**Supplementary information**
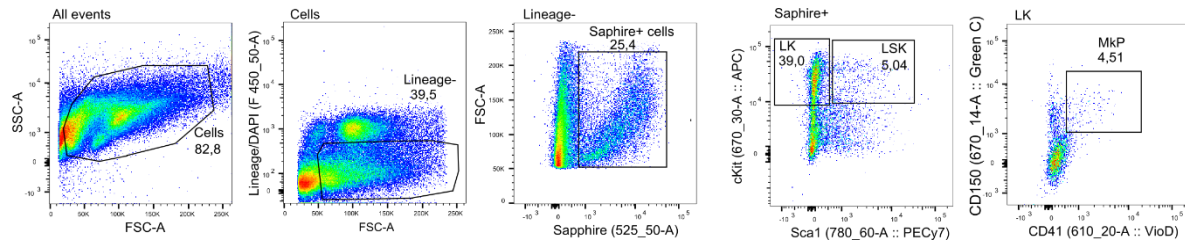
# Clonal tracing with somatic epimutations reveals dynamics of blood ageing

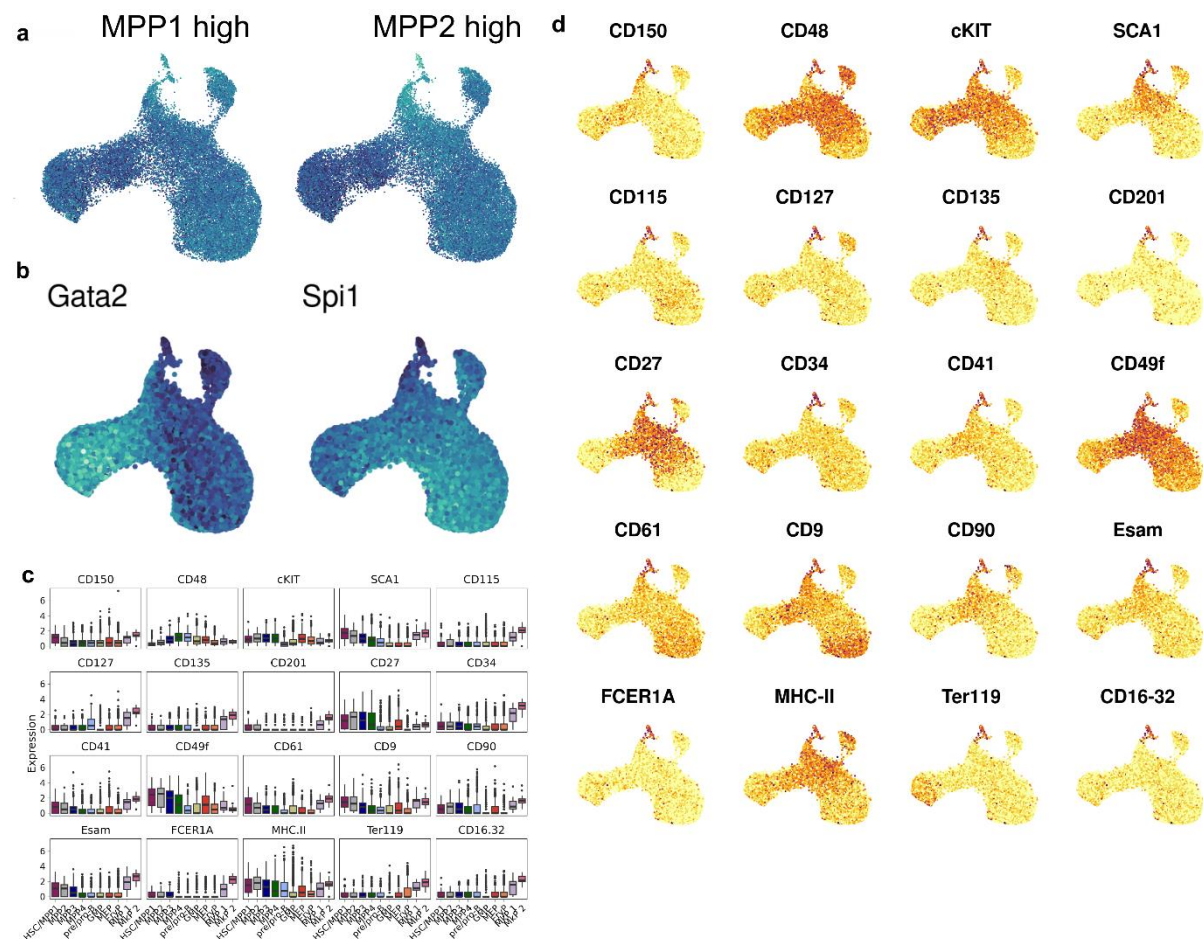In the format provided by the
authors and unedited

**Supplementary figures for:**

*Clonal tracing with somatic epimutations reveals dynamics of blood aging*

Michael Scherer, Indranil Singh, Martina Maria Braun, Chelsea Szu-Tu, Pedro Sanchez Sanchez, Dominik Lindenhofer, Niels Asger Jakobsen, Verena Körber, Michael