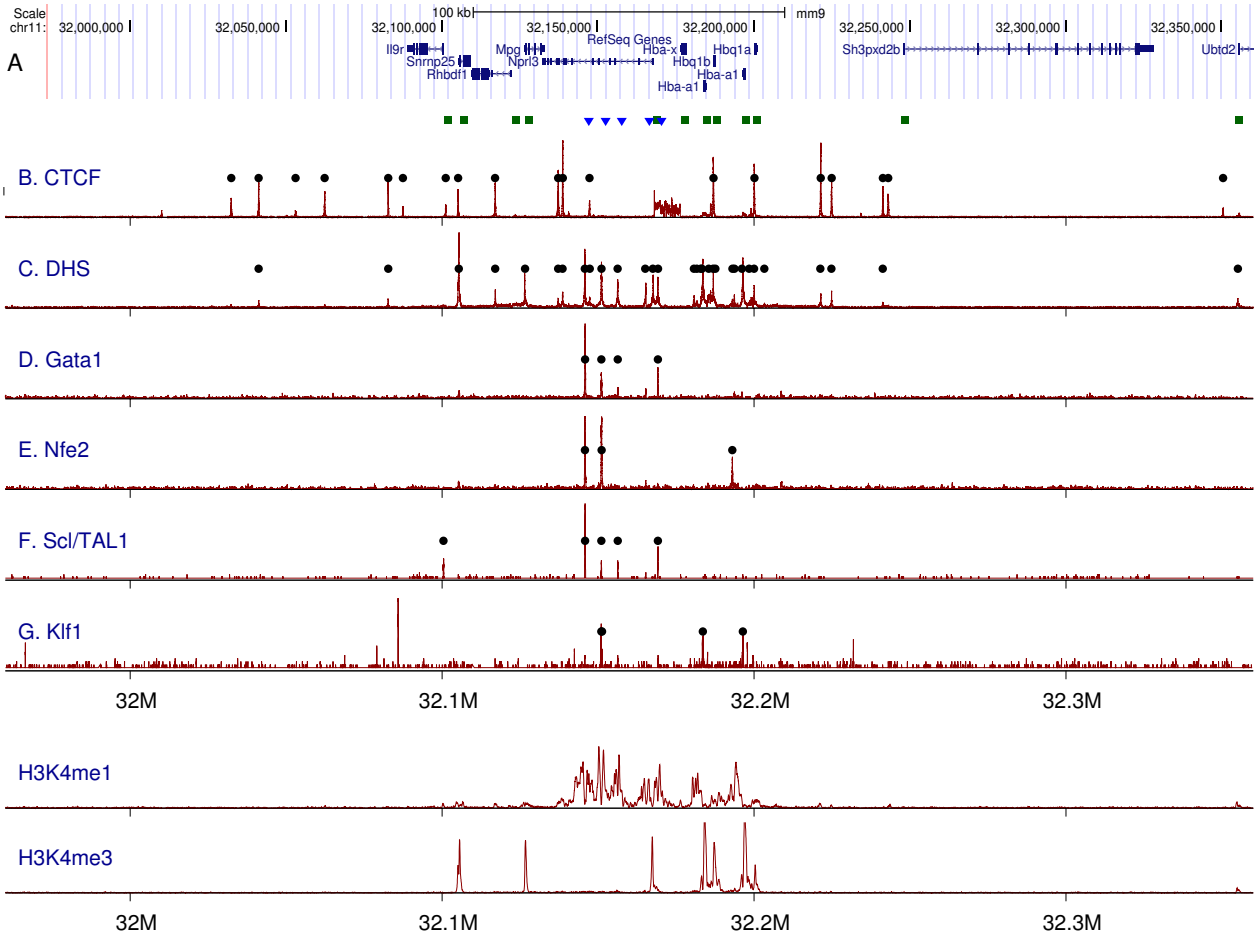

Additional file 3: Figure S2: **ChIP-seq and DNase-seq data are used as an input to the model.** (A) Genome browser view of genes in a 400 kbp region of mouse chromosome 11 surrounding the α-globin locus which is treated in our simulations. Symbols below the browser indicate the positions of the known regulatory elements (blue triangles) and the gene promoters (green squares). (B) ChIP-seq data for CTCF binding across the same region from mouse erythroid (Ter119$^+$) cells. Red lines show the pile-up of reads, and black points indicate the positions of binding sites identified by peak-calling (see Additional file 2: Supplementary Methods for details). Data from Ref. (14). (C) Similar plot showing DNase-seq data from the same cell type, identifying the positions of DNase-1 hypersensitive sites (DHS). Data from Ref. (56). (D)-(G) Plots showing ChIP-seq data, again from the same cell type, for four TFs thought to be key players in globin regulation. Data from Ref. (14) (GATA1 and NFe2), Ref. (50) (Scl/Tal1), and Ref. (57) (Klf1). (Note that where available, control data were used in the peak calling, meaning that peaks seen in the pile-up of reads which did not show significant enrichment above the control were not called.) Since there are DHS located at the binding sites of each of these proteins, we reduce the complexity of our model (and the need for assumptions about the interaction between TFs) by using these as a proxy for protein binding sites. (H)-(I) ChIP-seq data showing relevant histone modifications: monomethylation and trimethylation of H3K4 (associated with enhances and promoters respectively). Data from Ref. (56). All plots are aligned according to the horizontal axis.