





# Implications of stress-induced gene expression for hematopoietic stem cell aging studies

Received: 27 March 2023

Accepted: 15 December 2023

Published online: 16 January 2024

 Check for updates

Anna Konturek-Ciesla , Rasmus Olofzon , Shabnam Kharazi & David Bryder  

A decline in hematopoietic stem cell (HSC) function is believed to underlie hematological shortcomings with age; however, a comprehensive molecular understanding of these changes is currently lacking. Here we provide evidence that a transcriptional signature reported in several previous studies on HSC aging is linked to stress-induced changes in gene expression rather than aging. Our findings have strong implications for the design and interpretation of HSC aging studies.

Lifelong blood cell production emanates from bone marrow (BM) hematopoietic stem cells (HSCs). During aging, the hematopoietic output from HSCs declines<sup>1</sup>. To reveal the molecular mechanisms for this, several studies conducted genome-wide RNA expression profiling of young and aged candidate HSCs (cHSCs). Recently, a meta-analysis of previous data revealed a core aging signature (AS) of the most consistent gene expression changes, while highlighting discrepancies among different studies and datasets<sup>2</sup>.

In HSC research, cHSCs are typically extracted from the BM using fluorescence-activated cell sorting (FACS) and such procedures take time because of the scarcity of HSCs. Also, when isolating HSCs and other hematopoietic stem and progenitor cells (HSPCs) using dye uptake and exclusion techniques, cells are exposed to elevations in temperature before FACS. Limited information is available on how different experimental procedures might affect the molecular profiles of HSCs.

In this study, by reanalyzing transcriptomic HSC aging data from multiple previous studies, we identified a stress-associated signature that was uncoupled from aging. We found that some cell isolation procedures could evoke this response strongly in primary HSPCs. We discuss its implications for previous interpretations on HSC aging.

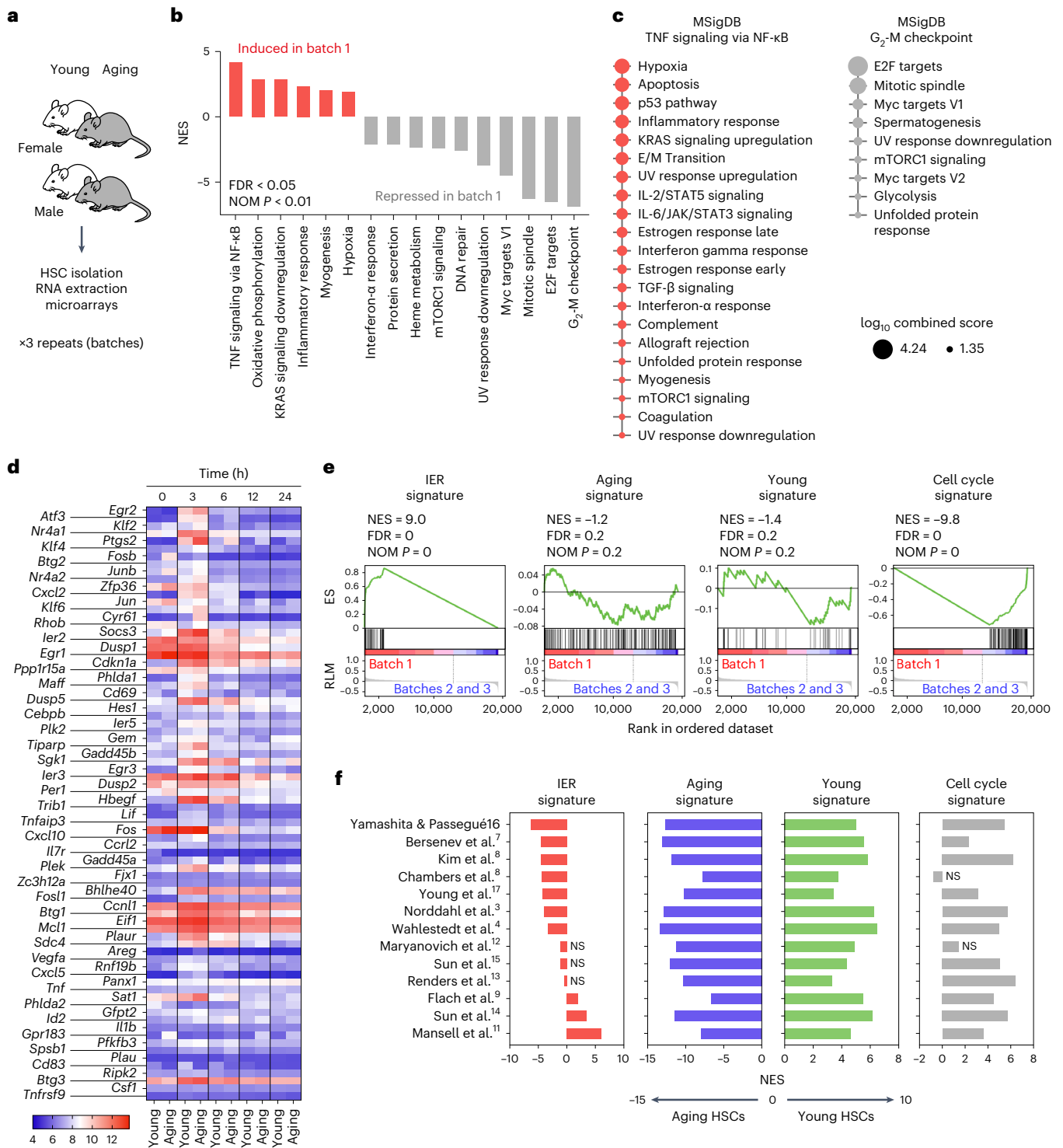
Two of our previous microarray datasets were part of a meta-analysis on HSC aging<sup>2</sup> and correlated well with the AS. However, while published separately as one for female and one for male cells<sup>3,4</sup>, they were generated simultaneously by isolating cHSCs in three batches, each with samples from young and aged male and female cells (Fig. 1a). This setup allowed us to directly assess potential batch effects. When comparing batch 1 to the other two batches for Molecular Signatures Database

(MSigDB) Hallmark gene sets<sup>5</sup>, we observed that genes expressed higher in batch 1 included several immediate early response (IER) genes (Supplementary Table 1) that associated particularly to ‘tumor necrosis factor (TNF) signaling via nuclear factor kappa-light-chain-enhancer of activated B cells (NF- $\kappa$ B)’, with downregulated genes relating strongly to ‘G<sub>2</sub>-M checkpoint’ (Fig. 1b and Supplementary Table 2). We defined an IER signature (IERsig) based on these upregulated genes (Supplementary Table 1). The IER genes were also positively associated with many other gene sets coupled to stress and negatively to cell cycle pathways (Fig. 1c and Supplementary Table 2). It is unknown to us why the samples of our first batch associated more with these expression signatures. Regardless, we next assessed the IERsig after brief culture of young and aged cHSCs to evaluate whether aging impacted on its induction. The IERsig genes *Fos* and *Jun* presented with high expression already after 0.5 h of culture, with other assessed IER genes peaking at 1–3 h and declining thereafter (Fig. 1d and Extended Data Fig. 1a). This agrees with *Fos* and *Jun* being primary IER genes that decline rapidly after their induction<sup>6</sup>. Notably, the induction of the IERsig was independent of age (Fig. 1d and Extended Data Fig. 1a).

Using gene set enrichment analysis (GSEA) for the IERsig and cell cycle gene sets from our batch analysis (Fig. 1b), and AS and young signature sets from Flohr Svendsen et al.<sup>2</sup>, we assessed genome-wide transcription in 13 reported datasets on murine HSC aging<sup>3,4,7–17</sup> (Supplementary Tables 2 and 3). These gene sets were benchmarked by comparing batch 1 to the other two batches (Fig. 1e). In all investigated studies, aged cHSCs displayed a strong AS, with the young signature associating with the young samples. Furthermore, in general, cell cycle

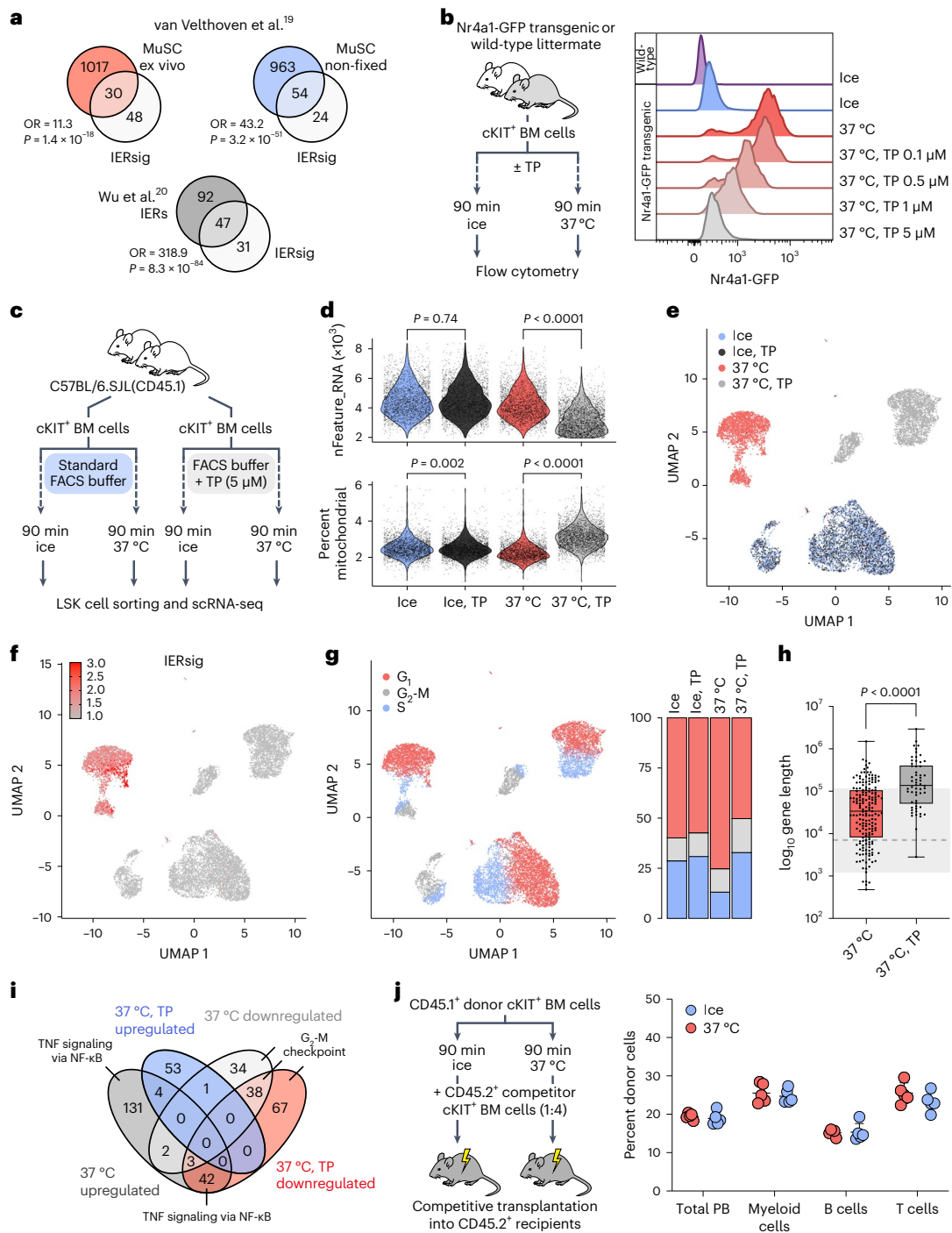
Division of Molecular Hematology, Lund Stem Cell Center, Institution for Laboratory Medicine, Lund University, Lund, Sweden.

