Supplement

By Maria Herberg, Ingmar Glauche, Thomas Zerjatke, Maria Winzi, Frank Buchholz, and Ingo Roeder

Intracellular network models

Two different network models have been used to describe the transcriptional regulation of the pluripotency factors Oct4, Sox2, Nanog and Rex1, and Fgf4/Erk signalling. The parameters of the fluctuation model (given in Table 1) were taken from our previous publication [1]. Therein, the parameters have been adjusted such that model simulations fit bimodal Nanog and Rex1 expression profiles under self-renewing conditions and during the onset of differentiation. The repression rate p is the only parameter that required a readjustment in the agent-based model description because of the differential cell cycle times (cf. Table 1 in the main text). The parameters of the oscillation model (given in Table 1) were taken from another publication by us [2]. Therein, model parameters were chosen to fit the bimodal Nanog distribution under LIF/serum conditions. In the present study, the oscillation model has been extended by Rex1 and the former hypothetical factor X has been replaced by Fgf4/Erk signalling. We used the same parameters to model Rex1 regulation as in the fluctuation model.

Agent-based model

Cell division in the agent-based model has been implemented in two different ways. In one scenario, cell cycle times of daughter cells are drawn from the normal distributions shown in Figure 1A depending on the Rex1 concentration of the mother cell (maternally inherent cell cycle times). In the other scenario, division probabilities of mESCs depend on their current Rex1 expression and change dynamically as shown in Figure 1B (dynamic dIvision probabilities). The scenario on fixed, predetermined cell cycle times has been used throughout the study, apart from the last section in Results (*Analysis of cellular genealogies reveals correlation structures between single mESCs*), in which both models have been compared.

Parameter estimation based on cell counts in Rex1-sorted LIF/serum populations

Following an initial phase of increased cell death (most likely due to cellular stress induced by the sorting procedure), the proportion of dying cells stabilises around

	value			value	
parameter	fluctuation	oscillation	parameter	fluctuation	oscillation
$s_{1,2}$	75^*	75^*	d_{OS}	$0.01^\#$	$0.01^\#$
s_3	0.1^*	0.1^*	d_O	$1^\#$	$1^\#$
s_4	40^*	30^*	d_S	$1^\#$	$1^\#$
s_5	15^*	15^*	d_N	$1^\#$	$1^\#$
s_6	140^*	140^*	d_R	$1^\#$	$1^\#$
s_7	2^*	10.08^*	d_E	$1^\#$	$0.2^\#$
k	0.1	0.1	σ_{OS}	0.03	0.01
k_N	0.1	0.5	σ_N	0.115	0.01
p	$15^{*(1)}$	$4.4^{*(1)}$	σ_R	0.03	0.01
	$20^{*(2)}$	-	σ_E	0.03	0.01

Table 1. *Parameter sets for the intracellular fluctuation model and the oscillation model (* molecules/min; $^\#$ 1/min; $^{(1)}$ assuming identical cell cycle times; $^{(2)}$ assuming differential cell cycle times).*

