Fate mapping of hematopoietic stem cells reveals two pathways of native thrombopoiesis

Mina N. F. Morcos, Congxin Li, Clara M. Munz, Alessandro Greco, Nicole Dressel, Susanne Reinhardt, Katrin Sameith, Andreas Dahl, Nils B. Becker, Axel Roers, Thomas Höfer & Alexander Gerbaulet

Hematopoietic stem cells (HSCs) produce highly diverse cell lineages. Here, we chart native lineage pathways emanating from HSCs and define their physiological regulation by computationally integrating experimental approaches for fate mapping, mitotic tracking, and single-cell RNA sequencing. We find that lineages begin to split when cells leave the tip HSC population, marked by high Sca-1 and CD201 expression. Downstream, HSCs either retain high Sca-1 expression and the ability to generate lymphocytes, or irreversibly reduce Sca-1 level and enter into erythro-myelopoiesis or thrombopoiesis. Thrombopoiesis is the sum of two pathways that make comparable contributions in steady state, a long route via multipotent progenitors and CD48^{hi} megakaryocyte progenitors (MkPs), and a short route from HSCs to developmentally distinct CD48^{-/lo} MkPs. Enhanced thrombopoietin signaling differentially accelerates the short pathway, enabling a rapid response to increasing demand. In sum, we provide a blueprint for mapping physiological differentiation fluxes from HSCs and decipher two functionally distinct pathways of native thrombopoiesis.
immunophenotypic HSCs (lineage<sup>−</sup> Sca-1<sup>+</sup> CD117<sup>+</sup> (LSK) CD48<sup>−</sup>/CD150<sup>+</sup>; Supplementary Fig. 1) are a heterogeneous population with regard to proliferative activity, lineage fate and repopulation capacity after transplantation. In particular, numerous HSC transplantation studies have described, to varying degrees, lineage fate biases of HSCs (reviewed in ref. 1). The production rate of mature cells from HSCs is massively enhanced after transplantation compared to physiological conditions<sup>4</sup>, and it remains unknown whether lineage differentiation pathways also differ between the two settings<sup>5</sup>. Recent in vivo barcoding studies of HSCs in native hematopoiesis have revealed HSC clones with different lineage output<sup>5</sup>,<sup>8</sup> which appear to resemble those seen after HSC transplantation<sup>4</sup>,<sup>10</sup>. Among them are HSCs producing only megakaryocyte progenitors (MkPs) independent of multipotent progenitors (MPPs)<sup>7</sup>,<sup>9</sup>,<sup>10</sup>. These findings raise the questions of the developmental relationships of distinct HSC and progenitor subpopulations in native hematopoiesis and, importantly, how distinct lineage pathways are regulated under physiological conditions to match the demand in mature cells (Fig. 1a). These questions are intimately linked with another controversy that has arisen in the context of the recent HSC fate mapping studies, suggesting steady but infrequent input from HSCs to hematopoiesis<sup>2</sup>,<sup>11</sup>,<sup>12</sup> or more active contribution<sup>13</sup>,<sup>14</sup> Here, we measure and integrate extensive data on HSC differentiation, proliferation, and molecular heterogeneity, all obtained under physiological conditions, to address these fundamental questions. We identify a HSC source population with high expression of CD201 and Sca-1, from which increasingly committed HSCs emerge. Finally, we identify a hitherto elusive MkP subpopulation that directly links HSCs to megakaryocytes.

**Results**

Physiological hematopoiesis emerges from CD201<sup>hi</sup> Sca-1<sup>hi</sup> HSCs. In previous HSC fate mapping studies, the actual stem cell source that supplies hematopoiesis over the lifetime of the mouse, called tip HSCs, has remained elusive<sup>11</sup>,<sup>13</sup>,<sup>15</sup>. The defining criterion for such tip HSCs is that after label induction, constant labeling frequency is maintained by perfect self-renewal; conversely, if an HSC subpopulation emerges from another source, its labeling frequency will approach the labeling frequency of this source with time<sup>16</sup>. Indeed, in existing HSC fate mapping data, a several-fold increase of HSC labeling frequency over at least one year has been observed that is fueled by a yet undefined subpopulation of tip stem cells within the phenotypic HSCs<sup>11</sup>,<sup>13</sup>,<sup>15</sup>. To selectively label HSCs for fate mapping, we generated the Fgd5<sup>ZsGreen:CreERT2</sup>/R26<sup>LSL-tdRFP</sup> mouse model (Supplementary Fig. 2a), which combines an inducible Cre expressed in HSCs<sup>16</sup> (Supplementary Fig. 2b) with an excision reporter with high activation threshold<sup>19</sup>,<sup>20</sup> (Supplementary Fig. 2c). Indeed, tamoxifen (TAM) application labeled HSCs with high specificity (Fig. 1b and Supplementary Fig. 2d–g), thus minimizing the problem of substantial concomitant labeling of progenitors downstream of HSCs seen in several studies<sup>13</sup>–<sup>15</sup>

Within the abundant LS–K<sup>−</sup> and L–K<sup>−</sup> populations, sporadic RFP<sup>+</sup> cells were detected 10 days after TAM induction (Supplementary Fig. 2d, e), but such cells represented only a negligible fraction of these progenitor populations (Supplementary Fig. 2f).

To characterize HSC heterogeneity, we used surface levels of CD201 (also termed EPCR) and Sca-1, as either marker alone has previously been shown to correlate with repopulation potential of transplanted HSCs<sup>21</sup>–<sup>25</sup>. RFP-labeled cells were strongly enriched

---

**Fig. 1** Fate mapping of CD201<sup>hi</sup> Sca-1<sup>hi</sup> HSCs. a Subdivision of the immuno-phenotypic HSC population (LSK CD48<sup>−</sup>/CD150<sup>+</sup>) by CD201 (EPCR) and Sca-1 expression; ES-HSC, CD201 (EPCR)<sup>hi</sup> Sca-1<sup>hi</sup> HSCs. Relations between HSC populations in native hematopoiesis and how they connect to lineage pathways will be studied. b Bone marrow populations of Fgd5<sup>ZsGreen:CreERT2</sup>/R26<sup>LSL-tdRFP</sup> mice were analyzed 10–11 days after TAM induction for RFP expression (n = 20, 4 independent experiments, means and individual values are shown; one-way ANOVA with Sidak correction; ns, not significant; p values shown in graph); for flow cytometry phenotypes see supplementary Fig. 1a. c To map HSC fate over time, Fgd5<sup>ZsGreen:CreERT2</sup>/R26<sup>LSL-tdRFP</sup> mice (n = 18, 10, 12, 10, 8, and 12 mice/consecutive timepoint (1.5–92 weeks of chase)) were TAM-induced, groups of mice were sacrificed at indicated time points, and bone marrow populations and peripheral blood leukocytes were analyzed for RFP expression (pooled data from five independent experiments are shown). d Percentage of RFP<sup>+</sup> HSCs and ES HSCs (mean and SEM). RFP-labeling of ES HSCs did not significantly rise (two-sided F-test on linear regression), while labeling of total HSCs did significantly increase after the catch-up phase (two-sided paired t-test, p-values for time points 1, 2, and 3: <0.001, 0.0057, and 0.0034). e Fractions of labeled peripheral blood and bone marrow cells relative to ES HSCs at indicated time points; the labeling frequencies of the cell populations from each individual were normalized to its ES HSC labeling at indicated time points; the labeling recovery is maintained by perfect self-renewal; conversely, if an HSC subpopulation emerges from another source, its labeling frequency will approach the labeling frequency of this source with time.
among CD201hi Sca-1hi HSCs (ES HSCs, Fig. 1b, Supplementary Fig. 2d–g). These ES HSCs were strongly enriched for quiescent cells and exhibited a primitive immuno-phenotype as judged by lower expression, compared to non-ES HSCs, of CD34^hi, CD48^lo, and CD117^lo, while CD150^hi was higher expressed (Supplementary Fig. 2h).

