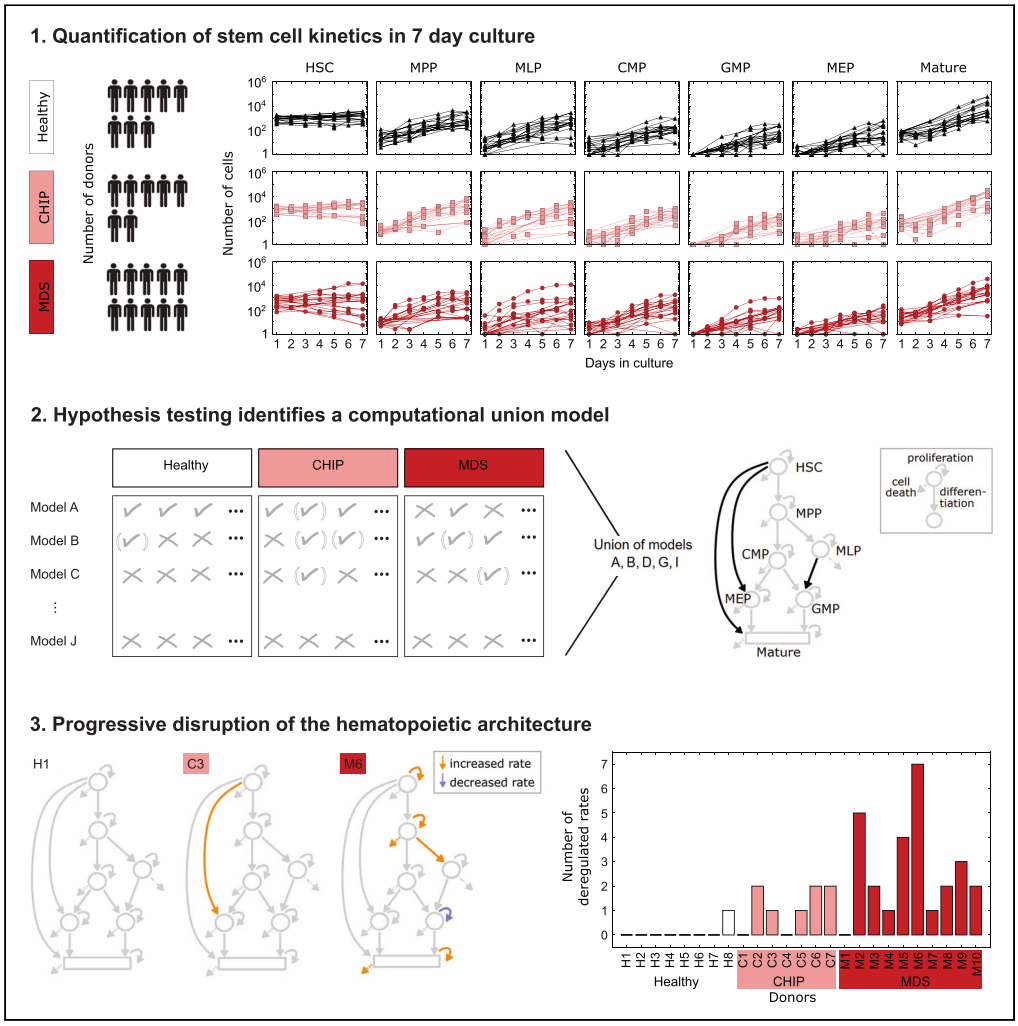
iScience

CellPress
OPEN ACCESS

Article

Progressive disruption of hematopoietic architecture from clonal hematopoiesis to MDS



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

carsten.marr@helmholtz-muenchen.de (C.M.)
katharina.goetze@tum.de (K.S.G.)

Highlights

We analyzed division- and time-resolved HSC kinetics in healthy, CHIP, and MDS

Computational modeling reveals a disrupted hematopoietic hierarchy in CHIP and MDS

Skewing of HSC toward MEP is observed in CHIP

Alterations in HSC kinetics is the defining feature of MDS

Buck et al., iScience 26, 107328
August 18, 2023
https://doi.org/10.1016/j.isci.2023.107328

iScience

CellPress
OPEN ACCESS

Article

Progressive disruption of hematopoietic architecture from clonal hematopoiesis to MDS

Michèle C. Buck,[1,10] Lisa Bast,[2,3,9,10] Judith S. Hecker,[1] Jennifer Rivière,[1] Maja Rothenberg-Thurley,[4] Luisa Vogel,[1] Dantong Wang,[2,3] Immanuel Andrä,[5] Fabian J. Theis,[2,3] Florian Bassermann,[1,6] Klaus H. Metzeler,[4,7] Robert A.J. Oostendorp,[1] Carsten Marr,[2,6,8,*] and Katharina S. Götze[1,6,11,*]

SUMMARY

Clonal hematopoiesis of indeterminate potential (CHIP) describes the age-related acquisition of somatic mutations in hematopoietic stem/progenitor cells (HSPC) leading to clonal blood cell expansion. Although CHIP mutations drive myeloid malignancies like myelodysplastic syndromes (MDS) it is unknown if clonal expansion is attributable to changes in cell type kinetics, or involves reorganization of the hematopoietic hierarchy. Using computational modeling we analyzed differentiation and proliferation kinetics of cultured hematopoietic stem cells (HSC) from 8 healthy individuals, 7 CHIP, and 10 MDS patients. While the standard hematopoietic hierarchy explained HSPC kinetics in healthy samples, 57% of CHIP and 70% of MDS samples were best described with alternative hierarchies. Deregulated kinetics were found at various HSPC compartments with high inter-individual heterogeneity in CHIP and MDS, while altered HSC rates were most relevant in MDS. Quantifying kinetic heterogeneity in detail, we show that reorganization of the HSPC compartment is already detectable in the premalignant CHIP state.

INTRODUCTION

Clonal hematopoiesis is an age-related condition characterized by acquired somatic mutations in hematopoietic stem/progenitor cells (HSPC) leading to clonal expansion of blood cells. While individuals with clonal hematopoiesis have normal blood counts and no evidence of hematologic disease,[1,2] detected mutations are often driver mutations in myeloid malignancies,[3] and clonal hematopoiesis predispose patients to develop acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS),[1,2] both diseases that increase with age and originate in HSPC. The term clonal hematopoiesis of indeterminate potential (CHIP) was coined to distinguish clonal hematopoiesis from malignant myeloid states such as myelodysplastic syndromes (MDS) or acute myeloid leukemia (AML), and is currently defined by a mutational burden (variant allele frequency, VAF) of $\geq$2%.[4] Myeloid-biased differentiation within the HSPC compartment in healthy individuals is a feature of aging,[5,6] and is also observed in MDS.[7–10] Recent analyses indicated that in CHIP a skewed differentiation potential[11] might contribute to changes in bone marrow (BM) composition. However, the underlying mechanisms leading to clonal expansion of mutated HSPC in CHIP, ranging from dysregulated differentiation and self-renewal rates to reorganization of the hematopoietic differentiation hierarchy are incompletely understood and may vary significantly depending on the type of mutation. We recently showed that clonal hematopoiesis is highly prevalent (50%) in the BM compartment of individuals without hematologic disease undergoing total hip replacement surgery, with a VAF ranging from 1.0% to 32.7%, (median 2.7%).[12] The availability of these BM samples allowed us to address the unresolved question whether the presence of a CHIP clone in the BM influences proliferation kinetics and cell fate decisions of HSPC compared to age-matched healthy controls and whether apparent changes are similar to those observed in MDS.

For healthy hematopoiesis, the differentiation hierarchy has long been defined by the classical model, in which hematopoietic stem cells (HSC) at the apex give rise to multipotent progenitors (MPP), which can differentiate into common myeloid progenitors (CMP) and multipotent lymphoid progenitors (MLP). CMP can further differentiate into megakaryocyte erythrocyte progenitors (MEP), or into granulocyte monocyte progenitors (GMP).[13–15] In recent years, the classical model of hematopoiesis has been

[1]Technical University of Munich (TUM), School of Medicine, Department of Medicine III, Munich, Germany

[2]Helmholtz Zentrum München–German Research Center for Environmental Health, Institute of Computational Biology, Neuherberg, Germany

[3]Technical University of Munich (TUM), Department of Mathematics, Chair of Mathematical Modeling of Biological Systems, Garching, Germany

[4]University Hospital, Ludwig-Maximilians-University, Department of Medicine III, Laboratory for Leukemia Diagnostics, Munich, Germany

[5]Technical University of Munich, Microbiology Institute, Munich, Germany

[6]German Cancer Consortium (DKTK), Heidelberg, Partner Site Munich, Germany

[7]University Hospital Leipzig, Department of Hematology and Cell Therapy, Leipzig, Germany

[8]Helmholtz Zentrum München–German Research Center for Environmental Health, Institute of AI for Health, Neuherberg, Germany

[9]Present address: Karolinska Institutet, Department of Medical Biochemistry and Biophysics, Laboratory of Molecular Neurobiology, Solna, Sweden

[10]These author contributed equally

[11]Lead contact

*Correspondence: carsten.marr@helmholtz-muenchen.de (C.M.), katharina.goetze@tum.de (K.S.G.)

https://doi.org/10.1016/j.isci.2023.107328

challenged by several studies proposing divergent hierarchies (summarized in[15–17]). By comparing these competing hierarchies quantitatively to HSPC kinetics, we recently showed that the classical model of hematopoiesis describes human healthy hematopoiesis best.[18] However, in CHIP and MDS, the plasticity of HSPC and transitions between cellular compartments may shift.

