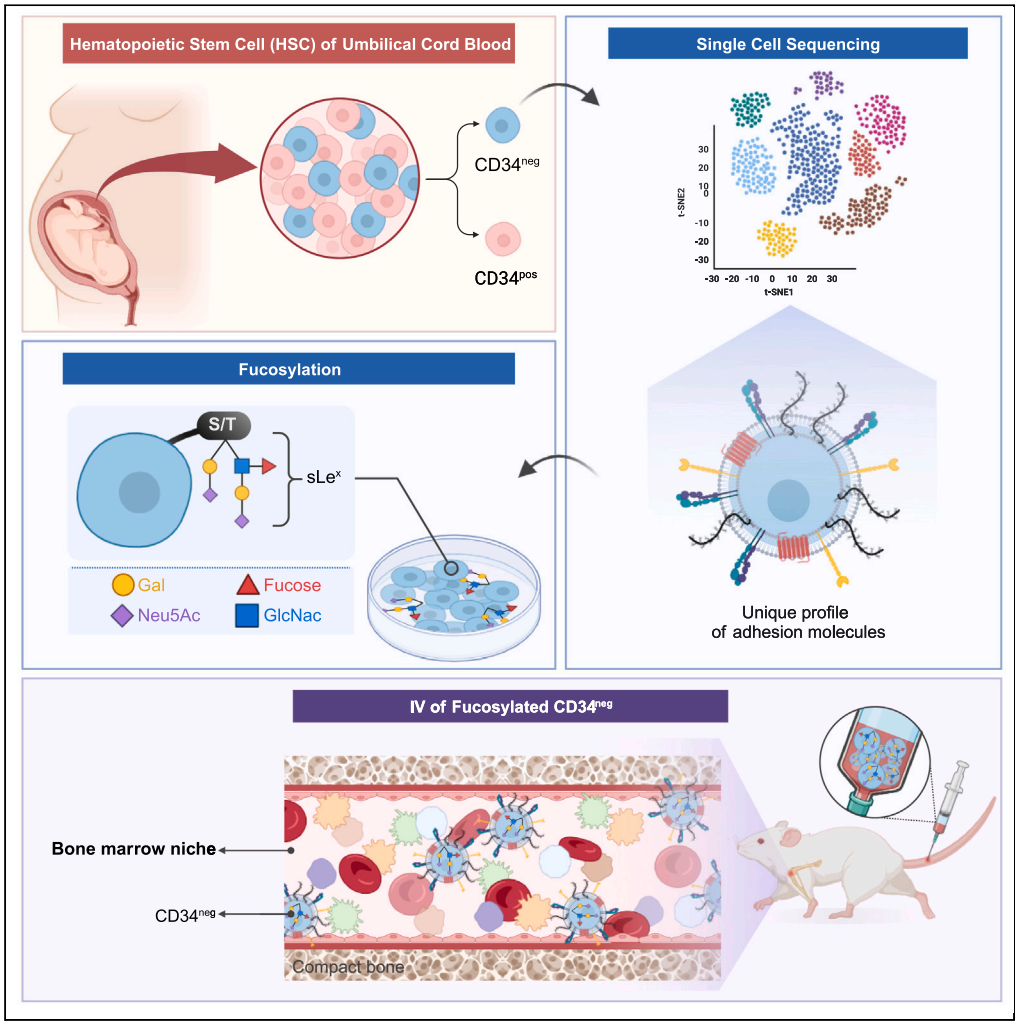
iScience

CellPress
OPEN ACCESS

Article

α1,3-fucosylation treatment improves cord blood CD34 negative hematopoietic stem cell navigation



Asma S. Al-Amoodi, Jing Kai, Yanyan Li, ..., Alfonso Saera-Vila, Satoshi Habuchi, Jasmeen S. Merzaban

jasmeen.merzaban@kaust.edu.sa

Highlights

CD34^{neg} stem cell migration to bone marrow improves after α1,3-fucosylation treatment

Engraftment of treated CD34^{neg} HSCs rises 7-fold with the creation of Sialyl Lewis X

CD34^{neg} HSCs have long-term blood cell regeneration capacity spanning several months

ScRNA-seq: CD34^{neg} HSCs surpass CD34^{pos} in adhesion molecules that interact with niche

Al-Amoodi et al., iScience 27, 108882
February 16, 2024
https://doi.org/10.1016/j.isci.2024.108882

iScience

CellPress
OPEN ACCESS

Article

α1,3-fucosylation treatment improves cord blood CD34 negative hematopoietic stem cell navigation

Asma S. Al-Amoodi,[1] Jing Kai,[1] Yanyan Li,[1] Jana S. Malki,[1] Abdullah Alghamdi,[1] Arwa Al-Ghuneim,[1] Alfonso Saera-Vila,[2] Satoshi Habuchi,[1] and Jasmeen S. Merzaban[1,3,4,*]

SUMMARY

For almost two decades, clinicians have overlooked the diagnostic potential of CD34^{neg} hematopoietic stem cells because of their limited homing capacity relative to CD34^{pos}HSCs when injected intravenously. This has contributed to the lack of appeal of using umbilical cord blood in HSC transplantation because its stem cell count is lower than bone marrow. The present study reveals that the homing and engraftment of CD34^{neg}HSCs can be improved by adding the Sialyl Lewis X molecule via α1,3-fucosylation. This unlocks the potential for using this more primitive stem cell to treat blood disorders because our findings show CD34^{neg}HSCs have the capacity to regenerate cells in the bone marrow of mice for several months. Furthermore, our RNA sequencing analysis revealed that CD34^{neg}HSCs have unique adhesion pathways, downregulated in CD34^{pos}HSCs, that facilitate interaction with the bone marrow niche. Our findings suggest that CD34^{neg} cells will best thrive when the HSC resides in its microenvironment.

INTRODUCTION

Transplanting hematopoietic stem cells (HSCs) has long been a routine treatment for blood disorders due to the regeneration capacity of these cells.[1] The success of such treatments mostly depends on the presence of HSCs that are positive for the characteristic stem cell marker, CD34, which are found in greater frequency in the bone marrow, mobilized peripheral blood (mPB) and umbilical cord blood than stem cells that are negative for this marker. The importance of the **CD34^{neg}** segment of HSCs has, therefore, tended to be sidelined due not only to its smaller quantity but also because, being more primitive, the cells have limited migratory capacity when injected into the blood stream,[2] the primary delivery method in hematopoietic stem cell transplantation (HSCT). Our research seeks to demonstrate the significant untapped potential of using **CD34^{neg}**HSCs more intentionally and effectively in transplantation. We will show that it is possible to improve the migration and engraftment of **CD34^{neg}**HSCs through intervening in their adhesion mechanisms, as well as demonstrate the benefits of ensuring their delivery to the bone marrow, particularly to prolonging cell regeneration.

Over the past 10 years, a growing body of research has lent support to the inclusion of **CD34^{neg}** cells in HSCT, beginning with the Bonnet et al.[3] study of 2013 that found **CD34^{neg}** cells had certain characteristics that **CD34^{pos}** cells lacked: namely, active Notch pathway signaling and repressed canonical Wnt pathway signaling, both of which allow **CD34^{neg}** cells to sustain a resting state for longer and, therefore, continue to multiply and reproduce new blood cells.[4] Prior to this study, **CD34^{neg}** had been largely overlooked since the early 2000s due to research by Gao[5] and others suggesting that the lower supply meant it was less practical to use clinically in repopulating blood cells compared with **CD34^{pos}**. Yet in the late 1990s, scientists including Bhatia pointed to the distinct cell regenerative capacity of **CD34^{neg}**HSCs despite their low frequency[6] while a study by Zanjani et al.[7] found that **CD34^{neg}**HSCs are capable of differentiating into **CD34^{pos}** progenitors and hematopoietic lineage cells. Nakamura et al.[8] added to this knowledge in 1999 by identifying the particular *in vitro* culture conditions that allowed **CD34^{neg}**HSCs to be maintained and multiplied. More recently, research from the Sonoda lab[9–11] as shown that when **CD34^{neg}**HSCs are injected directly into the bone marrow, they have a similar capacity to **CD34^{pos}** to give rise to hematopoietic cells in mouse blood after 8–12 weeks. The lab also brought the unique qualities of **CD34^{neg}**HSCs to light: Namely, that they are more primitive and quiescent than **CD34^{pos}** cells, and therefore able to repopulate the bone marrow for a longer period, giving them the qualifier of long-term HSCs (LT-HSCs).

To fully exploit these benefits offered by **CD34^{neg}**HSCs, the challenge lies in improving their delivery via the intravenous injection route that's become synonymous with HSCT. This led us to the following research questions: Can we improve the migration and engraftment of **CD34^{neg}**HSCs *in-vivo* by enhancing their adhesion mechanism? And furthermore, what insights can single-cell RNA sequencing reveal about the advantages of ensuring **CD34^{neg}**HSC transplantation? We chose to use umbilical cord blood for our experiments to highlight the

[1]Bioscience Program, Biological and Environmental Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia
[2]Sequentia Biotech SL, Barcelona, Spain
[3]KAUST Smart-Health Initiative, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[4]Lead contact
*Correspondence: jasmeen.merzaban@kaust.edu.sa
https://doi.org/10.1016/j.isci.2024.108882

potential of exploiting the most widely available source of HSCs, often disregarded by clinicians due to its lower concentration of stem cells than bone marrow, where the stem cell count is rich but harder to extract.[1]

The low adhesion capacity of **CD34^{neg}**HSCs is due to the absence on their surface of the carbohydrate ligand sialyl Lewis X (sLe^{x}), an essential navigation molecule that directs HSCs to the bone marrow by promoting binding to E-selectin, constitutively expressed on the bone marrow endothelium. When sLe^{x} was added to **CD34^{neg}**HSCs using α1,3 fucosylation, we found that: i) **CD34^{neg}**HSCs gained the ability to migrate *in vivo* to the bone marrow and, once they reach their new home, there was an about 7-fold improvement in engraftment compared with untreated cells; and ii) **CD34^{neg}**HSCs were able to regenerate blood cells in mouse models for several months. Single cell RNA-sequencing studies, meanwhile, underscored the importance of ensuring delivery of **CD34^{neg}**HSCs in transplantation. For one, **CD34^{neg}**HSCs naturally contain upregulated adhesion pathways that promote interaction with the bone marrow microenvironment, pathways that are downregulated in **CD34^{pos}** cells. This suggests that their presence in the bone marrow niche would contribute to blood cell differentiation and proliferation. Furthermore, erythrocyte/megakaryocyte genes are highly expressed on **CD34^{neg}**HSCs, supporting the notion that **CD34^{neg}** may play a role in the development of these precursors – for red blood cell formation and platelets, respectively – in the bone marrow. In sum, even though they are limited in number, **CD34^{neg}** cells can go a long way in improving regenerative medicine.

RESULTS

Two different methods effectively isolate CD34^{neg} hematopoietic stem cells based on cell markers

To isolate the population of **CD34^{neg}**HSCs within cord blood mononuclear cells, we used two methods that employ cell surface markers to sift out the more mature cells. These methods have been proven to leave only a concentration of these authentic stem cells that have the greatest capacity to generate hematopoietic cells *in vivo* when injected directly into the bone marrow.[6,9,12]

In the first method, we depleted more mature lineage-positive cells with the autoMACS separator, using the cocktail of monoclonal antibodies that target lineage-committed cells. What remained were lineage-negative (Lin^{neg}) fraction of cells enriched for stem and progenitor cells. These were then sorted according to the stem cell marker expression of CD38 and CD34 (Figure 1A). As illustrated in Figure 1A, Method 1 allowed us to divide the cells into two groups and, consistent with previous studies,[6,13,14] Lin^{neg}CD38^{neg}**CD34^{neg}** contained a higher number of cells than Lin^{neg}CD38^{neg}**CD34^{pos}**. The reason for this is, apart from HSCs and specialized endothelial cells,[15] most cells are negative for the CD34 marker, necessitating the removal additional lineage markers to narrow down a population of possible stem cells. The phenotypic purity of the sorted cells consistently exceeded 90% in post-sorting flow cytometric analysis (Figure 1B).

Method 2 involved staining the lineage-negative (Lin^{neg}) fraction of cells that were separated using autoMACS for 18 additional FITC conjugated lineage markers, as described in STAR Methods, as well as for characteristic stem cell markers CD34, CD133, and CD45 (Figure 1C). To further enrich a population of **CD34^{neg}**HSCs[9] from the autoMACS-separated Lin^{neg} fraction, we first identified the CD45^{pos} leukocytes, which accounted for ~40% of the total (Figure 1D). We next turned to identifying the 18Lin^{neg} fraction within this CD45^{pos} subset, which accounted for ~40%. Within this narrowed-down pool, we next focused on **CD34^{pos}** and **CD34^{neg}** populations, which made up 70.5 $\pm$ 12.4% (60.8%–83.5%) and 10.5 $\pm$ 5% (4.5%–16.8%) respectively. The **CD34^{pos}** cells were then further gated for CD38 and CD133 to generate the final sorted population of stem cells, namely, 18Lin^{neg}**CD34^{pos}**CD133^{pos} (73.7 $\pm$ 8.88% (63.7%–85%)). The **CD34^{neg}** population was gated for CD133, yielding a final sorted population of stem cells, 18Lin^{neg}**CD34^{neg}**CD133^{pos} (8.3 $\pm$ 3.87% (13.4%–29.7%)). As illustrated in Figure 1E, as high as 85% of **CD34^{pos}** cells expressed CD133^{pos}, while the ratio was less than 30% in **CD34^{neg}** cells.

To evaluate the effectiveness of the sorting, RT-qPCR analysis was conducted to assess the expression of CD34 in each population. As expected, the sorted **CD34^{pos}** cells displayed a higher expression of CD34 than the sorted **CD34^{neg}** cells (Figure 1F). Having succeeded at isolating cells via both methods, we then used these cells for downstream analysis.

Cord blood CD34^{pos} and CD34^{neg} hematopoietic stem cells display distinct homing molecule expression profiles

Various studies have observed that **CD34^{neg}**HSCs do not migrate and home efficiently in mouse models when they are injected intravenously.[11,16,17] Therefore, we set out to investigate whether there is a difference in the homing molecules apparent on **CD34^{pos}** versus **CD34^{neg}**HSCs. sLe^{x} (or its isomer sLe^{a}) is a sialofucosylated tetrasaccharide homing molecule found on glycoproteins and glycolipids that allows it to interact with selectins on the endothelium.[18]

As shown in Figure 2, the **CD34^{neg}**HSCs enriched using Method 1 (Lin^{neg}CD38^{neg}**CD34^{neg}**) (Figure 1A) and Method 2 (18Lin^{neg}**CD34^{neg}**CD133^{pos}) (Figure 1C), exhibited lower expression of sLe^{x} compared with the corresponding Lin^{neg}CD38^{neg}**CD34^{pos}** (~85%; $p < 0.001$, n = 3 independent experiments) and 18Lin^{neg}**CD34^{pos}**CD133^{pos} (~89%; $p < 0.0001$, n = 3 independent experiments), respectively (Figures 2A and 2E).

We previously reported that **CD34^{neg}** cells isolated using Method 1 lacked E-selectin binding.[13] We next analyzed the expression of E-selectin ligands,[19] chemokine receptors and integrins to characterize other essential homing molecules. Method 1 **CD34^{pos}** and **CD34^{neg}**HSCs expressed similar levels of CD44 and varying levels of CD43 and CD162/PSGL-1 (Figures 2B and 2F), while in the case of Method 2 HSCs, there was differential CD44 expression between **CD34^{pos}** and **CD34^{neg}**. There were no significant differences in the expression of chemokine receptor CXCR4, which is involved in the homing of HSCs to the bone marrow,[20,21] or in the integrins involved in HSC homing (β1, CD11a, CD49e, and CD49d)[22] (Figures 2C and 2E).

