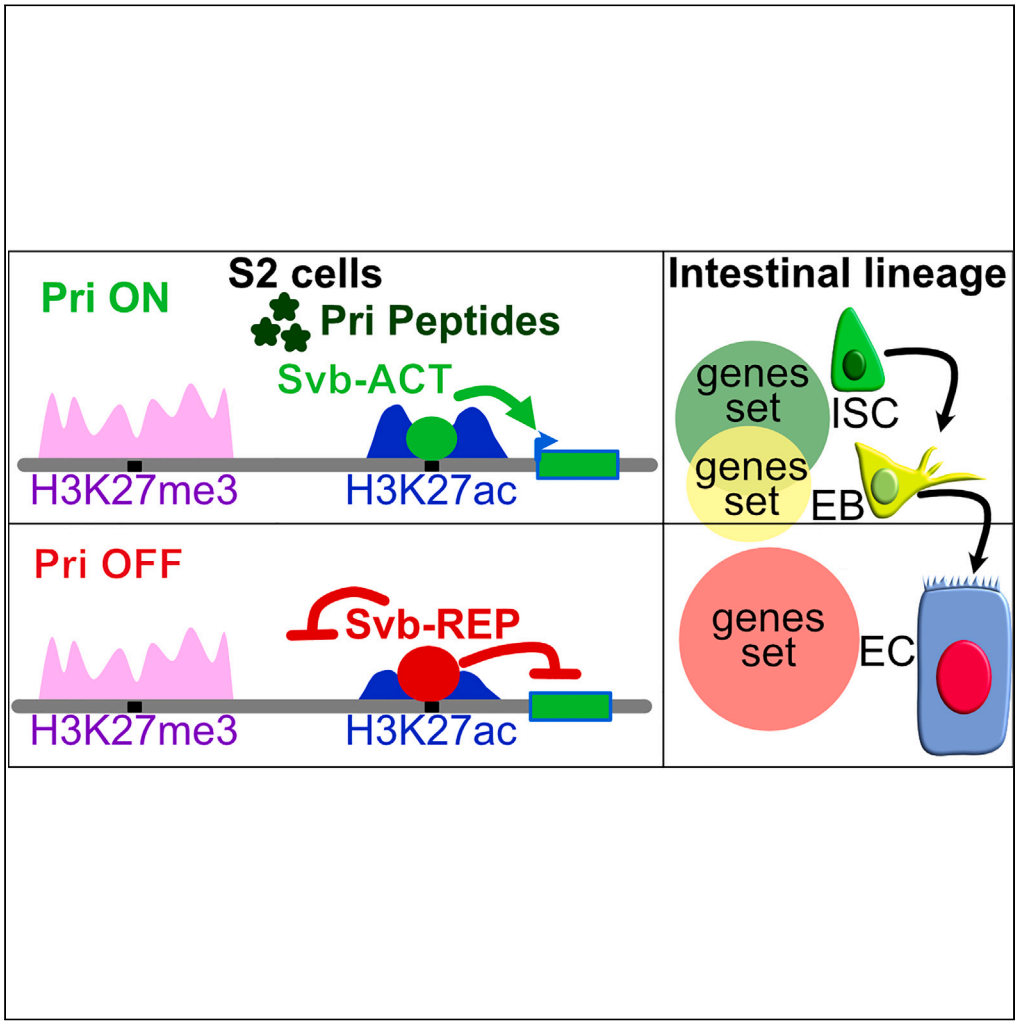
iScience

CellPress
OPEN ACCESS

Article

Crosstalk between chromatin and Shavenbaby defines transcriptional output along the Drosophila intestinal stem cell lineage



Alexandra Mancheno-Ferris, Clément Immarigeon, Alexia Rivero, ..., François Payre, Olivier Cuvier, Cédric Polesello

olivier.cuvier@univ-tlse3.fr (O.C.)
cedric.polesello@univ-tlse3.fr (C.P.)

Highlights

Constrained by chromatin, Svb-ACT and REP isoforms bind the same active enhancers

Svb-ACT and REP antagonistically control the same set of genes in a given cell type

Svb-REP represses transcription by deacetylating H3K27 at active enhancers

Chromatin dynamics along intestinal stem cell lineage defines Svb target genes

Mancheno-Ferris et al., iScience 27, 108624
January 19, 2024
https://doi.org/10.1016/j.isci.2023.108624

iScience

CellPress
OPEN ACCESS

Article

Crosstalk between chromatin and Shavenbaby defines transcriptional output along the Drosophila intestinal stem cell lineage

Alexandra Mancheno-Ferris,[1,2,7] Clément Immarigeon,[1,2,7] Alexia Rivero,[1,2] David Depierre,[1,3] Naomi Schickele,[1,3] Olivier Fosseprez,[1,3] Nicolas Chanard,[1,3] Gabriel Aughey,[4] Priscilla Lhoumaud,[1,3,5] Julien Anglade,[1,3] Tony Southall,[4] Serge Plaza,[1,6] François Payre,[1,2] Olivier Cuvier,[1,3,*] and Cédric Polesello[1,2,8,*]

SUMMARY

The transcription factor Shavenbaby (Svb), the only member of the OvoL family in *Drosophila*, controls the fate of various epithelial embryonic cells and adult stem cells. Post-translational modification of Svb produces two protein isoforms, Svb-ACT and Svb-REP, which promote adult intestinal stem cell renewal or differentiation, respectively. To define Svb mode of action, we used engineered cell lines and develop an unbiased method to identify Svb target genes across different contexts. Within a given cell type, Svb-ACT and Svb-REP antagonistically regulate the expression of a set of target genes, binding specific enhancers whose accessibility is constrained by chromatin landscape. Reciprocally, Svb-REP can influence local chromatin marks of active enhancers to help repressing target genes. Along the intestinal lineage, the set of Svb target genes progressively changes, together with chromatin accessibility. We propose that Svb-ACT-to-REP transition promotes enterocyte differentiation of intestinal stem cells through direct gene regulation and chromatin remodeling.

INTRODUCTION

Transcription factors (TFs) are primary determinants of cell phenotypes and behaviors, regulating gene expression via binding to DNA sequences located within enhancers and promoters *cis*-regulatory elements.[1–3] TF binding motifs being generally short (4–15 base pairs) and degenerate are very abundant in the genome. The accessibility of TF binding sites is influenced by many parameters and dimensions, including biophysical constraints, chromatin state, and organization into domains, and notably involving the presence/absence of epigenetic co-regulators. As a consequence, only a limited subset of TF binding sites are bound in a given cell type,[4–8] and substantial variations in binding events are often observed across tissues.[9,10] These cell-specific changes in binding site occupancy are often thought to have strong impact in transcriptional outputs,[11] though distinct tissues, developmental stages, and/or transcriptional factors may adopt distinct strategies. The underlying mechanisms and functional consequences currently remain to be fully elucidated.

TFs of the Ovo-like (OvoL) family are specific to metazoans[12] and display key functions during animal development. The founding member called Ovo/Shavenbaby (Svb) was initially identified in flies[13] for its dual role in germline and epidermal development. Mammalian species have evolved three paralogs referred to as OvoL1-3 that display partly redundant functions[14] and also contribute to germline, skin development, and diseases.[15] In addition, the deregulation of OvoL expression is associated to various cancers of epithelial origin,[16–18] and it has been proposed that OvoL factors act as guardians of epithelial integrity by preventing epithelial-mesenchymal transition.[17,19] While OvoL TFs share a highly conserved zinc-finger DNA-binding domain that recognizes the core sequence CNGTT,[20,21] they differ in their N-terminal regions and can exhibit distinct transcriptional activities.[18]

Drosophila encodes a unique Ovo/Svb gene, providing an attractive paradigm to dissect OvoL functions and investigate the mechanisms underlying tissue-specific activities. The somatic factor, Shavenbaby (Svb), is known for governing differentiation of epidermal cells in the embryo and adult derivatives.[22,23] The Svb protein is translated as a large transcriptional repressor, referred to as Svb-REP. Remarkably, under the action of Polished-rice (Pri, also known as Tarsal-less or Mille-pattes) peptides encoded by small open reading frames, the ubiquitin ligase

[1]Molecular, Cellular and Developmental biology department (MCD), Centre de Biologie Integrative (CBI), Université de Toulouse, CNRS, UPS, 31062 Toulouse, France
[2]Control of cell shape remodeling team, CBI, CNRS, UPS, 31062 Toulouse, France
[3]Chromatin Dynamics and Cell Proliferation team, CBI, CNRS, UPS, 31062 Toulouse, France
[4]Imperial College London, Sir Ernst Chain Building, South Kensington Campus, London SW7 2AZ, UK
[5]Institut Jacques Monod, Université Paris Cité/CNRS, 15 rue Hélène Brion, 75205 Paris Cedex 13, France
[6]Laboratoire de Recherche en Sciences Végétales, CNRS/UPS/INPT, 31320 Auzeville-Tolosane, France
[7]These authors contributed equally
[8]Lead contact
*Correspondence: olivier.cuvier@univ-tlse3.fr (O.C.), cedric.polesello@univ-tlse3.fr (C.P.)
https://doi.org/10.1016/j.isci.2023.108624

Ubr3 binds to Svb-REP and triggers proteasome-dependent processing resulting in the production of a shorter transcriptional activator, Svb-ACT.[24,25] In the embryonic epidermis, Svb-ACT triggers the expression of a battery of effector genes that are collectively responsible for the remodeling of epithelial cells.[26–28] We have recently reported that Svb is critically required for the maintenance of adult epithelial stem cells, which ensure the homeostasis of the renal[29] and intestinal systems.[30] The Svb transcriptional switch operated by Pri peptides is also at work to control the behavior of adult stem cells: SvbACT promotes intestinal stem cell (ISC) self-renewal and proliferation, while in contrast SvbREP dictates differentiation into enterocytes (ECs).[30]

Here we used genomic and bioinformatic approaches to decipher the role of the two Svb isoforms on gene expression. Using engineered cells in culture, we defined the molecular mode of action of Svb-ACT and REP. We show that, constrained by chromatin landscape, both Svb-ACT and Svb-REP bind to a similar array of open enhancers, mediating antagonistic effects on the expression of a common set of target genes. We further show that Svb-REP participates in remodeling local chromatin landscape to promote transcriptional repression, which can occur through distal enhancers. We developed unsupervised analyses of differential gene expression in response to Svb-ACT and Svb-REP allowing the identification of presumably direct Svb target genes. We next deployed these approaches to analyze the genome-wide transcriptional response to one or the other isoform in sorted populations of adult ISCs and of their progeny enteroblasts (EBs) and ECs. Strikingly, Svb target genes are different across the successive steps of the stem cell lineage. This change in the target genes repertoire during differentiation is accompanied by chromatin accessibility remodeling and concomitant accumulation of Svb-REP. This work sheds light on the possible crosstalk between peptide-induced, proteasome-mediated switch converting a repressor to a transcriptional activator and the chromatin landscape of its binding sites. Together, chromatin and Svb isoforms impose a dynamic and specific transcriptional output in differentiating cells.

RESULTS

Svb-REP and Svb-ACT bind common enhancers and antagonistically control gene expression

To map the genomic sites bound by Svb-ACT and Svb-REP, we made use of stable cell lines derived from S2 *Drosophila* cells, which do not express endogenous Svb and have been engineered to express Svb tagged with GFP, together with or without functional Pri peptides[24,25] (Figure 1A). Without Pri it stays in the repressor form, while the presence of Pri peptides triggers Svb maturation into the activator form. These cell lines are therefore referred to as Svb-REP and Svb-ACT cells, respectively. We performed chromatin immunoprecipitation sequencing (ChIP-seq) coupled to next-generation sequencing in Svb-ACT or Svb-REP cells (Figure 1A). ChIP-seq revealed 5,523 highly enriched for Svb-ACT binding (peak calling MACS q value = $10e^{-4}$) (Figure 1B). Consistent with a binding mediated by the zinc fingers domain, these regions feature DNA motifs matching CcGTT, the core Ovo/Svb binding motif ([28,31,32] and STAR Methods). The most enriched motif matches an ACCGTTA sequence, called F7-BS. The F7-BS motif was defined by a machine learning approach combining statistical analysis to phylogenetic information[33] as the best predictor of functional Svb binding sites *in vivo*.[28] Although peak calling detected a markedly smaller number of regions for Svb-REP (1,325 with MACS q value = $10e^{-4}$), regions bound by Svb-REP extensively overlapped with regions bound by Svb-ACT (Figures 1C and 1D). In addition, there was a clear enrichment for Svb-REP signal across all Svb-ACT peaks, as seen in heatmaps when compared to neighbor regions or a set of random genomic regions used as negative controls. Even when not detected as peaks, Svb-REP accumulation at Svb-ACT-bound sites was also visible along individual genes on genome browser views (Figures 1E and 1H). In most cases, Svb-REP-bound regions missed by peak calling corresponded to peaks of weak intensity in Svb-ACT cells, suggesting that the difference in the number of detected peaks may reflect variations in the depth of ChIP-seq data between Svb-ACT and REP samples. Accordingly, the ACT/REP signal ratio was similar between peaks detected in both Svb-ACT and Svb-REP, and peaks detected only in Svb-ACT condition (Figures 1E–1H). Besides quantitative differences, we therefore interpret these data to imply that Svb-ACT and Svb-REP bind to largely overlapping sets of regions across the genome, consistent with the fact that the two Svb isoforms share most of their sequence, including the DNA-binding zinc-finger domain.

