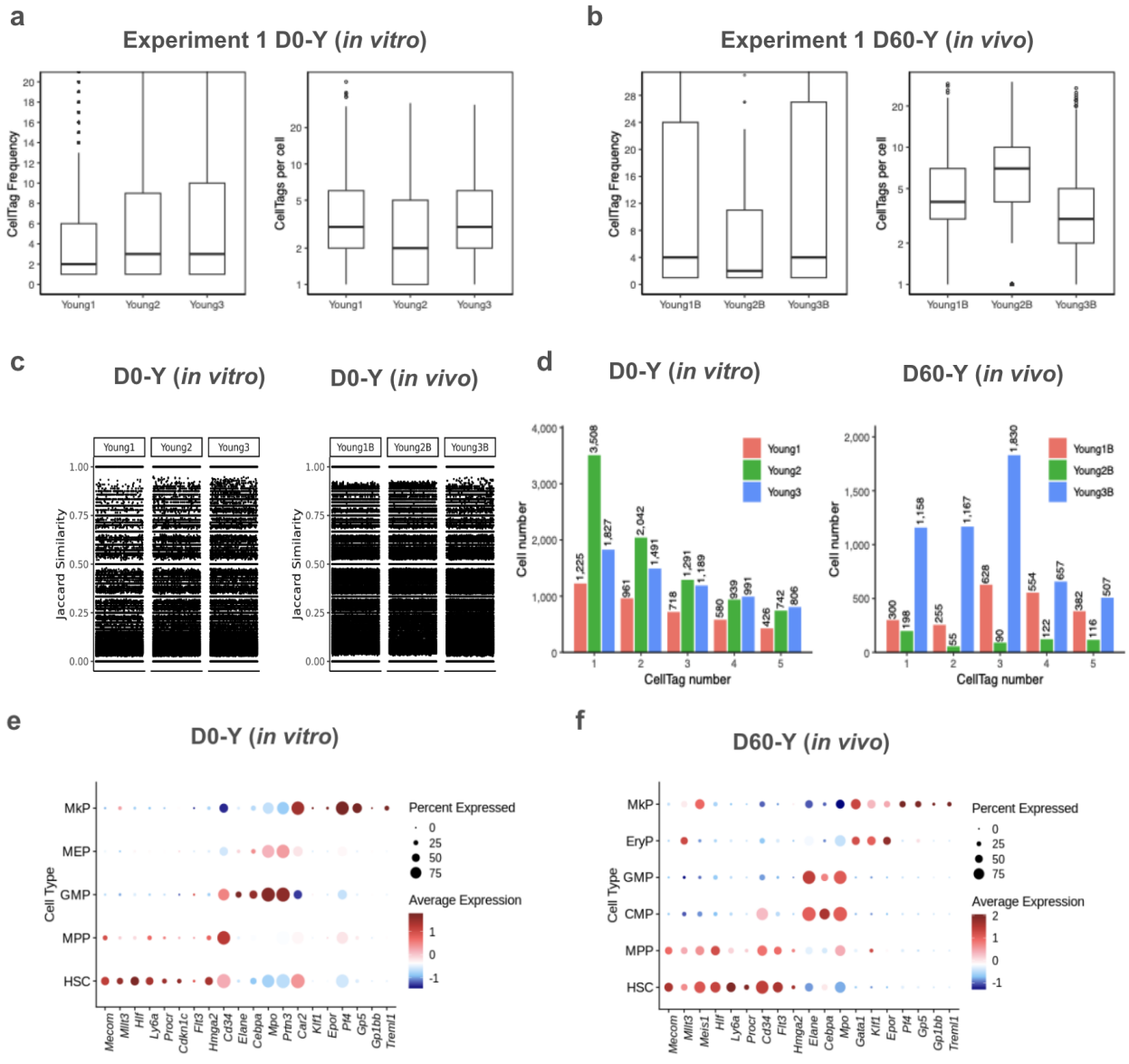
