iScience, Volume *24*

Supplemental Information

Computational modeling of stem and progenitor cell kinetics identifies plausible hematopoietic lineage hierarchies

Lisa Bast, Michèle C. Buck, Judith S. Hecker, Robert A.J. Oostendorp, Katharina S. Götze, and Carsten Marr

Supplemental Information

Supplemental Data Items

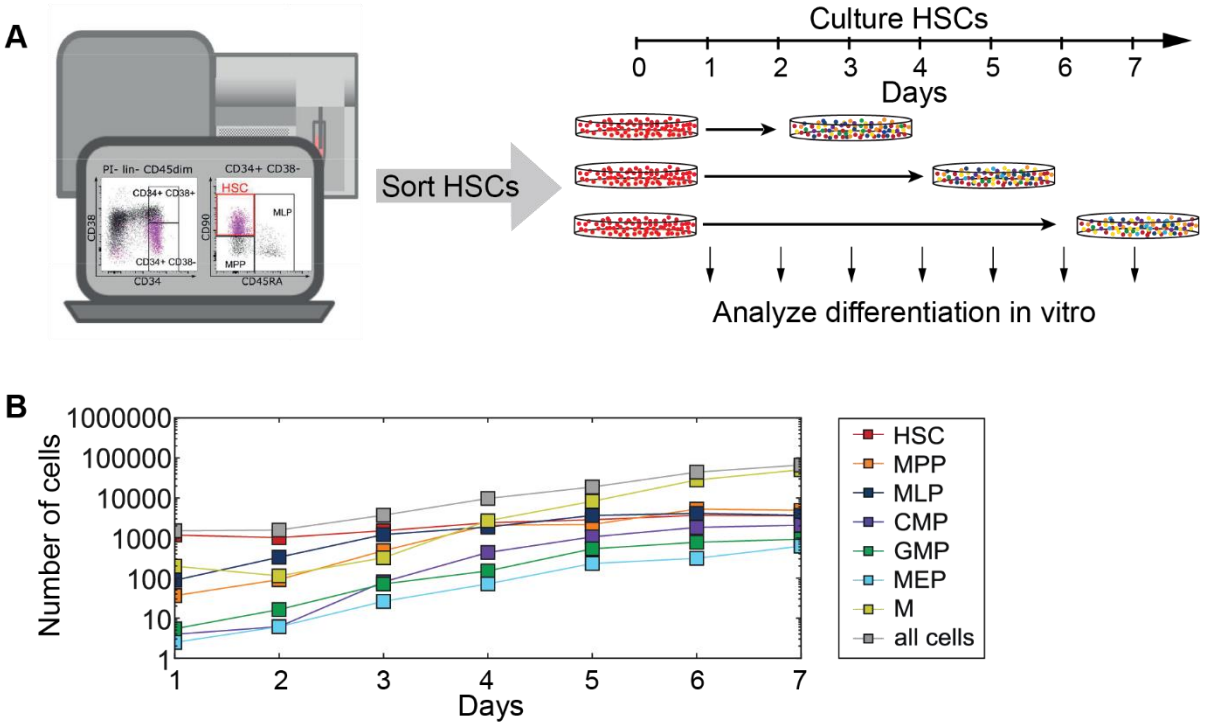

Figure S1. Measuring number of cells over time, Related to Figure 1 and Figure 3.
(A) Bone marrow cells from 10 donors were isolated. Each sample was purified by multiparameter immunophenotyping (FACS) for a HSC starting population which was then divided and cultured in several wells in parallel. Cell cultures were measured with FACS to observe cell abundances and division distribution of cultured bulk cells at subsequent measurement timepoints.
(B) Measured cell abundances for all seven cell type compartments (HSC, MPP, MLP, CMP, GMP, MEP, M) over time (connected dots) for one exemplary sample (ID 7).

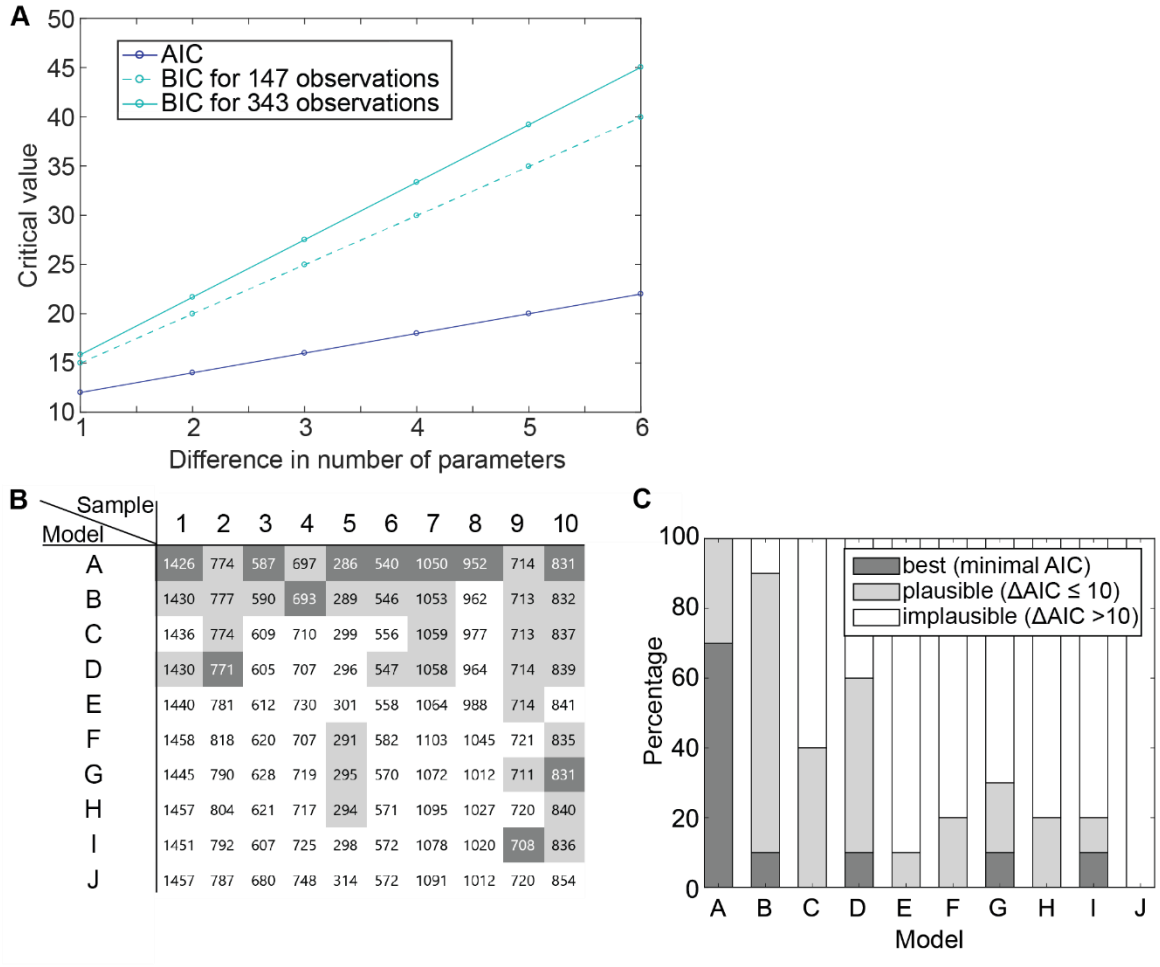

Figure S2. AIC identifies similar plausible models as BIC, Related to Figure 3.

(A) Critical values for BIC and AIC for rejecting a model over another less complex model for a difference in number of parameters of 1 to 6 are shown for the upper and lower boundary of observations. BIC shows larger critical values and is thus always a more conservative criterion compared to AIC.

(B) AIC values per model A-J (rows) for every individual sample (columns). Models were categorized into best (for the lowest AIC score), plausible (a difference to the lowest score ≤ 10) and implausible (a difference to the lowest score ≥ 10) for each sample.

(C) Relative frequency of a model to belong to one of the three categories (best, plausible, implausible) according to the AIC obtained from the 10 samples.