In fact, recent studies have shown a high level of subclonal diversity of HSC in MDS, suggesting a non-linear model of disease development characterized by parallel clonal evolution with distinct HSC subclones.[19] These leukemic HSC establish their own, unique hierarchical relationship by giving rise to further differentiated cells.[20] While it was initially assumed that leukemia-initiating cells are exclusively localized within the HSC compartment (defined as lin−CD34$^+$CD38$^-$CD90$^+$CD45RA− cells, see[20]), Ye et al.[21] showed that GMP are required for *in vivo* disease propagation in an AML mouse model, indicating that leukemic GMP have acquired self-renewal properties. These examples underscore the complex organization of the HSPC compartment in myeloid stem cell disorders, and the need to better characterize the hematopoietic hierarchy and its related kinetics. Most importantly, investigation of these characteristics in CHIP helps to uncover how clonal dominance is achieved over time and if the presence of a (usually small) CHIP clone is sufficient to reorganize the HSPC compartment toward a (pre-)leukemic stem cell compartment. Characterization of HSPC dynamics in CHIP is also a prerequisite for development of preventive or interventional approaches.

To this end, we assembled a unique cohort of BM samples from 7 individuals with CHIP, 10 MDS patients, and 8 age-matched healthy controls and combined division- and time-resolved FACS measurements of sorted and *in vitro* cultured HSC. Our *in vitro* approach allows us to observe cell type-specific abundances and division distributions at several time points. Equipped with a computational multi-compartment model with realistic waiting time distributions,[18] we are able to infer division, differentiation, and death rates for every observed progenitor cell type, allowing us to compare competing models of hematopoietic hierarchies and deconvolute kinetic alterations in CHIP and MDS. We show that alterations in kinetics and differentiation transitions can be readily detected in individuals with CHIP before any hematopoietic disorder becomes evident and that the hematopoietic architecture becomes progressively disrupted in the transition from CHIP to MDS.

RESULTS

Establishment of a cohort of CHIP and MDS bone marrow samples

Our CHIP cohort comprised BM from 7 individuals without evidence of hematologic disease-carrying mutations in *DNMT3A*, *SF3B1*, *NFE2*, *TP53*, or *KRAS* with VAF ranging from 1.3% to 27% (median 5%, Table 1). For comparison, we analyzed BM samples from 10 MDS patients: 5 low-risk MDS, 2 intermediate, 2 high, and 1 very high-risk MDS according to the revised international prognostic scoring system (IPSS-R[22]). Because MDS is a highly heterogeneous disease,[3,23] we focused mainly on MDS samples with mutations in either *ASXL1* or *SF3B1* (associated with poor vs. favorable prognosis, respectively)[24] to increase comparability and reduce variability between samples. These mutations are also among the most common mutations found in CHIP.[1,2,12] Additional mutations present in our MDS sample cohort were found in *TET2*, *DNMT3A*, *SRSF2*, *RUNX1*, *KRAS*, *SETBP1*, and *WT1* (Table 1). We complemented our cohort with 8 BM samples from healthy age-matched individuals without mutations (Table 1) that were partly reported in our previous study.[18]

Changes in HSPC compartment composition are detectable in CHIP and clearly skewed in MDS

As a starting point, we analyzed the composition of the HSPC compartment in unsorted mononuclear cells from BM samples of our cohort by multiparameter flow cytometry ("BM snapshot analysis") based on the expression of lineage markers and CD45, CD34, CD38, CD45RA, CD123, and CD90 (Thy-1)[14,25,26] as described in Bast et al. 2021. We assessed the abundance and fraction of HSC (lin-CD34$^+$CD38$^-$CD90$^+$CD45RA-), MPP (lin-CD34$^+$CD38$^-$CD90$^-$CD45RA-), MLP (lin-CD34$^+$CD38$^-$CD90$^+$CD45RA+) CMP (lin-CD34$^+$CD38$^+$CD45RA-CD123+), GMP (lin-CD34$^+$CD38$^+$CD45RA + CD123+), MEP (lin-CD34$^+$CD38$^+$CD45RA-CD123-), and CD34$^-$ compartments following the gating strategy presented in Figure 1A. By normalizing cell counts to the lineage negative (lin-) compartment,[18] we observed a significant reduction of HSPC proportions in MDS compared to healthy BM ($p = 0.038$, Kruskal-Wallis test followed by Dunn's multiple comparison test, Figure 1B). This difference was driven by significantly decreased CMP and GMP fractions in MDS (Figure 1C). Previous observations on the BM composition in MDS vary, with some groups showing similar results[7,8] and others demonstrating increased GMP and MPP compartment sizes.[9] However, gating strategies and normalizing of cell counts were not always the same, making comparison difficult. Interestingly in our cohort, all CHIP HSPC compositions fell between healthy controls and MDS samples (Figure 1C).

Table 1. Characteristics of the analyzed cohort

ID	Disease/ Entity	Age	Gender	IPSS-R score	Karyotype	Gene	Variant_DNA	Variant_Prot	VAF Bulk (%)	VAF HSC (%)
H1*	Healthy	57	m							
H2	Healthy	62	f							
H3*	Healthy	63	m							
H4*	Healthy	63	m							
H5	Healthy	63	f							
H6	Healthy	68	f							
H7*	Healthy	70	m							
H8*	Healthy	76	f							
C1	CHIP	58	f			DNMT3A	c.2393T>A	p.Leu798His	4	NA
C2	CHIP	65	f			DNMT3A	c.2401A>G	p.Met801Val	12	NA
C3	CHIP	70	f			DNMT3A	c.1015-2A>G	p = ? splice-site Variante	5	NA
C4	CHIP	76	m			TP53	c.536A>G	p.His179Arg	4,5	NA
						DNMT3A	c.1949T>G	p.Leu650Arg	1,3	NA
C5	CHIP	79	f			DNMT3A	c.2644C>T	p.Arg882Cys	27<	NA
						NFE2	c.578_581del	p.Asn193Ilefs*12	25	NA
C6	CHIP	68	f			SF3B1	c.1873C>T	p.Arg625Cys	15	NA
C7	CHIP	72	f			KRAS	c.40G>A	p.Val14Ile	2.43	NA
M1	MDS-RS-MLD	60	m	low risk (2)	46,XY	SF3B1	c.2098A>G	p.Lys700Glu	33	49
						DNMT3A	c.1906G>A	p.Val636Met	NA	61
						TET2	c.978delA	p.Lys326Asnfs*21	NA	45
						TET2	c.1630C>T	p.Arg544*	NA	15
M2	MDS-RS-MLD	77	m	low risk (2)	46,XY	TET2	c.2562del	p.Phe854Leufs*19	3.53	NA
M3	MDS-RS-MLD	54	m	low risk (3)	46,XY	SF3B1	c.1866G>T	p.Glu622Asp	50	41
M4	MDS/MPN-RS-T	68	f	low risk (3)	46,XX	SF3B1	c.2098A>G	p.Lys700Glu	50	50
						TET2	c.5620G>A	p.Glu1874Lys	50	39
M5	MDS-RS-MLD	77	m	low risk (3)	46,XY	ASXL1	c.1900A>T	p.Arg634*	43	NA
						SF3B1	c.2098A>G	p.Lys700Glu	45	NA
						TET2	c.3986T>G	p.Leu1329Arg	45	NA
						TET2	c.4210C>T	p.Arg1404*	49	NA
						WT1	c.1048T>C	p.Cys350Arg	51	NA
M6	MDS-RS-MLD	78	m	int risk (3.5)	46,XY	ASXL1	c.1900A>T	p.Arg634*	43	NA
						SF3B1	c.2098A>G	p.Lys700Glu	45	NA
						TET2	c.3986T>G	p.Leu1329Arg	43	NA
						TET2	c.4210C>T	p.Arg1404*	50	NA
						WT1	c.1048T>C	p.Cys350Arg	51	NA
M7	MDS-EB1	70	f	int risk (4.5)	46,XX	ASXL1	c.1900_1922del	p.Glu635Argfs*15	18	NA
						KRAS	c.34G>A	p.Gly12Ser	8	NA
M8	MDS-EB1	64	f	high risk (5)	45,XX,-7	ASXL1	c.1934dup	p.Gly646Trpfs*12	12	NA
						DNMT3A	c.939G>A	p.Trp313*	32	NA
						SETBP1	c.2608G>A	p.Gly870Ser	11	NA
M9	CMML-2	76	m	high risk (6)	46,XY	ASXL1	c.1772dupA	p.Tyr591*	50	NA
						RUNX1	c.1014dupC	plle339Hisfs*234	40	NA
						SRSF2	c.284C>A	p.Pro95His	45	NA
						TET2	c.3300delT	plLeu1101Serfs*5	50	NA
M10	MDS-EB2	75	f	very high (7)	complex	ASXL1	unknown	unknown	NA	NA

MDS, CHIP, and healthy age-matched sample information, including sample ID, age, gender, disease/entity (WHO 2017), IPSS-R-score, karyotype, and somatic mutations. Mutational status in MDS bulk material was analyzed for diagnostic purposes. For some MDS cases, sorted HSC were additionally sequenced. VAF: variant allele frequency. *: Samples previously analyzed in.[18] NA: not analyzed.

