

LIST OF SUPPLEMENTARY TABLES

Table S1: PU.1-interacting molecules in Scid.adh.2c2 cells. Related to Fig. 1.

Table S2: Runx1 consensus Position Weight Matrix (PWM) from DN1 and DN3 cells. Related to Fig. 3.

A. Position weight matrix and occurrence of sequences matching PWM at Group 1 and Group 2 sites

B. List of Runx motifs and Log-odds scores, genome-wide, to default cutoff score

C. Genes with additional Runx1 sites recognized in HOMER, using manually relaxed cutoff score

Table S3: Differentially expressed genes in Scid.Adh.2c2 cells, $p_{\text{adj}} < 1E-03$, sorted from highest to lowest \log_2 fold change value in response to wildtype PU.1. Related to Figs. 4, 5.

A. Genes showing significant response to transduction of PU.1 into Scid.adh.2c2 cells ($p_{\text{adj}} < 1E-03$; EdgeR).

B. Genes from A with positive change in response to PU.1; plus effects of sgRunx and sgSatb1 and presence/absence of different site classes.

C. Genes from A with negative change in response to PU.1; effects of sgRunx1 and Satb1 and presence/absence of different site classes.

Table S4: Gene expression effects in Scid.adh.2C2 cells transduced with Runx1DN. Related to Fig. 5.

Table S5: Genes in primary DN cells that are differentially expressed in response to PU.1 deletion: and effects of Satb1 and Runx1 deletion. Related to Fig. 7.

Table S6: List of Oligonucleotides. Related to STAR*Methods and Key Resources Table.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1, related to Figures 1 and 3: PU.1 introduction induces transdifferentiation of Scid.adh.2c2 cells to myeloid phenotype.

(A), Scid.adh.2c2 cells were infected with Myc-Flag-PU.1/hNGFR retrovirus and cultured for two days. Flow cytometric analysis of tagged PU.1 transduced cells (left) shows loss of T-lineage marker CD25 and gain of myeloid marker CD11b (Mac1). Experimental scheme is also indicated (right).

(B), Developmental patterns of expression of PU.1 and candidate interaction partners. The heatmap shows expression profiles for PU.1 (*Sfp1*), its binding partners (*Runx1*, *Satb1*), an Ets family member with an opposite expression pattern from PU.1, and other lymphoid transcription factors (*Gata3*, *Zfp1=FOG1*, *Ikzf1*, and *Tcf3=E2A*), during the multilineage precursor (MLP) to T cell developmental transition. Heatmap is taken from the Immunological Genome web site <www.immgen.org> (Heng *et al.*, 2008; Mingueneau *et al.*, 2013), using the MyGeneSet tool. Color scale indicates z scores with warm colors indicating higher relative levels and cold colors indicating lower relative levels.

(C), Binding patterns of multiple factors associated with *Runx1* occupancy sites in Scid.adh.2c2 cells. Tag count distributions for *Runx1*, *Satb1*, *GATA3*, *Fog1* and PU.1 ChIP, and ATAC signal around *Runx1* peaks using mock- or PU.1-introduced Scid.adh.2c2 cells are shown. All *Runx1* sites identified in the presence or absence of exogenous PU.1 were included in the analysis, as in

Fig. 3A for Satb1 peaks. PU.1 ChIP tag count distributions in primary DN1 and DN2a are also indicated to show relation to normal developmental sites of PU.1 binding in primary pro-T cells.

(D), Binding patterns of multiple factors associated with all PU.1 occupancy sites in PU.1-transduced Scid.adh.2c2 cells. Tag count distributions for ChIP and ATAC signal around PU.1 peaks are shown.

(E), Runx1 motif sequence logos from combined non-promoter Runx1 ChIP peak lists of DN1 and DN3 pro-T cells (Fig. 6A, Table S2) (left) and JASPAR reference motif (right) are shown.

(F), Site-quality comparisons between different classes of Runx1 sites occupied in Scid.adh.2c2 cells with and without forced expression of PU.1. Quality scores were established by calculating the preferred position weight matrix for Runx1 occupancy of non-promoter binding sites in DN1 and DN3 primary pro-T cells (E, left), and the occupied sites scored in Scid.adh.2c2 cells were ranked in terms of their log-odds similarities to this PWM. Violin plots show the distribution of motif log-odds similarity scores of Group 4, Group 5 and Group 6 peaks (Fig. S1C) against the PWM for Runx1 motif calculated from occupancy sites in DN1 and DN3 cells (panel E, left). Median, 25% and 75% percentiles are also shown. P-values were determined by Kruskal-Wallis statistical test.