 e-mail: [david.bryder@med.lu.se](mailto:david.bryder@med.lu.se)



**Fig. 1 | Identification of a batch-associated transcriptional signature that is unrelated to HSC aging.** **a**, Experimental design. RNA was extracted from purified cHSCs (lineage-SCA-1<sup>+</sup>cKIT<sup>+</sup>CD150<sup>+</sup>CD48<sup>-</sup>) from cohorts of young (2–4 months) and aged (21–24 months) mice of both sexes and processed for gene expression analysis using Affymetrix 430 2.0 arrays. The experiment involved three separate batches ( $n = 4,000$ – $10,000$  per cHSC/group in each replicate). Data are from Norddahl et al.<sup>3</sup> and Wahlestedt et al.<sup>4</sup>. **b**, GSEA for MSigDB Hallmark 2020 pathways that were associated with genes induced and repressed in batch 1. Shown are pathways with an FDR < 0.05 and nominal  $P < 0.01$  (NOM  $P$ ). **c**, MSigDB Hallmark 2020 pathway analysis of leading-edge genes extracted from the TNF signaling via NF- $\kappa$ B and G<sub>2</sub>-M checkpoint-associated gene sets. Shown are pathways with an adjusted  $P < 0.05$  (one-sided Fisher’s exact test with Benjamini–Hochberg correction). The size of the dots is proportional to the

Enrichr log<sub>10</sub> combined scores, which is a combined metric of  $P$  values and ORs (computed using the Enrichr software). **d**, Heatmap depicting the time course gene expression of the 78 IER genes associated with batch 1 in freshly isolated and in vitro-cultured HSCs. Data are from Beerman et al.<sup>33</sup>. **e**, GSEA signature enrichment plots for the IER, age-induced, age-repressed and cell cycle-related genes in batch 1 cHSCs compared to cHSCs from batches 2 and 3. **f**, GSEA results for the IER, age-induced, age-repressed and cell cycle-related genes in young and aged cHSCs from the indicated datasets. Unless indicated by NS (not significant), the bars depict the normalized enrichment scores (NES) with an FDR < 0.05 and NOM  $P < 0.01$ . NES > 0 indicates signature enrichment in young cHSCs, while NES < 0 indicates enrichment in aged cHSCs. **b, c, e, f**, The FDR method was applied to correct for multiple hypotheses testing. **b, c, f**, Detailed statistical information is provided in Supplementary Table 2. RLM, ranked list metric.



**Fig. 2 | The effect of cell isolation procedures on induction of the IER signature.** **a**, Top left, IERSig genes compared to differentially increased nascent transcripts (FDR < 0.1, fold change > 2) during muscle stem cell (MuSC) isolation (red circle). Top right, Overlap of IERSig genes with differentially upregulated genes in untreated versus in vivo-fixed MuSCs (FDR < 0.1, fold change > 2, blue circle). Data were derived from van Velthoven et al.<sup>19</sup> Bottom, IERSig genes compared to IER genes identified by Wu et al.<sup>20</sup> (gray circle). Each comparison displays the OR and *P* values (two-sided Fisher's exact test). **b**, Left, experimental design. Right, representative histograms depicting the Nr4a1-GFP signal in HSPCs incubated on ice or at 37 °C with increasing concentrations of TP (*n* = 3–4 mice per genotype analyzed in three independent experiments). **c**, Experimental design to evaluate the IER in lineage-SCA-1<sup>+</sup>cKIT<sup>+</sup> (LSK) HSPCs using scRNA-seq. **d**, Feature counts (ice ± TP, *P* = 0.074; 37 °C ± TP, *P* < 0.0001) and mitochondrial gene content (ice ± TP, *P* = 0.002; 37 °C ± TP, *P* < 0.0001) detected per cell (two-sided Mann–Whitney *U*-test). **e**, UMAP embedding of 12,556 LSK cells color-coded based on sample assignments (37 °C, *n* = 2,873 cells; 37 °C TP, *n* = 3,580

cells; ice, *n* = 3,048 cells; ice, TP *n* = 3,065 cells). **f**, UMAP embedding of LSK cells with aggregated IERSig expression depicted in red. **g**, Left, UMAP embedding of LSK cells color-coded according to cell cycle phase. Right, bar graphs depicting cell cycle phase distributions across samples. **h**, Distribution of gene lengths for genes upregulated in HSPCs incubated at 37 °C ± TP compared to cells kept on ice (*n* = 183 and *n* = 58 genes respectively, analyzed in one experiment). The box plots span the 25th to 75th centiles. The whiskers and center line indicate the minimum, maximum and median length. The background box represents the range of lengths for IERSig genes, with the dashed line indicating the median. *P* < 0.0001 (two-sided Mann–Whitney *U*-test). **i**, DEGs in HSPCs incubated at 37 °C ± TP compared to cells kept on ice. The MSigDB Hallmark 2020 pathways with the highest Enrichr combined score (adjusted *P* < 0.05, one-sided Fisher's exact test with Benjamini–Hochberg correction) for the indicated groups of genes are shown. **j**, Left, experimental design. Right, donor-derived chimerism in the indicated peripheral blood (PB) lineages 16 weeks after transplantation (*n* = 5 mice per group). The error bars denote the mean ± s.d.

genes were associated with young cHSCs. By contrast, the IERSig varied significantly between the datasets, being enriched in young or old samples, or showing no age-dependent enrichment (Fig. 1f, Extended Data Fig. 1b and Supplementary Table 2).

The literature presents a complex landscape when examining transcriptional features in HSCs associated with aging. One perspective posits that aging HSCs share transcriptional characteristics similar to those triggered by TNF<sup>16</sup>. However, this viewpoint may be influenced by the particular dataset generated, which showed a high correlation of the aged HSCs with the IERSig (Fig. 1f). Another line of research suggested that young HSCs may exhibit faster gene transcription than aging ones<sup>11</sup>, but a contrasting study posited an inverse relationship, indicating that transcription rates could actually be higher in aging HSCs<sup>2</sup>. The faster transcription observed in young HSCs in the former study could potentially be attributed to a strong induction of the IERSig in young cells in the particular dataset generated (Fig. 1f). The complexity deepens with work proposing reduced transforming growth factor- $\beta$  (TGF $\beta$ ) signaling in aging cHSCs<sup>14</sup>, which was not observed in a later meta-analysis<sup>2</sup>. The connection between the IERSig and TGF $\beta$  (Fig. 1c) and the specific genes showcased as downregulated with age in Sun et al.<sup>14</sup>, which represent prominent IER genes, is noteworthy. Finally, it is crucial to acknowledge that there are datasets that exhibit no discernible link with the IERSig in HSCs from either age group (Fig. 1f). Therefore, although RNA profiling studies have often detected the IERSig when examining murine HSC aging, its relationship to age is inconsistent. This variability urges cautious interpretation of preceding findings involving these genes.

The IERSig has been highlighted as an artifact during cell preparations outside hematopoiesis, which often require elevations in temperature<sup>18</sup>. Indeed, by comparing the IERSig to stress-associated genes identified in previous studies on muscle or neuronal cells<sup>19,20</sup>, we observed a significant overlap (Fig. 2a). We next used a green fluorescent protein (GFP)-reporter mouse line for *Nr4a1* (ref. 21), a prominent IER gene (Extended Data Fig. 1), and compared cKIT<sup>+</sup> Nr4a1-GFP BM cells kept on ice with those incubated at 37 °C for 90 min (Fig. 2b and Extended Data Fig. 2a). We also explored whether the IERSig could be mitigated with the transcriptional inhibitor triptolide (TP). Incubation at 37 °C led to a pronounced induction of GFP, which was blocked by TP in a dose-dependent manner (Fig. 2b). Hence, while cHSCs can be isolated with less induction of the IERSig, procedures that involve a rise in temperature lead to its direct induction.

To detail the IERSig further, we assessed the effects of temperature and TP using single-cell RNA sequencing (scRNA-seq) of HSPCs isolated from young mice (Fig. 2c and Extended Data Fig. 2b). We noted a significant reduction in RNA feature counts and a small yet steady elevation in mitochondrial gene content when incubating at 37 °C with TP, implying a decline in overall cell health after TP treatment (Fig. 2d). Using uniform manifold approximation and projection (UMAP), we distinguished cell clusters based on their treatment (Fig. 2e). Cells kept on ice formed a cluster regardless of TP treatment, while the samples incubated at 37 °C with or without TP formed separate clusters. Analysis of the IERSig revealed a potent and fairly uniform activation after incubation at 37 °C, which TP prevented. No noticeable impact from TP on the IERSig was observed when cells were kept on ice (Fig. 2f).

The projection of additional signatures identified in our batch analysis (Fig. 1b) on the scRNA-seq data revealed a homogenous distribution of the young-associated signature across samples (Extended Data Fig. 3a). The AS was associated with specific regions for each of the three clusters, but with no association as for treatment, and was inversely correlated to the cell cycle signature (Extended Data Fig. 3a). In agreement with the stress-induced downregulation of genes associated with the cell cycle pathways (Fig. 1b), further analysis revealed a pronounced decrease in the proportion of actively cycling cells after incubation at 37 °C (Fig. 2g).

Exploring the differentially expressed genes (DEGs) in cells incubated at 37 °C  $\pm$  TP (Supplementary Table 4) revealed that the genes upregulated in the absence of TP exhibited remarkably shorter lengths compared to those induced after TP treatment (Fig. 2h). In addition, by intersecting these DEGs, we identified the 'TNF signaling via NF- $\kappa$ B' as the most prominent pathway upregulated at 37 °C and that could be mitigated by transcriptional inhibition (Fig. 2i). Consistent with our previous observations, genes downregulated at 37 °C associated strongly with cell cycle progression (Fig. 2i and Extended Data Fig. 3b).

Finally, to assess the impact of IERSig on the in vivo function of HSPCs, we competitively transplanted cKIT<sup>+</sup> BM cells incubated on ice or at 37 °C for 90 min (Fig. 2j and Extended Data Fig. 3c). This revealed no functional consequences of IERSig induction on long-term HSC activity (Fig. 2j).

In this study, we highlighted variable activation of the IERSig across different research studies on HSC aging. Although some of these discrepancies may be due to variations in experimental design and methods across different studies (Supplementary Table 3), our findings underscore that HSPC exposure to stress triggers a swift induction of the IERSig, which is independent of the cells' age. Consequently, variations in handling of cells during isolation and the time elapsed until RNA extraction are probable factors contributing to the induction of the IERSig in primary HSPCs.

