

Hierarchically related lineage-restricted fates by multipotent hematopoietic stem cells

Joana Carrelha^{1,2}, Yiran Meng^{1,2}, Laura M. Kettyle^{3,4}, Tiago C. Luis^{1,2}, Ruggiero Norfo^{1,2}, Verónica Alcolea^{1,2}, Hanane Boukarabila^{1,2}♦, Francesca Grasso^{4,5}, Adriana Gambardella², Amit Grover², Kari Höglstrand^{3,4}, Allegra M. Lord^{3,4}, Alejandra Sanjuan-Pla²†, Petter S. Woll^{4,5}, Claus Nerlov^{2*}, Sten Eirik W. Jacobsen^{1-5*}

¹Haematopoietic Stem Cell Laboratory, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DS, UK. ²MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DS, UK. ³Department of Cell and Molecular Biology, Wallenberg Institute for Regenerative Medicine, Karolinska Institutet, Stockholm SE-171 77, Sweden. ⁴Department of Medicine Huddinge, Center for Hematology and Regenerative Medicine, Karolinska Institutet, Stockholm SE-171 77, Sweden. ⁵Karolinska University Hospital, Stockholm SE-171 77, Sweden.

♦Present address: Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center and University of Cincinnati, Cincinnati, OH 45229, USA.

†Present address: Hematology Research Group, IIS La Fe, Valencia 46026, Spain.

*These authors contributed equally to this work.

Rare multipotent hematopoietic stem cells (HSCs) in adult bone marrow (BM) with extensive self-renewal potential possess the ability to efficiently replenish all myeloid and lymphoid blood cells¹, securing long-term multilineage reconstitution following physiological and clinical challenges, including chemotherapy and hematopoietic transplantations²⁻⁴. HSC-transplantation remains the only curative treatment for many hematological malignancies, but inefficient blood-lineage replenishment remains a major cause of morbidity and mortality^{5,6}. Single cell transplantation has uncovered considerable heterogeneity among reconstituting HSCs⁷⁻¹¹, supported by findings in unperturbed hematopoiesis^{2-4,12} and suggested to reflect different propensities for lineage-fate decisions by distinct myeloid-, lymphoid- and platelet-biased HSCs^{7-10,13}. Other studies suggested that such lineage-bias might reflect generation within the phenotypic HSC compartment of unipotent or oligopotent self-renewing

progenitors, and implicated uncoupling of the defining HSC properties of self-renewal and multipotency^{11,14}. Here, highly sensitive tracking of progenitors and mature cells of the megakaryocyte/platelet, erythroid, myeloid, B and T cell lineages produced from singly transplanted HSCs revealed a highly organized, predictable and stable framework for lineage-restricted fates of long-term self-renewing HSCs. Most notably, a distinct class of HSCs adopts a fate towards effective and stable replenishment of a megakaryocyte/platelet-lineage tree but not other blood cell lineages, despite sustained multipotency, whereas no HSCs contribute exclusively to any other single blood-cell lineage. Single multipotent HSCs can also fully restrict towards simultaneous replenishment of megakaryocyte, erythroid and myeloid lineages without executing their sustained lymphoid lineage potential. Genetic lineage tracing supports an important role of platelet-biased HSCs also in unperturbed adult hematopoiesis. These findings uncover a limited repertoire of distinct HSC subsets, defined by a predictable and hierarchical propensity to adopt a fate towards replenishment of a restricted set of blood lineages, prior to loss of self-renewal and multipotency.

Previous studies of HSC lineage-bias transplanted CD45.2 HSCs into CD45.1 recipients⁷⁻⁹, allowing tracking of myeloid (M) and B- and T-lymphoid progeny, but not replenishment of mature erythrocytes (E) and platelets (P), since they lack CD45 expression. Since recent studies also tracking E and/or P suggested distinct replenishment of these lineages^{10,11}, we purified Lin⁻Sca1⁺c-Kit⁺CD34⁻CD150⁺CD48⁻ (LSK34⁻150⁺48⁻) cells, containing most if not all HSCs¹⁵⁻¹⁷, from BM of adult mice expressing enhanced green fluorescent protein (EGFP) from the *Gata1*-promoter (*Gata1*-EGFP)¹⁸ and tdTomato from the von Willebrand-factor promoter (*Vwf*-Tomato; Extended Data Fig. 1) to enrich platelet-primed *Vwf*^{pos} HSCs¹⁰. BM-ablated

CD45.1 mice were transplanted with a single CD45.2 *Vwf*-Tomato^{mid-high} (*Vwf*-Tomato^{pos}) LSK34⁻150⁺48⁻ cell in competition with CD45.1 BM-cells. Through highly specific and sensitive analysis of large numbers of peripheral blood (PB) cells, contributions >0.01% to each blood lineage could reliably be detected (Methods). Almost 40% of transplanted single cells contributed to $\geq 0.1\%$ of at least one lineage at 16-18 weeks post-transplantation, representing a candidate self-renewing long-term (LT)-HSC (Fig. 1a; Extended Data Fig. 2a). While approximately 50% replenished all 5 lineages (P, E, M, B and T), the other half showed no contribution ($\leq 0.01\%$) to one or more lineages at all analysis time points, despite other lineages being robustly and stably replenished (Fig. 1b-d; Extended Data Fig. 2, 3), establishing a stringent definition of lineage-restricted reconstitution (Methods). If blood lineage-restricted reconstitution patterns by HSCs were randomly distributed, numerous lineage-restriction combinations would be expected, but only a few closely related patterns were observed in more than 1,000 single HSC transplanted mice. Most notably, 10% of long-term reconstituting *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cells robustly and stably reconstituted platelets but no other lineages at any time-point (Fig. 1b-d; Extended Data Fig. 2b,d,e, 3a-b). No case was observed in which any other blood lineage was exclusively reconstituted (Fig. 1c-d; Extended Data Fig. 2e). In fact, the only additional stable lineage-restricted reconstitutions observed were P+E, P+E+M, and P+E+M+B, meaning that the only lineage invariably reconstituted was platelets (Fig. 1c,d; Extended Data Fig. 2c,d,e, 3c-e). Notably, also PE- and PEM-restricted replenishment consistently showed a strong P-bias (Fig. 1c,e; Extended Data Fig. 2d,f), although P-output was gradually reduced in mice with PEMB- versus PEM- versus P-restricted replenishment (Fig. 1c; Extended Data Fig. 2d; PEMB>PEM>P-restricted, $p < 0.05$ all time points; one-sided Wilcoxon-Mann-Whitney test). Parallel

analysis of spleens confirmed the lineage-restricted reconstitution (Extended Data Fig. 3g-h). Previous studies used c-Kit-deficient (W^{41}/W^{41}) recipients and/or competitor cells^{7,8,10}, which might not only provide competitive advantage for transplanted wild-type HSCs but potentially also for development of blood cell lineages. Herein, almost identical lineage-restricted patterns, at similar frequencies (no significant differences), were observed whether single V_{wf} -Tomato^{pos} LSK34⁻150⁺48⁻ cells were transplanted in competition with c-Kit-deficient (W^{41}/W^{41} ; Fig. 1d) or wild-type (Extended Data Fig. 2e) BM-cells. In contrast, lineage-bias was affected by type of competitor, with a consistent tendency for enhanced lymphoid output in the presence of WT competitor cells (Fig. 1e, Extended Fig. 2f-g). This is compatible with the intrinsic propensity for distinct HSCs to adopt lineage-restricted fates (lineage-restriction) being more resistant to extrinsic influences, than to different lineage biases.

The most primitive LSK150⁺48⁻ HSCs lack expression of CD229¹⁷. LSK150^{hi}48⁻CD229⁻ cells were enriched for V_{wf} -Tomato expression (Fig. 1f, Extended Fig. 4a), and transplanted single LSK34⁻150^{hi}48⁻CD229^{low/neg} cells were enriched for P-biased reconstitution when compared to LSK34⁻150^{hi}48⁻CD229^{hi} cells which were enriched for long-term lymphoid-biased reconstitution (Fig. 1g-h). No further distinct enrichment for V_{wf} -Tomato expression was observed based on CD41 expression (Extended Fig. 4b). Therefore, CD229 could partially replace the need for a transgenic V_{wf} -reporter to enrich platelet-biased HSCs.

Based on a different strategy for isolation and tracking of cells within the HSC-compartment, previous studies proposed the existence therein of lineage-restricted, oligopotent or unipotent, repopulating progenitors, including P-restricted progenitors, downstream of multipotent LT-HSCs¹¹. They reconstituted platelets much less

efficiently and more transiently than multipotent LT-HSCs, suggesting a much more restricted self-renewal potential, confirmed by few sustaining platelets in secondary recipients¹¹. The herein long-term platelet-restricted replenishment by *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ HSCs was far more robust and stable (more than 50-fold higher at 16-18 weeks; Fig. 1c; Extended Data Fig. 2b,d, 3a,b), clearly distinguishing them from described platelet-restricted progenitors. In support of P-restriction already taking place within the *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ LT-HSC compartment, long-term stable P-restricted reconstitution by single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ HSCs was invariably accompanied by robust long-term reconstitution of LSK150⁺48⁻ HSCs (mean 39%), virtually all being *Vwf*-Tomato^{pos} (mean 93%; Fig. 2a; Extended Data Fig. 5a). In contrast, in long-term multilineage reconstituted mice most reconstituted LSK150⁺48⁻ cells were *Vwf*-Tomato^{neg} (mean 7% *Vwf*-Tomato^{pos}; Fig. 2b; Extended Data Fig. 5b).

In mice with P-restricted and PEM-restricted reconstitution, BM and thymus lacked reconstitution of progenitors¹⁹⁻²¹ for lineages absent in PB, whereas robust progenitor-replenishment for reconstituted lineages was observed (Fig. 2a,c; Extended Data Fig. 5a,c). Subsets of multipotent progenitors (MPP) have been suggested to contain heterogeneous populations of lineage-restricted/biased progenitors in addition to true MPPs^{11,17,22,23}. LSKFlt3⁻CD150⁺CD48⁺ (150⁺48⁺) MPPs contain potent P-biased/restricted progenitors²³, LSKFlt3^{high}CD150⁻CD48⁺ lymphoid-primed MPPs (LMPPs)^{24,25}, LSKFlt3⁻CD150⁻CD48⁺ (150⁻48⁺) MPPs myeloid progenitor activity, and LSKFlt3⁻CD150⁻CD48⁻ (150⁻48⁻) MPPs multi-potent MPPs with considerable reconstitution potential^{17,23}. All MPP-compartments were robustly replenished in multilineage reconstituted mice, whereas long-term P-restricted reconstitution showed high chimerism for 150⁺48⁺ MPPs, with no detectable replenishment of 150⁻48⁺

MPPs, LMPPs or 150^+48^- MPPs. In PEM-restricted reconstitution 150^+48^+ and 150^+48^- MPPs were replenished whereas LMPPs and 150^+48^- MPPs remained undetectable in most mice (Fig. 2a-c; Extended Data Fig. 5). Whereas most 150^+48^+ cells in P-restricted reconstituted mice were Vwf -Tomato^{pos} (mean 87%), a small fraction (mean 3%) were Vwf -Tomato^{pos} in multilineage-reconstituted mice, suggesting that Vwf -Tomato^{pos} P-restricted reconstituting LSK 150^+48^- cells replenish a distinct platelet-primed subset of 150^+48^+ MPPs. This detailed BM-analysis of mice exclusively and robustly reconstituted with PB-platelets by a single platelet-primed LSK $34^-150^+48^-$ HSC tentatively identifies a platelet-restricted lineage tree, apparently independent of most implicated MPP-subsets.

Our detailed analysis in mice with P-restricted reconstitution failed to provide evidence for progenitor reconstitution beyond those potentially dedicated to the megakaryocyte-platelet lineage. This is compatible with long-term P-restricted LSK $34^-150^+48^-$ reconstituting cells representing unipotent progenitors with fully P-restricted potential as proposed in previous studies¹¹, although also compatible with self-renewing multipotent LSK $34^-150^+48^-$ platelet-primed LT-HSCs adopting a lineage-restricted fate exclusively toward P-restricted replenishment prior to loss of multipotency. To distinguish between these two fundamentally distinct possibilities, we sorted CD45.2⁺ LSK cells from mice with robust and stable P-restricted reconstitution four months after transplantation of a single Vwf -Tomato^{pos} LSK $34^-150^+48^-$ HSC, and assessed potential to produce granulocytes, monocytes/macrophages and T-lymphocytes *in vitro*²⁶. Remarkably, for all investigated mice with P-restricted reconstitution, definitive morphological, FACS and gene expression data unequivocally established that single Vwf -Tomato^{pos} LSK $34^-150^+48^-$ LT-HSCs exclusively producing platelets *in vivo*, also possess extensive

granulocyte, monocyte/macrophage and T-lymphoid potential *in vitro* (Fig. 2d-g, Extended Data Fig. 6). Therefore, Vwf -Tomato^{pos} LSK34⁻150⁺48⁻ cells executing potent P-restricted LT-reconstituting activity represent self-renewing multipotent HSCs that exclusively adopt a megakaryocyte-platelet lineage fate upon transplantation *in vivo*.

Upon extended analysis for up to 44 weeks, high and dominant P-reconstitution was sustained in all mice, although in 5 out of 13 mice with P-restricted reconstitution at 16-18 weeks, very low-level erythrocyte and/or myeloid reconstitution was also observed (Fig. 3a-b). Self-renewal potential of P-restricted Vwf -Tomato^{pos} HSCs, and whether P-restricted reconstitution potential is intrinsically programmed, was assessed in secondary transplantations. We consistently observed sustained high levels of P-restricted secondary reconstitution, although in several secondary recipients very low levels of erythroid, myeloid and even lymphocyte reconstitution was detected (Fig. 3c), providing further support for the multipotency of single Vwf -Tomato^{pos} HSCs exclusively adopting a fate towards platelet-replenishment in primary recipients. Secondary recipients also displayed highly selective, robust and stable reconstitution of the Vwf -Tomato^{pos} LSK150⁺48⁻ HSC, Vwf -Tomato^{pos} 150⁺48⁺ MPP, and MkP progenitor axis, as in corresponding primary recipients (Fig. 3d-e; Extended Data Fig. 7a). These experiments established that single P- or PE-restricted LT-HSCs, sustain self-renewal and distinct lineage-restricted fates, in secondary recipients.

To more unequivocally establish whether P- or PE-restricted self-renewing LT-HSCs sustain their multipotency, we purified Vwf -Tomato^{pos} LSK150⁺48⁻ cells from the BM of mice with robust P- or PE-restricted reconstitution 4 months after transplantation of a single Vwf -Tomato^{pos} LSK150⁺48⁻ HSC, and transplanted these

into secondary recipients (Extended Data Fig. 7b). Following an additional 4 months, sustaining P- and PE-restricted reconstitution (Fig. 3f), we established that the original single V_{wf} -Tomato^{pos} LSK150⁺48⁻ HSC transplanted had sustained its multipotency (Fig. 3g), despite continuing to strictly adopt a P- or PE-restricted fate *in vivo*. Since these secondary transplantations were performed using fresh BM competitor cells, rather than whole BM cells from the primary recipients as in most studies^{7,8}, they also rule out that the sustained lineage-restricted fate is explained by a distinct and constant composition of competing HSCs and progenitors.