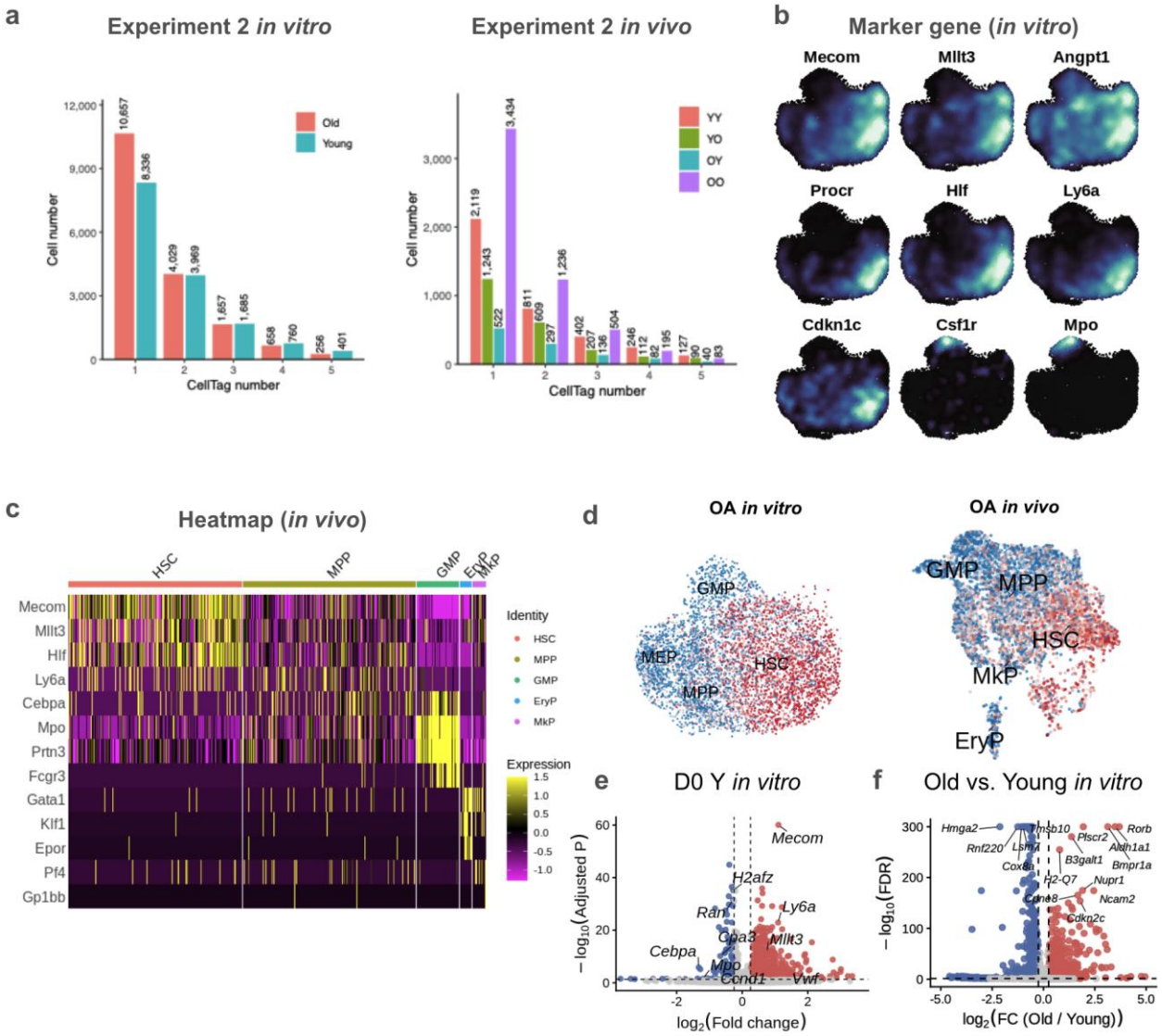
Extended Data

