





OPEN

# Hyaluronic acid-GPRC5C signalling promotes dormancy in haematopoietic stem cells

Yu Wei Zhang<sup>1,2,3</sup>, Julian Mess<sup>1,2,4,5</sup>, Nadim Aizarani<sup>1,2,3</sup>, Pankaj Mishra<sup>6</sup>, Carys Johnson<sup>7</sup>, Mari Carmen Romero-Mulero<sup>1,2</sup>, Jasmin Rettkowski<sup>1,2,4</sup>, Katharina Schönberger<sup>1,2,3</sup>, Nadine Obier<sup>1</sup>, Karin Jäcklein<sup>1</sup>, Nadine M. Woessner<sup>2,4,5,8</sup>, Maria-Eleni Lalioti<sup>1</sup>, Talia Velasco-Hernandez<sup>9</sup>, Katarzyna Sikora<sup>1</sup>, Ralph Wäsch<sup>10</sup>, Bernhard Lehnertz<sup>11</sup>, Guy Sauvageau<sup>11</sup>, Thomas Manke<sup>1</sup>, Pablo Menendez<sup>8,12,13</sup>, Sebastian Gottfried Walter<sup>14</sup>, Susana Minguet<sup>2,5,8</sup>, Elisa Laurenti<sup>7</sup>, Stefan Günther<sup>6</sup>, Dominic Grün<sup>1,5,15,16</sup> and Nina Cabezas-Wallscheid<sup>1,5</sup>  

**Bone marrow haematopoietic stem cells (HSCs) are vital for lifelong maintenance of healthy haematopoiesis. In inbred mice housed in gnotobiotic facilities, the top of the haematopoietic hierarchy is occupied by dormant HSCs, which reversibly exit quiescence during stress. Whether HSC dormancy exists in humans remains debatable. Here, using single-cell RNA sequencing, we show a continuous landscape of highly purified human bone marrow HSCs displaying varying degrees of dormancy. We identify the orphan receptor GPRC5C, which enriches for dormant human HSCs. GPRC5C is also essential for HSC function, as demonstrated by genetic loss- and gain-of-function analyses. Through structural modelling and biochemical assays, we show that hyaluronic acid, a bone marrow extracellular matrix component, preserves dormancy through GPRC5C. We identify the hyaluronic acid-GPRC5C signalling axis controlling the state of dormancy in mouse and human HSCs.**

Haematopoietic stem cells (HSCs) sit at the top of a hierarchically organized system and can reconstitute the entire blood system<sup>1,2</sup>. Many seminal studies have demonstrated important mechanisms regulating human HSCs<sup>3–9</sup>. For instance, Dick and colleagues have recently shown that TFEB-mediated endolysosomal activity maintains long-term-HSC (LT-HSC) quiescence<sup>10</sup>. Cell cycle profiles have been used to further distinguish and characterize heterogeneity within human LT-HSCs<sup>5,10–13</sup>. This unique approach has allowed researchers to identify key mechanisms governing the HSC state, such as how a quiescent reservoir of human HSCs resists regenerative stress to preserve key stem cell features<sup>12</sup>. Still, most studies have largely focused on cord blood (CB) HSCs, which represent a more primitive population than their counterparts in the adult bone marrow (BM) and have a distinctive transcriptomic profile<sup>10–12,14</sup>. Delineating human HSCs in adult BM is critical to answering fundamental unresolved questions, including whether dormant HSCs (dHSCs) exist, how the stem cell state is maintained and how aberrant haematopoiesis such as leukaemia develops.

Using long-term label-retaining cell assays, two subsets of HSCs were previously identified in the mouse system: dHSCs and active HSCs (aHSCs)<sup>15–17</sup>. Both rest in a non-cycling quiescent state, but dHSCs are defined by a deeper state of quiescence and harbour the highest stem cell potential<sup>17</sup>. Maintaining dormancy minimizes

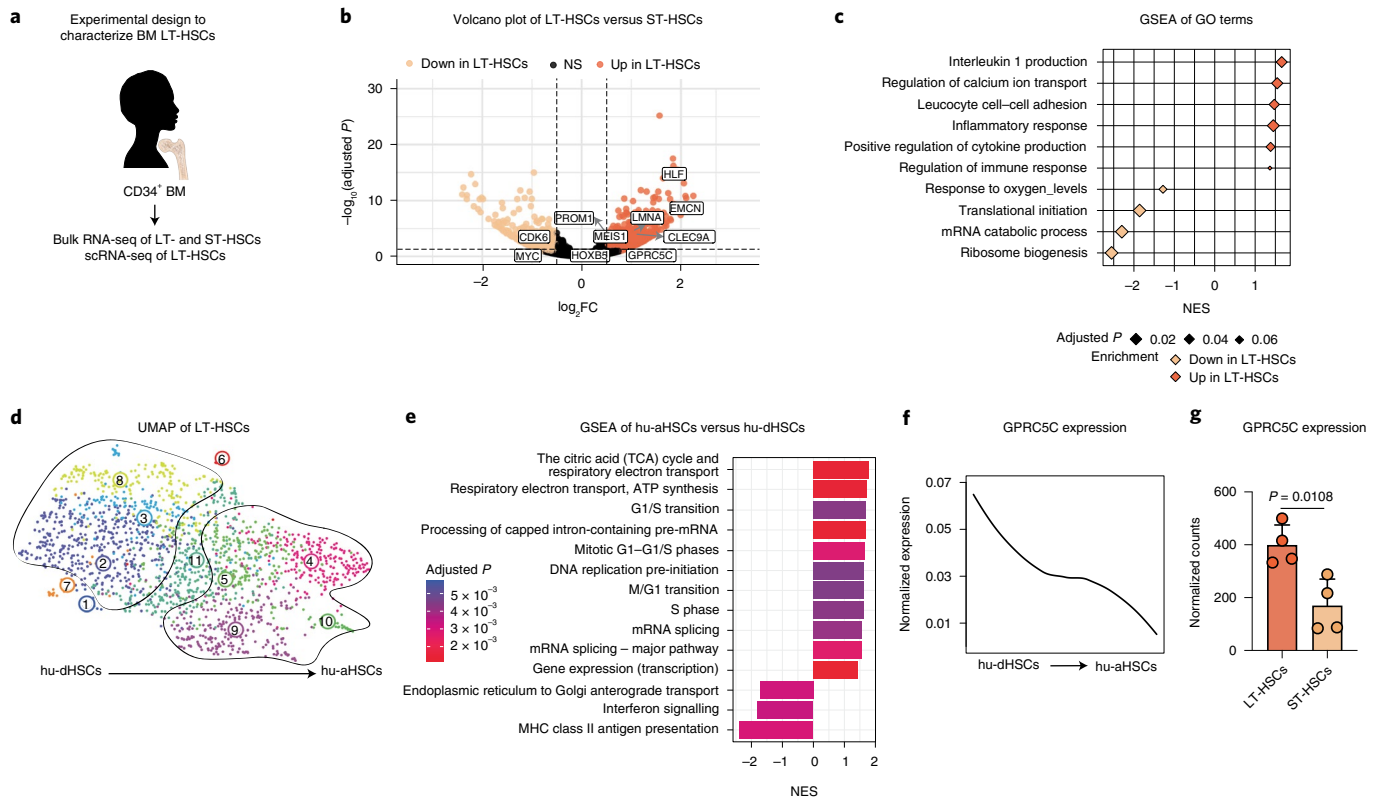
stress from cellular respiration and DNA damage, contributing to HSC longevity and functionality<sup>18</sup>. Still, the concept of dormancy has only been established in artificial hygienic conditions such as gnotobiotic mouse facilities.

In this Article, we used single-cell RNA sequencing (scRNA-seq) to identify a population of stem cells resembling mouse dHSCs among human BM LT-HSCs. We found that GPRC5C enriches for dormancy and is required to maintain stem cell features. We identified the hyaluronic acid (HA)–GPRC5C signalling axis controlling HSC state.

## Results

**Dormancy in human BM stem cells.** To understand adult HSC regulation, we performed genome-wide transcriptional profiling of healthy BM LT-HSCs (CD34<sup>+</sup>CD38<sup>–</sup>CD45RA<sup>–</sup>CD90<sup>+</sup>CD49f<sup>+</sup>) and short-term HSCs (ST-HSCs; CD34<sup>+</sup>CD38<sup>–</sup>CD45RA<sup>–</sup>CD90<sup>–</sup>CD49f<sup>–</sup>) (Fig. 1a and Extended Data Fig. 1a, b). As expected, we detected upregulation of stem cell- and quiescence-associated genes (HLF and CLEC9A) and downregulation of cell cycle-related genes (CDK6 and MYC) in LT-HSCs compared with ST-HSCs (Fig. 1b and Supplementary Table 1). Gene set enrichment analysis (GSEA) showed enrichment for categories associated with immune function in LT-HSCs (Fig. 1c and Extended Data Fig. 1c,d) and metabolic activity in ST-HSCs (Fig. 1c and Extended Data Fig. 1c).

<sup>1</sup>Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany. <sup>2</sup>Faculty of Biology, University of Freiburg, Freiburg, Germany. <sup>3</sup>International Max Planck Research School for Molecular and Cellular Biology (IMPRS-MCB), Freiburg, Germany. <sup>4</sup>Spemann Graduate School for Biology and Medicine (SGBM), Freiburg, Germany. <sup>5</sup>Centre for Integrative Biological Signalling Studies (CIBSS), Freiburg, Germany. <sup>6</sup>Pharmaceutical Bioinformatics, University of Freiburg, Freiburg, Germany. <sup>7</sup>Department of Haematology and Wellcome and MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. <sup>8</sup>Signalling Research Center BIOS, Freiburg, Germany. <sup>9</sup>Josep Carreras Leukemia Research Institute-Campus Clinic and Department of Biomedicine, School of Medicine, University of Barcelona, Barcelona, Spain. <sup>10</sup>Department of Hematology, Oncology and Stem Cell Transplantation, Faculty of Medical, University of Freiburg, Freiburg, Germany. <sup>11</sup>Institute for Research in Immunology and Cancer, University of Montreal, Montreal, Canada. <sup>12</sup>Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain. <sup>13</sup>Spanish Network for Cancer Research (CIBER-ONC)-ISCIII, Barcelona, Spain. <sup>14</sup>Department of Orthopedic Surgery, University Hospital of Cologne, Cologne, Germany. <sup>15</sup>Würzburg Institute of Systems Immunology, Max Planck Research Group at the Julius-Maximilians-Universität, Würzburg, Germany. <sup>16</sup>Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Würzburg, Germany.  e-mail: [cabezas@ie-freiburg.mpg.de](mailto:cabezas@ie-freiburg.mpg.de)

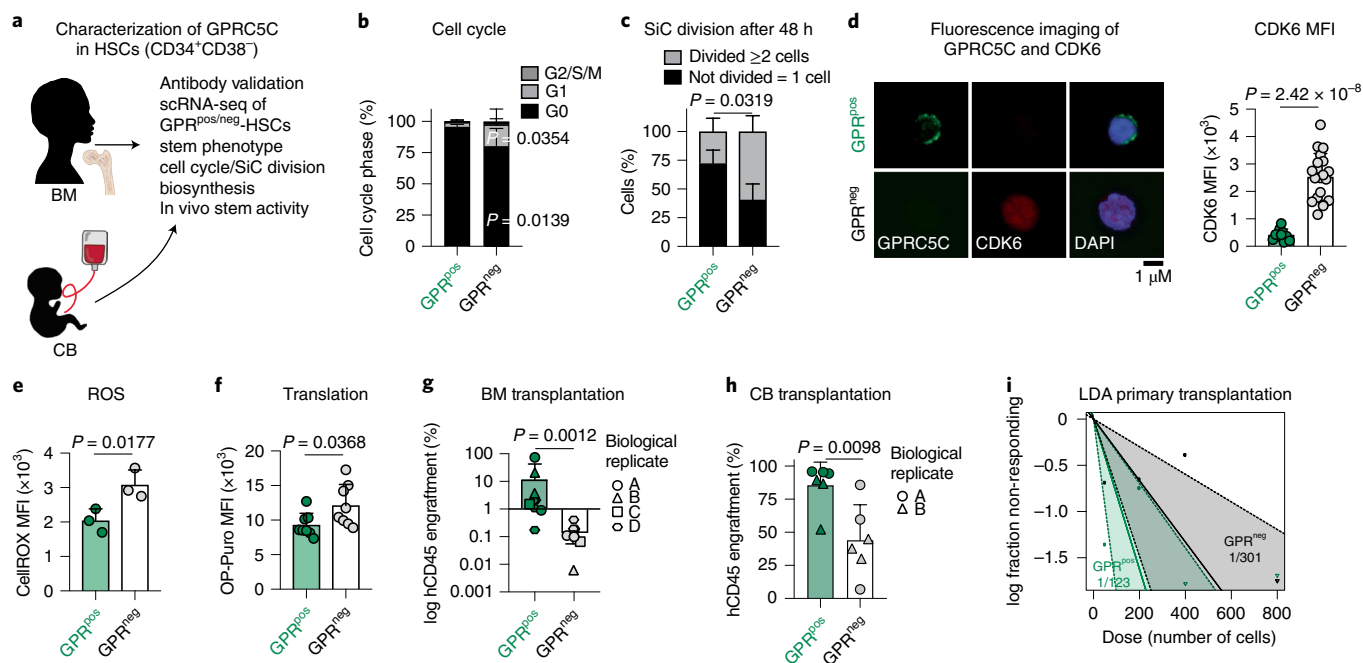


**Fig. 1 | Identification of a highly dormant HSC population in human BM.** **a**, Experimental design to characterize LT-HSCs from human BM donors. **b**, Volcano plot depicting DEGs between LT-HSCs and ST-HSCs from four BM donors. Benjamini-Hochberg (BH)-adjusted  $P$  values after Wald test.  $\log_2FC \geq 0.5$ ; adjusted  $P < 0.1$ . **c**, GSEA of GO terms in LT-HSCs compared with ST-HSCs. Normalized enrichment score (NES). **d**, UMAP of single-cell transcriptomes of human BM LT-HSCs from two donors highlighting VarID clusters. Numbers denote clusters. **e**, GSEA for genes that were differentially expressed between clusters 4, 5, 9 and 10 (hu-aHSCs) and 2, 3 and 8 (hu-dHSCs). BH-corrected  $P \leq 0.05$ . **f**, Smoothed pseudo-temporal normalized expression profile of GPRC5C in LT-HSCs. Smoothed profiles were computed by local regression. **g**, Bulk RNA-seq data depicting normalized counts of GPRC5C in LT-HSCs and ST-HSCs.  $n = 4$  biological replicates. Two-tailed  $t$ -test was performed. Data are presented as mean  $\pm$  s.d. The y axis of the smoothed pseudo-temporal expression profiles indicates normalized expression. GSEA was performed with BH-adjusted  $P$  values after adaptive multi-level splitting Monte Carlo approach. LT-HSCs (CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>CD90<sup>+</sup>CD49f<sup>+</sup>), ST-HSCs (CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>CD90<sup>-</sup>CD49f<sup>-</sup>) and HSCs (CD34<sup>+</sup>CD38<sup>-</sup>). Source numerical data are available in source data.

Next, we performed scRNA-seq of human BM LT-HSCs (Fig. 1a and Extended Data Fig. 1a) and found that cells formed a continuum (Fig. 1d)<sup>19</sup>. We pseudo-temporally ordered cells on the basis of their uniform manifold approximation and projection (UMAP)  $x$ -axis coordinate and then grouped pseudo-time profiles of all expressed genes into co-expression modules by applying a self-organizing map (SOM) (Extended Data Fig. 1e)<sup>20</sup>. Pathway enrichment analysis of the upregulated genes in modules 1, 2, 3 and 4, as defined by the SOM (corresponding to genes upregulated in clusters 9 and 10 of the UMAP), revealed an over-representation of terms associated with ‘translation’, suggesting these were more metabolically active HSCs (Extended Data Fig. 1f). GSEA of clusters 4, 5, 9 and 10 of the UMAP revealed an enrichment of cell cycle-related terms (Fig. 1e). Conversely, immune function-related terms were enriched in clusters 2, 3 and 8 of the UMAP. This pattern was consistent with processes previously observed when comparing human LT-HSCs versus ST-HSCs (Fig. 1c) and mouse dHSCs versus aHSCs<sup>21</sup>. We projected the ‘mouse dormant HSCs’ signature and found that clusters were organized along an ‘axis of dormancy’ (Extended Data Fig. 1g). Similar to mouse haematopoiesis, human HSC activation represented a continuum composed of intermediate cell states. The most dormant cells were located on the left-hand side of the UMAP (human dHSCs (hu-dHSCs)) expressing quiescence genes (MEG3 and CLEC9A), whereas the most active HSCs were on the

right-hand side expressing activation-associated genes (human aHSCs (hu-aHSCs); CDK6 and PCNA; Fig. 1d and Extended Data Fig. 1g,h)<sup>6</sup>. Further, we found a positive enrichment of two recently published quiescent CB LT-HSC gene signatures when projected onto our axis of dormancy (Extended Data Fig. 1i)<sup>10,11</sup>.

**GPRC5C enriches for a quiescent stem cell population.** To characterize dormancy in human stem cells, we screened for potential cell surface receptors. We decided to focus on GPRC5C because (1) its expression was enriched within hu-dHSCs (Fig. 1f), (2) it showed higher expression in LT-HSCs compared with ST-HSCs (Fig. 1g), and (3) GPRC5C had not been studied in humans<sup>21</sup>. Using a monoclonal human GPRC5C antibody, we investigated the role of GPRC5C in quiescence (Extended Data Fig. 2a–c). We analysed the cell cycle profile and found that GPRC5C<sup>pos</sup>-HSCs (CD34<sup>+</sup>CD38<sup>-</sup>) resided almost exclusively in the G<sub>0</sub> phase, while GPRC5C<sup>neg</sup>-HSCs were detected in both the G<sub>0</sub> and G<sub>1</sub> phases (Fig. 2a,b and Extended Data Fig. 2d). In vitro single-cell division analysis showed that GPRC5C<sup>neg</sup>-HSCs were twice as likely to undergo division compared with GPRC5C<sup>pos</sup>-HSCs (Fig. 2c). Consistent with CDK6 levels being linked to a quiescent but cell-cycle-primed state<sup>6</sup>, GPRC5C<sup>pos</sup>-HSCs expressed lower levels of CDK6 (Fig. 2d and Extended Data Fig. 2e,f). Using CellROX, we found lower reactive oxygen species (ROS) levels within the GPRC5C<sup>pos</sup>-HSCs



**Fig. 2 | Characterization of GPRC5C as a marker of quiescent stem cells.** **a**, Experimental design to characterize GPRC5C in human BM and CB. **b**, Fluorescence-activated cell sorting (FACS)-based analysis of cell cycle from BM GPRC5C<sup>pos/neg</sup>-HSCs.  $n = 5$  biological replicates. **c**, Single-cell (SiC) division assay of GPRC5C<sup>pos/neg</sup>-HSCs from BM donors quantified after 48 h in vitro culture.  $n = 3$  biological replicates. **d**, Left: fluorescence imaging of GPRC5C overlaid with intracellular CDK6 expression from BM HSCs. Right: quantification of the median fluorescence intensity (MFI) of CDK6 in GPRC5C<sup>pos/neg</sup>-HSCs from two BM donors.  $n = 10$  GPRC5C<sup>pos</sup> cells and  $n = 20$  GPRC5C<sup>neg</sup> cells. **e**, FACS measurement of CellROX Deep Red MFI in GPRC5C<sup>pos/neg</sup>-HSCs from BM donors.  $n = 3$  biological replicates. **f**, FACS measurement of OP-Puro MFI in GPRC5C<sup>pos/neg</sup>-HSCs from BM donors.  $n = 8$  biological replicates. **g**, GPRC5C<sup>pos/neg</sup>-HSCs from four BM donors were transplanted in NSBGW mice, which were killed at 36 weeks. Engraftment was assessed by percentage of human CD45-positive cells.  $n = 9$  biological replicates. **h**, 800 GPRC5C<sup>pos/neg</sup>-HSCs from two CB donors were transplanted in NSBGW mice, which were killed at 24 weeks. Engraftment was assessed by percentage of human CD45-positive cells.  $n = 6$  biological replicates. **i**, LDA to estimate the HSC frequency within GPRC5C<sup>pos/neg</sup>-HSCs from two CB donors at 24 weeks post-transplantation by extreme limiting dilution analysis (ELDA). All data presented as mean  $\pm$  s.d. Statistical significance was determined by two-tailed  $t$ -test (**d**, **e**, **f** and **h**), two-way analysis of variance (ANOVA) (**b** and **c**), or Mann-Whitney  $U$  test (**g**). NS, not significant.  $n$  indicates number of replicates. For all experiments, at least two independent experiments were performed. Source numerical data are available in source data.

(Fig. 2e and Extended Data Fig. 2g). We next performed index sorting of BM GPRC5C<sup>pos/neg</sup>-HSCs for mCEL-Seq2 scRNA-seq<sup>20,22</sup> (Extended Data Fig. 2h). Notably, only weak GPRC5C RNA expression was detected, and downstream analyses relied on the index sorting to identify GPRC5C<sup>pos/neg</sup>-HSCs (Extended Data Fig. 2i). The ‘mouse dormant HSCs’ signature was enriched in clusters 4 and 8, which were mainly represented by GPRC5C<sup>pos</sup>-HSCs (Extended Data Fig. 2j,k). Similarly, we found enrichment of stem-cell-related genes (MEG3 and HIF1a) complemented by lower expression of activation-related genes (CDK6 and CD74 (ref. <sup>23</sup>); Extended Data Fig. 2l). At the process level, immune-related terms were enriched, while translation-related processes were poorly represented as also shown by lower OP-Puro incorporation (a measurement of de novo protein synthesis)<sup>24</sup> in GPRC5C<sup>pos</sup>-HSCs (Fig. 2f and Extended Data Fig. 2m,n).

Phenotypically, GPRC5C<sup>pos</sup> cells in CD34<sup>+</sup>-enriched BM possessed twice the LT-HSCs compared with GPRC5C<sup>neg</sup> (Extended Data Fig. 3a). Functionally, in vivo reconstitution levels of BM GPRC5C<sup>pos</sup>-HSCs were higher compared with GPRC5C<sup>neg</sup>-HSCs (Fig. 2g and Extended Data Fig. 3b). We observed an increased frequency of CD19<sup>+</sup>-lymphoid cells in GPRC5C<sup>pos</sup>-HSC-engrafted animals, likely reflecting the larger grafts (Fig. 2g and Extended Data Fig. 3c). Single-cell differentiation assay of GPRC5C<sup>pos/neg</sup>-LT-HSCs showed no differences in the types of colonies produced (Extended Data Fig. 3d).

GPRC5C<sup>pos</sup>-HSCs from CB mimicked the properties of BM-derived GPRC5C<sup>pos</sup>-HSCs, including a higher frequency of G<sub>0</sub> cells, delayed entry into the cell cycle, lower CDK6 expression and enrichment of phenotypic LT-HSCs compared with CB GPRC5C<sup>neg</sup>-HSCs (Extended Data Fig. 3e–i). Functionally, peripheral blood (PB) and BM analysis from in vivo transplantation revealed that CB GPRC5C<sup>pos</sup>-HSCs had higher reconstitution capacity, with no lineage bias (Fig. 2h and Extended Data Fig. 3j,k). Limiting dilution analysis (LDA) revealed CB GPRC5C<sup>pos</sup>-HSCs contained twice as many HSCs compared with CB GPRC5C<sup>neg</sup>-HSCs (Fig. 2i and Extended Data Fig. 3l).