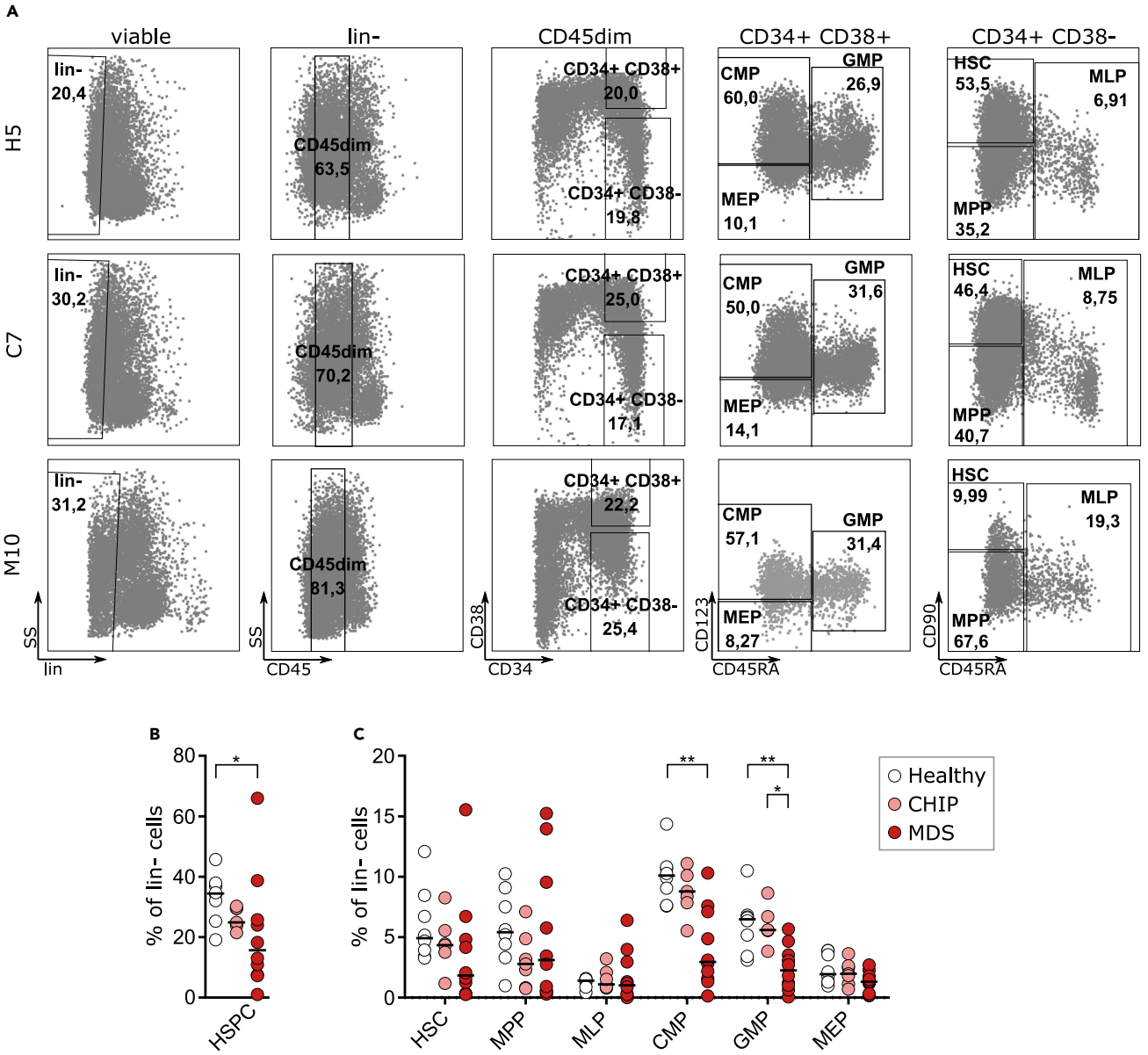


Figure 1. Percentages of HSPC are progressively reduced in CHIP and MDS bone marrow (BM) snapshot data

(A) Representative BM snapshot analysis of healthy, CHIP, and MDS. For CMP, GMP, and MEP fractions, cells were gated on lin-CD45dimCD34+CD38^{+} cells and further distinguished using CD123 and CD45RA. For HSC, MPP, and MLP fractions, cells were gated on lin-CD45dimCD34+CD38^{-} cells and further distinguished using CD90 and CD45RA.

(B) BM snapshots of healthy, CHIP, and MDS samples reveal a significantly reduced percentage of hematopoietic stem and progenitor cells (HSPC) in all lin- cells in MDS patients ($p = 0.038$, Kruskal-Wallis test followed by Dunn's multiple comparison test). Data are represented as individual points of independent samples, lines indicate median.

(C) The frequencies of CMP and GMP are significantly reduced in MDS BM ($p = 0.0031$ for CMP healthy vs. MDS, $p = 0.0074$ for GMP healthy vs. MDS, $p = 0.017$ for GMP CHIP vs. MDS, Kruskal-Wallis test followed by Dunn's multiple comparison test). For all compartments, a trend toward reduced frequencies from healthy over CHIP to MDS can be observed. Data are represented as individual points of independent samples, lines indicate median.

HSC from CHIP and MDS samples expand, differentiate, and divide in 7-day *in vitro* cultures

The observed changes in HSPC compartment composition in CHIP and MDS may be a result of altered kinetic rates (i.e., proliferation, differentiation, and cell death) in any one or several of the HSPC subpopulations as well as a reorganization of the hematopoietic hierarchy, with changes in differentiation transitions between cell types. To address this question, we employed a previously established workflow[18] in which

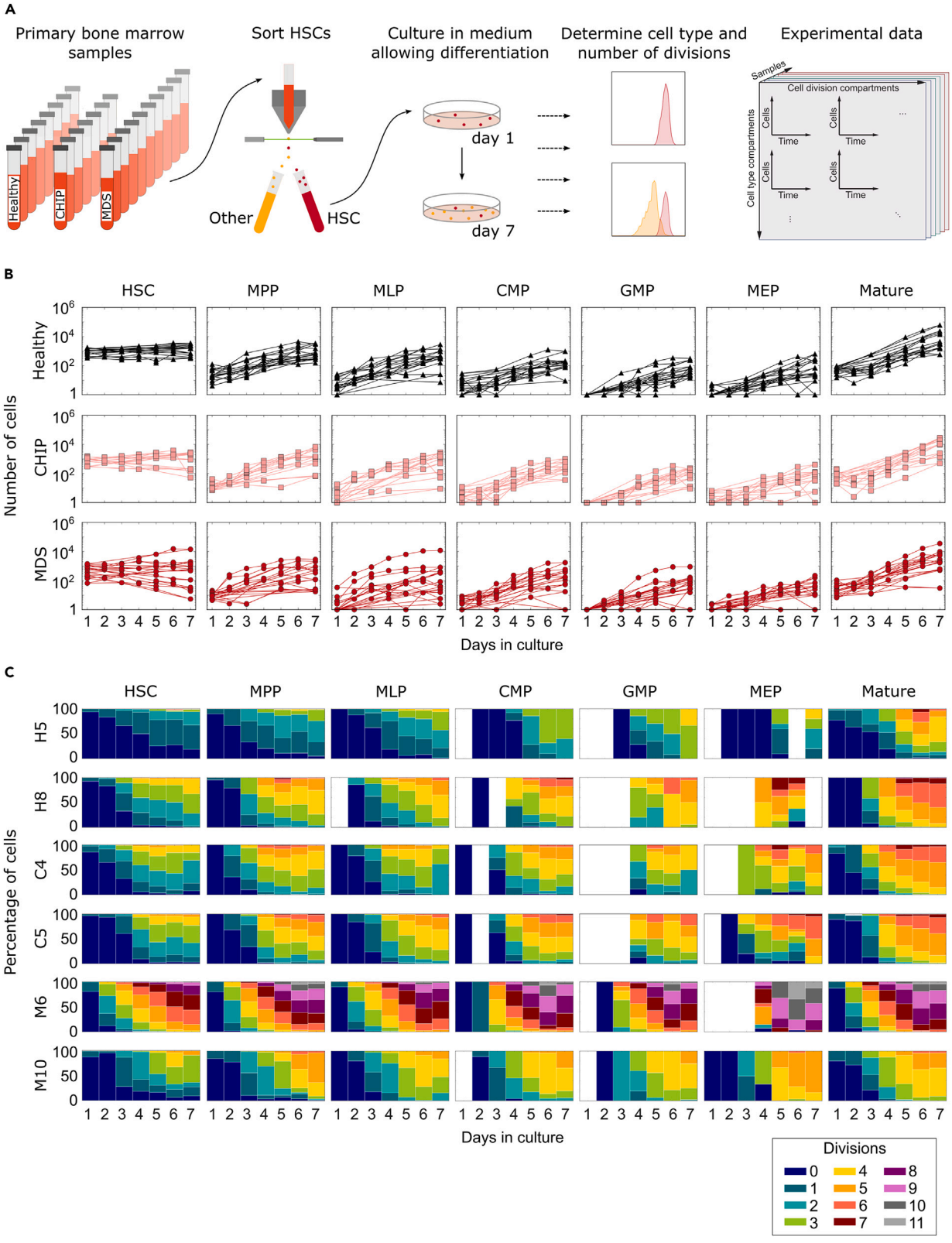

A
Primary bone marrow samples
Sort HSCs
Culture in medium allowing differentiation
Determine cell type and number of divisions
Experimental data
Healthy
CHIP
MDS
Other
HSC
day 1
day 7
Samples
Cell division compartments
Cell type compartments
Cells
Time
B
HSC
MPP
MLP
CMP
GMP
MEP
Mature
Healthy
CHIP
MDS
Number of cells
Days in culture
C
Percentage of cells
H5
H8
C4
C5
M6
M10
Days in culture
Divisions
0
1
2
3
4
5
6
7
8
9
10
11

