



**HAL**  
open science

## **Crosstalk between chromatin and the transcription factor Shavenbaby defines transcriptional output along the *Drosophila* intestinal stem cell lineage**

Alexandra Mancheno-Ferris, Clément Immarigeon, Alexia Rivero, David Depierre, Nicolas Chanard, Olivier Fosseprez, Gabriel Aughey, Priscilla Lhoumaud, Julien Anglade, Tony Southall, et al.

### ► **To cite this version:**

Alexandra Mancheno-Ferris, Clément Immarigeon, Alexia Rivero, David Depierre, Nicolas Chanard, et al.. Crosstalk between chromatin and the transcription factor Shavenbaby defines transcriptional output along the *Drosophila* intestinal stem cell lineage. 2023. hal-04289018

**HAL Id: hal-04289018**

**<https://cnrs.hal.science/hal-04289018v1>**

Preprint submitted on 16 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 1 **Crosstalk between chromatin and the transcription factor** 2 **Shavenbaby defines transcriptional output along the** 3 ***Drosophila* intestinal stem cell lineage.**

4  
5

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

10

11 <sup>1</sup> Molecular, Cellular and Developmental biology department (MCD), Centre de Biologie Intégrative (CBI),  
12 Université de Toulouse, CNRS, UPS, F-31062 Toulouse, France

13 <sup>2</sup> Control of cell shape remodeling team, CBI, CNRS, UPS, F-31062 Toulouse, France.

14 <sup>3</sup> Chromatin Dynamics and Cell Proliferation team, CBI, CNRS, UPS, F-31062 Toulouse, France.

15 <sup>4</sup> Imperial College London, Sir Ernst Chain Building, South Kensington Campus, London, SW7 2AZ

16 <sup>5</sup> Institut Jacques Monod, Université Paris Cité/CNRS, 15 rue Hélène Brion, 75205, Paris Cedex 13, France

17 <sup>6</sup> Laboratoire de Recherche en Sciences Végétales (LSRV), CNRS, UPS, 24 chemin de Borde Rouge, Auzeville,  
18 31326, Castanet-Tolosan, France.

19 <sup>7</sup> these authors contributed equally to this work

20 <sup>8</sup>Lead contact

21

22 \* correspondence: [olivier.cuvier@univ-tlse3.fr](mailto:olivier.cuvier@univ-tlse3.fr) (O.C.), [cedric.polesello@univ-tlse3.fr](mailto:cedric.polesello@univ-tlse3.fr) (C.P.)

23

24

## 25 **Summary**

26 The transcription factor Shavenbaby (Svb), the only member of the OvoL family in *Drosophila*,  
27 controls intestinal stem cell differentiation. Post-translational modification of Svb produces two  
28 protein isoforms, Svb-ACT and Svb-REP, which promote intestinal stem cell renewal or  
29 differentiation, respectively. Using engineered cell lines, we express either isoform to define  
30 their mode of action, and develop an unbiased method to identify Svb target genes in intestinal  
31 cells. Within a given cell type, Svb-ACT and Svb-REP antagonistically regulate the expression  
32 of a set of target genes, binding specific enhancers whose accessibility is constrained by.  
33 During intestinal differentiation, the set of target genes progressively changes, together with  
34 chromatin accessibility. Moreover, Svb-REP binding stabilizes three-dimensional enhancer-  
35 promoter loops, while influencing the local chromatin landscape to repress target genes. We  
36 propose that SvbACT-to-REP switch promotes enterocyte differentiation of intestinal stem  
37 cells through direct gene regulation and chromatin remodeling.

38

## 39 **Keywords**

40 OvoL, ovo/shavenbaby, *Drosophila*, intestinal stem cells, cell differentiation, transcription  
41 factors, enhancers, histone marks, chromatin organization, gene regulation