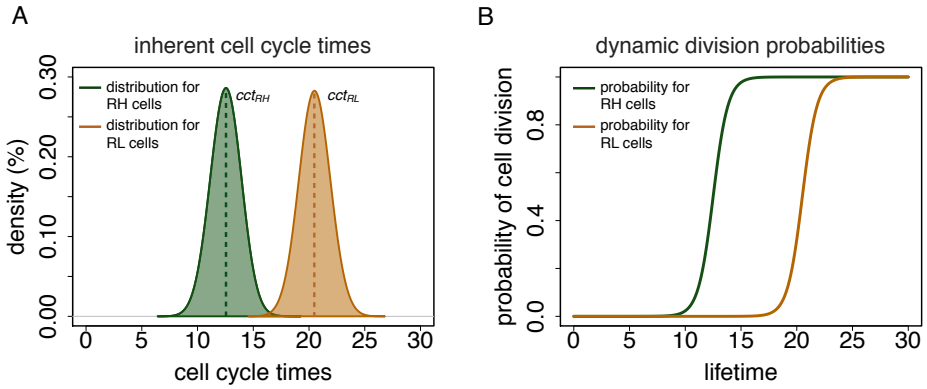

Figure 1. **Cell division scenarios to account for Rex1-related proliferation capacities**. (A) In the first scenario, each cell gets a predefined cell cycle time based on the Rex1 concentration of the mother cell just before division. If the mother cell is RH, both daughter cells draw cell cycle times from a normal distribution with mean cct_{RH} (green distribution). If the mother cell is RL, the daughter cells obtain cell cycle times from a normal distribution with mean cct_{RL} (orange distribution). The standard deviation is 1 for both distributions. (B) In the second division model, cells divide with a certain probability that depends on the current Rex1 concentration (green line for RH cells, orange line for RL cells) and increases with the lifetime of the cell. Cells can change between the two probability functions.

almost identical mean values of 9.18% in initially sorted RH subpopulations (sRH) and 9.35% in initially sorted RL subpopulations (sRL) similar to the unperturbed situation (Figure 2). Assuming a degradation time of 8 hours, a death rate d of 0.014 (cells per hour) in the agent-based model is required to account for a constant proportion of about 10% visible dead cells. We therefore assumed identical

death rates of $d = 0.014$ (cells per hour) for both subpopulations similar to the unperturbed situation.

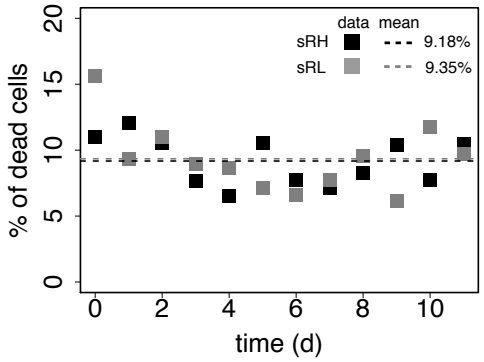

Figure 2. **The fraction of dying cells in sorted subpopulations**. The fraction of dying cells measured by flow cytometry reveals similar proportions in sRH (black squares) and sRL (grey squares) populations. Dotted lines show the respective mean proportions.

To estimate the Rex1-related mean cell cycle times cct_{RH} and cct_{RL} of mESCs cultured in LIF/serum conditions depending on the underlying mechanism controlling reversible state transitions, we performed two parameter screenings for each intracellular model. In particular, we first simulated mESC growth for a broad range of parameter sets (step size = 1 hour, screen 1 in Table 2) to reduce the parameter space to a reasonable size for the second screening (step size = 0.5 hours, screen 2 in Table 2). For each parameter combination (cct_{RH} and cct_{RL}) we simulated three replicates in screen 1 and six replicates in screen 2 and calculated the residual sum of squares (RSS) between measured and simulated cell counts for both sorted subpopulations defined as $RSS = RSS_{sRH} + RSS_{sRL}$ with:

$$\begin{aligned} RSS_{sRH} &= \sum_t (X^m_{sRH,t} - X^s_{sRH,t})^2 \\ RSS_{sRL} &= \sum_t (X^m_{sRL,t} - X^s_{sRL,t})^2 \end{aligned}$$

$X^m_{sRH,t}$ ... measured cell counts of living cells in sRH populations at time point t
$X^s_{sRH,t}$ ... simulated cell counts of living cells in sRH populations at time point t
$X^m_{sRL,t}$ ... measured cell counts of living cells in sRL populations at time point t
$X^s_{sRL,t}$ ... simulated cell counts of living cells in sRL populations at time point t

