CellPress
OPEN ACCESS

# Supplemental figures

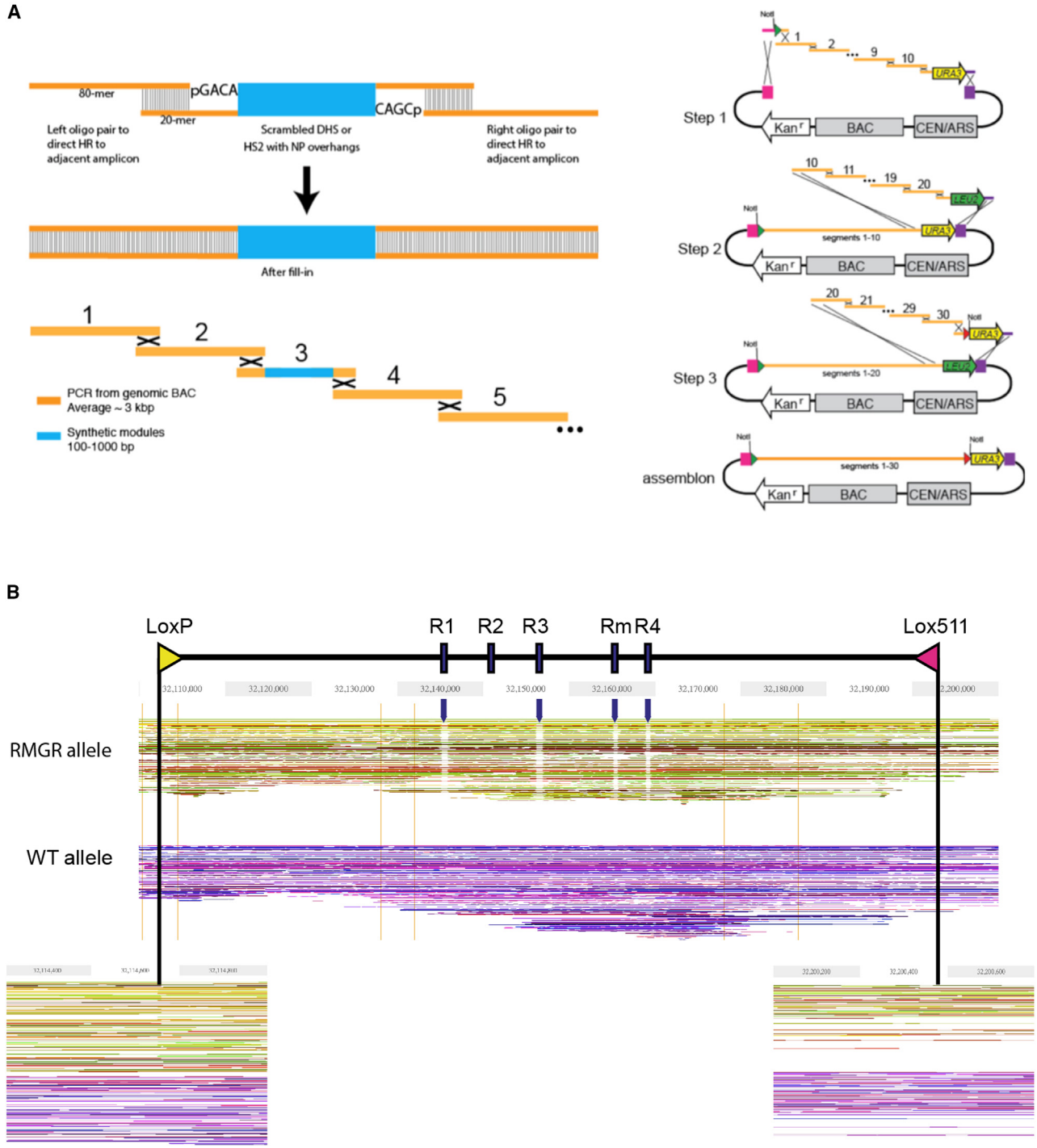


**Figure S1. Generation and validation of the R2-only *α-globin* model locus generated to test the sufficiency of the R2 enhancer element, related to Figure 1**