Kardorff, Lena Nitsch, Pauline Kautz, Julia Rühle, Agostina Bianchi, Luca Cozzuto, Robert Frömel, Sergi Beneyto-Calabuig, Caleb Lareau, Ansuman T. Satpathy, Renée Beekman, Lars M. Steinmetz, Simon Raffel, Leif S. Ludwig, Paresh Vyas, Alejo Rodriguez-Fraticelli and Lars Velten
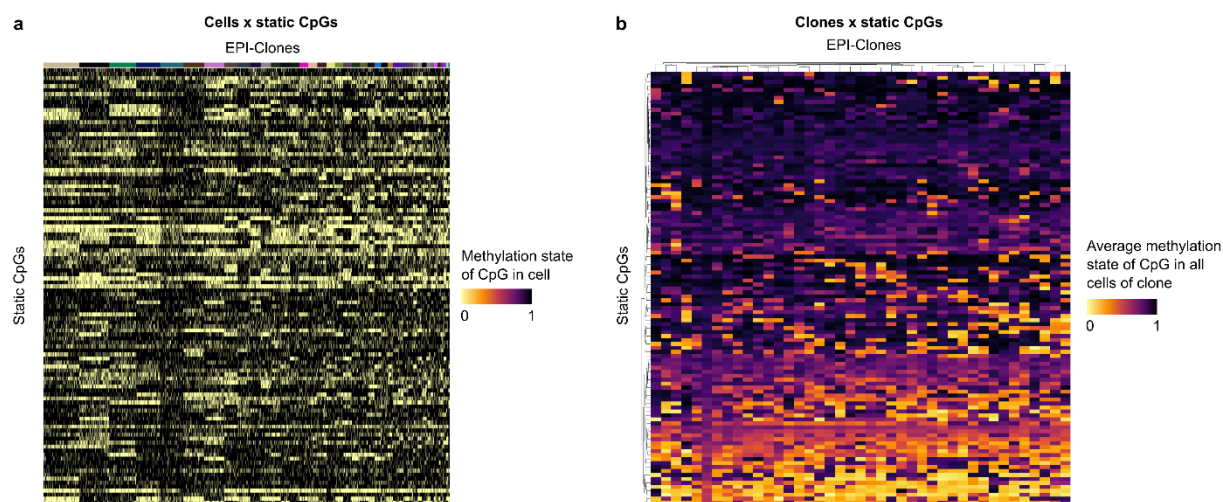
**Supplementary Fig. 1: FACS scheme employed for the main LARRY experiments (M.1, M.2).** This figure shows scatterplots of various markers (Lineage, cKIT, SCA-1, CD41, CD150).