Clonogenicity assays showed that Method 1 **CD34^{neg}** cells did not, in most cases, result in differentiated colonies (data not shown), as previously reported in other studies.[6,8,14] However, in some trials, a very low number of differentiated colonies, ~four to five times lower

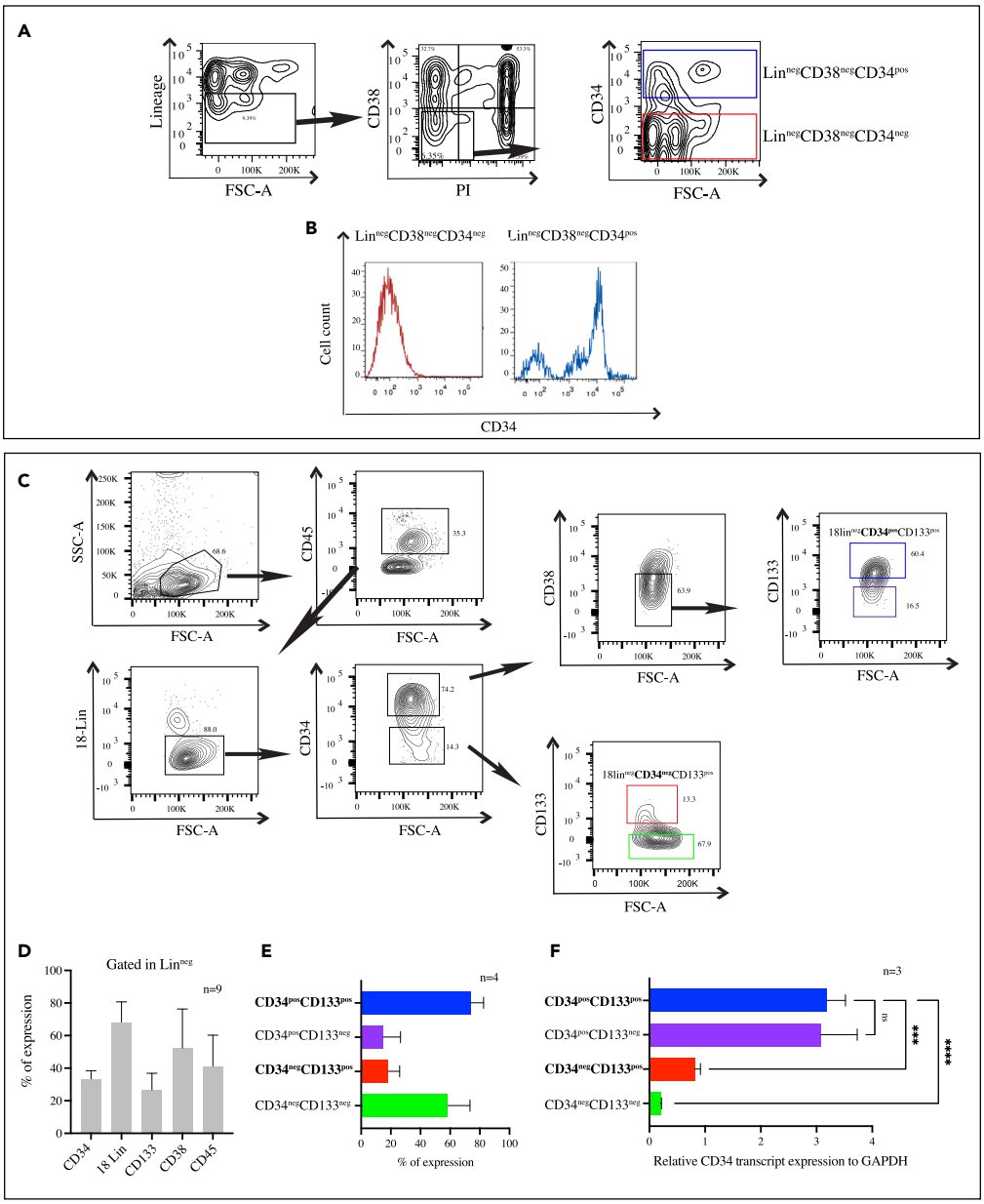


Figure 1. Purity of the CD34^{neg}HSC subsets isolated from human cord blood using two different methods

Overview of the gating strategy used to isolate HSCs. The lineage-depleted cells that were magnetically separated with cell surface markers against a cocktail of lineage markers (CD2, CD3, CD11b, CD14, CD15, CD16, CD19, CD56, CD123, and CD235a). The lineage-negative cells were then used for the two methods of HSC isolation.

(A) Layout of the gating strategy used to isolate Lin^{neg}**CD34^{neg}**CD38^{neg} and Lin^{neg}**CD34^{pos}**CD38^{neg} fractions by flow cytometric sorting (BD FACSAria III). The gate was set on the lymphocyte window for FSC/SSC. Left panel, cells in the lineage negative fraction were further analyzed for the expression of CD38 and PI viability staining (middle panel) to select double-negative gated. Following selection, these CD38 negative and PI negative cells were subdivided into two subpopulations based on CD34 expression: Lin^{neg}**CD34^{neg}**CD38^{neg} and Lin^{neg}**CD34^{pos}**CD38^{neg} (right panel). Data shown is representative of n = 5 independent experiments.

(B) The two populations from (A) were further stained with a CD34 antibody (clone: 581) and their purity was analyzed by flow cytometry.

(C) *Lin^{neg} fraction* stained with 18 lineage markers conjugated with FITC-(CD7, CD10, CD235a, CD33, CD20, CD24, CD4, CD66c, CD41, CD3, CD14, CD15, CD19, CD45RA, CD56, CD127, CD16 and CD2), and stained with BV510-CD45 and BV-421anti-CD34 (left panel). The cells were then subdivided into four sub-populations according to CD133 expression (right panel).

(D) The expression percentages for each gate in the 18-Lin^{neg} fraction analyzed with flow cytometry, n = 9.

Figure 1. *Continued*

(E) The percentages of the four sorted populations that that are either $CD34^{pos}CD133^{pos}$ and $CD34^{pos}CD133^{neg}$ in the **$CD34^{pos}$** subset, or $CD34^{neg}CD133^{pos}$ and $CD34^{neg}CD133^{neg}$ in the **$CD34^{neg}$** subset.

(F) RNA was isolated from the four sorted populations. SYBR green-based real-time qPCR was carried out using primers for CD34 and glyceraldehyde-3-phosphate (GAPDH). Results from qPCR were obtained from n = 3 independent experiments, ***p = 0.0002 and ****p < 0.0001, respectively using ordinary One-Way Anova.

than **$CD34^{pos}$** cells, resulted (Figure 2D). In contrast, CFU experiments utilizing Method 2 yielded differentiated colonies in both HSC cell populations, as seen in Figure 2I and described previously.[23]

Collectively, we determined that differences in sLe^x expression – namely the lack of it on **$CD34^{neg}$**HSCs – were likely the reason why **$CD34^{neg}$**HSCs struggle to home more than their positive counterparts do. Since we demonstrated that **$CD34^{neg}$** in Method 2 was superior to Method 1 in producing differentiated colonies *in vitro*, we opted to focus on Method 2 HSCs for all further analysis, apart from the *in-vivo* engraftment studies described later in discussion, which used both methods.

$CD34^{neg}$ cell type showed dramatic improvement in *in-vivo* regeneration potential after α1,3 fucosylation treatment

One of the ways of assessing how authentic a stem cell truly is involves seeing whether cell repopulation takes hold in immunodeficient mice once the cells are transplanted. To this end, we aimed to study the engraftment efficiency of human **$CD34^{neg}$**HSCs after sorting using both isolation methods. Given that both **$CD34^{neg}$**HSC populations lacked functional E-selectin ligands expressing sLe^x (Figures 2A and 2E), we used *ex vivo* α1,3 fucosylation to add the sugar necessary for sLe^x synthesis, as previously described.[24,25] There's been a growing research focus on improving the capacity of stem cells to bind to selectins via treatments that create sLe^x using recombinant human fucosyltransferase VI and VII (rhFT).[19,24,26–28]

Subsequently, we conducted short-term and long-term bone marrow engraftment assays using a xenotransplantation model that involved injecting human HSCs into immunodeficient NSG mice.[29,30] As illustrated in Figure S1, the detection limit for human cells ($humanCD45^{pos}$) in mouse bone marrow was around 0.1% of the total cells. We isolated **$CD34^{neg}$**HSCs from human cord blood samples using Method 1 to examine the α1,3 fucosylation effect. The cells were then either treated with the rhFTVI enzyme or the buffer control solution without the enzyme, and **$CD34^{pos}$** cells were used as a positive control (see Figure S2 schematic). At 8- and 12-week post-transplantation, the peripheral blood and bone marrow were collected and stained for the human CD45 antibody and evaluated by flow cytometry. Table S1 shows that Method 1 **$CD34^{neg}$**HSCs, untreated and treated with rhFTVI, did not engraft following IV injection and IF injection, a striking contrast to published work.[6] **$CD34^{pos}$** controls, by contrast, consistently showed robust levels of engraftment, confirming the performance of our xenograft model (Table S1).

Next, the engraftment analysis of Method 2 **$CD34^{neg}$**HSCs is summarized in Table 1. Although multiple attempts were made to optimize engraftment, such as increasing the number of cells transplanted, injecting cells directly into the interfemoral space, and using various methods of myeloablation (i.e., X-ray irradiation and busulfan treatment), human engraftment was rare and percentages of **$CD34^{neg}$**HSCs failed to exceed 0.2–1% in NSG mice. Moreover, α1,3 fucosylation treatment did not improve the engraftment, despite our best attempts to optimize the protocols. Again, confirming the utility of our model, engraftment occurred in all tissues when the positive control cells were used (4.5% in blood, 53% in bone marrow, and 10% in spleen) at 8 weeks post-transplantation, Figure S3, and at 16 weeks (10% in blood, 64% in bone marrow, and 51% in spleen).

In view of this, we decided to change the recipient mouse model to the NBSGW[31] strain (NOD.Cg-*Kit^{W-41J} Tyr^{+} $Prkdc^{scid}$ $Il2rg^{tm1Wjl}$*/ThomJ), which supports human hematopoietic stem cell engraftment without irradiation.[31] Figure 3A shows that rhFTVI enzyme α1,3 fucosylation enhanced HECA-452 reactivity in **$CD34^{neg}$**HSCs, indicating sLe^x structures were created. Method 2 was used to extract **$CD34^{neg}$**, which we intravenously implanted into non-irradiated NBSGW mice. As illustrated in Figure 3B, successful human engraftment was evident as early as 6 weeks post-transplantation for both untreated cells and fucosylated cells. At 9 weeks, rhFTVI-treated **$CD34^{neg}$** cells showed an increase in engraftment efficiency. At 12 weeks, the rhFTVI-treated group had a significant increase in engraftment (Figure 3B), indicating that α1,3 fucosylation gives **$CD34^{neg}$**HSCs long-term regenerative capacities. Representative flow cytometry analysis of donor cell engraftment in the recipient mice is illustrated in Figures 3C and Table 2. These findings demonstrate that transplanting **$CD34^{neg}$**HSCs into non-irradiated NBSGW mice with an intact bone marrow microenvironment is more effective than ablation.

Single cell RNA-seq suggests $CD34^{neg}$ hematopoietic stem cells have strong interactions with surrounding microenvironment

Multidimensional scRNA-seq analysis separates CD34-negative from CD34-positive HSCs in cord blood

The 10x Genomics Chromium platform[32] was used to evaluate the transcriptional heterogeneity of FACS-purified **$CD34^{pos}$** and **$CD34^{neg}$** subsets employing single-cell RNA sequencing (scRNA-seq). We generated a total of 5,063 single-cell RNA-seq profiles. Almost 2,400 **$CD34^{pos}$** cells and 2,300 **$CD34^{neg}$** cells were kept as high-quality libraries for further analysis. The number of genes and unique molecular identifiers (UMI) in each population are illustrated in Table S2.

Firstly, we analyzed the cellular heterogeneity using t-distributed stochastic neighbor embedding analysis (t-SNE) and uniform manifold approximation and projection (UMAP) of the entire multidimensional gene expression datasets for both HSC populations as shown in

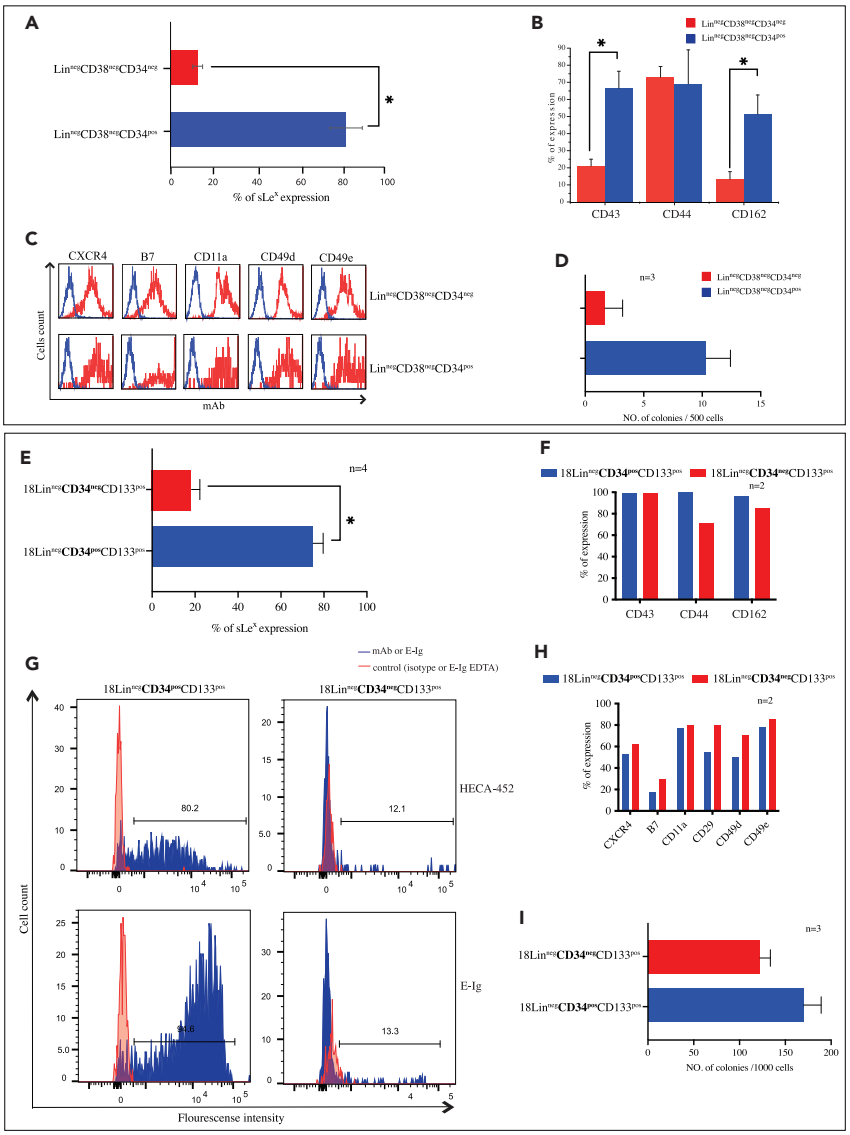


Figure 2. Characterization of homing molecules on CD34^{neg} HSCs obtained using the two Methods

(A) Lin^{neg}CD38^{neg}CD34^{neg} and Lin^{neg}CD38^{neg}CD34^{pos} cells obtained using method 1 were stained with HECA452-mAb and analyzed by flow cytometry to determine the expression of sLe$^{x/a}$ epitopes on their surface. These results are based on an average of n = 3 independent experiments. p-value is < 0.001.

(B) Flow cytometric analysis of E-selectin ligands (CD43, CD44, CD162) expressed on the two subsets of cells obtained by method 1. Results are expressed as an average percent of expression (above the isotype control) of n = 3 independent experiments. *p < 0.05 relative to CD34^{neg} subpopulation.

(C) Flow cytometric analysis of cell surface integrins and the stem cell chemokine receptor, on Lin^{neg}CD38^{neg} cord blood cells that either express CD34 (CD34^{pos}; lower panel) or not (CD34^{neg}; upper panel).

(D) Colony forming capacities of sorted Lin^{neg}CD38^{neg}CD34^{neg} and Lin^{neg}CD38^{neg}CD34^{pos}. Five hundred cells were cultured in methylcellouse at 37°C and 5% CO_2 in the presence of cytokines (SCF, IL-3, EPO and GM-CSF) for 12 to 14 days. The number of the colonies generated is shown.

(E) 18Lin^{neg}CD34^{neg}CD133^{pos} and 18Lin^{neg}CD34^{pos}CD133^{pos} obtained using method 2 were stained with HECA452-mAb and analyzed by flow cytometry to determine the expression of sLe^{x} epitopes on their surface, these results are based on an average of n = 3 independent experiments.

(F) Representative experiment of flow cytometry 18Lin^{neg}CD34^{neg}CD133^{pos} and 18Lin^{neg}CD34^{pos}CD133^{pos} cell subsets of sLe^{x} expression and E-selectin binding.

(G) Flow cytometric analysis of E-selectin ligands (CD43, CD44, CD162; upper panel) and (H) integrins and the chemokine receptor, expressed on the two subsets of cells obtained by method 2, 18Lin^{neg}CD34^{neg}CD133^{pos} and 18Lin^{neg}CD34^{pos}CD133^{pos}.

(I) Differences in colony-forming potential between 18Lin^{neg}CD34^{neg}CD133^{pos} and 18Lin^{neg}CD34^{pos}CD133^{pos} cells after sorting. This graphic displays the total number of colonies.

