TITLE PAGE

EX VIVO EXPANSION POTENTIAL OF MURINE HEMATOPOIETIC STEM CELLS: A RARE PROPERTY ONLY PARTIALLY PREDICTED BY PHENOTYPE

Qinyu Zhang[1], Rasmus Olofzon[1], Anna Konturek-Ciesla[1], Ouyang Yuan[1], David Bryder[1]

[1]Division of Molecular Hematology, Department of Laboratory Medicine, Lund Stem Cell Center, Faculty of Medical, Lund University, 221 84 Lund, Sweden.

Corresponding author:
David Bryder, DMH, BMC B12, Lund University, 221 84 Lund, Sweden +46706423951, David.Bryder@med.lu.se

Running title: HSC expansion *in vitro*

Key point:
Ex vivo self-renewal is an intrinsic property of rare candidate HSCs, with implications for assessments of HSC activity by transplantation.

ABSTRACT

The scarcity of hematopoietic stem cells (HSCs) restricts their use in both clinical settings and experimental research. Here, we examined a recently developed method for expanding rigorously purified murine HSCs *ex vivo.* After three weeks of culture, only 0.1% of cells exhibited the input HSC phenotype, but these accounted for almost all functional long-term HSC activity. Input HSCs displayed varying potential for *ex vivo* self-renewal, with alternative outcomes revealed by single cell multimodal RNA- and ATAC-seq profiling. While most HSC progeny offered only transient *in vivo* reconstitution, these cells efficiently rescued mice from lethal myeloablation. The amplification of functional HSC activity allowed for long-term multilineage engraftment in unconditioned hosts that associated with a return of HSCs to quiescence. Thereby, our findings identify several key considerations for *ex vivo* HSC expansion, with major implications also for assessment of normal HSC activity.

INTRODUCTION

Hematopoietic stem cells (HSCs) are the functional units in bone marrow transplantation (BMT) and gene therapy approaches via their ability for multilineage differentiation and long-term self-renewal[1]. However, restrictions in HSC numbers, the requirement of harsh host conditioning and challenges with how to genetically manipulate HSCs with retained function are barriers for BMT/gene therapy[2]. Due to the lack of reliable models for accurate assessment of human HSCs, mice continue to be the preferred experimental model for studying HSC biology. However, studying murine HSC biology also poses challenges due to the scarcity of HSCs.

Recently, culture conditions supporting efficient maintenance and expansion of murine HSCs were reported[3]. However, the expansion of functional *in vivo* HSC activity over a 28-day period (estimated to 236-899-fold) was not on par with the total cell expansion (>10^4-fold), demonstrating that large amounts of differentiated progeny were generated along with self-renewing HSC divisions[3]. Moreover, candidate HSCs (cHSCs) by phenotype represented only a minor portion of the cells at the culture end-point[3], although assessment of HSC activity based on this might be obscured by previous observations that HSCs change some phenotypes in culture[4]. While *in vitro* differentiation complicates assessments of HSC expansion, it might also be beneficial. For example, while HSCs can support lifelong hematopoiesis after bone marrow transplantation (BMT), they have limited capacity to rescue the host from lethal conditioning[5]. To address this limitation, there is a requirement for progenitor cells that can efficiently supply the host with mature blood cells, such as erythrocytes and platelets[6,7].

When transplanted, *in vivo* HSC activity is typically determined using both quantitative and qualitative parameters[8]. Most commonly, this is achieved by comparing to a known amount of competitor cells[9]. However, although the competitive repopulation assay (CRA) represents a mainstay in experimental HSC biology, it associates with shortcomings often neglected. First, while clinical BMT aims at transplanting large numbers of HSCs[10,11], experimental murine BMT typically use limited HSC grafts, which can strongly influence on HSC function[12]. Secondly, the CRA is from a quantitative perspective rather blunt, with sigmoidal rather than linear characteristics[9]. This limits the boundaries of the effects that can be measured. Thirdly, the CRA assesses HSC activity based on overall hematopoietic chimerism, which might reflect poorly on the actual HSC activity given the vastly different

turnover rates of different blood cells lineages[13]. At the same time, alternative assays such as limit dilution assays or single-cell transplantation assays are cumbersome, expensive and associated with ethical concerns. Newer developments in DNA barcoding technology have overcome many of these issues[14].

Another concern with BMT is the use of aggressive host conditioning to permit engraftment. While eradicating malignant cells is an important part of treatment regimens for leukemic patients, the fates of HSCs in conditioned hosts versus native hematopoietic contexts are dramatically different[15]. But transplantation into unconditioned recipients is inefficient and requires large amounts of transplanted HSCs to saturate available niches[16,17]. To achieve even limited HSC chimerism in unconditioned hosts can have strong long-term therapeutic effects[18] and many experimental approaches might benefit from avoiding host conditioning.

Here, we provide molecular and functional details on HSC differentiation and self-renewal following culture and define a set of parameters critical for *in vivo* assessment of HSC activity. Importantly, we find that HSC heterogeneity was a key influencing factor for HSC expansion even among rigorously purified cHSCs. This determinism has broad implications for experiments involving *in vitro* manipulation of HSCs, but also when assessing normal HSC function by BMT.

RESULTS

In vivo HSC activity is restricted to cells expressing high levels of EPCR

We first assessed the *in vivo* reconstituting activity of cHSCs with different phenotypes. The endothelial protein C receptor (EPCR or PROCR/CD201) is highly expressed on HSCs[19], while CD41 (ITGA2B) was suggested to be functionally relevant for HSC maintenance and homeostasis[20]. Based on expression of EPCR and CD41, we isolated 4 subfractions within BM SLAM LSK cells (Lin^- $Sca1^+cKit^+CD150^+CD48^{-/low}$) (Fig 1A) and competitively transplanted them at equal numbers into lethally irradiated recipients (Fig 1B). The reconstitution from the three $EPCR^{high}$ fractions was restricted to the PB myeloid lineages 4-week post-transplantation (Fig 1C), and most of these recipients displayed robust long-term (beyond 16 weeks) multilineage reconstitution with equivalent cHSC chimerism in the BM (Fig 1C-D). $EPCR^{high}CD41^-$ cells produced slightly more lymphoid cells than the $EPCR^{high}CD41^+$ cells at the endpoint (Fig S1), but the cHSC chimerism in the BM was very similar to other tested $EPCR^{high}$ fractions (Fig 1D). In contrast to Gekas et al.[20], we failed to observe that staining with the CD41 antibody (MWreg30 clone) compromises HSC engraftment. Transplantation of EPCR negative cells failed to associate with either long-term PB or BM cHSC chimerism, demonstrating lack of HSC activity within this fraction (Fig 1C-D).

Next, single $EPCR^+$ SLAM LSKs were index sorted (Fig S1B) and expanded *ex vivo* in an F12-Polyvinyl alcohol (PVA) based culture media[21]. At an early stage (8 days, Fig S1C), cultures initiated with SLAM LSK cells with higher EPCR expression proliferated less compared to those with lower EPCR expression. However, after longer expansion (25 days, Fig S1C), the EPCR higher cells generated on average a larger amount of progeny. To evaluate the HSC activity following *ex vivo* expansion and to correlate this directly to EPCR expression levels, the expanded cells from index sorted cultures of cHSCs were collected and transplanted into individual lethally irradiated mice (Fig S1D). While freshly and stringently isolated HSCs normally fail to rescue mice from lethal irradiation[5], we here entertained that culturing of HSCs in addition to inducing self-renewal, would also generate progenitors that could radioprotect the hosts. If so, survival could be used as a proxy for combined hematopoietic stem and progenitor cell (HSPC) activity. Indeed, mice receiving progeny from SLAM LSK EPCR higher cells were more efficiently radioprotected (Fig 1E). Thereafter, we initiated cultures with $EPCR^{high}$ or $EPCR^{lower}$ SLAM LSK cells

fractions (Fig S1E). Mice transplanted with cells expanded from EPCR^{high} cells radioprotected 5/5 recipients, while only two out of 5 animals receiving the progeny of EPCR^{low} cells survived long-term. As expected, these surviving mice displayed exclusive test-cell derived myelopoiesis long-term after transplantation (Fig 1F).

Taken together, these experiments established that HSC activity was restricted to EPCR^{high} HSCs, and such cells could effectively generate progeny *in vitro* that rescue mice from lethal irradiation. Therefore, we operationally defined our input cHSCs by their EPCR^{high} SLAM phenotype.

Phenotypic heterogeneity from *ex vivo* expanded cHSCs

The PVA-dependent cell culture system was reported to efficiently support HSC activity *ex vivo* over several weeks[3]. However, the functional HSC frequency did not appear on par with the total amount of cells generated by the end of the culture period, suggesting the generation also of many differentiated cells[3]. To detail this, we expanded aliquots of cHSCs *ex vivo* for 14-21 days, followed by multicolor phenotyping. From 50 cHSCs, an average of 0.5 million (14 days) and 13.6 million cells (21 days) were generated (Fig S2A) and which associated with a phenotypic heterogeneity that increased over time (Fig S2B). With the assumption that HSCs might retain their cell surface phenotype in cultures, we quantified EPCR^{high} SLAM cells and observed a decrease in their frequency over time, although this was numerically counteracted by the overall proliferation in cultures (Fig 2A). Thus, quantification of cHSCs translated into an average expansion of 54-fold (26- to 85-fold) and 291-fold (86- to 1,273-fold) at 14 and 21 days, respectively (Fig 2A).

To gain a deeper understanding of the cells generated in our *ex vivo* cultures, we collected cells from the entire culture expanded from 500 cHSCs and isolated nuclei for single cell multiome sequencing using a commercial platform from 10x Genomics. This approach combines single-cell RNA sequencing and single-cell ATAC sequencing, allowing for integrated analysis of these modalities.

Based on the transcriptomic signatures and distal motif identities, we were able to identify not only early hematopoietic stem and progenitor cells (HSPCs), but also many differentiated myeloid cells in the expansion cultures (Fig 2B), however with little evidence for lymphoid differentiation. Assessment of a condensed HSC signatures (Fig S2C) confirmed that the cHSCs were primarily located to cluster 5

(Fig 2D), with trajectory analyses suggesting multiple differentiation pathways from early HSPCs to differentiated myeloid progeny (Fig 2C).