Supplementary Figure 2, related to Figure 3: Satb1 and Runx1 are preferentially recruited to the sites where ATAC opened and directly bound by PU.1.

(A), Tag count distributions are shown for Satb1, Runx1 and PU.1 ChIP, and ATAC signals from mock- or PU.1-introduced Scid.adh.2c2 cells, specifically around the sites that become

ATAC accessible in response to PU.1-transduction. Sites are separated into those that are also bound (upper) and those that are unbound (lower) by PU.1 directly. .

(B), Average tag count distributions are indicated for Satb1, Runx1 and PU.1 ChIP peaks in the sites becoming ATAC accessible by PU.1-transduction, comparing those that are directly bound with those that are unbound by PU.1 in (A).

(C), The figure shows the numbers of Satb1 (left) or Runx1 (right) peaks in mock- or PU.1-introduced Scid.adh.2c2 cells co-occupied with ATAC opened peaks in (A, bound or unbound by PU.1).

Supplementary Figure 3, related to Figure 5: Regulatory impacts of CRISPR/Cas9-mediated knock out of Runx1 and Satb1 in Scid.adh.2c2 cells

(A), Target genes regulated by PU.1 in Satb1- or Runx1-dependent ways in Scid.adh.2c2 cells are also regulated by PU.1 in primary pro-T cells. Heat maps show the acute responses to PU.1 in primary pro-T cells of the same Satb1- or Runx1-responsive genes (sgSatb1 or sgRunx1/sgControl<0.75) identified in Fig. 5C. As described in our related study (Ungerbäck et al., op. cit.), exogenous PU.1 introduced into newly committed DN2b pro-T cells causes some cells to undergo partial reversal of differentiation (PU1WTHA25) and others, expressing higher levels of PU.1, to be driven out of the T-cell program (PU1WTHA44), with stronger overall gene expression differences. Hierarchical clustering analyses of the expression of the Satb1- or Runx1-responsive genes in both subsets of PU.1-introduced primary DN2 cells (Ungerbäck et al., op. cit.) are shown, compared to cells expressing empty vector instead of PU.1 (left). Experimental scheme is also indicated (right).

(B and C), Satb1- or Runx1-responsive genes are significantly enriched among PU.1-dependent (B) or PU.1-repressed (C) genes, as compared to all expressed genes. For comparison with Fig. 5C, the figures show hierarchical clustering analyses of all the expressed genes (RPKM>3 in sgControl/Mock) in the same Scid.adh.2c2 cells used for Fig. 5C (left). The percentages of genes showing a Satb1 or Runx1 response (sgSatb1 or $\text{sgRunx1}/\text{sgControl}<0.75$) among the PU.1-dependent (B) or PU.1-repressed (C) genes in Fig. 5C are shown, as compared with percentages among all expressed genes (right). P-value is determined by Fisher's exact test.

Supplementary Figure 4, related to Figure 5: Runx-dependent gene definition: confirmation with Runx1 dominant negative and correlation with PU.1 effect

Confirmation that genes repressed by introduction of PU.1 into Scid.adh.2c2 cells are frequently also genes that have a dose-dependent requirement for Runx1 activity in the absence of PU.1.

(A) Comparison of effects of sgRunx1 and Runx1DN for genes downregulated by sgRunx1. (B) Comparison of effects of sgRunx1 and exogenous PU.1, for same genes as in (B). (C) Same as A, but for genes upregulated by sgRunx1. The genes shown are those that were positively or negatively responsive to exogenous PU.1, and also potential Runx1 regulation candidates based on responses to sgRunx1 and Cas9. Values plotted are from Table S3 and Table S4.

Although Cas9-dependent deletion experiments using sgRNA against Runx1 (Fig. 5, Table S3) show detectable but relatively weak effects on these genes (blue bars in panels A-C), the same genes show strong and highly significant responses in the same directions in response to acute introduction of Runx1DN, which binds DNA but neither transactivates nor represses (dark red

bars in panels A & C). This confirms that these genes can be significantly affected by disruption of the Runx input.