Table 1. Engraftment efficiency of 18Lin^{neg}CD34^{neg}CD133^{pos}HSCs isolated using Method 2 in NSG mice

Cells	Cells injected	Mode of injection	No. mice	Percentage human cells in NSG recipient mouse bone marrow	
				8 weeks	16 weeks
Method 2: Interfemoral injection (IF)					
18Lin^{neg}**CD34^{pos}**CD133^{pos}	5000	IF	9	0.7, 3.5, 7, 15.8, 44.6, 15.8, 41.7, 50 (9/9)	0.2, 27, 4, 1, 10, 24, 2, 60, 55 (9/9)
18Lin^{neg}**CD34^{neg}**CD133^{pos}	1000	IF	5	0, 0, 0, 0, 0.1 (1/5)	0, 0, 0, 0, 0.1 (1/5)
Method 2: Different cell doses					
18Lin^{neg}**CD34^{pos}**CD133^{pos}	20,000	IF	1	77 (1/1)	83 (1/1)
	30,000	IF	2	70, 82 (2/2)	80, 83 (2/2)
	60,000	IF	1	85.9 (1/1)	85 (1/1)
18Lin^{neg}**CD34^{neg}**CD133^{pos}	1,300	IF	1	0.2 (1/1)	0 (0/1)
	2,000	IF	2	0.1,0 (1/2)	0.1, 0 (1/2)
	4,000	IF	2	0.2, 0 (1/2)	0.1, 0 (1/2)
Method 2: α1,3 fucosylation treatment					
rhFTVI-18Lin^{neg}**CD34^{neg}**CD133^{pos}	5,000	IF	4	0, 0, 0.1, 2.5 (2/4)	0, 0, 0.1, 1 (2/4)
	5,000	IV	4	0, 0, 0.3, 0.2 (2/4)	0, 0, 0.3, 1 (2/4)
Buffered- 18Lin^{neg}**CD34^{neg}**CD133^{pos}	5,000	IF	4	0, 0, 0.2, 0.1(2/4)	0, 0, 0.1, 0.1(2/4)
	5,000	IV	4	0, 0, 0.2, 0.1 (2/4)	0, 0, 0.2, 0.2 (2/4)
18Lin^{neg}**CD34^{pos}**CD133^{pos}	5,000	IV	2	35, 30 (2/2)	65, 60 (2/2)

Percentages of human CD45 donor cells engrafted in the recipient mouse post-transplantation of human CB-HSCs treated with or without the rhFTVI.

Figure 4A. The two HSC populations are largely segregated from one another in the t-SNE map and UMAP. Interestingly, dimensional reduction analysis also showed that the **CD34^{neg}** cells clustered into two distinct populations (Figure 4A).

Since the purification of HSCs is currently limited to fluorescence-activated cell sorting (FACS), we next evaluated the feasibility of flow cytometric cell sorting to enrich human **CD34^{pos}** and **CD34^{neg}**HSCs in a side-by-side comparison with sequencing data. Both phenotypes were purified from cord blood on the FACSAria platform, and the phenotypic composition of the purified subsets was analyzed based on CD45, CD34 and CD133 gene expression (Figure 4B). CD34 and CD45 could be used to accurately enrich the two populations as CD34 was highly expressed in **CD34^{pos}** and absent (under expressed) in **CD34^{neg}**HSCs. However, less than 40% of **CD34^{neg}**HSCs expressed CD133 (Figure 4C), while more than 85% of the **CD34^{pos}** subset expressed CD133. In summary, FACS sorting using the **CD34^{neg}**HSC phenotype resulted in a higher reduction of target cells compared with **CD34^{pos}**, but the **CD34^{neg}** subset still contained a significant proportion of HSCs for further analysis.

Cord blood CD34^{neg} and CD34^{pos} hematopoietic stem cells have unique transcriptional signatures

Seven transcriptionally distinct groups of single cells were identified within the two HSC populations using the unsupervised graph-based clustering tool; the clusters were mapped onto UMAP visualization (Figure 5A). As illustrated in Figure S4, 19 clusters were generated if we include clusters that contained fewer than 100 cells. Clusters 2 and 5 included the majority of **CD34^{pos}** cells and were selected for further analysis, while clusters 3, 8, and 6 and 7 were primarily populated by **CD34^{neg}** cells. However, based on CD34 and CD133 gene expression (Figure 4B), we inferred that clusters 3 and 8 were enriched with **CD34^{neg}**HSCs, and we used these two clusters for further analysis.

The top 50 genes from clusters 2, 5, 3, and 8 were compiled into a single list to identify any genes common to either HSC subset (Table S3). Figure 5B shows that 18% of the genes were only expressed in clusters 2 and 5, which correspond to **CD34^{pos}**. These genes are related to primitive HSCs (MSI2, HOPX, AVP, IFITM3) suggesting that **CD34^{pos}** cells express higher levels of these primitive genes. Clusters 3 and 8 (corresponding to **CD34^{neg}**) expressed different sets of primitive genes (MSI2, ALDH1, ETS2, GATA2, PBX1), as well as adhesion-related genes (ITGA2B, IL1B, and THBS1), and genes involved with the megakaryocyte/erythroid lineage (KLF1, ANK1, HBD, SPTA1, TAL1). Finally, HTR1F, RPS3, RPS4X, SMIM24, MEIS1, and ETS2 were the only six genes expressed in both **CD34^{pos}** and **CD34^{neg}** clusters (Figures 5B; Table S3). Interestingly, gene expression patterns were clearly distinguishable between both FACS-purified HSCs, suggesting considerable enrichment of each HSC population; only three genes were commonly expressed between clusters. Overall, this top 50 most highly expressed gene analysis showed that **CD34^{neg}** cells have a distinct transcriptomic profile compared with **CD34^{pos}**.

We analyzed the two populations to determine the number of genes that are differentially expressed according to a 3-fold criterion (Table S4), and this resulted in ~116 transcripts that were upregulated in **CD34^{neg}** and ~33 that were upregulated in **CD34^{pos}**. The top 20

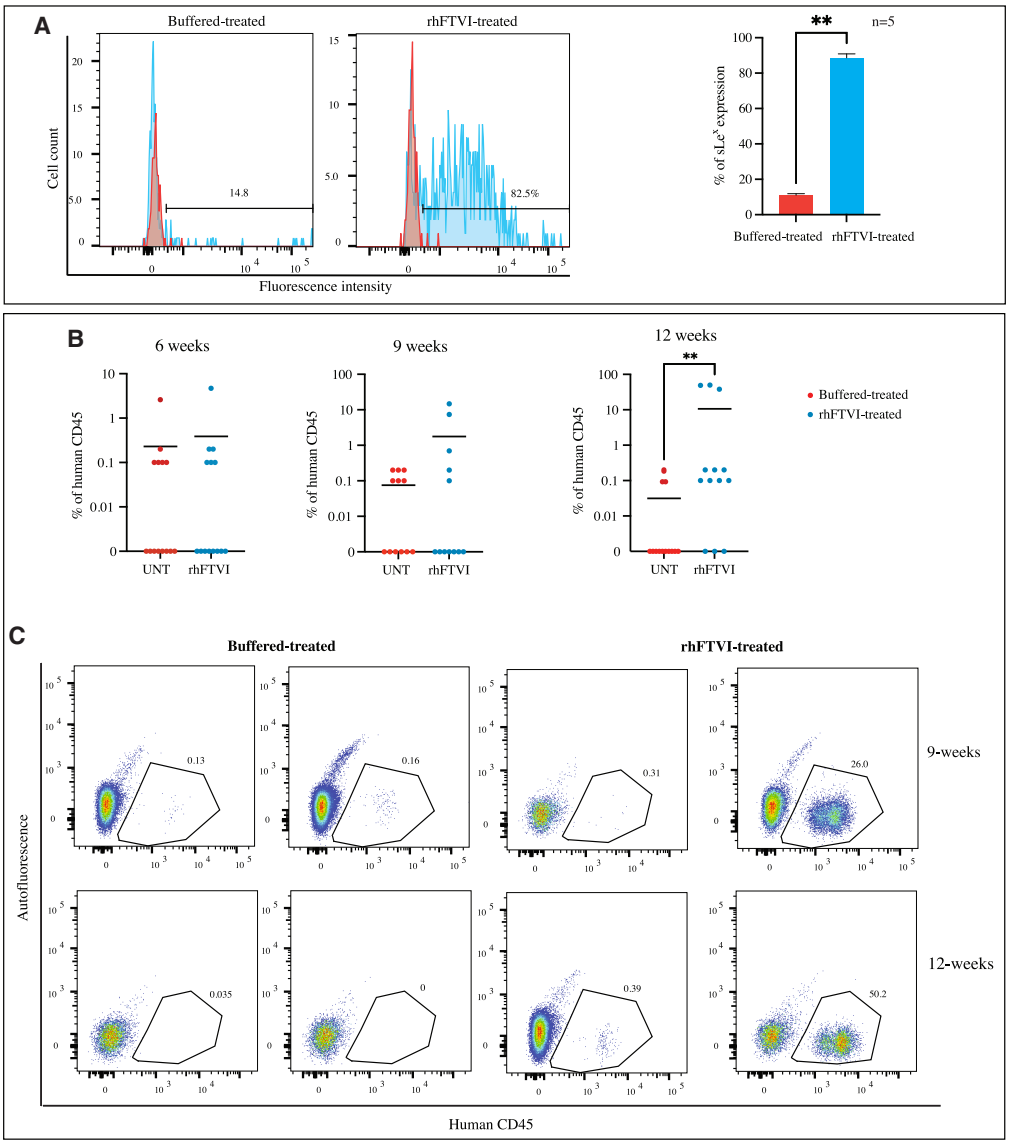


Figure 3. α1,3 fucosylation of 18Lin^{neg}CD34^{neg}CD133^{pos} HSCs improves engraftment in primary transplanted mice

(A) 18Lin^{neg}**CD34^{neg}**CD133^{pos} were eaither treated with rhFTVI buffer; blue or in buffer alone; red, and incubated for 30 min. Following treatment flow cytometric analysis analysis for sLe^{x} expression was determined (level panel) and the average percentage expression is represented in the (right panel). This is representative of n=5 independent experiements, **p = 0.002.

(B) Non irradiated NBSGW mice were transplanted with ~2000 18Lin^{neg}**CD34^{neg}**CD133^{pos} rhFTVI-treated (red) or buffer-treated (blue). Bone marrow from transplanted recipient mice was then investigated at the indicated short-term periods 6 and 9 weeks and longer-term periods at 12 weeks for the percentage of human donor cell contribution to total bone marrow cells. Each data point represents individual mouse (**p = 0.002).

(C) A representative analysis of human donor cell engraftment of two mice for both 18Lin^{neg}**CD34^{neg}**CD133^{pos} rhFTVI-treated (right panel), and two mice of buffer-treated 18Lin^{neg}**CD34^{neg}**CD133^{pos} HSCs (left panel).

differentially expressed genes in each population are illustrated in Figures 5C and 5D, and a heatmap of the same 20 genes by cluster is shown in Figure S5. Interestingly, genes involved in proliferation (SETBP1, MZB1, DUSP6) and lymphoid differentiation (IGHD, TRBC2, LTB) were enriched in **CD34^{pos}**HSCs, while those upregulated in **CD34^{neg}**HSCs are related to the development of RBCs and macrophages (KLF1, SLAMF7, IDO1, MPEG1, HBD, APOC1, ANK1, CD226, CD36), actin cytoskeleton reorganization, cell-cell contact and cell-matrix interaction (*GPR65, CLECL1, ITGA2B, THBS1, LGALS2*). The entire list of DEGs is shown in Table S4.

We next investigated whether genes that are highly differentially expressed between the HSC populations are transcription factors (TF), which are vital for gene expression regulation.[33,34] (Table 3). Key embryonic TFs, including GATA2, ETS1, and ERG,[35] were discovered to be highly expressed among these HSC populations. Genes such as MYC, MYB, TFDP1, STAT5A, STAT5B, TGF1, PBX1,[36–39] and KLF1, TAL1[40]

Table 2. Engraftment efficiency of 18Lin^{neg}CD34^{neg}CD133^{pos}HSCs isolated using Method 2 in NBSGW mice

	Cells	Cell Injection Dose	No. of mice	Percentage of human cells in recipient mice (w)		
				6	9	12
Exp. 1	Buffered-18Lin^{neg} CD34^{neg}CD133^{pos}	2000	6	0.1, 0.1, 2.5,0, 0, 0.1 (4/6)	0, 0.1, 0.2,0.2, 0, 0.2 (4/6)	0, 0, 0,0, 0.1, 0.2 (2/6)
	rhFTVI-18Lin^{neg} CD34^{neg}CD133^{pos}	2000	6	4.7, 0, 0.1, 0, 0.1, 0.3 (4/6)	0.1, 0, 0, 14.7, 0.7 (3/5)	0.1, 0.2, 0, 38, 0.2 (4/5)
Exp. 2	Buffered-18Lin^{neg} CD34^{neg}CD133^{pos}	1700	8	0.1, 0.1, 0, 0, 0, 0, 0, 0 (2/8)	0.2, 0.2, 0, 0, 0, 0, 0, 0 (2/8)	0.1, 0.1, 0.2, 0, 0, 0, 0, 0 (3/8)
	rhFTVI-18Lin^{neg} CD34^{neg}CD133^{pos}	1700	8	0.2, 0.2, 0.1, 0, 0.2, 0, 2, 0, 0 (4/8)	7.4, 0.2, 0.1, 0.1, 0, 0, 0, 0 (4/8)	0.1, 0.1, 0.1, 0.2, 49.8, 49, 0, 0 (6/8)

Percentages of human CD45 donor cells engrafted in the recipient mouse post-transplantation of human CB-HSCs treated with or without the rhFTVI.

were found to be overexpressed in **CD34^{neg}**HSCs, all important genes in HSC quiescence and self-renewal. Conversely, the upregulated transcription factors in **CD34^{pos}**HSCs were associated with cell proliferation (TCF7L2, TFEC, and BCL3) and differentiation (HOXA9, TCF4[41] and MAFF[42]) along the lymphoid and megakaryocyte/erythrocyte lineages. Our data showed that Runx3 was highly differentially expressed in **CD34^{pos}** versus **CD34^{neg}**HSCs, and that Runx1, Runx2, and Runx3 all play important roles in HSC maintenance during adult definitive hematopoiesis.[43]

Additionally, based on gene expression signature, pseudotime analysis suggests contrasting lineage differentiation potential among the HSC populations as shown in Figure S6A. ANK,[44] BLVRB, CD84[45] and CD41 were found to be significantly elevated in **CD34^{neg}** clusters, while KLF6 and Runx3 are elevated in **CD34^{pos}**. In addition, in order to identify which mature cells they would likely differentiate into based on the transcriptional patterns,[46,47] cell identification analysis showed that Megakaryocytes-Erythroid-Progenitor (MEP)[40,48] genes and erythroblast genes were enriched within **CD34^{neg}**HSCs, whereas Common-Myeloid-Progenitors (CMP) and natural killer progenitors appeared to be upregulated in CD34^{pos} (Figures S6B and S6C), Table S5.

CD34^{neg} hematopoietic stem cells are profoundly enriched with genes related to adhesion to the bone marrow niche through actin cytoskeleton and cell-extracellular matrix interaction pathways

To investigate which pathways are related to cell adhesion and migration, we performed gene enrichment analysis of the KEGG pathways.[49] We found that pathways for *cell adhesion molecules*, *hematopoietic cell lineage* and *leukocyte transendothelial migration* were abundant within both HSC populations (Figure 6A). Notably, pathways involved in adhesion and cell shape – such as *focal adhesion*, *regulation of actin cytoskeleton*, *adherence junction* and *extracellular matrix (ECM)-receptor interaction cell* – were upregulated in **CD34^{neg}**, while pathways related to cell differentiation and proliferation were highly expressed in **CD34^{pos}** (Figure 6A). These results suggest the crucial role that **CD34^{neg}**HSCs play regulating cell adhesion, specifically the upregulation of genes related to ECM-receptor interactions.

Table 4, Table S6 and Figures S7–S10, show distinct sets of genes in common pathways related to cellular adhesion and migration and hematopoietic cell lineage that were differentially expressed among the two HSCs. Of note, **CD34^{pos}**HSCs expressed higher levels of genes linked with lymphocyte marker. In contrast, **CD34^{neg}**HSCs expressed high levels of genes associated with erythro-myelo-megakaryocytic cells. Genes involved in adhesion were enriched in **CD34^{pos}** (CD99, CD44, CD34, SELL). Nonetheless, **CD34^{neg}** cells were shown to have a unique set of adhesion-related genes (ITGA2B, ITGB1, ALCAM, CD226). Interestingly, **CD34^{neg}** cells were enriched in genes related to actin cytoskeleton structure involved in focal adhesion or cell adherence junction pathways, and gene sets that dynamically regulate actin bundles were also upregulated. Next, Gene Ontology Enrichment Analysis (GOEA) was performed with biological process and cellular component categories (Figure 6B) to confirm that although the two populations shared some adhesion/migration pathways, they differed in gene sets and enrichment. For example, upregulated genes (*CDH1, ITGB1, CD84, PIK3CB, BSG,* and *CADM1*) were more prevalent in **CD34^{neg}**HSCs, leading to increased homophilic cell adhesion via plasma membrane adhesion molecules.