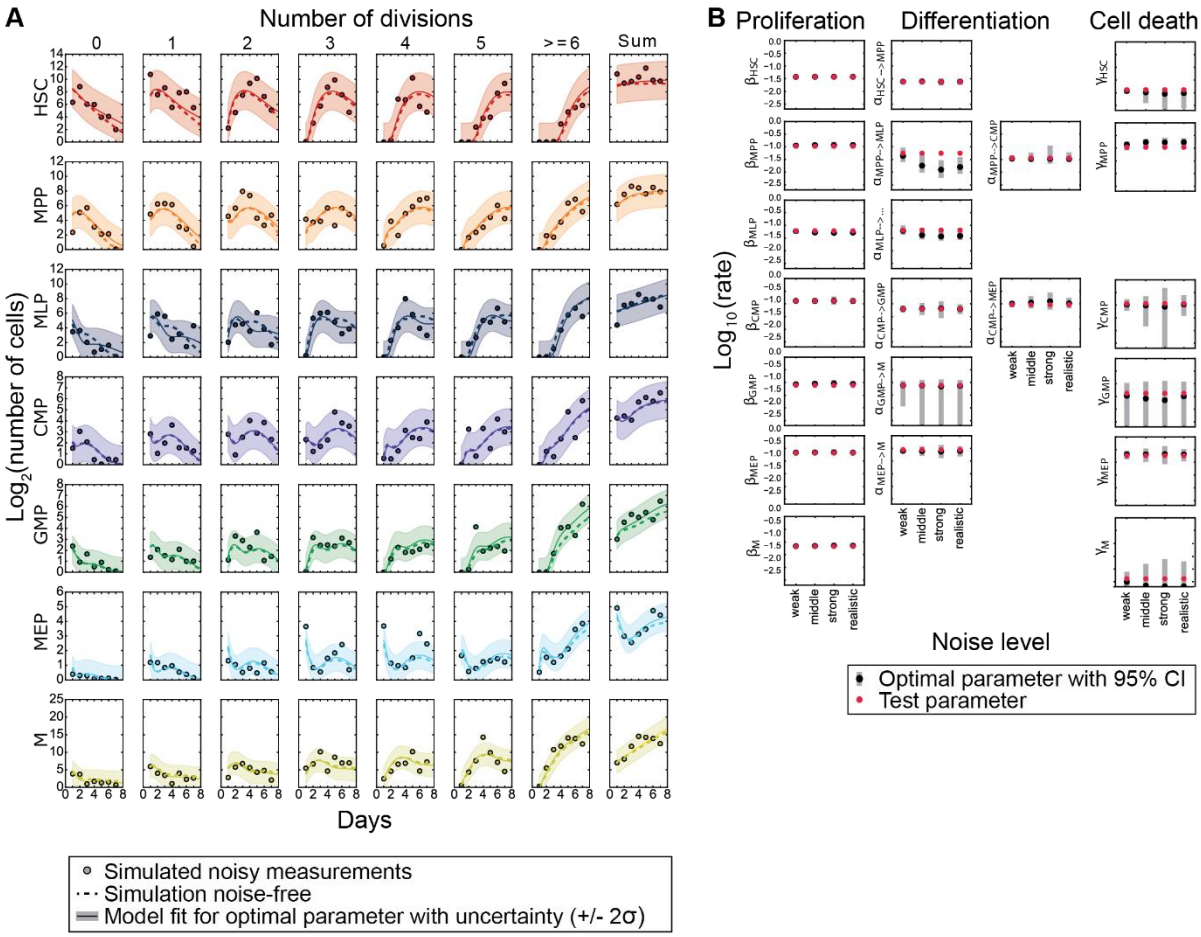

Figure S3. Exemplary model simulation and fitting, Related to Figure 4.

(A) Cell abundances of 49 compartments (first 7 columns) and sum over all divisions (last column) for each cell type. Model A was used to generate in silico data for a realistic test parameter (dots) and perturbed by cell compartment specific log-normal noise (shown for realistic noise level, $\sigma \in [0.6,1.1]$). Underlying (noise free) model observables (dashed line) and the model observables for the optimal parameter (solid line) deviate only slightly for the assumed noise level ($\pm 2\sigma$ error band).

(B) True parameter values (red dots) for which perturbed samples were simulated by using model A are for most rates contained in 95% confidence interval (grey boxes) of inferred parameter values (black dots) and deviate only slightly from the true value even for higher noise levels. The noise level for perturbation of simulated values was varied and set to σ= 0.4 for weak, σ= 0.8 for middle, σ=1.2 for strong, and $\sigma \in [0.6,1.1]$ for realistic noise.

Cell type(s)	Marker	Published in
HSC *	Lin$^-$ CD34$^+$ CD38$^-$ CD90$^+$ CD45RA$^-$	(Doulatov et al., 2010; Majeti et al., 2007)
HSC	Lin$^-$ CD34$^+$ CD38$^-$ CD45RA$^-$ CD90$^+$; CD49f$^+$	(Notta et al., 2011)
HSCs/ MPPs	CD133$^+$ CD34$^+$ CD45RA$^-$	(Görgens et al., 2013)
MPP *	Lin$^-$ CD34$^+$ CD38$^-$ CD90$^-$ CD45RA$^-$	(Doulatov et al., 2010; Majeti et al., 2007)
MPP	CD34$^+$ CD10$^-$ CD38$^-$ CD90$^-$ CD45RA$^-$	(Ostendorf et al., 2018)
CMP *	Lin$^-$ CD34$^+$ CD38$^+$ CD45RA$^-$ CD123$^+$	(Manz et al., 2002)
CMP	Lin$^-$ CD34$^+$ CD38$^+$ CD45RA$^-$ Flt3$^+$ CD7$^-$ CD10$^-$	(Doulatov et al., 2010)
GMP *	Lin$^-$ CD34$^+$ CD38$^+$ CD45RA$^+$ CD123$^+$	(Manz et al., 2002)
GMP	Lin$^-$ CD34$^+$ CD38$^+$ CD45RA$^+$ Flt3$^+$ CD7$^-$ CD10$^-$	(Doulatov et al., 2010)
MEP *	Lin$^-$ CD34$^+$ CD38$^+$ CD45RA$^-$ CD123$^-$	(Manz et al., 2002)
MEP	Lin$^-$ CD34$^+$ CD38$^+$ CD45RA$^-$ Flt3$^-$ CD7$^-$ CD10$^-$	(Doulatov et al., 2010)
Lymphoid/DC-restricted progenitors	CD34$^+$ Lin$^-$ CD10$^+$	(Galy et al., 1995)
MLP *	Lin$^-$ CD34$^+$ CD38$^-$ CD90$^{neg–lo}$ CD45RA$^+$	(Doulatov et al., 2010)
CLP	Lin$^-$, CD34$^+$, CD127$^+$	(Pang et al., 2011)
lympho-myeloid	CD133$^+$ CD34$^+$ CD45RA$^+$	(Görgens et al., 2013)
erythro-myeloid	CD133^{low} CD34$^+$ CD45RA$^-$	(Görgens et al., 2013)

Table S1: Commonly used gating strategies for identifying human hematopoietic stem and progenitor cells, Related to Figure 1. Strategies marked with * served as the basis for this study**.**

Parameter	Boundaries
$\beta_{S_j}, \gamma_{S_j}, \alpha_{S_{j1}\to S_{j2}}$, where $S_j, S_{j1}, S_{j2} \in \mathscr{S}$	$\left[\frac{1}{500}, 1\right] h^{-1}$
x_0	$\left[\min_r \left\{y^{\mathscr{D}}_{1,0}(t_1)\right\}, N_{input}\right]$, if $j = 1$ (for HSCs)
x_0	$\left[0, \max_r \left\{y^{\mathscr{D}}_{j,0}(t_1)\right\} + 0.1 * \max_r \left\{y^{\mathscr{D}}_{j,0}(t_1)\right\}\right]$, if $j > 1$ (other cell type compartments)

Table S2: Parameter boundaries used for fitting models A-J to experimental data, Related to Figure 2.