In summary, GPRC5C<sup>pos</sup>-HSCs are highly quiescent, possess low biosynthetic activity and have high long-term reconstitution capabilities.

**Loss of GPRC5C disturbs quiescence and stemness.** To explore the functionality of GPRC5C, we transduced human CD34<sup>+</sup> BM cells with a GPRC5C-targeting short hairpin RNA vector (Extended Data Fig. 4a–c). GPRC5C knockdown (KD) resulted in decreased cells in G<sub>0</sub> phase, a faster division rate and increased CDK6 expression (Extended Data Fig. 4d–f). CD34<sup>+</sup>EPCR<sup>+</sup>, a reliable marker of ex vivo expanded HSCs<sup>3</sup>, was reduced upon GPRC5C KD (Extended Data Fig. 4g). We performed RNA-seq analysis, projected the gene signature of GPRC5C-KD cells onto our axis of dormancy (Fig. 1d, scRNA-seq of LT-HSCs), and found

it was enriched for hu-aHSCs genes (Extended Data Fig. 4h,i). Similarly, GPRC5C KD upregulated genes associated with ‘mouse active HSCs’, and downregulated genes associated with ‘mouse dormant HSCs’ and ‘quiescent LT-HSCs’ from García-Prat et al.<sup>10</sup> (Extended Data Fig. 4j). GSEA indicated that ‘oxidative phosphorylation’ and ‘reactive oxygen species pathway’ were enriched in GPRC5C-KD cells, which was corroborated by MitoTracker Red, an indicator of mitochondrial activity (Extended Data Fig. 4k–n). We observed upregulation of translation-related terms and downregulation of immune-related processes upon GPRC5C KD (Extended Data Fig. 4k–m). Functionally, we performed in vitro serial colony-forming unit (CFU) assays. Primary plating of GPRC5C-KD cells exhibited increased colony-forming abilities, whereas secondary plating produced fewer colonies compared with control (Extended Data Fig. 4o). This is consistent with the notion that GPRC5C-KD cells have enhanced progenitor activity but reduced stem-like properties.

To validate our analysis on a purer HSC population and improve KD efficiency, we used BM LT-HSCs and targeted *GPRC5C* using two short hairpin RNAs (Fig. 3a). Consistent with our previous findings, we found accelerated in vitro single-cell division and a decrease in immature CD34<sup>+</sup>EPCR<sup>+</sup> cells upon GPRC5C KD (Fig. 3b–d). GSEA of GPRC5C-KD cells showed enrichment of terms related to biogenesis and processing pathways, indicating an enhanced metabolically active state in GPRC5C-KD HSCs (Fig. 3e).

Since GPRC5C expression is conserved between mice and humans, we used *Gprc5c*-knockout (KO) mice to perform experiments that are challenging or not feasible using primary human samples (Fig. 3f)<sup>25</sup>. *Gprc5c*-KO mice showed no differences in the frequency of phenotypic HSCs (Lineage<sup>−</sup>Sca1<sup>+</sup>cKit<sup>+</sup>[LSK] CD150<sup>+</sup> CD48<sup>−</sup>CD34<sup>−</sup>), multipotent progenitors (MPP) or differentiated blood cells compared with controls (Extended Data Fig. 5a–c). However, *Gprc5c*-KO HSCs showed downregulation of dormancy-associated genes (*Rarb* and *Epcr*) and upregulation of cell cycle-associated genes (*Pcna* and *Cdk6*) (Fig. 3g). Examination of the cell cycle profiles of *Gprc5c*-KO HSCs revealed a slight decrease of cells in the G<sub>0</sub> phase (Fig. 3h). In vitro single-cell cultivation of HSCs confirmed that *Gprc5c* KO had an accelerated division rate compared with wild-type (WT) control (Fig. 3i). Transcriptome analysis of *Gprc5c*-KO and WT HSCs indicated no compensatory upregulation of other members of GPCR class C group 5 (Extended Data Fig. 5d,e). GSEA analysis revealed downregulation of the ‘mouse dormant HSCs’ signature and upregulation of the ‘mouse active HSCs’ signature in *Gprc5c*-KO HSCs (Extended Data Fig. 5f). Subsequently, processes associated with ‘translation’ and ‘cell cycle’ were significantly enriched in *Gprc5c*-KO HSCs (Fig. 3j and Extended Data Fig. 5g). CellROX staining also revealed increased ROS levels in *Gprc5c*-KO HSCs (Extended Data Fig. 5h). These data indicate that loss of *Gprc5c* primes HSCs towards activation.

Functionally, we generated whole BM (WBM) competitive and HSC chimaeras (Extended Data Fig. 5i,j). From secondary WBM transplantations, we observed a significant reduction in the competitive engraftment potential of *Gprc5c*-KO cells compared with WT control (Fig. 3k). Similarly, we detected reduced reconstitution capabilities in *Gprc5c*-KO HSCs, but no changes in lineage distribution (Fig. 3l and Extended Data Fig. 5k,l). Possible bias from the initial engraftment and niche-specific effects were excluded by performing mobilization and reverse chimaera experiments (Extended Data Fig. 5m–p). Lastly, it has been shown that dHSCs are essential for proliferation-related stress<sup>18,26</sup>. Consistent with this, we found that *Gprc5c*-KO mice were more susceptible to serial injections of the chemotherapeutic agent fluorouracil (5-FU) (Extended Data Fig. 5q). Together, our results suggest that loss of *Gprc5c* leads to a loss of quiescence, self-renewal, and ultimately, an impaired ability to cope with proliferative stress conditions.

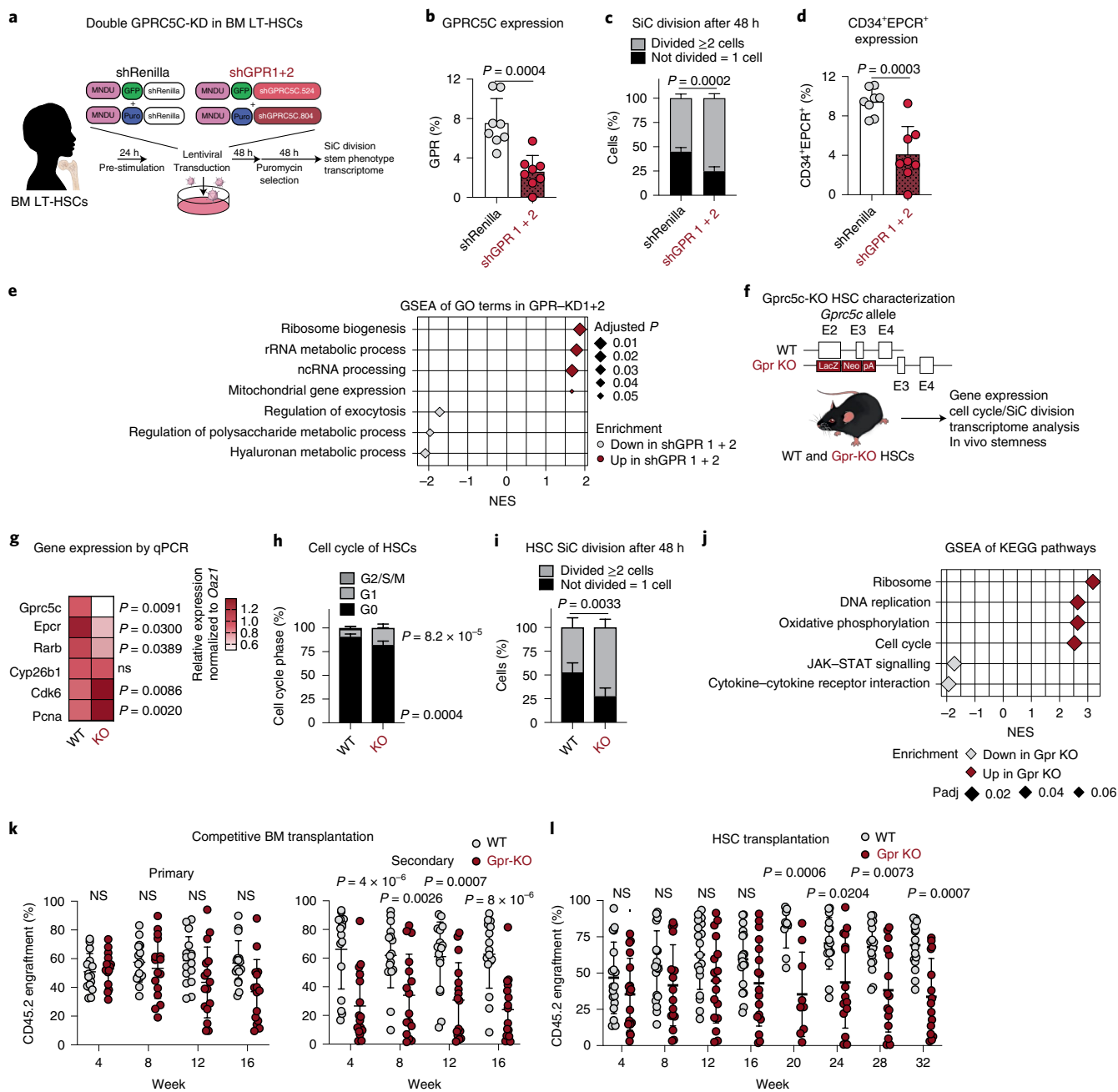
**Enhanced quiescence and stemness by overexpressing GPRC5C.** Next, we engineered GPRC5C overexpression (OE) in human CD34<sup>+</sup>CD38<sup>−</sup> mobilized peripheral blood (mPB) cells (Fig. 4a,b and Extended Data Fig. 6a). GPRC5C OE led to enhanced quiescence as indicated by an increase of G<sub>0</sub> cells, a slower first division rate and decreased CDK6 expression (Fig. 4c–e and Extended Data Fig. 6b). As expected, we observed an activation of HSCs upon culture, in contrast to freshly isolated HSCs, which are highly quiescent<sup>10</sup> (Fig. 2c). We detected maintenance of CD34<sup>+</sup>CD90<sup>+</sup> and CD34<sup>+</sup>EPCR<sup>+</sup> cells in GPRC5C-OE cells, consistent with enhanced HSC functionality (Fig. 4f,g and Extended Data Fig. 6c).

Transcriptome analysis revealed downregulation of genes linked to activation or cell cycle (*CDK6* and *MYC*) and upregulation of stemness-associated genes (*PROCR*) in GPRC5C-OE cells (Fig. 4h and Extended Data Fig. 6d). Interrogation of expression of the upregulated genes from GPRC5C-OE cells in the scRNA-seq data of LT-HSCs revealed that their expression was enriched in the hu-dHSC population (Fig. 1d and Fig. 4i). Similarly, GSEA revealed enrichment of genes related to ‘mouse dormant HSCs’ and ‘quiescent LT-HSCs’ from García-Prat et al.<sup>10</sup> (Extended Data Fig. 6e). Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis showed that terms associated with inflammatory signalling were enriched while ‘ribosome’ and ‘translational initiation’ were downregulated in GPRC5C-OE cells. The decrease in de novo translation was confirmed using OP-Puro (Fig. 4j,k and Extended Data Fig. 6f–h). We probed the expression of mitochondrial genes in the transcriptome and observed reduced mitochondrial DNA copy number (Extended Data Fig. 6i). In addition, we found decreased mitochondria potential in the GPRC5C-OE cells, as indicated by MitoTracker Red analysis (Fig. 4l and Extended Data Fig. 6j).

Serial CFU assay revealed that GPRC5C-OE cells had enhanced serial in vitro self-renewal capacity, which was confirmed by in vivo transplantation (Fig. 4m and Extended Data Fig. 6k). Lineage distribution revealed no difference, but GPRC5C OE showed a slight increase in the frequency of CD34<sup>+</sup> cells, consistent with enhanced HSC self-renewal capabilities (Fig. 4n). In secondary transplantations, all four GPRC5C-OE mice (compared with only one of four control mice) met the requirement of positive human engraftment (≥0.1% hCD45 cells) (Fig. 4o). In agreement, GPRC5C OE in BM cells also resulted in increased quiescence, lower ROS and enrichment of phenotypic HSC surface marker expression (Extended Data Fig. 6l–r).

We thus observed that ectopic GPRC5C reduced biosynthesis, but increased quiescence and enhanced the reconstitution ability of HSCs.

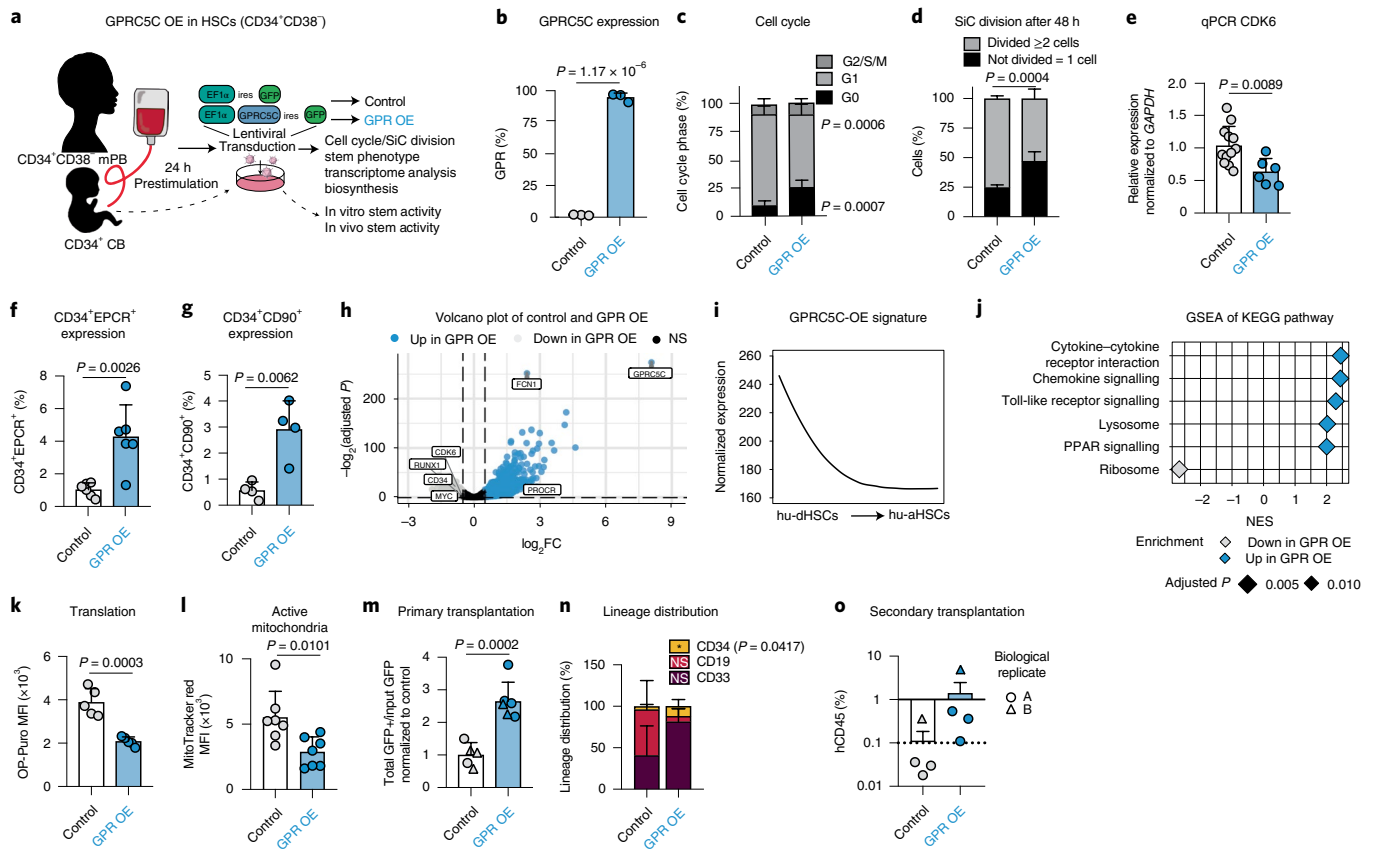
**HA activates GPRC5C signalling.** GPRC5C is a seven-transmembrane orphan receptor<sup>27</sup>. To predict potential regulators of dormancy, we performed in silico structural modelling of GPRC5C. Briefly, the full length of GPRC5C isoform 2 was modelled using a recombinant approach of homology and *ab initio* modelling (Fig. 5a and Methods). Like other members of the G-protein-coupled receptor (GPCR) superfamily, three distinct regions were identified in GPRC5C, namely: an extracellular domain (ECD), a seven-transmembrane domain (7TM), and an intracellular domain. However, unlike all other class C GPCRs, the ECD of GPRC5C lacks a characteristic venus fly trap (VFT) domain, which is the orthosteric binding site<sup>28–30</sup>. Instead, the ECD of GPRC5C consists of three  $\alpha$ -helices and two  $\beta$ -helices, along with the loop regions that connect the ECD to the 7TM, revealing a topology more closely related to class A than class C GPCRs. In silico molecular docking suggested that HA may interact with GPRC5C through the 7TM (Fig. 5a). Interestingly, HA activates class A GPCR  $\kappa$ -opioid receptors, although the mechanism of action is not yet understood<sup>31</sup>. HA is a major component of the BM niche and facilitates homing and binding of HSCs to endothelial cells<sup>32,33</sup>. In GPRC5C-OE cells, we found a significant increase in expression of GNA11 and GNA14, which are part of the G<sub>q</sub> alpha subunit (Extended Data Fig. 6s). Activation



**Fig. 3 | Loss of GPRC5C impairs stemness and quiescence.** **a**, Experimental design to knock down and assess the impact of decreased GPRC5C levels in BM LT-HSCs. **b**, FACS validation of GPRC5C protein levels.  $n = 8$  experiments. **c**, SiC division assay of GFP<sup>+</sup> BM cells quantified after 48 h in vitro culture.  $n = 4$  experiments for shGPR1+2 and  $n = 5$  experiments for shRenilla. **d**, Surface expression of CD34<sup>+</sup>EPCR<sup>+</sup> in GFP<sup>+</sup> BM cells.  $n = 8$  experiments. **e**, GSEA of GO terms in shGPR1+2 compared with shRenilla control. **f**, Experimental design to assess functional impact of Gprc5c KO in mouse HSCs. Deletion of Gprc5c was achieved by the insertion of LacZ downstream of the translation initiation site in exon 2 of both alleles. **g**, Differential gene expression. Heat map representing median RNA expression from qPCR data (normalized to housekeeping gene *Oaz1* and WT control).  $n = 8$  biological replicates. **h**, FACS-based HSC cell cycle analysis with Ki-67 and DAPI.  $n = 5$  biological replicates. **i**, SiC division assay of Gprc5c-KO and WT control HSCs quantified after 48 h in vitro culture.  $n = 4$  biological replicates for Gpr KO and  $n = 5$  biological replicates for WT. **j**, GSEA of KEGG pathways in Gprc5c-KO compared with WT control HSCs. **k**, PB analysis of competitive WBM chimaeras. CD45.2 cell percentage outcome in PB is presented.  $n = 16$  biological replicates. **l**, PB analysis of HSC transplants. CD45.2 cell percentage outcome in PB is presented.  $n = 17$  biological replicates for Gpr KO and  $n = 18$  biological replicates for WT. All data presented as mean  $\pm$  s.d. Statistical significance was determined by two-tailed *t*-test (**b**, **d** and **g**) or two-way ANOVA (**c**, **h**, **i**, **k** and **l**). *n* indicates number of replicates. GSEA was performed with BH-adjusted *P* values after adaptive multi-level splitting Monte Carlo approach. For all experiments, at least two independent experiments were performed. At least two human BM donors were used (**b**–**d**). Source numerical data are available in source data.

of G<sub>q</sub>-coupled GPCRs leads to a transient increase in intracellular calcium concentrations<sup>34</sup>. To measure calcium concentrations upon potential agonist interaction, we overexpressed GPRC5C and the

respective empty vector control in AML3 cells and loaded cells with Indo-1, a fluorescent Ca<sup>2+</sup> indicator dye (Fig. 5b). We tested HA and its subunits *N*-acetyl-D-glucosamine (NAG) and glucuronic acid