Second, using Gene Set Enrichment Analysis (GSEA), we demonstrated that gene sets related to the regulation of leukocytes to vascular endothelial cells and leukocyte migration were significantly higher in **CD34^{pos}**HSCs (Table S7). As expected, the gene sets related to extracellular matrix receptor and focal adhesion pathways showed significantly higher expression in **CD34^{neg}**HSCs. Then, we analyzed the ~65 actin-related pathways in the GSEA. As predicted, we found many pathways significantly enriched in **CD34^{neg}** (Table S7). These results raise the possibility that **CD34^{neg}**HSCs likely reside in the bone marrow niche due to a high concentration of adhesion-associated molecules related to ECM interaction that help them establish their niche's long-term repopulation activity.

Interestingly, cell-matrix adhesion, focal adhesion assembly and regulation of actin cytoskeleton are the pathways with the most genes upregulated in **CD34^{neg}**HSCs (Figure 7A). We confirmed these results at the protein level using confocal microscopy (Figure 7B). As shown in Figure 7C, vinculin (VCL) expression was considerably higher in **CD34^{neg}**HSCs compared to **CD34^{pos}** (n = 30 cells). Although β-actin transcripts appeared to be higher in **CD34^{neg}** (Figure 7D), no statistically significant difference was observed at the protein level (Figure 7C). Other actin cytoskeleton-related genes, including VASP, MYL12A, TLN1, and FLNA (Figure 7D), were also upregulated in **CD34^{neg}**HSCs. To

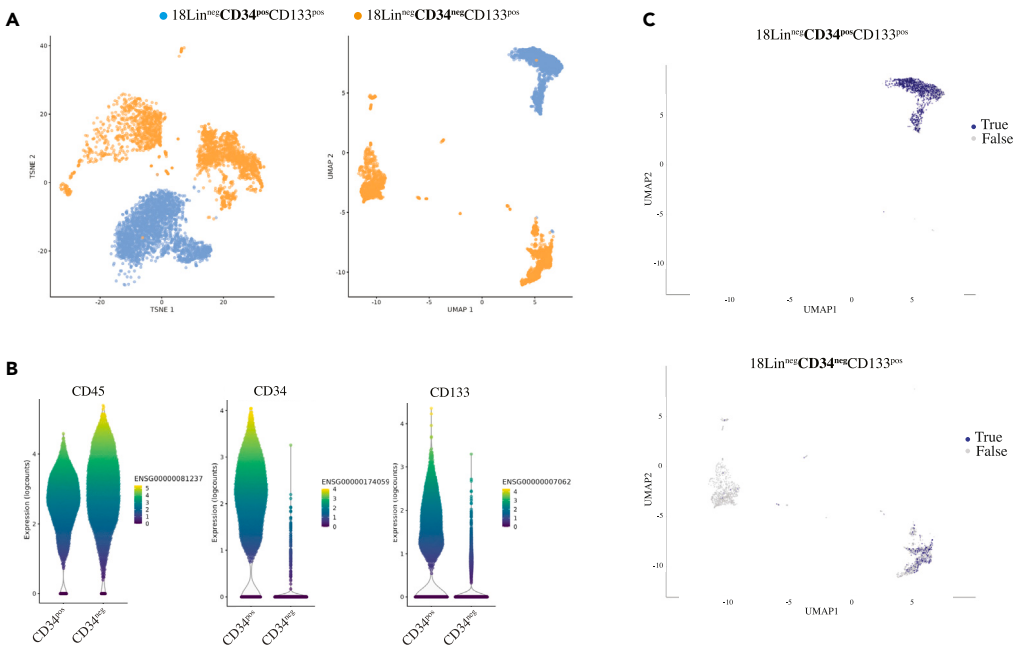


Figure 4. Dimensional reduction analysis revealed single cell CB-HSCs populations

(A) *t*-SNE (left) and UMAP (right) representation of the normalized gene expression values 18Lin^{neg}**CD34^{pos}**CD133^{pos} and 18Lin^{neg}**CD34^{neg}**CD133^{pos} HSCs. Each point in the plot represents a cell, dots of the same colors correspond to the same experimental group.

(B) Expression level of the CD45, CD34, and CD133 genes. These markers were used for sorting CB 18Lin^{neg}**CD34^{pos}**CD133^{pos} and 18Lin^{neg}**CD34^{neg}**CD133^{pos} populations.

(C) UMAP plot of the expression of CD34 and CD133 in 18Lin^{neg}**CD34^{pos}**CD133^{pos} HSCs (upper), and cells that show the negative expression of CD34 and highly expressed in CD133 in 18Lin^{neg}**CD34^{neg}**CD133^{pos} HSCs (lower).

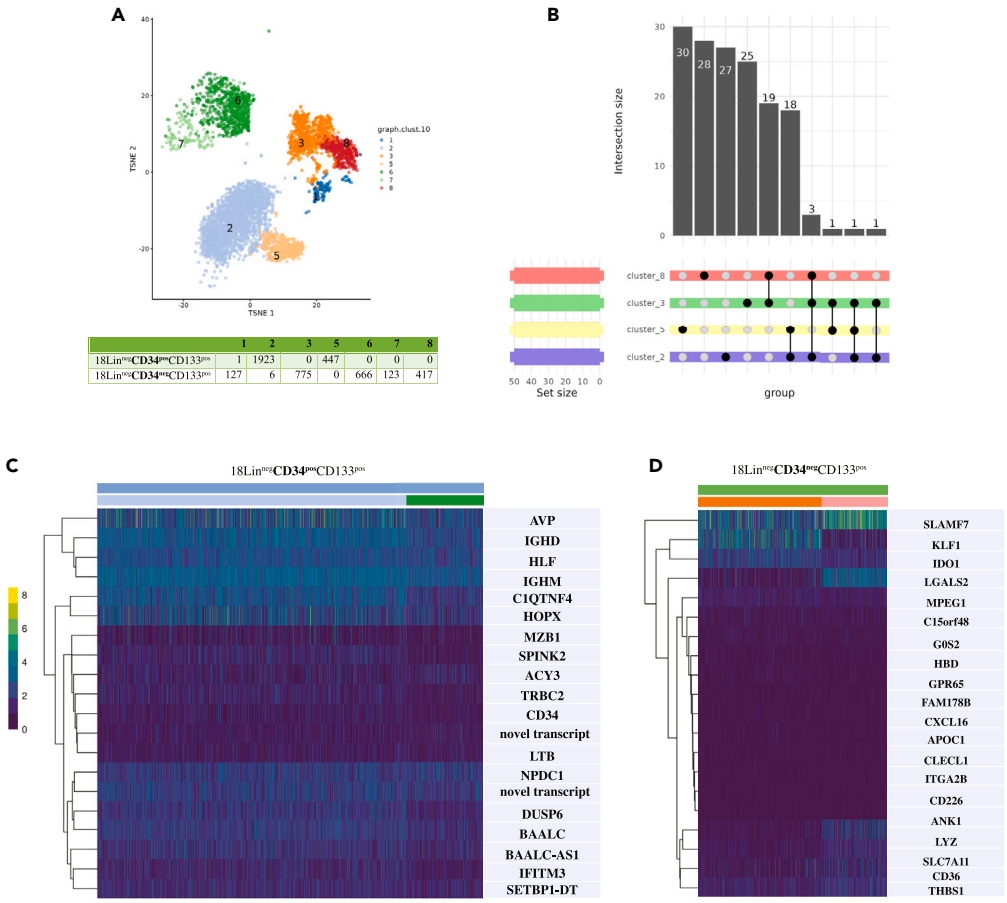
determine whether the upregulation of integrins in **CD34^{neg}**HSCs is related to adhesive interaction to ECM proteins, we used an adhesion assay to the ECM protein, fibronectin (Fn).[50] As illustrated in Figure 7E (*left panel*), **CD34^{neg}**HSCs have much higher levels of Fn binding compared to their **CD34^{pos}** counterparts. Moreover, to determine the contribution of β1α4 integrin in mediating Fn binding, **CD34^{neg}**HSCs were pretreated with either a blocking β1α4 integrin antibody or an isotype control (mouse IgG1) prior to the experiment on Fn binding. Figure 7E (*right panel*) shows that compared with cells blocked with the isotype control, cells incubated with β1α4 integrin lost their binding activity to Fn. These findings revealed that **CD34^{neg}**HSCs rely on β1α4 integrin to mediate their interaction with fibronectin.

DISCUSSION

The presence of short-term and long-term hematopoietic stem/progenitor cells (HSCs) in human cord blood and bone marrow has been highlighted in many studies.[4,6,7,11,14,23,51,52] Isolating these HSCs for the treatment of blood disorders has become a wide clinical practice using the CD34 surface protein marker. Clinical use of HSCs has tended to focus on **CD34^{pos}** cells, with **CD34^{neg}** cells often getting discarded because of their limited homing capacity when injected into the bloodstream. This has contributed to the underuse of umbilical cord blood in HSC transplantation because of the lower stem cell count compared with bone marrow. And yet, there is a body of research that has suggested that **CD34^{neg}** cells are more primitive stem cells,[4,11,53] meaning that they have the greatest potential for cell regeneration. By identifying ways to improve both homing and engraftment of **CD34^{neg}** cells through the addition of sialyl Lewis X, this study demonstrates the vast potential of using **CD34^{neg}**HSCs to treat blood disorders.

To test our theory, we began by applying two researched methods for isolating **CD34^{neg}**HSCs based on cell-surface markers, which allow for the sorting out of more primitive cell types from the billions of human hematopoietic mononuclear cells (MNCs) found in cord blood. The first method relied on CD34 and CD38 markers extracted from lineage-depleted MNCs,[6] while the second relied on a large number of lineage markers in addition to CD133 and CD34.[9] Our results mirrored published data.[6,9]

Our next step was to apply flow cytometry analysis, which revealed that **CD34^{neg}**HSCs enriched by either technique displayed much lower amounts of sLe^x and were less likely to bind E-selectin relative to **CD34^{pos}** cells. However, **CD34^{neg}** cells do express the chemokines and integrins necessary for bone marrow homing at levels comparable to **CD34^{pos}**. These findings highlight that the key reason **CD34^{neg}** cells have been less able to home when injected[9] is solely due to the absence of sLe^x, an epitope that can be easily added *ex vivo* to greatly increase the supply of stem cells reaching the bone marrow.[28,54–57] We found the two isolation methods differed significantly in how well **CD34^{neg}**HSCs differentiated to distinct colonies. Strong clonogenicity was demonstrated by Method 2, but Method 1 showed very little differentiation potential *in vitro*.



	1	2	3	5	6	7	8
18Lin^{neg}CD34^{pos}CD133^{pos}	1	1923	0	447	0	0	0
18Lin^{neg}CD34^{neg}CD133^{pos}	127	6	775	0	666	123	417

Figure 5. Different transcriptional signatures of CB-HSCs populations

(A) The seven distinct clusters identified by unsupervised clustering and characterized with differential gene expression and gene set enrichment analyses. Each dot represents a cell.

(B) Common transcripts expressed in four clusters mainly; 3, 8, 2, 5 with the schematic representation of intersections. Only the top 50 transcripts expressed in each cluster are represented.

(C) and (D) Heatmap of the most significant differentially expressed genes (DEGs) in CB-HSCs populations 18Lin^{neg}**CD34^{pos}**CD133^{pos} (cluster 2 and cluster 5) and 18Lin^{neg}**CD34^{neg}**CD133^{pos} HSCs (cluster 3 and cluster 8) ($p < 0.05$ and log fold change >0.5 in at least one cluster) in the different clusters revealed by DEG2 analysis.

One of our lines of experimentation was to use fucosyltransferase *ex vivo* treatment on **CD34^{neg}**HSCs. We then conducted *in vivo* experiments injecting the treated HSCs intravenously into NSG mice that had been subjected to myeloablation to destroy bone marrow cells. For the purposes of the *in vivo* studies, we isolated **CD34^{neg}**HSCs using both Method 1 and 2. In Method 1, both rhFTVI-treated and buffered cells failed to establish a human cellular presence in the bone marrow of mice 8 and 16 weeks after transplantation. We suspect that the failure to engraft could be because these **CD34^{neg}** cells need to be cultured – rather than simply isolated – on murine HESS-1 cells for several days to expand and differentiate the cells into **CD34^{pos}** before transplantation, as suggested by previous work.[8] It is important to note that we did detect the engraftment of positive controls using Method 1.

Turning to Method 2, we found that **CD34^{neg}**HSCs showed unremarkable human chimerism in NSG mice in comparison to positive cells, both after treatment with rhFTVI and in untreated cells. Human cell engraftment in both scenarios remained very low despite numerous attempts at altering the dose of transplanted cells, substituting X-ray for busulfan therapy, or injecting the cells intrafemorally. Rather, possible causes for the lack of engraftment could include that very high numbers of cells need to be injected for any notable level of repopulation to take hold in the *in-vivo* assay. In sum, it was challenging to replicate the *in-vivo* results demonstrated in several previously published reports[6,9] for **CD34^{neg}**HSCs using either approach in NSG mice. While **CD34^{neg}**HSCs may at first take longer to engraft in primary recipients than **CD34^{pos}**HSCs, their capacity is revealed more over time; they have been shown to be capable of repopulating blood cells for a longer period and maintain this capacity in serial transplants.[3]

Following the failure of engraftment in NSG mice, we turned next to a different mouse strain – NBSGW, known for successful HSC transplantation because it provides a proper microenvironment for supporting long-term self-renewal and differentiation of human HSCs, without

Table 3. Differentially expressed transcription factors in 18Lin^{neg}CD34^{neg}CD133^{pos} and 18Lin^{neg}CD34^{pos}CD133^{pos} HSCs

	Gene Name	Description	logFC
CD34^{pos}CD133^{pos}	HES1	hes family bHLH transcription factor 1	3.13
	ETS1	ETS proto-oncogene 1, transcription factor	3.13
	TCF7L2	transcription factor 7 like 2	2.26
	TCF4	transcription factor 4	1.78
	HOXA9	homeobox A9	1.78
	TFEC	transcription factor EC	1.63
	BCL3	BCL3 transcription coactivator	1.16
	ERG	ETS transcription factor ERG	1.13
	ZBTB8A	zinc finger and BTB domain containing 8A	1.12
	FOS	Fos proto-oncogene, AP-1 transcription factor subunit	1.05
	MAFF	MAF bZIP transcription factor F	0.96
	RUNX3	RUNX family transcription factor 3	0.82
	STAT1	signal transducer and activator of transcription 1	0.77
	SOX4	SRY-box transcription factor 4	0.56
	ELK3	ETS transcription factor ELK3	0.56
CD34^{neg}CD133^{pos}	KLF1	Kruppel like factor 1	12.70
	MYC	MYC proto-oncogene, bHLH transcription factor	3.03
	GATA2	GATA binding protein 2	3.10
	HLTF	helicase like transcription factor	1.87
	TAL1	TAL bHLH transcription factor 1, erythroid differentiation factor	1.72
	MYB	MYB proto-oncogene, transcription factor	1.67
	TADA3	transcriptional adaptor 3	1.59
	PHTF1	putative homeodomain transcription factor 1	1.26
	TFDP1	transcription factor Dp-1	1.19
	STAT5A	signal transducer and activator of transcription 5A	1.18
	STAT5B	signal transducer and activator of transcription 5B	1.14
	PBX1	PBX homeobox 1	1.65
	TGIF1	TGFB induced factor homeobox 1	1.17

the need for myeloablative conditioning.[31] This is important because myeloablative conditioning is known to be detrimental to the bone marrow architecture, while also raising the risk for hematological, gastrointestinal, and neurological adverse effects that could kill recipient animals.[58–60] For these experiments, we isolated **CD34^{neg}** using only Method 2 and observed that, when injected intravenously, notably better engraftment took place. Furthermore, *in vivo* analysis showed that the α1,3 fucosylation treatment of **CD34^{neg}**HSCs significantly increased their repopulation capacity in primary recipients up to 12 weeks after transplantation, whereas low engraftment was observed in untreated mice. We have previously shown the effectiveness of α1,3 fucosylation on the longevity of HSCs,[24] namely that fucosylated cells both regenerated the bone marrow of the initial recipient mouse and sustained and repopulated in a secondary recipient. The positive results in NBSGW mice demonstrate that an intact, not destroyed, bone marrow microenvironment better supports human HSC engraftment in the xenotransplantation model.