Parameter	**Individual ID (age, sex)**									
	1 (63, m)	**2 (57,m)**	**3 (70,m)**	**4 (63, m)**	**5 (51, m)**	**6 (26, m)**	**7 (29, m)**	**8 (44, m)**	**9 (24, m)**	**10 (76, w)**
$\alpha_{HSC \to MPP}$	-1.90 [-1.97, -1.83]	-1.72 [-1.79, -1.66]	-2.7 [-2.7, -2.52]	-1.47 [-1.57, -1.37]	-1.73 [-1.89, -1.57]	-1.61 [-1.74, -1.51]	-1.42 [-1.52, -1.32]	-1.47 [-1.54, -1.41]	-1.83 [-1.92, -1.73]	-1.62 [-1.74, -1.5]
β_{HSC}	-1.71 [-1.73, -1.68]	-1.64 [-1.67, -1.61]	-1.49 [-1.53, -1.46]	-1.45 [-1.5, -1.41]	-1.55 [-1.6, -1.49]	-1.32 [-1.38, -1.24]	-1.28 [-1.33, -1.22]	-1.35 [-1.38, -1.32]	-1.57 [-1.6, -1.54]	-1.52 [-1.55, -1.49]
γ_{HSC}	-2.48 [-2.70, -2.12]	-2.08 [-2.49, -1.67]	-1.43 [-1.5, -1.35]	-1.83 [-1.99, -1.72]	-1.51 [-1.72, -1.33]	-1.96 [-2.7, -1.7]	-2.7 [-2.7, -1.92]	-2.7 [-2.7, -2.28]	-2.07 [-2.57, -1.86]	-2.07 [-2.43, -1.72]
$\alpha_{MPP \to MLP}$	-1.33 [-1.50, -1.15]	-1.43 [-1.58, -1.27]	-2.7 [-2.7, -2.29]	-1.2 [-1.34, -1.04]	-2.7 [-2.7, -2.3]	-1.38 [-1.55, -1.24]	-0.97 [-1.21, -0.73]	-0.99 [-1.19, -0.79]	-1.34 [-1.57, -1.11]	-1.27 [-1.54, -1]
$\alpha_{MPP \to CMP}$	-1.14 [-1.32, -0.96]	-1.42 [-1.55, -1.28]	-2.7 [-2.7, -2.2]	-2.7 [-2.7, -2.53]	-2.7 [-2.7, -2.16]	-1.37 [-1.53, -1.12]	-1.32 [-1.49, -1.15]	-1.37 [-1.52, -1.23]	-1.49 [-1.72, -1.2]	-1.93 [-2.06, -1.8]
β_{MPP}	-1.26 [-1.31, -1.19]	-1.26 [-1.31, -1.2]	-1.28 [-1.32, -1.23]	-1.07 [-1.15, -0.98]	-0.85 [-1.05, -0.57]	-0.95 [-1.04, -0.85]	-1.21 [-1.44, -0.99]	-0.95 [-1.04, -0.86]	-1.22 [-1.3, -1.15]	-1.1 [-1.22, -0.98]
γ_{MPP}	-2.70 [-2.70, -2.29]	-2.7 [-2.7, -1.2]	-2.7 [-2.7, -2.11]	-1.32 [-1.67, -1]	-0.77 [-1.02, -0.49]	-2.7 [-2.7, -2.15]	-1.42 [-1.89, -0.96]	-1.3 [-1.95, -0.65]	-2.7 [-2.7, -1.37]	-1.03 [-1.42, -0.64]
α_{MLP}	-1.04 [-1.36, -0.77]	-1.06 [-1.3, -0.82]	-2.7 [-2.7, -1.91]	-0.79 [-0.99, -0.61]	-2.7 [-2.7, -1.91]	-2.7 [-2.7, -2.36]	-1.3 [-1.49, -1.11]	-1.36 [-1.53, -1.2]	-0.98 [-1.41, -0.67]	-1.19 [-1.46, -0.92]
β_{MLP}	-1.32 [-1.42, -1.21]	-1.39 [-1.59, -1.2]	-1.41 [-1.44, -1.36]	-1.08 [-1.23, -0.95]	-1.5 [-1.56, -1.44]	-1.13 [-1.2, -1.03]	-1.28 [-1.38, -1.19]	-1.15 [-1.24, -1.07]	-1.23 [-1.41, -1.05]	-1.27 [-1.41, -1.14]
$\alpha_{CMP \to GMP}$	-1.88 [-1.97, -1.78]	-1.48 [-1.76, -1.21]	-2.7 [-2.7, -2.27]	-1.71 [-1.86, -1.58]	-1.63 [-2.04, -1.43]	-1.51 [-1.68, -1.32]	-1.09 [-1.32, -0.85]	-1.19 [-1.4, -0.98]	-0.43 [-0.68, -0.17]	-2.7 [-2.7, 0]
$\alpha_{CMP \to MEP}$	-1.14 [-1.39, -0.42]	-0.63 [-0.88, -0.37]	-2.7 [-2.7, -1.93]	-1.77 [-1.87, -1.65]	-1.54 [-2.7, -1.29]	-0.84 [-1.24, -0.39]	-1.27 [-1.49, -1.05]	-1.02 [-1.21, -0.84]	-1.44 [-1.78, -0.79]	-2.7 [-2.7, 0]
β_{CMP}	-1.15 [-1.51, -0.88]	-0.99 [-1.18, -0.8]	-1.34 [-1.37, -1.31]	-1.19 [-1.25, -1.13]	-1.43 [-1.58, -1.3]	-0.41 [-0.7, -0.11]	-1.03 [-1.16, -0.9]	-0.85 [-0.99, -0.72]	-0.83 [-1.73, -0.45]	-1.13 [-1.28, -0.98]
γ_{CMP}	-0.36 [-1.62, -0.12]	-2.7 [-2.7, 0]	-2.7 [-2.7, -2.05]	-2.7 [-2.7, -2.17]	-2.7 [-2.7, -1.38]	-0.28 [-0.71, 0]	-2.7 [-2.7, -1.59]	-2.7 [-2.7, -1.25]	-0.5 [-2.7, 0]	-2.7 [-2.7, 0]
$\alpha_{GMP \to M}$	-2.70 [-2.70,	-2.7 [-2.7,	-2.7 [-2.7,	-1.26 [-1.52,	-2.7 [-2.7,	-2.7 [-2.7,	-2.7 [-2.7,	-2.7 [-2.7,	-0.5 [-0.76,	-2.7 [-2.7,

	-1.94]	0]	-1.63]	-1.08]	-0.44]	-1.18]	-1.28]	-0.43]	-0.1]	0]
β_{GMP}	-1.37 [-1.44, -1.22]	-1.2 [-1.5, -0.89]	-1.37 [-1.45, -1.31]	-1.02 [-1.13, -0.92]	-1.26 [-2.7, -0.4]	-1.05 [-1.16, -0.91]	-1.1 [-1.34, -0.85]	-0.99 [-1.24, -0.75]	-1.11 [-1.72, -0.57]	-1.44 [-1.5, -1.38]
γ_{GMP}	-2.70 [-2.70, -1.93]	-1.02 [-1.58, -0.46]	-1.81 [-2.7, -1.53]	-2.7 [-2.7, -1.21]	-0.82 [-1.18, -0.11]	-2.7 [-2.7, -1.11]	-0.88 [-1.16, -0.6]	-0.96 [-1.26, -0.67]	-2.7 [-2.7, 0]	-2.7 [-2.7, 0]
$\alpha_{MEP\rightarrow M}$	-0.68 [-1.28, -0.36]	0 [-0.25, 0]	-2.7 [-2.7, -2.19]	-2.7 [-2.7, -2.15]	-2.7 [-2.7, -0.64]	-0.8 [-1.36, -0.41]	-2.7 [-2.7, -1.07]	-0.6 [-0.87, -0.34]	-2.7 [-2.7, -0.02]	-1.58 [-2.4, -0.75]
β_{MEP}	-1.11 [-1.59, -0.77]	-0.41 [-0.86, 0]	-1.4 [-1.44, -1.35]	-1.24 [-1.32, -1.17]	-1.22 [-1.65, -0.54]	-0.98 [-1.49, -0.67]	-1.04 [-1.23, -0.84]	-0.74 [-0.92, -0.55]	-2.13 [-2.7, -1.31]	-1.27 [-1.41, -1.14]
γ_{MEP}	-1.11 [-2.70, 0]	-0.08 [-0.65, 0]	-2.7 [-2.7, -2.14]	-2.7 [-2.7, -2.05]	-0.95 [-1.29, -0.35]	-1.3 [-2.7, -0.44]	-0.85 [-1.12, -0.58]	-1.34 [-2.27, -0.41]	-2.7 [-2.7, -0.49]	-2.17 [-2.7, 0]
β_{M}	-1.48 [-1.52, -1.44]	-1.1 [-1.23, -0.98]	-1.19 [-1.23, -1.15]	-1.22 [-1.26, -1.16]	-1.2 [-1.25, -1.13]	-1.16 [-1.23, -1.09]	-1.32 [-1.36, -1.28]	-1.2 [-1.26, -1.13]	-1.04 [-1.13, -0.96]	-1.34 [-1.39, -1.29]
γ_{M}	-2.70 [-2.70, -2.31]	-2.7 [-2.7, 0]	-2.7 [-2.7, -2.18]	-2.7 [-2.7, -2.49]	-2.58 [-2.7, -1.72]	-2.7 [-2.7, -2.16]	-2.7 [-2.7, -2.17]	-2.7 [-2.7, -1.67]	-2.7 [-2.7, -1.79]	-2.7 [-2.7, -0.91]

