Low Kit expression identifies Hematopoietic Stem Cell subsets with enhanced lymphoid potential in mice and humans

Harold K. Elias[1,2], Sneha Mitra[3], Marina B. da Silva[1,14], Adhithi Rajagopalan[1,14], Brianna Gipson[1], Nicole Lee[1], Anastasia I. Kousa[1], Mohamed A.E. Ali[4], Simon Grassman[1], Xiaoqun Zhang[1], Susan DeWolf [6], Melody Smith[5], Hana Andrlova[1], Kimon V. Argyropoulos[1], Roshan Sharma[7], Teng Fei[8], Joseph C Sun[1], Cynthia E. Dunbar[9], Christopher Y Park[4], Christina S. Leslie[3], Avinash Bhandoola[10], and Marcel R.M. van den Brink[1,11,12,13]

Affiliations

[1]Department of Immunology, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

[2]National Institutes of Health (NIH), Bethesda, MD, 20892, USA, USA

[3]Computational and Systems Biology, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

[4]Department of Pathology, New York University Grossman School of Medicine, New York, NY, 10016, USA

[5]Division of Blood and Marrow Transplantation and Cellular Therapy, Department of Medicine, Stanford University School of Medicine, Stanford, CA, 94305

[6]Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

[7]Single-cell Analytics and Innovation Lab, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

[8]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

[9]Translational Stem Cell Biology Branch, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD, 20892, USA

[10]Laboratory of Genome Integrity, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD, 20892, USA

[11]Department of Immunology and Microbial Pathogenesis, Weill Cornell Medical College; New York, NY

[12]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA.

[13]Lead Contact

[14]These authors contributed equally

Correspondence: eliash@mskcc.org (HKE); vandenbm@mskcc.org (MRMvdB)

ABSTRACT

Hematopoietic stem cells (HSCs) with multilineage potential are critical for effective T cell reconstitution and restoration of the adaptive immune system after allogeneic Hematopoietic Cell Transplantation (allo-HCT). Our current understanding of lymphoid-poised HSC subsets is limited due to a lack of reliable markers. Through comprehensive molecular profiling of HSCs, we identify Kit^{lo} HSCs- a subset with enhanced lymphoid potential, which declines with age. Using a preclinical allo-HCT model, we demonstrate that Kit^{lo} HSCs exhibit superior lymphoid progenitor output, enhanced thymic recovery, and T-cell reconstitution resulting in improved T cell responses to infection post-HCT. Furthermore, Kit^{lo} HSCs with augmented BM lymphopoiesis mitigate age-associated thymic alterations, thus enhancing T-cell recovery in middle-aged hosts. Chromatin profiling revealed that Kit^{lo} HSCs exhibit higher activity of lymphoid-specifying transcription factors (TFs), including *Zbtb1*. Deletion of *Zbtb1* in Kit^{lo} HSCs, diminished their T-cell potential, while reinstating *Zbtb1* in megakaryocytic-biased Kit^{hi} HSCs rescued T-cell potential, *in vitro* and *in vivo*. Finally, we discover an analogous Kit^{lo} HSC subset with enhanced lymphoid potential in human bone marrow. Our results demonstrate that Kit^{lo} HSCs identify HSC subsets with enhanced lymphoid potential mediated by an underlying epigenetic program.

KEYWORDS: Hematopoietic stem cells; lymphoid differentiation; epigenetic priming; thymus; immune reconstitution; immunosenescence.

INTRODUCTION

T lymphocytes play a crucial role in the adaptive immune response. A highly diverse T cell pool is required to recognize and eliminate foreign pathogens while also maintaining self-tolerance. The thymus continuously produces T cells, but with advancing age, its regenerative capacity declines,[1] leading to decreased thymic output and compromised T cell diversity. Furthermore, the thymus is also vulnerable to acute damage from infections, cancer therapies, and conditioning regimens for allogeneic hematopoietic cell transplantation (allo-HCT), which results in prolonged T cell lymphopenia, thereby increasing the risk of infections and cancer relapse, contributing to transplant-related complications and mortality.[2] Likely for these reasons, early T-cell reconstitution is a positive prognostic indicator of allo-HCT outcomes.[3,4]

Post-transplant T cell reconstitution requires a steady supply of hematopoietic stem cells (HSCs)-derived lymphoid progenitors[5] as well as thymic recovery. Prior research has demonstrated that thymic integrity is dependent on thymocyte-stromal crosstalk, particularly in the thymic epithelial cell (TEC) compartment.[6-8] Age-associated reduced lymphoid potential originating at the level of HSCs,[9] coupled with thymic decline[1] contribute to delayed immune recovery. Therefore, identifying reliable markers to prospectively isolate HSC subsets that augment T lymphopoiesis and promote thymic regeneration is an unmet clinical need. Despite efforts to identify markers for HSCs with multilineage potential, the functional heterogeneity within phenotypic HSCs poses a challenge. More importantly, these markers are unknown within the context of the lymphoid-producing human hematopoietic stem cell hierarchy.

Single-cell transplantation with tracking and sequencing of progeny[10-15] have uncovered heterogeneity in self renewal capacity as well as lineage output among reconstituting HSCs-

highlighting HSCs with multilineage versus lineage-biased potential. These studies showed that, while only a fraction of HSCs generate "balanced" multilineage output, the majority exhibit a diverse range of lineage potential following transplantation- varying in their contributions and reconstitution kinetics for each lineage, thereby showcasing their inherent biases. These lineage biases are intrinsically stable and maintained following transplantation[16-18]. At the molecular level, these cell autonomous features are mediated by epigenetic configuration and are divergent from their transcriptional state.[13,19] Although this suggests the existence of a highly organized and predictable framework for lineage-restricted fates of long-term self-renewing HSCs, the molecular identities, gene regulatory networks and functional implications of these differences governing lymphoid fate decisions remain largely unexplored.

Here, we detail the T-cell potential within the long-term reconstituting Kit^{lo} HSC subset, which decreases with aging. We delineate the pivotal role of this population in orchestrating thymic regeneration and immune reconstitution in both young and aged hosts. Leveraging integrated single cell transcriptional and epigenetic profiling, we uncover the distinct epigenetic signature of Kit^{lo} HSCs, highlighting the presence of lymphoid-specifying transcription factors like Zbtb1 and their pivotal role in enhancing T-cell lymphoid potential. Additionally, we identify an analogous HSC subset in human bone marrow and demonstrate their enhanced lymphoid potential.

<u>METHODS</u>

Detailed description can be found in Supplemental Information

Mice and Cell lines

Young (6-10 weeks), aged (18-24 months), and middle-aged (14-16 months) C57BL/6J, B6.SJL-PtprcaPepcb/BoyJ, BALB/cJ, FVB-Tg (RAG2-EGFP)1Mnz/J, and B6(C)-Gt(ROSA)26Sor$^{em1.1(CAG\text{-}cas9*,\text{-}EGFP)Rsky/J}$ mice were acquired from Jackson Labs or National Institute on Aging. S17, mouse and human MS5-DLL4, and MS5 stromal cells were used for *in vitro* assays.

Peripheral Blood and Organ Analyses

Transplant chimerism was tracked in peripheral blood. Bone Marrrow, Thymi and spleens were analyzed by flow cytometry and ELISA at the indicated time points.

Single Cell Multiome ATAC and Gene Expression Cell Preparation

Single Cell Multiome ATAC + Gene Expression was performed with the 10X genomics system using Chromium Next GEM Single Cell Multiome Reagent Kit A (catalog no. 1000282) and ATAC Kit A (catalog no. 1000280) following Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Kits User Guide and demonstrated protocol - Nuclei Isolation for Single Cell Multiome ATAC + Gene Expression Sequencing.

Data and Code Availability

Multiome single cell RNA-and- ATAC sequencing data generated for this paper is accessible through Gene Expression Omnibus (GEO) Series accession number GSE246464. No new algorithms were developed for this paper. All analysis code and input for reproducing figures is available at GitHub repository (https://github.com/HElias01/AgingHSC-Multiome.git).

RESULTS

Reduced Lymphoid Output in Aged HSCs Impairs Thymic Recovery

We utilized a pre-clinical allo-HCT model to analyze immune reconstitution of young and old donors. To evaluate if HSCs with differences in lymphoid progenitor production could impact thymic recovery, we transplanted Lin−Sca-1+Kit+ (LSK) CD34−CD150+CD48− cells LT-HSCs from young adult (2-mo) or old (24-mo) C57BL/6 mice with rescue BM cells (CD45.1) into lethally irradiated young BALB/cJ recipients (**Figure S1A**). Because HSC-derived thymic progeny emerge within 5-6 weeks,[20] we established an 8-week harvest timepoint. This interval allowed sufficient time for HSC-derived progenitors to repopulate the thymus, thereby providing a better understanding of thymic regeneration dynamics and the efficacy of HSC transplantation in promoting thymic reconstitution.

Consistent with prior reports, eight weeks post-HCT, we found that young HSCs supported balanced lineage reconstitution, while old HSCs exhibited significantly impaired T-cell reconstitution in the peripheral blood (PB), despite comparable reconstitution patterns at four weeks (**Figures S1B-S1E**). Furthermore, we observed greater multilineage potential of young HSCs, including lymphoid and myeloid reconstitution (LMPP/MPP4 and CLP, **Figures S1F and S1G**), when compared to old HSCs. Furthermore, young HSC recipients had significantly greater thymic cellularity and reconstitution of precursor and mature thymocytes than recipients of old HSCs (**Figures S1H-S1J**). In conclusion, old HSCs have significantly less multi-lineage potential, but especially T lymphoid potential than young HSCs in young allo-HCT recipients.

Aging alters Multilineage Potential of HSCs

Recent reports have shown remarkable functional heterogeneity of HSC[21-26] subsets, highlighting discordance between current HSC phenotypic definitions to their molecular identities[11-13]. To identify HSCs with an enriched multilineage program and gain insights into their transcriptional and epigenetic framework, we performed multiome single-cell RNA and ATAC sequencing on hematopoietic stem cells (LT-HSCs: Lineage$^-$CD34$^-$CD48$^-$CD150$^+$Sca-1$^+$c-Kit$^+$) from 2-mo (young) or 24-mo (old) female mice.

To define HSC subsets, we performed unsupervised Leiden cluster analysis on the multiome scRNA-seq data from our young cohort (**Figure S2A-S2B**). Using nomenclature and signatures derived from previous studies [10,11,13,27-30], we mapped all unsupervised clusters into five main subsets (**Figure S2C**): Quiescent HSC (q-HSC), Platelet-biased HSC (Mgk-HSC), Multilineage HSC (MLin-HSC), Proliferative HSC (p-HSC), and a novel Intermediate HSC (Int-HSC) that showed overlapping features with Mgk-HSC, MLin-HSC and p-HSC subsets (**Data Table 1, Figure S2D**). To identify HSCs with multilineage potential, we compared multilineage enriched MLin-HSC versus low output q-HSC subsets in young HSCs and queried for previously described bonafide HSC markers (*Hlf, Kit, Neo1, Procr, CD244a, Pdzk1ip1, Ly9, Hoxb5*). *Hlf* and *Kit* had significantly higher expression in q-HSCs compared to MLin-HSCs (**Figure 1A**). *Hlf* has been previously implicated in HSC quiescence,[31-33] whereas *Kit* enriches for platelet-biased HSCs.[34] To further analyze the composition of HSCs with varying Kit expression, we defined Kit high and low HSC populations (Kit^{hi} and Kit^{lo} HSCs) based on the top and bottom 20% of *Kit* gene expression levels. We observed lower Kit expression in MLin-HSCs (**Figure 1B**) and increased representation of q-HSC and MLin-HSCs within the Kit^{lo} HSCs (**Figures 1C and S2E**), consistent with prior reports of increased self-renewal capacity and multipotency in Kit^{lo} HSCs[34,35].

To assess for age-associated changes, we first compared the transcriptional profile of young and old HSCs. In our differential gene analysis, we found that HSCs from young mice had higher expression of genes related to cell cycle (*Ccnd3, Cdc25a, Cdk1),* and cell polarization (*Arhgap15*); conversely, old HSCs were enriched for expression of genes related to quiescence (*Egr1, Junb, Jun, Jund, Fos*), platelet differentiation (*vWf, Clu, Itgb3, Fhl1 and Tgm2*), and myeloid differentiation (*Selp, Neo1*) (**Data Table 2**). Next to investigate age-related shifts in HSC cluster composition, we employed *ingest()* from scanpy to query the old HSC transcriptomic data using the "trained" young reference data for HSC subsets (**Figure S2F**). Subsequently, we generated a UMAP of our young and old datasets (**Figures S2G and 1D**) to visualize the post-ingestion results (**Figures 1E**). While all HSC subsets were present in young and old cohorts (**Figure 1F**), we notably observed an increased representation of MLin-HSCs, p-HSCs, and Int-HSCs in young HSCs. Conversely q-HSCs and Mgk-HSCs were predominantly in old HSCs (**Figures S2H- S2J**). These findings suggest shifts in the composition of HSC subsets with aging, indicating a reduction in multilineage potential accompanied by an increase in platelet bias.

Given the increased frequency of Mgk-HSCs in old HSCs and prior work[34] demonstrating that Kit^{hi} HSCs identify platelet-biased HSCs (**Figures S2F- S2I**), we hypothesized this could be attributed to an expansion in Kit^{hi} HSCs. We observed significantly higher *Kit* expression (**Figures 1G-1H and S2K**) in old HSCs, with an age-dependent reduction in the proportion of Kit^{lo} HSCs accompanied by an increase in the frequency of Kit^{hi} HSCs (**Figure 1I**). Further substantiating our transcriptional findings with flow cytometry analysis of young and old bone

marrow, we noted a significant expansion of Kit^{hi} HSCs, accompanied by a concomitant decrease in Kit^{lo} HSCs in old HSCs (**Figures S2L and 1J**), consistent with our prior report[36]. Together, these findings indicate that the frequency of multilineage Kit^{lo} HSCs declines with age.

Young Kit^{lo} HSCs Exhibit Enhanced *In Vitro* and *In Vivo* Lymphoid Potential

Next, we sought to determine the lymphoid potential of the HSC Kit subsets. To this end, we used the S17 lymphoid progenitor assay[37] to assess for lymphoid progenitor output of purified Kit^{hi} and Kit^{lo} HSCs from young mice (**Figure 2A**). Prior studies have elucidated the existence of functional heterogeneity among phenotypic Common Lymphoid Progenitors (CLPs) characterized by Ly6D expression,[38] resulting in the identification of two distinct subsets: Absolute Lymphocyte Progenitors (ALPs), also referred to as functional CLPs, and B-cell biased Lymphocyte Progenitors (BLPs). Following 12-day co-culture, we observed that young Kit^{lo} HSCs generated more CLPs and ALPs than Kit^{hi} HSCs (**Figure 2B**). To further examine their T-cell differentiation potential, we used murine artificial thymic organoids (mATO, BM-ATO)[39-41] (**Figure 2A**). Following an 8-week culture, we noted a substantial increase in overall ATO output from young Kitl^{o} HSCs in comparison to Kit^{hi} HSCs (**Figure 2C**).

To further evaluate the reconstituting potential of Kit^{lo} HSCs, we competitively transplanted Kit^{lo} or Kit^{hi} HSCs from young C57BL/6 (CD45.2) mice with rescue BM cells (CD45.1) into lethally irradiated young BALB/cJ recipients (**Figure 2D**). Eight weeks after allo-HCT, we found that Kit^{lo} HSCs generated significantly better B- and T-cell reconstitution, while myeloid reconstitution was comparable (**Figures 2E, S3A-S3B**). Concordant with a prior report of better self-renewal ability[34], we found that Kit^{lo} HSCs demonstrated increased LT HSC reconstitution (**Figure 2F**). In support of multilineage potential, we observed that Kit^{lo} HSC recipients had

higher lymphoid progenitor reconstitution (including LMPP and CLP, **Figures 2F and 2G**), along with comparable myeloid precursor output (including MPP3, CMP, GMP, MEP, **Figures 2F and S3C**).

To evaluate whether enhanced lymphoid precursor output could contribute to thymic reconstitution, we concurrently evaluated the recipient thymi. We found that Kit^{lo} recipients demonstrated enhanced thymic recovery, characterized by significantly higher thymic cellularity and reconstitution of precursor and mature thymocytes (**Figures 2H-2J**). Notably, in support of a critical role for thymocyte-stromal crosstalk in preserving thymic architecture[6-8], we also found higher thymic epithelial cell (TEC), specifically medullary TEC, and endothelial cell recovery in Kit^{lo} reconstituted thymi (**Figures 2K and S3D**). We next examined donor-derived lymphoid progenitors for differential expression of molecules, involved in homing to the thymus (CCR7, CCR9, and PSGL1)[42].We observed significantly higher expression of CCR7 on Kit^{lo}-derived lymphoid progenitors (CLP) (**Figure S3E**). In addition, we analyzed thymic supernatants and found no differences in levels of thymopoietic ligands, IL22[43] and RANKL[44] (**Figure S3F**). Finally, we found significantly higher secondary lymphoid organ (spleen) reconstitution in Kit^{lo} recipients (**Figures S3G and S3H**). To evaluate long-term reconstitution, we performed additional harvests at 20-weeks following HCT. We observed similar trends in higher BM lymphopoiesis, thymic reconstitution, and thymic stromal recovery in Kit^{lo} recipients (**Figure S4**).

To assess thymic output, we used the RAG2-GFP[45] model, which allows the analysis of recent thymic emigrants (RTEs) (**Figure 2L**). We observed higher numbers of PB GFP+ T cells in Kit^{lo} recipients **(Figure 2M**). Additionally, to assess whether this increased number PB T cells in Kit^{lo} recipients results in increased functionality, (**Figure 2N**), we analyzed response to ovalbumin

engineered *Listeria Monocytogenes* (LM-OVA) and observed a significant increase in OT1+ CD8+ T cells (**Figure 2O**).

A key element in the post-HCT recovery of thymic function is the influx of bone marrow-derived lymphoid progenitors, also known as thymic seeding progenitors (TSPs).[5] Furthermore, the significance of restoring the bone marrow-thymus axis in mitigating thymic involution has been underscored through various strategies aimed at augmenting BM-lymphopoiesis, including sex-steroid ablation[46,47] and administration of Ghrelin,[48] among other approaches.[49] We therefore hypothesized that robust BM lymphopoiesis driven by Kit^{lo} HSCs could potentially counteract age-related alterations within the thymic microenvironment, consequently ameliorating the decline in T-cell output. To this end, we competitively transplanted Kit^{lo} or Kit^{hi} HSCs from young C57BL/6 mice (CD45.2) with rescue BM cells (CD45.1) into lethally irradiated 14-month-old (Middle-aged) BALB/cJ recipients (**Figure 2P**). Sixteen weeks post-transplantation, we observed in Kit^{lo} vs Kit^{hi} recipients significantly higher numbers of peripheral T– and B-cells (**Figures 2Q, S3I, and S3J**). This was attributed to increased numbers of Kit^{lo} HSC-derived lymphoid progenitors (including LMPP/MPP4 and CLPs, **Figures S3K and S3L**), while numbers of myeloid precursors were comparable (MPP3, CMP, GMP, and MEP, **Figures S3K and S3M**). In our parallel analysis of the thymi of Kit^{lo} vs Kit^{hi} recipients, we observed increased numbers of thymocytes and TECs (**Figures 2R-2T and S3N**). Collectively, these findings establish the potential of Kit^{lo} HSCs for lymphopoiesis and to improve thymic recovery and peripheral T cell reconstitution in recipients of various ages.

Kit^{lo} HSCs are Epigenetically Programmed for Lymphoid Potential

HSCs can have a divergence between transcriptional and epigenetic lineage priming,[13,22,50] and epigenetic priming predicting differential lymphoid output.[13,19] To identify the gene regulatory network orchestrating the lymphoid potential differences in HSCs, we analyzed our single-cell ATAC-seq (scATAC-seq) component from our multiome dataset (**Figure S5D**). Through unsupervised clustering, we identified 6 clusters, including young and old HSCs (**Figures 3A and 3B**). Distinct HSC subsets identified via scRNA-seq analyses (**Figure 1B**) are situated in varied regions on the scATAC-seq UMAP (**Figures 3C and 3D**). Notably, MLin-HSC and Kit^{lo} HSC intersect with Cluster 3 from the scATAC-seq clusters (**Figure 3E**).

We characterized the differentially accessible regions (DARs) of chromatin that defined the HSC subsets and found 2816 DARs in cluster C3 (MLin-HSC and Kit^{lo} enriched) cluster compared to cluster C6 (q-HSC and Kit^{hi} enriched) (**Figure S5F**). Next to determine the differential activity and accessibility of lymphoid-specifying transcription factors (TFs) in the identified Kit^{hi} and Kit^{lo} HSC subsets, we analyzed chromatin accessibility variation within TF motifs using chromVAR and gene accessibility. Compared to Kit^{hi} HSCs, we found increased gene accessibility and enrichment for Zinc finger motifs, specifically *Zbtb1 and Zbtb7a,* in Kit^{lo} DARs, which are implicated in lymphoid differentiation program[13,51-53] (**Figures 3E-3F**).

In summary, we found that Kit^{lo} HSCs exhibit an epigenetic program, emphasizing an enrichment of lymphoid-specifying transcription factors, in support of their heightened lymphoid potential.

Zbtb1 **regulates T-cell potential in phenotypic HSCs**

Among members of the ZBTB gene family, *Zbtb7a* is a critical regulator of B versus T lineage fate decision[52], while *Zbtb1* plays a pivotal role in T-cell development[53-55]. Furthermore, given

the role of *Zbtb1* in lymphoid reconstitution[53], we hypothesized that Zbtb1 loss impairs the lymphoid potential of Kit^{lo} HSCs. To this end, we used the CRISPR-Cas9 system to generate *Zbtb1*-deficient Kit^{lo} HSCs (**Figure 4A**) and confirmed decrease in Zbtb1 by flow-cytometric analysis (**Figure S6A**). *Zbtb1*-deficient Kit^{lo} HSCs showed significantly reduced *in vitro* T-cell potential (**Figure 4B**) and impaired lymphoid reconstitution *in vivo* (**Figure 4C**). Next, we questioned whether reinstating *Zbtb1* expression in megakaryocytic-biased, Kit^{hi} HSCs is sufficient to rescue their lymphoid defect. To test our hypothesis, we overexpressed *Zbtb1* cDNA in Kit^{hi} HSCs (**Figure 4D and S6B**) and we observed a rescue of their T-cell potential when compared to Kit^{hi} controls, thus partially phenocopying Kit^{lo} HSCs (**Figures 4E**) as well as improved lymphoid reconstitution *in vivo* (**Figure 4F**).

