A novel view on stem cell development: Analysing the shape of cellular genealogies - Supplemental Data

Ingmar Glauche, Ronny Lorenz, Dirk Hasenclever and Ingo Roeder

1 Mathematical model of hematopoiesis

1.1 Model of stem cell self-renewal

In the single-cell based mathematical model of hematopoietic stem cell organisation each cell has an individual identity characterised by a set of properties, namely the current signaling context (A or Ω) the cell resides in, the position in the cell-cycle c and the affinity a. The vector of lineage propensities $\mathbf{x}$ represents an additional dimension of each simulated cell. The model is updated in discrete time steps, typically measuring one hour.

In (proliferation promoting) signaling context Ω a cell's affinity a, which relates to the its ability for long-term repopulation, decreases with constant rate $1/d$. After completion of a cell-cycle of length T_C (with the phases G1, S, G2/M), the cell divides and gives rise to two identical daughter cells, inheriting the properties of the parental cell. The lineage propensities x_i are updated according to the progressive control regime while the cell is in G1 phase (see section 1.2, below).

In (quiescence promoting) signaling context A the affinity a can be regained with constant rate r up to an upper limit $a_{\max} = 1$. For typical *in vitro* scenarios this rate is reduced to $r = 1$, indicating that the affinity a is maintained but not increased. Proliferation of cells is suppressed in A, assuming the cells to be in G0-phase. The lineage propensities x_i in signaling context A are updated according to the regressive control regime (see section 1.2, below).

Transition between the signaling contexts is modeled as a stochastic process: the probability for a change from A to Ω and vice versa depends on the cell's affinity a and the number of cells under the governance of the opposing signaling context. Details of transition function and their motivation are outlined in [1, 2].

After the affinity a of a cell under signaling context Ω has fallen below a lower limit $a_{\min}$, the cell's probability for transition to signaling context A declines to zero. Starting from this point the cell clone expands in a proliferative phase of t_{prolif} hours. In a subsequent maturation phase of t_{mature} hours the progeny survives in G1 without further divisions before the cells are finally deleted from the system. Cells in proliferative and maturation phase are counted in the pool of differentiating cells.

For the incorporation of cell death events (e.g. apoptosis), we assume that with a certain (low) probability p^{kill} every cell in G1-phase can be subject to induced cell death at each time step. This way, induced cell death acts in a random fashion.

1.2 Modeling intracellular lineage specification

For a mathematical representation of the intracellular lineage specification dynamics, each cell is characterised by a vector $\mathbf{x}(t) = (x_1(t), x_2(t), ..., x_N(t))$ coding for the N lineage propensities at time t. This lineage propensity vector $\mathbf{x}(t)$ is always normalised to 1 ($\sum_i x_i(t) = 1$), i.e. the $x_i(t)$ represent relative propensity levels. If a lineage propensity x_i exceeds a certain threshold x^{commit}, the cell is phenotypically assigned to a lineage fate X_i.

The lineage propensities x_i are updated as follows: In a stochastic process, which is repeated at every time step (typically measuring one hour), one lineage is chosen randomly with a probability equal to its lineage propensity $x_i(t)$. The chosen lineage propensity is modified according to $x_i(t+1) = x_i(t)(1 + m_i)$ in which m_i represent a context dependent, lineage specific reward. The subsequent normalisation of the lineage propensity vector $\mathbf{x}(t)$ to unity accounts for the concept of antagonistic interaction between different lineages.

The lineage specific rewards m_i are defined as linear functions of $x_i(t)$, in particular $m_i = b_i x_i + n_i$, in which the b_i and n_i differ for the progressive and the regressive control regime. Under the progressive control regime, advocated in the proliferative signaling context Ω, the reward functions have zero slope ($b_i^{\text{prog}} = 0$) and are restricted to the positive plane. This way $m_i = n_i^{\text{prog}}$ is positive and independent of x_i , leading to increasing divergence from the mean propensity level $x^* = 1/N$. In the regressive control regime, advocated in the quiescence promoting signaling context A, the reward function has negative slope ($b_i^{\text{reg}} < 0$) and a root at x^R which defines the corresponding intercept at $n_i^{\text{reg}} = -b_i^{\text{reg}} x^R$. The root at x^R should be chosen such that $x^R \approx x^* = 1/N$ in order to get convergence to the mean propensity level $x^* = 1/N$. In this setting the reward m_i is positive for $x_i < x^R$ and negative for $x_i > x^R$.