Table S3: Log_{10} transformed parameter values with their 95% profile likelihood-based confidence intervals in [cells/h] resulting from fitting model A to each of the 10 individuals, Related to Figure 3.

Transparent Methods

Derivation of the set of lineage hierarchy models tested

Based on the classical model of hematopoiesis (model A, Figure 1A) and recently reported experimental evidence, we derived nine alternative models, likewise containing compartments HSC, MPP, CMPs, MLP, MEP, GMP, and M, but with different direct differentiation transitions between them (Figure 1B-J). The restriction on 7 cell types is determined by the number of discernable populations in flow cytometry. Furthermore, our setup does not allow to depict myeloid and lymphoid differentiation past the HSPC compartment: Whereas conditions allowing in vitro differentiation of HSCs into mature myeloid cells are well established, it is much more difficult to culture mature lymphoid cells past the committed progenitor stage which requires a specific culture medium using a cloned stromal feeder layer (i.e. Whitlock-Witte culture).

Several studies in humans (Doulatov et al., 2010; Doulatov et al., 2012; Giebel et al., 2006; Reynaud et al., 2003; Goardon et al., 2011; Rossi et al., 2008; Hao et al., 2001) show that progenitor cells in the CD34+CD38− compartment, which are CD90+ (Thy1) and CD45RA+, correspond to multipotent lymphoid progenitor cells (MLP), and have lymphoid, macrophage, and dendritic potential. As these results suggest that MLPs can also differentiate to GMPs, we have incorporated this transition in models B, C, E and I (Figure 1B,C,E,I).

In a study investigating adult blood lineage commitment in mice (Adolfsson et al., 2005), the authors proposed a revised model of hematopoiesis. They identified a new cell type, the lymphoid-primed

multipotent progenitors (LMPPs), which are FLT3+ Lin^- $Sca\text{-}1^+$ $c\text{-}Kit^+$ cells (LSK Flt3+ cells), that possess B-cell, T-cell and granulocyte-monocyte (GM) potential but lack megakaryocyte-erythrocyte (MegE) potential. Mouse LSK cells include long-term HSCs, short-term HSCs, and MPPs. The existence of a distinct LSK subtype that does not have MegE potential may indicate that MEPs can directly arise from HSCs. Furthermore, loss of MegE potential in the newly defined LMPP compartment indicates a direct LMPP to GMP transition, without differentiation into CMPs first. In this model, HSCs can generate LMPPs with lymphocyte and GM potential, and CMPs with MegE and GM potential. These findings led on the one hand to the possible direct transition from the HSC to CMP compartment and from MPPs to GMPs in models F and H (Figure 1F,H), and on the other hand to a transition between the HSC and MEP compartment in models E, G and I (Figure 1E,G,I). The direct differentiation path from HSCs to MEPs (Figure 1E,G,I) was also supported by in vitro studies of Takano et al. (2004), who investigated colony forming units of LSK daughter and granddaughter cells. However, a separate study from (Forsberg et al., 2006) also investigated the lineage potential of FLT3+ LMPPs but found conflicting results, which instead support the classical model of hematopoiesis (model A, Figure 1A).
In another mouse study, a fraction of phenotypically defined HSCs was shown to express von Willebrand factor (vWF), a protein mainly expressed by platelets and endothelium (Månsson et al., 2007). The existence of a megakaryocyte-primed HSC subset was also experimentally investigated by Sanjuan-Pla et al. (2013) generating vWF-eGFP transgenic mice, isolating LSK $CD150^+$ $CD48^-$ $CD34^-$ HSCs with a high eGFP expression and transplanting them into irradiated mice. They found that vWF-eGFP+ HSCs were platelet biased, additionally contributing to other myeloid lineages whereas their lymphoid contribution was very marginal.
Models E and H furthermore include the direct differentiation path from MPPs to MEPs, which was suggested by Pronk et al. (2007) (Figure 1E,H). By studying the phenotypic, functional and molecular characteristics of myeloerythroid precursors, they identified MPPs which give rise to erythroid and megakaryocytic progeny through various intermediate stages.This finding is supported by human studies, in which BAH1 and CD71 were identified as erythroid and megakaryocytic differentiation markers within the CD34+ CD38- MPP compartment (Notta et al., 2016).

Sample collection and storage

Healthy BM samples were obtained from allogeneic donor BM filters or from femoral heads of patients undergoing hip replacement surgery. Written informed consent in accordance with the Declaration of Helsinki was obtained from all patients according to protocols approved by the ethics committee of the Technische Universität München (approval number 538/16).
Mononuclear cells were isolated by ficoll gradient density centrifugation. Cells were frozen in 10 % DMSO (Serva, Cat: 20385) and 90 % heat inactivated FCS (Merck, Cat: S0115), $5x10^7$ cells/ml) at -80°C using a freezing chamber maintaining a controlled freezing rate of approximately 1°C per minute. Samples were stored in a N2 biobank until further use.

Antibody staining and FACS-sorting

For the sorting procedure of HSCs, MNC cells were thawed and immediately placed into IMDM (1x) + GlutaMAX (Gibco, Cat: 31980-022). Dead cells were removed by density gradient centrifugation. MNC were washed with 2 ml PBS and centrifuged. For the ability to track cell divisions in later FACS analysis, pellets were mixed with 2 ml of 1 µM CellTrace™ Violet stain (ThermoFisher Scientific, Cat: C34557) in PBS (37°C) and incubated for 20 min at 37°C. The reaction was stopped by adding 10 ml ice-cold HF2 medium containing 1xHBSS (Gibco, Cat:14185-045), 2 % heat-inactivated FCS (Biochrom, Cat:S0115), 0.01 M HEPES (Gibco, Cat: 15630-056), and 100 U/ml Pen/Strep (Gibco, Cat: 15140-122). After incubating 5 min on ice, cells were centrifuged and antibody staining was performed. Cells were first incubated with biotin-coupled antibodies, including 1 µl of each anti-CD4 (BioLegend, Clone: RPA-T4, Cat: 300504), anti-CD8a (BioLegend, Clone: RPA-T8, Cat: 301004), anti-CD15 (BioLegend, Clone: H198, Cat: 323016), anti-CD19 (BioLegend, Clone: H1BT9, Cat: 302204), and anti-CD235a (eBioscience, Clone: HIR2, Cat: 13-9987-82). for 20 min, on ice in the dark and then centrifuged (1500 rpm, 5 min). Pellets were resuspended with 100 µl of fluorescence-coupled antibody mix, including 5 µl anti-CD34-FITC (BD, Clone: 581, Cat: 555821), 5µl anti-CD90-PE (eBioscience, Clone: 5E10, Cat: 12-0909-42), 5 µl anti-CD123-BV510 (BioLegend, Clone: 6H6, Cat: 306021), 2.5 µl anti-CD38-APC (BD, Clone: HB7, Cat: 345807), 2.5 µl anti-CD45RA-PE-Cy7 (BD, Clone: HI100, Cat: 560675), 1 µl CD45-PeCy5.5 (BioLegend, Clone: HI30, Cat: 304028), and 1 µl APC/Cy7-Streptavidin (BioLegend, Cat: 405208) and incubated for 40 min on ice and in the dark. Pellets were resuspended in 500 µl HF2 with 0.2 µg propidium iodide and filtered using a 40 µm cell strainer. The sorting procedure was performed on a BD FACSAria™ III equipped with 4 lasers (488 nm, 405 nm, 561 nm, 635 nm).