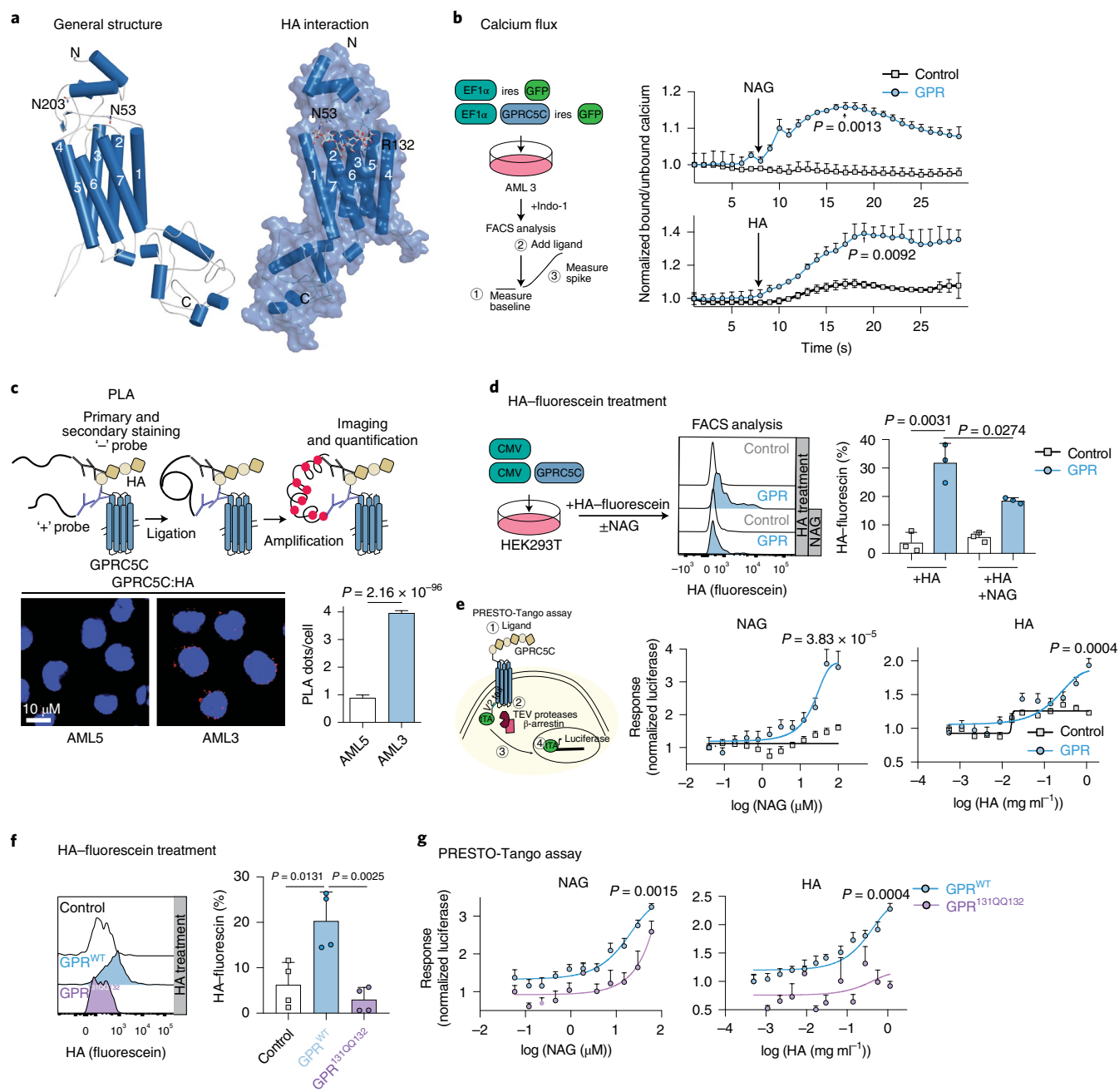
**Fig. 4 | GPRC5C OE enhances stemness.** **a**, Experimental design of GPRC5C OE in mPB and CB. Internal ribosomal entry site (ires). **b**, FACS validation of GPRC5C OE at protein level.  $n=3$  experiments. **c**, FACS-based cell cycle analysis of GPRC5C OE and control.  $n=6$  experiments. **d**, SiC division assay of GPRC5C OE and control quantified after 48 h in vitro culture.  $n=4$  experiments. **e**, qPCR of CDK6 expression. Normalized to housekeeping gene GAPDH and control.  $n=6$  experiments for GPRC5C OE and  $n=12$  experiments for control. **f**, Surface expression of CD34<sup>+</sup>EPCR<sup>+</sup> in GPRC5C OE and control.  $n=6$  experiments. **g**, Surface expression of CD34<sup>+</sup>CD90<sup>+</sup> in GPRC5C OE and control.  $n=4$  experiments. **h**, Volcano plot depicting DEGs between GPRC5C OE and control. Key genes associated with dormancy and stemness are highlighted. BH-adjusted  $P$  values after Wald test.  $\log_2FC \geq 0.5$ ; adjusted  $P < 0.1$ . **i**, Smoothed pseudo-temporal expression profile of the GPRC5C-OE signature in LT-HSCs. Smoothed profiles were computed by local regression. The y axis indicates normalized expression. **j**, GSEA of KEGG pathways in GPRC5C OE compared with control. Performed with BH-adjusted  $P$  values after adaptive multi-level splitting Monte Carlo approach. **k**, FACS measurement of OP-Puro MFI in GPRC5C OE and control.  $n=5$  experiments. **l**, FACS measurement of MitoTracker Red MFI in GPRC5C OE and control.  $n=7$  experiments. **m**, Transplantation of GPRC5C OE and control in CB cells from two biological replicates. Engraftment was measured from BM and calculated by the ratio of the total percentage GFP<sup>+</sup>hCD45<sup>+</sup> to the percentage input GFP post-transplantation and normalized to control.  $n=5$  control and  $n=6$  GPR-OE biological replicates. **n**, Lineage distribution of GPRC5C OE and control in CB cells. Distribution of myeloid (CD33), B-lymphoid (CD19) and immature surface phenotype (CD34) are presented.  $n=5$  control and  $n=6$  GPR-OE biological replicates. **o**, Secondary transplantation of 150,000 GFP<sup>+</sup>hCD45<sup>+</sup> cells.  $n=4$  biological replicates. All data presented as mean  $\pm$  s.d. Statistical significance was determined using two-tailed  $t$ -test (**b**, **e**, **f**, **g**, **l**, **m**, **n** and **k**) or two-way ANOVA (**c** and **d**).  $n$  indicates number of replicates. For all experiments, at least two independent experiments were performed. At least two human mPB or CB donors were used. Source numerical data are available in source data.

(GA) to elicit a spike in calcium. Addition of HA and NAG induced an increase in intracellular calcium levels, indicating that both activate GPRC5C signalling (Fig. 5b). In contrast, treatment with GA showed no changes in intracellular calcium (Extended Data Fig. 7a). As additional negative controls, we used glucosamine hydrochloride (a form of NAG lacking the amide group) and neuromedin B (NMB) (a known ligand of the NMB receptor), which we did not expect to interact with GPRC5C. As predicted, treatment with glucosamine hydrochloride and NMB did not alter intracellular calcium (Extended Data Fig. 7a).

To further support the interaction between HA and GPRC5C, we performed a proximity ligation assay (PLA; Fig. 5c). We used the AML3 and AML5 cell lines, since we found that both produce endogenous HA (due to expression of the HA-producing enzyme *HAS3*) but only AML3 cells express GPRC5C (Extended Data Fig. 7b,c),

avoiding the genetic modification of cells. We observed a significantly higher number of fluorescent dots in AML3 versus AML5 cells, indicating a proximity of less than 80 nm between GPRC5C and HA (Fig. 5c and Extended Data Fig. 7d). Next, we overexpressed GPRC5C and treated cells with an HA-conjugated fluorescein. GPRC5C-OE cells showed higher HA interaction compared with the control condition (Fig. 5d). To confirm NAG as the interacting subunit of HA, we designed a competitive interaction experiment. GPRC5C-OE cells were simultaneously treated with HA-conjugated fluorescein and non-fluorescent NAG. HA interaction with GPRC5C was reduced when NAG was present, suggesting that HA competes with NAG (Fig. 5d).

The desensitization of GPCRs depends on the recruitment of  $\beta$ -arrestin to the receptor, which uncouples the receptor from the G protein and promotes internalization and deactivation<sup>35</sup>.



**Fig. 5 | HA activates GPRC5C signalling.** **a**, Left: structural modelling of GPRC5C. Right: docked pose of the HA at the 7TM region of GPRC5C. **b**, Experimental design to determine intracellular calcium concentrations upon interaction with putative ligands. Normalized baseline calcium followed by flux in response to HA or NAG in control and GPRC5C-OE conditions in AML3.  $n = 3$  experiments. **c**, PLA between GPRC5C and HA, showing a distance smaller than 80 nm in AML3 cells but not AML5 cells. Plot representative of one of three independent experiments. Each independent experiment was performed in two technical replicates.  $n = 2250$  cells for AML5 and  $n = 2091$  cells for AML3. Presented as mean  $\pm$  s.e.m. **d**, HA-fluorescein treatment of HEK293T cells expressing GPRC5C and control. FACS plot representing HA retention in each condition.  $n = 3$  experiments. **e**, Left: schematic representation of PRESTO-Tango assay. Right: concentration-response curve of PRESTO-Tango construct for GPRC5C and control vector treated with NAG and HA in HTLA cells. Normalized to luciferase signal at the lowest concentration.  $n = 6$  experiments for NAG and  $n = 5$  experiments for HA. **f**, HA-fluorescein treatment of HEK293T cells expressing GPRC5C<sup>WT</sup>, GPRC5C<sup>131Q132</sup> and control. FACS plot representing HA retention in each condition.  $n = 4$  experiments. **g**, Concentration-response curve of PRESTO-Tango construct for GPRC5C WT and site-specific mutant (GPRC5C<sup>131Q132</sup>) at predicted HA interaction site in HTLA cells.  $n = 6$  experiments. All data presented as mean  $\pm$  s.d. unless otherwise stated. Statistical significance was determined using unpaired two-tailed  $t$ -test (**b**, **c**, **d**, **e**, **f** and **g**).  $n$  indicates number of replicates. For all experiments, at least two independent experiments were performed. Source numerical data are available in source data.

In the PRESTO-Tango assay,  $\beta$ -arrestin recruitment is quantified through luciferase<sup>36</sup> (Fig. 5e). We treated GPRC5C-OE HTLA cells with varying concentrations of HA, NAG and negative control

compounds such as GA, glucosamine hydrochloride and NMB. Quantification of luciferase revealed that HA and NAG treatments led to a dose-dependent response (Fig. 5e), while treatment with

GA, NMB and glucosamine hydrochloride did not elicit a luciferase response (Extended Data Fig. 7e). Together, these results suggest that HA and NAG induce GPRC5C activation.

To further dissect the underlying mechanism of GPRC5C activation by HA, we examined the putative mode of interaction. For some HA-binding proteins, a B(X<sub>7</sub>)B motif was described, where 'B' represents R (arginine) or K (lysine), and 'X' represents any non-acidic amino acid and at least one basic amino acid<sup>37–40</sup>. The 122KPDFSTCASRR132 motif connects TM2 and TM3 in GPRC5C. This motif is highly conserved between species and most closely resembles the B(X<sub>7</sub>)B motif (Extended Data Fig. 7f, black box). Further examination of contacting residues by molecular docking revealed that the interactions could be mediated through a strong network of hydrogen bonds with K122, along with a salt-bridge interaction with R132. To experimentally analyse the putative HA-binding site, we mutated 131RR132, two basic amino acids within the proposed B(X<sub>7</sub>)B domain, to 131QQ132 (Extended Data Fig. 7f). We transfected GPRC5C<sup>WT</sup> and GPRC5C<sup>131QQ132</sup> constructs and observed similar protein levels, as indicated by the Flag epitope conjugated at the N-terminus of GPRC5C (Extended Data Fig. 7g). GPRC5C<sup>131QQ132</sup> mutation led to a reduction in the ability to retain HA-conjugated fluorescein, and a reduced luciferase response compared with GPRC5C<sup>WT</sup> cells, further supporting the putative HA–GPRC5C interaction site (Fig. 5f,g). These results were additionally validated using a 131RR132 to 131AA132 GPRC5C mutant, which mimicked the loss of function previously observed in the GPRC5C<sup>131QQ132</sup> mutant (Extended Data Fig. 7h,i).

In eukaryotic cells, NAG is attached to an asparagine (N) as part of a process known as *N*-glycosylation. In many cases, *N*-glycosylation is critical for the structure and function of the protein and can modulate the ligand on-and-off rate<sup>41</sup>. The classical sequon for *N*-glycosylation is N-X-S/T (X ≠ P). This sequon was located at N<sup>203</sup>-S<sup>204</sup>-S<sup>205</sup> in GPRC5C. An additional putative C-containing glycosylation motif (N-X-C) was located at N<sup>53</sup>-K<sup>54</sup>-C<sup>55</sup>. This rare motif for the attachment of *N*-linked glycans is essential in some proteins<sup>42</sup>. To test whether glycosylation of these sites may be important for HA and NAG interaction, we generated GPRC5C<sup>N53L</sup> and GPRC5C<sup>N203L</sup> mutants. Using Flag-tag expression, we confirmed the presence of GPRC5C<sup>N53L</sup> on the cell membrane, while the GPRC5C<sup>N203L</sup> mutant was reduced on the cell surface (Extended Data Fig. 7j). Using HA-conjugated fluorescein, we showed that the GPRC5C<sup>N53L</sup> mutation reduced GPRC5C's ability to retain HA and dampened the luciferase response that was previously elicited by treatment with HA and NAG (Extended Data Fig. 7l,k). Similarly, *N*-glycosylation inhibitors such as PNGase and tunicamycin<sup>43,44</sup> hindered the ability of HA and NAG to elicit a luciferase response (Extended Data Fig. 7m,n). This suggests that NAG may partly modulate the activity of GPRC5C through the *N*-glycosylation of GPRC5C, which is critical for GPRC5C activation.

In summary, these data show that HA and its subunit NAG are regulators of GPRC5C activation.

**HA–Gprc5c signalling axis regulates HSCs.** To test whether the presence of HA and NAG could maintain quiescence and stemness, we sorted mouse Gprc5c<sup>pos/neg</sup>-HSCs on the basis of enhanced green fluorescent protein expression (EGFP) and treated them with HA or NAG for in vitro single-cell division, CFU assay, and in vivo transplantations (Fig. 6a). We observed that HA and NAG treatment induced quiescence, and enhanced in vitro and in vivo self-renewal capacity, in Gprc5c<sup>pos</sup>-HSCs specifically (Fig. 6b,c and Extended Data Fig. 8a). To exclude potential bias from HA<sup>pos</sup> cells, we sorted Gprc5c<sup>pos/neg</sup>-HSCs, which were HA<sup>neg</sup> and treated with HA and NAG; we nevertheless observed delayed division within the Gprc5c<sup>pos</sup>-HSCs specifically (Extended Data Fig. 8b).

Gprc5c-KO and WT HSCs were used to assess whether HA and NAG regulate HSCs through Gprc5c signalling (Fig. 6d). Treatment of WT HSCs with HA and NAG resulted in delayed division, increased CFU capability upon serial plating and enhanced in vivo reconstitution capacities upon transplantation (Fig. 6e,f and Extended Data Fig. 8c). However, in vitro HA and NAG treatment of Gprc5c-KO HSCs did not affect quiescence, in vitro self-renewal capacity or in vivo self-renewal capacity compared with WT HSCs (Fig. 6e,f and Extended Data Fig. 8c).

Next, we performed serial intraperitoneal injection of HA and NAG followed by poly I:C injection, which is known to drive HSCs into proliferation<sup>45</sup> (Fig. 6g). In vitro single-cell division analysis revealed that HSCs from WT mice treated with HA and NAG had delayed division compared with the control condition (Fig. 6h). Genes associated with HSC dormancy (Rarb and Epcr) were upregulated and genes associated with cell cycle (Pcna and Cdk6) were downregulated in HA- and NAG-treated mice (Fig. 6i). In contrast, we did not observe any changes in the division or gene expression of HSCs when Gprc5c-KO mice were subjected to the same conditions (Fig. 6h,i). These results suggest that HA and NAG treatment preserves mouse HSC quiescence via Gprc5c.

To test the conservation of HA-regulated quiescence in humans, we monitored the division kinetics of HA- and NAG-treated GPRC5C<sup>pos/neg</sup>-LT-HSCs from BM and CB. We found that treatment with HA and NAG resulted in delayed division in GPRC5C<sup>pos</sup>-LT-HSCs specifically (Fig. 6j and Extended Data Fig. 7d). Higher engraftment potential was observed in HA-treated GPRC5C<sup>pos</sup>-HSCs. Conversely, HA treatment had a limited effect on self-renewal capabilities in GPRC5C<sup>neg</sup>-HSCs (Fig. 6k).

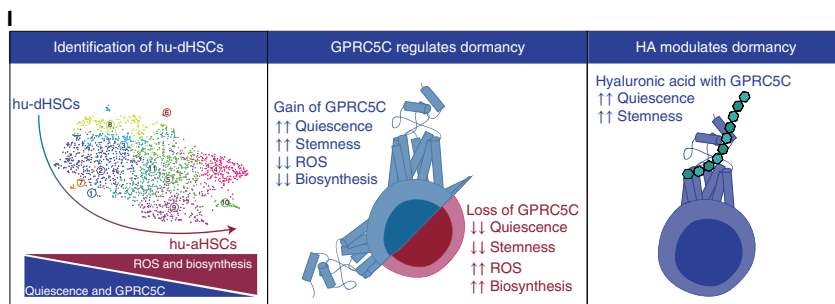
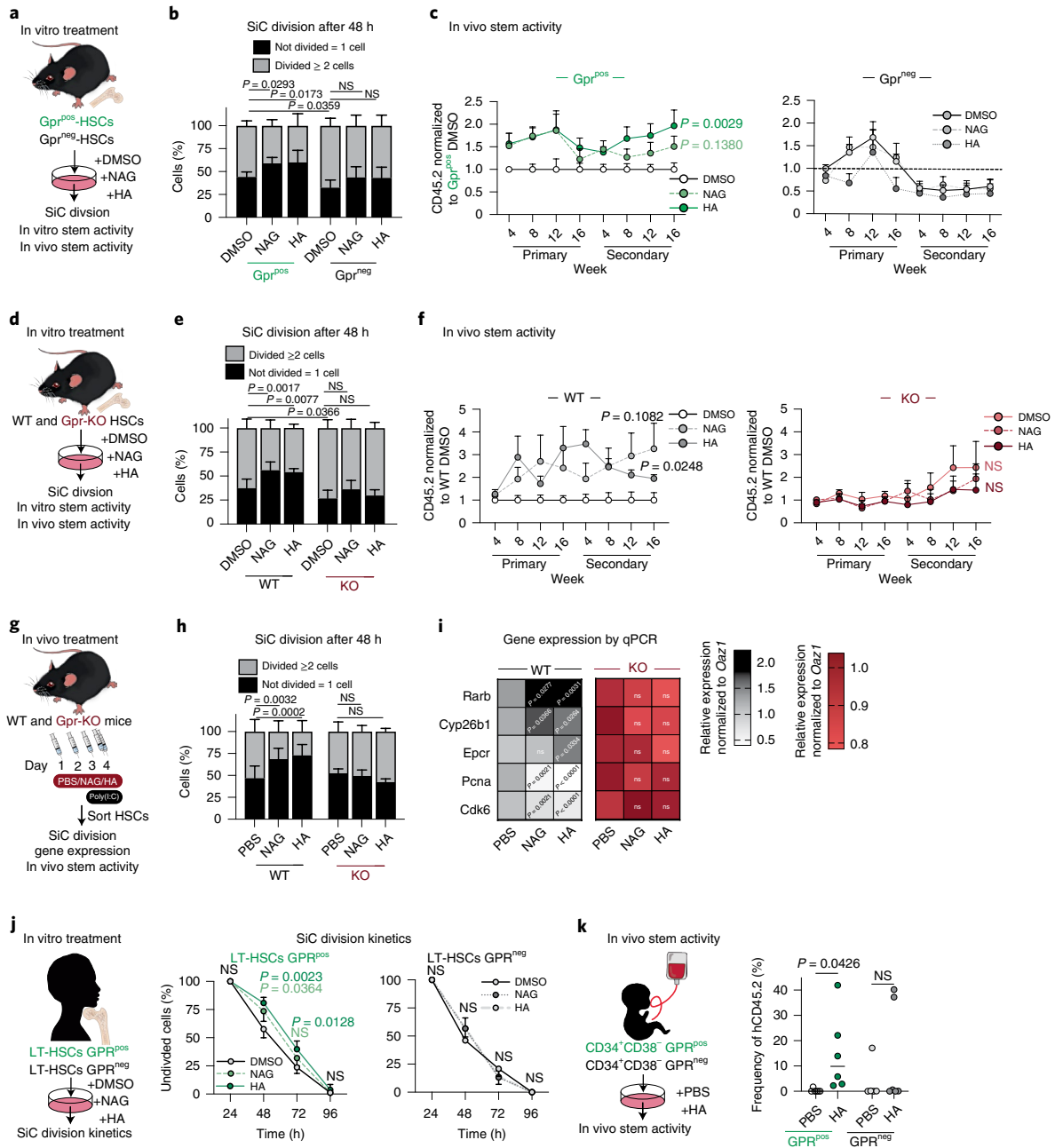
CD44 is the most well-described receptor for HA and has been shown to be critical in supporting mouse stem and progenitor cell homing to the BM niche<sup>46,47</sup>. To dismiss the possibility that HA-mediated quiescence and self-renewal is established via the CD44–HA signalling axis, we examined HA and NAG treatment in

**Fig. 6 | HA–Gprc5c signalling axis regulates human and mouse HSCs.** **a**, Experimental design to assess the effect of in vitro HA and NAG treatment on Gprc5c<sup>pos/neg</sup>-HSCs. **b**, SiC division assay of Gprc5c<sup>pos/neg</sup>-HSCs quantified after 48 h in vitro culture treatment with NAG, HA and DMSO control. *n* = 8 biological replicates. **c**, Transplantation of Gprc5c<sup>pos/neg</sup>-HSCs treated with NAG, HA and DMSO control. *n* = 7 biological replicates, except *n* = 5 biological replicates for Gprc5c<sup>pos</sup>-HSCs treated with NAG. **d**, Experimental design to assess the effect of in vitro HA and NAG treatment on Gprc5c-KO and WT control HSCs. **e**, SiC division assay of Gprc5c-KO and WT control HSCs quantified after 48 h in vitro culture treatment with NAG, HA and DMSO control. *n* = 6 biological replicates. **f**, Transplantation of Gprc5c-KO and WT control HSCs treated with NAG, HA and DMSO control. *n* = 9, 9, 6, 11, 12 and 17 biological replicates for WT-DMSO, WT-NAG, WT-HA, Gprc5c-KO-DMSO, Gprc5c-KO-NAG and Gprc5c-KO-HA, respectively. **g**, Experimental design to assess the effect of in vivo NAG and HA treatment in preventing viral stress from poly I:C. HSCs defined as Lineage<sup>-</sup>cKit<sup>+</sup>Epcr<sup>+</sup>CD150<sup>+</sup>CD48<sup>-</sup>CD34<sup>-</sup>. **h**, SiC division assay of in vivo treated Gprc5c-KO and WT control HSCs cells quantified after 48 h of in vitro culture. *n* = 6 biological replicates. **i**, Differential gene expression. Heat map representing median RNA expression from qPCR data (normalized to housekeeping gene Oaz1 and PBS control) *n* = 11 biological replicates. **j**, SiC division kinetics of BM GPRC5C<sup>pos/neg</sup>-LT-HSCs with in vitro culture treatment with NAG, HA and DMSO control. *n* = 4 biological replicates. **k**, Transplantation of GPRC5C<sup>pos/neg</sup>-HSCs from two CB donors treated with HA and control. *n* = 6 biological replicates. **l**, Representative summary of findings. All data presented as mean ± s.d., except **i** where median is presented. Statistical significance was determined using two-tailed *t*-test (**c**, **f**, **i** and **k**) or two-way ANOVA (**b**, **e**, **h** and **j**). *n* indicates number of biological replicates. For all experiments, at least two independent experiments were performed. Source numerical data are available in source data.



mouse CD44-KO HSCs (Extended Data Fig. 8e). Phenotypic characterization of CD44-KO BM cells revealed no differences in the frequency of HSCs or MPPs (Extended Data Fig. 8f). CD44-KO HSCs showed increased expression of cell cycle-associated genes (Cdk6 and Pcna) but no alteration in expression of dormancy-associated

genes (Gprc5c, *Rarb* and *Cyp26b1*) (Extended Data Fig. 8g). HSCs treated with NAG/HA had delayed primary division and showed enhanced in vitro self-renewal capacity irrespective of whether they were WT or CD44-KO HSCs (Extended Data Fig. 8h,i). These results are consistent with a previous report showing that deletion



of CD44 in adult BM HSCs has no impact on long-term reconstitution<sup>48</sup>, suggesting an alternative mechanism. Additionally, HA interacts with other receptors such as LYVE-1, TLR4 and RHAMM<sup>37,49,50</sup>. Thus, securing definite evidence of the HA–GPCRC5C interaction will only be possible by solving the crystal structure of GPCR5C.

Together, these results suggest that HSC quiescence is regulated by HA through its interaction with GPCR5C.

## Discussion

In this study, we found that human BM HSCs are organized on an axis of dormancy (Fig. 6l). This dormancy axis is characterized by gradual upregulation of processes such as translation, ribosomal biosynthesis, cell cycle and ROS. hu-dHSCs resembled mouse dHSCs, previously described only within the murine haematopoietic system, and quiescent populations within human CB<sup>5,10–13</sup>. We found that the orphan receptor GPCR5C enriches for HSC dormancy and plays an important role in regulating mouse and human HSC features. GPCR5C<sup>pos</sup> HSCs are highly quiescent, show high in vivo reconstitution capacities, and are regulated by the extracellular matrix molecule HA (Fig. 6l).