Previous studies have identified Zbtb1 targets which facilitate Notch-mediated T-cell differentiation in lymphoid progenitors[56] and support chromatin remodeling following DNA damage[57]. Given *Zbtb1* motif enrichment and enhanced T-cell differentiation in Kit^{lo} HSCs, we hypothesized that Zbtb1 targets mediating lymphoid differentiation gene programs are also enriched in Kit^{lo} HSCs. To this end, we queried Kit^{lo} enriched genes for Notch1-interacting proteins[58]. We observed ~50% overlap of Notch1-partners enriched in Kit^{lo} HSCs, of which 60% were Zbtb1-targets identified in a previous CHIP-seq study[56] (**Figure 4G**). Furthermore, we identified an additional subset of Zbtb1 targets (33%) that were not associated with the Notch1 interactome (**Data Table 3**). To further query for non-Notch related biological pathways, we performed pathway analysis and identified a strong enrichment for gene sets related to chromatin remodeling, RNA stability, and (**Figure 4H, Data Table 4**).

These data demonstrate that Zbtb1 contributes to an epigenetic regulatory program that facilitates T-cell differentiation in primitive Kit^{lo} HSCs likely through instructive and permissive mechanisms.

Old Kit^{lo} HSCs Exhibit Enhanced *In Vitro* and *In Vivo* Lymphoid Potential

To determine whether Kit^{lo} HSCs retain their lymphoid potential with age, we evaluated the lymphoid potential of old Kit subsets. Remarkably, even in the case of old mice, we noted a substantial increase in lymphoid progenitor and ATO output from old Kit^{lo} HSCs compared to Kit^{hi} HSCs (**Figures 5A-5B**), although modest when compared to the young HSCs. This finding is consistent with prior reports of proliferative defects in aged T-cell precursors contributing to decreased thymopoiesis.[59,60]

Next, to test their *in vivo* potential, we competitively transplanted either Kit^{lo} or Kit^{hi} HSCs from old C57BL/6 (CD45.2) mice with rescue BM (CD45.1) cells into lethally irradiated young BALB/cJ recipients (**Figure 5C**). At eight weeks post-transplantation, we noted significantly higher numbers of T cells in the peripheral blood of Kit^{lo} vs Kit^{hi} recipients. (**Figures 5D and S7A**). Consistent with our experiments using HSCs from young donors, we observed in the BM of Kit^{lo} vs Kit^{hi} recipients significantly higher numbers of LT HSCs, LMPPs and CLPs (**Figures 5E-5F**) with comparable myeloid potential (**Figures 5E and S7B**). Furthermore, mirroring our findings in HSCs from young mice, we observed in recipients of old Kit^{lo} HSCs enhanced overall thymic cellularity, thymocyte reconstitution, and recovery of TECs and endothelial cells (**Figures 5G-5I and S7C-S7D**), increased CD4+ and CD8+ T cells in the spleen (**Figures S7E and S7F**), and higher RTE output (**Figures 5J-5K**). At 20-weeks post-HCT, we observed similar trends of enhanced reconstitution in old Kit^{lo} HSC recipients, confirming their long-term

reconstituting potential (**Figure S7H-S7O**). Next to examine whether old HSC subsets with preserved lymphoid could augment aged thymic recovery, we competitively transplanted middle-aged recipients with old Kit^{lo} or Kit^{hi} donors (**Figure 5L**). We observed increased CLPs in the BM and peripheral T cells, but a modest, non-significant increase in thymic cellularity in Kit^{lo} recipients (**Figures 5M-5O, and S7G**). Similar to our *Zbtb1* functional studies with young Kit^{hi} HSCs, we observed a rescue in T-cell potential in *Zbtb1*-overexpressed old Kit^{hi} HSCs (**Figure 5P**).

Finally, to directly compare the per cell functionality of Kit^{lo} HSCs in the context of aging, we generated mixed chimeras combining Kit^{lo} HSCs cells from both young and old mice in an equal competition transplant (**Figure S8A**). Consistent with our *in vitro* findings (**Figures 5A-5B**), we found an overall decrease in lymphoid reconstitution within the PB, BM, and thymus with old Kit^{lo} HSCs when compared to their young counterparts (**Figures S8B-S8G**). Pathway analysis revealed decreased enrichment of chromatin remodeling related gene sets (**Figure S8H, Data Table 5**). Additionally, we noted decreased accessibility and expression of *Zbtb1 (***Figures S8I-S8K***)* in Old Kit^{lo} HSCs, indicating a impaired epigenetic program for lymphoid differentiation. These findings demonstrate that old Kit^{lo} HSCs exhibit preserved lymphoid potential and contribute to thymic recovery, albeit with reduced output.

Kit^{lo} Human BM HSCs Subsets show Higher Lymphoid Potential

Informed by a recent study that mapped the molecular profiles of human thymic seeding progenitors (TSPs) to their transcriptional counterparts within the HSC compartment,[61] we next sought to identify a comparable human HSC subset with lymphoid potential. Interrogating a CITE-Seq dataset[62] generated from young and old human bone marrow (BM) samples (**Figures**

6A-6C, S9A), we mapped our mouse Kit^{lo} gene signature onto human HSCs and noted a significant enrichment in young BM (**Figures 6D and 6E**). We observed lower protein expression of KIT (CD117) on young HSCs by using antibody-derived tag (ADT) reads (**Figure S9B**). To validate our in-silico findings, we next assessed KIT expression in human BM samples (young: 20-40 years; middle-age/ old: >40 years). In our FACS analysis, we found that phenotypic HSCs (pHSCs: CD34+CD38-CD10-CD45RA-CD90+) showed the lowest KIT expression (**Figure S9C**), with higher KIT levels on old pHSCs (**Figure 6F**). Moreover, we noted a statistically significant association between age and the composition of KIT subsets in pHSCs, indicating an increase in KIT^{hi} and a decrease in KIT^{lo} HSCs, aligning with our observations in mice. (**Figure 6G**).

To evaluate the T-cell potential of these human HSC subsets, we generated artificial human thymic organoids (huATOs) using purified KIT^{hi} or KIT^{lo} HSCs from human BM samples (**Figures S9D and 6H**; **Data Table 6**). In concordance with our mouse *in vitro* studies, following 8-10 weeks of culture, we observed significantly higher output of precursor and mature T-cells from KIT^{lo} HSCs- derived ATOs compared to KIT^{hi} HSCs (**Figure 6I**). To further evaluate the multilineage potential of KIT HSC subsets, we performed a lympho-myeloid differentiation assay, as previously described[63]. While we observed comparable myeloid differentiation potential, Kit^{lo} HSCs exhibited significantly more B cell and a non-significant increase in NK cell differentiation potential (**Figure 6J**).

Collectively, these results identify a previously uncharacterized subset of human BM HSCs, KIT^{lo} HSCs – which exhibit enhanced lymphoid potential.

DISCUSSION

Recent studies have uncovered significant heterogeneity within the HSC compartment and their role in hematopoiesis. This heterogeneity has implications for disease development and treatment. Of note, certain aspects of HSC heterogeneity challenge the established hematopoietic models, with lineage biases and lineage-restricted cells impacting self-renewal dynamics.[64,65] Decisions regarding lympho-myeloid fate are established at multiple levels of the hematopoietic hierarchy: HSCs,[16,17,66] MPP, LMPP, and more mature GMPs and CLPs. At the level of HSCs, the process of epigenetic imprinting confers a distinct advantage for lymphoid priming within multilineage HSCs, as well as maintaining these cell autonomous features imposed at steady state, even under conditions of stress.[13,19]

Here, we report on the enhanced lymphoid potential of the BM Kit HSC subset, Kit^{lo} HSCs, observed in both mice and humans. Through combining multiomic single-cell sequencing and functional analyses utilizing a preclinical allo-HCT model, we demonstrate augmented potential of Kit^{lo} HSCs for lymphopoiesis and to support thymic recovery, independent of age. During the early post-transplant period, we noted a relative increase in DN3 thymocytes relative to other thymocyte precursor subsets. This observation is consistent with previous studies reporting that DN3 thymocytes undergo compensatory expansion to support early post-transplant thymopoiesis, with minimal contribution from T-cell precursor thymocytes (ETP and DN2)[5,67,68]. Further investigation into mechanisms driving these compensatory processes and its implications for immune function is critical to our understanding of post-transplant immune reconstitution.

Mechanistically, Kit^{lo} HSCs exhibit differential expression and activity of lymphoid-specifying transcription factors (TFs), such as *Zbtb1*, indicative of their epigenetic program towards a

lymphoid fate. Studies in mouse models have shown that disruptions in *Zbtb1* result in T-cell developmental defects and severe combined immunodeficiency phenotypes[53,54], along with competitive disadvantages in B- and NK-cell compartments [53]. In our study, we identify *Zbtb1* as an essential transcription factor governing lymphoid specification in phenotypic HSCs and facilitating T-cell differentiation by reprogramming lineage biased Kit^{hi} HSCs. Additionally, we observed functional differences within Kit^{lo} HSCs in the context of aging, specifically a competitive disadvantage in Old Kit^{lo} HSCs, potentially attributable to their differences in *Zbtb1* expression. Consistent with prior research, our study suggests *Zbtb1* modulates Notch-dependent[56] and cellular survival programs[24,69]. However, further mechanistic studies are needed to elucidate the precise transcriptional mechanisms underlying this regulation.

KIT has previously been reported as a reliable cross-species marker to fractionate phenotypic HSCs in Cord Blood (CB)[70]. This study[70] showed reduced multipotency of KIT^{int} HSCs, contradicting a prior report that showed the long-term reconstituting potential of CD34+KIT^{lo} human BM cells[71]. In our study, combining transcriptional, immunophenotypic, and functional analyses of human BM samples we identified a KIT^{lo} subset within HSCs with enhanced lymphoid potential- demonstrating their translational relevance in human biology. In contrast to the CB study[70], our observation of multipotency in KIT^{lo} HSCs within the BM suggests an ontogeny specific relevance of Kit signaling. Although further functional characterization is warranted to definitively establish their clinical significance.

<u>Methods</u>

Animal use

Young mice (6-10 weeks of age): C57BL/6J (CD45.2/H-2Kb), B6.SJL-PtprcaPepcb/BoyJ (CD45.1), BALB/cJ (H-2Kd), and B6(C)-Gt(ROSA)26Sor$^{em1.1(CAG-cas9*,-EGFP)Rsky/J}$ (Rosa26^{Cas9} KI) mice were purchased from The Jackson Laboratory. Aged C57BL/6 (CD45.2/H-2Kb) mice (23-24 months of age) were obtained from National Institute of Aging (Baltimore, MD). Aged mice (ranging between 18-24 months of age) correlate with humans ranging between 56-69 years of age. We obtained middle-aged mice (ranging between 14-16 months of age) by initially purchasing young BALB/cJ (H-2Kd) mice from Jackson Laboratories (JAX) and subsequently allowing them to age under controlled conditions in our facility. Middle-aged mice (ranging between 14-16 months of age) correlate with humans ranging between 40-60 years of age. RAG2-EGFP-CD45.1 chimeric mice were generated by crossing FVB-Tg (RAG2-EGFP)1Mnz/J (JAX# 005688) and B6.SJL-PtprcaPepcb/BoyJ (CD45.1, JAX# 002014). Young OT-1-CD45.2 (Va2+Vb2+/ CD45.2/H2-Kb) were obtained from Dr. Andrea Schietinger. For consistency all experiments were carried out using only female mice. These mice were allowed to acclimatize in our vivarium for at least 10-14 days before experiments. Mice were euthanized with CO2 gas inhalation and were maintained under pathogen-free conditions according to an MSKCC IACUC-approved protocol.

Cell lines

To provide a microenvironment that supports lymphoid progenitor in vitro, we cocultured purified HSCs from young and old B6 bone marrow with the S17 stromal cells[37], which were originally obtained from K. Dorshkind (UCLA). To support in vitro T-cell differentiation, we co-cultured

purified HSCs from young and old B6 bone marrow and human BM with MS5-mDLL4 or MS5-hDLL4 stromal cells [39], which were obtained from Dr. Gay Crooks (UCLA). Subsequently, both cell lines were maintained in our laboratory as described in the original references. Details of the differentiation cultures are given below under HSC Lymphoid Differentiation Assays.

Cell Preparation and Staining

Femurs, tibias, pelvises, and the spine were dissected from euthanized mice and cleaned of muscles and connective tissue on ice in 1X PBS. Bone marrow cells were collected by crushing all the bones with a sterile mortar and pestle in FACS buffer: 1X PBS containing 2% Fetal Bovine Serum (Hyclone) and 2 mM EDTA. BM cells were enriched for c-Kit+ cells using mouse CD117 magnetic microbeads (Miltenyi Biotec) per manufacturer's protocol.

Bone Marrow analysis and Flow-cytometric Isolation of Hematopoietic Stem Cells

BM cells were stained with fluorochrome-conjugated antibodies as well as propidium iodide (Molecular Probes) or Zombie Aqua™ Fixable Viability Kit (Biolegend). All monoclonal antibodies were purchased from Biolegend and eBiosciences. The monoclonal antibodies included Phycoerythrin-Cyanine 7 (PE-Cy7) antibodies to lineage (Lin) markers (CD3ε, CD4, CD8a, B220, Gr-1, Mac-1 and Ter119), Brilliant Violet 711 (BV711) or Brilliant Violet 605 to Sca-1, APC/Cyanine7 to cKit, FITC antibody to CD34, PE antibody to CD150 (SLAM), Pacific Blue to CD48, APC to CD45.2 and PerCP/Cyanine5.5 to CD45.1. Cells were analyzed or sorted using a FACS LSR II UV or Aria II cell sorter (BD Biosciences). For all experiments (transplantation and in vitro assays) in which Kit^{hi}, Kit^{lo} or Kit HSCs were purified, they were double sorted to ensure >95% purity.

To calculate bone marrow cellularity and frequency of donor-derived hematopoietic precursors, two femurs and two tibias of primary recipient (BALB/cJ) were flushed into FACS buffer at 8- or 20-weeks post-transplant. Collected cells were then incubated with red blood cell lysis buffer (ACK lysis buffer, Thermo Fisher Scientific) for 8 minutes and then washed twice with PBS/2.5% FBS. Cells were resuspended and then stained in PBS/2.5% fetal calf serum with fluorochrome-conjugated antibodies (Biolegend, BD, eBiosciences) against lineage (Lin) markers (CD3ε, CD4, CD8a, B220, Gr-1, Mac-1 and Ter119) (PE-Cy5), Sca-1 (BV711), c-Kit (APC-Cyanine7), CD150/ SLAM (BV605), CD135/Flt3 (BV421 or PE-Cy5), CD127/IL7Ra (APC), CD48 (BV510), CD16/32 (AF-700), CD34 (FITC), CD45.2 (BUV395), CD45.1 (PE-Cy7 or AF-700), H-2Kd (PE) and H-2Kb (BV650 or BV786). For additional thymic adhesion molecule analyses, antibodies against CCR7 (BV421), CCR9 (FITC) and PSGL1 (PE) were used for chimerism studies. Following a 30-45 incubation on ice, cells were washed in in PBS/2.5% fetal calf serum. Finally, propidium iodide (Molecular Probes) was added as a viability dye before acquiring the data. Cells were analyzed or sorted using a FACS LSR II UV (BD Biosciences).

Transplantation experiments

A. For competitive transplantation assays, double-FACS sorted Kit^{hi}, Kit^{lo} HSCs cells from young or old CD45.2 mice were mixed with unfractionated BMMCs from young CD45.1 mice. Cells were transplanted via retro-orbital sinus injections into lethally irradiated young (6-8-week-old) or middle aged (14-16-month-old) BALB/cJ recipient mice (9 Gy, single dose, using a 79 X-Ray source) under isoflurane anesthesia, within 1h post-irradiation. **B. For Kit^{lo} equal competition experiment,** 375 double-FACS sorted Kit^{lo} HSCs cells from both young (CD45.1/CD45.2) and old CD45.2 mice were combined with 5x10^{5} unfractionated BMMCs from

young CD45.1 mice and transplanted into lethally irradiated 6-8-week-old BALB/cJ recipient mice. **C. For RTE experiments**, 750 double-FACS sorted Kit^{hi}, Kit^{lo} HSCs cells from young and old RAG2-GFP (CD45.1) mice were mixed with $5x10^5$ unfractionated BMMCs from young CD45.2 mice and transplanted into lethally irradiated 6-8-week-old BALB/cJ recipient mice. **C. OT-1 Transplant and Adoptive T-cell transfer for functional assessment**, 750 Kit^{hi} and Kit^{lo} HSCs were double FACS-sorted from 8–10-week-old OT-1 mice and competitively transplanted with $5x10^5$ unfractionated competitor BMMCs from (CD45.1) into lethally irradiated young (6-8-week-old) BALB/cJ (H-2Kd) recipients. Eight-week post-transplant (BMT), spleens of primary recipients were brought into single-cell suspension and 1/10 of splenocytes were adoptively transferred into young (8-10wk) C57BL/6 secondary recipients. 4 hours later, mice were infected with 5.000 CFU L. monocytogenes bacteria expressing chicken ovalbumin (LM-OVA) to evaluate functional CD8+T cell response to primary infection. On day 8 post-infection, spleens of infected mice were analyzed via flow cytometry. Counting beads were used to calculate absolute cell numbers.

Peripheral blood analysis

Peripheral blood samples were collected in 50 mM EDTA solution (Thermo Fisher Scientific) via retro-orbital sinus bleeds. Thereafter PB was incubated with red blood cell lysis buffer (ACK lysis buffer, Thermo Fisher Scientific) for 8 minutes and then washed twice with PBS/2.5% fetal bovine serum. Cells were resuspended and then stained in PBS/2.5% fetal calf serum with antibodies against CD3 (AF700), B220 (PE-Cy7), Gr-1 (APC), Mac-1 (APC), NK1-1 (BV605), CD45.2 (BUV395), CD45.1 (BV421), H-2Kd (PE) and H-2Kb (FITC) for chimerism studies for 30 mins followed by wash in in PBS/2.5% FBS. Finally, propidium iodide (Molecular Probes)

was added as a viability dye before acquiring the data. Cells were analyzed using a FACS LSR II UV and LSR Fortessa X50 (BD Biosciences). To generate absolute T-cell counts, complete blood counts, including differentials, were obtained using a ProCyte Dx Hematology Analyzer (IDEXX).

Thymi Harvests: FACS and ELISA analysis

All steps were performed on ice unless indicated. Primary recipient (BALB/cJ) thymi at 8 or 20 wks post-transplant were excised and enzymatically digested following an adapted protocol.[43] Briefly, thymi were mechanically dissociated into ca. 2 mm pieces. Tissue pieces were incubated with a digestion buffer (RPMI, 10% FCS, 62.5 um/mL liberase TM, 0.4 mg/ml DNase I) twice for 30 min at 37 C. Between incubation steps, supernatant containing dissociated cells was transferred to 50 mL conical tubes equipped with 100 um filter.

For thymic immunophenotypic analysis, cells were first incubated with Fc block solution (anti-CD16/CD32 antibody) for 10 min on ice. Solution was discarded and cells were stained in PBS/2.5% fetal calf serum with fluorochrome-conjugated antibodies against lineage (Lin) markers (CD19, CD11b, NK1-1, TCRγδ, Gr-1, Ter119) (Biotin), Streptavidin (PE-TexasRed), CD4 (PE-Cy7), CD8 (BV711), CD25 (BV510), CD44 (AF700), c-Kit (APC), CD135/Flt3 (BV421), CD45.2 (BUV395), CD45.1 (BV605), H-2Kd (PE) and H-2Kb (FITC). for chimerism studies. To characterize the CD45- compartment, thymic cells were stained with antibodies against UEA-1 (FITC), 6C3 (PE), EPCAM (BV605), PDGFRa (BV421), MHC-II (APC), CD31 (PE-Cy7), CD45 (BUV395) and Ter-119 (PE-Cy5.5) for 30 mins followed by wash in in PBS/2.5% FBS. Antibodies were purchased from either (Biolegend and eBiosciences), except Ulex europaeus agglutinin 1 (UEA-1), conjugated to FITC, was purchased from Vector

Laboratories (Burlingame, CA). Finally, 7-AAD (Molecular Probes) was added as a viability dye before acquiring the data. Flow Cytometric analysis was performed on FACS LSR II UV (BD Biosciences).

For all ELISA experiments, thymi were suspended by mechanical dissociation. The resultant supernatants were quantified using mouse cytokine (IL-22 and RANKL) specific ELISA kits from R&D Systems and read on a Spark Multimode Microplate Reader (TECAN).

Spleen Harvest and Analysis

All steps were performed on ice unless indicated. Primary recipient (BALB/cJ) spleens at 8- or 20-weeks post-transplant were excised and crushed, filtered through a 70uM strainer (pluriSelect), intermittently washed with PBS/2.5% fetal bovine serum and collected in a 50ml conical tube. Subsequently, spleen pellet was incubated in red blood cell lysis buffer (ACK lysis buffer, Thermo Fisher Scientific) for 8 minutes and then washed twice with PBS/2.5% fetal bovine serum. Prior to staining, an aliquot was set aside for counts with the Nexcelom Cellometer K2. For spleen immunophenotypic analysis, cells were first incubated with Fc block solution (anti-CD16/CD32 antibody) for 10 min on ice. Solution was discarded and stained in PBS/2.5% fetal calf serum with fluorochrome-conjugated antibodies against lineage (Lin) markers (CD19, CD11b, NK1-1, TCRγδ, Gr-1, Ter119) (Biotin), Streptavidin (PE-TexasRed), CD4 (PE-Cy7), CD8 (BV711), CD62L (BV510), CD44 (AF700), CD45.2 (BUV395), CD45.1 (BV605), H-2Kd (PE) and H-2Kb (FITC). for chimerism studies.

<u>HSC Lymphoid Differentiation Assays</u>

A. S17 Progenitor Assay

Lymphoid progenitor production from young and old B6 bone marrow was measured by mixing FACS purified 200 Kit^{hi} and Kit^{lo} HSCs with 50,000 S17 stromal cells[37] in 1.5 mL of methylcellulose (MC) medium. MC medium was prepared by supplementing a-MEM with 30% heat-inactivated FCS, 1% methylcellulose (STEMCELL Technologies), 5x10-5 M 2ME, 2 mM L-glutamine, 50 µg/mL gentamicin, 100 U/mL streptomycin, 100 µg/mL penicillin, 0.1 mM MEM vitamins, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 20 ng/mL stem cell factor, 20 ng/mL Flt3L ligand, and 50 ng/mL IL-7 (all from Biosource). The mixture was plated in non-tissue-culture treated 3.5-cm2 dishes (Becton Dickinson). Following 12 days of culture, the contents of the plates were harvested, cells were enumerated, and examined for production of lymphoid progenitors by flow cytometry. All cultures were placed at 37C, 5% CO2 humidified incubators until processing.

