



**HAL**  
open science

## **A latent subset of human hematopoietic stem cells resists regenerative stress to preserve stemness**

Kerstin B. Kaufmann, Andy G. X. Zeng, Etienne-Marie Coyaud, Laura Garcia-Prat, Efthymia Papalexi, Alex Murison, Estelle Laurent, Michelle Chan-Seng-Yue, Olga I. Gan, Kristele Pan, et al.

► **To cite this version:**

Kerstin B. Kaufmann, Andy G. X. Zeng, Etienne-Marie Coyaud, Laura Garcia-Prat, Efthymia Papalexi, et al.. A latent subset of human hematopoietic stem cells resists regenerative stress to preserve stemness. *Nature Immunology*, 2021, *Nature Immunology*, 22, pp.723-734. 10.1038/s41590-021-00925-1 . hal-04059078

**HAL Id: hal-04059078**

**<https://hal.univ-lille.fr/hal-04059078v1>**

Submitted on 5 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Figure #	Figure title One sentence only	Filename This should be the name the file is saved as when it is uploaded to our system. Please include the file extension. i.e.: <i>Smith_ED_Fig1.jpg</i>	Figure Legend If you are citing a reference for the first time in these legends, please include all new references in the main text Methods References section, and carry on the numbering from the main References section of the paper. If your paper does not have a Methods section, include all new references at the end of the main Reference list.
Extended Data Fig. 1	INKA1 and H4K16ac are mutually exclusive.	Kaufmann_Extended_Data_Fig1.tif	<p>(a) Representative sorting scheme to obtain LT<sup>-</sup>, ST<sup>-</sup>, IT90<sup>+</sup> IT90<sup>-</sup> HSCs or CD34<sup>+</sup>CD38<sup>-</sup> and CD34<sup>+</sup>CD38<sup>+</sup> cells from lineage depleted CB.</p> <p>(b) Confocal images of 2 individual ST-HSC to visualize heterogeneity and representative for n=3 independent CB pools analyzed. <i>Scale bar = 2 μm.</i></p> <p>(c) Confocal images of 4 individual, uncultured LT-HSC next to thereof derived images for INKA1 and H4K16ac channels after thresholding and binary conversion using ImageJ to visualize intracellular invers staining pattern in double dim cells of the heterogenous cell pool and representative for n=3 independent CB. <i>Scale bar = 4 μm.</i></p> <p>(d) Percentage INKA1<sup>+</sup> cells in HSC subsets according to INKA1<sup>+</sup> cell gate in Fig. 1d (immunostaining, Ø50 cells/population and CB, n=3 independent CB pools). <i>Mean±SD.</i></p> <p>(e) Confocal images and quantification (MFI) of prospectively isolated LT<sup>-</sup> and ST-HSC stained for INKA1 and CDK6 and representative for n=3 independent CB as summarized in right panel. Gating as used for Figure 1F is indicated. <i>Scale bar = 2 μm.</i></p>
Extended Data Fig. 2	Mobilized PB single cell RNA pseudotime analysis.	Kaufmann_Extended_Data_Fig2.tif	<p>(a) Cell cycle analysis (Ki-67, Hoechst, CDK6) of prospectively isolated HSC subsets (LT<sup>-</sup>, IT90<sup>+</sup>, IT90<sup>-</sup>, ST-HSC) from mPB (n=3) by flow cytometry. CB HSPC (CD34<sup>+</sup>CD38<sup>-</sup>) and Progenitors (Prog, CD34<sup>+</sup>CD38<sup>+</sup>) served as gating controls. <i>Mean±SD; two-sided paired t-test (**p=0.002).</i></p>

			(b-e) Single cell RNA-seq data from CD34 <sup>+</sup> CD38 <sup>-</sup> CD45RA <sup>-</sup> index-sorted mPB HSC (n=3, also stained and indexed for CD49f, CD90) visualized by diffusion mapping overlaid with (b) cell cycle score (Tirosh et al., 2016) and indicated transcript expression; (c) donor origin, (d) <i>ITGA</i> (CD49f) transcript expression, (e) CD49f and CD90 indexed protein surface expression and (f) retrospectively gated HSC subsets (LT-, IT90 <sup>+</sup> , IT90 <sup>-</sup> , ST-HSC).
Extended Data Fig. 3	Protein single cell pseudotime analysis.	Kaufmann_Extended_Data_Fig3.tif	<p>(a) Gating strategy for cell cycle phases of 2d with 20 hr EdU pulse cultured CB Lin<sup>-</sup>.</p> <p>(b,c) CB LT- and ST-HSC analyzed by immunostaining for Ki-67, H4K16ac, PAK4 and DAPI after 72 hrs of culture; plots in (h) show data combined from 3 independent CB pools (Ø50 cells/population and sample). <i>Scale bar = 4 µm.</i></p> <p>(d) Longitudinal confocal imaging of LT-HSC harvested at 5 timepoints of culture in high-cytokine conditions. Two individual LT-HSC are shown per timepoint. <i>Scale bar = 4 µm.</i></p> <p>(e) Individual cell fluorescence intensities (IntDen) of nuclear CDK6, H4K16ac, PAK4 and DAPI across pseudotime for HSC subsets (n=2913). Curve fit was performed with nonlinear regression excluding 1-2% of outliers (fifth order polynomial).</p> <p>(f) Pseudotime was split in early, mid and late fractions to visualize distribution of HSC subsets and Progenitors and marker levels across these fractions for all 3442 cells (histograms) or per population and CB (n=2-3, bar plots). Single cell distribution for each population across pseudotime is also shown in horizontal violin plot. <i>Mean±SD, two sided t-test (compared to “early”;</i> *<i>p</i><math>&lt;</math><i>0.05</i>; **<i>p</i><math>&lt;</math><i>0.01</i>; ***<i>p</i><math>&lt;</math><i>0.001</i>).</p>

Extended Data Fig. 4	Pharmacological PAK4 inhibition.	Kaufmann_Extended_Data_Fig4.tif	<p>(a) Brightfield images of HSPC (CD34<sup>+</sup>CD38<sup>-</sup>) and Progenitors (CD34<sup>+</sup>CD38<sup>+</sup>) after 5 d treatment with either DMSO or PF-3758309, reproducible across all independent repetitions (n&gt;9). <i>scale bar = 100 μm</i>.</p> <p>(b-e) CB Lin<sup>-</sup> cells were treated after 2d of prestimulation with either DMSO or PF-3758309 (PF). Concentration as indicated. After 5d of treatment cells were counted (c, n=2), cell cycle (b, d) combined with CD34 analysis was performed (pool of n=2).</p> <p>(f-g) CB Lin<sup>-</sup> cells treated after 2 days prestimulation for 5 d with 0.25 uM PF-3758309 or DMSO equivalent were subjected to AnnexinV and CD34 staining (n=3-4). <i>Mean±SD; two-sided ratio paired t-test.</i></p> <p>(h-j) CB HSPC and Progenitors treated for 4 and 6 d with 0.25 uM PF-3758309 or DMSO equivalent were subjected to cell cycle analysis (h), counting (i: HSPC; j: Progenitors). <i>Mean±SD, two-sided paired t-test (*p&lt;0.05; **p&lt;0.01; ***p&lt;0.001).</i></p> <p>(k) Cell counts of treated HSPC on day of injection (d5 treatment, d7 culture). <i>Mean±SD.</i></p> <p>(l) Secondary LDA (8 wk) in NSG and NSG-SGM3 mice of human CD45<sup>+</sup> cells from CB Lin<sup>-</sup> cells transplanted NSG grafts (24 wk).</p>
Extended Data Fig. 5	<i>INKA1</i> -OE mediated effects <i>in vitro</i> .	Kaufmann_Extended_Data_Fig5.tif	<p>(a-c) CTRL and <i>INKA1</i>-OE HSPC were sorted (BFP<sup>+</sup>) d6 post transduction and subjected to immunostaining (n=2) and INKA1 levels were determined (a); <i>two-sided t-test.</i> Colocalization was quantified in (b) by Coloc2/ImageJ and Spearman's rank correlation values for single cells with Costes' P-value = 1.00 (<i>**p=0.0041</i>); and in (c) by proximity ligation assay (PLA) with background set according to negative controls (IgG1: anti-PAK4 + rabbit IgG; IgG2 anti-INKA1 + mouse IgG); <i>two-sided Mann-Whitney-test (b,c; ***p&lt;0.0001).</i></p>

			<p>(d) Cell viability of transduced HSPC was determined via dye exclusion (PI, 7AAD; n=9).</p> <p>(e) CTRL and <i>INKA1</i>-OE LT- and ST-HSC (n=4-5) were sorted (BFP<sup>+</sup>) d6 post transduction and subjected to cell cycle analysis by flow cytometry.</p> <p>(f) Distribution of PAK4 protein levels in nucleus vs. cytoplasm according to confocal analysis of CTRL and <i>INKA1</i>-OE LT- and ST-HSC; <i>two-sided Mann-Whitney-test</i> (**<i>p</i>=0.0005).</p> <p>(g) Differentially expressed genes with adjusted <i>p</i>-value&lt;0.01 (according to <i>DESeq2</i>, **<i>p</i>=0.003, ***<i>p</i>-value range: &lt;10<sup>-4</sup>-10<sup>-26</sup>) in either LT- or ST-HSC transduced with <i>INKA1</i>-OE vs CTRL (log<sub>2</sub>FC).</p>
Extended Data Fig. 6	<i>INKA1</i> -OE mediated effects in <i>in vitro</i> differentiation assays.	Kaufmann_Extended_Data_Fig6.tif	<p>(a) CB LT- and ST-HSC were cultured, transduced (OE) and on day 3 ptd BFP<sup>+</sup> cells were sorted into methylcellulose for colony formation, scored and harvested 12 days later and replated (right) for 12 additional days (n=4). Colony type composition of primary and secondary colony formation assays (CFU-C) is shown in the lower panels; <i>Mean±SD, two-sided paired t-test</i>.</p> <p>(b) Total single cell colony output (d15) from stromal coculture assays per CB (n=3 independent pools, <i>Mean±SD</i>) and subfraction of CTRL and <i>INKA1</i>-OE transduced HSC subsets as indicated. <i>My</i>= myeloid (CD33<sup>+</sup>), <i>E</i>= erythroid (CD45<sup>-</sup>GlyA<sup>+</sup>/CD71<sup>+</sup>), <i>Mk</i> = megakaryocytic (CD41<sup>+</sup>), <i>NK</i> = CD56<sup>+</sup> cells. Upper right panel shows proliferative potential and lower panel total colony output per assay, population and condition as assessed by flow cytometry; <i>two-sided paired t-test</i>.</p>
Extended Data Fig. 7	<i>INKA1</i> -OE, sh <i>INKA1</i> and shPAK4 in xenograft assays.	Kaufmann_Extended_Data_Fig7.tif	<p>(a) Gating strategy for BM analysis from xenografts.</p> <p>(b) Human CD45 chimerism in injected femur (right, RF) and non-injected bone marrow of CB HSPC transduced</p>

			<p>with CTRL or <i>INKA1</i>-OE at various timepoints post-transplant (5 independent CB pools). Each cell represents individual mouse.</p> <p>(c) Relative engraftment in non-injected BM of transduced cells represented as <math>\log_2</math> ratios (%BFP<sup>+</sup> of CD45<sup>+</sup> <i>in vivo</i> output /%BFP<sup>+</sup> <i>in vitro</i> input) of 5 independent experiments, <i>two-sided t-test with Welch's correction</i> (**<i>p</i>=0.006 and <i>p</i>=0.008).</p> <p>(d) Lineage composition within BFP<sup>+</sup> xenografts at 20 wk post xenotransplantation of CTRL and <i>INKA1</i>-OE HSPC (5 experiments with 4-9 mice/group each).</p> <p>(e) Flow cytometric analysis of 20 wk BM grafts gated on CD45<sup>+</sup>BFP<sup>+</sup>CD34<sup>+</sup>CD19<sup>-</sup>CD38<sup>-</sup> population; <i>Mean±SD</i>.</p> <p>(f) Individual data point representation related to Figure 5e; <i>mice per group: 10/9(4w) - 5/5/5/5 (12w)- 9/7/9/8 (20w)</i>.</p> <p>(g) Human CD45<sup>+</sup> cell counts from injected femurs from NSG mice transplanted with CTRL or <i>INKA1</i>-OE HSPC at indicated timepoints and after single dose PBS or 5-FU treatment at 4 wk; <i>Mean±SEM, two-sided Mann-Whitney-test</i>.</p> <p>(h) Fold change (FC) of CD45<sup>+</sup> counts in injected femur at 12 wk relative to 4wk of CTRL and <i>INKA1</i>-OE HSPC transplanted and PBS/5FU treated mice; <i>Mean±SD, two-sided t-test</i> (*<i>p</i>=0.043).</p> <p>(i) Injected femurs (20w) were analyzed for %CD34<sup>+</sup> (left) and total pooled bone marrow was subjected to hierarchy analysis at 20w of CTRL and <i>INKA1</i>-OE HSPC after treatment with PBS or 5-FU at 4w. Pie chart represents CD34 compartment and radius is scaled to percentage CD34. <i>Mean±SD, two-sided t-test</i>.</p> <p>(j) Four shRNAs embedded in lentiviral vectors were tested each, for <i>INKA1</i> and <i>PAK4</i> knockdown, along with</p>
--	--	--	---

			<p>shRenilla as control in MOLM13 cells. Transduced cells were sorted (BFP<sup>+</sup>) and subjected to qRT-PCR for the targeted transcript. Vector copy number (VCN) was estimated by initial transduction efficiency assuming Poisson distribution to correct for VCN introduced bias in knockdown-efficiency. Highlighted in color are the shRNAs chosen for functional assays; <i>Mean±SD, 3 technical replicates.</i></p> <p>(k) Summary of secondary LDA analysis in NSG-SGM3 mice at 8w of shRNA transduced primary grafts (20w). Mice were considered engrafted at human CD45 chimerism &gt;0.1.</p> <p>(l) Lineage composition of BFP<sup>+</sup> NSG grafts from shRNA transduced HSPC as indicated (20 wk). <i>Mean±SD.</i></p>
Extended Data Fig. 8	INKA1 and PAK4 protein interactions identify SIRT1 as INKA1 interactor.	Kaufmann_Extended_Data_Fig8.tif	<p>(a) RNA expression (GSE125345) of PAK4/INKA1 interactome as identified by BioID across human CB HSPC and Progenitors.</p> <p>(b) Pathway (GO processes) enrichment map (FDR&lt;0.05, selected) of BioID interactome via Cytoscape.</p> <p>(c) Proximity ligation assay of HSPC and Progenitors including staining controls (90 cells, 6 variants as in scheme; 3 independent CB pools). Background was set according to controls. Four PLA signal positive cells are shown enlarged.</p> <p>(d) Shown are confocal images of two separate immunostainings of extended chromatin fibers generated from CB HSPC (CD34<sup>+</sup>CD38<sup>-</sup>) that were either stained for H4K16ac and SIRT1 (left) or SIRT1 and INKA1 (right). <i>Results were reproducible in 2 independent repetitions. Scale bar =2 μm.</i></p> <p>(e) Cultures of transduced (OE) CB HSPC (CD34<sup>+</sup>CD38<sup>-</sup>; n=3) were treated with DMSO or 25 μM EX-527 starting 3 days after transduction. EdU was added on d4 and cells</p>

			were analyzed 16 hrs later by flow cytometry for %EdU <sup>+</sup> cells and DNA synthesis rate by Edu MFI within transduced/BFP <sup>+</sup> cells; <i>Mean±SD, two-sided paired t-test</i> (* <i>p</i> <0.05; ** <i>p</i> =0.004).
Extended Data Fig. 9	INKA1 and PAK4 protein interactions implicate CD112 in HSPC biology.	Kaufmann_Extended_Data_Fig9.tif	<p>(a,b) Absolute and relative <i>Nectin2</i> and <i>AFDN</i> expression from GSE125345(a) and <i>AFDN</i> expression in mPB scRNAseq dataset across diffusion pseudotime (b).</p> <p>(c) CD112 levels in CB mononuclear cells (MNC) gated on CD34<sup>+</sup>CD19<sup>+</sup> and CD34<sup>+</sup>CD19<sup>-</sup> subpopulations and CD112 vs CD19 expression in CD34<sup>+</sup> cells from 20w xenografts.</p> <p>(d,e) Flow cytometric analysis of CD34 and CD112 surface expression on CB Lin<sup>-</sup> cells and CB MNC. <i>FMO+IgG = fluorescence minus one staining without anti-CD112-APC but IgG-APC; two-sided t-test</i> (*<i>p</i>&lt;0.05, **<i>p</i>&lt;0.01, ***<i>p</i>&lt;0.001).</p> <p>(f-g) Single cell index-sorting (indexed intensities: CD49f, CD90 and CD112) of 2800 CB CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> cells onto MS5 stromal layers followed by 18 days <i>in vitro</i> differentiation and analysis via high-throughput flow cytometry (n=4). LT- and ST-HSCs were retrospectively gated according to initially indexed CD49f and CD90 levels and respective colony output (f) and colony output per individual CB pool is shown in (g).</p> <p>(h) Cultured and transduced CB CD34<sup>+</sup>CD38<sup>-</sup> cells were analyzed for CD112 levels 5-6 d post-transduction of subgated cLT-(CD34<sup>+</sup>CD45RA<sup>-</sup>CD90<sup>+</sup>), cST-(CD34<sup>+</sup>CD45RA<sup>-</sup>CD90<sup>-</sup>) HSC, cProgenitors (CD34<sup>+</sup>CD45RA<sup>+</sup>) or total cHSPC (CD34<sup>+</sup>). <i>c = cultured, two-sided paired t-test</i> (*<i>p</i>&lt;0.05, **<i>p</i>=0.004).</p> <p>(i) Confocal images of BFP<sup>+</sup> sorted CB LT- and ST-HSC 6d after transduction followed by immunostaining (n=3; Ø45 cells/ sample); <i>two-sided Mann-Whitney-test</i>.</p>



			(j) Confocal images of immunostaining of CB LT- and ST-HSC (total 74 cells from 3 CB) and thereof derived Spearman's rank correlation values (calculated by Coloc2/ImageJ) for the indicated parameter pairs, only non-nuclear cell area was considered.
Extended Data Fig. 10	CD112 separates primed from non-primed LT-HSC.	Kaufmann_Extended_Data_Fig10.tif	<p>(a,b) Post-sort analysis of LT- and ST-HSC sorted for CD112<sup>lo</sup> and CD112<sup>hi</sup> either directly after sorting on the same instrument (a) or in parallel with CellRox staining after 30 min cytokine free culture(b; see g,h) on a FACS analyzer.</p> <p>(c) Representative plots of flow cytometric analysis of LT- and ST-HSC sorted for CD112<sup>lo</sup> and CD112<sup>high</sup> directly or after 42 hrs culture for CDK6 and Ki-67.</p> <p>(d) Quantification of flow cytometric analysis of LT- and ST-HSC sorted for CD112<sup>lo</sup> and CD112<sup>hi</sup> directly or after 42 hrs culture for CDK6 and Ki-67 (n=3; <i>paired t-test</i>).</p> <p>(e-g) LT- and ST-HSC sorted for CD112<sup>lo</sup> and CD112<sup>hi</sup> (n=4-7) were analyzed directly via immunostaining for INKA1 (e; total of 1587 single cells from 7 independent CB pools, <i>whiskers indicate 10-90 percentile, Kruskal-Wallis test</i>) or for CellROX labeling after 30 min cytokine-free culture and assessed by flow cytometry (f) and fluorescence microscopy (g); <i>box-and-whisker plots with median as centre and either 10-90% percentile (e) or minimal to maximal values are represented by whiskers (f). Two-sided t-test (f,g; *p&lt;0.05, **p&lt;0.01, ***p&lt;0.001)</i></p> <p>(h,i) LT-HSC sorted for CD112<sup>lo</sup> (n=2) and CD112<sup>hi</sup> (n=3) were subjected to ATAC-seq. The number of 500bp windows overlapping called peaks exclusively in two CD112<sup>hi</sup> or CD112<sup>lo</sup> replicates but not in the other is shown in (h) and (i) is a heatmap of -log P-Value of enrichment for known motifs enriched in either CD112<sup>hi</sup></p>

			<p>or CD112<sup>lo</sup>, relative to a background of all called peaks. CD112<sup>hi</sup> versus CD112<sup>lo</sup> LT-HSC and INKA1-OE vs CTRL LT-HSC.</p> <p>(k) GSEA results for CD112<sup>hi</sup> versus CD112<sup>lo</sup> RNAseq data (3472 GO biological processes; gene set size 25-500; FDR&lt;0.1). <i>nES</i> = <i>normalized enrichment score</i>; *FDR&lt;0.05; **FDR&lt;0.01.</p> <p>(l) Expression of leading-edge genes from mTORC1 and GOBP gene sets (2-3 independent CB pools).</p> <p>(m) Lineage composition of transplanted CD112<sup>lo</sup> and CD112<sup>hi</sup> LT-HSC. Each column represents an individual mouse.</p> <p>(n) Summary of secondary LDA from 12 wk primary grafts into NSG-SGM3 for additional 8 wk for 3 independent CB pools. Same cell numbers per dose were transplanted from initially transplanted CD112<sup>lo</sup> and CD112<sup>hi</sup> LT-HSC per CB but varied for individual CB cohorts.</p>
--	--	--	--

1  
2

Item	Present?	Filename	A brief, numerical description of file contents. i.e.: <i>Supplementary Figures 1-4, Supplementary Discussion, and Supplementary Tables 1-4.</i>
Supplementary Information	No		
Reporting Summary	Yes	NI-A30897C_reporting_summary.pdf	

3  
4

Type	Number	Filename	Legend or Descriptive Caption
------	--------	----------	-------------------------------

	If there are multiple files of the same type this should be the numerical indicator. i.e. "1" for Video 1, "2" for Video 2, etc.	This should be the name the file is saved as when it is uploaded to our system, and should include the file extension. i.e.: <i>Smith_Supplementary_Video_1.mov</i>	Describe the contents of the file
Supplementary Table	1	Kaufmann_Supplementary_Tables.xlsx	Supplementary Table 1 - High Confidence interactors of INKA1 and PAK4 according to proximity-dependent Biotin Identification (BioID) Assay  Supplementary Table 2 - Differential expressed genes (289) between CD112 <sup>lo</sup> and CD112 <sup>hi</sup> LT-HSC (statistics calculated by DESeq2).  Supplementary Table 3 - Detailed information of commercially available reagents and other materials/resources used.

