of at least ~12 hours to have a diffusion length that is over 1 mm (see Supplementary text). We sought to infer the half-lives of all the secreted molecules that are important for determining the survival-versus-extinction fate of a population. What is important is actually not the individual half-life of each molecule but rather the effective half-life of all molecules combined. Note that we established that more than one secreted factor determines the survival-versus-extinction fate of a population (Fig. 5B). To determine the effect half-life of all molecules combined, we used 46C cells that were induced to differentiate towards the

neuroectoderm lineage with Retinoic Acid (RA). We used two populations: (1) high-density population (5172 cells/cm^2); and (2) low-density population (862 cells/cm^2). Combining the results of (A) and (B) enabled us to infer the lower bound on the effective half-life, as we now explain. **(A)** After 2 days of differentiation with N2B27, we transferred the high-density population's medium to an empty plate that had no cells. We incubated the medium without any cells in a 37°C incubator for various amounts of time before transferring it to a low-density population that was just ending its second day of differentiation. After 4 days of incubation in the transferred medium (so a total of 6 days of differentiation), we measured the fold-change in density of the low-density population (black points). We plotted the results here as a function of the amount of time the medium spent in the incubator without any cells before we transferred it to the low-density population. The result shows that ageing the medium for 96 hours in 37°C before transferring it to the low-density population still results in rescuing of the low-density population (fold change in population density > 1). The fold change achievable does decrease as the medium's age increases, from ~4-fold (for unaged medium) to ~2-fold (for medium aged for ~96 hours). n = 3; error bars are s.e.m. **(B)** In a parallel experiment, we took the medium of the high-density population after two days of differentiation. Then, we diluted it by different amounts into a fresh differentiation medium (one that never harbored any cells). We incubated a 2-days-old low-density population into the diluted medium and then measured the fold-change in its density after four days (so a total of 6 days of differentiation). Plotted here is the fold-change in the population density as a function of how much of the medium from the high-density population was mixed with the fresh medium (e.g., 0.2 on the horizontal axis means a dilution by 1 in 5). The red line shows the maximum dilution that is allowed for still rescuing the low-density population. Any higher dilution causes the low-density population to have a fold-change in density that is less than one. n = 3; error bars are s.e.m.

Combining the results of (A) and (B), we can conclude that more than 1/5 of the secreted molecules remain in the medium after four days of ageing in (A) since, for otherwise, the results in (B) tell us that the fold-change in (A) for medium that was aged for 96 hours should be less than 1, which is not the case. In fact, using the same reasoning, we can say that the effective, combined half-life of the secreted molecules is at least two days. To see, this, note that a half-life of one day would mean that after four days, we would have 1/16 of the molecules degraded after four days since $1/2^4 = 1/16$. But $1/16 < 1/5$, which would mean that the low-density population should have become extinct in the medium that we aged for 96 hours in (A). This is not the case. Repeating the calculation by assuming that the effective, combined half-life is two days leads to: $1/2^2 = 1/4 > 1/5$, which is consistent with the data in (A). In summary, the effective, combined half-life of the secreted molecules is at least two days.

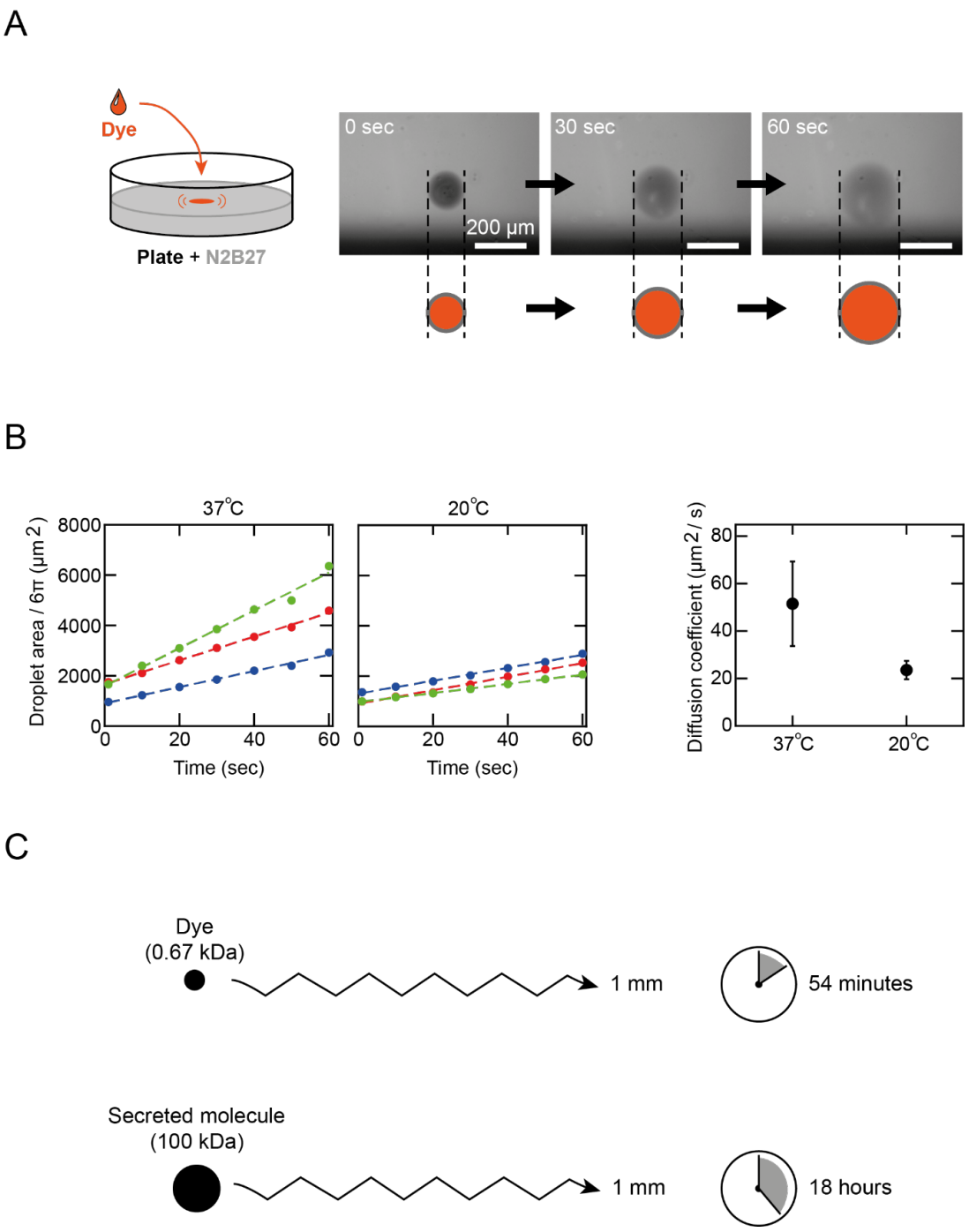


Figure S26 – Diffusion alone, without any other mechanism of transport, explains the long-range (millimeters-scale) spreading of secreted survival-promoting factors. To further support the idea that diffusion alone spreads the cell-secreted factors in our experiments, we determined how fast a droplet of a dye molecule of a known weight spreads in a differentiation medium without any cells, under the same incubation conditions as our cell cultures. **(A)** We used a gel loading dye (DNA Gel Loading Dye 6X, Thermo Scientific, #R0611) which consists of two molecules: bromophenol blue (669.96 Da) and xylene cyanol

