Cytometry
PART A
Journal of the International Society for Advancement of Cytometry

Image-Based Quantification and Mathematical Modeling of Spatial Heterogeneity in ESC Colonies

Maria Herberg,[1]* Thomas Zerjatke,[1] Walter de Back,[2] Ingmar Glauche,[1] Ingo Roeder[1]

[1]Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[2]Center for Information Services and High Performance Computing, Technische Universität Dresden, Dresden, Germany

Received 26 June 2014; Revised 11 September 2014; Accepted 6 November 2014

Grant sponsor: DFG priority program: SPP1356 Pluripotency and Cellular Reprogramming, Grant number: RO 3500/2-1

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Maria Herberg; Institute for Medical Informatics and Biometry, Fetscherstrasse 74, D-01307 Dresden, Germany.
E-mail: maria.herberg@tu-dresden.de

Published online 00 Month 2015 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.22598

Abstract

Pluripotent embryonic stem cells (ESCs) have the potential to differentiate into cells of all three germ layers. This unique property has been extensively studied on the intracellular, transcriptional level. However, ESCs typically form clusters of cells with distinct size and shape, and establish spatial structures that are vital for the maintenance of pluripotency. Even though it is recognized that the cells' arrangement and local interactions play a role in fate decision processes, the relations between transcriptional and spatial patterns have not yet been studied. We present a systems biology approach which combines live-cell imaging, quantitative image analysis, and multiscale, mathematical modeling of ESC growth. In particular, we develop quantitative measures of the morphology and of the spatial clustering of ESCs with different expression levels and apply them to images of both *in vitro* and *in silico* cultures. Using the same measures, we are able to compare model scenarios with different assumptions on cell–cell adhesions and intercellular feedback mechanisms directly with experimental data. Applying our methodology to microscopy images of cultured ESCs, we demonstrate that the emerging colonies are highly variable regarding both morphological and spatial fluorescence patterns. Moreover, we can show that most ESC colonies contain only one cluster of cells with high self-renewing capacity. These cells are preferentially located in the interior of a colony structure. The integrated approach combining image analysis with mathematical modeling allows us to reveal potential transcription factor related cellular and intercellular mechanisms behind the emergence of observed patterns that cannot be derived from images directly.

Key terms

quantitative measures; spatial modeling; embryonic stem cells; image quantification; transcription factor heterogeneity; cell–cell interactions

Mouse embryonic stem cells (ESCs) are cell lines derived from early embryos before the formation of the tissue germ layers (1). These cells are pluripotent as they retain the potential to differentiate into all cell types of an adult individual. However, the maintenance of ESC pluripotency requires appropriate culture conditions preventing the onset of differentiation. A commonly used ESC medium includes serum factors and the cytokine leukemia inhibitory factor (LIF) (2). Under such conditions, ESCs have a high cellular turnover and tend to self-organize as tightly packed aggregates, which presumably allow the establishment of cell–cell interactions similarly to the *in vivo* situation during embryonic development. However, LIF/serum conditions do not result in homogeneous populations of ESCs but promote a population-intrinsic, reproducible heterogeneity. In particular, it has been demonstrated that the expression levels of transcription factors (TFs) such as Nanog and Rex1 are highly variable and associated with different cell fate propensities (3–5). Flow cytometry measures of both Nanog and Rex1 reporter cell lines show a stable bimodal distribution with a fraction of Nanog/Rex1-high (about 70–80%) and a fraction of

ISAC
International Society for Advancement of Cytometry

Nanog/Rex1-low cells (about 20–30%), which is reestablished after cell sorting (3,4,6). Rex1 is a downstream target of Nanog and a sensitive marker for undifferentiated ESCs (5,7). Thus, sorted Rex1-high (RH) and Rex1-low (RL) cells exhibit distinct differentiation propensities (4,5).

The molecular mechanisms underlying this intrinsic heterogeneity are still under debate (8–11). A negative feedback loop mediated by FGF4/Erk signaling is repeatedly proposed as key mechanism to induce heterogeneous expression patterns of Nanog and Rex1 (11–13). We have previously demonstrated that a negative interaction between Nanog and FGF4/Erk together with a transcriptional background noise can lead to the establishment of two coexisting attractor states (i.e., bistability) and allow single ESCs to reversibly switch between them (11). However, to our knowledge none of the published conceptual or quantitative models explicitly accounts for the spatial arrangement of ESCs thereby neglecting intercellular interactions with neighboring cells or influences resulting from changes in the local cell density (14,15). It is well known from micropatterning approaches that the density and the clustering of mouse and human ESCs directly regulate the distribution of signaling molecules and the endogenous activation of different TFs (15,16). Moreover, it has been demonstrated that an appropriate initial seeding density is crucial for the self-renewing capacity of ESCs (17). However, a comprehensive analysis of the spatial arrangement of cultured ESCs and its relationship to a cell's intrinsic (i.e., transcriptional) state is missing.

The limited understanding of these interactions results to a large extent from the experimental challenges related with the simultaneous acquisition of both transcriptional and structural information of living cells. While common experimental strategies based on flow cytometry, qRT-PCR or RNA-Sequencing cannot preserve the spatial integrity of ESCs, live-cell imaging of cultured cell colonies offers a promising alternative. Complemented by fluorescence-based reporter systems, it becomes feasible to simultaneously monitor temporal pattern formation with respect to characteristic gene expression. However, for a rigorous interpretation of the resulting data, two steps are necessary. First, quantitative measures must be established to reliably and statistically assess spatial structures and correlations within images. Second, a reference system is required to which the occurrence of certain morphological structures or the spatial correlation of characteristic gene expressions can be compared. We take the view that mathematical modeling of cellular and intracellular ESC organization is a valuable tool as such a reference system allows exploring the consequences that functional changes at the cellular level (e.g., in adhesion or in cellular interaction between neighboring cells) may have on the population level. Comparing such *in silico* colonies to their *in vitro* counterparts on the basis of quantitative image-derived measurements is a suitable strategy to evaluate the consistency of functional models with experimental data. In this work, we specifically analyze ESC colony structures regarding morphological and transcriptional properties. In particular, we develop measures on preprocessed fluorescence images to quantitatively describe TF heterogeneity among cultured Rex1GFPd2 ESCs in LIF/serum conditions. Complementary, we apply an image-based modeling approach to reveal potential mechanisms linking biophysical properties (e.g., cell adhesion) of ESCs to intracellular processes. We use the modeling environment *Morpheus* (18) to first develop a spatial model of ESC growth that consistently reproduces distributions of morphological properties derived from the live-cell imaging analysis. This cellular model is complemented by integrating a previously developed intracellular network model, describing the dynamics of TFs involved in ESC pluripotency (11). The resulting multiscale model can be used to test the applicability of the measures to *in silico* data and to compare models of different cell adhesions and feedback regulations among single ESCs.