Analysis of cell compartments on days 1 to 7, cultured cells were harvested, centrifuged and antibody staining was performed as described for the sorting procedure. Additionally, 50 µl of Flow-Count Fluorospheres (Beckman Coulter, Cat: 7547053) were added. FACS analysis was performed on a BeckmanCoulter CyAn, equipped with 405nm, 488nm, and 633 nm lasers. Compensation and gating was performed using the FlowJo V10 software (FlowJo LLC, Ashland, OR). Cell divisions were estimated by the decreasing intensity of the CellTrace™ Violet fluorophore.

Cell culture

Sorted HSCs (Lin-CD34+CD38-CD90+CD45RA-) were cultured at a concentration of $2.5x10^3$ cells/ml in serum-free medium (80% IMDM(1x)+GlutaMAX (Gibco, Cat: 31980-022) and 20 % BIT9500 (StemCell Technologies, Cat: 09500)) freshly supplemented with 10µM 2-Mercaptoethanol (Gibco, Cat: 31350-010), 8 µg/ml Ciprofloxacin (CiproHEXAL 200mg/100ml), 4 µg/ml LDL (StemCell Technologies, Cat: 02698), 100 ng/ml SCF (R&D Systems, Cat: 255-SC), 100ng/ml FLT3-Ligand (R&D Systems, Cat 308-FK), 25 ng/ml TPO (R&D Systems, Cat: 288-TP), 10 ng/ml IL3 (R&D Systems, Cat: 203-IL), 10ng/ml IL6 (R&D Systems, Cat: 206-IL) , 50 ng/ml GM-CSF (R&D Systems, Cat: 215-GM), 50 ng/ml G-CSF (Filgrastim, Hexal), and 2U/ml erythropoietin (Janssen, PZN: 00878122). Cells were cultured at 37°C with 5 % CO_2.

Computational approach

To assess the plausibility of a set of previously suggested lineage hierarchies for healthy human hematopoiesis, we derived a computational modeling approach. Based on a selection of previously suggested lineage hierarchies, we mechanistically modeled the cell differentiation dynamics. In the following, we consider the cell types hematopoietic stem cells (HSCs), multipotent progenitor cells (MPPs), common myeloid progenitors (CMPs), multipotent lymphocyte progenitors (MLPs), megakaryocyte erythrocyte progenitors (MEPs), granulocyte monocyte progenitors (GMPs) and mature and late progenitors (M) as species $\mathcal{S} = \{HSC, MPP, MLP, CMP, GMP, MEP, M\}$, for which division distributions and counts were experimentally observed at time points $t_0, ..., t_{n_t}$ (Figure S1B-C). We consider a set of 10 biologically motivated lineage hierarchies (Figure 1A-J) and derive a mathematical model for each of them (Figure 2A). The 10 models are then used to analyse which lineage hierarchies are plausible and which ones can be rejected based on our experimental data.

Mechanistic models of cell differentiation dynamics

To mechanistically model human blood cell production, we compiled lineage hierarchies suggested in the literature. Each hierarchy is a biochemical reaction network in which the species $\mathcal{S}$ are given by the cell types observed in the experiment. Cell differentiation, proliferation, and death are defined as cell type specific reactions. For MLPs the outflux reaction is defined as net differentiation and describes differentiation combined with cell death to ensure structural parameter identifiability. Note that each model considers the same proliferation and cell death reactions but a different set of differentiation reactions (Table 2). These can be derived from the respective model scheme of the set of plausible lineage hierarchies (Figure 1A-J) and are given by the cell types observed in the experiment. Cell differentiation, proliferation, and death are defined as cell type specific reactions (Table 1), where reactions $R_1 - R_7$ and $R_{22} - R_{27}$ which describe proliferation and cell death are present in all models. The model complexity, which corresponds to the number of reaction rates varies between models A-J (Table 2). The respective ODE systems derived from the reactions above describe the evolution of the cell concentrations over time for each compartment of a particular model. Differentiation and cell death reduces and proliferation increases the number of cells within the compartment proportionally to the cell concentration of this compartment at time t. The ODE system for model A without considering division compartments or intermediate states is given by

$$
\begin{aligned}
\dot{x}_1 &:= \frac{d[HSC]}{dt} = -(\alpha_{HSC\to MPP} - \beta_{HSC} + \gamma_{HSC})[HSC] \\
\dot{x}_2 &:= \frac{d[MPP]}{dt} = \alpha_{HSC\to MPP}[HSC] - (\alpha_{MPP\to CMP} + \alpha_{MPP\to MLP} - \beta_{MPP} + \gamma_{MPP})[MPP] \\
\dot{x}_3 &:= \frac{d[MLP]}{dt} = \alpha_{MPP\to MLP}[MPP] - (\alpha_{MLP\to\ldots} - \beta_{MLP})[MLP] \\
\dot{x}_4 &:= \frac{d[CMP]}{dt} = \alpha_{MPP\to CMP}[MPP] - (\alpha_{CMP\to GMP} + \alpha_{CMP\to MEP} - \beta_{CMP} - \gamma_{CMP})[CMP] \\
\dot{x}_5 &:= \frac{d[GMP]}{dt} = \alpha_{CMP\to GMP}[CMP] - (\alpha_{GMP\to M} - \beta_{GMP} + \gamma_{GMP})[GMP] \\
\dot{x}_6 &:= \frac{d[MEP]}{dt} = \alpha_{CMP\to MEP}[CMP] - (\alpha_{MEP\to M} - \beta_{MEP} + \gamma_{MEP})[MEP] \\
\dot{x}_7 &:= \frac{d[M]}{dt} = \alpha_{GMP\to M}[GMP] + \alpha_{MEP\to M}[MEP] - (\beta_M + \gamma_{MEP})[M]
\end{aligned} \tag{1}
$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0$.
In general, the ODE system is given by

$$
\dot{x}_j := \frac{dS_j}{dt} = \sum_{i\in I_j} \alpha_{i\to j} \cdot S_i(t) + \left(\beta_j - \gamma_j - \sum_{o\in O_j} \alpha_{j\to o}\right) \cdot S_j(t), \tag{2}
$$

$\forall j = 1, \ldots, |\mathcal{S}|$ where I_j is the set of influx compartments and O_j the set of outflux compartments of the respective species $S_j \in \mathcal{S}$ and the initial condition is given by $\mathbf{x}(0) = \mathbf{x}_0$.
Incorporating the information of the number of cell divisions N_{div}, the ODE system is expanded by introducing additional states which indicate not only the cell type but also the number of divisions occurring within the time interval of interest
$[t_0, t_i^{obs}]$ (Figure 2A). Hence, each ODE describes the evolution of cell abundances of species $S_j \in \mathcal{S}$ that divided $n_{div} \in \{0, \ldots, N_{div}\}$ times over time, which is denoted by $\frac{d[S_{j,n_{div}}]}{dt}$. This leads to an ODE system of $N_{div} \cdot n_c$ equations, where $n_c = 7$ is the number of cell type compartments. It is given by

$$
\begin{aligned}
\dot{x}_{(j-1)\cdot N_{div}+1} &:= \frac{dS_{j,0}(t)}{dt} = \sum_{i\in I_j} \alpha_{i\to j} \cdot S_{i,0}(t) + \left(\beta_j - \gamma_j - \sum_{o\in O_j} \alpha_{j\to o}\right) \cdot S_j(t) \\
\dot{x}_{(j-1)\cdot N_{div}+n_{div}+1} &:= \frac{dS_{j,n_{div}}(t)}{dt} = 2\beta_j \cdot S_{j,n_{div}-1}(t) + \sum_{i\in I_j} \alpha_{i\to j} \cdot S_{i,n_{div}}(t) - \left(\beta_j + \gamma_j + \sum_{o\in O_j} \alpha_{j\to o}\right) \cdot S_{j,n_{div}}(t) \\
\dot{x}_{j\cdot N_{div}} &:= \frac{dS_{j,N_{div}}(t)}{dt} = 2\beta_j \cdot S_{j,N_{div}-1}(t) + \sum_{i\in I_j} \alpha_{i\to j} \cdot S_{i,N_{div}}(t) + \left(\beta_j - \gamma_j - \sum_{o\in O_j} \alpha_{j\to o}\right) \cdot S_{j,N_{div}}(t)
\end{aligned} \tag{3}
$$