5  
6

Parent Figure or Table	Filename This should be the name the file is saved as when it is uploaded to our system, and should include the file extension. i.e.: <i>Smith_SourceData_Fig1.xls</i> , or <i>Smith_Unmodified_Gels_Fig1.pdf</i>	Data description i.e.: Unprocessed Western Blots and/or gels, Statistical Source Data, etc.
Source Data Fig. 1	Kaufmann_SourceData_Fig1.xlsx	Statistical Source Data
Source Data Fig. 2	Kaufmann_SourceData_Fig2.xlsx	Statistical Source Data
Source Data Fig. 3	Kaufmann_SourceData_Fig3.xlsx	Statistical Source Data
Source Data Fig. 4	Kaufmann_SourceData_Fig4.xlsx	Statistical Source Data
Source Data Fig. 5	Kaufmann_SourceData_Fig5.xlsx	Statistical Source Data

Source Data Fig. 6	Kaufmann_SourceData_Fig6.xlsx	Statistical Source Data
Source Data Fig. 7	Kaufmann_SourceData_Fig7.xlsx	Statistical Source Data
Source Data Fig. 8	Kaufmann_SourceData_Fig8.xlsx	Statistical Source Data
Source Data Extended Data Fig. 1	Kaufmann_SourceData_Extended_Data_Fig1.xlsx	Statistical Source Data
Source Data Extended Data Fig. 2	Kaufmann_SourceData_Extended_Data_Fig2.xlsx	Statistical Source Data
Source Data Extended Data Fig. 3	Kaufmann_SourceData_Extended_Data_Fig3.xlsx	Statistical Source Data
Source Data Extended Data Fig. 4	Kaufmann_SourceData_Extended_Data_Fig4.xlsx	Statistical Source Data
Source Data Extended Data Fig. 5	Kaufmann_SourceData_Extended_Data_Fig5.xlsx	Statistical Source Data
Source Data Extended Data Fig. 6	Kaufmann_SourceData_Extended_Data_Fig6.xlsx	Statistical Source Data
Source Data Extended Data Fig. 7	Kaufmann_SourceData_Extended_Data_Fig7.xlsx	Statistical Source Data
Source Data Extended Data Fig. 8	Kaufmann_SourceData_Extended_Data_Fig8.xlsx	Statistical Source Data
Source Data Extended Data Fig. 9	Kaufmann_SourceData_Extended_Data_Fig9.xlsx	Statistical Source Data
Source Data Extended Data Fig. 10	Kaufmann_SourceData_Extended_Data_Fig10.xlsx	Statistical Source Data

7

8  
9

10 **A latent subset of human hematopoietic stem cells resists regenerative stress**  
11 **to preserve stemness**

12

13 Kerstin B. Kaufmann<sup>1</sup>, Andy G. X. Zeng<sup>1,2</sup>, Etienne Coyaud<sup>1</sup>, Laura Garcia-Prat<sup>1</sup>, Efthymia Papalexi<sup>3,4</sup>,  
14 Alex Murison<sup>1</sup>, Estelle M. N. Laurent<sup>1</sup>, Michelle Chan-Seng-Yue<sup>1,5</sup>, Olga I. Gan<sup>1</sup>, Kristele Pan<sup>1</sup>, Jessica  
15 McLeod<sup>1</sup>, H el ena Boutzen<sup>1</sup>, Sasan Zandi<sup>1,6</sup>, Shin-ichiro Takayanagi<sup>1,7</sup>, Rahul Satija<sup>3,4</sup>, Brian Raught<sup>1,8</sup>,  
16 Stephanie Z. Xie<sup>1</sup> and John E. Dick<sup>1,2\*</sup>

17

18 <sup>1</sup> Princess Margaret Cancer Centre; University Health Network; Toronto, ON, M5G 0A3; Canada

19 <sup>2</sup> Department of Molecular Genetics; University of Toronto; Toronto, ON, M5S 1A8; Canada

20 <sup>3</sup> New York Genome Center; New York, NY, 10013; USA

21 <sup>4</sup> Center for Genomics and Systems Biology; New York University; New York, NY, 10003; USA

22 <sup>5</sup> PanCuRx Translational Research Initiative; Ontario Institute for Cancer Research; Toronto, ON, M5G  
23 0A3; Canada

24 <sup>6</sup> Department of Laboratory Medicine and Pathobiology; University of Toronto; Toronto, ON, M5S 1A8;  
25 Canada

26 <sup>7</sup> Cell Therapy Project; R&D Division; Kirin Holdings Company, Limited; Tokyo, 194-8533; Japan

27 <sup>8</sup> Department of Medical Biophysics; University of Toronto; Toronto, ON, M5S 1A4; Canada

28

29

30 \*Corresponding author: john.dick@uhnresearch.ca

31 **ABSTRACT**

32 Continuous supply of immune cells throughout life relies on the delicate balance of the hematopoietic  
33 stem cell (HSC) pool between long-term maintenance and meeting the demands of both normal blood  
34 production and unexpected stress conditions. Here we identified distinct subsets of human long-term (LT-  
35 )HSCs that responded differently to regeneration-mediated stress: an immune checkpoint ligand CD112<sup>lo</sup>  
36 subset that exhibited a transient engraftment restraint (termed latency) before contributing to  
37 hematopoietic reconstitution and a primed CD112<sup>hi</sup> subset that responded rapidly. This functional  
38 heterogeneity and CD112 are regulated by INKA1, through direct interaction with PAK4 and SIRT1,  
39 inducing epigenetic changes and defining an alternative state of LT-HSC quiescence that serves to  
40 preserve self-renewal and regenerative capacity upon regeneration-mediated stress. Collectively, our data  
41 uncovered the molecular intricacies underlying HSC heterogeneity and self-renewal regulation and point  
42 to latency as an orchestrated physiological response that balances blood cell demands with preserving a  
43 stem cell reservoir.

44

45

46

47

## 48 INTRODUCTION

49 Continuous and life-long blood cell production hinges on the ability of rare multipotent LT-HSCs  
50 to balance two fundamental stem cell properties: self-renewal versus capacity to generate differentiated  
51 progeny. This balance ensures not only continuous hematopoietic output but also long-term maintenance  
52 of the LT-HSC pool. Maintenance occurs because HSC reside predominantly in a quiescent state that can  
53 last for weeks to years before intrinsic and extrinsic cues activate some LT-HSC to either self-renew or  
54 generate short-term (ST-)HSC, while others remain restrained in quiescence<sup>1-3</sup>. However, serial  
55 challenges, such as transplantation or inflammation, can lead to HSC attrition, exhaustion and biases in  
56 lineage output manifesting as cytopenias, late graft failure following HSC transplantation and  
57 compromised immune function upon aging<sup>4-6</sup>. Understanding *in vivo* HSC dynamics and the mechanisms  
58 that ensure HSC longevity on a molecular, cellular and population level is crucial for enabling cell  
59 therapies that permit both rapid regeneration and sustained long-term hematopoiesis.

60  
61 Although murine studies provided considerable insights into cellular and molecular mechanisms  
62 of HSC quiescence, poor understanding remains how quiescence control integrates with the fate decisions  
63 HSCs must make to either remain quiescent, differentiate and/or self-renew<sup>7,8</sup>. In humans, the longer  
64 lifespan and massively higher cellular output make dynamic integration of these fundamental stem cell  
65 properties even more acute. We particularly lack understanding of the processes underlying the activation  
66 and/or resistance of human HSC to adapt to varying hematopoietic demands during homeostasis or  
67 emergency hematopoiesis triggered by inflammation, acute blood loss or transplantation. Although  
68 human immunophenotypic CD90<sup>+</sup>CD49f<sup>-</sup> LT- and CD90<sup>-</sup>CD49f<sup>-</sup> ST-HSCs are both multipotent and  
69 possess engraftment capacity, ST-HSCs have markedly reduced self-renewal potential and faster  
70 quiescence exit kinetics<sup>3,9</sup>. Despite these functional differences between human LT- and ST-HSC subsets,  
71 published studies have not uncovered clear transcriptomic differences at the population level or by single  
72 cell (sc)RNA-seq and instead pointed to the continuous nature of early hematopoietic commitment<sup>10,11</sup>.  
73 Functional screens offer an alternative approach to identify mechanisms that distinguish LT-HSC from



74 ST-HSC. As normal HSC and leukemia stem cells (LSC) in acute myeloid leukemia (AML) share many  
75 similarities in stemness properties, we recently undertook a functional stemness screen of potential LSC  
76 regulators. We revealed *INKA1* (also known as *FAM212A*, *C3orf54*) as a regulator of LSC self-renewal<sup>12</sup>.  
77 This study linked human LSC quiescence and latency to INKA1-mediated inhibition of p21(Cdc42/Rac)-  
78 activated kinase 4 (PAK4) and to reduction of H4K16 acetylation (H4K16ac). However, this signaling  
79 axis was not studied in normal hematopoiesis. PAK4 plays a role in numerous cellular functions<sup>13</sup> and  
80 only a single study in human cell lines and a limited number of studies in frogs, fish and mice established  
81 that *Inka1* is an endogenous inhibitor of *Pak4* and essential for neural crest development<sup>14,15</sup>. H4K16ac is  
82 associated with activated transcription of single genes, but this chromatin mark is not strongly correlated  
83 with changes in global gene expression as assessed in HeLa cells and murine HSC<sup>16,17</sup>. Here, we  
84 uncovered a mechanistic link between PAK4 and H4K16ac, and its role in human hematopoiesis,  
85 demonstrating that a molecular network with INKA1 at its center governs the long-term regenerative  
86 capacity of human HSC. Moreover, we show that human LT-HSC are not homogeneous but composed of  
87 two cellular entities distinguished on the basis of CD112 expression, each with distinct alternative states  
88 of quiescence and repopulation kinetics.

89

## 90 **RESULTS**

### 91 **INKA1 and H4K16ac delineated HSC heterogeneity**

92 To explore the relevance of INKA1 and PAK4 in normal human hematopoiesis, we first  
93 interrogated published RNA expression datasets<sup>18,19</sup>. *INKA1* was highly expressed in HSC and progenitor  
94 enriched CD34<sup>+</sup> cells compared to mature CD34<sup>-</sup> populations, whereas *PAK4* showed low expression  
95 across the hematopoietic hierarchy (Fig. 1a). Analysis of protein expression and chromatin acetylation by  
96 confocal analysis of four CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> HSC subsets sub-fractionated by CD90/CD49f  
97 expression from cord blood (CB) showed heterogeneity in INKA1 and H4K16ac abundance at single-cell  
98 resolution. The majority of cells (85.5±8.2%) in these four populations were high in H4K16ac with low to  
99 background signals for nuclear INKA1 (INKA1<sup>lo</sup>; Fig. 1b-d and Extended Data Fig. 1a-c). Cells with

100 concomitant faint staining for both markers further revealed intranuclear spatial separation when analyzed  
101 after conversion into binary images (Extended Data Fig. 1c) suggesting these cells were in transition to  
102 either becoming INKA1<sup>hi</sup> or H4K16ac<sup>hi</sup> and that these molecules actively excluded one another. High  
103 nuclear signal for INKA1 (INKA1<sup>hi</sup>) was detected in a minor population (10-20%) with small nuclear  
104 diameter and low or negative staining for H4K16ac (Fig. 1e and Extended Data Fig. 1d). To determine if  
105 INKA1<sup>hi</sup> cells were in an alternative state of stem cell quiescence, we used CDK6 as a definitive marker  
106 of the activated/poised stem cell state<sup>3</sup> and compared INKA1 versus CDK6 content in LT- and ST-HSCs.  
107 Immunofluorescence analysis of LT-HSCs showed that INKA1<sup>hi</sup> cells had significantly lower CDK6  
108 protein abundance than INKA1<sup>lo</sup> cells; differences not seen in ST-HSCs (Fig. 1f and Extended Data Fig.  
109 1e). Hence, the inverse nuclear staining patterns of INKA1 and H4K16ac delineated at least two distinct  
110 populations within LT-HSCs that might be linked to alternative states of stem cell quiescence.

111

#### 112 ***INKA1* and *CDK6/PAK4* segregated quiescent single cells**

113 To gain more insight into the heterogeneous subpopulations within the LT-HSC pool and  
114 determine whether they varied in stem cell activation signatures, we undertook single-cell transcriptomic  
115 analysis of HSCs from granulocyte-colony-stimulating factor (G-CSF)-mobilized peripheral blood  
116 (mPB). Since mPB cells are activated to exit the bone marrow *in vivo*, we hypothesized that they may  
117 result in improved transcriptional resolution of HSC activation dynamics. We confirmed that 95.7±2.9%  
118 of CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> HSCs from mPB were in G<sub>0</sub> (Ki67<sup>-</sup>, Extended Data Fig. 2a) and that  
119 phenotypically defined LT- and ST-HSCs recapitulated the differences in CDK6 protein abundance as  
120 previously described for CB (Fig. 2a)<sup>3</sup>. Flow cytometric detection of CDK6 further partitioned Ki67<sup>-</sup> G<sub>0</sub>  
121 cells into canonically quiescent (CDK6<sup>-</sup>) and primed (CDK6<sup>+</sup>) fractions (47.8±14.2%; Fig. 2b).

122

123 Single-cell index-sorting (signals indexed: CD34, CD38, CD45RA, CD90, CD49f) followed by  
124 full-length transcript scRNA-seq of CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> mPB cells (from 3 donors) generated 1191  
125 single-cell transcriptomes. To identify a trajectory of HSC activation, we applied diffusion pseudotime

126 analysis while leveraging predictions of future cell states through RNA Velocity<sup>20,21</sup>(Fig. 2c and Extended  
127 Data Fig. 2b-f). We observed a gradual increase in CDK6 expression as well as cell cycle programs<sup>22</sup>  
128 from early to late pseudotime, confirming that we captured the transition from canonically quiescent to  
129 activated/poised HSC. Superimposing *INKA1* and *PAK4* expression onto the diffusion map revealed that  
130 the highest *INKA1* expression occurred early in pseudotime and thereafter declined, while *PAK4*  
131 expression increased late in pseudotime (Fig. 2d and Extended Data Fig. 2b). Despite early enrichment of  
132 *ITGA6* transcripts (encoding CD49f), surface expression of CD49f and CD90 was homogenous across  
133 pseudotime, implying that all 4 HSC subsets followed the same trajectory of activation (Extended Data  
134 Fig. 2d-f). Collectively, these data established that LT-HSCs with high *INKA1* are *CDK6*<sup>lo</sup> and fall on the  
135 non-primed end of the quiescence spectrum.

136

#### 137 **H4K16ac and PAK4 demarcated HSC cell cycle activation**

138 To examine whether H4K16ac and PAK4 played a role in HSC cell cycle priming, we  
139 interrogated cell cycle (Ki67) and H4K16ac status after cytokine stimulation by flow cytometry in  
140 lineage-depleted CB cells (CB Lin<sup>-</sup>) and by immunostaining in LT- and ST-HSCs (also including PAK4).  
141 H4K16ac was significantly lower in CB Lin<sup>-</sup> cells remaining in G<sub>0</sub> following cytokine exposure as  
142 compared to those that progressed into cell cycle. The signal peaked in late S-phase and subsequently  
143 declined in G<sub>2</sub>/M phase (Fig. 2e,f and Extended Data Fig. 3a). In agreement, higher abundance of  
144 H4K16ac and PAK4 were observed in Ki67<sup>hi</sup> vs. Ki67<sup>lo</sup>LT- and ST-HSCs (Fig. 2g and Extended Data  
145 Fig. 3b,c) suggesting upregulation upon cell cycle entry. To further resolve the longitudinal dynamics of  
146 H4K16ac and PAK4 expression during quiescence exit and entry into the first cell cycle of the most  
147 primitive HSC subsets, we performed multi-timepoint (6-100 h) single-cell protein imaging analysis  
148 under high-cytokine culture conditions. We used CDK6 as a G<sub>0</sub> exit marker<sup>3</sup> and CD34<sup>+</sup>CD38<sup>+</sup>  
149 progenitors were used as an activated control. We observed a significant increase in the signal intensities  
150 for CDK6, PAK4 and H4K16ac at late timepoints (66 and 100 h) when compared to 6 hrs for both LT-  
151 and ST-HSCs. However, LT-HSCs showed an overall lower CDK6 abundance than ST-HSC at 66 hrs in

152 accordance with the expected delay in upregulation of CDK6<sup>3</sup> (Fig. 2h and Extended Data Fig. 3d). We  
153 captured a similar delay of PAK4 and H4K16ac upregulation in LT-HSCs as compared to ST-HSC.

154

155 To account for the asynchronous cell cycle priming and progression in the imaging study, we  
156 applied diffusion pseudotime analysis<sup>21</sup> to our multidimensional single-cell protein imaging data (6  
157 parameters: nuclear intensity for H4K16ac, CDK6, PAK4, DAPI; total cell intensity for CDK6, PAK4;  
158 n= 3454 cells). This method ordered single cells according to their phenotype independently of the actual  
159 timepoint harvested. Accordingly, at all timepoints of analysis CD34<sup>+</sup>CD38<sup>+</sup> progenitors as well as HSCs  
160 from all four CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> subsets intermixed along the generated pseudotime trajectory that  
161 showed gradual upregulation of CDK6, H4K16ac and PAK4 along with DAPI/DNA intensity (Fig. 2i and  
162 Extended Data Fig. 3e,f). LT-HSCs were enriched earlier in pseudotime, in accordance with their known  
163 delayed cell cycle entry, whereas activated progenitor cells were skewed towards the late end of the  
164 trajectory (Extended Data Fig. 3f)<sup>3</sup>. The upregulation of PAK4 protein abundance and H4K16 acetylation  
165 preceded the upregulation of CDK6 in this single cell immunofluorescence-based pseudotime analysis  
166 (Fig. 2i). Overall, these data suggested that PAK4 and H4K16ac might indicate early cell-cycle priming  
167 and progression in HSCs.

168

### 169 **PAK4-inhibition induced G<sub>0</sub> maintains regenerative potential**

170 As INKA1 is an endogenous PAK4 inhibitor, we used pharmacological inhibition to obtain  
171 independent evidence that cell cycle progression depended on PAK4 activity. Such drug studies are  
172 difficult to undertake on highly purified HSC subpopulations so we used combined hematopoietic stem  
173 and progenitor cells (HSPC) subsets. Treatment of cytokine-stimulated CB CD34<sup>+</sup>CD38<sup>-</sup> HSPC and  
174 CD34<sup>+</sup>CD38<sup>+</sup> progenitors for 5 days with the PAK4 inhibitor PF-3758309, known to compete with  
175 INKA1<sup>14</sup>, resulted in a dose-dependent reduction of cell expansion along with an accumulation of G<sub>0</sub> cells  
176 (Ki67<sup>-</sup>; Fig. 3a,b and Extended Data Fig. 4a-d). Cell viability was only mildly reduced and CD34 surface  
177 expression was retained (Extended Data Fig. 4e-g). PF-3758309 treatment without prior cytokine-

178 stimulation had similar effects on cell counts and cell cycle (Extended Data Fig. 4h-j). *Ex vivo* treatment  
179 with PF-3758309 followed by limiting-dilution xenograft assays (LDA) in primary engrafted NSG mice  
180 did not demonstrate any change in HSC frequency or lineage output compared to DMSO controls (n=6;  
181 Fig. 3c,d and Extended Data Fig. 4k). Subsequently, HSC frequency in 4 of the 6 primary xenograft  
182 cohorts was assessed by serial LDA using human cytokine (SCF, GM-CSF, IL3) expressing NSG-SGM3  
183 mice as a more sensitive secondary recipient as demonstrated by direct comparison with NSG mice  
184 (Extended Data Fig. 4l). We found a trend towards a 2.0-fold increase in HSC frequency following initial  
185 PAK4 inhibitor treatment ( $p=0.071$ , Fig. 3e,f). These results suggested that PAK4 inhibition transiently  
186 stalled cytokine-stimulated HSPC in G<sub>0</sub> thereby maintaining their stem cell regenerative capacity.

### 187 188 ***INKA1*-OE conferred selective resistance to activation**

189 To gain more mechanistic insight into how *INKA1* functions as an endogenous PAK4 inhibitor  
190 within HSC subsets and influences their priming status, we utilized an *INKA1* overexpressing (*INKA1*-  
191 OE) lentiviral vector that co-expressed blue fluorescent protein (BFP) to track transduced cells. By  
192 focussing on BFP<sup>+</sup> *INKA1*-OE and control transduced CB CD34<sup>+</sup>CD38<sup>-</sup> cells, we established that *INKA1*  
193 protein was elevated and that PAK4 and *INKA1* were colocalized (Extended Data Fig. 5a-d). *INKA1*-OE  
194 in LT-HSCs, but not in ST-HSCs, reduced CDK6 intensity, H4K16ac and the relative abundance of  
195 PAK4 in the nucleus and increased the percentage of cells retained in G<sub>0</sub> (2.6±1.5 fold; n=5; Fig. 4a,b and  
196 Extended Data Fig. 5e,f) compared to controls. Gene set enrichment analysis (GSEA) of RNA-seq data  
197 further showed that *INKA1*-OE vs. controls preserved established HSC stemness signatures and induced a  
198 transcriptionally suppressed cell state in LT-HSCs, as evidenced by negative enrichment of multiple stem  
199 cell associated and inflammatory signaling pathways, stress responses, and cell cycle related gene sets  
200 (Fig. 4c,d and Extended Data Fig. 5g). To assess whether *INKA1*-OE interfered with HSC function we  
201 performed both single cell stroma-based differentiation assays and colony-forming assays (CFU-C).  
202 *INKA1*-OE did not change lineage outcome, however, *INKA1*-OE in LT-HSCs resulted in decreased  
203 primary CFU-C output without differences in secondary replating assays (Extended Data Fig. 6a,b).

204 Collectively, these data demonstrated that *INKA1*-OE functioned selectively in LT-HSCs by suppressing  
205 stem cell activation programs and inducing a transient restraint in colony formation without affecting their  
206 potential to generate multilineage hematopoietic progeny.

207

### 208 ***INKA1*-OE and *PAK4*-KD transiently restrained repopulation**

209 To directly test the functional impact that *INKA1*-OE might have on the *in vivo* properties of  
210 HSCs we utilized the xenograft assay. We previously observed that *INKA1*-OE in CB CD34<sup>+</sup>CD38<sup>-</sup>  
211 HSPCs induced delayed repopulation, indicative of a transient proliferative disadvantage of the  
212 transduced (BFP<sup>+</sup>) fraction at the 4 wk timepoint that is overcome by 20 wk<sup>12</sup>. To explore the effects of  
213 *INKA1* on stemness maintenance and activation, we carried out a more detailed analysis of CB HSPCs  
214 transduced with *INKA1*-OE or control vectors using serial xenotransplantation assays. We observed no  
215 differences in relative engraftment in the injected femurs at 2 wk (Fig. 5a and Extended Data Fig. 7a-c),  
216 whereas *INKA1*-OE resulted in a relative reduction of 4-fold ( $\log_2\text{BFPratio}=2$ ) of gene-marked (BFP<sup>+</sup>)  
217 cells at 4 wk, a timepoint when the majority of LT-HSCs are normally activated and cycling<sup>3</sup>.  
218 Examination of primary xenografts at 20 wk showed that the relative engraftment was comparable in both  
219 groups with no overt lineage skewing (Extended Data Fig. 7d). Secondary transplantation into NSG-  
220 SGM3 recipients was used to model successive rounds of stem cell activation and forced stress  
221 hematopoiesis due to the increased human pro-inflammatory cytokine exposure in NSG-SGM3 mice.  
222 When 2 wk primary samples were transplanted, only *INKA1*-OE cells engrafted (Fig. 5b). Similarly,  
223 *INKA1*-OE cells isolated from 4 wk primary xenografts generated grafts in all secondary recipients,  
224 indicative of self-renewal, while only half of control recipients were engrafted (Fig. 5b). Analysis of 20  
225 wk primary xenografts revealed a dissociation between HSC function and immunophenotype, the *INKA1*-  
226 OE group had a 3.6-fold reduction of phenotypic LT-HSCs but a 4.5-fold increase in the frequency of  
227 functional HSCs as measured by serial LDA, as compared to controls (Fig. 5b-e and Extended Data Fig.  
228 7e). Collectively, these data indicated that the transient restraint induced by *INKA1*-OE at early timepoints

229 allowed them to retain their self-renewal and regenerative capacity at later timepoints despite their  
230 exposure to regeneration-mediated stress.

231  
232 As *INKA1*-OE cells exhibited transient restraint during regeneration, we investigated whether the  
233 restrained HSCs were cycling more slowly *in vivo*. Xenotransplanted NSG mice were treated after 4  
234 weeks of repopulation with 5-Fluorouracil (5-FU), a treatment known to deplete proliferating cells while  
235 inducing quiescent HSCs to become activated<sup>23</sup>. A single dose of 5-FU resulted in higher total CD45<sup>+</sup> and  
236 BFP<sup>+</sup> cells counts in the *INKA1*-OE group at the 12-wk timepoint and a slight decline at 20 wk compared  
237 to 5-FU treated controls (Fig. 5e and Extended Data Fig. 7g,h). There was an expansion of the phenotypic  
238 HSC compartment of 5-FU treated, *INKA1*-OE mice, but depletion in control groups (Fig. 5f and  
239 Extended Data Fig. 7i). Serial LDA analysis showed higher HSC frequency within 5-FU versus PBS-  
240 treated *INKA1*-OE grafts, while repopulation in 5-FU treated control groups was abolished (Fig. 5g,h).  
241 These data established that the transient restraint enforced by INKA1 was linked to reduced cycling *in*  
242 *vivo*. In addition to *INKA1*-OE protection from 5-FU-mediated stem cell depletion, exposure of *INKA1*-  
243 OE cells to a second challenge (5-FU) during the period of restrained regeneration resulted in activation  
244 of HSCs and stem cell expansion without exhaustion of self-renewal potential.

245  
246 To evaluate whether HSPC engraftment capacity depended on INKA1 and was mediated via  
247 PAK4 inhibition, we transduced CB CD34<sup>+</sup>CD38<sup>-</sup> HSPCs either with control, two *INKA1*-knockdown  
248 (KD) and two *PAK4*-KD vectors and performed *in vivo* assays in analogy to *INKA1*-OE studies. *INKA1*-  
249 KD resulted in 2-fold reduction in the relative proportion of BFP<sup>+</sup> cells at 4 wk and 20 wk in primary  
250 recipients (Fig. 5i and Extended Data Fig. 7j). Upon secondary transplantation, the residual cells  
251 harvested at 20 wk and injected at the same cell doses performed as poorly as controls (Fig. 5j). By  
252 contrast, both of the *PAK4*-KD groups recapitulated the *INKA1*-OE phenotype, showing delayed  
253 repopulation in primary xenograft assays and increased HSC frequency upon secondary transplantation,  
254 with the shPAK4-3 cohort showing a 7.6-fold increase (Fig. 5i,j and Extended Data Fig. 7k). *PAK4*-KD

255 resulted in a higher abundance of lymphoid cells at the expense of myeloid and erythroid cells at 20 wk  
256 (Extended Data Fig. 7l). This result contrasts with the balanced output generated upon *INKA1*-OE, further  
257 implicating an additional INKA1 function beyond simply inhibiting PAK4. Collectively, *INKA1*-OE or  
258 *PAK4*-KD induced a reversible restraint of hematopoietic output that was coincident with enhancement of  
259 their long-term self-renewal and regenerative potential.

260

### 261 **Interactome linked INKA1 via SIRT1 to H4K16 deacetylation**

262 To gain further mechanistic insight into how INKA1 and PAK4 function and how they are linked  
263 to H4K16 acetylation, we studied their broader interactome via proximity-dependent biotin labelling  
264 (BioID)<sup>24</sup>. We generated inducible N- and C-terminal BirA\* (BirA R118G)-fusions for both proteins. 87  
265 high confidence interactors were identified that were either biotinylated by INKA1 fusions only (34),  
266 PAK4 fusions only (30), or both (23)(Fig. 6a, Supplementary Table 1). The transcripts of the majority of  
267 these retrieved proteins were also found highly expressed in CB LT-/ST-HSCs (GSE125345; Fig. 6a and  
268 Extended Data Fig. 8a).