Genes chosen for in this analysis were all those that have both (1) substantial responses to PU.1, positive or negative ($p_{\text{adj}} < 0.01$) and (2) some indication of a response to Runx1 disruption alone, positive or negative. For sgRunx1 responses a very permissive criterion ($p_{\text{adj}} < 0.5$) was used to collect a large number of possible candidates (Table S3). The same genes were then evaluated for their responses to acute transduction with Runx1 DN in Scid.adh.2c2 cells in completely independent experiments (Table S4). The questions addressed were: (1) how often do these permissively chosen potential Runx1 targets actually respond to Runx1DN, and (2) how much do the directions of the genes' responses to Runx1 predict the directions of their responses to PU.1? The genes shown are sorted according to the magnitudes of their responses to sgRunx1, and downregulated genes (Runx1-dependent) are shown in panels (A) and (B) while upregulated genes (Runx1-repressed) are shown in panel (C). Specifically, bars in the figure show \log_2 fold changes in expression of the indicated genes for the following perturbations compared to their levels in the corresponding controls: (A)-(C) sgRunx1 (no added PU.1) vs. control guide RNA (no added PU.1) = blue bars; (A) & (C) Runx1DN transduction vs. empty vector transduction for 48 hr = dark red bars; (B) PU1 transduction + control guide RNA vs. mock transduction + control guide RNA = green bars. Dark red bars in A, C are shown for all genes where effect of Runx1DN had $p_{\text{adj}} < 0.05$; most had $p_{\text{adj}} < 1E-10$. All values were calculated (EdgeR) from whole genome transcriptome data in two independent experiments of each type.

Summary of results: Results show, first, that despite the highly permissive criterion used to select this set of sgRunx1-responsive genes (blue bars), the responses to sgRunx1 agree well with

responses to the more potent Runx1DN (dark red bars), whether for activation or repression (panels A, C). Second, this analysis confirms that most of the Runx1-dependent genes that are PU.1 responsive in any direction (blue bars) are repressed by PU.1 (green bars in panel B). Fisher's exact tests for both relationships strongly indicate dependence (both $p < 0.00001$).

Supplementary Figure 5, related to Figure 6: Correlation of Runx1 peak redirections in transduced Scid.adh.2c2 cells and primary pro-T cells in normal development

(A), Venn diagrams show categorization of Runx1 peaks in primary DN1 and DN3 cells (Fig. 6A) (left). Note shift of Runx1 binding sites between the PU.1-expressing DN1 stage and the PU.1-nonexpressing DN3 stage. Violin plots show the distribution of motif log-odds similarity scores of DN1_specific, Shared and DN3_specific Runx1 peaks (right) against PWM for Runx1 motif in DN1 and DN3 (Fig. S1E). Median, 25% and 75% percentiles are shown with P-values (K-W statistical test). The lower median quality of sites successfully occupied by Runx1 in DN1 stage is notable because Runx1 levels themselves are lower at this stage (Fig. S1B), supporting the need for interaction with PU.1 or some other DN1-stage factor(s) to recruit Runx1 to these sites.

(B), Scatter plots of the Runx1 peaks on promoter (left) or non-promoter (right) sites in primary DN1 and DN3 cells are shown. Dots represent the log₂ values of normalized Runx1 ChIP tag count in DN1 (x-axis) and DN3 (y-axis). Runx1 peaks co-occupied with PU.1 in DN1 cells are indicated in yellow.

(C and D), Dynamics of Runx1 site responses to PU.1 in Scid.adh.2c2 cells resemble reversal of natural shifts of Runx1 sites in normal pro-T cell development. Tag count distributions for

Runx1 and PU.1 ChIP, and ATAC signal around peaks in Group 1 (C) and Group 2 (D) in Fig. 3A using mock- or PU.1-introduced Scid.adh.2c2 cells and primary DN1 or DN3 cells are shown.

(E), Average tag count distributions for Runx1 and PU.1 ChIP, and ATAC signals in (C and B) are quantitated: Group 1 sites from (C), Group 2 sites from (D).

(F), Impacts of PU.1 on gene expression tend to be opposite for genes with DN1-unique Runx1 sites as compared to genes with DN3-unique Runx1 sites. Cumulative distributions of expression changes in wild-type PU.1 (PU.1WTHA25)-introduced-transduced DN cells (Ungerbäck et al., op. cit.) for four groups of genes bound by different category of Runx1 peaks in DN1 and DN3 in (A) and differentially expressed in PU.1-introduced cells (FDR<0.1) are shown. The number of genes in each group and p-values (K-S tests for comparisons with ‘no peaks’) are indicated. Only genes are shown that have sites of the indicated class without other classes. Note that genes with DN1-specific Runx1 sites in normal development tend to be up-regulated in acute response to PU.1 while genes with DN3-specific Runx1 sites tend to be down-regulated by PU.1. Bracket shows comparison between the fractions repressed vs. activated in the different gene groups.