Figure 2. HSC derived from CHIP, MDS, and healthy samples expand and differentiate in 7-day *in vitro* cultures

(A) Workflow for investigating HSC *in vitro* behavior in healthy (n = 8), CHIP (n = 7), and MDS (n = 10) BM samples. After FACS-sorting of HSC, cells were cultured in serum-free medium allowing differentiation followed by time-resolved measurements of cell number, cell type, and cell division status by FACS.

(B) Cell counts of HSC, MPP, MLP, CMP, GMP, MEP, and mature cell compartments during 7-day culture. Individual samples are represented as points along the 7-day period.

(C) Cell divisions measured by FACS in all cell types reveal similar patterns in exemplary healthy (H5 and H8) and CHIP (C4 and C5) samples, but hyperproliferation in MDS (sample M6).

HSC are stringently sorted as lin-CD45dimCD34+CD38$^-$CD45RA-CD90$^+$ cells, additionally stained with CellTrace division marker, cultured for 7 days in a defined medium optimized for differentiation, and analyzed by multiparameter immunophenotyping on subsequent days (Figure 2A). This approach allowed us to measure the abundances of 7 cell types for all 25 BM samples in a time-resolved manner and track the number of cell divisions between the experimental start (day 1) and the respective observed time points, until the end of the experiment (day 7) (Figure S1). Cell numbers and division distributions of the 7 cell types showed that HSC from all samples (healthy, CHIP, and MDS) expanded, differentiated, and divided in our 7-day culture conditions (Figures 2B and 2C). Abundance of HSC, MPP, MLP, CMP, GMP, and MEP and mature (i.e., CD34 negative) cells were comparable in healthy, CHIP, and the majority of MDS samples (Figure 2B). In addition, cell division patterns in healthy and CHIP were very similar (Figure 2C), while some MDS samples showed accelerated cycling in all HSPC compartments (see Figure 2C for exemplary kinetics from sample M6).

To verify the plausibility of our 7-day *in vitro* culture system as a model system for analysis of HSPC kinetics in CHIP and MDS, we additionally compared a group of young samples (<55 years old, Table S1) to our aged healthy individual samples (>55 years old, Table 1), as they are expected to show less inter-individual heterogeneity than MDS samples. Similar to snapshot analysis of unprocessed BM, 7-day cultures of aged HSC show a tendency for increased HSPC proportions (Figures S2A and S2B). Furthermore, when analyzing cell production after 7-day culture *in vitro*, the cell yield from older HSC is significantly lower than from young HSC, correctly reflecting the hallmarks of aging BM which is characterized by hypocellularity and reduced blood cell production (Figure S2C).

Increasing heterogeneity within the hematopoietic hierarchy in CHIP and MDS HSPC

To investigate if the hematopoietic architecture is altered in CHIP and MDS, we performed a systematic model comparison based on the experimentally generated single-cell data (see Figure 3A for exemplary model fits). Analogous to samples from age-matched healthy individuals,[18] we fitted models of 10 different hematopoietic hierarchies (models A-J, Figure S3) to every individual sample and identified all plausible models, as well as the best performing model according to Bayesian Information Criterion (BIC) values (Figure 3B and Table S6). Similar to our analysis of healthy HSC samples,[18] model A is plausible for all individuals with CHIP. However, model A performs best in only 3 of 7 CHIP samples, with model B and I performing best in the other 4 CHIP samples, reflecting an increased heterogeneity compared to age-matched healthy samples. Of note, there was no apparent correlation between mutation type and best performing hierarchy model in the CHIP samples: While 5 out of 7 CHIP samples carried *DNMT3A* mutations (samples C1-5), assignment to the best performing model was spread among the three most plausible models (A, B, and I). However, 2 samples with *DNMT3A* mutations (C3 and C5) exhibited an increased rate of differentiation toward the MEP compartment (Figure 3E), corresponding to union model I and in line with recently published transcriptomic data.[27] The heterogeneity between samples was even more prevalent in MDS. Model A was no longer plausible in 6 out of 10 samples and performed best in only 3 samples. In addition to model A, models B, D, and G were now identified as the best performing models (Figure 3B). Again, there was no discernible correlation between mutation and best performing model for our MDS samples. Although all samples in which model B fit best carried *SF3B1* mutations (M1, M5, and M6, see Table 1), the small size of our cohort does not allow for any definite conclusions.

Union model reveals rate changes and increasing number of deregulated rates in CHIP and MDS HSPC

Since model A was not sufficient to describe the hematopoietic tree in the majority of MDS cases, we merged transitions from all models performing best in at least one case of CHIP or MDS (model A, B, D, G, and I from[18]) into a new union model. This model includes transitions between the MLP and the GMP compartment, the HSC and mature compartment, and the HSC and MEP compartment (Figure 3C). We

next estimated proliferation, differentiation, and cell death rates for HSC, MPP, MLP, CMP, GMP, MEP, and $CD34^-$ mature cells from the experimental *in vitro* generated data from CHIP, MDS, and healthy HSC (Figure S4 and Tables S3–S5). To identify accelerated and decelerated rates in CHIP and MDS donor samples, we calculated a 90% confidence interval (CI) based on parameter estimates and uncertainties of age-matched healthy samples (STAR Methods, Figure S4). We find altered proliferation, differentiation, and cell death rates in various cell types (Figures 3D–3F) and an increasing number of altered rates within CHIP and MDS (Figure 3G). Interestingly, 2 out of 7 CHIP samples and 6 out of 10 MDS samples showed increased HSC differentiation rates toward the MEP or mature compartment (Figures 3D, 3E, and 3F). In summary, CHIP and MDS samples show increased heterogeneity in terms of hematopoietic hierarchy and HSPC kinetics compared to healthy controls.

Next, we wanted to verify if kinetic alterations detected by computational modeling were supported by further functional assays and clinical data. Therefore, HSC were sorted in single-cell mode into 96-well plates and cultured for 7 days. The cell number of each well was counted repeatedly. To deconvolute changes in HSC cell division, we reported the percentage of cells which divided at least once for each analyzed time point. HSC from MDS patients divide earlier than CHIP and healthy HSC, but also show a higher heterogeneity between samples (Figure S5). Most importantly, MDS M2 and M6, which were identified to have increased HSC proliferation by computational modeling (see Figure 3F), are among the most rapid dividers in our single-cell assay (Figure S5).

Our approach depicts HSC intrinsic behavior, excluding effects mediated by the BM niche and monitors only effects appearing within 7 days. We, therefore, asked if the observed specific rate changes mirror the patients' blood values. The additional transition from the HSC compartment to the MEP compartment is found in 4 out of 10 MDS samples. Higher differentiation into MEPs and bypassing of the other myeloid lineage compartments could result in lower white blood cell counts and/or higher red blood cell or platelet values. Indeed, in samples with an HSC to MEP transition we observed the highest hemoglobin (M8 and M9) and the lowest leukocyte levels (M7 and M9, Figure S6).

Clustering individual samples pinpoints the most disrupted transitions in CHIP and MDS

To analyze whether the deregulated rates in CHIP and MDS segregate into discernable subgroups, we performed a weighted principal component analysis (PCA) on the parameter values estimated from MDS patients. Healthy and CHIP samples were projected onto the principal component (PC) space and two ellipses were calculated to highlight 85% and 95% confidence areas of healthy age-matched samples (see STAR Methods). Focusing on PC1 to PC4, 7 out of 10 MDS cases are outside the 95% confidence area at least once, and 9 out of 10 MDS cases are outside the 85% confidence area at least once (Figure 4A). The one sample where this is not the case (M4) was classified as MDS/MPN with thrombocytosis.[28,29] CHIP samples mainly overlap with age-matched healthy control samples. Not a single CHIP sample lies outside the 95% confidence area and only 1 out of 7 CHIP samples lies outside the 85% healthy confidence area. Although MDS samples are distinct from CHIP and healthy samples in our PCA, there is no clear pattern regarding the IPSS-R score, the number of mutations, or specific mutations including *SF3B1*, *ASXL1*, and *TET2* (Figures 4B, 4C, and S7).