(538.61 Da). For simplicity, our calculations below will assume that the dye consists of only the heavier molecule, bromophenol blue. We injected a single, 0.5-μL droplet of the dye at the center of a 6-cm diameter plate that contained 5-mL of transparent N2B27 medium, either at room temperature (20°C) or pre-warmed at 37°C. We used a wide-field microscope to make a time-lapse movie with a bird's eye view and snapshots every 10-seconds. Three snapshots (at 0, 30 and 60 seconds; all done at 37°C) of a single droplet of the dye shows the droplet expanding. Scale bar = 200 μm. **(B)** We determined the diffusion constant D of the dye in two ways: using the time-lapse movie and from theory. In the plots, three different colors represent three independent experiments. To determine *D* from the movies, we tracked the visible droplet boundary over time in a movie to plot the droplet area over time (shown in the two plots here at 37°C and 20°C). As shown, the droplet area linearly increased over time, which is consistent with pure diffusion (pure Brownian motion) since the area of a droplet is proportional to the mean squared displacement of a particle. Specifically, for a particle that undergoes a pure three-dimensional diffusion (Brownian motion), its mean squared displacement $\langle R^2 \rangle$ at time *t* is: $\langle R^2 \rangle = 6Dt$. Let *A* be the 2-dimensionally projected area of the droplet. Then, $\langle R^2 \rangle = \frac{A}{\pi}$ and hence, $D = \frac{A}{6\pi t} = \frac{A_{slope}}{6\pi}$, where A_{slope} is the slope of the linear fits to the droplet area as shown in the two plots here. From these fits, the experimentally determined diffusion constants D_{exp} at 37°C and 20°C are 51.5 ± 17.8 μm^2/s and 23.5 ± 3.9 μm^2/s respectively (*n* = 3; error bars are s.e.m.). As a comparison, we determined the diffusion constant D from theory - via the Stokes-Einstein equation which states, $D = \frac{kT}{6\pi\eta r_{dye}}$ where k is the Boltzmann constant, T is temperature, η is the medium's dynamic viscosity, and r_{dye} is the radius of the dye molecule. For water, η = 0.000692 kg/m·s at 37°C and η = 0.001003 kg/m·s at 20°C (from BioNumbers). We conservatively estimated r_{dye} by noting that bromophenol blue consists of ~10 carbon-carbon bonds which would mean that the dye molecule's diameter is 10 x 0.126 nm. For simplicity, we assume that $r_{dye} = 1\ nm$. The Stokes-Einstein equation then states that the dye's diffusion constants D_{theory} at 37°C and 20°C are 328.1 μm^2/s and 214.0 μm^2/s respectively. Hence, $D_{exp} < D_{theory}$. The fact that our analysis relies on the visible (by eye) boundary of the expanding droplet would underestimate the D_{exp} since the dye must be spreading at least as fast as the boundary does. More importantly, if there were significant convection currents in the liquid medium, then D_{exp} would be much larger than the measured value. This argues against there being any significant liquid convection in our cell-culture media. In other words, the dye-based experiment strongly indicates that cell-secreted survival-promoting factors spread out by pure diffusion rather than by convection currents which, according to the dye, are negligible in our cell-culture conditions.

Furthermore, note that $\frac{D_{exp}(37^oC)}{D_{exp}(20^oC)}$ = ~2.2x and $\frac{D_{theory}(37^oC)}{D_{theory}(20^oC)}$ = ~1.5x; the experimental and theoretical values for pure diffusion closely match (proportional to a factor on the order of one). **(C)** Based on Brownian motion in 3 dimensions with the experimentally determined diffusion constant, the dye molecule has a mean squared displacement of $\langle R^2 \rangle$ = 1 mm^2 after $t_{dye} = \frac{\langle R^2 \rangle}{6D}$ = 54 minutes (with $D = D_{exp}(dye, 37^oC)$ = 51.5 μm^2/s). The same calculation, but now based on the Stokes-Einstein estimate of the diffusion constant would yield t_{dye} = 8.5 minutes for a 1 mm^2 mean squared displacement (with $D_{theory}(dye, 37^oC)$ = 328.1 μm^2/s). Hence the theory predicts a faster spreading of dye than experimentally observed - again, arguing against liquid convection or any other mechanism besides diffusion helping to spread the dye. A secreted molecule of 100 kDa would have a mean squared displacement of 1 mm^2 after time, $t_{secreted} = \frac{\langle R^2 \rangle}{6D_{secreted}} = \frac{\langle R^2 \rangle \cdot r_{secreted}}{6 \cdot D_{exp}(dye, 37^oC) \cdot r_{dye}}$ = 18 hours, where we estimate the radius of the molecule to be $r_{secreted}$ = 20 nm (see Supplementary Text). Hence the two days taken to observe appreciable amount of FGF4 and other survival-promoting factor(s) traveling millimeters and accumulating is consistent with the $t_{secreted}$ calculated here (i.e., if $t_{secreted}$ were much larger than two days, then we should not be observing the survival-factors travelling by millimeters within two days).

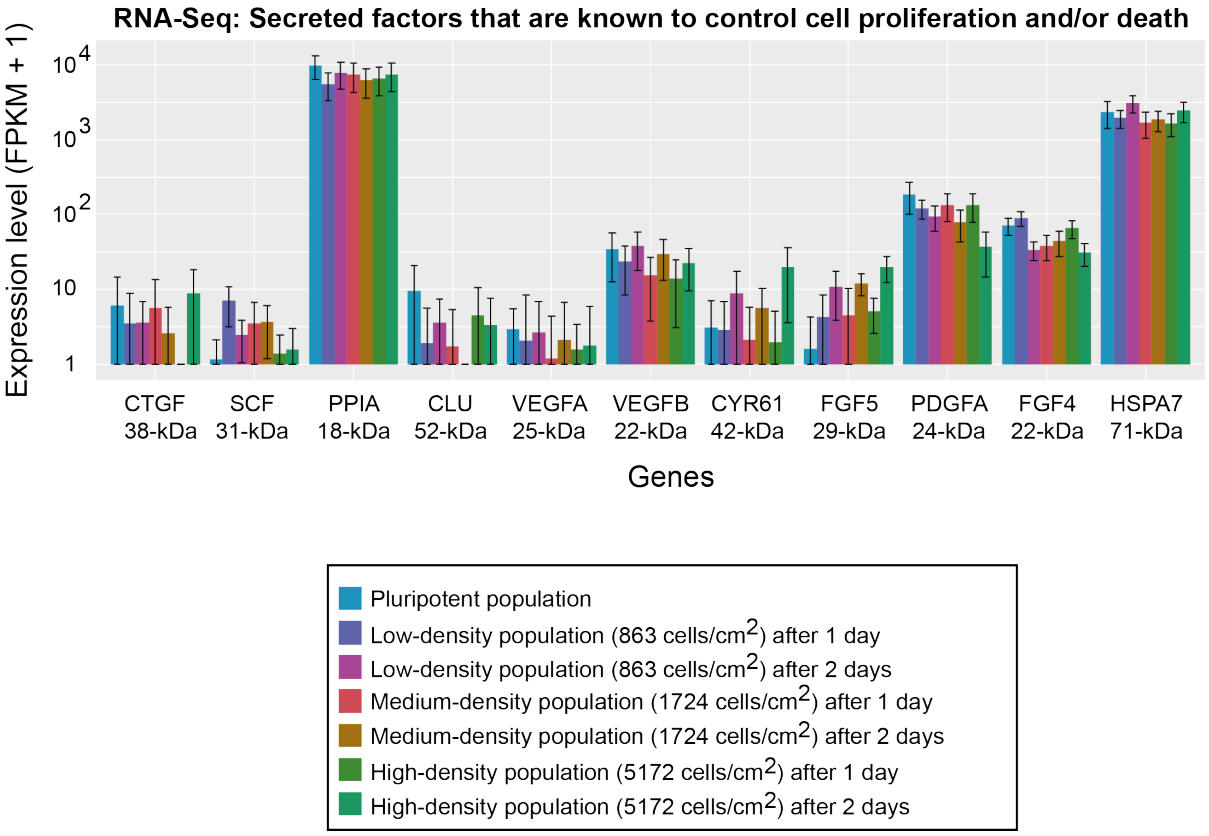
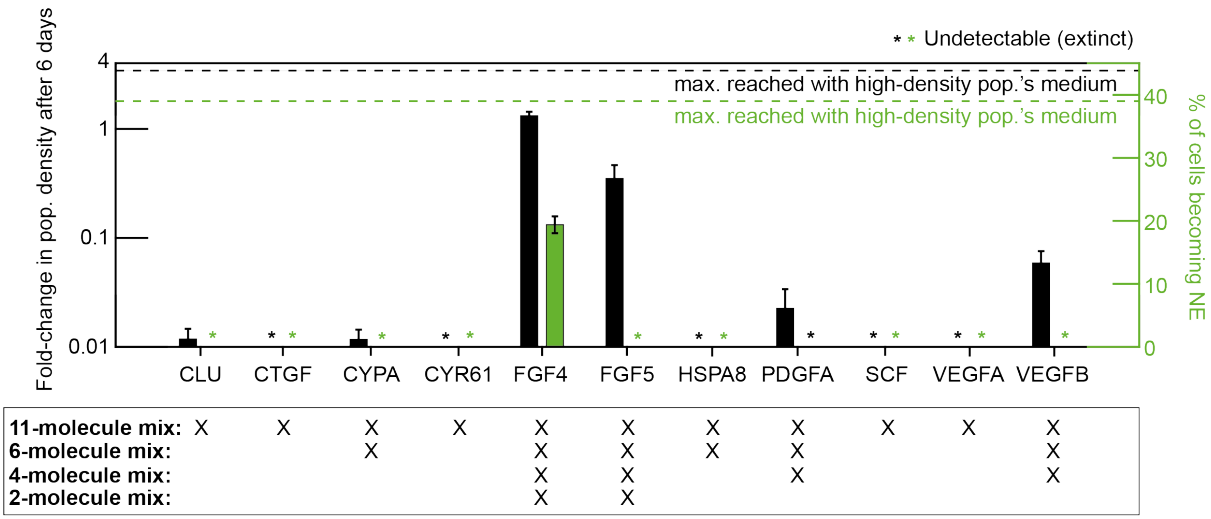