B. Mouse and Human ATO Assay

Prior to setting up the ATO experiments, the MS5 stromal cell lines were thawed and expanded in culture for at least 7 days. The cell lines and ATOs were set up and maintained as initially described.[39]

For the mouse ATO system, FACS purified 100 Kit^{hi} and Kit^{lo} HSC cells from young and old mice were combined with 150k MS5-mDll4 stromal cells to form each murine ATO. The media (DMEM/F-12, 50x B27, 100x Glutamax, 100x Pen/Strep, and 1000x Ascorbic acid) was changed twice a week, with beta-mercaptoethanol (1000x) and cytokines (5ng/mL murine IL-7, 5ng/mL Flt3L, and 5ng/mL SCF) added fresh each time.

For the human ATO system, healthy human bone marrow mononuclear cells were purchased from BIOIVT (Johnson City, TN) for FACS sorting of Kit^{hi} and Kit^{lo} HSCs within phenotypic

HSCs ($CD34^+CD38^-CD10^-CD45RA^-CD90^+$). The sorted cells were then aggregated with MS5-hDll4 (50-150 HSCs and 150k MS5-hDll4 cells per ATO) and seeded on culture inserts (Millipore Sigma). The human ATO media (RPMI 1640, 25x B27, 100x Glutamax, 100x Pen/Strep, 1000x Ascorbic Acid) was refreshed twice a week, supplemented with 2.5ng/mL recombinant human IL-7, 5ng/mL recombinant human Flt3L (Peprotech), and 5ng of human SCF was used only for the first week of differentiation.

For analysis via FACS, the ATO was mechanically disrupted, filtered through a 70uM strainer (pluriSelect), and counted with the Nexcelom Cellometer K2 prior to staining. Harvest timepoints varied, dependent on the initial primary cell population (mATO: 6-8-weeks; huATO: 8-10-weeks).

C. Lympho-myeloid Differentiation Assay

As previously described[63], MS-5 cells were seeded onto 0.1% gelatin-coated 24-well plate at an initiating density of 2×10^4/well in α-MEM medium (Thermo Fisher) supplemented with 10% FBS (Gibco), 1% penicillin-streptomycin (Hyclone) and 1% GlutaMAXTM (Gibco). Twenty-four hours after plating of MS-5 stroma, 200 cells were added into each well in the presence of 0.1 μM DuP-697 (Cayman Chemicals), 20 ng/ml SCF (Peprotech), 10 ng/ml G-CSF (Peprotech), 10 ng/ml FLT3L (Peprotech), 10 ng/ml IL-2 (Peprotech), 10 ng/ml IL-15 (Sino Biological). Cultures were maintained for 4 weeks with weekly half media changes (with 2x cytokines). Cocultures were transferred onto fresh MS-5 stroma every two weeks through 40 μm filter to remove the stromal cells. All the cells in each well were harvested and analyzed by flow cytometry at week 4. The antibodies used to read the lineage outputs in this assay were

anti-human CD45-BV480, anti-human CD14-BUV805, anti-human CD15-BV421, anti-human CD19-FITC, anti-human CD56-BUV496, and anti-human CD235a-PE-Cy7.

Lentiviral Production

For deletion experiments, scramble and Zbtb1 guide RNAs (sgRNA#1: GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC; sgRNA#2: GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC) were cloned into U6-EFS-mCherry (VectorBuilder), while Control and mouse Zbtb1 [NM_178744.3] were individually cloned into CMV-EFS-EGFP (VectorBuilder) for over-expression experiments and utilized for lentivirus generation. Lentivirus was generated by cotransfecting the above plasmids containing either sgRNAs or cDNA, pMD2.G (Addgene #12259), and psPAX2 (Addgene#12260) into HEK 293□T cells. SFEM Media was replaced 6h post-transfection, and the virus suspension was harvested 36-48 h post-media change. The supernatant was collected and used for lentiviral transductions.

HSC Lentiviral Transduction

As previously outlined sorted Kit^{lo} and Kit^{hi} HSCs from young, aged B6, and young $Rosa26^{Cas9}$ KI mice were cultured overnight in SFEM (STEMCELL Technologies) media supplemented with murine cytokines (50□ng/ml SCF, 10□ng/ml TPO) in single-wells of a 96-well plate (TC treated). The cells were transduced with lentiviral suspensions by spinfection in retronectin (Takara Bio)-coated plates for 90 mins. Post spinfection, cytokine-supplemented fresh media was added to the transduced wells. Twenty-four hours later, cells were either sorted for RFP and GFP positivity (for deletion experiments), or GFP expression (for over-expression experiments)

using the BD Aria instrument. KO and OE was confirmed by flow cytometric analysis with antibodies against Zbtb1 (AlexaFluor647, Signalway Antibody Cat#C47476-AF647).

Single Cell Multiome ATAC and Gene Expression Cell Preparation

Single Cell Multiome ATAC + Gene Expression was performed with the 10X genomics system using Chromium Next GEM Single Cell Multiome Reagent Kit A (catalog no. 1000282) and ATAC Kit A (catalog no. 1000280) following Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Kits User Guide and demonstrated protocol - Nuclei Isolation for Single Cell Multiome ATAC + Gene Expression Sequencing. Briefly, cells (viability 95%) were lysed for 4min and resuspended in Diluted Nuclei Buffer (10x Genomics, PN- 2000207). Lysis efficiency and nuclei concentration was evaluated on Countess II automatic cell counter by trypan blue staining. Nuclei were loaded per transposition reaction, with a targeting recovery between 1,000 and 10,000 nuclei after encapsulation. After transposition reaction nuclei were encapsulated and barcoded. Next-generation sequencing libraries were constructed following User Guide, which were sequenced on an Illumina NovaSeq 6000 system.

<u>SEQUENCING DATA PROCESSING</u>

Single Cell RNA-seq

i) Preprocessing and downstream data analysis

FASTQ files were processed using the 10x Cell Ranger package (v7.01). The Cell Ranger generated filtered_feature_bc_matrix.h5 files were processed following the guidelines on the shunPykeR GitHub repository(https://github.com/kousaa/shunPykeR), an assembled pipeline of publicly available single cell analysis packages put in coherent order, that allows for data analysis in a reproducible manner and seamless usage of Python and R code. Genes that were not

expressed in any cell and ribosomal and hemoglobin genes were removed from downstream analysis. Each cell was then normalized to a total library size of 10,000 reads and gene counts were log-transformed using the log(X+1) formula, in which log denotes the natural logarithm. Principal component analysis (components =20) was applied to reduce noise prior to data clustering. To select the optimal number of principal components to retain for each dataset, the knee point (eigenvalues smaller radius of curvature) was used. Leiden clustering[72] (resolution = 0.9) was used to identify clusters within the PCA-reduced data.

Quality of the single cells (**Fig. S2A-S2B**) was computationally assessed based on total counts, number of genes, mitochondrial and ribosomal fraction per cell, with low total counts, low number of genes (≤1000) and high mitochondrial content (≥0.2) as negative indicators of cell quality. Cells characterized by more than one negative indicator were considered as "bad" quality cells. Although cells were negatively sorted prior to sequencing for the CD45 marker, a small amount of non-hematopoietic cells (expressing no Ptprc), were detected within our dataset. To remove bad quality cells and contaminants in an unbiased way, we assessed them in a cluster basis rather than individually. Leiden clusters with a "bad" quality profile and/or a high number of contaminating cells were removed. Finally, cells marked as doublets by scrublet[73] were also filtered out. Overall, a total of 2325 cells, representing ~8.4% of all our data, was excluded from further analysis (see Figure S2 for per sample metrics). After removal of these cells, we calculated highly variable genes (HVG=3000) and re-performed PCA with unsupervised clustering analysis (components =20), followed by batch effect correction across all samples using harmony[74] to assist annotation of cell type subsets within the dataset.

ii) Defining Kit^{lo} and Kit^{hi} HSC subsets

Owing to the significant drop-out effect in scRNA-seq data,[75] it is virtually impossible to ascertain if a cell has low/mid (and sometimes high) Kit expression versus no expression. In order to methodically deal with this, we performed _data denoising and_ imputation using the MAGIC (Markov Affinity-based Graph Imputation of Cells) method,[76] to denoise distinguish between dropouts and genuinely low gene expression values. Post-MAGIC imputation, we proceeded to set the thresholds for *Kit*, specifically, we calculated the 20th and 80th percentile expression value across all cells. Subsequently, cells were classified into Kit^{lo} (*Kit* expression below the 20th percentile), Kit^{hi} (*Kit* expression above the 80th percentile), and Kit^{mid} (*Kit* expression between the 20th and 80th percentile). This approach helped to enhance the quality of the data by denoising and imputing the missing values, thereby providing a more accurate representation of the gene expression landscape.

iii) Differential expression analysis

Differential expression analysis for comparisons of interest was performed with MAST (Model-based Analysis of Single-cell Transcriptomics) using the likelihood ratio test.[77] In all cases, differentially expressed genes were considered statistically significant if the FDR-adjusted p-value was less than 0.05. To select maximally specific genes per cell type/subset, we ran MAST for all pairwise cluster comparisons (each versus the rest) and retained the top 10 DEGs for a given cluster (FDR<=0.05, sorted by decreasing coefficient). To generate the unique cluster signatures per se, we used the *sc.tl.score_genes*() function from *scanpy* [78] that calculates averaged scores based on the cluster specific genes (scores are subtracted with a randomly sampled reference gene set).

iv) Differential abundance (DA) analysis

DA HSC subset across age groups was identified by sampling neighborhoods of cells from a k-nearest neighbors (k-NN) graph and looking for enrichment of either age in each neighborhood as implemented in MiloR.[79] MiloR is a graph-based statistical method to compute differential cellular abundances in neighborhoods of cells. The 15 batch-corrected latent dimensions from *scanpy* were used for MiloR (v.0.99.19) k-NN graph construction (k = 35) and neighborhood indexing (proportion = 0.1). DA testing was performed with generalized linear models, including age as covariates (neighborhoods significant if spatial corrected FDR < 0.25).

Single- Cell ATAC sequencing

Preprocessing, dimensionality reduction, clustering

Single-cell ATAC-seq from young and old mice were aligned to the mm10 genome and we processed the cellranger output file, fragments.tsv, with ArchR (v.1.0.2)[80] was used for downstream analysis. We performed QC filtering on scATAC-seq using ArchR with the default parameters of *createArrowFiles*() (**Fig. S5**). We retained cells containing at least 1,000 and at most 100,000 fragments. Next, we filtered out the cells that did not pass QC in the corresponding scRNA-seq data, thus retaining 11493 cells common to both scRNA-seq and scATAC-seq. We then performed dimensionality reduction using iterative latent semantic indexing (LSI) on the top 25,000 variable features from the tile matrix to get a reduced dimensionality of 30 components with *addIterativeLSI*() function in ArchR. To generate visualizations, we employed the *addUMAP*() function in ArchR with the following settings: nNeighbors = 20; minDist = 0.1. For clustering, we utilized the *addClusters*() function in ArchR, specifying the

parameters as follows: method = 'Seurat'; knnAssign = 10, and maxClusters = 10. Next, we used Harmony to perform batch effect correction.

Peak-calling and TF motif accessibility scoring

Each individual sample was pseudobulked for peak calling with MACS2[81] peak caller and iterative peak overlapping removal within ArchR (using default settings). We subsequently added motif annotations using *addMotifAnnotations*() with the CisBP motif database and computed chromVAR[82] deviations for each single cell with *addDeviationsMatrix*(). To identify differentially accessible motifs within each group of interest, we applied the *rank_genes_groups*() function in *scanpy*. With the following settings: method = 'wilcoxon' and corr_method = 'benjamini-hochberg' and performed the analysis on the chromVAR zscore matrix.

Processing Human Single-Cell CITE-seq

We analyzed the young human bone marrow CITE-seq from Sommarin et al.[62] using Seurat.[83] Authors provided the peak matrix aligned to the hg38 genome. We used the sample BM_34 for which the authors provided the cell type annotations. We processed the scRNA-seq and CITE-seq data for young and old bone marrow using Seurat. The hash-tagged cells were demultiplexed using *HTODemux*() from Seurat. The cell type annotations were provided by the authors for young bone marrow samples labeled yBM1_hpc and yBM2_hpc. We predicted cell type annotations for the old bone marrow cells by label transfer using *FindTransferAnchors*() and *TransferData*() functions. We obtained the Kit^{lo} gene signature from our mouse scRNAseq data through differential gene expression analysis. To extend this signature to our human bone

marrow HSPC dataset, we used Ensembl[84] Biomart to identify human orthologues of the mouse genes. Subsequently, we scored the Kit^{lo} gene signature in human scBM HSCs using Seurat's (AddModuleScore) function.

Data and Code Availability

Multiome single cell RNA-and- ATAC sequencing data generated for this paper is available upon request from the corresponding author.

Statistical Analysis

Data were processed in GraphPad Prism 10.0 software. Statistical comparisons between 2 groups were performed with either the nonparametric unpaired Mann-Whitney U test, two-way ANOVA, ComBat batch effect correction[85] or Wilcoxon rank sum test with continuity correction as indicated.

ACKNOWLEDGEMENTS

We thank Dr. Andrea Schietinger for the gift of mice; Ronan Chaligné, Kalina Belcheva, and Anthony Michaels for help; Dr. Gay Crooks, and Julia Gerscheimer (UCLA) for sharing the mATO and huATO systems and assistance with troubleshooting; Kenneth Dorshkind (UCLA) for providing S17 cells; Parashar Dhopala and Göran Karlsson (Lund University) for providing the human BM CITE-seq raw matrices.[62]

FUNDING SOURCES

This research was supported by National Cancer Institute award numbers, R35-CA284024, P01-CA023766, R01-CA228308, and P30 CA008748 MSK Cancer Center Support Grant/Core

Grant; National Heart, Lung, and Blood Institute (NHLBI) award number R01-HL164902; National Institute on Aging award number P01-AG052359, and Tri Institutional Stem Cell Initiative. Additional funding was received from The Lymphoma Foundation, The Susan and Peter Solomon Family Fund, The Solomon Microbiome Nutrition and Cancer Program, Cycle for Survival, Parker Institute for Cancer Immunotherapy, Paula and Rodger Riney Multiple Myeloma Research Initiative, Starr Cancer Consortium, and Seres Therapeutics. HKE receives research support from NHLBI. SM receives research support from Marie-Josée Kravis fellowship in quantitative biology. SG receives research support from CRI / Donald J. Gogel Postdoctoral Fellowship (CRI #3934). SDW reports research funding from the MSK Leukemia SPORE Career Enhancement Program (NIH/NCI P50 CA254838-01), the MSK Gerstner Physician Scholar Program, and the Parker Institute for Cancer Immunotherapy. MS receives research support from the following funding sources: Burroughs Wellcome Fund (PDEP), Damon Runyon Cancer Research Foundation (Clinical Investigator Award), V Foundation (V Scholar), Emerson Collective, and the National Institutes of Health (NHLBI K08 HL156082-01A1). KA receives research support from the Gerstner Physician Scholar program.

AUTHOR CONTRIBUTIONS

HKE, MRMvdB conceived and designed the study. HKE, MRMvdB acquired funding. HKE, MBdS, AR, BG, NL, MAEA, SG planned research activities. HKE curated data and performed initial analysis. AIK and SM performed initial computational analysis of single-cell RNA- seq data, and single-cell ATAC- seq and human CITE-seq data, respectively. RS provided code and assisted HKE with downstream data/sequence computation and statistical analyses. HKE analyzed flow cytometric data. HKE, MBdS, AR, BG, NL, MAEA, XZ, HA, SG, MS, SDW, KV, CYP, CED, and JCS planned and performed experiments. CSL, AB, MRMvdB supervised

study. HKE, SM, and MBdS prepared figures and visualized data. HKE, MRMvdB wrote the initial draft. All authors reviewed and approved the manuscript.

DECLARATION OF INTERESTS

H.K.Elias, and M.R.M. van den Brink are co-inventors on patent applications related to this work. MRMvdB has received research support and stock options from Seres Therapeutics and stock options from Notch Therapeutics and Pluto Therapeutics; he has received royalties from Wolters Kluwer; has consulted, received honorarium from or participated in advisory boards for Seres Therapeutics, WindMIL Therapeutics, Rheos Medicines, Merck & Co, Inc., Magenta Therapeutics, Frazier Healthcare Partners, Nektar Therapeutics, Notch Therapeutics, Forty Seven Inc., Priothera, Ceramedix, Lygenesis, Pluto Therapeutics, GlaskoSmithKline, Da Volterra, Vor BioPharma, Novartis (Spouse), Synthekine (Spouse), and Beigene (Spouse); he has IP Licensing with Seres Therapeutics and Juno Therapeutics; and holds a fiduciary role on the Foundation Board of DKMS (a nonprofit organization). The Walter and Eliza Hall Institute of Medical Research receives milestone and royalty payments related to venetoclax. Employees are entitled to receive benefits related to these payments.

REFERENCES

1. Hale JS, Boursalian TE, Turk GL, Fink PJ. Thymic output in aged mice. *Proc Natl Acad Sci U S A*. 2006;103(22):8447-8452.
2. van den Brink M, Uhrberg M, Jahn L, DiPersio JF, Pulsipher MA. Selected biological issues affecting relapse after stem cell transplantation: role of T-cell impairment, NK cells and intrinsic tumor resistance. *Bone Marrow Transplant*. 2018;53(8):949-959.
3. de Koning C, Prockop S, van Roessel I, et al. CD4+ T-cell reconstitution predicts survival outcomes after acute graft-versus-host-disease: a dual-center validation. *Blood*. 2021;137(6):848-855.

4. van Roessel I, Prockop S, Klein E, et al. Early CD4+ T cell reconstitution as predictor of outcomes after allogeneic hematopoietic cell transplantation. *Cytotherapy*. 2020;22(9):503-510.
5. Zlotoff DA, Zhang SL, De Obaldia ME, et al. Delivery of progenitors to the thymus limits T-lineage reconstitution after bone marrow transplantation. *Blood*. 2011;118(7):1962-1970.
6. van Ewijk W, Shores EW, Singer A. Crosstalk in the mouse thymus. *Immunol Today*. 1994;15(5):214-217.
7. Poliani PL, Facchetti F, Ravanini M, et al. Early defects in human T-cell development severely affect distribution and maturation of thymic stromal cells: possible implications for the pathophysiology of Omenn syndrome. *Blood*. 2009;114(1):105-108.
8. Pala F, Oguz C, Corsino C, et al. Profound abnormalities in thymic epithelial cells in Rag1 hypomorphic mice: implications for immune reconstitution after stem cell transplantation. *The Journal of Immunology*. 2022;208(1_Supplement):167.102-167.102.
9. Elias HK, Bryder D, Park CY. Molecular mechanisms underlying lineage bias in aging hematopoiesis. *Semin Hematol*. 2017;54(1):4-11.
10. Rodriguez-Fraticelli AE, Weinreb C, Wang SW, et al. Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. *Nature*. 2020;583(7817):585-589.
11. Dong F, Hao S, Zhang S, et al. Differentiation of transplanted haematopoietic stem cells tracked by single-cell transcriptomic analysis. *Nat Cell Biol*. 2020;22(6):630-639.
12. Carrelha J, Meng Y, Kettyle LM, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*. 2018;554(7690):106-111.
13. Meng Y, Carrelha J, Drissen R, et al. Epigenetic programming defines haematopoietic stem cell fate restriction. *Nat Cell Biol*. 2023;25(6):812-822.
14. Morita Y, Ema H, Nakauchi H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med*. 2010;207(6):1173-1182.
15. Yamamoto R, Wilkinson AC, Ooehara J, et al. Large-Scale Clonal Analysis Resolves Aging of the Mouse Hematopoietic Stem Cell Compartment. *Cell Stem Cell*. 2018;22(4):600-607 e604.
16. Benz C, Copley MR, Kent DG, et al. Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell*. 2012;10(3):273-283.
17. Challen GA, Boles NC, Chambers SM, Goodell MA. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell*. 2010;6(3):265-278.
18. Naik SH, Perie L, Swart E, et al. Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature*. 2013;496(7444):229-232.
19. Yu VWC, Yusuf RZ, Oki T, et al. Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell*. 2016;167(5):1310-1322 e1317.
20. Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*. 1997;91(5):661-672.
21. Flohr Svendsen A, Yang D, Kim K, et al. A comprehensive transcriptome signature of murine hematopoietic stem cell aging. *Blood*. 2021;138(6):439-451.
22. Grover A, Sanjuan-Pla A, Thongjuea S, et al. Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat Commun*. 2016;7:11075.