(A) Short segments of DNA are generated by PCR from a target backbone or ordered as synthetic constructs with matching overhangs to allow for homologous recombination. Homologous recombination is performed in yeast with blocks of up to 10 segments in a single step. Alternating selection for the *URA3* and *LEU2*

*(legend continued on next page)*

genes is used to perform subsequent assembly steps. The resulting BAC contains the assembled locus, modules necessary for replication in both yeast and *E. Coli* (BAC/CEN/ARS, respectively), and selectable markers including the kanamycin resistance gene (Kanr).
(B) Top: schematic of the RMGR region with enhancer elements and RMGR lox exchange sites annotated. Middle: phased linked-reads for the R2-only RMGR and WT alleles, visualized with proprietary Loupe software from 10X Genomics. Enhancer element deletions are clearly visible by sequencing dropout in the RMGR allele, as indicated by black arrows. Bottom: junctions where lox sites remain in the R2-only allele are seen as mapped read ends in the RMGR allele, but not in the WT allele where there are no lox sites.

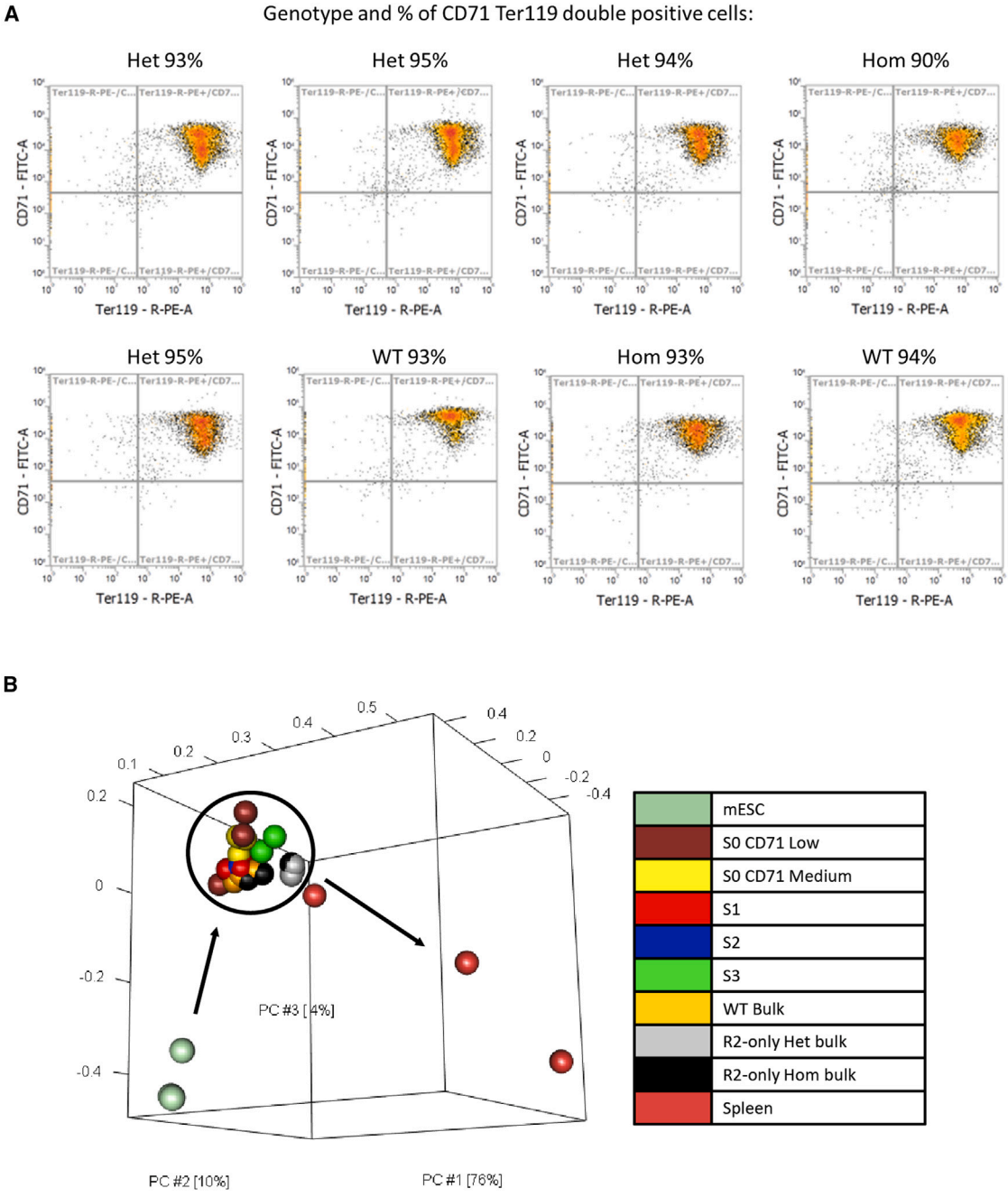


**Figure S2. Fetal liver erythroid cell population characterization from WT and R2-only heterozygous and homozygous E12.5 mouse embryos, related to Figure 2**

(A) Immunophenotypic characterization of the erythroid populations by FACS using standard erythroid-specific markers (CD71 and Ter119) derived from E12.5 WT and heterozygous and homozygous R2-only derived FL cells. Percentage of CD71+/Ter119+ cells is displayed and shows equivalent proportions in WT, R2-only heterozygous, and R2-only homozygous embryos.

(B) Principal component analysis (PCA) comparing genome-wide ATAC-seq peaks in WT, R2-only homozygous, and R2-only heterozygous FL erythroid cells alongside mESCs, developmentally staged FL erythroid cells (S1, S2, and S3 representing erythroid cells from early progenitors to more differentiated erythroid cells, respectively), and adult spleen erythroid cells. ATAC-seq peaks clustering together (circled in black) show equivalent genome-wide open chromatin landscape in all the FL erythroid cells analyzed, irrespective of their genotype, indicating no abnormal erythropoiesis in the analyzed populations.