We present our findings primarily in the context of HSC aging, but we have also observed that samples from the reference sets on defined HSPCs, including BloodSpot<sup>22</sup>, Immunological Genome Project (ImmGen)<sup>23</sup>, Gene Expression Commons<sup>24</sup> and data in the Tabula Muris Senis project<sup>25</sup>, all associate with the IERSig to various degrees (Extended Data Fig. 4a). Furthermore, in a recent extensive large-scale RNA-seq dataset on murine immune cells (ImmGen ULI)<sup>26</sup>, we observed that while a high IERSig associated with samples exposed to inflammatory stimuli (including thioglycolate and lipopolysaccharide), probably due to an elevated immune activation, a markedly upregulated IERSig was also consistent in samples extracted from organs that require enzymatic treatment at elevated temperatures (Extended Data Fig. 4b). Crucially, the IERSig in all these evaluated datasets was more likely to show variability than random gene sets (Extended Data Fig. 4). Hence, while some transcriptional features, in this study related to HSC aging, can still be extracted from cells with a simultaneous IERSig, this requires for the most part previous knowledge. Also, the identity of all the genes comprising the IER in HSPCs and their relationship to subsequent waves of transcription are unclear. For instance, given the central role of quiescence in HSC biology, the correlation between induction of the IERSig with reductions in cell cycle-associated genes<sup>18,27</sup> is noteworthy.

While methods have been suggested to counteract the expression of IER genes caused by preparation, such as adding transcription inhibitors to isolation buffers<sup>20</sup>, we found that this can have its own drawbacks by affecting RNA quality (Fig. 2d) and the induction of other genes with specific attributes (Fig. 2h,i). Likewise, and in agreement with what has previously been observed for MuSCs<sup>19</sup>, in vivo cell fixation before isolation<sup>18,19</sup> led to changes in cell surface epitopes and substantial losses of recoverable cells (Extended Data Fig. 5a), with fixation also leading to poor RNA yield and integrity (Extended Data Fig. 5b). Instead, as a currently more effective approach for evaluating gene expression in primary HSPCs, we advocate for careful maintenance of primary HSPCs on ice during isolation, coupled with batch processing and a subsequent monitoring of the IERSig in the ensuing gene expression profiles.

## Methods

### Mice

Young (2–6 months) and aged (20 months) C57BL/6NTac (strain no. B6, Taconic Bioscience), C57BL/6.SJL (bred in house) and Nr4a1-GFP (strain no. 016617, The Jackson Laboratory) male and female mice were used. Mice were housed in a controlled environment with 12-h

light–dark cycles with chow and water provided ad libitum. All animal procedures were approved by the Lund University Ethics Committee (no. 16468-20).

### Reagents

All commercially available reagents used in this study are listed in Supplementary Table 5.

### Flow cytometry analyses and cell sorting

BM cells were isolated from the tibia, femur and pelvis into ice-cold FACS buffer (2% FCS/PBS) with or without 5  $\mu$ M TP (Tocris)<sup>28</sup> and filtered through 70- $\mu$ m cell strainers. Cells were cKIT-enriched by anti-cKIT-APC staining, followed by incubation with anti-APC MicroBeads and magnetic separation on LS columns (Miltenyi Biotec). Aliquots of cells were resuspended in culture medium (DMEM with high glucose, 10 mM HEPES, 2% FCS)  $\pm$  TP (0.1, 0.5, 1 and 5  $\mu$ M) and incubated on ice or at 37 °C for 90 min. After incubation, cells were stained with biotinylated antibodies against B220, Gr-1, TER119, CD3, NK1.1 and Sca-1 Pacific Blue (Sony Biotechnology) for 30 min on ice in the dark. For scRNA-seq, cells were additionally stained with oligo-conjugated hashing (HTO) antibodies (TotalSeq-A0301, -A0302, -A0303, and -A0304, BioLegend). Secondary staining was performed with streptavidin-Brilliant Violet 605 (Sony Biotechnology). After staining, LSK HPSCs were FACS-purified on a BD FACS Aria III instrument (BD Biosciences) and subjected to single-cell profiling (10x Genomics) or analyzed on an LSRFortessa X20 analyzer with the FACS Diva software v.9.0 (BD Biosciences). For PB analysis using flow cytometry, blood samples were sedimented with 1% Dextran T500 (Sigma-Aldrich) for 30 min at 37 °C; the remaining erythrocytes were lysed using ammonium chloride solution (STEM-CELL Technologies) for 3 min at room temperature. Cells were stained with CD19-PE-Cy7, TER119-PerCP-Cy5.5, CD11B-APC, NK1.1-Pacific Blue, CD3-Alexa Fluor 700, CD45.1-Brilliant Violet 650 and CD45.2-Brilliant Violet 785 (Sony Biotechnology). For the analysis of in vivo-fixed HSPCs, unfractionated BM cells were stained with PE-Cy5-conjugated B220, Gr-1, TER119, CD3, NK1.1, Sca-1 Pacific Blue, CD48-FITC, CD150-PE-Cy7, CD135-PE (Sony Biotechnology) and cKIT-APC-eFluor 780, CD201-APC (eBioscience) antibodies. Before analysis or sorting, cells were stained with propidium iodide (1:1,000, Invitrogen) to exclude dead cells. Data analysis was done using FlowJo v.10.5.3 (FlowJo LLC).

### Transplantation experiment

cKIT-enriched BM cells derived from C57BL/6.SJL-CD45.1 donor mice were isolated into ice-cold FACS buffer and incubated on ice or at 37 °C for 90 min. Competitor cKIT-enriched BM cells derived from C57BL/6N-CD45.2 mice were kept on ice throughout the whole experiment. After incubation, donor and competitor cells were mixed at a 1:4 ratio and transplanted into lethally irradiated (9.5 Gy) C57BL/6N-CD45.2 recipient mice via intravenous injections. Recipient mice received antibiotic prophylaxis (ciprofloxacin-supplemented water, 125 mg l<sup>-1</sup>, Krka) for 2 weeks beginning on the day of irradiation. Donor-derived reconstitution was monitored in PB using flow cytometry.

### Analysis of in vivo-fixed HSPCs

Young (2 months old) C57BL/6 female mice were anesthetized with ketamine (MDS Animal Health)-xylazine (Bayer Animal Health) (10% KX, 10  $\mu$ l g<sup>-1</sup>); in vivo fixation was performed using transcardiac perfusion with PBS followed by 4% paraformaldehyde (PFA). BM cells were isolated from the tibia, femur and pelvis into ice-cold FACS buffer (2% FCS/PBS), filtered through 70- $\mu$ m cell strainers and processed as described in the 'Flow cytometry analyses and cell sorting' section.

### Quantitative PCR with reverse transcription

cKIT-enriched BM cells were isolated from young (2 months) and aging (20 months) C57BL/6 male mice and incubated at 37 °C for up to 16 h in DMEM High Glucose (Gibco) supplemented with 10% FCS, 0.1 mM

2-mercaptoethanol (Invitrogen), 1 $\times$  penicillin-streptomycin-glutamine (Invitrogen), 50 ng ml<sup>-1</sup> stem cell factor, 10 ng ml<sup>-1</sup> TPO and 10 ng ml<sup>-1</sup> Flt3L (all from PeproTech). After incubation, RNA was extracted using the Direct-zol RNA MicroPrep Kit with on-column DNase I treatment (Zymo Research). Reverse transcription was carried out using the qScript cDNA Supermix according to the manufacturer's instruction (Quantabio). The resulting complementary DNA (cDNA) was used for quantitative PCR with reverse transcription (RT-qPCR) with SsoAdvanced SYBR Green Supermix (BioRad Laboratories) and the following primers: *Actb*: 5'-CCACAGCTGAGAGGGAAATC-3' (forward), 5'-CTTCTC-CAGGGAGGAAGAGG-3' (reverse); *Dusp1*: 5'-ACCATCTGCCTTGCTTAC-CTC-3' (forward), 5'-CTCCGCCTCTGCTTCACAAA-3' (reverse); *Egr1*: 5'-CCTATGAGCACCTGACCACA-3' (forward), 5'-GAAGCGCCAGTATAG-GTGA-3' (reverse); *Fos*: 5'-ATGGGCTCTCCTGTCAACAC-3' (forward), 5'-GCTGTAACCGTGGGGATAA-3' (reverse); *Jun*: 5'-GGAACGACCTTC-TACGACGAT-3' (forward), 5'-GGGTACTGTAGCCGTAGGC-3' (reverse); *Nr4a1*: 5'-TTGAGTTCCGGCAAGCCTACC-3' (forward), 5'-GTGTACCCGTC-CATGAAGGTG-3' (reverse); *Zfp36*: 5'-CCCTCACCTACTTCGCTAC-3' (forward), 5'-ACTTGTGGCAGAGTTCGGTTT-3' (reverse). Analyses were performed in two independent experiments with  $n = 2$  biological replicates per group.

### Analysis of RNA integrity from fixed HSPCs

cKIT-enriched BM cells were isolated from young (2 months old) C57BL/6.SJL female mice and split into four groups (approximately 900,000 cells per sample). Untreated cells were kept on ice throughout the whole procedure. Cells from remaining samples were spun down, resuspended in 2% PFA in PBS and incubated at room temperature for 15 min. Cells were spun down at 400g for 5 min after adding wash buffer (PBS supplemented with 1% BSA and 40 U ml<sup>-1</sup> SUPERase<sup>•</sup>In, Invitrogen) and were subsequently resuspended in the wash buffer. For reverse crosslinking, cells were incubated either at room temperature for 5 min in wash buffer supplemented with 125 mM glycine (Serva) or at 56 °C for 60 min in wash buffer supplemented with 40 U ml<sup>-1</sup> proteinase K (Roche), followed by incubation on ice for at least 5 min. RNA was isolated using the Direct-zol RNA MicroPrep Kit with on-column DNase I treatment (Zymo Research). RNA integrity was measured using the RNA 6000 Pico Kit and a 2100 Bioanalyzer instrument (Agilent Technologies).

### scRNA-seq and bioinformatics analysis

scRNA-seq was performed in a single batch. BM cells were isolated from a pool of young (2–3-month-old) C57BL/6.SJL female mice ( $n = 5$ ), split into four groups and processed simultaneously. cDNA libraries were generated using the Chromium Next GEM Single Cell 3' Reagent Kit v.3.1 (10x Genomics). Briefly, LSK cells (40,000 cells for each group) were sorted into ice-cold FACS buffer, resuspended at 1,250 cells per microliter in FACS buffer and 54,000 cells were loaded onto the Chromium Controller (10x Genomics). The cDNA sequencing library was generated according to the manufacturer's instructions (10x Genomics). The HTO-derived sequencing library was prepared according to CITE-seq\_and\_Hashing\_protocol\_190213 (<https://cite-seq.com/protocols>). Final cDNA-derived and HTO-derived libraries were pooled and sequenced on an Illumina NovaSeq 6000 instrument using the S2 Reagent Kit v.1.5 (100 cycles). Preparation of libraries and sequencing were done at the Center for Translational Genomics at Lund University.

After sequencing libraries were processed using Cell Ranger v.7.0.0. The FASTQ files were aligned to the mouse reference genome (mm10) to create unique molecular identifier count tables of gene expression. HTO demultiplexing was performed with cellhashr (<https://github.com/BimberLab/cellhashr>). Quality control and downstream analyses were performed using Seurat<sup>29</sup> v.4.1.1. The filtering thresholds used for cell barcode exclusion were as follows: a high number of mitochondrial transcripts (>6%); low-quality or empty droplets (the number of detected genes below 2,000 or above 9,000

and the number of detected transcripts below 5,000 or above 60,000); and the number of HTO transcripts below 20 or above 400. Motivated by the simultaneous processing of all cells and the near complete overlap between two of the sample groups (the two groups where cells were kept on ice; Fig. 2e), no additional sample integration was performed. Identification of highly variable genes, data normalization and dimensionality reduction were performed using Seurat<sup>29</sup> with default parameters.