23. Herault L, Poplineau M, Mazuel A, Platet N, Remy E, Duprez E. Single-cell RNA-seq reveals a concomitant delay in differentiation and cell cycle of aged hematopoietic stem cells. *BMC Biol*. 2021;19(1):19.
24. Kim YW, Zara G, Kang H, et al. Integration of single-cell transcriptomes and biological function reveals distinct behavioral patterns in bone marrow endothelium. *Nat Commun*. 2022;13(1):7235.
25. Konturek-Ciesla A, Dhapola P, Zhang Q, et al. Temporal multimodal single-cell profiling of native hematopoiesis illuminates altered differentiation trajectories with age. *Cell Rep*. 2023;42(4):112304.
26. Safi F, Dhapola P, Warsi S, et al. Concurrent stem- and lineage-affiliated chromatin programs precede hematopoietic lineage restriction. *Cell Rep*. 2022;39(6):110798.
27. Giladi A, Paul F, Herzog Y, et al. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat Cell Biol*. 2018;20(7):836-846.
28. Cabezas-Wallscheid N, Buettner F, Sommerkamp P, et al. Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell*. 2017;169(5):807-823 e819.
29. Lauridsen FKB, Jensen TL, Rapin N, et al. Differences in Cell Cycle Status Underlie Transcriptional Heterogeneity in the HSC Compartment. *Cell Rep*. 2018;24(3):766-780.
30. Cabezas-Wallscheid N, Klimmeck D, Hansson J, et al. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell*. 2014;15(4):507-522.
31. Wahlestedt M, Ladopoulos V, Hidalgo I, et al. Critical Modulation of Hematopoietic Lineage Fate by Hepatic Leukemia Factor. *Cell Rep*. 2017;21(8):2251-2263.
32. Komorowska K, Doyle A, Wahlestedt M, et al. Hepatic Leukemia Factor Maintains Quiescence of Hematopoietic Stem Cells and Protects the Stem Cell Pool during Regeneration. *Cell Rep*. 2017;21(12):3514-3523.
33. Lehnertz B, Chagraoui J, MacRae T, et al. HLF expression defines the human hematopoietic stem cell state. *Blood*. 2021;138(25):2642-2654.
34. Shin JY, Hu W, Naramura M, Park CY. High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. *J Exp Med*. 2014;211(2):217-231.
35. Grinenko T, Arndt K, Portz M, et al. Clonal expansion capacity defines two consecutive developmental stages of long-term hematopoietic stem cells. *J Exp Med*. 2014;211(2):209-215.
36. Elias HK, Shin JY, Ali MA, et al. Identification of Old HSCs with Preserved Self-Renewal and Long-Term Reconstitution Potential. *Blood*. 2019;134:2479.
37. Montecino-Rodriguez E, Kong Y, Casero D, Rouault A, Dorshkind K, Pioli PD. Lymphoid-Biased Hematopoietic Stem Cells Are Maintained with Age and Efficiently Generate Lymphoid Progeny. *Stem Cell Reports*. 2019;12(3):584-596.
38. Inlay MA, Bhattacharya D, Sahoo D, et al. Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev*. 2009;23(20):2376-2381.
39. Seet CS, He C, Bethune MT, et al. Generation of mature T cells from human hematopoietic stem and progenitor cells in artificial thymic organoids. *Nat Methods*. 2017;14(5):521-530.
40. Montel-Hagen A, Sun V, Casero D, et al. In Vitro Recapitulation of Murine Thymopoiesis from Single Hematopoietic Stem Cells. *Cell Rep*. 2020;33(4):108320.

41. Zhou W, Gao F, Romero-Wolf M, Jo S, Rothenberg EV. Single-cell deletion analyses show control of pro-T cell developmental speed and pathways by Tcf7, Spi1, Gata3, Bcl11a, Erg, and Bcl11b. *Sci Immunol*. 2022;7(71):eabm1920.
42. Zlotoff DA, Sambandam A, Logan TD, Bell JJ, Schwarz BA, Bhandoola A. CCR7 and CCR9 together recruit hematopoietic progenitors to the adult thymus. *Blood*. 2010;115(10):1897-1905.
43. Dudakov JA, Hanash AM, Jenq RR, et al. Interleukin-22 drives endogenous thymic regeneration in mice. *Science*. 2012;336(6077):91-95.
44. Lopes N, Vachon H, Marie J, Irla M. Administration of RANKL boosts thymic regeneration upon bone marrow transplantation. *EMBO Mol Med*. 2017;9(6):835-851.
45. Yu W, Nagaoka H, Jankovic M, et al. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature*. 1999;400(6745):682-687.
46. Goldberg GL, Sutherland JS, Hammet MV, et al. Sex steroid ablation enhances lymphoid recovery following autologous hematopoietic stem cell transplantation. *Transplantation*. 2005;80(11):1604-1613.
47. Goldberg GL, Alpdogan O, Muriglan SJ, et al. Enhanced immune reconstitution by sex steroid ablation following allogeneic hemopoietic stem cell transplantation. *J Immunol*. 2007;178(11):7473-7484.
48. Dixit VD, Yang H, Sun Y, et al. Ghrelin promotes thymopoiesis during aging. *J Clin Invest*. 2007;117(10):2778-2790.
49. Spadaro O, Youm Y, Shchukina I, et al. Caloric restriction in humans reveals immunometabolic regulators of health span. *Science*. 2022;375(6581):671-677.
50. Mansson R, Hultquist A, Luc S, et al. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity*. 2007;26(4):407-419.
51. Shin B, Hosokawa H, Romero-Wolf M, et al. Runx1 and Runx3 drive progenitor to T-lineage transcriptome conversion in mouse T cell commitment via dynamic genomic site switching. *Proc Natl Acad Sci U S A*. 2021;118(4).
52. Maeda T, Merghoub T, Hobbs RM, et al. Regulation of B versus T lymphoid lineage fate decision by the proto-oncogene LRF. *Science*. 2007;316(5826):860-866.
53. Siggs OM, Li X, Xia Y, Beutler B. ZBTB1 is a determinant of lymphoid development. *J Exp Med*. 2012;209(1):19-27.
54. Punwani D, Simon K, Choi Y, et al. Transcription factor zinc finger and BTB domain 1 is essential for lymphocyte development. *J Immunol*. 2012;189(3):1253-1264.
55. Zhang X, Lu Y, Cao X, Zhen T, Kovalovsky D. Zbtb1 prevents default myeloid differentiation of lymphoid-primed multipotent progenitors. *Oncotarget*. 2016;7(37):58768-58778.
56. Koizumi M, Kama Y, Hirano KI, et al. Transcription factor Zbtb1 interacts with bridging factor Lmo2 and maintains the T-lineage differentiation capacity of lymphoid progenitor cells. *J Biol Chem*. 2022;298(11):102506.
57. Kim H, Dejsuphong D, Adelmant G, et al. Transcriptional repressor ZBTB1 promotes chromatin remodeling and translesion DNA synthesis. *Mol Cell*. 2014;54(1):107-118.
58. Yatim A, Benne C, Sobhian B, et al. NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. *Mol Cell*. 2012;48(3):445-458.
59. Min H, Montecino-Rodriguez E, Dorshkind K. Reduction in the developmental potential of intrathymic T cell progenitors with age. *J Immunol*. 2004;173(1):245-250.

60. Srinivasan J, Vasudev A, Shasha C, et al. The initial age-associated decline in early T-cell progenitors reflects fewer pre-thymic progenitors and altered signals in the bone marrow and thymus microenvironments. *Aging Cell*. 2023.
61. Cordes M, Cante-Barrett K, van den Akker EB, et al. Single-cell immune profiling reveals thymus-seeding populations, T cell commitment, and multilineage development in the human thymus. *Sci Immunol*. 2022;7(77):eade0182.
62. Sommarin MNE, Dhapola P, Safi F, et al. Single-Cell Multiomics Reveals Distinct Cell States at the Top of Human Hematopoietic Hierarchy. *bioRxiv*. 2021:2021.2004.2001.437998.
63. Karamitros D, Stoilova B, Aboukhalil Z, et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat Immunol*. 2018;19(1):85-97.
64. Notta F, Zandi S, Takayama N, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*. 2016;351(6269):aab2116.
65. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*. 2017;19(4):271-281.
66. Dykstra B, Kent D, Bowie M, et al. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell*. 2007;1(2):218-229.
67. Paiva RA, Sousa AGG, Ramos CV, et al. Self-renewal of double-negative 3 early thymocytes enables thymus autonomy but compromises the beta-selection checkpoint. *Cell Rep*. 2021;35(2):108967.
68. Maillard I, Schwarz BA, Sambandam A, et al. Notch-dependent T-lineage commitment occurs at extrathymic sites following bone marrow transplantation. *Blood*. 2006;107(9):3511-3519.
69. Cao X, Lu Y, Zhang X, Kovalovsky D. Zbtb1 Safeguards Genome Integrity and Prevents p53-Mediated Apoptosis in Proliferating Lymphoid Progenitors. *J Immunol*. 2016;197(4):1199-1211.
70. Cosgun KN, Rahmig S, Mende N, et al. Kit regulates HSC engraftment across the human-mouse species barrier. *Cell Stem Cell*. 2014;15(2):227-238.
71. Kawashima I, Zanjani ED, Almaida-Porada G, Flake AW, Zeng H, Ogawa M. CD34+ human marrow cells that express low levels of Kit protein are enriched for long-term marrow-engrafting cells. *Blood*. 1996;87(10):4136-4142.
72. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep*. 2019;9(1):5233.
73. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst*. 2019;8(4):281-291 e289.
74. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16(12):1289-1296.
75. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nature methods*. 2014;11(7):740-742.
76. van Dijk D, Sharma R, Nainys J, et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell*. 2018;174(3):716-729 e727.
77. Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol*. 2015;16:278.
78. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1):15.

79. Dann E, Henderson NC, Teichmann SA, Morgan MD, Marioni JC. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat Biotechnol*. 2022;40(2):245-253.
80. Granja JM, Corces MR, Pierce SE, et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet*. 2021;53(3):403-411.
81. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
82. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 2017;14(10):975-978.
83. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587 e3529.
84. Cunningham F, Allen JE, Allen J, et al. Ensembl 2022. *Nucleic Acids Res*. 2022;50(D1):D988-D995.
85. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8(1):118-127.
86. Fang Z, Liu X, Peltz G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics*. 2023;39(1).

FIGURE LEGENDS

Figure 1: Aged HSCs Exhibit Loss of Multilineage Potential

(**A-H**) Multiome single-cell RNA and ATAC sequencing was performed on HSC (Lineage-Sca-1+cKit+ CD34-CD48-Flt3-CD150+) cells isolated from 2-mo (young) or 24-mo (old) female C57BL/6 mice. nTOTAL =12,350 cells. (**A**) Volcano plot of DGE analysis between MLin-HSC vs. q-HSC subsets in young HSCs, highlighting bonafide HSC markers. (**B**) Violin plot for imputed *Kit* gene expression by annotated HSC subsets (as described in **Fig. S2C**. (**C**) Frequency of computationally defined HSC subtypes in Kit^{hi} and Kit^{lo} subsets in young HSCs. (**D-E**) Combined Uniform manifold approximation and projection (UMAP) of young and old HSCs after Harmony batch correction annotated by age **(D)** and annotated HSC subsets (q-HSC: Quiescent HSCs; Mgk-HSCs: Platelet-biased HSCs; Mlin-HSC: Multilineage HSCs; p-HSC: Proliferative HSCs; Int-HSC: Intermediate HSCs) **(E)**. **(F)** Frequency of HSC subtypes in young and old HSCs. **(G)** Heatmap showing gene expression of bonafide HSC markers ordered by

increasing Kit expression. **(H)** Violin plot for imputed Kit gene expression by age. Statistical analysis was performed using the Wilcoxon test. **(I)** Scaled change in frequency for each Kit HSC subset with age. **(J)** FACS analysis showing frequency of Kit^{hi} and Kit^{lo} HSCs by age. All data are from n=6 mice/group. Error bars represent mean ± SEM. **P<0.01, ***P<0.001. P values calculated by two-way ANOVA.

Figure 2: Young Kit^{lo} HSCs Exhibit Enhanced *In Vitro* and *In Vivo* Lymphoid Potential

(A) Experimental schema to evaluate *in vitro* lymphoid progenitor and T-cell differentiation potential of Kit^{hi} (red) and Kit^{lo} HSCs (blue) from 2-mo (young) C57BL/6 mice using the S17 lymphoid assay and murine artificial thymic organoid (mATO) respectively. Enumeration of absolute number of lymphoid progenitor cells following 14 days in S17 lymphoid assay (B) and absolute number of T-cell subsets following 8-weeks of culture in mATOs **(C).** Aggregated data across 4-5 independent experiments. **(D)** Experimental schema for competitive allogeneic HCT (allo-HCT) using Kit^{hi} (red) and Kit^{lo} HSCs (blue) from 2-mo (young) C57BL/6 mice with competitor bone marrow (BM) cells from B6.SJL-PtprcaPepcb/BoyJ mice transplanted into lethally irradiated 7-week-old (young) BALB/cJ recipients. **(E-K)** Eight weeks after competitive HCT, **(E)** frequency of donor-derived chimerism (CD45.2/ H-2Kb) of mature lineages in the peripheral blood (PB, CD45+, Myeloid cell: Gr1+CD11b+; B-cell: B220+; T-cell: CD3+), (**F-G)** Enumeration of absolute number of donor-derived cells in the BM for LSK cells (LT HSC: $Lineage^{-}Sca\text{-}1^{+}cKit^{+}$ $CD34^{-}CD48^{-}Flt3^{-}CD150^{+}$ **;** ST HSC: $Lineage^{-}Sca\text{-}1^{+}cKit^{+}$ $CD48^{-}Flt3^{-}$ $CD150^{-}$**;** MPP2: $Lineage^{-}Sca\text{-}1^{+}cKit^{+}$ $CD48^{+}Flt3^{-}CD150^{-}$; MPP3: $Lineage^{-}Sca\text{-}1^{+}cKit^{+}$ $CD48^{+}Flt3^{-}CD150^{+}$; MPP4/ LMPP: $Lineage^{-}Sca\text{-}1^{+}cKit^{+}$ $Flt3^{+}CD150^{-}$) **(F),** Common Lymphoid Progenitors cells (CLP: $Lineage^{-}IL7Ra^{+}Flt3^{+}Sca^{mid/lo}$ Kit^{lo}) **(G)**. **(H-K)** Post-HCT thymi analysis

for total thymic cellularity **(H)**, **(I-J)** enumeration of absolute number of donor-derived cells for T-cell precursors (ETP: Lineage$^-$ CD4$^-$ CD8$^-$CD44$^+$ CD25$^-$Kit$^+$;DN2: Lineage$^-$ CD4$^-$ CD8$^-$CD44$^+$ CD25$^+$; DN3: Lineage$^-$ CD4$^-$ CD8$^-$CD44$^-$ CD25$^+$) **(I)**, Mature T-cells (DP: Lineage$^-$ CD4$^+$ CD8$^+$; SP4: Lineage$^-$ CD4$^+$ CD8$^-$; SP8: Lineage$^-$ CD4$^-$ CD8$^+$**) (J),** and analysis of thymic CD45- compartment for Thymic Epithelial Cells (TEC: CD45$^-$ EpCAM$^+$), Cortical TEC (cTEC: CD45$^-$ EpCAM$^+$UEA-1^{lo} 6C3^{hi} MHCII$^{hi/lo}$), and Medullary TEC (mTEC: CD45$^-$ EpCAM$^+$UEA-1^{hi} 6C3^{lo} MHCII$^{hi/lo-}$) (**K**). All data are from n=10-11 mice/group. **(L)** Experimental schema to evaluate thymic function following competitive allo-HCT using young HSCs. **(M)** Frequency of donor-derived Recent Thymic Emigrants (RTE: CD3+GFP+) in the PB at 8 weeks post-HCT. All data are from n=5-6 mice/group. **(N)** Experimental schema to investigate functional response of differential Kit-expressing HSC-derived T-cells in secondary recipients to *L. monocytogenes* expressing chicken ovalbumin (LM-OVA) infection. **(O)** Spleen analysis of secondary recipients for frequency of OT1-derived chimerism of CD8+ T cells (OT1+CD8+: Va2$^+$Vb5$^+$H2-K^{b}CD8$^+$) 7 days post-infection. All data are from n=9-10 mice/group. **(P)** Experimental schema for competitive allogeneic HCT (allo-HCT) using Kit^{hi} (red) and Kit^{lo} HSCs (blue) from 2-mo (young) C57BL/6 mice with competitor bone marrow (BM) cells from B6.SJL-PtprcaPepcb/BoyJ mice transplanted into lethally irradiated 14-mo (Middle-aged) BALB/cJ recipients. **(Q)** Frequency of donor-derived T-cell chimerism at the indicated time-points. **(R-T)** Sixteen weeks after competitive HCT enumeration of total thymic cellularity **(R)**, donor-derived T-cell precursor thymocyte subsets as defined in Figure 3F **(S)**, donor-derived mature T-cell thymocyte subsets as defined in Figure 3G **(T).** All data are from n=6-7 mice/group. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. P values calculated by

nonparametric unpaired Mann-Whitney U test. Panels **A**, **D**, **L**, **N**, **and P** were created using BioRender.

Figure 3: Kit^{lo} HSCs are Epigenetically Programmed for Lymphoid Potential

(A) UMAP of 6 unsupervised HSC clusters identified by tile-based scATAC-seq analysis after Harmony batch correction. **(B-D)** Annotated by age **(B)**, HSC subsets **(C)**, and Kit subsets **(D)** on the scATAC-seq UMAP. **(E)** scATAC-seq cluster composition by HSC subsets (**left**) and Kit subsets (**right**). (**F-G**) Matrix plots showing motif enrichment identified by ChromVAR (**F**) and gene accessibility (**G**) for lymphoid-specifying transcription factors for Kit subsets in young and old HSCs.

Figure 4: *Zbtb1* regulates T-cell potential in phenotypic HSCs

(A) Experimental schema for T-cell differentiation assay and competitive allogeneic HCT with *Zbtb1*-deficient Kit^{lo} HSCs generated with 2-mo young Rosa26Cas9-eGFP KI mice. **(B)** Frequency of T-cell subsets following 6-weeks of culture in mATOs. Aggregated data across 3 independent experiments. (C) Frequency of donor-derived lineage chimerism of control or *Zbtb1* KO Kit^{lo} HSCs in the PB at 6 weeks following competitive HCT. All data are from n=5-6 mice/group. (D) Experimental schema for T-cell differentiation assay and competitive allogeneic HCT with *Zbtb1*-OE Kit^{hi} HSCs from young mice. (E) Frequency of T-cell subsets following 6-weeks of culture in mATO of young Kit^{hi} HSCs. Aggregated data across 3 independent experiments. (F) Frequency of donor-derived lineage chimerism of control or *Zbtb1* OE young Kit^{hi} HSCs in the PB at 6-weeks following competitive HCT. All data are from n=5-6 mice/group. (G) Previously identified Zbtb1-targets[56] and Notch1-interacting proteins[58] enriched

in Kit^{lo} HSCs. (H) Pathway enrichment analysis for non-Notch1 Zbtb1-targets enriched in Kit^{lo} HSCs by *gseapy*()[86] using GO_Biological_Process_2023 libraries. Bubbleplot showing representative pathways. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001. P values calculated by nonparametric unpaired Mann-Whitney U test. Panels **A** and **D** were created using BioRender.

Figure 5: Old Kit^{lo} HSCs Exhibit Enhanced *In Vitro* and *In Vivo* Lymphoid Potential

(A-B) Evaluation of *in vitro* lymphoid progenitor and T-cell differentiation potential of Kit^{hi} (red) and Kit^{lo} HSCs (blue) from or 24-mo (old) C57BL/6 mice using the S17 lymphoid assay and murine artificial thymic organoid (mATO) respectively, as shown in Figure 2A. Enumeration of absolute number of lymphoid progenitor cells following 14 days in S17 lymphoid assay **(A)** and absolute number of T-cell subsets following 8-weeks of culture in mATOs **(B).** Aggregated data across 4-5 independent experiments. (**C)** Experimental schema for competitive allogeneic HCT (allo-HCT) using Kit^{hi} (red) and Kit^{lo} HSCs (blue) from 22–24-mo (old) C57BL/6 mice with competitor bone marrow (BM) cells from B6.SJL-PtprcaPepcb/BoyJ mice transplanted into lethally irradiated young BALB/cJ recipients. **(D-I)** Eight weeks after competitive HCT enumeration of absolute number of donor-derived T-cells in the PB **(D)**, LSK cell subsets as defined in Figure 3C (**E)**, CLP cells as defined in Figure 3D **(F)**, total thymic cellularity **(G)**, donor-derived T-cell precursor thymocyte subsets as defined in Figure 3F **(H)**, TEC compartment as defined in Figure 3H **(I).** All data are from n=9-10 mice/group. (**J)** Experimental schema to evaluate thymic function following competitive allo-HCT using old HSCs. **(K)** Frequency of donor-derived Recent Thymic Emigrants (RTE: CD3^{+}GFP^{+}) in the PB at 8 weeks post-HCT. All data are from n=5 mice/group. (**L)** Experimental schema for

competitive allogeneic HCT (allo-HCT) using Kit^{hi} (red) and Kit^{lo} HSCs (blue) from 22-24-mo (old) C57BL/6 mice with competitor bone marrow (BM) cells from B6.SJL-PtprcaPepcb/BoyJ mice transplanted into lethally irradiated middle-aged BALB/cJ recipients. **(M-O)** Eight weeks after competitive HCT enumeration of absolute number of donor-derived T cells in PB **(M)**, donor-derived CLPs in the BM as defined in Figure 3D **(N)**, and thymic cellularity **(O)**. All data are from n=5-6 mice/group. (**P**) Frequency of T-cell subsets following 6-weeks of culture in mATO of old Kit^{hi} HSCs following *Zbtb1* OE. Aggregated data across 3 independent experiments. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. P values calculated by nonparametric unpaired Mann-Whitney U test. Panels **C, J,** and **L** were created using BioRender.

Figure 6 KIT^{lo} Human BM HSC Subsets show Higher Lymphoid Potential

(A-E) Young and old human BM CITE-seq dataset generated and published by Sommarin et al.[62] **(A-C)** UMAP of CD34+ young and old human BM annotated with HSC subsets **(A)**, and age **(B)**. **(C)** HSC cluster composition by age. **(D)** Kit^{lo} gene signature (top 50 marker genes from our mouse data converted to human orthologues) was overlaid on human HSC UMAP. **(E)** Violin plot for Kit^{lo} gene score by age. Statistical analysis was performed using the Wilcoxon test. **(F-G)** FACS analysis of human BM samples. **(F)** Violin plot of KIT protein expression on phenotypic HSCs (pHSC: CD34+CD38-CD10-CD45RA-CD90+) by age (young: 20-40yrs; MA/old: >40yrs). **(G)** Scatterplot showing distribution of KIT^{hi} and KIT^{lo} HSC subsets within pHSCs in human BM samples with varying age. **(H)** Experimental schema to evaluate *in vitro* T-cell and lympho-myeloid potential of human BM HSC subsets based on differential KIT expression using human artificial thymic organoids (hATO) and MS5 assay respectively. **(I)**

Following 8-10-weeks of culture, enumeration of CD19+ B-cells and T-cell subsets within CD34+ T-cell precursors (Early Thymic Progenitors, ETP: $CD34^{+}CD1a^{-}CD7^{-}$; $CD1a^{neg}$- ProT: $CD34^{+}CD1a^{-}CD7^{+}$; $CD1a^{pos}$-ProT: $CD34^{+}CD1a^{+}CD7^{+}$), and mature T-cells (DP: $CD34^{-}CD5^{+}CD7^{+}CD4^{+}CD8^{+}$; SP4: $CD34^{-}CD5^{+}CD7^{+}CD4^{+}CD8^{-}$; SP8: $CD34^{-}CD5^{+}CD7^{+}CD4^{-}CD8^{+}$). Aggregated data represent mean from 8 independent BM donors ± SEM. **(J)** Following 4-weeks of culture, enumeration of Lineage output (Neutrophil-Granulocytes: hCD45+CD15+; monocyte-macrophages: hCD45+CD14+; B-cells: hCD45+CD19+; NK-cells: hCD45+CD56+) after culturing 200 KIT^{lo} and KIT^{hi} HSC cells on MS-5 stroma, as previously described[63]. Aggregated data represent mean from 5 independent BM donors ± SEM. **(L)** Model illustrating age-related loss of lymphoid potential is attributed to a decreased in the frequency of Kit^{lo} HSCs in aged BM. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. P values for panel **G** was based on generalized estimating equation logistic model adjusted for age, KIT^{lo} and KIT^{hi} groups and their interaction, the matched data structure captured by an exchangeable working correlation structure. P values calculated either by nonparametric paired Mann-Whitney U test (Panels **I** and **J**), unpaired Mann-Whitney U test (Panel **F**). Panels **H** and **L** were created using BioRender.