For a biological motivation of the assumed lineage specification dynamics and for further technical details we refer to [3].

2 Model parameters

2.1 Culture scenarios

Model description and parameters are provided for the three general classes of simulation scenarios presented in the results section.

Growth Scenario. For the growth scenario 400 model systems are initialised, each with an individual cell. These individual cells are tracked for 300 hours, and subsequently the cellular genealogies are derived. Model parameters are chosen such that the system can establish a homeostatic situation. The parameters are given in Supplementary Table 1.

Homeostatic scenario. Initialising a model system with a single cell, cell numbers reach a homeostatic situation after about 500 hours (compare Figure 3 in the main publication). At time point t=700 hours, all stem cells (i.e. cells with $a > 0.1$) of a particular realisation are uniquely marked and subsequently tracked for the next 300 hours. For the particular, randomly chosen realisation 399 cells with affinity $a > 0.1$ have been found and their genealogies have been reconstructed. Model parameters are identical to the growth scenario. The differences in the genealogies result from the differences in the initial configuration (expansion of a single cell in an empty model system vs. homeostasis in a "filled" system). Parameters are given in Supplementary Table 1.

Differentiation scenario. For the differentiation scenario, all stem cells (i.e. cells with $a > 0.1$) of a particular realisation are uniquely marked in the homeostatic situation at time point t=1500. In the next step the differentiation rate is increased to $d = 1.1$ and the regeneration rate is reduced to $r = 1.0$ for all cells. This way, the cells rapidly decrease their affinity a, lose the potential for self-renewal and undergo terminal differentiation. 390 cells with $a > 0.1$ have been found in the particular realisation. These cells have been tracked during the next 300 h, and genealogies have been reconstructed. Parameters of the simulations are given in Supplementary Table 1.

2.2 Distribution of the topological measures

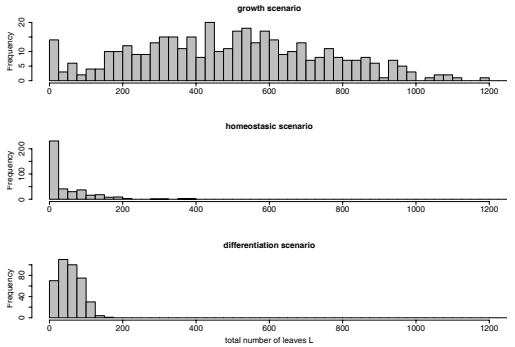
For a more detailed perception of the shape of the distributions that have been reported using the boxplots in Figure 4 of the main text, we provide corresponding histograms in Supplementary Figure 1.

References

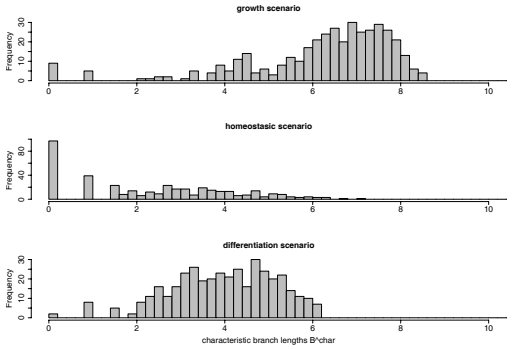
[1] I. Roeder and M. Loeffler. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. *Exp. Hematol.*, 30(8): 853–861, 2002.

[2] I. Roeder, M. Horn, I. Glauche, A. Hochhaus, M. C. Mueller, and M. Loeffler. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nat Med*, 12(10):1181–1184, 2006.

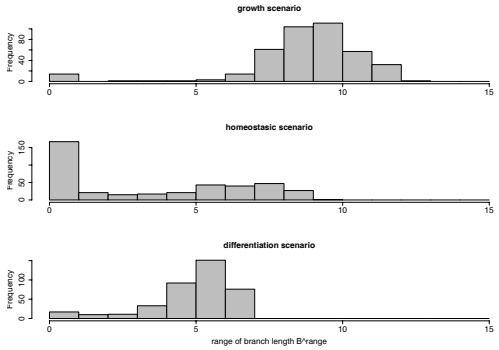
[3] I. Glauche, M. Cross, M. Loeffler, and I. Roeder. Lineage specification of hematopoietic stem cells: Mathematical modeling and biological implications. *Stem Cells*, 25 (7):1791–1799, 2007.