Since the EPCR expression also condensed to cluster 5 (Fig S2D), we further investigated whether the expression of EPCR marks cHSCs *ex vivo* by collecting *ex vivo* expanded EPCR^{+} cells for single-cell multiome sequencing. These cells displayed less heterogeneity than what we observed for the total culture (Fig 2B), which included a more even distribution of the condensed HSC signature (Fig 2E) as well as the expression of EPCR (Fig S2E). Here, the most pronounced separator for the interrogated cells attributed to the cell cycle position (Fig 2F).

To explore if other markers could provide additional phenotypic information for genuine HSCs in cultures, we applied an Fgd5 reporter mouse model previously shown to selectively mark HSCs *in situ*[22]. While the Fgd5 reporter only marked a subset of cells in culture (Fig S2F), we noted heterogeneous expression of CD48 within the Fgd5 positive fraction (Fig S2F-G) and that we failed to clearly demarcate in our single-cell multimodal data.

Collectively, these results demonstrated robust expansion of cHSCs using the PVA-based culture system, albeit with additional parallel generation of a large number of more differentiated cells. While the identities of most of the cells in cultures could be clearly deduced by their combined gene expression and chromatin accessibility profiles, this approach failed to inform on the heterogeneous expression pattern of CD48 on cHSCs (Fig S2).

Functional HSC activity associates to a minor LSK SLAM EPCR^{high} fraction within *ex vivo* cultures

Given our inability to distinctly identify cHSCs by their molecular profiles, we next explored the possibility to prospectively isolate HSCs from the cHSC cultures and instead rely on their functional ability to long-term repopulate lethally irradiated hosts for readout. cHSCs were isolated and cultured *ex vivo* for 21 days and the expanded cells were separated into two equal portions: one portion was kept unfractionated, while three subpopulations of CD150^{+}LSKs: EPCR^{high}CD48$^{-/low}$, EPCR^{high}CD48^{+} and EPCR^{-} were sorted from the other portion. The fractions were competitively transplanted into lethally irradiated mice, such that each recipient received fractions

equivalent to the expansion from 50 cHSCs (EE50) along with 500,000 competitor WBM cells (Fig 3A).

Stable and very high long-term multilineage reconstitution was observed in all recipients of unfractionated cultured cells. By contrast, test-cell derived reconstitution was not recovered from EPCR^{-} cells (Fig 3B and Fig S3). EPCR^{high}CD48^{+} cells produced very high reconstitution levels short-term after transplantation, but the chimerism provided by these cells dropped considerably over time (Fig 3B and Fig S3). This differed strikingly from recipients receiving the minor fraction of EPCR^{high}CD48$^{-/low}$ cells, which presented with robust long-term multilineage reconstitution (Fig 3B and Fig S3). While almost no BM cHSC chimerism was observed from EPCR^{high}CD48^{+} cells, EPCR^{high}CD48$^{-/low}$ cells produced very high chimerism in all evaluated progenitor fractions, including for cHSCs (Fig 3C).

These data established that functional *in vivo* HSC activity following culture is restricted to the minor fraction of SLAM LSK EPCR^{high} cells.

Quantification of HSC activity in *ex vivo* expansion cultures

Expanded cHSCs presented with a vastly higher reconstitution activity compared to freshly isolated cHSCs, and mostly fell out of range for accurate quantification (Fig 3B). We therefore assessed the impact of lowering the amount of input cHSCs and enhancing the amount of competitor cells. Aliquots of 10 cHSCs were cultured and expanded for 21 days. Cells from each culture (EE10) were next mixed with 2 or 20 million (2M or 20M) competitor WBM cells, followed by transplantation (Fig 4A, ‘individual’). In parallel, we also included a group in which EE50 cHSCs were mixed with 10 million WBM cells and split among 5 lethally irradiated recipients (Fig 4A, ‘pooled’). This allowed us to assess the variability among input cHSCs,

Although robust reconstitution was observed in most recipients of ‘individual’ cultures (Table 1, Fig 4B-C, Fig S4A-B), the levels of test-cell derived cells varied extensively among these recipients. To quantify the HSC activity in each recipient, we calculated the repopulating units (RUs)[9] for each lineage and recipient separately at the experimental endpoint (Table 1). Short-lived myeloid cells has been proposed as a better indicator of ongoing HSC activity[23], and in agreement with this, we observed a high correlation between the RU^{myeloid} and RUs for cHSCs in the BM (Table 1). A much more consistent HSC activity was observed in between recipients

of ‘pooled’ cells (Table 1, Fig 4B-C). This confirmed that the differential chimerism in recipients of limited cHSC numbers relates to an unresolved cellular heterogeneity of input cHSCs. RUs per initial cHSC were calculated for freshly isolated BM cHSCs (the three EPCR^{high} groups from Fig 1B) and following 3-week culture (‘pooled’ group). This revealed 461- and 70-fold increases in PB myeloid and BM cHSC chimerism in recipients of *ex vivo* expanded cells, respectively. This is an estimate of the increase of functional HSC activity following *ex vivo* culture (Fig S4G).

We next assessed HSC clonality using a lentiviral barcoding approach. Five aliquots of 1,000 cHSCs were transduced with a barcode library[24] 48 hours after culture and then expanded for 19 days. Half of the expanded cells from each well were transplanted into one ‘parental’ recipient, respectively, while the other half was mixed with each other, with each ‘daughter’ recipient receiving 1/5 of the mixed cells (Fig 4D-i). As a proxy for the HSC activity, barcodes were extracted from BM test-derived myeloid cells 16-weeks post-transplantation. Following stringent filtering, we retrieved 223 unique barcodes in the ‘parental’ recipients, with a highly variable contribution (0.1 - 26.0%, Fig 4E-G). Representation of the same barcode in ‘parental’ and ‘daughter’ recipients demonstrates a shared clonal/HSC origin, where more robustly expanding HSCs should have a higher chance of reconstituting more ‘daughter’ recipients. Our data agreed well with this presumption (Fig 4E), and the most actively shared clones also associated with larger outputs in ‘daughter’ recipients (Fig 4F).

While our initial barcode experiments were designed to assess the *in vitro* expansion potential of cHSCs, a next set of experiments were executed to assess the clonal HSC activity following expansion. For this, cHSCs were *ex vivo* expanded for 13 days, provided with barcodes[25] overnight, and transplanted into 4 lethally irradiated recipients (Fig 4D-ii). 16-weeks post-transplantation, we recovered 573 unique barcodes. The median clone sizes of HSCs post-culture were significantly smaller than the pre-culture clones (Fig 4G), which was expected as more clones (573 vs 223) were evaluated in the ‘post-culture’ setting. However, the frequency of dominantly contributing clones was significantly less abundant (Fig 4H), demonstrating a more even contribution from individual clones in this setting. This is in line with the interpretation that expanded HSCs are functionally more equivalent than input cHSCs.

Finally, we assessed how the culture system supports the *in vivo* activity of fetal liver (FL) cHSCs. E14.5 FL-cHSCs were sorted using the same immuno-phenotype as BM cHSCs and cultured *ex vivo*. We observed an average 188-fold increase of cHSCs following 21-days culture (Fig S4D). We used the 'pooled' approach for functional evaluation and competed EE10 cells with either 2 or 20 million WBM cells (Fig S4C). While *ex vivo* expanded FL-cHSCs contributed considerably to multilineage reconstitution (Fig S4E), we observed compared to adult cHSCs 5.7- and 6.6-fold lower reconstitution from FL-cHSCs within PB myeloid cells and cHSCs, respectively (Fig S4F-G).

Taken together, these experiments demonstrated robust increases in HSC activity following culture of cHSCs, albeit slightly less for FL-cHSCs. However, clonal barcode assessments revealed substantial variation in *ex vivo* expansion potential of even stringently purified input cHSCs.

***Ex vivo* expanded cHSCs returns to a quiescent state following engraftment in unconditioned recipients**

Successful engraftment in an unconditioned setting requires large numbers of HSCs/HSPCs[16], which reportedly can be accommodated by PVA-cultured HSCs[3]. To test this, we transplanted EE100 CD45.1$^+$ cHSCs into lethally irradiated or unconditioned CD45.2$^+$ recipients (Fig S5A). In lethally irradiated hosts, cultured cells reconstituted >90% of PB cells, including with robust contribution of donor cells to the myeloid lineages. By contrast, most of the unconditioned recipients lacked long-term reconstitution (Fig S5B). Given that CD45 mismatching might be a sufficient immunological barrier to prevent HSPC reconstitution[26], we instead transplanted expanded cells into unconditioned F1 CD45.1$^+$/CD45.2$^+$ hosts (Fig 5A). Clear long-term multilineage engraftment was observed in all these hosts (Fig 5B).

In the BM, cHSCs exist for the most part in a quiescent state, which contrasts the situation in cultures (Fig S5C). To investigate to what extent *ex vivo* expanded cHSCs could return to quiescence following transplantation, we expanded 100 CD45.2$^+$ cHSCs for 21 days and stained all their progeny with the proliferation-tracking dye Cell Trace Violet (CTV) prior to transplantation into CD45.1$^+$/CD45.2$^+$ hosts. As a control, we transplanted CTV labelled unmanipulated CD45.2$^+$ cKit-enriched BM cells (Fig 5C). In these animals, splenic CD4$^+$ enriched cells were co-transplanted as a control for non-dividing cells (Fig S5D). Reconstitution and cHSC

fates were monitored biweekly for up to 8-weeks post-transplant. Although the phenotypic cHSC contribution in recipients receiving *ex vivo* expanded cells outnumbered the contribution from unmanipulated cells two-weeks post-transplantation, the levels became comparable later (Fig 5D). Recipients from both groups showed high levels of CTV signal two-weeks post-transplantation, with the majority of donor EPCR^{+} cHSCs having undergone fewer than 3 cell divisions. Over time, increasing numbers of cHSCs had proliferated within both groups and with similar patterns (Fig 5E-F, Fig S5E), with a slightly higher number of undivided cHSCs retrieved from *ex vivo* expanded HSCs at all time points (Fig 5E). The most quiescent cHSCs associated with higher EPCR levels (Fig S5F-G), again attesting to the usefulness of EPCR to predict robust HSC activity.