Supplemental Figure Legends

Figure S1: (A) Experimental schema for competitive allogeneic HCT (allo-HCT) using 2-mo (young) and 24-mo (old) 750 HSCs [$Lin^{-}c\text{-}Kit^{+}Sca1^{+}$ ($L^{-}S^{+}K^{+}$) $CD150^{+}$ $CD34^{-}CD48^{-}$] from C57BL/6 mice with competitor bone marrow (BM) cells from B6.SJL-PtprcaPepcb/BoyJ mice transplanted into lethally irradiated 7-week-old (young) BALB/cJ recipients. **(B-E)** Frequency of donor-derived chimerism of all hematopoietic cells mature lineages **(B),** T-cell **(C),** Myeloid **(D),**

B-cell **(E)** in the PB at the indicated time points. **(F-G)** Eight weeks after competitive HCT enumeration of absolute number of LMPP cells (LMPP/MPP4: Lineage$^-$Sca-1$^+$cKit$^+$ Flt3$^+$CD150$^-$) (**F**), CLP cells (CLP: Lineage$^-$IL7Ra$^+$Flt3$^+$Sca$^{mid/lo}$ Kit^{lo}) **(G)**. **(H-J)** Post-HCT thymi analysis for total thymic cellularity **(H)**, enumeration of absolute number of donor-derived cells for T-cell precursors (ETP: Lineage$^-$ CD4$^-$ CD8$^-$CD44$^+$ CD25$^-$Kit$^+$; DN2: Lineage$^-$ CD4$^-$ CD8$^-$CD44$^+$ CD25$^+$; DN3: Lineage$^-$ CD4$^-$ CD8$^-$CD44$^-$ CD25$^+$) **(I)**, Mature T-cells (DP: Lineage$^-$ CD4$^+$ CD8$^+$; SP4: Lineage$^-$ CD4$^+$ CD8$^-$; SP8: Lineage$^-$ CD4$^-$ CD8$^+$**)** **(J)**. All data are from n=9-10 mice/group. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. P values calculated by nonparametric unpaired Mann-Whitney U test. Panel **A** was created using BioRender.

Figure S2: (A) UMAP of young HSCs after Harmony batch correction annotated by unsupervised Leiden cluster analysis (**B**), manually annotated HSC subsets (q-HSC: Quiescent HSCs; Mgk-HSCs: Platelet-biased HSCs; Mlin-HSC: Multilineage HSCs; p-HSC: Proliferative HSCs; Int-HSC: Intermediate HSCs) (**C**). (**D**) Dot plot showing marker genes for each HSC subset. Circle size and color indicate percentage and expression level, respectively. (**E**) Composition of defined HSC subtypes in Kit^{hi} and Kit^{lo} subsets in young HSCs as shown by scaled change in frequency.

(**F**) UMAP of old HSCs annotated by HSC subsets identified by *scanpy ingest*()[78] using young HSCs, as reference dataset. **(G)** UMAP of combined young and old HSCs after Harmony batch correction annotated by sample. (**H**) Composition of defined HSC subtypes in old and young HSCs as shown by scaled change in frequency. **(I)** A neighborhood graph of the results from Milo differential abundance testing. Nodes are neighborhoods, colored by their log fold change

across ages. Non-differential abundance neighborhoods (FDR 10%) are colored white, and sizes correspond to the number of cells in each neighborhood. Graph edges depict the number of cells shared between neighborhoods. The layout of nodes is determined by the position of the neighborhood index cell in the UMAP in panel **Figure 1D. (J)** Beeswarm plot of the distribution of log fold change across age in neighborhoods containing cells from different cell type clusters. Differential abundance neighborhoods at FDR 35% are colored. HSCs detected as differentially abundant annotated by HSC subsets (as defined in **Figure 1A**). **(J) (K)** Histogram for magic imputed *Kit* gene expression by age. **(L)** Representative FACS plots showing the gating strategy to define Kit^{hi} and Kit^{lo} HSC subsets in BM from young and old mice.

Figure S3: (A-H) Following allo-HCT (**as shown in Figure 2D**), frequency of donor-derived chimerism of hematopoietic cells **(A)**, mature lineages: Myeloid cells, B-cells, and T-cells **(B)**, in the PB at the indicated time points. **(C-D)** Enumeration of absolute number of donor-derived MP subset cells in the BM as defined in Figure S1H (**C)**, thymic non-TEC compartment (endothelial cells: $CD45^-$ $EpCAM^-$Ter-119 $PDGFR\alpha^-$ $CD31^+$; fibroblasts: $CD45^-$ $EpCAM^-$Ter-119^-CD31^- $PDGFR\alpha^+$) **(D)**. **(E)** Relative expression of thymic-adhesion molecules (PSGL1, CCR7, and CCR9) on donor-derived CLPs (Statistical analysis was performed using ComBat[85] to perform batch-effect correction for the mean MFI across two experiments), **(F)** Absolute amount of intrathymic thymopoietic ligands, IL-22 (left), RANKL (right) was measured by ELISA after allo-HCT. **(G-H)** Frequency of donor-derived **(G)** CD4+ and **(H)** CD8+ T-cell (naïve: CD62L+CD44- ; Central Memory: CD62L+CD44+; Effector Memory: CD62L-CD44+) subsets in the spleen. All data are from n=10-11 mice/group. **(I-O)** Following allo-HCT (**as shown in Figure 2P**)**,** frequency of donor-derived chimerism of mature lineages, myeloid (**I**) and B-cells

(**J**) in the PB at the indicated timepoints. **(K-O)** Enumeration of absolute number of donor-derived LSK **(K) and** MP subset cells (**M**) in the BM as defined in Figures 2F and S1H, respectively**,** thymic TEC **(N) and** non-TEC **(O)** compartments as defined in Figures 2K and S3D**.** Analyses were performed at 16 weeks following competitive HCT. All data are from n=9-10 mice/group.

Figure S4: Competitive allo-HCT using Kit subsets from young mice, as described in Figure 2D. Twenty weeks after competitive HCT, (**A**) frequency of donor-derived chimerism of mature lineages in the peripheral blood as described in Figure 2E.Enumeration of absolute number of donor-derived cells in the BM for LSK cell subsets as defined in Figure 2F (**B**), MP cell subsets as defined in Figure S1H (**C**), CLP cells as defined in Figure 2G (**D**), total thymic cellularity (**E**), donor-derived mature T-cells (**F**) and T-cell precursor thymocyte subsets (**G**) as defined in Figures 3I-3J, thymic stromal compartment (CD45- cells) as defined in Figures 2K and S3D (**H**). All data are from n=8 mice/group. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001. P values calculated by nonparametric unpaired Mann-Whitney U test.

Figure S5: (A-E) Single-cell ATAC seq pre-processing and QC analysis. **(A)** Comparison of TSS enrichment score and number of fragments. The dashed lines depict the cutoffs used. **(B)** Ridge plots showing distribution of TSS enrichment scores per sample (**top**) and distribution of unique nuclear fragments per sample (**bottom**). **(C)** Plots showing fragment size distribution (**top**) and TSS enrichment profile (**bottom**) across samples. (**D**) scATAC-seq UMAP without batch correction colored by sample. **(E)** Volcano plots of differentially accessible (DA) peaks (fold change> 2, FDR <0.01) for HSC subset Cluster 3 vs. Cluster 6 (Wilcoxon rank-sum test).

Figure S6: Zbtb1 expression assessed by flow cytometric analysis following knock out (KO) in Rosa26Cas9 KI Kit^{lo} HSCs **(A)** and OE of Zbtb1 cDNA in Kit^{hi} HSCs **(B)**.

Figure S7: (**A-F**) Following allo-HCT (**as shown in Figure 3C**)**,** frequency of donor-derived chimerism of mature lineages in the PB **(A)**. **(B-F)** Enumeration of absolute number of donor-derived MP subset cells in the BM as defined in Figure S1H **(B),** donor-derived mature T-cells as defined above **(C)**, thymic non-TEC compartment as defined in Figure S3D **(D),** and donor-derived **(E)** CD4+ T-cell and **(F)** CD8+ T-cell subsets in the spleen as defined above. All data are from n=9-10 mice/group. All analyses were performed at 8 weeks following competitive HCT. **(G)** Following allo-HCT (**as shown in Figure 3L**), frequency of donor-derived chimerism of mature lineages in the peripheral blood. Analyses were performed at 8 weeks following allo-HCT. All data are from n=5-6 mice/group. **(H-O)** Competitive allo-HCT using Kit subsets from old mice, as described in Figure 3C. Twenty weeks after competitive HCT, (**H**) frequency of donor-derived chimerism of mature lineages in the peripheral blood as described in Figure 2E.Enumeration of absolute number of donor-derived cells in the BM for LSK cell subsets as defined in Figure 2F (**I**), MP cell subsets as defined in Figure S1H (**J**), CLP cells as defined in Figure 2G (**K**), total thymic cellularity (**L**), donor-derived mature T-cells (**M**) and T-cell precursor thymocyte subsets (**N**) as defined in Figures 3I-3J, thymic stromal compartment (CD45- cells) as defined in Figures 2K and S3D (**O**). All data are from n=6-7 mice/group. Error bars represent mean ± SEM. Error bars represent mean ± SEM. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. P values calculated by nonparametric unpaired Mann-Whitney U test.

Figure S8: Equal Competition allo-HCT using Kit^{lo} HSCs from young and old mice. (**A**) Experimental schema for equal competition allogeneic HCT (allo-HCT) using Kit^{lo} HSCs from 2-mo young (blue) and 22-24-mo old (light-blue) C57BL/6 mice combined with competitor bone marrow (BM) cells from B6.SJL-PtprcaPepcb/BoyJ mice transplanted into lethally irradiated 7-week-old (young) BALB/cJ recipients. Twenty weeks after competitive HCT, (**B**) frequency of donor-derived chimerism of mature lineages in the peripheral blood as described in Figure 3B. Enumeration of absolute number of donor-derived cells in the BM for LSK cell subsets as defined in Figure 2F (**C**), MP cell subsets as defined in Figure S1H (**D**), CLP cells as defined in Figure 2G (**E**), donor-derived T-cell precursor thymocyte subsets as defined in Figure 3I (**F**) and mature T-cells as defined in Figure 3J (**G**). All data are from n=9 mice. Error bars represent mean ± SEM. (**H**) Pathway enrichment analysis performed on differentially expressed genes comparing old versus young Kit^{lo} HSCs by *gseapy*()[86] using Hallmark_2020 and GO_Biological_Process_2023 libraries. Bubbleplot showing representative pathways enriched in young Kit^{lo} HSCs. (**I-K**) Matrix plots showing motif enrichment identified by ChromVAR (I), gene accessibility score (**J**), and gene expression (**K**) for Kit^{lo} subset in young and old HSCs. *P<0.05, **P<0.01, ***P<0.001. P values calculated by nonparametric unpaired Mann-Whitney U test.

Figure S9: (**A**) UMAP of CD34+ young and old human BM CITE-seq[62] annotated with 15 clusters. **(B)** Violin plot for CD117 ADT-expression by age. Statistical analysis performed using Wilcoxon test. **(C)** Representative FACS plots of human BM samples showing gating strategy to define HSPC subsets. **(D)** Representative FACS plots of human BM samples showing gating strategy to define KIT^{hi} and KIT^{lo} HSCs in human BM.

1287

Figure 1

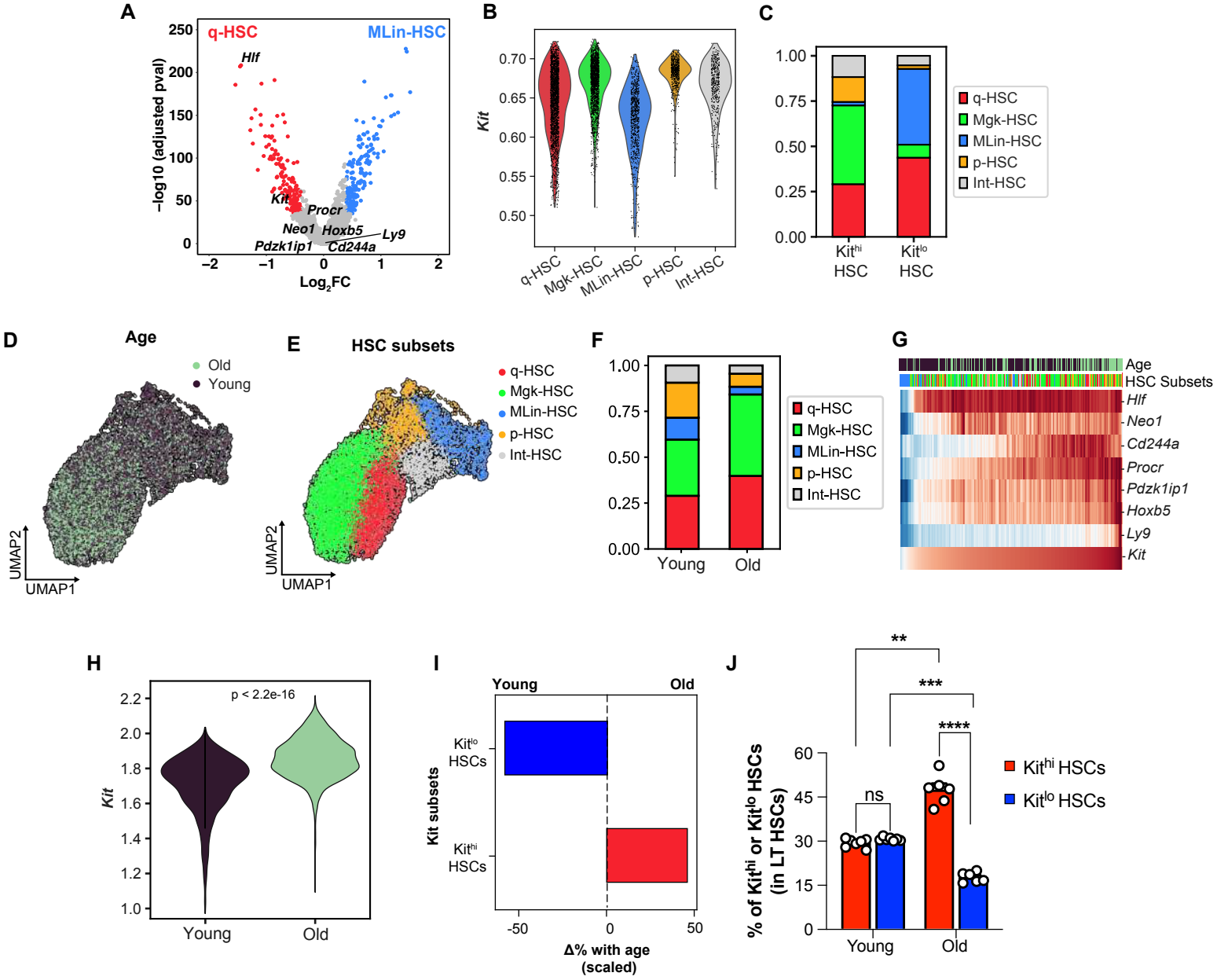

A
q-HSC
MLin-HSC
Hlf
Kit
Procr
Neo1
Hoxb5
Ly9
Pdzk1ip1
Cd244a
−log10 (adjusted pval)
Log2FC
B
Kit
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
C
Kithi HSC
Kitlo HSC
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
D
Age
Old
Young
UMAP2
UMAP1
E
HSC subsets
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
UMAP2
UMAP1
F
Young
Old
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
G
Age
HSC Subsets
Hlf
Neo1
Cd244a
Procr
Pdzk1ip1
Hoxb5
Ly9
Kit
H
p < 2.2e-16
Kit
Young
Old
I
Young
Old
Kit subsets
Kitlo HSCs
Kithi HSCs
Δ% with age (scaled)
J
**
***
****
ns
% of Kithi or Kitlo HSCs (in LT HSCs)
Kithi HSCs
Kitlo HSCs
Young
Old

Figure 2

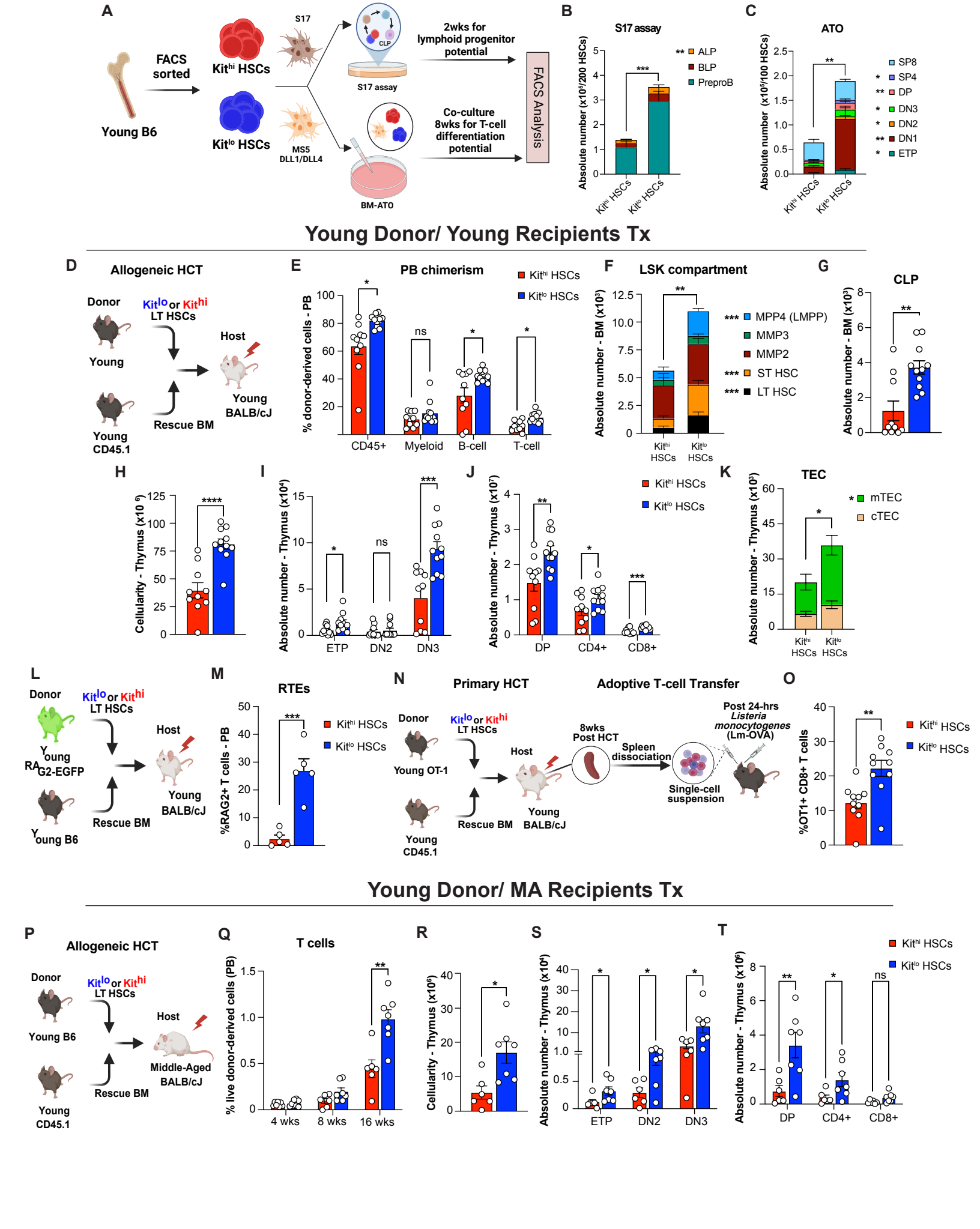

A
Young B6
FACS sorted
Kit^hi HSCs
Kit^lo HSCs
S17
MS5 DLL1/DLL4
CLP
S17 assay
BM-ATO
2wks for lymphoid progenitor potential
Co-culture 8wks for T-cell differentiation potential
FACS Analysis
B
S17 assay
Absolute number (x10^5/200 HSCs)
ALP
BLP
PreproB
C
ATO
Absolute number (x10^5/100 HSCs)
SP8
SP4
DP
DN3
DN2
DN1
ETP
Young Donor/ Young Recipients Tx
D
Allogeneic HCT
Donor
Kit^lo or Kit^hi LT HSCs
Young
Host
Young BALB/cJ
Rescue BM
Young CD45.1
E
PB chimerism
% donor-derived cells - PB
CD45+
Myeloid
B-cell
T-cell
F
LSK compartment
Absolute number - BM (x10^3)
MPP4 (LMPP)
MMP3
MMP2
ST HSC
LT HSC
G
CLP
Absolute number - BM (x10^3)
H
Cellularity - Thymus (x10^6)
I
Absolute number - Thymus (x10^4)
ETP
DN2
DN3
J
Absolute number - Thymus (x10^7)
DP
CD4+
CD8+
K
TEC
Absolute number - Thymus (x10^3)
mTEC
cTEC
L
Donor
Kit^lo or Kit^hi LT HSCs
Young RAG2-EGFP
Host
Young BALB/cJ
Rescue BM
Young B6
M
RTEs
%RAG2+ T cells - PB
N
Primary HCT
Adoptive T-cell Transfer
Donor
Kit^lo or Kit^hi LT HSCs
Young OT-1
Host
Young BALB/cJ
Rescue BM
Young CD45.1
8wks Post HCT
Spleen dissociation
Single-cell suspension
Post 24-hrs Listeria monocytogenes (Lm-OVA)
O
%OT1+ CD8+ T cells
Young Donor/ MA Recipients Tx
P
Allogeneic HCT
Donor
Kit^lo or Kit^hi LT HSCs
Young B6
Host
Middle-Aged BALB/cJ
Rescue BM
Young CD45.1
Q
T cells
% live donor-derived cells (PB)
4 wks
8 wks
16 wks
R
Cellularity - Thymus (x10^6)
S
Absolute number - Thymus (x10^4)
ETP
DN2
DN3
T
Absolute number - Thymus (x10^6)
DP
CD4+
CD8+