Figure S27 – RNA-Seq: Expression levels of secreted factors that are known to control cell proliferation. To identify secreted factors other than FGF4 that might contribute to determining a population survival, we performed RNA-Seq to detect expression of any secreted factors that are known to control cell proliferation and/or death. Data shown for 46C cells induced to differentiate towards the neuroectoderm (NE) lineage with Retinoic Acid (RA). We performed RNA-Seq on four populations: (1) pluripotent population prior to differentiation; (2) low-density (863 cells/cm^2) population; (3) high-density (5172 cells/cm^2) population; and (4) medium-density (1724 cells/cm^2) population that is near the threshold density. For the three differentiating populations, we collected their cells on the first and second day after triggering differentiation. Expression levels (FPKMs) of secreted factors that are known to control proliferation and/or apoptosis in ES cells and that fall within the range of molecular weights that the membrane-filter experiments identified (50 – 300 kDa with +/-50% error) (Fig. 4A-B). Shown are the following genes: *Ctgd*, *Scf*, *Ppia*, *Clu*, *Vegfa*, *Vegfb*, *Cyr61*, *Fgf5*, *Pdgfa*, *Fgf4* and *Hspa8*. Below each gene name is the molecule's weight (kDa) according to two online resources: Uniprot and ExPASy. *n* = 3 for all plots; Error bars are s.e.m.

A



B

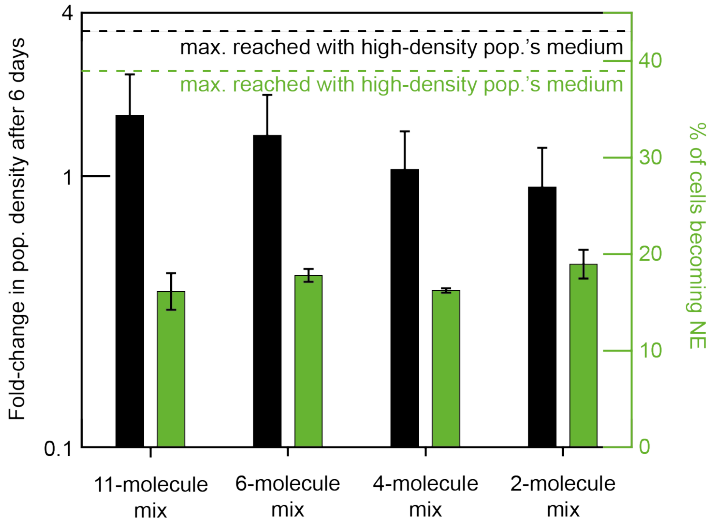


Figure S28 – Out of all the extracellular factors that we added one-by-one into differentiation medium, only FGF4 rescues low-density populations from extinction. The RNA-Seq (Fig. S27) revealed that 11 secreted factors that are known to control cell proliferation/and or death were highly expressed in differentiating, high-density populations. We thus reasoned that one or combinations of these factors may be the secreted molecule(s) that determine the survival-versus-extinction fate of a population. **(A)** We tested these molecules by adding them one-by-one into the medium of a low-density population (863 cells/cm^2) that would ordinarily become extinct. We used 46C cells that were induced to

differentiate towards the neuroectoderm lineage (NE) with Retinoic Acid (RA). We added the following molecules individually, each at a saturating concentration: version of recombinant mouse FGF4 used in Fig. S21C (200 ng/mL), recombinant human FGF5 (200 ng/mL), recombinant mouse PDGFA (100 ng/mL), recombinant mouse VEGFB 186 (100 ng/mL), recombinant mouse VEGFA (100 ng/mL), recombinant human CYR61 (500 ng/mL), recombinant human CTGF (500 ng/mL), recombinant mouse CLU (200 ng/mL), recombinant human HSPA8 (500 ng/mL), recombinant human CYPA (1000 ng/mL), and recombinant mouse SCF (2000 ng/mL). After 6 days in a medium containing one of these molecules, we measured the fold-change in density (black bars) and differentiation efficiency (green bars) of the low-density population. n = 3; error bars are s.e.m. These results show that only the recombinant mouse FGF4 causes the fold-change in population density to be higher than one. All the other factors resulted in the low-density population either approaching extinction (fold change much less than 1) or becoming extinct (indicated with an asterisk). The black dashed line marks the maximum fold-change in population density achieved when the low-density population grows in the medium of a high-density population. The green dashed line marks the maximum differentiation efficiency achieved when the low-density population grows in the medium of a high-density population. The box beneath the plot shows which signaling factors were mixed together and then given to the low-density population in (B). **(B)** Results obtained by giving combinations of the 11 factors together to the low-density population, with the ingredients of the mixture indicated in the box below (A). Giving all 11 factors together at once yielded the highest growth (~2-fold increase in population density; black bar), which was only ~2 folds lower than the black dashed line (maximum fold-change achieved by incubating the low-density population in the medium of a high-density population). But, with the 11 molecules added together at once, the differentiation efficiency (green) remained rather low at ~20% compared to the ~40% (green dashed line) that we get from incubating the low-density population in the medium of a high-density population. As we progressively reduced the number of signaling factors in the mixture from 11 to 2, we observed only a modest decrease in population growth, down to about ~1 fold. Importantly, FGF4 was included in all these mixtures.