## 42 Introduction

43 Transcription factors (TFs) are primary determinants of cell phenotypes and behaviors,  
44 regulating gene expression via binding to DNA sequences located within enhancers and  
45 promoters cis-regulatory elements (Takahashi and Yamanaka, 2006) (Bulyk, 2003; Rohs *et*  
46 *al.*, 2009). TF binding motifs being generally short (4-15 base pairs) and degenerate, they are  
47 very abundant in the genome. The accessibility of TF binding sites is influenced by many  
48 parameters and dimensions, including biophysical constraints, chromatin state and  
49 organization into domains, and notably involving the presence/absence of epigenetic co-  
50 regulators. As a consequence, only a limited subset of TF binding sites are bound in a given  
51 cell type (Buenrostro *et al.*, 2018; Cuvier and Fierz, 2017)(Guertin and Lis, 2010; Slattery *et*  
52 *al.*, 2014; Wang *et al.*, 2012), and substantial variations in binding events are often observed  
53 across tissues (Arvey *et al.*, 2012; Spitz and Furlong, 2012). These cell-specific changes in  
54 binding site occupancy are often thought to have strong impact in transcriptional outputs (Sen  
55 *et al.*, 2019), though distinct tissues, developmental stages and/or transcriptional factors may  
56 adopt distinct strategies. The underlying mechanisms and functional consequences currently  
57 remain to be fully elucidated.

58

59 Transcription factors of the Ovo-like (OvoL) family are specific to metazoans (Kumar *et al.*,  
60 2012) and display key functions during animal development. The founding member called  
61 Ovo/Shavenbaby (Svb) was initially identified in flies (Mevlinio *et al.*, 1995) for its dual role  
62 in germline and epidermal development. Mammalian species have evolved three paralogs  
63 referred to as OvoL1-3 that display partly redundant functions (Teng *et al.*, 2007) and also  
64 contribute to germline, skin development and diseases (Tsuji *et al.*, 2017). In addition, the  
65 deregulation of OvoL expression is associated to various cancers of epithelial origin (Roca *et*  
66 *al.*, 2013; Wang *et al.*, 2017; Watanabe *et al.*, 2014) and it has been proposed that OvoL  
67 factors act as the guardians of epithelial integrity by preventing epithelial-mesenchymal  
68 transition (Hong *et al.*, 2015; Wang *et al.*, 2017). While OvoL TFs share a highly-conserved  
69 zinc finger DNA binding domain that recognizes the core sequence CNGTT (Lee and Garfinkel,  
70 2000; Lu and Oliver, 2001), they differ in their N-terminal regions and can exhibit distinct  
71 transcriptional activities (Watanabe *et al.*, 2014).

72 *Drosophila* encodes a unique Ovo/Svb gene, providing an attractive paradigm to dissect OvoL  
73 functions and investigate the mechanisms underlying tissue-specific activities. The somatic  
74 factor, Shavenbaby (Svb), is known for governing differentiation of epidermal cells in the  
75 embryo and adult derivatives (Chanut-Delalande *et al.*, 2014; Payre *et al.*, 1999). The Svb  
76 protein is translated as a large transcriptional repressor, referred to as Svb-REP. Remarkably,

77 under the action of Polished-rice (Pri, a.k.a. Tarsal-less or Mille patte) peptides encoded by  
78 small open reading frames, the ubiquitin ligase Ubr3 binds to Svb-REP, and triggers  
79 proteasome-dependent processing resulting in the production of a shorter transcriptional  
80 activator, Svb-ACT (Kondo *et al.*, 2010; Zanet *et al.*, 2015). In the embryonic epidermis, Svb-  
81 ACT triggers the expression of a battery of effector genes that are collectively responsible for  
82 the remodeling of epithelial cells (Chanut-Delalande *et al.*, 2006; Fernandes *et al.*, 2010;  
83 Menoret *et al.*, 2013). We have recently reported that Svb is critically required for the  
84 maintenance of adult epithelial stem cells, which ensure the homeostasis of the renal (Bohère  
85 *et al.*, 2018) and intestinal systems (Al Hayek *et al.*, 2021). In addition, the Svb transcriptional  
86 switch operated by Pri peptides is also at work to control the behavior of adult stem cells :  
87 SvbACT promotes intestinal stem cell self-renewal and proliferation, while in contrast SvbREP  
88 dictates differentiation into enterocytes (Al Hayek *et al.*, 2021).