Cell cycle analysis was carried out using the CellCycleScoring function from Seurat. For the analysis of gene signatures, the aggregated expression of the respective genes was calculated using the AddModuleScore function from Seurat; cells were classified based on the aggregated expression using a modified version of the CellCycleScoring function that output only two classifications. The aggregated values for genes were calculated using a bin normalization approach for each individual gene and visualized on the UMAP embeddings through the FeaturePlot function.

DEGs were analyzed using Seurat's FindMarkers function with default parameters and testing performed only on highly variable genes (two-sided Wilcoxon rank-sum test). DEG testing between cells incubated on ice with or without TP yielded three DEGs with nonsignificant *P* values; therefore, these samples were treated as one ('ice') for subsequent comparisons. As the dataset showed clear separation based on cell cycle phase classification, the samples were first compared cell cycle phase-wise. The DEGs between 37 °C with TP versus ice, 37 °C without TP versus ice across three cell cycle phases (G<sub>1</sub>/S/G<sub>2</sub>-M) were next used for MSigDB pathway analysis using Enrichr (<https://maayanlab.cloud/Enrichr/>).

All analysis steps except for the BCL to FASTQ conversion and the Cell Ranger run can be found in a Snakemake pipeline in the NCBI repository, with included conda (<https://zenodo.org/record/4774217>) environment specifications and version numbers of all packages used.

### Analysis of published microarray and bulk RNA-seq data

The analysis was performed using R v.4.1.0, v.4.1.3 and v.4.3.2. For the microarray data, raw.CEL files were retrieved from the Gene Expression Omnibus (GEO) and RMA-normalized using the affy v.1.76.0 or oligo v.1.62.2 packages. For bulk RNA-seq data analysis, count tables were downloaded from the GEO and data were normalized in R using the DESeq2 (ref. 30) v.1.38.3. When raw gene counts were not available, the normalized count tables provided by the authors were used. Genes with average expression counts less than the total number of samples were filtered out. GSEA<sup>31</sup> was performed using normalized data of the input files and the following gene sets: MSigDB Hallmark provided by the software<sup>5</sup> (MSigDB v.2022.1.Mm); custom-generated gene sets for IER and cell cycle; and aging and young signatures retrieved from Flohr Svendsen et al.<sup>2</sup>. All analyses were performed using the GSEA software v.4.3.2 (<https://www.gsea-msigdb.org/gsea/index.jsp>). The false discovery rate (FDR) method was applied to correct multiple hypotheses testing. Mining of gene lists was performed using the online tool Enrichr<sup>32</sup> (<https://maayanlab.cloud/Enrichr/>) with default parameters. Graphs were generated using Prism v.9.5.1 (GraphPad Software).

For the analysis of stress-associated genes from van Velthoven et al.<sup>19</sup> and Wu et al.<sup>20</sup>, data were downloaded from the GEO or using the supplementary tables provided by the authors. The DEGs with an FDR < 0.1 and fold change > 2 were used for comparison between IERsig genes identified in this study. The odds ratios (ORs) were computed in R by calculating the hypergeometric distribution, where the number of IERsig genes was 78 and the total number of genes was 19,570 (defined based on the total number of annotated genes on Affymetrix 430.2 microarrays from which the IERsig gene list was curated). The calculations of *P* values for the hypergeometric distributions were performed in R using a Fisher's exact test.

For the analysis of the IERsig in the reference datasets on murine HSPCs, the R packages dplyr v.1.1.3 and ggplot2 v.3.4.4 were used.

Briefly, the fraction of the IERsig in each sample was calculated by dividing the sum of gene expression levels for the IERsig genes by the total sum of gene expression levels in each sample. Data were visualized as dot plots with the fraction of IERsig across samples. To assess how the IERsig compared to randomly sampled gene sets, bootstrapping was used with a total of 10,000 iterations for each dataset. For reproducibility, a random seed (123) was set. In each iteration, a set of genes (RNA-seq data) or probe sets (microarrays), corresponding to the size of the IERsig, were randomly sampled from the gene expression data. The fraction of the bootstrapped genes in each sample was calculated in the same manner as for the IERsig. Descriptive statistics were computed for both the bootstrapped results and the IER signature fractions. To assess how the IER signature compared to the bootstrapped values, the centile rank of the coefficient of variation (CV) of the IER signature within the distribution of bootstrap CVs was computed and converted to a percentage. To visualize the results, histograms were generated to display the distribution of bootstrap CVs and the CV of the IER signature.

### Statistics and reproducibility

Data were analyzed using R v.4.1.0, v.4.1.3 and v.4.3.2, Microsoft Excel v.16.54 and Prism v.9.5.1 (GraphPad Software). All experiments were repeated as indicated; *n* indicates the number of independent biological replicates. No statistical methods were used to predetermine sample sizes but our sample sizes are similar to those reported in previous publications<sup>3,4</sup>. For the comparisons presented in Fig. 2d,h and Extended Data Fig. 5a, normality and equal variances were formally tested and data met the assumptions of the statistical tests used. Data collection and analysis were not performed blinded to the conditions of the experiments. For the transplantation and in vivo fixation experiments, no randomization method was applied to allocate animals to the experimental groups. Animals (*n* = 3–5) in the same cage received the same treatment. For the analyses of Nr4a1-GFP transgenic mice and the RT-qPCR experiments, no randomization method was applied and mice per sample were assigned to the experimental groups based on genotype or age. No animals or data were excluded from the analyses. For the flow cytometry analyses and RT-qPCR experiments, the analyses were performed in two or three independent experiments; no inconsistent results were observed. All results were successfully reproduced. The specific statistical test used for each experiment is indicated in the corresponding figure legend.

### Reporting summary

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

### Data availability

The original scRNA-seq data have been deposited in GEO under the accession no. [GSE224590](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE224590). The published datasets used for the analysis were retrieved from the GEO using the following accession numbers: HSC aging datasets ([GSE27686](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE27686), [GSE44923](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE44923), [GSE55525](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE55525), [GSE6503](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6503), [GSE48893](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48893), [GSE39553](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39553), [GSE47817](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE47817), [GSE127522](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE127522), [GSE128050](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128050), [GSE151333](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE151333), [GSE156807](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156807), [GSE109546](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109546), [GSE157455](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE157455) and [GSE165982](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165982)); datasets on muscle and neuronal cells ([GSE97399](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97399), [GSE103976](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE103976), [GSE15907](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE15907)); reference sets on murine HSPCs: [GSE14833](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE14833) and [GSE6506](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6506) (bloodspot microarray); [GSE34723](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE34723) (gene expression commons microarray); [GSE15907](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE15907) (ImmGen microarray); [GSE109125](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109125) (ImmGen ULI RNA-seq); and [GSE132042](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132042) (Tabula Muris Senis RNA-seq data from BM cells). Other data generated in this study are available in the Source data files. Source data are provided with this paper.

### Code availability

The code for the bioinformatics analysis, alongside information on software and package versions, is available at [https://github.com/razofz/DB\\_AKC\\_citeseq](https://github.com/razofz/DB_AKC_citeseq) and in the Supplementary Software file.

## References

- de Haan, G. & Lazare, S. S. Aging of hematopoietic stem cells. *Blood* **131**, 479–487 (2018).
- Flohr Svendsen, A. et al. A comprehensive transcriptome signature of murine hematopoietic stem cell aging. *Blood* **138**, 439–451 (2021).
- Norddahl, G. L. et al. Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. *Cell Stem Cell* **8**, 499–510 (2011).
- Wahlestedt, M. et al. An epigenetic component of hematopoietic stem cell aging amenable to reprogramming into a young state. *Blood* **121**, 4257–4264 (2013).
- Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
- Tullai, J. W. et al. Immediate-early and delayed primary response genes are distinct in function and genomic architecture. *J. Biol. Chem.* **282**, 23981–23995 (2007).
- Bersenev, A. et al. Lnk deficiency partially mitigates hematopoietic stem cell aging. *Aging Cell* **11**, 949–959 (2012).
- Chambers, S. M. et al. Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. *PLoS Biol.* **5**, e201 (2007).
- Flach, J. et al. Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
- Kim, K. M. et al. Taz protects hematopoietic stem cells from an aging-dependent decrease in PU.1 activity. *Nat. Commun.* **13**, 5187 (2022).
- Mansell, E. et al. Mitochondrial potentiation ameliorates age-related heterogeneity in hematopoietic stem cell function. *Cell Stem Cell* **28**, 241–256 (2021).
- Maryanovich, M. et al. Adrenergic nerve degeneration in bone marrow drives aging of the hematopoietic stem cell niche. *Nat. Med.* **24**, 782–791 (2018).
- Renders, S. et al. Niche derived netrin-1 regulates hematopoietic stem cell dormancy via its receptor neogenin-1. *Nat. Commun.* **12**, 608 (2021).
- Sun, D. et al. Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* **14**, 673–688 (2014).
- Sun, X. et al. Nicotinamide riboside attenuates age-associated metabolic and functional changes in hematopoietic stem cells. *Nat. Commun.* **12**, 2665 (2021).
- Yamashita, M. & Passegué, E. TNF- $\alpha$  coordinates hematopoietic stem cell survival and myeloid regeneration. *Cell Stem Cell* **25**, 357–372 (2019).
- Young, K. et al. Decline in IGF1 in the bone marrow microenvironment initiates hematopoietic stem cell aging. *Cell Stem Cell* **28**, 1473–1482 (2021).
- van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
- van Velthoven, C. T. J., de Morree, A., Egner, I. M., Brett, J. O. & Rando, T. A. Transcriptional profiling of quiescent muscle stem cells in vivo. *Cell Rep.* **21**, 1994–2004 (2017).
- Wu, Y. E. et al. Detecting activated cell populations using single-cell RNA-seq. *Neuron* **96**, 313–329 (2017).
- Moran, A. E. et al. T cell receptor signal strength in T<sub>reg</sub> and iNKT cell development demonstrated by a novel fluorescent reporter mouse. *J. Exp. Med.* **208**, 1279–1289 (2011).
- Bagger, F. O. et al. BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic Acids Res.* **44**, D917–D924 (2016).
- Heng, T. S. P. et al. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
- Seita, J. et al. Gene Expression Commons: an open platform for absolute gene expression profiling. *PLoS ONE* **7**, e40321 (2012).
- Klumpe, S. et al. A modular platform for automated cryo-FIB workflows. *eLife* **10**, e70506 (2021).
- Yoshida, H. et al. The cis-regulatory atlas of the mouse immune system. *Cell* **176**, 897–912 (2019).
- Land, R. H. et al. The orphan nuclear receptor NR4A1 specifies a distinct subpopulation of quiescent myeloid-biased long-term HSCs. *Stem Cells* **33**, 278–288 (2015).
- Chen, F., Gao, X. & Shilatifard, A. Stably paused genes revealed through inhibition of transcription initiation by the TFIID inhibitor triptolide. *Genes Dev.* **29**, 39–47 (2015).
- Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
- 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).
- 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).
- Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
- Beeraman, I., Seita, J., Inlay, M. A., Weissman, I. L. & Rossi, D. J. Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* **15**, 37–50 (2014).