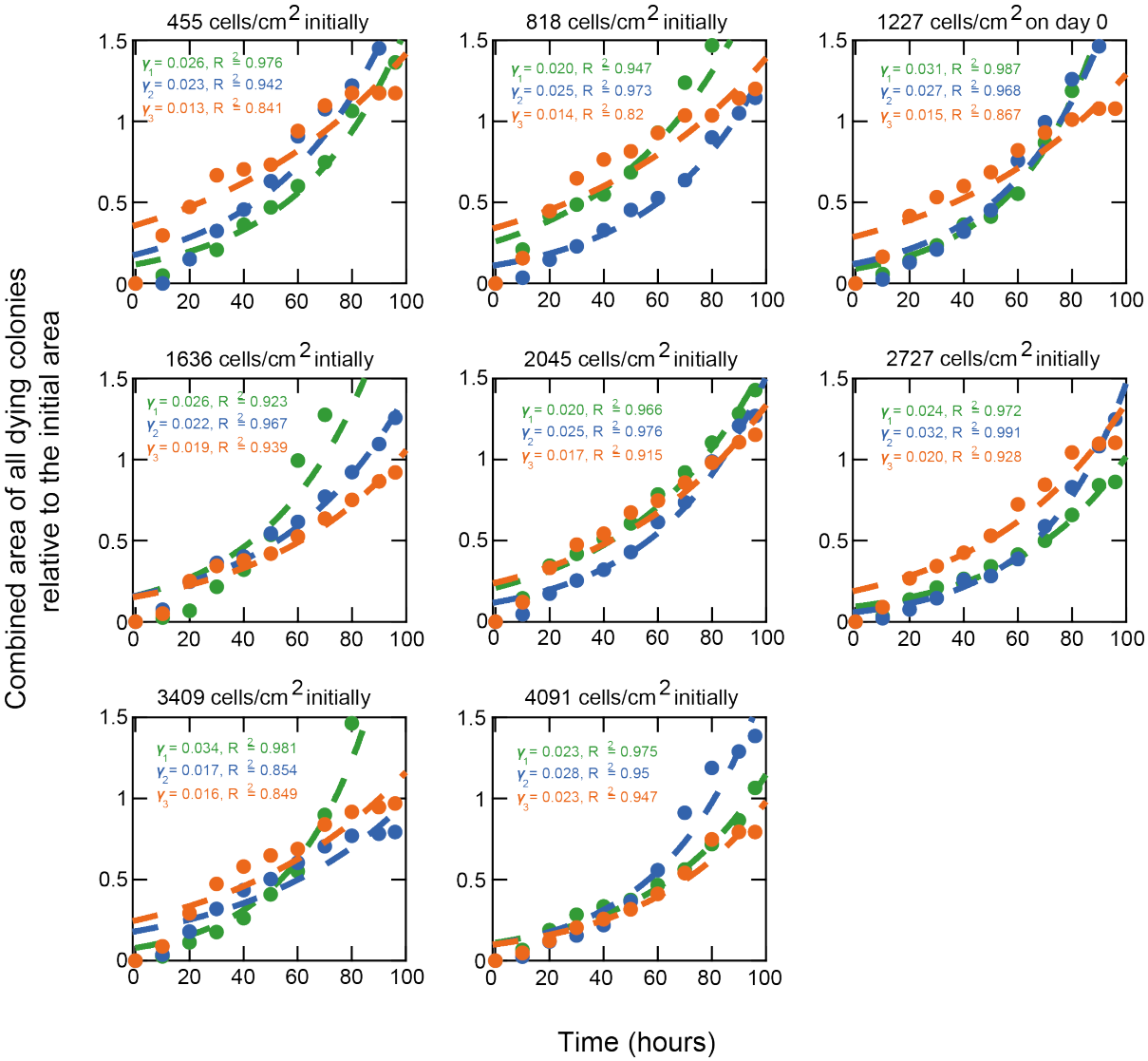


Figure S29 – Determining the rate of cell death with time-lapse microscopy. We sought to verify that the rate of cell death is independent of the initial population density. This is an ingredient of our mathematical model. With time-lapse microscopy, we measured the growth of individual microcolonies over four days for a wide range of initial population densities. We used E14 cells undergoing unguided differentiation. Each box shows a population of a different starting density (indicated above each box). For each population density, we tracked microcolonies in 17 fields of view. Each field of view has a dimension of 1.40 mm x 0.99 mm. We examined all densities in triplicates (each color represents an independent replicate). Dying colonies typically lifted off the plate (and thus disappeared from the field of view) or started to display clearly visible apoptotic bodies. From these movies, we inferred the death rate of cells by taking the cumulative sum of the last recorded areas of each colony just before it died, as a way to measure how much colony area is "lost" (i.e., deaths) in time. We then divided this combined area of all dead cells by the combined initial area. We determined the death rate by fitting a single exponential function - with the death rate γ (in 1/hours) - to the

data points (dashed curves represent the fits to data points of corresponding color; each color is a single replicate; n = 3 for each initial density) (see Fig. S30A for a demonstration of this procedure). The values of γ and corresponding R^2 are indicated in each plot. Error bars are s.e.m. This figure complements Fig. S30.

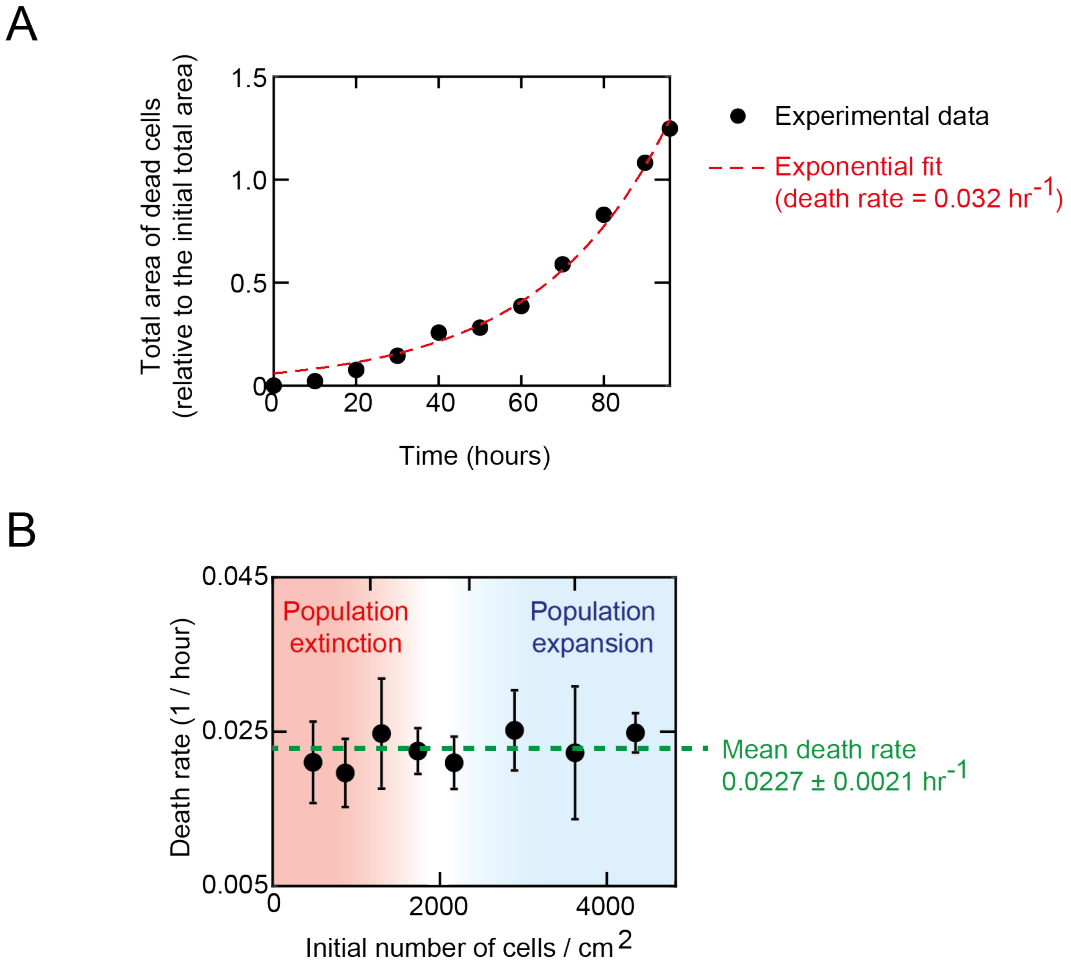


Figure S30 – Rate of cell death is independent of initial population density. (A) Example that shows how we inferred the death rate from measuring the area of dead cells per unit time for each initial population density. Detailed explanation is in the caption of Fig. S29. **(B)** Death rate extracted in Fig. S29 for many different initial population densities. This plot shows that the death rate of differentiating cells is constant (0.0227 ± 0.0021 hr^{-1}). It is independent of initial population density. n = 3; Error bars are s.e.m.