We induced a large cohort (n = 82) of Fgd5^ZsGreen:CreER(T2)/R26^SL-tdRFP mice to follow label propagation over time (Fig. 1c). To quantify label propagation from ES HSCs through the hematopoietic system, we calculated the labeling frequencies of downstream bone marrow and peripheral blood populations relative to ES HSC labeling at final analysis of a given individual, thus normalizing for variability of initial RFP labeling frequency in ES HSCs (Supplementary Fig. 2j, k). All progenitor and mature cell populations slowly increased their labeling frequencies, trailing HSCs. Labeling of MPPs increased with considerable delay and did not equilibrate with HSC labeling within 21 months (Fig. 1e), which implies that progression of cells from ES HSCs to circulating mononuclear BM cells. We reasoned that obtaining these data for all cell populations in a stem cell differentiation pathway will also yield the pathway topology. To examine this, we generated a toy model for a generic branching lineage topology in silico (Supplementary Fig. 4b) and simulated fate mapping and mitotic tracking data. We then constructed all principally possible topologies linking stem cells, progenitors, and mature cells (Supplementary Fig. 4c), and asked which topology is compatible with the simulated data. Statistical model selection based on the simulated data with added noise indeed recovered the true pathway and rejected all others (Supplementary Fig. 4d, e). By contrast, using the in silico fate mapping data alone did not single out the correct model (Supplementary Fig. 4f). This example with computationally simulated data suggests that lineage pathways downstream of tip HSCs can be inferred from integrating experimental data on stem cell fate mapping, mitotic history, and cell numbers.

The myeloid lineage diverges within phenotypic HSCs. We applied this inference approach to our experimental fate mapping data, complementing them with measurements of H2B-GFP label dilution, and stem and progenitor cell numbers measured relative to lineage-negative (lin^-) cells. In addition to subdivision of HSCs by CD201 and Sca-1 expression, we stratified MPPs and HPCs by Sca-1 expression. This yielded a graph of all possible pathway topologies, with unidirectional differentiation steps (as implied by transplantation studies) and the possibility of reversible Sca-1 loss (Fig. 2a and Supplementary Fig. 4g, h). We systematically enumerated 144 pathways (subgraphs) that connect ES HSCs to common-myeloid progenitors (CMPs) and CLPs (Supplementary Fig. 5a) and translated these into kinetic equations describing cell abundance, RFP label propagation, and H2B-GFP dilution (Supplementary Methods, section 1). The resulting pathway models were ranked by their ability to fit the experimental data, using the bias-corrected Akaike information criterion (AICc). Remarkably, a single model was substantially better than the next-ranked models (ΔAICc > 2, Fig. 2b and Supplementary Fig. 5b), showing that high Sca-1 levels were either inherited from precursor (e.g., Sca-1hi HSC) to progeny (e.g., Sca-1hi MPP) or lost irreversibly (Fig. 2c). This model fit all the experimental data (Fig. 2d–f; for parameter values see Supplementary Table 1 and Supplementary Fig. 5c). The next best ranking models (within ΔAICc 10, indicated by orange bars in Fig. 2b; Supplementary Fig. 5b) were variants of the best model, all showing parallel Sca-1hi and Sca-1lo differentiation pathways and irreversible conversion from Sca-1hi to Sca-1lo cells at every level of the stem and progenitor cell hierarchy (Supplementary Fig. 6a, b). Models that either lacked inheritance of Sca-1 level (Fig. 2b, green bars, non-parallel) or allowed Sca-1hi to Sca-1lo transitions (Fig. 2b, blue bars, non Sca-1-high-to-low), or both (Fig. 2b, gray bars), did not account for the experimental data. These key features of model selection were recapitulated when only 50% of the experimental data set was used (Supplementary Fig. 6c), implying that the data robustly discriminate between the alternative models.

Combining fate mapping with mitotic tracking yields lineage paths. We asked whether the detailed and time-resolved HSC fate mapping data (Fig. 1e) contain information on lineage pathways. In the case of lymphoid development, label from ES HSCs clearly arrived in progenitors before it entered mature cells (e.g., common-lymphoid progenitors (CLPs) and B cells, respectively, in Fig. 1e). Applying this principle to our data ruled out certain topologies linking stem cells, progenitors, and mature cells (Fig. 2f). Therefore, tracing RFP^+ ES HSCs allowed us to infer the behavior of the total ES HSC population. Taken together, these data show that ES HSCs are a self-renewing HSC population that replenish the remainder of HSCs during physiological hematopoiesis, thus placing ES HSCs upstream of other phenotypic HSCs in the hematopoietic hierarchy.

To quantify label propagation from ES HSCs through the hematopoietic system, we calculated the labeling frequencies of downstream bone marrow and peripheral blood populations relative to ES HSC labeling at final analysis of a given individual, thus normalizing for variability of initial RFP labeling frequency in ES HSCs (Supplementary Fig. 2j, k). All progenitor and mature cell populations slowly increased their labeling frequencies, trailing HSCs. Labeling of MPPs increased with considerable delay and did not equilibrate with HSC labeling within 21 months (Fig. 1e), which implies that progression of cells from ES HSCs to circulating mononuclear BM cells. We reasoned that obtaining these data for all cell populations in a stem cell differentiation pathway will also yield the pathway topology. To examine this, we generated a toy model for a generic branching lineage topology in silico (Supplementary Fig. 4b) and simulated fate mapping and mitotic tracking data. We then constructed all principally possible topologies linking stem cells, progenitors, and mature cells (Supplementary Fig. 4c), and asked which topology is compatible with the simulated data. Statistical model selection based on the simulated data with added noise indeed recovered the true pathway and rejected all others (Supplementary Fig. 4d, e). By contrast, using the in silico fate mapping data alone did not single out the correct model (Supplementary Fig. 4f). This example with computationally simulated data suggests that lineage pathways downstream of tip HSCs can be inferred from integrating experimental data on stem cell fate mapping, mitotic history, and cell numbers.

The myeloid lineage diverges within phenotypic HSCs. We applied this inference approach to our experimental fate mapping data, complementing them with measurements of H2B-GFP label dilution, and stem and progenitor cell numbers measured relative to lineage-negative (lin^-) cells. In addition to subdivision of HSCs by CD201 and Sca-1 expression, we stratified MPPs and HPCs by Sca-1 expression. This yielded a graph of all possible pathway topologies, with unidirectional differentiation steps (as implied by transplantation studies) and the possibility of reversible Sca-1 loss (Fig. 2a and Supplementary Fig. 4g, h). We systematically enumerated 144 pathways (subgraphs) that connect ES HSCs to common-myeloid progenitors (CMPs) and CLPs (Supplementary Fig. 5a) and translated these into kinetic equations describing cell abundance, RFP label propagation, and H2B-GFP dilution (Supplementary Methods, section 1). The resulting pathway models were ranked by their ability to fit the experimental data, using the bias-corrected Akaike information criterion (AICc). Remarkably, a single model was substantially better than the next-ranked models (ΔAICc > 2, Fig. 2b and Supplementary Fig. 5b), showing that high Sca-1 levels were either inherited from precursor (e.g., Sca-1hi HSC) to progeny (e.g., Sca-1hi MPP) or lost irreversibly (Fig. 2c); this model fit all the experimental data (Fig. 2d–f; for parameter values see Supplementary Table 1 and Supplementary Fig. 5c). The next best ranking models (within ΔAICc 10, indicated by orange bars in Fig. 2b; Supplementary Fig. 5b) were variants of the best model, all showing parallel Sca-1hi and Sca-1lo differentiation pathways and irreversible conversion from Sca-1hi to Sca-1lo cells at every level of the stem and progenitor cell hierarchy (Supplementary Fig. 6a, b). Models that either lacked inheritance of Sca-1 level (Fig. 2b, green bars, non-parallel) or allowed Sca-1hi to Sca-1lo transitions (Fig. 2b, blue bars, non Sca-1-high-to-low), or both (Fig. 2b, gray bars), did not account for the experimental data. These key features of model selection were recapitulated when only 50% of the experimental data set was used (Supplementary Fig. 6c), implying that the data robustly discriminate between the alternative models.