Materials and Methods

Cell Culture and Imaging

In Rex1GFPd2 ESCs [described in (19)], a destabilized GFP protein is expressed from the Rex1 locus. This construct ensures a comparable half-life of 2 h between the GFP and Rex1 protein, which is essential to quantitatively monitor the dynamic behavior of ESCs over short time scales. Rex1GFPd2 cells are cultured without feeders on plastic coated with 0.1% gelatin in LIF/serum conditions Dulbecco's modified Eagle's medium (DMEM; Hi-Glucose) supplemented with 10% fetal calf serum (FCS), 0.1 mM 2-mercaptoethanol, 0.1 mM MEM nonessential amino acids, and 100 units/ml LIF). Cells are seeded at a low density of 1.4×10^4 cells cm^{-2} and splitted after 2 days in culture to facilitate monolayer growth. Flow cytometry analyses were performed using a BD FACSCalibur cytometer. Data were analyzed using FlowJo software.

Live-cell images are obtained using a DeltaVision imaging system with a high-sensitive AMCCD camera. Bright field and wide field fluorescent images are taken simultaneously 48 hours after seeding at a magnification of 10×. Digital images are stored in 16 bit TIFF format and have a constant size of 512×512 pixels to ensure comparability between all images.

Image Processing

Bright field images. In a preprocessing step, these images are adjusted to a common mean intensity and corrected for uneven illumination by subtracting the image background. Colonies are segmented using *ilastik* (20), a publicly available software tool that uses a machine learning approach (Fig. 1A). To avoid background segmentation at the colony borders, segmentation areas (masks) are shrunken by applying an erosion filter. Sixty four images containing a total of 1,975 colonies have been used for the analysis. One hundred single cells have been segmented manually to determine the mean cell size of 110 pixels.

Fluorescent images. To account for the uneven illumination (e.g., due to inhomogeneities of the culture dish surface and the light source, Fig. 1B), the images were background corrected. As the colonies often cover large areas of the images, a background estimation using Gaussian filtering is not

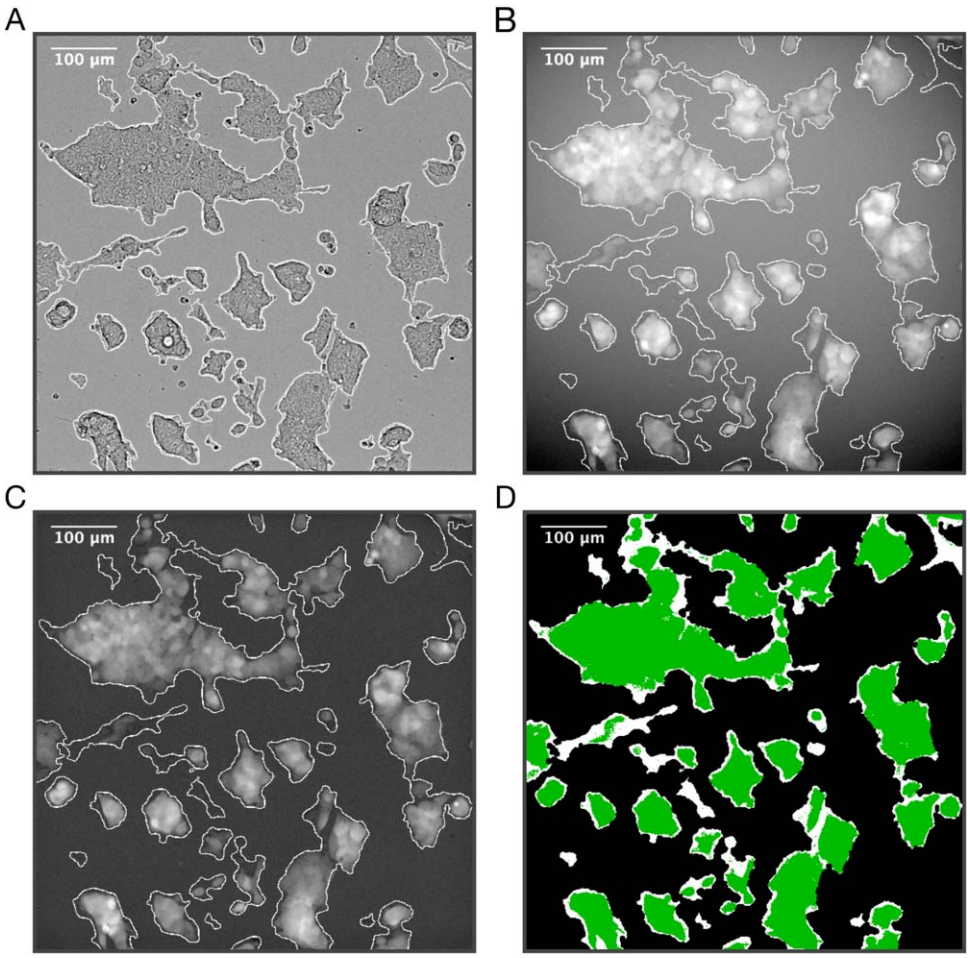

Figure 1. Image processing. (**A**) Segmentation of ESC colony structures in bright field images using the software tool *ilastik* (20). (**B**) Projecting colony masks [see (A)] to fluorescence images. (**C**) Fluorescence image after background correction. (**D**) Binarization of fluorescence image to visualize and separate clusters of RH (green) and RL expression (white). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

applicable. Also, a calculation of the position dependent gain directly from the images as described in (21) is not possible as time series would be necessary. Instead, we use the following approach: foreground pixels obtained by the segmentation of the bright field image are masked within the fluorescent image. To avoid that foreground pixels are left at the colony borders, the masks are enlarged using a dilation filter. The resulting image only contains background pixels and black holes at colony positions. These holes are iteratively filled with the mean intensity of local background pixels starting from the border of the holes and continuing to the innermost pixels. The resulting background image is smoothened by a Gaussian filter with a large kernel size and subtracted from the original fluorescent image (Fig. 1C). To visualize and quantify clusters of cells with high expression levels, fluorescence intensities can be binarized as shown in Figure 1D (technical details below).

Measures on Spatial Heterogeneity

To characterize ESC colonies and quantify spatial patterns of TF expression, we apply different measures on images of Rex1GFPd2 ESC cells describing the morphology, the distribution of fluorescence levels, and the clustering of RH cells of a colony. In the following, the respective measures are explained in detail.