Principal components (PC) 1 to 4 account for approximately 90% of the total variance in CHIP and for about 80% in MDS (Figures S8A and S8B). The most contributing rates in CHIP are MPP proliferation, CMP to MEP differentiation, and HSC apoptosis, whereas the most contributing rates in MDS are HSC to MPP differentiation, MLP to GMP differentiation, and HSC apoptosis (Figures S8C and S8D). Notably, in MDS all of the 5 rates linked to HSC proliferation, differentiation (HSC to MPP, MEP, and mature cells), or cell death are in the top 8 most contributing rates.

DISCUSSION

Most insights into human hematopoiesis rely on snapshot analyses from primary tissue. Results are thus necessarily affected by cell source differences, preprocessing biases, and lab-specific experimental protocols, leading to partially conflicting reports of HSPC proportions in MDS.[7–9] In contrast, our standardized *in vitro* culturing approach makes BM samples comparable in terms of cell numbers, and particularly, with respect to differentiation kinetics. Computational profiling i.e., fitting of a mathematical model to stem and progenitor cell compartments and the number of divisions since experiment start, reveals the preferred hematopoietic hierarchy for each sample and rate differences with cell-type resolution. Moreover, it allows us

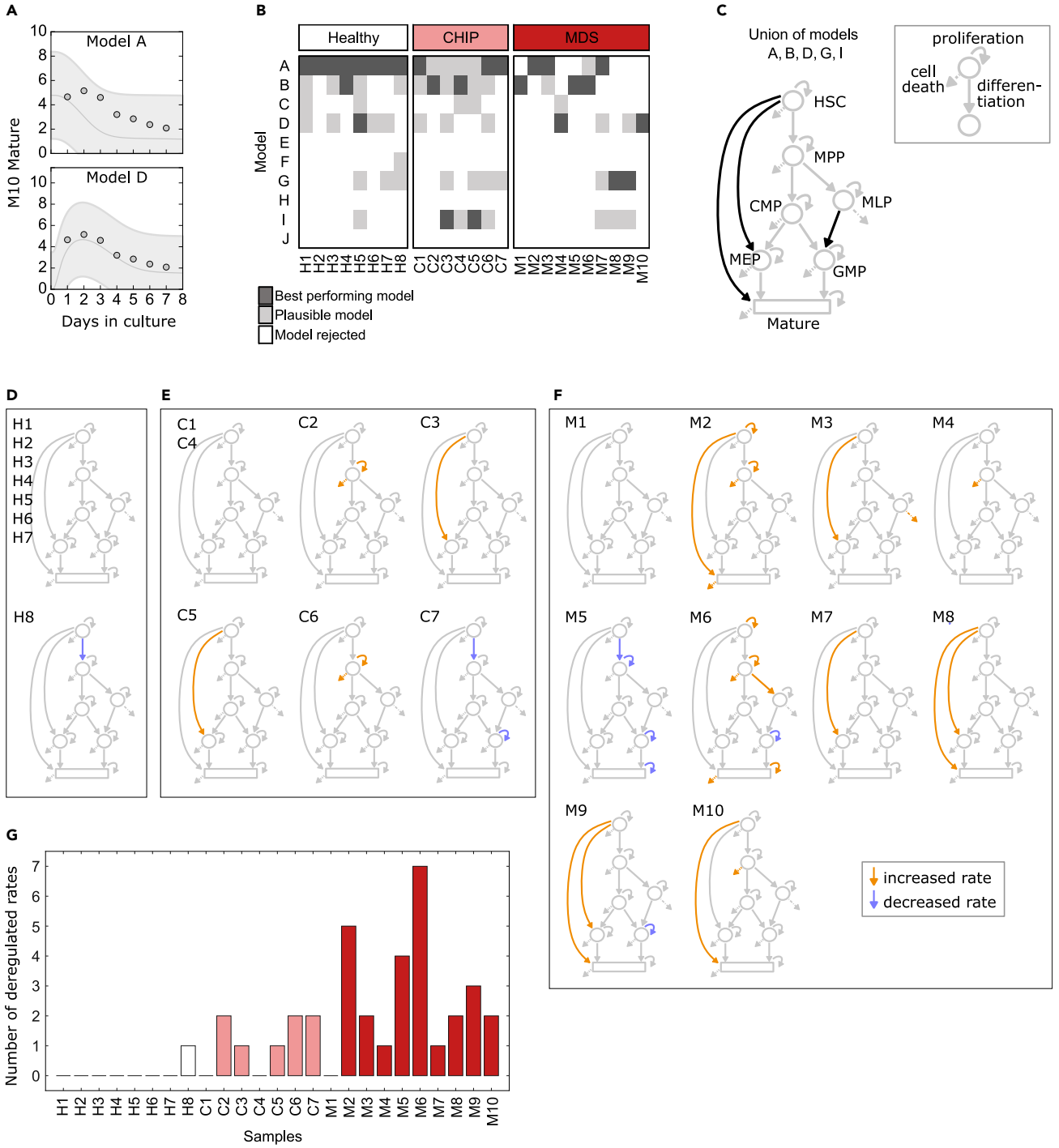


Figure 3. Hierarchy heterogeneity and number of deregulated rates increases from clonal hematopoiesis to MDS

(A) Mature (CD34$^-$) cells from sample M10 which underwent 0 divisions fitted with model A (rejected) and model D (best performing).

(B) Systematic model comparison reveals best performing models according to BIC values are model A (5 CHIP and 3 MDS individuals), model B (2 CHIP and 3 MDS individuals), model D (2 MDS individuals), model G (2 MDS individuals) and model I (2 CHIP individuals). For two healthy individuals (H4 and H5), two models perform similarly well.

(C) Union of best performing models in at least one healthy, MDS, or CHIP sample defines a new hematopoietic hierarchy model.

(D) Detected up-regulated (orange) and down-regulated (blue) rates for each healthy individual.

(E) Detected up-regulated (orange) and down-regulated (blue) rates for each CHIP individual.

(F) Detected up-regulated (orange) and down-regulated (blue) rates for each MDS individual.

(G) A higher number of altered rates per individual can be found in CHIP and MDS.

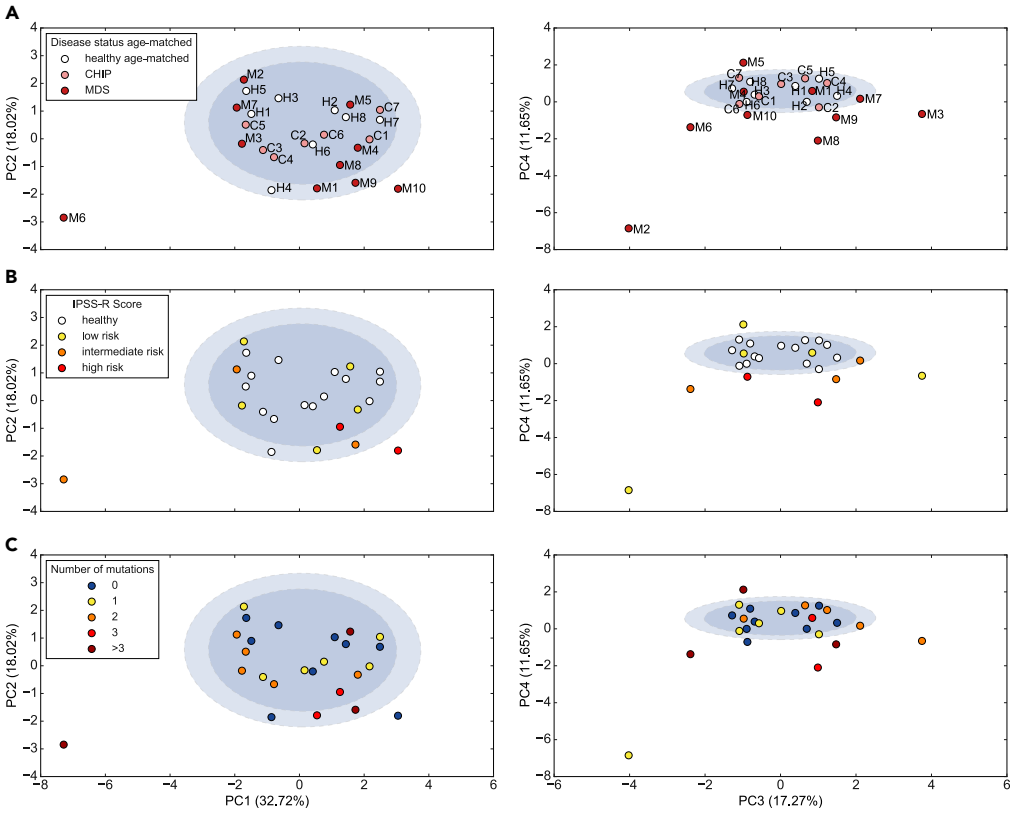


Figure 4. MDS samples can be distinguished from CHIP and healthy samples using PCA

Weighted PCA was performed on proliferation, differentiation, and cell death rates of each cell compartment estimated from MDS patients. Depicted are PC1, PC2, PC3, and PC4. Dots show the location of each sample. Blue ellipses indicate 95% and 85% confidence areas based on rates of healthy donors.

(A) Donor samples (dots) with their ID in PCA space spanned by the first four PCs. Highlighted are their disease status (healthy, CHIP, and MDS).

(B) Donor samples in PC space with their respective IPSS-R score.