The selected model (Fig. 2c) predicts that erythro-myeloid commitment requires downregulation of Sca-1 expression either in HSCs or in progenitors (Fig. 2g). Specifically, multiple possible routes lead to CMPs. Based on the differentiation rates inferred from the experimental data, we computed the probability that a
Fig. 2 Statistical inference of HSC differentiation pathways. a Model scheme, with subdivision of HSCs and progenitors according to Sca-1 expression and, for HSCs, CD201 expression. Straight arrows denote cell differentiation (vertical), or change in Sca-1 expression (horizontal), or both (diagonal); curved arrows denote cell proliferation. All possible differentiation are superimposed. b Model selection by bias-corrected Akaike information criterion (AICc) against experimental data for HSC fate mapping, mitotic history, and relative cell numbers. Models are classified (and color-coded accordingly) by two features: parallel differentiation pathways and irreversible loss of Sca-1. The best-fitting model has \( \Delta \text{AICc} = 0 \). None of the remaining 143 models is within \( \Delta \text{AICc} < 2 \). Further models with acceptable evidence (\( \Delta \text{AICc} < 10 \)) are all variants of the best-fitting model (see Supplementary Fig. 5b). c Best model selected by experimental data. Arrow width signifies values of the corresponding differentiation or proliferation rates, while node size quantifies cell number (relative to lin\(^-\) cells) in each compartment at 100 days of chase; see Supplementary Table 1. Fits of the best model for RFP labeling frequency propagation (d), compartment size (e), and H2B-GFP dilution (f) (experimental data: mean ± SEM; mice in d and e are the same as in Fig.1c–e; for f, \( n = 3, 4, 4, 2, 2, 4, 4 \) mice/consecutive timepoint (0–21 weeks of chase) were analyzed; model fit: straight line, best fit; gray-shaded areas, 95% confidence bounds). g Inferred lineage pathways to CMPs and their relative contributions to generation of CMPs (showing the probability that a CMP was generated by each pathway; error bars, 95% prediction bands). h As predicted by the model, transplanted Sca-1\(^{hi}\) HSCs, but not Sca-1\(^{lo}\) HSCs, yielded lymphoid progeny. 100 Sca-1\(^{hi}\) or Sca-1\(^{lo}\) donor HSCs were competitively transplanted and contribution to peripheral blood neutrophils, B and T cells was determined (data re-analyzed from Morcos et al.\(^{25}\)). Lineage bias of donor-derived leukocytes was calculated (Neutrophil chimerism/[B + T cell chimerism])/2; \( n = 4 \) recipient mice for Sca-1\(^{hi}\) HSCs and \( n = 3 \) for Sca-1\(^{lo}\) HSCs in primary and secondary transplantation; mean and SEM are shown). i Ratio of cell fluxes between Sca-1\(^{lo}\) HPC-1 → CMP and Sca-1\(^{hi}\) HPC-1 → CLP for young and old mice (error bars, 95% prediction bands).
CMP was generated by any particular pathway (Supplementary Methods, section 2). The majority (~2/3) of CMPs could be traced back to CD201−/lo Sca-1lo HSCs (Fig. 2g, pathways 1 and 5), downstream followed by massive selective proliferation in the Sca-1lo pathway. Hence, in unperturbed hematopoiesis, a major commitment to erythro-myeloid differentiation occurs already at the HSC stage. This prediction is further supported by the outcome of competitive transplantation of immuno-phenotypic HSCs (LSK CD48−/lo CD150−) purified according to their Sca-1 expression level, which yielded lower and preferentially myeloid reconstitution of recipients transplanted with Sca-1lo HSCs, while Sca-1hi donor HSCs resulted in durable engraftment and a balanced output (Fig. 2h).

In transplantation experiments, aged HSCs produce more myeloid than lymphoid output compared to HSCs from young mice35. To study whether this myeloid bias occurs during physiological hematopoiesis, we induced RFP label in HSCs of mice >1.5 years of age by TAM administration (Supplementary Fig. 7a–f), traced label propagation as described above for mice labeled as young adults (Fig. 1), and generated mitotic tracking data with H2B-GFP dilution for aged mice (Supplementary Fig. 7g–i). As previously described, aged mice exhibited increased HSC frequency36,37 and thrombocythemia37 (Supplementary Fig. 7b, c). The share of ES HSCs among the total HSC population was significantly larger in aged animals (Supplementary Fig. 7e), while the labeling frequency of ES HSCs in old and young animals remained constant over time (Supplementary Fig. 7d). The fundamental dynamics of label propagation from HSCs to downstream progenitor compartments were similar between aged and young mice (Supplementary Fig. 7f), except for faster label increase in total HSCs, MKPs, and platelets in aged animals.

Using the mathematical model (Fig. 2c), we inferred the differentiation, proliferation and death rates of aged HSPCs from the combined fate mapping, mitotic tracking, and cell number data, based on an accurate model fit within experimental error (Supplementary Fig. 7j). We found that the differentiation rates from CD201−/lo Sca-1hi HSCs to Sca-1lo MPPs and from Sca-1lo HPC-1 to CMPs significantly increased (Supplementary Table 1), as did the respective cell fluxes (Supplementary Fig. 7k); by contrast, the faster label propagation from ES HSCs to the remainder of the HSCs (Supplementary Fig. 7k) was accounted for by the increase in ES HSC number (Supplementary Fig. 7l). As a result of these changes with aging, the flux of CMP production relative to CLP production was increased more than five-fold in aged mice, revealing an accentuated myeloid bias with aging in physiological hematopoiesis (Fig. 2i).

Finally, we noticed an extraordinarily high loss rate of CD201−/lo Sca-1lo HSCs (Fig. 2c, arrow labeled “High loss rate”), which is unlikely to be explained solely by cell death as the loss rate then would be higher than the differentiation rate to Sca-1lo MPPs. This puzzling observation will be addressed in the following.

CD201−/lo Sca-1lo HSCs feed thrombopoiesis via CD48−/lo MKPs. MPPs have been thought to supply all lineage pathways38,39, but recent work indicated that platelets can be generated directly from HSCs in mice7,9,10 and humans40. Indeed, when comparing labeling frequencies in committed progenitors and mature cell types relative to MPPs in our own data, only platelets showed significantly faster and higher label acquisition than both MPPs and CMPs (Fig. 3a and Supplementary Fig. 8a, b), arguing against thrombopoiesis proceeding solely via MPPs. By contrast, CD201−/lo Sca-1lo HSCs had higher labeling frequency than platelets throughout the entire chase period (Fig. 1e). These data suggest that the high loss rate of CD201−/lo Sca-1lo HSCs in the model fit (Fig. 2c) was actually caused by their yet unidentified differentiation into megakaryocyte progenitors, implying two distinct pathways of thrombopoiesis (Fig. 3b). To test this idea, we performed further fate mapping experiments that included MKPs. We found that MKPs had consistently lower RFP labeling frequencies than platelets (Fig. 3e). Thus, the fate mapping data rule out a homogeneous MKP population that is supplied by the conventional pathway via MPPs and the direct pathway from HSCs. As MKPs are well established as precursors of megakaryocytes (Mks)11,14,42, we searched for solutions for this paradox. The simplest working model, as confirmed by mathematical analysis of RFP label propagation, suggests that the conventional and the direct thrombopoiesis pathways are active simultaneously and proceed through distinct MKP subpopulations (Supplementary Fig. 8c, d).