## Acknowledgements

M. Wahlestedt and G. Norddahl generated the microarray data in refs. 3 and 4. The Nr4a1-GFP transgenic mice were kindly provided by B. Johansson Lindbom. We thank N. Liu and I. Lundgaard for help with the in vivo fixation procedures. We thank M. Eldeeb, Q. Zhang, O. Yuan, A. Rundberg-Nilsson and M. Wahlestedt for discussions and critical reading of the manuscript. We thank V. Anthonydhasan and P. Dhapola for suggestions on the bioinformatics analyses. We thank V. Nordberg for help with the literature review and discussions. This work was supported by grants from the Swedish Cancer Society (grant no. 211470Pj), the Pediatric Leukemia Foundation (grant no. PR2022-0091), the Swedish Research Council (grant no. 2022-00932) to D.B. and from the Royal Physiographic Society of Lund Foundation (grant no. 42335) to A.K.-C. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

A.K.-C. and D.B. designed the research, analyzed the data, created the figures, wrote the manuscript and acquired the funding. A.K.-C. and S.K. performed the experiments. R.O., A.K.-C. and D.B. performed the bioinformatics analysis. D.B. supervised the project.

## Funding

Open access funding provided by Lund University.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s43587-023-00558-z>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43587-023-00558-z>.

**Correspondence and requests for materials** should be addressed to David Bryder.

**Peer review information** *Nature Aging* thanks Gerald de Haan, Karl Lenhard Rudolph and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

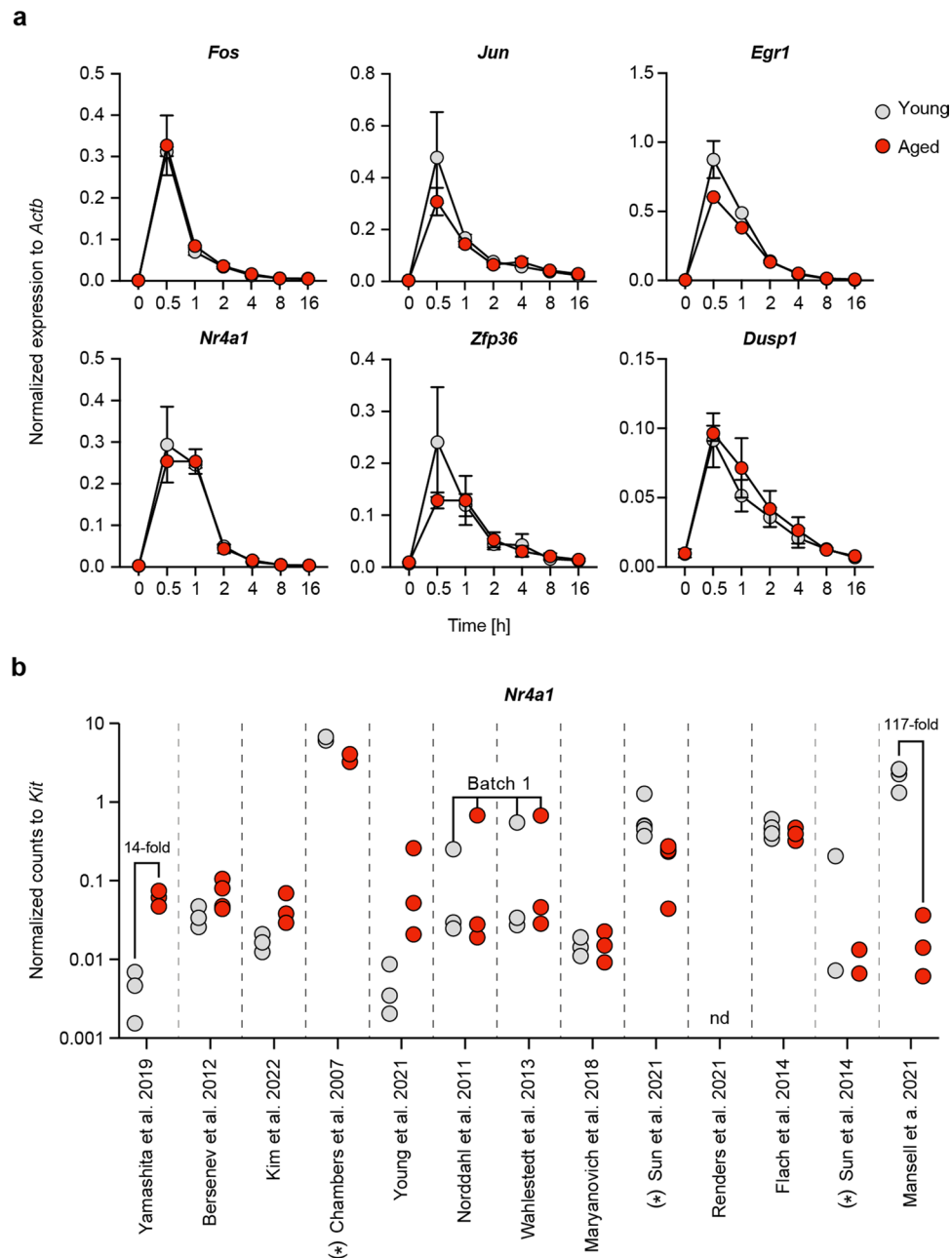
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**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/>.

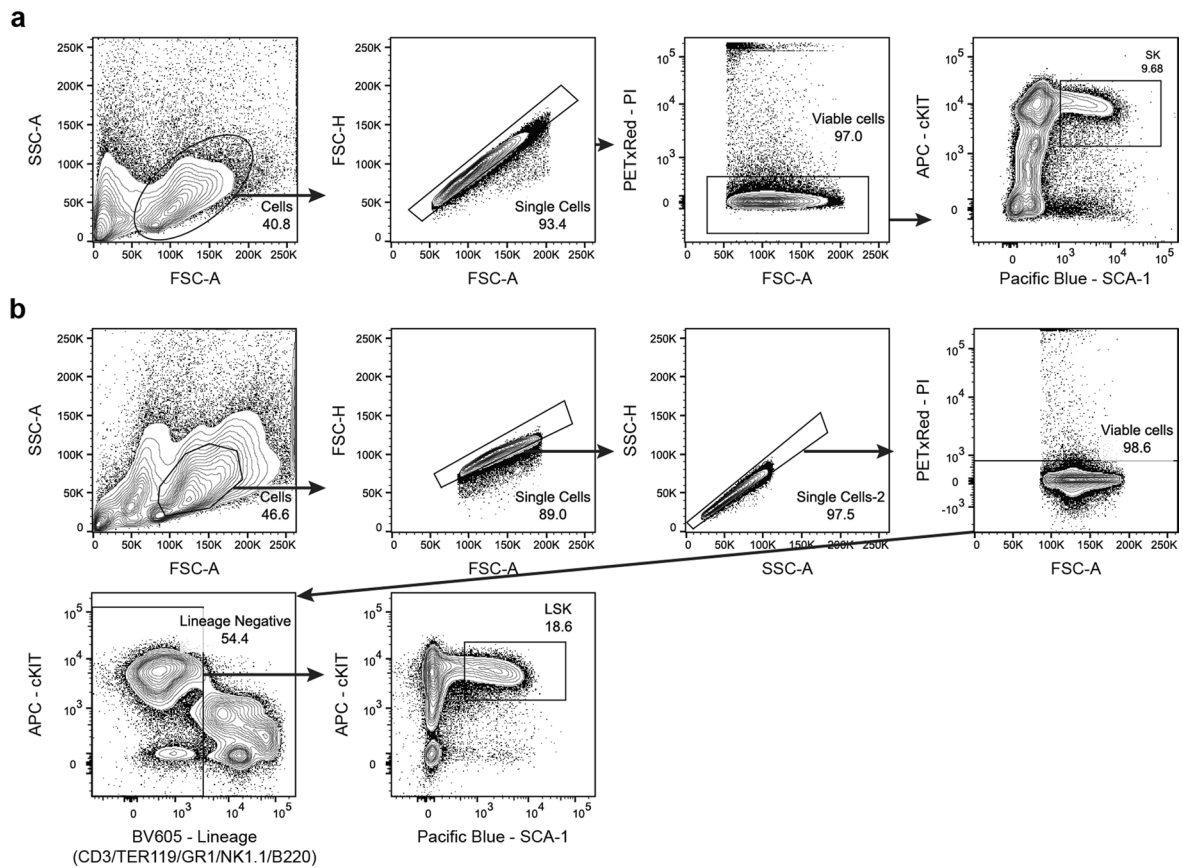
© The Author(s) 2024



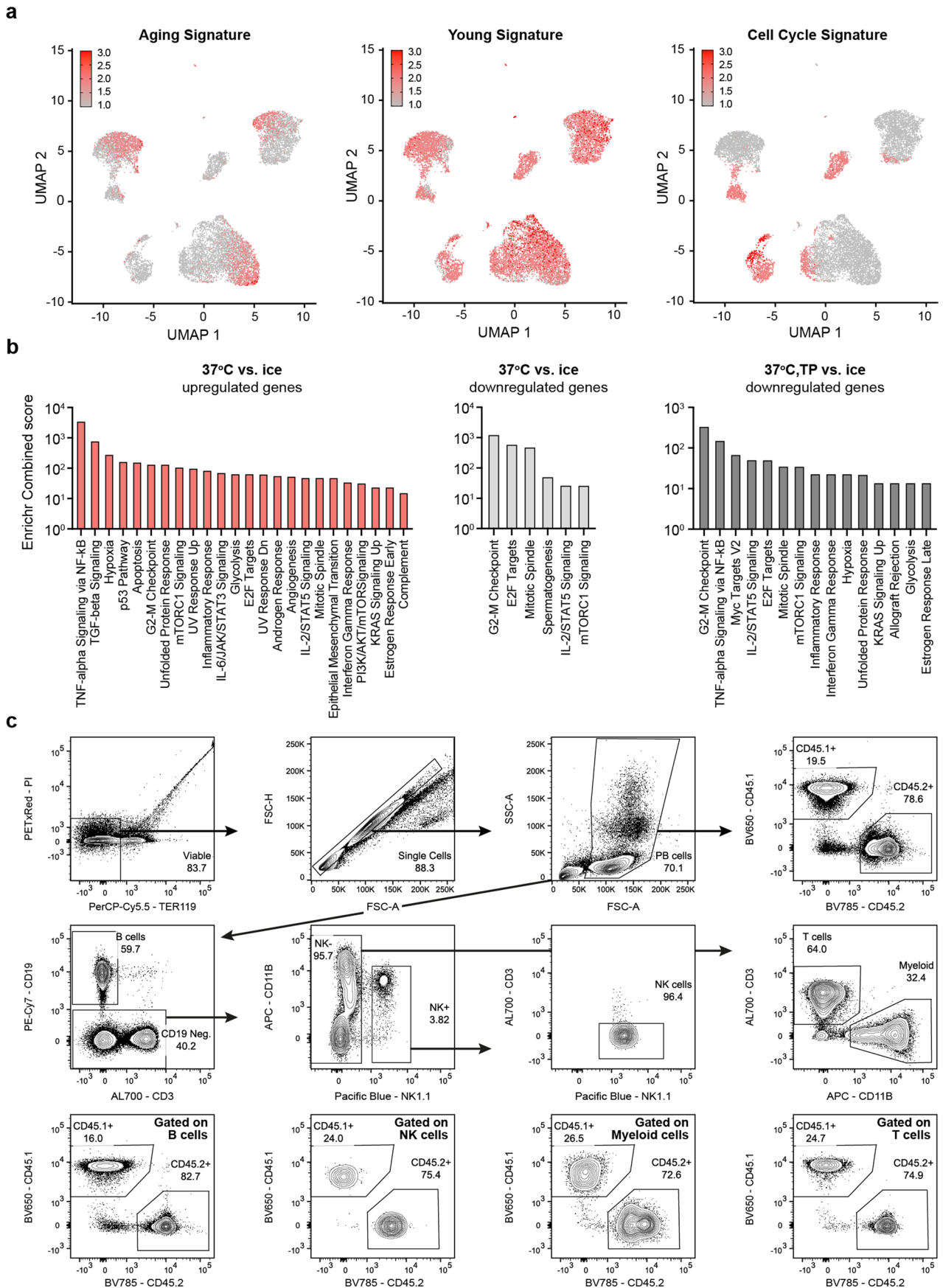


**Extended Data Fig. 1 | Validation of the IER signature. a)** Time course expression of selected IERSig genes in *ex vivo* cultured HSPCs from young (2 months) and aged (20 months) mice. Error bars denote mean ± SEM (n = 2 mice/group analyzed in two independent experiments). Related to Fig. 1d. **b)** Expression of *Nr4a1* in young and aged cHSCs in the indicated datasets. The *Nr4a1* read counts were normalized to *Kit*. The asterisks (\*) mark datasets that

were generated using cell isolation procedures that involve incubations at 37 °C (see also Supplementary Table 2). Fold differences in *Nr4a1* expression in young and aged cHSCs in Yamashita et al.<sup>16</sup> and Mansell et al.<sup>11</sup> are indicated above each comparison. Samples derived from batch 1 in Norddahl et al.<sup>3</sup> and Wahlestedt et al.<sup>4</sup> are also indicated. nd, not detected.