Figure 3

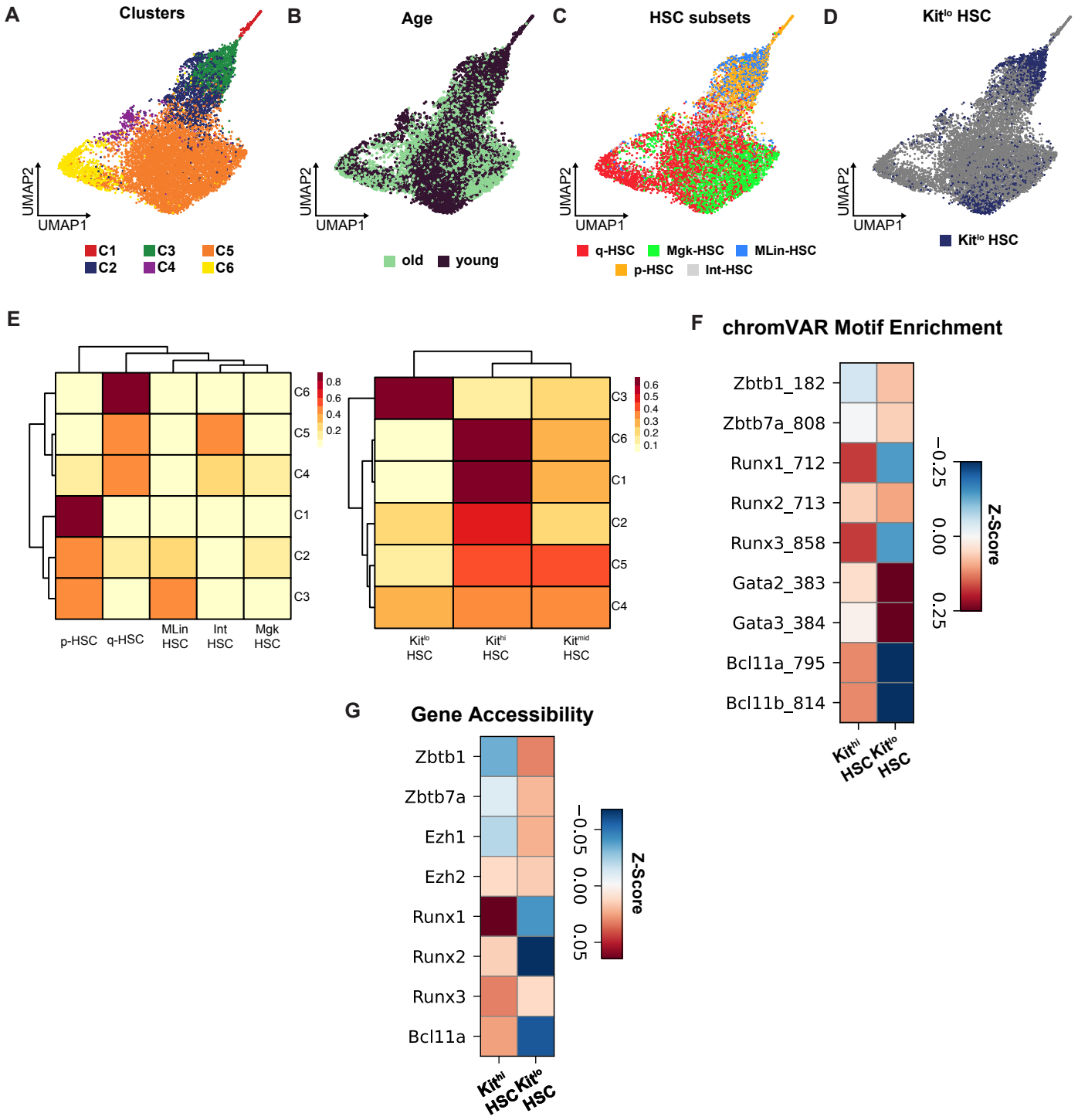

A
Clusters
UMAP2
UMAP1
C1
C2
C3
C4
C5
C6
B
Age
UMAP2
UMAP1
old
young
C
HSC subsets
UMAP2
UMAP1
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
D
Kit^{lo} HSC
UMAP2
UMAP1
Kit^{lo} HSC
E
C6
C5
C4
C1
C2
C3
p-HSC
q-HSC
MLin HSC
Int HSC
Mgk HSC
0.8
0.6
0.4
0.2
C3
C6
C1
C2
C5
C4
Kit^{lo} HSC
Kit^{hi} HSC
Kit^{mid} HSC
0.6
0.5
0.4
0.3
0.2
0.1
F
chromVAR Motif Enrichment
Zbtb1_182
Zbtb7a_808
Runx1_712
Runx2_713
Runx3_858
Gata2_383
Gata3_384
Bcl11a_795
Bcl11b_814
Kit^{hi} HSC
Kit^{lo} HSC
−0.25
0.00
0.25
Z-Score
G
Gene Accessibility
Zbtb1
Zbtb7a
Ezh1
Ezh2
Runx1
Runx2
Runx3
Bcl11a
Kit^{hi} HSC
Kit^{lo} HSC
−0.05
0.00
0.05
Z-Score

Figure 4

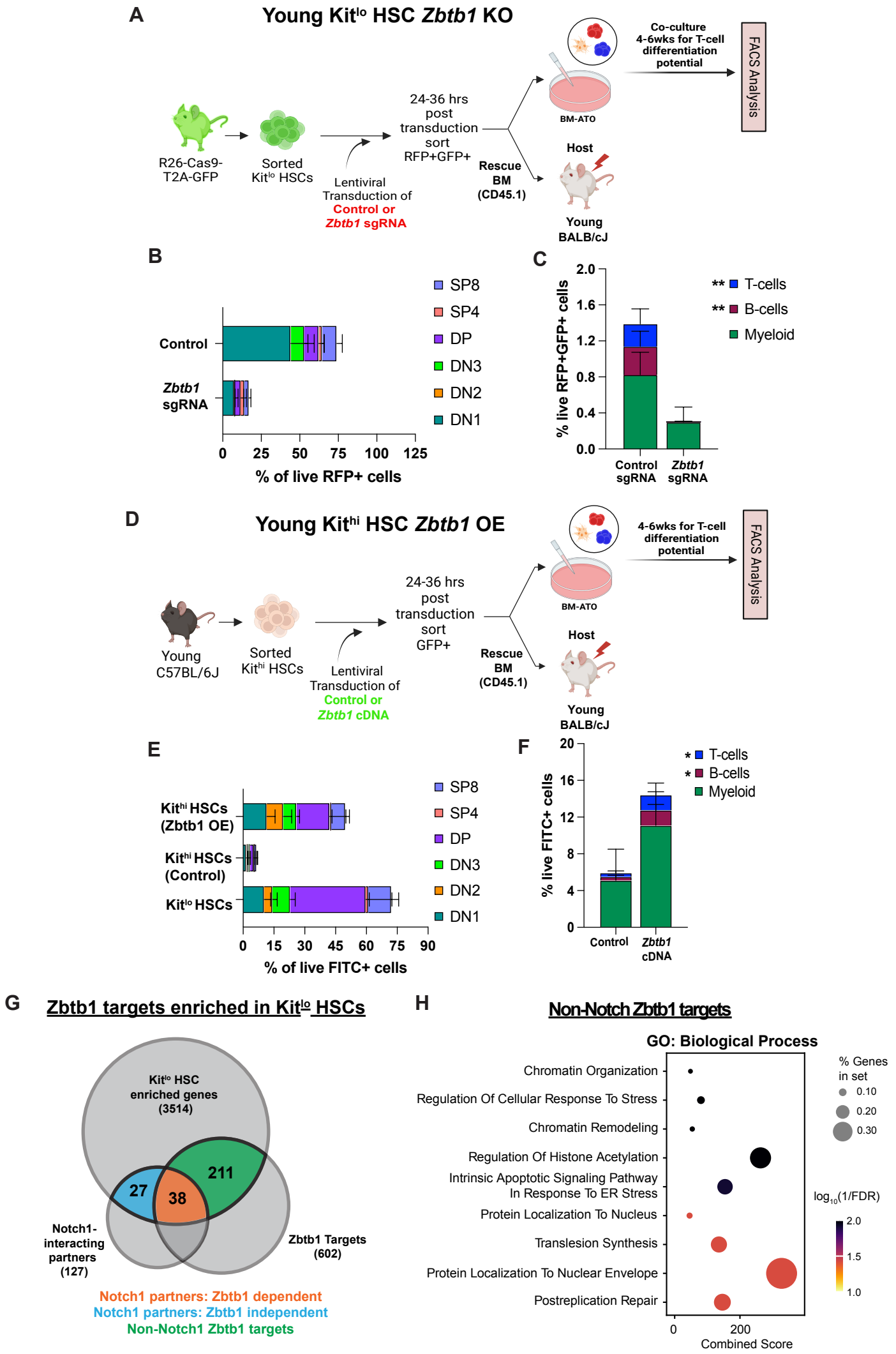

A
Young Kit^{lo} HSC Zbtb1 KO
R26-Cas9-T2A-GFP
Sorted Kit^{lo} HSCs
Lentiviral Transduction of Control or Zbtb1 sgRNA
24-36 hrs post transduction sort RFP+GFP+
Rescue BM (CD45.1)
BM-ATO
Co-culture 4-6wks for T-cell differentiation potential
FACS Analysis
Host
Young BALB/cJ
B
Control
Zbtb1 sgRNA
SP8
SP4
DP
DN3
DN2
DN1
0 25 50 75 100 125
% of live RFP+ cells
C
% live RFP+GFP+ cells
2.0 1.6 1.2 0.8 0.4 0.0
** T-cells
** B-cells
Myeloid
Control sgRNA
Zbtb1 sgRNA
D
Young Kit^{hi} HSC Zbtb1 OE
Young C57BL/6J
Sorted Kit^{hi} HSCs
Lentiviral Transduction of Control or Zbtb1 cDNA
24-36 hrs post transduction sort GFP+
Rescue BM (CD45.1)
BM-ATO
4-6wks for T-cell differentiation potential
FACS Analysis
Host
Young BALB/cJ
E
Kit^{hi} HSCs (Zbtb1 OE)
Kit^{hi} HSCs (Control)
Kit^{lo} HSCs
SP8
SP4
DP
DN3
DN2
DN1
0 15 30 45 60 75 90
% of live FITC+ cells
F
% live FITC+ cells
20 16 12 8 4 0
* T-cells
* B-cells
Myeloid
Control
Zbtb1 cDNA
G
Zbtb1 targets enriched in Kit^{lo} HSCs
Kit^{lo} HSC enriched genes (3514)
27
38
211
Notch1-interacting partners (127)
Zbtb1 Targets (602)
Notch1 partners: Zbtb1 dependent
Notch1 partners: Zbtb1 independent
Non-Notch1 Zbtb1 targets
H
Non-Notch Zbtb1 targets
GO: Biological Process
Chromatin Organization
Regulation Of Cellular Response To Stress
Chromatin Remodeling
Regulation Of Histone Acetylation
Intrinsic Apoptotic Signaling Pathway In Response To ER Stress
Protein Localization To Nucleus
Translesion Synthesis
Protein Localization To Nuclear Envelope
Postreplication Repair
0 200
Combined Score
% Genes in set
0.10
0.20
0.30
log$_{10}$(1/FDR)
2.0
1.5
1.0

Figure 5

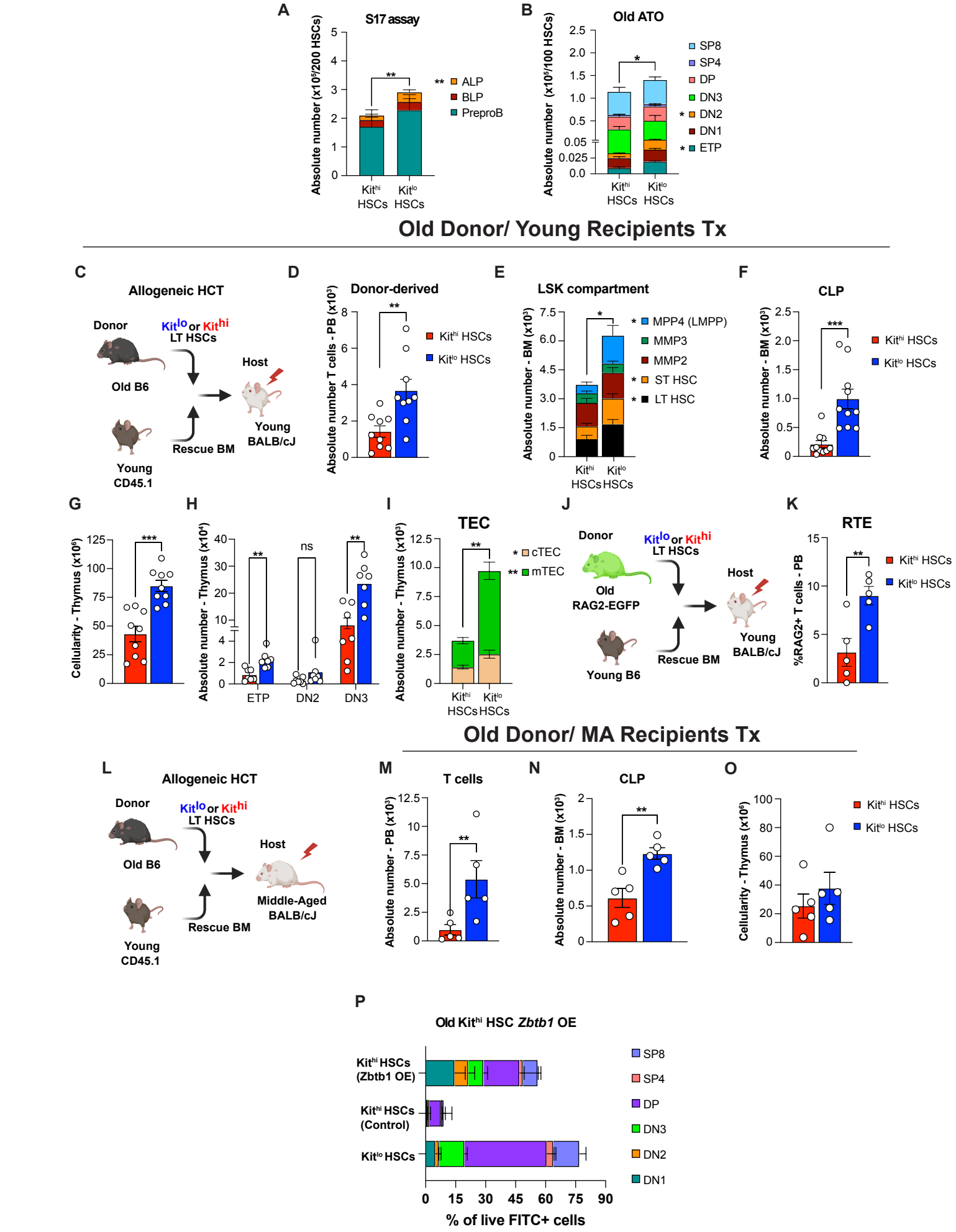

A
S17 assay
Absolute number (x10⁵/200 HSCs)
** ALP
BLP
PreproB
Kit^hi HSCs
Kit^lo HSCs
B
Old ATO
Absolute number (x10⁵/100 HSCs)
SP8
SP4
DP
DN3
* DN2
DN1
* ETP
Old Donor/ Young Recipients Tx
C
Allogeneic HCT
Donor
Kit^lo or Kit^hi LT HSCs
Old B6
Host
Young BALB/cJ
Rescue BM
Young CD45.1
D
Donor-derived
Absolute number T cells - PB (x10³)
Kit^hi HSCs
Kit^lo HSCs
E
LSK compartment
Absolute number - BM (x10³)
* MPP4 (LMPP)
MMP3
MMP2
* ST HSC
* LT HSC
F
CLP
Absolute number - BM (x10³)
G
Cellularity - Thymus (x10⁶)
H
Absolute number - Thymus (x10⁴)
ETP
DN2
DN3
ns
I
TEC
Absolute number - Thymus (x10³)
* cTEC
** mTEC
J
Donor
Old RAG2-EGFP
Young B6
Host
Young BALB/cJ
Rescue BM
K
RTE
%RAG2+ T cells - PB
Old Donor/ MA Recipients Tx
L
Allogeneic HCT
Donor
Old B6
Young CD45.1
Rescue BM
Middle-Aged BALB/cJ
M
T cells
Absolute number - PB (x10³)
N
CLP
Absolute number - BM (x10³)
O
Cellularity - Thymus (x10⁶)
P
Old Kit^hi HSC Zbtb1 OE
Kit^hi HSCs (Zbtb1 OE)
Kit^hi HSCs (Control)
Kit^lo HSCs
SP8
SP4
DP
DN3
DN2
DN1
% of live FITC+ cells

Figure 6

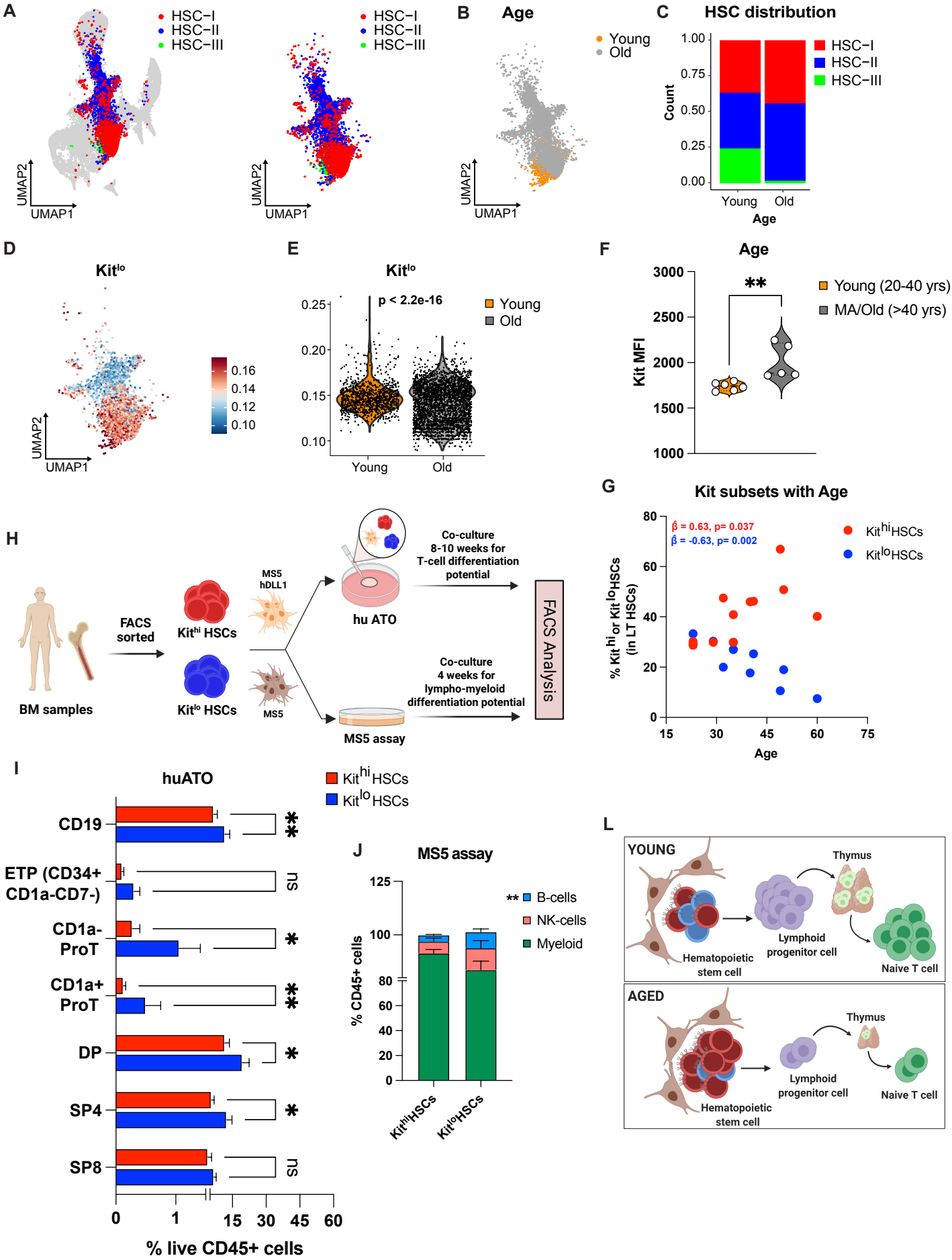

A
HSC−I
HSC−II
HSC−III
UMAP2
UMAP1
B
Age
Young
Old
C
HSC distribution
Count
Young
Old
Age
D
Kit^lo
0.16
0.14
0.12
0.10
E
Kit^lo
p < 2.2e-16
Young
Old
F
Age
Kit MFI
**
Young (20-40 yrs)
MA/Old (>40 yrs)
G
Kit subsets with Age
β̂ = 0.63, p= 0.037
β̂ = -0.63, p= 0.002
Kit^hi HSCs
Kit^lo HSCs
% Kit^hi or Kit^lo HSCs (in LT HSCs)
Age
H
BM samples
FACS sorted
Kit^hi HSCs
Kit^lo HSCs
MS5 hDLL1
MS5
hu ATO
MS5 assay
Co-culture 8-10 weeks for T-cell differentiation potential
Co-culture 4 weeks for lympho-myeloid differentiation potential
FACS Analysis
I
huATO
Kit^hi HSCs
Kit^lo HSCs
CD19
ETP (CD34+ CD1a-CD7-)
CD1a- ProT
CD1a+ ProT
DP
SP4
SP8
ns
% live CD45+ cells
J
MS5 assay
% CD45+ cells
B-cells
NK-cells
Myeloid
Kit^hi HSCs
Kit^lo HSCs
L
YOUNG
Thymus
Hematopoietic stem cell
Lymphoid progenitor cell
Naive T cell
AGED
Thymus
Hematopoietic stem cell
Lymphoid progenitor cell
Naive T cell