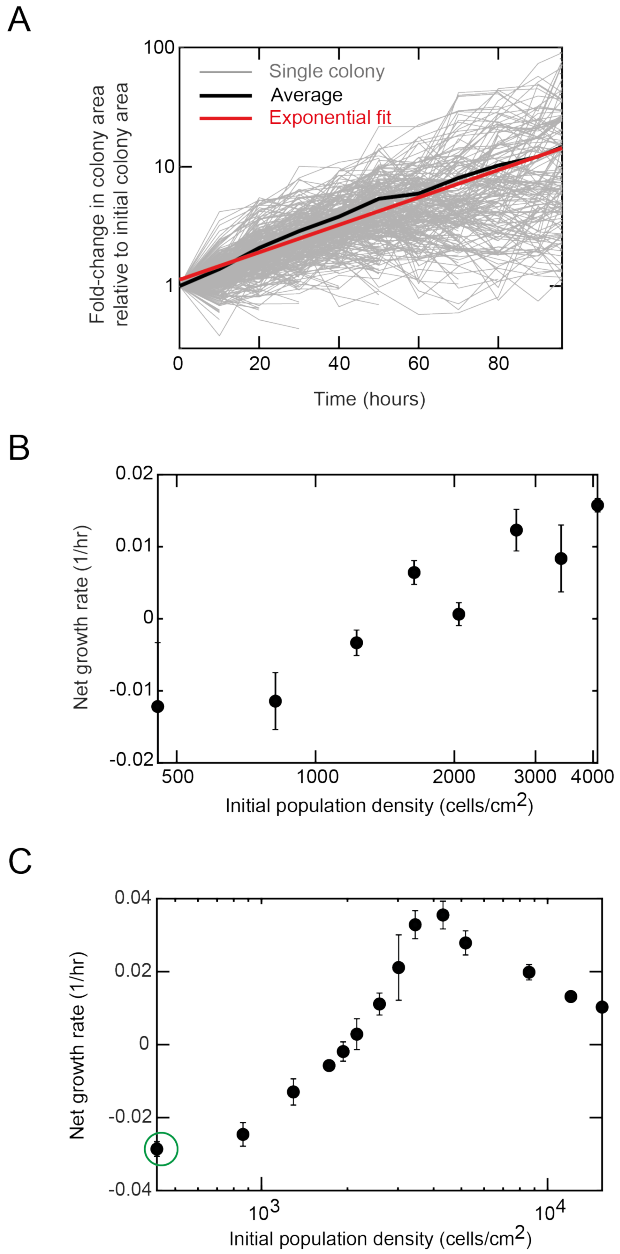


Figure S31 – Density-dependent net growth rate (= growth rate - death rate) measured in two ways. Net growth rate is negative if the death rate is higher than the growth rate (i.e., for a population that becomes extinct). It is positive if the growth rate is larger than the death rate (i.e., for a population that grows). We measured the cells' net growth rates in two different ways. In one, we used time-lapse microscopy data to measure colony growths. In the other, we manually counted cells -and thus the population density over time as in Fig. 1C. **(A)** Example showing the microscope-based method of determining the net growth rate for a given initial population density. As shown in Fig. S8 and described in the caption there, we measured the area of each colony over time for four days, imaging 17 fields of view for each starting population density. Shown here is an example for E14 cells. Each grey curve shows the area of a single microcolony over time. Each colony's area is normalized to its initial area (thus all

curves here start at a value of one on the vertical axis). Black curve is the average of all the grey curves. Red line is an exponential curve (line in this semi-log plot) that we fitted to the black curve (i.e., fitted to the population average). The slope of the red line is the net growth rate for this population. **(B)** Using the microscope-based method outlined in (A) for every initial population density, we obtained the growth rate (black dots) as a function of the initial population density. n = 3; Error bars are s.e.m. **(C)** Net growth rates determined by manual counting of cells (i.e., data in Fig. S3) rather than from the microscopy data. n = 3; Error bars are s.e.m.

To determine the (not net) growth rate for each initial population density, we added the constant death rate (determined in Fig. S30) to the net growth rate (determined in (B-C)). We found that the maximum possible growth rate is 0.05219 hr^{-1}, which we used as one of the parameter values in our mathematical model. From (C), we found that the lowest possible net growth rate is -0.026 hr^{-1} (indicated with green arrow in (C)). This value nearly matches the death rate (0.023 ± 0.002 hr^{-1}; found in Fig. S30B), a consequence that we explore in our mathematical model (see Supplementary text).