To characterize the chromatin landscape allowing Svb binding, we investigated whether Svb-bound regions displayed recognizable histone mark patterns. In a first step, we made use of dataset available in *Drosophila* cells devoid of Svb.[34–38] Then we performed our own ChIP-seq analysis in S2 cells. As detailed in the following, the two complementary approaches gave similar results. To allow higher sensitivity in peak detection, we used Svb-ACT peaks detected with relaxed parameter (MACS q value $10e^{-1}$, n = 6,939). A strong enrichment in histone 3 acetylated on K27 (H3K27ac) was found across regions to which Svb can bind (Figures 2A and 2B). Notably, H3K27ac enrichment directly correlated with Svb-ChIP signal (see heatmap intensity "gradients" in Figure 2A). H3K27ac marks open chromatin, in particular active enhancers,[37,39,40] which are also enriched in histone 3 methylated on K4.[41,42] Consistently, regions bound by Svb were enriched in H3K4me1 and H3K4me3 that accumulated on both sides of the peak (Figures 2A and 2B). In contrast, H3K27me3 that marks repressive chromatin[43] was depleted at Svb binding regions (Figure 2A), as best seen from the dataset taken from the literature. Although a similar behavior was also observed in our experimental conditions, the latter dataset of weaker quality was not retained for further analysis. Therefore, these results obtained with different sets of data (integrated or experimental) suggest that Svb binds active enhancers. We further tested this hypothesis by analyzing Svb binding to all regions acting as enhancers in naive S2 cells (Figures 2C and 2D), as defined by short for self-transcribing active regulatory region sequencing (STARR-seq).[44] We found a significant overlap between Svb-bound regions and functional enhancers. Moreover, enhancers associated with high Svb signal display stronger enhancer activity as defined by STARR-seq signal (brown) and higher H3K27ac (blue), compared to enhancers that cannot be bound by Svb (Figures 2C and 2D). Taken together, these results indicate that Svb preferentially binds to potent enhancers.

To analyze the respective influence of Svb isoforms on gene expression, we next performed RNA sequencing (RNA-seq) in parental S2, Svb-ACT, and Svb-REP cells. When compared to controls, both Svb-ACT and Svb-REP significantly changed gene expression (Figure 3A).

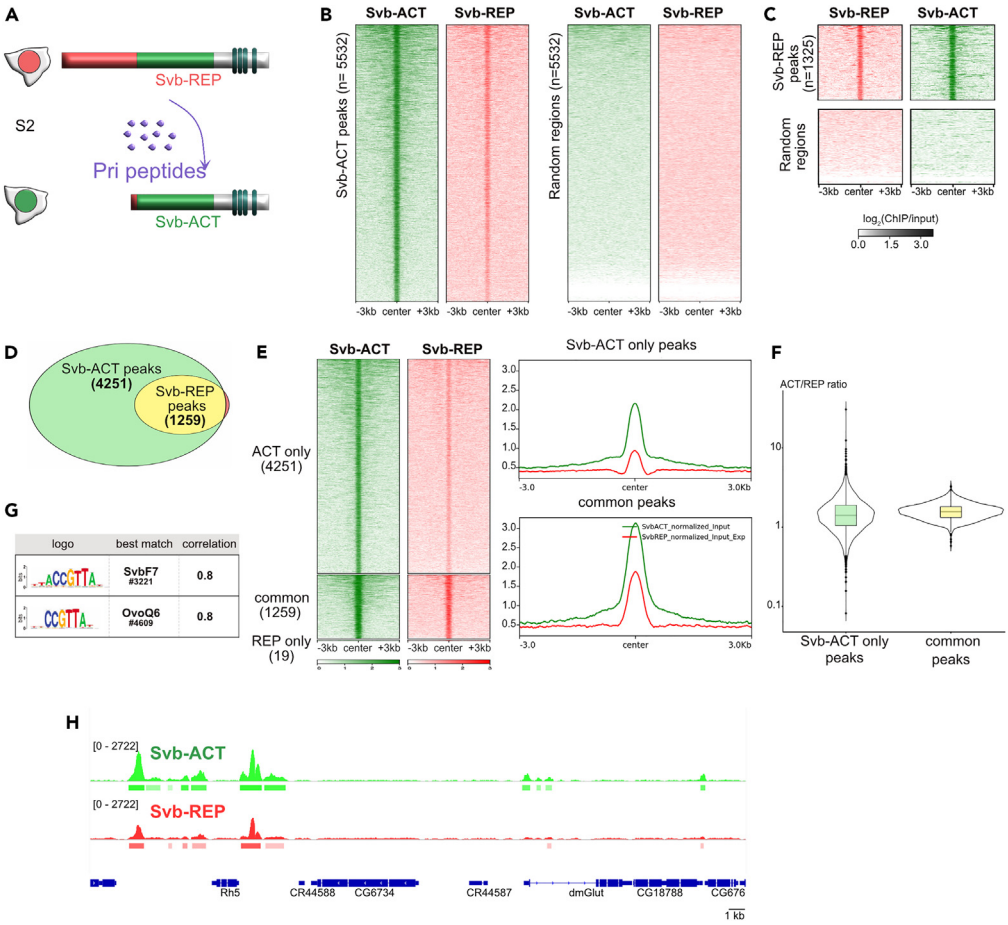


Figure 1. Svb activator and repressor isoforms bind common genomic region

(A) Cultured S2 cells provide an exquisite model system for studying the mode of action of individual Svb isoforms. Pri small peptides trigger Svb-REP to Svb-ACT proteasome-dependent post-translational processing.

(B) Heatmap of Svb-ACT (green) and Svb-REP (red) enrichment across all detected Svb-ACT peaks of the *Drosophila* genome (n = 5,532 detected by MACS with a q value of $10e^{-4}$). Each row represents a 6 kb window centered on peak midpoint, sorted by the intensity of Svb-ACT ChIP-seq signal. ChIP-seq signal intensity for a similar set of random genomic regions is show on the right.

(C) Heatmap showing Svb-ACT (green) and Svb-REP (red) ChIP signal along detected Svb-REP peaks (n = 1325, detected by MACS with a q value of $10e^{-4}$, top) and random regions (bottom).

(D) Venn diagram comparing the sets of MACS peaks detected for Svb-bound regions in Svb-ACT and Svb-REP cells. For the common population, overlapping regions are considered as single regions, which explains the slight difference in the total number of regions between the Venn diagram and the Svb-ACT and REP MACS peaks.

(E) Left panels show heatmap of Svb-ACT (green) and Svb-REP (red) ChIP signal at genomic regions detected as MACS peaks only in Svb-ACT cells (n = 4251, top), in both Svb-ACT and Svb-REP cells (n = 1259, middle), or only in Svb-REP cells (n = 19, bottom). Right panels represent the intensity of Svb-ACT and Svb-REP signal displayed as metaprofiles integrating signal of all peaks detected only in Svb-ACT cells (top), and all peaks detected in Svb-ACT and Svb-REP cells.

(F) Quantification of the ratio between Svb-REP and Svb-ACT signal, in regions corresponding to MACS peaks common to Svb-ACT and Svb-REP (n = 1259, red), and MACS peaks only detected in Svb-ACT cells (n = 4251, green). For both sets of bound regions, no significant difference was observed, p value = 0. 12.

(G) DNA motifs being the most enriched in Svb-ACT peaks as defined by RSAT suite, the correlation score reflects result of a Pearson test between the motif detected and logos present in the RSAT database.[32]

(H) Genome browser view of the *Rhr* locus showing respective intensity of Svb-ACT and Svb-REP ChIP-seq signal.

In vivo assays in embryonic epidermal cells have shown the antagonistic activity of Svb-ACT and Svb-REP on individual direct target genes,[24,26,27] leading to the upregulation or downregulation of those, respectively. Accordingly, in S2 cells, the genes upregulated upon Svb-ACT expression often display Svb binding peaks, whether we consider Svb-ACT peaks or Svb-REP peaks (Figure S2). This reinforces the idea that both isoforms bind a common set of target genes and regulate their transcription antagonistically as it is the case in embryo. Therefore, we selected all differentially expressed genes (DEGs, p value<0.05) between Svb-ACT and Svb-REP and focused on those responding to both isoforms, i.e., being upregulated by Svb-ACT and downregulated by Svb-REP (log_2(FC) > 0 in Svb-ACT and log_2(FC) < 0 in Svb-REP conditions, Figure 3A and Tables S1, S2, and S3). This supervised approach defined a group of 285 putative target genes. To

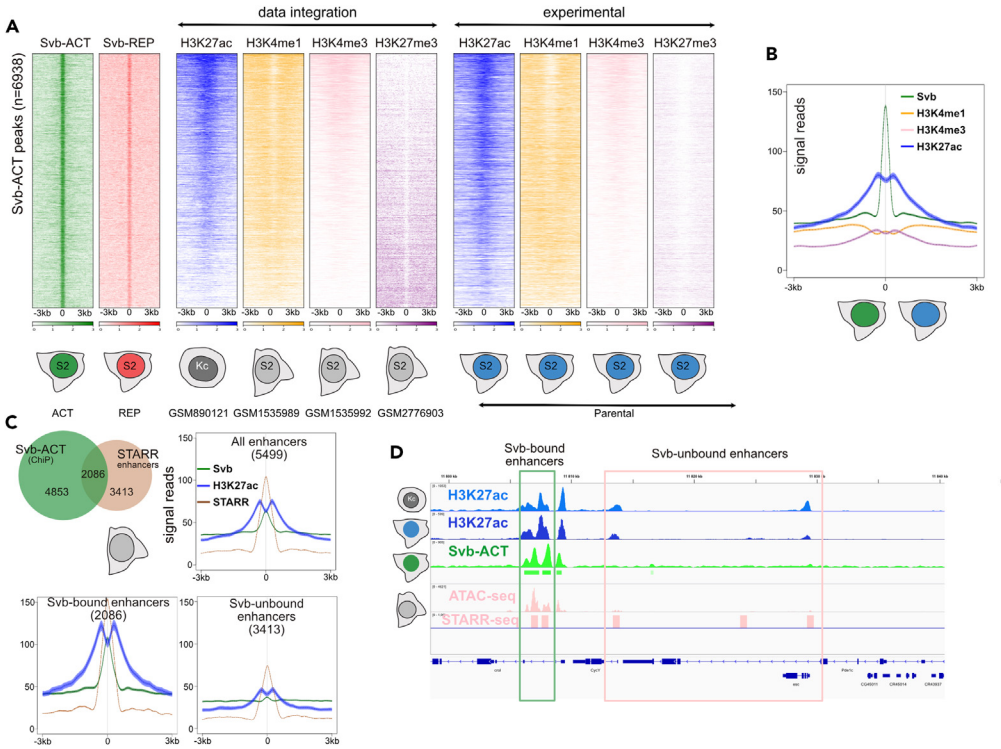


Figure 2. Svb binds chromatin regions with enhancer properties

(A) Heatmaps of Svb-ACT, H3K27ac (blue), H3K4me1 (yellow), H3K4me3 (pink), and H3K27me3 (purple) ChIP-seq enrichment across all peaks bound by Svb (n = 6939, MACS q value $10e^{-1}$). The first four panels (data integration) show ChIP-seq intensity for histone marks taken from literature (H3K27ac #GSM890121, H3K4me1 #GSM1535989, H3K4me3 #GSM1535992 and H3K27me3 #GSM2776903). To ensure proper comparison, raw data were re-processed using the procedures we standardized. The heatmaps on the right show results of ChIP-seq experiments we performed for the same histone marks in control S2 cell lines, devoid of both *svb* and *pri*. Each row represents a 6 kb window centered on Svb peak midpoint, sorted by Svb-ACT enrichment.

(B) Metaprofiles of histone marks from our experimental ChIPseq centered on Svb-ACT peaks. Because of poorer signal/noise ratio, data for H3K27me3 were not kept for analysis.

(C) Intersection between Svb-ACT peaks and active enhancers identified by STARR-seq.[44] The Venn diagram shows the overlap between Svb-ACT-bound regions (green, n = 6939) and active enhancers (brown, n = 5,499). Metaprofiles of 6 kb centered on Svb peak midpoint show the average enrichment of Svb-ACT (green), STARR signal (brown), and H3K27ac (blue) for all active enhancers, or Svb-bound and Svb-unbound enhancers.

(D) Genome browser view of a characteristic genomic region, showing the behavior of H3K27ac (GSM890121, Kc cells), H3K27ac (this study, S2 cells), and ATAC-seq (GSE103177, S2 cells) around STARR-seq identified enhancers (S2 cells), bound or not by Svb-ACT.

further assess the influence of Svb-REP and Svb-ACT on gene expression, we developed an unsupervised approach based on hierarchical clustering and k-means methods to identify Svb direct target genes solely based on RNA-seq data. Hierarchical clustering[45] defines 5 groups of DEGs (Figure 3B), one being very similar to the supervised set (278/285 genes in common, Figures 3C and 3E). We next used k-means clustering to challenge the robustness of each hierarchical cluster gene sets and calculated a "robustness score" that represents the proportion of individual genes staying clustered in a same group throughout 15 k-means iterations (r-score, Figure 3D, see STAR Methods). The cluster comprising putative Svb targets displayed the highest r-score, and 195 genes were common to all iterations and methods (Figures 3E and S1F). Notably, 86% of those are associated to Svb binding peaks (Figure 3F), further suggesting that this set of 195 genes represents the core direct targets of Svb in S2 cells (see Figure 3G for a representative target gene). Our method to define Svb target genes using RNA-seq data clustering was stringent as genes need to be both activated by Svb-ACT and repressed by Svb-REP, leading to a highly selective set of target genes. Indeed, many genes associated with Svb peaks do not behave as target genes showing that only a fraction of Svb binding events leads to changes in gene expression.