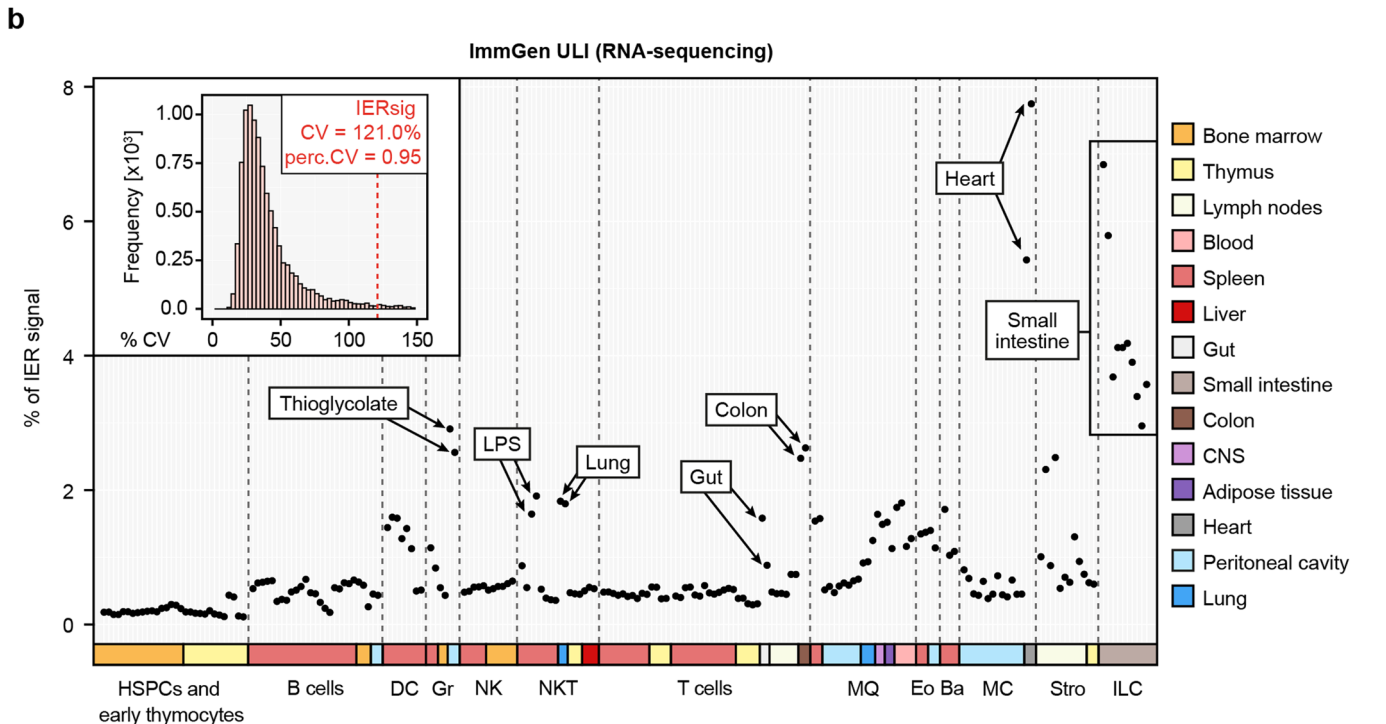
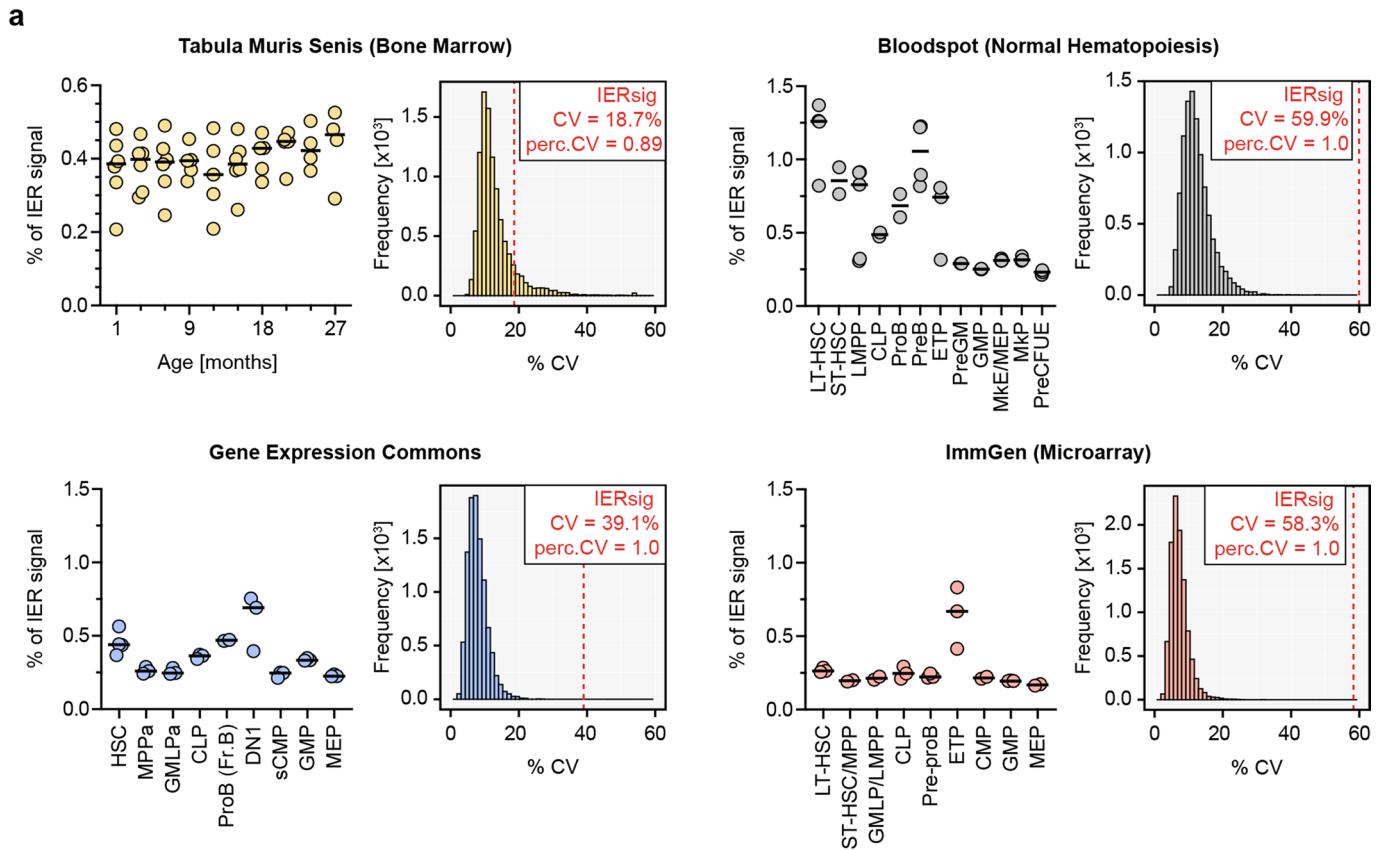
**Extended Data Fig. 2 | The gating strategy for flow cytometry analysis and FACS sorting. a)** The gating strategy for the analysis of Nr4a1-GFP signal in bone marrow SCA-1+cKIT+ HSPCs. Related to Fig. 2b. **b)** The gating strategy for bone marrow LSK cell sorting for single cell RNA sequencing experiment. Related to Fig. 2c.



Extended Data Fig. 3 | See next page for caption.

**Extended Data Fig. 3 | The effect of cell isolation procedures on induction of the IER signature.** **a)** UMAP embeddings of LSK cells with the aggregated expression of the age-induced, age-repressed and cell cycle-related genes depicted in red. Related to Fig. 2f. **b)** The MSigDB Hallmark 2020 pathways that associate with genes induced and repressed in HSPCs incubated at 37 °C ± TP compared to cells kept on ice. Shown are Enrichr combined scores (computed

