

## **SUPPLEMENTARY DATA LEGENDS**

### **Supplementary DATA 1: List of guide genes as determined by ICGS algorithm.**

List of genes on which the heatmap shown in [Fig. 3b](#) is based.

### **Supplementary DATA 2: List of genes differentially expressed between 49f<sup>+</sup> Subset1 and 49f<sup>+</sup> Subset2.**

Shown are genes with FDR <0.05 as determined by DESeq2. Of note, CD34 and CLEC9A expression were respectively 1.47 fold higher in 49f<sup>+</sup> Subset2 and 1.9 fold higher in 49f<sup>+</sup> Subset1, but with FDR >0.05.

### **Supplementary DATA 3: List of genes differentially expressed between Subset1 and Subset2.**

Shown are genes with FDR <0.05 as determined by DESeq2. Of note, CD34 and CLEC9A expression were respectively 1.92 fold higher in Subset2 and 5.67 fold higher in Subset1.

### **Supplementary DATA 4: lineage composition of human grafts in mice transplanted with Subset1, Subset2, 49f<sup>+</sup> Subset1 and 49f<sup>+</sup> Subset2 cells.**

Shown is the percentage of human engraftment and relative graft composition as percentage of My, Ly and Ery cells from engrafted mice at the indicated doses and weeks from the indicated groups.