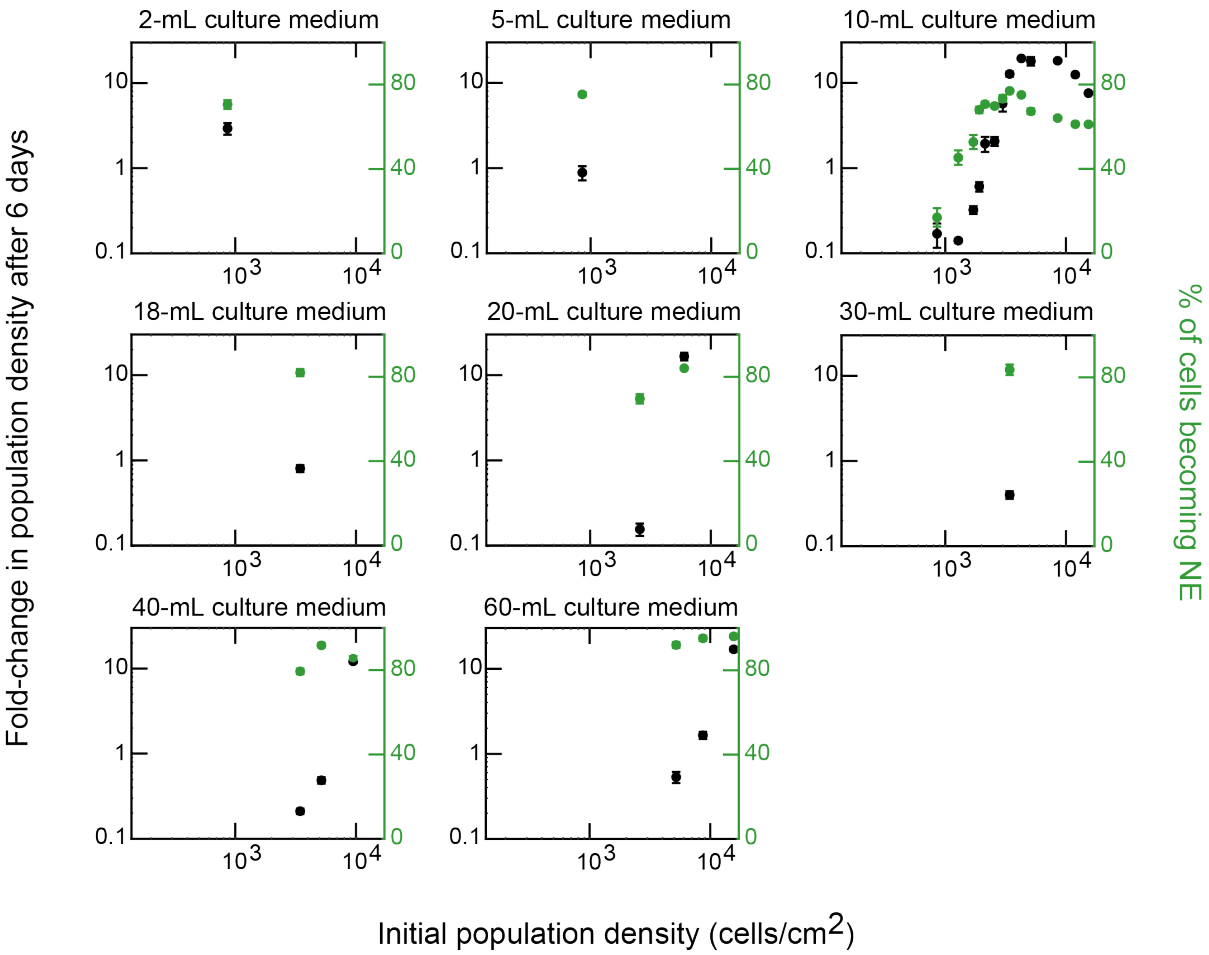


Figure S32 – Volume of liquid medium at the start of differentiation determines survival-vs-extinction fate of differentiating populations. To experimentally test the model's prediction (i.e., to verify that the model-produced phase diagram in Fig. 5F is correct as is the model's prediction in Fig. 5E), we incubated 46C populations of different starting densities in different volumes of liquid medium (and thus in different heights of liquid medium). We induced the cells to differentiate toward the neuroectoderm (NE) lineage with Retinoic Acid (RA). Each box shows a different volume of liquid medium. Horizontal axis of each box is the initial population density. Black points are the fold changes in the population density (relative to the initial density) after six days. Green points are the percentages of cells that entered the NE lineage (Sox1-GFP positive cells measured with flow cytometer). n = 3; Error bars are s.e.m.; n = 3. As seen in Fig. 5F, these results match the phase diagram produced by the model (i.e., data points of the right type fall on the right side (above or below) the model-produced phase boundary in Fig. 5F).

Table S1 - List of primers used for RT-qPCR.

Amotl2_FWD	CAGAGGGACAATGAGCGATTGC
Amotl2_REV	TCACGCTTGGAAGAGGTCCTCA
Bcl2_FWD	TGTGGATGACTGAGTACCTGAAC
Bcl2_REV	GAGAAATCAAACAGAGGTCGCATG
Cyr61_FWD	GGATGAATGGTGCCTTGC
Cyr61_REV	GTCCACATCAGCCCCTTG
Bax_FWD	GGGGCCTTTTTGCTACAGGG
Bax_Rev	AAAGATGGTCACTGTCTGCC
Bbc3_FWD	GTGTGGAGGAGGAGGAGTG
Bbc3_REV	TCGATGCTGCTCTTCTTGTC
Mdm2_FWD	GAAGGAGGAAACGCAGGACA
Mdm2_REV	CCTGGCAGATCACACATGGT
Gapdh_FWD	TGACCTCAACTACATGGTCTACA
Gapdh_REV	CTTCCCATTCTCGGCCTTG
FGF4_Fwd	AAGGCACCTGCCCTGTTCTG
FGF4_Rev	GGGAGCTAGCTGGCTGAAGAAA
FGF5_Fwd	AACTCCATGCAAGTGCCAAAT
FGF5_Rev	CGGACGCATAGGTATTATAGCTG
FGFR1_Fwd	TGGACCGCATTGTGGCCTTGAC
FGFR1_Rev	TCAGCGCCGTTTGAGTCCACTG
FGFR2_Fwd	TGCACGCAGGATGGACCTCTCT
FGFR2_Rev	TGCTCCTCGGGGACACGGTTAA
FGFR3_Fwd	AAGCCAGCCAGCTGCACACA
FGFR3_Rev	TCAAACGGCACGGAGAGGTCCA
FGFR4_Fwd	TAATGAGGGAGTGCTGGCACGC
FGFR4_Rev	AGGGTCGTGGCTGAAAACCGAG

Table S2 - Ingredients for the differentiation medium (N2B27) and their molecular weights.

We obtained this information from ATCC, Ying et al. (*27*), Mittal and Voldman (*47*), and Brewer et al., *J Neurosci Res* (1993). Most components of differentiation medium (N2B27) are smaller than the smallest filter size that we used, and if larger then not vital for growth of ES cells (see Mittal and Voldman (*47*)).

Abbreviations: MW = molecular weight, Da = Daltons, N/A = not applicable (i.e., not included in mixture), CONF = confidential, propriety information (Invitrogen).

Compound Name	MW (Da)	Conc. in DMEM/F-12 (µM)	Conc. in Neurobasal (µM)	Conc. In N2 supplement (µM)	Conc. In B-27 minus vitamin-A supplement (µM)
Inorganic salts					
CaCl2 (anhydrous)	111	1000	1800	N/A	N/A
CuSO4 (anhydrous)	160	0.01	N/A	N/A	N/A
Fe(NO3)3·9H2O	404	0.13	0.25	N/A	N/A
FeSO4·7H2O	278	1.5	N/A	N/A	N/A
MgSO4 (anhydrous)	120	700	812	N/A	N/A
KCl	75	4000	5333	N/A	N/A
NaHCO3	84	14,300	26,000	N/A	N/A
NaCl	58	120,000	51,300	N/A	N/A
Na2HPO4 (anhydrous)	142	500	N/A	N/A	N/A
NaH2PO4 ·H2O	120	500	1,000	N/A	N/A
ZnSO4·7H2O	288	1.5	N/A	N/A	N/A
Amino Acids					
L-Alanine	89	50	22.5	N/A	N/A
L-Arginine·HCl	210	700	400	N/A	N/A
L-Asparagine·H2O	150	50	5	N/A	N/A
L-Aspartic Acid	133	50	N/A	N/A	N/A
L-Cystine·HCl·H2O	294	60	N/A	N/A	N/A
L-Cystine·2HCl	312	100	10	N/A	N/A
L-Glutamic Acid	147	50	N/A	N/A	N/A
L-Glutamine	146	2,500	500	N/A	N/A
Glycine	75	250	400	N/A	N/A
L-Histidine·HCl·H2O	209	150	200	N/A	N/A
L-Isoleucine	131	415	800	N/A	N/A
L-Leucine	131	450	800	N/A	N/A
L-Lysine·HCl	146	625	1,000	N/A	N/A
L-Methionine	149	115	200	N/A	N/A
L-Phenylalanine	165	215	400	N/A	N/A
L-Proline	115	150	67	N/A	N/A
L-Serine	105	250	400	N/A	N/A
L-Threonine	119	450	800	N/A	N/A