Figure S1

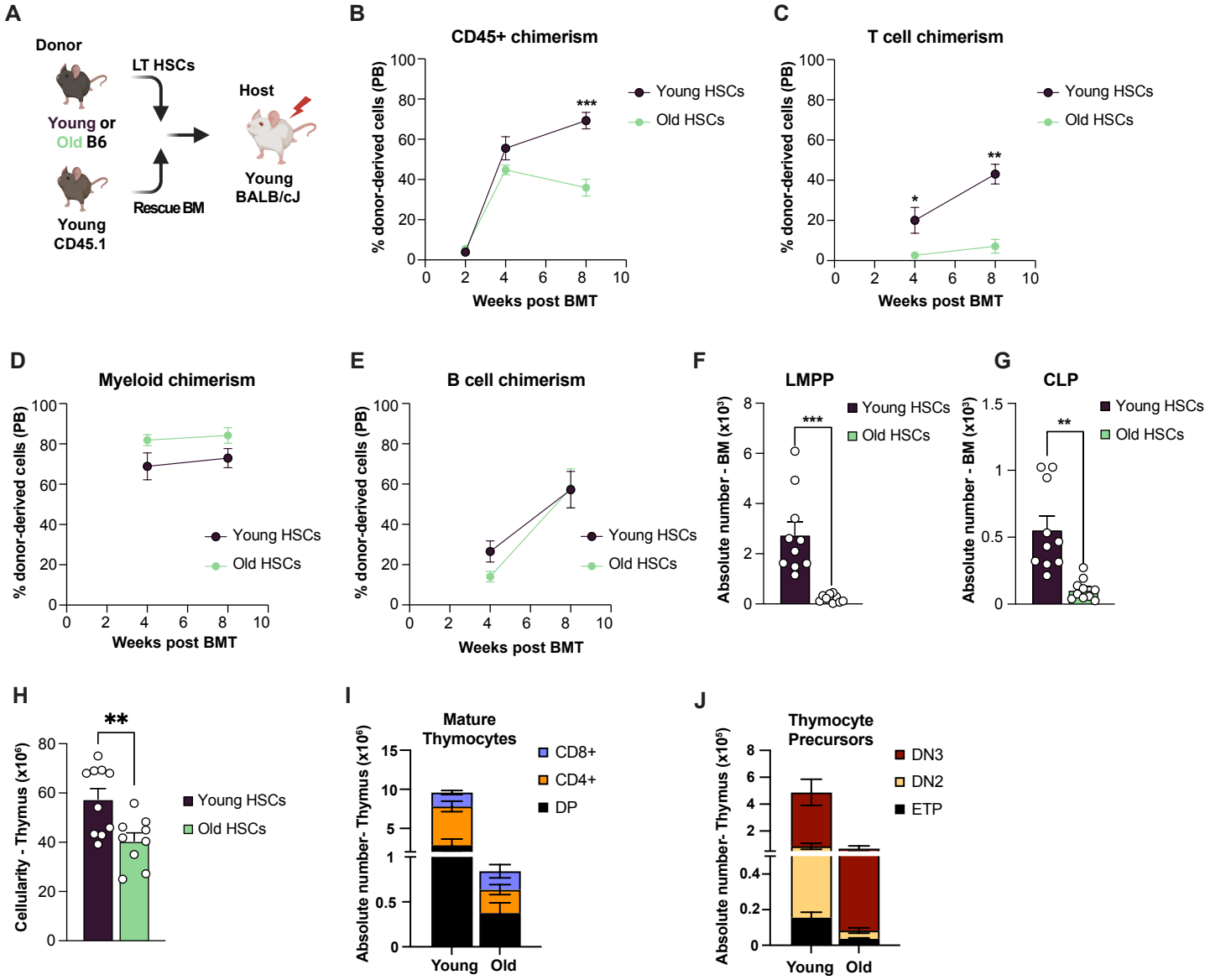

A
Donor
Young or
Old B6
Young
CD45.1
LT HSCs
Rescue BM
Host
Young
BALB/cJ
B
CD45+ chimerism
% donor-derived cells (PB)
Weeks post BMT
Young HSCs
Old HSCs
***
C
T cell chimerism
% donor-derived cells (PB)
Weeks post BMT
Young HSCs
Old HSCs
*
**
D
Myeloid chimerism
% donor-derived cells (PB)
Weeks post BMT
Young HSCs
Old HSCs
E
B cell chimerism
% donor-derived cells (PB)
Weeks post BMT
Young HSCs
Old HSCs
F
LMPP
Absolute number - BM (x10³)
Young HSCs
Old HSCs
***
G
CLP
Absolute number - BM (x10³)
Young HSCs
Old HSCs
**
H
Cellularity - Thymus (x10⁶)
Young HSCs
Old HSCs
**
I
Mature
Thymocytes
Absolute number- Thymus (x10⁶)
CD8+
CD4+
DP
Young
Old
J
Thymocyte
Precursors
Absolute number- Thymus (x10⁵)
DN3
DN2
ETP
Young
Old

Figure S2

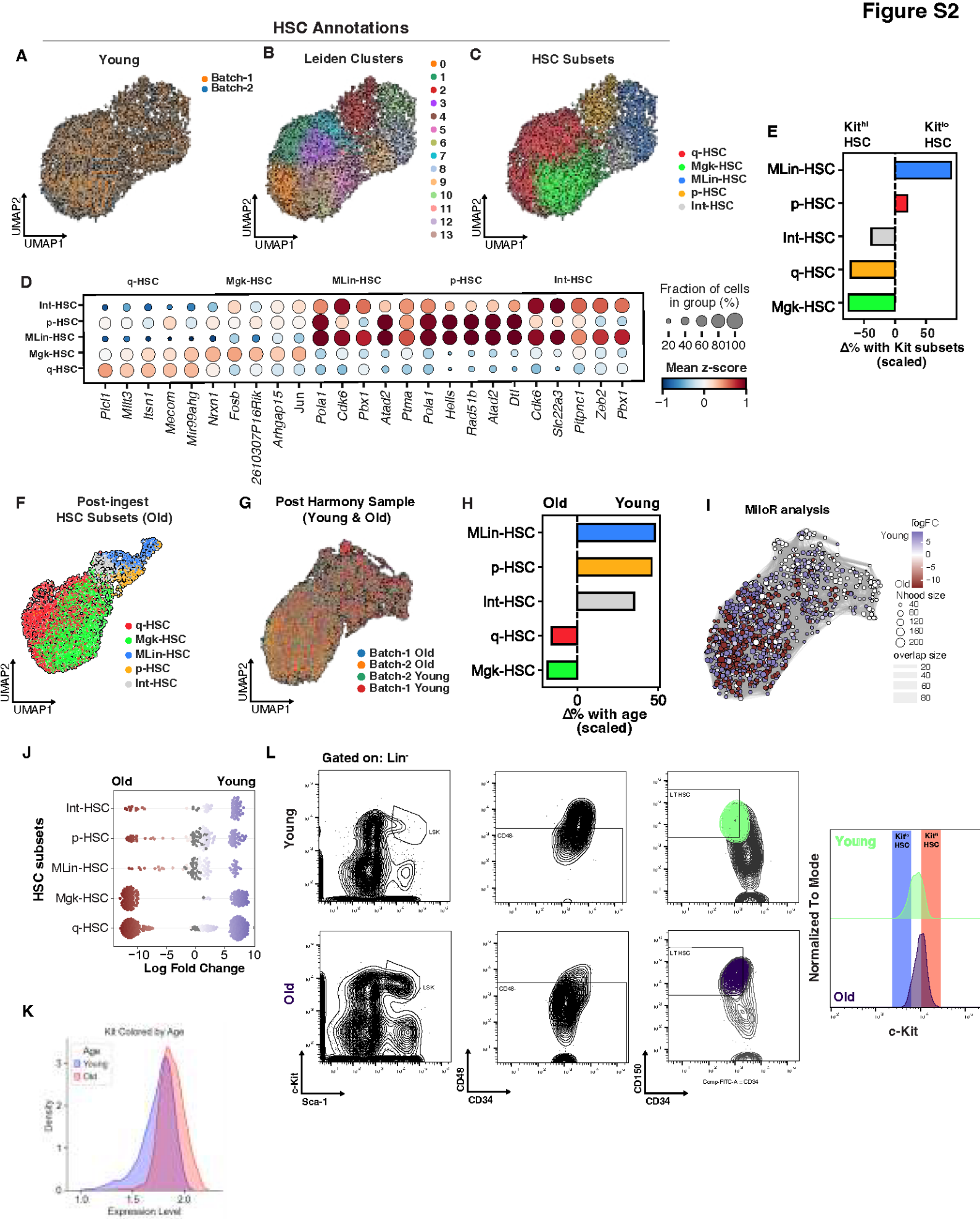
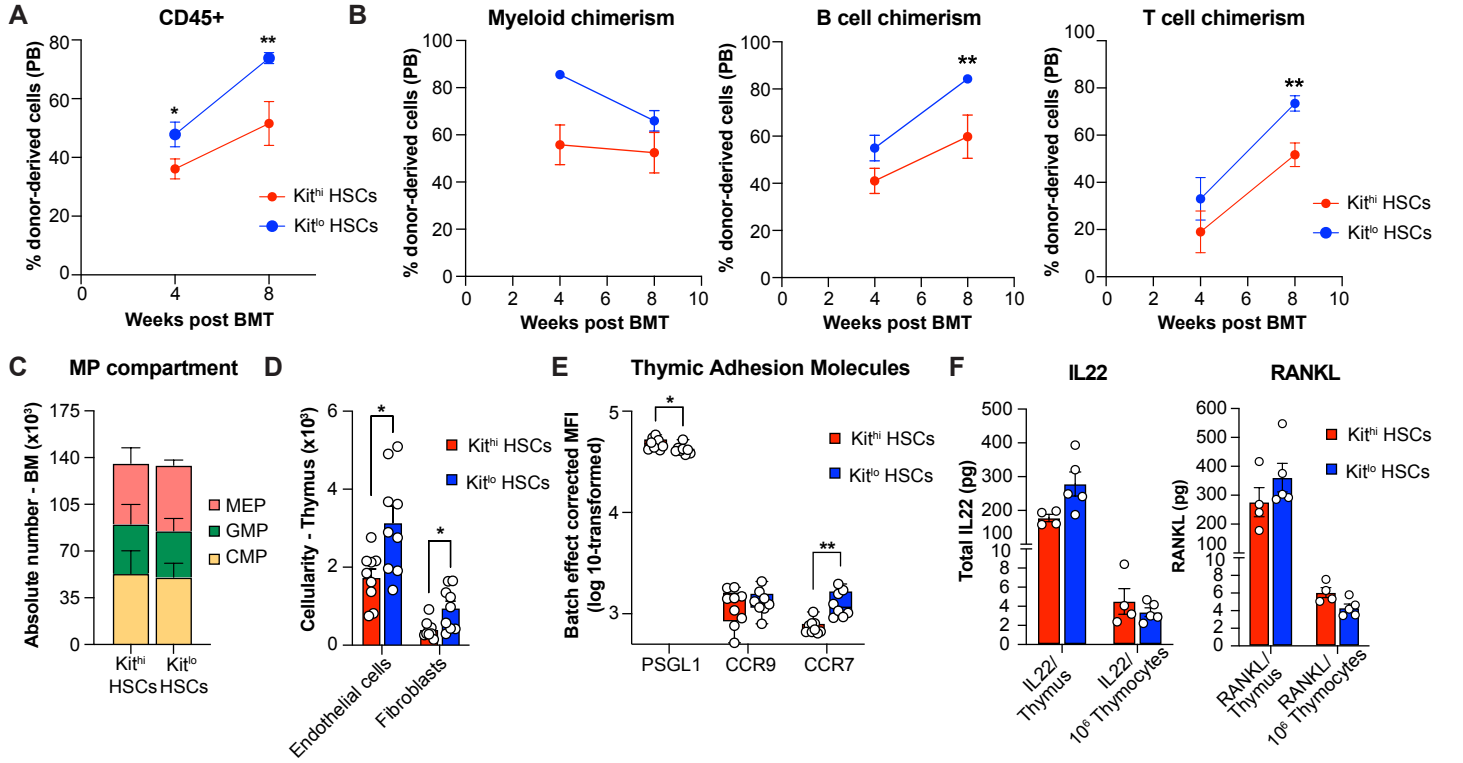
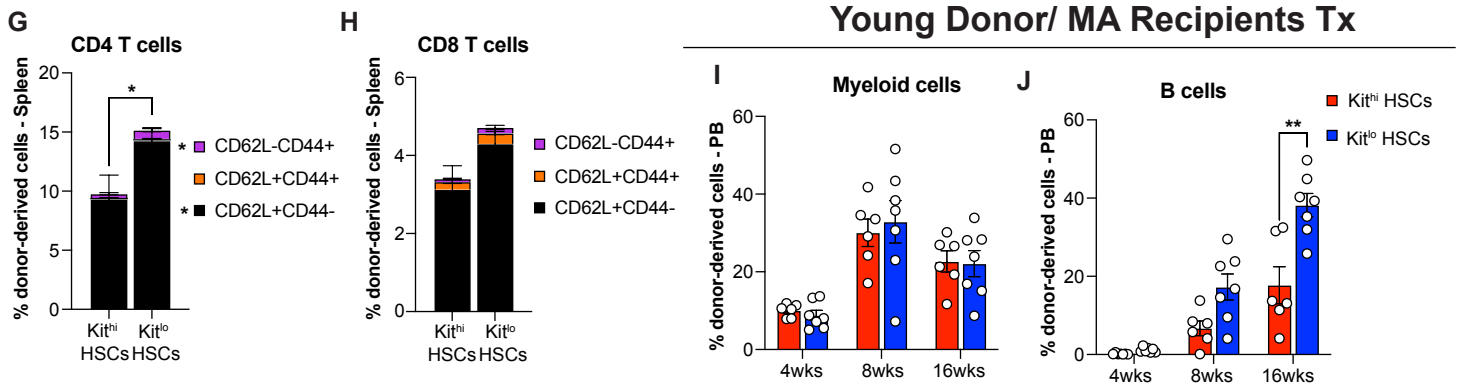

HSC Annotations
A
Young
Batch-1
Batch-2
UMAP2
UMAP1
B
Leiden Clusters
0
1
2
3
4
5
6
7
8
9
10
11
12
13
C
HSC Subsets
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
D
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
Plcl1
Mlt3
Itsn1
Mecom
Mir99ahg
Nrxn1
Fosb
2610307P16Rik
Arhgap15
Jun
Pola1
Cdk6
Pbx1
Atad2
Ptma
Pola1
Hells
Rad51b
Atad2
Dtl
Cdk6
Slc22a3
Pitpnc1
Zeb2
Pbx1
Fraction of cells in group (%)
20 40 60 80 100
Mean z-score
−1
0
1
E
Kit^hi HSC
Kit^lo HSC
MLin-HSC
p-HSC
Int-HSC
q-HSC
Mgk-HSC
−50
0
50
Δ% with Kit subsets (scaled)
F
Post-ingest HSC Subsets (Old)
q-HSC
Mgk-HSC
MLin-HSC
p-HSC
Int-HSC
UMAP2
UMAP1
G
Post Harmony Sample (Young & Old)
Batch-1 Old
Batch-2 Old
Batch-2 Young
Batch-1 Young
UMAP2
UMAP1
H
Old
Young
MLin-HSC
p-HSC
Int-HSC
q-HSC
Mgk-HSC
0
50
Δ% with age (scaled)
I
MiloR analysis
logFC
Young
Old
Nhood size
40
80
120
160
200
overlap size
20
40
60
80
J
Old
Young
HSC subsets
Int-HSC
p-HSC
MLin-HSC
Mgk-HSC
q-HSC
−10
−5
0
5
10
Log Fold Change
K
Kit Colored by Age
Age
Young
Old
Density
1.0
1.5
2.0
Expression Level
L
Gated on: Lin-
Young
Old
LSK
CD48-
LT HSC
c-Kit
Sca-1
CD48
CD34
CD150
CD34
Comp-FITC-A :: CD34
Young
Old
Kit^lo HSC
Kit^hi HSC
Normalized To Mode
c-Kit