The physiological relevance, if any, of platelet-biased HSCs in steady-state unperturbed hematopoiesis remains unclear, as recent HSC fate mapping studies were not designed to assess this, including clonal tracking by barcoding in which platelets were not assessed since they do not harbour DNA used to retrieve barcodes^{2-4,12}. A V_{wf} -CreERT2 mouse line selectively labelling a fraction of LSK34⁻48⁻ cells with high CD150 expression (Fig. 3h), selectively, robustly and increasingly with time labelled platelets in adult mice (Fig. 3i). Up to 2 months after a short Tamoxifen-pulse, almost exclusively platelets were labelled, subsequently along with a small fraction of erythrocytes and myeloid cells, but virtually no lymphocytes, compatible with the labelling and significant activity of herein identified P- and PEM-biased HSCs in steady-state hematopoiesis.

Single cell index-sorting²⁷ was performed to assess lineage reconstitution as a function of V_{wf} -Tomato expression in LSK34⁻150⁺48⁻ BM cells (Fig. 4a). V_{wf} -Tomato^{mid-high} LSK34⁻150⁺48⁻ fractions showed a higher frequency of repopulating cells than V_{wf} -Tomato^{neg} cells (Fig. 4b). In agreement with the proposed existence of lymphoid-biased HSCs^{7-9,13,28}, a large fraction of mice transplanted with single V_{wf} -Tomato^{neg} LSK34⁻150⁺48⁻ cells showed robust long-term lymphoid-biased or even

lymphoid-restricted PB-reconstitution (Fig. 4c-d), independently of level of CD150 expression (Extended Data Fig. 8a-c). However, in contrast to P-, PE or PEM-restricted reconstituting LT-HSCs (which never showed evidence of replenishment of any other lineages), lymphoid-restricted/biased long-term PB reconstitution was always preceded by transient multi-lineage reconstitution (Fig. 4d-e; Extended Data Fig. 8d,e), more compatible with persistence of long-lived lymphocytes²⁹ or lymphoid progenitors produced by short-term multipotent HSCs than the ongoing activity of a lymphoid-biased HSC. In further support of this, and in agreement with previous studies,⁷⁻⁸ long-term lymphoid-biased or lymphoid-restricted reconstitution was not associated with detectable replenishment of the LSK150⁺48⁻ HSC compartment (Extended Data Fig. 9), and also accompanied by reduced PB and progenitor reconstitution in secondary recipients (Fig. 4f, Extended Data Fig. 10a-c). Thus, while *Vwf*-Tomato^{neg} LSK34⁻150⁺48⁻ short-term repopulating cells might play an important role in replenishment of lymphocytes, detailed PB kinetics and HSC and progenitor reconstitution analysis, fail to support programming of lymphoid-fate restriction at the LT-HSC-level, as here demonstrated for P-, PE-, PEM- and PEMB-restricted HSCs. Excluding lymphoid-biased/restricted and other short-term repopulating cells, distinct patterns of PB-lineage reconstitution were observed among LSK34⁻150⁺48⁻ LT-HSCs with different *Vwf*-Tomato levels (Fig. 4g-k; Extended Data Fig. 10d). Robust long-term multilineage reconstitution was observed in all fractions, with a clear tendency towards increasing platelet-bias with increasing levels of *Vwf*-Tomato expression. P-restricted reconstitution was exclusively a property of *Vwf*-Tomato^{pos} and never seen with *Vwf*-Tomato^{neg} cells. *Vwf*-Tomato^{mid} and *Vwf*-Tomato^{low} P-restricted reconstitution was at lower levels than from *Vwf*-Tomato^{high} cells (Extended Data Fig.

8a). Therefore, multipotent HSCs robustly adopting a P-restricted fate predominantly have a platelet-primed V_{wf} -Tomato^{high} phenotype.

In conclusion, highly sensitive tracking of PB-replenishment of P-, E-, M-, B- and T-cell lineages, and their progenitors, in mice transplanted with single cells establish the existence of four distinct and closely related stages of self-renewing HSCs in adult BM that stably adopt lineage-restricted fates despite remaining multipotent, therefore establishing that a uni-lineage P-restricted fate determination can occur already in self-renewing LT-HSCs with sustained multipotency. This suggests that the programme required for multipotency might also be required for the self-renewal capacity of LT-HSCs adopting a stable maintenance of a P-restricted fate. PEMB-, PEM-, PE- and P-restricted multipotent HSCs all reside in the LSK34⁺150⁺48⁻ V_{wf} -Tomato^{pos} platelet-primed LT-HSC compartment¹⁰, and typically show a distinct and stable platelet-bias in their reconstitution. Although this is compatible with a hierarchy of multipotent HSCs gradually restricting their fate towards exclusive megakaryocyte-platelet production, the hierarchical or non-hierarchical relationship between multipotent HSCs which adopt distinct lineage-restricted fates, remains to be established. Findings in mouse^{10,11} and human³⁰ showing that platelet levels increase and decrease more rapidly than other blood-cell lineages suggest that their progenitors are short-lived, and must therefore be continuously replenished at a high rate, which could explain the need for HSCs dedicated to platelet-replenishment. Despite having transplanted more than 1,000 mice, we found no evidence in support of the existence of LT-HSCs with a fate restricted to any other blood cell lineages. The sustained and serially transplantable P-restricted blood lineage replenishment suggests that the P-restricted fate of multipotent V_{wf} -Tomato^{high} HSCs is intrinsically

programmed, most likely through epigenetic mechanisms. While the LT-HSCs adopting a P-restricted fate sustains multipotency, it remains to be established at what subsequent stage in the progenitor hierarchy established by P-restricted HSCs, multipotency is eventually lost¹¹. Unraveling these mechanisms, and developing approaches to expand P-restricted HSCs, is likely to help ameliorate extensive needs for platelet transfusions following chemotherapy and BM transplantation.

REFERENCES

- 1 Osawa, M., Hanada, K.-I., Hamada, H. & Nakauchi, H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* **273**, 242-245 (1996).
- 2 Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322-327 (2014).
- 3 Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* **518**, 542-546 (2015).
- 4 Sawai, C. M. *et al.* Hematopoietic stem cells are the major source of multilineage hematopoiesis in adult animals. *Immunity* **45**, 597-609 (2016).
- 5 Seggewiss, R. & Einsele, H. Immune reconstitution after allogeneic transplantation and expanding options for immunomodulation: an update. *Blood* **115**, 3861-3868 (2010).
- 6 Pineault, N. & Boyer, L. Cellular-based therapies to prevent or reduce thrombocytopenia. *Transfusion* **51 Suppl 4**, 72S-81S (2011).
- 7 Müller-Sieburg, C. E., Cho, R. H., Thoman, M., Adkins, B. & Sieburg, H. B. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood* **100**, 1302-1309 (2002).
- 8 Dykstra, B. *et al.* Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell* **1**, 218-229 (2007).
- 9 Challen, G. A., Boles, N. C., Chambers, S. M. & Goodell, M. A. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF- β 1. *Cell Stem Cell* **6**, 265-278 (2010).
- 10 Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, 232-236 (2013).
- 11 Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154**, 1112-1126 (2013).
- 12 Pei, W. *et al.* Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* **548**, 456-460 (2017).
- 13 Wang, J. *et al.* Per2 induction limits lymphoid-biased haematopoietic stem cells and lymphopoiesis in the context of DNA damage and ageing. *Nat Cell Biol* **18**, 480-490 (2016).
- 14 Haas, S. *et al.* Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. *Cell Stem Cell* **17**, 422-434 (2015).
- 15 Kiel, M. J., Yilmaz, O. H., Iwashita, T., Terhorst, C. & Morrison, S. J. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109-1121 (2005).
- 16 Yilmaz, O. H., Kiel, M. J. & Morrison, S. J. SLAM family markers are conserved among hematopoietic stem cells from old and reconstituted mice and markedly increase their purity. *Blood* **107**, 924-930 (2006).

- 17 Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, 102-116 (2013).
- 18 Drissen, R. *et al.* Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat Immunol* **17**, 666-676 (2016).
- 19 Nagasawa, T. Microenvironmental niches in the bone marrow required for B-cell development. *Nat Rev Immunol* **6**, 107-116 (2006).
- 20 Bhandoola, A., von Boehmer, H., Petrie, H. T. & Zúñiga-Pflücker, J. C. Commitment and developmental potential of extrathymic and intrathymic T cell precursors: plenty to choose from. *Immunity* **26**, 678-689 (2007).
- 21 Pronk, C. J. *et al.* Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell* **1**, 428-442 (2007).
- 22 Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118-1129 (2008).
- 23 Pietras, E. 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).
- 24 Mansson, R. *et al.* Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* **26**, 407-419 (2007).
- 25 Adolfsson, J. *et al.* Identification of Flt3⁺ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* **121**, 295-306 (2005).
- 26 Böiers, C. *et al.* Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. *Cell Stem Cell* **13**, 535-548 (2013).
- 27 Wilson, N. K. *et al.* Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell* **16**, 712-724 (2015).
- 28 Eaves, A. C. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* **125**, 2605-2613 (2015).
- 29 Sprent, J. & Tough, D. F. Lymphocyte life-span and memory. *Science* **265**, 1395-1400 (1994).
- 30 Vellenga, E. *et al.* Autologous peripheral blood stem cell transplantation in patients with relapsed lymphoma results in accelerated haematopoietic reconstitution, improved quality of life and cost reduction compared with bone marrow transplantation: the Hovon 22 study. *British Journal of Haematology* **114**, 319-326 (2001).

FIGURE LEGENDS

Figure 1 | Stable long-term lineage-restricted reconstitution patterns by single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cells. **a**, Reconstituted mice (mean+s.e.m.; ≥ 1 lineage $\geq 0.1\%$ at 16-18wks) transplanted with single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ HSC and *W⁴¹/W⁴¹* BM support. n=292 mice, 15 experiments. **b**, Analysis of platelet, erythroid-, myeloid-, B- and T-cell reconstitution in P-restricted patterns 16-18wks post-transplantation, representative of >30 mice in >40 single-cell transplantation experiments. **c**, Reconstitution kinetics (mean \pm s.e.m.). P-restricted, n=12 mice; PE-restricted, n=3; PEM-restricted, n=18; PEMB-restricted, n=22; Multilineage, n=54. **d**, Distribution of lineage-restricted reconstitution patterns (lineage positive if $>0.01\%$ in ≥ 1 analysis point; n=109 mice). **e**, Distribution of lineage-bias within restriction patterns in **d** (all plotted in **c**). P-bi frequency: Multi vs PEMB *p=0.019, PEMB vs PE *p=0.024, **p=0.004, ***p=0.0003, ****p<0.001. **f**, *Vwf*-Tomato expression in LSK34⁻150^{hi}48⁻ cells with different CD229 expression. Plots representative of 6 mice in 4 experiments. Percentage of parent gate and median fluorescence intensity (MFI). **g**, Reconstituted mice (mean+s.e.m.) transplanted with single CD229^{low/neg} (n=71 mice) or CD229^{high} (n=70) LSK34⁻150^{hi}48⁻ cells in 3 experiments. **h**, Distribution of lineage-bias by CD229^{low/neg} (n=41 mice) or CD229^{high} (n=33) single cells. *p=0.013; **p=0.002. Statistical comparisons: two-tailed Fisher's exact test (95% CI).

Figure 2 | Single reconstituting LSK34⁻150⁺48⁻ cells establish lineage-restricted hematopoietic hierarchies but remain multipotent. a-c, Reconstitution percentages (mean±s.e.m.) of HSPCs 16-44wks post-transplantation in P-restricted (**a**, n=8 mice, except Pro-B n=5), Multilineage (**b**, n=6, except Pro-B n=3) and PEM-restricted mice (**c**, n=7, except Pro-B n=4 and DP n=6) in 12 experiments. Orange: all mice positive. Grey: no positive mice. Pink: positive/negative mice, frequency and mean±s.e.m. of positives shown. **d**, GM *in vitro* potential of donor-derived LSK cells (mean±s.e.m. per group, 1 cell/well, 60-240 wells/mouse, 20 experiments). Multi, n=10 mice; PEM-restricted, n=5; PE-restricted n=3; P-restricted, n=9. Fraction of positive mice above bars. P vs Multi *p=0.023; PE vs PEM *p=0.047; **p=0.003. **e**, T-cell *in vitro* potential, donor-derived LSK cells (mean±s.e.m., 10 cells/well, 7-72 wells/mouse, 21 experiments). Multi, n=12 mice; PEMB-restricted, n=8; PEM-restricted, n=7; PE-restricted, n=4; P-restricted, n=8. Fraction of positive mice above bars. P vs PEMB *p=0.026; PEM vs Multi and PEMB vs Multi *p=0.011; ****p<0.0001. **f**, FACS-profile of T-cell progenitors generated *in vitro*, representative of 21 experiments. **g**, Gene expression heatmap of *in vitro* T-cell progenitors. CD90.2⁺CD25⁺ samples: Multi, n=3 mice; PEMB-restricted, n=3; P-restricted, n=3. CD4⁺CD8⁺ samples: Multi, n=4 mice; PEMB-restricted, n=2; PEM-restricted, n=1. 2-8 wells/mouse, 5 experiments. Controls: DN2 (CD90.2⁺CD25⁺) thymocytes, n=4 mice; DP (CD4⁺CD8⁺) thymocytes, n=5 mice; CD19⁺CD11b⁻ spleen B-cells, n=4 mice. Mean Ct values per mouse group, normalized to mean Ct of *Hprt1/B2m*. Statistical comparison of means: two-tailed t-test (95% CI).