L-Tryptophan	204	45	80	N/A	N/A
Vitamins					
D-Biotin (B7)	244	0.01	CONF	N/A	N/A
Choline Chloride	140	64	28	N/A	N/A
Folic Acid (B9)	441	6	9	N/A	N/A
myo-Inositol (B8)	180	70	40	N/A	N/A
Niacinamide (B3)	122	16.5	33	N/A	N/A
D-Pantothenic Acid (B5)	219	10	8	N/A	N/A
Pyridoxine·HCl (B6)	202	10	20	N/A	N/A
Riboflavin (B2)	376	0.6	1	N/A	N/A
Thiamine·HCl (B1)	373.5	6	10	N/A	N/A
Vitamin B-12	1,355	0.5	0.25	N/A	N/A
Proteins					
Insulin	5,800	N/A	N/A	4.3	CONF
Serotransferrin	77,000	N/A	N/A	1.33	CONF
BSA	66,000	N/A	N/A	0.75	CONF
SOD-1 (Dimer)	32,000	N/A	N/A	N/A	CONF
Catalase (Tetramer)	240,000	N/A	N/A	N/A	CONF
Other					
D-Glucose	180	17,500	25,000	N/A	N/A
HEPES	238	15,000	10,000	N/A	N/A
Hypoxanthine	136	17.5	N/A	N/A	N/A
Linoleic Acid	280	0.16	N/A	N/A	N/A
Phenol Red, Sodium Salt	354	23	23	N/A	N/A
Putrescine·2HCl	161	0.5	N/A	N/A	N/A
Pyruvic Acid·Na	111	500	225	N/A	N/A
DL-Thioctic Acid	206	0.5	N/A	N/A	N/A
Thymidine	242	1.5	N/A	N/A	N/A
Progesterone	314	N/A	N/A	0.02	CONF
Sodium Selenite	173	N/A	N/A	0.02	CONF
Putrescine	88	N/A	N/A	180	CONF
Lecithin	770	N/A	N/A	N/A	CONF
Linolenic Acid	278	N/A	N/A	N/A	CONF
Phosphatidylcholine	776	N/A	N/A	N/A	CONF
Retinol	286	N/A	N/A	N/A	CONF
Retinyl Acetate	328	N/A	N/A	N/A	CONF
Sodium Selenite	173	N/A	N/A	N/A	CONF
T3 (Triodo-I-Thyronine)	650	N/A	N/A	N/A	CONF
DL-α-Trocopherol (Vitamin E)	431	N/A	N/A	N/A	CONF
DL-α-Trocopherol Acetate	473	N/A	N/A	N/A	CONF
L-Tyrosine·2Na·2H2O	263	200	400	N/A	N/A
L-Valine	117	450	800	N/A	N/A

Supplementary Text

1. Calculations show that diffusion can occur on a 1-mm scale for secreted molecules

1.1. Using dimensional analysis to derive conditions required for millimeter-scale diffusion

Our experiments show that the cells secrete molecules that diffuse on the order of 1 mm. In this section, we use calculations to show that physics of diffusion also allows for biomolecules to diffuse on the order of 1 mm.

At first, a millimeter-scale diffusion may seem counterintuitive. This is because one often does not think about two cells, 1 mm apart, signaling to each other with diffusible molecules. Aside from the traveling axion potentials among neurons, one often does not think of millimeter scale communication among mammalian cells that are relatively immobile like ES cells. But physics allows for this for the proteins with the sizes that we experimentally found. Consider the general form of reaction-diffusion equation for the concentration *c* of a secreted molecule at position $\vec{r}$ and time *t*:

$$\frac{\partial c}{\partial t} = \mathrm{D}\nabla^2 c(\vec{r}, t) - \gamma c(\vec{r}, t)$$

where *D* and γ are the molecule's diffusion constant and degradation rate respectively. This equation lacks the molecule's source (i.e., secreting cell) for simplicity and because we are interested in how far the molecule diffuses, which is independent of the secretion rate. A molecule's diffusion length *L* is the characteristic (typical) distance that it travels before degrading. We can deduce *L* from dimensional analysis. In the next section (section 1.2), we give the full, time-dependent solution to above equation which yields the same conclusion as the one we give in this section. The only parameters in the above equation are *D* and γ. *D* has the dimension, length^2/time, and γ has the dimension, 1/time. Hence it follows that $L = \sqrt{D/\gamma}$, which is a well-known result. We can arrive at the same result through a more complicated way - by solving the reaction-diffusion equation above with a constantly secreting cell. But that is an overkill. We also have

$$\gamma = \frac{\log(2)}{\tau}$$

where τ is the molecule's half-life. Next, as is common for biomolecules, we can estimate *D* by using the Stokes-Einstein relation for diffusing, spherical particles:

$$D = \frac{kT}{6\pi\eta r}$$

where *k* is the Boltzmann constant, *T* is temperature, η is the dynamic viscosity of the medium in which the molecule is diffusing, and *r* is the radius of the spherical particle that is diffusing. The ES cells grew at 37 ^{O}C and in an aqueous medium. Thus, for an order of magnitude estimate, we can use the dynamic viscosity of water at 40 ^{O}C: $\eta = 0.653 \times 10^{-3} N \cdot s/m^2$ [http://www.engineersedge.com]. This value nearly stays the same at 30 ^{O}C and is thus insensitive to temperature for our purpose. Then, at 37 ^{O}C, we have

$$D = \frac{3.5 \times 10^{-19}\ m^3/s}{r}$$

Before estimating the *r*, we can estimate the maximum diffusion constant D_{max} for our molecules of interest. Our experiments with molecular filters showed that the secreted protein(s) must be in the neighborhood of ~100 kDa (specifically, between 50 to 300 kDa (Fig. 4A-B); and more precisely, between 25 kDa and 450 kDa if we take the conservative estimate of $\pm$50% error in the filter-pore sizes). Each amino acid has a mass of ~110 Da and end-to-end length of ~0.35 nm. Thus, our proteins of interest (of 100 kDa) would consist of ~1000 amino acids which, when joined stretched end-to-end, have a length of 350 nm (clearly a gross overestimate of the *r* for the secreted proteins). With this overestimate, we have $D_{max} = 1\ \mu m^2/s$, which is the conservative, lower bound given for the diffusion constant of large biomolecular machines such as ribosomes. In fact, the conventional values assigned to the diffusion constant of typical proteins inside a cytoplasm fall in the range of 5 - 50 $\mu m^2/s$ (from BioNumbers). Since the cytoplasm is a highly crowded environment that limits diffusion, we can expect a higher value for the protein of our interest that diffuses in the liquid medium.

We now estimate the *D*. The typical value of *r* for proteins is ~5 nm (from BioNumbers). In contrast, *r* ~ 30 nm for eukaryotic ribosomes (from BioNumbers), which are macromolecular complexes that are certainly heavier and larger than our proteins of interest as our experiments revealed. As a conservative estimate (i.e., to estimate a reasonable lower bound on the diffusion length), suppose that *r* ~ 20 nm for our proteins of interest, four times larger than the value of *r* for typical proteins (~5 nm). Then according to above equation, we have

$$D \approx 20\ \mu m^2/s$$

which is still reasonable in the ES cell's relatively non-viscous liquid medium (recall that 5 - 50 $\mu m^2/s$ for a crowded cytoplasm). Note that

$$\tau = \frac{L^2 \log(2)}{D}$$

Hence, for a diffusion length *L* of 1 mm, we need τ ~ 12 hours. Note that this is a conservative, overestimate of the secreted molecule's half-life since we underestimated the diffusion constant by assuming that the molecule is larger than half the radius of a ribosome (i.e., assumed *r* ~ 20 nm). A more realistic estimate such as, for example, *r* ~ 10 nm, would yield τ ~ 6 hours. A more typical value that one assigns to *r* for proteins is *r* ~ 5 nm, which would yield τ ~ 3 hours. In summary, both the conservative and more realistic estimates yield protein half-lives, which are required for a near millimeter-scale diffusion, that are well within the typically cited values of protein half-lives. Crucially, we measured the half-lives of the secreted molecules, including FGF4, in the supernatants of our cell cultures (Figs. S24-S25) and found that they are longer than the calculated half-lives here, meaning that the secreted molecules can, in fact, reach further than the 1-mm diffusion length we used in our calculation here.