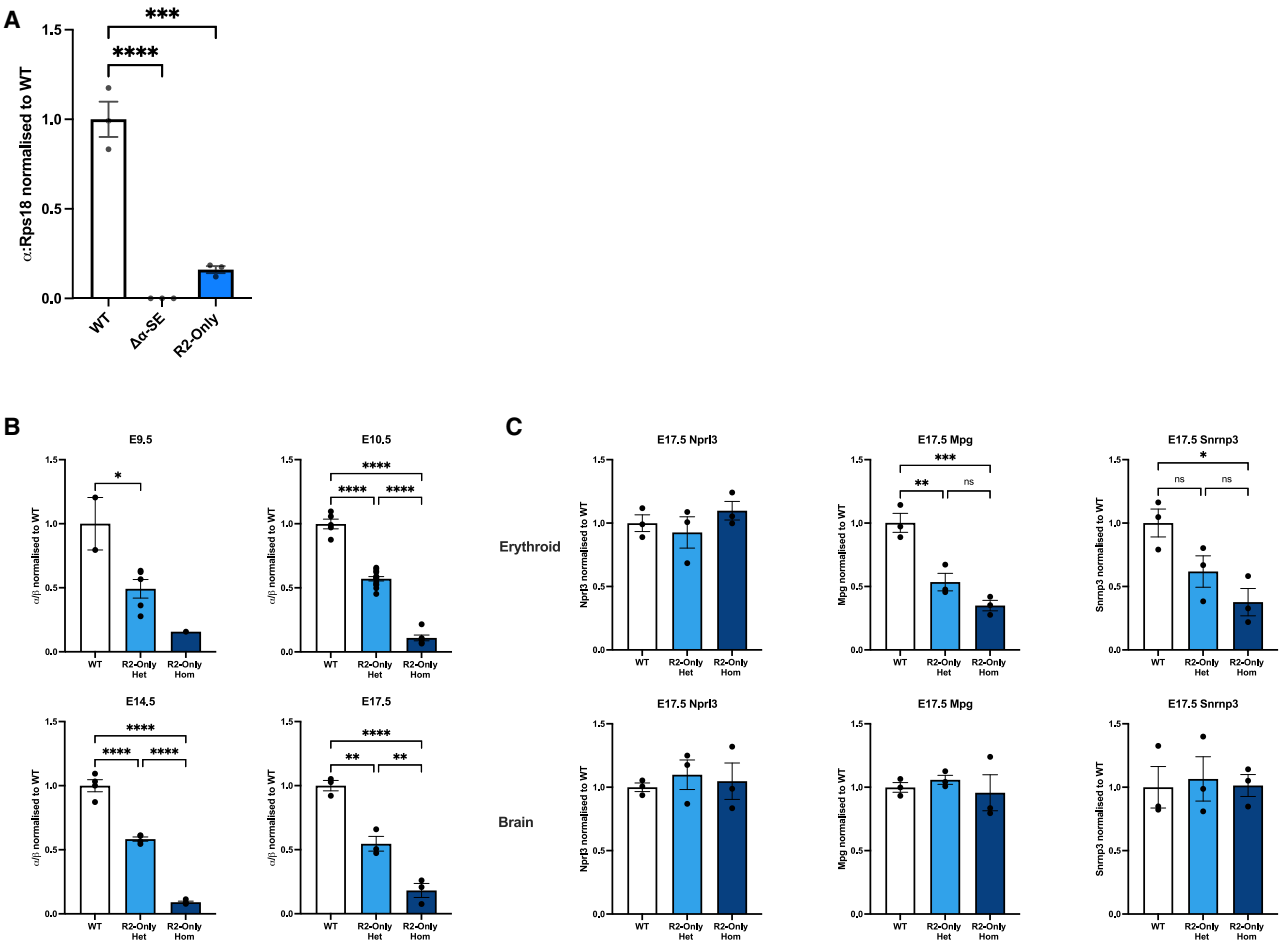


**Figure S3. R2-only mouse-derived erythroid cells differentiate but exhibit specific gene expression perturbations at various stages of embryonic development, related to Figures 1 and 2**

(A) $\alpha$-*globin* gene expression in WT, $\Delta\alpha$-SE, and R2-only EB-derived erythroid cells ($n \geq 3$) assayed by RT-qPCR. Expression normalized to RPS18 and displayed as a proportion of WT expression. Dots, biological replicates; error bars, Standard Error. Statistical analysis was performed using one-way ANOVA with a Tukey post-hoc test: ****$p \leq 0.00001$, ***$p \leq 0.0001$.

(B) RT-qPCR comparing $\alpha$-*globin* expression in FL erythroid cells from WT, R2-only heterozygous, and R2-only homozygous E9.5, E10.5, E14.5, and E17.5 littermates. Expression normalized to $\beta$-*globin* and displayed as a proportion of WT expression. Dots, biological replicates; error bars, Standard Error. Statistical analysis was performed using one-way ANOVA with a Tukey post-hoc test: ****$p \leq 0.00001$, ***$p \leq 0.0001$, **$p \leq 0.001$, *$p \leq 0.01$.

(C) RT-qPCR comparing Nprl3, Mpg, and Snrnp25 expression in WT, R2-only heterozygous, and R2-only homozygous littermates. Expression normalized to RPS18 and displayed as a proportion of WT expression. Dots, biological replicates; error bars, Standard Error. Upper, E17.5 erythroid cells; lower, matched E17.5 brain tissue. Statistical analysis was performed using one-way ANOVA with a Tukey post-hoc test: ****$p \leq 0.00001$, ***$p \leq 0.0001$, **$p \leq 0.001$, *$p \leq 0.01$.