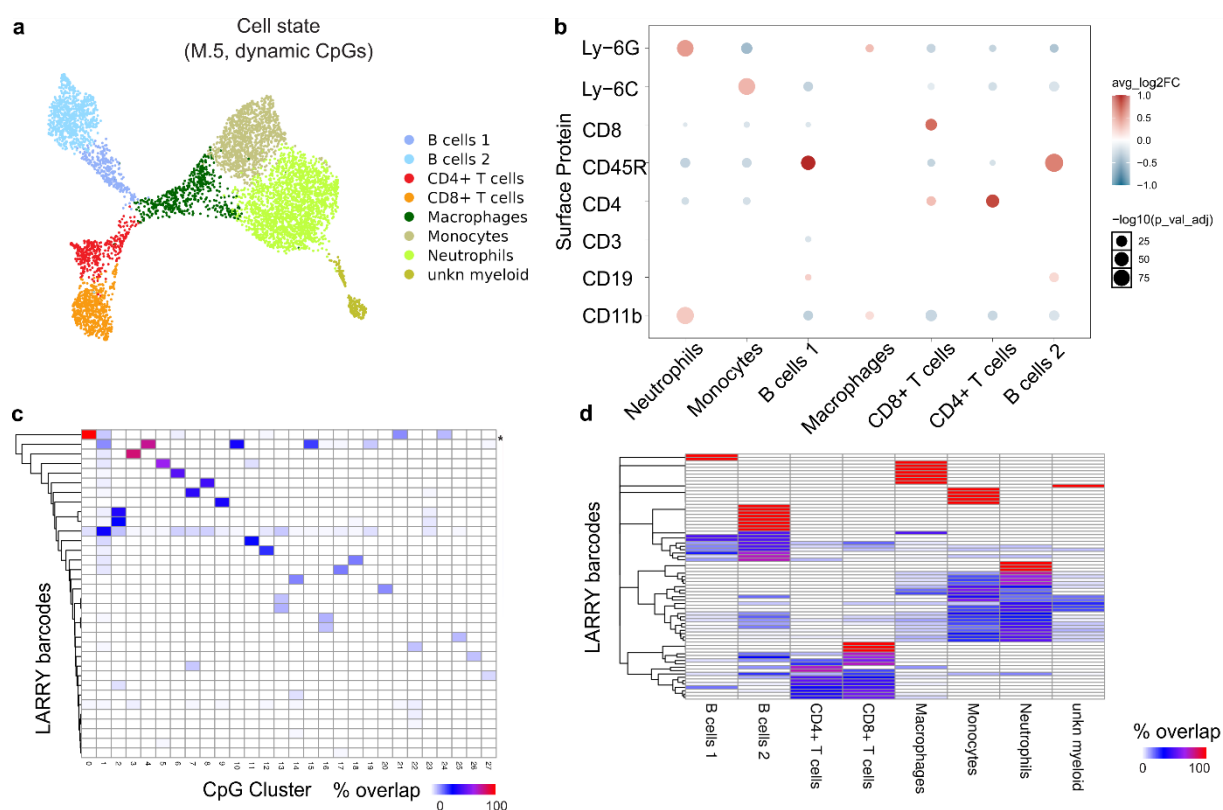


**Supplementary Fig. 2: ScTAM-seq delivers a high-resolution map of murine hematopoiesis (experiment M.1-M.3). A.** Relative methylation state of cells in amplicons specifically methylated in MPP1/MPP2 in bulk data. The UMAP is the integrated UMAP across experiments M.1-M.3 as shown in Figure 1c. **B.** Relative methylation state of cells in amplicons in the vicinity

of Gata2/Spi1 binding sites. **C.** Expression of 20 surface proteins across all the cell types as boxplots. The expression values were normalized using the centered-log-ratio normalization. **D.** The integrated UMAP generated from all CpGs as in Figure 1c is colored by the expression of 20 surface proteins measured using oligonucleotide-tagged antibodies.
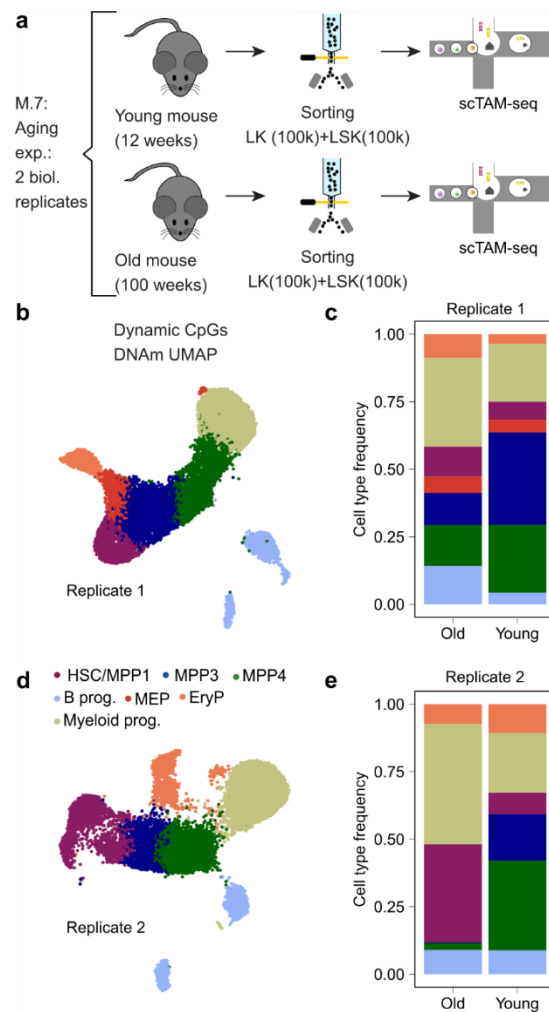


**Supplementary Fig. 3: DNA methylation states of static CpGs in cells. A.** DNA methylation heatmap indicating the methylation state of static CpGs (rows) in all cells of experiment M.1. Cells are grouped by EPI-clones. **B.** Average DNA methylation state of each EPI-clone for the static CpGs for experiment M.1. The color indicates the mean methylation value of all cells in a specific EPI-clone (column) for the static CpGs.