89 Here we use genomic and bioinformatic approaches to study the molecular action of the two  
90 Svb isoforms on gene expression, across the intestinal stem cell lineage. Using engineered  
91 cells in culture, we define the molecular mode of action of Svb-ACT and REP. We show that,  
92 constrained by chromatin landscape, both Svb-ACT and Svb-REP bind to a same array of  
93 open enhancers, mediating antagonistic effects on the expression of a common set of target  
94 genes. We develop unsupervised analyses of differential gene expression in response to Svb-  
95 ACT and Svb-REP allowing the identification of Svb target genes in adult intestinal stem cells  
96 (ISCs) and their progeny enteroblasts and enterocytes. Strikingly, Svb target genes are  
97 different in each cell type of the lineage. Interestingly, this change in the target genes repertoire  
98 during differentiation is accompanied by chromatin accessibility remodeling, and concomitant  
99 accumulation of SvbREP. We further show in cultured cells that Svb-REP participates in  
100 remodeling local chromatin landscape, playing an active role in transcriptional repression,  
101 which can occur through distal enhancers. This work sheds light on the possible crosstalk  
102 between the chromatin landscape co-evolving with the proteasome mediated maturation of a  
103 transcription factor, which may serve to optimize transcriptional output to differentiation.

104

## 105 **Results**

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

107 To map the genome-wide set of sites bound by Svb-ACT and Svb-REP, we made use of stable  
108 cell lines derived from cultured S2 *Drosophila* cells, since they do not express endogenous  
109 Svb and have been engineered to express one or the other Svb form (Kondo *et al.*, 2010;  
110 Zanet *et al.*, 2015). We performed chromatin immunoprecipitation coupled to next-generation  
111 sequencing (ChIP-seq) in Svb-ACT::GFP or Svb-REP::GFP cells (Figure 1A). ChIP-seq

112 revealed 6,939 regions highly enriched for Svb-ACT binding (Figure 1B). Consistent with a  
113 binding mediated by the zinc fingers domain, these regions feature DNA motifs matching  
114 CcGTT, the Ovo/Svb binding matrixes ((Castro-Mondragon *et al.*, 2017; Menoret *et al.*, 2013;  
115 Nguyen *et al.*, 2018) and methods). The most enriched motif matches a ACCGTTA sequence  
116 (hereafter referred to as F7-BS, Figure 1D). F7-BS - defined by a machine learning approach  
117 combining statistical analysis to phylogenetic information (Rouault *et al.*, 2010) is the best  
118 predictor of functional Svb binding sites *in vivo* (Menoret *et al.*, 2013).

119 Although peak calling detected a markedly smaller number of regions for Svb-REP (1,325),  
120 they extensively overlap with Svb-ACT bound regions and there is clear enrichment for Svb-  
121 REP across all Svb-ACT peaks (Figure 1B, Figure S1A-F)). We thus conclude that both  
122 isoforms bind common sites genome-wide.

123 To characterize the chromatin landscape of Svb-bound regions, we interrogated whether these  
124 regions display recognizable patterns of histone marks using the large set of data available in  
125 *Drosophila* cells (Herz *et al.*, 2012; Huang *et al.*, 2017; Kellner *et al.*, 2012; Li *et al.*, 2015;  
126 Schauer *et al.*, 2017). Considering all Svb-ACT peaks (allowing peak detection at higher  
127 sensitivity), a strong enrichment in histone3 acetylated on K27 (H3K27ac) across sites bound  
128 by Svb was found (Figure 1B-C). Notably, H3K27ac enrichment directly correlated with Svb-  
129 ChIP signal (see the intensity “gradients” in Figure 1B). H3K27ac marks open chromatin, in  
130 particular active enhancers (Creyghton *et al.*, 2010; Cubenas-Potts *et al.*, 2017; Li *et al.*, 2015)  
131 which are also enriched in histone 3 methylated on K4 (Heintzman *et al.*, 2009; Sethi *et al.*,  
132 2020). Consistently, Svb-bound sites were consistently enriched in H3K4me1 and H3K4me3  
133 that accumulated on both sides of the peak (Figure 1B-C). Accordingly, H3K27me3 marks of  
134 repressive chromatin (Liu *et al.*, 2020) were depleted (Figure 1B-C). Therefore, these data  
135 suggest that Svb binds active enhancers. We challenged this hypothesis by analyzing Svb  
136 binding to all enhancers defined by STARR-seq ((Arnold *et al.*, 2013) Figure 1E). This revealed  
137 that enhancers associated with high Svb signal display stronger enhancer activity (brown) and  
138 higher H3K27ac (blue), compared to Svb unbound enhancers (Figure 1E), confirming that Svb  
139 binds active enhancers.