Class C GPCRs are defined by two unique structural features: first, class C GPCRs contain a large ECD comprising a VFT and a cysteine-rich domain; second, they form constitutive dimers through the VFT<sup>29</sup>. GPCR5C is unique because its protein sequence is much shorter compared with other members of class C. Our in silico modelling revealed that the region of GPCR5C spanning the extracellular matrix contains 90 amino acids, making it structurally difficult to form the VFT module (typically between 500 and 600 amino acids). GPCR5C, along with GPCR5A and GPCR5B, is classified as a class C GPCR owing to sequence homology. However, its topology has revealed that it may be more closely related to class A GPCRs<sup>51</sup>. Given the distinctive differences revealed by the in silico model of GPCR5C compared with other members of class C GPCRs, its reclassification or subclassification may be warranted. Using in silico modelling and mutational cell-based assays, we have proposed a potential model of HA interaction with the 7TM domain of GPCR5C. However, it remains unclear if a mutation in the HA-binding domain of GPCR5C would result in loss of dormancy in human HSCs. Investigation of the functionality of these mutations would require CRISPR–Cas9 technologies, which are still challenging to use in human LT-HSCs owing to the limited effectiveness of long-term cultivation of human HSCs.

Some uncertainty remains regarding the proposed binding mode of HA since the in silico modelling is dependent on existing structures; however, there are few experimentally determined atomic models of the GPCR class C family. To address the precise mechanism by which GPCR5C binds to HA and NAG would require further biophysical assays, such as co-crystallization or cryo-electron microscopy. However, isolation and crystallization of membrane proteins has proven notoriously difficult. Further, GPCRs are often purified in the presence of their orthosteric or allosteric ligands, particularly antagonists and inverse agonists, to maintain their conformationally stabilized form<sup>52</sup>. The orthosteric ligand for GPCR5C is not yet known, and identification would require intensive high-throughput screening. Furthermore, we cannot discount the possibility that other known HA receptors, such as TLR4, may also play a role in regulating dormancy. Nonetheless, our extensive cell-based assays show that GPCR5C is activated by NAG and HA, thus allowing us to identify a signalling axis that regulates HSC dormancy.

In summary, we have found that maintaining a deep state of quiescence, or dormancy, is regulated by the GPCR5C–HA signalling axis in both mouse and human HSCs.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary

information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41556-022-00931-x>.

Received: 29 April 2021; Accepted: 2 May 2022;

Published online: 20 June 2022

## References

- Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
- Till, J. E. & McCulloch, E. A. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res.* **14**, 213–222 (1961).
- Fares, I. et al. EPCR expression marks UM171-expanded CD34<sup>+</sup> cord blood stem cells. *Blood* **129**, 3344–3351 (2017).
- Frobel, J., Rahmig, S., Franzen, J., Waskow, C. & Wagner, W. Epigenetic aging of human hematopoietic cells is not accelerated upon transplantation into mice. *Clin. Epigenetics* **10**, 67 (2018).
- Knapp, D. J. H. F. et al. Single-cell analysis identifies a CD33<sup>+</sup> subset of human cord blood cells with high regenerative potential. *Nat. Cell Biol.* **20**, 710–720 (2018).
- Laurenti, E. et al. CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell* **16**, 302–313 (2015).
- Lehnertz, B. et al. H3K27M/I mutations promote context-dependent transformation in acute myeloid leukemia with RUNX1 alterations. *Blood* **130**, 2204–2214 (2017).
- Takayama, N. et al. The transition from quiescent to activated states in human hematopoietic stem cells is governed by dynamic 3D genome reorganization. *Cell Stem Cell* **28**, 488–501 (2020).
- Xie, S. Z. et al. Sphingolipid modulation activates proteostasis programs to govern human hematopoietic stem cell self-renewal. *Cell Stem Cell* **25**, 639–653.e7 (2019).
- García-Prat, L. et al. TFEB-mediated endolysosomal activity controls human hematopoietic stem cell fate. *Cell Stem Cell* **28**, 1838–1850.e10 (2021).
- Belluschi, S. et al. Myelo-lymphoid lineage restriction occurs in the human haematopoietic stem cell compartment before lymphoid-primed multipotent progenitors. *Nat. Commun.* **9**, 4100 (2018).
- Kaufmann, K. B. et al. A latent subset of human hematopoietic stem cells resists regenerative stress to preserve stemness. *Nat. Immunol.* **22**, 723–734 (2021).
- Knapp, D. J. H. F. et al. A topological view of human CD34(+) cell state trajectories from integrated single-cell output and proteomic data. *Blood* **133**, 927–939 (2019).
- Cesana, M. et al. A CLK3–HMGA2 alternative splicing axis impacts human hematopoietic stem cell molecular identity throughout development. *Cell Stem Cell* **22**, 575–588.e7 (2018).
- Foudi, A. et al. Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nat. Biotechnol.* **27**, 84–90 (2009).
- Takizawa, H., Regoes, R. R., Boddupalli, C. S., Bonhoeffer, S. & Manz, M. G. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J. Exp. Med.* **208**, 273–284 (2011).
- Wilson, A. et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
- Walter, D. et al. Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature* **520**, 549–552 (2015).
- Grün, D. Revealing dynamics of gene expression variability in cell state space. *Nat. Methods* **17**, 45–49 (2020).
- Herman, J. S., Sagar & Grün, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods* **15**, 379–386 (2018).
- Cabezas-Wallscheid, N. et al. Vitamin A-retinoic acid signaling regulates hematopoietic stem cell dormancy. *Cell* **169**, 807–823.e19 (2017).
- Hashimshony, T. et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-seq. *Genome Biol.* **17**, 77 (2016).
- Becker-Herman, S. et al. CD74 is a regulator of hematopoietic stem cell maintenance. *PLoS Biol.* **19**, e3001121 (2021).
- Signer, R. A. J., Magee, J. A., Salic, A. & Morrison, S. J. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature* **509**, 49–54 (2014).
- Sano, T. et al. Comparative characterization of GPCR5B and GPCR5CLacZ knockin mice; behavioral abnormalities in GPCR5B-deficient mice. *Biochem. Biophys. Res. Commun.* **412**, 460–465 (2011).
- Trumpp, A., Essers, M. & Wilson, A. Awakening dormant haematopoietic stem cells. *Nat. Rev. Immunol.* **10**, 201–209 (2010).
- Rajkumar, P. et al. Identifying the localization and exploring a functional role for Gpr5c in the kidney. *FASEB J.* **32**, 2046–2059 (2018).
- Chun, L., Zhang, W.-h & Liu, J.-f Structure and ligand recognition of class C GPCRs. *Acta Pharmacol. Sin.* **33**, 312–323 (2012).

29. Møller, T. C., Moreno-Delgado, D., Pin, J.-P. & Kniazeff, J. Class C G protein-coupled receptors: reviving old couples with new partners. *Biophysical Rep.* **3**, 57–63 (2017).
30. Wu, H. et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* **344**, 58–64 (2014).
31. Zavan, B. et al. Hyaluronic acid induces activation of the  $\kappa$ -opioid receptor. *PLoS ONE* **8**, e55510 (2013).
32. Nilsson, S. K. et al. Hyaluronan is synthesized by primitive hemopoietic cells, participates in their lodgment at the endosteum following transplantation, and is involved in the regulation of their proliferation and differentiation in vitro. *Blood* **101**, 856–862 (2003).
33. Goncharova, V. et al. Hyaluronan expressed by the hematopoietic microenvironment is required for bone marrow hematopoiesis. *J. Biol. Chem.* **287**, 25419–25433 (2012).
34. Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* **3**, 639–650 (2002).
35. Lefkowitz, R. J. & Shenoy, S. K. Transduction of receptor signals by  $\beta$ -arrestins. *Science* **308**, 512–517 (2005).
36. Kroeze, W. K. et al. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat. Struct. Mol. Biol.* **22**, 362–369 (2015).
37. Yang, B., Yang, B. L., Savani, R. C. & Turley, E. A. Identification of a common hyaluronan binding motif in the hyaluronan binding proteins RHAMM, CD44 and link protein. *EMBO J.* **13**, 286–296 (1994).
38. Ziebell, M. R. & Prestwich, G. D. Interactions of peptide mimics of hyaluronic acid with the receptor for hyaluronan mediated motility (RHAMM). *J. Comput. Aided Mol. Des.* **18**, 597–614 (2004).
39. Amemiya, K., Nakatani, T., Saito, A., Suzuki, A. & Munakata, H. Hyaluronan-binding motif identified by panning a random peptide display library. *Biochim. Biophys. Acta Gen. Subj.* **1724**, 94–99 (2005).
40. Zhao, J. et al. Characterization of a motif for specific binding to hyaluronan in chicken SPACR. *J. Neurochem.* **106**, 1117–1124 (2008).
41. Lee, S.-M. et al. Calcitonin receptor N-glycosylation enhances peptide hormone affinity by controlling receptor dynamics. *J. Mol. Biol.* **432**, 1996–201 (2020).
42. Yasuda, D., Imura, Y., Ishii, S., Shimizu, T. & Nakamura, M. The atypical N-glycosylation motif, Asn-Cys-Cys, in human GPR109A is required for normal cell surface expression and intracellular signaling. *FASEB J.* **29**, 2412–2422 (2015).
43. Blom, D., Hirsch, C., Stern, P., Tortorella, D. & Ploegh, H. L. A glycosylated type I membrane protein becomes cytosolic when peptide: N-glycanase is compromised. *EMBO J.* **23**, 650–658 (2004).
44. Duksin, D. & Mahoney, W. C. Relationship of the structure and biological activity of the natural homologues of tunicamycin. *J. Biol. Chem.* **257**, 3105–3109 (1982).
45. Essers, M. A. G. et al. IFN $\alpha$  activates dormant haematopoietic stem cells in vivo. *Nature* **458**, 904–908 (2009).
46. Legras, S. et al. CD44-mediated adhesiveness of human hematopoietic progenitors to hyaluronan is modulated by cytokines. *Blood* **89**, 1905–1914 (1997).
47. Misra, S., Hascall, V. C., Markwald, R. R. & Ghatak, S. Interactions between hyaluronan and its receptors (CD44, RHAMM) regulate the activities of inflammation and cancer. *Front. Immunol.* **6**, 201 (2015).
48. Cao, H. et al. The role of CD44 in fetal and adult hematopoietic stem cell regulation. *Haematologica* **101**, 26–37 (2016).
49. Termeer, C. et al. Oligosaccharides of Hyaluronan activate dendritic cells via toll-like receptor 4. *J. Exp. Med.* **195**, 99–111 (2002).
50. Lawrance, W., Banerji, S., Day, A. J., Bhattacharjee, S. & Jackson, D. G. Binding of hyaluronan to the native lymphatic vessel endothelial receptor LYVE-1 is critically dependent on receptor clustering and hyaluronan organization. *J. Biol. Chem.* **291**, 8014–8030 (2016).
51. Bräuner-Osborne, H. et al. Cloning and characterization of a human orphan family C G-protein coupled receptor GPRC5D1. *Biochim. Biophys. Acta Gene Struct. Expr.* **1518**, 237–248 (2001).
52. Milić, D. & Veprintsev, D. B. Large-scale production and protein engineering of G protein-coupled receptors for structural studies. *Front. Pharm.* **6**, 66 (2015).

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



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

© The Author(s) 2022

## Methods

**Mouse models.** All mice were bred in-house in the animal facility at the Max Planck Institute of Immunobiology and Epigenetics in individually ventilated cages. Mice were killed by cervical dislocation according to German guidelines. Animal procedures were performed according to the protocols approved by the German authorities, Regierungspräsidium Freiburg (the killing of animals for scientific purposes according to §4 (3) of the German Animal Protection Act); animal protocol numbers: G-18/142, G-19/74 and G-20/124.

**Isolation of mouse HSCs.** Murine BM cells were isolated from the femur, tibia, hip bone and vertebrae by gentle crushing with mortar and pestle in PBS. Red blood cell lysis was performed with ACK Lysing Buffer (Thermo Fisher Scientific) for 5 min at room temperature (RT). Dynabeads Untouched Mouse CD4 Cells Kit (Invitrogen) was used for lineage negative enrichment according to the manufacturer's protocol. Briefly, the BM was stained with 1:4 dilution of the Lineage Cocktail for 30–60 min at 4°C on a rotating wheel. Labelled cells were then incubated for 20 min with 400 µl of washed Dynabeads coated with polyclonal sheep anti-rat IgG per sample. Depletion of lineage cells was performed using a magnet. Lineage-depleted BM cells were stained with lineage markers (Gr1, CD11b, B220, Ter119, CD4 and CD8a), ckit, Sca1, CD150, CD48 and CD34 and sorted on an Aria Fusion II for downstream analysis.

**Mouse transplantation.** *HSC transplants.* In total, 100 HSCs (LSK CD150<sup>+</sup>CD48<sup>-</sup>CD34<sup>-</sup>) from Gprc5c KO and WT controls were sorted and transplanted by tail-vein injection into fully irradiated CD45.1 recipients (4.5 + 5 Gy).

*Competitive WBM chimaeras.* Two sets of chimaeras were generated: CD45.1/2 WT + CD45.2 WT and CD45.1/2 WT + CD45.2 Gprc5c KO. In every case, 10<sup>6</sup> and 10<sup>6</sup> WBM cells were mixed. Mixed WBM were transplanted into recipients as described above.

*Reverse transplantation.* A total of 10<sup>6</sup> WBM from WT CD45.1 mice were isolated and transplanted into recipients, either Gprc5c-KO or littermate controls, as described above.

*Secondary transplants.* WBM cells were isolated after 16 weeks post-transplantation, and 10<sup>6</sup> WBM cells were re-transplanted into fully irradiated CD45.1 recipients.

For all transplanted mice, the contribution of donor cells was monitored in PB obtained from the submandibular vein at 4, 8, 12 and 16 weeks post-transplantation.

**Homing.** For analysis of homing capacity, 5 × 10<sup>6</sup> WBM cells (WT or Gprc5c-KO CD45.2) were transplanted into fully irradiated CD45.1 recipients (4.5 + 5 Gy) and killed 48 h later. CD45.1/2 staining was performed, and engraftment was analysed by flow cytometry analysis.

**5-FU injections.** Once a week, WT and Gprc5c-KO mice were intraperitoneally injected with 150 mg kg<sup>-1</sup> 5-FU (Sigma) dissolved in PBS, with a final concentration of 20 mg ml<sup>-1</sup>. Body weight was checked weekly. Mice were killed when 20% of their initial weight was lost.

**NAG and HA in vivo treatments.** WT and Gprc5c-KO mice were intraperitoneally injected with either PBS, 100 mg kg<sup>-1</sup> NAG or 60 mg kg<sup>-1</sup> HA for four consecutive days. Additionally, 50 µg of polyI:polyC (poly I:C) was injected intraperitoneally on day 4. On day 5, mice were killed and BM cells were isolated for further analysis.

**NBSGW transplantations.** *Endogenous GPRC5C.* In total, 7,000 BM or 800 CB GPRC5C<sup>pos/neg</sup>-CD34<sup>+</sup>CD38<sup>-</sup> cells were sorted into PBS. Before transplantation, the number of cells was validated by haemocytometer and then injected intravenously into 7- to 10-week-old female NBSGW mice that were irradiated with 1.5 Gy. Engraftment was monitored by PB obtained from the submandibular vein. After 24 or 36 weeks, mice were killed and BMs (hip bone, tibia and femur) were analysed for human chimaerism.

*GPRC5C-OE.* A total of 10<sup>4</sup> CD34<sup>+</sup> CB cells were transduced after 24 h of pre-stimulation. The transduction efficiency was assessed by flow cytometry after 2 days of cultivation. Total culture was injected intravenously into 7- to 10-week-old female NBSGW mice as stated above. After 24 weeks, mice were killed and BM was analysed for human chimaerism along with GFP. Primary engraftment was calculated by total hCD45<sup>+</sup>GFP<sup>+</sup> engraftment normalized to the initial transduction efficiency of GFP.

For purification of human cells from xenotransplanted mice, BM from individual mice was enriched for human cells with the Mouse Cell Depletion Kit (Miltenyi Biotec) according to the manufacturer's protocol, and 140,000 GFP<sup>+</sup> cells were sorted and re-transplanted. A mouse was considered engrafted if the percentage of hCD45<sup>+</sup> was greater than or equal to 0.1%.

**Cell cycle analysis (Ki-67 or CDK6).** Stained cells were fixed with BD Cytofix/Cytoperm Buffer (BD Biosciences) for 10 min at 4°C. Next, cells were stained with intracellular Ki-67 at 1:100 (BD Biosciences) or CDK6 at 1:500 (Abcam) in PermWash solution (BD Biosciences) for 45–90 min at 4°C. Before cell cycle analysis on the BD LSRFortessa Cell Analyzer (BD Biosciences) using BD FACSDiva v8.0.3, cells were stained in the dark with Hoechst 33342/DAPI (Invitrogen) for 30 min at RT.

**Immunofluorescence quantification of CDK6.** GPRC5C<sup>pos/neg</sup>-CD34<sup>+</sup>CD38<sup>-</sup> were sorted by the FACSaria Fusion II flow cytometer (BD Biosciences) directly on polylysine-coated slides and were fixed over 10 min with BD Cytofix/Cytoperm Buffer (BD Biosciences) at RT. Fixed cells were permeabilized with 0.2% Triton (Sigma) for 10 min and blocked for 20 min with 10% goat serum (Life Technologies). Cells were stained with 1:500 of CDK6 Alexa Fluor 647 (Abcam) for 1 h and washed with PBS. Slides were visualized on the LSM 780 confocal microscope (Zeiss) using ZEN blue v2.5, and fluorescence quantification was performed with ImageJ.

**CellROX/MitoTracker staining.** Cells were incubated at 37°C with CellROX DeepRed at 1/500 (Invitrogen) or MitoTracker Red at 50 nM (Invitrogen) in their corresponding media for 30 min. Cells were subsequently washed three times in PBS and stained for FACS analysis on the BD LSRFortessa Cell Analyzer (BD Biosciences).

**O-propargyl-puromycin.** Stained cells were fixed for 15 min at RT with BD Cytofix/Cytoperm Buffer (BD Biosciences) and permeabilized with 0.1% Triton for 15 min at RT. The copper-catalysed azide-alkyne cycloaddition was performed using an Alexa594-azide (Life Technologies, 5 mM final concentration) and the Click-iT Cell Reaction Buffer Kit (Life Technologies) according to the manufacturer's instructions.

**Single-cell differentiation assay.** Anonymized BM CD34<sup>+</sup> samples were purchased from Lonza Group Ltd with informed consent from healthy donors. BM samples were stored in accordance with regulated procedures and the Human Tissue Act, and with protocols approved by the relevant research and ethics committees (covered in Cambridge by the Cambridge Blood and Stem Cell Biobank, IRAS ref. 149581).

To isolate cell populations from BM CD34<sup>+</sup> cells, frozen samples were thawed by drop-wise addition of pre-warmed Iscove's modified Dulbecco's medium (IMDM; Life Technologies) + 0.1 mg ml<sup>-1</sup> DNase (Sigma) + 50% foetal bovine serum (FBS, Life Technologies). Cells were counted before re-suspension in an antibody mix with a panel of cell-surface markers (5 × 10<sup>6</sup> cells ml<sup>-1</sup> in Stemspan antibody mix) (CD34, CD38, CD45RA, CD90, CD49f, Zombie, CD19, biotinylated hyaluronan and streptavidin) for 30 min at 4°C, followed by washing with 1 ml of 3% PBS/FCS. BM LT-HSCs were sorted using single-cell purity and index sorting to allow retrospective correlation between cell-surface marker expression from each single cell with in vitro functional lineage output. GPRC5C<sup>pos/neg</sup> BM LT-HSCs were determined retrospectively by index analysis after single-cell sorting.

MS5 stromal cells were imported from Prof. Katsuhiko Itoh at Kyoto University and were tested to support myeloid (My), erythroid (Ery), megakaryocytic (Meg) and natural killer (NK) differentiation with appropriate control populations. MS5 cells were plated at passage 10–13 at 3,000 cells per well in flat 96-well plates in 100 µl Myelocult H5100 medium (StemCell Technologies) + 1% Pen/Strep (Life Technologies) 1 day before sorting BM LT-HSCs. On the day of the sort, the medium was changed to 100 µl per well MEM cytokine medium to support the growth of My, Ery, Meg and NK colonies. MEM medium composition included StemPro base medium supplemented with nutrients (0.035%) (Life Technologies) and the following cytokines: SCF 100 ng ml<sup>-1</sup>, Flt3-L 20 ng ml<sup>-1</sup>, TPO 100 ng ml<sup>-1</sup>, IL-6 50 ng ml<sup>-1</sup>, IL-3 10 ng ml<sup>-1</sup>, IL-11 50 ng ml<sup>-1</sup>, GM-CSF 20 ng ml<sup>-1</sup>, IL-2 10 ng ml<sup>-1</sup> and IL-7 20 ng ml<sup>-1</sup> (all Miltenyi Biotec); erythropoietin (EPO) 3 units ml<sup>-1</sup> (Eprex, Janssen-Cilag); h-LDL 50 ng ml<sup>-1</sup> (StemCell Technologies); 1% L-glutamine (Life Technologies); and 1% penicillin-streptomycin (Life Technologies). Indicated populations were index-sorted as single cells (one cell per well) and cultured for 3 weeks at 37°C and 5% CO<sub>2</sub>.

**PLA.** Each sample of 80,000 cells was rested on diagnostic microscope slides (Thermo Fisher Scientific) at 37°C for 45 min. Cells were fixed with 4% paraformaldehyde for 20 min at RT and blocked for 1 h. Blocked cells were stained with the Duolink kit (Sigma Aldrich) according to the manufacturer's instructions. The antibody combination used was mouse anti-GPRC5C (R&D systems) and rabbit anti-HA (LS Bio). Nuclei were stained with DAPI (Roth). Images were taken at 60× magnification with a confocal microscope (Nikon C2) and analysed with BlobFinder.

**Single-cell differentiation assay analysis.** All single-cell-derived colonies were collected into 96 U-bottom plates using a plate filter to prevent carryover of MS5 cells. Cells were then stained with 50 µl per well of antibody mix (CD66, CD15, CD14, CD56, CD71, CD45, GlyA, Cd41a and CD11b), incubated for 20 min in the dark at RT, and then washed with 3% PBS/FBS (100 µl per well). The lineage

composition and the size of colonies formed was assessed by high-throughput flow cytometry. A single cell was defined as giving rise to a true colony if the sum of cells detected in the CD45<sup>+</sup> and GlyA<sup>+</sup> gates was  $\geq 30$  cells and the colony output was determined. Ery colonies were identified as CD45<sup>+</sup>GlyA<sup>+</sup>  $\geq 30$  cells, Meg colonies as CD45<sup>+</sup>CD41<sup>+</sup>CD15<sup>+</sup>CD14<sup>-</sup>  $\geq 20$  cells, My colonies as [(CD45<sup>+</sup>CD14<sup>+</sup>) + (CD45<sup>+</sup>CD15<sup>+</sup>)]  $\geq 30$  cells and NK colonies as CD45<sup>+</sup>CD56<sup>+</sup>  $\geq 30$  cells. If a true colony was found but did not meet any threshold for lineage determination, the colony was designated as 'undifferentiated'. Colony analysis was performed in RStudio v1.25.

**Single-cell division.** Single cells were FACS-sorted into individual wells on 72-well Terasaki plates (Greiner Bio-One) with 10  $\mu$ l of StemPro or Stemspan with complete cytokines. Cells were evaluated for division under a light microscope 48 or 72 h after sorting.