Extended Data Figure 1



Extended Data Fig. 1 | Robust CellTag labelling and clustering enable lineage-resolved single-cell analysis of HSC. a,b, Distribution of CellTag frequencies (number of cells sharing each barcode) and CellTags per cell demonstrating efficient and uniform barcode labelling; **a,** *in vitro* (Day 0); **b,** *in vivo* (Day 60). **c,** Distribution of pairwise Jaccard similarity between cells supporting robust clone identification. **d,** Distribution of CellTag barcodes per clone indicating stable clone size representation. **e,f,** Dot plots of canonical haematopoietic marker gene expression across scRNA-seq clusters confirming accurate cell type annotation; **e,** Day 0; **f,** Day 60.

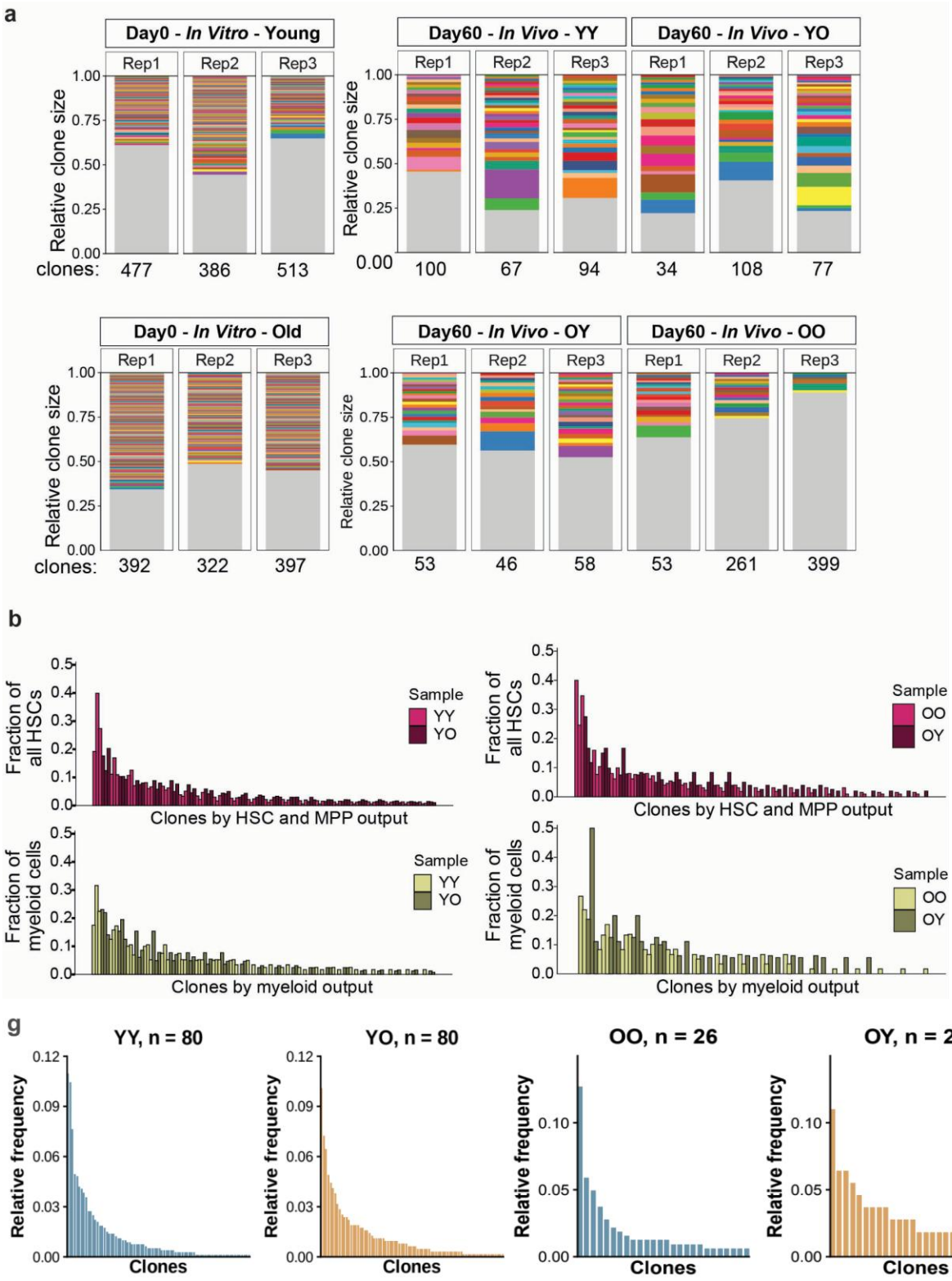
Extended Data Figure 2



Extended Data Fig. 2 | CellTag-based lineage tracing preserves clonal structure and transcriptional identity across transplantation conditions. a, Distribution of CellTag numbers per clone across young/old donor × young/old host transplantation conditions, shown separately for in vitro (left) and in vivo (right) samples, demonstrating consistent barcode representation across conditions. **b,** UMAP visualization of scaled expression of selected previously identified marker genes[1] across single cells at Day 0 (in vitro), supporting transcriptional state annotation. **c,** Heatmap showing expression of marker genes across scRNA-seq clusters at Day 60 (*in vivo*), confirming cell identity across transplantation conditions. **d,** Single-cell maps showing clonal HSC output activity (OA) values in vitro (left) and in vivo (right) in Experiment 2, illustrating functional

heterogeneity of HSC clones. **e,** Volcano plot showing differential gene expression between low-output and high-output HSC in young donors at Day 0 (in vitro) from Experiment 2. **f,** Volcano plot showing differential gene expression between young and old HSC at Day 0 (in vitro) from Experiment 2, identifying age-associated transcriptional differences. Genes with adjusted $p < 0.05$ and $|\log_2$ fold change$| > 0.25$ are highlighted.