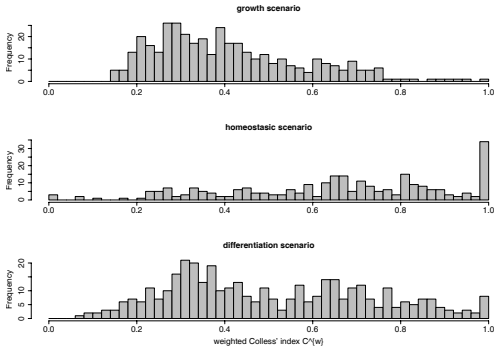
(a) Number of leaves L



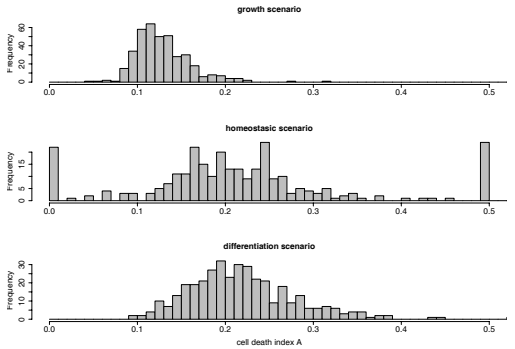
(b) Characteristic branch length B^{char}



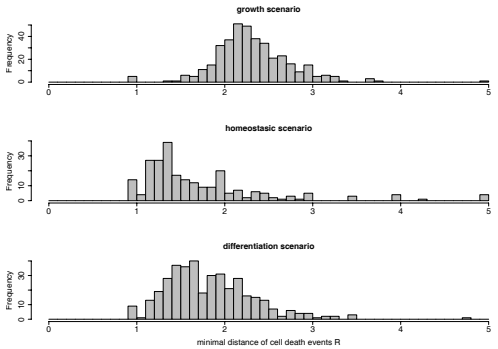
(c) Difference in branch lengths B^{diff}



(d) Weighted Colless' index C



(e) Cell death index A



(f) Minimal topological distance between characteristic events R

Supplementary Figure 1: **Characteristic measures of tree shape.** Shown are the histograms that correspond to the boxplots in Figure 4.

parameter	growth scenario	homeostatic scenario	differentiation scenario
d	1.035	1.035	1.1
r	1.07	1.07	1.0
$a_{\min}$	0.1	0.1	0.1
$f_\alpha(0)$	0.5	0.5	0.5
$f_\alpha(N_\mathrm{A}/2)$	0.3	0.3	0.3
$f_\alpha(N_\mathrm{A})$	0.01	0.01	0.01
$f_\alpha(\infty)$	0	0	0
N_A	1300	1300	1300
$f_\omega(0)$	0.5	0.5	0.5
$f_\omega(N_\Omega/2)$	0.3	0.3	0.3
$f_\omega(N_\Omega)$	0.1	0.1	0.1
$f_\omega(\infty)$	0	0	0
N_Ω	280	280	280
T_C	24 hours	24 hours	24 hours
T_S	8 hours	8 hours	8 hours
$T_{G_2/M}$	4 hours	4 hours	4 hours
t_{prolif}	220 hours	220 hours	220 hours
t_{mature}	175 hours	175 hours	175 hours
N	3	3	3
progressive control			
n_i $(i = 1...N)$	0.15	0.15	0.15
b_i $(i = 1...N)$	0	0	0.15
regressive control			
n_i $(i = 1...N)$	0.1	0.1	0.1
b_i $(i = 1...N)$	-0.6	-0.4	-0.4
t_{sim}	300 hours	300 hours	300 hours
tracked cell pool	400 independent simulations each initialised with one cell	one simulation, tracking all 399 cells with $a > a_{\min}$ from homeostatic situation at time point 700	one simulation, tracking all 390 cells with $a > a_{\min}$ from homeostatic situation at time point 1500

Supplementary Table 1: **Model parameters.** d: differentiation coefficient; r: regeneration coefficient; $a_{\min}$: lower limit for the maintenance of the self-renewal ability; $f_{\alpha/\omega}$: transition characteristic for change from GE-Ω/A to GE-A/Ω; $N_{\mathrm{A}/\Omega}$: scaling factor of transition characteristics (for further details see [2] and the corresponding supplemental notes); T_C: cell cycle duration; T_S: S-phase duration; $T_{G_2/M}$: duration of G_2 and M-phase; t_{prolif}: duration of proliferation phase; t_{mature}: duration of maturation phase; N number of lineages; n_i and b_i define the lineage specific reward function $m_i = -b_i x_i + n_i$, separately for the progressive and the regressive control regime; t_{sim}: observation period