1. Colony morphology
   The morphology of ESC colonies is described by:
   a. the area A, that is, the number of image pixels covered by a colony, and
   b. the circularity C, defined as $C = 4\pi A/p^2$, where A denotes the area and p the perimeter length of the colony. This measure equals 1 for a circle and decreases to 0 for increasingly irregular shapes.
2. Fluorescence intensity
   The fluorescence levels and their spatial distribution within a colony are measured by:
   a. the mean fluorescence intensity I, that is, the mean of the intensity of all pixels of a colony,
   b. the mean distance between distinct intensity values D: This measure quantifies the "spottiness" of intensity values within each colony. For 500 sampled

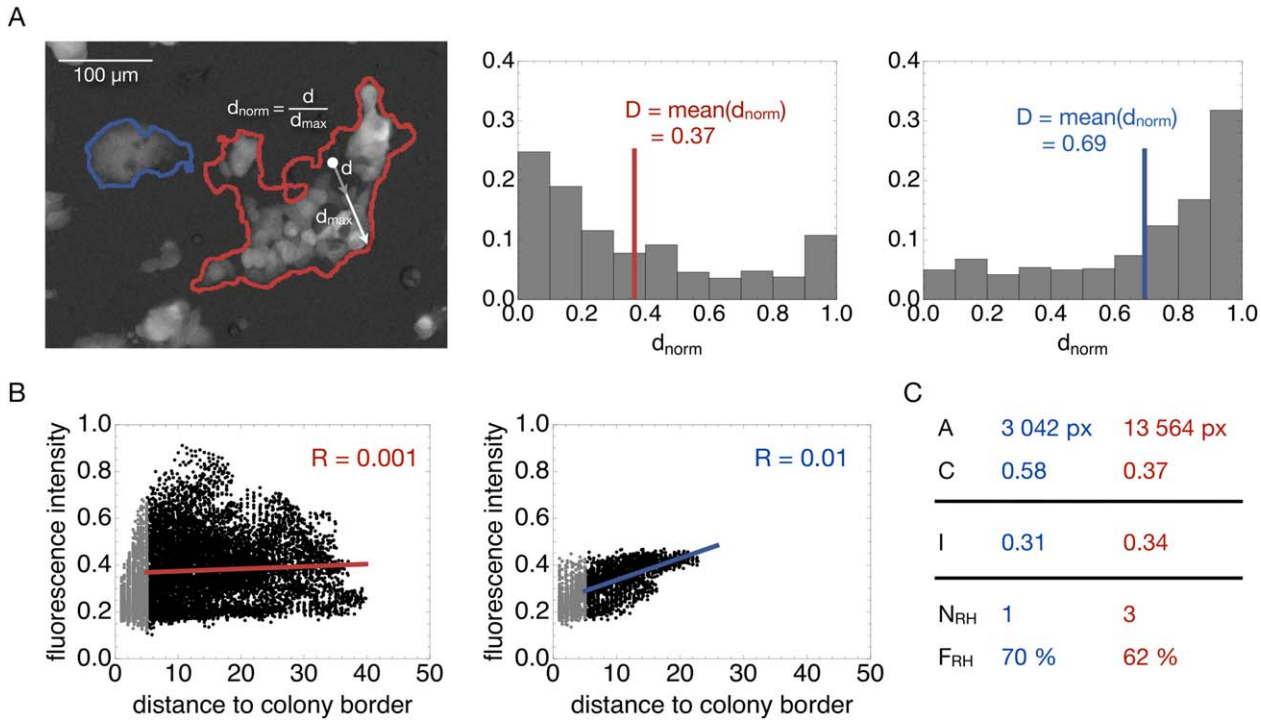

Figure 2. Measures on the spatial distribution of fluorescence intensities. (**A**) The calculation of the distance between distinct intensity values d_{norm} is illustrated for one example pixel (out of 500 randomly sampled pixels) in the red outlined colony. The distance d to the nearest pixel having a distinct intensity value (i.e., exceeding a chosen threshold) in a randomly chosen direction is evaluated and normalized by the distance to the colony border in the same direction d_{max}. The histograms show distributions of 500 such distances d_{norm} for the two colored example colonies. For every colony the mean distance D is calculated as a characteristic measure. (**B**) The distance to the colony border and the fluorescence intensity are correlated for all pixels per colony. Then, the regression coefficient R, that is, the slope of the regression line is determined. (**C**) Summary table of quantitative measures for the two outlined colonies: area A, circularity C, mean fluorescence intensity I, number of RH clusters N_{RH}, and fraction of RH clusters per colony F_{RH}. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

pixels per colony, we randomly choose a direction from a uniform distribution in the interval $[0,2\pi]$ and calculate the distance to the nearest pixel in that direction with a fluorescence intensity deviating more than a given threshold. This threshold is defined on the distribution of all colonies' fluorescence intensity values. We take the distance between 5 and 95% quantile as a robust measure for the distribution range. To capture relevant differences in Rex1 expression levels, $^1/_4$ of this range turned out to be an appropriate threshold. The obtained distance per sampled pixel (denoted by d) is normalized by the distance to the colony border in the chosen direction, called d_{max}, that is, $d_{norm} = d/d_{max}$ (Fig. 2A). Averaging over all 500 pixels leads to a single measure per colony termed D, which equals 1 for homogeneous colonies and tends to 0 for colonies with a high number of small clusters with different fluorescence intensities (examples in Fig. 2A).

c. the distance to the colony border versus fluorescence intensity: to examine the localization of fluorescence intensities within a colony, we correlate each pixel's fluorescence level to its shortest distance to the colony border. A linear regression is fitted and the regression coefficient R, that is, the slope of the regression line, is taken as a single measure per colony (Fig. 2B). Positive coefficients indicate increasing fluorescence toward the colony center, whereas zero/negative values indicate no/anticorrelations, respectively. The standardized effect size is determined as $r = \text{mean}(R)/\text{sd}(R)$, with sd denoting the standard deviation. To avoid artifacts caused by improper segmentation of the colony border, we neglect all fluorescence intensities within a distance of 5 pixels to the border when calculating the linear regression (grayed points in Fig. 2B).

3. Cell clustering

The fluorescence intensity distribution of all colony pixels is bimodal on a log scale (cf. Fig. 5C). To determine the clustering of cells with high and low Rex1 expression, we use the minimum between the two peaks to determine a threshold for binarization. We define the following measures on the binarized images to describe the spatial distribution of RH clusters (Fig. 2C):

a. the number of distinct RH clusters per colony N_{RH}, which is counted automatically for each colony, and
b. the spatial fraction of RH clusters per colony F_{RH}, that is, the area covered by fluorescent clusters divided by the total area of the colony.

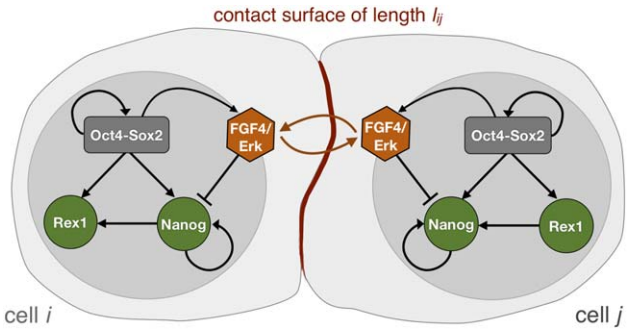

Figure 3. Scheme of the interaction model. Each individual cell contains a network model composed of the TFs Oct4-Sox2, Nanog and Rex1 and FGF4/Erk signaling as described in (11). In adjacent cells, denoted cell *i* and cell *j*, the repression of the TF Nanog is regulated by contact-mediated FGF4/Erk signaling that depends on the length of the contact surface l_{ij} [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Multiscale Modeling and Simulation

Complementary to the *in vitro* data analysis, we use the modeling environment *Morpheus* (18) to construct a multiscale simulation model of ESC organization. This modeling environment facilitates the integration of cell-based and differential equation-based models, for example, to represent feedbacks between biophysical interactions among cells and their intracellular transcriptional regulation. Simulation results can be stored as images in multichannel TIFF format, facilitating "1-to-1" integration in the image analysis workflow described above.