Supplementary Figure 6, related to Figure 7: CRISPR/Cas9-mediated knock out of *Spi1* (PU.1), *Runx1* and *Satb1* in primary DN cells

(A), Phenotype of Cas9, Bcl2-doubly transgenic progenitor cells prior to transduction with sgRNA against *Spi1*, *Runx1*, or *Satb1*. Flow cytometry analysis of BM-derived precursors after 4 days of OP9-DL1 culture (before retrovirus infection) shows that cells are phenotypically Kit+ DN1 cells.

(B), Flow cytometric analysis of sgRNA transduced BM-derived precursors after 8 days of OP9-DL1 culture shows high frequency of transduction in CD45⁺ Lin⁻ cells.

(C), Inhibition of early pro-T cell proliferation by Cas9-mediated deletion of *Spil*. Cell numbers of CD45⁺ and CFP⁺ cells in Fig. 7A are shown. P-values were determined by Student *t*-test.

(D), Successful removal of PU.1 protein in PU.1-specific sgRNA-transduced pro-T cells from mice with transgenic Cas9. PU.1 protein expression in sgControl and sgPU.1 introduced primary DN cells (in Fig. 7A) was determined by intracellular staining followed by FACS analysis.

(E), Functional collaboration between PU.1 and Runx1 shown by divergent responses in primary pro-T cells to loss of PU.1 by genes with different classes of Runx1 sites. Genes with DN1-specific or DN3-specific Runx1 sites, unique or overlapping with PU.1 sites, are compared for responses to loss of PU.1. Cumulative distributions of expression changes in PU.1-deficient primary DN cells are shown for four groups of genes, bound by DN1-specific (left) and DN3-specific (right) Runx1 peaks depicted in Fig. 6A and differentially expressed in sgPU.1-introduced cells (FDR<0.1). The number of genes in each group and p-values (K-S tests for comparisons with 'PU.1_no_Runx1 peaks') are indicated.

(F), Satb1- or Runx1-responsive genes are significantly enriched in PU.1-dependent or PU.1-repressed genes (Fig. 7B) compare to all expressed genes. For comparison with Fig. 7B, the figures show hierarchical clustering analyses of the all expressed genes (RPKM>3 in sgControl) in the same cells used for Fig. 7B (left). The percentages of Satb1- or Runx1-responsive genes (sgSatb1 or sgRunx1/sgControl<0.75) among the PU.1-dependent and PU.1-repressed genes in Fig. 7B and all expressed genes are shown (right). P-value is determined by Fisher's exact test.