**Primary cell culture.** Human CD34<sup>+</sup> cells were cultured in HSC medium, consisting of StemSpan ACF (StemCell Technologies) supplemented with 100 ng ml<sup>-1</sup> human stem cell factor (PeproTech), 100 ng ml<sup>-1</sup> FMS-like tyrosine kinase 3 ligand (PeproTech), 50 ng ml<sup>-1</sup> thrombopoietin (PeproTech), 10  $\mu$ g ml<sup>-1</sup> human low-density lipoprotein (StemCell Technologies) and 1 $\times$  GlutaMax (Thermo Fisher Scientific) with or without UM171 at 35 nM.

Primary murine haematopoietic cells were cultured in StemPro-34 SFM (Life Technologies) with supplement and 50 ng ml<sup>-1</sup> SCF, 25 ng ml<sup>-1</sup> TPO, 30 ng ml<sup>-1</sup> Flt3-Ligand (all PeproTech), 100 U ml<sup>-1</sup> penicillin–streptomycin and 2 mM L-glutamine (Thermo Fisher).

**CFU assays.** *Mouse CFU.* In total, 200–400 HSCs were either FACS-sorted directly into 1 ml of MethoCult M3434 (StemCell Technologies) or into 384-well plates containing 100  $\mu$ l Complete Stem Cell Media was supplemented with HA (250–500  $\mu$ g ml<sup>-1</sup>), NAG (20  $\mu$ M) or DMSO control. After 48–72 h of in vitro treatment, total cultured HSCs were collected into 1 ml of MethoCult M3434 (StemCell Technologies). Approximately 5 days after primary plating, the number of colonies was quantified. Then, 10<sup>4</sup> cells were used for secondary, tertiary and quaternary plating. Approximately 5 days after each replating, colonies were quantified and replated.

*Human CFU.* A total of 10<sup>5</sup> GFP<sup>+</sup> cells were sorted into 1 ml of MethoCult H4435 (StemCell Technologies). Seven days after primary plating, the colony numbers were quantified. After 14 days, 10<sup>4</sup> cells were replated and quantified.

**HA–fluorescein binding assay.** HEK293T cells were transfected with 30  $\mu$ g ml<sup>-1</sup> of the GPRC5C-Tango plasmids using JetPrime transfection reagent. After 24–36 h transfection, 60,000–80,000 cells were treated with 8  $\mu$ g HA conjugated to fluorescein with or without 100  $\mu$ M NAG for 60 min at RT with shaking. Cells were washed, and fluorescein was quantified using the LSR II flow cytometer (BD Bioscience).

**qRT–PCR analysis.** For real-time quantitative PCR (qPCR), the total RNA of 1,000–2,000 cells was isolated with the Arcturus PicoPure RNA isolation kit (Applied Biosystems) according to the manufacturer's guidelines. First-strand synthesis was reverse transcribed using the SuperScript VILO cDNA Synthesis Kit (Invitrogen) according to the manufacturer's guidelines. For qPCR analysis, Fast SYBR Green Master Mix or TaqMan probes (Thermo Fisher) with TaqMan Fast Advanced Master Mix (Thermo Fisher) were used on a StepOne Real-Time PCR System (Applied Biosystems) or QuantStudio Flex 6 Real-Time PCR System (Applied Biosystems). RNA expression was normalized to expression of *Oaz1* or *GAPDH* housekeeping genes and presented as relative quantification (ratio 2<sup>- $\Delta\Delta$ CT</sup>).

**Virus production and primary cell infection.** Lentiviral particles were produced in HEK293T cells. In each 150 mm tissue culture dish, 15  $\mu$ g of the lentiviral vector was transiently packaged with 9  $\mu$ g psPAX2 and 6  $\mu$ g pMD2.G with JetPrime (Polyplus). Viral supernatant was collected at 48 h post-transfection. Virus concentration was achieved by ultracentrifugation at 30,000 rpm through a 20% sucrose cushion for 120 min at 4°C. For infection of CD34<sup>+</sup> cells, each well of non-tissue-culture-treated 96-well plates (Corning) was coated with 6.4  $\mu$ g RetroNectin reagent (Takara Clontech) in 100  $\mu$ l PBS and incubated overnight at 4°C. The coated wells were washed and blocked with PBS, containing 2% BSA, for 30 min at RT. Viral particles (multiplicity of infection 50 for OE and 100 for KD) were added to the RetroNectin-coated wells and centrifuged at 2,000g for 2 h at 32°C. The supernatant was removed, and wells were washed once with PBS containing 2% BSA. Then, 0.5–1 $\times$ 10<sup>4</sup> CD34<sup>+</sup> cells (24 h pre-stimulated in HSC medium with UM171) were placed into lentivirus-preconditioned wells. At 16 h post-transduction, cells were washed from the RetroNectin with HSC medium and placed onto a fresh plate.

**Construction of lentiviral shGPRC5C.** Short hairpin RNA vectors targeting GPRC5C were designed according to Pelosoff et al.<sup>53</sup>. Lentiviral cloning was performed according to the protocol described in Fares et al.<sup>1</sup>.

Template from Pelosoff et al.:

*shGPRC5C.804.* TGCTGTTGACAGTGAGCGCGTGGATCGTCATGTATACTT-ATAGTGAAGCCACAGATGTATAAGTATACATGACGATCCACATGCCTACTGCCTCGGA

*shGPRC5C.528.* TGCTGTTGACAGTGAGCGCAGAGGTCATCATCAATACAGATAGTGAAGCCACAGATGTATCTGTATTGATGATGACCTCTATGCCTACTGCCTCGGA

**Calcium flux assay.** AML3 cells were infected with viral particles to overexpress GPRC5C or control. Cells were resuspended in IMDM with 1% FBS at 10<sup>6</sup> cells ml<sup>-1</sup>. Per millilitre of cell suspension, 7.5  $\mu$ l Indo-1 (Invitrogen) was added and incubated at 37°C for 45 min. Following incubation, cells were washed twice with 1% FBS IMDM and then incubated for 30 min at 37°C for de-esterification. The baseline of calcium was acquired on the LSR Fortessa II (BD Bioscience) for 30 s with the sample heater at 37°C. The cells were stimulated with compounds and acquired for an additional 90 s. The ratio of indo-1-bound to indo-1-unbound fluorescence emission signals was calculated as a quantitative parameter for change in intracellular calcium levels. All values were normalized to the baseline of each sample.

**Isolation of human CD34<sup>+</sup> cells.** All experiments were performed according to the ethical protocols outlined in Bonn (AZ 264/18) and Mannheim (AZ 2020-639N).

CD34<sup>+</sup> cells were isolated using EasySep Human Cord Blood CD34 Positive Selection Kit II (StemCell Technologies). Cell pellets were collected and resuspended in buffer (PBS, 2 mM EDTA, 2% FBS and 1  $\mu$ l ml<sup>-1</sup> DNase) at 5  $\times$  10<sup>8</sup> cells ml<sup>-1</sup>. Immunomagnetic labelling and separation were performed according to the manufacturer's instructions. Isolated cells were immediately stained with CD34/CD38 and sorted for CD34<sup>+</sup>CD38<sup>-</sup> for downstream experiments.

**PRESTO-Tango assay.** PRESTO-Tango assay was performed according to the protocol in Kroeze et al.<sup>36</sup>, with slight modifications. Briefly, HTLA cells were cultured in 10% FBS DMEM with 2  $\mu$ g ml<sup>-1</sup> puromycin and 100  $\mu$ g ml<sup>-1</sup> hygromycin B. Cells were transfected with 30  $\mu$ g ml<sup>-1</sup> of the GPRC5C-Tango plasmids using JetPrime. After 24 h of transfection, cells were plated at 60,000–80,000 cells per well into a 96-well plate. Then, 3.5 $\times$  concentration of the drug was prepared in 20 mM HEPES and 1 $\times$  HBSS, pH 7.4, and 40  $\mu$ l of the drug population was added to each well. After 24 h of incubation, cells were washed with PBS and transferred into a 96-well plate (Greiner) with 40  $\mu$ l of Bright-Glo solution (Promega). Following a 5-min incubation at RT, each well was measured for 5 s on a Centro LB 963 Microplate Luminometer (Berthold).

**In silico modelling.** First, the three-dimensional coordinates of the 7TM regions of the protein were generated using homology modelling. In this approach, the sequences of the 7TM regions were retrieved from UniProtKB (ID: Q9NQ84-2), and suitable three-dimensional templates were identified using blastp. The three-dimensional structure of the 7TM was generated using Prime (Schrödinger Suite 2020-3). In the next step, we used the iterative threading assembly refinement (I-TASSER) server with an *ab initio* modelling protocol to generate the remaining residues of the ECD (amino acids 1–50) and the cytoplasmic domain (amino acids 301–441) (refs. <sup>54,55</sup>). The best-ranked *ab initio* models were selected on the basis of the in-built ranking method and were added to the 7TM model using Modeller (Webb and Sali, 2016). The highest-ranked model was selected and energy-minimized in Prime using the OPLS3e force field (Schrödinger Suite 2020-3).

In the next step, the full-length model was subjected to 1  $\mu$ s molecular dynamics simulations using the Desmond Molecular Dynamics System (Schrödinger Suite 2020-3). The system was built by adding highly aligned palmitoyl-oleoyl-phosphatidyl-choline bilayers to the full-length GPRC5C, which was placed in an orthorhombic water cell while applying the SPC solvent model and the OPLS3e force field. The overall system was neutralized by adding an appropriate number of Cl<sup>-</sup> ions and was subjected to equilibration using the default Desmond relaxation protocol. The final state was subjected to 1  $\mu$ s molecular dynamics in an *NpT* (constant temperature and constant pressure) ensemble at 300 K per 1 atm. The Nosé–Hoover chain and Martyna–Tobias–Klein with 1.0 ps and 2.0 ps relaxation time, respectively, and isotropic coupling were used. Coulomb interactions were evaluated using the smooth particle-mesh Ewald summation as a long-range method with 9 Å short-range cut-off. The representative structures were extracted from the overall timescale based on protein stability and subjected to further evaluation.

A representative structure was then selected from the overall trajectory based on the protein stability and subjected to extra-precision (XP) Glide docking (Schrödinger Suite 2020-3) against HA using the default parameters. The best binding pose was selected on the basis of the highest XP Glide score. The images were created using PyMol (Schrödinger Suite 2020-3).

**scRNA amplification and library preparation.** scRNA-seq was performed according to the mCEL-Seq2 protocol<sup>20</sup>. CD34<sup>+</sup>CD38<sup>-</sup>GPRC5C<sup>pos/NEG</sup> cells were sorted into 384-well plates containing 240 nl of unique primer mix and 1.2  $\mu$ l mineral oil (Sigma-Aldrich). Sorted plates were centrifuged at 2,200g for 5 min

at 4°C and stored at -80°C. Then, 160 nl reverse transcription reaction mix and 2.2 µl second-strand reaction mix (New England Biolabs) was added to each well to convert RNA into complementary DNA. cDNA from each plate was pooled and transferred to DNA LoBind microcentrifuge tubes (Eppendorf). AMPure/RNAClean XP beads (Beckman Coulter) (0.8× of total volume) were used for cDNA clean-up. In vitro transcription was done with the Transcription Kit (Thermo Fisher) at 37°C for 13 h. To remove primers, samples were incubated with ExoSP-IT (Thermo Fisher) at 94°C for 3 min, and then a final RNA clean-up was performed with 0.8× RNAClean XP beads. Libraries were sequenced on a HiSeq 3000 (Illumina) and NovaSeq 6000 sequencing system (Illumina) (paired-end multiplexing run, high-output mode) at a depth of ~150,000–200,000 reads per cell.

LT-HSCs scRNA-seq was performed on the 10x Genomics platform using the Chromium Single Cell 3' Reagent Kit V3 (10x Genomics) according to the manufacturer's instructions. Briefly, 4,000 LT-HSCs cells were sorted and pooled from two CD34<sup>+</sup> BM specimens (StemCell Technologies) into 0.4% BSA in PBS. Cells were loaded according to the manufacturer's instructions, aiming for a targeted cell recovery of 2,000–3,000 cells. The quality of the obtained cDNA library upon adapter ligation and index PCR (13 cycles) was assessed by Bioanalyzer fragment analysis (High Sensitivity DNA Kit, Agilent). Libraries were sequenced on HiSeq 2500 and 3000 sequencing systems (Illumina) (paired-end multiplexing run, high output mode) at a depth of ~500,000 reads per cell.

**Bulk RNA-seq.** The RNA of 1,000–3,000 freshly sorted cells was extracted using Arcturus PicoPure kit (Applied Biosystems) and DNase treatment was performed (Qiagen). Subsequently, cDNA libraries were prepared using SMART-Seq v4 Ultra Low Input RNA kit (Takara Bio) with 12 cycles of amplification. Further, NEBNext Ultra II FS DNA kit and NEBNext Multiplex Oligos (New England BioLabs) were used to generate uniquely dual-barcoded sequencing libraries from cDNA libraries. To this end, 5 ng of cDNA library was fragmented for 22.5 min, adaptors were ligated and libraries were amplified for eight cycles. RNA-seq libraries were sequenced at a depth of 55 million reads, 100 bp paired end on a NovaSeq platform (Illumina).

**Bulk RNA-seq analysis method: low-level processing.** Raw fastq files were mapped against the hg38 or the mm10 reference genome by the mRNA-seq tool from the bioinformatics pipeline snakePipes<sup>56</sup>. 'Alignment' mode was used for mapping of sequenced reads using STAR (version STAR\_2.7.4a)<sup>57</sup>, followed by the quantification of expression counts by featureCounts<sup>58</sup>. DeepTools QC (version 3.3.2) (ref. <sup>59</sup>) was used for quality checking. Genes with an average expression that was higher than 100 counts among all samples were selected for further analysis. The differential expression was generated with DESeq2 (ref. <sup>60</sup>); results with a false discovery rate <0.1 and a log<sub>2</sub> fold change (FC) threshold ≥0.5 were considered significant.

**Bulk RNA-seq analysis method: downstream analysis.** The relative expression of differentially expressed genes (DEGs) based on the variance-stabilized read counts as calculated by DESeq2 package was visualized with principal component analysis and heat maps. Volcano plots representing the results of differential expression were generated with EnhancedVolcano. GSEA was performed with the fgsea R package<sup>61</sup> for evaluation of expression in pairwise comparisons of previously published HSCs activation/quiescence–dormancy gene signatures<sup>11,21</sup>, the hallmark gene set (h.all.v7.2.symbols.gmt<sup>62</sup>), the KEGG pathway database gene set (c2.cp.kegg.v7.2.symbols.gmt<sup>63</sup>) and the GO Biological Processes (c5.go.bp.v7.2.symbols.gmt) obtained from the Molecular Signatures Database<sup>64</sup>, considering significant pathways at a false discovery rate <0.1. The GSEA enrichment profile of concrete signatures was plotted using gseaplot2 function.

**scRNA-seq analysis: quantification of transcript abundance.** Paired-end reads from the mCEL-Seq2 and 10x human HSC data were aligned to the same reference genome (GRCh38) using the snakePipes scRNA-seq workflow (version 2.2.3) in Gruen mode and in STARsolo mode, respectively<sup>56</sup>. Gencode version 31 gtf was used for feature counting.

**Genotyping of patients in the 10x dataset and splitting cells by patient.** To assign cells to the corresponding patient, genotyping procedure was used, as specified in Xu et al.<sup>65</sup>. Briefly, bam files produced by STARsolo were filtered for high-quality reads using samtools view parameters q 10 F 3844. Reads were assigned to genes with featureCounts using bam format as output. Gene-assigned bam files were de-duplicated per gene and per cell with umitools. Genetic variants were called on de-duplicated bams with freeBayes (with parameters -iXu -C 2 -q 1), and high-quality variants (QUAL >30) were filtered for with vcfFilter. scSplit<sup>65</sup> was run on the filtered vcf file, requesting two genotypes and setting a ceiling on the expected fraction of doublets at 20% (with parameters -n 2 -d 0.2).

Seurat object output by snakePipes (mode: STARsolo) was annotated with genotypes obtained from scSplit and split into three objects, containing counts for cells assigned to either of the two patients in the cell mixture, or cells predicted to be doublets. Unnormalized count matrices for the two patients were used in further analysis, while the doublet cells were discarded.

**scRNA sequencing data analysis.** Overall, 1,827 cells passed the quality control threshold of >1,000 transcripts for the 10x Genomics human LT-HSC data. For normalization, the total transcript counts in each cell were normalized to 1, and were multiplied by the minimum total transcript count across all cells that passed the quality control threshold (>1,000 transcripts per cell). VarID<sup>19</sup> was run with a vector of batch variables indicating the batches/patients, and the following parameters: mintotal = 1,000, minexpr = 5, minnumber = 5, large = TRUE, regNB = TRUE. For the mCEL-Seq2 datasets, we recovered 639 cells (mintotal = 1,000). VarID was run with the default parameters.

**Cell ordering and generation of SOMs.** Cells were ordered in ascending order based on the coordinates of the first UMAP dimension, that is the cells' x-axis coordinates from the UMAP. SOMs were generated using the FateID package based on the ordering inferred from the UMAP<sup>20</sup>. Only genes with more than two counts after size normalization in at least a single cell were included for the SOM analysis. In brief, smooth profiles were derived by applying local regression on normalized transcript counts after ordering cells. Next, a one-dimensional SOM with 200 nodes was computed on these profiles after z-transformation.

**Pathway enrichment analysis and GSEA.** Symbol gene IDs were first converted to Entrez gene IDs using the clusterProfiler package<sup>66</sup>. Pathway enrichment analysis and GSEA<sup>67,68</sup> were implemented using the ReactomePA package. Pathway enrichment analysis was performed on genes taken from the different modules in the SOMs. GSEA was performed using the DEGs between clusters 3, 2 and 8 and clusters 9, 4, 5 and 10 inferred by the diffeXpn function from the RaceID package.

**Human quiescent CB and dormant mouse HSC signatures.** Quiescent CB genes were selected from the list of DEGs between quiescent and active HSCs from the bulk RNA-seq data generated in Belluschi et al.<sup>11</sup>. Genes with a log<sub>2</sub>FC >0.5 were selected for the quiescent signature, and their expression in the human 10x dataset was aggregated and plotted.

Dormant mouse HSC genes were selected from the list of DEGs between dormant and active mouse HSCs from the bulk RNA-seq data generated in Cabezas-Wallscheid et al.<sup>21</sup>. Human–mouse orthologous genes with a log<sub>2</sub>FC >1 were selected for the dHSC signature, and their expression in the human 10x dataset was aggregated and plotted.

**Statistics and reproducibility.** Two-tailed *t*-tests were used to compare between groups, and the Mantel–Cox test was used to determine statistical significance of the survival curve. At least two independent experiments were performed for all experiments, unless otherwise stated. Data are shown as mean ± standard deviation (s.d.) or mean ± standard error of the mean (s.e.m.) as indicated. GraphPad Prism was used for statistical analysis.

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

## Data availability

The raw data are deposited in ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) and are available under the accession numbers E-MTAB-9862, E-MTAB-9863, E-MTAB-9874, E-MTAB-9892, E-MTAB-9922 and E-MTAB-9967. Source data are provided with this paper. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

## References

- Peloso, R. et al. Prediction of potent shRNAs with a sequential classification algorithm. *Nat. Biotechnol.* **35**, 350–353 (2017).
- Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* <https://doi.org/10.1186/1471-2105-9-40> (2008).
- Zhang, J., Yang, J., Jang, R. & Zhang, Y. GPCR-I-TASSER: a hybrid approach to G protein-coupled receptor structure modeling and the application to the human genome. *Structure* **23**, 1538–1549 (2015).
- Bhardwaj, V. et al. snakePipes: facilitating flexible, scalable and integrative epigenomic analysis. *Bioinformatics* **35**, 4757–4759 (2019).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
- Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- Sergushichev, A. A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. Preprint at *bioRxiv* <https://doi.org/10.1101/060012> (2016).
- Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

63. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
64. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
65. Xu, J. et al. Genotype-free demultiplexing of pooled single-cell RNA-seq. *Genome Biol.* **20**, 290 (2019).
66. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
67. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
68. Mootha, V. K. et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).

## Acknowledgements

We thank W. Kroeze and M. Harms for sharing the HTLA cell line; S. Hobitz, K. Schuldes, L. Bischer and A. Würch from the Max Planck Institute of Immunology and Epigenetics (MPI-IE) Flow Cytometry Core Facility for their assistance; and D. Ruf, K. Eichmann, C. Zedlick, H. Kwiatkowski and all members of the MPI-IE Laboratory Animal Core Facility for excellent animal welfare and husbandry. We thank the MPI-IE Deep Sequencing Core Facility for their assistance. We thank F. A. Schildberg from the University Hospital Bonn, D. Wider from the University Medical Center Freiburg and B. Hauptrock from the Medical Center of Mainz for supportive logistics. This work was supported by the Max Planck Society, the ERC-Stg-2017 (VitASTEM, 759206), the Behrens-Weise-Foundation, the German Research Foundation (DFG) under the German Excellence Strategy (CIBSS-EXC-2189, project ID 390939984), SFB1425 (Project #422681845), SFB992 (Project #192904750; B07), SFB1479 (P05), the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Actions Grant (agreement 813091) and the José Carreras Leukämie-Stiftung all to N.C.-W. D.G. was supported by the DFG under Germany's Excellence Strategy (CIBSS-EXC-2189, project ID 390939984), and by the ERC-2018-COG (ImmuNiche-818846). E.L. is supported by a Sir Henry Dale Fellowship from Wellcome and the Royal Society (107630/Z/15/Z). Research in E.L.'s laboratory is supported by Wellcome (203151/Z/16/Z), BBSRC (BB/P002293/1), EHA, BIRAX,

the Royal Society and core support grants from Wellcome and the MRC to the Wellcome-MRC Cambridge Stem Cell Institute (203151/Z/16/Z). C.J. is supported by an MRC iCASE PhD studentship. For the purposes of Open Access, E.L. has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. S.M. thanks CIBSS-EXC-2189 and SFB1425. P.M. acknowledges CERCA/Generalitat de Catalunya and Fundació Josep Carreras-Obra Social la Caixa for institutional support and the Spanish Ministry of Economy and Competitiveness (PID2019-108160RB-I00) and José Carreras Leukämie-Stiftung. T.V.-H. was supported by a Marie Skłodowska-Curie Fellowship (grant agreement 792923).

## Author contributions

Conceptualization: Y.W.Z. and N.C.-W.; methodology: Y.W.Z., J.M., N.A., P.M., C.J., M.C.R.-M., J.R., K. Schönberger, N.O., K.J., N.M.W., M.-E.L., T.V.-H., K. Sikora, R.W., B.L., G.S., T.M., P.M., S.G.W., S.M., E.L., S.G., D.G. and N.C.-W.; investigation: Y.W.Z. and J.M.; writing—original draft: Y.W.Z., N.C.-W., P.M. and S.G.; supervision: N.C.-W. All authors participated in editing the manuscript.

## Funding

Open access funding provided by Max Planck Society.

## Competing interests

The authors declare no competing interests.