The multiscale model of ESC growth is composed of two coupled submodels representing cellular and intracellular dynamics. For the intracellular part, we use an established mathematical model representation previously developed by us in (11,22) and illustrated in Figure 3. In brief, the dynamics of transcriptional regulation are modeled by a set of coupled stochastic differential equations describing the temporal changes of the TFs Oct4, Sox2, Nanog, and Rex1 as well as FGF4/Erk signaling. In this model, Nanog is repressed by FG4/Erk signaling and affected by a transcriptional background noise, thus, leading to the heterogeneous expression patterns of Nanog and Rex1 in LIF/serum conditions.

The cellular part of the model is described using a cellular Potts model (CPM) that accounts for cell shape, intercellular adhesion, cell division, and cell death (23). In a CPM, cells are represented as discrete entities with spatial domains on a lattice sharing an index $\sigma = \{1, 2, \ldots, N\}$ with the special index $\sigma = 0$ representing the medium. Cell shapes are updated to minimize the effective energy function: $E = \Sigma_i J_\sigma + \Sigma_c[\lambda_A(a_\sigma - A_\sigma)^2 + \lambda_P(p_\sigma - P_\sigma)^2]$. Here, the first term sums the cell–cell adhesion energy J_σ for all cell–cell interfaces *i*, while the second term sums over cells *c* and constrains cell area *a* and cell perimeter *p* to their target values *A* and *P*. *P* is normalized to the circumference of a circle $\sqrt{4A\pi}$. Cell shape changes and motility result from local sampling of cellular protrusions and retractions of lattice nodes at the boundary. The probability of accepting a trial depends on the change in effective energy ΔE, according to the Metropolis algorithm: $P(\Delta E) = \begin{cases} 1 & \text{if } \Delta E \leq 0 \\ e^{-\Delta E} & \text{otherwise} \end{cases}$.

Cell divisions are modeled by dividing a cell domain in half using a random orientation to obtain two daughter cells that inherit all properties from their mother cell. The probability of cell division per time step depends on Rex1 expression. Differential proliferation rates of $p_{\text{high}} = 0.058$ for RH cells and $p_{\text{low}} = 0.034$ for RL cells, as well as the apoptosis rate $a = 0.012$ have been determined by model-based analysis of sorted cell populations.

Depending on the model scenario (cf. paragraph Image-based modeling of ESC growth in Results), adhesion energies J_σ between ESCs are homogeneous, that is, identical for all cells, or heterogeneous with an explicit coupling to the cell's Rex1 expression. In the latter case, the expression of homophilic adhesion molecules such as cadherins are assumed to be linearly proportional to the logarithm of Rex1 concentration $J_\sigma = J + \omega \log(\text{Rex1})$ whereby adhesion between a pair of adjacent cells is determined by the cell with lowest expression.

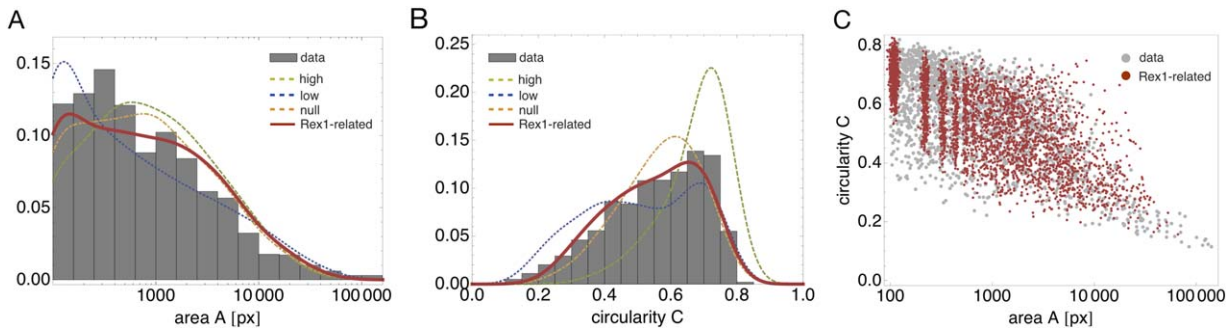

Figure 4. Morphology quantification of ESC colonies. (**A**) Area *A* for real ESC colonies (gray histogram, median 576 px, interquartile range 1,724 px) and simulated colonies with different cell–cell adhesions (green, high adhesion; blue, low adhesion; yellow, no adhesion; red, TF-related adhesion). (**B**) Circularity C for real ESC colonies (gray histogram, median 0.58, interquartile range 0.23) and simulated colonies [same color-code as in (A)]. (**C**) Correlation between A and C for real ESC colonies (gray dots) and for simulated colonies with TF-related adhesions (red dots).

Keeping the above described parameters constant, the optimal parameter set for the adhesion energies J_σ has been identified by a systematic screen within a reduced parameter space and a subsequent comparison of area and circularity distributions of simulated colonies with the experimental data using a least-square method. Nanog repression is modeled by a contact-mediated FGF4/Erk signal using the average Erk expression in adjacent cells $E_n^i = \frac{\Sigma_j E \cdot l_{ij}}{\Sigma_j l_{ij}}$ where cell j is adjacent to cell i and shares a cell–cell contact of length l_{ij} (24); Fig. 3).

A lattice with Neumann boundary conditions of size 512 × 512 pixels is used, matching the resolution of live-cell images. The cell target area A_σ is drawn from a normal distribution with mean 110 and standard deviation of 5 pixels, estimated from live-cell images. The cellular parameters $\lambda_A = 1$, $\lambda_P = 5$, and $P_\sigma = 0.8$ remain unchanged throughout simulation and across model scenarios. The initial cell number is drawn from a normal distribution with mean 100 and standard deviation of 25 cells. Cells are updated in 2,880 discrete time steps mimicking 48 h in an experimental setting.

During simulation, images of simulated ESC cultures are saved every hour (i.e., 60 time steps) with single cells color-coded according to their Rex1 expression. For a comparison with the imaging data, 75 simulations were performed for each model scenarios to obtain a similar number of experimental and simulated cell colonies. Parameter sets and full specification of the described model (i.e., model configuration files in *Morpheus* XML format) are available as Supporting Information.

Software

Image preprocessing and analysis was performed using *Mathematica 8* (25), images were segmented with *ilastik* (20). The mathematical model was implemented in *Morpheus* (18). Analyses on simulated TF concentrations were done with statistic software *R* (26).

Results

Morphological Characterization of ESC Colonies

ESCs cultured in LIF/serum are typically described as heterogeneous with respect to both their morphology and the expression of several TFs (3–5,27). However, this appraisal is either based on a visual appearance of cell/colony structures using live-cell microscopy or on flow cytometry measurements of isolated ESCs. To provide an objective and reproducible quantification of the spatial heterogeneity among cultured ESCs *in situ*, we simultaneously capture bright field and fluorescence images of a Rex1GFPd2 reporter cell line and apply different measures on their morphology and expression patterns. The image processing steps and the measures are described in detail in Materials and Methods.