$\forall j = 1, \ldots, |\mathcal{S}|$ and $n_{div} = 0, 1, \ldots, N_{div}$. The waiting time for a differentiation, proliferation, or death event is anti-proportional to the corresponding reaction rate and follows an exponential distribution $T \sim \exp(r)$, where $r \in \{\alpha_{(.)}, \beta_{(.)}, \gamma_{(.)}\}$. This is in contrast with the observation that the considered processes (differentiation, proliferation, and cell death) require a minimum time to be completed. To more accurately describe transition times between cell states, we introduced intermediate states and further expanded the model (Figure 2A). By introducing intermediate states, the waiting time to stay in a particular state corresponds to the sum of exponentially distributed waiting times of its n_{IS} intermediate states and is thereby per definition Erlang(n_{IS},r) distributed (Matis and Wehrly, 1990).
The model allows to describe up to N_{div} division compartments per cell type compartment and if cells divide more often (more than N_{div} times), they accumulate in the N_{div}-compartment of the respective species S_j. In total it consists of

$$
N_{eq} = \begin{cases} N_{div} \cdot \left(\left(n_{IS} \cdot \sum_{j=1}^{|\mathcal{S}|} (n_j^{out} + 2)\right) + 1\right), & n_{IS} > 1 \\ N_{div} \cdot |\mathcal{S}|, & n_{IS} = 1 \end{cases} \tag{4}
$$

equations, where n_j^{out} is the number of outfluxes of compartment S_j and n_{IS} is the number of intermediate states within each compartment S_j.
Each ODE describes the time evolution of the number of cells of species $S_j \in \mathcal{S}$ that divided $n_{div} \in \{0, ..., N_{div}\}$ times and are in the k-th proliferation intermediate state, the l-th differentiation and the m-th cell death intermediate state, which is denoted by $\frac{d[S^{(k,l,m)}_{j,n_{div},i_o}(t)]}{dt}$ where $n_{div} = 0, ..., N_{div}, \quad i_o = 1, ..., n_j^{out}, \quad k, l, m = 1, ..., n_{IS}$ and $j = 1, ..., |\mathcal{S}|$.
The ODE system is given by

$$
\begin{aligned}
\frac{d[S^{(0,0,0)}_{j,n_{div},1}]}{dt} &:= \begin{cases}
n_{IS} \cdot \left(\sum_{i \in I_j} \alpha_{i \to j} \cdot S^{(0,n_{IS},0)}_{i,n_{div},1}(t) \right. \\
\quad \left. -(\beta_j + \sum_{o \in O_j} \alpha_{j \to o} + \gamma_j) \cdot S^{(0,0,0)}_{j,n_{div},1}(t) \right) & , \text{ if } n_{div} = 0 \\
n_{IS} \cdot \left(\sum_{i \in I_j} \alpha_{i \to j} \cdot S^{(0,n_{IS},0)}_{i,n_{div},1}(t) \right. \\
\quad -(\beta_j + \sum_{o \in O_j} \alpha_{j \to o} + \gamma_j) \cdot S^{(0,0,0)}_{j,n_{div},1}(t) \\
\quad \left. +\beta_j \cdot S^{(n_{IS},0,0)}_{j,n_{div}-1,1}(t) \right) & , \text{ if } n_{div} \in \{1, ..., N_{div}-1\} \\
n_{IS} \cdot \left(\sum_{i \in I_j} \alpha_{i \to j} \cdot S^{(0,n_{IS},0)}_{i,n_{div},1}(t) \right. \\
\quad -(\beta_j + \sum_{o \in O_j} \alpha_{j \to o} + \gamma_j) \cdot S^{(0,0,0)}_{j,n_{div},1}(t) \\
\quad \left. +\beta_j \cdot S^{(n_{IS},0,0)}_{j,n_{div}-1,1}(t) + 2 \cdot \beta_j \cdot S^{(n_{IS},0,0)}_{j,n_{div},1}(t) \right) & , \text{ if } n_{div} = N_{div}
\end{cases} \\
&= \dot{x}_{(j-1)\cdot(N_{div}\cdot(2\cdot n_{IS}+1)+\sum_{c=1}^{j-1} n_c^{out}\cdot n_{IS}\cdot N_{div}+(n_{div}+1)} \\
\frac{d[S^{(k,0,0)}_{j,n_{div},1}(t)]}{dt} &:= n_{IS} \cdot \beta_j \left(S^{(k-1,0,0)}_{j,n_{div},1}(t) - S^{(k,0,0)}_{j,n_{div},1}(t) \right) \\
&= \dot{x}_{(j-1)\cdot(N_{div}\cdot(2\cdot n_{IS}+1)+\sum_{c=1}^{j-1} n_c^{out}\cdot n_{IS}\cdot N_{div}+(n_{div}+k+1)} \\
\frac{d[S^{(0,0,m)}_{j,n_{div},1}]}{dt} &:= n_{IS} \cdot \gamma_j \left(S^{(0,0,m-1)}_{j,n_{div},1}(t) - S^{(0,0,m)}_{j,n_{div},1}(t) \right) \\
&= \dot{x}_{(j-1)\cdot(N_{div}\cdot(2\cdot n_{IS}+1)+\sum_{c=1}^{j-1} n_c^{out}\cdot n_{IS}\cdot N_{div}+(n_{div}+m+n_{IS}+1)} \\
\frac{d[S^{(0,l,0)}_{j,n_{div},i_o}]}{dt} &:= n_{IS} \cdot \alpha_{j \to i_o} \left(S^{(0,l-1,0)}_{j,n_{div},i_o}(t) - S^{(0,l,0)}_{j,n_{div},i_o}(t) \right) \\
&= \dot{x}_{(j-1)\cdot(N_{div}\cdot(2\cdot n_{IS}+1)+\sum_{c=1}^{j-1} n_c^{out}\cdot n_{IS}\cdot N_{div}+(n_{div}+l+i_o+2\cdot n_{IS}+1)}
\end{aligned} \tag{5}
$$

and initial condition $\mathbf{x}(0) = \mathbf{x}_0$.
Note that for both model extensions, the number of states increases, but the number of parameters stays constant.

Parameter inference

Our models contain between 29 and 35 unknown parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_{n_\theta})$, which are the reaction rates $\beta_{S_j}, \gamma_{S_j}, \alpha_{S_{j1} \to S_{j2}}$, where $S_j, S_{j1}, S_{j2} \in \mathcal{S}$ and the initial conditions are given by $\mathbf{x}(0) = \mathbf{x}_0(\theta)$, where

$$
\mathbf{x}_0(\theta) = S^{(k,l,m)}_{j,n_{div},i_o}(0) = \begin{cases} \theta_j, & \text{if } k = l = m = 0 \text{ and } n_{div} = i_o = 1 \\ 0, & \text{otherwise} \end{cases} \tag{6}
$$

for $j = 1, ..., |\mathcal{S}|$, and $i_{div} = 0, ..., 6$. These parameters are estimated by minimizing the weighted difference between observed and modeled cell counts by applying maximum likelihood estimation.
Let $\mathcal{M}(\theta)$ be a particular model consisting of dynamics $\dot{\mathbf{x}} = f(\mathbf{x}, \theta)$ and model observations $y^{\mathcal{M}} = h(x, \theta)$:

$$\mathscr{M}(\theta): \left\{ \begin{array}{l} \dot{\mathbf{x}} = f(\mathbf{x},\theta) = \left\{ \frac{d[S^{(k,l,m)}_{j,n_{div},i_o}](t)}{dt} \right\}, \ \mathbf{x}(0) = \mathbf{x}_0(\theta), \\ \mathbf{y}^{\mathscr{M}} = h(x,\theta) = \left\{ \sum_{k,l,m=0}^{n_{IS}} \sum_{i_o=1}^{n^{out}_j} [S^{(k,l,m)}_{j,n_{div},i_o}](t) \right\} \end{array} \right\}, \tag{7}$$