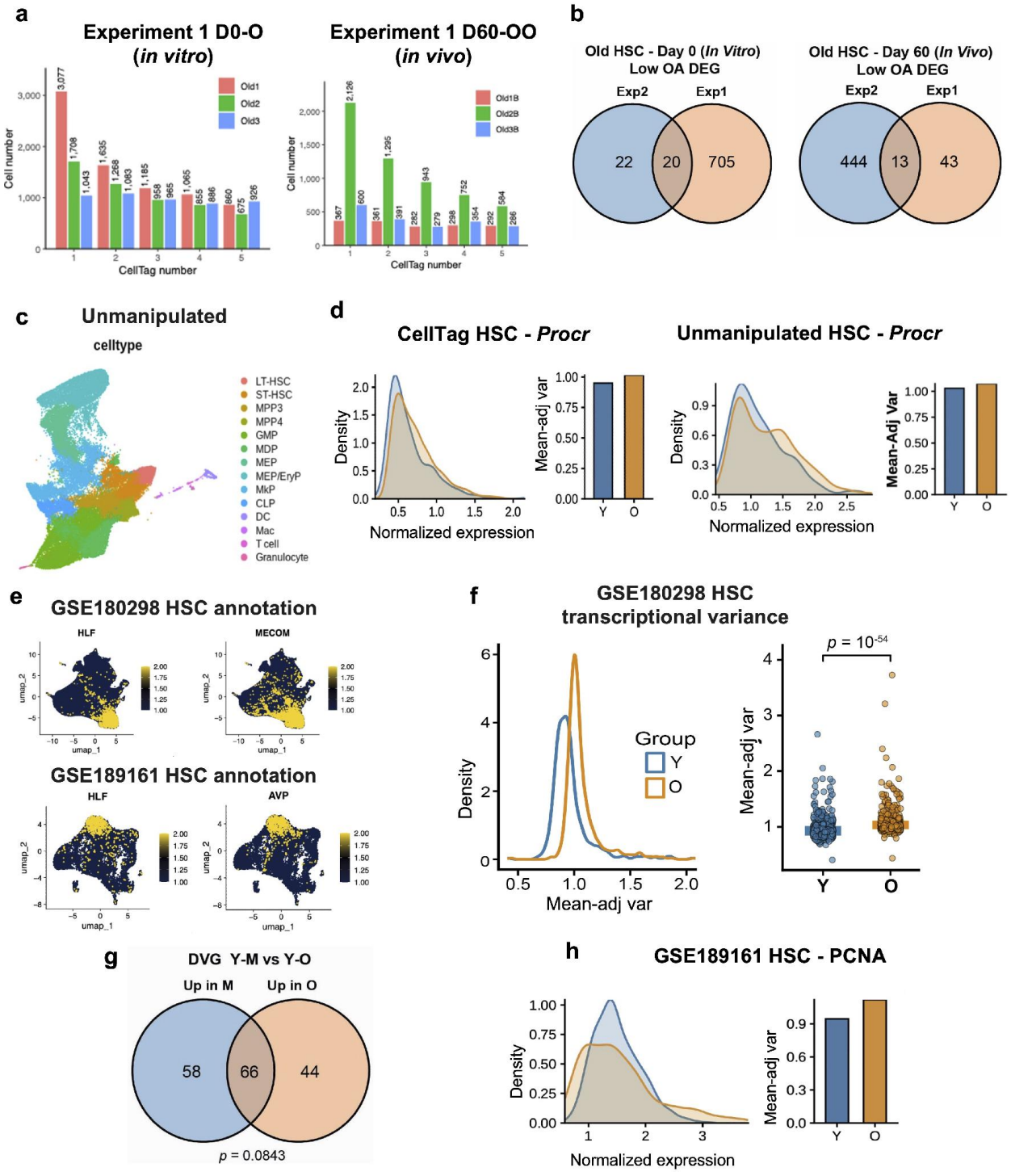
Extended Data Figure 3



Extended Data Figure 3 | Preservation of clonal architecture in different combinations of donor-host age transplantations. a, Comparison of clone sizes for young/old donor × young/old

host transplantation mice (3 biological replicates). Clones with a relative size less than 1% are shown in grey. Clone numbers for each group are indicated at the bottom of the bars. **b,** Comparison of HSC output and myeloid output for the 20 clones with the highest HSC/myeloid output between young and old hosts (3 replicates) across transplantation conditions (YY, YO, OO, OY). Statistical significance was assessed using the Wilcoxon rank-sum test between young and old hosts. **c,** Distribution of clone sizes (relative frequency) across transplantation conditions (YY, YO, OO, OY), showing preservation of clonal architecture across donor–host age combinations.