Figure 3 | Long-term persistence of lineage-restricted reconstitution patterns by multipotent HSCs. **a**, Frequency of strict long-term maintenance of P- and PEM-restricted reconstitution by single *Vwf*-Tomato^{pos} HSCs, observed at all analysis points for a total 20-44wks (P-restricted, n=14 mice; PEM-restricted, n=10; 12 experiments). **b**, Reconstitution kinetics (mean±s.e.m.) in P- and PEM-restricted mice maintained (n=9 each) or changed (n=5) analysed ≥20wks post-transplantation. **c**, Reconstitution (mean±s.e.m of positive mice) 16-18wks post-transplantation in primary and secondary recipients. P-restricted, n=3 primary with 6 secondary, 1-3/donor (left; 2 experiments). Multilineage, n=4 primary with 11 secondary, 1-4/donor (right; 4 experiments). Frequency positive mice above bars. **d-e**, Reconstitution percentages (mean±s.e.m.) of HSPCs 44wks post-transplantation of a single HSC. P-restricted (**d**) and Multilineage (**e**), and for each a corresponding secondary recipient 26wks post-transplantation. Orange: all mice positive. Grey: no positive mice. White: not analysed. **f**, PE- and P-restricted reconstitution, 16-18wks post-transplantation, in primary recipients of a single *Vwf*-Tomato^{pos} HSC and their secondary recipients of donor-derived *Vwf*-Tomato^{pos} HSCs (3 experiments). **g**, GM (1 cell/well, 19-37 wells/mouse) and T cell (10 cells/well, 6-19 wells/mouse) *in vitro* potentials of donor-derived LSK cells from secondary recipients in **f**, 19-21wks post-transplantation (mean±s.e.m., n=5 mice). Fraction of positive mice above bars. **h**, HSC labelling 2-3wks post-Tamoxifen treatment of *Vwf*-CreERT2/R26-Tomato mice (mean±s.e.m % of LSK34-150+48-, n=3). **i**, Blood labelling (mean±s.e.m post-Tamoxifen treatment of *Vwf*-CreERT2/R26-Tomato mice (5 experiments; n=14 mice at 4 and 8wks; n=11 at 12 and 16-18wks). Statistical comparison of means with two-tailed t-test: 4 and 8wks, P vs each lineage, all p<0.0001; 12wks, P vs E p=0.003, P vs

M p=0.004, P vs B p=0.0007, P vs T p=0.00005; 16-18wks, P vs E p=0.0004, P vs M p=0.0005, P vs B and P vs T p<0.0001.

Figure 4 | Lymphoid-biased reconstitution independently of long-term reconstituting HSCs. **a**, V_{wf} -Tomato gating in post-sort analysis of index-sorted LSK34⁺150⁺48⁺ cells. Mean percentages \pm s.e.m., n=14 mice, 14 experiments. **b**, Reconstituted mice (mean+s.e.m.; any lineage \geq 0.1% at 16-18wks) transplanted with a single HSC and WT BM support in 14 experiments. V_{wf}^{neg} , n=107 mice; V_{wf}^{low} , n=43; V_{wf}^{mid} , n=68; V_{wf}^{high} , n=174. *p=0.023; **p=0.007. **c**, Frequency of lymphoid-bias in **b**. V_{wf}^{neg} vs V_{wf}^{low} *p=0.030; V_{wf}^{low} vs V_{wf}^{high} *p=0.021; ****p<0.0001. **d**, Reconstitution kinetics (mean \pm s.e.m.). L-restricted: lymphoid \geq 0.1% and absence of non-lymphoid lineages 16-18wks (n=19 mice). L-biased: lymphoid reconstitution \geq 0.1% and \geq 1 non-lymphoid lineage at 16-18wks (n=5). **e**, Frequency of detection of PB lineages (\geq 0.1% reconstitution) in \geq 1 analysis point (mice in **d**). **f**, Reconstitution kinetics (mean \pm s.e.m.) in secondary recipients of BM from L-restricted (left; n=7 primary mice and 21 secondary mice in 3 experiments, 2-4 secondary/donor) and L-biased mice (right; n=6 primary and 15 secondary in 2 experiments, 2-3 secondary/donor). **g**, Reconstituted mice (mean+s.e.m) with M and/or P \geq 0.1% at 16-18wks, excluding L-biased reconstitution (same mice in **b**). **p=0.0098; ***p=0.0001; ****p<0.0001. **h**, Distribution of restriction patterns in mice in **g**. Ratio of patterns with vs without lymphoid output: V_{wf}^{neg} vs V_{wf}^{high} **p=0.006; V_{wf}^{mid} vs V_{wf}^{high} **p=0.002. **i**, Distribution of lineage-bias in mice in **g**. P-bi frequency: V_{wf}^{neg} vs V_{wf}^{high} **p=0.009; V_{wf}^{mid} vs V_{wf}^{high} **p=0.006. **j-k**, Distribution of lineage restriction (**j**) and bias (**k**) in total LSK34⁺150⁺48⁺ LT-HSCs (combining data in **a,g,h,i**). Statistical comparisons: two-tailed Fisher's exact test (95% CI).

METHODS

Animals

All mice were bred and maintained, and all experimental procedures were performed, in accordance with UK Home Office regulations. All experiments were approved by the Oxford University Clinical Medicine Ethical Review Committee.

All mouse strains were backcrossed for more than six generations onto a C57BL/6J background. Both male and female adult mice (7 to 18 weeks old) were used in experiments. When multiple experimental groups were analysed, mice were allocated so that each group was evenly matched for age range and frequency of mice of each sex. Index-sorted single cells were transplanted randomly within each cohort of recipient mice. No statistical methods were used to predetermine the experimental sample size. The frequency of reconstitution patterns observed in preliminary single cell transplantation experiments was used to estimate the minimum number of single-cell transplanted mice required in order to observe ≥ 5 mice with each of the peripheral blood reconstitution patterns identified. Additional mice were transplanted with single cells in order to obtain enough biological replicates of relevant reconstitution patterns for all lines of experiments. Investigators were not blinded to experimental group allocation or when assessing experimental outcomes.

B6.SJL-Ptprc^a Pepc^b/BoyJ mice (CD45.1) were used as primary and secondary transplantation recipients, and in some experiments (CD45.1 x C57BL/6J)F1 mice (CD45.1/2) were also used as secondary recipients. As specified, wild type CD45.1 or *Kit*^{W-41J/W-41J}-*Gpi1*^{a/a} mice³¹ with CD45.1 allotype (*W*⁴¹/*W*⁴¹) were used as donors of unfractionated bone marrow (BM) support cells in primary competitive transplantation experiments. The *Vwf*-tdTomato transgene was generated by bacterial artificial chromosome (BAC) recombinant engineering at the EMBL Genome

Engineering Core Facility. To validate the hematopoietic expression pattern of *Vwf*-tdTomato, its co-expression with *Vwf*-EGFP was analysed after being crossed to similarly constructed and previously published *Vwf*-EGFP mice¹⁰. To produce single cell donors for primary transplantations in which both platelet and erythroid reconstitution could be evaluated, *Vwf*-tdTomato mice were crossed to the previously published *Gata1*-EGFP mice¹⁸, generating *Vwf*-tdTomato/*Gata1*-EGFP double reporter mice, heterozygous for both transgenic alleles and expressing the CD45.2 allotype. A *Vwf*-CreERT2 line was generated by Cyagen Biosciences by knock-in of an IRES-CreERT2-P2A-EGFP cassette into 3'UTR of the endogenous *Vwf* gene in C57BL/6 mice. For fate mapping experiments, this *Vwf*-CreERT2 line was crossed to B6;129S6-Gt(ROSA)26Sor^{tm9(CAG-tdTomato)Hze}/J (R26R-tdTomato). EGFP expression was not detectable in these mice.

Single cell transplantations

Single cell transplantations were performed as previously published¹⁰, using a FACSARIAII, FACSARIAIII or FACSARIA Fusion cell sorter (BD Biosciences). BM cell suspensions were prepared by crushing leg, sternum and spine bones into PBS (Gibco) with 5% fetal calf serum (FCS, Gibco) and 2mM ethylenediaminetetraacetic acid (EDTA, Sigma-Aldrich). Single phenotypically-defined Lin⁻Sca1⁺c-Kit⁺CD34⁻CD150⁺CD48⁻ BM hematopoietic stem cells (HSCs, LSK34⁻150⁺48⁻) were sorted from *Vwf*-tdTomato/*Gata1*-EGFP (CD45.2) adult mice (antibody details in Supplementary Table 1) into Iscove's Modified Dulbecco's Medium (IMDM, Gibco) with 20% BIT-9500 Serum Substitute (Stem Cell Technologies), 100U/mL Penicillin and 0.1mg/mL Streptomycin (100x Pen/Strep, PAA Laboratories), 2mM L-Glutamine (PAA Laboratories) and 0.1mM 2-Mercaptoethanol (Sigma-Aldrich). Single cells

were deposited by an automated cell deposition unit (ACDU). Accurate deposition of single fluorescent beads was assessed before and after single cell sorting to ensure that a single cell was deposited into all wells. Single sorted HSCs were mixed with 1×10^6 W^{41}/W^{41} or 2×10^5 wild-type CD45.1 unfractionated BM support cells, and transplanted by intravenous injection into lethally irradiated (10 Gy) CD45.1 mice. Single transplanted HSCs were sorted as $V_{wf}\text{-Tomato}^{\text{pos}}$ ($V_{wf}\text{-Tomato}^{\text{mid-high}}$) or, in order to determine the significance of $V_{wf}\text{-tdTomato}$ expression level, were index sorted and based on the fluorescence intensity of $V_{wf}\text{-tdTomato}$ for each cell were grouped into V_{wf}^{neg} , V_{wf}^{low} , V_{wf}^{mid} or V_{wf}^{high} categories. The same indexing strategy was used to identify $\text{CD150}^{\text{high}}$ and $\text{CD150}^{\text{low}}$ transplanted single cells. FACS reanalysis confirmed that mean sort purity (defined as fulfilment of all sort gating criteria upon reanalysis) was 95%.

Secondary transplantations

In bulk BM secondary transplantation experiments, 5×10^6 or 30×10^6 unfractionated BM cells from single-cell reconstituted primary recipient mice were resuspended in PBS 1% FCS and transplanted by intravenous injection into lethally irradiated (10 Gy) CD45.1 or CD45.1/2 recipients. Secondary transplantations into lethally irradiated CD45.1 mice were also performed with donor-derived LSK34-150+48-HSCs sorted into the IMDM-based single-cell transplantation media described above, together with fresh CD45.1 BM support cells.

Peripheral blood reconstitution analysis

PB samples were collected from tail vein into lithium heparin-coated microvettes (Sarstedt). A small aliquot of unfractionated PB was used for analysis of erythroid cells, separately or mixed with platelet solution. Platelets were separated by centrifugation of PB samples at 100xg for 10 minutes at room temperature. For leukocyte separation, samples were incubated 1:1 with 2% w/v Dextran (Sigma-Aldrich), for 20-30 minutes at 37°C. Excess erythrocytes were lysed from the leukocyte preparation through incubation with ammonium chloride solution (Stem Cell Technologies) for 2 minutes at room temperature, and leukocyte samples were pre-incubated with Fc-block prior to staining. Anti-mouse antibody staining for reconstitution analysis by flow cytometry was carried out for 15-20 minutes at 4°C in PBS 1% FCS 2mM EDTA (antibody details in Supplementary Table 1), and samples were analysed using LSRII, LSR Fortessa or LSR Fortessa X-20 flow cytometers (BD Biosciences).

Reconstituted erythroid cells were identified as Ter119⁺ *Vwf*-Tomato^{neg} *Gata1*-GFP⁺. Reconstituted platelets were identified as Ter119⁻CD150⁺CD41⁺ *Vwf*-Tomato^{pos} *Gata1*-GFP⁺. To investigate if aggregation between platelets derived from the transplanted single HSC and derived from the support cells might result in an overestimation of platelet reconstitution levels in transplanted mice, we mixed *Vwf*-Tomato^{pos} *Gata1*-GFP⁺ and wild-type PB to obtain platelet ratios of approximately 1:1 (n=6), 1:4 (n=6) and 1:10 (n=3). The mixed PB samples were processed for flow cytometry as detailed above, and for the expected percentages of approximately 50%, 25% and 10% *Vwf*-Tomato^{pos} *Gata1*-GFP⁺ platelets we as predicted observed 54.1% ± 5.4%, 27.6% ± 3.9% and 7.0 ± 1.0%, respectively. We also transplanted CD45.1 recipients with 20x10⁶ unfractionated BM cells in which the ratio of *Vwf*-EGFP to

Vwf-Tomato cells was 1:100 or 1:2 (n=3 recipients for each). With the 1:100 ratio we observed $0.2\% \pm 0.09$ of Tomato+GFP⁺ aggregates among labeled platelets, and with the 1:50 ratio we observed $1.3\% \pm 0.39$, indicating that aggregation between platelets is not substantial.

Leukocyte cell populations in PB were defined as follows: myeloid cells, NK1.1⁻CD4⁻CD8⁻CD19⁻CD11b⁺; B cells, NK1.1⁻CD4⁻CD8⁻CD11b⁻CD19⁺; CD4 T cells, NK1.1⁻CD19⁻CD11b⁻CD8⁻CD4⁺TCR β ⁺; CD8 T cells, NK1.1⁻CD19⁻CD11b⁻CD4⁻CD8⁺TCR β ⁺. Contribution to leukocytes in PB and spleen by CD45.2⁺ transplanted single cells was identified by co-expression of CD45.2 (and lack of CD45.1 expression) and mature lineage markers (antibody details in Supplementary Table 1), and calculated based on the frequency of CD45.2⁺ cells among total CD45⁺ events for each lineage, excluding CD45 negative events and CD45.1⁺CD45.2⁺ double-positive artifacts.

PB reconstitution in primary transplant recipients was analysed by flow cytometry at 8-10, 12-14 and 16-18 weeks post-transplantation for all mice, and also at 3-5 weeks and 20-44 weeks in specific extended kinetics experiments. Mice were considered to be reconstituted by a potential long-term HSC when donor contribution to at least one PB lineage was $\geq 0.1\%$ at 16-18 weeks post-transplantation.

Reconstitution patterns generated by potential long-term HSCs were considered to be restricted when reconstitution of one or more PB lineages was consistently below the detection level ($\leq 0.01\%$; see detection thresholds section below) at all time points (at least 3) of PB analysis up to 18 weeks post-transplantation. Therefore, for a transplanted HSC to be classified as being P-restricted, no reconstitution of neither E, M, B or T cells was observed at any time point in PB up to 16-18 weeks. Likewise, patterns in which each PB lineage was detectable ($>0.01\%$) in at least one time point

up to 18 weeks post-transplantation were considered to be Multilineage. According to these criteria, four lineage-restricted patterns were found to be produced by $V_{Wf}^{\text{mid-high}}$ LSK150⁺48⁺34⁻ cells: PEMB-, PEM-, PE- and P-restricted. When transplanting V_{Wf}^{neg} and V_{Wf}^{low} cells, another relevant pattern was also found, and designated LM-restricted. This pattern was defined by reconstitution of lymphoid (B and/or T cells) and at least one, but not all, myelo-erythroid lineages at one or more analysis points (3-5, 8-10, 12-14, 16-18 weeks).

Lineage-bias was defined as a sustained threefold higher lineage output compared to the remaining PB lineages, observed at 12-14 as well as 16-18 weeks post-transplantation. Patterns in which no single lineage or combination of lineages was threefold higher than all other lineages were then considered to be without a specific bias (“no bias”), and so were the patterns with no concordance of lineage-bias observed at 12-14 weeks and 16-18 weeks. The L-bias category includes mice in which lymphoid lineage (B and/or T) reconstitution is threefold higher than platelets and myeloid cells, but not necessarily threefold higher than erythroid cells due to their slower decline after the HSC exhaustion that underlies L-bias output.