269

270 Pathway enrichment analysis of the interactome confirmed enrichment for GTPase signaling, cell  
271 polarity, protein localization, signal transduction, cell cycle and regulation of lysine acetylation (Extended  
272 Data Fig. 8b). One of the INKA1 interactors was the NAD<sup>+</sup>-dependent deacetylase SIRT1, a known  
273 regulator of many biological processes that functions through targeting H4K16ac<sup>25</sup> and other histone and  
274 non-histone proteins for deacetylation. We confirmed nuclear and cytoplasmic INKA1 interaction with  
275 SIRT1 in CB CD34<sup>+</sup>CD38<sup>-</sup> cells (Fig. 6b and Extended Data Fig. 8c). Staining of chromatin fibers from  
276 CD34<sup>+</sup>CD38<sup>-</sup> cells further confirmed chromatin association of SIRT1 and INKA1 that was inverse to  
277 H4K16ac (Extended Data Fig. 8d). *INKA1*-OE in CD34<sup>+</sup>CD38<sup>-</sup> cells showed increased relative  
278 redistribution of SIRT1 into the nucleus, while pharmacological SIRT1 inhibition with EX-527  
279 accelerated DNA synthesis and partly rescued reduced EdU incorporation upon *INKA1*-OE (Fig. 6c and  
280 Extended Data Fig. 8e). Thus, the interaction of INKA1 with SIRT1 provided an explanation both for the



281 inverse staining patterns of INKA1 and H4K16ac in HSC subsets and the suppression of H4K16ac upon  
282 *INKA1*-OE. Broadly, the regulation of SIRT1 localization and activity by INKA1 reveals a novel function  
283 beyond PAK4 inhibition.

284

### 285 **INKA1/PAK4 interactome identified CD112 as an HSPC marker**

286 Among other high confidence INKA1/PAK4 interactors we found an adhesion and immune  
287 regulatory molecule, CD112 (PVRL2, Nectin-2; Fig. 6a) and a protein connecting Nectins to the actin  
288 cytoskeleton, MLLT4 (Afadin). *Nectin2* (transcript for CD112) and *Afadin* showed a very distinct  
289 HSC/HSPC-specific transcript expression pattern (Extended Data Fig. 9a,b). CD112 and CD34 exhibited  
290 similar surface expression on CB mononuclear cells (MNC) although CD112 did not capture lymphoid-  
291 primed CD19<sup>+</sup>CD34<sup>+</sup> cells (Fig. 6d,e and Extended Data Fig. 9c,d). Across the CD34<sup>+</sup> hierarchy, CD112  
292 and CD34 expression was consistently higher in the more primitive fractions, with the exception of  
293 lymphoid-primed progenitors (MLP), which was CD112<sup>lo</sup>CD34<sup>hi</sup> (Extended Data Fig. 9e). In stromal  
294 differentiation assays we observed no lineage restriction for single cells (CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>) with  
295 either low or high CD112 expression according to initial index-sorting (Extended Data Fig. 9f,g).

296

297 To validate the BioID results, we tested the impact of *INKA1*-OE on surface expression of  
298 CD112. *INKA1*-OE reduced CD112 surface expression on culture-derived LT-HSCs (cLT-HSCs)<sup>19</sup> and  
299 cST-HSCs, although total protein abundance was not altered (Fig. 6f and Extended Data Fig. 9h,i). We  
300 captured individual cST-HSCs with extremely high intracellular CD112 that were almost exclusive to  
301 *INKA1*-OE suggesting induction of receptor internalization (Extended Data Fig. 9i). Using the late  
302 endosome/lysosome marker LAMP1, we confirmed cytoplasmic colocalization with CD112 and INKA1  
303 in LT- and ST-HSCs (Extended Data Fig. 9j). Extracellular CD112 immunostaining indicated that cells  
304 with high expression of CD112 on the surface are INKA1<sup>lo</sup> and vice versa (Fig. 6g). Finally, our scRNA-  
305 seq pseudotime analysis provided independent support for this inverse relationship; *INKA1* expression  
306 was enriched early and CD112/*Nectin2* was enriched late in pseudotime (Fig. 6h). Together, these data

307 support a direct functional connection between INKA1 and CD112 and predict that they may mark  
308 alternative states of stem cell quiescence, with primed LT-HSCs being *INKA1*<sup>lo</sup> and CD112/*Nectin2*<sup>hi</sup>.

309

### 310 **CD112<sup>hi</sup>/CD112<sup>lo</sup> marked alternative states of LT-HSCs**

311 To determine if the relative abundance of CD112 can discriminate primed (high in CDK6 and  
312 H4K16ac) from non-primed (*INKA1*<sup>hi</sup>) cells, we sorted CD112<sup>hi</sup> and CD112<sup>lo</sup> cell fractions from purified  
313 LT-HSC and ST-HSC populations, that are >90% in G<sub>0</sub><sup>3</sup> (Fig. 7a and Extended Data Fig. 10a,b). Cell  
314 cycle analysis revealed that the CD112<sup>hi</sup>LT-HSC fraction had less non-primed (%CDK6<sup>lo</sup>Ki67<sup>-</sup>) cells and  
315 less cells in G<sub>0</sub> vs. CD112<sup>lo</sup> LT-HSCs (91.1±5.3% vs. 95.9±2.2% Ki67<sup>-</sup>; Fig. 7b and Extended Data Fig.  
316 10c). These differences were still evident upon short-term culture (42 hrs; Extended Data Fig. 10c,d). In  
317 accordance, CD112<sup>hi</sup>LT-HSCs exhibited the highest acetylation of H4K16 and the lowest proportion of  
318 *INKA1*<sup>hi</sup> cells compared to the other 3 populations, whereas CD112<sup>lo</sup>LT-HSCs showed the highest  
319 *INKA1* abundance (Fig. 7c,d and Extended Data Fig. 10e). No significant differences were seen between  
320 CD112<sup>hi</sup> and CD112<sup>lo</sup>ST-HSCs for all parameters assessed (Fig. 7b-e and Extended Data Fig. 10e-g).  
321 These data suggest that CD112 enables prospective enrichment of LT-HSCs in alternative states of  
322 quiescence in a context specific manner reminiscent of the LT-HSC specific phenotypes induced by  
323 *INKA1*-OE.

324

325 Quiescent LT-HSCs generate energy preferentially by glycolysis, in contrast to activated HSPC  
326 that rely on oxidative phosphorylation that yields reactive oxygen species (ROS)<sup>5</sup>. CD112<sup>hi</sup>LT-HSC and  
327 both ST-HSC-derived populations showed similar ROS signals (CellROX; Fig. 7e and Extended Data  
328 10f,g). By contrast CD112<sup>lo</sup>LT-HSCs produced less ROS indicating that CD112<sup>lo</sup>LT-HSCs are also  
329 metabolically less primed for activation. We have recently established chromatin accessibility signatures  
330 of quiescent (LT/HSPC signature) vs. activated states (Act/HSPC signature) of human HSPCs<sup>26</sup>. We  
331 generated chromatin accessibility signatures of CD112<sup>hi</sup> and CD112<sup>lo</sup>LT-HSCs by ATAC-seq and  
332 projected them onto our published single cell ATAC-seq data from CB HSPCs<sup>26</sup>. The CD112<sup>lo</sup> signature

333 was highly concordant with the LT/HSPC signature and inversely correlated with the Act/HSPC signature  
334 (Fig. 7f,g and Extended Data Figure 10h,i). By contrast, the more primed, CD112<sup>hi</sup> population generated  
335 an epigenetic signature that overlapped with both established signatures. Hence, CD112 surface  
336 expression also delineated alternative chromatin states within quiescent LT-HSCs. These distinct states  
337 were also revealed from GSEA of RNA-seq data comparing CD112<sup>hi</sup> and CD112<sup>lo</sup>LT-HSC. CD112<sup>hi</sup>LT-  
338 HSC upregulated several prototypical stem cell response pathways that were downregulated in LT-HSC  
339 upon *INKA1*-OE (Extended Data Fig. 10j) or included mTORC1 signaling (Fig. 7h)<sup>27</sup>. Overall, we  
340 retrieved 289 differentially expressed genes between these two distinctly quiescent LT-HSC subsets (Fig.  
341 7i and Supplementary Table 2). Further GSEA analysis revealed that CD112<sup>lo</sup> cells were enriched for the  
342 highly interconnected biological processes of cell adhesion, cytoskeleton regulation and GTPase signaling  
343 (Fig. 7j and Extended Data Fig. 10k,l). In summary, fractionation of LT-HSCs on the basis of CD112  
344 expression identified pre-existing, molecularly-distinct subsets that fall on opposite ends of the  
345 quiescence spectrum, thereby revealing previously undescribed heterogeneity within the LT-HSC subset.

346

### 347 **CD112<sup>lo</sup> captured LT-HSC with delayed repopulation kinetics**

348 To determine whether the quiescence-associated phenotype of CD112<sup>lo</sup>LT-HSCs is functional  
349 and associated with delayed activation and post-transplant stem cell latency similar to what was observed  
350 with *INKA1*-OE, we assessed the repopulation kinetics of CD112<sup>hi</sup> and CD112<sup>lo</sup>LT-HSCs from 3 CB  
351 pools. Human CD45 chimerism in mouse cohorts transplanted with CD112<sup>lo</sup>LT-HSCs tended to be lower  
352 than that of their CD112<sup>hi</sup>LT-HSC-transplanted counterparts at 4 wk ( $p=0.102$ ; Fig. 8a,b) and reaching  
353 significance at 12 wk ( $p=0.041$ ). The percentage of CD45<sup>+</sup> cells for both LT-HSC subsets increased from  
354 4 to 12 wk resulting in high median engraftment (>60% for 2 of 3 CB). Median engraftment for  
355 CD112<sup>lo</sup>LT-HSC remained stable thereafter to 20 wk, whereas it significantly decreased from 12 to 20 wk  
356 for CD112<sup>hi</sup>LT-HSC. Similar to *INKA1*-OE experiments, no overt mature lineage shift was detected  
357 (Extended Data Fig. 10m). In secondary LDA, CD112<sup>lo</sup>LT-HSCs from all 3 CB pools had a combined  
358 3.6-fold higher HSC frequency compared to their CD112<sup>hi</sup>LT-HSC counterparts (Fig. 7c,d and Extended

359 Data Fig. 10n). These data suggested that CD112<sup>hi</sup>LT-HSCs were primed for activation and able to  
360 generate hematopoietic output more rapidly; however, they subsequently showed signs of exhaustion in  
361 the primary recipients and possessed limited secondary transplantation capacity. By contrast, the delayed  
362 repopulation (latency) of low-primed (CD112<sup>lo</sup>) LT-HSCs was associated with superior retention of  
363 functional HSC, recapitulating the *in vivo* phenotype induced by both *INKA1*-OE or *PAK4*-KD. Thus,  
364 these data established the human LT-HSC pool was composed of at least two distinct cellular subsets  
365 marked by alternative states of quiescence. In the face of regeneration-mediated stress, these subsets  
366 showed differing physiological responses, one subset met acute demand while the other preserved  
367 stemness.

368

## 369 **DISCUSSION**

370 Our study provides key insights into the molecular underpinnings orchestrating the complex *in*  
371 *vivo* dynamics that human HSCs undergo during transplantation to permit rapid hematopoietic  
372 repopulation while also ensuring preservation of a stem cell reservoir. We identified functionally,  
373 transcriptionally and epigenetically distinct LT-HSC subsets that are positioned along a spectrum of  
374 varying depths of G<sub>0</sub> quiescence. The distinct G<sub>0</sub> states are defined by dichotomy between *INKA1*/*SIRT1*  
375 and H4K16ac and along a spectrum of *PAK4*, *CDK6* and *CD112* protein expression; a molecular network  
376 directly linking cell surface to chromatin state. The G<sub>0</sub> state that a LT-HSC resides in determines its long-  
377 term fate upon regeneration-mediated challenge. Activated LT-HSCs contribute rapidly to repopulation  
378 but lose self-renewal capacity and exhaust, while the most quiescent LT-HSCs exhibit latency in their  
379 repopulation kinetics but retain self-renewal and long-term regenerative capacity. The latency phenotype  
380 we uncovered illustrates the complexity of human HSC dynamics *in vivo*, and argues that latency, in  
381 addition to self-renewal, quiescence and differentiation, is a fundamental stem cell hallmark that  
382 preserves HSC pool size during regeneration-mediated demand.

383

384 Our discovery of functionally-distinct LT-HSC subsets is consistent with HSC clonal contribution  
385 dynamics from post-transplant gene therapy patients: Initial reconstitution occurs from transient clones,  
386 while long-term hematopoiesis originates from clones that only become activated after a 1-2-year latency  
387 phase<sup>2,28</sup>. Our genetic models, combining *ex vivo* lentiviral transduction with xenografting, are analogous  
388 to gene therapy protocols and to previous clonal tracking studies of mice and xenografts<sup>8,29</sup> that pointed to  
389 functional heterogeneity manifesting in two engraftment waves. Our study establishes that these distinct  
390 fates are neither stochastic nor induced by *ex vivo* culture but wired into distinct molecular states of  
391 human LT-HSCs. These can be prospectively isolated into latent (CD112<sup>lo</sup>) and primed (CD112<sup>hi</sup>) LT-  
392 HSC subsets even before exposure to regeneration-mediated stress. The resistance of latent LT-HSCs  
393 (CD112<sup>lo</sup> or *INKA1*-OE) to activation early during reconstitution highlights that the regeneration-  
394 mediated stress imposed by transplantation is in itself not a sufficient stressor to activate latent LT-HSC.  
395 However, certain stimuli including longer times of repopulation, secondary transplantation or 5-FU  
396 treatment lift this restraint and trigger self-renewal of latent LT-HSC. All these stimuli increase demand  
397 on ST-HSCs to generate continuous hematopoietic output leading to their exhaustion<sup>9,23</sup>. Thus, we  
398 speculate that this loss of immediate progeny either triggers direct activation of latent LT-HSC or lifts  
399 their restraint; a concept predicted by murine *in vivo* dynamics modeling<sup>30</sup>. The transient lineage output  
400 restraint we observed agrees with reported properties of murine low-output HSCs<sup>8</sup> but contrasts with  
401 transient myeloid-restricted “latent-HSCs” from aged mice<sup>31</sup>. Overall, these studies suggest that stimuli  
402 stronger than those experienced during primary transplantation might be needed to read out full stem cell  
403 potential.

404

405 While our earlier study<sup>3</sup> found differences in G<sub>0</sub> exit kinetics between LT- and ST-HSCs, here we  
406 uncovered distinct quiescence states prior to LT-HSC activation. The increased purity achieved by CD112  
407 enabled detection of changes in mTORC activity along the trajectory of LT-HSC cell priming that were  
408 previously obscured if only “bulk” of LT- versus ST-HSCs were compared<sup>3</sup>. Our findings are reminiscent  
409 to the mTorC1 activation properties of G<sub>Alet</sub> muscle stem cells that are primed to respond rapidly to

410 stress<sup>27</sup>. Embryonic diapause is a reversible H4K16ac-negative dormant state that requires mTOR  
411 signaling and the H4K16 acetyltransferase Mof for activation<sup>32,33</sup>. Thus, the molecular quiescence and  
412 priming programs we have uncovered in LT-HSC might be conserved across some species and tissues.  
413 The distinct latent and primed molecular signatures we ascribed to distinct CD112 immunophenotypes  
414 shed light on RhoGTPase signaling, stress responses and cell adhesion; processes all found relevant for  
415 murine HSC function<sup>34,35</sup>. CD112/PVRL2 is a ligand for the immune checkpoint receptors TIGIT and  
416 CD226<sup>36</sup> suggesting that cross talk between HSCs and T- and NK-cells might play a role in quiescence  
417 induction, HSC recruitment and/or immune surveillance<sup>37</sup>. Furthermore, CD112 upregulation upon cell  
418 cycling is similar to MHC-I in tissue stem cells<sup>38</sup> suggesting a role in quality control ensuring HSC pool  
419 integrity given their inefficient genomic repair during dormancy<sup>39</sup>. Conversely, INKA1-mediated CD112  
420 downregulation may facilitate immune privilege of quiescent HSCs<sup>40</sup>.

421  
422         Identifying the H4K16 deacetylase SIRT1 as an INKA1 interactor suggests a mechanism  
423 explaining the dichotomy between INKA1 and H4K16ac. SIRT1 is a stress sensor that regulates the  
424 cellular adaptation to metabolic stress while also exerting anti-inflammatory activity via negative  
425 regulation of mTOR<sup>41,42</sup>. H4K16ac, in addition to deacetylation by SIRT1, is also regulated by pH  
426 changes, depends on acetyl-CoA and is involved in the regulation of DNA damage response during the  
427 cell cycle<sup>17,43,44</sup>. Thus, H4K16ac may indicate the metabolic changes occurring upon HSC activation and  
428 cell cycle progression while also ensuring HSC quality control. The molecular interplay that arises from  
429 the complex INKA1-interactome we identified, points to additional mechanistic roles for INKA1 beyond  
430 PAK4 inhibition and emphasizes the need for studies across all blood cell lineages and other tissue  
431 hierarchies.

432  
433         Collectively, our work has uncovered a potential pivot point for the balance between varying  
434 hematopoietic demands and long-term sustainability of the hematopoietic system and identified latency as  
435 a fundamental property of stem cells that integrates quiescence control with HSC fate decisions.

436 **REFERENCES**

- 437 1 Bernitz, J. M., Kim, H. S., MacArthur, B., Sieburg, H. & Moore, K. Hematopoietic Stem Cells Count and  
438 Remember Self-Renewal Divisions. *Cell* **167**, 1296-1309 e1210, doi:10.1016/j.cell.2016.10.022 (2016).
- 439 2 Scala, S. *et al.* Dynamics of genetically engineered hematopoietic stem and progenitor cells after autologous  
440 transplantation in humans. *Nat Med* **24**, 1683-1690, doi:10.1038/s41591-018-0195-3 (2018).
- 441 3 Laurenti, E. *et al.* CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell*  
442 **16**, 302-313, doi:10.1016/j.stem.2015.01.017 (2015).
- 443 4 Takizawa, H. *et al.* Pathogen-Induced TLR4-TRIF Innate Immune Signaling in Hematopoietic Stem Cells  
444 Promotes Proliferation but Reduces Competitive Fitness. *Cell Stem Cell* **21**, 225-240 e225,  
445 doi:10.1016/j.stem.2017.06.013 (2017).
- 446 5 Walter, D. *et al.* Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells.  
447 *Nature* **520**, 549-552, doi:10.1038/nature14131 (2015).
- 448 6 Muller, L., Di Benedetto, S. & Pawelec, G. The Immune System and Its Dysregulation with Aging. *Subcell*  
449 *Biochem* **91**, 21-43, doi:10.1007/978-981-13-3681-2\_2 (2019).
- 450 7 Cabezas-Wallscheid, N. *et al.* Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell  
451 Dormancy. *Cell* **169**, 807-823 e819, doi:10.1016/j.cell.2017.04.018 (2017).
- 452 8 Rodriguez-Fraticelli, A. E. *et al.* Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis.  
453 *Nature* **583**, 585-589, doi:10.1038/s41586-020-2503-6 (2020).
- 454 9 Notta, F. *et al.* Isolation of single human hematopoietic stem cells capable of long-term multilineage  
455 engraftment. *Science* **333**, 218-221, doi:10.1126/science.1201219 (2011).
- 456 10 Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*  
457 **19**, 271-281, doi:10.1038/ncb3493 (2017).
- 458 11 Laurenti, E. *et al.* The transcriptional architecture of early human hematopoiesis identifies multilevel control  
459 of lymphoid commitment. *Nat Immunol* **14**, 756-763, doi:10.1038/ni.2615 (2013).
- 460 12 Kaufmann, K. B. *et al.* A stemness screen reveals C3orf54/INKA1 as a promoter of human leukemia stem  
461 cell latency. *Blood* **133**, 2198-2211, doi:10.1182/blood-2018-10-881441 (2019).
- 462 13 Dart, A. E. & Wells, C. M. P21-activated kinase 4--not just one of the PAK. *Eur J Cell Biol* **92**, 129-138,  
463 doi:10.1016/j.ejcb.2013.03.002 (2013).
- 464 14 Baskaran, Y. *et al.* An in cellulo-derived structure of PAK4 in complex with its inhibitor Inka1. *Nat Commun*  
465 **6**, 8681, doi:10.1038/ncomms9681 (2015).
- 466 15 Luo, T. *et al.* Inca: a novel p21-activated kinase-associated protein required for cranial neural crest  
467 development. *Development* **134**, 1279-1289, doi:10.1242/dev.02813 (2007).
- 468 16 Grigoryan, A. *et al.* LaminA/C regulates epigenetic and chromatin architecture changes upon aging of  
469 hematopoietic stem cells. *Genome Biol* **19**, 189, doi:10.1186/s13059-018-1557-3 (2018).
- 470 17 McBrian, M. A. *et al.* Histone acetylation regulates intracellular pH. *Mol Cell* **49**, 310-321,  
471 doi:10.1016/j.molcel.2012.10.025 (2013).

- 472 18 Hay, S. B., Ferchen, K., Chetal, K., Grimes, H. L. & Salomonis, N. The Human Cell Atlas bone marrow  
473 single-cell interactive web portal. *Exp Hematol* **68**, 51-61, doi:10.1016/j.exphem.2018.09.004 (2018).
- 474 19 Xie, S. Z. *et al.* Sphingolipid Modulation Activates Proteostasis Programs to Govern Human Hematopoietic  
475 Stem Cell Self-Renewal. *Cell Stem Cell* **25**, 639-653 e637, doi:10.1016/j.stem.2019.09.008 (2019).
- 476 20 La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494-498, doi:10.1038/s41586-018-0414-6  
477 (2018).
- 478 21 Haghverdi, L., Buttner, M., Wolf, F. A., Büttner, F. & Theis, F. J. Diffusion pseudotime robustly  
479 reconstructs lineage branching. *Nat Methods* **13**, 845-848, doi:10.1038/nmeth.3971 (2016).
- 480 22 Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq.  
481 *Science* **352**, 189-196, doi:10.1126/science.aad0501 (2016).
- 482 23 Randall, T. D. & Weissman, I. L. Phenotypic and functional changes induced at the clonal level in  
483 hematopoietic stem cells after 5-fluorouracil treatment. *Blood* **89**, 3596-3606 (1997).
- 484 24 Coyaud, E. *et al.* Global Interactomics Uncovers Extensive Organellar Targeting by Zika Virus. *Mol Cell*  
485 *Proteomics* **17**, 2242-2255, doi:10.1074/mcp.TIR118.000800 (2018).
- 486 25 Chen, Y. *et al.* Quantitative acetylome analysis reveals the roles of SIRT1 in regulating diverse substrates  
487 and cellular pathways. *Mol Cell Proteomics* **11**, 1048-1062, doi:10.1074/mcp.M112.019547 (2012).
- 488 26 Takayama, N. *et al.* The Transition from Quiescent to Activated States in Human Hematopoietic Stem Cells  
489 Is Governed by Dynamic 3D Genome Reorganization. *Cell Stem Cell* **28**, 488-501,  
490 doi:10.1016/j.stem.2020.11.001 (2021).
- 491 27 Rodgers, J. T. *et al.* mTORC1 controls the adaptive transition of quiescent stem cells from G0 to G(Alert).  
492 *Nature* **510**, 393-396, doi:10.1038/nature13255 (2014).
- 493 28 Biasco, L. *et al.* In Vivo Tracking of Human Hematopoiesis Reveals Patterns of Clonal Dynamics during  
494 Early and Steady-State Reconstitution Phases. *Cell Stem Cell* **19**, 107-119, doi:10.1016/j.stem.2016.04.016  
495 (2016).
- 496 29 Belderbos, M. E. *et al.* Donor-to-Donor Heterogeneity in the Clonal Dynamics of Transplanted Human Cord  
497 Blood Stem Cells in Murine Xenografts. *Biol Blood Marrow Transplant* **26**, 16-25,  
498 doi:10.1016/j.bbmt.2019.08.026 (2020).
- 499 30 Klose, M., Florian, M. C., Gerbaulet, A., Geiger, H. & Glauche, I. Hematopoietic Stem Cell Dynamics Are  
500 Regulated by Progenitor Demand: Lessons from a Quantitative Modeling Approach. *Stem Cells* **37**, 948-957,  
501 doi:10.1002/stem.3005 (2019).
- 502 31 Yamamoto, R. *et al.* Large-Scale Clonal Analysis Resolves Aging of the Mouse Hematopoietic Stem Cell  
503 Compartment. *Cell Stem Cell* **22**, 600-607 e604, doi:10.1016/j.stem.2018.03.013 (2018).
- 504 32 Hussein, A. M. *et al.* Metabolic Control over mTOR-Dependent Diapause-like State. *Dev Cell* **52**, 236-250  
505 e237, doi:10.1016/j.devcel.2019.12.018 (2020).
- 506 33 Khoa, L. T. P. *et al.* Histone Acetyltransferase MOF Blocks Acquisition of Quiescence in Ground-State ESCs  
507 through Activating Fatty Acid Oxidation. *Cell Stem Cell* **27**, 441-458, doi:10.1016/j.stem.2020.06.005  
508 (2020).
- 509 34 Reddy, P. N. *et al.* p21-activated kinase 2 regulates HSPC cytoskeleton, migration, and homing via CDC42  
510 activation and interaction with beta-Pix. *Blood* **127**, 1967-1975, doi:10.1182/blood-2016-01-693572 (2016).