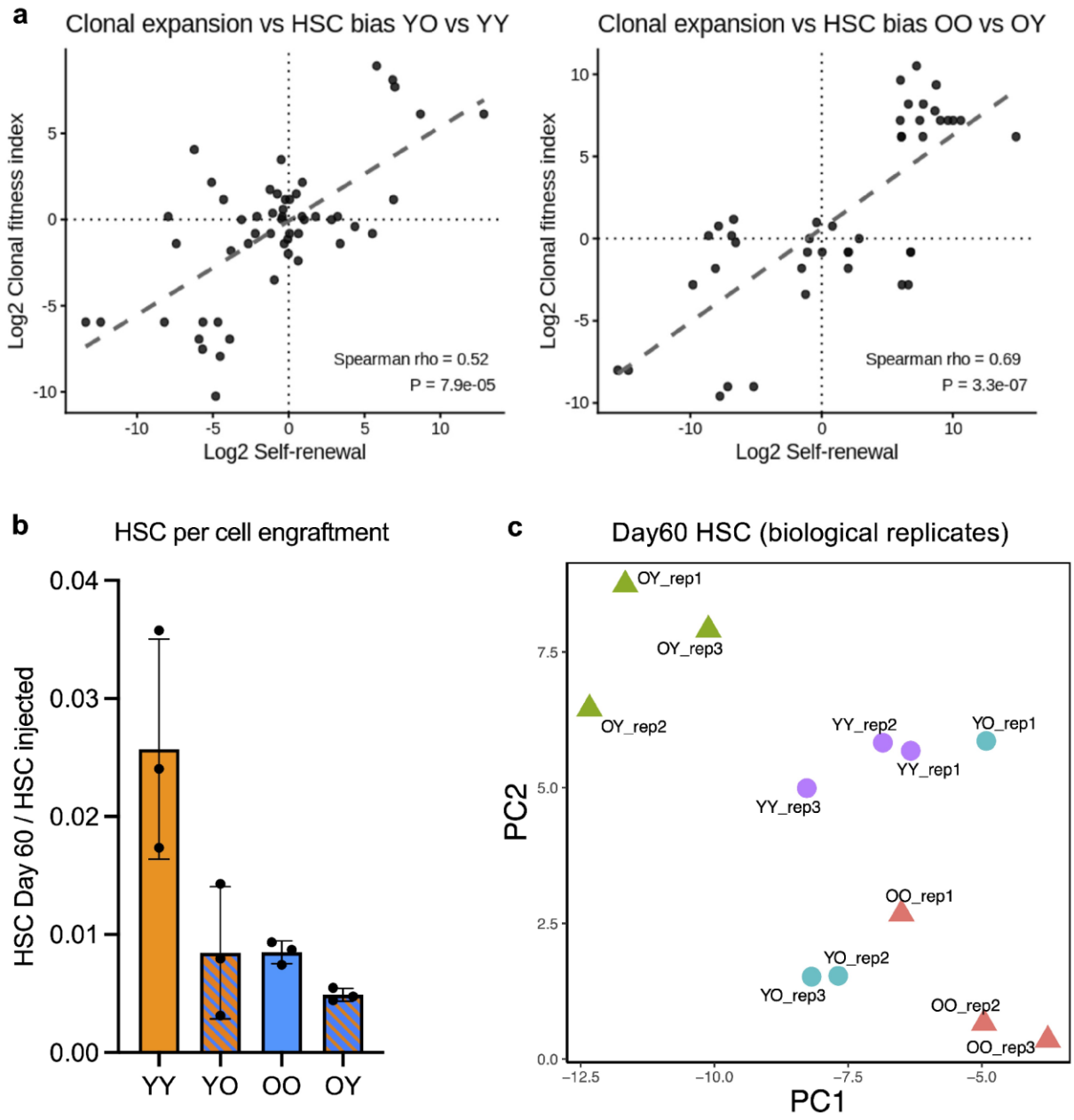
Extended Data Figure 4



Extended Data Fig. 4 | Ageing-associated transcriptional variability is reproducible across datasets and independent of mean expression. a, Distribution of CellTag barcodes per clone in old donor transplantation experiments (Experiment 1), showing consistent clonal representation in vitro (Day 0) and in vivo (Day 60). **b,** Venn diagram showing the overlap of low output DVG in

old HSC from Experiment1 vs Experiment 2 *in vitro* and day 0 (left) and *in vivo* at day 60 (right). **c,** UMAP of scRNA-seq profiles from unmanipulated haematopoietic cells across age groups, demonstrating preservation of major cell states. **d,** Distribution of *Procr* expression in HSC in CellTag Experiment 2 (Day 0, left) and unmanipulated HSC (right), with quantification of variability differences between young and old samples. **e,** Expression of canonical HSC markers in human bone marrow scRNA-seq datasets (GSE180298 and GSE189161), confirming conservation of transcriptional programs across species. **f**, Inter-cellular heterogeneity between young (Y, 18-20 y, n = 3), and old (O, >65 y, n = 5) HSC from GSE180298, left, density plot of mean adjusted variance for all genes, right, plot of mean adjusted variance for aging-associated DVG (n = 510), Statistical significance was assessed using the Wilcoxon rank-sum test **g**, Venn diagram showing overlap of up-DVG in middle age and old HSC from GSE189161. **h,** Example of an age-associated differentially variable gene (DVG; PCNA) in human bone marrow data, illustrating independence of expression variance from mean expression.