In initial experiments we observed that no primary recipients with $\leq 0.01\%$ reconstitution of any of the PB lineages at 8-10 weeks post-transplantation were reconstituted by a long-term HSC, since in these mice (n=26) reconstitution did not reach $\geq 0.1\%$ for any lineage at 16-18 weeks. Therefore, in further primary transplantation experiments recipients with $\leq 0.01\%$ reconstitution of all lineages at ≥ 8 weeks post-transplantation were not analysed further. Secondary recipients were always analysed at 8-10, 12-14 and 16-18 weeks post-transplantation, and also at 3-5 weeks if transplanted with BM cells from Lymphoid-restricted and Lymphoid-biased primary reconstituted recipients. Recipient mice that did not survive or had to be

sacrificed before 16-18 weeks after primary or secondary transplantation were eliminated from the analysis (4.5% of a total of 528 mice). Secondary recipients without detectable reconstitution ($>0.01\%$) of any myelo-erythroid lineage (platelets, erythroid and myeloid cells) at 16-18 weeks post-transplantation and/or at the terminal analysis time point (18-28 weeks post-transplantation) were not considered to be HSC-reconstituted and therefore eliminated from secondary reconstitution analysis, except if being recipient of BM cells from Lymphoid-restricted and Lymphoid-biased primary reconstituted recipients.

Reconstitution analysis of hematopoietic stem and progenitor cells

Reconstitution analysis of lymphoid progenitors in thymus and hematopoietic stem and progenitor cell (HSPC) populations in BM was performed using LSRII, LSR Fortessa or LSR Fortessa X-20 flow cytometers (BD Biosciences). Except in the myeloid progenitor panel, which involves staining with fluorophore-conjugated CD16/32, BM, spleen and thymus cells were pre-incubated with Fc-block prior to anti-mouse antibody staining (antibody details in Supplementary Table 1). Reconstitution was calculated by quantifying the contribution of transplanted CD45.2⁺ cells to each cell population, except in the case of CFU-E in which *Gata1*-GFP⁺ cells were quantified instead due to the low expression level of CD45.1 and CD45.2 in these erythroid progenitors, which have been shown to be ~100% *Gata1*-GFP⁺ ¹⁸.

BM HSPC populations were defined as follows. HSC: Lin⁻Sca1⁺c-Kit⁺(LSK)Flt3⁻CD150⁺CD48⁻; 150⁺48⁺ MPP: LSKFlt3⁻CD150⁺CD48⁺; 150⁻48⁺ MPP: LSKFlt3⁻CD150⁻CD48⁻; LMPP: LSKFlt3^{high}CD150⁻CD48⁺; MkP: Lin⁻Sca1⁻c-Kit⁺(LK)CD150⁺CD41⁺; MEP: LKCD41⁻CD16/32⁻CD150⁺CD105⁻; CFU-E:

LKCD41⁻CD16/32⁻CD150⁻CD105⁺; PreGM: LKCD41⁻CD16/32⁻CD150⁻CD105⁻; GMP: LKCD41⁻CD150⁻CD16/32⁺; Pro-B: Lin⁻B220⁺CD19⁺c-Kit⁺ (in some experiments further defined as IgM⁻). Thymus T cell progenitors were defined as Lin⁻CD4⁺CD8⁺. We observed that the size of the classically-defined LMPP population (25% highest Flt3⁺ cells within LSK)²⁵ varied significantly between donor-derived CD45.2⁺ LSK and support-derived CD45.1⁺ LSK in different reconstituted mice, and therefore an LMPP gate equivalent to the 25% highest Flt3-expressing LSK cells in 7-18 week old non-transplanted mice was used as described before²⁵.

Detection thresholds for reconstitution analysis by flow cytometry

PB from non-transplanted CD45.1 and CD45.1/2 mice was screened alongside experimental samples to evaluate the minimum threshold of reconstitution detection. In the CD45.1 and CD45.1/2 controls (n≥10 for each), ≥10,000 events were recorded for each PB lineage. The detection threshold for *Vwf*-Tomato^{pos} *Gata1*-GFP⁺ platelets and *Vwf*-Tomato^{neg} *Gata1*-GFP⁺ erythroid cells was found to be 0.01% for both recipient types. The detection threshold for CD45.2 false positive events in the leukocyte lineages was found to also be 0.01% in CD45.1 recipients, and 0.03% in CD45.1/2 recipients. Reconstitution levels above these threshold levels were only considered reliable if also ≥5 donor-derived events could be recorded within a lineage gate. Reconstitution detection threshold for myeloid, B and T cells in spleen was also found to be 0.01%, based on analysis of CD45.1 controls (n=2).

As in PB, minimum reconstitution detection thresholds for progenitors were based on HSPC analysis of non-transplanted CD45.1 and CD45.1/2 controls (n≥4 for each), and defined as follows for both recipient types. 150⁺48⁺ MPP and 150⁻48⁺ MPP, >0.2%; HSC and LMPP, >0.1%; PreGM, >0.03%; MEP and Pro-B, >0.02%; GMP,

CFU-E and thymus DP, >0.01%. The thresholds indicated for HSPC reconstitution were frequently higher than these minimum threshold values due to the rarity of some HSPC populations in transplanted mice. In some cases, as specified, the number of events that could be recorded by flow cytometry was too low to reach the minimum threshold levels, additionally because reconstitution levels above the minimum thresholds were only considered reliable if ≥ 5 donor-derived events could be recorded. When averaging reconstitution percentages across several biological replicates, mean detection thresholds were calculated based on mean number of events recorded for each HSPC population.

Analysis of *Vwf*-tdTomato/*Vwf*-EGFP mice

To validate the novel *Vwf*-Tomato reporter mouse line, co-expression of *Vwf*-Tomato and *Vwf*-GFP in *Vwf*-tdTomato/*Vwf*-EGFP double reporter mice was analysed in PB platelets and erythroid cells, as well as BM LSK34⁻150⁺48⁻, MEP, MkP, PreCFU-E, CFU-E, PreGM and GMP. PB and BM HSPC analysis was performed as detailed above, with the addition of a c-Kit/CD117 enrichment step for HSPCs. Briefly, 2×10^8 unfractionated BM cells were resuspended in 200 μ L PBS 5% FCS 2mM EDTA and incubated with 5 μ L anti-mouse CD117 MicroBeads (Miltenyi Biotec) for 20 minutes at 4°C, then processed through MACS LS columns (Miltenyi Biotec) according to manufacturer instructions.

Fate mapping analysis of *Vwf*-CreERT2/R26R-tdTomato mice

The *Vwf*-CreERT2 mouse line was generated by Cyagen Biosciences. Cre recombinase activation was induced in 8-18 weeks old *Vwf*-CreERT2/R26R-tdTomato mice by Tamoxifen administration. Tamoxifen powder (Sigma T5648) was

dissolved in a sterile solution of 10% ethanol and 90% corn oil, for a final concentration of 20mg/mL. This solution was administered to the mice by oral gavage, 4mg per day for 5 consecutive days.

Peripheral blood from Tamoxifen-induced *Vwf*-CreERT2/R26R-tdTomato mice was processed similarly to that of transplanted mice. In order to control for potential aggregation of labelled and non-labelled cells, a *Gata1*-EGFP CD45.1 blood spike-in was introduced to experimental samples (n=14) at two different time points for each sample. CD41 staining was included to exclude platelets adhering to white blood cells.

***In vitro* myeloid lineage potential assays and validation**

Granulocyte-macrophage (GM) lineage potential was evaluated *in vitro* in X-VIVO15 liquid medium containing Gentamycin and L-Glutamine (Lonza) and supplemented with 10% FCS (GE Healthcare HyClone), 0.1mM 2-Mercaptoethanol (Sigma-Aldrich), and the following cytokines: 2ng/mL mouse stem cell factor (mSCF, PeproTech), 5ng/mL human FLT3 ligand (hFL, Immunex), 5ng/mL human thrombopoietin (hTPO, PeproTech), 5ng/mL mouse interleukin 3 (mIL-3, PeproTech), 10ng/mL human granulocyte colony-stimulating factor (hG-CSF, Neopogen) and 10ng/mL mouse GM colony-stimulating factor (mGM-CSF, Immunex). Donor-derived (CD45.2⁺) LSK cells were bulk sorted from single-cell transplanted CD45.1 mice and manually plated into 4 wells of round-bottom 96-well plates (Corning Costar) at a density of 25-50 cells per well, or into 60-240 wells of 60-well Terasaki microplates (Thermo Fisher Scientific) at a density of 1 cell per well. Culture confluency in Terasaki microplates was scored at day 10-14, and presence of mature myeloid cells was confirmed at day 13-18 by morphological

analysis of cytopsin slides (prepared as described below), and gene expression analysis by multiplex quantitative PCR (described below) of 4 wells per biological replicate (TaqMan probe details in Supplementary Table 2). Myeloid cells generated by 25-50 donor-derived LSK cells were also analysed by flow cytometry for co-expression of CD45.2, CD11b and Ly6G.

***In vitro* T cell lineage potential assay and validation**

T cell lineage potential was evaluated *in vitro* in OP9-DL1 stromal co-cultures, as previously described³². GFP⁺ OP9-DL1 stromal cells (cell line provided by A. Cumano, Institut Pasteur) were maintained in adherent cultures in 75cm² flasks (Corning) with Opti-MEM Glutamax medium (Gibco) supplemented with 10% FCS, 0.1mM 2-Mercaptoethanol, 100U/mL Penicillin and 0.1mg/mL Streptomycin. Expected properties of OP9-DL1 cell line were validated: distinct morphology, GFP expression, and capacity to robustly promote T cell differentiation from hematopoietic stem and progenitor cells. OP9-DL1 cell stocks were regularly tested for mycoplasma contamination by PCR.

One or two days before sorting the cells of interest for the assay, adherent cells were trypsinized (Trypsin EDTA, PAA Laboratories) and monolayers of OP9-DL1 stromal cells were prepared by seeding $1.5\text{-}2 \times 10^3$ cells per well in flat-bottom 96-well plates (Corning Costar). Donor-derived (CD45.2⁺) LSK cells were bulk sorted and manually plated at a density of 10 cells per well into 7-72 wells of pre-prepared confluent monolayers (~80%) of OP9-DL1 stromal cells. Cultures were maintained for 4-5 weeks in culture medium supplemented with the following cytokines: 5ng/mL mSCF and 5ng/mL hFL during first 1-2 weeks of culture, then mSCF only; or 5ng/mL mSCF and 5ng/mL hFL throughout, with addition of 4-5ng/mL hIL-7 during final week.

Cultures were analysed by flow cytometry for detection of CD45.2⁺CD90.2⁺CD25⁺ and CD45.2⁺CD4⁺CD8⁺ T cell progenitors using LSRII, LSR Fortessa or LSR Fortessa X-20 cytometers (antibody details in Supplementary Table 1). T cell gene expression analysis was performed by multiplex quantitative PCR (TaqMan probe details in Supplementary Table 2) on samples of 25-50 donor-derived T cell precursors sorted from 2-8 wells from each biological replicate using a FACS Aria Fusion sorter (BD Biosciences).

***In vitro* lineage potential assays for secondary recipients**

Lethally-irradiated (10Gy) CD45.1 mice (n=3 recipients per donor) were transplanted with donor-derived (CD45.2⁺) LSK34-150+48- *Vwf-Tomato*^{pos} cells sorted at 18-26 weeks post-transplantation from primary recipients with PE- or P-restricted patterns (n=3 donors). Each secondary recipient received 250-1,000 donor-derived sorted cells together with 2x10⁵ wild-type CD45.1 support cells. 19-21 weeks after secondary transplantation, donor-derived LSK cells were sorted from the secondary recipients with maintained PE- or P-restricted patterns and tested in GM and T cell *in vitro* lineage potential assays as described in the previous section.

GM potential: bulk sorted CD45.2⁺ LSK cells were manually plated at a density of 1 cell per well into 20-37 wells of 60-well Terasaki microplates. Cell confluency per well was scored at day 13, and presence of mature myeloid cells was confirmed by morphological analysis.

T cell potential: bulk sorted CD45.2⁺ LSK cells were manually plated at a density of 10 cells per well into 6-19 wells of pre-prepared monolayers of OP9-DL1 stromal cells. Cultures were maintained for 4 weeks before flow cytometry analysis, and

identity of phenotypic CD45.2⁺ T cell progenitors was confirmed by gene expression analysis.

Cytospins and cell morphology analysis

Myeloid cultures grown *in vitro* were collected into PBS with 20% FCS (Gibco) and centrifuged onto polylysine-coated microscope slides (VWR) by a Shandon Cytospin 4 Cytocentrifuge (Thermo Scientific). Slides were stained with May Grünwald (Sigma-Aldrich) and Giemsa (Fluka) solutions according to manufacturer instructions, and mounted with Pertex (Cellpath). Morphology of stained cells was analysed with an Olympus BX60 upright compound microscope, images were captured with INFINITY imaging software (Lumenera) and edited with ImageJ software (NIH, public domain).

Multiplexed quantitative PCR

Multiplex quantitative PCR analysis was performed using the BioMark 48.48, 96.96 or 192.24 Dynamic Array platform (Fluidigm) and TaqMan Gene Expression Assays (Thermo Fisher Scientific) as previously described^{32,33}. Cells were sorted or cultures were transferred into PCR tubes containing 10µL or 15µL of RT/pre-amplification buffer, respectively. CellsDirect One-Step qRT-PCR kit (Invitrogen) was used for cDNA synthesis and pre-amplification of target genes. 10µl of pre-amplification buffer consisted of 2.5µl TaqMan assay mix containing all assays at 0.2x dilution, 5µL CellsDirect 2x Reaction mix (Invitrogen), 1.2µL CellsDirect RT/Taq mix (Invitrogen), 1.2µL TE buffer and 0.1µL SUPERase-In RNase Inhibitor (Ambion). Targeted cDNA pre-amplification was performed during 22 cycles and pre-amplified product was diluted 1:5 in TE buffer before processing with Dynamic Array protocol

according to manufacturer instructions. Details of TaqMan Gene Expression Assays (Life Technologies) in Supplementary Table 2. Controls with no template or with no RT polymerase were also analysed in each experiment, as well as positive and negative control cell samples.

Heatmaps were generated in Microsoft Excel using data analysed by the ΔC_t method, using the mean of housekeeping genes Hypoxanthine Guanine Phosphoribosyltransferase 1 (*Hprt1*) and Beta-2-microglobulin gene (*B2m*) for normalization. Only samples expressing *Hprt1*, *B2m*, and the pan-hematopoietic CD45 gene (*Ptpnc*) were included for further analysis.

Statistics

Statistical comparisons of continuous data with assumed normal distribution (mean percentages) were performed with parametric unpaired two-tailed Student's *t*-test. Statistical comparisons of categorical data (frequencies) were performed with nonparametric two-tailed Fisher's exact test. Statistical comparison of platelet reconstitution levels between patterns was performed with one-sided Wilcoxon Mann-Whitney test (normal distribution not assumed based on Shapiro-Wilk normality test). Only statistically significant differences ($p < 0.05$) were indicated in the figures. For all statistical comparisons: * $p < 0.05$, ** $p < 0.01$, ***, $p < 0.001$, **** $p < 0.0001$.