Another key area of focus was using transcriptional analysis with single-cell RNA sequencing to characterize the adhesion pathways of **CD34^{pos}** and **CD34^{neg}**HSCs and understand how the two cell types may be functionally different in their adhesion and migration capacities. According to initial analysis, **CD34^{neg}**HSCs displayed a unique expression profile for the most highly expressed genes. Moreover, when we analyzed the differentially expressed genes, we found high levels of Erythrocyte/Megakaryocyte genes on **CD34^{neg}**HSCs, supporting the notion that **CD34^{neg}** may play a role in the development of these precursors – for red blood cell formation and platelets, respectively – in

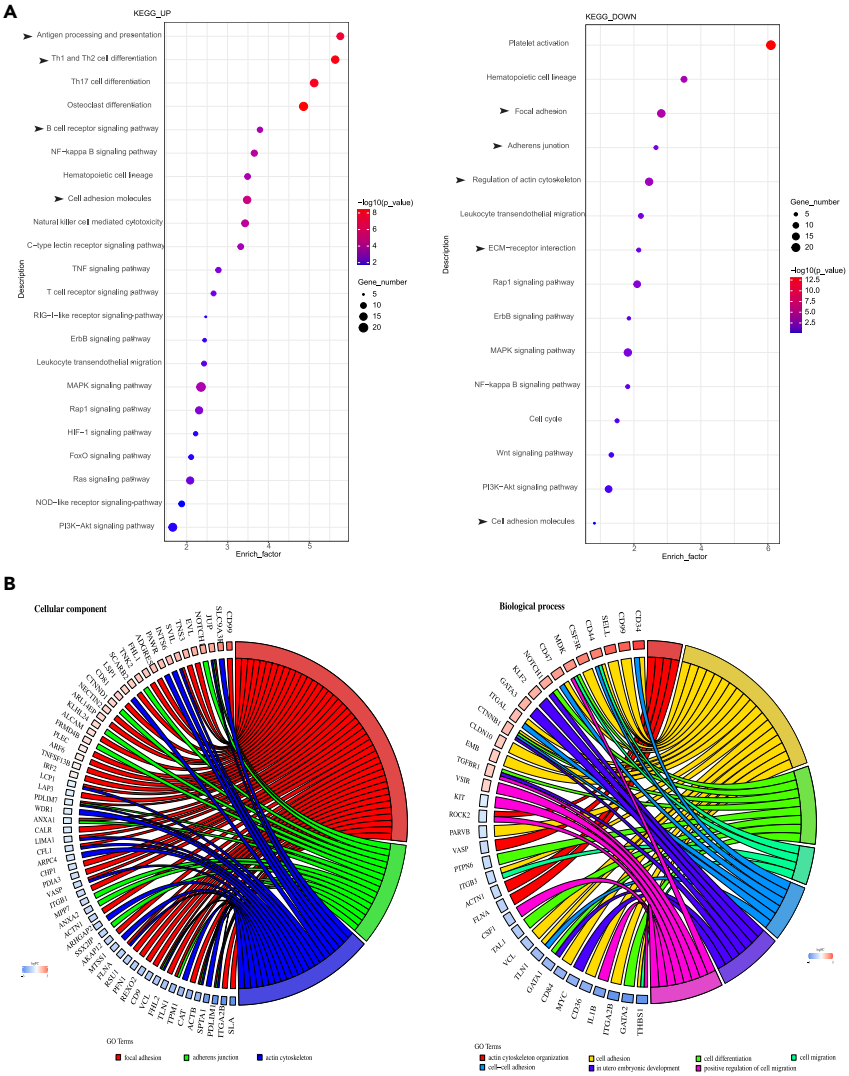


Figure 6. Gene enrichment analysis of differentially expressed genes in CB-HSCs populations

(A) Categorization of common genes were upregulated (left) or downregulated (right) in 18Lin^{neg}**CD34^{pos}**CD133^{pos} cells or based on KEGG pathways annotation.

(B) Cellular component and Biological processes represented by the differentially expressed genes.

the bone marrow, as previously reported from microarray data.[11] One observation was how they differed from another study that used Hoechst dye efflux to isolate **CD34^{neg}**HSCs and found higher levels of lymphoid lineage genes than we identified.[61] Going more deeply, our research revealed that transcription factors responsible for megakaryocyte and erythrocyte differentiation were upregulated in **CD34^{neg}**HSCs compared with **CD34^{pos}**HSCs, including KLF and GATA2 factors. GATA2 controls the proliferation and development of early hematopoietic precursors to red blood cells. Deleting it in mice has caused severe anemia and death.[62] Our pathway enrichment analysis also confirmed that erythrocyte/megakaryocyte genes are uniquely expressed in **CD34^{neg}**HSCs as they were absent in **CD34^{pos}**HSCs.

Now we will turn to analyzing the expression of well-known homing molecules on HSC populations: E-selectin ligands, MHC, cadherins and integrins. Our RNA sequencing analysis showed that CD44 (HCELL),[63–65] an E-selectin ligand found on human HSCs, neutrophils and T cells,[63] was upregulated in **CD34^{pos}**HSCs. L-selectin, which mediates the recruitment of naive T Cells to lymph nodes among other functions,[66] is highly upregulated in **CD34^{pos}**HSCs, although the role L-selectin plays on these cells remains unclear.[39] However, one study has shown that Sell$^{-/-}$ mice where L-selectin is absent display abnormality in lymphocyte migration.[67] For cell-cell contact that is integral for HSCs recognition and activation in the microenvironment, a lot of MHC classes were significantly expressed in **CD34^{pos}**HSCs versus **CD34^{neg}**HSCs, while cadherin, which mediates cell-cell adhesion, was under-expressed in both. **CD34^{neg}**HSCs, meanwhile, showed an increase in homophilic cell attachment through plasma membrane adhesion molecules (PIK3CB, CD84, NECTIN2, BSG), implying the potential for cell-cell communication.

Table 4. The gene sets that enriched 18Lin^{neg}CD34^{neg}CD133^{pos} and 18Lin^{neg}CD34^{pos}CD133^{pos} HSCs related to KEGG adhesion pathways

	GeneName	CD34^{pos}	CD34^{neg}	logFC	p-value	FDR	NCBI
Cell Adherence Junction	*TGFBR1*	1.11	0.86	0.37	9.42E-11	4.91E-10	7046
	NECTIN2	1.21	0.61	0.98	0.00E+00	0.00E+00	5819
	IGF1R	3.00	1.91	0.65	0.00E+00	0.00E+00	3480
	TCF7L2	2.43	0.51	2.26	0.00E+00	0.00E+00	6934
	TGFBR2	1.67	1.18	0.50	0.00E+00	0.00E+00	7048
	CTNND1	2.54	1.77	0.52	0.00E+00	0.00E+00	1500
	VCL	1.00	2.60	1.38	0.00E+00	0.00E+00	7414
	ACTN1	0.20	1.01	2.35	0.00E+00	0.00E+00	87
	ACTB	15.09	64.62	2.10	0.00E+00	0.00E+00	60
	PTPN6	0.78	1.95	1.32	0.00E+00	0.00E+00	5777
	SSX2IP	0.45	1.54	1.78	0.00E+00	0.00E+00	117178
Hematopoietic Cell Lineage	*CD37*	7.60	5.91	0.36	0.00E+00	0.00E+00	951
	CSF3R	2.20	0.28	2.96	0.00E+00	0.00E+00	1441
	CD34	4.24	0.13	4.98	0.00E+00	0.00E+00	947
	ITGA2B	0.01	3.07	7.73	0.00E+00	0.00E+00	3674
	CD9	0.19	1.24	2.74	0.00E+00	0.00E+00	928
	TFRC	2.06	5.16	1.32	0.00E+00	0.00E+00	7037
	IL1B	0.96	5.42	2.49	0.00E+00	0.00E+00	3553
	CD36	0.01	1.55	7.41	0.00E+00	0.00E+00	948
	CD55	2.04	5.33	1.38	0.00E+00	0.00E+00	1604
Leukocyte transendothelial migration	*CD99*	7.75	2.72	1.51	0.00E+00	0.00E+00	4267
	PIK3R1	4.12	1.89	1.13	0.00E+00	0.00E+00	5295
	PRKCB	1.61	0.83	0.97	0.00E+00	0.00E+00	5579
	RHOH	2.47	1.09	1.18	0.00E+00	0.00E+00	399
	CTNND1	2.54	1.77	0.52	0.00E+00	0.00E+00	1500
	VCL	1.00	2.60	1.38	0.00E+00	0.00E+00	7414
	PIK3CB	0.76	2.14	1.50	0.00E+00	0.00E+00	5291
	ACTN1	0.20	1.01	2.35	0.00E+00	0.00E+00	87
	ACTB	15.09	64.62	2.10	0.00E+00	0.00E+00	60
	MYL12A	4.08	13.56	1.73	0.00E+00	0.00E+00	10627
	VASP	1.36	3.44	1.34	0.00E+00	0.00E+00	7408
	ITGB1	1.80	4.74	1.39	0.00E+00	0.00E+00	3688
	RASSF5	0.78	1.82	1.22	0.00E+00	0.00E+00	83593
Cell Adhesion Molecules	*CD99*	7.75	2.72	1.51	0.00E+00	0.00E+00	4267
	CD34	4.24	0.13	4.98	0.00E+00	0.00E+00	947
	SELL	2.89	0.74	1.97	0.00E+00	0.00E+00	6402
	NECTIN2	1.21	0.61	0.98	0.00E+00	0.00E+00	5819
	CD40LG	0.01	1.63	6.94	0.00E+00	0.00E+00	959
	ITGB1	1.80	4.74	1.39	0.00E+00	0.00E+00	3688
	CD226	0.01	2.07	7.70	0.00E+00	0.00E+00	10666
	ALCAM	1.06	2.64	1.32	0.00E+00	0.00E+00	214

(*Continued on next page*)

Table 4. *Continued*

	GeneName	CD34^{pos}	CD34^{neg}	logFC	p-value	FDR	NCBI
Focal Adhesion	*BIRC2*	3.58	2.19	0.707	0.00E+00	0.00E+00	329
	AKT3	1.86	0.80	1.21	0.00E+00	0.00E+00	10000
	IGF1R	3.00	1.91	0.64	0.00E+00	0.00E+00	3480
	PIK3R1	4.11	1.88	1.12	0.00E+00	0.00E+00	5295
	PRKCB	1.61	0.82	0.96	0.00E+00	0.00E+00	5579
	ITGA2B	0.014	3.073	7.72	0.00E+00	0.00E+00	3674
	CFL1	9.024	20.040	1.15	0.00E+00	0.00E+00	1072
	VCL	1.00	2.60	1.37	0.00E+00	0.00E+00	7414
	ACTB	15.08	64.62	2.09	0.00E+00	0.00E+00	60
	THBS1	0.045	7.00	7.26	0.00E+00	0.00E+00	7057
	ITGB1	1.80	4.73	1.39	0.00E+00	0.00E+00	3688

An important finding of our RNA sequencing study was identifying highly expressed genes in **CD34^{neg}**HSCs related to their crucial role in interacting with the bone marrow microenvironment. Previous research on the bone marrow niche, while intensive, has revealed little about the factors that mediate the interaction between the HSCs and the surrounding environment consisting of proteins, cells, vessels and nerves that support the settlement and renewal of HSCs.[68,69] Our sequencing data findings indicate **CD34^{neg}**HSCs use integrins α5β1[70] to bind to ECM proteins such as fibronectin, laminin and collagen and, in addition, integrins α4β1[70] that bind with other adhesion molecules on the surface of niche cells including VCAM and ICAM.[50,71] A noteworthy finding was the markedly raised expression of integrin ITGA2B (CD41) in **CD34^{neg}**HSCs. It has previously been established that CD41 is expressed by quiescent HSCs and has a function in adult hematopoiesis; animals in which this integrin is absent (integrin II defective animals Itga2b−/−) have displayed pancytopenia, a malfunction of platelet formation, and enhanced apoptosis.[72] Moreover, CD41^{pos} HSCs possessed long-term repopulation capacity on serial transplantations and showed a marked myeloid bias compared with CD41^{neg} HSCs, which yielded a more lymphoid-biased progeny.[72] Another ECM interaction example in our study was the strong expression of CD36 in **CD34^{neg}**HSCs, which interacts with collagen in the bone marrow niche. A further finding was the upregulation in **CD34^{neg}**HSCs of cytokine interleukin 1 (IL-1), known for inducing the adhesion molecule VCAM1 in endothelial cells and the bone marrow stem cell niche.[73]

Our RNA sequencing also highlighted two important pathways in **CD34^{neg}**HSCs– actin cytoskeleton regulation and Cell-ECM – that play a role in determining how the HSCs will behave in the environment. For instance, we found that vinculin is upregulated in **CD34^{neg}**HSCs. Vinculin is a crucial component in regulating HSC repopulation capacity due to its role in long-term HSC reconstitution, independent of integrin functions.[61,74] In another example, we showed high transcript levels of thrombospondin-1 (THBS1) in **CD34^{neg}**HSCs, a homing molecule that helps maintain a stem cell's quiescence. When THBS1 is missing, studies have shown that cells acquire more self-renewal capacity.[75] Moreover, it has been demonstrated that phosphotyrosine phosphatase (PTP) has a role in HSC retention in the bone marrow niche.[61,76] Our results show enhanced expression of one of the PTPs, PTPN6, in **CD34^{neg}**HSCs. Moreover, the cytoskeletal protein talin is upregulated in **CD34^{neg}**HSCs, talin being shown to have a crucial role[61] in the HSCs' ability to adhere to the ECM.

Our study also involved performing *in-vitro* assays to offer additional experimental evidence correlating the transcript and protein levels. Our first finding was that vinculin protein expression was dramatically higher in **CD34^{neg}**HSCs compared to **CD34^{pos}**HSCs. Further, we proved that the formation of adhesive structures that link cells to the ECM first requires integrin activation[71] as α4β1. The cytoskeletal motility apparatus of HSCs could have a migratory function important from embryogenesis to adulthood, but this has yet to be investigated in greater detail.

Our present study makes two important contributions to understanding the biology of **CD34^{neg}**HSCs. We demonstrated that deficiency in **CD34^{neg}**HSC homing is caused by the lack of the sLe^{x} epitope which, when added *ex vivo*, successfully improves the HSC's homing and engraftment capacity. Second, we shed light on the differences between **CD34^{neg}** and **CD34^{pos}**HSCs by zooming in to the transcription level. Here, the biggest takeaway was that **CD34^{neg}**HSCs showed a differentiation bias toward erythropoiesis lineage cells, or RBCs, underscoring the potential diagnostic use of **CD34^{neg}**HSCs in RBC-related diseases. In addition, we offer insight into the **CD34^{neg}**HSC expression of adhesion molecules directly related to interaction with the bone marrow microenvironment. Given that **CD34^{neg}**HSCs are more primitive in the stem cell hierarchy than their positive counterparts, their potential for generating a long-term pool of new blood cells seems to depend on the existence of a vibrant niche environment, an important discovery for future research into how to exploit this underused stem cell, abundantly available in cord blood banks, more fully in transplantation.