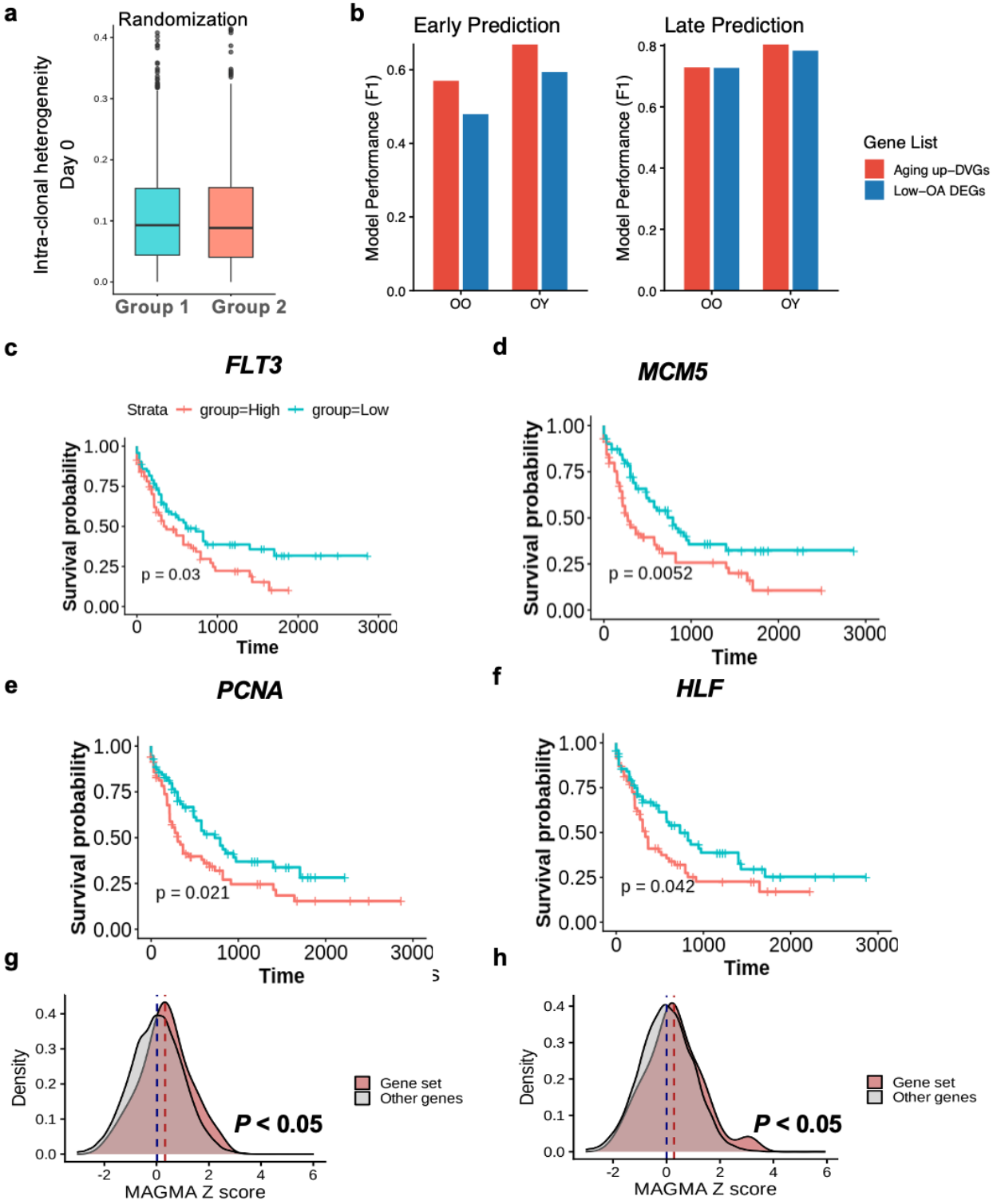
Extended Data Figure 5



Extended Data Fig. 5 | Heterochronic transplantations modulate HSC clonal expansion dynamics. a, Relationship between differential self-renewal and differential clone fitness index (CFI) across shared clones, showing consistency of clonal behaviour across environments (Experiment 2). **b**, Number of engrafted immunophenotypic HSC per five million bone marrow cells normalized to number of immunophenotypic HSC injected by donor and host age. **c**, Replicate-level pseudo-bulk in PC1-PC2 PCA space of HSC transcriptomes at day 60. Each point represents one biological replicate (n = 12 mice; three per group: OO, OY, YO, YY). Colours, transplant group; shapes, donor age (triangle, old; circle, young). PERMANOVA on Euclidean

distances using PC1-PC30 showed significant global separation among groups (R^2 = 0.72, P = 1.0×10^{-4}) and a significant donor × host-age interaction (R^2 = 0.28, p = 1.0×10^{-4}), supporting an asymmetric, donor-age–dependent transcriptional response to mismatched microenvironments. PERMDISP indicated homogeneous within-group dispersion (F = 0.10, p = 0.95). Pairwise comparisons did not reach significance after Benjamini–Hochberg correction (all p adj = 0.10), reflecting the limited resolution of permutation-based tests with three replicates per group.

Extended Data Figure 6



Extended Data Fig. 6 | Variability-associated gene programs predict clonal behaviour and are linked to human disease risk. a, Null model showing no association between intra-clonal heterogeneity at Day 0 and clone size at Day 60 under random assignment, demonstrating

specificity of observed relationships. **b,** Random forest model performance (macro F1 scores) from 5-fold cross-validation for predicting self-renewal status using in vitro (Day 0) gene expression across different gene sets; left, early prediction; right, late prediction. **c-f,** Survival analysis of the TCGA acute myeloid leukaemia (AML) cohort based on the expression of individual genes derived from the DVG feature set, demonstrating their clinical relevance. **g,h,** Enrichment of age-associated DVG in genome-wide association studies (GWAS); **g,** TET2 clonal haematopoiesis (CHIP); **h,** myelodysplastic syndrome (MDS), showing overlap between variability-associated genes and disease-associated loci.