To gain unbiased insight into putative differentiation trajectories from HSCs to MKPs, we performed single-cell RNA sequencing (scRNAseq) of index-sorted HSCs, MPPs, HPC-1, and LS K progenitors including MKPs, using Smart-seq243. These populations occupied distinct but partially overlapping regions in the transcriptional landscape computed by PHATE34 (Fig. 3d, upper left panel). The different kinds of progenitor cells were well identified by the phenotypic markers used, as indicated by preferential expression of myeloid, erythroid, and megakaryocyte lineage gene modules3 in myeloid progenitors, erythroid progenitors, and MKPs, respectively (Fig. 3d). Moreover, Sca-1 was most highly expressed in HSCs at the apex of the transcriptional landscape (Fig. 3e), which also expressed high levels of Wvef and Isg21h (Fig. 3f) as previously described45,46. We noted that MKPs showed heterogeneous surface expression of CD48 (Fig. 3g), with CD48−/lo MKPs being significantly closer to Sca-1lo HSCs in transcript space (Fig. 3h). To extract putative developmental trajectories between HSC and progenitor subpopulations, we applied the graph abstraction method PAGA47 (Fig. 3i). PAGA recovered the topology inferred from fate mapping (Fig. 3i, red lines) with high confidence (Fig. 3j). Remarkably, for thrombopoiesis, only CD48−/lo MKPs were directly connected to CD201−/lo Sca-1lo HSCs (Fig. 3i, green line) whereas CD48hi MKPs were connected to HPC-1 via LS K progenitors (Fig. 3i, blue lines). Thus, single-cell transcriptome data point to distinct origins of CD48−/lo and CD48hi MKPs. Moreover, the significant correlation of pseudotime and Mk score in each MKP subpopulation individually suggests that each subset matured in parallel in a continuous process (Fig. 3k).

PAGA also inferred a putative link between CD48−/lo and CD48hi MKPs, indicating that they could be connected. To further address the origins of the MkP subpopulations, we performed in vivo fate mapping. Indeed, fate mapping label was rapidly acquired by CD48−/lo MKPs in induced Fdg2ZaGreen:CreERT2/ R26Lsl-tdRFP mice, whereas CD48hi MKPs acquired less label at slower kinetics (Fig. 4a, b). The CD48−/lo MKP labeling quickly exceeded its own initial labeling as well as of all other progenitors, suggesting that CD48−/lo MKPs directly acquire label from HSCs. Consistent with pseudotime analysis (Fig. 3k), the complete lack of label equilibration between the two MkP subsets implies that they are located in two distinct pathways (rather than CD48−/lo MKPs being upstream of CD48hi MKPs). The LSK CD48hi CD150+ population (HPC-2), which has been implied in thrombopoiesis7,15,39, could not be upstream of CD48−/lo MKPs as evidenced by their consistently lower RFP labeling, but might give rise to CD48hi MKPs (Supplementary Fig. 8e). As predicted by PAGA, CD48−/lo MKPs were developmentally close to HSCs, as they were the only cell type among all LS K progenitors which retained or inherited significant amounts of H2B-GFP in chased R26ERT2/Col1A1H2B-GFP mice (Fig. 4c). Functional characterization of both MKPs by culture and transplantation (Fig. 4d, e and...
Supplementary Fig. 8f–h) confirmed their predominant thrombocytic potential with some residual oligopotent behavior. While most cultivated CD48 hi MkP gave rise to a single Mk, significantly more CD48 lo MkP proliferated before acquiring a megakaryocytic morphology. Likewise, transplantation demonstrated enhanced platelet generation from CD48 lo MkPs. While both MkP populations similarly expressed megakaryocytic markers (Supplementary Fig. 8i–k), CD48 hi MkPs expressed higher levels of myeloid genes and CD48 lo MkP harbored an erythroid signature (Supplementary Fig. 8j–k), which was further evidenced by their robust erythroid potential upon transplantation (Fig. 4e).

To summarize, combining the data from RNAseq, fate mapping, and mitotic tracking, we identified a pathway of thrombopoiesis proceeding directly from CD201–/lo Sca-1–/lo HSCs to CD48–/lo MkPs that is used in parallel with the conventional pathway of platelet generation via MPPs and CD48 hi MkPs (Fig. 4f).

Modeling label propagation through these pathways now gave excellent agreement with the experimental data (Fig. 4g). We used the model to infer how many megakaryocytes are produced per day through these distinct pathways (Fig. 4h). These pathway fluxes are the product of differentiation rate (which is higher for CD48–/lo MkPs, consistent with transplantation outcome (cf. Fig. 4e)) and precursor number (which is higher for CD48 hi MkPs), resulting in equal contributions of both pathways to steady-state thrombopoiesis in young mice (Fig. 4f–i). The higher differentiation rate of CD48–/lo MkPs, compared to CD48 hi MkPs is sustained by both their own proliferation and influx from CD201–/lo Sca-1–/lo HSCs (Supplementary Table 1). Moreover, the
Fig. 3 Direct and conventional thrombopoiesis pathways proceed via distinct progenitors. a RFP labeling of cell populations from Fgd5ZsGreen::CreERT2/R26SL-tdRFP mice (same as in Fig. 1c–e) as shown as fraction of label in the MPPs of the same animal at final analysis. b Model scheme with direct route to MkPs (LSL-K CD41+/CD150+). c Fgd5ZsGreen::CreERT2/R26SL-tdRFP mice were TAM-induced and indicated cell populations were analyzed at 10, 59, and 120 days (n = 5 animals/time point, means and SEM are shown). d PHATE embedding of single-cell transcriptomics of index-sorted hematopoietic stem and progenitor cells (n = 820 cells total) isolated from 6 R26rtTA/Col1A1H2B-GFP animals, colored by cell type and expression of lineage markers (blue: low; yellow: high). e, hSCs, MPPs, and HPCs-1 show a decrease of Sca-1 surface expression in proximity of MkPs and myeloid-committed progenitors. f PHATE plots showing Vwf and Itga2b expression. g Lower surface expression of CD48 in MkPs residing closer to HSCs in the transcriptional landscape. h Proximity in transcriptional space of MKP subsets to HSCs (left), and of HSC subsets to CD48+/MKPs (right), placing CD48+/MKPs close to Sca-1+/HSCs. (two-sided Wilcoxon test; left plot: n = 135 cells, right plot: n = 279 cells; p-value ≤ 10–5 for both comparisons; boxplot description: central line indicates median, upper (lower) hinges indicate third (first) quartiles, top (bottom) whisker extends to largest value within a 1.5× inter-quartile range from top (bottom) hinge). i PAGA embedding of surface marker-defined populations, displaying transcriptional proximity of CD201/−/Sca-1−/HSC to CD48+/−/MKP (green link), and a distinct differentiation path to CD48+/MKP (blue) via myeloid (MyP, LS−K CD41+ CD150−) and erythroid progenitors (EryP, LS−K CD41− CD150−). Red PAGA links were previously inferred by statistical model selection (see Fig. 2c). Cell populations were identified via indexed surface markers, except for lymphoid-biased HPCs-1, which were identified by their transcriptional signature. j Confidence score for PAGA links. k Comparison between Mk score and pseudotime shows parallel Mk maturation trajectories in CD48−/lo and CD48+/hi MKP subsets. Gray areas represent 95% confidence bounds for linear regression, R indicates the Pearson correlation coefficient, p-value is relative to a two-sided alternative hypothesis.