Therefore, *ex vivo* expanded cHSCs can efficiently engraft immunologically matched hosts without conditioning. Under such conditions, expanded cHSCs can rapidly return to quiescence following transplantation.

DISCUSSION

Although many efforts have aimed to *in vitro* expand murine HSCs, the results have for the most part been disappointing[27]. A recent study suggested that the substitution of undefined serum products with the synthetic polymer PVA is one key determinant towards this goal[3]. Here, we detailed this culture system further. We focused on the requirements of candidate input HSCs, the reproducibility of the system, the cellular output and the functional *in vivo* performance of the *in vitro* replicating HSCs. Several observations that we believe are valuable not only for future applications of the PVA culture system, but also for HSC biology in general, emerged from our work.

While the PVA-system supports HSC activity even from unfractionated BM cell preparations[28], the long-term repopulating HSCs (LT-HSCs) in cruder input cell preparations is very low. While having its own advantages, this makes it difficult to track the fates of cHSCs. Therefore, we here consistently initiated cultures with more stringently purified cHSCs. As a starting point, we used cHSCs isolated based on a current standard SLAM phenotype ($Lin^{-}Sca^{+}cKit^{+}CD48^{-}CD150^{+}$)[29]. In agreement with previous reports[30], we observed that CD201/EPCR further enriches for LT-HSCs, a knowledge that has so far not been adopted as a standard in the field[29]. This applied both to cHSCs isolated directly from the BM and following culture. In fact, despite that HSCs change many phenotypes in culture[4], the LT-HSC activity following culture was remarkably enriched not only in the EPCR positive fraction, but also for other phenotypic attributes of unmanipulated LT-HSCs. Importantly, such cHSCs represented only a very minor fraction in cultures (0.6% and 0.1% of 2- and 3-week cultures, respectively) but contained all LT-HSC activity, with the massive overall expansion in cultures leading to substantial net expansions of such cells (100- to 1,200-fold).

While we gained insights into the dynamics of *ex vivo* differentiation of cHSCs into mature lineages by simultaneously measuring ATAC signal and RNA transcripts in individual nuclei using single-cell multiome sequencing, the limited sensitivity of this method hindered the identification of more homogeneous populations of bona fide HSCs within the sorted EPCR+ cells. Therefore, we assessed the functional *in vivo* HSC activity in most of our subsequent work. While our clonal barcoding experiments unequivocally demonstrated HSC self-renewal, the large spectra in expansion potential of individual cHSCs was noteworthy. This variation was reduced when assessing HSC activity following culture and aligns with the interpretation that

the offspring of cHSCs that self-renew robustly in cultures exhibit less differences. Nonetheless, the demonstrated *in vitro* self-renewal refutes alternative views in which the culture process would mainly act to enhance the reconstitution capacity of individual HSCs, for instance by altering their homing properties and/or by the conversion of more differentiated progenitors into cells with LT-HSC activity. Therefore, our data support the view that this (or other) culture systems might not support all types of LT-HSCs previously alluded to[31]. Further support to this notion can be derived from our data on FL-cHSCs, which did not expand better than adult HSCs in the PVA system, despite that FL HSCs repopulate mice more efficiently than adult HSCs when unmanipulated[32]. In a broader sense, this associates also with the general concept of self-renewal and that has recently been highlighted in studies on native hematopoiesis, where distinct but only slowly contributing LT-HSCs co-exist with more active progenitors that can presumably also execute self-renewal divisions[33-35]. It appears that many aspects of this structure are re-created in cultures and, as we show, this requires as input the normally very slowly dividing and presumably most primitive LT-HSCs.

While the parallel self-renewal and differentiation in HSC cultures represent a fundamental difference compared to other *in vitro* stem cell self-renewal systems (e.g., for ES/iPS cells), where preferential self-renewal can be achieved, the differentiation of HSCs can be taken advantage of. We observed that even single cultured HSCs can enhance the speed of donor reconstitution and alleviate the effects of myeloablative conditioning, and the parallel self-renewal and differentiation can also be harnessed to approach HSC fate decisions at a molecular level[36]. This was exemplified by our single-cell multimodal experiments, which allowed for determinations of the differentiation trajectories of cHSCs in the culture system and their associated molecular features. However, as the expansion of genuine LT-HSC activity is far from unlimited, a shortage of HSCs is by all likelihood still an obstacle for larger-scale screening efforts such as genome-wide CRISPR screens and larger small molecule screens.

During our studies, we encountered several aspects of the CRA that are often neglected. First, the HSC activity when transplanted at limited numbers highlighted large variations in between recipients, which is in line with previous extensive single cell transplantation experiments[37-41]. While raised previously[42], taking into account not only the repopulating activity but also clonal aspects has not reached wide use.

Emerging barcoding technologies, by which many clones can be evaluated in a single mouse, makes this easier. Secondly, the progeny from a relatively small number of cultured cHSCs (10-50 cells) vastly outcompeted “standard” doses of competitor cells, making quantification unreliable/impossible. This contrasted results from freshly isolated cHSCs transplanted at the same doses used to initiate cultures. While this concern could be overcome by adjusting the graft composition (e.g. enhancing the number of competitor cells), extensive variation still existed among cultures that could be eliminated by using larger amounts of input cells and splitting the contents into individual recipients. This should have relevance also for normal HSC biology by highlighting intrinsic rather than stochastic regulation as the central determinant for persistent HSC function. A third consideration relates to the definition of ongoing HSC activity. We and others have previously established that this is best mirrored by the HSC contribution to myelopoiesis[43,44]. When we assessed cultured $EPCR^{+}CD48^{+}$ cells (with predominantly transient multilineage activity), we observed prominent long-term T cell contribution but limited myeloid reconstitution. This should not be confused with the concept of HSC lineage-bias, but rather reflects that the pool of T cells *in vivo* can be maintained by homeostatic proliferation and thus do not require continuous input from HSCs. Thus, short-lived non-self-renewing multipotent progenitors, which are generated in large numbers in culture, can contribute extensive (lymphoid) offspring in the long-term but should not be mistaken for ongoing HSC activity. This consideration is valid also for competitor grafts in the CRA (in the form of WBM cells), which contain numerous progenitor cells that can rapidly and persistently contribute to lymphopoiesis. Thus, as opposed to the classic CRA, where the overall donor contribution is assessed[9], the contribution to short-lived lineages reflects better the ongoing HSC activity, with the contribution to lymphoid lineages serving as a qualitative parameter for multipotency. In agreement with previous viral barcoding studies on HSCs[14] we observed that when cultured HSCs are barcoded and assessed long-term, there is a strong correlation between barcodes in BM progenitors and myeloid cells, while the pool of mature B cells contain clones that cannot be recovered in other lineages (data not shown).

Successful repopulation of non-conditioned hosts is based on the suggestion that large numbers of HSCs are needed to saturate those few BM niches available for engraftment[18,45]. Avoiding toxic myeloablation has apart from evident clinical implications also a precedence in experimental murine HSC biology, where lethal

myeloablation enforces vigorous HSC proliferation following transplantation[12], but which contrasts native contexts[12,33,34]. We observed that the CD45.1/CD45.2 differences are sufficiently immunogenic to mediate long-term graft failure in the non-conditioned setting, which in hindsight has been alluded to previously[18,26,46]. Thus, by matching hosts for CD45-isoforms, we could successfully obtain low-level (2-5%) long-term multilineage chimerism in adult non-myeloablated hosts from only 100 cHSCs. Intriguingly, despite being activated to proliferate *in vitro*, many cHSCs rapidly returned to a quiescent state *in vivo*, with the most dormant cHSCs exhibiting the most stringent HSC phenotype (EPCR^{high} SLAM LSKs). The possibility for non-conditioned transplantation should complement genetic lineage tracing models aimed at exploring both normal and malignant hematopoiesis in more native/physiological contexts[12,33,34].

In summary, we here characterized and detailed murine HSCs as they self-renew and differentiate *in vitro*. Although several aspects of the PVA system remain to be explored, a powerful *in vitro* self-renewal system has been a long-sought goal in HSC biology where HSCs have been notoriously difficult to uphold in an undifferentiated/self-renewal state. The merger of this system with recent advances in transplantation biology, genome engineering and single-cell techniques holds promise for many exciting discoveries of relevance to both basic and more clinically oriented hematopoietic research.

MATERIALS AND METHODS

Mice

Adult (8-10-week-old) C57Bl/6N ($CD45.2^+$) female mice were purchased from Taconic. Transgenic Fgd5-ZsGreen-2A-CreERT2 mice[22] (JAX:027789) was a kind gift from Derrick Rossi. Mice were maintained in the animal facilities at the Biomedical Center of Lund University and kept in environment-enriched conditions with 12-hour light-dark cycles and water and food provided *ad libitum*. All experimental procedures were approved by a local ethical committee (permits M186-15 and 16468/2020). All efforts were made to reduce the number of experimental animals and suffering.

BM transplantation

All mice used as recipients were 8 - 12 weeks of age. For conditioned recipients, mice were lethally irradiated (950 cGy) at least 4 h prior transplantation. The conditioned mice received prophylactic ciprofloxacin (HEXAL, 125 mg/L in drinking water) for 2 weeks beginning on the day of irradiation. All transplantations were performed through intravenous tail vein injection. The amount of transplanted test cells and competitor/support cells is described in the results section.

In vitro HSC culture

In vitro HSC cultures were performed using F12-PVA based culture conditions as previously described[3]. In brief, cHSCs (HSCs ($EPCR^{high}CD150^+CD48^{-/low}$ LSKs) were sorted into 96-well flat-bottom plates (Thermo Scientific) coated with 100 ng/ml fibronectin (Sigma) for >1 hour at 37°C prior to sorting. HSCs were sorted into 200 µl HSC media (Supplemental Table 1) and expanded at 37°C with 5% CO_2 for up to 21 days. The first media changes were performed 5 days (for ≥ 50 initial HSCs), 10 days (for 10 initial HSCs) or 15 days (for single-cell cultures) after initiation of cultures and then every 2 days. For pre-cultures associated with lentiviral transduction, the first media change was performed 24h after transduction and then as above. Cells were split at a 1:2-1:4 ratio into new fibronectin coated plates when reaching 80-90% confluency (normally every 4 days after the first split). In order to expand cHSCs to reach a workable number as well as avoiding vast numbers of differentiated cells generated after very long-term of culture (f.i. up to 4 weeks), the

cHSCs were kept expanding *ex vivo* for 3 weeks as a standard protocol. Following 14- and/or 21-day expansion, cellularity was assessed using an Automated Cell Counter (TC20, BioRad) and used for flow cytometric analyses and/or transplantation. BM cells collected from wild type animals were pooled together for sorting to initiate HSC cultures. Each initial well was considered as a technical replicate and treated/analyzed separately.