Supplementary Methods

Extended CellTag clone filtering and quality control

CellTag barcodes were subjected to additional filtering steps beyond initial clone calling to ensure robustness of downstream clonal analyses. Only clones containing at least two cells were retained for any analysis involving intra-clonal comparisons. To minimize the influence of barcode noise and sampling artifacts, clones with extremely low abundance at the post-transplant timepoint were excluded. Specifically, clones with a relative frequency below 0.005 at the post-transplant timepoint were removed from analyses of self-renewal, clonal fitness, and heterogeneity.

For longitudinal analyses, only clones detected at both Day 0 (pre-transplant) and the post-transplant endpoint were considered. This restriction ensured that all reported clonal behaviors reflected true persistence and expansion rather than de novo barcode detection or dropout. Clone sizes were normalized by the total number of cells recovered per sample to enable comparisons across mice and transplantation conditions.

Definition and quantification of clonal behaviours

Clonal self-renewal

Clonal self-renewal was quantified using output activity, defined as the ratio of non-HSC output to HSC persistence for each clone.

Clonal fitness index

Clonal fitness was defined as the fold-change in clone size between Day 0 and the post-transplant timepoint. Clone sizes were normalized by the total number of cells recovered per sample, and fitness values were log2-transformed for analysis. Only clones detected at both timepoints and containing at least two cells at Day 0 were included in fitness analyses. Clones were stratified into high- and low-fitness groups based on the distribution of log2 fold-change values.

Clonal size

Clonal size was defined as the number of cells per clone at a given timepoint. For descriptive analyses, clones were categorized based on relative frequency as small (<0.01), medium (0.01–

0.05), or large (>0.05). Clones with fewer than two cells were excluded from any analysis involving transcriptional variability or state composition.

Quantification of intra-clonal transcriptional heterogeneity

Intra-clonal transcriptomic heterogeneity was quantified using two complementary approaches. First, transcriptional dispersion within each clone was measured as the average pairwise Euclidean distance between cells in the top 30 principal components derived from scRNA-seq data. Second, compositional heterogeneity was assessed using normalized Shannon entropy calculated from the distribution of annotated cell states within each clone.

To ensure statistical reliability, heterogeneity metrics were computed only for clones containing at least two cells, and in some analyses restricted to clones with at least five cells. When comparing heterogeneity across clones of different sizes, entropy and dispersion metrics were normalized by clone size.

Differential gene expression analyses

Differential gene expression analyses were performed using MAST. For clonal analyses, gene expression at Day 0 was compared between groups of clones stratified by subsequent clonal behaviour, including self-renewal status and clonal fitness. Genes with adjusted *P* values below the specified threshold were considered differentially expressed. Gene set enrichment analyses were performed using curated pathway databases and gene ontology annotations to interpret biological programs associated with distinct clonal outcomes.

Differential gene expression variance analyses

To identify genes exhibiting differential cell-to-cell expression variability independent of mean expression, we performed differential gene expression variance analyses across age groups, clonal outcome groups, and cellular states.

For each gene, expression values were first normalized at the single-cell level using the same normalization procedure applied for differential expression analyses. Variance was computed within defined cell populations (for example, HSC only) after controlling for differences in mean expression. Genes with low average expression less than 0.1 in log-normalized scale and genes that expressed in less than 10 percent of the total cells in both comparison groups were excluded to avoid inflation of variance estimates due to technical noise.

Differential variance between groups was assessed by comparing gene-wise variance estimates using a regression-based framework that models variance as a function of biological condition while accounting for mean–variance dependence. Statistical significance was determined using likelihood ratio tests, and *p* values were corrected for multiple hypothesis testing using the Benjamini–Hochberg procedure.

Genes with significantly increased or decreased variance were referred to as differentially variable genes (DVG). DVG were analysed separately from differentially expressed genes to distinguish changes in transcriptional heterogeneity from shifts in mean expression. Gene set enrichment analysis was performed on DVG to identify biological pathways associated with altered transcriptional variability during aging and across clonal outcomes.

Functional analysis of differentially variable gene

Gene set enrichment analyses were conducted using R package clusterProfiler (3.10.1)[2] We leveraged rank-based GSEA and over-representation analysis (ORA) to test for pathway enrichment of differential expressed gene and differential variable gene. We leveraged PROGENy (1.17.3)[3] to infer per-cell pathway activity scores in the human single-cell dataset. Spearman correlation analysis between cell-wise DVG module scores and pathway activity was performed using three DVG gene sets (human-specific, mouse-specific and shared) applied to the human dataset. A more positive correlation coefficient (ρ) indicates higher expression of the pathway-related genes mapped to input DVG as an indication of the pathway enrichment.