Taken together, these data show that Svb-ACT and Svb-REP display antagonistic transcriptional activities, acting on a common core set of approximately 200 genes defined by a combination of experimental and bioinformatic supervised and unsupervised approaches.

Chromatin landscape defines Svb-binding sites

Having defined the set of Svb target genes in S2 cells, we next asked how it relates to genes regulated by Svb in the embryonic epidermis,[28] focusing on direct target genes (i.e., being downregulated in embryos lacking Svb, or Pri, and displaying a Svb binding peak in embryos). No

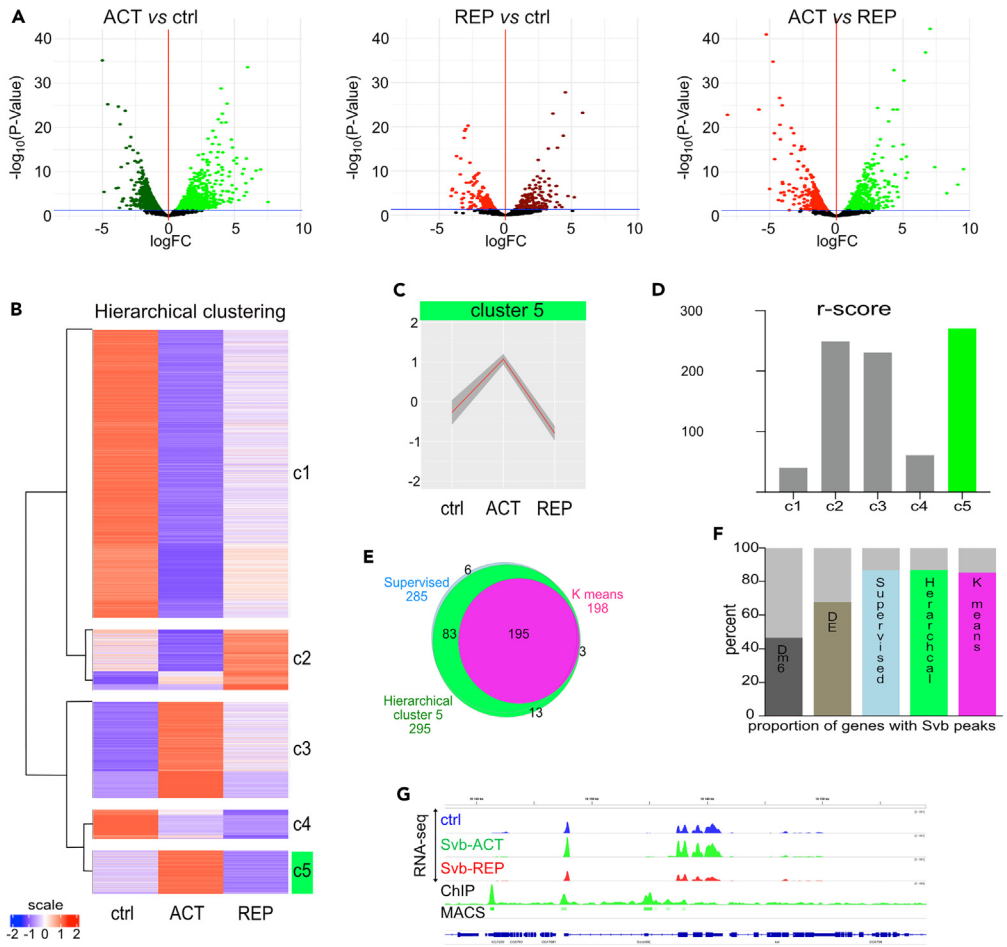


Figure 3. Identification of Svb direct target genes in S2 cells

(A) Volcano plots show differentially expressed (DE) genes upon Svb-ACT or Svb-REP expression in S2 cells.

(B) Hierarchical clustering of DE genes across RNA-seq experimental conditions.

(C) Metaprofile of all genes of cluster 5 in the hierarchical clustering (B).

(D) r-score defines the robustness of each gene cluster from hierarchical clustering, based on the conservation of same individual genes through 15 iterations of k-means clustering. See Figure S1 and STAR Methods for details. c5 is highlighted in green.

(E) Venn diagram showing a robust set of 195 target genes commonly retrieved through supervised (p value <0.05 in Svb-ACT vs. Svb-REP, AND $\log_2$FC > 0 in Svb-ACT vs. ctrl, AND $\log_2$FC < 0 Svb-REP vs. ctrl), hierarchical and k-means clustering methods.

(F) Diagram showing the proportion of genes with SvbACT ChIP peaks in the whole genome (Dm6), all differentially expressed (DE) genes, or Svb putative target genes defined by the supervised method, hierarchical clustering or k-means clustering.

(G) Genome browser screenshot of the *Soc36E* locus displaying typical Svb target gene behavior (basal expression in blue, Svb-ACT upregulation in green; Svb-REP downregulation in red, Svb ChIP in green).

significant overlap between the two sets of genes was observed (Figure 4A), showing that the transcriptional response to Svb strongly depends on the cell type. Moreover, this divergence was also observed when comparing Svb ChIP peaks between embryonic and S2 cells (Figure 4B), in which Svb binds mostly distinct genomic regions. These results therefore suggested that the difference in the sets of genes regulated by Svb across cell types was driven by changes in Svb binding sites occupancy.

To get a deeper look at the chromatin context of Svb binding, we focused on evolutionarily conserved F7-BS motifs that often mediate transcription of Svb direct targets *in vivo*[28] and are also highly enriched within Svb peaks in S2 cells (Figure 4B). There are 2,354 evolutionarily conserved F7-BS in the *Drosophila* genome, 26% of those being bound by Svb in S2 cells (Figure 4Da + b). We analyzed chromatin accessibility around F7-BS using available Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAC-seq) data from S2 cells[46] and observed that F7-BS that can be bound by Svb in S2 cells are open, while non-bound sites are often in a closed conformation (Figure 4C). Of note, the binding status of F7-BS in embryos has no effect on the chromatin accessibility in S2 cells (a vs. b and c vs. d). We further analyzed histone post-translational modifications (using published datasets) to inspect the chromatin landscape around F7-BS. This analysis showed that Svb can bind to F7-BS (a and b) that bear hallmarks of active enhancers: higher H3K27ac, H3K4me1, and H3K4me3 and lower

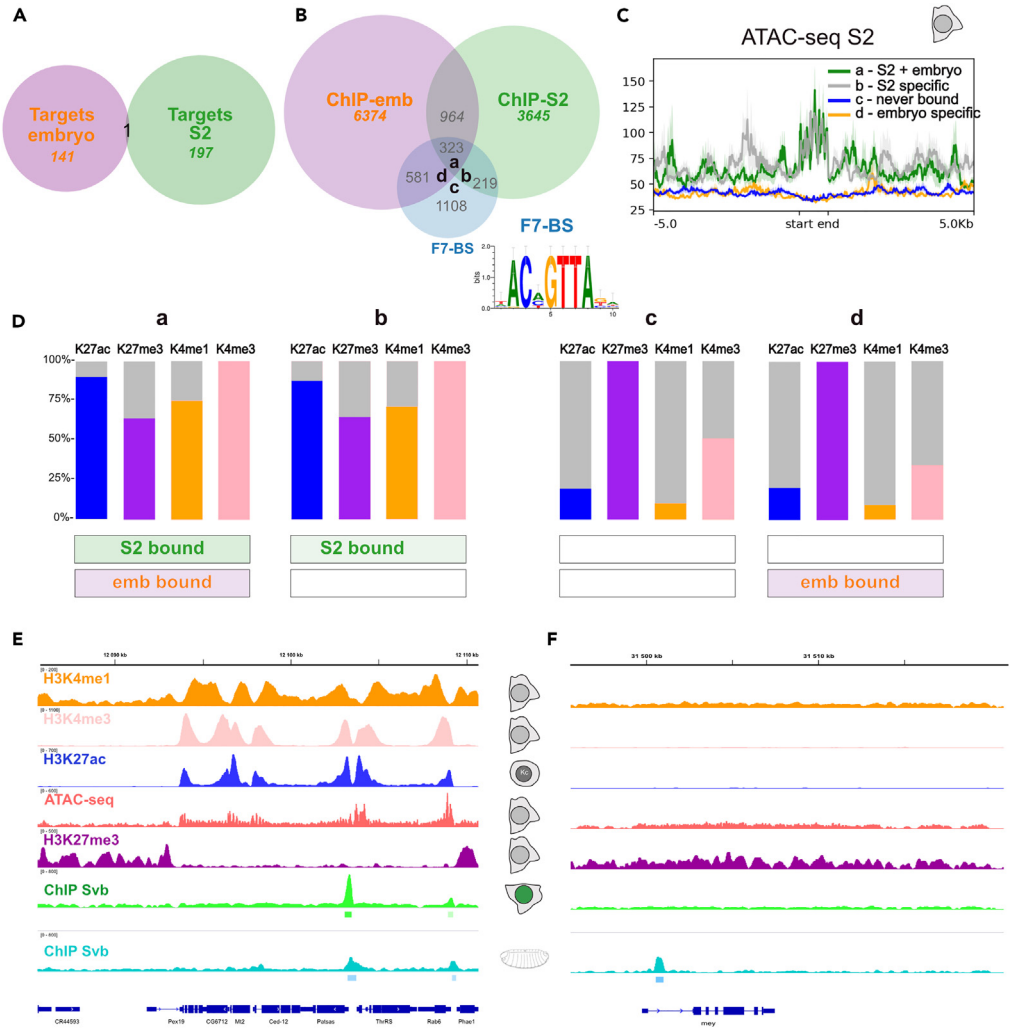


Figure 4. Chromatin landscape and accessibility define Svb binding

(A) Venn diagram showing the intersection between the sets of Svb target genes in S2 cells (n = 198) and in embryonic epidermal cells (n = 142,[28]).

(B) Venn diagram showing the intersections between Svb-bound regions in S2 cells (green), in embryonic epidermal cells (purple), and genomic F7-BS regions (blue). F7-BS position weight matrix is shown below. Each F7-BS is thus bound or not by Svb in either cell type, defining 4 possible behavior: a, b, c, and d mark the 4 groups of F7-BS further investigated in panels (C) and (D).

(C) ATAC-seq metaprofiles from S2 cells[46] show chromatin accessibility over 10 kb around F7-BS for the clusters defined in B.

(D) Enrichment of the indicated histones marks around F7-BS motifs in the different sets defined in (B).

(E and F) Snapshot of genome browser showing Svb ChIP (from S2 cells in green and from embryos in cyan) at two genomic loci. *ThrRS* is bound by Svb in both S2 cells and embryos (E). For *mey* (F), while Svb binds to an enhancer driving *mey* gene expression in embryonic epidermal cells,[28] no Svb binding was detected in S2 cells. Histone marks (H3K4me1 in orange, H3K4me3 in pink, H3K27ac in blue, H3K27me3 in purple) and ATAC-seq (red) allow visualization of the chromatin organization of corresponding regions in S2 cells. When compared to *ThrRS*, the data show a repressive/closed configuration of chromatin at the *mey* locus correlating with the and absence of Svb binding in this cell type.

H3K27me3, as compared to non-bound F7-BS (Figures 4D–4F). Again, this is independent of Svb ability to bind these sites in another context (e.g., in embryo, a vs. b and c vs. d), showing that the repertoire of Svb binding sites is strongly dependent on chromatin landscape.

Together, these results indicate that the Svb TF regulates different sets of target genes in S2 cells and embryonic epidermal cells. Specific patterns of histone marks highlight how Svb-responsive genes may be gated by distinct pool of sites that are accessible to Svb binding, hence defining a subset of cell-type-specific target genes regulated by Svb.

SvbREP influences the chromatin landscape

Since chromatin accessibility constrains Svb binding, we wondered whether Svb might in turn affect chromatin organization. Svb-ACT and REP have distinct subnuclear localization: Svb-ACT distributes diffusely throughout the nucleus, while Svb-REP accumulates in foci[24,25], raising

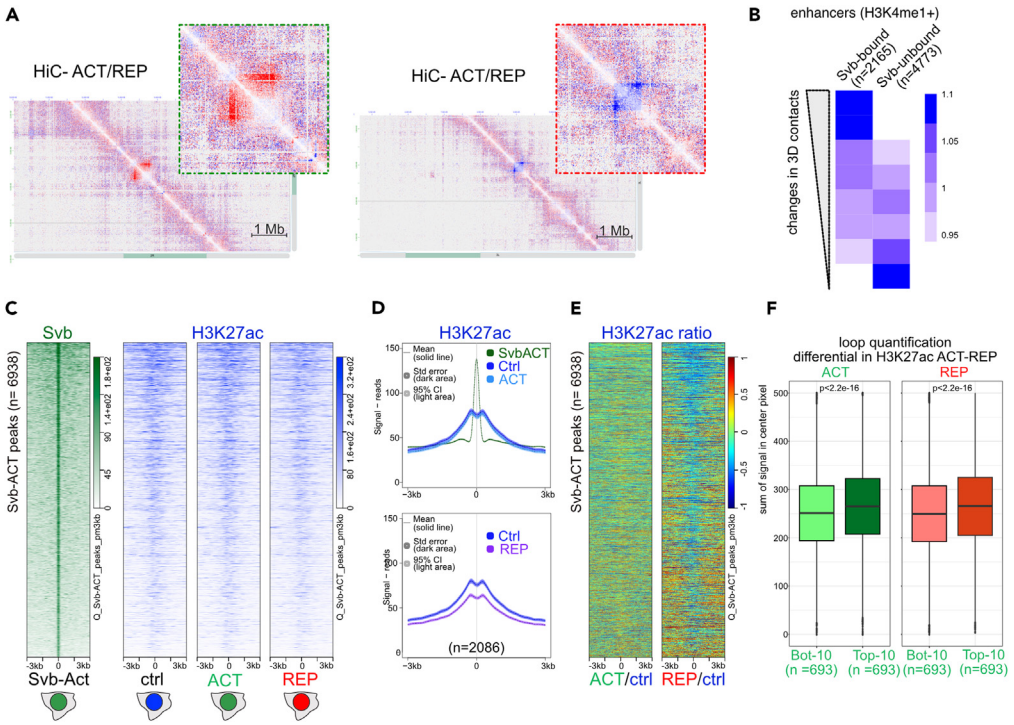


Figure 5. The SvbREP-to-ACT switch is accompanied by dynamic remodeling of the chromatin landscape

(A) 2D plot showing long-range interactions probed using chromosome conformation capture (3C/Hi-C) in cells expressing Svb-ACT compared to Svb-REP. Enlarged areas (dotted squares) highlight possible 3D long-range contacts. The red and blue regions display stabilized interactions in Svb-ACT or -REP conditions, respectively.