(C) Donor samples in PC space with their respective number of mutations.

to speculate about the observed decrease of HSPCs within all lin- cells from healthy, over CHIP, to MDS (Figure 1). Our results suggest that the increased *in vitro* differentiation rates observed in half of the CHIP and MDS samples (Figures 3D–3F), in addition to a limited self-renewing capacity of HSPC, may contribute to emerging impaired blood production in CHIP and MDS.

Our results moreover show deviating hierarchical structures for most MDS patients, with a median of 2 deregulated rates per patient (range 0–8), demonstrating a reorganization of the classic hematopoietic tree with disease progression. As baseline, we used plausible hierarchies inferred from a cohort of healthy age-matched controls, which differ slightly from the ones inferred in[18] due to differences in the cohort's age distribution. Strikingly, a transformation toward disorderly hematopoiesis can already be observed in CHIP, however at a lower frequency and with fewer alterations per sample (median of 1 and range 0–2). This reorganization in MDS and possibly in CHIP is potentially driven by pre-leukemic stem cells (leukemic stem cells [LSC]), which are shown to be localized within the lin-CD34^{+}CD38^{-}CD90^{+}CD45RA-compartment.[20]

Indeed, we could show that mutations present in bulk BM of our MDS patients were also detected at the HSC level, and with a similarly high VAF (samples M1, M3, and M4, all from samples with SF3B1 mutations, Table 1), suggesting a substantial percentage of LSC in MDS HSC. While we did not sequence the HSC population in our CHIP samples separately, it has been previously shown that the *DNMT3A* mutation originates at the multipotent HSC level (corresponding to our sorted lin-CD34^{+}CD38^{-}CD90^{+}CD45RA- population) while *TET2* has been described to occur in myeloid restricted stem cells.[11,30] We and others have shown that the *ASXL1*

mutation also arises in HSC.[11,31] Thus, we can extrapolate that a significant percentage of HSC are clonal in our cultures.

By correlating mutations with altered rates, we found no specific kinetic pattern to reappear in our 10 MDS samples. However, we noted that the vast majority (8 out of 10) of MDS samples showed alterations on the HSC level. This was different in CHIP, where only 3 out of 7 samples showed altered HSC rates. As baseline, we again used parameter estimates inferred from the cohort of healthy, age-matched controls. Across all samples, 14 altered rates involve the HSC compartment. Ten of those were differentiation rates from the HSC compartment toward the MEP or mature compartment, which are both transitions not described by the classical model of hematopoiesis (model A in Figure S3). Skewing of myeloid HSC cell fate toward megakaryocyte-erythroid progenitors has been described using single-cell transcriptomics for *DNMT3A* R882 mutated HSPC.[27] In our cohort, we profiled 5 CHIP samples with a *DNMT3A* mutation. Of these, 3 had *DNMT3A* as a single mutation, and one case (C5) contained the R882 mutation commonly associated with progression to myeloid disease. Indeed, HSC from our sample C5 show an increased rate of differentiation toward MEP. Interestingly, HSC from sample C3 containing a splice variant of *DNMT3A* at a VAF of 5% also show this same kinetic pattern, suggesting that skewing toward MEP differentiation in *DNMT3A* mutated HSC may not be restricted to the R882 variant. In total, our data demonstrate that reorganization of the HSPC compartment starts at the apex of the hierarchy in CHIP as well as MDS, and alterations in HSC kinetics are most important in MDS.

For achieving clonal dominance, mutated HSCs need to have a growth advantage or increased fitness level.[32,33] The alterations we found in CHIP and MDS samples cannot completely explain this phenomenon. However, this comes as no surprise, as our approach is based on 7-day measurements, whereas clonal dominance is achieved over years or as has been proven for myeloproliferative neoplasms even over the lifespan of an individual.[32–34] In our cell culture, HSCs are supported to divide and differentiate with limited maintenance of self-renewal properties. We thus focus on cell-intrinsic driven alterations, whereas *in vivo*, the hematopoietic system is additionally regulated by the microenvironment. Notably, the kinetic changes we observe come from a mixture of mutated and non-mutated HSC. It remains to be seen if the cell-intrinsic alterations affect only clonal HSC or also non-mutated cells. Our work thus offers an alternative perspective on the fascinating phenomenon of clonal expansion, supplementing previous work based on snapshot data[11] and longitudinal analysis over years[33] with insights into differentiation hierarchy alterations and changes of cell type kinetics of blood stem cells. Future studies should focus on extrinsic effects, such as stem and progenitor trafficking,[35] regulation of quiescence through interaction with the BM niche,[36] or the block of BM to blood egress, which would contribute to the observed MDS cytopenia in the presence of hypercellular BM.

Detailed insights into differentiation pathways and rates are critical for understanding how hematopoiesis yields sufficient blood cells throughout the lifetime of an individual. Improving our understanding of cell-type specific differentiation, proliferation, and cell death may help to predict how hematologic diseases develop and how they respond to therapy. Our computational approach provides access to these parameters and may facilitate comparisons of benign and leukemic kinetics to healthy hematopoiesis to identify cell types and rates most affected in leukemogenesis for each patient individually. It might thus also help to uncover targetable cell intrinsic disease mechanisms in the future.

Limitations of the study

Our model enables us to track the intrinsic behavior of HSC *in vitro* and correlate this to disease states and mutational subtypes, but we have so far only profiled a small cohort which will require further confirmation. Importantly, the influence of extrinsic factors such as the BM microenvironment on the kinetics of diseased HSC cannot be examined in our system. In the future, dissecting the contribution of clonal (i.e., mutated) and non-mutated HSC to the observed disturbed kinetics in CHIP and MDS will require linking our computational approach to sequencing data.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107328.

ACKNOWLEDGMENTS

This work was supported by research grants from the Deutsche Forschungsgemeinschaft SFB 1243 for projects A09 (K.S.G. and R.A.J.O.), A17 (F.J.T. and C.M.), A04 (F.B.), and A06 (K.H.M.), and the FOR 2033 project B3 (K.S.G. and R.A.J.O.) as well as the Deutsche Jose Carreras Leukämie Stiftung (DJCLS, R14/18, K.S.G.) and the German Cancer Consortium joint funding program (DKTK CHOICE). C.M. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement No. 866411) and from the BMBF project TIDY (031L0170B). K.S.G has received funding from the ERC under the European Union's Horizon 2020 Marie Sklodowska-Curie Innovative Training Network (MSCA-ITN, Grant agreement No. 953407). L.B. was financially supported by the Joachim Herz Stiftung. We are grateful to Marc Schmidt-Supprian, Mark van der Garde, Christian Müller, and Elmar Spiegel for critical comments and discussion. We would like to thank Martin Nolde and Dominikus Hausmann for supplying femoral heads, and Torsten Haferlach for mutational information on selected patients.

AUTHOR CONTRIBUTIONS

M.C.B. designed, established and performed experiments, analyzed data and drafted the manuscript. L.B. derived and implemented mathematical models, performed parameter inference, analyzed data and modeling results and drafted the manuscript. J.S.H. performed bone marrow biopsies and collected clinical patient data. J.R. performed antibody staining and provided critical comments. M.R.T. performed sequencing analyses. L.V. isolated MNC from bone marrow samples and performed sample biobanking. D.W. reviewed codes. I.A. assisted with flow cytometric cell sorting. F.J.T., and F.B. provided conceptual advice and critical comments. K.H.M. performed sequencing analyses and provided conceptual advice. R.A.J., C.M., and K.S.G. designed experiments, analyzed data, and wrote the manuscript. All authors read and agreed with the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: September 11, 2022
Revised: May 9, 2023
Accepted: July 5, 2023
Published: July 10, 2023

REFERENCES

1. Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. N. Engl. J. Med. *371*, 2488–2498.

2. Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N. Engl. J. Med. *371*, 2477–2487.

3. Sperling, A.S., Gibson, C.J., and Ebert, B.L. (2017). The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. Nat. Rev. Cancer *17*, 5–19.

4. Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P., and Ebert, B.L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood *126*, 9–16.

5. Kuranda, K., Vargaftig, J., de la Rochere, P., Dosquet, C., Charron, D., Bardin, F., Tonnelle, C., Bonnet, D., and Goodhardt, M. (2011). Age-related changes in human hematopoietic stem/progenitor cells. Aging Cell *10*, 542–546.

6. Pang, W.W., Price, E.A., Sahoo, D., Beerman, I., Maloney, W.J., Rossi, D.J., Schrier, S.L., and Weissman, I.L. (2011). Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. Proc. Natl. Acad. Sci. USA *108*, 20012–20017.

7. Will, B., Zhou, L., Vogler, T.O., Ben-Neriah, S., Schinke, C., Tamari, R., Yu, Y., Bhagat, T.D., Bhattacharyya, S., Barreyro, L., et al. (2012). Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations. Blood *120*, 2076–2086.

8. Pang, W.W., Pluvinage, J.V., Price, E.A., Sridhar, K., Arber, D.A., Greenberg, P.L., Schrier, S.L., Park, C.Y., and Weissman, I.L. (2013). Hematopoietic stem cell and progenitor cell mechanisms in myelodysplastic syndromes. Proc. Natl. Acad. Sci. USA *110*, 3011–3016.