Cell preparation

Mice were euthanized by cervical dislocation after which bones (both right and left tibias, femurs, and pelvis) or spleens were extracted. Fetal cells were extracted from livers at E14.5. Bones were crushed using a mortar and pestle and BM cells were collected in ice-cold phosphate-buffered saline (PBS, Gibco) with 1% fetal bovine serum (Sigma-Aldrich) (FACS buffer), filtered (100 µm mesh) and washed. FL or spleen cells were brought into single cell suspension by grinding through a 100 µm mesh. BM or FL cells were cKit-enriched by anti-cKit-APCeFluor780 (eBioscience) staining and spleen cells were CD4-enrihced by anti-CD4-APC-Cy7 staining, followed by incubation with anti-APC conjugated beads (Miltenyi Biotec). Magnetic separation was performed using LS columns according to manufacturer's instructions (Miltenyi Biotec). cKit-enriched BM or FL cells were washed and stained with fluorescently labeled antibodies for analysis or sorting.

Peripheral blood (PB) chimerism analysis after transplantation was done as previously described[35]. In brief, blood was drawn from tail vein into FACS buffer containing 10 U/ml Heparin (Leo). After incubating with an equal volume of 2% dextran (Sigma) in PBS at 37 °C, the upper phase was collected and erythrocytes were lyzed using an ammonium chloride solution (STEMCELL Technologies) for 3 min at room temperature, followed by washing.

For analysis and/or sorting after *ex vivo* expansion, cultured cells were resuspended by pipetting and collected using FACS buffer. An aliquot was taken for cell counting using an Automated Cell Counter (TC20, BioRad).

Flow cytometry analysis and FACS sorting

Cells were kept on ice when possible, with the FACS buffer kept ice-cold. Staining, analysis and sorting were performed as previously described[12]. In brief, cells were resuspended in Fc-block (1:50, 5 x 10^6 cells/50 µl, InVivoMab) for 10 min and then

for 20 min with a twice concentrated antibody staining mixture (Supplemental Tables 2-10) at 4 °C in dark. In case biotinylated lineage markers were included, a secondary staining with streptavidin BV605 was performed for 20 min (1:400, 5 x 10^6 cells/100 µl, Sony) at 4 °C in dark. After a final wash, the cells were resuspended in FACS buffer containing propidium iodide (1:1000, Invitrogen).

All FACS experiments were performed at the Lund Stem Cell Center FACS Core Facility (Lund University) on Aria III and Fortessa X20 instruments (BD). Bulk populations were sorted using a 70-micron nozzle, 0.32.0 precision mask, and a flow rate of 2 - 3K events/s. HSCs for single cell culture were index sorted. FACS analysis was performed using FlowJo version 10.8.0 (Tree Star Inc.).

Multiome single-cell sequencing

Library preparation: Multiome sequencing experiments were performed at the Center for Translational Genomics (Lund University) using the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Bundle kit according to manufacturer's instructions (10x Genomics). 40,000 viable cells or $EPCR^+$ viable cells were sorted from *ex vivo* expanded cultures for multiome single-cell sequencing. Data has been uploaded to GEO under accession number GSE234906 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE234906)~~ and the secure token for reviewer access is **ixwtcooannanloz**~~.

Bioinformatic analysis:

Data was mainly analyzed with the Seurat[47] and Signac[48] R[49] packages, as well as the SAILERX Python package[50] for the joint-modality dimensionality reduction. All accompanying code for the single-cell multiome analysis post Cell Ranger can be found at this Github repository (https://github.com/razofz/DB_QZ_multiome), including conda specifications for the environments used with exact versions of all packages.

i. Preprocessing

The count matrices were generated with the Cell Ranger ARC tool (v2.0.0), aligned to the reference genome mm10. The aggregated dataset (containing data for both samples) were then loaded into Signac/Seurat in R where non-standard chromosomes were removed and the Seurat object split into the two samples, here

named Diverse and Immature, which corresponds to whole culture or sorted EPCR$^+$ cells, respectively. The data were then filtered on UMI, gene and peak counts, as well as nucleosome signal and TSS enrichment, with separate thresholds for the two samples. The ATAC features were classified as either distal or proximal peaks, and split accordingly. Proximal peaks were defined as a ± 2 kbp region around the transcription start site (TSS). The TSS positions were retrieved from GREAT's website, and GREAT v4 mm10 version was used. All peaks not classified as proximal received the classification "distal". Motifs were then identified for distal and proximal peaks separately[51].

ii. Downstream processing and joint-modality dimensionality reduction

The samples were processed separately. Since SAILERX uses pre-existing dimensionality reduction and clusters for the RNA modality, the RNA data was normalized and scaled, highly variable genes (HVGs) were found, principal component analysis (PCA) was run and clustering was performed, all with the Seurat standard functions with default parameters. After that, relevant data was extracted (metadata including RNA clusters, PCA embeddings, a whitelist of non-filtered out cell barcodes, ATAC peak count matrix) and with it an hdf5 file similar to the example files provided by the SAILERX repository was created (https://github.com/uci-cbcl/SAILERX). A SAILERX model was then trained on the generated hdf5 files (still sample-wise) according to their provided instructions. The resulting embeddings from the SAILERX joint-modality dimensionality were then exported and imported into the respective Seurat objects. In Seurat, the embeddings were used for graph construction, clustering and constructing UMAPs. Scores for cell cycle gene signatures were calculated and visualized (Fig 2E for the Immature sample) using Seurat's corresponding standard functions with default parameters.

iii. Differential gene expression

Differential gene expression testing was performed on the resulting clusters using Seurat's FindAllMarkers function with the HVGs found above. This analysis was used to calculate HSC signature score as described below (v. HSC signature).

iv. ATAC processing and cluster annotation

Motif activity scores were identified with chromVAR[52] for distal and proximal peaks respectively, still sample-wise. Differentially active motifs (DAMs) were then identified for each cluster. Annotation of clusters in terms of cell type was then performed manually with the help of DEGs and DAMs (Fig 2B), guided by preexisting data on the expression profiles of cell types using the Enrichr platform (https://maayanlab.cloud/Enrichr/).

v. HSC signature

A 13 gene condensed HSC signature list (Fig S2C) was obtained by mining genes associated to cluster 5 with publicly available gene expression data of HSCs from the ImmGen consortia and Bloodspot (https://servers.binf.ku.dk/bloodspot/). This signature was then used to calculate a score with Seurat's AddModuleScore function. A classification for expression of this gene signature was performed with a slightly modified version of Seurat's CellCycleScoring function, where instead of three classifications (S, G2M or Undecided/G1) two classifications were used (expressed or not expressed). This classification was then used to create the contours used in Fig 2D-E, and the signature score used for coloring the points in the same plots.

vi. Trajectory inference in the Diverse sample

Trajectory inference was performed for the Diverse sample with the slingshot package[53]. As initiating cluster, the cluster 5 was assigned (the earliest cluster), and end clusters chosen were clusters 4, 1, 6 and 7. The slingshot function of slingshot was used with otherwise default parameters. A joint UMAP with the four resulting trajectories was constructed (Fig 2C). To circumvent the problem of interfering color scales, only the top 20% expressing cells for each marker was chosen for coloring, and overlapping cells between these sets were excluded.

Assessment of HSC heterogeneity via DNA barcoding

Lentiviral barcode libraries were purchased from AddGene (No. 115645 and No. 83993) and amplified according to instructions[24,25]. AddGene 83993 is a Crispri gRNA library that was repurposed here for the purpose of lentiviral barcoding (e.g. no Cas9 was co-expressed). These libraries were kind gifts from Samantha Morris and Jonathan Weissman.

For pre-culture labelling, five wells containing 1000 CD45.1$^+$ HSCs each were sorted and cultured for 48 h. Cells were then transduced with the AddGene 115645 library (containing >5737 unique barcodes) overnight with approximately 35% transduction efficiency, which was followed by regular culture procedures. After 19 days, expanded cells were collected from each initial wells separately. Half of the expanded cells were transplanted into individual 'parental' CD45.2$^+$ recipients and the rest were mixed together and transplanted into five 'daughter' CD45.2$^+$ recipients. Each 'parental' or 'daughter' recipient received an estimated expansion equivalent to 500 cHSCs. For post-culture labelling, 3000 cHSCs were sorted from CD45.1$^+$ mice and expanded for 13 days. Thereafter, the cells were transduced with the AddGene 83993 library (containing >10090 unique barcodes) overnight at an approximate density of 100,000 cells/well. Half of the cultured and transduced cells were transplanted into 5 CD45.2$^+$ WT recipients. Each recipient received an estimated expansion equivalent of 300 cHSCs, of which approximately 15% were transduced according to assessments of transduction efficiencies. Barcode labelled cells were collected and analyzed from four recipients. For both pre- and post-culture labeling experiments, recipients were lethally irradiated and provided with 500,000 unfractionated CD45.2+ WBM cells each as support. At the experimental end points (16 weeks after transplantation), barcode labelled BM myeloid cells were sorted into 200 µl RNA lysis buffer (Norgen Biotek Corp) ("pre-culture" barcoding) or pelleted for genomic DNA extraction ("post-culture" barcoding) and proceeded to library preparation and sequencing analysis.