For the assessment of the colonies' morphological heterogeneity, we provide histograms of colony area and circularity in Figures 4A and 4B (gray bars), showing that ESC colonies largely differ in both area and circularity. After 48 h in culture, colony areas range from 100 pixels (roughly corresponding to a single cell) to 10^5 pixels (more than 1,000 cells). The distribution is asymmetric with 50% of the colonies having a size smaller than 600 pixels (Fig. 4A). The circularity of the colonies ranges from 0.1 (highly irregular) to 0.8 (roundish) with a median of 0.58 (Fig. 4B). Moreover, the distribution of gray dots in Figure 4C indicates an anticorrelation between area and circularity with small colonies having a higher circularity compared to large colonies. Especially, midsize colonies (i.e., 10^3 px $<$ area $<$ 10^4 px) show a large heterogeneity in their circularity, potentially due to the merging of spatially extended cell colonies under LIF/serum conditions as previously demonstrated (27). We use this quantitative information on the morphology of ESC colonies to establish a spatial model describing colony formation.

Image-Based Modeling of ESC Growth

To simulate the spatio-temporal development of ESC colonies, we model proliferative and spatially extended cell agents using a cellular Potts model (CPM) (23). Although some model parameters are experimentally accessible (e.g., the proliferation and apoptosis rate) or can be determined directly from high magnification live-cell images (e.g., the size of a single cell, see M&M), other parameters such as the adhesion strength of individual cells within a colony are not directly measurable. To estimate these values, we keep experimentally accessible parameters at their measured values while simulating ESC growth with different assumptions on cell–cell adhesions. First, we assume that the adhesion strength between all ESCs is homogeneous, and we systematically vary the corresponding model parameter (i.e., the adhesion energies J_σ, cf. M&M) from high (i.e., attracting) to low (i.e., repelling). Subsequently, we compare area and circularity distributions of the simulated colonies to our experimental data (example distributions are shown in Figs. 4A and 4B). While homogenously high adhesions among cells (dashed green lines) lead to midsize and predominantly circular colonies, low adhesions (dotted blue lines) result in smaller and mainly irregular colonies. Assuming no adhesion, the proportion of circular colonies is underestimated (dashed-dotted yellow lines). In summary, none of the homogeneous scenarios sufficiently reproduce the circularity distribution obtained from live-cell images leading to the assumption that cell–cell adhesion may be (indirectly) related to the expression of Rex1. To test this hypothesis, the adhesiveness of cells is coupled to the Rex1 expression of the intracellular submodel (see M&M). More precisely, we assume that cells with a low Rex1 expression are less adhesive than RH cells. With this assumption, RH cells form compact and spherical colonies, while RL cells are more likely to dissociate and organize in loosely coupled structures. The morphology of simulated colonies in terms of area and circularity (red lines in Figs. 4A and 4B) closely mimics the morphological structures seen in real ESC colonies and captures the negative correlation between these measures (red dots in Fig. 4C). However, it remains unclear whether such a Rex1-related adhesion also has a functional implication, for example., through the differential establishment of regulatory cell–cell interactions. In the following, we compare the occurrence of spatial expression patterns between *in vitro* and *in silico* ESC

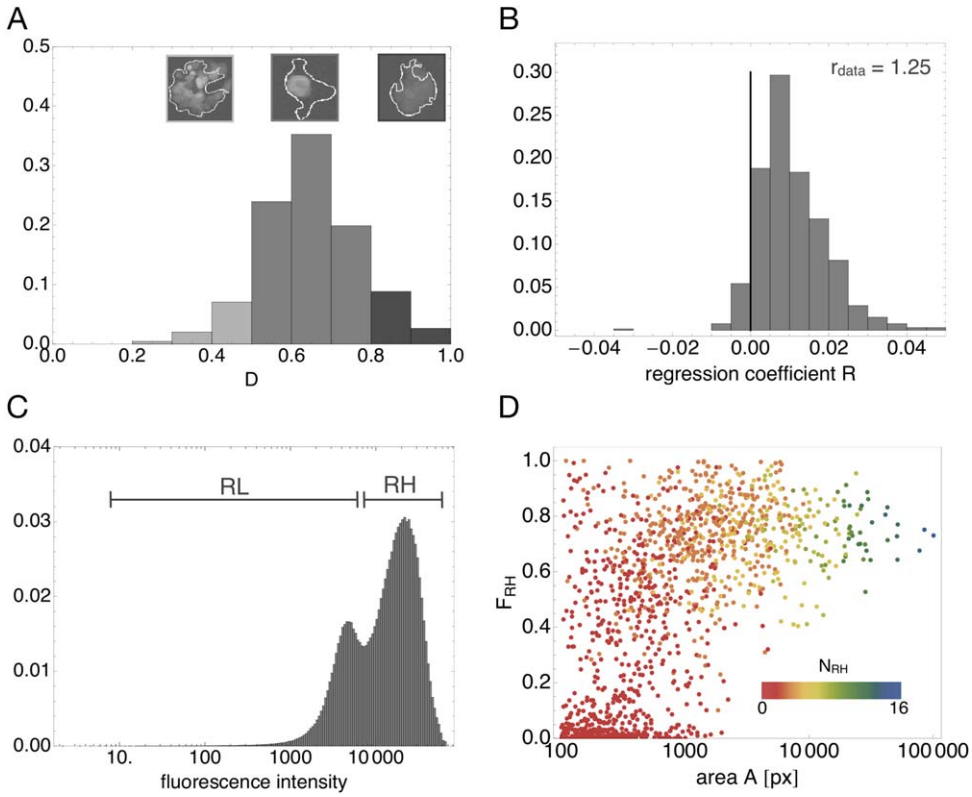

Figure 5. Measures on spatial fluorescence distributions in ESC colonies. (**A**) Distribution of the mean distance *D* and representative images of a spotted (light gray), a clustered (gray) and a homogenous (dark gray) colony (cf. Fig. 2A). (**B**) Histogram of distance to border versus intensity regression coefficients *R* for ESC colonies. *R* is close to 0 (vertical, black line) if there is no correlation between the intensity values and the distance to the border of a colony (cf. Fig. 2B). (**C**) Bimodal distribution of the fluorescence intensities in the image data. (**D**) Correlation between the fraction of highly fluorescent clusters F_{RH} and the area *A*. The colors give the number of high Rex1 clusters per colony (color code on a log scale). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

colonies to study consequences of such (hypothetical) functional interactions.