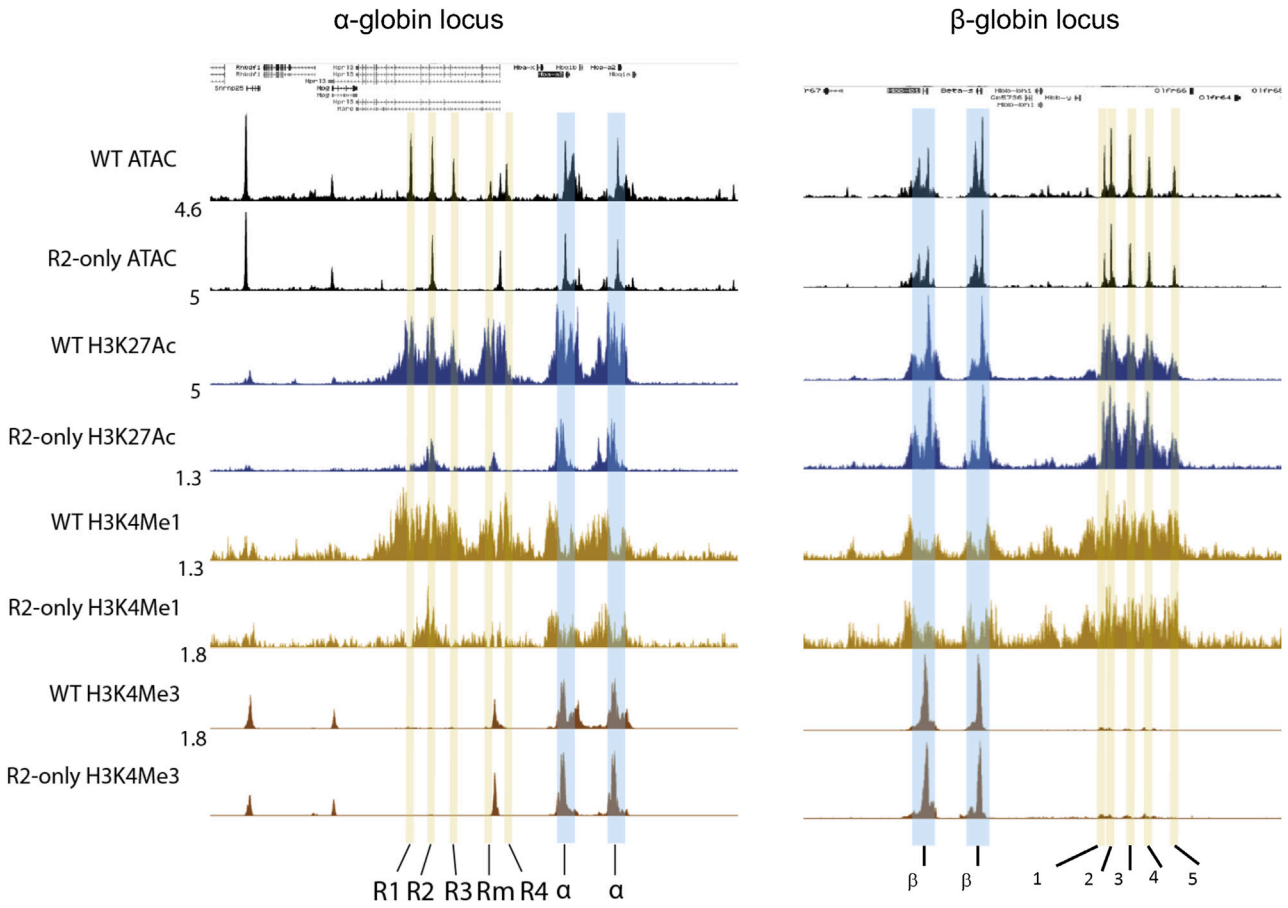


**Figure S4. R2 retains the hallmarks of an active enhancer in R2-only erythroid cells, related to Figure 3**

UCSC tracks represent data from WT (n = 3) and R2-only homozygous (n = 3) FL erythroid cells at the *α-globin* locus, starting at the top with ATAC-seq (black) and followed by ChIPmentation of the following chromatin marks: H3K27Ac (navy); H3K4Me1 (yellow); H3K4Me3 (brown). Tracks = merged biological replicates. All tracks are cpm normalized. Left, *α-globin* locus with highlighted regulatory elements (R1–R4 in green) and promoters (corresponding to *Hba-a1* and *Hba-a2*, the adult α-globin expressed in definitive FL erythroid cells, blue), coordinates = chr11: 32,090,000–32,235,000 (mm9); right, *β-globin* locus with highlighted regulatory elements (HS1–HS5 in green) and promoters (corresponding to *Hbb-b1* and *Hbb-b2*, the adult *β-globin* expressed in definitive FL erythroid cells, blue), shown as an unaffected control, coordinates = chr7:110,936,000–111,070,000 (mm9).