Figure S1

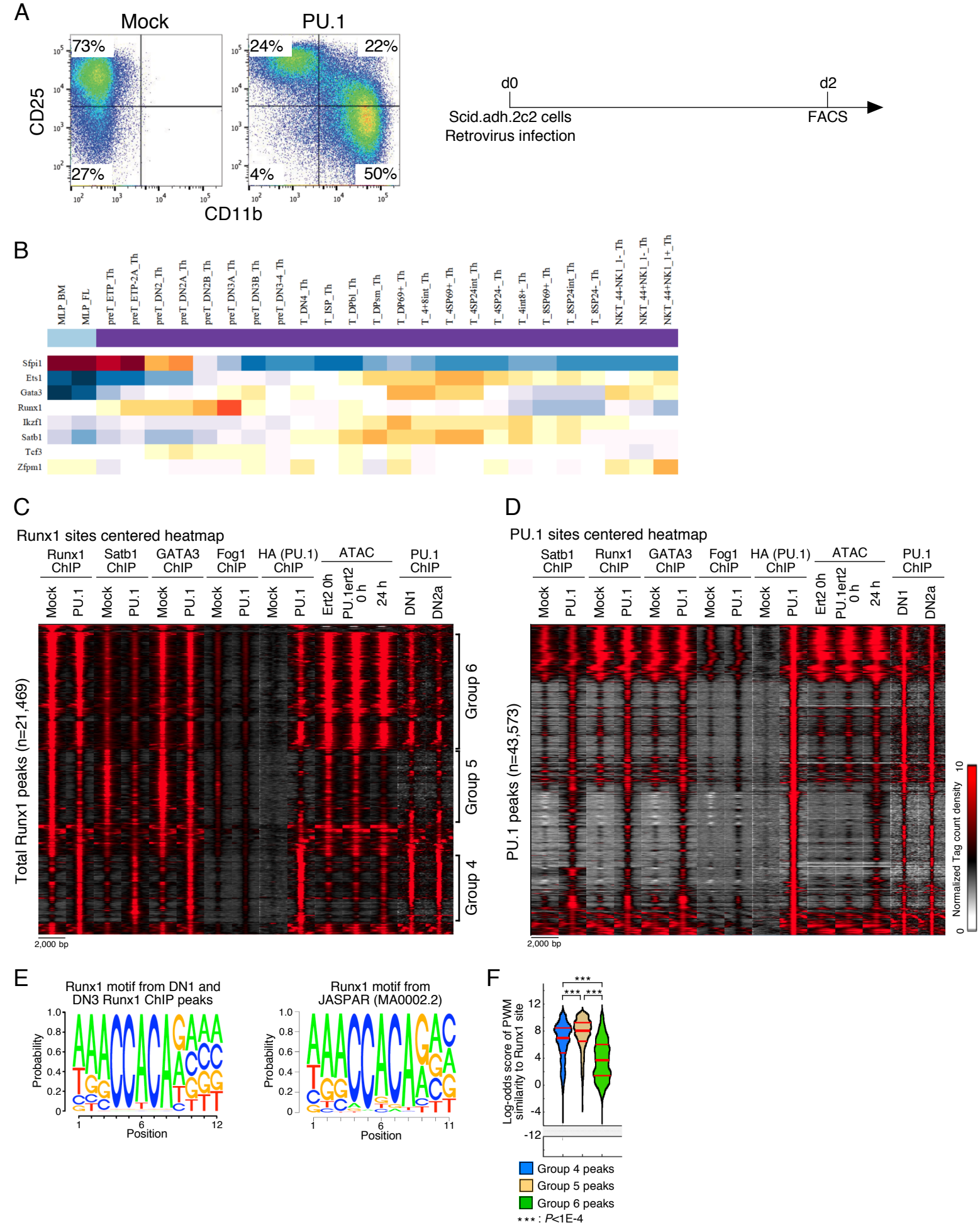
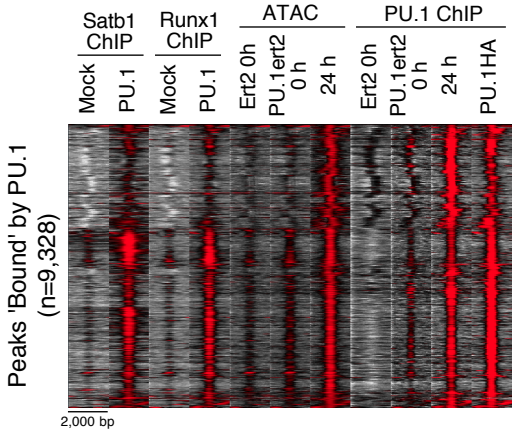
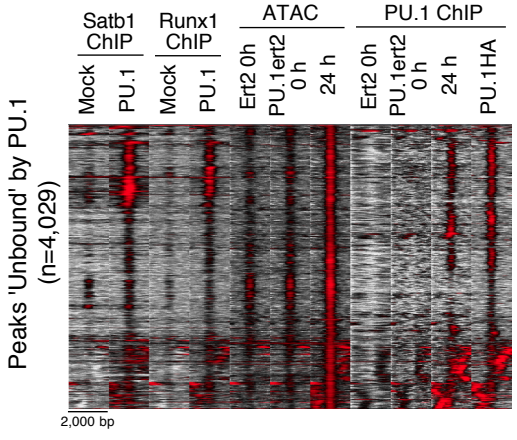


Figure S2

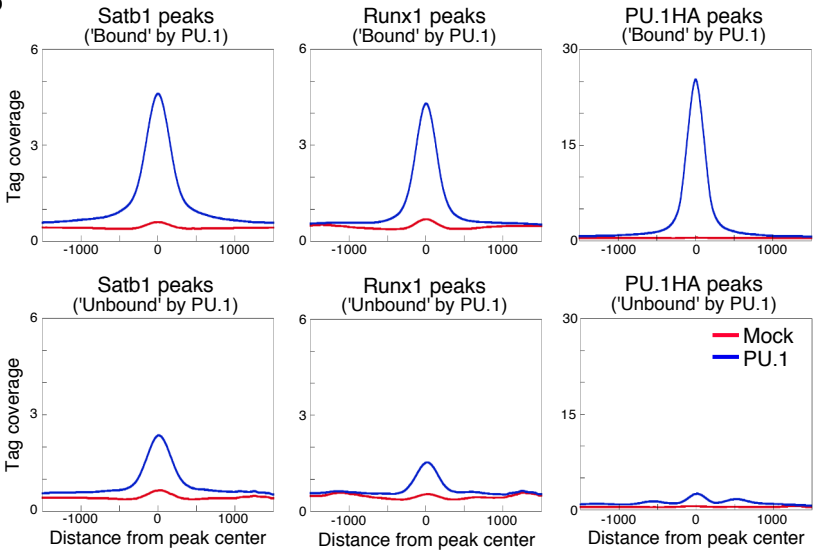
A ATAC opened sites centered heatmap (Peaks 'Bound' by PU.1)