Fluorescence Pattern in ESC Colonies

Due to the activity of FGF4/Erk signaling, LIF/serum conditions promote cultures in which the expression of Rex1 is heterogeneous and obeys a bimodal distribution with a fraction of RH and a fraction of RL cells (4,11). To quantify the spatial heterogeneity of Rex1 levels in ESC colonies, we first determine the mean distances *D* between distinct expression levels (Fig. 5A, cf. M&M). Colonies with a very short mean distance of $D < 0.5$ (light gray bars, 9.5% of all colonies) are visually characterized by a number of fluorescent spots with different intensities. Most colonies appear with a mean distance *D* between 0.5 and 0.8 (gray bars, 79.0%), thus containing only a few, mostly coherent fluorescent clusters with similar expression levels. Large mean distances $D > 0.8$ (dark gray bars, 11.5%) indicate homogeneous colonies. Besides quantifying the "spottiness" of ESCs colonies, we are furthermore interested in the position of high and low expressing cells with respect to the colony border. Figure 5B shows the distribution of the regression coefficients *R* for all colonies, which characterize the relation between fluorescence intensities and distance to the colony border (cf. Fig. 2B). A slope of zero (see black, vertical line in Fig. 5B) implies that the intensity is independent of the distance, while a slope greater than zero indicates a positive correlation. The shift of the distribution to the right reveals that ESCs with a higher Rex1 expression are more likely located in the interior of a colony structure than close to its border (effect size $r = 1.25$). This effect remains even if pixels in the middle of the colonies (which are generally more compact and suspect to superposition of fluorescence signals) are excluded from the analysis.

It has been demonstrated that RH cells (separated by flow cytometry) have a higher self-renewal potential compared to RL cells (5,19). Here, we specifically aim to analyze the spatial clustering of ESCs with high fluorescence intensity. Therefore, we binarize the fluorescence images based on their bimodal intensity distribution shown in Figure 5C (cf. M&M). If we use the position of the minimum between the two fluorescence peaks as cutoff to separate high and low intensities, the pixel-based RH fraction comprises about 70%. This proportion is closely similar to the fraction of ESCs defined as RH by flow cytometry analysis (4,11). In the following, we apply two measures, namely the number N_{RH} and the spatial fraction F_{RH} of RH clusters, to characterize the binarized images. As shown in the scatter plot in Figure 5D, the spatial fraction of high cells F_{RH} is only weakly correlated

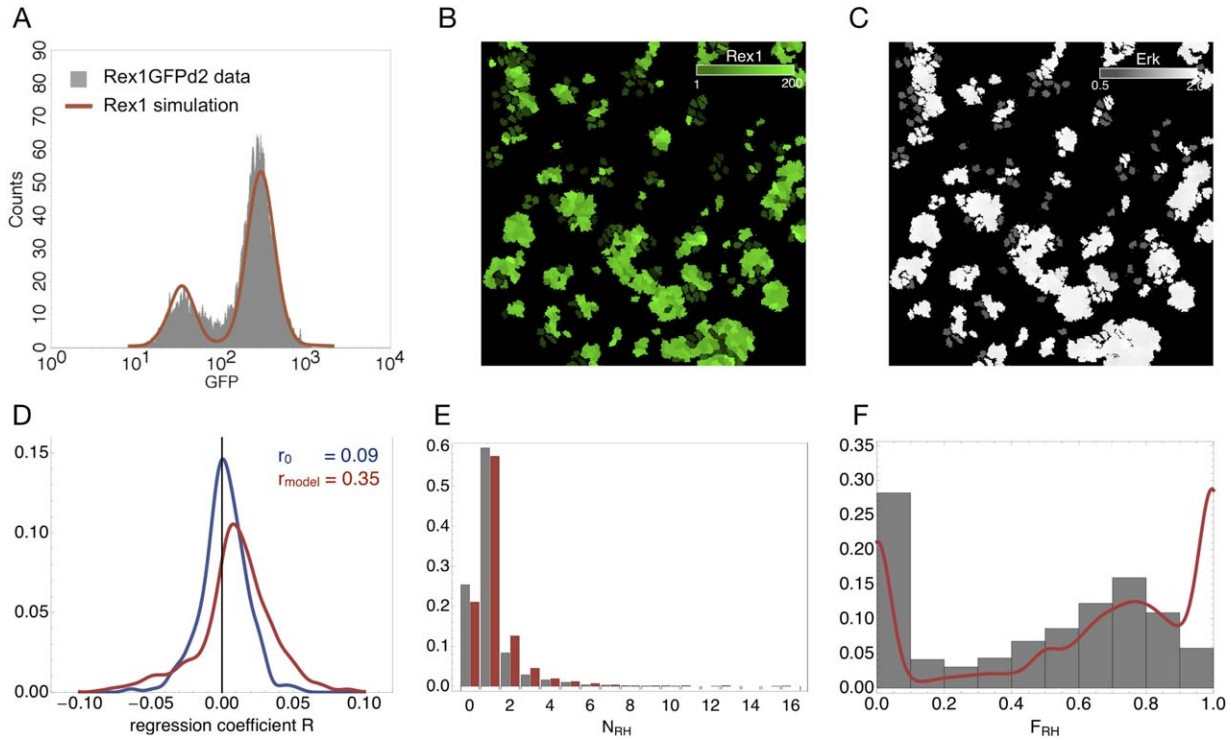

Figure 6. Comparison of model results with experimental data. (**A**) Flow cytometry measurements of Rex1GFPd2 cells reveal a bimodal distribution (gray histogram), which is consistently reproduced by simulated cell populations (red line). (**B**) Image of Rex1 expression levels of simulated ESCs. (**C**) Image of Erk levels received by simulated ESCs. (**D**) Distributions of regression coefficients *R* for the proposed ESC model (red line) and for a reference scenario (blue line), in which TF-related cell adhesions and proliferation rates are neglected. For comparison, see *in vitro* data in Fig. 5B. (**E**) The number of RH clusters per colony for live-cell (gray bars) and simulated (red bars) ESC colonies. (**F**) Distributions of spatial fractions of RH clusters in live-cell (gray histogram) and simulated (red line) ESC colonies.

with the area A of the respective colony. High fractions ($F_{RH} > 0.8$) can be found in all colonies independent of their size. However, colonies with a low spatial fraction ($F_{RH} < 0.2$) are rather small ($A < 800$ px). The color code in Figure 5D illustrates the number of highly fluorescent clusters per colony termed N_{RH}. While most of the colonies contain a small number of clusters (red/yellow), only a few large colonies are fragmented, that is, 2% of colonies consist of more than five clusters (green/blue).

In the following, we apply the same measures to simulated ESC colonies to analyze the consistency of our functional model with these observations.

Comparison of *In Vitro* and *In Silico* Cultures

The growth factor FGF4 is expressed by ESCs and is a potent autocrine activator of the differentiation-inducing Erk signaling (28,29). Erk itself acts as a repressor of Nanog transcription (11). Thus, in the simulation model, Nanog transcription is repressed by the cell's intrinsic Erk concentration and the Erk level of the neighboring cells, normalized by their contact surface (see M&M). To compare transcriptional and spatial patterns of the resulting *in silico* colonies with experimental data, we perform a large number of model simulations. As demonstrated by the distribution in Figure 6A (red line), a simulated ESC population in the spatial model establishes a bimodal Rex1 distribution similar to the experimental data. Images of simulated ESC colonies are shown in Figures 6B and 6C. ESCs in Figure 6B are colored according to their intrinsic Rex1 expression (RH: light green, RL: dark green). Figure 6C illustrates that isolated cells are only affected by their own, intrinsic Erk expression (Fig. 6C, Erk: dark gray), while cells located within a colony receive additional negative cues from their neighbors (Fig. 6C, Erk: light gray).