Student's *t*-test and Fisher's exact test were performed with GraphPad Prism software and GrapPad QuickCalcs online tool: www.graphpad.com/quickcalcs. Wilcoxon-Mann-Whitney test was performed with online tool: ccb-compute2.cs.uni-saarland.de/wtest/?id=www/www-ccb/html/wtest ³⁴.

Data availability

Source data for all figures is available in the online version of the paper. Further details are available from the corresponding authors upon reasonable request.

ADDITIONAL REFERENCES (METHODS ONLY)

- 31 Benveniste, P. *et al.* Intermediate-term hematopoietic stem cells with extended but time-limited reconstitution potential. *Cell Stem Cell* **6**, 48-58 (2010).
- 32 Tehranchi, R. *et al.* Persistent malignant stem cells in del(5q) myelodysplasia in remission. *N Engl J Med* **363**, 1025-1037 (2010).
- 33 Luis, T. C. *et al.* Initial seeding of the embryonic thymus by immune-restricted lympho-myeloid progenitors. *Nat Immunol* **17**, 1424-1435 (2016).
- 34 Marx, A., Backes, C., Meese, E., Lenhof, H. P. & Keller, A. EDISON-WMW: Exact dynamic programming solution of the Wilcoxon-Mann-Whitney test. *Genomics Proteomics Bioinformatics* **14**, 55-61 (2016).

EXTENDED DATA FIGURE LEGENDS

Extended Data Figure 1 | Characterization of *Vwf*-Tomato and *Vwf*-GFP co-expression in PB and BM. a-c, Flow cytometry analysis (1 experiment) of *Vwf*-Tomato and *Vwf*-GFP co-expression in PB platelets and erythroid cells (**a**), in c-Kit enriched BM LSK34⁻150⁺48⁻ cells showing representative gating strategy used in single-cell transplantation sorts (**b**), and in myeloid and erythroid progenitors (**c**). **d**, FACS profile of BM LSK34⁻150⁺48⁻ cells in *Vwf*-tdTomato/*Gata1*-EGFP mice (2 experiments). Mean percentages of parent gates, n=3 mice for all plots.

Extended Data Figure 2 | Stable long-term lineage-restricted reconstitution in recipients of single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cells with wild-type BM support cells. a, Reconstituted mice (mean±s.e.m.) transplanted with single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cell and WT BM support. n=58 transplanted mice, 7 experiments. No statistically significant difference in frequency of reconstituted mice between *W⁴¹/W⁴¹* (Fig.1a) and WT support (p=0.46). **b-c**, Analysis of platelet, erythroid-, myeloid-, B- and T-cell contribution 16-18wks post-transplantation in P-restricted (**b**) and PEM-restricted mice (**c**) (plots representative of >40 single cell transplantation experiments). **d**, Reconstitution kinetics (mean±s.e.m.). P-restricted, n=3 mice; PEM-restricted, n=5; PEMB-restricted, n=6; Multilineage, n=11. **e**, Distribution of lineage-restricted reconstitution patterns in mice in **a**. No statistically significant difference in frequency of each pattern between mice transplanted with *W⁴¹/W⁴¹* (Fig.1d, n=109 mice) and WT support (n=25). Multi p=0.662; PEMB p=0.785; PEM p=0.769; PE p=1.0; P-restricted p=1.0. **f**, Distribution of lineage-bias within reconstitution patterns in **e**. Statistically significant differences in P-bi frequency between patterns are indicated above bars (*p=0.015; ***p=0.0002); and for each pattern between mice transplanted with *W⁴¹/W⁴¹* (Fig.1e) and WT support within bars (PEMB: *p=0.022, ***p=0.0007; Multi: L-bi *p=0.026, PEM-bi *p=0.044). **g**, Overall lineage-bias distribution in **e** (WT support, n=25 mice) and Fig.1e (*W⁴¹/W⁴¹* support, n=109). L-bi *p=0.034; No bias *p=0.028; PEM-bi **p=0.004. Statistical comparisons: two-tailed Fisher's exact test (95% CI).

Extended Data Figure 3 | Analysis of PB lineage-restricted reconstitution by single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cells and spleen lymphocyte reconstitution analysis in lineage-restricted reconstituted mice. **a-d**, Flow cytometry analysis of PB platelet, erythroid, myeloid, B and T lymphocyte reconstitution, 16-18wks post-transplantation of a single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cell (representative of >40 single cell transplantation experiments). P-restricted (**a-b**), PE-restricted (**c**), PEM-restricted (**d**), PEMB-restricted (**e**) and Multilineage (**f**) stably reconstituted mice. **g**, Flow cytometry analysis of myeloid and lymphoid reconstitution in spleen 23wks post-transplantation, corresponding to the P-restricted PB reconstitution pattern in **Extended Data Fig. 2b**. **h**, Reconstitution (mean+s.e.m.) of PB platelets and spleen lymphocytes 16-44wks post-transplantation in mice with P-restricted (n=8 mice), PEM-restricted (n=7) and Multilineage (n=10) reconstitution. 13 experiments. Frequency of positive mice and mean reconstitution in positive mice shown.

Extended Data Figure 4 | *Vwf*-Tomato co-expression with CD229 and CD41 in BM HSCs. **a**, *Vwf*-Tomato expression in LSK34⁻150^{hi}48⁻ cells with different CD229 expression levels in two different mice, representative of 4 experiments. Percentages of parent gate and median fluorescence intensities (MFI) shown. **b**, Expression of CD41 and *Vwf*-Tomato (mean % of parent gate ± s.e.m., n=3 mice in 1 experiment) in LSK34⁻150^{hi}48⁻ cells with different levels of CD229 expression.

Extended Data Figure 5 | HSPC reconstitution analysis in lineage-restricted mice transplanted with single *Vwf*-Tomato^{pos} LSK34⁻150⁺48⁻ cells. **a-d**, Flow cytometry reconstitution analysis of BM and thymus HSPCs in a representative P-restricted reconstituted mouse at 23wks post-transplantation (**a**), and representative Multilineage (**b**) and PEM-restricted (**c**) reconstituted mice 16wks post-transplantation. Plots representative of 12 HSPC analysis experiments.

Extended Data Figure 6. Multipotency of single reconstituting LSK34⁻150⁺48⁻ cells with *in vivo* lineage-restricted output. **a**, *In vitro*-derived granulocytes (top row) and monocytes/macrophages (bottom row) generated by donor-derived LSK cells sorted from single-HSC transplanted mice with long-term P-restricted reconstitution (each column shows results for one mouse; cytopins representative of 9 P-restricted mice analysed). Scale-bar=50µm. **b**, FACS of granulocytes (CD11b⁺Ly6G⁺) and monocytes/macrophages (CD11b⁺Ly6G⁻) generated by 25-50

donor-derived LSK from 3 P-restricted mice (representative of 4 mice analysed in 3 experiments). **c**, Gene expression heatmap of GM cells generated *in vitro*. Multi, n=1 mouse. P-restricted, n=2 mice. 1 experiment, 4 GM wells/mouse. Control: ~1,000 fresh PB leukocytes, n=1. Mean Ct values per mouse group, normalized to mean Ct of *Hprt1/B2m*.

Extended Data Figure 7. Long-term persistence of platelet-restricted and platelet-biased reconstitution patterns by multipotent HSCs. **a**, Reconstitution of HSPC hierarchy in BM and thymus, 26wks post-transplantation, of two secondary recipients of a mouse with sustained P-restricted reconstitution at 44wks post-transplantation (additional recipient of the donor in **Fig. 3d**). **b**, Sorting strategy for secondary transplantation of *Vwf*-Tomato^{pos} HSCs from the mice in **Figure 3f** (representative gating of mouse 2). Percentages of parent gates shown.

Extended Data Figure 8 | Reconstitution of PB mature lineages in mice transplanted with single LSK34⁻150⁺48⁻ cells with different levels of *Vwf*-Tomato expression and CD150 levels. **a**, Reconstituted mice (mean+s.e.m.) transplanted with single *Vwf*^{neg} HSCs with low (n=57) or high CD150 levels (n=77) in 11 experiments. No statistically significant difference, p=0.24. LSK34⁻150⁺48⁻ cells were index sorted and CD150^{hi} *Vwf*^{neg} cells defined as having CD150 expression levels overlapping with CD150 expression in LSK34⁻150⁺48⁻ *Vwf*^{mid-high} cells (see Fig.4a), and CD150^{lo} *Vwf*^{neg} cells defined as having lower CD150 expression levels than LSK34⁻150⁺48⁻ *Vwf*^{mid-high} cells. **b**, Distribution of restriction patterns generated by single *Vwf*^{neg} CD150^{lo} (n=18 mice) and CD150^{hi} cells (n=17). No statistically significant differences. LM: p=0.60; Multi p=0.40; PEMB p=0.60. **c**, Distribution of lineage-bias in **e**. No statistically significant differences. No bias p=0.23; PE-bi p=0.49; L-bi p=0.12. **d-e**, Flow cytometry analysis of PB platelet, erythroid-, myeloid-, B- and T-cell reconstitution 4wks (top panels) and 18wks post-transplantation (bottom panels) in L-restricted (**b**) and L-biased (**c**) reconstituted mice (representative of 14 single cell transplantation experiments of *Vwf*^{neg} HSCs). Statistical comparisons: two-tailed Fisher's exact test (95% CI).

Extended Data Figure 9 | HSPC reconstitution analysis in L-restricted and L-biased reconstituted mice transplanted with a single LSK34⁻150⁺48⁻ cell. a-b, Flow cytometry analysis of PB platelet, erythroid-, myeloid-, B- and T-cell reconstitution, and HSPC hierarchy reconstitution, in L-restricted (**a**, representative of 3 experiments) and L-biased (**b**, representative of 2 experiments) reconstituted mice at 23wks post-transplantation. **c**, Reconstitution percentages (mean±s.e.m.) of HSPCs 22-23wks post-transplantation in L-restricted (n=6 mice; except Pro-B, PreGM, GMP, MEP, CFU-E and MkP, n=3) and L-biased (n=3) reconstitution patterns. Orange: all mice positive. Grey: no positive mice. Pink: positive/negative mice; frequency and mean±s.e.m. of positives shown.

Extended Data Figure 10 | Reconstitution analysis of secondary recipients of BM cells from single-cell transplanted mice with L-restricted and L-biased reconstitution. a-b, Flow cytometry analysis of PB platelet, erythroid-, myeloid-, B- and T-cell reconstitution 8wks post-transplantation of secondary recipients of BM cells from primary mice with L-restricted (**a**, representative of 3 experiments) and L-biased (**b**, representative of 2 experiments) reconstitution generated by a single LSK34⁻150⁺48⁻ cell. PB and BM HSPC reconstitution analysis of the primary recipients is shown in **Extended Data Fig. 9a-b**. **c**, Reconstitution percentages (mean±s.e.m) of HSPCs in L-biased primary mice 21-29wks post-transplantation (n=3) and their secondary recipients 17wks post-transplantation (n=9, 3/donor). Orange: all mice positive. Grey: no positive mice. Pink: positive/negative mice; frequency and mean±s.e.m. of positives shown. **d**, Reconstitution kinetics (mean±s.e.m.) of mice in **Fig. 4h**, 14 experiments. Multilineage: V_{wf}^{neg} , n=6 mice; V_{wf}^{low} , n=2; V_{wf}^{mid} , n=12; V_{wf}^{high} , n=12. P-restricted: V_{wf}^{low} , n=1; V_{wf}^{mid} , n=1; V_{wf}^{high} , n=8.

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A.J. Mead, D. Atkinson, A. Giustacchini and N. Ashley for expert assistance with the Fluidigm array platform [WIMM Single Cell Core Facility is supported by the MRC MHU (MC_UU_12009), the Oxford Single Cell Biology Consortium (MR/M00919X/1) and the WT-ISSF (097813/Z/11/B#) funding, the WIMM Strategic Alliance awards G0902418 and MC_UU_12025]; P. Sopp and S-A. Clark for expert flow cytometry technical support and cell sorting services [WIMM FACS Core Facility is supported by the MRC HIU, MRC MHU (MC_UU_12009), NIHR Oxford BRC and the John Fell Fund (131/030 and 101/517), the EPA fund (CF182 and CF170) and by WIMM Strategic Alliance awards (G0902418 and MC_UU_12025)]; the Biomedical Services at University of Oxford for animal technical support; the EMBL Monterotondo Gene Expression Service and Transgenic Core Facility for generating the *Vwf*-tdTomato BAC and the corresponding transgenic mouse line; N. Iscove (Ontario Cancer Institute, University Health Network) for *W⁴¹/W⁴¹* mice; A. Cumano (Institut Pasteur) for OP9-DL1 stromal cells; R. Drissen and S. Duarte (Weatherall Institute of Molecular Medicine) for discussions and assistance with the preliminary phase of the studies; A. Hillen (Karolinska Institutet), B. Wu and T. Bouriez-Jones (Weatherall Institute of Molecular Medicine) for technical assistance.

This work was supported by Marie Curie Early Stage Researcher Fellowship (J.C.), the Medical Research Council UK (G0801073 and MC_UU_12009/5 to S.E.W.J. and G0701761, G0900892 and MC_UU_12009/7 to C.N.), the Swedish Research Council (S.E.W.J.), the Knut och Alice Wallenberg Foundation (WIRM; S.E.W.J.), the Tobias Foundation (S.E.W.J.), StratRegen KI (S.E.W.J.), Bloodwise (project grant 15006 to C.N.) and a BBSRC Project Grant (BB/M024350/1 to C.N.).

Author contributions S.E.W.J. and C.N. conceptualized the research, with input from A.S-P and J.C. S.E.W.J., C.N., J.C., Y.M., L.M.K. and P.S.W. designed the experiments and analysed the data. J.C. and Y.M. performed all experiments except fate mapping, with assistance from T.C.L., A.Ga. and A.Gr. (single cell transplantations), L.M.K., R.N. and V.A. (peripheral blood reconstitution analysis), H.B. (blood and progenitor reconstitution analysis) and F.G. (CD229/CD41 analysis). L.M.K performed fate mapping experiments with assistance from K.H. and A.M.L. S.E.W.J., C.N. and J.C. wrote the manuscript, which was subsequently reviewed and approved by all authors.

Author information The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.E.W.J. (sten.jacobsen@imm.ox.ac.uk; sten.eirik.jacobsen@ki.se) or C.N. (claus.nerlov@imm.ox.ac.uk).