- 511 35 Yang, L. *et al.* Rho GTPase Cdc42 coordinates hematopoietic stem cell quiescence and niche interaction in  
512 the bone marrow. *Proc Natl Acad Sci U S A* **104**, 5091-5096, doi:10.1073/pnas.0610819104 (2007).
- 513 36 Andrews, L. P., Yano, H. & Vignali, D. A. A. Inhibitory receptors and ligands beyond PD-1, PD-L1 and  
514 CTLA-4: breakthroughs or backups. *Nat Immunol* **20**, 1425-1434, doi:10.1038/s41590-019-0512-0 (2019).
- 515 37 Kalia, V., Penny, L. A., Yuzefpolskiy, Y., Baumann, F. M. & Sarkar, S. Quiescence of Memory CD8(+) T  
516 Cells Is Mediated by Regulatory T Cells through Inhibitory Receptor CTLA-4. *Immunity* **42**, 1116-1129,  
517 doi:10.1016/j.immuni.2015.05.023 (2015).
- 518 38 Agudo, J. *et al.* Quiescent Tissue Stem Cells Evade Immune Surveillance. *Immunity* **48**, 271-285 e275,  
519 doi:10.1016/j.immuni.2018.02.001 (2018).
- 520 39 Beerman, I., Seita, J., Inlay, M. A., Weissman, I. L. & Rossi, D. J. Quiescent hematopoietic stem cells  
521 accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* **15**, 37-50,  
522 doi:10.1016/j.stem.2014.04.016 (2014).
- 523 40 Hirata, Y. *et al.* CD150(high) Bone Marrow Tregs Maintain Hematopoietic Stem Cell Quiescence and  
524 Immune Privilege via Adenosine. *Cell Stem Cell* **22**, 445-453 e445, doi:10.1016/j.stem.2018.01.017 (2018).
- 525 41 Ryall, J. G. *et al.* The NAD(+)-dependent SIRT1 deacetylase translates a metabolic switch into regulatory  
526 epigenetics in skeletal muscle stem cells. *Cell Stem Cell* **16**, 171-183, doi:10.1016/j.stem.2014.12.004 (2015).
- 527 42 Wang, Y. *et al.* Histone Deacetylase SIRT1 Negatively Regulates the Differentiation of Interleukin-9-  
528 Producing CD4(+) T Cells. *Immunity* **44**, 1337-1349, doi:10.1016/j.immuni.2016.05.009 (2016).
- 529 43 Cluntun, A. A. *et al.* The rate of glycolysis quantitatively mediates specific histone acetylation sites. *Cancer*  
530 *Metab* **3**, 10, doi:10.1186/s40170-015-0135-3 (2015).
- 531 44 Sharma, G. G. *et al.* MOF and histone H4 acetylation at lysine 16 are critical for DNA damage response and  
532 double-strand break repair. *Mol Cell Biol* **30**, 3582-3595, doi:10.1128/MCB.01476-09 (2010).

533

534

### 535 **Acknowledgments**

536 We thank the obstetrics units of Trillium Health and William Osler Health Partners for CB; The UHN-  
537 SickKids Flow cytometry facility for cell sorting; M. Minden and the Leukemia Tissue Bank at Princess  
538 Margaret Cancer Centre/University Health Network for providing primary mPB samples. We thank E.  
539 Laurenti and all members of the Dick lab, in particular J.C.Y. Wang, for critical feedback. This research  
540 was supported by the Deutsche Forschungsgemeinschaft (K.B.K). and is part of the University of  
541 Toronto's Medicine by Design initiative which receives funding from the Canada First Research  
542 Excellence Fund. Work in J.E.D.'s laboratory is supported by funds from the Princess Margaret Cancer  
543 Centre Foundation, the Canadian Institutes of Health Research (Foundation: 154293 (J.E.D)), Operating

544 Grants # 154293 and # 89932 (J.E.D), International Development Research Centre, Canadian Cancer  
545 Society (Grant #703212 (J.E.D)), Terry Fox Research Institute Program Project Grant, Ontario Institute  
546 for Cancer Research through funding provided by the Government of Ontario, a Canada Research Chair,  
547 and the Ontario Ministry of Health and Long Term Care.

548

#### 549 **Author Contributions**

550 K.B.K. conceived the study, performed research, analyzed data, and wrote the manuscript; A.G.X.Z.,  
551 A.M. and M.C.S.Y. analyzed data; E.C., E.P. and S.Z.X. performed research and analyzed data; L.G.P.,  
552 E.M.N.L., O.I.G., K.P., J.M., H.B., S.Z. and S.T. performed research; R.S. and B.R. supervised specific  
553 experiments; A.G.X.Z., O.I.G and S.Z.X. edited the manuscript; S.Z.X. provided conceptual input; J.E.D.  
554 wrote the manuscript, secured funding and supervised the study.

555

#### 556 **Competing Interests Statement**

557 J.E. Dick served on the SAB at Trillium Therapeutics, reports receiving a commercial research grant from  
558 Celgene, and has ownership interest (including patents) in Trillium Therapeutics Inc.; S.T. is employed by  
559 Kirin Holdings Company, Ltd.; all other authors declare no competing interests.

560

561

#### 562 **FIGURE LEGENDS**

563

564 **Fig. 1: INKA1 and H4K16ac show inverse abundance and visualize heterogeneity in HSC subsets.**

565 (a) *INKA1/FAM212A* and *PAK4* transcript expression in CB RNA-seq data (upper, GSE125345) and as  
566 means of inferred subpopulation from bone marrow (BM) single cell data (lower, Hay *et al* 2018).

567 Corresponding subpopulations (delineated by dotted lines) are listed in middle panel, designation as used  
568 in text. *Prog* = *Progenitors*.

569 (b) Gating strategy to obtain four HSC subsets (detailed in Extended Data 1a). *IT* = *intermediate*.

570 (c) Confocal images of 2 individual CD90<sup>+</sup>CD49f<sup>+</sup> LT-HSCs showing the two extremes of the mutually  
571 exclusive abundance of INKA1 and H4K16ac representative for n=3 independent CB as summarized in  
572 (d). *Scale bar = 2 μm.*

573 (d) Mean fluorescence intensity (MFI) of single cells for nuclear (nucl.) INKA1 and H4K16ac by  
574 confocal analysis (Ø50 cells/population and CB, 3 CB pools). *INKA1<sup>hi</sup> cell gate (dashed) as for Fig. 1e*  
575 *and Extended Data Fig. 1d).*

576 (e) Nuclear diameter of INKA1<sup>hi</sup> and INKA1<sup>lo</sup> (remainder) cells from 3 CB and 4 HSC subsets combined,  
577 *two-sided Mann-Whitney-test (\*\*p<0.0001).*

578 (f) Immunostaining of LT-and ST-HSCs for INKA1 and CDK6 (120 cells/population, 3 CB pools),  
579 CDK6 MFI is shown for cells fractionated into INKA1<sup>hi</sup> (in analogy to (d)) and the remainder (INKA1<sup>lo</sup>);  
580 *two-sided t-test (\*\*p=0.00003).*

581

582 **Fig. 2: Single cell transcriptome and protein analysis validates that *INKA1* vs its target *PAK4* and**  
583 **H4K16ac are on opposing ends of cell cycle priming and progression.**

584 (a) Normalized nuclear CDK6 intensity in LT- and ST-HSCs (CB and mPB; n=4; Ø50 cells/population  
585 and sample; immunostaining). *two-sided Mann-Whitney-test (\*\*p=0.0036, \*\*\*p<0.0001).*

586 (b) Percentage CDK6<sup>±</sup> LT- and ST-HSCs within G<sub>0</sub> (Ki67<sup>-</sup>) from mPB (n=3, flow cytometry),  
587 *Mean±SD, two-sided paired t-test (\*p=0.020, \*\*p=0.004).*

588 (c) Diffusion mapping of scRNAseq data from CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> mPB cells using RNAvelocity.  
589 *DC = diffusion component, superimposed scaled color gradient: early (0.0) to late (1.0) pseudotime.*

590 (d) Indicated transcript expression against diffusion pseudotime, curves fitted by nonlinear regression.

591 (e,f) CB Lin<sup>-</sup> cells cultured for 2 d with 20 hr EdU pulse to subgate cell cycle phases (n=5; flow  
592 cytometry), *paired t-test(e), one-way ANOVA(f); minimal to maximal values are represented in box-and-*  
593 *whisker plot with median as centre.*

594 (g) CB LT- and ST-HSCs analyzed for Ki67, H4K16ac, PAK4 and DAPI after 72 hrs of culture (high  
595 cytokine conditions; n=3; Ø50 cells/population/sample; immunostaining), *one-way ANOVA* (\*\* $p < 0.01$ ;  
596 \*\*\* $p < 0.001$ ).

597 (h) Excerpt of confocal analysis of cultured LT- and ST-HSCs harvested at 5 timepoints (Ø50  
598 cells/population/timepoint/sample; n=3). *One-way ANOVA with Welch's correction; significance level #*  
599 *and \* as compared to 6 h timepoint or pair-wise (LT- vs ST-HSCs) for the same timepoint, respectively*  
600 *# $p < 0.05$ , ## $p < 0.001$ , ### $p < 0.001$ ; \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ); Scale bar: 4  $\mu\text{m}$ .*

601 (i) Protein diffusion pseudotime analysis of 3442 cells from 5 timepoints, four CD45RA<sup>-</sup> HSC subsets  
602 (defined by CD90 and CD49f; n=3) and Progenitors (n=2) by immunostaining. Timepoints or marker  
603 levels (gradients) as indicated. Longitudinal marker levels of HSC subsets across pseudotime were  
604 inferred by nonlinear regression (upper right).

605

606 **Fig. 3: PAK4 inhibition induces G<sub>0</sub> arrest *in vitro* maintaining stem cell regenerative capacity.**

607 Sorted CB CD34<sup>+</sup>CD38<sup>-</sup> HSPC and CD34<sup>+</sup>CD38<sup>+</sup> progenitors were cultured, after 2 days medium was  
608 supplemented with 0.25  $\mu\text{M}$  PF-3758309 and 5 days later cells were analyzed for cell counts, viability  
609 and cell cycle or injected into NSG mice.

610 (a) Fold expansion according to input and cell count on day 7 of culture (3 CB pools; HSPC:  
611 CD34<sup>+</sup>CD38<sup>-</sup>; Prog: CD34<sup>+</sup>CD38<sup>+</sup>), *Mean  $\pm$  SD, two-sided t-test* (\* $p = 0.048$ , \*\*\* $p = 0.00045$ ).

612 (b) Cell cycle analysis of 5 d treated cultures (n=4), *Mean  $\pm$  SD, two-sided paired t-test* (\*\* $p < 0.01$ ).

613 (c) HSC frequencies as estimated by primary LDA from 5 d treated cultures transplanted into a total of  
614 105 NSG (16-20 wk, 6 CB donors/pools, each dot representative of 6-10 injected mice, hCD45  
615 engraftment > 0.1%). *Initial cell culture input (resulting progeny injected/mouse): 160-500 (low), 800-*  
616 *2500 (mid), 4000-12500 (high dose). Mean frequency indicated by fraction numbers. Frequency was*  
617 *calculated either according to the initial cell doses cultured irrespective of the actual cell number*

618 *transplanted or according to the actual number of progeny after treatment transplanted. TX=*  
619 *transplantation.*  
620 (d) Lineage composition (human CD45<sup>+</sup>) of engrafted primary NSG mice (n=9-12). *Mean±SD.*  
621 (e,f) Individual assay (e) and combined (f) HSC frequency estimates according to secondary LDA (8 wk)  
622 in NSG-SGM3 mice upon transplantation of sorted hCD45<sup>+</sup> cells from NSG grafts generated from  
623 cultured and treated HSPC. *∅: in 1 of 4 assays no mouse engrafted in DMSO cohort resulting in a*  
624 *frequency of 1/∞. CD45<sup>+</sup> cell dose transplanted into secondary recipients (n=86): 180K-540K (high),*  
625 *60K-180K (mid), 20K-60K (low).*

626

627 **Fig. 4: *INKAI*-OE LT-HSCs resist activation *in vitro*.**

628 (a,b) Immunostaining of prospectively isolated, transduced (OE) and BFP<sup>+</sup>-sorted CB LT- and ST-HSCs  
629 (n=3, day 6 post-transduction), *one-sided Kruskal-Wallis-test; minimal to maximal values are represented*  
630 *in box-and-whisker plot with median as centre (\*p<0.005, \*\*\*p<0.001).*  
631 (c,d) Gene sets enriched upon *INKAI*-OE vs CTRL in LT-HSCs or ST-HSCs with FDR q-value <0.1 at  
632 least in one dataset (RNA-seq, n=3, d6 ptd). *Kolmogorov-Smirnov statistic \*FDR < 0.10; \*\*FDR*  
633 *<0.05; \*\*\*FDR<0.01.*

634

635 **Fig. 5: *INKAI*-OE and *PAK4*-KD preserve long-term stem cell regenerative potential under**  
636 **regeneration-mediated stress.**

637 Progeny of 10,000 CB Lin<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup> cells were transduced with OE or knock-down (KD) vectors and  
638 transplanted into NSG mice 1 day later. *Input: mean of transduction efficiency at day 3 and 6.*

639 (a) Relative engraftment of transduced cells (log<sub>2</sub> ratios: %BFP<sup>+</sup> of CD45<sup>+</sup> *in vivo* output injected  
640 femur/%BFP<sup>+</sup> *in vitro* input) of 1 to 5 independent experiments, *Mean±SD, one-way ANOVA with*  
641 *Tukey's post hoc test (\*\*p=0.0078 and p=0.0027, \*\*\*p<0.0001).*

642 (b,c) Scored engraftment ( $>0.1\%$  hCD45<sup>+</sup>) of secondary transplants from bone marrow (BM) grafts  
643 harvested and CD45<sup>+</sup>BFP<sup>+</sup> sorted ((injected cell dose indicated) at multiple timepoints (2w, 4w, 20w) ;  
644 20w grafts were assessed in secondary LDA (n=3; c).  
645 (d) Flow cytometric analysis of pooled BM from 20 w primary xenografts (n=4) gated on  
646 CD45<sup>+</sup>BFP<sup>+</sup>CD34<sup>+</sup>CD19<sup>-</sup>CD38<sup>-</sup> population, *Mean $\pm$ SD, two-sided paired t-test.*  
647 (e,f) NSG mice (n=80) were transplanted with CTRL or *INKA1*-OE CB CD34<sup>+</sup>CD38<sup>-</sup> cells. At 4 w 150  
648 mg/kg 5-Fluorouracil (5-FU) or PBS equivalent was injected intraperitoneally. Multiplicity of infection  
649 was adjusted to result at comparable %BFP<sup>+</sup> at 4 w for CTRL and *INKA1*-OE; untreated xenografts at 4 w  
650 served as baseline.  
651 (e) BFP<sup>+</sup> cell counts of input and in BM of injected femurs as inferred by total cell count/femur x %BFP<sup>+</sup>.  
652 *Mean $\pm$ SEM of 5-10 mice/timepoint and condition, two-sided Mann-Whitney-test.*  
653 (f) Total BM (20w) of 5-FU or PBS treated mice was subjected to hierarchy analysis.  
654 (g,h) Secondary LDA analysis of CD45<sup>+</sup>BFP<sup>+</sup> cells from 20 w primary grafts (+8wk;  $>0.1\%$  hCD45<sup>+</sup>).  
655 (i,j) Relative %BFP<sup>+</sup> engraftment (primary xenografts; i, *Mean $\pm$ SD, two-sided t-test, \* $p<0.05$ ;*  
656 *\*\* $p=0.0010$ ) and HSC frequency by secondary LDA (8 wk) for KD vectors (j).*

657

658 **Fig. 6: INKA1/PAK4 interactome identifies SIRT1 as link between INKA1 and H4K16ac and**  
659 **CD112 as a HSPC marker.**

660 (a) INKA1-PAK4 protein interaction network by BioID. *Edge width indicates SpecSum/abundance of*  
661 *biotinylated peptides; node color resembles differential RNA expression in CB LT- vs. ST-HSCs ( $\log_2$ ;*  
662 *GSE125345). \*previously reported PAK4 interactors.*

663 (b) CB CD34<sup>+</sup>CD38<sup>-</sup> HSPC and CD38<sup>+</sup>CD38<sup>+</sup> progenitors were subjected to proximity ligation assay  
664 (PLA) for INKA1 and SIRT1 (n=3,  $\emptyset$ 100 cells/ sample). *Scale bar =6  $\mu$ m. Mean $\pm$ SD, two-sided t-test*  
665 *(\*\* $p=0.005$ ).*

666 (c) CTRL and INKA1-OE CB CD34<sup>+</sup>CD38<sup>-</sup> HSPC were BFP<sup>+</sup> sorted 6d after transduction followed by  
667 immunostaining for SIRT1 and INKA1 and confocal microscopy. Relative distribution of SIRT1 in

668 nucleus was calculated dividing nuclear intensity by total cell intensity (n=3; Ø40 cells/ sample). *Scale*  
669 *bar =5 µm; two-sided t-test with Welch's correction (\*\*p=0.008).*

670 (d, e) Flow cytometric analysis of CB Lin<sup>-</sup> cells (n=11, d) and CB mononuclear cells (MNC, n=4, e) for  
671 CD112 MFI across subsets. *Mean±S, two-sided t-test (\*\*p<0.0001).*

672 (f) Cultured and transduced HSPC were analyzed 6 days later for CD112 levels of subgated cLT-  
673 (CD34<sup>+</sup>CD45RA<sup>-</sup>CD90<sup>+</sup>), cST-(CD34<sup>+</sup>CD45RA<sup>-</sup>CD90<sup>-</sup>) HSCs, cProgenitors (CD34<sup>+</sup>CD45RA<sup>+</sup>) or total  
674 cHSPCs (CD34<sup>+</sup>) *c= cultured.*

675 (g) LT-HSCs preincubated with anti-CD112 before fixation to only stain for cell surface expression were  
676 subjected to immunostaining and quantified intensity for total INKA1 and surface CD112 were plotted  
677 against each other (n=68 from 3 CB). *Scale bar =2 µm.*

678 (h) *Nectin2/CD112* RNA expression levels against scRNAseq diffusion pseudotime.

679

680 **Fig. 7: CD112 surface expression delineates alternatively quiescent states of LT-HSCs.**

681 (a) Gating strategy for sorting CD112<sup>hi</sup> and CD112<sup>lo</sup> CB CD90<sup>+</sup>CD49f<sup>+</sup> LT- and CD90<sup>-</sup>CD49f<sup>-</sup> ST-HSCs.

682 (b) LT- and ST-HSCs sorted for CD112<sup>hi</sup> and CD112<sup>lo</sup> and subsequently analyzed by intracellular flow  
683 cytometry for CDK6 and Ki67; *two-sided paired t-test.*

684 (c,d) Immunostaining of CD112<sup>hi</sup> and CD112<sup>lo</sup> sorted LT- and ST-HSCs (n=5-7; Ø50 cells/population and  
685 sample), *one-way ANOVA and unpaired t-test; minimal to maximal values are represented in box-and-*  
686 *whisker plot with median as centre (c; \*\*\*p<0.0001)and two-sided Mann-Whitney test (d; \*\*p=0.0067).*

687 (e) Flow cytometric analysis of CellROX labelling.

688 (f,g) LT-HSCs sorted for CD112<sup>lo</sup>(n=2) and CD112<sup>hi</sup>(n=3) were subjected to ATAC-seq. (f) shows  
689 UMAP 2D projection of the enrichment of all 10 NMF signatures previously identified<sup>26</sup> in CD34<sup>+</sup>CD38<sup>-</sup>  
690 or CD34<sup>+</sup>CD38<sup>+</sup> CB single cells (combined shown), *colours indicate the enrichment (Z-Score) of the*  
691 *CD112<sup>lo</sup> or CD112<sup>hi</sup> LT-HSC specific regions. Red indicates enrichment, blue depletion. Pearson's*  
692 *correlation between the single-cell enrichment of the CD112<sup>lo</sup> or CD112<sup>hi</sup> LT-HSC specific regions and*  
693 *the LT/HSPC (yellow) and Act/HSPC (green) signatures<sup>26</sup> is shown in (g).*

694 (h) Hallmark GSEA of CD112<sup>hi</sup> vs. CD112<sup>lo</sup> LT-HSCs and *INKA1* expression (bulk RNA-seq).  
695 *Kolmogorov-Smirnov statistic*.  
696 (i) Differential gene expression for CD112<sup>hi</sup> vs. CD112<sup>lo</sup> LT-HSCs (RNA-seq, 2-3 independent CB  
697 pools).  
698 (j) Enrichment map visualization of GSEA results for CD112<sup>hi</sup> versus CD112<sup>lo</sup> RNA-seq data (3472 GO  
699 biological processes; gene set size 25-500; FDR<0.1).

700

701 **Fig. 8: CD112<sup>hi</sup> LT-HSCs are primed whereas CD112<sup>lo</sup> demarcates latent LT-HSCs.**

702 Prospectively isolated CD112<sup>hi</sup> and CD112<sup>lo</sup>LT-HSCs from 3 independent CB pools were transplanted  
703 separately at equal cell doses (as indicated) into NSG mice (n=90).

704 (a, b) Engraftment kinetics according to human CD45<sup>+</sup> chimerism in injected femurs, two-sided *paired t-*  
705 *test; minimal to maximal values are represented in box-and-whisker plot with median as centre*  
706 *(\*p<0.05)*.

707 (c) HSC frequency of individual CB pools estimated by secondary LDA of sorted hCD45<sup>+</sup> cells from 12  
708 w primary xenografts transplanted into NSG-SGM3 mice and analyzed after 8 w, *two-sided ratio paired*  
709 *t-test (\*p=0.0476)*.

710 (d) Combined secondary LDA analysis from 3 CB pools. *Chisquare likelihood ratio test statistic*.

711

712

713

714 **METHODS**

715

716 **Materials and Reagents**

717 Detailed information of commercially available reagents and other materials used in this study is  
718 summarized in Supplementary Table 3.

719



720 **Human Specimen Collection**

721 Human CB samples were obtained with informed consent from William Osler, Trillium and Credit Valley  
722 Hospitals according to procedures approved by the University Health Network (UHN) Research Ethics  
723 Board. Mononuclear cells (MNC) from single CB units or independent pools of male and female CB units  
724 (~4-15 units) were obtained by centrifugation on Lymphoprep medium, and after ammonium chloride  
725 lysis MNC were depleted of lineage positive cells (lineage depletion) by negative selection with the  
726 StemSep Human Progenitor Cell Enrichment Kit according to the manufacturer's protocol. Resulting  
727 lineage negative (Lin-) CB cells were stored viably at -80°C or -150°C. Human G-CSF mobilized  
728 peripheral blood samples (mPB, female and male, age: 40-63 years) were obtained with informed consent  
729 from the Leukemia Tissue Bank at Princess Margaret Cancer Centre/University Health Network  
730 according to procedures approved by the UHN Research Ethics Board. Following thawing and MNC  
731 isolation, CD34+ cells from mPB were enriched by positive selection with the CD34 Microbead kit  
732 (Miltenyi) and LS column purification with MACS magnet technology (Miltenyi).

733

734 **Mice**

735 Animal experiments were done in accordance with institutional guidelines approved by the University  
736 Health Network (UHN) Animal care and we complied with all relevant ethical regulations for animal  
737 testing and research. NOD.Cg PrkdcscidII2rgtm1Wjl /SzJ (NSG) and NOD.Cg-  
738 *Prkdc<sup>scid</sup>Il2rg<sup>tm1Wjl</sup>*Tg(CMV-IL3,CSF2,KITLG)1Eav/MloySzJ (NSG-SGM3) mice were housed in a  
739 controlled environment with 12:12 light-dark cycle with a 30 min transition, room temperature of 21–23  
740 °C and humidity of 30-60%, and had ad libitum access to dry laboratory food and water .at the animal  
741 facility (ARC) at Princess Margaret Cancer Centre, They were kept in a room designated only for  
742 immunocompromised mice with individually ventilated racks equipped with complete sterile micro-  
743 isolator caging (IVC), on corn-cob bedding and supplied with environmental enrichment in the form of a  
744 red house/tube and a cotton nestlet.

745 Cages are changed every <7 days under a biological safety cabinet. Health status is monitored using a  
746 combination of soiled bedding sentinels and environmental monitoring.