Thrombopoietin signaling enhances direct thrombopoiesis. The differentiation rate of HSCs is thought to respond to the need for mature cell types after injury or other challenges, but a direct assay for quantifying HSC differentiation in physiological hematopoiesis has been lacking. We reasoned that HSC fate mapping may provide such an assay. To examine this, we exposed TAM-induced Fgd5ZsGreen::CreERT2/R26SL-tdRFP and DOX-pulsed R26rtTA/Col1A1H2B-GFP mice to the myeloablative drug 5-Fluorouracil (5-FU), and investigated how ES HSCs contribute to hematopoietic recovery (Fig. 5a, b). 5-FU subpopulations of CD48−/lo MkPs is equivalent to ~3 additional cell divisions in all HSC and MPP subpopulations (Supplementary Fig. 10b, c). In CD48+/hi MkPs, 5-FU diluted the H2B-GFP label to the range of the background control. By contrast, CD48−/lo MkPs still exhibited H2B-GFP labeling after recovery from 5-FU treatment, which further supports their direct emergence from HSCs. Taken together, these experiments show that HSC fate mapping provides an assay for changes in HSC output in response to perturbation.

We used this assay to study the effects of thrombopoietin signaling, which is a physiological regulator of thrombopoiesis as well as HSC maintenance. To this end, we treated Fgd5ZsGreen::CreERT2/R26SL-tdRFP and R26rtTA/Col1A1H2B-GFP mice with the thrombopoietin receptor (Mpl) agonist, romiplostim, for 5 consecutive days (Fig. 5d, and Supplementary Fig. 10a, d–f). As previously reported, this regimen expanded all LSK subpopulations except MPPs (Supplementary Fig. 10f) and raised platelet counts (Fig. 5e). To understand how this is achieved, we studied label propagation from ES HSCs to platelets. RFP labeling frequencies of total HSCs including ES HSCs were not affected by romiplostim treatment (Fig. 5f–g). However, we observed strong increases in RFP-labeling frequencies in downstream progenitors, continuing a trend shown by previous studies employing a different Cre driver (5-FU–/Col1A1H2B-GFP) (Fig. 5g and Supplementary Fig. 10a). In addition, romiplostim administration to R26rtTA/Col1A1H2B-GFP mice induced ~1 additional divisions in HSCs and progenitors (Supplementary Fig. 10d, e). To understand how the acceleration of differentiation and enhanced cell division conspire to increase platelet production, we fitted the mathematical model to the experimental data (Fig. 4f and Supplementary Fig. 11a–c). The flux through the direct pathway was increased at all stages from ES HSCs to megakaryocytes, resulting in an almost 10-fold increase of the pathway flux (Fig. 5h). By comparison, the flux through the conventional pathway was only moderately upregulated (Fig. 5b, i). Taken together, thrombopoietin signaling enhances platelet production by channeling CD201−/lo Sca-1− HSCs into the direct thrombopoiesis pathway (Fig. 5j).

Discussion

We have demonstrated that stem cell fate mapping charts physiological differentiation pathways when combined with proliferation assays and mathematical inference. We thus pinpoint tip stem cells of native hematopoiesis and show how lineage pathways diverge immediately downstream of tip HSCs. Specifically, we find that thrombopoiesis proceeds via two distinct progenitors that previously were thought to be a homogeneous population. Our identification of MkP subpopulations has been enabled by closely linking single-cell transcriptomics with fate mapping and mitotic tracking. Moreover, the accuracy of predictions derived from computational model selection attest to the utility of this tool for interrogating complex, multi-faceted data. Our results on HSC self-renewal and differentiation rates are consistent with previous studies employing a different Cre driver (Tie2MeriCreMer) to label HSCs for fate mapping. Surface markers of stem and progenitor cells have been selected for their utility in prospective isolation of these cells, and less emphasis has been given to their molecular function. Interestingly, however, Sca-1, marking tip-HSCs together with CD201, is an interferon-stimulated gene (ISG). Constitutive expression of ISGs, with protective function, is a hallmark of stem cells also beyond the hematopoietic system. CD201 has been implied in HSC retention to the niche. Our functional data on tip HSCs imply that...
the gene expression program associated with high co-expression of Sca-1 and CD201 deserves closer scrutiny. Our identification of a short pathway of thrombopoiesis, leading from Sca-1<sup>lo</sup> HSCs via CD48<sup>–/lo</sup> MkPs to megakaryocytes, strengthens the recently advanced view that thrombopoiesis is special among all adult hematopoietic lineage pathways by emerging directly from HSCs, and provides a specific progenitor for this pathway<sup>7,9,40</sup>. Contrary to the classical view<sup>15,38</sup>, but in agreement with Rodriguez-Fraticelli et al.<sup>7</sup>, this pathway does not proceed via MPP. Our fate mapping and single cell transcriptome data uncovered that each MkP subset matured independently and that the cell flux of CD48<sup>–/lo</sup> MkP was accelerated as this two-fold smaller population had higher proliferative potential and contributed equally to platelet production.

While platelets, neutrophils, and erythrocytes make up the major share of hematopoietic cell production in steady state, their amplification strategies are fundamentally different: Neutrophil and erythrocyte numbers are achieved by massive progenitor proliferation, whereas platelet numbers are obtained via prolonged shedding from megakaryocytes after enormous cell growth. This special amplification strategy is linked to the evolution of anucleated platelets in mammalian hematopoiesis<sup>54</sup>. We now find that these two strategies for amplifying cell numbers are already manifest at the HSC level. The major cell flux from Sca-
1° HSCs supplies thrombopoiesis directly, while myelopoiesis is fed by a much smaller flux from Sca-1° HSCs and amplified by selective proliferation of Sca-1° progenitors downstream. Moreover, the short thrombopoiesis pathway can produce platelets more rapidly than the conventional pathway, as CD48°/lo MkPs differentiate faster into megakaryocytes than their CD48 hi counterparts. This feature may explain the higher flux through the short pathway under conditions of increased demand for platelet production during aging and upon applying a thrombopoietin receptor agonist. These findings would be consistent with the view that the conventional pathway may reflect a residual ability of myeloid progenitors to give rise to platelets. This is further supported by our finding that the differentiation flux from heterogeneous CMPs to myeloid and erythroid lineages is much larger than to the conventional thrombopoiesis pathways.