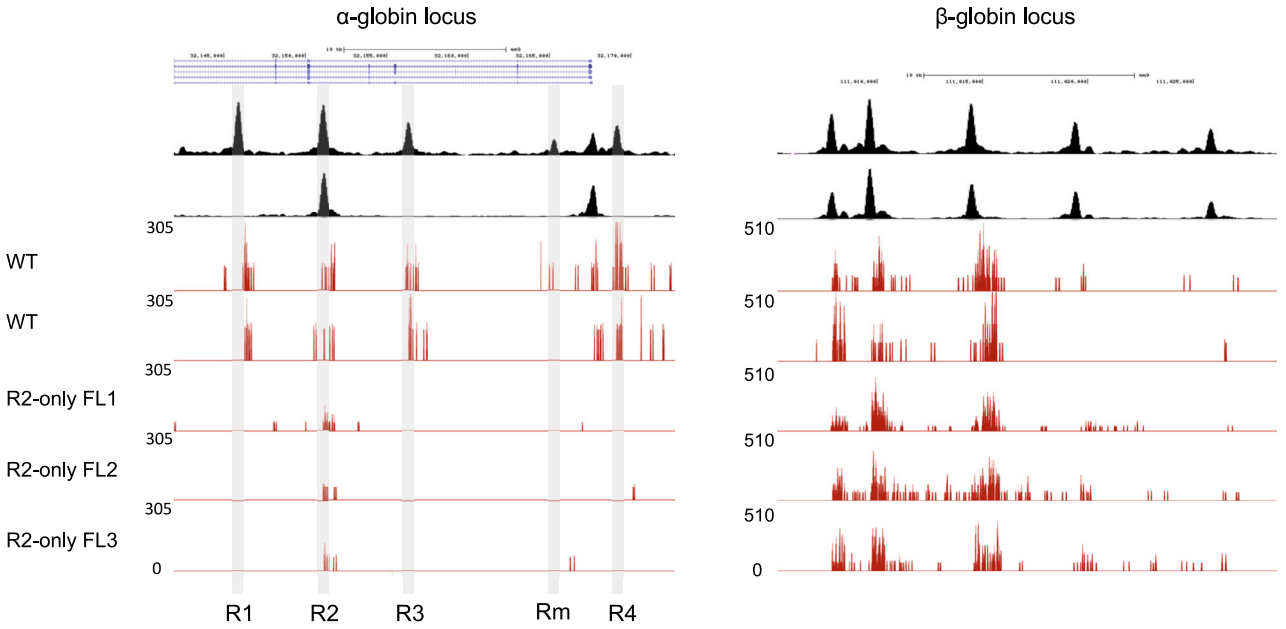


**Figure S5. R2 eRNA transcription is reduced in R2-only erythroid cells, related to Figure 4**

ATAC-seq in WT (n = 3), R2-only heterozygous (n = 3), and R2-only homozygous (n = 3) FL erythroid cells; Poly-A minus RNA-seq tracks from WT (n = 2) and R2-only homozygous (n = 3) FL. Left, *α-globin* locus showing reduced eRNA transcripts in 3 homozygote-derived samples; right, *β-globin* locus as an unaffected control. All tracks rpkm normalized.