STAR★METHODS

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

- KEY RESOURCES TABLE

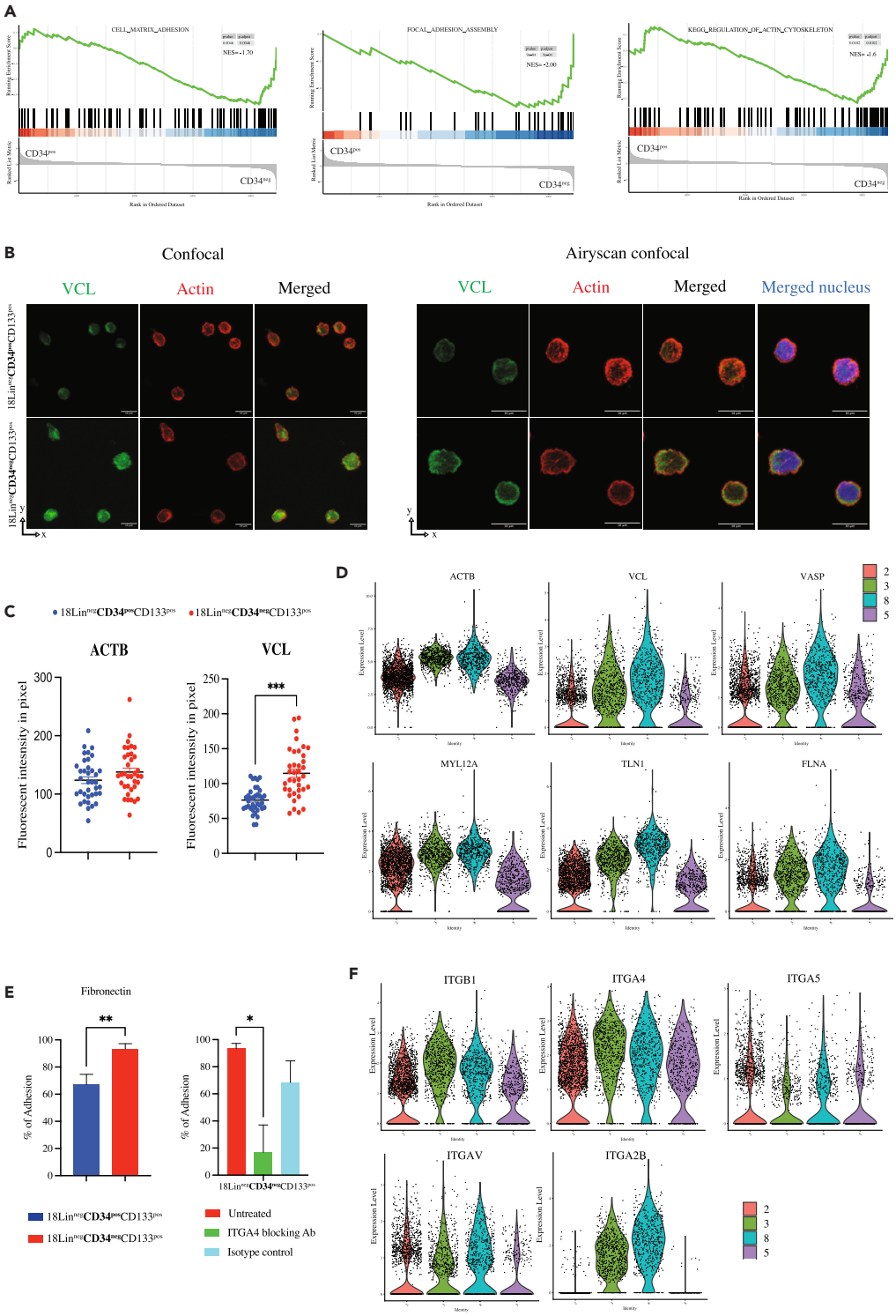


Figure 7. The enrichment of adhesion genes that correlated to actin cytoskeleton or cell-extracellular matrix in 18Lin^{neg}CD34^{neg}CD133^{pos} HSCs

(A) Gene sets were obtained from MSigDB database (Broad Institute) and then compared against the EPC and HSC signatures. Plots depict enriched gene sets in EPC. Black lines indicate the position of each gene across the ranked dataset. The normalized enrichment score (NES) is indicated in each plot.

(B) and (C) Fluorescence characterization of VLC and actin expression in CB-HSCs populations. 2D projection of top view of the 3D reconstructed fluorescence image of VLC (green), actin cytoskeleton (red), and nucleus (blue). The cells were fixed and immunolabeled for VLC, actin cytoskeleton, and nucleus using AF-488 dye-conjugated antibody, AF-647 dye-conjugated antibody, and DAPI, respectively, (***p < 0.001). Scale bar: 10 μm.

Figure 7. *Continued*

(D) Expression analysis of actin-related genes pathway. Expression counts for cluster 2 and 5 represent 18Lin^{neg}**CD34^{pos}**CD133^{pos} and cluster 3 and 8 represent 18Lin^{neg}**CD34^{neg}**CD133^{pos} for highly expressed genes in actin-cytoskeleton organization pathway.

(E) Adhesion capacity on fibronectin of CB-HSCs populations. Binding to fibronectin was determined for both 18Lin^{neg}**CD34^{pos}**CD133^{pos} and 18Lin^{neg}**CD34^{neg}**CD133^{pos} (left panel) and the contribution of 4 integrin in mediating Fn binding was determined by incubating cells with specific blocking antibodies or isotype control prior to the adhesion assay (n = 3), **p = 0.006 and *p = 0.017.

(F) Expression analysis of integrin genes. Expression counts for cluster 2 and 5 represent 18Lin^{neg}**CD34^{pos}**CD133^{pos} and cluster 3 and 8 represent 18Lin^{neg}**CD34^{neg}**CD133^{pos} for highly expressed genes in actin-cytoskeleton organization pathway.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108882.

ACKNOWLEDGMENTS

For research involving animals, the Animal Resource Core Lab (ARCL) at the King Abdullah University of Science and Technology (KAUST) was an invaluable resource, and Ms. Simona Spinelli and Mr. Stefano Pietro provided superb instruction and support. Recombinant human FTVI was developed in a collaboration with Dr. Jae Man Lee and Dr. Takahiro Kusakabe during a Competitive Research Grant (OCRF-2014-CRG3-2276) from KAUST. We appreciate Ms. Umm Habiba's help with lab administration. We thank Ms. Andrea Divlin from the Academic Writing Services at KAUST and Daliah Merzaban for editing the article. We thank Amani Ageeli for preparing the abstract image using Biorender. We also thank Dr. Dunia Jawdat and Dr. Walid Mashaqbeh from the Cord Blood Bank, King Abdullah International Medical Research Center (KAIMRC; Riyadh) for their insight and guidance in processing cord blood samples. The entire Cell Migration and Signaling Laboratory deserves recognition for their support and insightful comments.

Funding

This work was supported by a King Abdullah University of Science and Technology Faculty Baseline Research Funding Program to J.S.M.

AUTHOR CONTRIBUTIONS

A.S.A. designed, performed, and analyzed experiments and wrote the article. J.K performed bioinformatic analysis. Y.L. conducted *in vivo* mouse experiments. J.S.M.[1] performed adhesion assay work. A.A. performed the work related to confocal imaging and S.H. supervised this work. A.A-G. supported *in vivo* analysis of tissues. A.S-V. performed bioinformatic analysis. J.S.M.[2] conceived, designed, and analyzed this study and wrote the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 24, 2023
Revised: November 24, 2023
Accepted: January 8, 2024
Published: January 12, 2024

REFERENCES

1. Zhong, X.Y., Zhang, B., Asadollahi, R., Low, S.H., and Holzgreve, W. (2010). Umbilical cord blood stem cells: what to expect. Ann Ny Acad Sci *1205*, 17–22.
2. Jenq, R.R., and van den Brink, M.R.M. (2010). Allogeneic haematopoietic stem cell transplantation: individualized stem cell and immune therapy of cancer. Nat. Rev. Cancer *10*, 213–221.
3. Anjos-Afonso, F., Currie, E., Palmer, H.G., Foster, K.E., Taussig, D.C., and Bonnet, D. (2013). CD34(-) cells at the apex of the human hematopoietic stem cell hierarchy have distinctive cellular and molecular signatures. Cell Stem Cell *13*, 161–174.
4. Anjos-Afonso, F., Currie, E., Palmer, H.G., Foster, K.E., Taussig, D.C., and Bonnet, D. (2013). CD34(-) Cells at the Apex of the Human Hematopoietic Stem Cell Hierarchy Have Distinctive Cellular and Molecular Signatures. Cell Stem Cell *13*, 161–174.
5. Gao, Z., Fackler, M.J., Leung, W., Lumkul, R., Ramirez, M., Theobald, N., Malech, H.L., and Civin, C.I. (2001). Human CD34+ cell preparations contain over 100-fold greater NOD/SCID mouse engrafting capacity than do CD34- cell preparations. Exp. Hematol. *29*, 910–921.
6. Bhatia, M., Bonnet, D., Murdoch, B., Gan, O.I., and Dick, J.E. (1998). A newly discovered class of human hematopoietic cells with SCID-repopulating activity. Nat. Med. *4*, 1038–1045.
7. Zanjani, E.D., Almeida-Porada, G., Livingston, A.G., Flake, A.W., and Ogawa, M. (1998). Human bone marrow CD34- cells engraft in vivo and undergo multilineage expression that includes giving rise to CD34+ cells. Exp. Hematol. *26*, 353–360.
8. Nakamura, Y., Ando, K., Chargui, J., Kawada, H., Sato, T., Tsuji, T., Hotta, T., and Kato, S. (1999). Ex vivo generation of CD34(+) cells from CD34(-) hematopoietic cells. Blood *94*, 4053–4059.
9. Wang, J., Kimura, T., Asada, R., Harada, S., Yokota, S., Kawamoto, Y., Fujimura, Y., Tsuji, T., Ikehara, S., and Sonoda, Y. (2003). SCID-repopulating cell activity of human cord blood-derived CD34- cells assured by intra-bone marrow injection. Blood *101*, 2924–2931.
10. Sonoda, Y. (2008). Immunophenotype and functional characteristics of human primitive CD34-negative hematopoietic stem cells: the significance of the intra-bone marrow injection. J. Autoimmun. *30*, 136–144.
11. Sumide, K., Matsuoka, Y., Kawamura, H., Nakatsuka, R., Fujioka, T., Asano, H., Takihara, Y., and Sonoda, Y. (2018). A revised road map for the commitment of human cord blood CD34-negative hematopoietic stem cells. Nat. Commun. *9*, 2202.
12. Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., and Dick, J.E. (2011). Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science *333*, 218–221.
13. AbuSamra, D.B., Aleisa, F.A., Al-Amoodi, A.S., Jalal Ahmed, H.M., Chin, C.J., Abuelela, A.F., Bergam, P., Sougrat, R., and Merzaban, J.S. (2017). Not just a marker: CD34 on human hematopoietic stem/progenitor cells dominates vascular selectin binding along with CD44. Blood Adv. *1*, 2799–2816.
14. Gallacher, L., Murdoch, B., Wu, D.M., Karanu, F.N., Keeney, M., and Bhatia, M. (2000). Isolation and characterization of human CD34(-)Lin(-) and CD34(+)Lin(-) hematopoietic stem cells using cell surface markers AC133 and CD7. Blood *95*, 2813–2820.
15. Lin, G., Finger, E., and Gutierrez-Ramos, J.C. (1995). Expression of CD34 in endothelial cells, hematopoietic progenitors and nervous cells in fetal and adult mouse tissues. Eur. J. Immunol. *25*, 1508–1516.
16. Lange, C., Li, Z., Fang, L., Baum, C., and Fehse, B. (2007). CD34 modulates the trafficking behavior of hematopoietic cells in vivo. Stem Cell. Dev. *16*, 297–304.
17. Wang, J., Kimura, T., Asada, R., Harada, S., Yokota, S., Kawamoto, Y., Fujimura, Y., Tsuji, T., Ikehara, S., and Sonoda, Y. (2003). SCID-repopulating cell activity of human cord blood–derived CD34− cells assured by intra–bone marrow injection. Blood, The Journal of the American Society of Hematology *101*, 2924–2931.
18. Abuelela, A.F., Sakashita, K., and Merzaban, J.S. (2014). Cell surface enzymatic engineering-based approaches to improve cellular therapies. In Micro-and Nanoengineering of the Cell Surface (Elsevier), pp. 175–213.
19. Merzaban, J.S., Burdick, M.M., Gadhoum, S.Z., Dagia, N.M., Chu, J.T., Fuhlbrigge, R.C., and Sackstein, R. (2011). Analysis of glycoprotein E-selectin ligands on human and mouse marrow cells enriched for hematopoietic stem/progenitor cells. Blood *118*, 1774–1783.
20. Lapidot, T., and Kollet, O. (2002). The essential roles of the chemokine SDF-1 and its receptor CXCR4 in human stem cell homing and repopulation of transplanted immune-deficient NOD/SCID and NOD/SCID/B2m(null) mice. Leukemia *16*, 1992–2003.
21. Moll, N.M., and Ransohoff, R.M. (2010). CXCL12 and CXCR4 in bone marrow physiology. Expert Rev. Hematol. *3*, 315–322.
22. Peled, A., Kollet, O., Ponomaryov, T., Petit, I., Franitza, S., Grabovsky, V., Slav, M.M., Nagler, A., Lider, O., Alon, R., et al. (2000). The chemokine SDF-1 activates the integrins LFA-1, VLA-4, and VLA-5 on immature human CD34(+) cells: role in transendothelial/stromal migration and engraftment of NOD/SCID mice. Blood *95*, 3289–3296.
23. Abe, T., Matsuoka, Y., Nagao, Y., Sonoda, Y., and Hanazono, Y. (2017). CD34-negative hematopoietic stem cells show distinct expression profiles of homing molecules that limit engraftment in mice and sheep. Int. J. Hematol. *106*, 631–637.
24. Al-Amoodi, A.S., Li, Y., Al-Ghuneim, A., Allehaibi, H., Isaioglou, I., Esau, L.E., AbuSamra, D.B., and Merzaban, J.S. (2022). Refining the migration and engraftment of short-term and long-term HSCs by enhancing homing-specific adhesion mechanisms. Blood Adv. *6*, 4373–4391.
25. Sackstein, R. (2012). Engineering cellular trafficking via glycosyltransferase-programmed stereosubstitution. Ann. N. Y. Acad. Sci. *1253*, 193–200.
26. Xia, L., McDaniel, J.M., Yago, T., Doeden, A., and McEver, R.P. (2004). Surface fucosylation of human cord blood cells augments binding to P-selectin and E-selectin and enhances engraftment in bone marrow. Blood *104*, 3091–3096.
27. Popat, U., Mehta, R.S., Rezvani, K., Fox, P., Kondo, K., Marin, D., McNiece, I., Oran, B., Hosing, C., Olson, A., et al. (2015). Enforced fucosylation of cord blood hematopoietic cells accelerates neutrophil and platelet engraftment after transplantation. Blood *125*, 2885–2892.
28. Sackstein, R., Merzaban, J.S., Cain, D.W., Dagia, N.M., Spencer, J.A., Lin, C.P., and Wohlgemuth, R. (2008). Ex vivo glycan engineering of CD44 programs human multipotent mesenchymal stromal cell trafficking to bone. Nat. Med. *14*, 181–187.
29. Larochelle, A., Vormoor, J., Hanenberg, H., Wang, J.C., Bhatia, M., Lapidot, T., Moritz, T., Murdoch, B., Xiao, X.L., Kato, I., et al. (1996). Identification of primitive human hematopoietic cells capable of repopulating NOD/SCID mouse bone marrow: implications for gene therapy. Nat. Med. *2*, 1329–1337.
30. Guenechea, G., Gan, O.I., Dorrell, C., and Dick, J.E. (2001). Distinct classes of human stem cells that differ in proliferative and self-renewal potential. Nat. Immunol. *2*, 75–82.
31. Waskow, C., Madan, V., Bartels, S., Costa, C., Blasig, R., and Rodewald, H.R. (2009). Hematopoietic stem cell transplantation without irradiation. Nat. Methods *6*, 267–269.
32. Povinelli, B.J., Rodriguez-Meira, A., and Mead, A.J. (2018). Single cell analysis of normal and leukemic hematopoiesis. Mol. Aspect. Med. *59*, 85–94.
33. Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631–644.
34. Aggarwal, R., Lu, J., Pompili, V.J., and Das, H. (2012). Hematopoietic stem cells: transcriptional regulation, ex vivo expansion and clinical application. Curr. Mol. Med. *12*, 34–49.