(B) Genome-wide analysis of long-range contacts between enhancer and promoters in Svb-ACT cells compared to Svb-REP cells. Long-range contacts were measured for all putative enhancer-promoter pairs in Svb-ACT versus Svb-REP conditions (see STAR Methods); net variations were quantified and ranked from highest to lowest changes in 3D contacts. Enhancer as identified by STARR-seq[44,47] were classified as bound or unbound by Svb by intersecting their coordinates with the Svb peaks as identified in this study. A Fisher's exact test was used to assess the enrichment of Svb-bound and Svb-unbound enhancers in each quantile.

(C) Heatmap showing both Svb-ACT binding (green) and H3K27ac abundance (blue) at Svb-bound regions in control, Svb-ACT, and Svb-REP cells. All heatmaps are ranked by Svb ChIP intensity (MACS p value $10e^{-4}$, 6939 peaks from Svb-ACT cells).

(D) Metaprofiles overlay of (C) (ACT vs. ctrl and REP vs. ctrl).

(E) Heatmap showing the ratio of H3K27ac ChIP signal in ACT vs. control (left) and REP vs. control.

(F) Boxplots quantifying the levels of 3D contacts between enhancers with nearby TSSs in cells expressing Svb-ACT or Svb-REP (left and right boxplot, respectively). Note that the 3D contacts are higher for enhancers with the highest differences in H3K27ac levels (top 10%; SvbACT>Svb-REP; dark green or red) as compared to those with the least differences (bottom 10%; SvbACT = Svb-REP; light green or red), as validated statistically for cells expressing Svb-ACT or Svb-REP (p value by Wilcoxon pairwise test).

the possibility that Svb isoforms could affect higher-order chromatin organization. We addressed this question with chromosome conformation capture (Hi-C) experiments performed in Svb-ACT or Svb-REP S2 cells. Inspecting long-range contacts at high resolution (5 kb) using 2D plots, we detected long-range contacts that are more frequent in Svb-ACT or in Svb-REP condition, which could reflect the formation or stabilization of specific loops (Figure 5A).

To further test the specificity of these 3D interactions, we considered the net changes in enhancer-promoter interactions when comparing Svb-ACT and REP conditions, genome wide. The enhancer-promoter pairs were then ranked according to such net changes in 3D contacts between the two conditions. For each decile from the list, we estimated the enrichment for Svb-bound or Svb-unbound enhancers (Figure 5B and STAR Methods). Svb-bound enhancers were enriched among the enhancers on which the most significant change in 3D contacts was scored in Svb-REP compared to Svb-ACT, unlike for unbound enhancers (Figure 5B). These data thus show that Svb impinges on chromatin organization, as monitored by changes in long-range enhancers-promoters 3D contacts.

Next, we assessed whether Svb affected the chromatin landscape at a local scale. First we confirm that, among accessible chromatin regions defined by ATAC-seq, Svb can bind only to regions that display high H3K27ac mark, a key mark of active enhancers,[39] while unbound regions have no detectable H3K27ac (Figure S3A). We analyzed the level of H3K27 acetylation upon Svb-ACT or Svb-REP expression and found that H3K27ac levels stayed unchanged upon Svb-ACT binding (Figures 5C–5E and S3A). On the contrary, the binding of Svb-REP induced a local decrease in H3K27ac (Figures 5C–5E, S3A, and S3B). The amplitude of this effect depends on the intensity of Svb peaks, as revealed by the

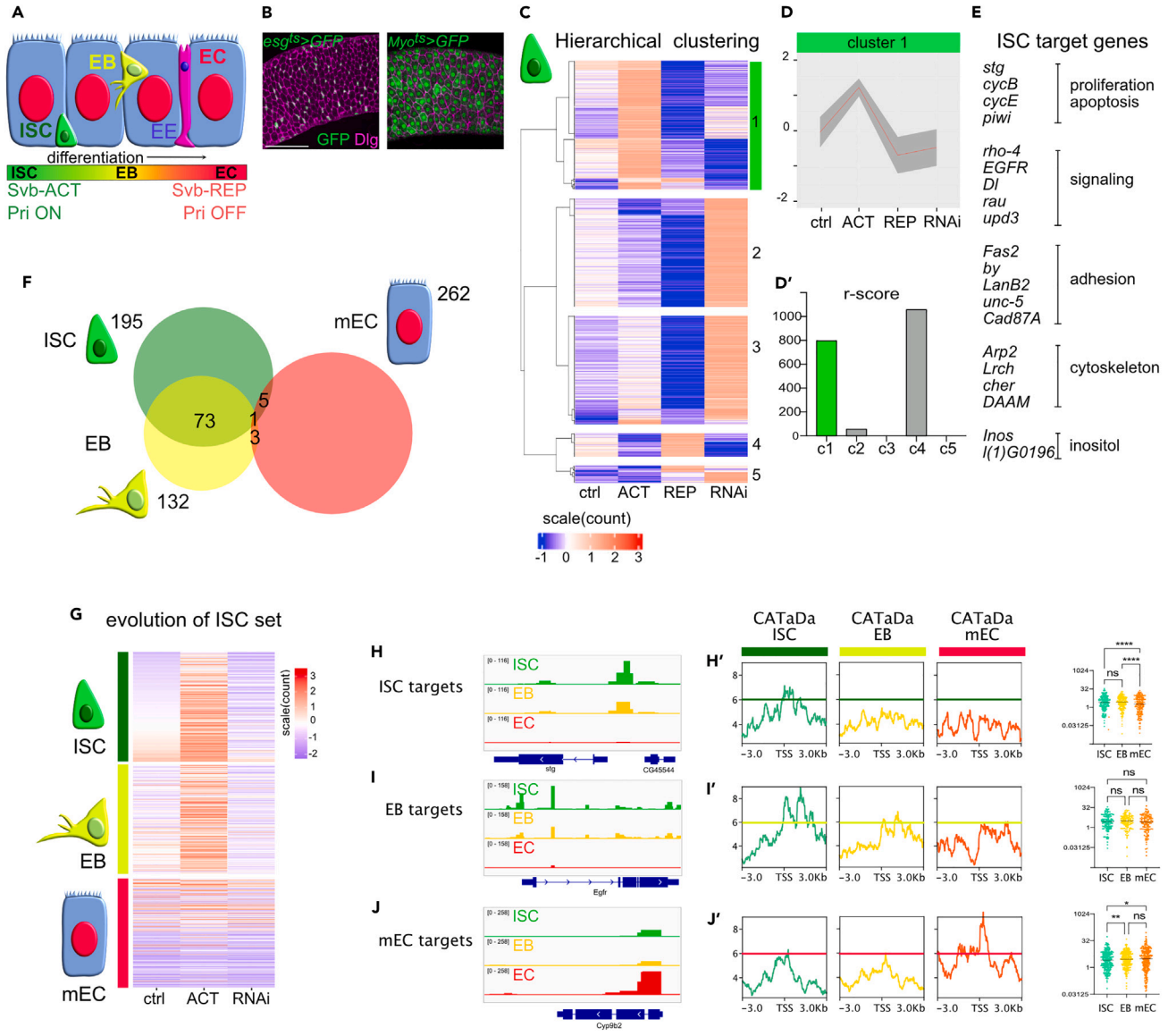


Figure 6. Evolution of Svb target genes through the lineage of intestinal stem cells

(A) Schematic representation of the intestinal stem cell lineage. Svb-ACT is the main form in ISCs and Svb-REP in ECs. ISC in green, EB in yellow, EC in blue, and EE in purple.

(B) Cell-type-specific drivers have been used to manipulate Svb expression in adult fly progenitors (ISCs and EBs, *esg:GAL4, Gal80^{ts}*) or ECs (*Myo1A:GAL4,Gal80^{ts}*). Driver expression is visualized thanks to a *UAS>GFP* transgene. The cell contour is visualized using Dlg protein labeling.

(C) Hierarchical clustering of RNA-seq data from sorted ISCs reveals DE genes upon manipulation of Svb function (Table S5). Four experimental conditions were used: *control* (ctrl), *UAS-SvbACT* (ACT), *UAS-Svb^{3Kmut}* (REP), and *UAS-Svb RNAi* (RNAi). Cluster 1 comprises genes activated by SvbACT and repressed by SvbREP, which are also downregulated upon Svb loss of function. Scale bar represents 100 μm.

(D) Metaprofile of cluster 1 corresponding to putative Svb target genes in ISCs. (D') Robustness analysis (r-score) of each hierarchical cluster.

(E) Subset of genes from cluster 1 with their associated functions.

(F) Venn diagram showing how Svb target gene sets evolve during intestinal cell differentiation. Note the similarity between Svb target genes in ISCs and EBs, as opposed to Svb targets in ECs. Svb target genes were identified in ISCs, EBs, and early and mature ECs (eECs and mECs, respectively) using the same pipeline (see Figure S5, Tables S1 and S2).

(G) Heatmaps showing the expression of the 195 ISC target genes in ctrl, Svb-ACT, and *svb-RNAi* conditions, in ISC, EB, and mEC (top to bottom). For progenitors (ISCs and EBs) and ECs, transgenes were expressed thanks to *esg^{ts}* and *Myo^{ts}* respectively.

(H–J) CATaDa data[50] allow studying the dynamics of chromatin accessibility through the lineage. (H, I, J) are IGV genome browser screenshots of representative loci of an ISC target gene (*stg*, H), an EB target gene (*egfr*, I) and a mEC target gene (*Cyp9b2*, J), showing CaTADA signal in ISC (green), in EB (yellow), and in EC

Figure 6. *Continued*

(red). Averaged CATaDa signal profile (in reads per million), centered on the TSS of all ISC target genes (H'), all EB target genes (I') and all mEC target genes (J') are shown, showing accessibility in ISC (green), EB (yellow) and EC cells (red) (left to right of each panel). The plots on the right present for each gene the CATaDa signal around the TSS (1kb on both side). Statistical analysis confirms the general trend that TSS of Svb target genes in ISCs close during intestinal cell differentiation, while TSS of Svb targets in ECs become open (Kruskal-wallis test with Dunn's multiple comparisons).

REP/control ratio (Figure 5E). We next wondered whether this effect is specific to Svb-bound regions and analyzed enhancers considered as "non-bound" by Svb. The level of H3K27ac at these enhancers is low; however, one can still see a decrease in H3K27ac. To determine whether this decrease could be due to weak Svb binding or if it was independent of Svb, we analyzed the relationship between Svb binding and H3K27ac in all enhancers defined by STARR-seq (Figure S3C). In this aim, we divided the two categories of enhancers (Svb-"bound" versus "unbound") in two groups depending on the level of H3K27 acetylation. Then we analyzed the Svb-ChIP signal in these 4 groups. This analysis reveals that, among the "non-bound" enhancers, the top 50% acetylated enhancers display detectable Svb binding which could explain the observed reduction of H3K27ac upon SvbREP expression on these enhancers. In addition, we also observe a stronger Svb-ChIP signal on "bound" enhancers with the highest level of H3K27ac (Figure S3C). Together these results suggest that Svb binding globally correlates with the level of H3K27 acetylation on enhancers. In turn, Svb-REP binding directly impinges on H3K27 acetylation of active enhancers.

We next tested whether Svb-bound regions that displayed the strongest changes in the levels of H3K27ac were those involved in 3D contacts (Figure 5E). To readily test this, we used aggregation plots as previously developed by us and others[48,49] to measure the frequency of long-range contacts between Svb-bound regions (defined by ChIP-seq) and gene transcription start sites (TSSs) (Figure S4).We found that Svb-binding sites showing notable changes in H3K27ac levels corresponded to regions of higher 3D contacts with neighboring TSSs, in contrast to regions of low Svb binding (Figure 5F). However, it should be noted that this analysis did not reveal a clear link between differential H3K27ac level and differential 3D contacts with TSS at Svb-bound regions (Figure 5F). It is therefore most likely that changes in 3D contacts at Svb sites increase or decrease as a function of context, as illustrated in Figure 5B.

Overall, these results show that Svb binding to enhancers correlates with the level of H3K27ac and that Svb isoforms impact the chromatin landscape through changes in long-range 3D contacts and local histone modification. Considering the mechanism of Svb processing by proteasomal maturation of the Svb-REP into Svb-ACT isoform, this suggests that Pri peptides trigger a rapid switch in chromatin state that may have functional impact on transcription.