9. Ostendorf, B.N., Flenner, E., Flörcken, A., and Westermann, J. (2018). Phenotypic characterization of aberrant stem and progenitor cell populations in myelodysplastic syndromes. PLoS One *13*, e0197823.

10. Ganan-Gomez, I., Yang, H., Ma, F., Montalban-Bravo, G., Thongon, N., Marchica, V., Richard-Carpentier, G., Chien, K., Manyam, G., Wang, F., et al. (2022). Author Correction: Stem cell architecture drives myelodysplastic syndrome progression and predicts response to venetoclax-based therapy. Nat. Med. *28*, 1097.

11. Arends, C.M., Galan-Sousa, J., Hoyer, K., Chan, W., Jäger, M., Yoshida, K., Seemann, R., Noerenberg, D., Waldhueter, N., Fleischer-Notter, H., et al. (2018). Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. Leukemia *32*, 1908–1919.

12. Hecker, J.S., Hartmann, L., Rivière, J., Rivière, J., Ksienzyk, B., Buck, M.C., Van Der Garde, M., Fischer, L., Winter, S., Rauner, M., et al. (2021). CHIP & HIPs: Clonal hematopoiesis is common in hip arthroplasty patients and associates with autoimmune disease. Blood *138*, 3278. https://doi.org/10.1182/blood.2020010163.

13. Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193–197.

14. Manz, M.G., Miyamoto, T., Akashi, K., and Weissman, I.L. (2002). Prospective isolation of human clonogenic common myeloid progenitors. Proc. Natl. Acad. Sci. USA *99*, 11872–11877.

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

16. Haas, S., Trumpp, A., and Milsom, M.D. (2018). Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. Cell Stem Cell *22*, 627–638.

17. Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature *553*, 418–426.

18. Bast, L., Buck, M.C., Hecker, J.S., Oostendorp, R.A.J., Götze, K.S., and Marr, C. (2021). Computational modeling of stem and progenitor cell kinetics identifies plausible hematopoietic lineage hierarchies. iScience *24*, 102120.

19. Chen, J., Kao, Y.-R., Sun, D., Todorova, T.I., Reynolds, D., Narayanagari, S.-R., Montagna, C., Will, B., Verma, A., and Steidl, U. (2019). Myelodysplastic syndrome progression to acute myeloid leukemia at the stem cell level. Nat. Med. *25*, 103–110.

20. Woll, P.S., Kjällquist, U., Chowdhury, O., Doolittle, H., Wedge, D.C., Thongjuea, S., Erlandsson, R., Ngara, M., Anderson, K., Deng, Q., et al. (2014). Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. Cancer Cell *25*, 794–808.

21. Ye, M., Zhang, H., Yang, H., Koche, R., Staber, P.B., Cusan, M., Levantini, E., Welner, R.S., Bach, C.S., Zhang, J., et al. (2015). Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia. Cell Stem Cell *17*, 611–623.

22. Greenberg, P.L., Tuechler, H., Schanz, J., Sanz, G., Garcia-Manero, G., Solé, F., Bennett, J.M., Bowen, D., Fenaux, P., Dreyfus, F., et al. (2012). Revised international prognostic scoring system for myelodysplastic syndromes. Blood *120*, 2454–2465.

23. Bejar, R., Levine, R., and Ebert, B.L. (2011). Unraveling the molecular pathophysiology of myelodysplastic syndromes. J. Clin. Oncol. *29*, 504–515.

24. Haferlach, T., Nagata, Y., Grossmann, V., Okuno, Y., Bacher, U., Nagae, G., Schnittger, S., Sanada, M., Kon, A., Alpermann, T., et al. (2014). Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia *28*, 241–247.

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

26. Doulatov, S., Notta, F., Eppert, K., Nguyen, L.T., Ohashi, P.S., and 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.

27. Nam, A.S., Dusaj, N., Izzo, F., Murali, R., Myers, R.M., Mouhieddine, T.H., Sotelo, J., Benbarche, S., Waarts, M., Gaiti, F., et al. (2022). Single-cell multi-omics of human clonal hematopoiesis reveals that DNMT3A R882 mutations perturb early progenitor states through selective hypomethylation. Nat. Genet. *54*, 1514–1526.

28. Patnaik, M.M., and Tefferi, A. (2017). Refractory anemia with ring sideroblasts (RARS) and RARS with thrombocytosis (RARS-T): 2017 update on diagnosis, risk-stratification, and management. Am. J. Hematol. *92*, 297–310.

29. Khoury, J.D., Solary, E., Abla, O., Akkari, Y., Alaggio, R., Apperley, J.F., Bejar, R., Berti, E., Busque, L., Chan, J.K.C., et al. (2022). The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia *36*, 1703–1719. https://doi.org/10.1038/s41375-022-01613-1.

30. Buscarlet, M., Provost, S., Zada, Y.F., Bourgoin, V., Mollica, L., Dubé, M.P., and Busque, L. (2018). Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for DNMT3A. Blood *132*, 277–280.

31. Hartmann, L., Hecker, J.S., Rothenberg-Thurley, M., Rivière, J., Jentzsch, M., Ksienzyk, B., Buck, M.C., van der Garde, M., Fischer, L., Winter, S., et al. (2022). Compartment-specific mutational landscape of clonal hematopoiesis. Leukemia *36*, 2647–2655.

32. Watson, C.J., Papula, A.L., Poon, G.Y.P., Wong, W.H., Young, A.L., Druley, T.E., Fisher, D.S., and Blundell, J.R. (2020). The evolutionary dynamics and fitness landscape of clonal hematopoiesis. Science *367*, 1449–1454.

33. Robertson, N.A., Latorre-Crespo, E., Terradas-Terradas, M., Lemos-Portela, J., Purcell, A.C., Livesey, B.J., Hillary, R.F., Murphy, L., Fawkes, A., MacGillivray, L., et al. (2022). Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness

effects. Nat. Med. *28*, 1439–1446. https://doi.org/10.1038/s41591-022-01883-3.

34. Williams, N., Lee, J., Mitchell, E., Moore, L., Baxter, E.J., Hewinson, J., Dawson, K.J., Menzies, A., Godfrey, A.L., Green, A.R., et al. (2022). Life histories of myeloproliferative neoplasms inferred from phylogenies. Nature *602*, 162–168.

35. Hoggatt, J., Mohammad, K.S., Singh, P., Hoggatt, A.F., Chitteti, B.R., Speth, J.M., Hu, P., Poteat, B.A., Stilger, K.N., Ferraro, F., et al. (2013). Differential stem- and progenitor-cell trafficking by prostaglandin E2. Nature *495*, 365–369.

36. Goncalves, K.A., Silberstein, L., Li, S., Severe, N., Hu, M.G., Yang, H., Scadden, D.T., and Hu, G.-F. (2016). Angiogenin Promotes Hematopoietic Regeneration by Dichotomously Regulating Quiescence of Stem and Progenitor Cells. Cell *166*, 894–906.

37. Rothenberg-Thurley, M., Amler, S., Goerlich, D., Köhnke, T., Konstandin, N.P., Schneider, S., Sauerland, M.C., Herold, T., Hubmann, M., Ksienzyk, B., et al. (2018). Persistence of pre-leukemic clones during first remission and risk of relapse in acute myeloid leukemia. Leukemia *32*, 1598–1608.

38. Villaverde, A.F., Barreiro, A., and Papachristodoulou, A. (2016). Structural Identifiability of Dynamic Systems Biology Models. PLoS Comput. Biol. *12*, e1005153.

39. Stapor, P., Weindl, D., Ballnus, B., Hug, S., Loos, C., Fiedler, A., Krause, S., Hroß, S., Fröhlich, F., and Hasenauer, J. (2018). PESTO: Parameter EStimation TOolbox. Bioinformatics *34*, 705–707.

40. Fröhlich, F., Weindl, D., Schälte, Y., Pathirana, D., Paszkowski, Ł., Lines, G.T., Stapor, P., and Hasenauer, J. (2021). AMICI: high-performance sensitivity analysis for large ordinary differential equation models. Bioinformatics *37*, 3676–3677.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
healthy human BM samples	patients undergoing total hip arthroplasty (TUM 538/16), see Table 1	N/A
human MDS BM samples	patients undergoing routine diagnostic BM evaluation (TUM 538/16), see Table 1	N/A
Software and algorithms		
GraphPad Prism	GraphPad Inc, La Jolla, CA	Version 6.01,
Inkscape	https://inkscape.org	Inkscape 1.1 (c68e22c387, 2021-05-23)
MATLAB	The MathWorks, Inc.,	R2017a
MATLAB toolbox STRIKE-GOLDD	Villaverde, et al.[38]	https://github.com/afvillaverde/strike-goldd_2.1
MATLAB toolbox PESTO	Stapor, et al.[39]	https://github.com/ICB-DCM/PESTO/
MATLAB toolbox AMICI	Fröhlich, et al.[40]	https://github.com/ICB-DCM/AMICI
Python (using libraries pandas (1.1.5), numpy (1.19.2), seaborn (0.11.0), matplotlib (3.3.2), scipy (1.5.2), scikit-learn (0.23.2), statsmodels (0.12.1), h5py (2.10.0)	Python Software Foundation	version 3.6.12, https://www.python.org/
JupyterLab	Project Jupyter	version 2.2.6, https://jupyterlab.readthedocs.io/en/stable/
Deposited data		
All original code	[18]/this paper	https://github.com/marrlab/HematopoieticDisorderAnalysis
FACS raw data	this paper	Data will be shared by the lead contact upon request
FACS-derived data tables	this paper	https://github.com/marrlab/HematopoieticDisorderAnalysis/blob/master/MATLAB/parameter_inference/data/2020_08_data.xlsx

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Katharina Götze (katharina.goetze@tum.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Raw FACS files from de-identified human donor and patient BM will be shared by the lead contact upon request.
- FACS-derived data which were used to fit models A-J and the union model have been deposited at https://github.com/marrlab/HematopoieticDisorderAnalysis/blob/master/MATLAB/parameter_inference/data/2020_08_data.xlsx (see key resources table). They are publicly available as of the date of publication.