Comparing the model results with the experimental data using the quantitative measures described above, it turns out that both morphological and fluorescence patterns can consistently be explained by the proposed multiscale model. Interestingly, in the model RH cells are also located in the interior of a colony as demonstrated by the asymmetric distribution of the regression coefficients R in Figure 6D (effect size $r_{model} = 0.35$). To investigate the biophysical properties causing this appearance, we simulate ESC growth with different assumptions on cell proliferation and cell adhesion and analyze their distinct impact on the overall appearance of ESC colonies. In particular, for a setting in which all cells have a homogeneous low adhesion and identical proliferation rates, we observe that the cell's fluorescence intensity is independent of the location within the colony, indicated by symmetric distribution of regression coefficients R (effect size $r_0 = 0.09$, blue line in Fig. 6D). Adding either TF-related cell adhesions or proliferation rates, the distribution is shifted to the right ($r_1 = 0.22$, $r_2 = 0.27$, data not shown). As demonstrated in Figure 6E, the

clustering of cells/pixels with high Rex1 expression is almost identical between *in vitro* (gray bars) and *in silico* (red bars) cultures. The majority of ESC colonies has one high cluster, while only 2% possess more than five clusters. One notable difference between simulated and real colonies is found in the spatial fraction F_{RH} (Fig. 6F). Although the shape of both distributions is comparable, the amount of homogeneous RH colonies (i.e., $F_{RH} > 0.8$) is greater in the model than in the real data (model: 42%, data: 17%). Analyzing the images, we conclude that this divergence results from the fact that the borders of live-cell colonies appear less fluorescent than borders of simulated colonies. Therefore, completely homogenous RH colonies are rarely measured in the *in vitro* data; an effect that the model does not yet account for.

Discussion

Live-cell imaging of fluorescent reporter cell lines allows visualizing the heterogeneity of crucial TFs of ESC pluripotency grown in standard LIF/serum conditions (3,4,30). Using an image-based modeling approach, we developed a reference framework to which the *in vitro* image data can be compared and which allows the exploration of mechanisms underlying the functional organization of individual ESCs in spatially extended colonies.

As a prerequisite for the comparison of *in vitro* and *in silico* data, we established several quantitative measures for the overall colony appearance and applied them to microscopy image data of cultured Rex1GFPd2 ESCs. Although there are a number of standard measures of image textures (e.g., signal entropy, cooccurrence matrices) available, in a particular biological context these measures are generally more difficult to interpret and do not appear well-suited to describe biological patterns. With the measures presented in this article, we found that ESC colonies are highly variable in both morphology (i.e., size and circularity) and spatial fluorescence patterns. The variability of fluorescence intensities (i.e., the "spottiness") is predominant in large colonies, which can be a result of their origin (i.e., the merging of smaller substructures) or an indication of complex (e.g., opposing) cell interactions and cell signaling. For midsize and large colonies, we found that fluorescence levels decrease toward the colony border, which is measured by the regression coefficients R. Cells in the center of a colony might occasionally overlap due to limited space and compression. Thus, we additionally calculated the coefficients R neglecting the image pixels in the center. The overall effect of decreasing fluorescence levels toward the colony border appears robust, thus indicating that cells with low Rex1 expression and a higher propensity for differentiation are mainly located at the periphery of a colony. The binarization of the fluorescence intensities to separate these functionally distinct cells reveals that most colonies contain only one compact cluster of RH cells, which occupies a large spatial fraction. Analyzing the mean intensity of small colonies (area $A < 200$ px), 80% are classified as RL. Among large colonies (area $A > 10^4$ px), only about 30% have a RL appearance similar to the proportion observed in ESC populations analyzed by flow cytometry.

By construction and simulation of a mathematical multiscale model of ESC growth, we found that the circularity of *in vitro* colonies can most suitably be reproduced by *in silico* ESCs, in which adhesion properties depend on the individual cell's Rex1 concentration. The assumption that RH cells adhere strongly to each other, while RL have a very low adhesiveness, is widely consistent with the spatial patterns observed in live-cell images. As an example, the low adhesion of RL cells in our model leads to an overrepresentation of these cells in small/single-cell colonies, which is similar to the experimental data. Moreover, Rex1-related adhesions can account for the fact that RH cells are mainly located in the interior of a colony. This effect is further enhanced through the higher proliferation rate of RH cells. Notably, ESCs grown in defined 2*i* conditions [i.e., growth media supplemented by inhibitors of Gsk3 and the Erk pathway (31)] display neither transcriptional heterogeneity nor a propensity for differentiation (32). Analyzing the morphology of ESC colonies in 2*i*, we previously demonstrated that these cells form tightly packed and spherical colonies indicating rather high adhesion strengths between them (27). This appearance is consistent with our above conclusion that undifferentiated (i.e., RH) cells are more adhesive. However, it remains currently unclear whether the adhesiveness results from the intracellular cell state or whether adhesive cells express higher levels of pluripotency factors due to the establishment of important cell–cell interactions.

The modeling environment *Morpheus* can account for such feedback mechanisms between cell–cell adhesion and the transcriptional cell state by explicitly modeling intercellular signaling between adjacent cells (33). Here, we analyzed the case that intracellular pluripotency regulation of ESCs is regulated by a negative FGF4/Erk signaling mediated by adjacent cells within a colony. We demonstrate that this coupling is fully consistent with the experimentally observed colony structures evaluated from microscopy images. Furthermore, the model reliably reproduces the bimodal distribution of Rex1, which is a hallmark of ESC cultures in LIF/serum and which we also detect in the image data (cf. Fig. 5C). However, other intercellular processes, such as the biochemical communication via Wnt signaling or biomechanical interactions might also play an essential role in the formation of colonies and the overall maintenance of pluripotency in ESC populations.

A possible strategy to further investigate the relationship between cellular and intracellular properties of ESCs is a temporal analysis of pattern formation in evolving colonies using time-lapse microscopy. We previously demonstrated that an automated tracking of ESC colonies and a quantification of morphological features over time is feasible (27). This colony tracking framework can now be extended by quantitative measures of spatial fluorescence to analyze temporally extended processes such as cell differentiation. As a good example, White et al. recently studied spatial patterns associated with ESC differentiation by embryoid body formation and found that associated patterns can be explained by competing influences between neighboring Oct4+ and Oct4− cells (34).