RNA was isolated according to the protocol of Single Cell RNA Purification Kit (Norgen Biotek Corp) and cDNA were synthesis using qScript cDNA SuperMix (Quantabio). Genomic DNA was extracted using PureLink Genomic DNA Mini Kit (Invitrogen). The 8- (pre-culture) and 21-base pair (bp) barcodes (post-culture) were amplified by PCR using Q5 High-Fidelity DNA Polymerase (New England Biolabs). Primers used are listed in Supplemental Table 11. Adapters containing index sequences (Supplemental Table 10) were added by PCR for 7 cycles. After each PCR amplification step, the products were cleaned up using SPRIselect beads (Beckman Coulter). Purified libraries were run on Agilent High Sensitivity DNA Kit chip (Agilent Technologies) to verify the expected size distribution, quantified by Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), and pooled at equimolar concentrations. Pooled libraries were loaded on an Illumina MiSeq Reagent Nano Kit

v2 flow cell following protocols for single-end 100-cycle sequencing. FASTQ files were further assessed with FastQC (v.0.11.2), including read count, base quality across reads, and guanine and cytosine (GC) content per sequence.

After extracting the reads of each barcode, the frequency in each recipient was calculated and barcodes with a frequency less than 0.1% were considered background reads. This is an arbitrary cut-off. Barcodes present in more than one 'parental' recipient were excluded from analysis (7 % and 2% of barcodes pre- and post-culture, respectively) to avoid the possibility of including HSCs transduced independently with the same barcode or potential over-represented barcodes. Reads were normalized to 10^6 total reads per recipient followed by statistical analyses.

Cell Trace Violet labelling for *in vitro* and *in vivo* HSC proliferation tracing

CD45.2$^+$ cultured cells, cKit-enriched BM cells or CD4-enriched spleen cells were collected, washed and pelleted. Cells were resuspended to a final concentration of 5 x 10^6 cells/ml and labelled with 1 µM CTV (Invitrogen) in PBS for 10 min at 37°C. The reaction was stopped by adding same volume of ice-cold FACS buffer and washing again with FACS buffer. For *in vitro* proliferation tracing, 10^5 CTV labeled cKit-enriched BM cells were plated into each well containing F12-PVA culture media (Supplemental Table 1). At each time point, half of the cultured cells were collected and stained for evaluating cHSC CTV signals (Supplemental Table 8) while the remaining cells were kept in culture for continue expansion. For *in vivo* proliferation tracing, CTV labeled EE100 cultured cells or 5 x 10^6 cKit-enriched BM cells were mixed with 2 x 10^6 CTV labeled CD4-enriched spleen cells and transplanted into one CD45.1$^+$/CD45.2$^+$ unconditioned recipients. For analysis, BM cells were cKit-enriched and stained accordingly (Supplemental Table 9). CD4 positive spleen cells were collected from the same recipient as a control for positive signal/undivided cells (Supplemental Table 10).

Statistical test

Statistical analyses were performed using Microsoft Excel and Graphpad Prism 9. Significance was calculated by Mann-Whitney tests if not otherwise specified. Statistical significances are described in figure legends for the relevant graphs. In all legends, n denotes biological replicates.

ACKNOWLEDGMENTS

We acknowledge Dr. Shabnam Kharazi for scientific discussions and technical support, Mr. Yun Sheng for advice on analysis for barcode sequencing and Prof. Hiromitsu Nakauchi and Dr. Adam C. Wilkinson for advice on cell cultures. The work was supported by grants from the Tobias Foundation, the Swedish Cancer Foundation, Barncancerfonden and the Swedish Research Council to D.B. and the Royal Physiographic Society of Lund foundation to Q.Z and A.K-C.

AUTHORSHIP CONTRIBUTIONS

Q.Z. designed the research, performed experiments, analyzed data, and wrote the manuscript; R.O. performed bioinformatic analysis and wrote the manuscript; A.K-C. and O.Y. performed experiments and analyzed data; D.B. supervised the project, was responsible for funding acquisition, designed the research, analyzed data and wrote the manuscript.

CONFLICT OF INTEREST DISCLOSURES

The authors report no conflict of interest.

REFERENCES

1 Konturek-Ciesla, A. & Bryder, D. Stem Cells, Hematopoiesis and Lineage Tracing: Transplantation-Centric Views and Beyond. *Frontiers in Cell and Developmental Biology* **10** (2022). https://doi.org/10.3389/fcell.2022.903528

2 Walasek, M. A., van Os, R. & de Haan, G. Hematopoietic stem cell expansion: challenges and opportunities. *Ann N Y Acad Sci* **1266**, 138-150 (2012). https://doi.org/10.1111/j.1749-6632.2012.06549.x

3 Wilkinson, A. C. *et al.* Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature* (2019). https://doi.org/10.1038/s41586-019-1244-x

4 Zhang, C. C. & Lodish, H. F. Murine hematopoietic stem cells change their surface phenotype during ex vivo expansion. *Blood* **105**, 4314-4320 (2005). https://doi.org/10.1182/blood-2004-11-4418

5 Nakauchi, H., Takano, H., Ema, H. & Osawa, M. Further characterization of CD34-low/negative mouse hematopoietic stem cells. *Ann N Y Acad Sci* **872**, 57-66; discussion 66-70 (1999). https://doi.org/10.1111/j.1749-6632.1999.tb08453.x

6 Yang, L. *et al.* Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* **105**, 2717-2723 (2005). https://doi.org/10.1182/blood-2004-06-2159

7 Na Nakorn, T., Traver, D., Weissman, I. L. & Akashi, K. Myeloerythroid-restricted progenitors are sufficient to confer radioprotection and provide the majority of day 8 CFU-S. *J Clin Invest* **109**, 1579-1585 (2002). https://doi.org/10.1172/JCI15272

8 Purton, L. E. & Scadden, D. T. Limiting factors in murine hematopoietic stem cell assays. *Cell Stem Cell* **1**, 263-270 (2007). https://doi.org/10.1016/j.stem.2007.08.016

9 Harrison, D. E., Jordan, C. T., Zhong, R. K. & Astle, C. M. Primitive hemopoietic stem cells: direct assay of most productive populations by competitive repopulation with simple binomial, correlation and covariance calculations. *Exp Hematol* **21**, 206-219 (1993).

10 Barnett, D. *et al.* Guideline for the flow cytometric enumeration of CD34+ haematopoietic stem cells. Prepared by the CD34+ haematopoietic stem cell working party. General Haematology Task Force of the British Committee for Standards in Haematology. *Clin Lab Haematol* **21**, 301-308 (1999). https://doi.org/10.1046/j.1365-2257.1999.00253.x

11 Trebeden-Negre, H. *et al.* Delayed recovery after autologous peripheral hematopoietic cell transplantation: potential effect of a high number of total nucleated cells in the graft. *Transfusion* **50**, 2649-2659 (2010). https://doi.org/10.1111/j.1537-2995.2010.02746.x

12 Sawen, P. *et al.* Mitotic History Reveals Distinct Stem Cell Populations and Their Contributions to Hematopoiesis. *Cell Rep* **14**, 2809-2818 (2016). https://doi.org/10.1016/j.celrep.2016.02.073

13 Bryder, D., Sasaki, Y., Borge, O. J. & Jacobsen, S. E. Deceptive multilineage reconstitution analysis of mice transplanted with hemopoietic stem cells, and implications for assessment of stem cell numbers and lineage potentials. *J Immunol* **172**, 1548-1552 (2004). https://doi.org/10.4049/jimmunol.172.3.1548

14 Lu, R., Neff, N. F., Quake, S. R. & Weissman, I. L. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in

conjunction with viral genetic barcoding. *Nat Biotechnol* **29**, 928-933 (2011). https://doi.org/10.1038/nbt.1977

15 Busch, K. & Rodewald, H. R. Unperturbed vs. post-transplantation hematopoiesis: both in vivo but different. *Curr Opin Hematol* **23**, 295-303 (2016). https://doi.org/10.1097/MOH.0000000000000250

16 Brecher, G., Ansell, J. D., Micklem, H. S., Tjio, J. H. & Cronkite, E. P. Special proliferative sites are not needed for seeding and proliferation of transfused bone marrow cells in normal syngeneic mice. *Proc Natl Acad Sci U S A* **79**, 5085-5087 (1982). https://doi.org/10.1073/pnas.79.16.5085

17 Bhattacharya, D. *et al.* Niche recycling through division-independent egress of hematopoietic stem cells. *J Exp Med* **206**, 2837-2850 (2009). https://doi.org/10.1084/jem.20090778

18 Bhattacharya, D., Rossi, D. J., Bryder, D. & Weissman, I. L. Purified hematopoietic stem cell engraftment of rare niches corrects severe lymphoid deficiencies without host conditioning. *J Exp Med* **203**, 73-85 (2006). https://doi.org/10.1084/jem.20051714

19 Balazs, A. B., Fabian, A. J., Esmon, C. T. & Mulligan, R. C. Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. *Blood* **107**, 2317-2321 (2006). https://doi.org/10.1182/blood-2005-06-2249

20 Gekas, C. & Graf, T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood* **121**, 4463-4472 (2013). https://doi.org/10.1182/blood-2012-09-457929

21 Wilkinson, A. C., Ishida, R., Nakauchi, H. & Yamazaki, S. Long-term ex vivo expansion of mouse hematopoietic stem cells. *Nat Protoc* (2020). https://doi.org/10.1038/s41596-019-0263-2

22 Gazit, R. *et al.* Fgd5 identifies hematopoietic stem cells in the murine bone marrow. *J Exp Med* **211**, 1315-1331 (2014). https://doi.org/10.1084/jem.20130428

23 Domen, J., Cheshier, S. H. & Weissman, I. L. The role of apoptosis in the regulation of hematopoietic stem cells: Overexpression of Bcl-2 increases both their number and repopulation potential. *J Exp Med* **191**, 253-264 (2000). https://doi.org/10.1084/jem.191.2.253

24 Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219-224 (2018). https://doi.org/10.1038/s41586-018-0744-4

25 Horlbeck, M. A. *et al.* Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife* **5** (2016). https://doi.org/10.7554/eLife.19760

26 van Os, R. *et al.* Immunogenicity of Ly5 (CD45)-antigens hampers long-term engraftment following minimal conditioning in a murine bone marrow transplantation model. *Stem Cells* **19**, 80-87 (2001). https://doi.org/10.1634/stemcells.19-1-80

27 Wilkinson, A. C., Igarashi, K. J. & Nakauchi, H. Haematopoietic stem cell self-renewal in vivo and ex vivo. *Nature reviews. Genetics* **21**, 541-554 (2020). https://doi.org/10.1038/s41576-020-0241-0

28 Ochi, K., Morita, M., Wilkinson, A. C., Iwama, A. & Yamazaki, S. Non-conditioned bone marrow chimeric mouse generation using culture-based enrichment of hematopoietic stem and progenitor cells. *Nat Commun* **12**, 3568 (2021). https://doi.org/10.1038/s41467-021-23763-z

29 Challen, G. A., Pietras, E. M., Wallscheid, N. C. & Signer, R. A. J. Simplified murine multipotent progenitor isolation scheme: Establishing a consensus approach for multipotent progenitor identification. *Exp Hematol* **104**, 55-63 (2021). https://doi.org/10.1016/j.exphem.2021.09.007

30 Kent, D. G. *et al.* Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood* **113**, 6342-6350 (2009). https://doi.org/10.1182/blood-2008-12-192054

31 Haas, S., Trumpp, A. & Milsom, M. D. Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. *Cell Stem Cell* **22**, 627-638 (2018). https://doi.org/10.1016/j.stem.2018.04.003

32 Rebel, V. I., Miller, C. L., Eaves, C. J. & Lansdorp, P. M. The repopulation potential of fetal liver hematopoietic stem cells in mice exceeds that of their liver adult bone marrow counterparts. *Blood* **87**, 3500-3507 (1996).