where $j = 1, ..., |\mathcal{S}|$, $k, l, m = 1, ..., n_{IS}$, $i_o = 1, ..., n^j_{out}$, $i_{div} = 1, ..., N_{div}$ and $w \in \{1, 2\}$ and let

$$\mathscr{D} = \left\{ t_s, y^{\mathscr{D}}_{j,n_{div}}(t_s) \right\}_{s=1,..,n_t,\ j=1,...,|\mathscr{S}|,\ n_{div}=1,...,N_{div}} \tag{8}$$

be the data (Figure S1B-C). Here $y^{\mathcal{D}}_{j,i_{div}}(t_s)$ denotes the vector of observed cell counts of species $j = 1, ..., |\mathcal{S}|$ that divided $i_{div} = 0, ..., N_{div}$ times at time t_s of a particular individual. For parameter estimation, we assumed the observations $y^{\mathcal{D}}_{j,i_{div}}(t_s)$ are subject to multiplicative log-normally distributed measurement noise

$$\begin{aligned} y^{\mathscr{D}}_{j,n_{div}}(t_s) &= y^{\mathscr{M}}_{j,n_{div}}(t_s) \cdot \nu, \text{ with } \nu \sim log\mathcal{N}(0, \sigma^2_{j,n_{div}}) \\ \log(y^{\mathscr{D}}_{j,n_{div}}(t_s)) &= \log(y^{\mathscr{M}}_{j,n_{div}}(t_s)) + \varepsilon, \text{ with } \varepsilon \sim \mathcal{N}(0, \sigma^2_{j,n_{div}}) \end{aligned} \tag{9}$$

due to counting errors (i.e. technical error of the FACS machine, see Methods) or false cell type assignment while processing raw FACS data by gating.

In order to assess how well $\mathcal{M}$ fits the experimental data for a certain set of parameters θ, the log-likelihood $\ell_D(\theta)$ is calculated according to the assumed multiplicative log-normally distributed measurement noise

$$\ell_D(\theta) = -\frac{1}{2} \sum_{s=1}^{n_t} \sum_{n_{div}=1}^{N_{div}} \sum_{j=1}^{|\mathscr{S}|} \log(2\pi\sigma^2_{j,n_{div}}) + \left(\frac{\left(\log(\mathbf{y}^{\mathscr{D}}_{j,n_{div}}(t_s) + 1) - \log(\mathbf{y}^{\mathscr{M}}_{j,n_{div}}(t_s, \theta) + 1) \right)^2}{\sigma^2_{j,n_{div}}} \right). \tag{10}$$

In order to estimate the unknown parameter vector, the optimization problem

$$\theta^{ML} = \underset{\substack{\theta \\ \text{subject to } \mathscr{M}}}{\text{argmax}} \ \ell_D(\theta) \tag{11}$$

is solved using multi-start local hierarchical optimization (Loos et al., 2018) with trust-region-reflective algorithm and $n_{MS} = 1000$ multi starts. With the hierarchical optimization approach $\sigma^2_{j,i_{div}}, j = 1, ..., |\mathcal{S}|, i_{div} = 1, ..., n_{div}$ is analytically calculated each time the log-Likelihood function is evaluated. The noise parameter is therefore not part of the parameter vector. The starting values $(\theta^{start}_i)_{i=1,...,n_{MS}}$ (initial parameter vectors) are determined according to latin hypercube sampling (Eliáš and Vořechovský, 2016; Loos et al., 2018). The resulting optimal parameter is observed at the highest ℓ_D value. To ensure that the optimization procedure converged, we checked if this best log-likelihood value is observed several times for different starting values. The boundaries for parameter values can be found in Table S2. We used the MATLAB toolboxes AMICI (Fröhlich et al., 2017) for model definition and PESTO (Stapor et al., 2018) for parameter inference.

Structural identifiability of candidate models

A model is structurally identifiable if it is possible to determine parameter values from measurements of the model output. A structural identifiability analysis was performed using a method introduced by (Villaverde and Banga, 2017) and the MATLAB toolbox STRIKE-GOLDD (Villaverde et al., 2019; Villaverde and Banga, 2017). This method evaluates the change of the observables along the model dynamics by calculating Lie derivatives. If the change in the Lie derivatives with parameters $\theta_m,\ m = 1, ..., n_\theta$ leads to linear dependent vectors, at least one parameter is structurally non-identifiable. This analysis can thereby reveal which parameters are identifiable and which ones are non-identifiable for the different hierarchies if one assumes ideal noise-free data with a large sample size (Table 3).

Practical identifiability of inferred parameters

To determine the 95% confidence intervals for each parameter and assess the practical identifiability, we calculated the profile likelihood and used it to calculate confidence intervals for the parameters

(Kreutz et al., 2013). We defined a parameter as practically identifiable from a specific sample if its 95% profile-likelihood-based confidence interval is non-overlapping with its lower and upper parameter boundaries (Supplemental Table 2).

Computation time calculation
For every lineage hierarchy and sample, optimization of 1000 multi starts was run in parallel on a machine with two Intel® Xeon® Silver 4214 12-Core (2,2 GHz, 3,2 GHz Turbo, 16,5M L3 Cache) processors (in total 24 workers) and the time for performing parameter inference was tracked.

Model selection
The parameter estimation was performed individually for all individuals and models A-J. We compared and ranked the different models based on their Bayesian Information Criterion (BIC) value. The BIC of model $\mathcal{M}_q(\boldsymbol{\theta})$
is defined as

$$BIC_q := -2\log(P(\mathcal{D}|\theta_q^{ML})) + n_{\theta_q} \cdot \log(n_{obs}), \tag{12}$$

where n_{θ_q} describes the number of parameters of model $\mathcal{M}_q(\boldsymbol{\theta})$ and n_{obs} the number of observations used for model fitting (Neath and Cavanaugh, 2012). The Akaike Information Criterion (AIC, (Akaike, 1992)) of $\mathcal{M}_q(\boldsymbol{\theta})$ is given by

$$AIC_q := -2\ell_{\mathcal{D}}(\theta_q^{\hat{ML}}) + 2n_\theta. \tag{13}$$

Calculation of the AIC or BIC scores provides a ranking of all considered models in which the best performing model is the one with the lowest score

$$\mathcal{M}_{\text{rank 1}} = \mathcal{M}_q, \text{ if } C_q = \min_{m\in\{1,\dots,n_M\}}(C), \tag{14}$$

To derive the set of plausible and implausible models the differences

$$\Delta_j^C := C_j - \min_{m\in\{1,\dots,n_M\}}(C), \tag{15}$$

With C being the AIC or BIC score and j the index of the respective model, are calculated. Model $\mathcal{M}_j$ can be rejected if $\Delta_i^{Score} > 10$ (Guthery et al., 2003). Consequently, for the comparison of two models $\mathcal{M}_1$ and $\mathcal{M}_2$, where $\mathcal{M}_2$ is of higher or equal complexity than $\mathcal{M}_1$ the null hypothesis $\mathcal{H}_0$: "Model $\mathcal{M}_1$ is true." can be rejected if $\Delta^C_{\mathcal{M}_1,\mathcal{M}_2} = C_{\mathcal{M}_1} - C_{\mathcal{M}_2} > 10$. As $\Delta^C_{\mathcal{M}_1,\mathcal{M}_2}$ are given by

$$\begin{aligned}\Delta^{AIC}_{\mathcal{M}_1,\mathcal{M}_2} &= AIC_{\mathcal{M}_1} - AIC_{\mathcal{M}_2} = T(\mathcal{M}_1,\mathcal{M}_2) - 2(n_\theta^{\mathcal{M}_2} - n_\theta^{\mathcal{M}_1})\\ \Delta^{BIC}_{\mathcal{M}_1,\mathcal{M}_2} &= T(\mathcal{M}_1,\mathcal{M}_2) - log(n_{obs})(n_\theta^{\mathcal{M}_2} - n_\theta^{\mathcal{M}_1}),\end{aligned} \tag{16}$$

for AIC and BIC, model $\mathcal{M}_1$ can be rejected if the respective test statistic $T(\mathcal{M}_1,\mathcal{M}_2)$ exceeds critical values of

$$\begin{aligned}v^{AIC}_{crit} &= 10 + 2(n_\theta^{\mathcal{M}_2} - n_\theta^{\mathcal{M}_1}),\\ v^{BIC}_{crit} &= 10 + log(n_{obs})(n_\theta^{\mathcal{M}_2} - n_\theta^{\mathcal{M}_1}).\end{aligned} \tag{17}$$