Identification of Svb target genes across the ISC lineage

We next investigated the transcriptional response to Pri-induced switch in Svb isoforms within the physiological context of adult ISCs leading to EC differentiation (Figure 6A). Pri is only expressed in ISCs and early EB progenitors[30] triggering the production of Svb-ACT that acts to promote their maintenance and proliferation. In later stages of the lineage, Svb-REP triggers and accompanies differentiation into ECs.[30] Svb-REP accumulates during differentiation (Figure 6A) and is required in mature ECs (mECs) to maintain their epithelial properties.[30] Thus, the transition from Svb-ACT to REP isoforms has a drastic effect on ISC and EC behavior, but the target genes mediating Svb function(s) remained unknown.

To identify genes regulated by Svb in each cell type from the lineage, we profiled the transcriptome of four cell types purified from posterior midguts: ISCs, EBs, early EC (eEC), and mEC (see STAR Methods). These cells were isolated by fluorescence-activated cell sorting (FACS) from dissected guts of control animals, upon RNAi-mediated *svb* depletion and expression of Svb-ACT or Svb-REP^{3Kmut} (an engineered form of Svb-REP insensitive to Pri and thus not convertible into Svb-ACT[25]). The manipulation of Svb function in the adult gut was specifically targeted either to progenitor cells (ISC and EB) or to ECs (eEC and mEC), using cell-type-specific drivers and temporal control (Figure 6B; Jiang et al. and Micchelli et al.[51,52]). To identify primary targets of Svb, we purified cells at 6 days of induction, a time point where the phenotypes are still limited, with Svb-ACT initiating proliferation of progenitor cells and Svb-REP promoting EC differentiation (Figure S5). We also verified that our RNA-seq datasets were consistent with available transcriptomic data[53] and indeed represented distinct cell types along the ISC to EC differentiation route (ISC, EB, and EC Figure S6A). Building on the unbiased method developed in cultured cells (Figure 3 and Tables S1, S2, and S3), we then identified the set of genes whose expression is likely directly regulated by Svb. In each cell type, we identified one robust cluster of 132–265 genes antagonistically regulated by Svb-ACT versus Svb-REP isoforms, as expected for direct target genes (Figures 6C–6G and S6B). The fact that their expression was also downregulated upon *svb* knockdown in progenitors confirms that these genes are normally activated by Svb-ACT in ISCs/EBs (Figures 6C, 6G, and S6B). On the contrary, Svb target genes in mECs tend to be upregulated upon *svb* knockdown, confirming that Svb acts mostly as a repressor in these cells (Figure S6B and Tables S1, S2, and S3).

In agreement with the function of Svb and Pri in ISCs/EBs,[30] we identified Svb target genes controlling cell cycle, proliferation, and survival of stem cells (*string, cycE, cycB, piwi* …), as well as signaling pathways known to regulate stemness (*egfr, Dl*, reviewed in Hung et al., Kohlmaier et al., Sousa-Victor et al., and Boumard et al.[54–56]) (Figure 6E). Our results show that these genes can be antagonistically regulated by Svb isoforms, which reinforces the importance of the transition between Svb-ACT and REP in the ISC lineage: the repression of such regulators by Svb-REP might favor cell cycle exit and commitment toward differentiation.

Interestingly, the sets of Svb target genes were markedly different in progenitors versus differentiated cells from the intestinal lineage (Figures 6F and S6C, and Tables S1, S2, S3, S4, and S5). Focusing on the ISC targets (195 genes), we found that their expression and regulation evolve during differentiation: 73 remain regulated by Svb in EB, 25 behave as Svb target genes in eEC, and only 5 genes are targets of Svb in mEC. Thus, the change in Svb target genes occurs gradually (Figures 6F and S6C). Accordingly, the expression of Svb target genes in ISCs tends to decrease during their differentiation, as well as the ability of Svb-ACT to activate their transcription (Figure 6G).

It has been shown that chromatin accessibility in ISCs evolves as they progressively differentiate into EBs and ECs.[50] We analyzed chromatin accessibility for Svb target genes in ISCs, EBs, and mECs using data obtained by Chromatin Accessibility Targeted DamID (CATaDa), a technique based on driving methyltransferase expression in specific cell populations along the lineage.[50] We observed that the chromatin at Svb target genes in ISCs and EBs is mostly accessible in ISCs, still open in EBs, but closed in differentiated ECs (e.g., *stg* and *egfr*, Figures 6H, 6I, and S6D). On the contrary, Svb target genes in ECs tend to be more accessible in differentiated ECs (e.g., Cyp9b2 [Figure 6J] and *Mal-A1* [Figure S6E]). We tested whether this trend was global by averaging CATaDa signal from entire gene sets corresponding to cell-specific Svb target genes (Figures 6H′, 6I′, and 6L′). This confirms that, as cells differentiate, the TSSs of Svb target genes in ISCs and EBs close down, while those of mECs become accessible.

All together these results show that Svb target gene sets evolve gradually during intestinal cell differentiation, concomitant with changes in chromatin accessibility, and Svb-ACT to Svb-REP isoform transition. This suggests that, as observed in cultured cells, the chromatin state governs Svb-accessible regions thereby constraining its transcriptional output.

DISCUSSION

How binding site occupancy sculpts tissue-specific response of genome expression to a given TF remains poorly understood. OvoL1,2,3 factors encode transcription activating or repressive isoforms that are involved in a growing number of contexts, during development, differentiation, or physiology, in somatic stem cells, as well as in various human pathologies. Using the simpler *Drosophila* as a paradigm of OvoL functions, we report detailed characterization of changes in genome expression in response to Svb-ACT and Svb-REP isoforms across distinct contexts, in particular along the lineage of ISCs in which they are playing a dual role.

Svb-ACT and REP bind a number of overlapping genomic regions which display characteristics of active enhancers (high levels of H3K27ac, H3K4me1, and me3; low levels of H3K27me3; significant activity in STARR-seq). Of note, the level of fixation of Svb correlates with the level of H3K27 acetylation (Figure S3C). A limited fraction of these enhancers allows Svb isoforms to antagonistically control the transcription of a core set of about 200 target genes, which are strikingly distinct in different cell types. We show that chromatin accessibility is a condition for Svb binding, providing a first explanation for the divergence of Svb target genes between cell types. In the ISC lineage, Svb target genes evolve during differentiation of ISCs into ECs, as does chromatin accessibility.[50] Focusing on the set of Svb target genes in ISCs, we find that chromatin closes locally during differentiation, which prevents ectopic expression of these genes in ECs, even upon Svb-ACT misexpression (Figures 6G–6J). This is reminiscent of the situation in S2 cells, where embryonic enhancers are closed and unresponsive to Svb, suggesting that the mechanism discovered in S2 cells is relevant to the intestinal lineage. Thus, chromatin dynamics, even in the same lineage, may explain why the sets of Svb target genes can be strikingly distinct in different cell types: while some target genes get closed and unregulatable, other binding sites open, allowing regulation of different target genes.

Interestingly, Svb isoforms can affect chromatin 3D organization at specific loci (Figure 5A), promoting differential long-range contacts between bound regions and gene promoters, as seen in S2 cells (Figure 5B). In addition, Svb-REP favors modification of local chromatin landscape at bound enhancers, promoting deacetylation of H3K27. This suggests a mechanism of active repression, potentially through specific recruitment of chromatin-modifying enzymes mediated by the N-terminus domain of Svb-REP which is missing in Svb-ACT. Considering that Pri peptides induce proteasome degradation of this domain (the molecular switch from Svb REP-to-ACT isoforms), this work highlights that small peptides and the nuclear proteasome can impact genome organization and expression.

In the intestinal lineage, Svb-REP is expressed in both progenitors and differentiated EC. Pri expression in ISCs allows Svb-ACT maturation, while Svb-REP accumulates in differentiating cells devoid of Pri. In ISCs, we identify Svb target genes (*egfr, stg* …) whose expression is critical for stemness and proliferation. The progressive accumulation of Svb-REP occurring in EBs could contribute to downregulate these genes, consistent with cell-cycle arrest and differentiation. This can explain the phenotypes observed for Svb-ACT or Svb-REP in ISCs, promoting, respectively, proliferation and differentiation (Al Hayek et al.,[30] Figure S6).

Notably, the EGF Receptor (EGFR) pathway was previously shown to directly control *svb* transcription,[30] forming a feedback loop that could reinforce the asymmetry between ISC daughter cells. In ISCs, high Svb-ACT would maintain high EGFR, and high Dl, while cells starting to accumulate Svb-REP would decrease EGFR and Dl transcription, triggering differentiation.

SvbREP accumulation during differentiation is concomitant with transcriptional silencing of Svb target genes in ISCs. As in S2 cells, Svb-REP likely participates in deacetylation of H3K27 in the intestinal lineage. This could mediate direct gene repression through inhibition of enhancer activity. In addition, this histone modification could foreshadow the chromatin compaction that occurs during differentiation, which mediates long-term silencing of "stemness" genes.

Our study highlights a crosstalk between chromatin dynamics and Svb isoforms, defining a cell-specific transcriptional output which could ultimately control cell fate acquisition.

Limitation of the study

For technical reasons, Svb ChIP-seq and Hi-C studies have not been carried out in the fly intestine and were performed in S2 cells. Consequently, Svb target genes were defined in the ISC lineage from RNA-seq data (performed for each intestinal cell type and in several genetic backgrounds), following the methodology developed and validated in S2 cells. The analysis of Hi-C data revealed 3D chromatin modifications upon expression of either Svb-ACT or Svb-REP isoforms. However, how these 3D modifications directly impinge on the transcriptional regulation of individual Svb target genes remains to be further explored. This is in part due to the limitations of the Hi-C data to translate variations captured in 3D, between each gene and the full complement of its enhancers, into changes in gene expression.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Material availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - S2 cells lines
  - *Drosophila* lines
- METHOD DETAILS
  - Cell culture
  - ChIP experiments
  - RNA extraction, sequencing on S2 cells
  - *Drosophila* genetics
  - Gut dissection and cell dissociation
  - Sample preparation for FACS
  - Sample preparation for straining mature EC
  - RNAseq, sequencing and quality check on intestinal cells (ISC, EB, eEC and mEC)
  - Hi-C
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - RNAseq analysis
  - ChIPseq analysis
  - Clustering
  - Gene set enrichment analysis (GSEA)
  - Hi-C analysis
  - CATaDa analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108624.

ACKNOWLEDGMENTS

We thank Marion Aguirrebengoa (BigA platform) for her bioinformatic expertise, Emmanuelle Näser from the Cell sorting service of the TRI genotoul platform, and Hugues Parrinello from the MGX sequencing platform. We thank all the members of the Payre laboratory for discussions on the manuscript. This work was supported by ANR (Chrononet, to F.P.'s team; Helico, O.C.'s team), Fondation pour la Recherche Médicale (FRM, grant DEQ20170336739 to F.P.'s team and DEQ20160334940 to O.C.'s team), and an FRBT grant. A.M.-F., A.R., and D.D. were supported by fellowships from Ligue contre le Cancer. C.I. was supported by FRM and Université Toulouse 3 Paul Sabatier. We would like to thank the MCD unit for covering the publication fees.

AUTHOR CONTRIBUTIONS

A.M.-F. performed ChIP-seq experiments for histone marks in S2 cells, with the help of J.A., and most computational analyses with the help of A.R. (unsupervised clustering and GSEA), D.D. (Histones ChIP-seq), and C.I. (intestinal cells RNA-seq analyses). C.I. and C.P. dissected guts of different genotypes; C.I. did FACS purification and prepared samples for intestinal cells RNA-seq. P.L. and S.P. prepared S2 cell samples for RNA-seq and Svb::GFP ChIP-seq. G.A. and T.S. performed CATaDa analysis. O.F., N.C., N.S., and O.C. performed and analyzed HiC. O.C., S.P., F.P., and C.P. designed the project. F.P., O.C., and C.P. supervised the project. C.I., F.P., O.C., and C.P. wrote the paper with contribution from all authors.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: January 26, 2023
Revised: July 5, 2023
Accepted: November 30, 2023
Published: December 3, 2023