To estimate the effect of differentially variable gene on genetic susceptibility of clonal-haematopoiesis (CH), MAGMA (v1.10)[4] was applied to test for differential variable gene (DVG) enrichment in CHIP, CHIP-DNMT3A, CHIP-TET2[5], MDS[6], and AML[6] GWAS. Mouse genes were converted to human orthologs using the MGI mapping file[7], and corresponding human Entrez IDs were used for MAGMA gene-set analysis. GWAS datasets included in this analysis were obtained from the NHGRI-EBI catalogue.

Machine learning-based prediction of clonal outcomes

Random forest supervised machine learning models were used to predict clonal fitness and self-renewal outcomes from Day 0 transcriptional profiles. Feature matrices were constructed using either highly variable genes or curated gene signatures derived from differential expression

analyses. Individual cells were used as the unit of cross-validation. Clone-level cross-validation was not feasible due to the limited number of clones available for model training.

Model performance was evaluated using held-out test sets and permutation-based null models. Feature importance and model stability were assessed across cross-validation folds to ensure robustness of predictive signals.

Statistical analysis

All statistical analyses were performed using R and Python. Continuous variables were compared using non-parametric or parametric tests as appropriate, and categorical variables were compared using contingency-based tests. Multiple hypothesis testing was corrected using the Benjamini–Hochberg procedure unless otherwise stated. Exact statistical tests, sample sizes, and significance thresholds are reported in the corresponding figure legends.

Analytical tool of clonal heterogeneity

To facilitate the estimation of clonal heterogeneity, we developed scCloneVar, an analytical R package that enables clonal statistical assessments, particularly at the clone-level and cell-level transcriptional heterogeneity. scCloneVar assigns each clone intra- and inter-clonal Euclidean distances derived from embeddings of high-dimensional single-cell data and identifies differential variable genes (DVG) using a mean-adjusted approach[8]. In addition, scCloneVar supports quantification of clonal behaviour and visualizing of both global and local heterogeneity patterns and related functional implications.

Extended Data Table 1 | Mouse details and aging cohorts.

Extended Data Table 2 | Experiment 1 sorting and in vitro (Day 0) scRNAseq.

Extended Data Table 3 | Experiment 1 in vitro (Day 0) sorting for scRNAseq.

Extended Data Table 4 | Experiment 1 in vivo (Day 60) sorting for scRNAseq.

Extended Data Table 5 | Experiment 2 sorting and QC.

Extended Data Table 6 | Experiment 2 in vitro (Day 0) sorting for scRNAseq.

Extended Data Table 7 | Experiment 2 in vivo (Day 60) sorting for scRNAseq.

Extended Data Table 8 | Clone-level statistics across aging contexts.

Extended Data Table 9 | Numbers of differential gene expression associated with clonal output activity.

Extended Data Table 10 | Gene set and pathway enrichment associated with differential variability programs (Experiment 2).

Extended Data Table 11 | Gene set and pathway enrichment associated with differential variability programs (shared by Experiment 2 and Unmanipulated datasets).

Extended Data Table 12 | Overlap of differentially variable genes across mouse and human datasets.

Extended Data Table 13 | Gene set and pathway enrichment associated with differential variability programs (Human_GSE189161).

Extended Data Table 14 | Clonal self-renewal and clonal fitness index (CFI) metrics of old mice.

Extended Data Table 15 | Clonal self-renewal and clonal fitness index (CFI) metrics of young mice.

Extended Data Table 16 | Predictive model features and performance metrics for clonal fitness classification.

Extended Data references

1 Rodriguez-Fraticelli, A. E. *et al.* Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. *Nature* **583**, 585-589 (2020). https://doi.org/10.1038/s41586-020-2503-6

2 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16**, 284-287 (2012). https://doi.org/10.1089/omi.2011.0118

3 Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun* **9**, 20 (2018). https://doi.org/10.1038/s41467-017-02391-6

4 de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015). https://doi.org/10.1371/journal.pcbi.1004219

5 Kar, S. P. *et al.* Genome-wide analyses of 200,453 individuals yield new insights into the causes and consequences of clonal hematopoiesis. *Nat Genet* **54**, 1155-1166 (2022). https://doi.org/10.1038/s41588-022-01121-z

6 Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association tool for biobank-scale data. *Nat Genet* **53**, 1616-1621 (2021). https://doi.org/10.1038/s41588-021-00954-4

7 Baldarelli, R. M., Smith, C. L., Ringwald, M., Richardson, J. E. & Bult, C. J. Mouse Genome Informatics: an integrated knowledgebase system for the laboratory mouse. *Genetics* **227** (2024). https://doi.org/10.1093/genetics/iyae031

8 Eling, N., Richard, A. C., Richardson, S., Marioni, J. C. & Vallejos, C. A. Correcting the Mean-Variance Dependency for Differential Variability Testing Using Single-Cell RNA Sequencing Data. *Cell Syst* **7**, 284-294.e212 (2018). https://doi.org/10.1016/j.cels.2018.06.011