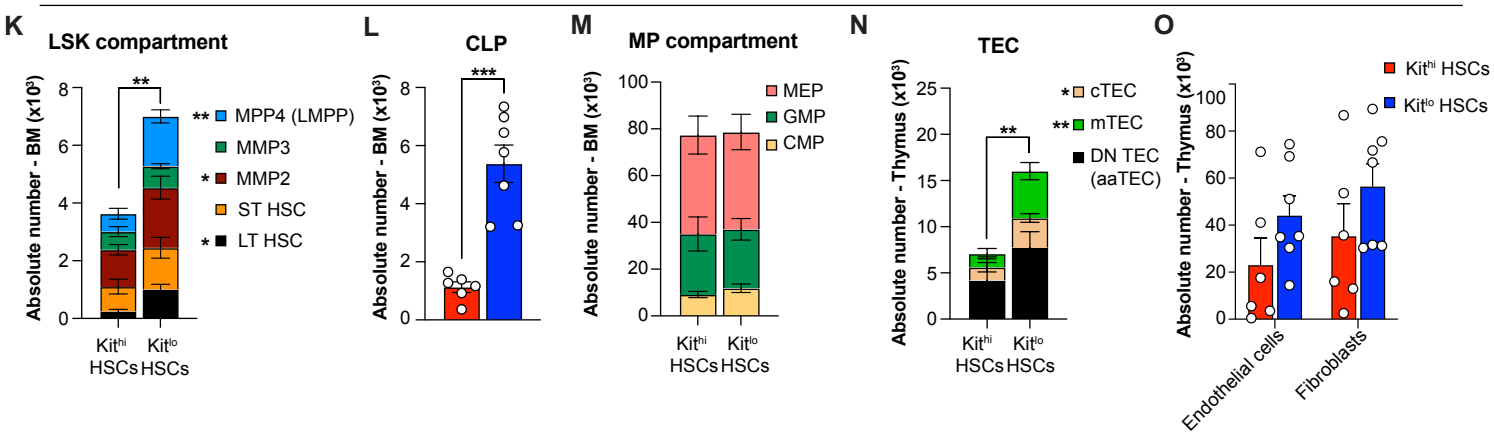
Figure S3

Young Donor/ Young Recipients Tx





Young Donor/ MA Recipients Tx



Young 20-week Post HCT

Figure S4

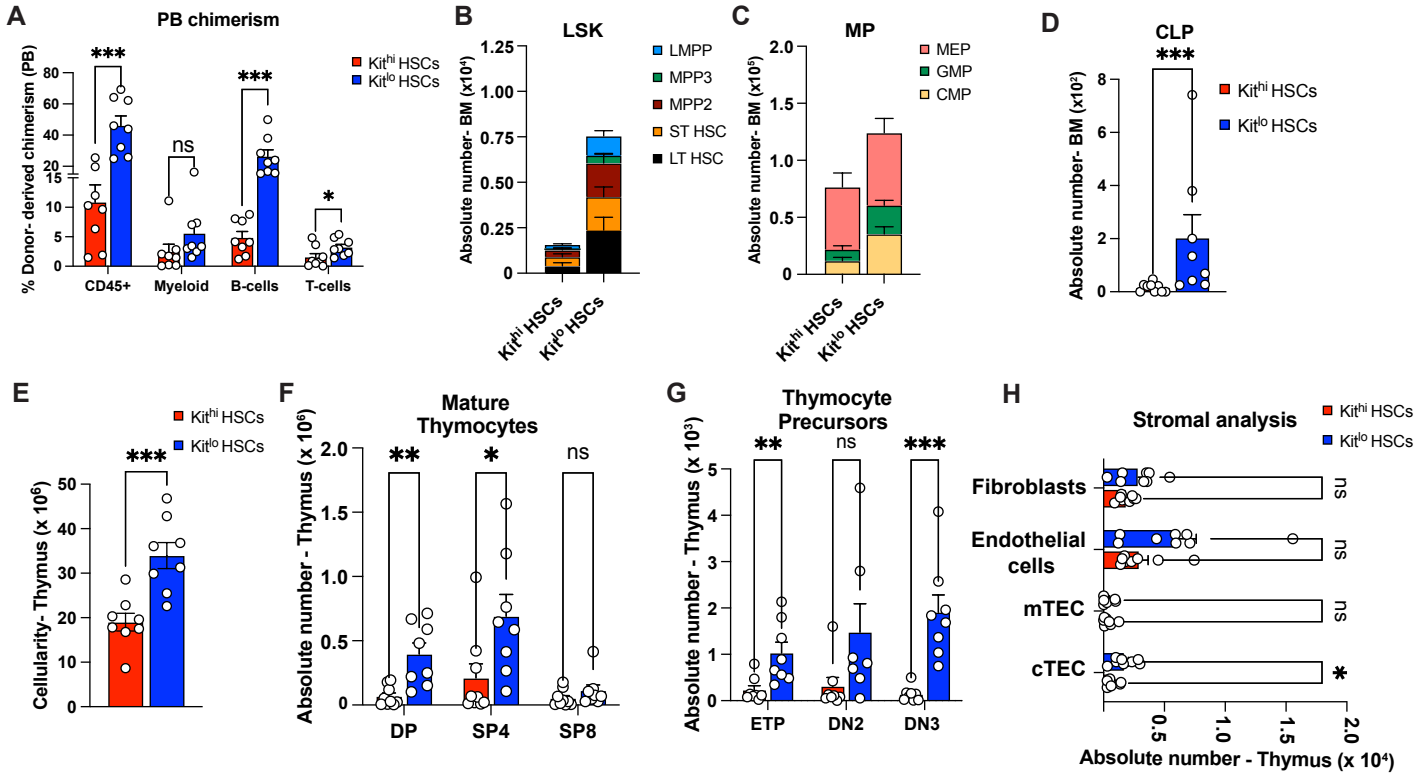


Figure S5

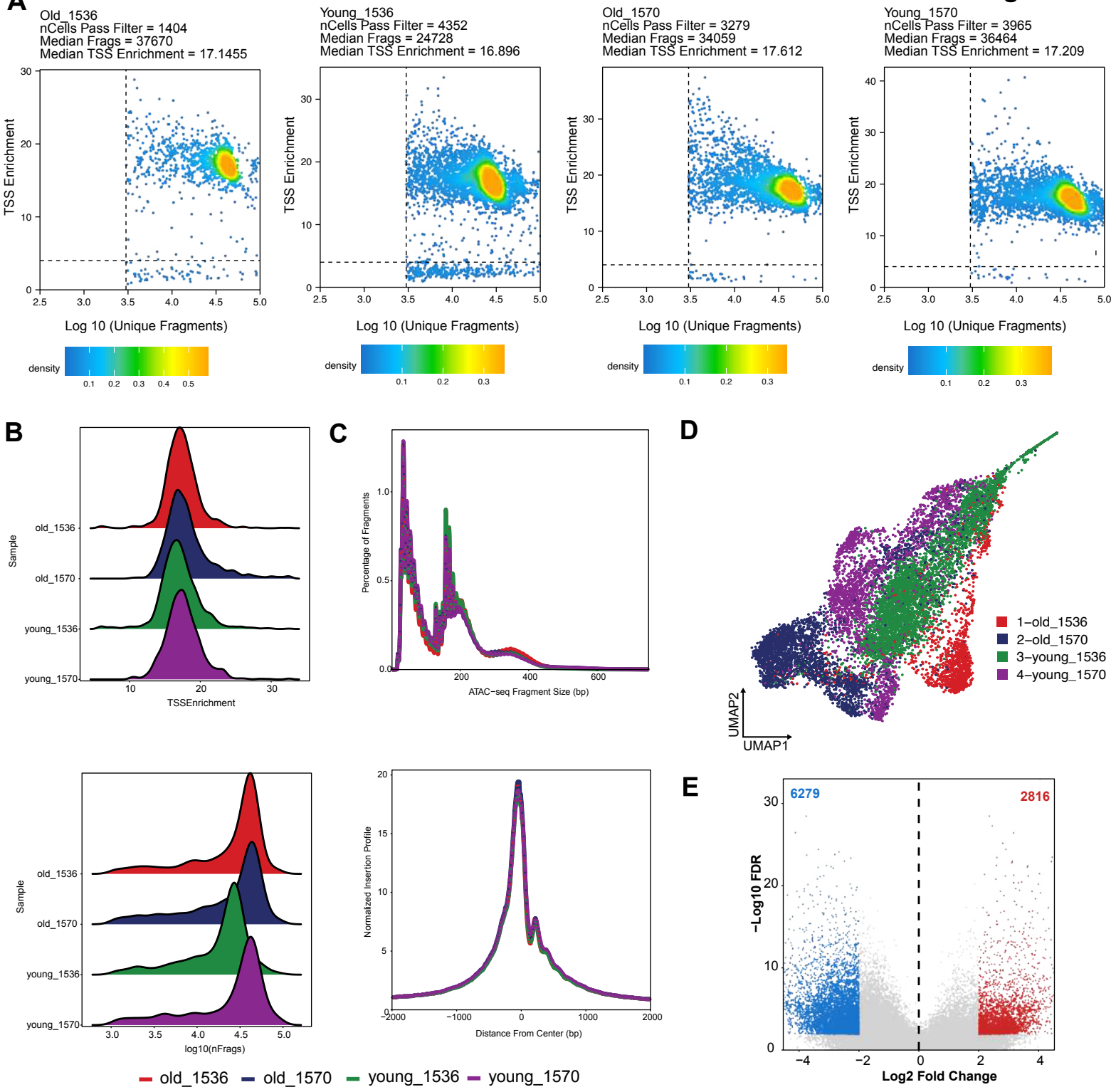

A
Old_1536
nCells Pass Filter = 1404
Median Frags = 37670
Median TSS Enrichment = 17.1455
Young_1536
nCells Pass Filter = 4352
Median Frags = 24728
Median TSS Enrichment = 16.896
Old_1570
nCells Pass Filter = 3279
Median Frags = 34059
Median TSS Enrichment = 17.612
Young_1570
nCells Pass Filter = 3965
Median Frags = 36464
Median TSS Enrichment = 17.209
TSS Enrichment
Log 10 (Unique Fragments)
density
B
Sample
old_1536
old_1570
young_1536
young_1570
TSSEnrichment
C
Percentage of Fragments
ATAC−seq Fragment Size (bp)
D
UMAP2
UMAP1
1−old_1536
2−old_1570
3−young_1536
4−young_1570
log10(nFrags)
Normalized Insertion Profile
Distance From Center (bp)
E
6279
2816
−Log10 FDR
Log2 Fold Change
old_1536
old_1570
young_1536
young_1570

Figure S6

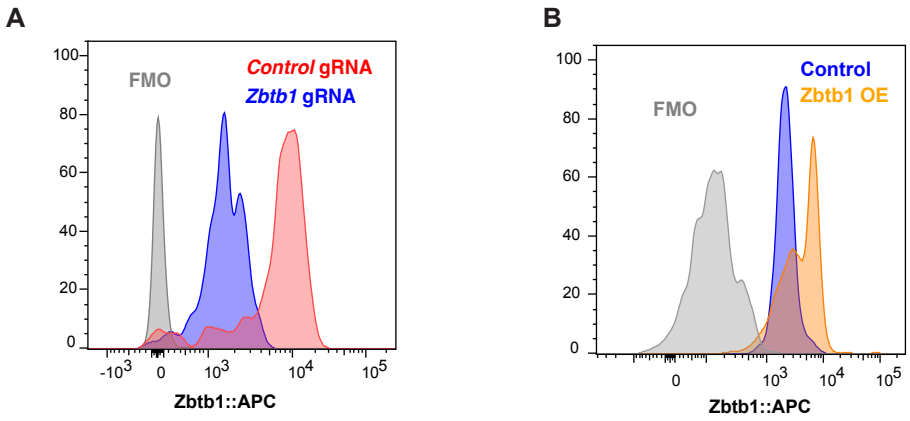
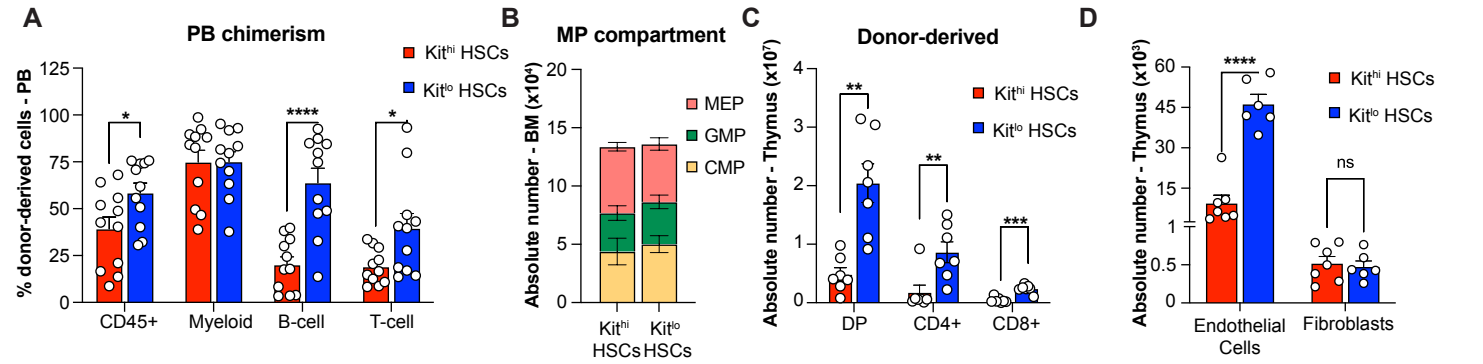
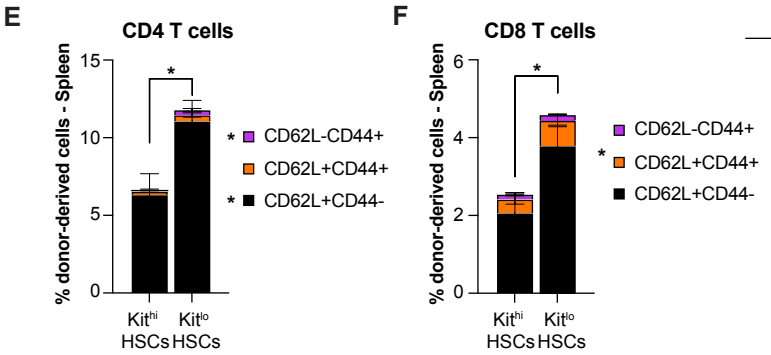

A
FMO
Control gRNA
Zbtb1 gRNA
100
80
60
40
20
0
-10^3
0
10^3
10^4
10^5
Zbtb1::APC
B
FMO
Control
Zbtb1 OE
100
80
60
40
20
0
0
10^3
10^4
10^5
Zbtb1::APC

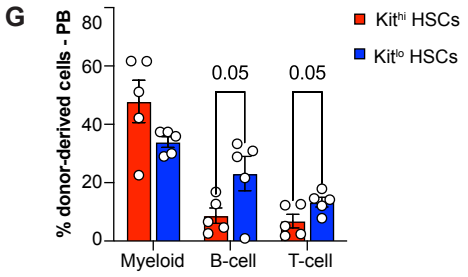
Figure S7

Old Donor/ Young Recipients Tx

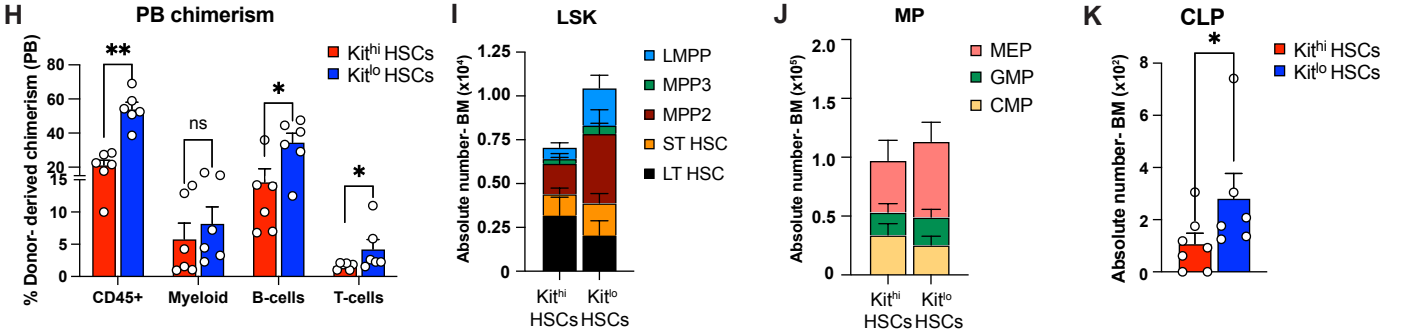




Old Donor/ MA Recipients Tx



Old 20-week Post HCT



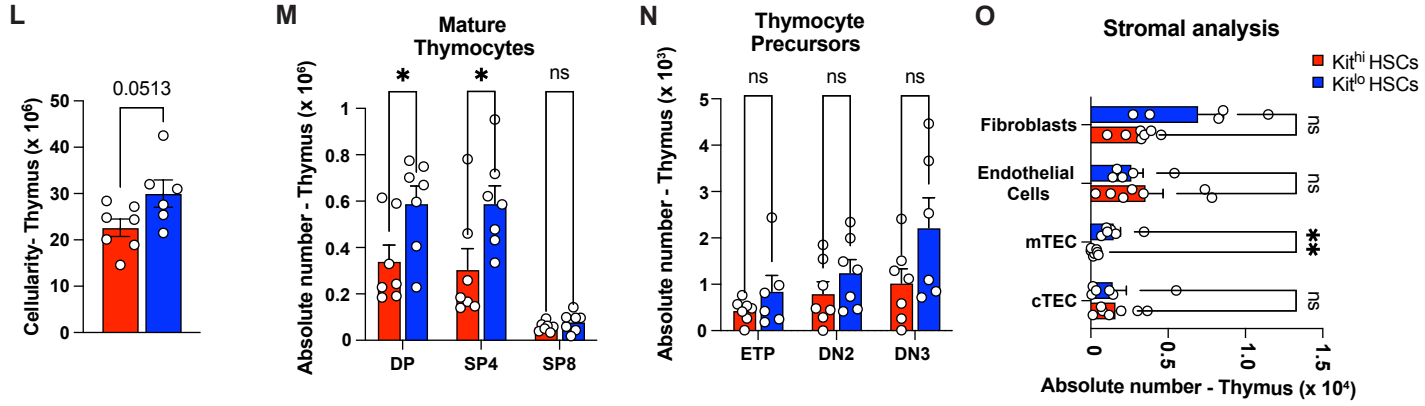


Figure S8

Equal Competition Tx

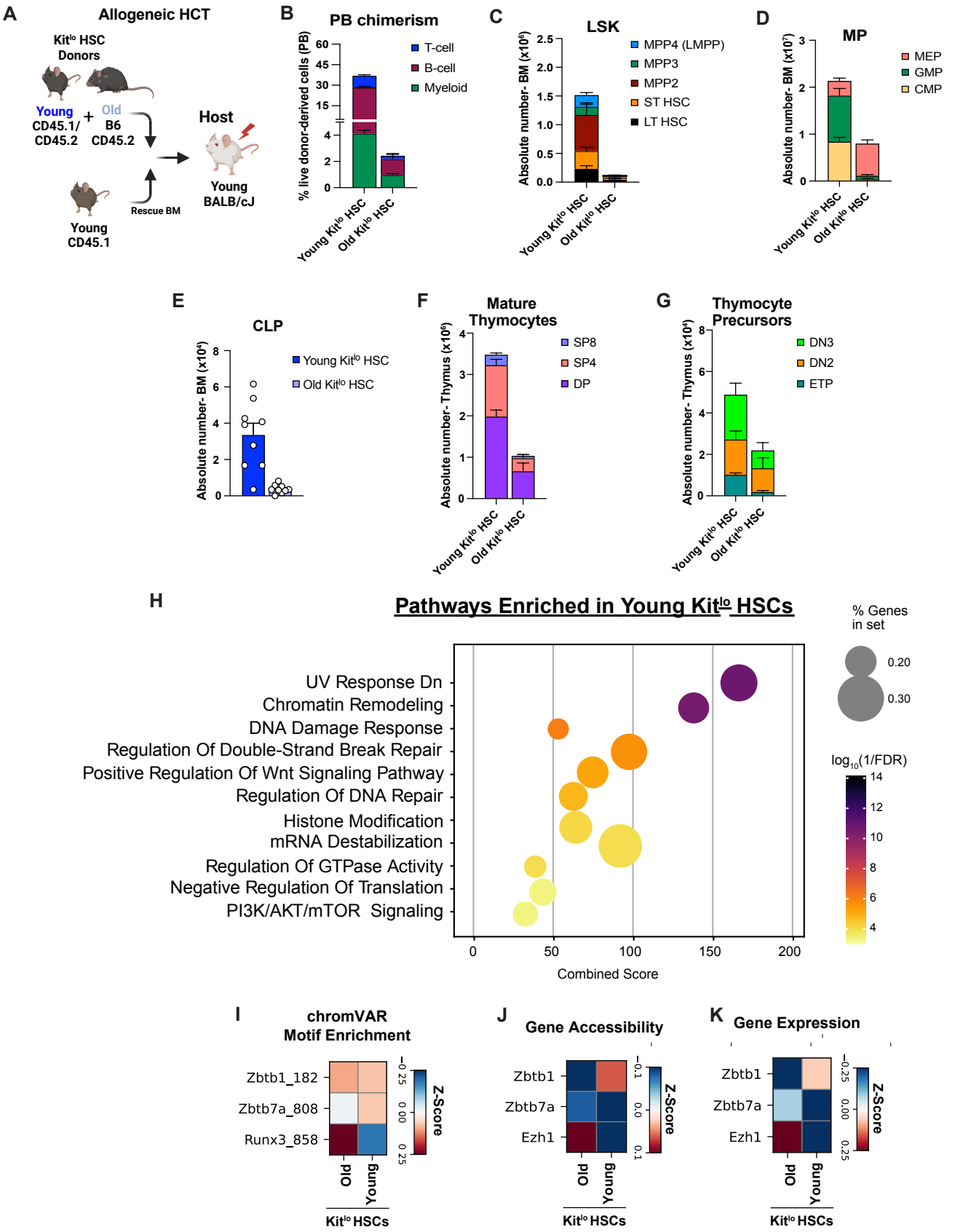
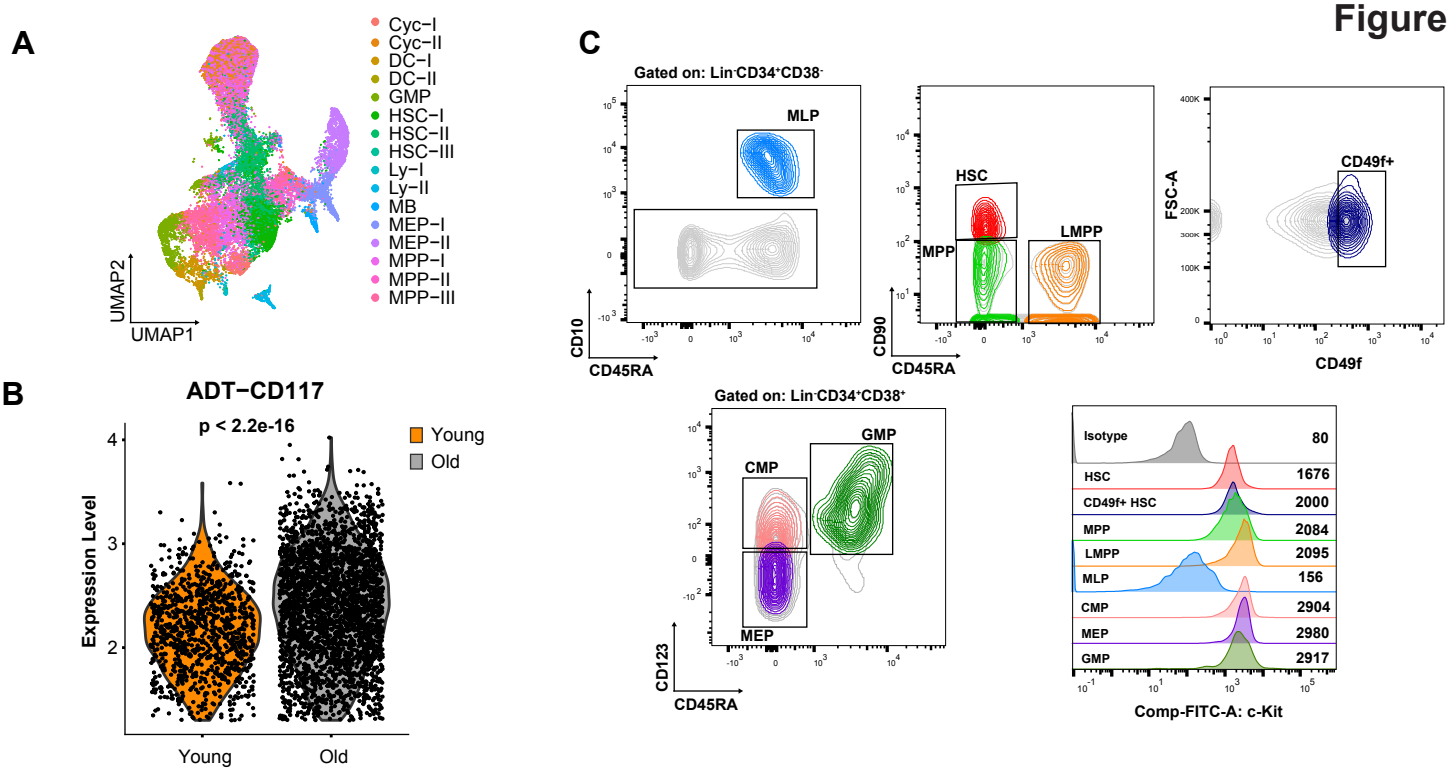
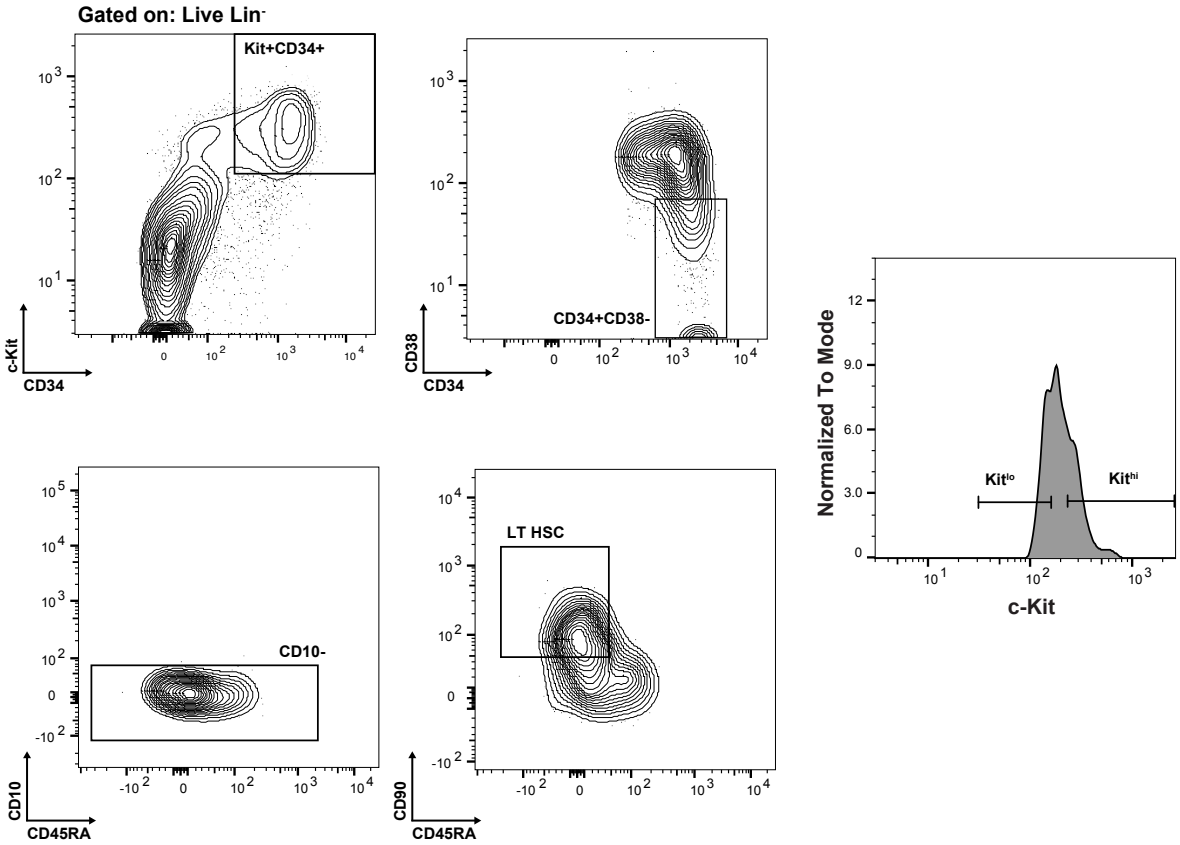


Figure S9



D Gating Strategy to Fractionate Kit Subsets in Human BM