747

#### 748 **Flow Cytometric Analysis and Sorting**

749 Primary human cells were thawed by dropwise addition of X-VIVO 10 + 50% fetal calf serum  
750 supplemented with DNase (100 µg/mL final concentration) and resuspended at a density of  $5 \times 10^6$   
751 cells/mL. Cells were then incubated (30 min, 4°C) with the following antibodies: FITC- or BV421-anti-  
752 CD45RA (1/50), PE- or APC-anti-CD90 (1/100), PECy5-anti-CD49f (1/50), PECy7-anti-CD38 (1/100),  
753 APCCy7-anti-CD34 (1/200) and if applicable APC-anti-CD112(1/100) for sorting and/or flow cytometric  
754 analysis or anti-Nectin2 (goat) for subsequent surface immunostaining via confocal analysis. After  
755 washing cells were resuspended at  $0.5-10 \times 10^6$  cells/mL in PBS, 5% FBS supplemented with 0.1 µg/L  
756 propidium iodide (PI) and sorted on a BD FACS Aria III instrument.

757 For flow cytometric analysis of CB MNC in addition to SYTOX blue the following antibodies were used:  
758 V500-anti-CD45 (1/100), APC-Cy7-anti-CD34 (1/200), PE-Dazzle-CD38 (1/100), APC-anti-CD112  
759 (1/100), BV786-anti-CD33 (1/100), PE-anti-CD19(1/100), PE-anti-GlyA (1/100), FITC-anti-CD3  
760 (1/100). Cell preparation for intracellular staining of sorted cells was performed with BD  
761 Cytofix/Cytoperm as described before<sup>3</sup> and incubated with antibodies as applicable (FITC or PE-anti-  
762 Ki67 (1/30), AF647-anti-CDK6 (1/30), AF488-anti-H4K16ac (1/100)) in PermWash solution (BD)  
763 overnight. If applicable, Edu Click-IT reaction was performed according to manufacturer's instructions  
764 and prior to flow cytometry analysis cells were stained with Hoechst 33342 (1:2000; Thermo Fisher) or  
765 DAPI (1 µg/mL; Sigma). Other flow cytometry-based assays were performed according to the  
766 manufacturer's protocol: AnnexinV-FITC or -APC Kit (BD) and CellROX staining (ThermoFisher) in X-  
767 Vivo 10, 1% BSA, L-Glutamine, Pen/Strep. Cells were sorted on BD FACS Aria III or analyzed on BD  
768 Canto II, BD Fortessa or BD Celesta instruments. Positive gates were set according to unstained or  
769 isotype controls. Data was analyzed by FlowJo.

770

771 **Immunostaining**

772 Sorted cells were spun onto Poly-L-Lysine (Sigma)-coated slides (Ibidi, 200 xg, 10 min), fixed with 4%  
773 formaldehyde (Sigma) and permeabilized with 0.5% Triton X-100 (Sigma) before blocking (PBS, 10%  
774 FBS, 5% BSA; 30 min, RT). Slides were incubated with primary antibodies (Anti-INKA1(1/100), Anti-  
775 PAK4 (1/50), AF647-Anti-CDK6 (1/50), AF488-Anti-H4K16ac(1/100), Anti-SIRT1(1/100), Anti-  
776 Nectin2 (1/100), AF647-Anti-Ki67 (1/100)) in blocking solution O/N at 4°C. If necessary Anti-  
777 FAM212A and Anti-PAK4 antibodies were directly labeled using the Zip Alexa Fluor 555 Rapid  
778 Antibody Labeling Kit (Thermo Fisher) at a 0.1x antibody concentration but otherwise according to  
779 manufacturer's instructions to allow co-staining with other directly labeled rabbit or mouse antibodies,  
780 respectively. As applicable, secondary antibodies (1/400) were added (PBS, 0.025% Tween 20, Sigma,  
781 1.5 h, RT). After washing, nuclei were stained with 1 µg/mL DAPI (Sigma) and slides were mounted  
782 (FluoromountG). Single cell images were captured by a Zeiss LSM700 Confocal Microscope (oil,  
783 63x/1.4NA, Zen 2012) and processed and analyzed with ImageJ/Fiji and FlowJo10 or Python. In case of  
784 high inter-experimental variability between independent stainings data were normalized by z-score  
785 transformation. Pixel Intensity Spatial Correlation Analysis of colocalization was carried out using the  
786 Coloc2/Fiji plugin for ImageJ Software to obtain Spearman's rank correlation values for a region of  
787 interest incorporating either an individual cell, its nucleus or its non-nuclear area. The Coloc2 plugin was  
788 run with a PSF of 4.1-4.7 depending on channel combination and with 50 Costes' randomizations. For  
789 pseudotime analysis using single cell immunofluorescence data, fluorescence values were log transformed  
790 and scaled (CDK6, nuclear CDK6, DAPI, H4K16ac, PAK4, and nuclear PAK4). Diffusion mapping  
791 embedding was subsequently calculated as well as diffusion pseudotime<sup>21</sup> within scanpy. For the  
792 CellROX assay sorted cells were washed, resuspended in cytokine-free medium (X-Vivo 10, 1% BSA, L-  
793 Glutamine, Pen/Strep) supplemented with 5 µM CellROX Green and then incubated for 30 min at 37°C in  
794 Poly-L-Lysine coated chambered glass coverslips/ µ-slides (Ibidi). Subsequently, slides were spun (200  
795 xg, 10 min), medium was removed, and cells were fixed, washed, permeabilized and blocking was  
796 performed as before. Slides were then incubated with AF647-anti-CDK6 (1/50) antibody in blocking

797 solution for 1 hr at RT. After washing, nuclei were stained with 1  $\mu\text{g}/\text{mL}$  DAPI, chambers were washed,  
798 covered with FluoromountG and images were captured by a Zeiss LSM700 Confocal Microscope  
799 (20x/0.8 NA, Zen 2012).

800

### 801 **Proximity Ligation Assay**

802 Proximity ligation assay (PLA) was performed after sample preparation for immunostaining as described  
803 followed by using the Duolink® In Situ Orange Starter Kit Mouse/Rabbit (Millipore Sigma) according to  
804 manufacturer's instructions with few adaptations: Before primary antibody incubation overnight in PBS,  
805 10% FBS, 5% BSA, slides were blocked with PBS, 10% FBS, 5% BSA for 5 hrs at 4°C. After two  
806 washes with Wash Buffer A (Kit), slides were incubated with the Duolink® Blocking solution for 3 hrs at  
807 37°C. For all subsequent steps manufacturer's instructions were followed. Images were captured by a  
808 Zeiss LSM700 Confocal Microscope (20x/0.8 NA, Zen 2012) and processed and analyzed with  
809 ImageJ/Fiji. Background signal was determined according to several control HSPC/Progenitor specimens  
810 incubated with either no primary antibody (AB), no primary or secondary AB, anti-SIRT1(1/100) as only  
811 primary AB, anti-FAM212A (1/100) as only primary AB, anti-SIRT191/100) or anti-PAK4 (1/50) in  
812 combination with rabbit IgG (anti-V5; 1/100), anti-FAM212A in combination with mouse IgG (anti-  
813 BrdU; (1/100 or 1/50; respectively)).

814

### 815 **Chromatin fiber assay**

816 For visualizing epigenetic information and proteins bound to chromatin extended chromatin fibers were  
817 prepared as described elsewhere<sup>45,46</sup> with few modifications: CB HSPC CD34<sup>+</sup>CD38<sup>-</sup> were resuspended in  
818 PBS at  $2.5 \times 10^5$  cells/mL, 8  $\mu\text{l}$  were spotted as a line on a glass slide and air-dried before 150  $\mu\text{l}$  lysis  
819 buffer (500 mM NaCl, 500 mM Urea, 25 mM Tris, 1% Triton X-100, pH 7.5) was added. After 15 min at  
820 RT slides were placed in a 25° angle to extend fibers by gravity. For fixation slides were treated for 10  
821 min with 4% PFA and then washed for 30 min in PBS. After 1 hr of blocking (PBS, 10% FBS, 5% BSA)  
822 primary antibody incubation followed (90 min, RT; 1/50 in blocking solution). Subsequently slides were

823 washed in PBS, 0.025% Tween (10 min, RT) and incubated with secondary antibodies (1/400 in PBS,  
824 0.025% Tween; 90 min, RT). After 1x washing in PBS, 0.025% Tween (10 min), slides were washed 2x  
825 with PBS and incubated with 5  $\mu\text{g}/\text{mL}$  Hoechst 33342 in PBS (30 min, RT). Slides were then washed  
826 (PBS), mounted with FluoromountG and analyzed with a Zeiss LSM700 Confocal Microscope (oil,  
827 63x/1.4NA, Zen 2012).

828

### 829 **Cell Culture**

830 For the time course experiment ~1000 sorted cells were cultured in Terasaki plates in 25  $\mu\text{l}$ . To align  $G_0$   
831 exit kinetics with previous data<sup>3</sup> cells were cultured under high cytokine conditions: StemPro-34 SFM  
832 media (Thermo Fisher) with the supplied supplement, 1x L-Glutamine, 1x Pen/Strep, 0.02% Human LDL  
833 and the following cytokines: FLT3L (20 ng/mL), GM-CSF (20 ng/mL), SCF (100 ng/mL), TPO  
834 (100 ng/mL), EPO (3 U/mL), IL-3 (10 ng/mL), IL-6 (50 ng/mL). Cells were harvested at indicated  
835 timepoints and subjected to immunostaining. For all other primary cell cultures cells were cultured at a  
836 density of  $0.4\text{-}5 \times 10^5$  cell/mL in 96-well round-bottom plates with low-cytokine condition media: X-VIVO  
837 10 media, 1 % BSA supplemented with 1x Pen/Strep, L-Glutamine and the following cytokines: SCF  
838 (100 ng/ml), Flt3L (100 ng/ml), TPO (50 ng/ml) and IL7 (10 ng/ml). For experiments with the PAK4  
839 inhibitor PF-3758309 (solvent: DMSO), the SIRT1 inhibitor EX-527 (final concentration 25  $\mu\text{M}$ ; solvent:  
840 DMSO) or EdU (final concentration 10  $\mu\text{M}$ ) 1 volume of complete medium with 2x reagent or equivalent  
841 DMSO concentration as indicated was added to 1 volume of cell suspension. For PF-3758309 treated  
842 cells after 3 days a half medium exchange was performed with complete medium supplemented with 1x  
843 inhibitor concentration. For LDA experiments initial cell input was: 160-500 (low dose), 800-2500 (mid  
844 dose), 4000-12500 (high dose).

845

### 846 **Lentiviral Vector Production**

847 VSV-G pseudotyped lentiviral vector particles were produced, and titers were determined as previously  
848 described<sup>12</sup>. The negative control (CTRL) vector for overexpression encodes *gp91<sup>phox</sup>P415H* (catalytic

849 inactive gp91<sup>phox</sup>/*CYBB*) and for knockdown a shRNA directed against Renilla luciferase (shCTRL)<sup>12</sup>.

850 Additional shRNA were generated as described before<sup>12</sup>.

851

### 852 **Lentiviral Transduction**

853 If for a particular assay not stated otherwise, after 16-20 hours of pre-stimulation in low cytokine

854 condition medium cells were transduced with OE or KD and control vectors at matching multiplicity of

855 infection<sup>12</sup> aiming at mid-range (20-40%) transduction efficiencies but without lentiviral preparation

856 exceeding 20% of total culture volume. Transduction efficiency (%BFP<sup>+</sup>) was determined at day 3 and

857 day 6 post-transduction by flow cytometry on a BD Celesta and for xenotransplantation assays the mean

858 of both timepoints served as initial input estimate.

859

### 860 **Quantitative RT-PCR for *PAK4* and *INKA1***

861 RNA from FACS-sorted fractions was isolated with TRIzol (Invitrogen) or for low cell numbers (<20,000

862 cells) with PicoPure RNA Isolation Kit (Thermo Fisher) according to manufacturer's protocol.

863 Quantitative RT-PCR analysis was performed using the QuantiFast SYBR Green PCR Kit (Qiagen) on

864 the LightCycler 480 Instrument II (Roche). All signals were quantified using the  $\Delta$ Ct method and were

865 normalized to the levels of GAPDH.

866

### 867 **Methylcellulose Colony Forming Unit Assay**

868 LT-HSC and ST-HSC were sorted and cultured in low cytokine medium for 4 hours and then transduced

869 with CTRL or INKA1-OE lentiviral vectors. At day 3 post-transduction, 200 BFP<sup>+</sup> cells were sorted

870 directly into 2 mL methylcellulose (H4034, Stem Cell Technologies), supplemented with FLT3 Ligand

871 (10 ng/mL) and IL6 (10 ng/mL) and plated onto 2x35 mm dishes as duplicates. Colonies were allowed to

872 differentiate for 12-14 days and then morphologically assessed in a blind fashion by a second investigator.

873 Subsequently, colonies from replicate plates were pooled, resuspended in PBS/5% FBS, washed and 0.5%

874 of progeny from LT-HSC and 1% of progeny of ST-HSC was added to fresh methylcellulose as above,  
875 replated and scored after 12-13 days.

876

### 877 **Single Cell Stromal Assays**

878 Single cell *in vitro* assays were set up as with low passage murine MS-5 stroma cells<sup>47</sup> seeded at a density  
879 of 1500 cells per 96-well and grown for 2-4 days in H5100 media (StemCell Technologies, 05150). One-  
880 day prior to co-culture initiation, the H5100 media was removed and replaced with 100 µl erythro-  
881 myeloid differentiation media: StemPro-34 SFM media (Thermo Fisher) with the supplied supplement, 1x  
882 L-Glutamine, 1x Pen/Strep, 0.02% Human LDL and the following cytokines: FLT3L (20 ng/mL), GM-  
883 CSF (20 ng/mL), SCF (100 ng/mL), TPO (100 ng/mL), EPO (3 U/mL), IL-2 (10 ng/mL), IL-3  
884 (10 ng/mL), IL-6 (50 ng/mL), IL-7 (20 ng/mL), and IL-11 (50 ng/mL).

885

### 886 **Single cell stromal assays: *INKA1*-OE**

887 LT-, IT90+, IT90- and ST-HSC were sorted from CB Lin- and cultured in low cytokine medium for 4  
888 hours before transduction with CTRL or *INKA1*-OE lentiviral vectors. On day 3 post-transduction, cells  
889 were washed, resuspended in PBS, 2.5% FBS, 0.1 µg/L PI and PI<sup>BFP+</sup> single cells were deposited onto  
890 seeded MS5 stromal layers in 96 well plates using a BD FACSAria III. Single cells were co-cultured to  
891 proliferate and differentiate for 15 days (18 days post-transduction) with an addition of 100 µl erythro-  
892 myeloid differentiation media per well on day 8. Wells with hematopoietic cell content were marked 1  
893 day prior and the total number of wells with colonies was used to estimate single cell cloning efficiency.  
894 On the day of final analysis cultures were trypsinized, blocked with FBS, filtered (8027, Pall), stained and  
895 flow cytometric analyzed on a BD Celesta using the high-throughput sampler (HTS) unit. The following  
896 antibodies were used: V500-anti-CD45 (1/100), BV786-anti-CD33 (1/100), FITC-anti-CD56 (1/100), PE-  
897 anti-GlyA (1/100), PECy5-anti-CD14 (1/100), APC-anti-CD71 (1/100) and APC-Cy7-anti-CD41 (1/200).  
898 Erythroid cells were defined as CD45 negative and positive for CD71 and/or GlyA.

899

900 **CD112 Index-sorted Single Cell Stromal Assay**

901 On the day of the experiment, CB Lin<sup>-</sup> cells were stained as above to sort and deposit Sytox Blue<sup>-</sup>  
902 CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> single cells onto seeded MS5 stromal layers in 96 well plates while single cell  
903 fluorescence intensities for PECy5-anti-CD49f, PE-anti-CD90 and APC-anti-CD112 were indexed using  
904 a BD FACSAria III. Single cells were cultured for 18 days with an addition of 100 µl erythro-myeloid  
905 differentiation media per well at day 8. Wells with hematopoietic cell content were marked 1 day prior  
906 and the total number of wells with colonies was used to estimate single cell cloning efficiency. On day 18  
907 of co-culture cells hematopoietic cells were resuspended, filtered (8027, Pall), stained and flow  
908 cytometric analyzed on a BD Celesta using the HTS unit. The following antibodies were used: V500-anti-  
909 CD45(1/100), BV786-anti-CD33 (1/100), BV605-anti-CD56 (1/100), PE-anti-GlyA (1/100), PECy5-anti-  
910 CD11b (1/200), FITC-anti-CD71 (1/100), APC-Cy7-anti-CD41(1/200) and APC-anti-CD34(1/100).  
911 Clonal lineage output was scored and aligned with indexed single cell fluorescence intensity for CD112 at  
912 initial single cell deposition generating an .fcs file for visualization in FlowJo10.

913

914 **Xenotransplantation**

915 Aged and gender matched 8-16 wk old male and female recipient NOD.Cg-PrkdcscidII2rgtm1Wjl/SzJ  
916 (NSG) or NOD.Cg-PrkdcscidII2rgtm1WjlTg(CMVIL3,CSF2,KITLG)1Eav/MloySzJ (NSG-SGM3) mice  
917 were sublethally irradiated (224-225 cGy) 24 hours prior to intra-femoral injection. Cells were either  
918 injected directly after sorting, the progeny of 10<sup>4</sup> CD34<sup>+</sup>CD38<sup>-</sup> transduced cells one day after transduction  
919 or the progeny of initially plated CD34<sup>+</sup>CD38<sup>-</sup> cells treated in culture with PF-3758309 or DMSO. For  
920 analysis, mice were euthanized, injected femur and other long bones (non-injected femur, tibiae) were  
921 flushed separately in Iscove's modified Dulbecco's medium (IMDM) and 5% of cells were analyzed by  
922 flow cytometry (BD Celesta). For primary transplant analysis of initially *ex vivo* cultured cells, mice were  
923 euthanized at indicated timepoints (2, 4, 12, 20 weeks) after transplantation and human chimerism was  
924 assessed along with BFP and the following antibodies in addition to staining with Fixable Viability Dye  
925 eFluor 506 (1/500): PECy5-anti-CD45(1/100), APCCy7-anti-CD34(1/200), BV786-anti-CD33(1/100),



926 PE-anti-CD19(1/100), PE-anti-GlyA(1/100), FITC-anti-CD71(1/100). For experiments with sorted and  
927 injected LT-HSC CD112<sup>high</sup> and CD112<sup>low</sup> fractions the antibody panel was: V500-anti-CD45(1/100),  
928 FITC-anti-CD34(1/50), BV786-anti-CD33(1/100), PE-anti-CD19(1/100), PE-anti-GlyA(1/100),  
929 APCCy7-anti-CD41(1/200) and APC-anti-CD112(1/100);in addition to SYTOX Blue. Sick and mis-  
930 injected mice were excluded from analysis. For purification of human cells from xenotransplanted mice,  
931 fresh or thawed BM from individual or from pools of 2-5 mice were either directly analyzed by flow  
932 cytometry and/or flow sorted for serial transplantation, or after mouse cell depletion (Mouse Cell  
933 Depletion Kit, Miltenyi) or CD34<sup>+</sup> enrichment (human CD34 Microbead kit, Miltenyi) according to  
934 manufacturer's protocol. CD34<sup>+</sup> enriched cells were analyzed by flow cytometry with the following  
935 antibodies: FITC-anti-CD45RA (1/50), APC-anti-CD90 (1/100) or PE-anti-CD90(1/100), PECy5-anti-  
936 CD49f (1/50), PECy7-anti-CD38(1/200), APCCy7-anti-CD34(1/200), PE-anti-CD19(1/100) or BV711-  
937 anti-CD19(1/100), V500-anti-CD45(1/100) and if applicable with APC-anti-CD112(1/100). For  
938 secondary transplantation human double CD45<sup>+</sup> (APC-anti-CD45(1/100); FITC-anti-CD45(1/100); and if  
939 applicable BFP<sup>+</sup>) cells were sorted from each pooled sample and injected at indicated doses into irradiated  
940 (224-225 cGy) NSG-SGM3 mice and engraftment assessed at 8 weeks post-transplant. Injected femur and  
941 non-injected femurs were isolated, flushed and analyzed separately with the following reagents: PECy5-  
942 anti-CD45(1/100), APC-anti-CD45 (1/100), PE-anti-CD19(1/100), PE-anti-GlyA(1/100), BV786-anti-  
943 CD33(1/100), FITC-anti-CD3(1/100), APCCy7-anti-CD34 (1/200) and if applicable SYTOX Blue or  
944 BFP. A mouse was considered engrafted if human %CD45<sup>++</sup> > 0.1. For LDA experiments HSC frequency  
945 was estimated using the ELDA software (<http://bioinf.wehi.edu.au/software/elda/>).

946

#### 947 **5-Fluorouracil injections**

948 Male NSG mice xenotransplanted with transduced HSPC 4 weeks earlier (as above) were weighted and  
949 accordingly i.p. injected with a single dose of 150 mg/kg clinical grade 5-Fluorouracil (Hospira) or an  
950 equivalent volume of PBS. Sub-cohort were then analyzed 8 and 16 weeks later (corresponding to 12- and

951 20-weeks post-transplant, respectively) as described above. Mice in parallel xenotransplanted from the  
952 same transduced HSPC pool and euthanized at 4 weeks post-transplant served as baseline control.

953

#### 954 **Single cell RNAseq and processing**

955 CD34<sup>+</sup> enriched mPB samples from three healthy donors (age 58-63; 2x male, 1x female) were thawed  
956 and stained (V500-anti-CD45(1/100), FITC-anti-CD45RA(1/50), APC-anti-CD90(1/100), PECy5-anti-  
957 CD49f(1/50), PECy7-anti-CD38(1/100), APCCy7-anti-CD34(1/200), PI) to sort and deposit PI  
958 CD45<sup>+</sup>CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> single cells into 96 well plates while single cell fluorescence intensities for  
959 PECy5-anti-CD49f and PE-anti-CD90 were additionally indexed using a BD FACSAria III. Single cells  
960 were immediately lysed when sorted into wells with a mix of 5× Maxima Transcription Buffer, dNTPs,  
961 and NXGen RNase inhibitor (Lucigen), custom 3'end reverse transcriptase primers<sup>48</sup> and water (detailed  
962 protocol deposited: <https://dx.doi.org/10.17504/protocols.io.nkgdctw>). Plates were spun and immediately  
963 transferred onto dry ice and stored at -80 until further processing. A modified version of the SMARTseq2  
964 protocol<sup>49</sup> was performed with a custom modification in which a 12-base cell barcode was included in the  
965 3'-end reverse transcriptase primer. This enabled us to perform multiplexed pooling before library  
966 preparation with the Nextera XT DNA sample prep kit (Illumina).The resulting libraries of 1536 single  
967 cell and simultaneously processed 192 CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup> BM cells were sequenced in a total of 4  
968 lanes in Rapid Run mode on an Illumina HiSeq2500 instrument (Single index, paired end, 26bp read 1,  
969 32bp read 2, 8bp index read with custom barcodes). Reads were aligned with STAR 2.6.0c. Default  
970 parameters were used except for the following: sjdbOverhang 30. RNA velocity was run on the bam files  
971 using the run\_smartseq2 function in velocity 0.17.8 with default settings<sup>20</sup>. Cells with total read counts  
972 or total genes 4 median absolute deviations (MADs) below the median were removed, as were cells with a  
973 proportion of mitochondrial genes 4 MADs above the median. Denoising was performed using dca 0.2.2  
974 and batch correction was applied using an mnnorrect implementation within scanpy 1.3.3<sup>50,51</sup>. After  
975 preprocessing, mPB data was projected onto a diffusion map embedding and diffusion pseudotime<sup>21</sup> was  
976 calculated using default parameters. Imputed gene expression was subsequently compared against

977 pseudotime. To assign cells as LT-HSC, ST-HSC, IT90+ and IT90-, 40<sup>th</sup> percentile cutoffs were applied  
978 for both CD90 and CD49f surface intensities as measured through index sorting.

979

### 980 **Bulk RNA-seq processing and analysis**

981 CD112<sup>high</sup> and CD112<sup>low</sup>LT-HSC, total LT-HSC and ST-HSC were sorted from 3 independent CB pools  
982 and either directly resuspended and frozen (-80°C) in PicoPure RNA Isolation Kit Extraction Buffer or  
983 after transduction followed by a second sort (PI-BFP<sup>+</sup>) on day 6 post-transduction (average cell count:  
984 10<sup>4</sup>; range 0.25-3x10<sup>4</sup>). RNA was isolated using the PicoPure RNA Isolation Kit (Thermo Fisher)  
985 according to manufacturer's instructions. Samples that passed quality control according to integrity  
986 (RIN>7) and concentration as verified on a Bioanalyzer pico chip (Agilent Technologies) were subjected  
987 to further processing by the Center for Applied Genomics, Sick Kids Hospital: cDNA conversion was  
988 performed using SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara) and libraries were  
989 prepared using Nextera XT DNA Library Preparation Kit (Illumina). Equimolar quantities of libraries  
990 were pooled and sequenced with 3-4 cDNA libraries per lane on a High Throughput Run Mode Flowcell  
991 with v4 sequencing chemistry on the Illumina HiSeq 2500 following manufacturer's protocol generating  
992 paired-end reads of 125-bp in length to reach depth of 60-80 million reads per sample. Reads were then  
993 aligned with STAR v2.5.2b against hg38 and annotated with ensembl v90. Default parameters were used  
994 except for the following: chimSegmentMin 12; chimJunctionOverhangMin 12; alignSJDBoverhangMin  
995 10; alignMatesGapMax 100000; alignIntronMax 100000; chimSegmentReadGapMax parameter 3;  
996 alignSJstitchMismatchNmax 5 -1 5 5. Read counts were generated using HTSeq v0.7.2 and general  
997 statistics were obtained from picard v2.6.0. Differential gene expression was performed using DESeq2  
998 v1.22.2 following recommended practices. Genes were ranked using the formula  $-\text{sign}(\log\text{FC}) * -$   
999  $\log_{10}(\text{pvalue})$ . The rank files from each comparison were used in pathway analysis (GSEA) using 2000  
1000 permutations and default parameters-if not specified otherwise, against pathway collections from  
1001 databases (MsigDB; HALLMARK, GOBP from <http://baderlab.org/GeneSets>, version February 2020) or  
1002 retrieved from the literature (HSC\_120\_COMMON<sup>52</sup>). The gene set aHSC/MPP\_UP\_(vs\_dHSC) was

1003 derived from published murine RNAseq data<sup>7</sup> combining genes with logFC > 1 and FDR < 0.01 in either  
1004 activated HSC (aHSC) or MPP versus dormant HSC (dHSC). For the *INKA1*-OE datasets pathways were  
1005 excluded from analysis if gene set enrichment was driven by a single gene. EnrichmentMap version 3.1.0  
1006 in Cytoscape was used to visualize enriched pathway gene sets at FDR < 0.1 with an overlap coefficient  
1007 set to 0.5 .