The conditions and results of the parameter screenings are summarised in Table 2. Parameter combinations with the lowest error in the second screen have been used for the agent-based fluctuation model (cct_{RH}=12.5h and cct_{RL}=20.5h) and

	cct_{RH}		cct_{RL}		
	range	value	range	value	RSS
fluctuation model					
screen 1	10.0h-30.0h	12.0h	10.0h-40.0h	21.0h	1.04
screen 2	11.0h-14.0h	**12.5h**	14.0h-34.0h	**20.5h**	0.77
	12.0h-12.5h		16.5h-24.5h		$< 1.18^*$
oscillation model					
screen 1	10.0h-20.0h	11.0h	10.0h-80.0h	70.0h	1.49
screen 2	11.0h-12.0h	**11.0h**	62.0h-78.0h	**74.5h**	1.06
	11.0h-11.0h		70.5h-76.5h		$< 1.46^*$

Table 2. *Results of the two parameter screenings to estimate the Rex1-related mean cell cycle times cct_{RH} and cct_{RL} depending on the intracellular network model (* 5% quantile).*

for the oscillation model (cct_{RH}=11.0h and cct_{RL}=74.5h), respectively. In order to assess the sensitivity of the simulation results we determined the parameter range for which error rates are lower than the 5% quantile of the error distribution shown in Figure 3.

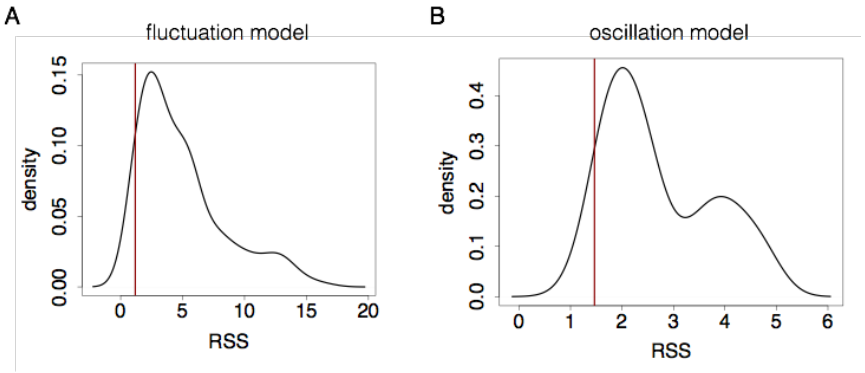

Figure 3. **Error distributions for cell cycle time estimates.** Distribution of the RSS using (A) the fluctuation model and (B) the oscillation model. Red lines show the limits of the respective 5% quantiles (i.e. 1.18 and 1.46).

Due to the interactions in the oscillation and in the fluctuation model, Nanog and Rex1 concentrations correlate in individual mESCs as shown in Figure 4.

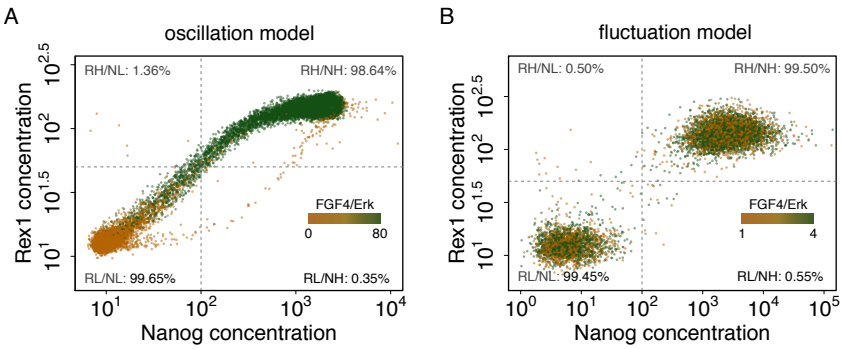

Figure 4. **Correlation between Nanog and Rex expression in simulated mESCs.** Each dot indicates Nanog and Rex1 levels in one cell at a certain time point. The colour code illustrates the corresponding Fgf4/Erk level of that cell. In both models almost 100% of the cells are in the same expression state with respect to Nanog and Rex1 expression. However, in the oscillation model Fgf4/Erk depends on the expression of Nanog (A), while it is independent in the fluctuation model (B).