(Peaks 'Unbound' by PU.1)



B



C

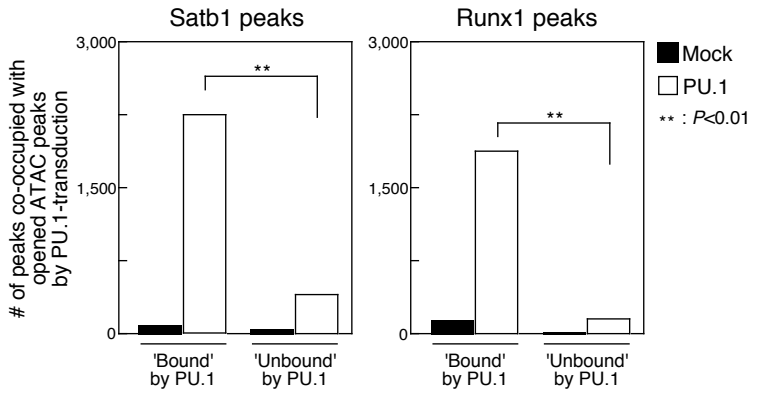


Figure S4

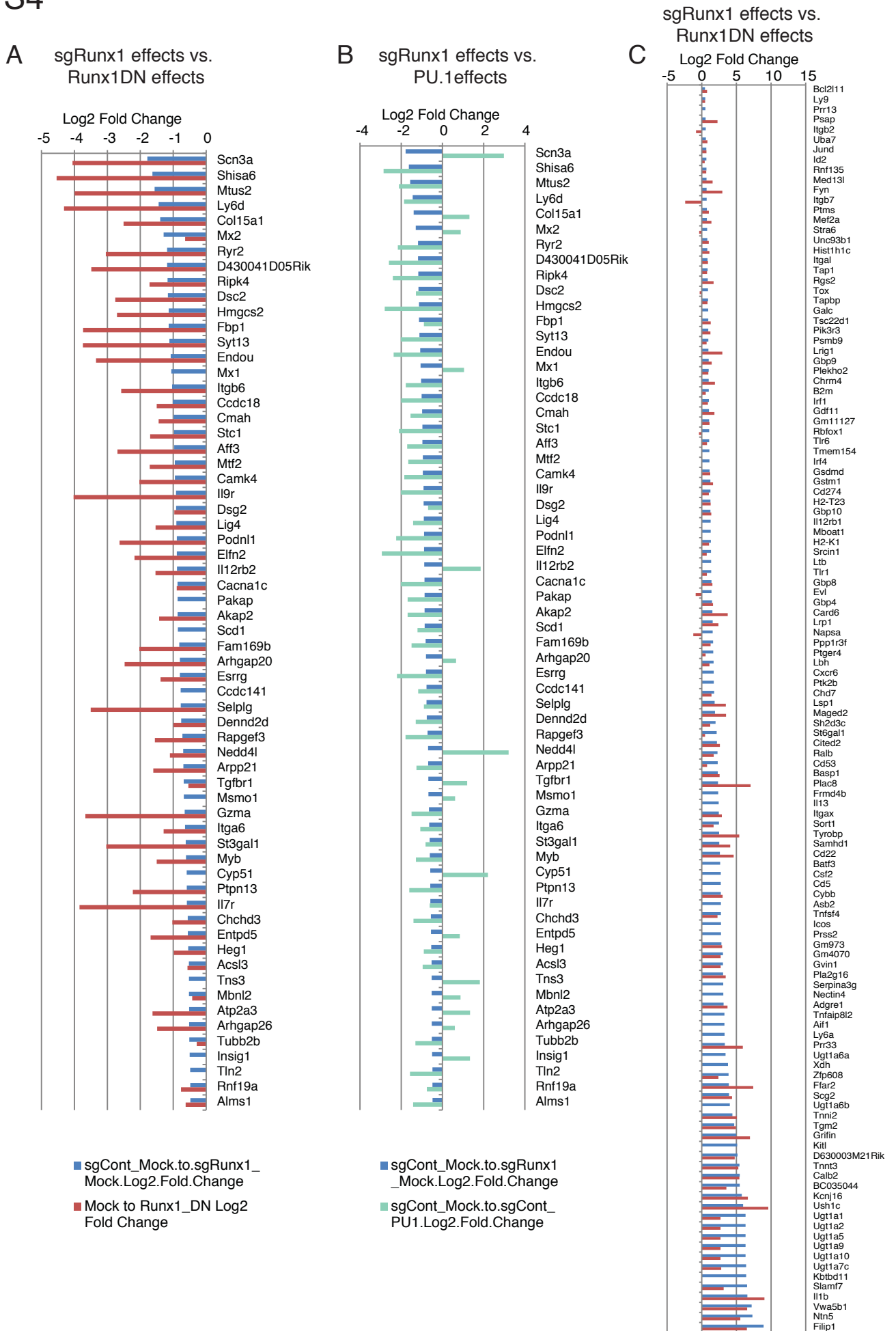