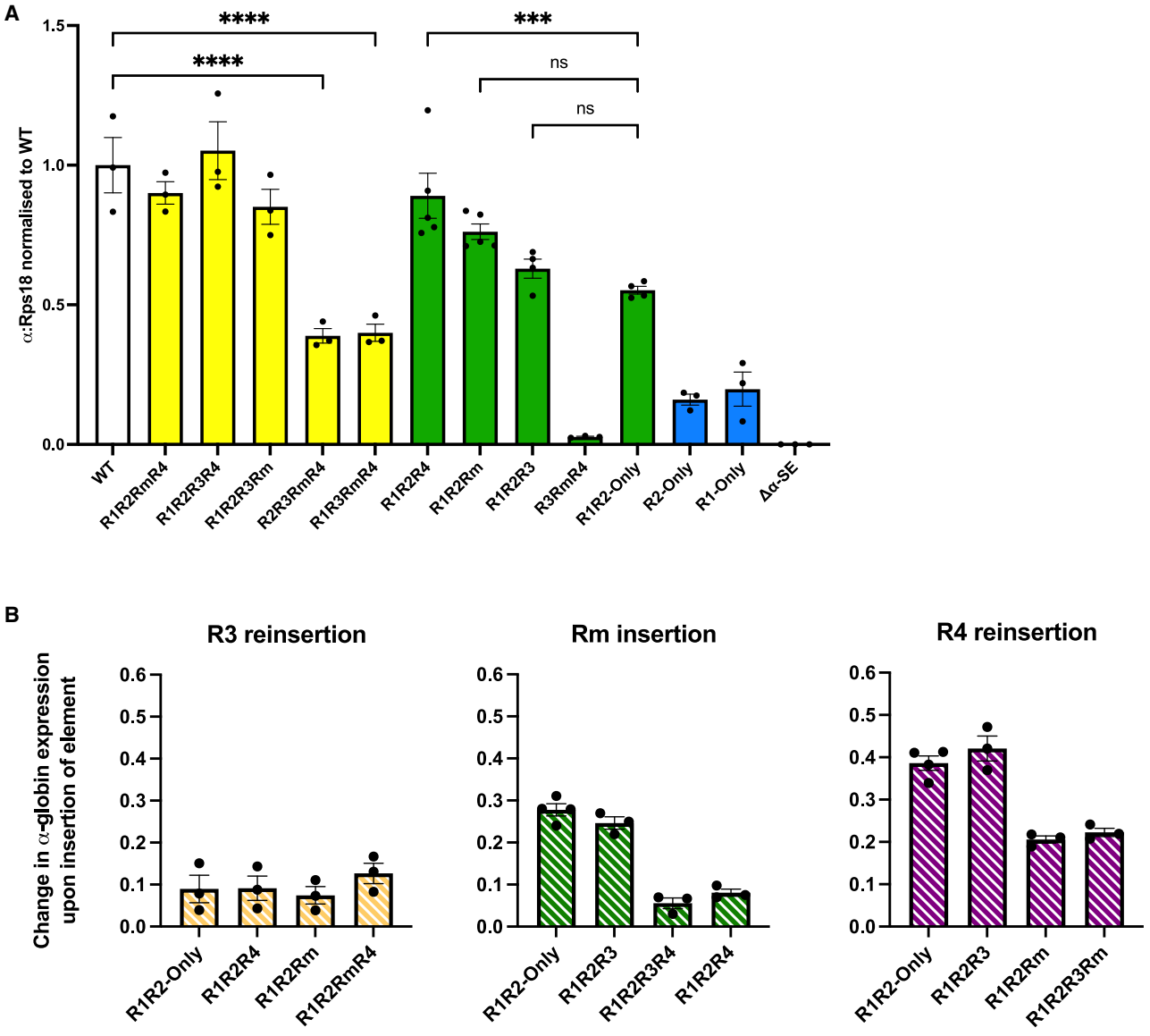


**Figure S6. The R1 and R2 enhancers rely on R3, Rm, and R4 in order to exert their full potential, related to Figure 5**

(A) α-*globin* gene expression in enhancer titration EB-derived erythroid cells (n ≥ 3) assayed by RT-qPCR. Expression normalized to *RPS18* and displayed as a proportion of WT expression. Dots, biological replicates; error bars, Standard Error. Statistical analysis was performed using one-way ANOVA with a Tukey post-hoc test: ***p ≤ 0.0001, ****p ≤ 0.00001.

(B) Percentage increase in α-*globin* expression following R3-reinsertion, Rm-reinsertion, or R4-reinsertion in each corresponding genetic background (R1R2-only, R1R2R4, R1R2Rm, R1R2RmR4; R1R2-only, R1R2R3, R1R2R3R4, R1R2R4; or R1R2-only, R1R2R3, R1R2Rm, R1R2R3Rm, respectively), as assayed by RT-qPCR in EB-derived erythroid cells (n ≥ 3). Dots, biological replicates; error bars, Standard Error.