REFERENCES

1. Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.
2. Bulyk, M.L. (2003). Computational prediction of transcription-factor binding site locations. Genome Biol. *5*, 201.
3. Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. Nature *461*, 1248–U81.
4. Guertin, M.J., and Lis, J.T. (2010). Chromatin landscape dictates HSF binding to target DNA elements. PLoS Genet. *6*, e1001114.
5. Slattery, M., Zhou, T.Y., Yang, L., Machado, A.C.D., Gordan, R., and Rohs, R. (2014). Absence of a simple code: how transcription factors read the genome. Trends Biochem. Sci. *39*, 381–399.
6. Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Res. *22*, 1798–1812.
7. Buenrostro, J.D., Corces, M.R., Lareau, C.A., Wu, B., Schep, A.N., Aryee, M.J., Majeti, R., Chang, H.Y., and Greenleaf, W.J. (2018). Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell *173*, 1535–1548.e16.
8. Cuvier, O., and Fierz, B. (2017). Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells. Nat. Rev. Genet. *18*, 457–472.
9. Arvey, A., Agius, P., Noble, W.S., and Leslie, C. (2012). Sequence and chromatin determinants of cell-type-specific transcription factor binding. Genome Res. *22*, 1723–1734.
10. Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. *13*, 613–626.
11. Sen, S.Q., Chanchani, S., Southall, T.D., and Doe, C.Q. (2019). Neuroblast-specific open chromatin allows the temporal transcription factor, Hunchback, to bind neuroblast-specific loci. Elife *8*.
12. Kumar, A., Bhandari, A., Sinha, R., Sardar, P., Sushma, M., Goyal, P., Goswami, C., and Grapputo, A. (2012). Molecular phylogeny of OVOL genes illustrates a conserved C2H2 zinc finger domain coupled by hypervariable unstructured regions. PLoS One *7*, e39399.
13. Mevelninio, M., Terracol, R., Salles, C., Vincent, A., and Payre, F. (1995). Ovo, a Drosophila Gene Required for Ovarian Development, Is Specifically Expressed in the Germline and Shares Most of Its Coding Sequences with Shavenbaby, a Gene Involved in Embryo Patterning. Mech. Dev. *49*, 83–95.
14. Teng, A., Nair, M., Wells, J., Segre, J.A., and Dai, X. (2007). Strain-dependent perinatal lethality of Ovol1-deficient mice and identification of Ovol2 as a downstream target of Ovol1 in skin epidermis. Biochim. Biophys. Acta *1772*, 89–95.
15. Tsuji, G., Hashimoto-Hachiya, A., Kiyomatsu-Oda, M., Takemura, M., Ohno, F., Ito, T., Morino-Koga, S., Mitoma, C., Nakahara, T., Uchi, H., et al. (2017). Aryl hydrocarbon receptor activation restores filaggrin expression via OVOL1 in atopic dermatitis. Cell Death Dis. *8*, e2931.
16. Roca, H., Hernandez, J., Weidner, S., McEachin, R.C., Fuller, D., Sud, S., Schumann, T., Wilkinson, J.E., Zaslavsky, A., Li, H., et al. (2013). Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer. PLoS One *8*, e76773.
17. Wang, Z.H., Li, Z., Hu, M., Yang, Q.J., Yan, S., Wu, R.S., Li, B.A., and Guo, M. (2017). Ovol2 gene inhibits the Epithelial-to-Mesenchymal Transition in lung adenocarcinoma by transcriptionally repressing Twist1. Gene *600*, 1–8.
18. Watanabe, K., Villarreal-Ponce, A., Sun, P., Salmans, M.L., Fallahi, M., Andersen, B., and Dai, X. (2014). Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by Ovol2 transcriptional repressor. Dev. Cell *29*, 59–74.
19. Hong, T., Watanabe, K., Ta, C.H., Villarreal-Ponce, A., Nie, Q., and Dai, X. (2015). An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States. PLoS Comput. Biol. *11*, e1004569.
20. Lee, S., and Garfinkel, M.D. (2000). Characterization of Drosophila OVO protein DNA binding specificity using random DNA oligomer selection suggests zinc finger degeneration. Nucleic Acids Res. *28*, 826–834.
21. Lu, J., and Oliver, B. (2001). Drosophila OVO regulates ovarian tumor transcription by binding unusually near the transcription start site. Development *128*, 1671–1686.
22. Chanut-Delalande, H., Hashimoto, Y., Pelissier-Monier, A., Spokony, R., Dib, A., Kondo, T., Bohère, J., Niimi, K., Latapie, Y., Inagaki, S., et al. (2014). Pri peptides are mediators of ecdysone for the temporal control of development. Nat. Cell Biol. *16*, 1035–1044.
23. Payre, F., Vincent, A., and Carreno, S. (1999). Ovo/svb integrates Wingless and DER pathways to control epidermis differentiation. Nature *400*, 271–275.
24. Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., Kobayashi, S., Payre, F., and Kageyama, Y. (2010). Small peptides switch the transcriptional activity of Shavenbaby during Drosophila embryogenesis. Science *329*, 336–339.
25. Zanet, J., Benrabah, E., Li, T., Pélissier-Monier, A., Chanut-Delalande, H., Ronsin, B., Bellen, H.J., Payre, F., and Plaza, S. (2015). Pri sORF peptides induce selective proteasome-mediated protein processing. Science *349*, 1356–1358.
26. Chanut-Delalande, H., Fernandes, I., Roch, F., Payre, F., and Plaza, S. (2006). Shavenbaby couples patterning to epidermal cell shape control. PLoS Biol. *4*, e290.
27. Fernandes, I., Chanut-Delalande, H., Ferrer, P., Latapie, Y., Waltzer, L., Affolter, M., Payre, F., and Plaza, S. (2010). Zona pellucida domain proteins remodel the apical compartment for localized cell shape changes. Dev. Cell *18*, 64–76.
28. Menoret, D., Santolini, M., Fernandes, I., Spokony, R., Zanet, J., Gonzalez, I., Latapie, Y., Ferrer, P., Rouault, H., White, K.P., et al. (2013). Genome-wide analyses of Shavenbaby target genes reveals distinct features of enhancer organization. Genome Biol. *14*, R86.
29. Bohère, J., Mancheno-Ferris, A., Al Hayek, S., Zanet, J., Valenti, P., Akino, K., Yamabe, Y., Inagaki, S., Chanut-Delalande, H., Plaza, S., et al. (2018). Shavenbaby and Yorkie mediate Hippo signaling to protect adult stem cells from apoptosis. Nat. Commun. *9*, 5123.
30. Al Hayek, S., Alsawadi, A., Kambris, Z., Boquete, J.-P., Bohère, J., Immarigeon, C., Ronsin, B., Plaza, S., Lemaitre, B., Payre, F., et al. (2021). Steroid-dependent switch of OvoL/Shavenbaby controls self-renewal versus differentiation of intestinal stem cells. EMBO J. *40*, e104347.
31. Castro-Mondragon, J.A., Jaeger, S., Thieffry, D., Thomas-Chollier, M., and van Helden, J. (2017). RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription factor binding motif collections. Nucleic Acids Res. *45*, e119.
32. Nguyen, N.T.T., Contreras-Moreira, B., Castro-Mondragon, J.A., Santana-Garcia, W., Ossio, R., Robles-Espinoza, C.D., Bahin, M., Collombet, S., Vincens, P., Thieffry, D., et al. (2018). RSAT 2018: regulatory sequence analysis tools 20th anniversary. Nucleic Acids Res. *46*, W209–W214.
33. Rouault, H., Mazouni, K., Couturier, L., Hakim, V., and Schweisguth, F. (2010). Genome-wide identification of cis-regulatory motifs and modules underlying gene coregulation using statistics and phylogeny. Proc. Natl. Acad. Sci. USA *107*, 14615–14620.
34. Herz, H.-M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y.-H., Mickey, K., Voets, O., Verrijzer, C.P., and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Genes Dev. *26*, 2604–2620.
35. Huang, C., Yang, F., Zhang, Z., Zhang, J., Cai, G., Li, L., Zheng, Y., Chen, S., Xi, R., and Zhu, B. (2017). Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 Trithorax group protein function in Drosophila. Nat. Commun. *8*, 1649.
36. Kellner, W.A., Ramos, E., Van Bortle, K., Takenaka, N., and Corces, V.G. (2012). Genome-wide phosphoacetylation of histone H3 at Drosophila enhancers and promoters. Genome Res. *22*, 1081–1088.
37. Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.-T., Cubeñas-Potts, C., Hu, M., Lei, E.P., Bosco, G., et al. (2015). Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Mol. Cell *58*, 216–231.
38. Schauer, T., Ghavi-Helm, Y., Sexton, T., Albig, C., Regnard, C., Cavalli, G., Furlong, E.E., and Becker, P.B. (2017). Chromosome topology guides the Drosophila Dosage Compensation Complex for target gene activation. EMBO Rep. *18*, 1854–1868.
39. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA *107*, 21931–21936.
40. Cubenas-Potts, C., Rowley, M.J., Lyu, X., Li, G., Lei, E.P., and Corces, V.G. (2017). Different enhancer classes in Drosophila bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. Nucleic Acids Res. *45*, 1714–1730.
41. Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z.,

Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108–112.
42. Sethi, A., Gu, M., Gumusgoz, E., Chan, L., Yan, K.K., Rozowsky, J., Barozzi, I., Afzal, V., Akiyama, J.A., Plajzer-Frick, I., et al. (2020). Supervised enhancer prediction with epigenetic pattern recognition and targeted validation. Nat. Methods *17*, 807–814.
43. Liu, J., Ali, M., and Zhou, Q. (2020). Establishment and evolution of heterochromatin. Ann. N. Y. Acad. Sci. *1476*, 59–77.
44. Arnold, C.D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M., and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. Science *339*, 1074–1077.
45. Johnson, S.C. (1967). Hierarchical clustering schemes. Psychometrika *32*, 241–254.
46. Ibrahim, M.M., Karabacak, A., Glahs, A., Kolundzic, E., Hirsekorn, A., Carda, A., Tursun, B., Zinzen, R.P., Lacadie, S.A., and Ohler, U. (2018). Determinants of promoter and enhancer transcription directionality in metazoans. Nat. Commun. *9*, 4472.
47. Zabidi, M.A., Arnold, C.D., Schernhuber, K., Pagani, M., Rath, M., Frank, O., and Stark, A. (2015). Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. Nature *518*, 556–559.
48. Liang, J., Lacroix, L., Gamot, A., Cuddapah, S., Queille, S., Lhoumaud, P., Lepetit, P., Martin, P.G.P., Vogelmann, J., Court, F., et al. (2014). Chromatin immunoprecipitation indirect peaks highlight long-range interactions of insulator proteins and Pol II pausing. Mol. Cell *53*, 672–681.
49. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.
50. Aughey, G.N., Estacio Gomez, A., Thomson, J., Yin, H., and Southall, T.D. (2018). CATaDa reveals global remodelling of chromatin accessibility during stem cell differentiation *in vivo*. Elife *7*, e32341.
51. Jiang, H., and Edgar, B.A. (2009). EGFR signaling regulates the proliferation of Drosophila adult midgut progenitors. Development *136*, 483–493.
52. Micchelli, C.A., and Perrimon, N. (2006). Evidence that stem cells reside in the adult Drosophila midgut epithelium. Nature *439*, 475–479.
53. Hung, R.-J., Hu, Y., Kirchner, R., Liu, Y., Xu, C., Comjean, A., Tattikota, S.G., Li, F., Song, W., Ho Sui, S., et al. (2020). A cell atlas of the adult Drosophila midgut. Proc. Natl. Acad. Sci. USA *117*, 1514–1523.
54. Kohlmaier, A., Fassnacht, C., Jin, Y., Reuter, H., Begum, J., Dutta, D., and Edgar, B.A. (2015). Src kinase function controls progenitor cell pools during regeneration and tumor onset in the Drosophila intestine. Oncogene *34*, 2371–2384.
55. Sousa-Victor, P., Ayyaz, A., Hayashi, R., Qi, Y., Madden, D.T., Lunyak, V.V., and Jasper, H. (2017). Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. Cell Rep. *20*, 2527–2537.
56. Boumard, B., and Bardin, A.J. (2021). An amuse-bouche of stem cell regulation: Underlying principles and mechanisms from adult Drosophila intestinal stem cells. Curr. Opin. Cell Biol. *73*, 58–68.
57. Jiang, H., Patel, P.H., Kohlmaier, A., Grenley, M.O., McEwen, D.G., and Edgar, B.A. (2009). Cytokine/Jak/Stat Signaling Mediates Regeneration and Homeostasis in the Drosophila Midgut. Cell *137*, 1343–1355.
58. Ray, S., Rosenberg, M.I., Chanut-Delalande, H., Decaras, A., Schwertner, B., Toubiana, W., Auman, T., Schnellhammer, I., Teuscher, M., Valenti, P., et al. (2019). The mlpt/Ubr3/Svb module comprises an ancient developmental switch for embryonic patterning. Elife *8*.
59. Andrews, S. (2015). FastQC.
60. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinforma. Oxf. Engl. *29*, 15–21.
61. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinforma. Oxf. Engl. *31*, 166–169.
62. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinforma. Oxf. Engl. *26*, 139–140.
63. Su, S., Law, C.W., Ah-Cann, C., Asselin-Labat, M.-L., Blewitt, M.E., and Ritchie, M.E. (2017). Glimma: interactive graphics for gene expression analysis. Bioinforma. Oxf. Engl. *33*, 2050–2052.
64. Sievert, C. Interactive Web-Based Data Visualization with R, Plotly, and Shiny
65. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.
66. Wang, M., Zhao, Y., and Zhang, B. (2015). Efficient Test and Visualization of Multi-Set Intersections. Sci. Rep. *5*, 16923.
67. Wiwie, C., Baumbach, J., and Röttger, R. (2015). Comparing the performance of biomedical clustering methods. Nat. Methods *12*, 1033–1038.
68. Ramey, J.A. (2017). Evaluation of Clustering Algorithms.
69. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2016). Fast Gene Set Enrichment Analysis. Bioinformatics.
70. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma. Oxf. Engl. *25*, 1754–1760.
71. Feng, J., Liu, T., and Zhang, Y. (2011). Using MACS to identify peaks from ChIP-Seq data. Curr. Protoc. Bioinforma. *Chapter 2*. Unit 2.14.
72. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165.
73. Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J.A., Delerce, J., Jaeger, S., Blanchet, C., Vincens, P., Caron, C., et al. (2015). RSAT 2015: Regulatory Sequence Analysis Tools. Nucleic Acids Res. *43*, W50–W56.
74. Thomas-Chollier, M., Sand, O., Turatsinze, J.V., Janky, R., Defrance, M., Vervisch, E., Brohee, S., and van Helden, J. (2008). RSAT: regulatory sequence analysis tools. Nucleic Acids Res. *36*, W119–W127.
75. Köster, J., and Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. Bioinformatics *28*, 2520–2522.
76. Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinf. *11*, 237.
77. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.
78. Morgan, M. (2019). AnnotationHub: Client to Access AnnotationHub Resources.
79. Carlson, M. (2016). org.Dm.eg.db: Genome Wide Annotation for Fly.
80. Chen, H. (2018). VennDiagram: Generate High-Resolution Venn and Euler Plots.
81. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag).
82. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849.
83. Rau, A., Gallopin, M., Celeux, G., and Jaffrézic, F. (2013). Data-based filtering for replicated high-throughput transcriptome sequencing experiments. Bioinformatics *29*, 2146–2152.
84. Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2021). Cluster: Cluster Analysis Basics and Extensions.
85. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008.
86. Lhoumaud, P., Hennion, M., Gamot, A., Cuddapah, S., Queille, S., Liang, J., Micas, G., Morillon, P., Urbach, S., Bouchez, O., et al. (2014). Insulators recruit histone methyltransferase dMes4 to regulate chromatin of flanking genes. EMBO J. *33*, 1599–1613.
87. Heurteau, A., Perrois, C., Depierre, D., Fosseprez, O., Humbert, J., Schaak, S., and Cuvier, O. (2020). Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes. Genome Biol. *21*, 193.
88. Carroll, T.S., Liang, Z., Salama, R., Stark, R., and de Santiago, I. (2014). Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. Front. Genet. *5*, 75.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti GFP	Acris Antibodies	Cat#TP401
Anti Dlg	DSHB	Cat#4F3; RRID: AB_528203
H3K4me1	abcam	ab8895; RRID: AB_306847
H3K4me3	abcam	ab8580; RRID: AB_306649
H3K27ac	abcam	ab4729; RRID: AB_2118291
H3K27me3	abcam	ab6002; RRID: AB_305237
GFP trap	Chromotek	GTA-20; RRID: AB_2631357
AlexaFluor-488	ThermoFischer Scientific	Cat#Z25302
Alexa-555	ThermoFischer Scientific	Cat#Z25205
Chemicals, peptides, and recombinant proteins		
TrypLE (10X)	Gibco	A12177-01
PAF (16%)	Electron Microscopy Sciences	Cat#15710
Critical commercial assays		
For S2: RNeasy Kit	Qiagen	Cat#74104
For Gut: MasterPure RNA purification kit	Biosearch technologies	Cat#MPY03100
Arima-HiC Kit	Arima Genomics	
Library Preparation using KAPA Hyper Prep Kit	Roche molecular system	KK8500
Deposited data		
ChipSeq- S2 Svb ACT and REP	This paper	GSE199514
ChipSeq- S2 Histones	This paper	GSE199512
RNASeq- S2 Ctrl, ACT et REP	This paper	GSE199511
RNASeq- Gut ISC	This paper	GSE199510, GSE220561
RNASeq- Gut EB	This paper	GSE220323
RNASeq- Gut, eEC	This paper	GSE220558
RNASeq- Gut: mEC	This paper	GSE220560
HiC-S2 cells Svb ACT and REP	This paper	GSE221863
Experimental models: Cell lines		
Drosophila melanogaster: S2 lines FS and 1B	Kondo et al.[24]	PMID: 20647469
Experimental models: Organisms/strains		
Drosophila melanogaster: esg-Gal4, UAS-GFP, tubP-Gal80ts	Jiang et al.[57]	PMID: 19563763
Drosophila melanogaster: MyoIA-Gal4, UAS-GFP, tubP-Gal80ts	Jiang et al.[57]	PMID: 19563763
Drosophila melanogaster :Canton S	Bloomington DSC	BL#64349
Drosophila melanogaster: w1118; P{GD9026}v41584	VDRC	# 41584
Drosophila melanogaster : w;; P{UAS-SvbACT::GFP}	Ray et al.[58]	PMID: 30896406
Drosophila melanogaster : w;; P{UAS-Svb3Kmut::GFP}	Ray et al.[58]	PMID: 30896406
Software and algorithms		
FastQC v0.11.5	Andrews.[59]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
STAR v2.5.2b	Dobin et al.[60]	https://github.com/alexdobin/STAR
HTSeq-count v0.6.0	Anders et al.[61]	https://htseq.readthedocs.io/en/release_0.11.1/index.html
EdgeR v3.34.1	Robinson et al.[62]	https://bioconductor.org/packages/release/bioc/html/edgeR.html
Glimma v2.0.0	Su et al., 2017.[63]	https://bioconductor.org/packages/release/bioc/html/Glimma.html
Plotly v4.9.3	Sievert.[64]	https://plotly.com/r/
GenomicRanges v1.42.0	Lawrence et al.[65]	https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
superExactTest	Wang et al.[66]	https://github.com/mw201608/SuperExactTest/
Stats v4.1.1	Wiwie et al.[67]	https://cran.r-project.org/web/packages/STAT/index.html
Clusteval v0.1	Ramey, 2017.[68]	https://github.com/ramhiser/clusteval
Fgsea v1.18.0	Korotkevich et al.[69]	https://bioconductor.org/packages/release/bioc/html/fgsea.html
bwa v0.7.17-r1188	Li, H., & Durbin, R.,[70]	https://github.com/lh3/bwa
MACS2 v2.2.7.1	Feng et al.[71]	https://pypi.org/project/MACS2/
deeptools2 v 3.5.0	Ramírez et al.[72]	https://deeptools.readthedocs.io/en/develop/index.html
RSAT	Medina-Rivera et al.[73] Thomas-Chollier et al.[74]	http://rsat.sb-roscoff.fr
R	R Core Team, 2021	https://www.r-project.org
RStudio	RStudio Team, 2020	https://posit.co/products/open-source/rstudio/
Snakemake	Köster and Rahmann, 2012.[75]	https://snakemake.readthedocs.io/en/v5.4.5/index.html
ChIPpeakAnno v3.24.2	Zhu et al.[76]	https://bioconductor.org/packages/release/bioc/html/ChIPpeakAnno.html
Bedtools v2.29.2	Quinlan et al.[77]	https://bedtools.readthedocs.io
AnnotationHub v2.16.1	Morgan.[78]	https://bioconductor.riken.jp/packages/3.9/bioc/html/AnnotationHub.html
GenomicAlignments	Lawrence et al.[65]	https://bioconductor.org/packages/release/bioc/html/GenomicAlignments.html
TxDb.Dmelanogaster.UCSC.dm6.ensGene	Team BC, Maintainer BP ,2019	https://bioc.ism.ac.jp/packages/3.12/data/annotation/html/TxDb.Dmelanogaster.UCSC.dm6.ensGene.html
Org.Dm.eg.db	Carlson.[79]	https://bioconductor.riken.jp/packages/3.4/data/annotation/html/org.Dm.eg.db.html
VennDiagram	Chen.[80]	https://CRAN.R-project.org/package=VennDiagram.
ggplot2	Wickham.[81]	https://ggplot2-book.org
ComplexHeatmap	Gu et al.[82]	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html
HTSfilter	Rau et al.[83]	https://bioconductor.org/packages/release/bioc/html/HTSFilter.html
Cluster	Maechler et al.[84]	https://cran.r-project.org/web/packages/cluster/index.html
kentUtils		https://github.com/ENCODE-DCC/kentUtils
Samtools	Danecek et al.[85]	http://www.htslib.org

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Cédric Polesello (cedric.polesello@univ-tlse3.fr).

Material availability

This study did not generate new unique reagents.

Data and code availability

- RNA-seq, ChiP-seq and Hi-C data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- Code availability: This paper does not report original code.
- Scripts used on this paper are available on https://github.com/PayreLab/manchenoferris_immarigeon_etal and https://github.com/CuvierLab/Hic.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

S2 cells lines

We used *Drosophila* S2 cells stable cell lines co- expressing the copper-inducible constructs pMT-Svb::GFP and pMT-pri ("1B", or "ACT"), or pMT- Svb::GFP and pMT-pri1-4fs ("FS", or REP"), which encodes a frame-shifted variant of Pri as a negative control.[25]

Drosophila lines

w ; esg-Gal4, UAS-GFP , tubP-Gal80ts.[57]
w ; MyoIA-Gal4, UAS-GFP, tubP-Gal80ts.[57]
Canton S (BL#64349)
w1118; P{GD9026} (VDRC v41584).
*w;; P{UAS-SvbACT::GFP}*and *w;; P{UAS-Svb3Kmut::GFP.*[58]
Only virgin females were used for RNAseq. Flies were collected and aged for 3 days at 22°C, then placed at 29°C for 6 days.

METHOD DETAILS

Cell culture

Drosophila S2 cells were grown in Schneider medium supplemented with 10% fetal calf serum and 1% penicillin/streptomycin (Invitrogen). The expression of pMT plasmids was induced by CuSO4 at the final concentration of 1 mM during 24h.

ChIP experiments

ChIP experiments were performed essentially as previously described[86] with the following specifications: cells were cross-linked with 0.8% of formaldehyde for 10 min at room temperature. Cross-linking was stopped by adding glycine (2M) and cells were washed twice in cold PBS 1X and NaBu 10mM. Cells were resuspended in 500 µL of ChIP permeation buffer (PBS + 0.2% Triton + 10 mM NaBu), incubated for 20 min at room temperature; cells were washed with Lysis Buffer (140 mM NaCl 5M, 15 mM HEPES pH 7.6, 1 mM pH 8 EDTA, 1% Triton, 0.1% Sodium Deoxycholate, 0.5 mM DTT,10nM Sodium Butyrate, 25X protease inhibitor), and resuspended in Lysis Buffer, 1% SDS, 0.5% Sodium N-lauroylsarcosinate. Chromatin was sonicated in Bioruptor (Diagenode) using high-power settings, intervals of 30s burst/pause, for 15 cycles, to obtain fragments of ≈ 250 pb. Sonicated chromatin was diluted 10 times with LB no SDS on protein low bind tubes (Eppendorf), and the chromatin fraction cleared by centrifugation (16000 g, 10°C, 5 min) was then used for immunoprecipitation. Beads were washed and blocked with Lysis Buffer 0.1% SDS, 0.5% Sodium N-lauroylsarcosinate and BSA 0.1 mg/mL. Antibodies were complexed with beads at 4°C overnight in Lysis Buffer 0.1% SDS, 0.5% Sodium N-lauroylsarcosinate, using protein low bind tubes. GFP-TRAP beads (Chromotek), anti H3K27ac (abcam ab4729), anti H3K4me1 (abcam ab8895) anti H3K4me3 (abcam 8580) and anti H3K27me3 (abcam ab6002) antibodies were used for ChIP.

Chromatin was pre-clarified O/N at 4°C before the immunoprecipitation step that was carried out at 4°C for 4h. 10% of each sample was used as input. Chromatin-bead complexes were eluted twice at 70°C during 20 min, first with 10 mM EDTA, 1% SDS, 50 mM Tris-Cl pH 8 and then with TE, 0.67% SDS. Reversing cross-link was performed overnight at 65°C, and DNA purified after phenol-chloroform extraction and ethanol precipitation was resuspended in 1x TE.

RNA extraction, sequencing on S2 cells

RNAseq experiments were made on control, 1B and FS cells (GSE199511). Cells were harvested and RNA extraction was made with RNeasy Kit (Qiagen). Bank of reads and sequencing were done by HiSeq 200 (Illumina) at BGI to obtain 29-34M of reads (single-end 50nt) per replicates.

Drosophila genetics

See complete genotypes and references in the Reagents and Tools table. To purify progenitor cells (either ISC or EB), *esg*^{ts} virgin females were crossed with the following males: control (*Canton-s*), *UAS-svb-RNAi*, *UAS-Svb-ACT*, or *UAS-Svb-REP*^{3Kmut}. To purify enterocytes (either early or mature EC), *MyoIA*^{ts} virgin females were crossed to males of the same genotype. Adult progeny (only virgin female) was collected, aged at room temperature (22°c) for 72h then transferred at 29°C to induce UAS expression. Flies were kept for 6 days at 29°C, changing the food tube every 48h hours to avoid bacterial proliferation and intestinal stress.

Gut dissection and cell dissociation

Batches of >20 virgin females were CO2-anaethesized, put on ice, their midgut were dissected in PBS. Malpighian tubules and hindgut were removed and only the posterior part of the midgut is kept (R3, R4, R5). PBS was replaced by cold 1mM PBS1X-EDTA. Posterior midguts were shrunk into pieces using a razorblade (around 30-60 seconds on a glass plate), then transferred into a fresh tube. Samples were spin 5minutes at 1000rpm 4°C, and supernatant discarded except 50µL. 200uL TrypLE 10X was added (8X final with 0.2mM EDTA) and the guts were incubated at 37°C with intense rocking for 10 minutes (ISC, EB, mEC) or 15 minutes (eEC). Digestions were stopped with 750µL SSM (Serum Supplemented Medium = Schneider's medium +10% Fetal Bovine Serum + 1% Pen/Strep. Note that samples were then treated differently either for FACS (ISC, EB, eEC) or for straining (mEC). In both cases, mechanical trituration was performed by pipetting up and down using low binding, flame-rounded narrow tips (homemade).

Sample preparation for FACS

The cell solution was strained into a cold falcon tube (70µM nylon cell strainer, pre-wet) with 1mL SSM). The Eppendorf tube and the strains were rinsed with 3x1mL PBS-EDTA, so that the samples were 5mL total. The quality of the dissociation and the health of the cells was checked using microscope examination. Samples were spun at 700rcf, 4°C, for >10 minutes. Supernatant was discarded except 100/200µL, to which we added 200µL PBS + 3µL NGS. Propidium Iodide was added to label dead cells, then cells of interest were sorted using FACS (Cell Sorter BD FACSAria Fusion in BSL3) directly into a tube containing 300µL 1X Tissue and Cell Lysis solution (MasterPure RNA purification kit) + 1µg Proteinase K (50µg/µL). Cells were vortexed and placed on ice just after sorting. FACS was used to sort early EC (eEC) from *MyoIA*^{ts} flies, and to sort both ISC and EB from *esg*^{ts} flies (based on a size criterion within GFP+ cells).