140 To analyze the respective influence of Svb isoforms on gene expression, we next performed  
141 RNA-sequencing in control, Svb-ACT and Svb-REP cells. When compared to controls, both  
142 Svb-ACT and Svb-REP significantly changed gene expression (Figure S1G). *In vivo* assays in  
143 embryonic epidermal cells have shown the antagonistic activity of Svb-ACT and Svb-REP on  
144 individual direct target genes (Chanut-Delalande *et al.*, 2006; Fernandes *et al.*, 2010; Kondo  
145 *et al.*, 2010), leading to the upregulation or downregulation of those, respectively. We expected  
146 direct target genes of Svb in S2 cells to display a similar behavior. Therefore, we selected all  
147 differentially expressed genes (DEG, p-value<0.05) between Svb-ACT and Svb-REP, and

148 focused on those responding to both isoforms ( $\log_2(\text{FC}) > 0$  in Svb-ACT and  $\log_2(\text{FC}) < 0$  in Svb-  
149 REP conditions, [Figure S1G](#)). This supervised approach defined a group of 285 putative target  
150 genes. Consistent with the prediction that these genes contain direct targets, 86% of those are  
151 associated with Svb binding peaks ([Figure 1J](#)). To further tackle the influence of Svb REP and  
152 ACT on gene expression, we develop an unsupervised approach based on hierarchical  
153 clustering and k-means to identify Svb direct target genes solely based on RNAseq data.  
154 Hierarchical clustering ([Johnson, 1967](#)) defines 5 groups of DE genes ([Figure 1F](#)), one being  
155 very similar to the supervised set (278/285 genes in common, [Figure 1G,I](#)). We used 15  
156 iterations of k-means clustering to challenge the robustness of each hierarchical cluster gene  
157 sets, and developed the “robustness score” method to estimate the proportion of genes that  
158 cluster together throughout K-means iterations (r-score, [Figure 1H](#), see methods). The cluster  
159 comprising putative Svb targets got the highest r-score and 195 genes were common to all  
160 iterations and methods ([Figure 1I](#)), 86% of these common genes being associated to Svb  
161 binding peaks ([Figure 1J](#)). This set of 195 genes is referred hereafter as S2 Svb target genes.  
162 Taken together, these data show that Svb-ACT and Svb-REP display antagonistic  
163 transcriptional activities, acting on a common set of approximately 200 genes validated through  
164 an unsupervised approach (see [Figure 1K](#) for a representative target gene).

165

### 166 [Chromatin landscape defines Svb binding sites](#)

167 Having defined the set of Svb target genes in S2, we next asked whether it reflected the  
168 regulation of genes by Svb in embryonic epidermal cells ([Menoret et al., 2013](#)), focusing on  
169 direct target genes (*i.e.*, being differentially expressed and displaying a binding peak in  
170 embryo). No significant overlap between the two sets of genes was observed ([Figure 2A](#)). This  
171 suggested that Svb could behave differently depending on the cell type. Indeed, this  
172 divergence is also observed when comparing Svb-ChIP peaks between embryonic and S2  
173 cells ([Figure 2B](#)).

174 To get a deeper look at the chromatin context of Svb binding, we focused on evolutionarily  
175 conserved F7-BS motifs that often mediate transcription of Svb direct targets *in vivo* ([Menoret  
176 et al., 2013](#)) and are highly enriched within Svb peaks in S2 cells ([Figure 1D](#)). There are 2354  
177 conserved F7-BS in the *Drosophila* genome, 26% of those being bound by Svb in S2 cells  
178 ([Figure 2Ba+b](#)). Chromatin profiles of histone post-translational modifications were generated  
179 to inspect the chromatin landscape around F7-BS. This analysis showed that Svb-bound F7-  
180 BS (a and b) displayed significantly higher H3K27ac, H3K4me1 and H3K4me3 and lower  
181 H3K27me3, as compared to non-bound F7-BS ([Figure 2C, c and d](#)). Of note, the binding status  
182 of F7-BS in embryo has no effect on the chromatin landscape in S2 cells (a vs b and c vs d):