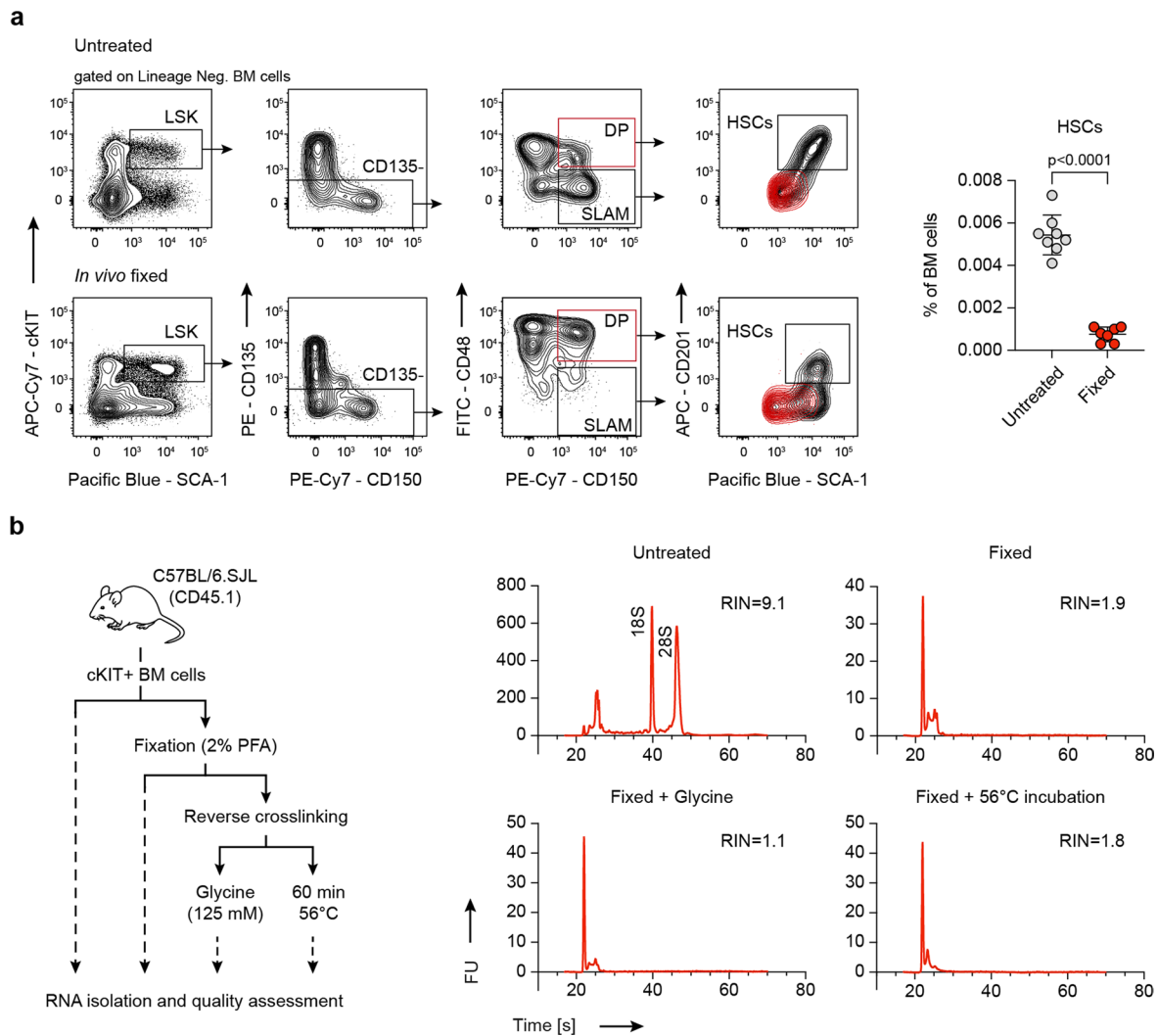
by Enrichr software) for pathways with an adjusted p-value < 0.05 (one-sided Fisher's exact test with Benjamini-Hochberg correction). No significant pathways were associated with genes upregulated in HSPCs incubated at 37 °C + TP. Detailed statistical information is provided in Supplementary Table 2. Related to Fig. 2i. **c)** Representative flow cytometry analysis of donor-derived chimerism in indicated peripheral blood lineages. Related to Fig. 2j.



**Extended Data Fig. 4 | Expression of the IER signature in indicated datasets.**

**a)** Dot plots depicting the cumulative expression of IERsig genes in BM samples across various age groups and different HSPC subsets in indicated datasets. Data points indicate individual mice and bars denote median. Histograms display the distribution of bootstrap coefficient of variation (CV) and the CV of the IERsig in the respective datasets. Data derive from the Tabula Muris Senis<sup>25</sup>, Immunological Genome Project (ImmGen, microarray)<sup>23</sup>, Bloodspot Mouse Normal Hematopoietic System<sup>22</sup> and Gene Expression Commons<sup>24</sup> datasets. **b)** Aggregated expression of IERsig genes in distinct hematopoietic cell

populations isolated from various organs. Samples derived from mice exposed to inflammatory stimuli (Thioglycolate and lipopolysaccharide, LPS), or from selected organs that require enzymatic treatment are indicated. The histogram in the top-left corner represents the distribution of bootstrap coefficient of variation (CV) and the CV of the IERsig. Data derive from Immunological Genome Project (ImmGen ULI RNA sequencing)<sup>26</sup>. Ba, basophils; DC, dendritic cells; Eo, eosinophils; Gr, granulocytes; ILC, innate lymphoid cells; NK, natural killer cells; NKT, natural killer T cells; MC, mast cells; MQ, macrophages; Stro, stromal cells



**Extended Data Fig. 5 | The effect of fixation on HSPC phenotypes and RNA quality. a** Evaluation of HSC phenotypes (*left*) and frequencies (*right*) in C57BL/6 mice following *in vivo* fixation with paraformaldehyde (n = 8). Untreated mice served as controls (n = 7). LSK CD135-CD48+CD150+ population (DP, red) was used as internal negative control to set the gate for CD201. Analyses were

performed in three independent experiments. Error bars denote mean  $\pm$  SD,  $p = 1.4 \times 10^{-8}$  (two-sided Student's *t*-test). **b** Analysis of RNA recovery and integrity following isolation from untreated and *ex vivo* fixed HSPCs. Experimental strategy (*left*) and electropherograms representing RNA profiles (*right*). RNA was 10x diluted prior to analysis. FU, arbitrary fluorescence units

## Reporting Summary

Nature Portfolio 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 Portfolio 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

Published datasets were retrieved from Gene Expression Omnibus.  
Flow cytometry data were collected using BD LSRFortessa X-20 and BD FACS Aria III instruments and FACS Diva software v9.0 (BD Bioscience).  
Single cell RNA sequencing data were collected using Illumina NovaSeq6000 instrument.

Data analysis

For the analysis of published transcriptomic datasets, the following software were used: Cell Ranger v7.0.0, R v4.1.0, v4.1.3 and v4.3.2, Seurat, v4.1.1, affy v1.76.0, oligo v1.62.2, DESeq2 v1.38.3, MSigDB v2022.1.Mm, GSEA v4.3.2, dplyr v1.1.3, ggplot2 v3.4.4, Enrichr (<https://maayanlab.cloud/Enrichr/>), GraphPad Prism v9.5.1, Microsoft Excel v16.54. The FASTQ files were aligned to the mouse reference genome (mm10). All code for the single cell RNA sequencing analysis, with accompanying conda environments with package version specifications, can be found at [https://github.com/razofz/DB\\_AKC\\_citeseq](https://github.com/razofz/DB_AKC_citeseq) and in the Supplementary Software file.  
Flow cytometry data were analyzed using FlowJo v10.5.3 (Treestar).

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Original single-cell RNA-sequencing data is deposited in GEO under the accession number GSE224590. Published datasets used for the analysis were retrieved from GEO using accession numbers: HSC aging datasets: GSE27686, GSE44923, GSE55525, GSE6503, GSE48893, GSE39553, GSE47817, GSE127522, GSE128050, GSE151333, GSE156807, GSE109546, GSE157455 and GSE165982; datasets on muscle and neuronal cells: GSE97399, GSE103976, GSE15907; reference sets on murine HSPCs: GSE14833 and GSE6506 (Bloodspot microarray), GSE34723 (Gene Expression Commons microarray), GSE15907 (ImmGen microarray) and GSE109125 (ImmGen ULI RNA-sequencing) and GSE132042 (Tabula Muris Senis RNA-sequencing data from BM cells). Other data generated in this study is available in the Source Data files.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

## 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

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size	No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications (Norddahl et al. 2011, Wahlestedt et al. 2013). For mice experiments, the number of mice is indicated in the corresponding figure legends. For single cell RNA sequencing, 40,000 LSK bone marrow cells were isolated from each sample and pooled prior to sequencing. In total 56,000 cells were loaded on 10x Chromium platform. For transplantation experiment, 250,000 donor and 1,000,000 competitor cKIT+ bone marrow cells were mixed and transplanted into each recipient mouse. For quantitative RT-PCR experiments, 5,000,000 cKIT+ bone marrow cells were isolated from young and aged mice and split into samples for time course analyses (700,000 cells per each sample). For ex vivo HSPC fixation, 4,000,000 cKIT+ bone marrow cells were isolated from young mice and split into four samples (~900,000 cells per each sample) for subsequent RNA integrity analyses.
Data exclusions	No data were excluded.
Replication	For flow cytometry analysis of Nr4a1-GFP signal in bone marrow cells, a total of n= 3-4 mice/genotype were analyzed in three independent experiments. Quantitative RT-PCR analyses were performed in two independent experiments. For HSPC analysis following in vivo fixation, a total of n=7-8 mice/group were analyzed in three independent experiments. For all experiments no inconsistent results were observed. The results were successfully replicated. Transplantation analysis was performed in one replicate with n=5 recipients/group.
Randomization	Adult (8-30 weeks old) and aged (20 months) female and male mice were used. For transplantation experiment, female mice were used due to housing and maintenance convenience. For transplantation and in vivo fixation experiments, a simple randomization procedure was applied to cages of mice, such that n = 3-5 animals per cage were assigned to each group and all animals within a cage received the same treatment. For analyses of Nr4a1-GFP transgenic mice and quantitative real-time PCR experiments, no randomization method was applied and the mice/samples were assigned to experimental groups based on genotype or age.



## 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

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

CD11b – APC (Sony Biotechnology, Cat# 1106060, clone M1/70), dilution 1:800  
 CD117 - APC (Sony Biotechnology, Cat# 1129060, clone 2B8), dilution 1:200  
 CD117 - APCeFluor780 (eBioscience, Cat# 47-1171-82, clone 2B8), dilution 1:200  
 CD135 - PE (Sony Biotechnology, Cat# 1276530, clone A2F10), dilution 1:100  
 CD150 - PE/Cy7 (Sony Biotechnology, Cat# 1179570, clone TC15-12F12.2), dilution 1:200  
 CD19 - PE/Cy7 (Sony Biotechnology, Cat# 1177600, clone 6D5), dilution 1:200  
 CD201 - APC (eBioscience, Cat# 17-2012-82, clone eBio1560), dilution 1:200  
 CD3 - Alexa Fluor 700 (Sony Biotechnology, Cat# 1101080, clone 17A1), dilution 1:200  
 CD3 - Biotin (Sony Biotechnology, Cat# 1101220, clone 17A1), dilution 1:200  
 CD3 - PE/Cy5 (Sony Biotechnology, Cat# 1101550, clone 145-2C11), dilution 1:200  
 CD45.1 - Brilliant Violet 650 (Sony Biotechnology, Cat# 1153680, clone A20), dilution 1:100  
 CD45.2 - Brilliant Violet 785 (Sony Biotechnology, Cat# 1149195, clone 104), dilution 1:100  
 CD45R/B220 - Biotin (Sony Biotechnology, Cat# 1116020, clone RA3-6B2), dilution 1:200  
 CD45R/B220 - PE/Cy5 (Sony Biotechnology, Cat# 1116050, clone RA3-6B2), dilution 1:200  
 CD48 - FITC (Sony Biotechnology, Cat# 1117020, clone HM48-1), dilution 1:200  
 NK1.1 - Biotin (BD Bioscience, Cat# 553163, clone PK136), dilution 1:200  
 NK1.1 - PE/Cy5 (Sony Biotechnology, Cat# 1143580, clone PK136), dilution 1:200  
 NK1.1 - Pacific Blue (Sony Biotechnology, Cat# 1143610, clone PK136), dilution 1:200  
 Ly6A/E (Sca-1) - Pacific Blue (Biolegend, Cat# 122520, clone E13-161.7), dilution 1:200  
 Ly6G/Ly6C (Gr-1) - Biotin (Sony Biotechnology, Cat# 1142020, clone RB6-8C5), dilution 1:400  
 Ly6G/Ly6C (Gr-1) - PE/Cy5 (Sony Biotechnology, Cat# 1142050, clone RB6-8C5), dilution 1:400  
 Streptavidin - Brilliant Violet 605 (Sony Biotechnology, Cat# 2626145), dilution 1:400  
 Ter119 - Biotin (Sony Biotechnology, Cat# 1181020, clone TER119), dilution 1:400  
 Ter119 - PE/Cy5 (Sony Biotechnology, Cat# 1181050, clone TER119), dilution 1:400  
 Ter119 - PerCP/Cy5.5 (Sony Biotechnology, Cat# 1181140, clone TER119), dilution 1:400  
 Hashtag 1 - TotalSeq-A0301 (Biolegend, Cat# 155801, clone M1/42; 30-F11), 0.5 ug/sample (up to 10 mln cells)  
 Hashtag 2 - TotalSeq-A0302 (Biolegend, Cat# 155803, clone M1/42; 30-F11), 0.5 ug/sample (up to 10 mln cells)  
 Hashtag 3 - TotalSeq-A0303 (Biolegend, Cat# 155805, clone M1/42; 30-F11), 0.5 ug/sample (up to 10 mln cells)  
 Hashtag 4 - TotalSeq-A0304 (Biolegend, Cat# 155807, clone M1/42; 30-F11), 0.5 ug/sample (up to 10 mln cells)