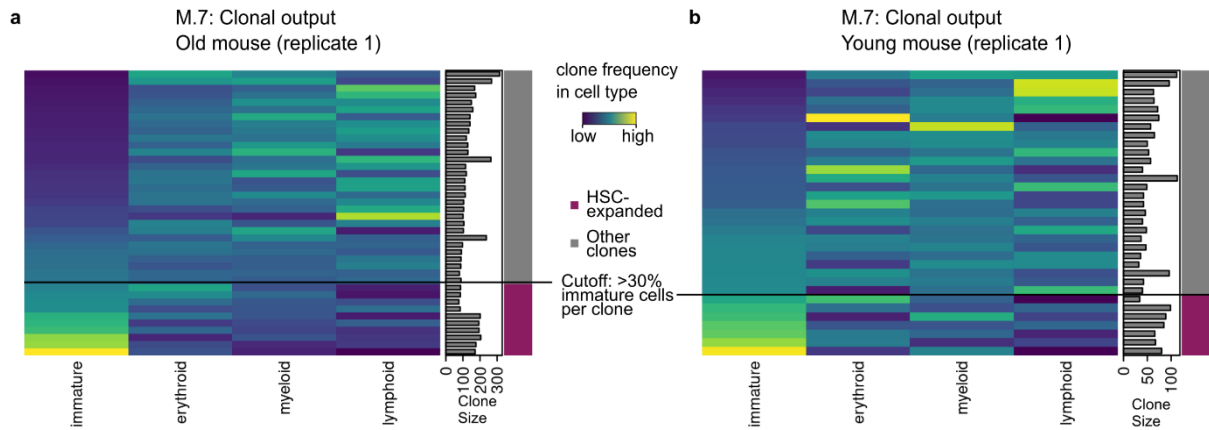


**Supplementary Fig. 4: Application of EPI-clone on mature immune cells (experiment M.5). A.** DNA methylation UMAP of the dynamic CpGs defined on experiment M.1, for the mature
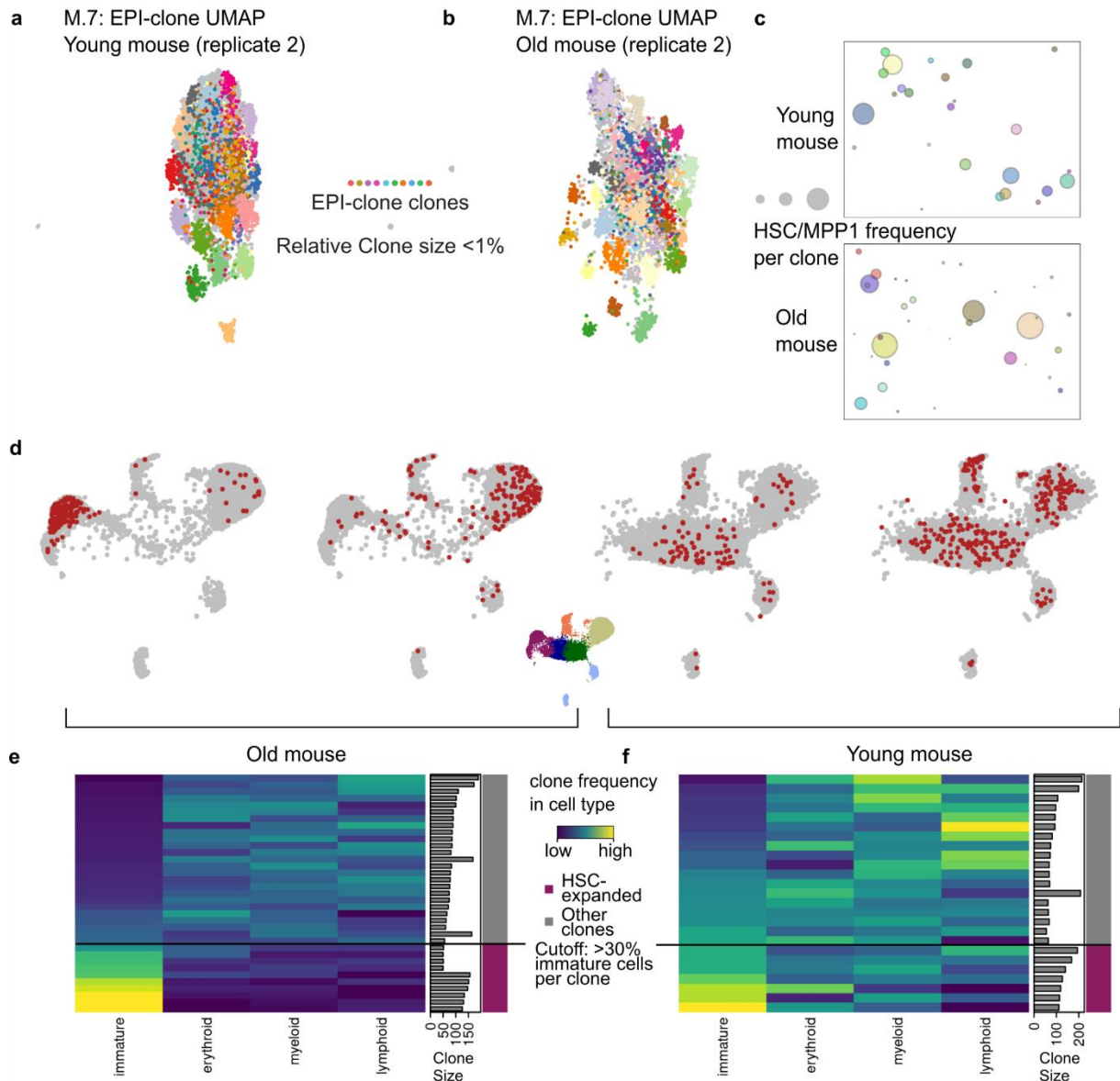
immune cells (experiment M.5). Colors indicate different cell type annotation identified by surface protein expression. **B.** Differential surface protein expression between the different cell type clusters identified in panel a. Shown is the average log2 fold change between the clusters together with the adjusted p-value of a Wilcoxon test. The surface protein data has been normalized using SCTransform. **C.** Heatmap visualizing the overlap between EPI-clone clusters (x-axis) and the LARRY barcodes (y-axis). **D**. Heatmap representing the distribution of different cell