183 whereas F7-BS bound only in S2 cells also displayed the features of active enhancers (b),  
184 sites bound by Svb only in the embryonic epidermis exhibited closed chromatin configuration  
185 in S2 cells (d). A representative embryonic Svb target gene is shown (Figure 2D, see  
186 repressive chromatin landscape and absence of Svb-ChIP peak in S2). We challenge this  
187 observation using available ATAC-seq data from S2 cells, and confirm that chromatin is open  
188 and accessible at Svb-bound F7-BS (Figure 2E a and b) whereas Svb-unbound sites are in a  
189 closed conformation (Figure 2E c and d), again independently of Svb binding in embryo.

190 Together, these results indicate that the Svb transcription factor regulates different sets of  
191 target genes in S2 cells and embryonic epidermal cells. Specific patterns of histone marks  
192 highlight how such cell-type specific genes may be gated through accessibility of distinct pool  
193 of sites that are accessible to Svb binding, hence defining a subset of cell-type specific target  
194 genes regulated by Svb.

195

#### 196 Identification of Svb target genes across the intestinal stem cell lineage

197 While the absence of endogenous expression of Svb in cultured cells was helpful to address  
198 the binding and transcriptional effect of each isoform, the extent to which this experimental  
199 model was representative of the choice of target genes *in vivo* remained to be established. We  
200 next investigated its activity in adult intestinal stem cells (ISCs) lineage leading to enterocyte  
201 (EC) differentiation (Figure 3A). We previously described that Svb-ACT acts in ISC to promote  
202 their maintenance and proliferation, while Svb-REP triggers and accompanies differentiation  
203 into ECs (Al Hayek *et al.*, 2021). Svb-REP accumulates during differentiation (Figure 3A), and  
204 is required in mature EC to maintain their epithelial properties (Al Hayek *et al.*, 2021). Thus,  
205 the switch from Svb-ACT to REP isoforms has a drastic effect on ISC and EC behavior, but  
206 the target genes mediating this function remained unknown.

207 To identify genes regulated by Svb in each cell type from the lineage, we profiled the  
208 transcriptome of four cell types from posterior midguts: ISCs, enteroblasts (EBs), early EC  
209 (eEC) and mature EC (mEC) (see methods). These cells were purified from dissected guts of  
210 control animals, upon RNAi-mediated *svb* depletion, and expression of Svb-ACT or Svb-  
211 REP<sup>3Kmut</sup> (a engineered form insensitive to Pri and thus not convertible into Svb-ACT (Zanet *et*  
212 *al.*, 2015). Note that manipulation of Svb function was specifically targeted either to progenitor  
213 cells (ISC and EB) or to ECs (eEC and mEC), using specific drivers allowing spatial and  
214 temporal control (Figure 3B, (Jiang and Edgar, 2009; Micchelli and Perrimon, 2006)).

215 We first verified that our RNA-seq datasets are consistent with available transcriptomic data  
216 and indeed represent distinct cell types along the ISC to EC differentiation route (ISC, EB,  
217 early EC and mature EC, Figure S2A). Then, building on the unbiased method developed using

218 cultured cells (Figure 1), we identify the set of genes whose expression is likely directly  
219 regulated by Svb. In each cell type, we identify one robust cluster of 132-265 genes  
220 antagonistically regulated by Svb as expected for direct target genes (Figure 3C to G, Figure  
221 S2B). The fact that their expression was also downregulated upon *svb* knockdown in  
222 progenitors confirms that these genes are normally activated by Svb-ACT in ISCs (Figure 3C  
223 and G, Figure S2B). On the contrary, Svb target genes in mature ECs tend to be upregulated  
224 upon *svb* knockdown, confirming that Svb acts mostly as a repressor in these cells (Figure  
225 S2B/Table S1).