33 Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322-327 (2014). https://doi.org/10.1038/nature13824

34 Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* **518**, 542-546 (2015). https://doi.org/10.1038/nature14242

35 Sawen, P. *et al.* Murine HSCs contribute actively to native hematopoiesis but with reduced differentiation capacity upon aging. *Elife* **7** (2018). https://doi.org/10.7554/eLife.41258

36 Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367** (2020). https://doi.org/10.1126/science.aaw3381

37 Yamamoto, R. *et al.* Large-Scale Clonal Analysis Resolves Aging of the Mouse Hematopoietic Stem Cell Compartment. *Cell Stem Cell* **22**, 600-607 e604 (2018). https://doi.org/10.1016/j.stem.2018.03.013

38 Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154**, 1112-1126 (2013). https://doi.org/10.1016/j.cell.2013.08.007

39 Benz, C. *et al.* Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell* **10**, 273-283 (2012). https://doi.org/10.1016/j.stem.2012.02.007

40 Carrelha, J. *et al.* Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature* **554**, 106-111 (2018). https://doi.org/10.1038/nature25455

41 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). https://doi.org/10.1016/j.stem.2013.05.014

42 Ema, H. & Nakauchi, H. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood* **95**, 2284-2288 (2000).

43 Norddahl, G. L. *et al.* Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. *Cell Stem Cell* **8**, 499-510 (2011). https://doi.org/10.1016/j.stem.2011.03.009

44 Domen, J. & Weissman, I. L. Hematopoietic stem cells and other hematopoietic cells show broad resistance to chemotherapeutic agents in vivo when overexpressing bcl-2. *Exp Hematol* **31**, 631-639 (2003). https://doi.org/10.1016/s0301-472x(03)00084-5

45 Czechowicz, A., Kraft, D., Weissman, I. L. & Bhattacharya, D. Efficient transplantation via antibody-based clearance of hematopoietic stem cell niches. *Science* **318**, 1296-1299 (2007). https://doi.org/10.1126/science.1149726

46 Xu, H., Exner, B. G., Chilton, P. M., Schanie, C. & Ildstad, S. T. CD45 congenic bone marrow transplantation: evidence for T cell-mediated immunity. *Stem Cells* **22**, 1039-1048 (2004). https://doi.org/10.1634/stemcells.22-6-1039

47 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021). https://doi.org/10.1016/j.cell.2021.04.048

48 Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat Methods* **18**, 1333-1341 (2021). https://doi.org/10.1038/s41592-021-01282-5

49 Team, R. C. R: A language and environment for statistical computing. . (2022). https://doi.org/https://www.R-project.org/

50 Cao, Y. *et al.* Integrated analysis of multimodal single-cell data with structural similarity. *Nucleic acids research* **50**, e121 (2022). https://doi.org/10.1093/nar/gkac781

51 Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic acids research* **48**, D87-D92 (2020). https://doi.org/10.1093/nar/gkz1001

52 Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14**, 975-978 (2017). https://doi.org/10.1038/nmeth.4401

53 Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018). https://doi.org/10.1186/s12864-018-4772-0

TABLES

Table 1. RUs for each lineage in PB and for BM cHSCs of each recipient.

Recipients		*PB*	*B cells*	*T cells*	*Myeloid* cells	*BM cHSCs*
Individual BM EE10 + 2M	***#1***	6	6	6	4	12
	#2	31	35	19	27	75
	#3	21	23	17	35	70
	#4	98	144	46	57	21
	#5	1	1	2	0.03	0.4
Pooled BM EE10 + 2M	***#1***	101	78	173	583	49
	#2	110	84	120	516	82
	#3	87	68	103	248	130
	#4	101	70	149	384	187
	#5	65	127	186	331	112

** 1 RU equals to the average reconstitution activity of 1 x10 ^{5} WBM cells.*

Figures

Figure 1

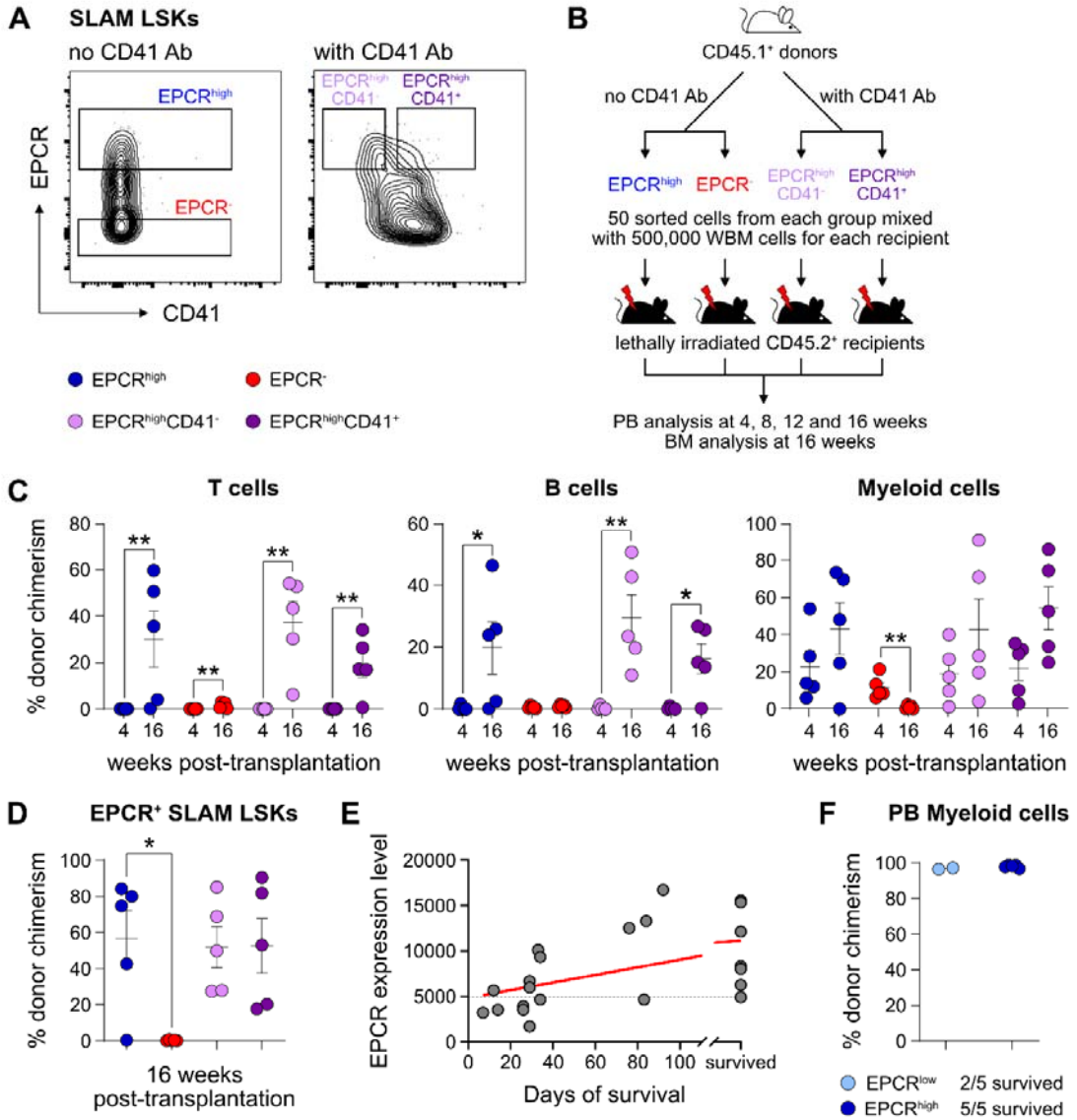


Figure 1. EPCR expression within the BM SLAM LSK fraction is a high-confidence predictor of transplantation-associated HSC activity.

(A) Expression patterns of EPCR and CD41 within the BM LSK SLAM fraction. Gates depict the assessed cell fractions.

(B) Strategy used to assess the correlation of EPCR and CD41 expression to the *in vivo* HSC activity.

(C) Test-cell derived chimerism in PB 4 and 16 weeks post-transplantation.