Statistical analysis of cellular genealogies

To investigate whether different assumptions on intracellular or cellular properties result in detectable correlation structures between sister cells, simulated cellular genealogies as shown in Figure 4 have been evaluated statistically. In particular, the approach described in [3] has been adapted to measure (i) correlations in the cell cycle times of sister cells, and (ii) correlations in the number of state transitions. In the following we briefly introduce the procedure.

(i) Correlations in cell cycle times between sister cells: for all pairs of sister cells with completed cell cycles the difference between their cell cycle times was calculated and averaged. The resulting observed mean difference is denoted as $\bar{d}_o$. Assuming that cell cycle times of sister cells are independent, the expected distribution of mean differences can be calculated by the following randomisation procedure: within each genealogy and each generation, the cells' relationship is destroyed by randomly choosing pairs of dividing cells. For all random pairs of cells, the mean difference in the cell cycle times is calculated. Repeating this randomisation step very often, a null distribution is derived, to which the observed value $\bar{d}_o$ can be compared (cf. Figure 6) by computing an empirical p-value.

(ii) Correlations in the number of state transitions: for all pairs of dividing sister cells, the difference between their number of state transitions has been calculated and averaged. Thereby, the number of state transitions has been normalised by the cell's lifetime and short fluctuations, defined as less than 3 simulated time steps (i.e. corresponding to 3h in an experimental setting), have been neglected. Comparing the mean difference of the number of state transitions to the null distribution of randomly shuffled cell pairs, a significant p-value indicates that state transitions are not random, but intrinsically anchored and heritable.

Figure 5 illustrates simulated cellular genealogies either using maternally inherent cell cycle times (upper row) or dynamic division probabilities (bottom row).

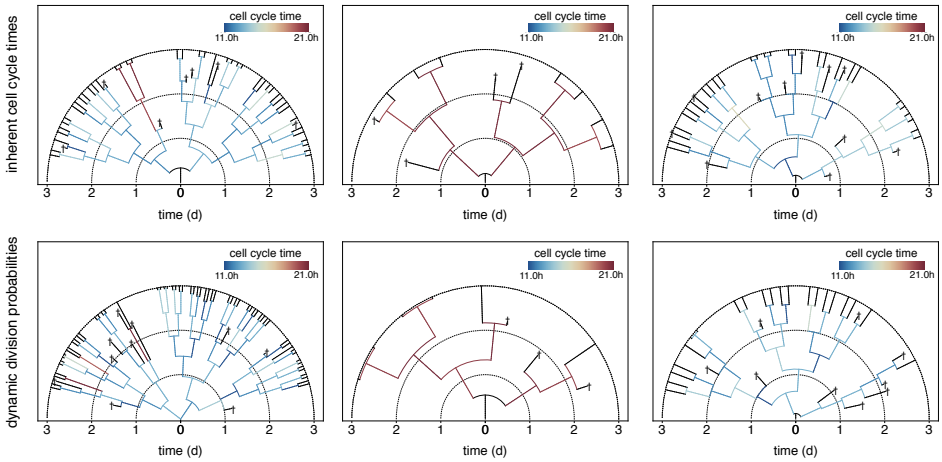

Figure 5. **Cellular genealogies depending on the cell division scenario.** In these genealogies, cells are depicted as branches starting with the founding cell at the root. A dividing cell is connected to its daughter cells via an arc. Dead cells are marked by crosses. Time is depicted radially. The colour code illustrates individual cell cycle times. Black branches indicate that the effective cycle time of this particular cell is not measurable, e.g. due to cell death or incomplete cell tracks.