Figure 1

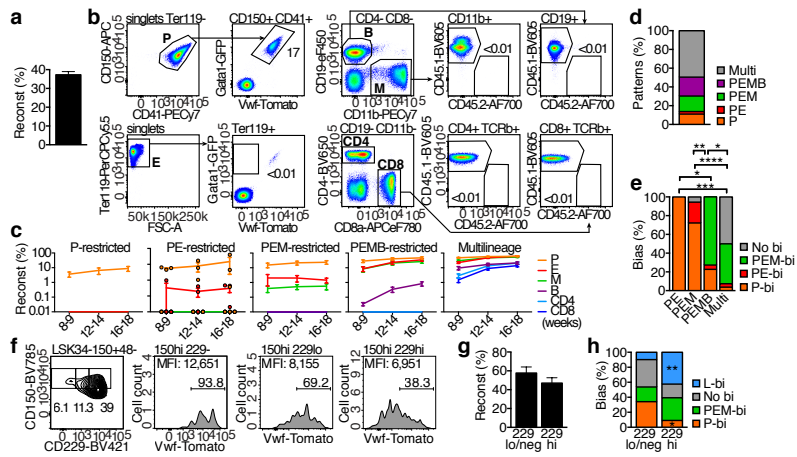


Figure 2

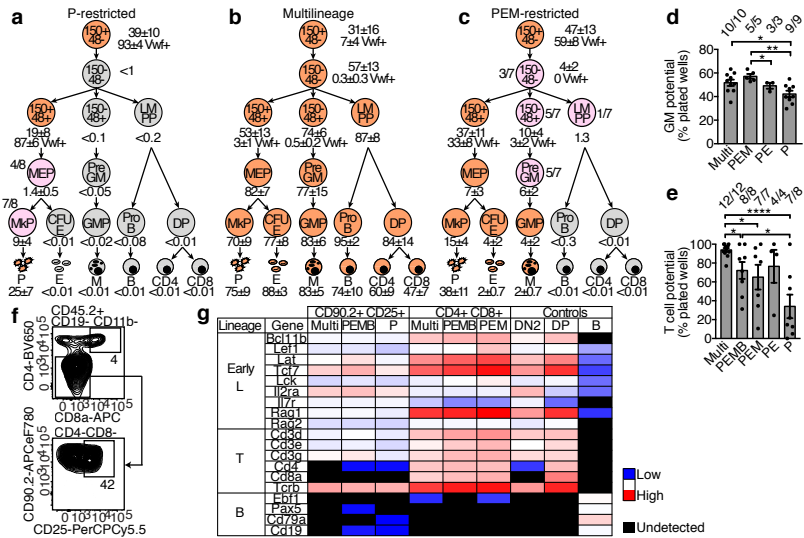
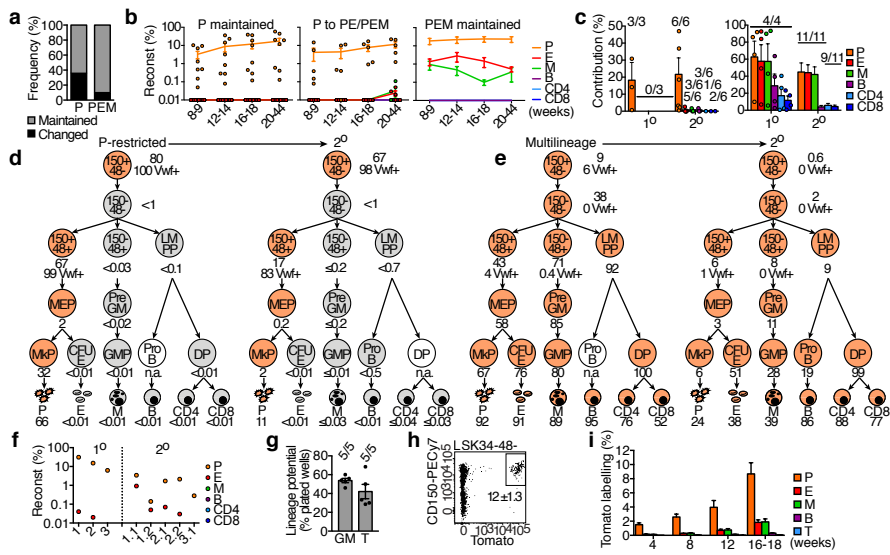
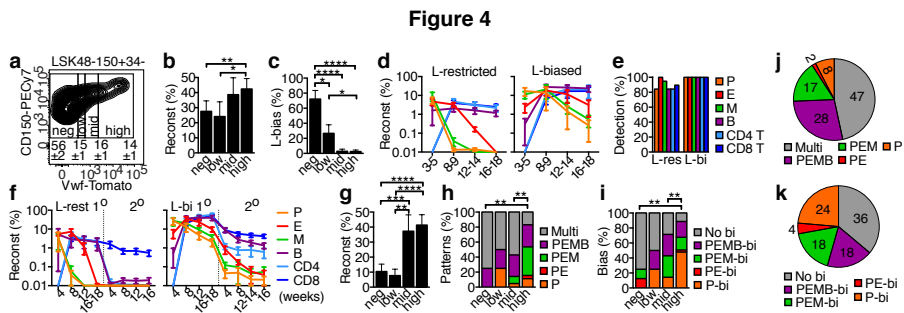
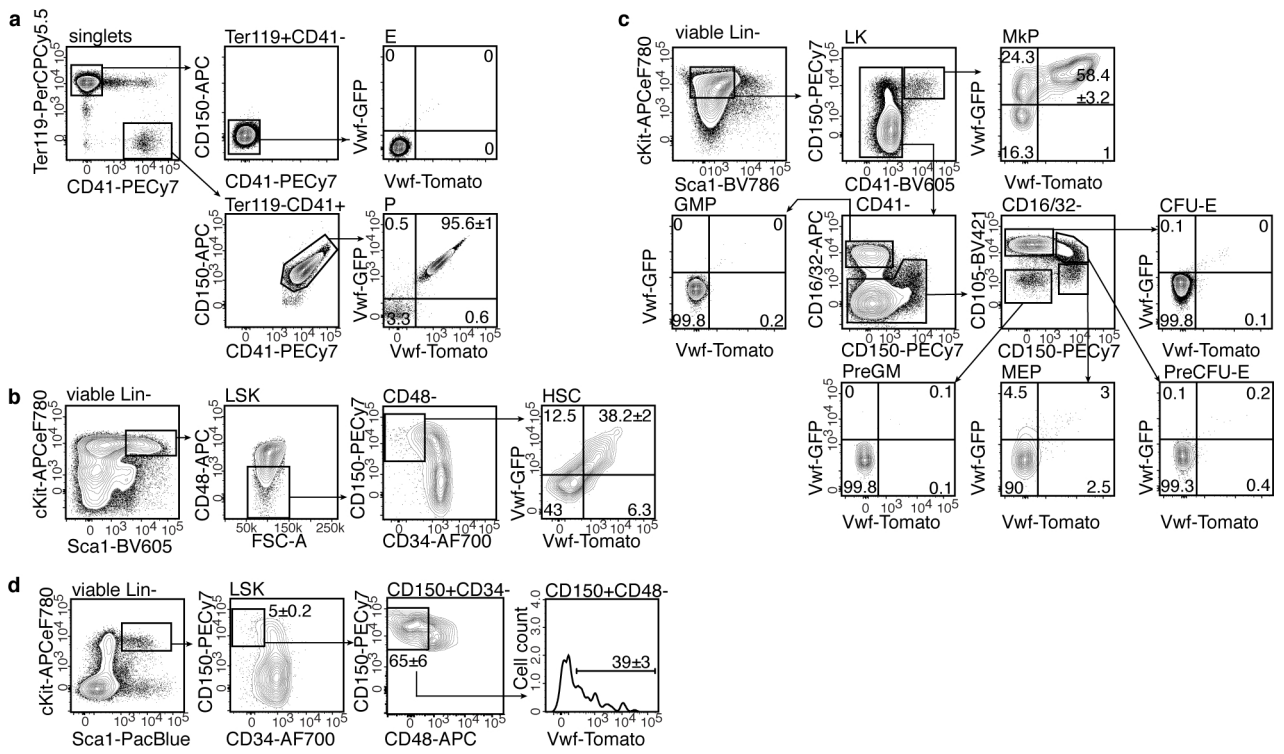