1008

### 1009 **Proximity-dependent Biotin Identification (BioID) Assay**

1010 Coding sequences for *INKA1* and *PAK4* were PCR amplified from pENTR223\_C3orf54\_STOP of  
1011 pDONR221\_PAK4 (primers: ASCI-INKA1\_fw, Not1\_INKA1\_rev, ASCI-PAK4\_fw, Not1\_PAK4\_rev),  
1012 respectively. Thereby, AscI and NotI restriction sites were introduced and used to clone both genes into  
1013 pcDNA5/FRT/TO-FLAG-BirA\*-[MCS] and pcDNA5/FRT/TO-[MCS]-FLAG-BirA\* to generate N- and  
1014 C-terminal BirA\* fusion constructs. Transfection, induction, mass spectrometric analysis and data  
1015 processing for the BioID assay was performed as described previously in HEK293<sup>24</sup>. Briefly, high  
1016 confidence interactors were identified fulfilling these criteria: iProphet>0.9, at least two unique peptides  
1017 in each of 4 runs, and a Bayesian FDR< 1% comparing the two highest peptide counts of each given ID  
1018 between 20 controls (FlagBirA\* only construct) and 4 bait protein sample analysis. Data was loaded into  
1019 Cytoscape for network visualization and in Cytoscape STRING enrichment (GO processes) was used for  
1020 pathway analysis to subsequently generate the enrichment map of selected pathways with FDR<0.05.

1021

### 1022 **ATAC-seq processing and analysis**

1023 Library preparation for ATAC-Seq was performed on 2000 cells with Nextera DNA Sample Preparation  
1024 kit (Illumina), according to a previously reported protocol<sup>53</sup>. 4 ATAC-seq libraries were sequenced per  
1025 lane in HiSeq 2500 System (Illumina) to generate paired-end 50-bp reads. Reads were mapped against the  
1026 hg19 human reference genome using BWA with default parameters. All duplicate reads, and reads  
1027 mapped to mitochondria, chrY, an ENCODE blacklisted region or an unspecified contig were removed  
1028 (ENCODE Project Consortium 2012). MACS 2.0.10 was used to call peaks with a uniform extension size

1029 of 147 bp. A set of 500 bp windows over hg19, overlapping by 250bp was generated and for each sample  
1030 the windows overlapping a called peak were identified. Windows were considered to be present in each  
1031 condition if they overlapped a peak in 2 replicates. Windows only considered present in one condition  
1032 were considered specific to that condition. Motif enrichment was performed using HOMER against a  
1033 catalogue of all peaks called in any hematopoietic population was produced by merging all called peaks  
1034 which overlapped by at least one base pair using bedtools. As described in *Takayama et al.*<sup>26</sup>, cells from  
1035 three sorted CD34+CD38-CD45RA- and three CD34+CD38+ populations were processed on the 10X  
1036 Genomics single cell ATAC-seq platform, and 12,414 single cells were retained based on default  
1037 cellranger QC criterion and chromVAR depth filtering. Cellranger-reanalyze (1.1, 10x Genomics) was  
1038 subsequently used to map reads over the sites identified in the bulk ATAC-Seq catalogue, and read counts  
1039 were binarized. chromVAR (with default settings) was used to calculate the enrichment of each of the  
1040 hematopoietic signatures identified in this paper in each single cell, as well as for the windows considered  
1041 specific to the CD112<sup>low</sup> or CD112<sup>high</sup> LT-HSC subsets.

1042

### 1043 **Statistical Analysis**

1044 GraphPad Prism or MS Excel were used for all statistical analyses except RNA-seq, which was analyzed  
1045 with R and Python. Unless otherwise indicated in Fig. legend, mean  $\pm$  SD values are reported in the  
1046 graphs and minimal to maximal values are represented in box-and-whisker plots. Statistical significance  
1047 (\* $p < 0.05$ , \*\* $p < 0.01$  and \*\*\* $p < 0.001$ , n.s. represents non-significance) was determined if not  
1048 otherwise indicated in Figure legends with two-sided Student's t-tests with or without Welch's correction  
1049 as applicable. If normal-distribution assumptions were not valid, statistical significance was evaluated  
1050 using the Mann-Whitney test (two-tailed) for single comparisons or Kruskal-Wallis test for multiple  
1051 comparisons. A normal distribution of the data was tested using the Kolmogorov-Smirnov test if the  
1052 sample size allowed. One-way ANOVA was used to compare means among three or more independent

1053 groups. Applicable post hoc tests were used to compare all pairs of treatment groups when the overall P  
1054 value was <0.05.

1055

#### 1056 **Data availability**

1057 The accession number for RNA-seq data generated for this manuscript is GEO: GSE148885, and GEO:  
1058 GSE125345 for previously published RNAseq data used in this study. Raw RNA-seq and ATACseq data  
1059 is available on EGA: EGAD00001006541 (Bulk RNA-Seq: EGAS00001004768; scRNAseq:  
1060 EGAS00001004769; ATACseq: EGAS00001005121). For GSEA GOBP were retrieved from  
1061 <http://baderlab.org/GeneSets> (version February 2020), the Hallmark gene set v.7.0 from  
1062 [ftp.broadinstitute.org://pub/gsea/gene\\_sets/h.all.v7.0.symbols.gmt](ftp.broadinstitute.org://pub/gsea/gene_sets/h.all.v7.0.symbols.gmt) and Eppert HSC-R from  
1063 [http://www.gsea-msigdb.org/gsea/msigdb/cards/EPPERT\\_HSC\\_R.html](http://www.gsea-msigdb.org/gsea/msigdb/cards/EPPERT_HSC_R.html).

1064

#### 1065 **Code availability**

1066 Code generated during this study is available at GitHub [https://github.com/andygxzeng/HSC\\_Pseudotime](https://github.com/andygxzeng/HSC_Pseudotime).

1067

#### 1068 **Methods-only References**

1069 45 Wooten, M. *et al.* Superresolution imaging of chromatin fibers to visualize epigenetic information on  
1070 replicative DNA. *Nat Protoc* **15**, 1188-1208, doi:10.1038/s41596-019-0283-y (2020).

1071 46 Sullivan, B. A. Optical mapping of protein-DNA complexes on chromatin fibers. *Methods Mol Biol* **659**, 99-  
1072 115, doi:10.1007/978-1-60761-789-1\_7 (2010).

1073 47 Itoh, K. *et al.* Reproducible establishment of hemopoietic supportive stromal cell lines from murine bone  
1074 marrow. *Exp Hematol* **17**, 145-153 (1989).

1075 48 Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene  
1076 expression data. *Nat Biotechnol* **33**, 495-502, doi:10.1038/nbt.3192 (2015).

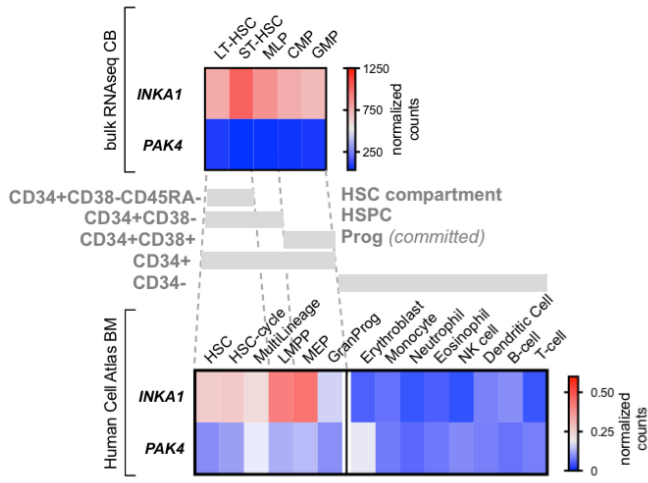
1077 49 Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171-181,  
1078 doi:10.1038/nprot.2014.006 (2014).

1079 50 Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing  
1080 data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* **36**, 421-427, doi:10.1038/nbt.4091  
1081 (2018).

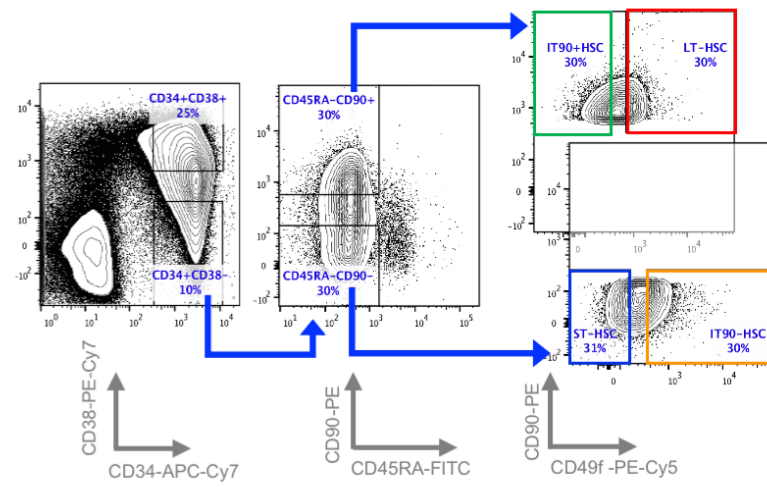
- 1082 51 Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S. & Theis, F. J. Single-cell RNA-seq denoising using a  
1083 deep count autoencoder. *Nat Commun* **10**, 390, doi:10.1038/s41467-018-07931-2 (2019).
- 1084 52 Fares, I. *et al.* EPCR expression marks UM171-expanded CD34(+) cord blood stem cells. *Blood* **129**, 3344-  
1085 3351, doi:10.1182/blood-2016-11-750729 (2017).
- 1086 53 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native  
1087 chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and  
1088 nucleosome position. *Nat Methods* **10**, 1213-1218, doi:10.1038/nmeth.2688 (2013).