types in clones defined through LARRY barcodes (y-axis). The color code indicates the frequency of each of the cell types in the clones defined by LARRY barcodes.
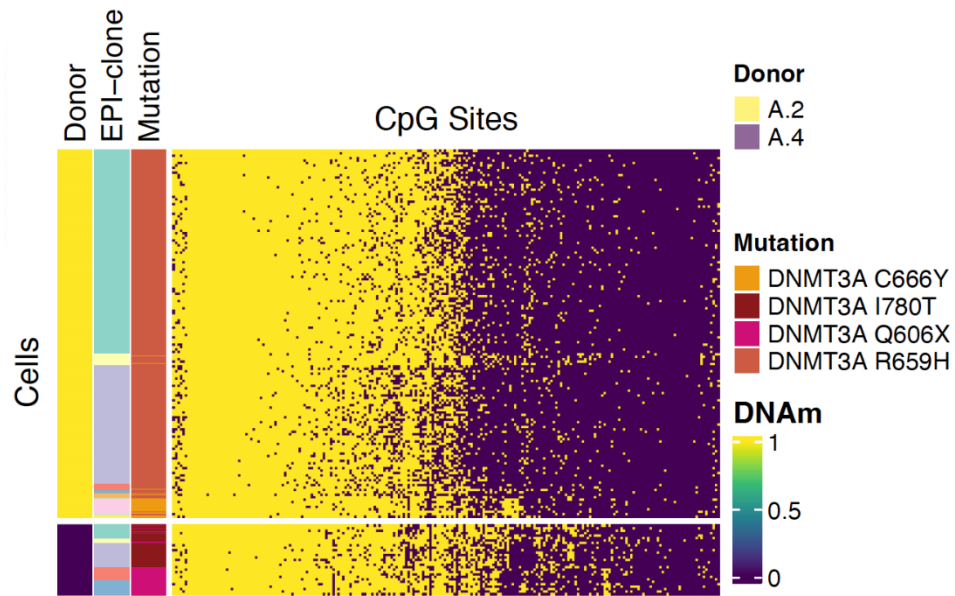


**Supplementary Fig. 5: Cell type distributions in the old versus the young mouse (experiment M.7). A.** Overview of the experimental design for M.7: aging experiment**. B, D.** UMAP based on the dynamic CpGs identified in experiment M.1, depicting the cell types identified in the aging experiment (combing young and old mouse, **B**: replicate 1, **D**, replicate 2). **C, E.** Comparison of cell type distributions in the old versus the young mice for the aging experiments (**C**: replicate 1, **E**: replicate 2). The scTAM-seq schematic shown in A was adapted from ref. [5], BioMed Central, under a Creative Commons licence CC BY 4.0.