1.2. Using solution to reaction-diffusion equation to confirm the conclusions that we derived with dimensional analysis

In this section, we use the steady-state solution to the three-dimensional reaction-diffusion equation to confirm our conclusions in the previous section, which were based on dimensional analysis. Assuming the ES cell to be a spherical cell of radius *R* that secretes a molecule at a constant rate η, equally in all directions, the three-dimensional reaction-diffusion equation is

$$\frac{\partial c}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(Dr^2 \frac{\partial c}{\partial r}\right) - \gamma c + \frac{\eta}{4\pi R^2}\delta(r - R)$$

where *c* is the concentration, *r* is radial distance from the center of the ES cell, and δ is the Dirac delta function. As one can verify by substituting into the reaction-diffusion equation, its steady-state solution is

$$c(r) = \frac{c_R R}{r} exp\left(-\frac{r - R}{\lambda}\right)$$

where $\lambda = \sqrt{D/\gamma}$ and

$$c_R = \frac{\eta}{4\pi R^2}\frac{\gamma}{\lambda}\frac{1}{1 + \frac{\lambda}{R}}$$

Note that c_R is the steady-state concentration on the surface of the cell. The exponential term in the steady-state solution tells us that the characteristic length (i.e., diffusion length) is exactly $\sqrt{D/\gamma}$, which exactly matches the diffusion length that we derived from dimensional

analysis. Thus, the conclusions based on the dimensional analysis holds true. Note that this was not guaranteed since there could have been a small, numerical factor that scaled the λ down by several orders of magnitude.

2. Description of the mathematical model

In this section we describe a stochastic model that recapitulates all the main features of our experimental data. Our model is a slightly modified version of a model that we previously built to describe cooperative growth of yeast cells at high temperatures (Laman Trip and Youk, *Nat Microbiol* 2020 (*51*)). In fact, the dynamics of differentiating ES cell populations (Fig. 1C) are strikingly similar to the population dynamics of yeast cells at high temperatures (see Fig. 2 in (*51*)) which also exhibits a survival-versus-extinction fate that depends on the initial population density (our yeast work examined how yeast populations grew as a function of temperature and the initial population density). The only modification from the yeast work is that we now consider a molecule that degrades whereas the model for yeast considered an infinitely stable molecule (glutathione). In our model, differentiating ES cells constantly secrete a molecule. This molecule can be any molecule that promotes cell proliferation. For concreteness and based on our experiment, we will let the molecule be FGF4 but it can be any other molecule (note that our filter-based experiments (Figs. 4A-B) established that a molecule that is heavier than FGF4 is also involved in determining the survival-versus-extinction fate). In our model, we assume for simplicity that cells secrete FGF4 at a constant rate. FGF4 extracellularly accumulates. The probability of a cell replicating nonlinearly increases as the extracellular concentration of FGF4 increases (Fig. 5B - blue curve). For simplicity, we assume that the probability of a cell dying is constant, as we experimentally verified (Fig. 5B).

Let N_t be a stochastic variable that represents the number of alive cells at time *t*. We take time in discrete integer steps. At each timestep, cells die with a constant probability P_γ and replicate with probability $P_\mu(t)$ which depends on the extracellular concentration of FGF4 (denoted M_t) at time *t.* The probability of replication $P_\mu(t)$, as a function of M_t, follows the Hill equation, with the Hill coefficient set to one for simplicity, and the maximum value μ. It is:

$$P_\mu(t) = \mu \frac{M_t}{K + M_t}$$

where *K* is a constant.

FGF4 degrades with rate *d*. Our experiments establish a long half-life of FGF4 - it is at least 80 days (Fig. S24) - and a long, half-life for the mixture of all secreted molecules that

determines the survival-versus-extinction fate (at least two days (Fig. S25)). At each time step t, the number of cells that replicate is sampled from a binomial distribution with $P_\mu(t)$ being the probability of one cell replicating. The number of cells that die at time t is sampled from a binomial distribution with P_γ being the probability of one cell dying. We assume that the molecule is well-mixed since our experiments established that the population effectively acts as a single entity that either dies or lives, due to the molecules that diffuse over several millimeters. Combining all the elements above, our model is completely described by the following set of stochastic equations:

$$N_{t+1} = N_t + Binom\left(N_t, P_\mu(t)\right) - Binom\left(N_t, P_\gamma\right)$$

$$P_\mu(t) = \mu \frac{M_t}{K + M_t}$$

$$P_\gamma = const$$

$$M_{t+1} = \frac{N_t}{V} + d\, M_t$$

where V is the volume of liquid medium, and N_0 is the initial number of cells. Since cells start without any extracellular FGF4 (since the differentiation medium initially has no FGF4 or any molecules that determine the survival-versus-extinction fate), we must have $P_\mu(0)$=0, which indeed follows from above equation. Note that we measure the FGF4 concentration (M_t and M_{t+1}) in units of secretion rate - the additional concentration of FGF4 generated in one time step - as the secretion rate per cell is 1/V in the above equation.

The above set of equations has five parameters, four of which are directly fixed by our experimental data. Hence, we cannot freely tune the four parameters - μ, P_γ, V, and d - and hence our model is highly constrained. We measured the maximum possible growth rate μ in two ways: (1) time-lapse microscopy in which we tracked the area of each colony over time for four days (Fig. S8 and Fig. S31); and (2) by counting the number of cells over time in a population (Fig. 1C, and Fig. S1 and Fig. S3). Both methods yielded a similar value: $\mu = 0.0519$. To directly read off the value of the death rate P_γ from the experiments, we used the fact that a population of a very low initial density never grows and thus its rate of decline in cell number is P_γ (green circle in Fig. S31C) which we found to be: $P_\gamma = 0.0256$. We determined V by directly measuring the volume of each liquid medium. Finally, we determined d by measuring the half-life of the mixture of all secreted molecules (Fig. S25) and the half-life of extracellular FGF4 alone (Fig. S24). We used the half-life of FGF4, which leads to d = 0.99 if we let each time step to represent 1 hour. With these constraints by the experiments, only one parameter, K, remains free for us to tune in the above set of equations. We chose the

value for *K* such that the threshold density - at which the population can either survive or shrink towards extinction (Fig. 1C - middle row) - matches the experimentally determined threshold density. We thus let $K = 485000$.

With the parameter values assigned as described above, our model recapitulates all the main features that we experimentally observed (Fig. 5). We ran stochastic simulations that are based on the equations of our model, with the above-mentioned parameter values, for a wide range of initial density N_0. For sufficiently low initial densities, we have $P_\mu(t) < P_\gamma$ for all times since cells are not able to accumulate enough FGF4 (cells continue to die and as this occurs, the cells' efforts to accumulate FGF4 is even further hindered). Thus, such a population goes extinct. For sufficiently high initial densities, we eventually have $P_\mu(t) > P_\gamma$ after some time. Here, M_t becomes sufficiently high - it goes above the necessary threshold concentration at which the probability of replicating equals the death rate - which results in the population expanding towards the carrying capacity. For a particular set of initial population densities, we eventually have $P_\mu(t) \approx P_\gamma$ for some stretch of time. This results in the population remaining nearly constant in cell numbers for some time. However, the population eventually gets pushed stochastically to either extinction or the carrying capacity (Fig. 5). This population-level stochasticity occurs due to the stochastic actions by just a few cells in the population - whether or not a few cells stochastically divide while $P_\mu(t) \approx P_\gamma$ and thus just slightly increases $P_\mu(t)$ above P_γ.

By running simulations for various initial densities at liquid volumes, we determined the phase-boundary that separates the survival phase and extinction phase (boundary curve in Fig. 5F). To get the boundary curve, we ran the simulation eight times for each pair of initial density and liquid volume. We then determined how many of these simulations, with the same initial condition, led to survival (and thus eventually reaching the carrying capacity) and how many of them led to a population extinction. In the extinction phase (red region in Fig. 5F), 100% of the simulations caused the population to become extinct. In the survival phase (blue region in Fig. 5F), 100% of the simulations caused the population to survive and eventually reach the carrying capacity. These determinations then allowed us to identify the boundary between the two phases - the phase boundary - which is shown as a curve in Fig. 5F. At this boundary, the probability of dying is equal to the probability of replicating, leading to equal numbers of cells replicating and dying, until stochastic fluctuations determine the population's fate (survival or extinction). We found that this boundary nearly linearly increases with the liquid volume in the regime of media volume that we could experimentally access.