**Fig. 1**

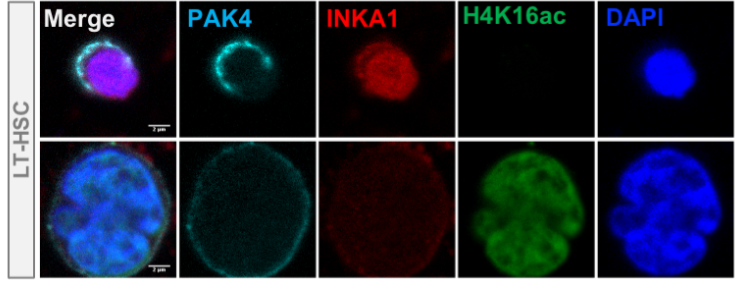
**a**



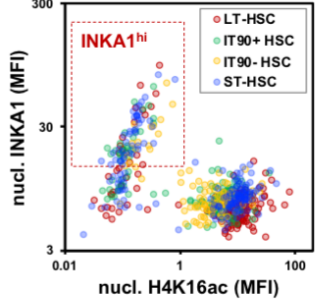
**b**



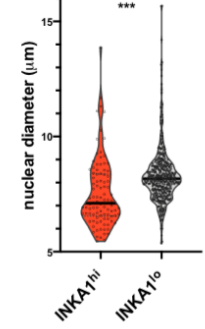
**c**



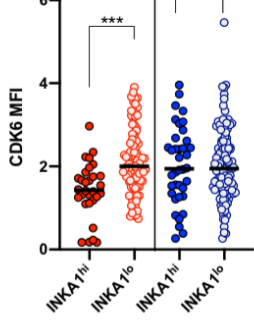
**d**



**e**

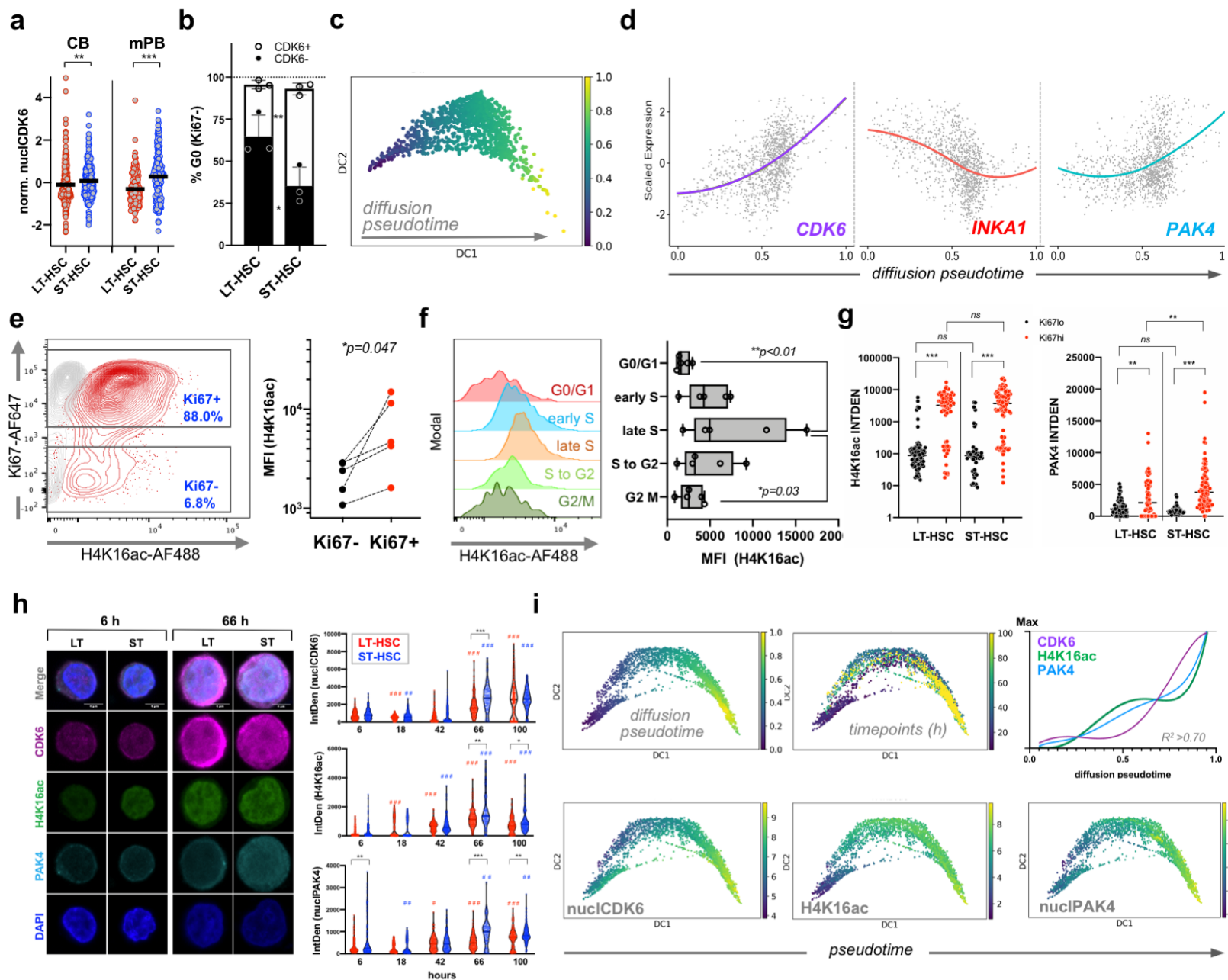


**f**

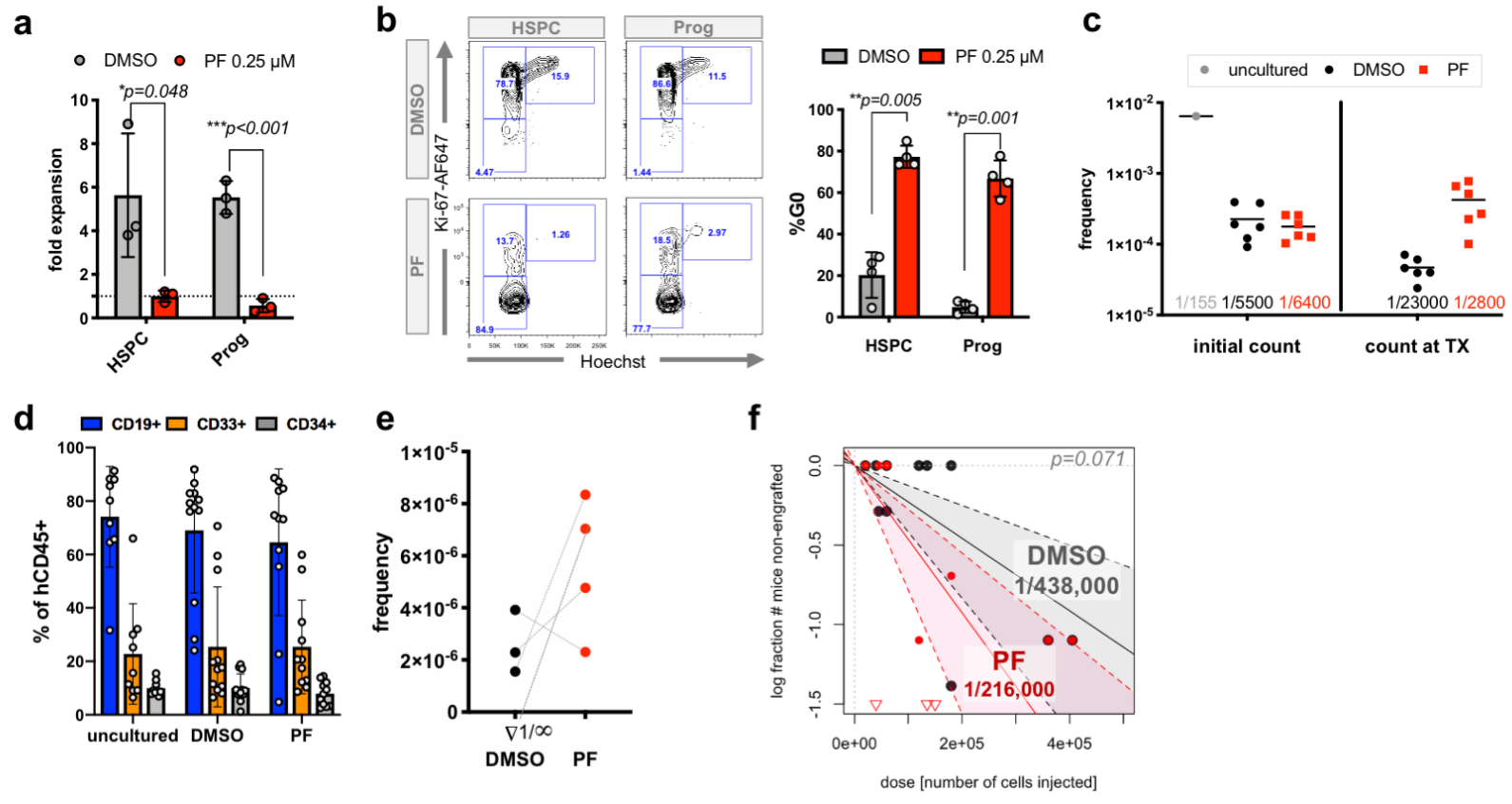




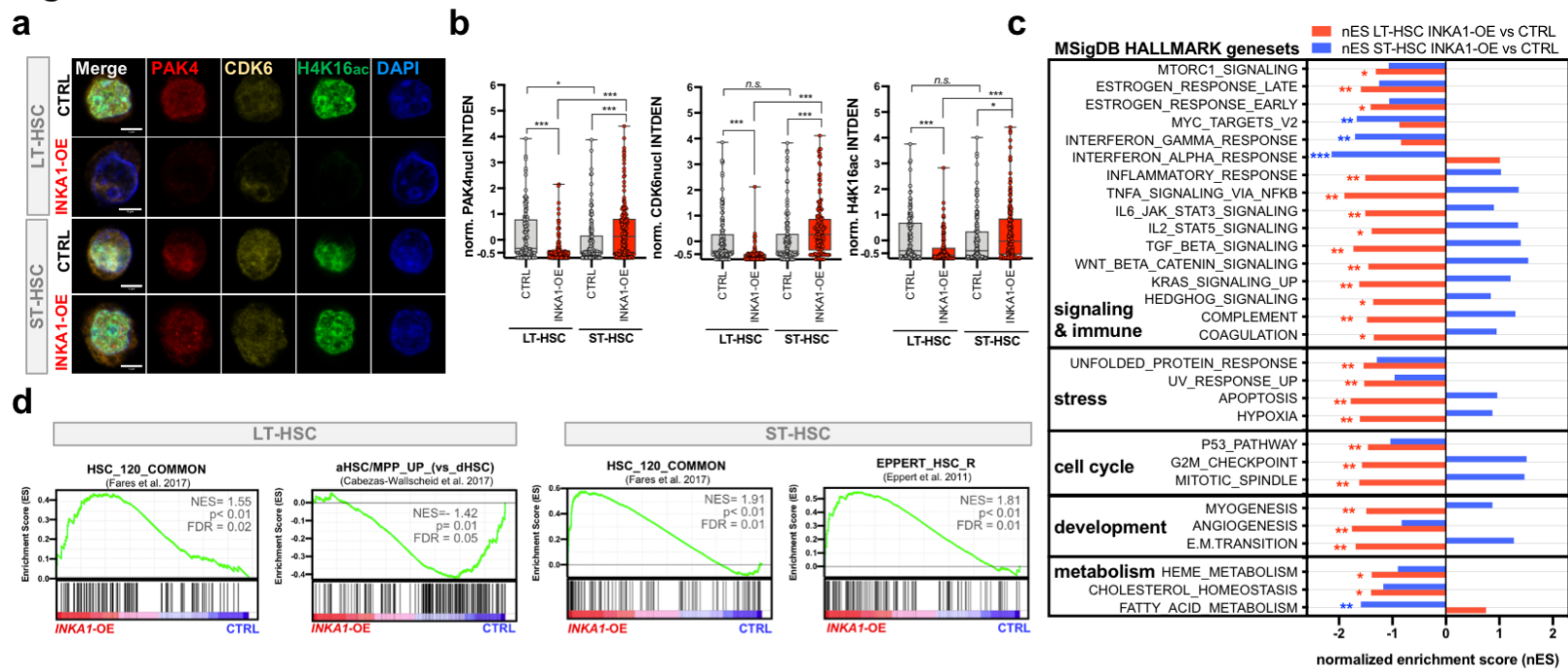
**Fig. 2**



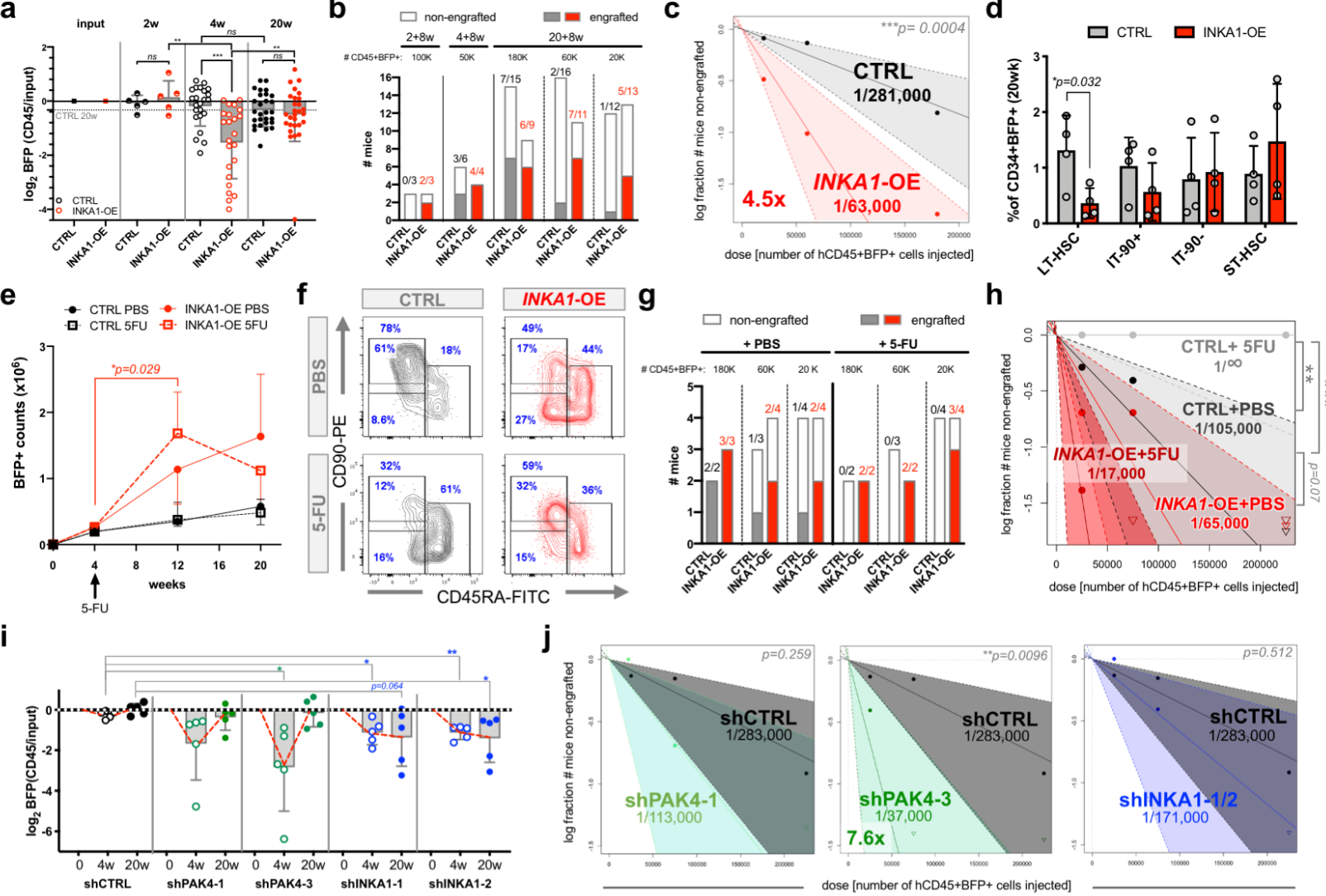
**Fig. 3**



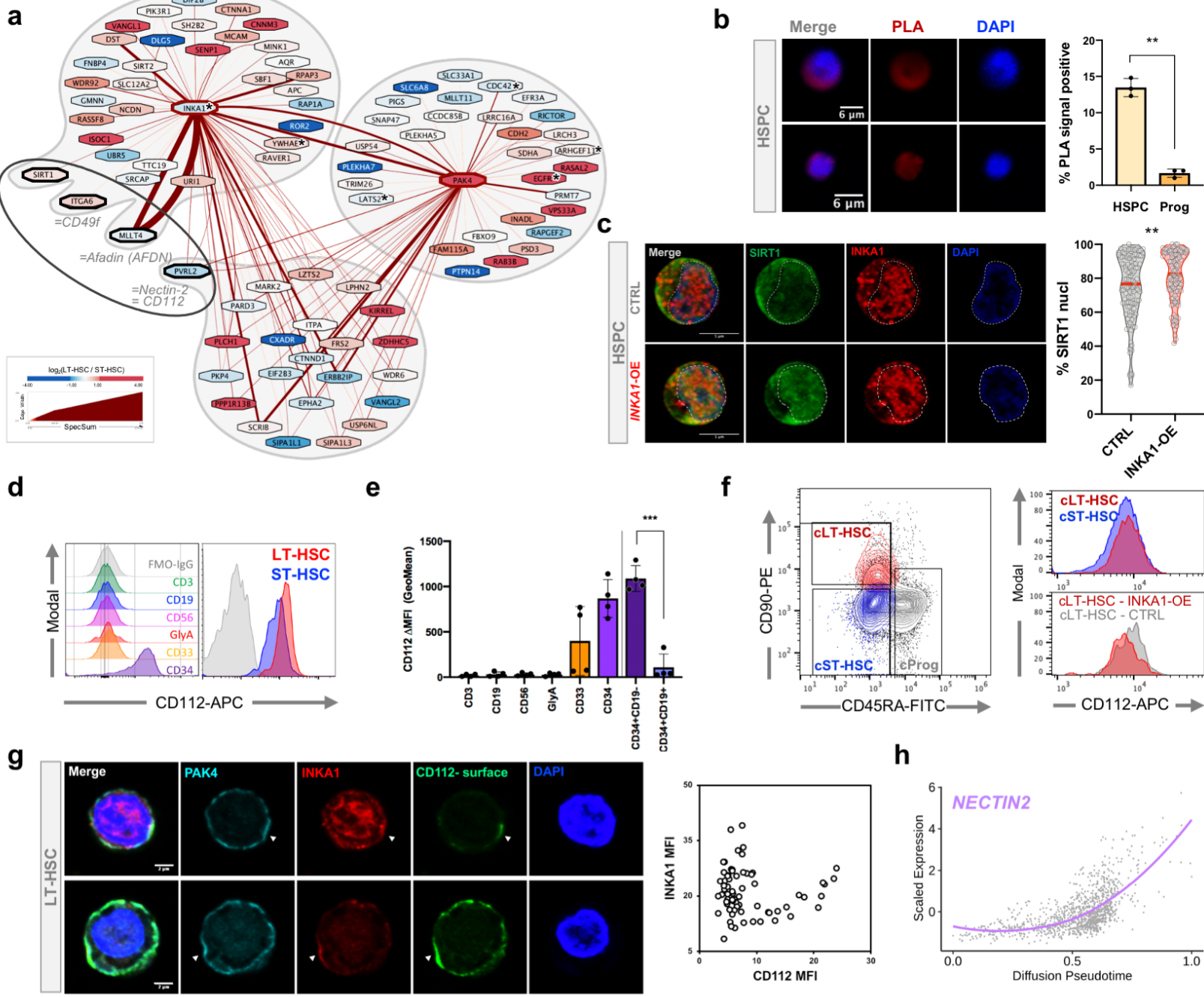
**Fig. 4**



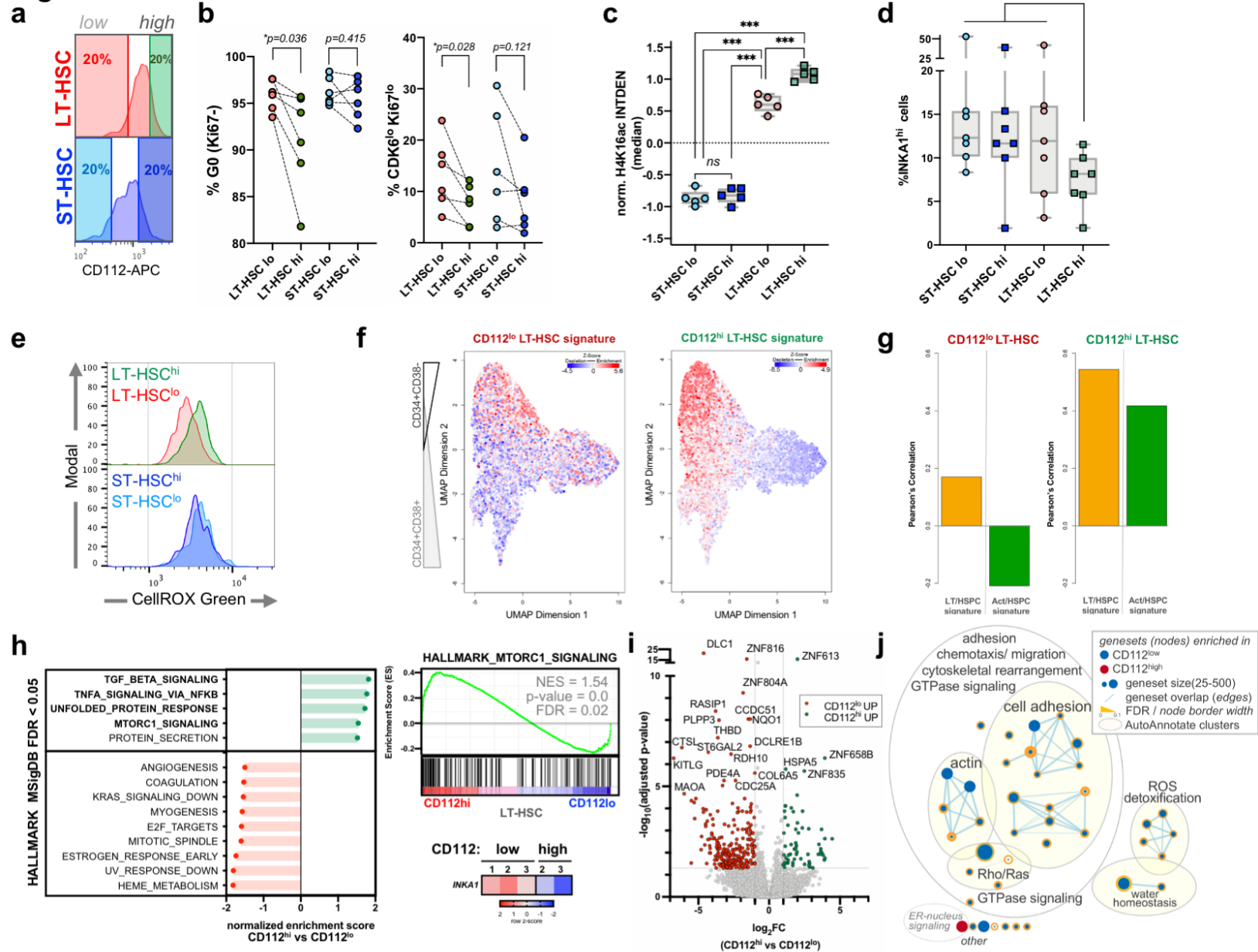
**Fig. 5**



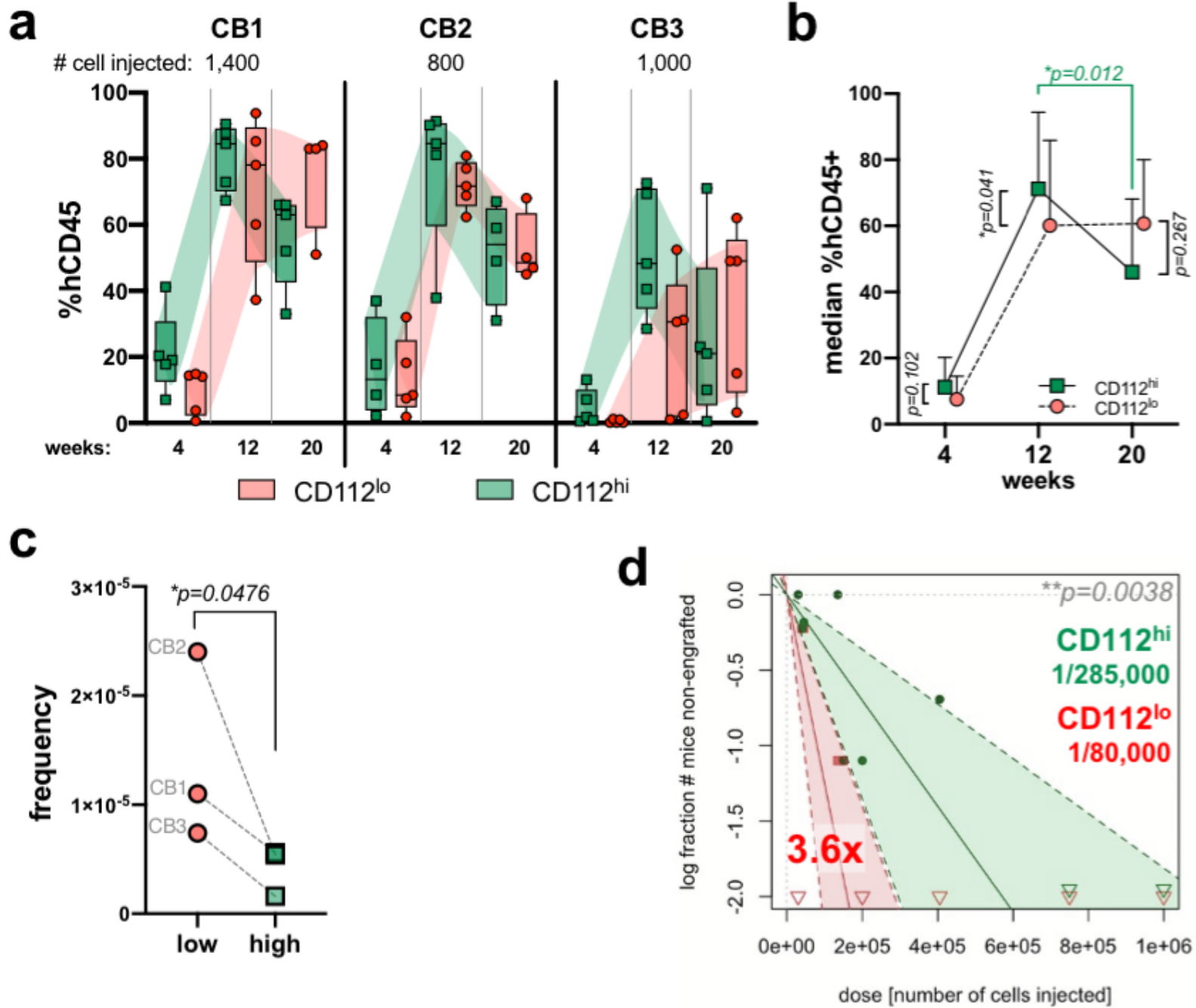
**Fig. 6**



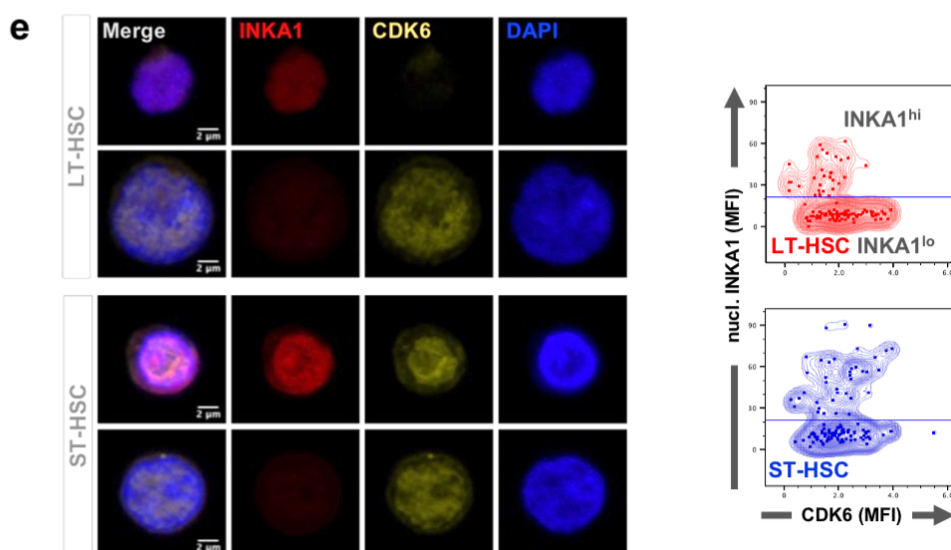
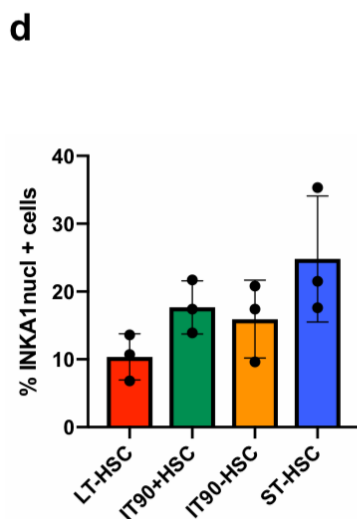
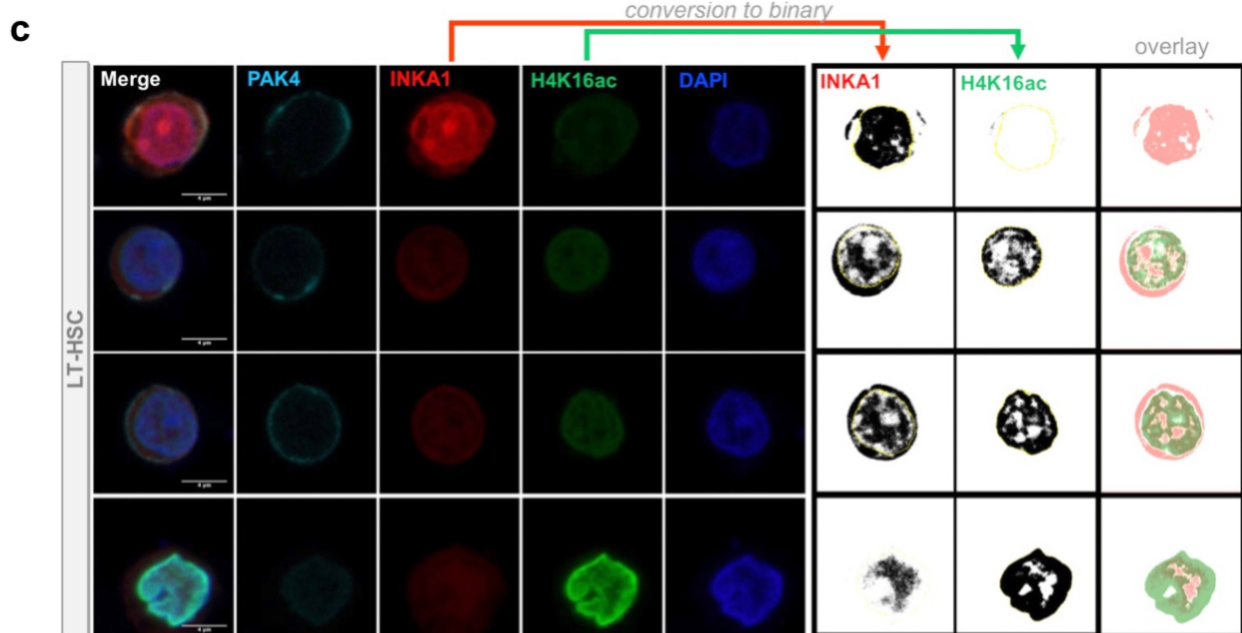
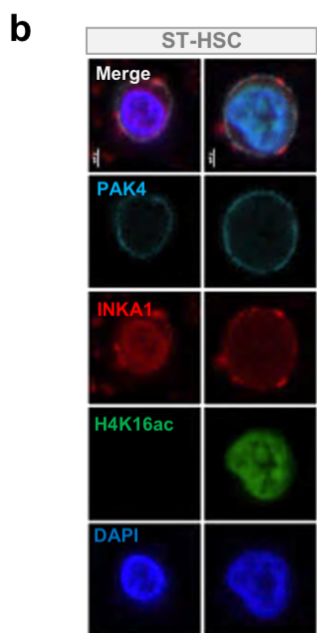
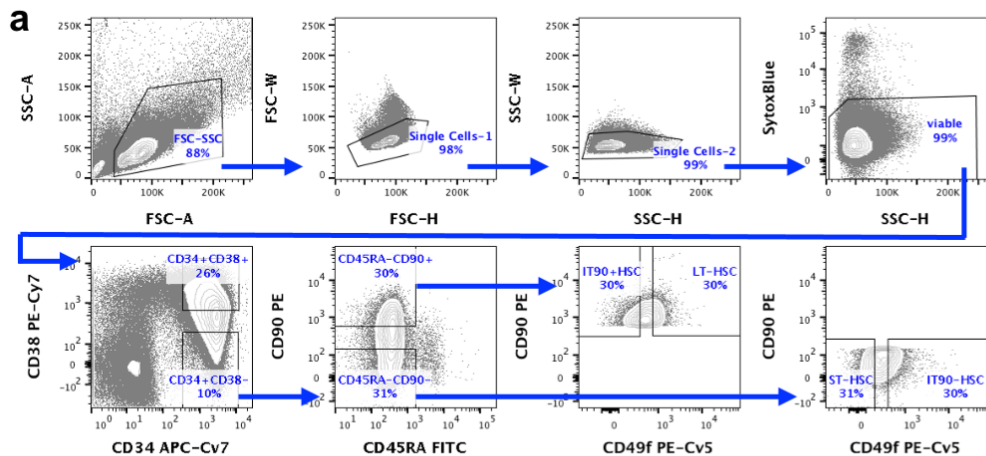
**Fig. 7**