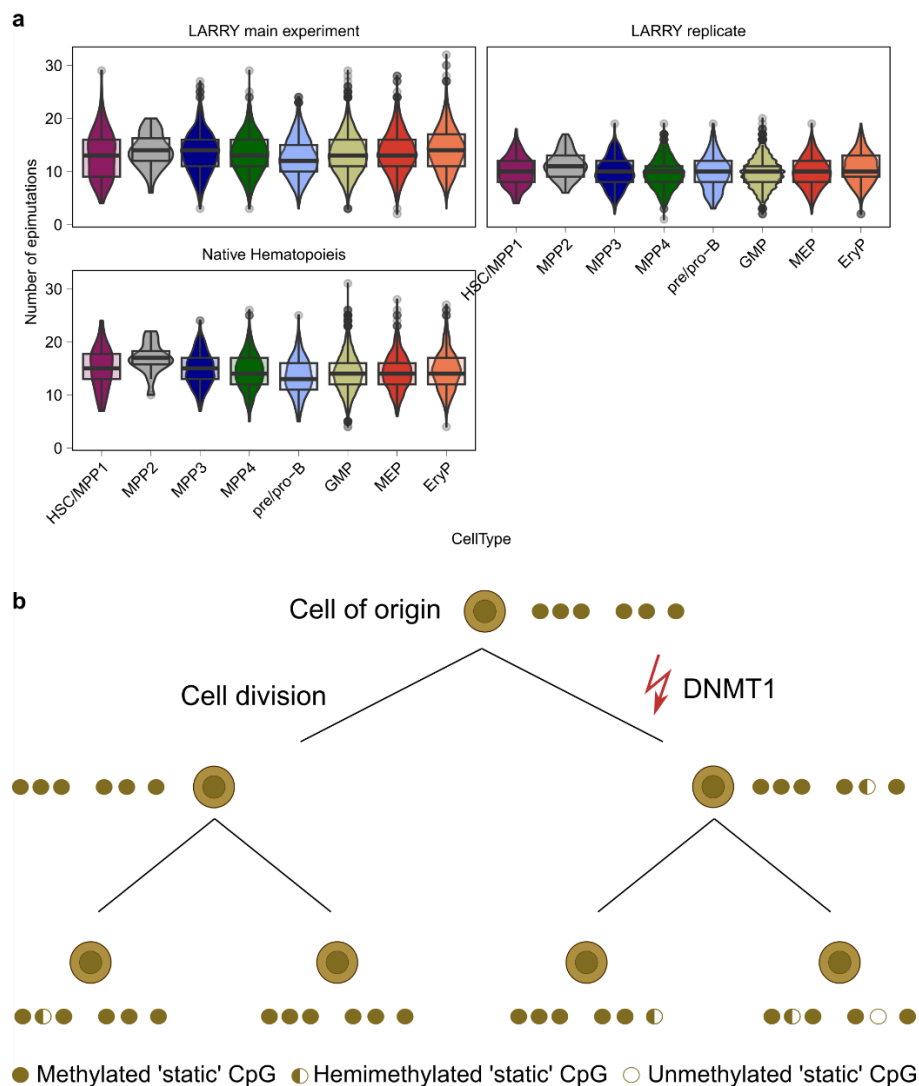
**Supplementary Fig. 6: Clonal output for the different cell type clusters for the aging experiment (experiment M.7; data shown for replicate 1). A, B.** Characterizing clonal behavior of all clones (rows) across the different cell types (columns) as a heatmap from aging replicate 1. The color indicates the relative frequency of the clones toward four main differentiation trajectories. HSC-expanded clones are determined by setting a cutoff of at least 30% immature cells in the clone. Immature=HSC/MPP1, MPP2; myeloid=MPP4, myeloid progenitors; lymphoid=B cell progenitors; erythroid=MEP, erythroid progenitor. This is shown for the old (A) and young (B) mouse separately.