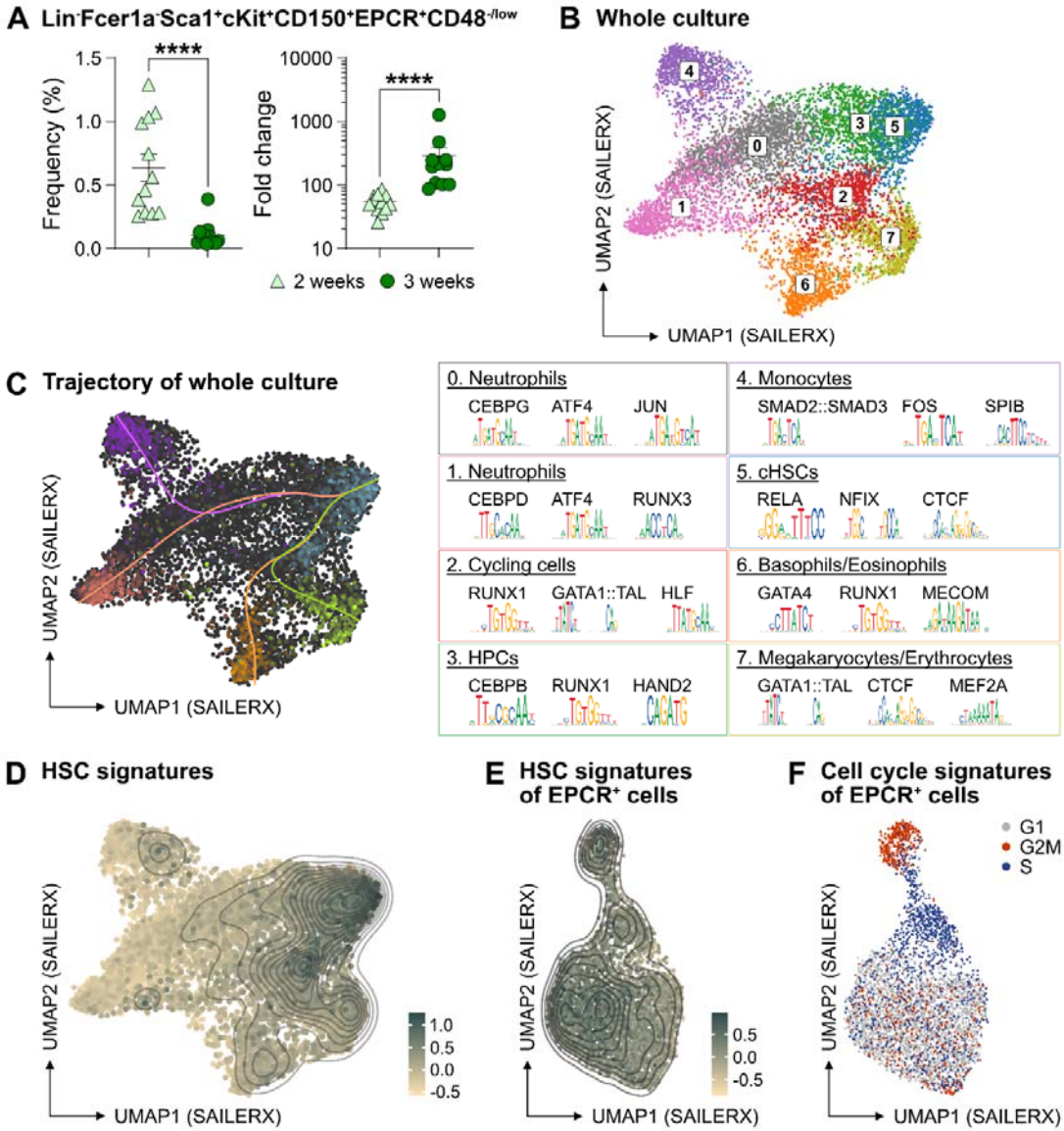
(D) Test-cell derived chimerism in BM $EPCR^{+}$ SLAM LSKs 16 weeks post-transplantation.

(E) Correlation between duration of radioprotection and EPCR expression levels. BM LSK SLAM cells were co-stained with EPCR and index-sorted at one cell per well, cultured for 21 days, and the content of each well transplanted to individual recipients (n = 22). Correlation to mortality of individual mice was made by assessing which well was transplanted into which mouse and coupling this to the index-sorting information. The grey dash line indicated the separation of EPCR higher (>4900) or lower expression (<4900) cHSCs.

(F) Donor contribution in PB myeloid cells. Mice were transplanted with *ex vivo* expanded cells from either 50 SLAM LSK $EPCR^{low}$ (n = 5) or 50 SLAM LSK $EPCR^{high}$ cells (n = 5) and assessments made 16 weeks after transplantation.

All data points depict values in individual recipients. Error bars denote SEM. The asterisks indicate significant differences. *, $p < 0.05$; **, $p < 0.01$. In E) a regression line was generated based on an end-point survival of 150 days (the time at which the experiment was terminated).

Figure 2



885

886

Figure 2. The phenotypic properties of HSC expansion cultures.

(A) Frequency and fold change of phenotypic cHSCs (EPCR^{high} SLAM LSKs) in *ex vivo* cultures after 14 or 21 days of culture. Data points depict values from individual cultures initiated from 50 cHSCs. Error bars denote SEM. The asterisks indicate significant differences. ****, $p < 0.0001$.

(B) UMAP (based on SAILERX dimensionality reduction) of single-cell multiome profiling of cells expanded *ex vivo* for 21 days. Cell type annotations were derived using marker gene signatures and distal motif identities.

(C) Trajectory analysis of lineage differentiation for cells expanded *ex vivo* (left), with the top 3 scoring TF motifs of each cluster (right).

(D) Expression of HSC signature on whole culture.

(E) Expression of HSC signature of EPCR^{+} cells sorted from *ex vivo* cultures.

(F) Cell cycle phase classification of EPCR^{+} cells sorted from *ex vivo* cultures.