Figure S5

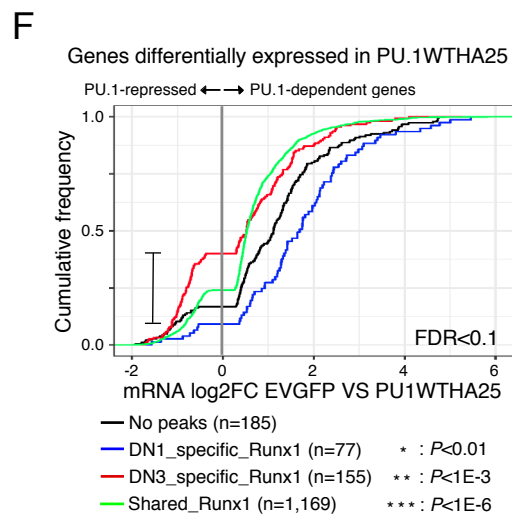
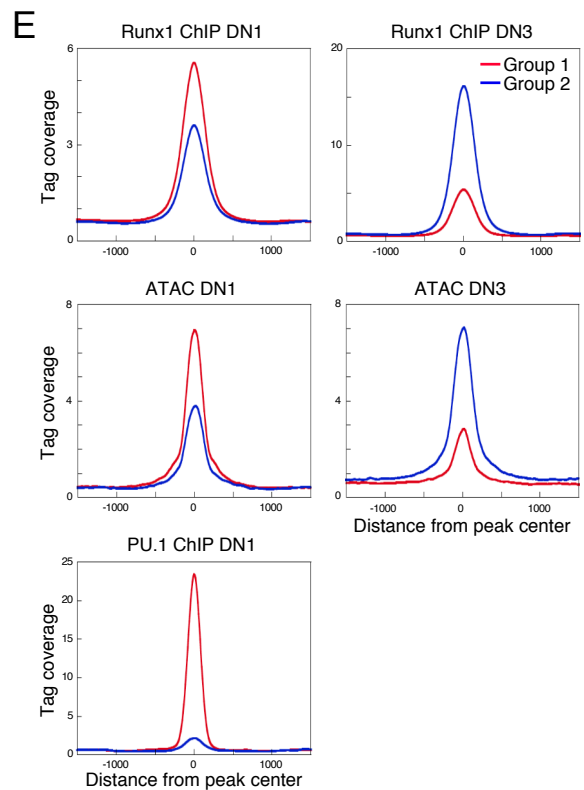
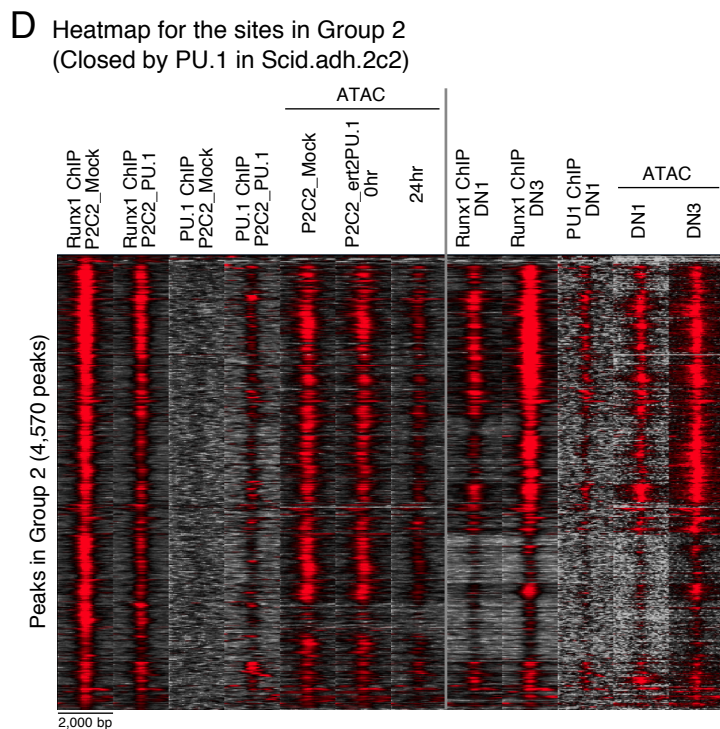