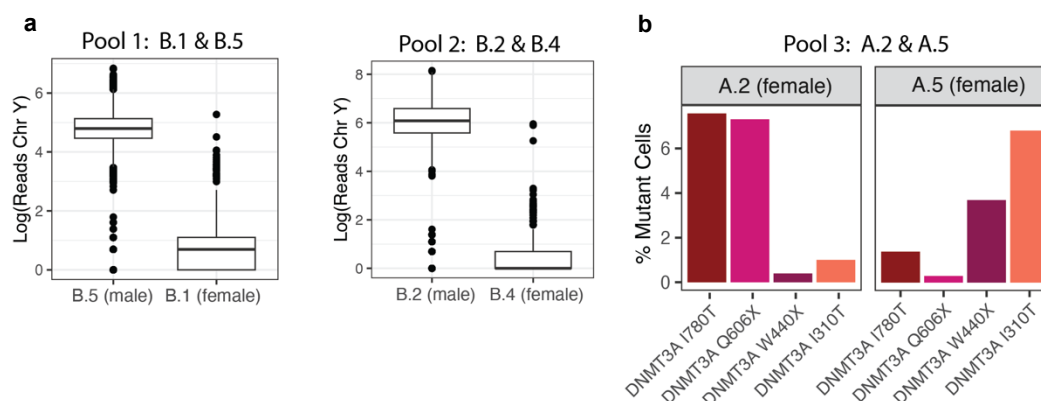
**Supplementary Fig. 7: Investigating murine aging in a biological replicate (experiment M.7; data shown for replicate 2). A.** DNA methylation UMAP based on the static CpGs for a native, young (12 weeks) mouse. **B.** DNA methylation UMAP based on the static CpGs and the associated EPI-clone clusters for the old mouse. **C**. Bubble plot visualizing the frequency (measured as the square of the HSC/MPP1 frequency) of HSC/MPP1 cells per clone for the old versus the young mouse. **D.** Visualizing clones in the differentiation UMAP (DNA methylation) for the old (left) and young (right) mouse from replicate 2. Two examples of clones are shown for both mice. **E, F.** Characterizing clonal behavior of all clones (rows) across the different cell types (columns) as a heatmap from aging replicate 2. The color indicates the relative frequency of the clones toward four main differentiation trajectories. HSC-expanded clones are determined by setting a cutoff of at least 30% immature cells in the clone Immature=HSC/MPP1, MPP2; myeloid=MPP4, myeloid progenitors; lymphoid=B cell progenitors; erythroid=MEP, erythroid progenitor. This is shown for the old (E) and young (F) mouse separately.

**Supplementary Fig. 8: Single CpG level DNA methylation for donors with DNMT3A mutations.** Cells carrying DNMT3A mutations (rows) are grouped by donor and EPI-Clone.

**Supplementary Fig. 9: Potential origin of epimutations. A.** Number of epimutations per cell type for the experiment M.1-M.3. An epimutation is defined as a CpG that shows a DNA methylation state different from the default (i.e. most prevalent across all cells) DNA methylation state of this CpGs according to all profiled cells. **B.** Model for the origin of epimutations. From a fully methylated default state in the cell of origin, epimutations randomly occur as loss of methylation. One potential explanation is the insufficient function of DNMT1 for copying the methylation state to the daughter strand, causing a methylation loss on one of the strands. In a subsequent cellular division, DNA methylation is then lost on both strands.



**Supplementary Fig. 10: Validation of SNP-based demultiplexing.** For donors that were pooled on a single Mission Bio tapestri run, vireo[53] was used for donor deconvolution based on germline SNVs in autosomes. **A.** In the case of male-female pools, chromosome Y was used to validate this deconvolution. **B.** in the case of a female-female pool, somatic variants previously described as specific to the patients[43] were used. The two variants on the left were specific to donor A.2, the two variants on the right were specific to donor A.5.