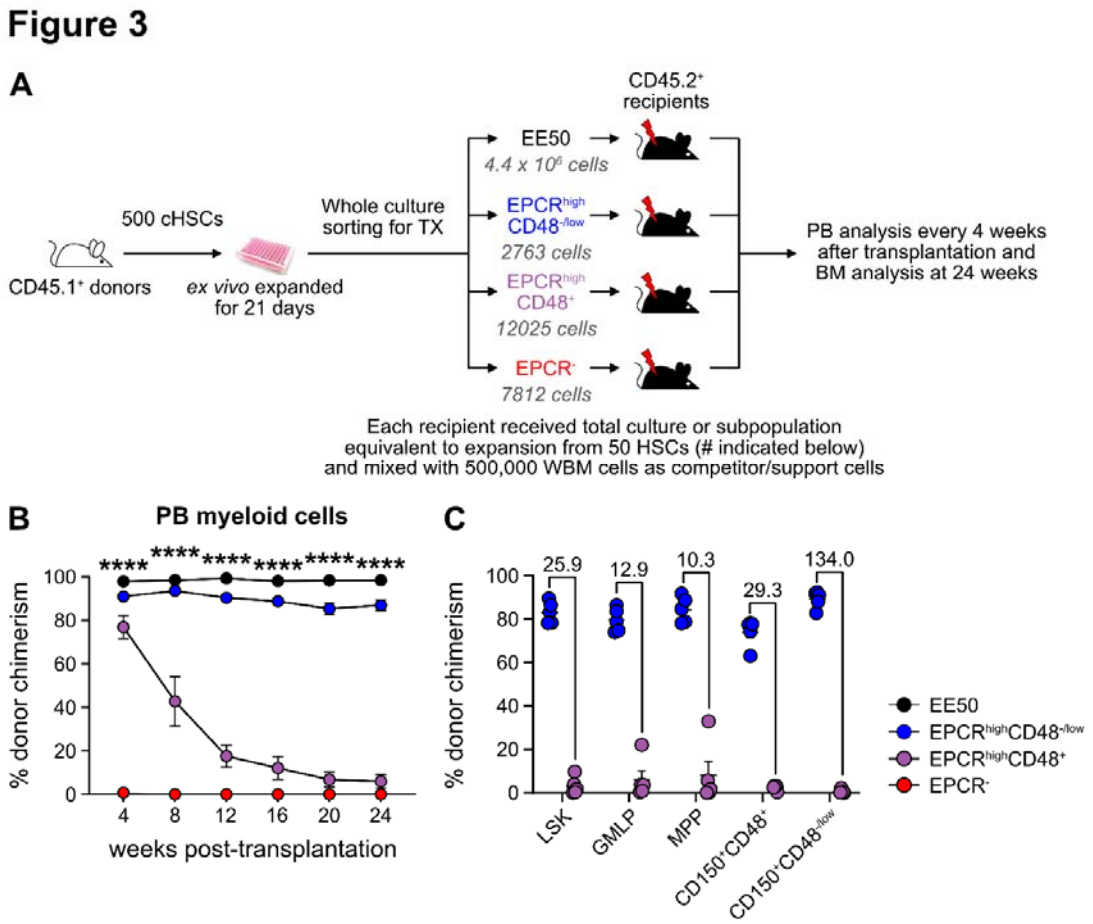


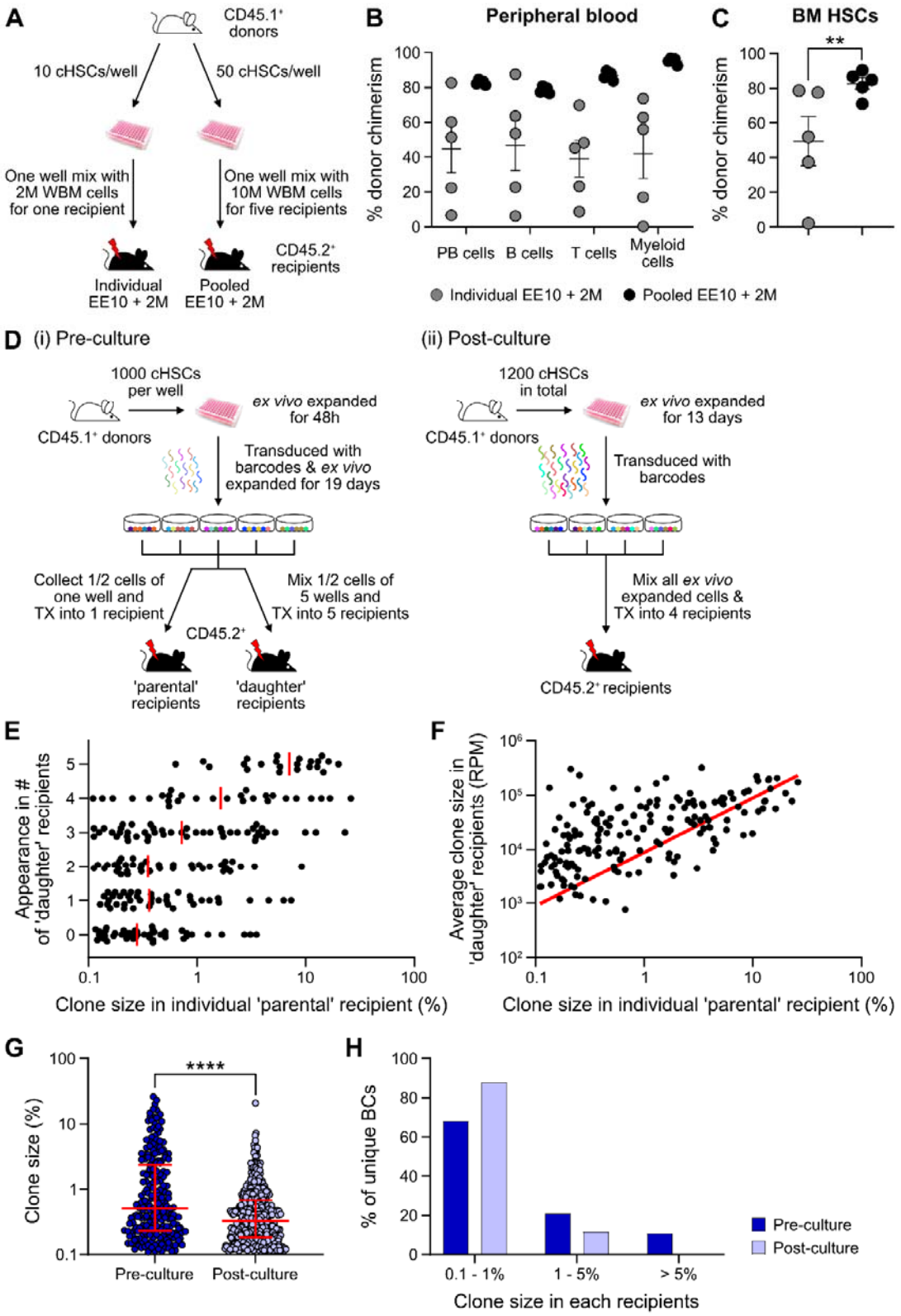
Figure 3. HSC activity can be prospectively isolated from cHSC cultures and associates to a minor EPCR$^+$ SLAM LSK fraction.

(A) Strategy to assess the *in vivo* HSC activity of subfractions from *ex vivo* cultures.

(B) Test-cell derived chimerism in PB myeloid lineages over 24-weeks post-transplantation. Data represent mean values ± SEM (n = 5 for each group). A one-way ANOVA test was applied and the asterisks indicate significant differences among the four groups. ****, $p < 0.0001$.

(C) Test-cell derived chimerism in BM progenitor subsets 24-weeks post-transplantation. Numbers indicate fold differences between the EPCR$^+$ CD48$^{-/low}$ and EPCR$^+$CD48$^+$ fractions, and data points depict chimerism levels in individual recipients.

Figure 4



914

915

Figure 4 Quantification of repopulating activity from expanded cHSCs.

(A) Competitive transplantation strategies used to assess the repopulation of *ex vivo* expanded cHSCs.

(B) Test-cell derived PB reconstitution 16-weeks post-transplantation. Symbols denote individual mice and means ± SEM.

(C) BM cHSCs chimerism 16 weeks post transplantation. Symbols denote individual mice, and means ± SEM.

(D) Barcode approaches used to assess the clonal HSC contribution of *ex vivo* expanded HSCs before (i) or after (ii) *ex vivo* expansion.

(E) Clone sizes of unique barcodes in 'parental' recipients and their appearance in 'daughter' recipients, demonstrating extensive variation in expansion capacity among individual HSCs. Red lines indicate the median clone size in 'parental' recipients.

(F) Clone sizes of unique barcodes detected in BM myeloid cells of 'parental' recipients and their corresponding contribution in 'daughter' recipients. The red line denotes the correlation/linear regression.

(G) Clone sizes in 'parental' recipients transplanted with 'pre-culture' barcoded cells, or in recipients of 'post-culture' barcoded cells. Median clone sizes are shown with interquartile ranges.

(H) Frequency of barcodes and their clone sizes in recipients of pre- or post-cultured barcoded HSCs. BCs, barcodes.

All data points depict values in individual recipients or barcodes. Asterisks indicate significant differences. *, $p < 0.05$; ****, $p < 0.0001$.

Figure 5

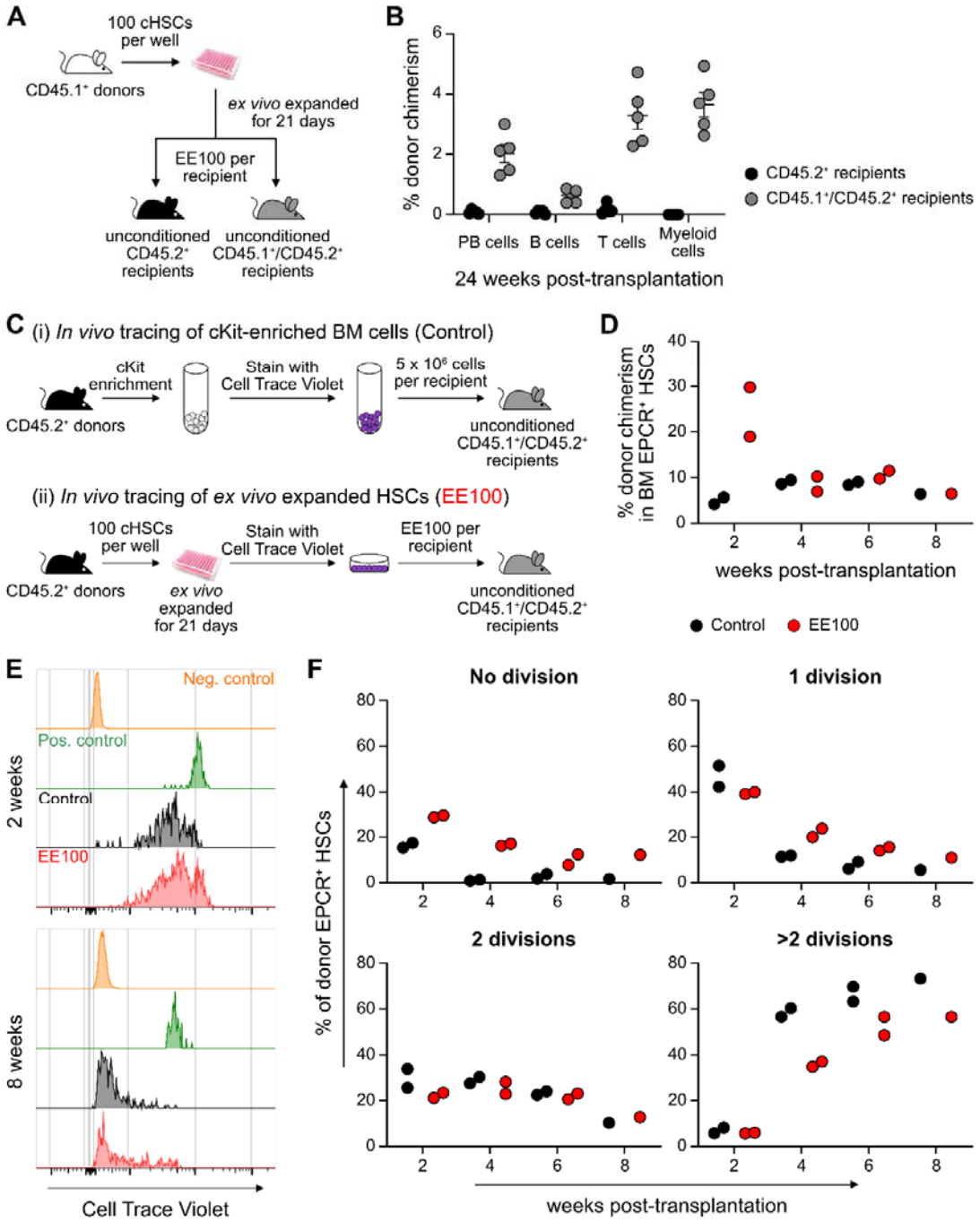

A
100 cHSCs per well
CD45.1⁺ donors
ex vivo expanded for 21 days
EE100 per recipient
unconditioned CD45.2⁺ recipients
unconditioned CD45.1⁺/CD45.2⁺ recipients
B
% donor chimerism
PB cells
B cells
T cells
Myeloid cells
24 weeks post-transplantation
CD45.2⁺ recipients
CD45.1⁺/CD45.2⁺ recipients
C
(i) In vivo tracing of cKit-enriched BM cells (Control)
cKit enrichment
Stain with Cell Trace Violet
5 x 10⁶ cells per recipient
CD45.2⁺ donors
unconditioned CD45.1⁺/CD45.2⁺ recipients
(ii) In vivo tracing of ex vivo expanded HSCs (EE100)
100 cHSCs per well
Stain with Cell Trace Violet
EE100 per recipient
CD45.2⁺ donors
ex vivo expanded for 21 days
unconditioned CD45.1⁺/CD45.2⁺ recipients
D
% donor chimerism in BM EPCR⁺ HSCs
weeks post-transplantation
Control
EE100
E
Neg. control
Pos. control
Control
EE100
2 weeks
8 weeks
Cell Trace Violet
F
No division
1 division
2 divisions
>2 divisions
% of donor EPCR⁺ HSCs
weeks post-transplantation

Figure 5. Cultured cHSCs allow for transplantation into non-conditioned syngeneic recipients.

(A) Strategy to assess the ability of *ex vivo* expanded cHSCs to engraft unconditioned recipients.

(B) Test-cell derived PB reconstitution 24-weeks post-transplantation. Symbols denote individual mice and means ± SEM.

(C) Strategy used to assess the *in vivo* proliferation of *ex vivo* expanded cHSCs.

(D) BM cHSC chimerism 2-8 weeks post transplantation.

(E) Representative CTV label profiles of donor $EPCR^+$ cHSCs compared to negative control signal (host $EPCR^+$ cHSCs) and positive signal (donor $CD4^+$ spleen cells) at 2 or 8 weeks post-transplantation.

(F) Donor $EPCR^+$ cHSCs were evaluated for the number of cell divisions they had undergone through 8 weeks post-transplantation.

All data points depict values in individual recipients.