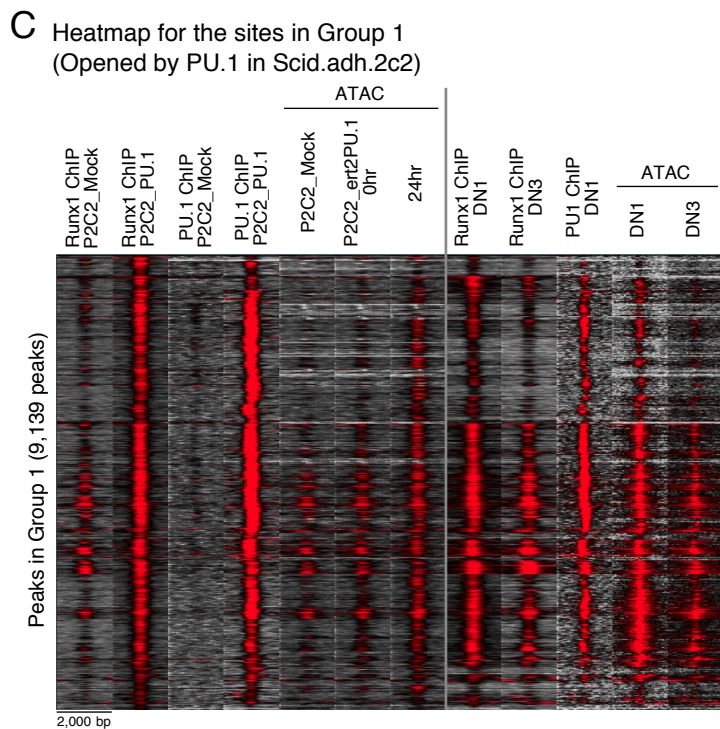
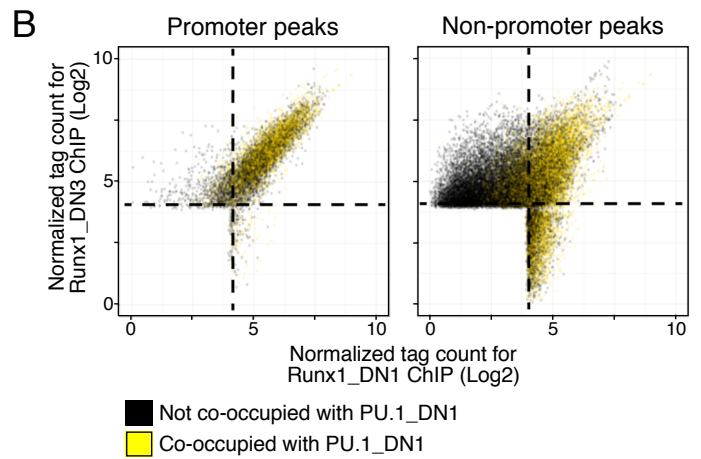
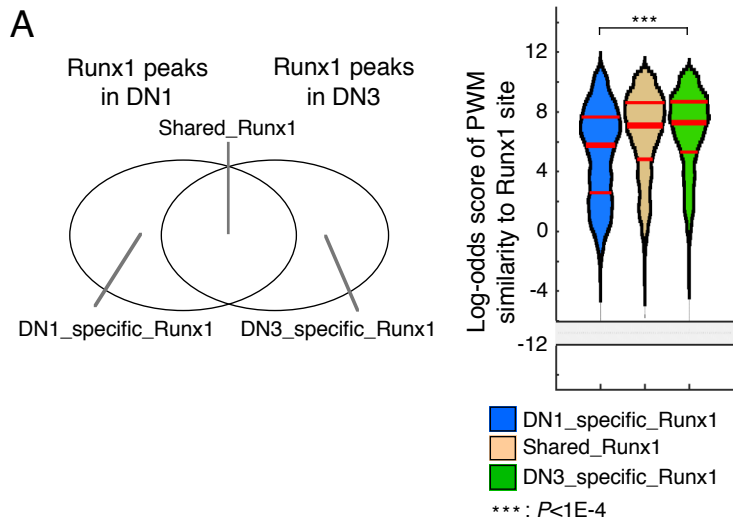


Figure S6