35. Taoudi, S., Bee, T., Hilton, A., Knezevic, K., Scott, J., Willson, T.A., Collin, C., Thomas, T., Voss, A.K., Kile, B.T., et al. (2011). ERG dependence distinguishes developmental control of hematopoietic stem cell maintenance from hematopoietic specification. Genes Dev. *25*, 251–262.
36. Wilson, A., Murphy, M.J., Oskarsson, T., Kaloulis, K., Bettess, M.D., Oser, G.M., Pasche, A.C., Knabenhans, C., MacDonald, H.R., and Trumpp, A. (2004). c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. Gene Dev. *18*, 2747–2763.
37. Sheng, Y., Ma, R., Yu, C., Wu, Q., Zhang, S., Paulsen, K., Zhang, J., Ni, H., Huang, Y., Zheng, Y., and Qian, Z. (2021). Role of c-Myc haploinsufficiency in the maintenance of HSCs in mice. Blood *137*, 610–623.
38. Maurer, B., Kollmann, S., Pickem, J., Hoelbl-Kovacic, A., and Sexl, V. (2019). STAT5A and STAT5B-Twins with Different Personalities in Hematopoiesis and Leukemia. Cancers *11*, 1726.
39. Ivetic, A., Hoskins Green, H.L., and Hart, S.J. (2019). L-selectin: A Major Regulator of Leukocyte Adhesion, Migration and Signaling. Front. Immunol. *10*, 1068.
40. Kang, Y., Kim, Y.W., Yun, J., Shin, J., and Kim, A. (2015). KLF1 stabilizes GATA-1 and TAL1 occupancy in the human beta-globin locus. Biochim. Biophys. Acta *1849*, 282–289.
41. (2019). TCF4 promotes erythroid developmentF.E.M. In 't Hout, J. van Duren, D. Monteferrario, E. Brinkhuis, N. Mariani, T.M. Westers, D. Chitu, G. Nikoloski, A.A. van de Loosdrecht, and B.A. van der Reijden, et al., eds. *69*, 17–21.e11.
42. Shivdasani, R.A. (2001). Molecular and transcriptional regulation of megakaryocyte differentiation. Stem Cell. *19*, 397–407.
43. de Bruijn, M., and Dzierzak, E. (2017). Runx transcription factors in the development and function of the definitive hematopoietic system. Blood *129*, 2061–2069.
44. Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., et al. (2015). Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell *163*, 1663–1677.
45. Dumon, S., Walton, D.S., Volpe, G., Wilson, N., Dassé, E., Del Pozzo, W., Landry, J.R., Turner, B., O'Neill, L.P., Göttgens, B., and Frampton, J. (2012). Regulation at the Onset of Definitive Hematopoiesis and Commitment to Differentiation. PLoS One *7*, e43300. ARTN e43300.
46. Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D.B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S., et al. (2014). Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. Cell Stem Cell *15*, 507–522.
47. Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A.M., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell *121*, 295–306.
48. Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., et al. (2016). Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell *164*, 325.
49. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29.
50. Domingues, M.J., Cao, H., Heazlewood, S.Y., Cao, B., and Nilsson, S.K. (2017). Niche Extracellular Matrix Components and Their Influence on HSC. J. Cell. Biochem. *118*, 1984–1993.
51. Takahashi, M., Matsuoka, Y., Sumide, K., Nakatsuka, R., Fujioka, T., Kohno, H., Sasaki, Y., Matsui, K., Asano, H., Kaneko, K., and Sonoda, Y. (2014). CD133 is a positive marker for a distinct class of primitive human cord blood-derived CD34-negative hematopoietic stem cells. Leukemia *28*, 1308–1315.
52. Okuno, Y., Iwasaki, H., Huettner, C.S., Radomska, H.S., Gonzalez, D.A., Tenen, D.G., and Akashi, K. (2002). Differential regulation of the human and murine CD34 genes in hematopoietic stem cells. Proc. Natl. Acad. Sci. USA *99*, 6246–6251.
53. Tajer, P., Pike-Overzet, K., Arias, S., Havenga, M., and Staal, F.J.T. (2019). Ex vivo expansion of hematopoietic stem cells for therapeutic purposes: lessons from development and the niche. Cells *8*, 169.
54. Sackstein, R. (2009). Glycosyltransferase-programmed stereosubstitution (GPS) to create HCELL: engineering a roadmap for cell migration. Immunol. Rev. *230*, 51–74.
55. Sackstein, R. (2012). Re:"Ex vivo fucosylation improves human cord blood engraftment in NOD-SCID IL-2Rγnull mice". Exp. Hematol. *40*, 518–520.
56. Sackstein, R. (2018). The first step in adoptive cell immunotherapeutics: assuring cell delivery via glycoengineering. Front. Immunol. *9*, 3084.
57. Merzaban, J.S., Imitola, J., Starossom, S.C., Zhu, B., Wang, Y., Lee, J., Ali, A.J., Olah, M., Abuelela, A.F., Khoury, S.J., and Sackstein, R. (2015). Cell surface glycan engineering of neural stem cells augments neurotropism and improves recovery in a murine model of multiple sclerosis. Glycobiology *25*, 1392–1409.
58. Geissler, E.N., Mcfarland, E.C., and Russell, E.S. (1981). Analysis of Pleiotropism at the Dominant White-Spotting (W) Locus of the House Mouse - a Description of 10 New W-Alleles. Genetics *97*, 337–361.
59. Shultz, L.D., Lyons, B.L., Burzenski, L.M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S.D., King, M., Mangada, J., et al. (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. J. Immunol. *174*, 6477–6489.
60. McIntosh, B.E., Brown, M.E., Duffin, B.M., Maufort, J.P., Vereide, D.T., Slukvin, I.I., and Thomson, J.A. (2015). Nonirradiated NOD,B6.SCID Il2rgamma-/- Kit(W41/W41) (NBSGW) mice support multilineage engraftment of human hematopoietic cells. Stem Cell Rep. *4*, 171–180.
61. Ohmori, T., Kashiwakura, Y., Ishiwata, A., Madoiwa, S., Mimuro, J., Furukawa, Y., and Sakata, Y. (2010). Vinculin Is Indispensable for Repopulation by Hematopoietic Stem Cells, Independent of Integrin Function. J. Biol. Chem. *285*, 31763–31773.
62. Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature *371*, 221–226.
63. Ali, A.J., Abuelela, A.F., and Merzaban, J.S. (2017). An Analysis of Trafficking Receptors Shows that CD44 and P-Selectin Glycoprotein Ligand-1 Collectively Control the Migration of Activated Human T-Cells. Front. Immunol. *8*, 492.
64. AbuSamra, D.B., Al-Kilani, A., Hamdan, S.M., Sakashita, K., Gadhoum, S.Z., and Merzaban, J.S. (2015). Quantitative Characterization of E-selectin Interaction with Native CD44 and P-selectin Glycoprotein Ligand-1 (PSGL-1) Using a Real Time Immunoprecipitation-based Binding Assay. J. Biol. Chem. *290*, 21213–21230.
65. Katayama, Y., Hidalgo, A., Chang, J., Peired, A., and Frenette, P.S. (2005). CD44 is a physiological E-selectin ligand on neutrophils. J. Exp. Med. *201*, 1183–1189.
66. Grailer, J.J., Kodera, M., and Steeber, D.A. (2009). L-selectin: role in regulating homeostasis and cutaneous inflammation. J. Dermatol. Sci. *56*, 141–147.
67. Arbonés, M.L., Ord, D.C., Ley, K., Ratech, H., Maynard-Curry, C., Otten, G., Capon, D.J., and Tedder, T.F. (1994). Lymphocyte homing and leukocyte rolling and migration are impaired in L-selectin-deficient mice. Immunity *1*, 247–260.
68. Crane, G.M., Jeffery, E., and Morrison, S.J. (2017). Adult haematopoietic stem cell niches. Nat. Rev. Immunol. *17*, 573–590.
69. Domingues, M.J., Cao, H., Heazlewood, S.Y., Cao, B., and Nilsson, S.K. (2017). Niche extracellular matrix components and their influence on HSC. J. Cell. Biochem. *118*, 1984–1993.
70. Grassinger, J., Haylock, D.N., Storan, M.J., Haines, G.O., Williams, B., Whitty, G.A., Vinson, A.R., Be, C.L., Li, S., Sørensen, E.S., et al. (2009). Thrombin-cleaved osteopontin regulates hemopoietic stem and progenitor cell functions through interactions with alpha9beta1 and alpha4beta1 integrins. Blood *114*, 49–59.
71. Lee-Thedieck, C., and Spatz, J.P. (2014). Biophysical regulation of hematopoietic stem cells. Biomater. Sci. *2*, 1548–1561.
72. Gekas, C., and Graf, T. (2013). CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. Blood *121*, 4463–4472.
73. Cook-Mills, J.M., Marchese, M.E., and Abdala-Valencia, H. (2011). Vascular cell adhesion molecule-1 expression and signaling during disease: regulation by reactive oxygen species and antioxidants. Antioxidants Redox Signal. *15*, 1607–1638.
74. Humphries, J.D., Wang, P., Streuli, C., Geiger, B., Humphries, M.J., and Ballestrem, C. (2007). Vinculin controls focal adhesion formation by direct interactions with talin and actin. J. Cell Biol. *179*, 1043–1057.
75. Kaur, S., and Roberts, D.D. (2016). Divergent modulation of normal and neoplastic stem cells by thrombospondin-1 and CD47 signaling. Int. J. Biochem. Cell Biol. *81*, 184–194.
76. Ni, F., Yu, W.M., Wang, X., Fay, M.E., Young, K.M., Qiu, Y., Lam, W.A., Sulchek, T.A., Cheng, T., Scadden, D.T., and Qu, C.K. (2019). Ptpn21 Controls Hematopoietic Stem Cell Homeostasis and Biomechanics. Cell Stem Cell *24*, 608–620.e6.
77. Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve

quantification accuracy. Genome Res. *27*, 491–499.

78. Wingett, S.W., and Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. F1000Res. *7*, 1338.
79. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
80. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.
81. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.
82. Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. *36*, 421–427.
83. McCarthy, D.J., Campbell, K.R., Lun, A.T.L., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics *33*, 1179–1186.
84. Germain, P.L., Lun, A., Garcia Meixide, C., Macnair, W., and Robinson, M.D. (2021). Doublet identification in single-cell sequencing data using scDblFinder. F1000Res. *10*, 979.
85. Miao, Z., Deng, K., Wang, X., and Zhang, X. (2018). DEsingle for detecting three types of differential expression in single-cell RNA-seq data. Bioinformatics *34*, 3223–3224.
86. Al-Amoodi, A.S., Sakashita, K., Ali, A.J., Zhou, R., Lee, J.M., Tehseen, M., Li, M., Belmonte, J.C.I., Kusakabe, T., and Merzaban, J.S. (2020). Using Eukaryotic Expression Systems to Generate Human alpha1,3-Fucosyltransferases That Effectively Create Selectin-Binding Glycans on Stem Cells. Biochemistry *59*, 3757–3771.
87. Huygen, S., Giet, O., Artisien, V., Di Stefano, I., Beguin, Y., and Gothot, A. (2002). Adhesion of synchronized human hematopoietic progenitor cells to fibronectin and vascular cell adhesion molecule-1 fluctuates reversibly during cell cycle transit in ex vivo culture. Blood *100*, 2744–2752.
88. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.
89. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.
90. Wang, T., Li, B., Nelson, C.E., and Nabavi, S. (2019). Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. BMC Bioinf. *20*, 40.
91. Mabbott, N.A., Baillie, J.K., Brown, H., Freeman, T.C., and Hume, D.A. (2013). An expression atlas of human primary cells: inference of gene function from coexpression networks. BMC Genom. *14*, 632.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CD7	Iotest	A07755
CD10	Iotest	IM2720U
CD235a	Merck	FCMAB217f
CD33	invitrogen	11-0339-42
CD20	Iotest	IM1455U
CD24	Sigma	SAB4700624; RRID:AB_10897982
CD4	Miltenyi Biotec	130-080-501; RRID:AB_244326
CD66c	Beckman	IM2039U
CD41	Iotest	IM0649U
CD3	Miltenyi Biotec	130-080-401; RRID:AB_244231
CD14	Miltenyi Biotec	130-081-701
CD15	Miltenyi Biotec	130-081-101; RRID:AB_244217
CD19	Biorad	MCA6058f
CD45RA	southern Biotech	9630–02
CD56	Biorad	MCA2693f
CD127	invitrogen	11-1278-42
CD16	Biorad	MCA2537T
CD2	Dako	F0767
CD34	BD	555824; RRID:AB_398614
CD133	Miltenyi Biotec	130-113-106; RRID:AB_2725935
CD38	Miltenyi Biotec	130-113-427; RRID:AB_2733813
CD45RA	BD	564552; RRID:AB_2738841
CD90	BD	562556; RRID:AB_2737651
CD49f	BD	740793; RRID:AB_2740456
PE- streptavidin	BD	554061; RRID:AB_10053328
Anti-human Vinculin primary conjugated AF-488	Abcam	ab196454
Rhodamine Phalloidin	Thermo Scientific	R415
PE-conjugated anti-human IgG Fc	BioLegend	366904; RRID:AB_2876689
Biological samples		
Cord blood units	KAIMRC	Different donors
Chemicals, peptides and recombinant proteins		
Human serum albumin	Sigma-Aldrich	A9080
Fetal bovine serum (FBS)	Gibco (Invitrogen)	16000044
Ficol-paque	GE Healthcare	17-1440-02
Hank's Balanced Salt Solution (HBSS)	Gibco (Invitrogen)	14170-088
7-AAD	BD	559925
MethoCult H4433	Stem Cell Technologies	04435
HEPES	Gibco (Invitrogen)	15630080
GDP-fucose	Sigma	G4401
Triton X-100	Thermo Fisher Scientific	BP151-100

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bovine Serum Albumin (BSA)	ThermoFisher	A7979
Prolong DAPI mounting medium	Thermo Fisher Scientific	P36931
Phosphate-buffered saline (DPBS)	Gibco (Invitrogen)	14190144
1640-RPMI media	Gibco (Invitrogen)	21875091
Versene dissociation buffer	Gibco (Invitrogen)	15040033
Single Cell 3′ GEM, Library and Gel Bead Kit V3	10× Genomics	PN-1000075
DynaBeads® MyOne Silane Beads	Thermo Fischer Scientific	37002D
Fucosylatransferase VI	KAUST	Lab-made
Deposited data		
scRNA-seq	NCBI	GSE237832
Experimental models: Organisms/strains		
NSG	Charles River	614
NBSGW	JAX	026622
Software and algorithms		
UMI-tools v1.1.2	(Smith T. et al., 2021)[77]	github.com/CGATOxford/UMI-tools
FASTQC v0.11.9	(Andrews S. et al., 2020)[78]	github.com/s-andrews/FastQC
Trimmomatic v0.39	(Bolger A. M. et al., 2021)[79]	usadellab.org/cms/?page=trimmomatic
STAR v2.7.9a	(Dobin A. et al., 2021)[80]	github.com/alexdobin/STAR
FeaureCounts v2.0.3	(Liao Y. et al., 2014)[81]	subread.sourceforge.net/
Scran v1.22.1	(Lun A. T. L. et al., 2021)[82]	bioconductor.org/packages/3.14/bioc/html/scran.html
Scater v1.22.0	(McCarthy D. J. et al., 2021)[83]	bioconductor.org/packages/3.14/bioc/html/scater.html
ScDblFinder v1.8.0	(Germain P. et al., 2021)[84]	bioconductor.org/packages/3.14/bioc/html/scDblFinder.html
DEsingle v1.14.0	(Miao Z. et al., 2018)[85]	bioconductor.org/packages/3.14/bioc/html/DEsingle.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jasmeen Merzaban (jasmeen.merzaban@kaust.edu.sa).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The accession numbers for the datasets are listed in the key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethic statement

All experimental procedures were conducted in accordance with the Guide for the Care and Use of Laboratory Animals (NIH publication no.85-23, revised 1996), Implementing Regulations of the Law of Ethics of Research on Living Creatures (KSA National Committee of BioEthics - Third Edition) and were conducted under the authority of the King Abdullah University of Science and Technology (KAUST) Institutional Animal Care and Use Committee (IACUC protocol number: 17IACUC20) KAUST is AAALAC International accredited institution. Cord blood (CB) units were purchased from the Cord Blood Bank at King Abdullah International Medical Research Center (KAIMRC), Riyadh, Saudi Arabia.Animals.

Animals

Male and female mice were acquired from Charles River Labs and were bred and maintained in isolator cages with autoclaved food and water under specific pathogen-free conditions within the KAUST Animal Resource Core Lab. NOD.Cg-Prkdc^{scid} Il2rgt^{m1Wjl}/SzJ (NSG) mice aged 6–12 weeks were irradiated with a single dose of 2.7 Gy utilizing an Xstrahl Cabinet Irradiator 12 hours prior to transplantation. NOD.Cg-Kit^{W-41J} Tyr^{+}Prkdc^{scid}Il2rg^{tm1Wjl}/ThomJ recipient mice were also utilized in several experiments (NBSGW).[58–60]

Cells

Cord blood (CB) units were purchased from King Abdullah International Medical Research Center's (KAIMRC) Cord Blood Bank. CB buffy coat-enriched cells were thawed in MEM-α medium containing 10% ACD-A (anticoagulant citrate dextrose solution) and 5% human serum albumin (Sigma), centrifugated at 300g, resuspended in 10 mL PBS containing 0.1 mg/mL DNAse, incubated at room temperature for 10 minutes, centrifuged, washed twice with MEM medium containing 10% ACD-A, resuspended in 30 mL of MEM-α culture medium with 5% FBS and incubated at 37°C for 12 to 24 hours in a hypoxic chamber (Stem cells Technologies) containing water in a petri dish. The following day, the cells were rinsed with PBS/ACD-A, exposed to a density gradient using Ficol-paque (GE Healthcare) in LeucoSep™ centrifuge tubes (Thomas Scientific) for 12 minutes at room temperature with brake-off, then the mononuclear layer was transferred to a separate tube, washed twice with PBS/ACD, filtered using a BD Falcon 70 μm cell strainer to obtain single cells, which were enumerated, stained, sorted, and subjected to flow cytometric analysis.