Although we successfully demonstrated that the proposed quantitative measures can be applied to both *in vitro* and *in silico* images, several challenges remain. For example, although the bimodal distributions of intensity values (cf. Figs. 5C and 6A) allow to define and compare clusters with high and low expression, the observed intensity values cannot be directly mapped: the distribution of fluorescence values for the *in vitro* images results from pixel-wise intensities, while the readout of flow cytometry measures and model simulations are cell-wise intensities, that is, one single value per cell. Within our modeling framework, the concentration of TFs is calculated for an individual cell and assigned to all pixels of this spatially extended cell object. Thus, we do not account for potential intracellular variations of the intensity values. Furthermore, *in vitro* images are always "blurry" due to the point-spread function of the microscope and the proximity of adjacent cells, leading to a superposition of fluorescence signals. This effect has been neglected in the simulated images so far. Vice versa, it is currently not possible to resolve single cells within the bright field images as the colonies are often tightly packed. Therefore, we are currently unable to assign a single intensity value to individual cells. As a consequence of these structural differences between *in vitro* and *in silico* images, it is not reasonable to compare them with respect to the "spottiness" of the colonies, which measures the mean distance between distinct intensity values (cf. Fig. 5A and M&M). Special attention is furthermore required at colony borders. To avoid artifacts due to a segmentation of background pixels, automatically detected colony structures are slightly shrunken and pixels close to the edge are excluded for the analysis of position versus intensity, measured by the regression coefficients *R*.

In conclusion, our approach demonstrates that the development of appropriate simulation models is an important strategy to establish references for the comparison with experimentally obtained images. This comparison inevitably requires well-defined and statistically founded measures of spatial patterning. Correspondence between the observed data on one side and the conceptual models on the other side cumulates evidence for the appropriateness of the functional assumptions underlying the numerical implementations. Furthermore, such image-based quantitative models can be used to speculate about the impact of certain interactions and their potential perturbations, thereby providing the basis for further, hypothesis-driven experimental approaches.

Acknowledgment

The authors thank Tüzer Kalkan for the Rex1GFPd2 embryonic stem cells and Frank Buchholz for providing microscopy techniques and cell culture facilities.

Literature Cited

1. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981;292:154–156.
2. Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev 1998;12:2048–2060.
3. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A. Nanog safeguards pluripotency and mediates germline development. Nature 2007;450:1230–1234.
4. Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H. Identification and characterization of subpopulations in undifferentiated ES cell culture. Development 2008;135:909–918.
5. Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Francis Stewart A, Smith A, et al. The transcriptional and epigenomic foundations of ground state pluripotency. Cell 2012;149:590–604.
6. Kalmar T, Goodell MA, Lim C, Hayward P, Muñoz-Descalzo S, Nichols J, Garcia-Ojalvo J, Martinez Arias A. Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol 2009;7:e1000149.
7. Shi W, Wang H, Pan G, Geng Y, Guo Y, Pei D. Regulation of the pluripotency marker Rex-1 by Nanog and Sox2. J Biol Chem 2006;281:23319–23325.
8. Karwacki-Neisius V, Goke J, Osorno R, Halbritter F, Ng JH, Weisse AY, Wong FC, Gagliardi A, Mullin NP, Festuccia N, et al. Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. Cell Stem Cell 2013;12:531–545.
9. Navarro P, Festuccia N, Colby D, Gagliardi A, Mullin NP, Zhang W, Karwacki-Neisius V, Osorno R, Kelly D, Robertson M, et al. OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. EMBO J 2012;31:4547–4562.
10. Miyanari Y, Torres-Padilla ME. Control of ground-state pluripotency by allelic regulation of Nanog. Nature 2012;483:470–473.
11. Herberg M, Kalkan T, Glauche I, Smith A, Roeder I. A model-based analysis of culture-dependent phenotypes of mESCs. PLoS One 2014;9:e92496.
12. Chickarmane V, Olariu V, Peterson C. Probing the role of stochasticity in a model of the embryonic stem cell - Heterogeneous gene expression and reprogramming efficiency. BMC Syst Biol 2012;6:98.
13. Silva J, Smith A. Capturing pluripotency. Cell 2008;132:532–536.
14. Zhou J, Zhang Y, Lin Q, Liu Z, Wang H, Duan C, Wang Y, Hao T, Wu K, Wang C. Embryoid bodies formation and differentiation from mouse embryonic stem cells in collagen/matrigel scaffolds. J Genet Genomics 2010;37:451–460.
15. Peerani R, Onishi K, Mahdavi A, Kumacheva E, Zandstra PW. Manipulation of signaling thresholds in "engineered stem cell niches" identifies design criteria for pluripotent stem cell screens. PLoS One 2009;4:e6438.
16. Warmflash A, Sorre B, Etoc F, Siggia ED, Brivanlou AH. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. Nat Methods 2014;11:847–854.
17. Fernandes TG, Fernandes-Platzgummer AM, da Silva CL, Diogo MM, Cabral JM. Kinetic and metabolic analysis of mouse embryonic stem cell expansion under serum-free conditions. Biotechnol Lett 2010;32:171–179.
18. Starruss J, de Back W, Brusch L, Deutsch A. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. Bioinformatics 2014; 30:1331–1332.
19. Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, Smith A. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat Cell Biol 2011;13:838–845.
20. Sommer C, Straehle C, Köthe U, Hamprecht FA. Ilastik: Interactive learning and segmentation toolkit. Chicago, IL: IEEE. Proceedings of IEEE International Symposium on Biomedical Imaging. 2011. pp. 230–233.
21. Schwarzfischer M, Marr C, Krumsiek J, Hoppe PS, Schroeder T, Theis FJ. Efficient fluorescence image normalization for time lapse movies. Heidelberg: Proceedings of Microscopic Image Analysis with Applications in Biology; 2011.
22. Glauche I, Herberg M, Roeder I. Nanog variability and pluripotency regulation of embryonic stem cells--Insights from a mathematical model analysis. PLoS One 2010; 5:e11238.
23. Graner F, Glazier JA. Simulation of biological cell sorting using a two-dimensional extended Potts model. Phys Rev Lett 1992;69:2013–2016.
24. Podgorski GJ, Bansal M, Flann NS. Regular mosaic pattern development: A study of the interplay between lateral inhibition, apoptosis and differential adhesion. Theor Biol Med Model 2007;4:43.
25. Wolfram Research. Wolfram Research, I. Mathematica 2010. Champaign, IL: Wolfram Research, Inc.; 2010.
26. Team RDC. R: A language and environment for statistical computing. Austria: R Foundation for Statistical Computing; 2010.
27. Scherf N, Herberg M, Thierbach K, Zerjatke T, Kalkan T, Humphreys P, Smith A, Glauche I, Roeder I. Imaging, quantification and visualization of spatio-temporal patterning in mESC colonies under different culture conditions. Bioinformatics 2012;28:i556–i561.
28. Burdon T, Stracey C, Chambers I, Nichols J, Smith A. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev Biol 1999;210:30–43.
29. Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development 2007;134:2895–2902.
30. Hayashi K, Lopes SMCdS, Tang F, Surani MA. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell 2008;3:391–401.
31. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. The ground state of embryonic stem cell self-renewal. Nature 2008;453:519–523.
32. Wray J, Kalkan T, Smith AG. The ground state of pluripotency. Biochem Soc Trans 2010;38:1027–1032.
33. de Back W, Zimm R, Brusch L. Transdifferentiation of pancreatic cells by loss of contact-mediated signaling. BMC Syst Biol 2013;7:77.
34. White DE, Kinney MA, McDevitt TC, Kemp ML. Spatial pattern dynamics of 3D stem cell loss of pluripotency via rules-based computational modeling. PLoS Comput Biol 2013;9:e1002952.