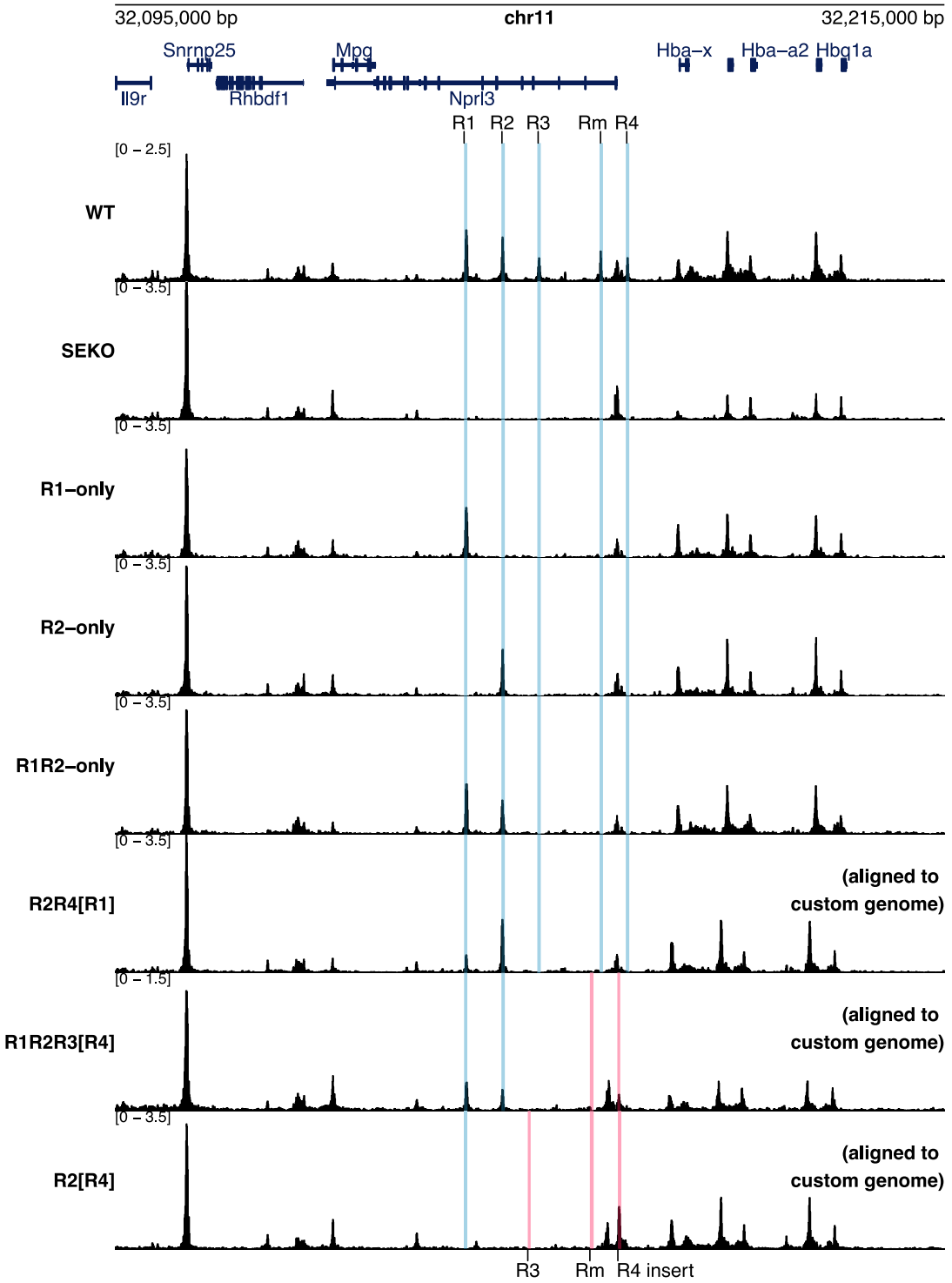


**Figure S7. Chromatin accessibility at the *α-globin* locus for hemizygous R1, R2, and R4 enhancer insertion models differentiated *in vitro*, related to Figure 6**

ATAC-seq tracks of the *α-globin* locus for hemizygous ΔSE (SEKO), R1-only, R2-only, R1R2-only, R2R4[R1], R1R2R3[R4], and R2[R4] models of the *α-globin* locus from CD71$^{+}$ EB-derived erythroid cells, compared to the track from WT hemizygous EB-derived erythroid cells. WT, ΔSE, R1-only, R2-only, and R1R2-only tracks are aligned to mm9 reference genome, and R2R4[R1], R1R2R3[R4], and R2[R4] ATAC tracks were aligned to their respective custom genomes derived

*(legend continued on next page)*

from the mm9 reference. R1R2R3[R4] and R2[R4] genomes incorporate deletions of R3 and R2, respectively; hence, the position of downstream elements is shifted. Enhancers in their WT positions are highlighted in blue or red if they have been shifted in the custom genome. Tracks are normalized by CPM and averaged over two independently targeted clones for each mutation, differentiated in parallel. R2[R4] is data from one clone. Plot was generated using PlotGardener.