## Additional information

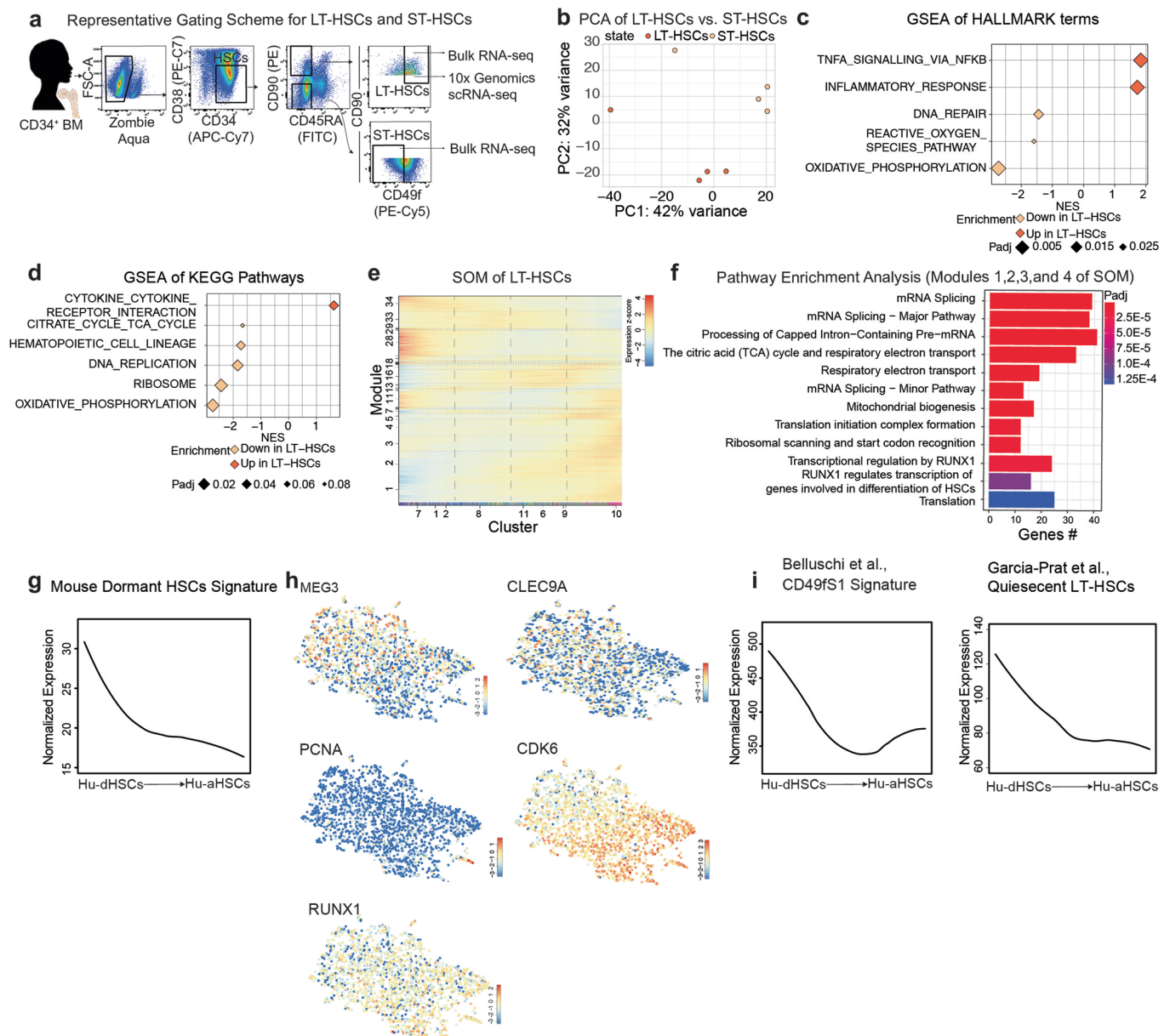
**Extended data** is available for this paper at <https://doi.org/10.1038/s41556-022-00931-x>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41556-022-00931-x>.

**Correspondence and requests for materials** should be addressed to Nina Cabezas-Wallscheid.

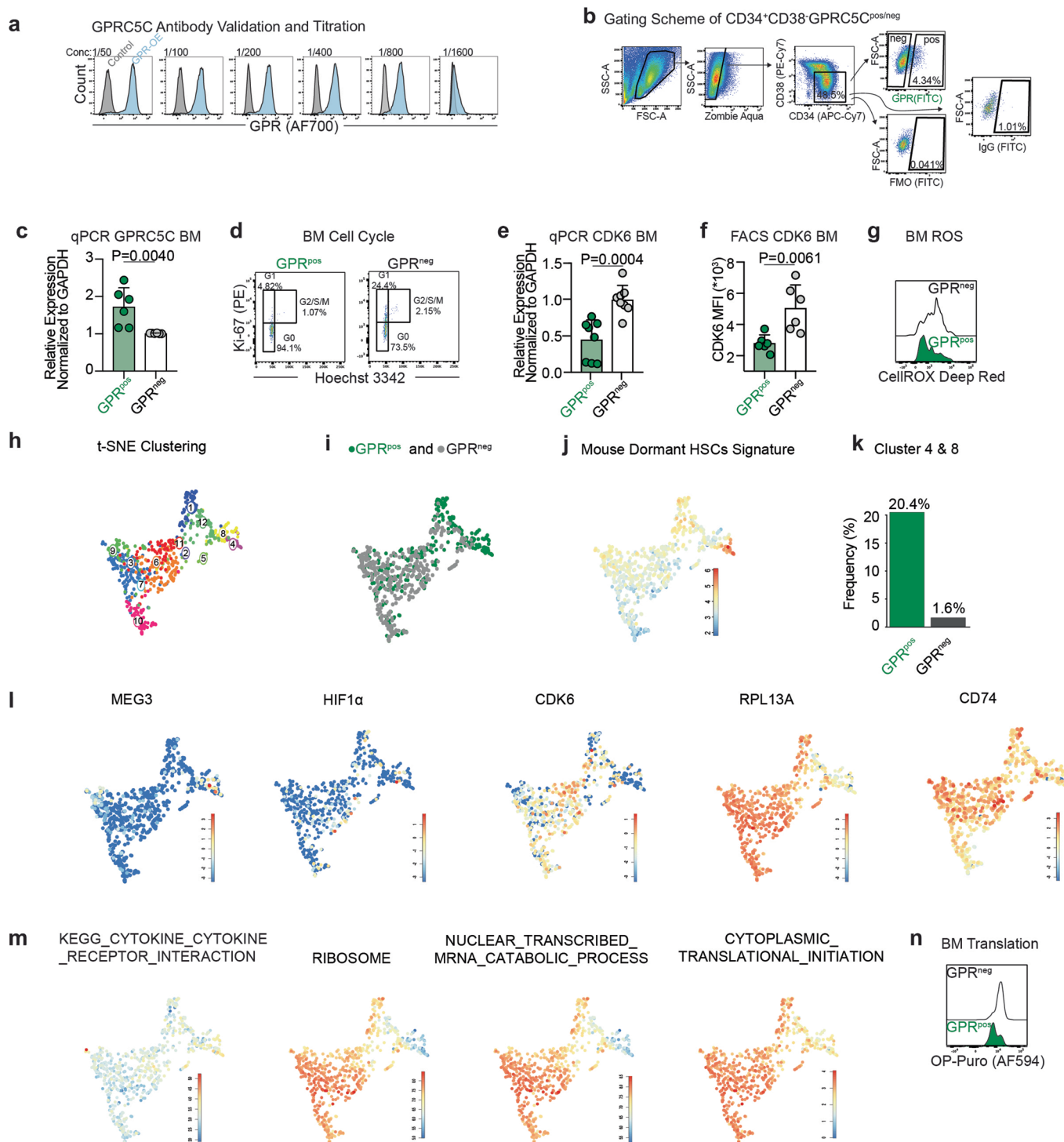
**Peer review information** *Nature Cell Biology* thanks Vijay Sankaran, Toshio Suda and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

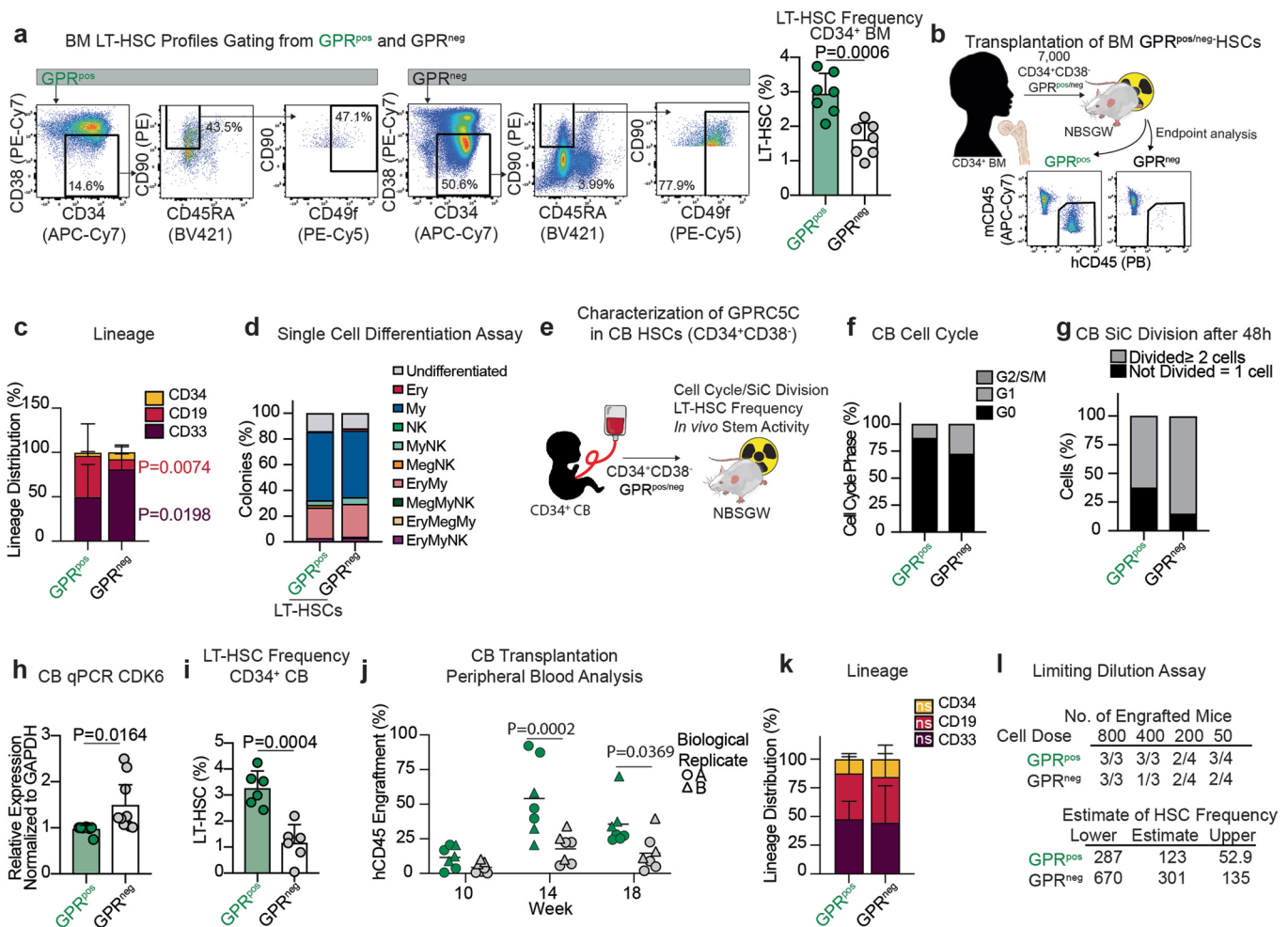


**Extended Data Fig. 1 | a.** Representative FACS gating for HSCs (CD34<sup>+</sup>CD38<sup>-</sup>), ST-HSCs (CD34<sup>+</sup>CD38<sup>-</sup>CD90<sup>-</sup>CD45RA<sup>-</sup>CD49f<sup>-</sup>), and LT-HSCs (CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>CD90<sup>+</sup>CD49f<sup>+</sup>) from CD34<sup>+</sup> BM donors used for bulk RNA-seq and scRNA-seq. **b.** Principal component analysis (PCA) of RNA-seq data from LT-HSCs and ST-HSCs. Based on top 5,000 variable genes after variance-stabilization of DESeq2 Wald test. **c.** GSEA of Hallmark terms in LT-HSCs compared to ST-HSCs. **d.** GSEA of KEGG Pathways in LT-HSCs compared to ST-HSCs. **e.** SOM of LT-HSCs from scRNA-seq. **f.** Pathway enrichment analysis from modules 1, 2, 3, and 4 of SOM of LT-HSCs. Hypergeometric test. **g.** Smoothed pseudo-temporal normalized expression profile of *Mouse dormant HSCs* signature in LT-HSCs. Smoothed profiles were computed by local regression. **h.** LT-HSC expression UMAPs of the stem and activation associated genes. The colour bars indicate log<sub>2</sub>-normalized expression. **i.** Smoothed pseudo-temporal normalized expression profile of the *CD49f1* signature from Belluschi et al., and *Quiescent LT-HSCs* from García-Prat et al., smoothed profiles were computed by local regression. Colour bars at the bottom of the SOMs show VarID clusters or RaceID clusters. The y-axis of the smoothed pseudo-temporal expression profiles indicates normalized expression. GSEA was performed with BH-adjusted p-values after adaptive multilevel splitting Monte Carlo approach.



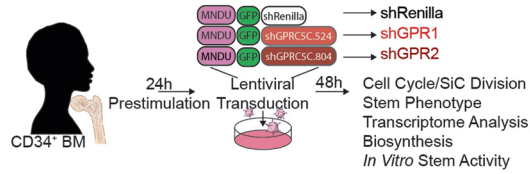


**Extended Data Fig. 2 | a.** Titration of monoclonal human GPRC5C antibody. HEK293T transfected with overexpressing GPRC5C or control plasmid and stained with decreasing concentration (left to right) of the GPRC5C antibody. **b.** Representative FACS scheme for gating of GPRC5C<sup>pos/neg</sup>-HSCs. GPRC5C<sup>pos</sup> cells are identified using a fluorescence minus one (FMO) or IgG control. **c.** qPCR validation of GPRC5C RNA expression from BM GPRC5C<sup>pos/neg</sup>-HSCs. Normalized to housekeeping gene *GAPDH* and GPRC5C<sup>neg</sup>.  $n=6$  biological replicates. **d.** Representative FACS plots of cell cycle profiles from BM GPRC5C<sup>pos/neg</sup>-HSCs. **e.** qPCR of *CDK6* expression from GPRC5C<sup>pos/neg</sup>-HSCs. Normalized to *GAPDH* gene and then GPRC5C<sup>neg</sup>-HSCs.  $n=8$  biological replicates. **f.** FACS measurement of CDK6 MFI in BM GPRC5C<sup>pos/neg</sup>-HSCs.  $n=6$  biological replicates. **g.** Representative FACS histogram of CellROX Deep Red MFI in BM GPRC5C<sup>pos/neg</sup>-HSCs. **h.** t-SNE representation of the clusters from BM GPRC5C<sup>pos/neg</sup>-HSCs. **i.** t-SNE depicting BM GPRC5C<sup>pos/neg</sup>-HSCs. **j.** Projection of the t-SNE of the *Mouse Dormant HSCs* signature. The colour bars indicate  $\log_2$ -normalized expression. **k.** The enrichment of GPRC5C<sup>pos</sup>-HSCs in clusters 4 and 8 is 20.4% compared to 1.6% for GPRC5C<sup>neg</sup>-HSCs. **l.** t-SNE of stem- and activation-associated genes. The colour bars indicate  $\log_2$ -normalized expression. **m.** t-SNE of KEGG and GO pathways. The colour bars indicate  $\log_2$ -normalized expression. **n.** Representative FACS histogram of OP-Puro MFI in BM GPRC5C<sup>pos/neg</sup>-HSCs. All data presented as mean  $\pm$  SD. Statistical significance was determined using two-tailed t-test (c, e, f) and at least two independent experiments were performed. ns, not significant.  $n$  indicates number of replicates. Source numerical data are available in source data.

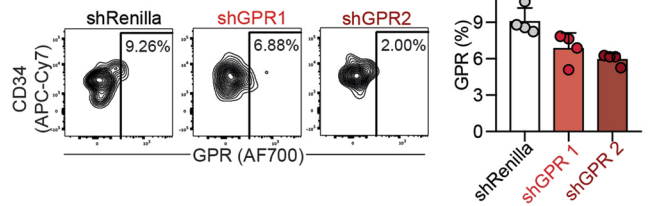


**Extended Data Fig. 3 | a.** Representative FACS gating scheme of LT-HSCs within GPRC5C<sup>pos/neg</sup> population from CD34<sup>+</sup> BM donors. Frequency of LT-HSCs within GPRC5C<sup>pos/neg</sup> cells from CD34<sup>+</sup> BM donors.  $n = 7$  biological replicates. **b.** Experimental design to assess the *in vivo* reconstitution ability of BM GPRC5C<sup>pos/neg</sup>-HSCs. **c.** Lineage distribution of GPRC5C<sup>pos/neg</sup>-HSCs. Data presented refer to the distribution of myeloid (CD33), B-lymphoid (CD19) and immature surface phenotype (CD34).  $n = 9$  biological replicates. **d.** Percentage of determined colony types derived from single BM GPRC5C<sup>pos/neg</sup>-LT-HSCs following culture under differentiating conditions for three weeks.  $n = 2$  biological replicates. **e.** Experimental design to characterize GPRC5C in human CB. **f.** FACS analysis of cell cycle from CB GPRC5C<sup>pos/neg</sup>-HSCs.  $n = 2$  biological replicates. **g.** SIC division assay of GPRC5C<sup>pos/neg</sup>-HSCs from CB donors quantified after 48 hours *in vitro* culture.  $n = 2$  biological replicates. **h.** qPCR of *CDK6* expression from CB GPRC5C<sup>pos/neg</sup>-HSCs. Normalized to housekeeping gene *GAPDH* and GPRC5C<sup>pos</sup>-HSCs.  $n = 9$  biological replicates. **i.** Frequency of LT-HSCs within GPRC5C<sup>pos/neg</sup> cells from CD34<sup>+</sup> CB donors.  $n = 6$  biological replicates. **j.** Assessment of reconstitution dynamics. Engraftment was assessed by percentage of human CD45 cells in the PB.  $n = 6$  biological replicates. **k.** Lineage contribution of CB GPRC5C<sup>pos/neg</sup>-HSCs. Data presented refer to the distribution of myeloid (CD33), B-lymphoid (CD19), and immature surface phenotype (CD34).  $n = 6$  biological replicates. **l.** Table showing the number of cells injected per dose, number of mice per dose, and the number of engrafted mice per dose. These data were used to calculate the HSC frequency seen in Fig. 2i. All data presented as mean  $\pm$  SD. Statistical significance was determined using two-tailed t-test (a, h, i, j) or two-way ANOVA (c and k) and at least two independent experiments were performed. ns, not significant.  $n$  indicates number of replicates. Source numerical data are available in source data.

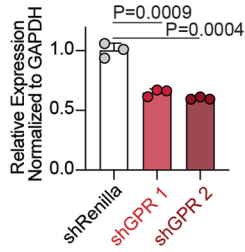
**a** GPRC5C-KD in CD34<sup>+</sup> BM



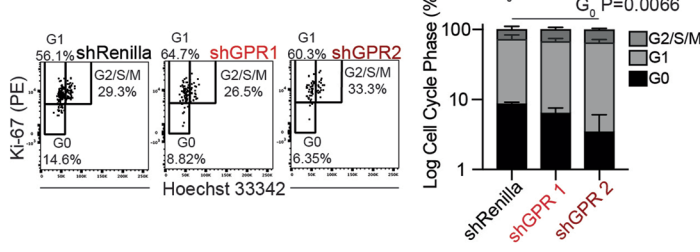
**b** GPRC5C Expression



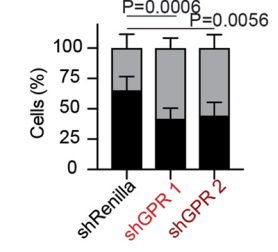
**c** GPRC5C qPCR



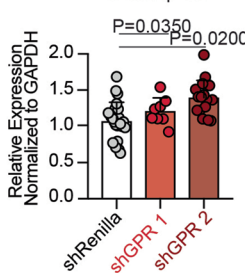
**d** Cell Cycle



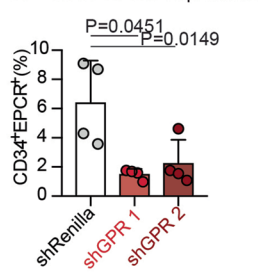
**e** SiC Division after 48h



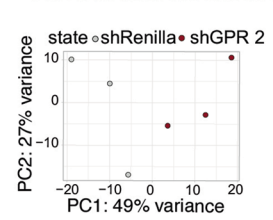
**f** CDK6 qPCR



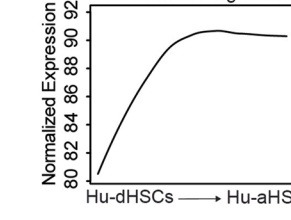
**g** CD34<sup>+</sup>EPCRC<sup>+</sup> Expression



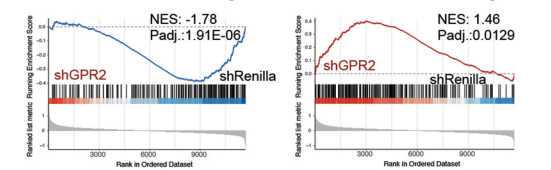
**h** PCA of shRenilla and shGPR 2



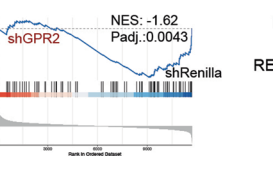
**i** GPRC5C-KD Signature



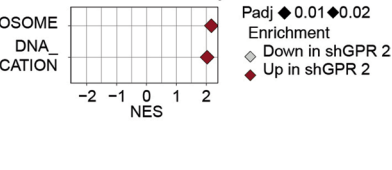
**j** Mouse Dormant HSCs Signature Mouse Active HSCs Signature