Figure 3

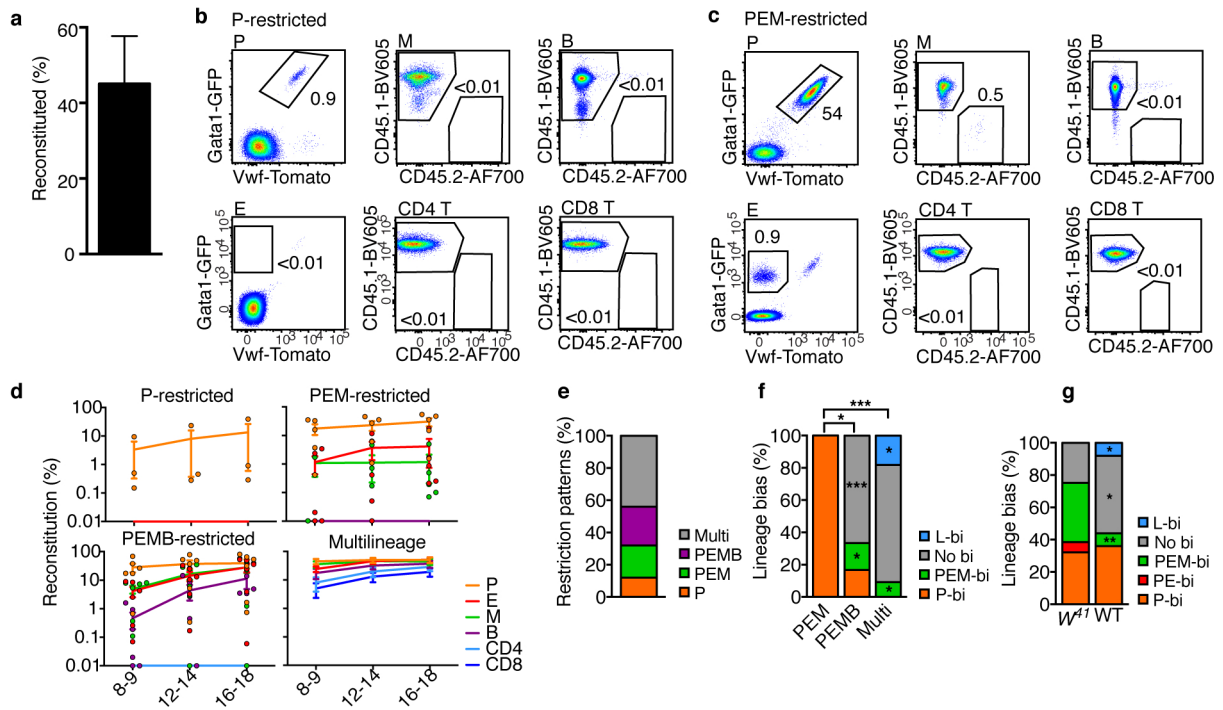




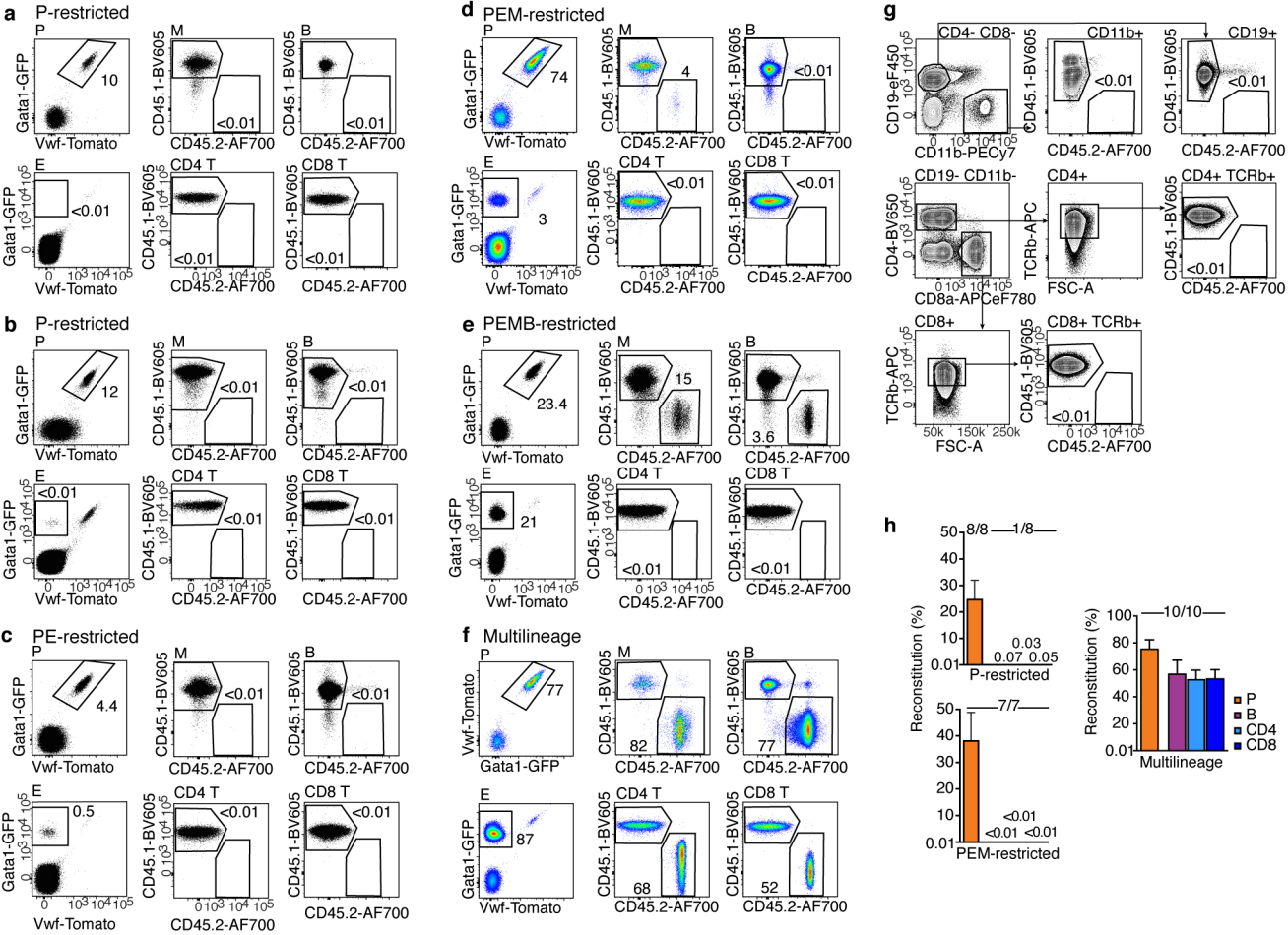
Extended Data Figure 1



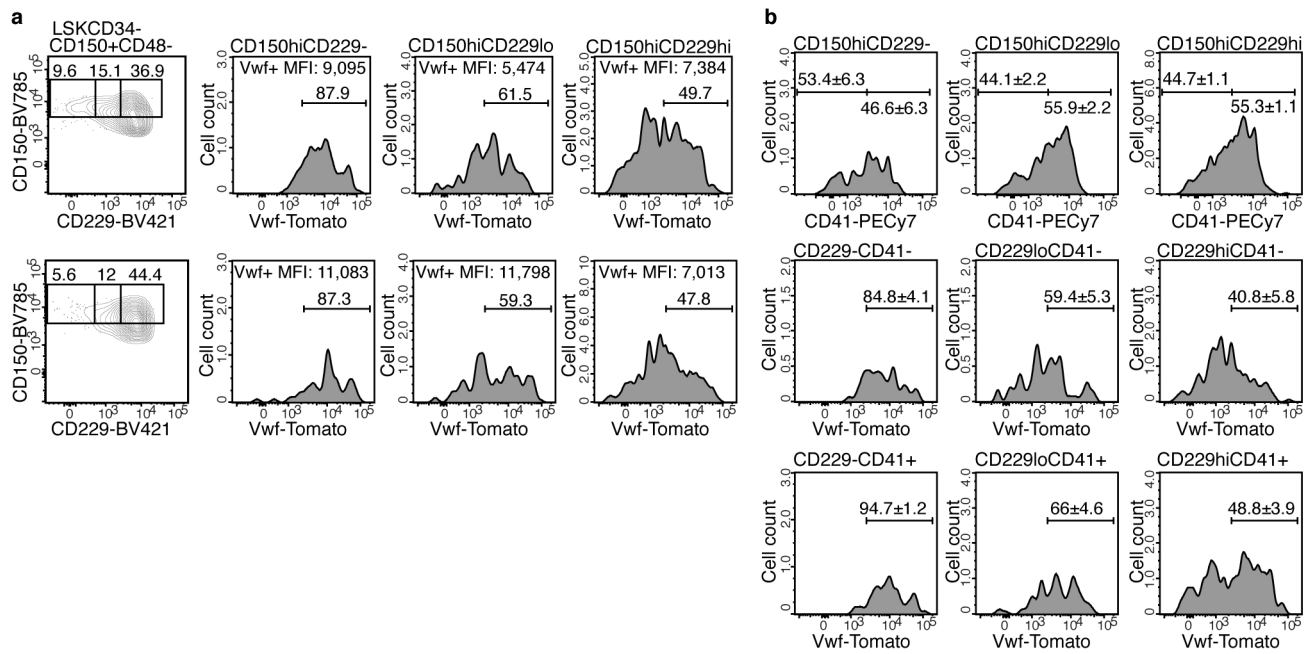
Extended Data Figure 2



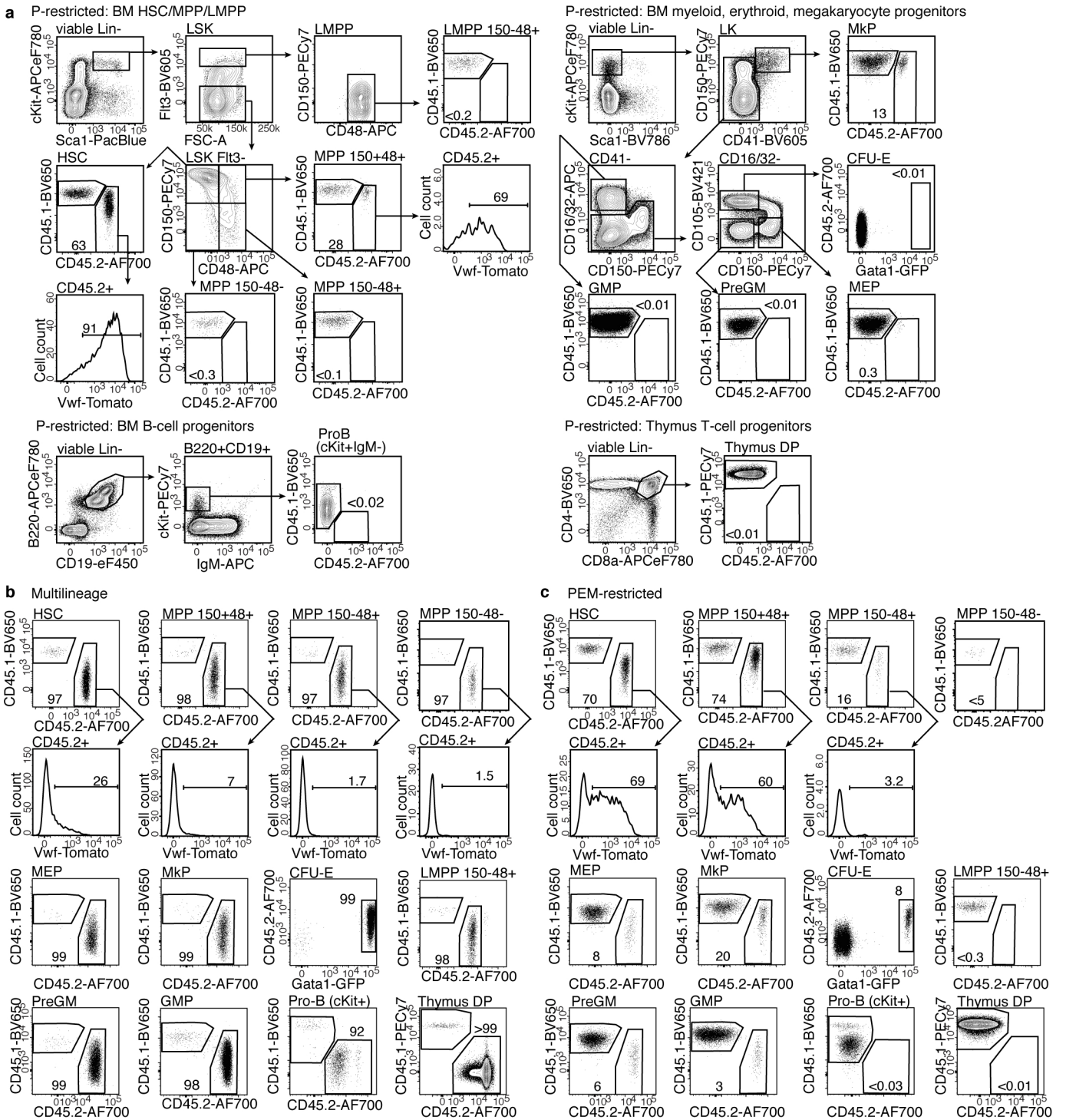
Extended Data Figure 3



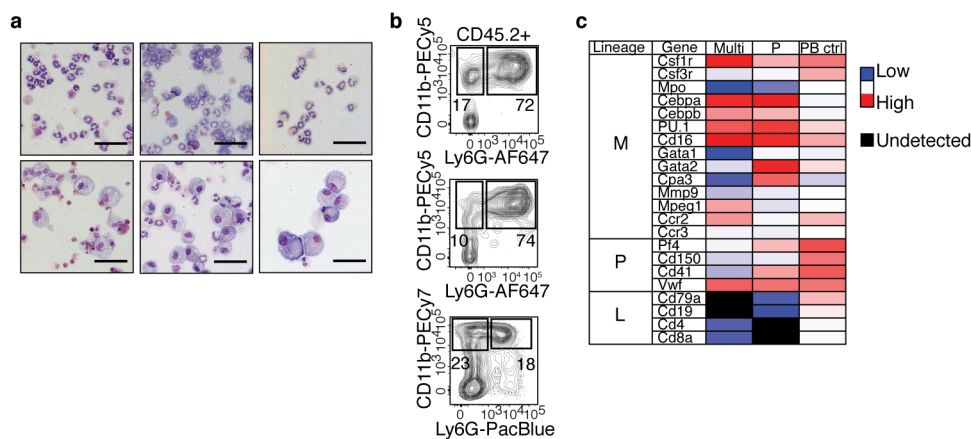
Extended Data Figure 4



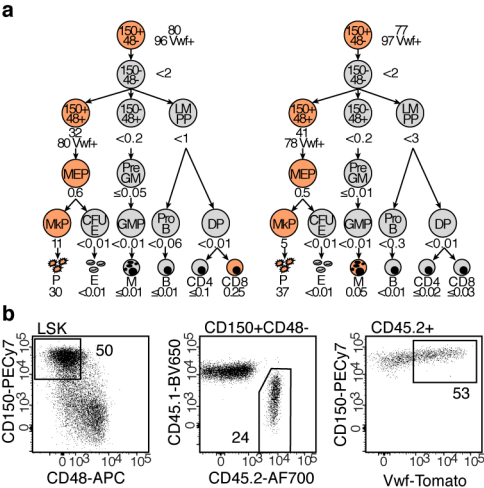
Extended Data Figure 5



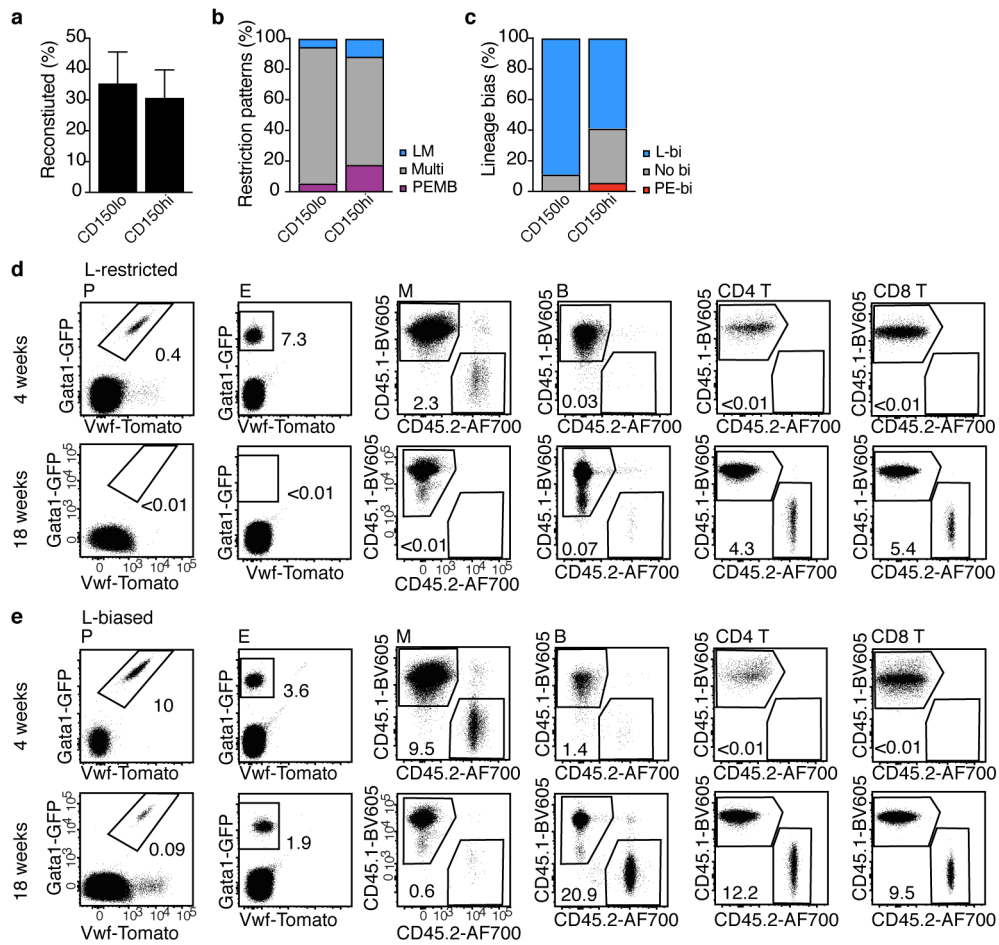
Extended Data Figure 6



Extended Data Figure 7



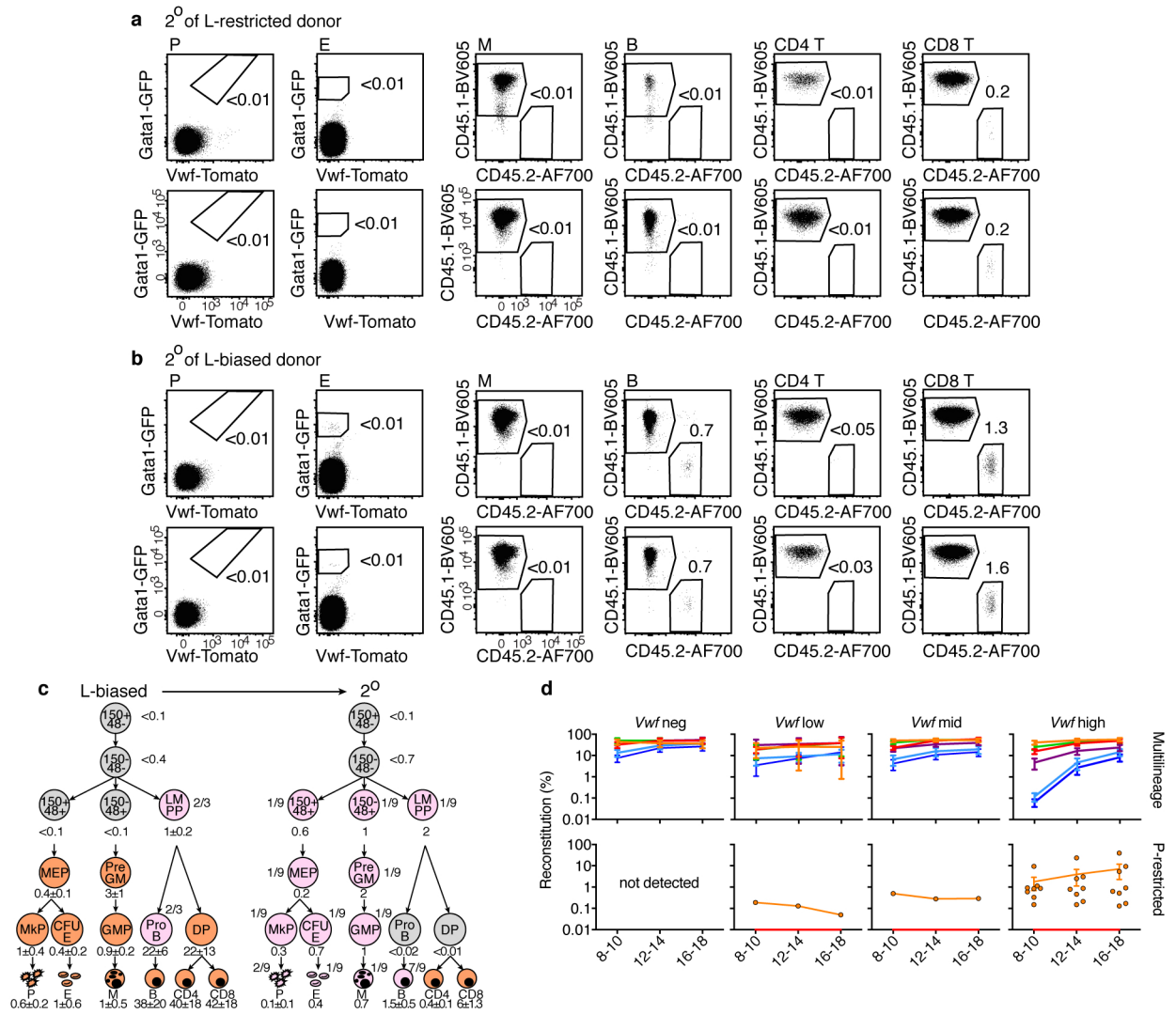
Extended Data Figure 8



Extended Data Figure 9



Extended Data Figure 10



Supplementary Table 1. Antibodies and viability dyes used in flow cytometry.