In silico analysis
To test the implementation, robustness and accuracy of our model selection approach, and if it is able to identify plausible models, we performed an in silico analysis. For each considered lineage hierarchy we implemented the extended model as described in (1) and simulated with a realistic test parameter 3 samples which correspond to the measured differentiation dynamics of 3 individuals which underlie different noise levels. Based on the inference result from fitting experimental data, we defined realistic model-specific test parameters $\theta^{test}_{\mathcal{M}_q}, q = 1,\dots,10$ and simulated data from each model $\mathcal{M}_q \in \{A,B,C,D,E,F,G,H,I,J\}$ with $\theta^{test}_{\mathcal{M}_q}$. We then performed MLE with all models on all in silico data sets and observed which models perform best, are plausible, or can be rejected. Subsequently we fitted each of the 10*3 samples with every considered model and observed if each test parameter lies within the 95% confidence interval of the corresponding optimized parameter, as well as the distance between

test and optimal parameter (Figure S3B), the model fit (Figure S3A, 4A,C) and the BIC scores of each model for each simulated data set for varying noise levels (Figure 4B,D).

Supplemental References

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., Jacobsen, S.E.W., 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell 121, 295–306.

Akaike, H., 1992. Information Theory and an Extension of the Maximum Likelihood Principle. Springer Series in Statistics. https://doi.org/10.1007/978-1-4612-0919-5_38

Doulatov, S., Notta, F., Eppert, K., Nguyen, L.T., Ohashi, P.S., Dick, J.E., 2010. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. Nat. Immunol. 11, 585–593.

Doulatov, S., Notta, F., Laurenti, E., Dick, J.E., 2012. Hematopoiesis: a human perspective. Cell Stem Cell 10, 120–136.

Eliáš, J., Vořechovský, M., 2016. Modification of the Audze–Eglājs criterion to achieve a uniform distribution of sampling points. Advances in Engineering Software. https://doi.org/10.1016/j.advengsoft.2016.07.004

Forsberg, E.C., Serwold, T., Kogan, S., Weissman, I.L., Passegué, E., 2006. New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. Cell 126, 415–426.

Fröhlich, F., Kaltenbacher, B., Theis, F.J., Hasenauer, J., 2017. Scalable Parameter Estimation for Genome-Scale Biochemical Reaction Networks. PLoS Comput. Biol. 13, e1005331.

Galy, A., Travis, M., Cen, D., Chen, B., 1995. Human T, B, natural killer, and dendritic cells arise from a common bone marrow progenitor cell subset. Immunity 3, 459–473.

Giebel, B., Zhang, T., Beckmann, J., Spanholtz, J., Wernet, P., Ho, A.D., Punzel, M., 2006. Primitive human hematopoietic cells give rise to differentially specified daughter cells upon their initial cell division. Blood 107, 2146–2152.

Goardon, N., Marchi, E., Atzberger, A., Quek, L., Schuh, A., Soneji, S., Woll, P., Mead, A., Alford, K.A., Rout, R., Chaudhury, S., Gilkes, A., Knapper, S., Beldjord, K., Begum, S., Rose, S., Geddes, N., Griffiths, M., Standen, G., Sternberg, A., Cavenagh, J., Hunter, H., Bowen, D., Killick, S., Robinson, L., Price, A., Macintyre, E., Virgo, P., Burnett, A., Craddock, C., Enver, T., Jacobsen, S.E.W., Porcher, C., Vyas, P., 2011. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell 19, 138–152.

Görgens, A., Radtke, S., Möllmann, M., Cross, M., Dürig, J., Horn, P.A., Giebel, B., 2013. Revision of the human hematopoietic tree: granulocyte subtypes derive from distinct hematopoietic lineages. Cell Rep. 3, 1539–1552.

Guthery, F.S., Burnham, K.P., Anderson, D.R., 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. The Journal of Wildlife Management. https://doi.org/10.2307/3802723

Hao, Q.L., Zhu, J., Price, M.A., Payne, K.J., Barsky, L.W., Crooks, G.M., 2001. Identification of a novel, human multilymphoid progenitor in cord blood. Blood 97, 3683–3690.

Kreutz, C., Raue, A., Kaschek, D., Timmer, J., 2013. Profile likelihood in systems biology. FEBS J. 280, 2564–2571.

Loos, C., Krause, S., Hasenauer, J., 2018. Hierarchical optimization for the efficient parametrization of ODE models. Bioinformatics 34, 4266–4273.

Majeti, R., Park, C.Y., Weissman, I.L., 2007. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell 1, 635–645.

Månsson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thorén, L., Adolfsson, J., Buza-Vidas, N., Qian, H., Soneji, S., Enver, T., Sigvardsson, M., Jacobsen, S.E.W., 2007. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. Immunity 26, 407–419.

Manz, M.G., Miyamoto, T., Akashi, K., Weissman, I.L., 2002. Prospective isolation of human clonogenic common myeloid progenitors. Proc. Natl. Acad. Sci. U. S. A. 99, 11872–11877.

Matis, J.H., Wehrly, T.E., 1990. Generalized stochastic compartmental models with Erlang transit times. J. Pharmacokinet. Biopharm. 18, 589–607.

Neath, A.A., Cavanaugh, J.E., 2012. The Bayesian information criterion: background, derivation, and

applications. Wiley Interdisciplinary Reviews: Computational Statistics. https://doi.org/10.1002/wics.199
Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., Dick, J.E., 2011. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science 333, 218–221.
Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., McPherson, J.D., Stein, L.D., Dror, Y., Dick, J.E., 2016. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science 351, aab2116.
Ostendorf, B.N., Flenner, E., Flörcken, A., Westermann, J., 2018. Phenotypic characterization of aberrant stem and progenitor cell populations in myelodysplastic syndromes. PLoS One 13, e0197823.
Pang, W.W., Price, E.A., Sahoo, D., Beerman, I., Maloney, W.J., Rossi, D.J., Schrier, S.L., Weissman, I.L., 2011. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. Proc. Natl. Acad. Sci. U. S. A. 108, 20012–20017.
Pronk, C.J.H., Rossi, D.J., Månsson, R., Attema, J.L., Norddahl, G.L., Chan, C.K.F., Sigvardsson, M., Weissman, I.L., Bryder, D., 2007. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. Cell Stem Cell 1, 428–442.
Reynaud, D., Lefort, N., Manie, E., Coulombel, L., Levy, Y., 2003. In vitro identification of human pro-B cells that give rise to macrophages, natural killer cells, and T cells. Blood 101, 4313–4321.
Rossi, D.J., Jamieson, C.H.M., Weissman, I.L., 2008. Stems cells and the pathways to aging and cancer. Cell 132, 681–696.
Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., Chowdhury, O., Stenson, L., Lutteropp, M., Green, J.C.A., Facchini, R., Boukarabila, H., Grover, A., Gambardella, A., Thongjuea, S., Carrelha, J., Tarrant, P., Atkinson, D., Clark, S.-A., Nerlov, C., Jacobsen, S.E.W., 2013. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature 502, 232–236.
Stapor, P., Weindl, D., Ballnus, B., Hug, S., Loos, C., Fiedler, A., Krause, S., Hroß, S., Fröhlich, F., Hasenauer, J., Wren, J., 2018. PESTO: Parameter EStimation TOolbox. Bioinformatics 34, 705–707.
Takano, H., Ema, H., Sudo, K., Nakauchi, H., 2004. Asymmetric division and lineage commitment at the level of hematopoietic stem cells: inference from differentiation in daughter cell and granddaughter cell pairs. J. Exp. Med. 199, 295–302.
Villaverde, A.F., Banga, J.R., 2017. Dynamical compensation and structural identifiability of biological models: Analysis, implications, and reconciliation. PLoS Comput. Biol. 13, e1005878.
Villaverde, A.F., Evans, N.D., Chappell, M.J., Banga, J.R., 2019. Input-Dependent Structural Identifiability of Nonlinear Systems. IEEE Control Syst. Lett. 3, 272–277.