- All original code has been deposited at https://github.com/marrlab/HematopoieticDisorderAnalysis and is publicly available as of the date of publication (see key resources table).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Healthy, CHIP and MDS sample collection and storage

MDS BM samples were obtained from patients undergoing routine clinical evaluation. The remaining BM samples (CHIP and healthy cohorts) were collected from femoral heads of patients undergoing hip replacement surgery.[12,18] 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 (TUM 538/16).

Age, gender and other characteristics of all human study participants are listed in Table 1. As far as possible individuals were age- and gender-matched. The mean age was 65.3 years for healthy donors, 69.8 years for individuals with CHIP, and 69.9 years for MDS patients. Concerning gender, the percentage of female individuals was 50% in the healthy cohort, 85% in the CHIP cohort and 40% in the MDS cohort. An influence of gender on the results of our study cannot be concluded or ruled out due to the limited cohort size.

METHOD DETAILS

Antibody staining and FACS sorting

For sorting of lin-CD45dimCD34+CD38-CD90$^+$CD45RA- HSCs and the analysis of cell compartments on day 1-7, samples were processed as described.[18]

Briefly, MNC 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). Compensation and gating were 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

Bulk-sorted HSC were cultured at concentration of 2.5x10^3 cells/ml in serum-free medium supplemented with 8 growth factors (SFM + 8GF). Single-HSC were sorted into the inner wells of a 96-well plate and cultured in 100 µl SFM + 8 GF. Outer wells were filled with H_2O. For a detailed description of medium and culture conditions see.[18] If there was a sufficient number of HSC after sorting, or if there was enough frozen BM material for repeated sorting, samples were analyzed in replicates.

Mutational analysis

MDS patient samples were sequenced for diagnostic purposes for the most common mutations using a 72 myeloid gene panel.[24] Healthy BM samples from hip replacement surgeries were analyzed for clonal hematopoiesis-associated mutations using a previously described NGS assay covering 68 genes recurrently altered in myeloid neoplasms.[37] For assessment of variant allele frequencies (VAF) in sorted HSC, DNA from a minimum of 200 cells was subjected to whole-genome amplification (GenomiPhi V2, GE Healthcare), and subsequently analyzed using the same 68-gene panel.

Terminology

In statistics, the term frequency (or absolute frequency) is used to describe the number of an observed event, while relative frequency (or ratio) refers to the proportion of a particular event within all events. Unfortunately, in biomedical research literature, these terms are not used as stringently. Here, to avoid confusion, we use the term 'numbers' to describe absolute cell counts n of a particular cell type, and 'proportion' to describe the number of counts *n* of a particular cell type in relation to all counts *N*, as *n/N*.

Multi-compartment model parameter inference

We modeled hematopoiesis as a biochemical reaction network of seven species S = {HSC, MPP, MLP, CMP, GMP, MEP, CD34-}, as previously published for models A-J in.[18] For the A,B,D,G,I union model (see Figure 3C) the ordinary differential equations (ODE) system of the multi-compartment model includes as all other previously described models (models A-J) reactions $R_1 - R_7$ (proliferation) and $R_{22} - R_{27}$ (cell death), but includes $R_8, R_{10} - R_{13}, \ R_{16} - R_{21}$ to describe differentiation between cell compartments (see Tables 1 and 2 in[18]). The ABDGI union model thus results in 24 rates (1 rate per reaction) and is more complex than model A (21 rates), which we identified as the best performing model in all 9 healthy age-matched individuals (Table S6). Parameter inference was conducted by maximum likelihood estimation for every individual separately (see Transparent Methods in[18]). For the newly defined union model, structural identifiability of parameters was analyzed with MATLAB toolbox STRIKE-GOLDD ([38], see Table S2). Practical identifiability was assessed by calculating profile likelihood-based confidence intervals (see Tables S3–S5).

Identification of increased or decreased rates

To identify considerably increased or decreased rates in MDS patients or CHIP individuals (Figures 3D, 3E, and S3C), we calculated 90% confidence intervals for rate estimates of healthy donors without mutations and used them as cutoff values. As our optimization approach does not reveal the full posterior distribution of each individual rate estimate, we used an approximation. We assumed that the posterior distribution of each individual rate corresponds to a log-normal distribution. As an approximation for the distribution parameters we used the individual rate estimates $r^H_{(i,j)}, i = 1, ..., 24, j = 1, ..., 8$, each belonging to a healthy donor, and their profile likelihood-based 95% confidence intervals $\left[CI^l_{(i,j)}, CI^u_{(i,j)}\right]$: $\hat{\mu}_{(i,j)} = ln(r^H_{(i,j)}) - \frac{\hat{\sigma}^2_{(i,j)}}{2}$ and $\hat{\sigma}_{(i,j)} = \frac{q_{0.975} - \sqrt{q^2_{0.975} - 2ln(\frac{CI^u_{(i,j)}}{r^H_{(i,j)}})} + q_{0.025} - \sqrt{q^2_{0.025} - 2ln(\frac{CI^l_{(i,j)}}{r^H_{(i,j)}})}}{2}$, where q_α is the α quantile of the standard normal distribution $\mathcal{N}(0,1)$.

By sampling *n* = 1000 random numbers from each individual $log\mathcal{N}(\hat{\mu}_{(i,j)}, \hat{\sigma}_{(i,j)}), j = 1, ..., 8$ distribution and calculating the 0.05 and 0.95 percentiles of the 5000 pooled random numbers, we approximated the 90% confidence interval for all *i* = 1, ..., 24 rate estimates of healthy donors for each proliferation, differentiation and death rate.

Weighted principal component analysis

To compress the information of rate estimates and investigate the variation in parameter values between individuals, we performed a weighted principal component analysis (PCA, Figure 4). We projected the 24

parameters of each MDS patient to a low dimensional space which is spanned by the n_{PC} principal components that explain most of the variance between the individuals. Let X be an n_{rates} x $n_{patients}$ matrix containing the n_{rates} rate estimates for $n_{patients}$ patients and U be an $n_{patients}$ x n_{PC} matrix containing the n_{PC} eigenvectors of the weighted covariance matrix of X with the highest absolute corresponding eigenvalues as rows. The weighted covariance matrix of X is calculated by $log\mathcal{N}(\hat{\mu}_{(i,j)}, \hat{\sigma}_{(i,j)}), j = 1, ..., 10$ where $W \in \mathbb{R}^{n_{rates} \times n_{patients}}$ is the matrix of weights for each observation in X and $\odot$ denotes the element-wise product. The weights take into account rate estimates, their 95% confidence intervals and data richness, i.e. how many data points were available to perform the parameter estimation. Let $CI^u_{(i,j)}$ and $CI^l_{(i,j)}$ be the upper and lower bounds of the profile likelihood-based confidence interval of parameter $r_{i,j}$ estimated based on data of individual j. The rate-specific weights are then given by $w^r_{(i,j)} := \frac{r_{max_i} - r_{min_i}}{CI^u_{(i,j)} - CI^l_{(i,j)}}$, which is higher for low parameter uncertainty as this corresponds to a small confidence interval in the denominator. The data-specific weights are given by $w^{\mathcal{D}}_{(i,j)} = \frac{n^{\mathcal{D}}_{obs,j}}{\sum_{j=1}^{n_{patients}} n^{\mathcal{D}}_{obs,j}} \quad \forall i = 1, ..., n_{rates}$, where $n^{D}_{obs,j}$ is the number of observations in data set D belonging to individual j. They are higher the more data points were observed and used to fit the respective rates for j. Upon normalization, the overall weight $w_{(i,j)}$ of parameter i and sample j is defined by $w_{(i,j)} = \frac{\frac{w^r_{(i,j)}}{\sum_{i=1}^{n_{rates}} \sum_{j=1}^{n_{patients}} w^r_{(i,j)}} + w^{\mathcal{D}}_{(i,j)}}{2}$. We transformed X to the principal component space by calculating $Y = X \cdot U$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details of experiments, including statistical tests and number of patient replicates can be found in the results section. The significance level (alpha) was set to 0.05. The significance level is depicted in the figures as: * $p < 0.05$, ** $p < 0.01$.

Before computational modeling, data quality of HSC sorting and FACS analyses was verified. Two MDS cases had a very low (<70%) HSC purity after sorting and were excluded from further analysis. Samples with very low counts of viable cells (<200 absolute cell counts) in FACS analysis were analyzed by computational modeling and later checked if the generated data could be fit by the mathematical model similar to higher quality samples. On this basis, we excluded three more MDS cases.