# Fig. 8

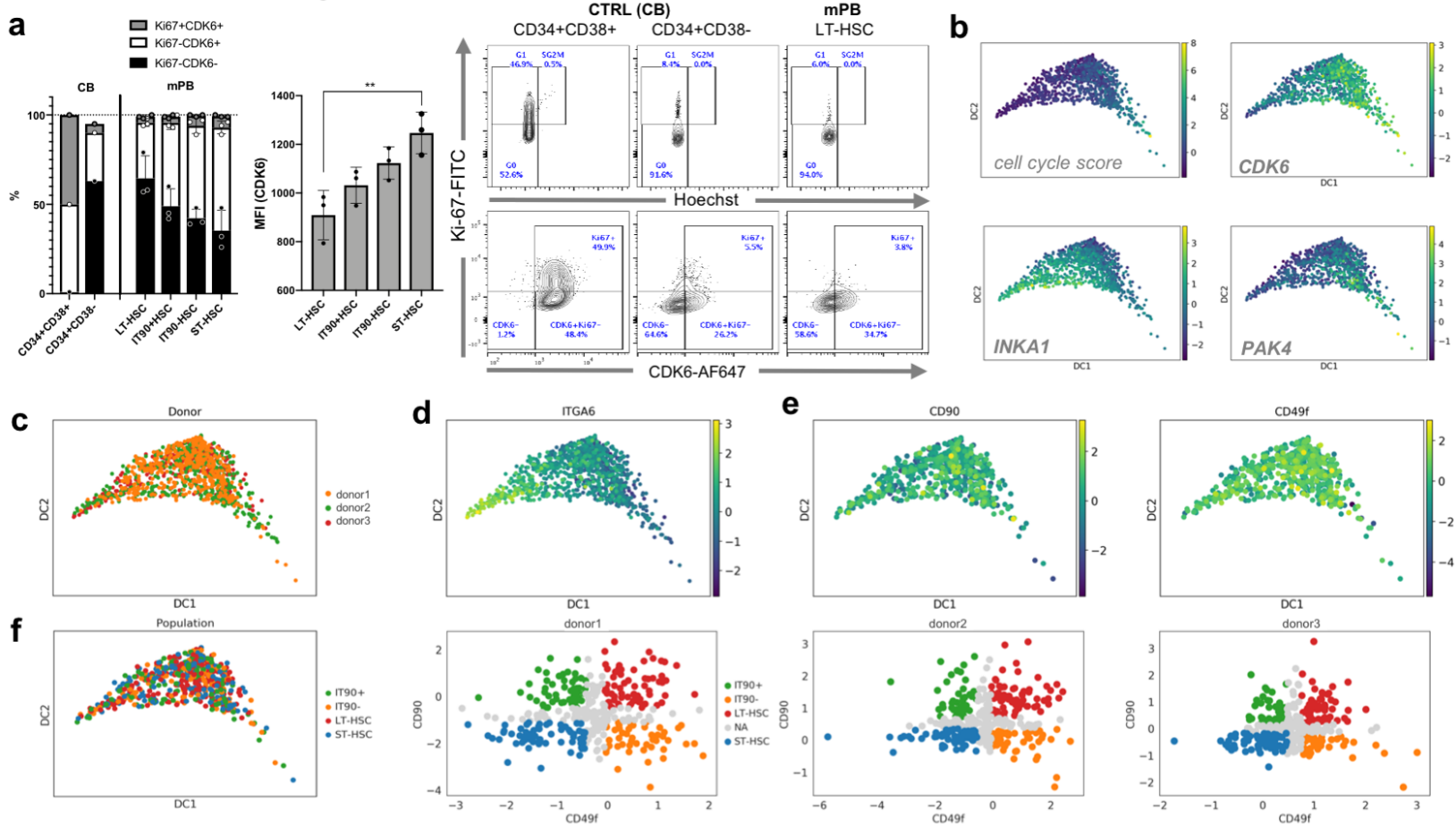


# Extended Data Fig.1

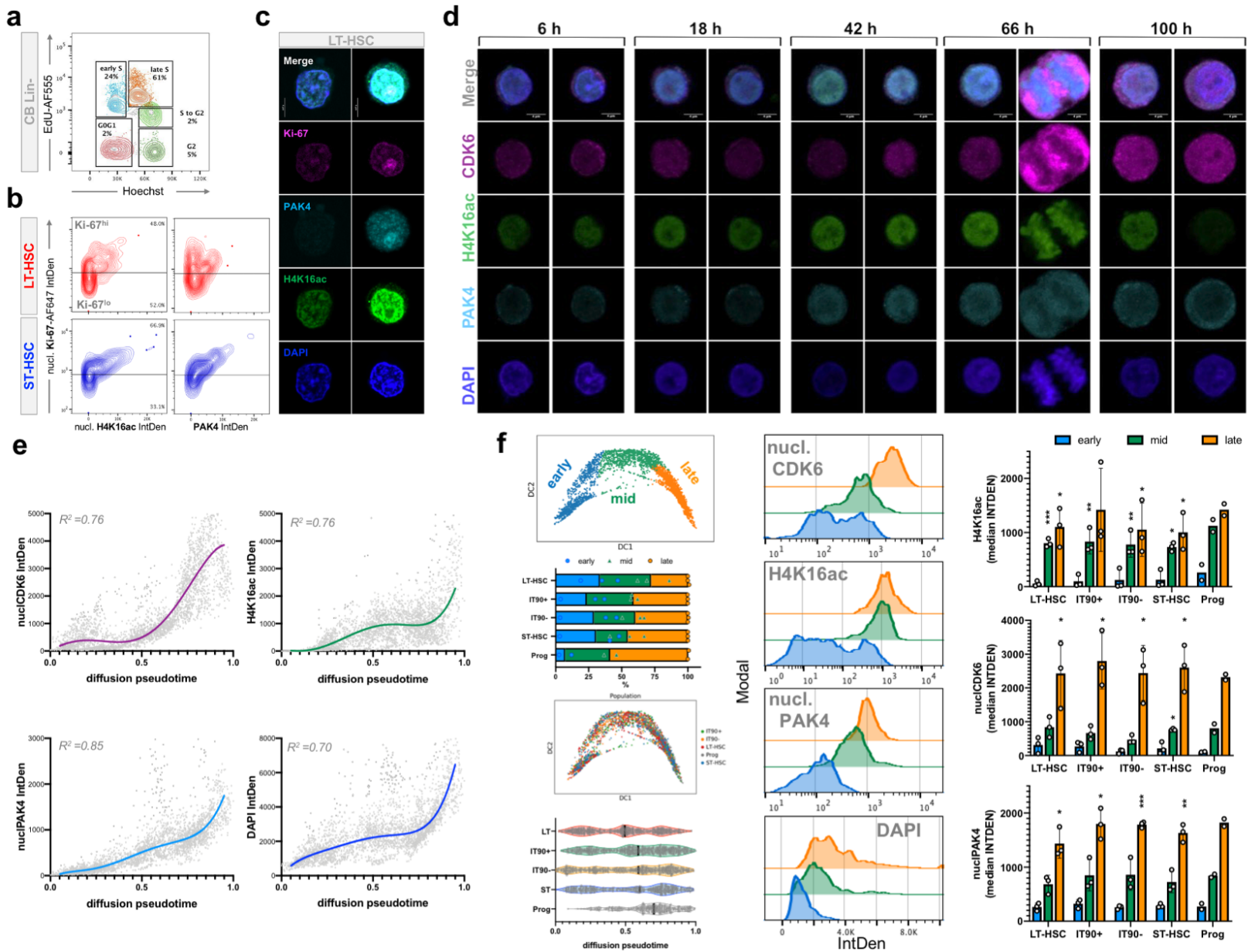




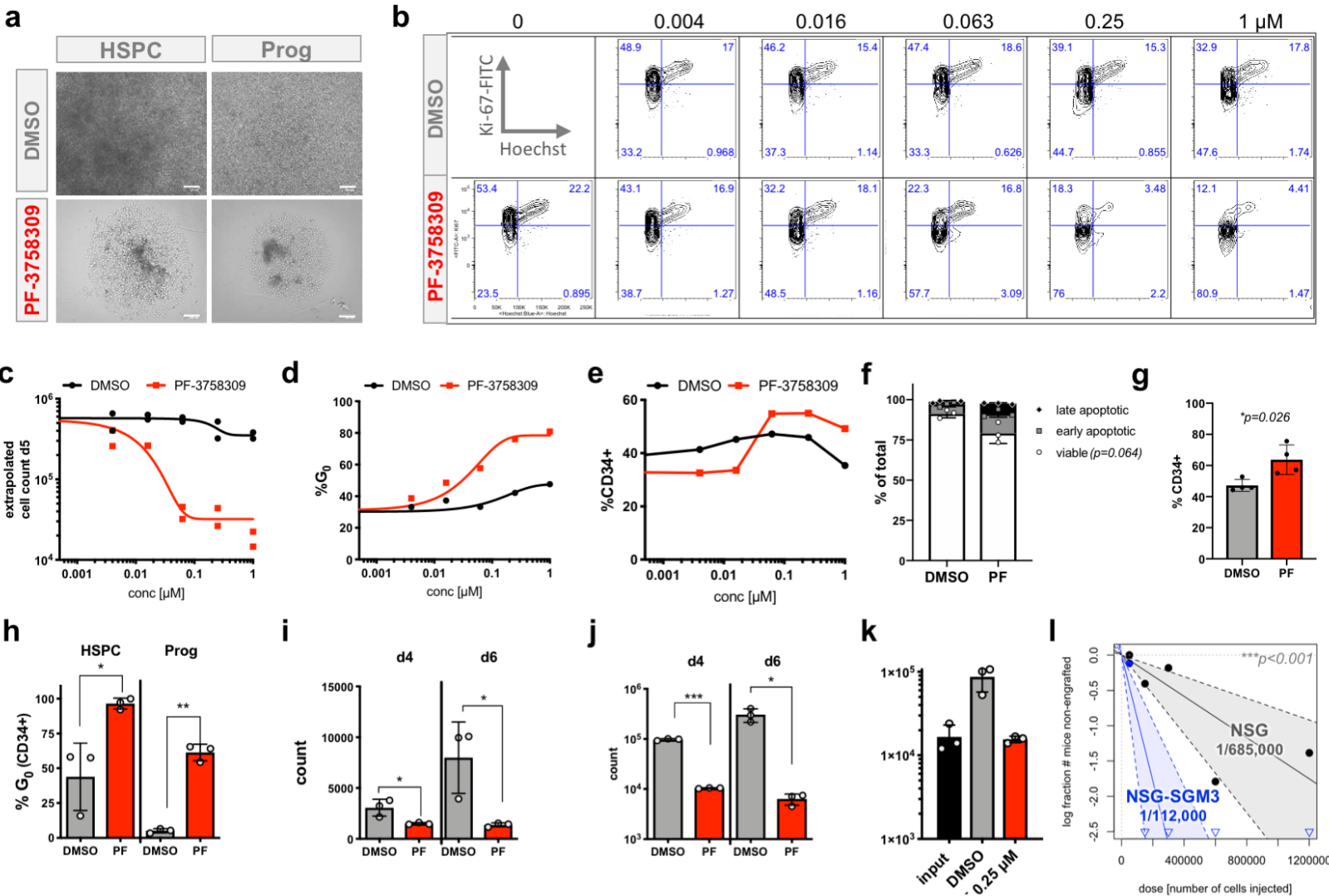
## Extended Data Fig. 2



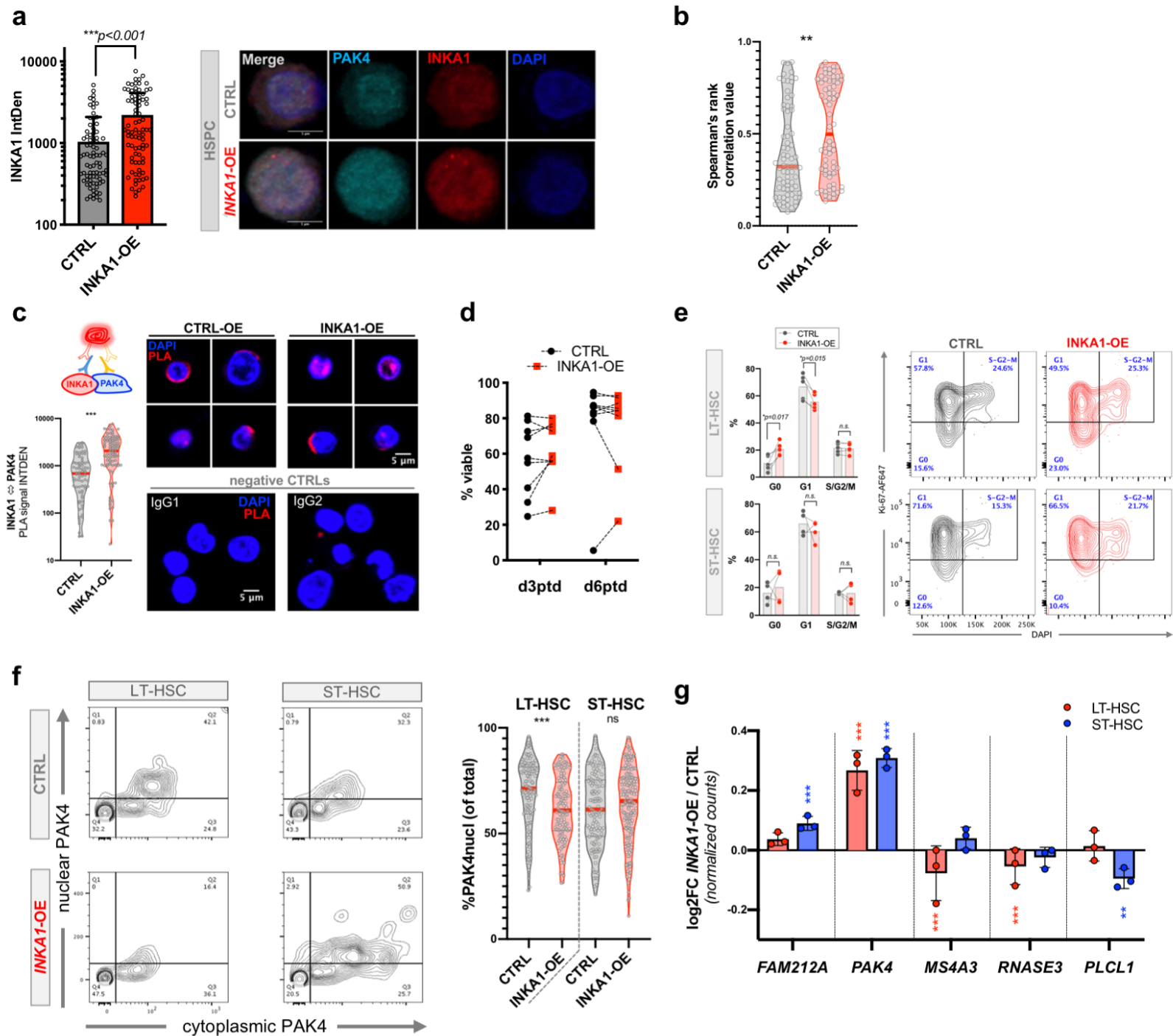
# Extended Data Fig. 3



# Extended Data Fig. 4

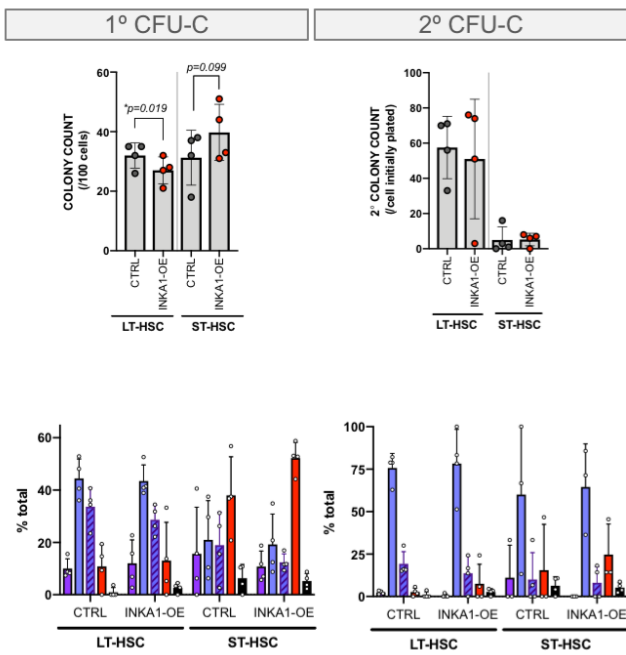


# Extended Data Fig. 5

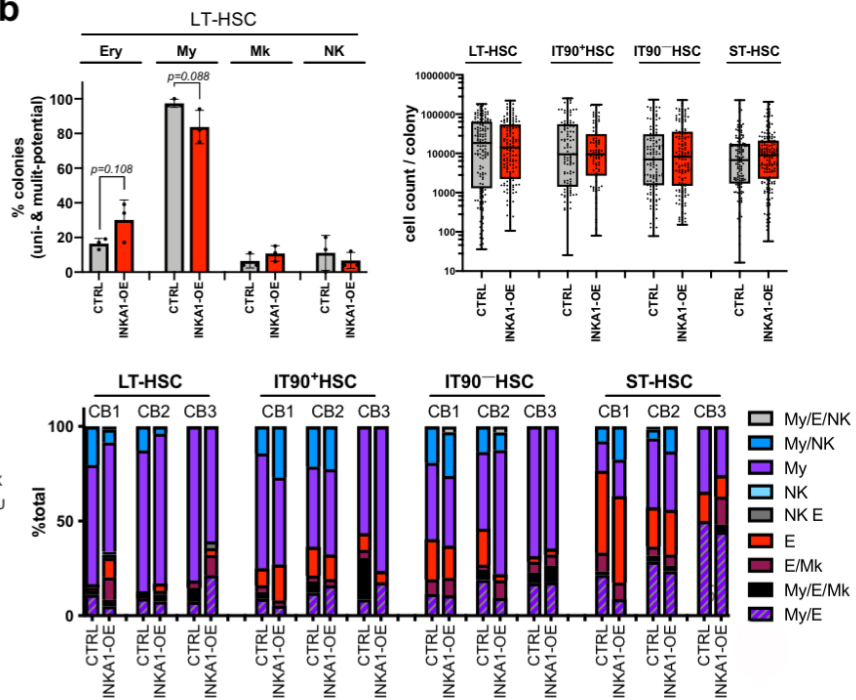


# Extended Data Fig. 6

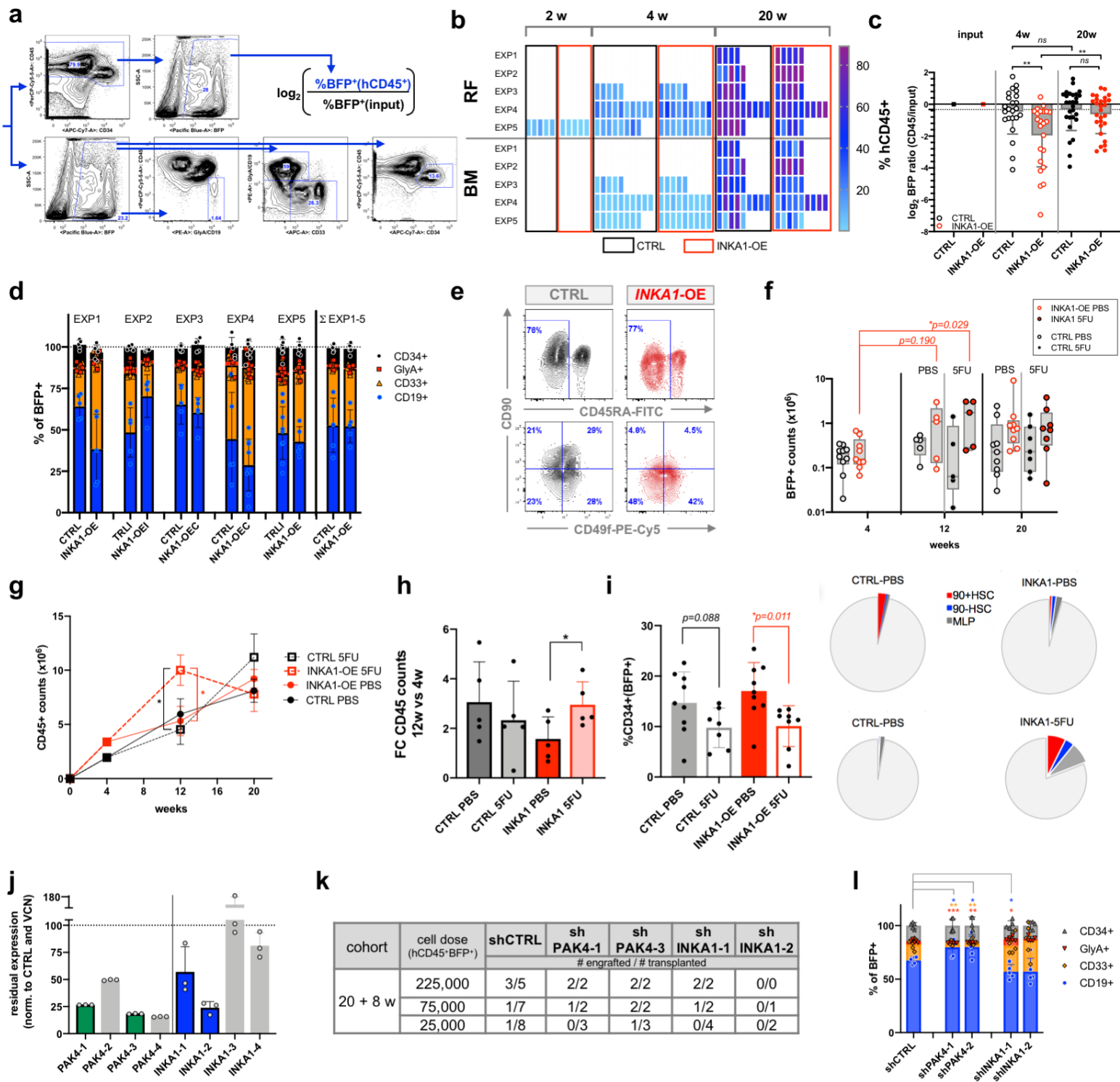
**a**



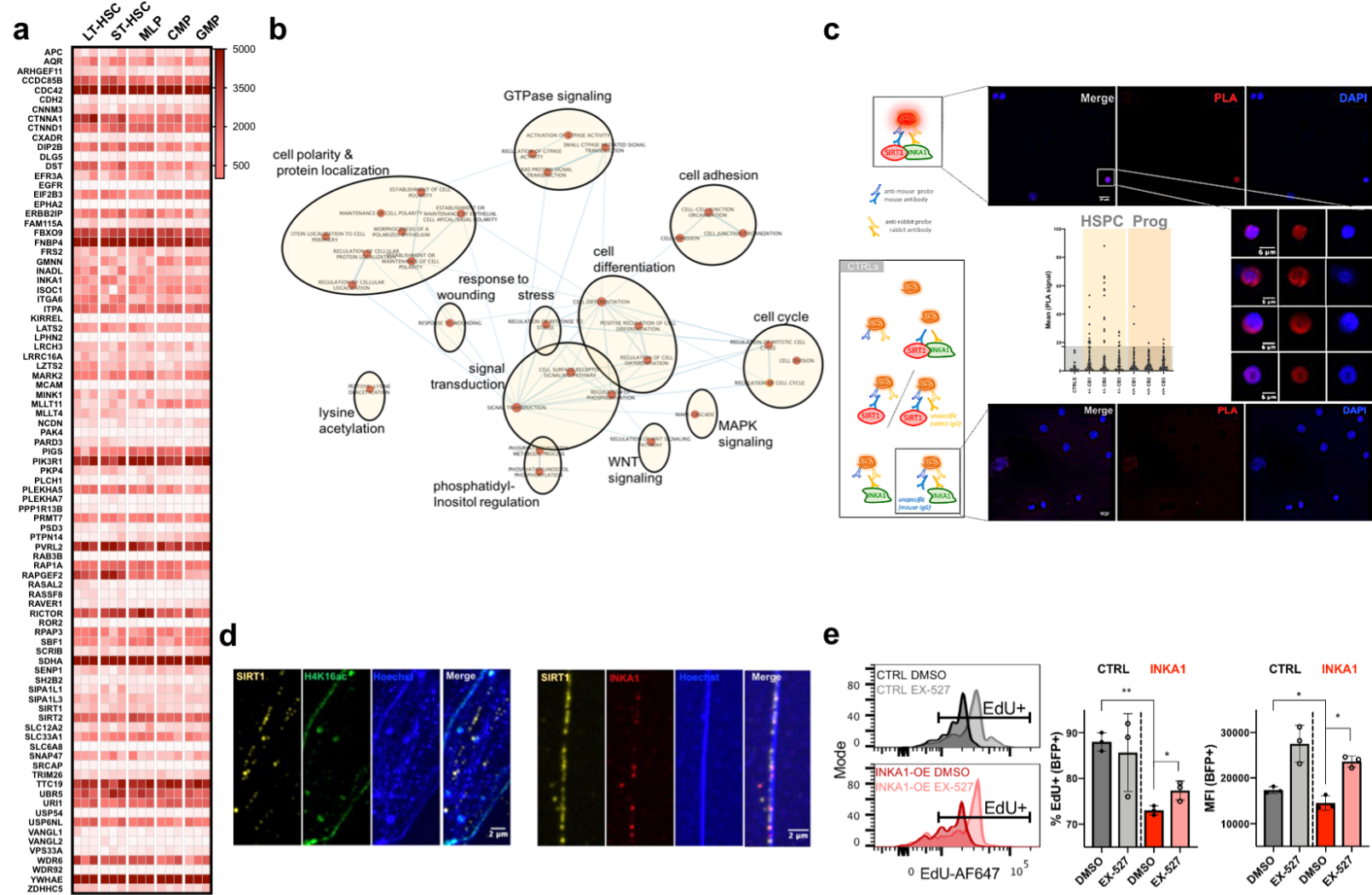
**b**



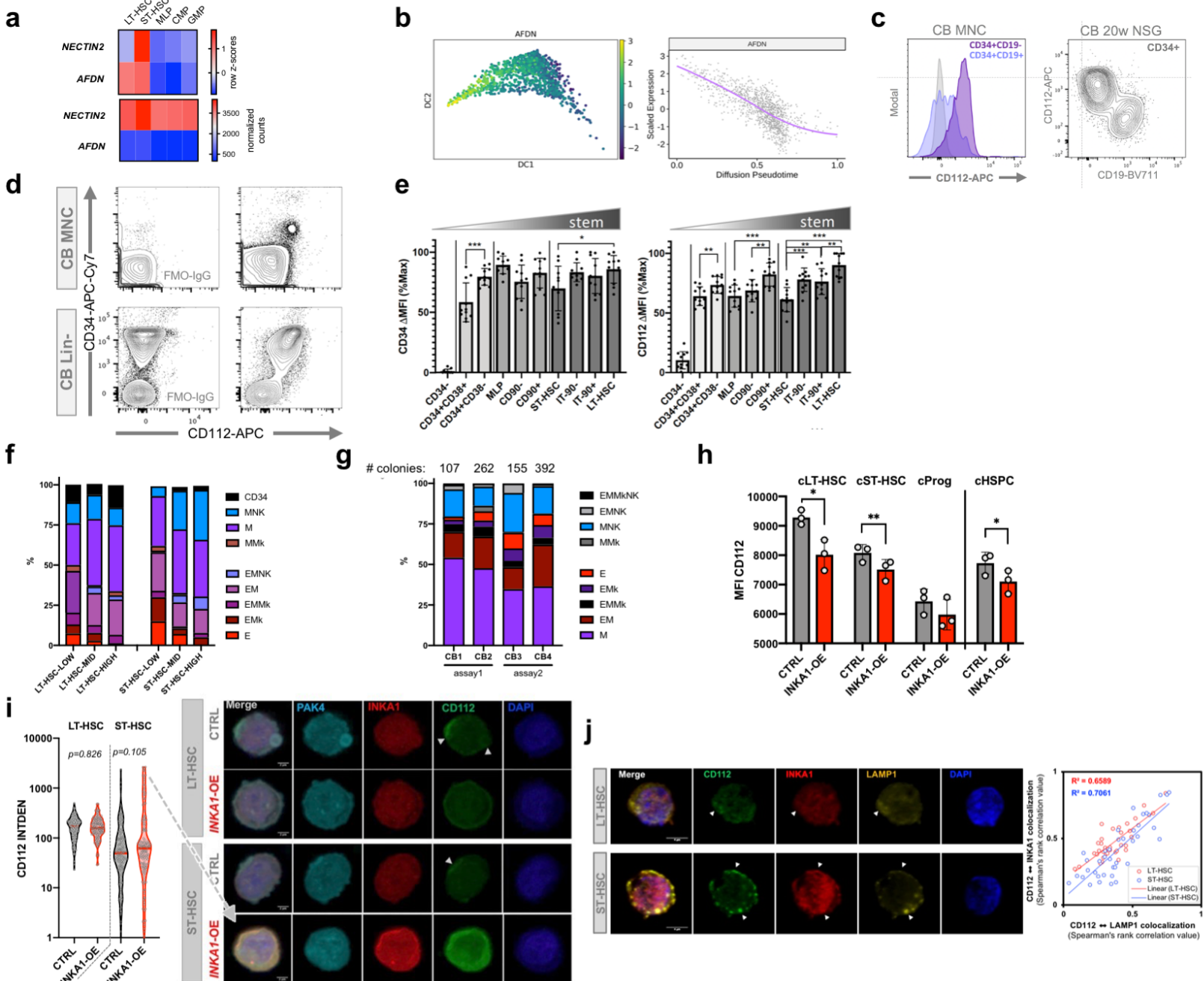
# Extended Data Fig. 7



# Extended Data Fig. 8



# Extended Data Fig. 9





# Extended Data Fig. 10