Our observation that activating a physiological regulator of thrombopoiesis further accentuates this effect (increasing the flux from Sca-1° HSCs to thrombopoiesis and accelerating, in a compensatory manner, proliferation of Sca-1° MPps) strengthens this interpretation. Thus, our data show that HSCs contribute very strongly to thrombopoiesis in native hematopoiesis, whereas they feed only infrequently into myelopoiesis, which is sustained by a much smaller flux from Sca-1° HSCs to thrombopoiesis and accelerating, in a compensatory manner, proliferation of Sca-1° MPps. More-
clone ebio1560, 1:100, eBioscience) and anti-CD41 (BV605, clone MWReg30, 1:100, Biolegend) for HSC and early progenitor analysis in Fgd5^{ZsGreen:CreERT2}/R26^{SL-tdRFP} –animals or 2. anti-Streptavidin (Horizon V500, 1:800, BD), anti-CD117(c-kit) (APC/eF780, clone 2B8, 1:1600, eBioscience), anti-Ly-6A/E (Sca-1) (PCP/Cy5.5, clone D7, 1:400, eBioscience), anti-CD34 (eF660, clone D7, 1:400, eBioscience), anti-CD127 (PE/Cy7, clone A7R34, 1:200, Biolegend), anti-CD135 (BV421, clone A2F10, 1:100, Biolegend) and anti-CD16/32 (AF700, clone 93, 1:100, eBioscience) for committed progenitor (CMP, GMP, MEP, CLP) analysis in Fgd5^{ZsGreen:CreERT2}/R26^{LSL-tdRFP} –animals, or 3. anti-Streptavidin (eF710, 1:200, eBioscience), anti-CD117(c-kit) (APC/eF780, clone 2B8, 1:1600, eBioscience), anti-Ly-6A/E (Sca-1) (PCP/Cy5.5, clone D7, 1:400, eBioscience), anti-CD150 (PE/Cy7, clone TC15-12F12.2, 1:200, Biolegend), anti-CD48 (PE, clone HM48-1, 1:400, eBioscience), anti-CD201 (EPCR) (APC, clone ebio1560, 1:100, eBioscience), anti-CD34 (eF450, clone RAM34, 1:25, eBioscience) and anti-CD41 (BV605, clone MWReg30, 1:100, Biolegend) for R26^{rtTA}/Col1A1^{H2B-GFP}–animals was performed.
**Fig. 5** Fate mapping of perturbed hematopoiesis. a-c Fgd5ZsGreen;CreERT2/R26SL-tdRFP mice were TAM-induced, treated with S-FU (n = 5) or left untreated (n = 5), sacrificed at 35 days after treatment, and bone marrow populations and peripheral blood leukocytes were analyzed for RFP expression. b Ratios of peripheral blood cell numbers in treated (S-FU) vs. untreated (untr.) animals 6 days after treatment (means and SD are shown). c Label in populations 35 days after treatment is shown as percentage of RFP+ cells in ES HSC (left graph) or normalized as fraction of label in ES HSC (right graph, mean with SEM, one-way ANOVA with Sidak correction). d-g Fgd5ZsGreen;CreERT2/R26SL-tdRFP mice were TAM-induced, treated with romiplostim (n = 4) or saline (NaCl, n = 4), and sacrificed 11 d after start of treatment and hematopoietic populations were analyzed for RFP expression. e Time-course of platelet numbers (means and SD are shown). f RFP labeling in ES HSCs (means and SEM are shown). g Label in bone marrow and blood populations is shown as fraction of label in ES HSC in respective populations (mean with SEM, one-way ANOVA with Sidak correction). h Graphical overview of altered cell fluxes (number of cells differentiated per day) by romiplostim application. Orange arrows indicate minor or non-significant changes. (cf. Supplementary Fig. 1). I Inferred ratio of cell fluxes via CD48+/lo MkP to cell fluxes via CD48hi MkP at steady state (NaCl, gray) or after romiplostim (orange) perturbation. J Cell fluxes from CD201-/lo Sca-1- HSCs into Sca-1+MPPs and CD48-/lo MkPs at steady state (NaCl, gray) or after romiplostim (orange) perturbation. In I and J, the error bars represent 95% confidence bounds.

For cell surface phenotype analysis of peripheral blood cells, cells either underwent erythrocyte lysis and were stained with anti-CD3 (APC, clone 145-2C11, 1:100, or PCP/eF710, clone eBio500A2, 1:400), anti-CD11b (APC/eF780, clone M1/70, 1:600, eBio), anti-CD45 (PCP/eF710, clone eBio1D3, 1:2000), and anti-Ly-6C/G (Gr1) (eF450, clone RB6-8C5, 1:1200), all from eBioScience, or were stained directly with anti-CD41 (APC, clone eBioMWRe30, 1:50, ebioscience) and anti-Ter119 (FITC, clone TER-119, 1:100, Biolegend).

Peripheral blood cells of transplanted B6.CD45.1/2 mice were stained with anti-CD3 (PCP/eF710, clone eBio500A2, 1:400), anti-CD11b (APC/eF780, clone M1/70, 1:600, eBio), anti-CD45 (PCP/eF710, clone eBio1D3, 1:2000), and anti-Ly-6C/G (Gr1) (eF450, clone RB6-8C5, 1:1200) from eBioScience, and anti-CD41 (APC, clone a20, 1:800) and anti-CD45.2 (FITC, clone a104, 1:400) from Biolegend.

Bone marrow cells of transplanted B6.CD45.1/2 mice were stained with anti-Streptavidin (Horizon V500, 1:900, BD), anti-CD11b (PCP/eF710, clone eBio421, 1:400, BD, anti-CD45.2 (FITC, clone a104, 1:400) and anti-CD45.2 (BV711, clone a104, 1:100) or with anti-Streptavidin (eF710, 1:200, ebioscience), anti-CD117 (c-kit) (APC/eF780, clone CB6.2B8, 1:1600, eBioscience), anti-CD117 (c-kit) (APC/eF780, clone CB6.2B8, 1:1600, eBioscience), anti-CD117 (c-kit) (PCP/eCy5.5, clone D7, 1:400, ebioscience), anti-CD48 (BV421, clone HM48-1, 1:400, BD), anti-CD45.2 (APC, clone a20, 1:800, Biolegend) and anti-CD45.2 (BV271, clone a104, 1:100) or with anti-Streptavidin (eF710, 1:200, ebioscience), anti-CD117 (c-kit) (APC/eF780, clone CB6.2B8, 1:1600, eBioscience), anti- Ly-6A/E (Sca-1) (PCP/eCy5.5, clone D7, 1:400, ebioscience), anti-CD150 (PE/Cy7, clone TC15-12F22, 1:200, Biolegend), anti-CD48 (BV421, clone HM48-1, 1:400, BD), anti-CD201/EPCR (PCP/eF710, clone eBio1560, 1:200, ebioscience), anti-CD48 (BV421, clone HM48-1, 1:400, BD), anti-CD201/EPCR (PCP/eF710, clone eBio1560, 1:200, ebioscience), anti-CD48 (BV421, clone HM48-1, 1:400, BD), anti-CD201/EPCR (PCP/eF710, clone eBio1560, 1:200, ebioscience) from Biolegend, anti-CD117 (c-kit) (APC, clone a20, 1:800, Biolegend) and anti-CD45.2 (FITC, clone a104, 1:400).

For double staining, antibodies, manufacturers, catalog numbers and clone numbers please refer to Supplementary Table 2. For a detailed overview of gating strategies and cell population abundance refer to Supplementary Fig. 1. Absolute numbers of lin− BM cells were determined employing a Miltenyi MACSquant flow cytometer.

**Single cell culture.** Single cells were deposited into 96 well U-bottom plate (TPP) using a BD FACs ARIA II cell sorter. Cells were cultured in Iscove’s Modified Dulbecco’s Medium (Gibco) supplemented with 10% FCS (FBS superior, Biochrom), 30 µl Penicillin/Streptomycin, 0.1 µg/ml PeproTech, 5 µl Trpophoetin (mpTPO 20 mg/ml PeproTech), Interleukin-3 (mlL 3-20 ng/ml PeproTech), Erythropoietin (rEPO, 5U/ml NeoRecomon, Roche). Culture wells were visually inspected and counted by light microscopy (PrimoVert, Zeiss) and cultured until reaching steady state (72 °C 5 min) with 1× concentrated KAPA HotStart ReadyMix and 300 nM dual indexing primers. After PCR, single cell libraries of samples were sequenced with AccuBlue Broad range chemistry, equimolarly pooled, and purified twice with 1× volume Sera-Mag SpeedBeads. This was followed by Illumina 30 bp paired-end sequencing of 60–80 million reads per sample. One out of three repeats could not be sequenced from the remaining two plates were used for outlier detection. Normalization was performed using the pooling method combined with scaling normalization within each plate to account for different coverages (multiBatchNorm function in the R limma package). Mutual nearest neighbor correction was also tested, confirming the results from the linear model (Supplementary Fig. 12). Feature selection was performed using variance modeling based on spike-in transcripts and selecting genes with a non-zero biological variance with an FDR threshold of 0.05, then excluding genes contained in the cell cycle entry in Gene Ontology, 10 principal components were retained and used to compute PHATE embedding and transcriptomic distance. Lineage modules were detected using KAPA HiFi HotStart ReadyMix and 300 nM dual indexing primers. After PCR, single cell libraries of samples were sequenced with AccuBlue Broad range chemistry, equimolarly pooled, and purified twice with 1× volume Sera-Mag SpeedBeads. This was followed by Illumina 30 bp paired-end sequencing of 60–80 million reads on the Illumina NovaSeq 6000 platform in S4 PE mode.