METHODS DETAILS

HSC enrichment

The initial enrichment phase was identical even though different sorting strategies were used with varied combinations of surface markers. As a means of negative selection, we used a lineage cell depletion cocktail (Miltenyi Biotec) including ten lineage markers (CD2, CD3, CD11b, CD14, CD15, CD16, CD19, CD56, CD123, and CD235a (Glycophorin A)) for lineage depletion. Briefly, cells were stained for 10 minutes with Biotin-Antibody Cocktail and rinsed, biotin microbeads were added for 15 minutes at 4°C. Using the depletion software on an AutoMacs pro separator, we divided the cells into a lineage negative fraction and a lineage positive fraction. PE- streptavidin was used to verify the depletion. The first approach involves staining the lineage-depleted fraction with 10 μg/mL FITC-CD38 and BV421-CD34 for 30 minutes, washed twice in FACS buffer (5% FBS, 2 mM EDTA, and HBSS) and stined with PI. After sorting, the purity of CD38^{neg}CD34^{pos} and CD38^{neg}CD34^{pos} cells was evaluated with CD34 antibody. The second technique involved staining the lineage-depleted fraction with FITC-conjugated lineage-specific monoclonal antibodies (CD7, CD10, CD235a, CD33, CD20, CD24, CD4, CD66c, CD41, CD3, CD14, CD15, CD19, CD45RA, CD56, CD127, CD16, and CD2), Pacific blue anti-CD34 and APC anti-CD133 for 30 minutes at 4C° (all antibodies are listed in Table S1). Prior to sorting on BD FACS Aria III, 7-AAD was added to the tubes, and the samples were washed twice using FACS buffer (5% FBS, 2 mM EDTA, and HBSS). The specificity of these lineage-specific mAbs utilized for cell sorting was verified in advance with fluorescence minus one. Human cord blood stem cells that were sorted into the 18Lin^{neg}CD34^{pos}CD133^{pos} and 18Lin^{neg}CD34^{neg}CD133^{pos} populations were employed in these studies.

Flow cytometric analysis

All flow cytometry was performed on a FACSCanto II platform or BD FACS Aria III. The data were analyzed using FlowJo software (BD) and the positive percentages were compared to fluorescence minus one (FMO) controls. To determine the expression of sLe^{x} on the surface of the cells, either rhFTVI-treated or buffer treated (negative control) HSC populations were placed in 96-well-plates, stained with 10 μg/mL anti-sLe^{x} antibody for 30 minutes at 4°C (HECA-452), resuspended in FACS buffer (10-mM EDTA, 5% FBS and HBSS) and washed twice with FACS buffer (200 μL/well). To detect E-selectin binding, 10 μg/mL of recombinant E-selectin human Ig (E-Ig) chimera was prepared in buffer (20 mM HEPES pH 7.5, 2 mM $CaCl_2$ and 5% FBS) and used to stain cells. A PE-conjugated anti-human IgG Fc (1:200 dilution in chimera buffer; BioLegend) secondary antibody was used to detect E-Ig. As a control for E-selectin staining, 20-mM EDTA was added to the chimera buffer. To detect cell surface proteins, the HSCs populations were stained with 10 μg/mL of primary conjugated antibodies against surface markers CXCR4, CD49e, CD49d, and CD29 in 100 μL FACS buffer (2 mM EDTA, 5% FBS and HBSS) for 30 minutes at 4°C, washed three times with FACS buffer, and analyzing surface-marker expression.

Bone marrow engraftment analysis

Six to 12-week-old female recipient mice (NSG) were irradiated with 2.7 Gy using an Xstrahl Cabinet Irradiator 12 hours before transplantation. Where indicated, human stem cell populations were treated with rhFTVI or buffer control (untreated), resuspended in HBSS buffer and 200 uL of suspension cells injected into recipient mice via tail vein (i.v.) or through intrafemoral injection (*i.f.*). The weight of transplanted mice was measured weekly. Following transplantation at 6, 8 and 12 weeks, 200 uL of blood collected from submandibular and ~30 uL bone marrow harvested using sterile U-100 syringes needle. The samples were collected in tubes containing RPMI medium with 10% FBS, centrifuged and stained for flow cytometry analysis. Engraftment efficiency was assessed by flow cytometry using mouse CD45 [30-F12—APC] and human CD45 [HI30—Pacific blue] antibodies, and expressed as a percentage of the donor cells, hematopoietic cells, and their produced blood cells present in the bone marrow and blood of the recipients.

Clonogenic progenitor cell assay

The colony forming unit assay was conducted for HSCs populations in MethoCult H4433 medium (Stem Cell Technologies). Briefly, ~1,500 cells were pelleted in 1 L of MethoCult, seeded into 35-mm low adherent culture dishes (Stem Cell Technologies), and incubated at 37°C in a 5% CO_2 humidified chamber for 14–21 days. Total colonies were counted under a brightfield microscope using a 20X objective lens and visually differentiated as erythroid burst-forming units (BFU-Es), granulocyte-macrophage colony forming units (CFU-GMs), megakaryocyte colony forming units (CFU-Ms), and granulocyte-erythroid-megakaryocyte-macrophage colony forming units (CFU-GEMMs).

Exofucosylation

The fucosylation treatment was performed as described previously.[86] Briefly, 18Lin^{neg}**CD34**^{neg} CD133^{pos} **HSCs** were harvested, washed twice with Hank's Balanced Salt Solution (HBSS), and resuspended at a density of in FTVI reaction buffer [25 mM HEPES (pH 7.5) (Gibco Invitrogen), 0.1% human serum albumin (Sigma-Aldrich), 0.5 mM GDP-fucose (Sigma), and 5 mM $MnCl_2$] and 1 ug purified rhFTVI enzyme in HBSS. Cells were incubated at 37°C for 30 min. Buffer only controls without the rhFTVI enzyme were used as a negative control. After the reaction, the cells were washed twice with HBSS and 10 mM EDTA and used immediately for experiments.

Confocal imaging

Freshly isolated CB 18Lin^{neg}**CD34**^{pos} CD133^{pos} and 18Lin^{neg}**CD34**^{neg} CD133^{pos} HSCs were fixed in 4% PFA for 15 minutes at RT, washed with PBS for 5 minutes, permeabilized in 0.1% Triton X-100 in PBS for 5 minutes at RT. Next, cells were blocked using 1% BSA (ThermoFisher) for 50 minutes at RT and immunolabeled with 200 µl of 10 µg/mL mouse anti-human Vinculin primary conjugated AF-488 antibody (Abcam) in 1% BSA (5 mg/mL) for 1 hour at RT. After washing the unbound antibodies, cells were immunolabeled with 200 µl of 10 µg/mL Rhodamine Phalloidin antibody (ThermoFisher) in 1% BSA (5 mg/mL) for 40 minutes at RT, washed with PBS three times between each step and centrifuged at 350 g for 5 minutes in a 15 ml canonical tube. The immunolabeled cells were suspended in Prolong DAPI mounting medium, then seeded on a poly-L-lysin coated glass coverslip surface and imaged directly. Confocal microscopy imaging was performed using an inverted spinning disk microscope (Zeiss LSM 880), equipped with Plan-Apochrmat 63x/1.4 Oil DIC M27 objective lens. 3D (1 µm z-step size) fluorescence images of the stained cells were captured. The frame size per field of view is kept at 225µm^2, and the pixel size equals 70 nm. Two color fluorescence images were acquired sequentially. 488 laser power set at 6% and detector gain set at 850 to image the Vinculin, while 561 laser power set at 4% and detector gain set at 800 to image the actin cytoskeleton. The pinhole was set at 1 AU for both channels. Fluorescence images were acquired using the ZEN 2009 software platform (Zeiss). The image acquisition parameters were kept consistent throughout the experiment. To measure the expression level of Vinculin and actin cytoskeleton expressed in 18Lin^{neg}**CD34**^{pos} CD133^{pos} and 18Lin^{neg}**CD34**^{neg} CD133^{pos} HSCs, we captured fluorescence images of cells through multiple z-planes to construct 3D images. Using ImageJ, 3D fluorescence images of cells were stacked at sum intensity projection and the fluorescence integrated intensity were measured per cell and calculated the mean intensity per pixel after subtracting the background signal.

Adhesion assay

Adhesion assays were performed using CB 18Lin^{neg}**CD34**^{pos} CD133^{pos} **and** 18Lin^{neg}**CD34**^{neg} CD133^{pos} **HSCs** either on fibronectin and control BSA coated plates based on protocols established by Huygen.[87] BSA control plate was prepared by coating a flat-bottom 96-well plate (Corning) in 1% bovine serum albumin (BSA; Gibco, Life Technologies) prepared in 1X phosphate-buffered saline (DPBS; Gibco, Life Technologies) followed by overnight incubation in 4-degrees. The following day, the BSA plate as well as a commercial pre-coated fibronectin 96-well plate (Corning) were blocked with 1%-BSA prepared in serum-free 1X-1640-RPMI media (Gibco, Life Technologies) for 30 minutes in 37C° followed by two washes with serum-free RPMI. Next, 4,000 cells were added in triplicate fibronectin wells for three conditions of untreated, integrin-blocked (human CD49D antibody), and isotype-blocked (mouse IgG2b antibody; BD Biosciences) samples. After a 2 hours 37C° incubation, suspension cells were collected by aspiration and two washes in 1%-BSA-PBS was performed. Adhesion cells were collected by incubated versene dissociation buffer (Gibco, Life Technologies) for 5 minutes at 37C° followed by vigorous pipetting. Cell numbers were counted by Countess3 Automated Cell Counter (Invitrogen) and percentage of adhesion was calculated as [(adherent cells)/(adherent cells + suspension cells)*100].

RNA extraction, library preparation and RNA-sequencing

18Lin^{neg}**CD34**^{pos}CD133^{pos} and 18Lin^{neg}**CD34**^{neg}CD133^{pos} populations were isolated by FACS and the numbers of viable cells were counted using Trypan blue. Single-cell RNA-Seq libraries were prepared using the Single Cell 3′ GEM, Library and Gel Bead Kit V3 (10× Genomics, Cat#PN-1000075) according to the manufacturer's instructions. Briefly, single cells were partitioned into individual gel beads-in-emulsion (GEMs) and the RNA obtained from lysed cells was barcoded through reverse transcription. Each cell is encapsulated in a gel bead that contains a unique 14- base pair (bp) molecular barcode, a 10-bp randomer to index molecules (unique molecular identifier, UMI), and an anchored 30-bp oligo-dT to prime polyadenylated RNA transcripts. DynaBeads® MyOne Silane Beads (Thermo Fischer Scientific, Cat# 37002D) were used to purify the resulting barcoded cDNA, which was subsequently amplified via PCR (12–14 cycles, depending on the quantity of cDNA available). Libraries were then checked and quantified using Agilent 2100 Expert Software. The libraries were sequenced on a NovaSeq 6000 SP flow cell (Illumina).

QUANTIFICATION AND STATISTICAL ANALYSIS

scRNA-seq data preprocessing

The Cell Ranger pipeline was used for initial processing of the sequencing data. For Cell barcoding and UMI extraction, Chromium single cell protocols were used to generate R1 reads with the cell barcode and UMI and R2 reads with the mRNA sequence. Using UMItools (https://github.com/CGATOxford/UMI-tools), true cell barcodes and UMIs were moved from the read sequence of R1 FASTQ file to the read name of R2 FASTQ file, which contains the gene sequence and was used in the rest of the analysis. A quality check was performed on the R2 FASTQ files. To remove adapters and low-quality bases from the reads, a trimming step was performed using the Trimmomatic software (http://www.usadellab.org/cms/?page=trimmomatic) using the following parameters: minimum read length was set to 35 bp and the quality score to 25. Then, Spliced Transcripts Alignment to a Reference (STAR) software (https://github.com/alexdobin/STAR) was used to align the reads to the human reference genome (GRCh38).

Gene expression quantification and cell quality control

Software feature Counts (http://subread.sourceforge.net/) was used to assign the high-quality mapped reads to the genomic features (i.e., genes) of the official annotation of the human reference genome. UMItools was used to perform PCR deduplication by collapsing the UMIs per gene and per cell. Quality control was based on three metrics: total UMIs (de-duplicated reads), number of detected genes, and percentage of mitochondrial reads detected. Three different thresholds were calculated for each metric: 1, visually chosen; 2, automatically selected threshold based on the median absolute deviation (MAD) from the median value of each metric across all cells; and 3, automatically selected threshold based on the median absolute deviation (MAD) from the median value of each initial sample independently (https://bioconductor.org/packages/3.14/bioc/html/scater.html). Cells were removed if they were discarded based on at least one metric threshold.

Normalization, cell clustering and doublet detection

Normalization was applied by calculating the library size as the total sum of counts across all genes for each cell. Counts were first pooled from many cells to increase the size of the counts for accurate size factor estimation. Pool-based size factors were then deconvolved into cell-based factors for normalization of each cell's expression profile. Next, gene counts were divided by the cell size factor and log transformed using the R packages scran (https://bioconductor.org/packages/3.14/bioc/html/scran.html) and scater (https://bioconductor.org/packages/3.14/bioc/html/scater.html). Three different approaches were used for dimension reduction clustering, Principal Component Analysis, *t*-stochastic neighbor embedding (*t*-SNE) and uniform manifold approximation and projection (UMAP).

Graph-based clustering of the single cells based on their gene expression pattern was performed using the scran R package. Briefly, clusters were constructed where each node is a cell and the edges are weighted based on the similarity between cells involved. Next, clusters were identified based on similarity with the nearest 10 cells; 19 clusters were identified using this approach. For further filtration, we removed the clusters that contained less than 100 cells, and about seven clusters were obtained for further analysis. Two approaches were used to detect doublets using the scDblFinder R package. The first approach detects doublets as clusters with expression profiles lying between two other clusters. The second doublet detection strategy involves artificial simulation of doublets from the single-cell expression data and then trains a classifier to identify putative doublet cells among the real cells.

Gene Set Enrichment Analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG)[88] pathway analysis and Gene Ontology (GO) term enrichment were analyzed by performing hypergeometric tests (R phyper function) for each individual term/pathway and FDR correction was applied (FDR <0.05). Further analysis of differentiating characteristics between $18Lin^{neg}CD34^{pos}CD133^{pos}$ and $18Lin^{neg}CD34^{neg}CD133^{pos}$ populations involved Gene Set Enrichment Analysis (GSEA), as previously described.[89] GSEA can identify whether the members of a gene set (a collection of which are housed in the Molecular Signatures Database [MSig-DB] - Broad Institute, Boston, MA) are enriched in an independent rank-ordered profile of genes that are differentially expressed between two experimental groups. Thus, GSEA is able to provide definitions of overrepresented biological functions without the implicit bias associated with cut-off-based analyses. After ranking all genes for which a test statistic could be computed based on a fold-change (comparing CBC3 to CBC1), the GSEA function was used (pre-ranked, "classic" mode with 10,00 permutations) to calculate the enrichment of focused gene sets among the differentially expressed genes.

Differential expression analysis

DEsingle[85] was selected for differential expression analysis based on a recently published benchmark.[90] Briefly, a Zero-Inflated Negative Binomial model was used to determine the their proportion of real and dropout zeros and to identify three types of differential expressed (DE) genes—DEs: genes with a significant difference in the proportion of real zeros but with no difference in the other cells; DEa: significantly differentially expressed genes with no difference on proportion of real zeros; and DEg: genes with significant difference in both the proportions of real zeros and the expression abundances. Single cell RNA-Seq data was deposited to Gene Expression Omnibus (GEO) under accession number GEO: GSE237832.

Identification of cell type

Two approaches were combined to infer cell types for each cluster. First, we directly examined the expression levels of a set of canonical markers for the target cell types (Table S2). Enrichment of these markers in certain clusters was considered a strong indication of the clusters representing the corresponding cell types. The second approach used the automatic cell type identification method with the SingleR package (version 1.99.2) in R. Built-in reference Human Primary Cell Atlas Data[91] was used to annotate each cell in HSCs populations with known labels based on similarity to the reference via SingleR() function. The label with the highest score was assigned to the test cell, followed by further fine-tuning to resolve closely related labels. Together, the top-ranked cell types were considered as the labels for each cluster and provided cell-type information.

Inference of the developmental trajectory for CB-HSCs

The cell state transitions for CB-HSCs were estimated using the Monocle (version 2) algorithm. The gene-cell matrix in the scale of UMI counts was provided as an input to Monocle, and then, its newCellDataSet function was called to create an object with the parameter expressionFamily = negbinomial.size. The cell trajectory within the integrated Seurat object which includes $CD34^{pos}$ and $CD34^{neg}$ cells was inferred using the default parameters of Monocle after dimension reduction and cell ordering.

Statistical analysis

All experiments were carried out in triplicate. Data are represented as the mean ± standard deviation (SD). Data were plotted and statistically analyzed using GraphPad Prism version 9.0 for Mac. Statistical analysis was performed using either a two-sided Fisher exact test, or Welch's t-test, or two-tailed Student's *t* test or one-way ANOVA. p-values <0.05 were considered significant. Specific statistical tests used, significance, number of animals used and other details are indicated in individual figure legends.