Figure 6 depicts simulated cellular genealogies using either the fluctuation model (A, top) or the oscillation model (B, top). A statistical analysis of the number of state transitions reveals differences in the correlation of switching events between both mechanisms.

In a population with rare state transitions (as predicted by the fluctuation model), in many sister pairs both cells will keep their state and therefore show no difference in the number of switching events. This is reflected by a reference distribution with rather low values (Figure 6A, bottom). In contrary, in a population with frequent state transitions (as predicted by the oscillation model), the mean difference in

the number of switching events between sister cells can be much higher, leading to a reference distribution with higher values (Figure 6B, bottom). While in the fluctuation model the observed mean difference in the number of switching events between sister cells ($\bar{d}_0$) is consistent with the random distribution and therefore with the assumption of independence between sister cells (p-value=0.24, Figure 6A, bottom), in the oscillation scenario the observed mean value $\bar{d}_0$ deviates from the random distribution, thus indicating a synchronisation of switching events between sister cells (p-value$<10^{-4}$).

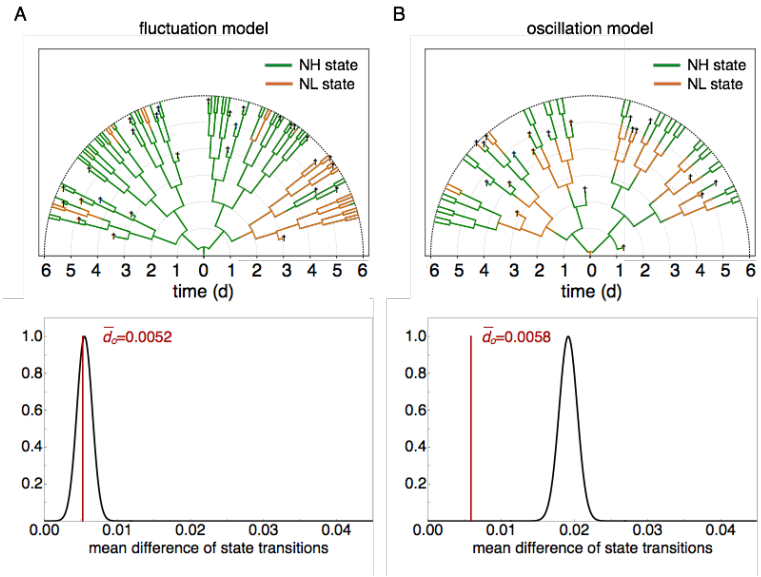

Figure 6. **Comparison of the number of state transitions between sister cells and random pairs dependent on the intracellular network model.** (A) The fluctuation model predicts rare, stochastic state transitions as shown by the simulated genealogy, in which cells are colour coded according to their Rex1 expression. Thus, the observed mean difference in the number of switching events between sister cells using this model (red line, $\bar{d}_0$=0.0052) does not differ from mean differences expected for random pairs (black distribution, p-value=0.24). (B) In contrast, in the oscillation model state transitions are more frequent such that the observed mean difference in the number of switching events between sister cells ($\bar{d}_0$=0.0052) differs significantly from the one expected for random pairs (p-value$< 10^{-4}$).

References

[1] Herberg, M., Kalkan, T., Glauche, I., Smith, A. & Roeder, I. 2014 A model-based analysis of culture-dependent phenotypes of mescs. PLoS ONE, **9**(3), e92 496.

[2] Glauche, I., Herberg, M. & Roeder, I. 2010 Nanog variability and pluripotency regulation of embryonic stem cells–insights from a mathematical model analysis. PLoS ONE, **5**(6), e11 238.

[3] Bach, E., Zerjatke, T., Herklotz, M., Scherf, N., Niederwieser, D., Roeder, I., Pompe, T., Cross, M. & Glauche, I. 2014 Elucidating functional heterogeneity in haematopoietic progenitor cells: a combined experimental and modelling approach. Experimental hematology, **42**(9), 826–837.