### Validation

The antibodies were validated for the specified applications by the manufacturer. Validations profiles and references can be found in the provided links. All antibodies have been used in test panels prior to experimental procedures to determine optimal dilutions.  
 CD11b – APC; species reactivity: Mouse, Human; tested application: FC, product citations: 32; <https://www.sonybiotechnology.com/us/catalog/product/view/id/360/s/apc-anti-mouse-human-cd11b/>  
 CD117 – APC; species reactivity: Mouse; tested application: FC, product citations: 7; <https://www.sonybiotechnology.com/us/apc-anti-mouse-cd117-c-kit-13>  
 CD117 - APCeFluor780; species reactivity: Mouse, Human; tested application: FC, FN, IV; product citations: 88; <https://www.thermofisher.com/antibody/product/CD117-c-Kit-Antibody-clone-2B8-Monoclonal/47-1171-82>  
 CD135 – PE; species reactivity: Mouse; tested application: FC, product citations: 3; <https://www.sonybiotechnology.com/se/pe-anti-mouse-cd135-7>  
 CD150 - PE/Cy7; species reactivity: Mouse; tested application: FC, product citations: 22; <https://www.sonybiotechnology.com/se/pe-cy7-anti-mouse-cd150-slam-7>  
 CD19 - PE/Cy7; species reactivity: Mouse; tested application: FC, product citations: 12; <https://www.sonybiotechnology.com/se/pe-cy7-anti-mouse-cd19-7>  
 CD201 – APC; species reactivity: Mouse; tested application: FC, product citations: 12; <https://www.thermofisher.com/antibody/product/CD201-EPCR-Antibody-clone-eBio1560-1560-Monoclonal/17-2012-82>  
 CD3 - Alexa Fluor 700; species reactivity: Mouse; tested application: FC, product citations: 9; <https://www.sonybiotechnology.com/>

us/alexa-fluor-reg-700-anti-mouse-cd3-6  
 CD3 – Biotin; species reactivity: Mouse; tested application: FC, product citations: 8; <https://www.sonybiotechnology.com/se/biotin-anti-mouse-cd3-7>  
 CD3 - PE/Cy5; species reactivity: Mouse; tested application: FC, product citations: 31; <https://www.sonybiotechnology.com/se/pe-cy5-anti-mouse-cd3-epsilon-5>  
 CD45.1 - Brilliant Violet 650; species reactivity: Mouse; tested application: FC, product citations: 21; <https://www.sonybiotechnology.com/se/brilliant-violet-650-trade-anti-mouse-cd45-1-7>  
 CD45.2 - Brilliant Violet 785; species reactivity: Mouse; tested application: FC, product citations: 12; <https://www.sonybiotechnology.com/us/brilliant-violet-785-trade-anti-mouse-cd45-2-3>  
 CD45R/B220 – Biotin; species reactivity: Mouse, Human; tested application: FC, product citations: 13; <https://www.sonybiotechnology.com/se/biotin-anti-mouse-human-cd45r-b220-7>  
 CD45R/B220 - PE/Cy5; species reactivity: Mouse, Human; tested application: FC, product citations: 13; <https://www.sonybiotechnology.com/se/pe-cy5-anti-mouse-human-cd45r-b220-7>  
 CD48 – FITC; species reactivity: Mouse; tested application: FC, product citations: 6; <https://www.sonybiotechnology.com/se/fitc-anti-mouse-cd48-7>  
 NK1.1 – Biotin; species reactivity: Mouse; tested application: FC, product citations: 11; <https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/biotin-mouse-anti-mouse-nk-1-1.553163>  
 NK1.1 - PE/Cy5; species reactivity: Mouse; tested application: FC, product citations: 15; <https://www.sonybiotechnology.com/se/pe-cy5-anti-mouse-nk-1-1-7>  
 NK1.1 - Pacific Blue; species reactivity: Mouse; tested application: IF, product citations: 15; <https://www.sonybiotechnology.com/us/pacific-blue-trade-anti-mouse-nk-1-1-6>  
 Ly6A/E (SCA-1) - Pacific Blue; species reactivity: Mouse; tested application: FC, product citations: 31; <https://www.biolegend.com/en-us/products/pacific-blue-anti-mouse-ly-6a-e-sca-1-antibody-3901>  
 Ly6G/Ly6C (Gr-1) – Biotin; species reactivity: Mouse; tested application: FC, product citations: 21; <https://www.sonybiotechnology.com/se/biotin-anti-mouse-ly-6g-ly-6c-gr-1-7>  
 Ly6G/Ly6C (Gr-1) - PE/Cy5; species reactivity: Mouse; tested application: FC, product citations: 22; <https://www.sonybiotechnology.com/se/pe-cy5-anti-mouse-ly-6g-ly-6c-gr-1-7>  
 TER119 – Biotin; species reactivity: Mouse; tested application: FC, product citations: 8; <https://www.sonybiotechnology.com/se/biotin-anti-mouse-ter-119-erythroid-cells-7>  
 TER119 - PE/Cy5; species reactivity: Mouse; tested application: FC, product citations: 7; <https://www.sonybiotechnology.com/se/pe-cy5-anti-mouse-ter-119-erythroid-cells-7>  
 TER119 - PerCP/Cy5.5; species reactivity: Mouse; tested application: FC, product citations: 8; <https://www.sonybiotechnology.com/se/percp-cy5-5-anti-mouse-ter-119-erythroid-cells-6>  
 Hashtag 1 - TotalSeq-A0301; species reactivity: Mouse; tested application: PG, product citations: 13; <https://www.biolegend.com/en-us/products/totalseq-a0301-anti-mouse-hashtag-1-antibody-16103>  
 Hashtag 2 - TotalSeq-A0302; species reactivity: Mouse; tested application: PG, product citations: 11; <https://www.biolegend.com/en-us/products/totalseq-a0302-anti-mouse-hashtag-2-antibody-16104>  
 Hashtag 3 - TotalSeq-A0303; species reactivity: Mouse; tested application: PG, product citations: 9; <https://www.biolegend.com/en-us/products/totalseq-a0303-anti-mouse-hashtag-3-antibody-16105>  
 Hashtag 4 - TotalSeq-A0304, species reactivity: Mouse; tested application: PG, product citations: 8; <https://www.biolegend.com/en-us/products/totalseq-a0304-anti-mouse-hashtag-4-antibody-16106>

## Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	Mus musculus, Nr4a1-GFP (JAX, #016617), 22-32 weeks old, female mice Mus musculus, C57BL/6NTac (Taconic Biosciences, #B6), 8-12 weeks old and 20 months old, males and female mice Mus musculus, C57BL/6.SJL - CD45.1 (in-house breeding), 8-12 weeks old, female mice All mice were housed in a controlled environment with 12-hour light-dark cycles with chow and water provided ad libitum.
Wild animals	This study did not involve wild animals.
Reporting on sex	For transplantation experiments, analyses of Nr4a1-GFP reporter cells and single cell RNA sequencing, the cells were isolated from female mice. For quantitative RT-PCR experiments, the cells were isolated from male mice. For in vivo and ex vivo fixation, the cells were isolated from female mice.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	All animal experiments performed in this study were approved by the Lund University Ethical Committee (#16468-20)

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

## 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

Bone marrow cells were isolated from tibia, femur, and pelvis into ice-cold FACS buffer (2% FBS/PBS) with or without 5  $\mu$ M Triptolide (Tocris). The cells were cKIT-enriched by anti-cKIT-APC staining, followed by incubation with anti-APC MicroBeads and magnetic separation on LS columns (Miltenyi Biotec). Aliquots of cells were resuspended in culture media (DMEM with high glucose/10 mM HEPES/2% FBS) with or without Triptolide (0.1, 0.5, 1 and 5  $\mu$ M) and incubated on ice or at 37°C for 90 min. Following incubation, the cells were stained with biotinylated antibodies against B220, Gr-1, TER119, CD3, NK1.1, SCA-1-Pacific Blue (Sony Biotechnology) for 30 min on ice in the dark. For single cell RNA sequencing, the cells were additionally stained with oligo-conjugated hashing antibodies (TotalSeq-A0301, -A0302, -A0303, and -A0304, BioLegend). The secondary staining was performed with streptavidin-BV605 (Sony Biotechnology). For peripheral blood analysis by flow cytometry, blood samples were sedimented with 1% Dextran T500 (Sigma-Aldrich) for 30 min at 37°C, and remaining erythrocytes were lysed using ammonium chloride solution (STEMCELL Technologies) for 3 min at room temperature. The cells were stained with CD19-PE-Cy7 (dilution 1:200), TER119-PerCP-Cy5.5 (dilution 1:400), CD11B-APC (dilution 1:800), NK1.1-Pacific Blue (dilution 1:200), CD3-Alexa Fluor700 (dilution 1:200), CD45.1-BV650 (dilution 1:100) and CD45.2-BV785 (dilution 1:100) (Sony Biotechnology). For analysis of in vivo-fixed HSPCs, unfractionated BM cells were stained with PE-Cy5-conjugated B220, Gr-1, TER119, CD3, NK1.1, SCA-1-Pacific Blue, CD48-FITC, CD150-PE-Cy7, CD135-PE (Sony Biotechnology) and cKIT-APC-eFluor780, CD201-APC (eBioscience) antibodies. Prior to analysis or sorting, cells were stained with PI (1:1000, Invitrogen) to exclude dead cells.

Instrument

LSRFortessa X-20 analyzer (BD Bioscience)  
BD FACSAria III cell sorter (BD Bioscience)

Software

BD Diva software v9.0 (BD Bioscience)  
FlowJo 10 software v10.5.3 (Treestar)

Cell population abundance

For single cell RNA sequencing, 40,000 Lin-SCA-1+cKIT+ bone marrow cells were sorted from each sample. Sort purity was determined by re-analysis of sorted cells.

Gating strategy

For cell sorting, single cells were gated using FSC-A/FSC-H and live nucleated cells were gated based on SSC-A/FSC-A and PI staining. Lineage-, SCA-1+ and cKIT+ gating was set based on internal reference populations. For peripheral blood analysis, single cells were gated using FSC-A/FSC-H and live nucleated cells were gated based on SSC-A/FSC-A, PI and TER119 staining. Cells were subsequently gated based on CD19, CD3, CD11B and NK1.1 expression. Donor- and competitor-derived cells were discriminated by gating CD45.1+ (donor) and CD45.2+ (competitor) cells within mature peripheral blood cell subsets. Gating strategy for all flow cytometry and FACS sorting experiments are included in Extended Data Figures 2, 3 and 5.

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