Sample preparation for straining mature EC

We noted that ISC and EB cells were easy to dissociate but ECs were not. Longer incubation in TrypLE and harsher trituration allowed dissociation of individual eEC, but this was accompanied with high mortality. We thus reasoned that it would be less stressful for ECs to proceed for mild dissociation of ISC and EB and recover patches of healthy enterocytes by simple filtering. For this procedure, digested samples (10 minutes in TRYPLE 8X) were recovered on a prewet 70µM nylon cell strainer. The strained was rinsed to remove dissociated cells, then turned upside down on a petri dish and rinsed with 5mL PBS-EDTA (5mL) then 2mL SSM. Cell patches were checked under fluorescent macroscope. Samples were transferred into a tube and spun at 700rcf at 4°C for 10 minutes. The supernatant was discarded except 150µL. This pellet was mixed with 150µL 2X Tissue and Cell Lysis solution (MasterPure RNA purification kit) + 1µg Proteinase K (50µg/µL).

RNAseq, sequencing and quality check on intestinal cells (ISC, EB, eEC and mEC)

RNAs were purified from each sample using MasterPure RNA purification kit, following manufacturer's recommendation (including the DNAse treatment). Final RNAs were resuspended in 10µL TE and added 1µL RiboGuard (RNases inhibitor). Construction of RNA banks and sequencing was made on the Montpellier GenomiX Platform. RNA banks were done with the Ovation SoLo RNAseq System kit (Nugen). Bank validation was done with the quantification of the complementary DNA with the Standard Sensitivity NGS kit on Fragment Analyzer and with qPCR (ROCHE Light Cycler 480). The sequencing was done on NovaSeq 6000 (Illumina) with NovaSeq Reagent Kits (100cycles). Single-reads of 100nt were sequenced. ISC= GSE199510, EB=GSE220323, eEC=GSE220558, mEC=GSE220560.

Hi-C

Hi-C data pertaining to this study (GSE221863) were generated using the genome-wide Hi-C kit from Arima Genomic S2 cells. Hi-C analysis at high resolution (1kb) was performed for cells expressing either Svb-REP or Svb-ACT and performed as previously described[87] using pipelines available through Github https://github.com/CuvierLab/Hi-C data for analysis of 3D genomic interactions in Drosophila S2 cells. 2D plots (Figure 5A) were obtained using Juicebox (https://www.aidenlab.org/juicebox/) at 5kb resolution, using the sqrt format coverage (sqrt).

QUANTIFICATION AND STATISTICAL ANALYSIS

RNAseq analysis

The quality of sequencing was measured with FastQC (v0.11.5). Mapping was done with STAR (v2.5.2b, default parameter)[60] on *Drosophila* genome dm6 (r6.13 flybase), reads were counted with HTseq-count (v0.6.0, -t gene -r pos -i gene_symbol)[61] and statistical analysis were performed on with R and RStudio with edgeR (v3.34.1)[62] using the negative binomial generalized log-linear model to the read counts for each gene (gmLRT) and decideTestDGE to determine differential genes after normalization with HTSFilter.[83] Graphs were obtained with Glimma

(v2.0.0)[63] and Plotly (v4.9.3).[64] Identification of direct target genes was performed with GenomicRanges (v1.42.0)[65] and superExactTest.[66] Clustering analysis are described in dedicated Clustering method section. (Figure 3 for S2 cells and Figure 6 for intestinal lineages).

ChIPseq analysis

Data of ChIPseq for H3K4me1 (GSE63518), H3K4me3, H3K27ac (GSE36374), H3K27me3 (GSE93100), H3K9me2(GSE47229), H4K16ac (GSE94115) were used to analyze the chromatin landscape of Svb binding sites. ChIP sequencing of control, FS (REP) and 1B (ACT) S2 cells (single-end 50 nt) were performed with HiSeq 2000 (Illumina) at BGI (GSE199513). Histones ChIP sequencing (paired-end 50 nt) was performed at Montpellier GenomiX platform on NovaSeq 6000 (Illumina) with NovaSeq Reagent Kits (300 cycles). The H3K27ac ChIPseq bank construction was also done by Montpellier Genomix platform with TruSeq® ChIP Sample Preparation (Illumina) (GSE199512).

For all conditions, reads were mapped on dm6 genome (r6.13 flybase) with bwa (v0.7.17-r1188)[70] with default parameters, unique reads were filtered and peak calling was done on replicates merge files with MACS2 (v2.2.7.1)[71] for histone marks of databases (–nomodel –broad –f BAM –g dm –B –q 0.00001) or for histone marks made in this study (–nomodel –f BAMPE –g dm –B –q 1 e-05 –broad) and (-f BAM –B –g dm –to-large –nomodel –q 0.0001 or –q 0.1) for Svb forms. To obtain the merge replicates files, we used samtools merge.[85] Metrics and quality of ChIPseq peaks were calculated with FastQC (v0.11.8)[59] and MultiQC (v1.14).[88] To standardize and make quicker the analysis Snakemake (v 5.4.5)[75] pipelines were done (available on Github: https://github.com/PayreLab/manchenoferris_immarigeon_etal). To compare the set of peaks, ChIPpeakAnno (v3.24.2)[76] packages and BEDtools[77] fisher (v2.29.2) were used. The ratio of Svb-Act and Svb-Rep ChIP-seq were calculated with a homemade python (v 3.6) script: on the common peaks population context the ratio is performed between the summit value of Svb-Act and Svb-Rep obtained with MACS2, on the Svb-ACT only population context as Svb-Rep are not detected the ratio is performed between Svb-Act MACS2 summit and Svb-Rep number of read between the start and end of the Svb-Act only detected peaks.

ChIPseq signal was normalized by taking into account the genome mapping with deeptools2 (v 3.5.0) bamCoverage[72] (–binSize 10 –normalizeUsing RPKM –effectiveGenomeSize 125464728), then we normalized for the noise due to the experiment with deeptools2 (v 3.5.0) bigwigCompare (-b1 experiment file -b2 input file –scaleFactors 1:1 –operation "log2"). Due to the fact that Svb-REP is less expressed than Svb-ACT in S2 cells the normalization for the noise considers the importance of this difference scaling Svb-REP peaks signal by the result of the ratio of logFC for svb in the two conditions: –scaleFactors 1.41:1. Heatmaps of binding enrichment were performed with computeMatrix (-bS 100 -a 3000 –b 3000) and plotHeatmap (deeptools2 (v 3.5.0)[72]). RSAT[73,74] was used to determine the motif enrichment on Svb peaks. Presence analysis of histones in presence of Svb were made with deeptools fisher[72] software. Coverage analysis was done with MultiBamSummary and plotCorrelation of deeptools2 (v 3.5.0)[72] software. All the analysis were performed with R and RStudio with GenomicRanges (v1.42.0), GenomicAlignments.[65] The annotation of the peaks was done with annotationHub (v2.16.1),[78] TxDb.Dmelanogaster.UCSC.dm6.ensGene (v3.12.0) and org.Dm.eg.db.[79] Target genes are defined by genes with a distance < 5kb. Graphical representations were done with the package VennDiagram,[80] Glimma[63] and ggplot2.[81] The population considered for each experiment is indicated both in the figures and in the legends to Figures 1, 2, 4, 5, S2, S3, and S4.

Clustering

Genes similarly expressed in all conditions (difference of reads < 50) were filtered-off. For each condition, the mean RNAseq count for the replicates was calculated, centered and reduced. Then the distance matrices were calculated with get_dist function from cluster package.[84] We used two different algorithms for clustering using stats() package of R.[67] The function hclust() for the hierarchical ascendant clustering was used with the Spearman distance and the Ward.D2 option. The function kmeans() was used for k-means clustering with default options. k-means clustering was repeated 15 times, 5 times with predefined centroids -corresponding to genes of the different clusters defined by hierarchical clustering- and 10 with random centroids. The different k-means clusters obtained by each iteration were compared with cluster_similarity (Clusteval version 0.1).[68] We defined that the clustering is similar if the score equals 1.

For the S2 cells: We used for this analysis the RNAseq data from the S2 cells lines. We obtained after filtration a matrix of 3556 differentially expressed genes (Figure 2).

For the ISC data: We used for this analysis the RNAseq data from the ISC cells obtained by FACS method. We obtained after filtration a matrix of 817 differentially expressed genes (Figure 6). We determined by hierarchical ascendant clustering 5 clusters and use this parameter for the k-means clustering k = 5.

For the EB data: We used for this analysis the RNAseq data from the EB cells obtained by FACS method. We obtained after filtration a matrix of 1461 differentially expressed genes. We determined by hierarchical ascendant clustering 5 clusters and use this parameter for the k-means clustering k = 5 (Figure 6).

For the eEC data: We used for this analysis the RNAseq data from the eEC cells obtained by FACS method. We obtained after filtration a matrix of 1694 differentially expressed genes. We determined by hierarchical ascendant clustering 5 clusters and use this parameter for the k-means clustering k = 5 (Figure 6).

For the mEC data: We obtained after filtration a matrix of 1186 differentially expressed genes. We determined by hierarchical ascendant clustering 4 clusters and use this parameter for the k-means clustering k =4 (Figure 6).

To determine the most robust group of genes we define a score of robustness or r-score. This r-score is calculated for each comparison between clusters of k-means and hierarchical algorithms. First, we define a percent of similarity between k-means clustering and hierarchical clustering such as:*((number of common gene)/((number of gene in kmeans cluster+number of gene in hierarchical cluster)-number of common gene))×100*.

Then for each k-means, if this percentage is greater than or equal to the average obtained with all the clustering groups (in percentage), the comparison obtained a score of 1. For a cluster, all of these scores were added together to obtain the match score. For each cluster, we also calculated the intersection, conservation and dispersion. intersection (i)= minimal number of genes common to all clustering (h-clustering and all k-means). conservation (c)= i/ number of genes in the h-clustering. dispersion (d)= higher number of genes corresponding to one cluster/i. Finally, the r-score = c/d * match score (Table S2).

Graphical representations were made with ComplexHeatmap.[82]

Gene set enrichment analysis (GSEA)

Gene set analysis was performed with fgsea software (v1.18.0).[69] Gene lists are those obtained after differential expression analysis (ACTvsCTL, REPvsCTL and ACTvsREP) and were sorted in descending logFC order. Enrichment was tested on gene lists associated with Svb-ACT act or Svb-REP peaks (Figure S2). The population considered for each experiment are shown in the legends of Figure S2.

Hi-C analysis

Aggregation analysis of Figure 5 was performed as previously in 1D/2D plots[48,49] to compare genomic contacts in cells expressing either Svb-REP or Svb-ACT. Hi-C data were aggregated onto pairs of enhancers identified by STARR-seq[44,47] and distant target transcription start sites. Enhancer-promoter pairs were then ranked depending on net variations of 3D contacts between Svb-REP and Svb-ACT and the ranked quantiles of pairs were assessed for enrichment depending on presence or absence of Svb binding or not using a Fisher exact test. 3D contacts were calculated by assessing the levels of contacts of any enhancer with any TSS that is in the proper range of distances, here between 40 and 200 kb.

Aggregation analysis of Figure S4 was based on differential quantification of H3K27ac REP-ACT (log2(Quantif_H3K27ac_REP)-log2(Quantif_H3K27ac_ACT)) for an HiC resolution of 1kb. Data were sorted and split in ten groups. Top 10 % correspond to group 1 with the most differential in H3K27ac associated with Svb-ACT (Svb-ACT>Svb-REP). Bottom 10 % correspond to group 10 with the least differential in H3K27ac (Svb-ACT=Svb-REP). Pair formation was performed between the active TSS and selected regions associated with group 1 or 10 with a minimum distance of 40kb and a maximum distance of 200kb. At first, the possible couples of Svb sites and active TSS sites, separated by a minimal distance of 40kb and a maximal distance of 200kb, are defined. Then, the sub-matrices of these couples are extracted from the Hi-C matrix. Finally, the sub-matrices are aggregated to obtain an APA plot. In other words, the APA plots represent the average signal of the interactions between Svb sites and active TSSs. To make the aggregated plots on R, the HicAggR package developed by the Cuvier Team (https://github.com/CuvierLab/HicAggR) was used. The matrices Hi-C are extracted with a resolution of 1kb. This resolution is defined by binning the genome into regular intervals or bins of a given length (1kb in this instance). Each sub-matrix is quantilized between 1 and 500, i.e. vectorized and sorted, and the values are scaled between 1 and 500. To do this, the percentiles of the values are calculated and then these percentiles are multiplied by 500, so that the maximum value of the sub-matrix obtains 500 and the minimum value 1. Unassigned values are omitted. Quantization serves as a metric for estimating the rank of each contact pixel relative to its sub-matrix. The signal of the pixel center (3x3), which represent the interaction between the two sites of interest, has been quantified and presented as Boxplot (Figure 5F). The p-value were calculated using a Wilcoxon test. The population considered for each experiment is shown in both Figures 5 and S4, with the corresponding legends.

CATaDa analysis

CATaDa chromatin accessibility data from ISCs, EBs, and ECs were obtained from[50] (Figure 6). Per-GATC fragment counts were RPM normalized and visualized in Integrative Genomics Viewer (IGV). GFF signal files were converted to bigwig format using kentUtils bedGraphToBigWig (https://github.com/ENCODE-DCC/kentUtils) and profiles of RPM chromatin accessibility scores across transcriptional start sites were plotted using deeptools computeMatrix and plotProfile functions. Data of chromatin accessibility at individual promoter regions for global comparison was extracted using 2kb regions centered on the TSS, means were compared using Kruskal-wallis test with Dunn's multiple comparison. The populations considered for each cell type are shown in Figure 6F.