Application	Antibody or viability dye	Conjugate	Clone	Company	Catalog number
Sorting of BM HSCs for single cell transplantation and analysis of CD229 and CD41 co-expression (lineage combinations I and II)	c-Kit	APC-eF780	2B8	eBioscience	47-1171-82
	Sca-1	Pacific Blue	E13-161.7	BioLegend	122520
		BV605	D7		108133
	CD150	PE-Cy7	TC15-12F12.2	BioLegend	115914
		BV785			115937
	CD48	APC	HM48-1	BioLegend	103412
	CD34	AF700	RAM34	eBioscience	56-0341-82
	CD229	Biotin	Ly9ab3	BioLegend	122903
	Streptavidin	BV421	n/a	BioLegend	405226
	CD41	PE-Cy7	MWReg30	eBioscience	25-0411-82
	B220 (I+II)		RA3-6B2	BioLegend	103210
	Ter119 (I+II)		TER-119	BioLegend	116210
	CD4 (I+II)		RM4-5	BioLegend	100514
	CD8a (I+II)		53-6.7	BioLegend	100710
	Gr-1 (I)		RB6-8C5	BioLegend	108410
	CD11b (I)		M1/70	BioLegend	101210
	CD5 (I)		53-7.3	BioLegend	100610
	IL7R α (II)		A7R34	eBioscience	15-1271-81
	7-AAD	n/a	n/a	Sigma	A9400-1MG
	DAPI	n/a	n/a	Invitrogen	D3571
Reconstitution analysis of PB platelets and erythrocytes	Ter119	PerCP-Cy5.5	TER-119	eBioscience	45-5921-82
	CD150	APC	TC15-12F12.2	BioLegend	115910
	CD41	PE-Cy7	MWReg30	eBioscience	25-0411-82
	NK1.1	PE-Cy5	PK136	BioLegend	108716
Reconstitution analysis of myeloid, B and T leucocytes in PB and spleen	CD11b	PE-Cy7	M1/70	BioLegend	101215
		Pacific Blue			<i>Discontinued</i>
	CD19	eF450	1D3	eBioscience	48-0193-82
	CD4	BV650	RM4-5	BioLegend	100546
	CD8a	APC-eF780	53-6.7	eBioscience	47-0081-82
	TCR β	APC	H57-597	eBioscience	17-5961-83
	CD45.1	PE-Cy7	A20	BioLegend	110730
	CD45.2	AF700	104	BioLegend	109822
	7-AAD	n/a	n/a	Sigma	A9400-1MG
	DAPI	n/a	n/a	Invitrogen	D3571
Reconstitution analysis of B and T leucocytes in spleen (alternative panel)	CD19	Pacific Blue	1D3	eBioscience	<i>Discontinued</i>
		eF450		eBioscience	48-0193-82
	CD4	BV650	RM4-5	BioLegend	100546
	CD8a	APC-eF780	53-6.7	eBioscience	47-0081-82
	TCR β	APC	H57-597	eBioscience	17-5961-83
	CD45.1	PE-Cy7	A20	BioLegend	110730
	CD45.2	AF700	104	BioLegend	109822
	Gr-1		RB6-8C5	BioLegend	108410
	CD11c		N418	BioLegend	117316
	NK1.1		PK136	BioLegend	108716
	Ter119		TER-119	BioLegend	116210
	CD11b		M1/70	BioLegend	101210
	7-AAD	n/a	n/a	Sigma	A9400-1MG
	c-Kit	APC-eF780	2B8	eBioscience	47-1171-82
	Sca-1	Pacific Blue	E13-161.7	BioLegend	122520
	CD150	PE-Cy7	TC15-12F12.2	BioLegend	115914
Reconstitution analysis of HSPCs in BM: HSC, MPPs, LMPP	CD48	APC	HM48-1	BioLegend	103412
	Flt3	Biotin	A2F10	eBioscience	13-1351-85
	Streptavidin	BV605	n/a	BioLegend	405229
	CD45.1	BV650	A20	BioLegend	110736
	CD45.2	AF700	104	BioLegend	109822
	Gr-1		RB6-8C5	BioLegend	108410
	B220		RA3-6B2	BioLegend	103210
	Ter119		TER-119	BioLegend	116210
	CD4		RM4-5	BioLegend	100514
	CD8a		53-6.7	BioLegend	100710
	CD11b		M1/70	BioLegend	101210
	CD5		53-7.3	BioLegend	100610
	7-AAD	n/a	n/a	Sigma	A9400-1MG
	c-Kit	APC-eF780	2B8	eBioscience	47-1171-82
	Sca-1	Pacific Blue	E13-161.7	BioLegend	122520
	CD150	PE-Cy7	TC15-12F12.2	BioLegend	115914
	CD48	APC	HM48-1	BioLegend	103412
	Flt3	Biotin	A2F10	eBioscience	13-1351-85
	Streptavidin	BV605	n/a	BioLegend	405229
	CD45.1	BV650	A20	BioLegend	110736
	CD45.2	AF700	104	BioLegend	109822
	Gr-1		RB6-8C5	BioLegend	108410
	B220		RA3-6B2	BioLegend	103210
	Ter119		TER-119	BioLegend	116210
	CD4		RM4-5	BioLegend	100514
	CD8a		53-6.7	BioLegend	100710
	CD11b		M1/70	BioLegend	101210
	CD5		53-7.3	BioLegend	100610
Reconstitution analysis of HSPCs in BM: PreGM, GMP, MEP, MKP, CFU-E	7-AAD	n/a	n/a	Sigma	A9400-1MG
	c-Kit	APC-eF780	2B8	eBioscience	47-1171-82
	Sca-1	BV786	D7	BD	563991
	CD150	PE-Cy7	TC15-12F12.2	BioLegend	115914
	CD16/32	APC	93	eBioscience	17-0161-81
	CD105	BV421	M1/718	BD	562780
	CD41	BV605	MWReg30	BD	563317
	CD45.1	BV650	A20	BioLegend	110736
	CD45.2	AF700	104	BioLegend	109822
	Gr-1		RB6-8C5	BioLegend	108410
	B220		RA3-6B2	BioLegend	103210
	Ter119		TER-119	BioLegend	116210
	CD4		RM4-5	BioLegend	100514
	CD8a		53-6.7	BioLegend	100710
	CD11b		M1/70	BioLegend	101210
	CD5		53-7.3	BioLegend	100610
Reconstitution analysis of Pro-B cells in BM	7-AAD	n/a	n/a	Sigma	A9400-1MG
	B220	APC-eF780	RA3-6B2	eBioscience	47-0452-82
		APC		eBioscience	17-0193-82
	CD19	eF450	1D3	eBioscience	48-0193-82
		Pacific Blue		eBioscience	<i>Discontinued</i>
	c-Kit	APC	2B8	BioLegend	105814
	IgM	APC	II/41	BD	550676
	CD45.1	BV650	A20	BioLegend	110736
	CD45.2	AF700	104	BioLegend	109822
	Gr-1		RB6-8C5	BioLegend	108410
	NK1.1		PK136	BioLegend	108716
	Ter119		TER-119	BioLegend	116210
	CD3e		145-2C11	BioLegend	100310
	F4/80		BM8	BioLegend	123112
	CD11c		N418	BioLegend	117316
	7-AAD	n/a	n/a	Sigma	A9400-1MG
Reconstitution analysis of T cell progenitors in thymus	DAPI	n/a	n/a	Invitrogen	D3571
	CD4	BV650	RM4-5	BioLegend	100546
	CD8a	APC-eF780	53-6.7	eBioscience	47-0081-82
	CD45.1	PE-Cy7	A20	BioLegend	110730
	CD45.2	AF700	104	BioLegend	109822
	Gr-1		RB6-8C5	BioLegend	108410
	B220		RA3-6B2	BioLegend	103210
	Ter119		TER-119	BioLegend	116210
	CD11b		M1/70	BioLegend	101210
	CD19		1D3	eBioscience	15-0193-83
	NK1.1		PK136	BioLegend	108716
	CD11c		N418	BioLegend	117316
Analysis of PB platelets and erythrocytes in Vwf-CreERT2 fate mapping	DAPI	n/a	n/a	Invitrogen	D3571
	Ter119	APC	TER-119	eBioscience	17-5921-81
	CD150	BV785	TC15-12F12.2	BioLegend	115937
	CD41	PE-Cy7	MWReg30	eBioscience	25-0411-82
Analysis of PB myeloid, B and T leucocytes in Vwf-CreERT2 fate mapping	NK1.1	Pacific Blue	PK136	BioLegend	108722
	CD11b	APC	M1/70	BioLegend	101212
	CD19	PE-Cy5	1D3	eBioscience	15-0193-83
	CD4	APC-eF780	RM4-5	eBioscience	47-0042-82
	CD8a	APC-eF780	53-6.7	eBioscience	47-0081-82
	CD45.1	BUV395	A20	BD	565212
	CD45.2	BUV737	104	BD	564880
	CD150	BV785	TC15-12F12.2	BioLegend	115937
	CD41	PE-Cy7	MWReg30	eBioscience	25-0411-82
	DAPI	n/a	n/a	Invitrogen	D3571
Analysis of donor LSK-derived myeloid cells generated in vitro	CD11b	PE-Cy7	M1/70	BioLegend	101215
		PE-Cy5		BioLegend	101210
	Ly8G	Pacific Blue	1A8	BioLegend	127612
		AF647		BioLegend	127610
	CD45.1	BV605	A20	BioLegend	110736
	CD45.2	AF700	104	BioLegend	109822
	7-AAD	n/a	n/a	Sigma	A9400-1MG
	DAPI	n/a	n/a	Invitrogen	D3571
Analysis of donor LSK-derived T cell precursors generated in vitro	NK1.1	Pacific Blue	PK136	BioLegend	108722
		PE-Cy5		BioLegend	108716
	CD11b	PE-Cy5	M1/70	BioLegend	101210
		PE		eBioscience	12-0112-83
		PE		BD	557399
	CD19	PE-Cy5	1D3	eBioscience	15-0193-83
		eF450		eBioscience	48-0193-82
	CD90.2	APC-eF780	53-2.1	eBioscience	47-0902-82
	CD25	PerCP-Cy5.5	PC61	BioLegend	102029
	CD4	BV650	RM4-5	BioLegend	100546
	CD8a	APC	53-6.7	eBioscience	47-0081-83
	CD45.1	PE-Cy7	A20	BioLegend	110730
	CD45.2	AF700	104	BioLegend	109822
	DAPI	n/a	n/a	Invitrogen	D3571

n/a: not applicable

Supplementary Table 2. TaqMan gene expression assays used in Fluidigm analysis of *in vitro* myeloid and lymphoid cultures.

Gene symbol	Gene name	TaqMan assay ID	Custom assay details
<i>B2m</i>	Beta-2-microglobulin	Mm00437762_m1	
<i>Bcl11b</i>	B-cell leukemia/lymphoma 11B	Mm00480516_m1	
<i>Cd4</i>	CD4 antigen	Mm00442754_m1	
<i>Cd8a</i>	CD8 antigen, alpha chain	Mm01182108_m1	
<i>Cd3d</i>	CD3 antigen, delta polypeptide	Mm00442746_m1	
<i>Cd3e</i>	CD3 antigen, epsilon polypeptide	Mm00599683_m1	
<i>Cd3g</i>	CD3 antigen, gamma polypeptide	Mm00438095_m1	
<i>Cd16 / Fcgr3</i>	Fc receptor, IgG, low affinity III	Mm00438882_m1	
<i>Cd19</i>	CD19 antigen	Mm00515420_m1	
<i>Cd79a / Mb1</i>	CD79A antigen (immunoglobulin-associated alpha)	Mm00432423_m1	
<i>Cebpa</i>	CCAAT/enhancer binding protein (C/EBP), alpha	Mm00514283_s1	
<i>Cebpb</i>	CCAAT/enhancer binding protein (C/EBP), beta	Mm00843434_s1	
<i>Cpa3</i>	Carboxypeptidase A3, mast cell	Mm00483940_m1	
<i>Csf1r</i>	Colony stimulating factor 1 receptor	Mm01266652_m1	
<i>Csf3r</i>	Colony stimulating factor 3 receptor (granulocyte)	Mm00432735_m1	
<i>Ebf1</i>	Early B cell factor 1	Mm00432948_m1	
<i>Gata1</i>	GATA binding protein 1	Mm01352636_m1	
<i>Gata2</i>	GATA binding protein 2	Mm00492301_m1	
<i>Hprt1</i>	Hypoxanthine guanine phosphoribosyl transferase 1	Mm01545399_m1	
<i>Il2ra</i>	Interleukin 2 receptor, alpha chain	Mm00434261_m1	
<i>Il7r</i>	Interleukin 7 receptor	Mm00434295_m1	
<i>Itga2b / Cd41</i>	Integrin alpha 2b	Mm00439741_m1	
<i>Lat</i>	Linker for activation of T cells	Mm00456761_m1	
<i>Lck</i>	Lymphocyte protein tyrosine kinase	Mm00802897_m1	
<i>Lef1</i>	Lymphoid enhancer binding factor 1	Mm00550265_m1	
<i>Mmp9</i>	Matrix metalloproteinase 9	Mm00442991_m1	
<i>Mpeg1</i>	Macrophage expressed gene 1	Mm012222137_g1	
<i>Mpo</i>	Myeloperoxidase	Mm01298424_m1	
<i>Pax5</i>	Paired box gene 5	Mm00435501_m1	
<i>Pf4</i>	Platelet factor 4	Mm00451315_g1	
<i>Ptpnc / Cd45</i>	Protein tyrosine phosphatase, receptor type C	Mm01293577_m1	
<i>Rag1</i>	Recombination activating gene 1	Custom assay	Fwd: 5'-TGTGGAGCAAGGTAGCTTAGC-3' Rev: 5'-TCATCGGGTGCAGAACTGAAG-3' Probe: 5'-CATGGCTGCCTCCTTG-3'
<i>Rag2</i>	Recombination activating gene 2	Mm00501300_m1	
<i>Sfp1 / PU.1</i>	Spleen focus forming virus (SFFV) proviral integration oncogene	Mm00488142_m1	
<i>Slamf1 / Cd150</i>	Signaling lymphocytic activation molecule family member 1	Mm00443316_m1	
<i>Tcf7</i>	Transcription factor 7, T-cell specific	Mm00493445_m1	
<i>Tcrb</i>	T cell receptor beta chain (constant regions 1 and 2)	Custom assay	Fwd: 5'-GCACAATCCTCGAAACCACT-3' Rev: 5'-CCCTGATGATAGGATGCTGAA-3' Probe: 5'-GTGGCCAGAGGGCTCACCCA-3'
<i>Vwf</i>	Von Willebrand factor homolog	Mm00550376_m1	