**Single-cell transcriptome analysis.** Raw reads were mapped to the mouse genome (mm10) and splice-site information from Ensembl release 8760 with gnap (version 2018-07-04)61. Uniquly mapped reads and gene annotations from Ensembl were used as input for featureCounts (version 1.6.2)62 to create counts per gene and cell. Outlier cells in quality control metrics (library size, number of detected genes, ERCC, and mitochondrial reads) were excluded. One out of three pipelines exhibited a large number of poor quality cells, thus quality control metrics from the remaining two plates were used for outlier detection. Normalization was performed using the pooling method combined with scaling normalization within each plate to account for different coverages (multiBatchNorm function in the R batchelor package). Batch effects were regressed out of the data using a linear model (removeBatchEffect method in the R limma package). Mutual nearest neighbor correction was also tested, confirming the results from the linear model (Supplementary Fig. 12). Feature selection was performed using variance modeling based on spike-in transcripts and selecting genes with a non-zero biological variance with an FDR threshold of 0.05, then excluding genes contained in the cell cycle entry in Gene Ontology, 10 principal components were retained and used to compute PHATE embedding and transcriptomic distance. Lineage modules were detected using the FindMarkers function (Widoloxon rank sum test) in the R package scen, scanned by plate of origin. Code is available from Github (https://github.com/hoever-lab/ThrombopoiesisRNAseq).

**Bulk transcriptome analysis.** Basic quality control of the resulting sequencing reads was performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and RNA-SeqC (version 1.1.8)64. Reads were aligned to the mm10 reference genome using GSNAP (version 2020-12-16)65; the Ensembl gene...
annotation version 98 was used for splice site detection. Uniquely aligned reads per gene were counted using featureCounts (version 2.0.1) and the respective Ensembl annotation. Normalization of raw read counts based on library size, and testing for differential gene expression between conditions was performed using the DESeq2 R package (version 1.30.1, R version 3.5.1). 763 genes with an absolute fold-change ≥ 1 and an adjusted p-value ≤ 0.05 were considered significantly differentially expressed. Lineage-specific gene sets were obtained from Msigdb using the MsigdbR package and querying the cell type collection (C8) for lineage-related keywords. Gene set enrichment analysis was performed using the clusterProfiler R package. 75 The analysis can be reproduced by using the code at https://github.com/hoefer-lab/Thrombopoiesis_RNAseq.

Mathematical modeling. We use ordinary differentiation equations to describe the dynamics of label propagation, compartment sizes, and H2B-GFP dilution (for details see Supplementary Methods). Computational inference was performed using the trust-region method to search for the local minimum of the negative log-likelihood. To find the global minimum, the initial parameter values were chosen by Latin hypercube sampling (5000 samples for each model). The tasks above were performed using Data2Ddynamic (D2D) framework, a Matlab-based open source package. 68 The source codes and data used in the modeling can be found at https://github.com/hoefer-lab/Thrombopoiesis_Modeling.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The scRNA sequencing data generated in this study have been deposited in the NCBI Geo database under accession number GSE139390. The bulk RNA sequencing data generated in this study have been deposited in the NCBI Geo database under accession number GSE143409. Source data are provided with this paper.

Code availability

The transcriptome analysis and modeling codes are available at Github (https://github.com/hoefer-lab/Thrombopoiesis_RNAseq and https://github.com/hoefer-lab/Thrombopoiesis_Modeling).

Received: 22 December 2020; Accepted: 11 July 2022; Published online: 03 August 2022

References

1. Haas, S., Trumpf, A. & Milosom, M. D. Causes and consequences of hematopoietic stem cell heterogeneity. Cell Stem Cell 22, 627–638 (2018).
2. Sun, J. et al. Clonal dynamics of native hematopoiesis. Nature 514, 322–327 (2014).
3. Rodríguez-Fraticelli, A. E. et al. Single-cell lineage tracing unveils a role for TCF15 in hematopoiesis. Nature 583, 585–589 (2020).
4. Laurenti, E. & Göttgens, B. From hematopoietic stem cells to complex differentiation landscapes. Nature 553, 418–426 (2018).
5. Pei, W. et al. Polylox barcoding reveals hematopoietic stem cell fate realized in vivo. Nature 548, 456–460 (2017).
6. Pei, W. et al. Resolving fates and single-cell transcriptomes of hematopoietic stem cell clones by PolyloxExpress barcoding. Cell Stem Cell https://doi.org/10.1016/j.stem.2020.07.018 (2020).
7. Rodríguez-Fraticelli, A. E. et al. Clonal analysis of lineage fate in native hematopoiesis. Nature 533, 212 (2018).
8. Bowling, S. et al. An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cell. Cell 181, 1410–1422.e1427 (2020).
9. Carrell, J. et al. Hierarchically related lineage-restricted fates of multipotent hematopoietic stem cells. Nature 554, 106–111 (2018).
10. Sawai, C. et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell 154, 1112–1126 (2013).
11. Busch, K. et al. Fundamental properties of unperturbed hematopoiesis from stem cells in vivo. Nature 518, 542–546 (2015).
12. Takahashi, M. et al. Reconciling flux experiments for quantitative modeling of normal and malignant hematopoietic stem/progenitor dynamics. Stem Cell Rep. 16, 741–753 (2021).
13. Sawai, C. et al. Hematopoietic stem cells are the major source of multilineage hematopoiesis in adult animals. Immunity 45, 597–609 (2016).
14. Chapple, R. H. et al. Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. Blood Adv. 2, 1200–1228 (2018).
15. Säwen, P. et al. Murine HSCs contribute actively to native hematopoiesis but with reduced differentiation capacity upon aging. Elife https://doi.org/10.7554/eLife.41528 (2018).
16. Höfer, T. & Rodewald, H. R. Differentiation-based model of hematopoietic stem cell functions and lineage pathways. Blood 132, 1106–1113 (2018).
17. Barle, M. et al. Hematopoietic stem cells self-renew symmetrically or gradually proceed to differentiation. Preprint at bioRxiv https://doi.org/10.1101/2020.08.06.239186 (2020).
18. Gao, Z. et al. Fgfl5 identifies hematopoietic stem cells in the murine bone marrow. J. Exp. Med. 211, 1315–1331 (2014).
19. Luche, H., Weber, O., Nageswara Rao, T., Blum, C. & Fehling, H. J. Faithful activation of an extra bright red fluorescent protein in “knock in” Cre reporter mice ideally suited for lineage tracing studies. Eur. J. Immunol. 37, 43–53 (2007).
20. Álvarez-Aznar, A. et al. Tamanoxifen-independent recombination of reporter genes limits lineage tracing and mosaic analysis using CreERT2 lines. Transgenic Res. 29, 53–68 (2020).
21. Balazs, A. B., Fabian, A. J., Panczyk, C. T. & Mulligan, R. C. Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. Blood 107, 2317–2321 (2006).
22. Kent, D. G. et al. Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. Blood 113, 6342–6350 (2009).
23. Rabe, J. L. et al. CD34 and EPCR coordinately enrich functional murine hematopoietic stem cells under normal and inflammatory conditions. Exp. Hematol. 81, 1–15.e16 (2020).
24. Wilson, N. et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. Cell Stem Cell 16, 712–724 (2015).
25. Morcos, M. N. F. et al. SCA-1 expression level identifies quiescent hematopoietic stem and progenitor cells. Cell Stem Cell 8, 1472–1478 (2017).
26. Shin, J. Y., Hu, W., Naramura, M. & Park, C. Y. High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. J. Exp. Med. 211, 217–231 (2014).
27. Morita, Y., Ema, H. & Nakauchi, H. Heterogeneity and hierarchy within the most primitive hematopoietic stem compartment. J. Exp. Med. 207, 1173–1182 (2010).
28. Upadhyaya, S. et al. Kinetics of adult hematopoietic stem cell differentiation in vivo. J. Exp. Med. https://doi.org/10.1084/jem.201810316 (2018).
29. Höfer, T., Busch, K., Klapproth, K. & Rodewald, H. R. Fate mapping and quantitation of hematopoiesis in vivo. Ann. Rev. Immunol. 34, 449–478 (2016). 0732-0582.
30. Morcos, M. N. F. et al. Continuous mitotic activity of primitive hematopoietic stem cells in adult mice. J. Exp. Med. https://doi.org/10.1084/jem.20191284 (2020).
31. Burnham, K. P. & Anderson, D. R. Model Selection and Multimodel Inference - A Practical Information-Theoretic Approach (Springer, 2002).
32. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & de Haan, G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. J. Exp. Med. https://doi.org/10.1084/jem.20111490 (2011).
33. de Haan, G., Nijhoff, W. & Van Zant, G. Mouse strain-dependent changes in frequency and proliferation of hematopoietic stem cells during aging: A correlation between lifespan and cycling activity. Blood 89, 1543–1550 (1997).
34. Grover, A. et al. Single-cell RNA sequencing reveals molecular and functional platelet bias of aged hematopoietic stem cells. Nat. Commun. 7, 11075 (2016).
35. Boyer, S. W., Schroeder, A. V., Smith-Berdan, S. & Forsberg, E. C. All hematopoietic cells develop from hematopoietic stem cells through Flk2/Frco-positive progenitor cells. Cell Stem Cell 9, 64–73 (2011).
36. Pietras, Eric M. et al. Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. Cell Stem Cell 17, 35–46 (2015).
37. Notta, F. et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science 351, aab2116 (2016).
38. Hornor, T. N., Miyamoto, T. & Weissman, I. L. Characterization of mouse clonogenic megakaryocyte progenitors. Proc. Natl Acad. Sci. USA 100, 205–210 (2003).
42. Pronk, C. J. H. et al. Elucidation of the phenotypic, functional, and molecular topography of a myeloid progenitor cell hierarchy. Cell Stem Cell 13, 1339–1349 (2016).

43. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. 9, 171–181 (2014).

44. Moon, K. R. et al. Visualizing structure and transitions in high-dimensional biological data. Nat. Biotechnol. 37, 1482–1492 (2019).

45. Sanjuan-Pla, A. et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature 502, 232–236 (2013).

46. Gekas, C. & Graf, T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. Blood 121, 4463–4472 (2013).

47. Wolf, F. A. et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. 20, 29 (2019).

48. Culmer, D. L. et al. Circulating and vein wall P-selectin promote venous thrombogenesis during aging in a rodent model. Thrombosis Res. 131, 42–48 (2013).

49. Dayal, S. et al. Hydrogen peroxide promotes aging-related platelet hyperactivation and thrombosis. Circulation 127, 1308–1316 (2013).

50. Kovtunyuk, L. V., Manz, M. G. & Takizawa, H. Enhanced thrombopoietin but not G-CSF receptor stimulation induces self-renewing hematopoietic stem cell divisions in vivo. Blood 127, 3175–3179 (2016).

51. Holmes, C. & Stanford, W. L. Concise review: Stem cell antigen-1: Expression, function, and enigma. Stem Cells 25, 1339–1347 (2007).

52. Wu, X. et al. Intrinsically immune shapes viral resistance of stem cells. Cell 172, 423–438 e425 (2018).

53. Gur-Cohen, S. et al. PAR1 signaling regulates the retention and recruitment of EPCR-expressing bone marrow hematopoietic stem cells. Nat. Med. 21, 1307–1317 (2015).

54. Martin, J. F. & Wagner, G. P. The origin of platelets enabled the evolution of eutherian placentaion. Biol. Lett. 15, 20190374 (2019).

55. Schoedel, K. et al. The bulk of the hematopoietic stem cell population is dispensable for murine steady-state and stress hematopoiesis. Blood 128, 2285–2296 (2016).

56. Sheikh, B. N. et al. MOZ (KAT6A) is essential for the maintenance of murine hematopoietic stem cells. Blood 128, 2307–2318 (2016).

57. Foudi, A. et al. Analysis of histone 2B-GFP retention reveals slowly cycling biological data. Nat. Biotechnol. 27, 84–90 (2008).

58. Lallemand, Y., Luria, V., Haffner-Krausz, R. & Lonai, P. Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. Transgenic Res. 7, 105–112 (1998).

59. Hu, Y. & Smyth, G. K. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J. Immunol. Methods 347, 70–78 (2009).

60. Zerbino, D. R. et al. Ensemble 2018. Nucleic Acids Res. 46, D754–D761 (2017).

61. Nacu, S. & Wu, T. D. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26, 873–881 (2010).

62. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2013).

63. Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol. 17, 75 (2016).

64. Delacla, D. S. et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 28, 1530–1532 (2012).

65. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).

66. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550 (2005).

67. Yu, G., Wang, L.-G., Han, Y., He, Q.-Y. clusterProfiler: an R Package for comparing biological themes among gene clusters. OMICS: J. Integrative Biol. 16, 1089–1098 (2012).

68. Raue, A. et al. Data2Dynamics: A modeling environment tailored to parameter estimation in dynamical systems. Bioinformatics 31, 3558–3560 (2015).

Acknowledgements T.H. was supported by the German Research Council (DFG) through SFB 873 (B11) and SFB 1129 (B11), Marie Sklodowska-Curie Actions H2020-MSA-ITN-2017 (Quantitative T-Cell Immunology and Immunotherapy, project number: 764698), and DKFZ core funding. A. Gerbaulet was supported by DFG (GE3038/1-1) and Fritz Thyssen Stiftung (Az.10.19.1.013MN). A.R. and A.D. were funded by DFG (RO2133/10-1) in the setting of FOR2577. M.N.F.M. and C.M.M. were supported by the Dresden International Graduate School for Biomedicine and Bioengineering, granted by DFG in the context of the Excellence Initiative. We thank Christa Haase, Madeleine Rickauer and Livia Schulze for technical assistance and all members of Höfer and Gerbaulet groups for discussion.

Author contributions A.G. and T.H. conceptualized the study. Experimental work and data analysis was carried out by M.N.F.M. and C.M.M. with assistance from N.D. T.H., C.L., and N.B.B. conceptualized and performed mathematical modeling and statistical inference. Sequencing was performed by S.R. and A.D. A. Greco and K.S. performed transcriptome analyses. A.R. advised the study. T.H., A.G., M.N.F.M., C.M.M., and C.L. wrote the manuscript. All authors read, commented, and agreed on the final version of the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Competing interests The authors declare no competing interests.

Additional information Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-31914-z.

Correspondence and requests for materials should be addressed to Thomas Höfer or Alexander Gerbaulet.

Peer review information Nature Communications thanks Leila Perté, Alejo Rodriguez-Fraticelli, and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022