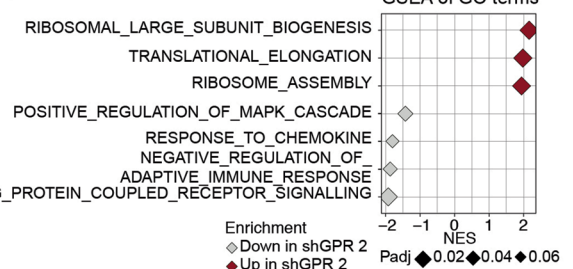
**k** GSEA of KEGG Pathways



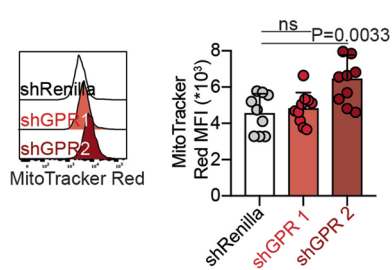
**l** GSEA of GO terms



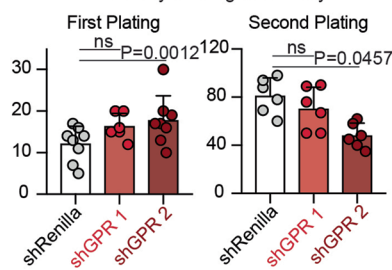
**m** GSEA of Hallmark



**n** Active Mitochondria

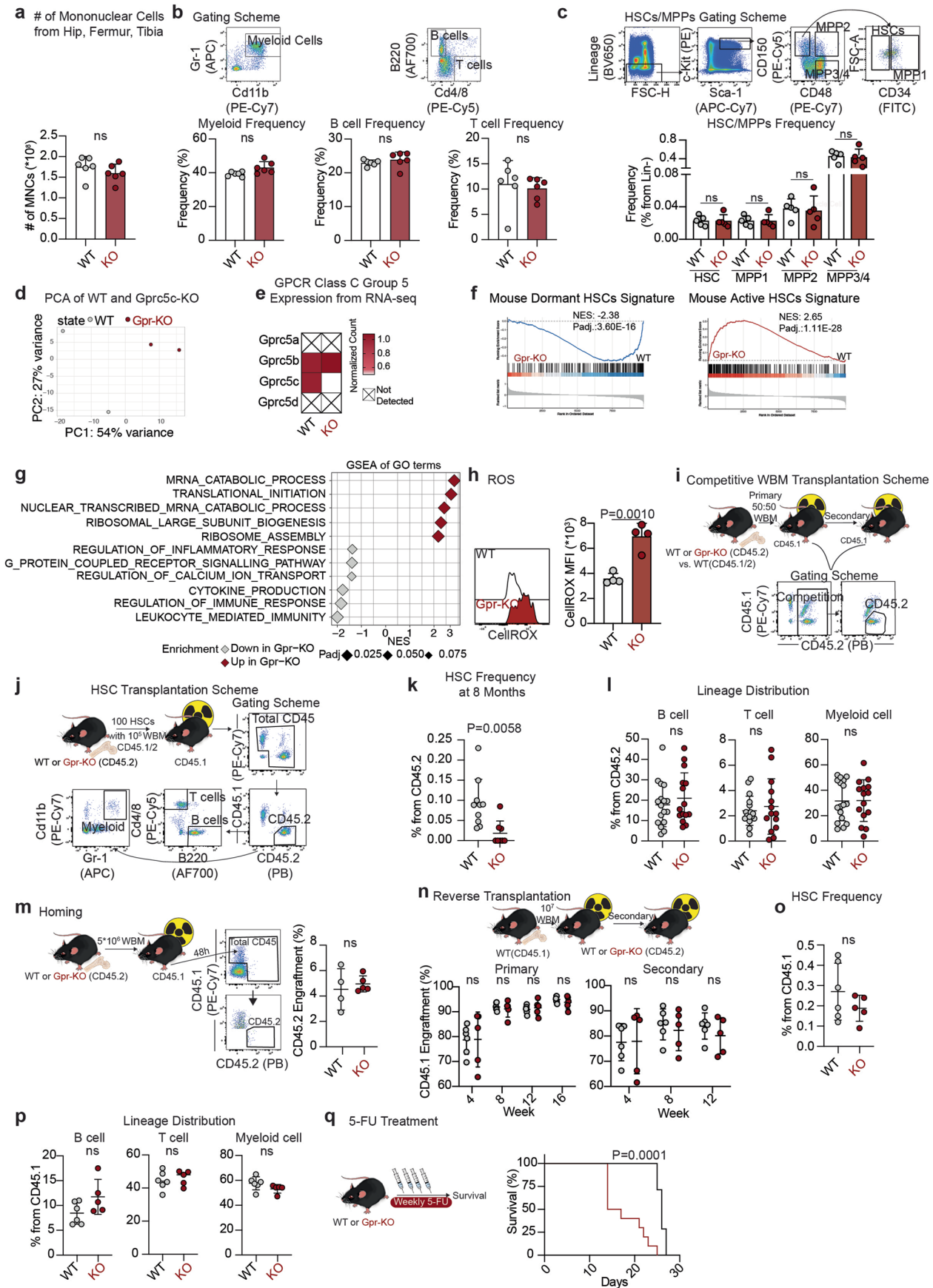


**o** Colony-forming Unit Assay



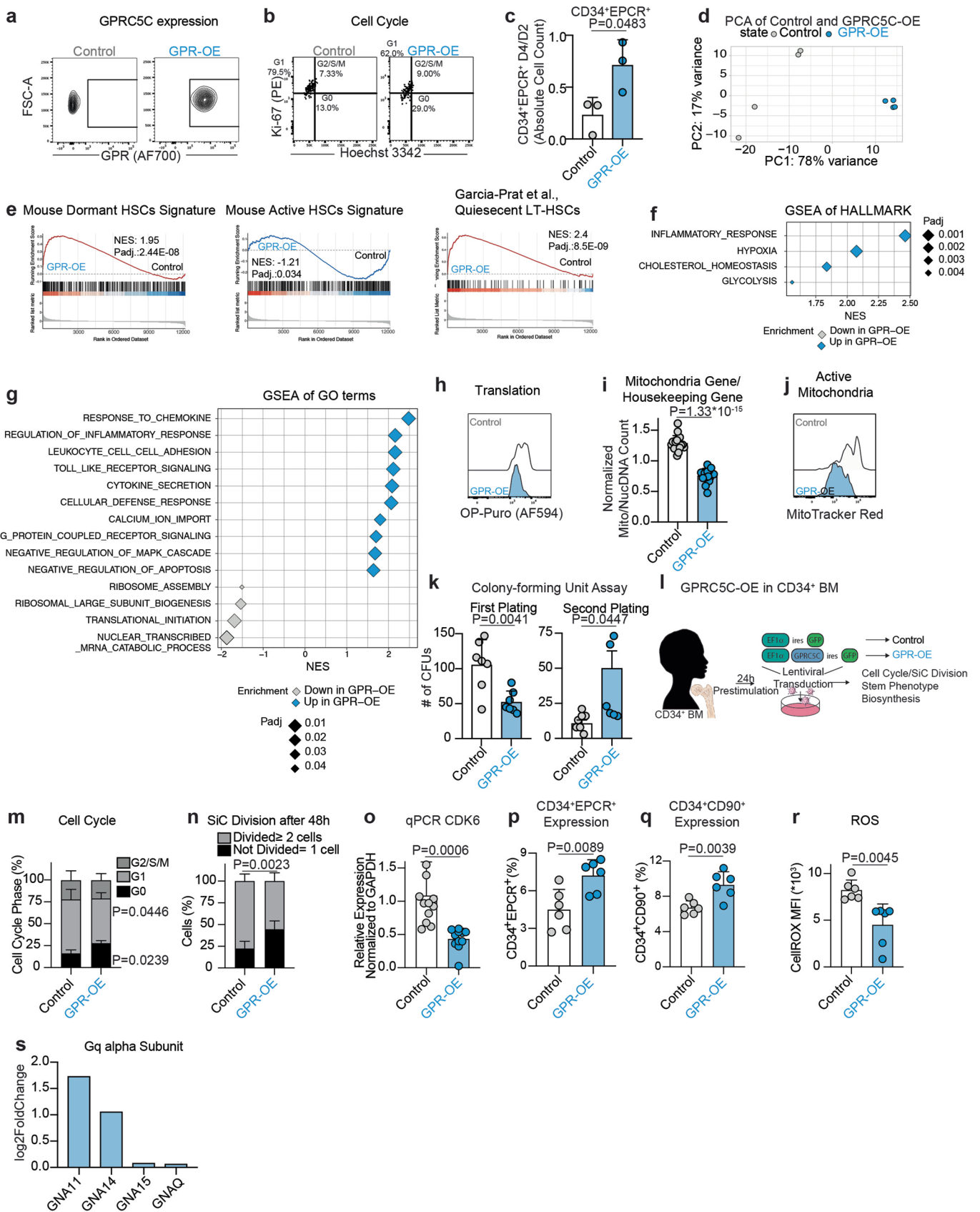
Extended Data Fig. 4 | See next page for caption.

**Extended Data Fig. 4 | a.** Experimental design to knockdown and assess the impact of decreased GPRC5C levels in CD34<sup>+</sup> BM. **b.** FACS validation of GPRC5C protein levels quantified in GFP<sup>+</sup> BM cells.  $n=4$  experiments. **c.** qPCR of *GPRC5C* expression in GFP<sup>+</sup> BM cells. Normalized to housekeeping gene *GAPDH* and shRenilla.  $n=3$  experiments. **d.** Cell cycle FACS analysis on GFP<sup>+</sup> BM cells. Cell cycle phase is represented on log<sub>10</sub> scale.  $n=4$  experiments. **e.** SiC division assay of GFP<sup>+</sup> BM cells quantified after 48 hours *in vitro* culture.  $n=4$  experiments. **f.** qPCR of *CDK6* expression in GFP<sup>+</sup> BM cells. Normalized to housekeeping gene *GAPDH* and shRenilla.  $n=8$  experiments. **g.** Surface expression of CD34<sup>+</sup>EPCR<sup>+</sup> in GFP<sup>+</sup> BM cells.  $n=4$  experiments. **h.** PCA of RNA-seq data from shGPR 2 and shRenilla control. Based on top 5,000 variable genes. log<sub>2</sub>FC threshold  $\geq 0.5$ , padj.  $< 0.1$ . **i.** Smoothed pseudo-temporal expression profile of the *GPR-KD* signature in LT-HSCs. Smoothed profiles were computed by local regression. The y-axis of the smoothed pseudo-temporal expression profiles indicates normalized expression. **j.** GSEA of *Mouse Dormant HSCs*, *Mouse Active HSCs*, and *Quiescent LT-HSCs* from García-Prat et al., signature in shGPR 2 and shRenilla control. **k.** GSEA of KEGG pathways in shGPR 2 compared to shRenilla control. **l.** GSEA of GO terms in shGPR 2 compared to shRenilla control. **m.** GSEA of Hallmarks terms in shGPR 2 compared to shRenilla control. **n.** FACS measurement of MitoTracker Red MFI in GFP<sup>+</sup> BM.  $n=10$  experiments. **o.** CFU assay of GFP<sup>+</sup> cells from GPRC5C-KD and shRenilla control. Primary plating  $n=8$  experiments, except  $n=6$  experiments for shGPR1. Secondary plating  $n=6$  experiments. All data presented as mean  $\pm$  SD. Statistical significance was determined using two-tailed t-test, except d and e (two-way ANOVA). GSEA performed with BH-adjusted p-values after adaptive multilevel splitting Monte Carlo approach. ns, not significant.  $n$  indicates the number of biological replicates. For all experiments, at least two independent experiments were performed. At least two human BM donors were used. Source numerical data are available in source data.



Extended Data Fig. 5 | See next page for caption.

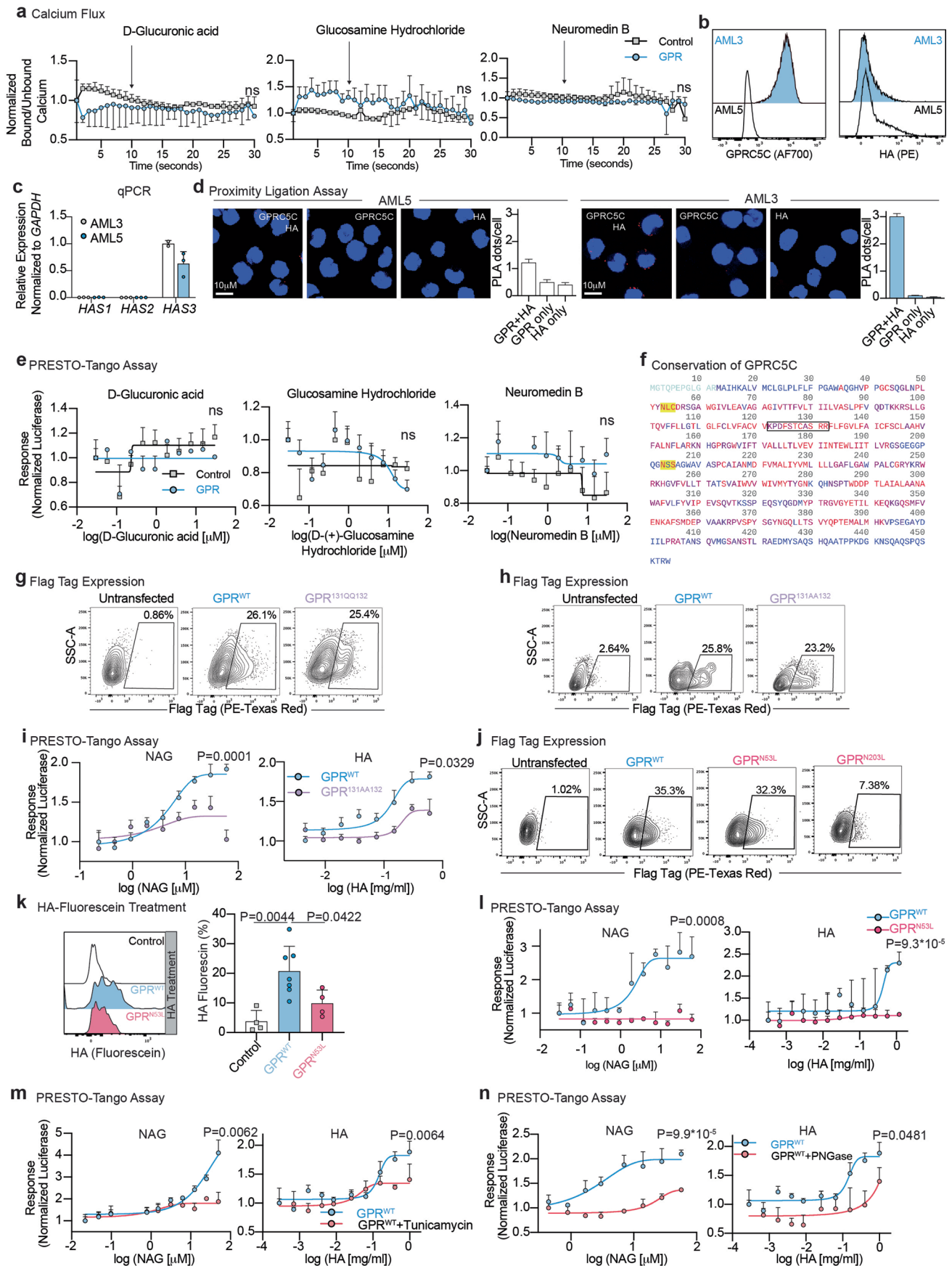
**Extended Data Fig. 5 | a.** Absolute number of mononuclear cells (MNC) isolated from hip, femur, and tibia.  $n=6$  biological replicates. **b.** FACS gating scheme for myeloid, B cells and T cells. Frequency of described populations from Gprc5c-KO and WT control BM.  $n=6$  biological replicates. **c.** Upper panel: FACS gating scheme for HSCs (LSK-CD150<sup>+</sup>CD48<sup>+</sup>CD34<sup>-</sup>), MPP1 (LSK-CD150<sup>+</sup>CD48<sup>+</sup>CD34<sup>+</sup>), MPP2 (LSK-CD150<sup>+</sup>CD48<sup>+</sup>), and MPP3/4 (LSK-CD150<sup>-</sup>CD48<sup>+</sup>). Lower panel: frequency of described populations from Gprc5c-KO and WT control BM.  $n=5$  biological replicates. **d.** PCA of RNA-seq data from Gprc5c-KO and WT control HSCs. Based on top 5,000 variable genes.  $\log_2FC$  threshold  $\geq 0.5$ ,  $padj < 0.1$ . **e.** Differential gene expression. Heatmap representing mean of normalized counts (normalized to WT).  $n=2$  biological replicates. **f.** GSEA of gene signatures in Gprc5c-KO and WT control HSCs. **g.** GSEA of GO terms in Gprc5c-KO compared to WT control HSCs. **h.** FACS measurement of CellROX Deep Red MFI in Gprc5c-KO compared to WT control HSCs.  $n=4$  biological replicates. **i.** Experimental design for competitive WBM transplantation. 50:50 ratio transplant of CD45.1/2 WT: CD45.2 WT or Gprc5c-KO WBM into CD45.1 recipient mice. Representative FACS gating scheme to quantify engraftment. **j.** Experimental design for HSC transplantation and FACS gating scheme to quantify engraftment and lineage distribution. **k.** Endpoint frequency of BM HSCs in CD45.2 cells of Gprc5c-KO ( $n=8$  biological replicates) and WT control ( $n=10$  biological replicates). **l.** Lineage distribution of BM Gprc5c-KO ( $n=14$  biological replicates) and WT control ( $n=16$  biological replicates) from HSC transplantations. **m.** Left panel: experimental design and FACS gating scheme of homing assay. Right panel: frequency of CD45.2 cells engraftment 48 hours post transplantation.  $n=4$  biological replicates. **n.** Upper panel: experimental design for reverse chimeras. Lower panel: PB analysis of reverse chimeras. Engraftment is represented by CD45.1 cell percentage engraftment.  $n=5$  biological replicates for Gprc5c-KO and  $n=6$  biological replicates for WT mice. **o.** Endpoint frequency of BM HSCs from CD45.1 cells transplanted in Gprc5c-KO ( $n=5$  biological replicates) and WT control ( $n=6$  biological replicates) mice. **p.** Lineage distribution of BM CD45.1 cells transplanted in Gprc5c-KO ( $n=5$  biological replicates) and WT control ( $n=6$  biological replicates) mice. **q.** Survival curve of Gprc5c-KO ( $n=10$  biological replicates) and WT control ( $n=7$  biological replicates) mice treated with 5-FU. All data presented as mean  $\pm$  SD. Statistical significance was determined using two-tailed t-test, except panel q: Mantel-Cox test. GSEA performed with BH-adjusted p-values after adaptive multilevel splitting Monte Carlo approach. ns, not significant.  $n$  indicates the number of biological replicates. Source numerical data are available in source data.



Extended Data Fig. 6 | See next page for caption.

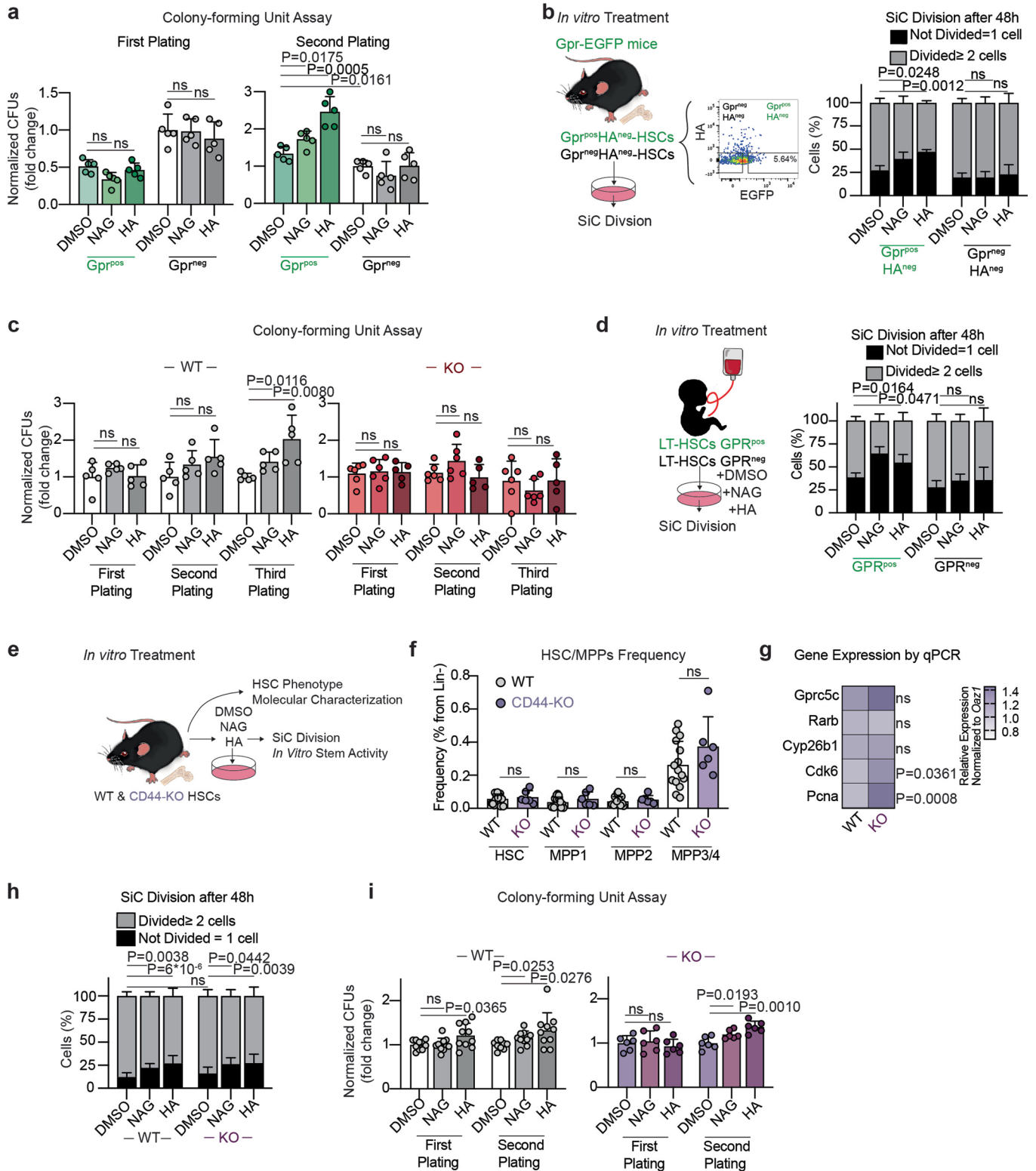
**Extended Data Fig. 6 | a.** Representative FACS plot for validation of GPRC5C-OE in GFP<sup>+</sup> mPB cells. **b.** Representative FACS plot of cell cycle profiles from GFP<sup>+</sup> mPB cells. **c.** Ratio of the absolute number of CD34<sup>+</sup>EPCR<sup>+</sup> cells from day 4 compared to day 2 *in vitro*.  $n = 3$  experiments. **d.** PCA of RNA-seq data from GPRC5C-OE and control. Based on top 5,000 variable genes.  $\log_2FC$  threshold  $\geq 0.5$ ,  $p_{adj.} < 0.1$ . **e.** GSEA of quiescent signatures in GPRC5C-OE compared to control. **f.** GSEA of Hallmarks terms in GPRC5C-OE compared to control. **g.** GSEA of GO terms in GPRC5C-OE compared to control. **h.** FACS histogram of OP-Puro MFI in GPRC5C-OE and control. **i.** RNA-seq data showing ratio read count of mitochondria genes normalized to housekeeping genes (*GAPDH*, *ACTB*, *PPIA*).  $n = 19$  genes averaged from 4 technical replicates. **j.** FACS histogram of MitoTracker Red MFI in GPRC5C-OE and control. **k.** CFU assay of GFP<sup>+</sup> cells from GPRC5C-OE and control.  $n = 7$  experiments, except for secondary plating GPR-OE, where  $n = 6$  experiments. **l.** Experimental design to overexpress and assess the impact of the gain of GPRC5C expression in BM CD34<sup>+</sup>. **m.** FACS-based cell cycle analysis of GPRC5C-OE and control.  $n = 6$  experiments. **n.** SiC division assay of GPRC5C-OE and control quantified after 48 hours *in vitro* culture.  $n = 6$  experiments. **o.** qPCR of *CDK6* expression. Normalized to housekeeping gene *GAPDH* and control.  $n = 12$  experiments for WT and  $n = 11$  experiments for GPR-OE. **p.** Surface expression of CD34<sup>+</sup>EPCR<sup>+</sup> in GPRC5C-OE and control.  $n = 6$  experiments. **q.** Surface expression of CD34<sup>+</sup>CD90<sup>+</sup> in GPRC5C-OE and control.  $n = 4$  experiments. **r.** FACS measurement of CellROX Deep Red MFI in GPRC5C-OE and control.  $n = 6$  experiments. **s.**  $\log_2FC$  of proteins of Gq alpha subunit. All data presented as mean  $\pm$  SD. Statistical significance was determined using two-tailed t-test, except m and n (two-way ANOVA). GSEA performed with BH-adjusted p-values after adaptive multilevel splitting Monte Carlo approach. ns, not significant.  $n$  indicates the number of biological replicates. For all experiments, at least two independent experiments were performed. At least two human BM donors were used. Source numerical data are available in source data.





Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | a.** Normalized ratio of baseline calcium level measured by Indo-1 followed by calcium flux in response to D-glucuronic acid (GA), glucosamine hydrochloride (G), and neuromedin B (NMB) in GPRC5C-OE ( $n=3$  experiments) and control ( $n=4$  experiments) in AML3 cells. **b.** HA and GPRC5C expression in AML3 and AML5. **c.** *HAS1*, *2*, and *3* expression in AML3 and AML5.  $n=3$  experiments, except for *HAS3* expression in AML3  $n=2$  experiments. **d.** Technical PLA controls of GPRC5C and HA in AML3 and AML5. Plot is representative of one out of three independent experiments. Each independent experiment was performed in two technical replicates.  $n=543, 551, 463, 990, 503,$  and  $523$  cells for AML3: GPR + HA, GPR only, HA only and AML5: GPR + HA, GPR only, and HA only, respectively. Data presented as mean  $\pm$  SEM. **e.** Concentration-response curve of PRESTO-Tango construct of GPRC5C and control vector treated with GA, G, and NMB in HTLA cells. Normalized to luciferase response at the lowest concentration.  $n=4$  experiments. **f.** Residue conservation of GPRC5C based on a multiple sequence alignment of 250 homologous proteins extracted from UniRef90. Red: highly conserved; purple: conserved; blue: not conserved; light blue: no conservation information available; yellow background: sequon for N-glycosylation used for mutation analysis within the ECD of the protein; black box: the putative BX<sub>7</sub>B domain. **g.** FACS validation of surface protein levels of GPRC5C<sup>WT</sup> and site-specific mutant (GPRC5C<sup>131QQ132</sup>) in HTLA cells. **h.** FACS validation of surface protein levels of GPRC5C<sup>WT</sup> and site-specific mutant (GPRC5C<sup>131AA132</sup>) in HTLA cells. **i.** Concentration-response curve of PRESTO-Tango construct for GPRC5C<sup>WT</sup> and site-specific mutant (GPRC5C<sup>131AA132</sup>) at predicted HA interaction site in HTLA cells.  $n=6$  experiments. **j.** FACS validation of surface protein levels of GPRC5C<sup>WT</sup> and site-specific mutants (GPRC5C<sup>N53L</sup> and GPRC5C<sup>N203L</sup>) in HTLA cells. **k.** HA-fluorescein-treated cells quantified for GFP retention in HEK293T cells expressing GPRC5C<sup>WT</sup> and GPRC5C<sup>N53L</sup>. FACS plots representing HA retention in each condition.  $n=4$  experiments, except for GPRC5C<sup>WT</sup> where  $n=7$  experiments. **l.** Concentration-response curve of PRESTO-Tango construct for GPRC5C<sup>WT</sup> and site-specific mutant (GPRC5C<sup>N53L</sup>) at predicted HA interaction site in HTLA cells.  $n=6$  experiments. **m.** Concentration-response curve of PRESTO-Tango construct for GPRC5C<sup>WT</sup> treated with tunicamycin in HTLA cells.  $n=3$  experiments. **n.** Concentration-response curve of PRESTO-Tango construct for GPRC5C<sup>WT</sup> treated with PNGase in HTLA cells.  $n=3$  experiments. All data presented as mean  $\pm$  SD, except where otherwise stated. Statistical significance was determined using two-tailed t-test. ns, not significant.  $n$  indicates the number of replicates. For all experiments, at least two independent experiments were performed. Source numerical data are available in source data.



Extended Data Fig. 8 | See next page for caption.

**Extended Data Fig. 8 | a.** Serial CFU assay of  $Gprc5c^{pos/neg}$ -HSCs cultured in NAG, HA, or DMSO control for 72 hours.  $n = 6$  biological replicates. **b.** Left panel: experimental design to assess the effect of *in vitro* HA, NAG, and DMSO treatment on  $HA^{neg}Gprc5c^{pos/neg}$ -HSCs. Right Panel: SiC division assay of  $HA^{neg}Gprc5c^{pos/neg}$ -HSCs quantified after 48 hours *in vitro* culture treatment with NAG, HA, and DMSO.  $n = 5$  biological replicates. **c.** Serial CFU assay of  $Gprc5c$ -KO and WT control HSCs cultured in NAG, HA, and DMSO control for 72 hours.  $n = 5, 3, 5, 6, 6,$  and  $5$  biological replicates for WT-DMSO, WT-NAG, WT-HA, Gpr-KO-DMSO, Gpr-KO-NAG, and Gpr-KO-HA, respectively. **d.** SiC division assay of  $GPRC5C^{pos/neg}$ -LT-HSCs with *in vitro* culture treatment for 48 hours with NAG, HA, and DMSO control.  $n = 4$  biological replicates. **e.** Experimental design to assess the effect of *in vitro* NAG, HA, and DMSO treatment on CD44-KO and WT control mice. **f.** Frequency of HSC/MPP populations from 8-week-old CD44-KO ( $n = 6$  biological replicates) and WT control ( $n = 12$  biological replicates) mice. **g.** Differential gene expression. Heatmap representing median RNA expression from qPCR data (normalized to housekeeping gene *Oaz1* and WT control).  $n = 7$  biological replicates. **h.** SiC division assay of CD44-KO and WT control HSCs quantified after 48 hours *in vitro* culture treatment with NAG, HA and DMSO.  $n = 15, 15, 13, 8,$  and  $12$  biological replicates for WT-DMSO, WT-NAG, WT-HA, CD44-KO-DMSO, CD44-KO-NAG, and CD44-KO-HA, respectively. **i.** Serial CFU assay of CD44-KO ( $n = 6$  biological replicates) and WT control ( $n = 10$  biological replicates) HSCs cultured *in vitro* with NAG, HA, and DMSO for 48 hours. All data presented by mean  $\pm$  SD. Statistical significance was determined using two-tailed *t*-test (a, c, f, g, i) or two-way ANOVA (b,d, h). ns, not significant. *n* indicates number of biological replicates. For all experiments, at least two independent experiments were performed. Source numerical data are available in source data.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

**Data collection** BD FACSDiva v8.0.3 (Flow cytometers and sorters, BD Bioscience), ZEN blue v2.5 (Zeiss international), Berthold Microplate Luminometer (Berthold), confocal microscope (Nikon C2).

**Data analysis**

Bulk RNA-seq analysis method: low level processing  
Raw fastq files were mapped against the hg38 or the mm10 reference genome by the mRNA-seq tool from the bioinformatics pipeline snakePipes. "Alignment" mode was used in this tool for the mapping of sequenced reads using STAR (STAR\_2.7.4a), followed by the quantification of expression counts by featureCounts. DeepTools QC (3.3.2) was used for quality checking. Genes with an average expression higher than 100 counts along all samples were selected for further analysis. The differential expression was generated with DESeq; considering significant the results with a FDR < 0.1 and a log<sub>2</sub>FC threshold ≥ 0.5.

Bulk RNA-seq analysis method: Downstream analysis  
The relative expression of DEGs based on the variance stabilized read counts calculated by DESeq2 package was visualized in PCA and heatmaps. Volcano plots representing the results of differential expression were generated with EnhancedVolcano. Gene set enrichment analysis (GSEA) was performed with fgsea R package for evaluation of expression in the pairwise comparisons of previously published gene signatures of HSCs Activation/ Quiescence-Dormancy, the hallmark gene set (h.all.v7.2.symbols.gmt), the KEGG pathway database gene set (c2.cp.kegg.v7.2.symbols.gmt) and the GO Biological Processes (c5.go.bp.v7.2.symbols.gmt,) obtained from the Molecular Signatures Database, considering significant pathways at an FDR < 0.1. The GSEA enrichment profile of concrete signatures was plotted using gseaplot2 function.

Single-cell RNA-seq analysis

Quantification of transcript abundance

Paired-end reads from the mCEL-Seq2 and 10x human HSC data were aligned to the same reference genome (GRCh38) using snakePipes scRNAseq workflow in mode Gruen and in mode STARsolo, respectively. Gencode version 31 gtf was used for feature counting.

Genotyping of patients in the 10x dataset and splitting cells by patient:

To assign cells to the corresponding patient, genotyping procedure as specified on was used. Briefly, bam files produced by STARsolo were filtered for high quality reads using samtools view parameters `q 10 F 3844`. Reads were assigned to genes with featureCounts using bam format as output. Gene-assigned bam files were deduplicated per gene and per cell with umi tools. Genetic variants were called on deduplicated bams with freeBayes (with parameters `-iXu -C 2 -q 1`), and high quality variants (`QUAL > 30`) were filtered for with vcfFilter. scSplit was run on the filtered vcf file requesting two genotypes and setting a ceiling on the expected fraction of doublets at 20% (parameters `-n 2 -d 0.2`).

Seurat object output by snakePipes mode STARsolo was annotated with genotypes obtained from scSplit and split into three objects, containing counts for cells assigned to either of the two patients in the cell mixture, or cells predicted to be doublets. Unnormalized count matrices for the two patients were used in further analysis, while the doublet cells were discarded.

Single-cell RNA sequencing data analysis

Overall, 1,827 cells passed the quality control threshold of >1,000 transcripts for the 10x Genomics human LT-HSC data. For normalization, the total transcript counts in each cell were normalized to 1 and multiplied by the minimum total transcript count across all cells that passed the quality control threshold (>1,000 transcripts per cell). VarID19 was run with a vector of batch variables indicating the batches/patients, and the following parameters: `mintotal = 1000`, `minexpr = 5`, `minnumber = 5`, `large=TRUE`, `regNB=TRUE`. For the mCEL-Seq2 datasets, we recovered 639 cells from patient 1048 (`mintotal=1000`). VarID was run with the default parameters.

Cell ordering and generation of self-organizing maps

Cells were ordered in ascending order based on the coordinates of the first UMAP dimension, i.e., the cells' x-axis coordinates from the UMAP. SOMs were generated using the FateID package on the basis of the ordering inferred from the UMAP. Only genes with >2 counts after size normalization in at least a single cell were included for the SOM analysis. In brief, smooth profiles were derived by applying local regression on normalized transcript counts after ordering cells. Next, a one-dimensional SOM with 200 nodes was computed on these profiles after z-transformation. Neighboring nodes were merged if the Pearson's correlation coefficient of the average profiles of these nodes exceeded 0.85. The remaining aggregated nodes represent the gene modules shown in the SOM figures.

Pathway enrichment analysis and gene set enrichment analysis.

Symbol gene IDs were first converted to Entrez gene IDs using the clusterProfiler package. Pathway enrichment analysis and GSEA were implemented using the ReactomePA package. Pathway enrichment analysis was done on genes taken from the different modules in the SOMs. GSEA was done using the differentially expressed genes between clusters 3,2,8 and clusters 9,4,5,10 inferred by the diffexpnb function from the RaceID package.

Human Quiescent CB and dormant mouse HSC signatures

Quiescent CB genes were selected from the list of differentially expressed genes between quiescent and active. HSCs from the bulk RNA-seq data generated in Belluschi et al.. Genes with a  $\log_2$  FoldChange > 0.5 were selected for the quiescent signature, and their expression in the human 10x dataset was aggregated and plotted.

Dormant mouse HSC genes were selected from the list of differentially expressed genes between dormant and active mouse HSCs from the bulk RNA-seq data generated in Cabezas-Wallscheid et al.,. Human-mouse orthologous genes with a  $\log_2$  FoldChange > 1 were selected for the dormant HSC signature, and their expression in the human 10x dataset was aggregated and plotted.

We used the following tools for analysis of bulk RNA-seq:

- snakePipes mRNA-seq workflow in mode Alignment and STAR for preprocessing
- R package DESeq2 for differential expression; considering significant the results with a  $FDR < 0.1$  and a  $\log_2FC$  threshold  $\geq 0.5$ .
- R package EnhancedVolcano for volcano plots representation
- R packages fgsea and enrichplot for GSEA

We used the following tools for analysis of single-cell RNA-seq:

- snakePipes scRNAseq workflow in mode Gruen and STARsolo for preprocessing of mCEL-Seq2 dataset
- STARsolo, samtools, featureCounts and umi tools for preprocessing of 10x dataset
- freeBayes for genetic variants calling
- VarID for batch removal
- FateID package for SOMs generation
- R package clusterProfiler for symbol gene IDs to Entrez gene IDs conversion
- R package ReactomePA for pathway enrichment and GSEA

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA-seq data has been deposited on ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) under the following accession numbers: E-MTAB-9862, E-MTAB-9863, E-MTAB-9874, E-MTAB-9892, E-MTAB-9922, E-MTAB-9967

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined based on extensive experience with similar experiments in our laboratory (Cabezas-Wallscheid et al., 2014 and 2017; Sommerkamp et al., 2020 and 2021; Renders et al., 2021)
Data exclusions	Sample exclusion was done only as a result of premature mouse death or if clear errors in pre-processing occurred.
Replication	All attempts of replication were successful. Key experiments were performed at least twice, exact number of independent experiments with various biological replicates can be found in the figure legends.
Randomization	All samples/ mice were analysed and allocated randomly.
Blinding	No blinding experiments were needed, since values were quantitative comparisons as determined by software and measurement.

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

## Methods

n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern		

## Antibodies

Antibodies used

Human Antibodies  
 APC-Cy7 CD34 BioLegend 343514 1/100,  
 PE CD90 BioLegend 328110 1/50,  
 BV605 CD90 BD Biosciences 562685 1/50,  
 PE-Cy7 CD38 BioLegend 303516 1/100,  
 PE CD201 BioLegend 351904 1/50,  
 BV421 CD45RA BD Biosciences 562885 1/100,  
 FITC GPRC5C R&D Systems FAB6594G-100UG 1/50 Clone:577315,  
 AF700 GPRC5C R&D Systems FAB6594N-100UG 1/50 Clone: 577315,  
 FITC CD45RA BD Biosciences 555488 1/100,  
 PE-Cy5 CD49f BD Biosciences 551129 1/200,  
 APC CD201 BioLegend 351906 1/50,  
 PE Ki67 BioLegend 350504 1/200,  
 PE-Cy5 Ki67 BioLegend 350519 1/200,  
 AF488 Ki67 BioLegend 350507 1.200,  
 AF647 CDK6 Abcam ab198946 1/500,  
 PE CD33 BD Biosciences 555450 1/200,  
 PE-Cy7 CD19 BD Biosciences 557835 1/200,  
 BV421 CD34 BD Biosciences 744904 1/100,  
 BV421 CD66b BD Biosciences G10F5 1/100 ,  
 BV650 CD15 BioLegend 323033 1/200,  
 BV711 CD14 BioLegend 367139 1/1000,  
 FITC CD56 BioLegend 362545 1/200,  
 PE-Cy5 CD45 BioLegend 304009 1/300,  
 PE CD71 BioLegend 334105 1/1000,  
 PE-Cy7 GlyA BioLegend 334619 1/300,  
 APC CD41a BioLegend 303709 1/200,  
 APC-Cy7 CD11b BioLegend 101225 1/300,  
 Hyaluronan Binding protein Amsbio AMS.HKD-BC40 1/50.  
 Hyaluronic Acid LS Bio C295845 1/50  
 Flag Tag clone M2 Sigma F3165

Mouse Antibodies  
 BV650 CD8a BioLegend 100742 1/1000,  
 BV650 CD11b BioLegend 101259 1/1000,  
 BV650 Ly-6G/Ly-6C (Gr-1) BioLegend 108442 1/1000,  
 BV650 TER-119/Erythroid Cells BioLegend 116235 1/1000,  
 BV650 CD45R/B220 BioLegend 103241 1/1000,  
 BV650 CD4 BD Biosciences 563232 1/1000,  
 PE/Cy7 anti-mouse CD8a BioLegend 100722 1/1000,  
 PE/Cy7 anti-mouse/human CD11b BioLegend 101216 1/1000,  
 PE/Cy7 anti-mouse Ly-6G/Ly6C (Gr-1) BioLegend 108416 1/1000,  
 PE/Cy7 anti-mouse TER-119/Erythroid Cell BioLegend 116221 1/1000,  
 PE/Cy7 anti-mouse/human CD45R/B220 BioLegend 103222 1/1000,  
 PE/Cy7 anti-mouse CD4 BioLegend 100422 1/1000,  
 Brilliant Violet 711 anti-mouse CD117 (c-kit) BioLegend 105835 1/1000,  
 Brilliant Violet 421 anti-mouse CD117 (c-kit) BioLegend 105828 1/1000,  
 PE anti-mouse CD117 (c-kit) BioLegend 105808 1/100,  
 APC/Cy7 anti-mouse Ly-6A/E (Sca-1) BioLegend 108126 1/500,  
 PE/Cy5 anti-mouse 150 (SLAM) BioLegend 115912 1/500,  
 Pacific Blue anti-mouse CD48 BioLegend 103418 1/500,  
 APC-eFlour780 CD45 Invitrogen 47-0451-82 1/200,  
 PE Streptavidin BioLegend 405203 1/2000,  
 CD34-AF700 BioLegend 560518 1/50,  
 CD34-FITC BioLegend 553733 1/50,  
 CD45.2-Pacific Blue BioLegend 109820 1/500,  
 CD45.1-PE/Cy7 BioLegend 110730 1/500,  
 CD4-PE/Cy5 BioLegend 100410 1/2000,  
 CD8a-PE/Cy5 BioLegend 100710 1/1000,  
 CD11b-APC/Cy7 BioLegend 101226 1/1000,  
 Gr1-APC BioLegend 108412 1/1000,  
 B220-AF700 BioLegend 103434 1/500

For flow cytometry antibodies, Lot numbers cannot reasonably be provided, as multiple different lots have been used over the course of this study. Depending on the individual experimental setting flow cytometry antibodies from the same clone but coupled to different fluorochromes were used.

#### Validation

All flow cytometry antibodies were already established and commonly used in our laboratory and have been published multiple times by us and other groups. CDK6 was validated on more vs. less quiescent cells as described in Cabezas-Wallscheid et al., 2017, Cell. GPRC5C (R&D) was extensively validated as shown in Extended Data Figure 2.

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	AML3 (Max Planck), HEK293T (Takarabio), AML5 (DSMZ)
Authentication	Common and well-established cell lines. None of the cell lines were authenticated.
Mycoplasma contamination	Not detected
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified lines were used

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	NOD,B6.SCID Il2ry-/-KitW41/W41 Jackson 26497 C57BL/6J (CD45.2) MPI-IE B6Ly5.1 (CD45.1) MPI-IE B6Ly5.1(CD45.1/2) MPI-IE Gpr5c LacZ Knockin Mice (Gpr5c-KO) Sano et al. B6.129(Cg)-Cd44tm1Hbg/J Jackson 005085  Male and female animals were used. Mice were between 8-weeks to 12-weeks of age.
Wild animals	No wild animals were used in this study
Field-collected samples	No field collected samples were used in this study
Ethics oversight	All mice were bred in-house in the animal facility at the MPI-IE in individually ventilated cages (IVCs). According to German guidelines, mice were euthanized by cervical dislocation. Animal procedures were performed according to the protocols



Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

#### Sample preparation

**Cell cycle analysis (Ki-67 or CDK6)**  
Stained cells were fixed with BD Cytofix/Cytoperm Buffer (BD Biosciences) for 10 minutes at 4 °C. Next, cells were stained with intracellular Ki-67 at 1:100 (BD Biosciences) or CDK6 at 1:500 (Abcam) in PermWash solution (BD Biosciences) for 45–90 minutes at 4 °C. Prior to cell cycle analysis on the BD LSRFortessa Cell Analyzer (BD Biosciences), cells were stained in the dark with Hoechst 33342/DAPI (Invitrogen) for 30 minutes at RT.

**CellROX/MitoTracker staining**  
Cells were incubated at 37 °C with CellROX DeepRed at 1/500 (Invitrogen) or MitoTracker Red at 50 nM (Invitrogen) in their corresponding media for 30 minutes. Cells were subsequently washed 3 times in PBS and stained for FACS analysis on the BD LSRFortessa Cell Analyzer (BD Biosciences).

**O-propargyl-puromycin**  
Stained cells were fixed for 15 minutes at RT with BD Cytofix/Cytoperm Buffer (BD Biosciences) and permeabilized with 0.1% Triton for 15 minutes at RT. The copper-catalysed azide-alkyne cycloaddition (CuAAC) was performed using an Alexa594-azide (Life Technologies, 5 mM final concentration) and the Click-iT Cell Reaction Buffer Kit (Life Technologies) according to the manufacturer's instructions.

**HA-fluorescein binding assay**  
HEK293T cells were transfected with 30 µg/ml of the GPRC5C-Tango plasmids using JetPrime transfection reagent. After 24 to 36 hours' transfection, 60,000 to 80,000 cells were treated with 8 µg HA conjugated to fluorescein with or without, 100 µM NAG for 60 minutes at RT with shaking. Cells were washed and fluorescein was quantified using the LSR II flow cytometer (BD Bioscience).

**Calcium flux assay**  
AML3 cells were infected with viral particles to overexpress GPRC5C or control. Cells were resuspended in IMDM with 1% FBS at 106 cells/ml. 7.5 µl Indo-1 (Invitrogen) was added per millilitre of cell suspension and incubated at 37 °C for 45 minutes. Following incubation, cells were washed twice with 1% FBS IMDM and then incubated for 30 minutes at 37 °C for de-esterification. The baseline of calcium was acquired on the LSR Fortessa II (BD Bioscience) for 30 seconds with the sample heater at 37 °C. The cells were stimulated with compounds and acquired for an additional 90 seconds. The ratio of indo-1-bound to indo-1-unbound fluorescence emission signals was calculated as a quantitative parameter for change in intracellular calcium levels. All values were normalized to the baseline of each sample.

#### Instrument

For cell sorting: FACS Aria Fusion (Becton Dickinson)  
For analysis: LSR II, LSR Fortessa (Becton Dickinson)

#### Software

Analysis was performed with FlowJo, statistical analysis with Graphpad Prism.

#### Cell population abundance

Population abundance is reported on Figures.

#### Gating strategy

**Phenotype Gating:**  
FSC-A v. SSC-A (cell gate excluding debris) --> FSC-W v. FSC-H (doublet exclusion) --> SSC-W v. SSC-H (doublet exclusion) --> more detailed gating strategy for each individual experiment is indicated in Supplemental Information

**Cell Cycle Gating:**  
FSC-A v. SSC-A (cell gate excluding debris) --> DAPI-A v. DAPI-H (doublet exclusion) --> gating in interested population --> Ki-67 and DAPI-H.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.