226 In agreement with the function of Svb in ISC (Al Hayek *et al.*, 2021), we identify genes  
227 controlling cell cycle, proliferation and survival of stem cells (*string*, *cycE*, *cycB*, *piwi*...) as well  
228 as signaling pathways known to control stemness (*egfr*, *Di*) (Figure 3E) (reviewed in (Boumard  
229 and Bardin, 2021)(Kohlmaier *et al.*, 2015; Sousa-Victor *et al.*, 2017)). This discovery reinforces  
230 the importance of the switch between Svb-ACT and REP isoforms in ISC progeny: the  
231 downregulation of such regulators by Svb-REP might favor cell cycle exit and commitment  
232 towards differentiation.

233 Interestingly, the set of Svb target genes is markedly different in progenitors versus  
234 differentiated cells from the lineage (Figure 3F, Figure S2D), reminiscent of the situation  
235 observed between embryonic and cultured S2 cells (Figure 2). Focusing on the 195 ISC target  
236 genes, we show that their expression and regulation evolves during differentiation: 73 remain  
237 Svb targets in EB, 25 are also targets in early EC, and only 5 are targets in mature EC. Thus,  
238 the change in Svb target gene set occurs gradually (Figure 3F, Figure S2D). Accordingly, ISC  
239 target genes' expression tends to decrease during differentiation, as well as the ability of Svb-  
240 ACT to activate their transcription (Figure 3G, Figure S2C).

241 It was described that ISC chromatin accessibility evolves as they differentiate into EB and EC  
242 (Aughey *et al.*, 2018). We analyzed chromatin accessibility at Svb target genes' loci in ISC, EB  
243 and mEC using CATaDa (Chromatin Accessibility Targeted DamID), a technique based on the  
244 Dam methyltransferase expression specifically in one of these cell types (Aughey *et al.*, 2018).  
245 We observed that the chromatin at ISC and EB target genes is most accessible in ISC, still  
246 open in EB and closed in enterocytes (e.g. *stg* and *egfr*, Figure 3H,I). On the contrary, EC  
247 target genes tend to be more accessible in differentiated EC (e.g. *Cyp9b2* Figure 3J). We  
248 tested whether this trend was global by averaging CATaDa signal from entire gene sets  
249 corresponding to cell-specific target genes (Figure 3H',I',L'). This confirms that, as cells  
250 differentiate, the TSSs of target genes in ISC and EB close down while those of mEC become  
251 accessible.



252 All together these results show that Svb target gene sets evolve gradually during intestinal cell  
253 differentiation, concomitant with changes in chromatin accessibility. This suggests that, as  
254 observed in cultured cells, the chromatin state governs Svb-accessible regions thereby  
255 constraining its transcriptional output. These results prompted us to test whether Svb isoforms  
256 whose ratio changes during differentiation (as Svb-REP accumulates), could influence  
257 chromatin organization directly, surrounding its binding sites.

258

#### 259 [SvbREP modifies local chromatin state](#)

260 To address the putative effect of Svb on chromatin landscape, we used our cultured cell system  
261 and profiled the changes in H3K27 acetylation, a key mark of active enhancers ([Creyghton et](#)  
262 [al., 2010](#)), in control, Svb-ACT and Svb-REP conditions. We focused on Svb-bound regions  
263 and confirmed H3K27ac enrichment at Svb peaks ([Figure 4A](#)), as previously observed with  
264 public datasets ([Figure 1](#)). Interestingly, we observed a decrease in H3K27ac in presence of  
265 Svb-REP ([Figure 4A, B](#)). The amplitude of this effect depends on the intensity of Svb peaks,  
266 as revealed by the REP/control ratio ([Figure 4C](#)), suggesting that the Svb-REP directly  
267 impinges H3K27 acetylation.

268 We deployed Hi-C experiments to ask whether these local changes may affect 3D chromatin  
269 organization. Inspecting long-range contacts at high resolution (1kb) using 2D plot suggested  
270 that specific 3D contacts may actually form in the Svb-REP condition, which could reflect the  
271 formation or stabilization of specific loops ([Figure 4D](#)). We wondered whether these contacts  
272 might occur between Svb bound enhancers and target promoters. To readily test this at  
273 genome-wide scales, we used aggregation plots as previously developed by us and others  
274 ([Liang et al., 2014](#); [Rao et al., 2014](#)). Aggregate Hi-C plots depending on presence or absence  
275 of Svb suggest that, Svb-REP stabilizes such enhancer-promoter interactions ([Figure 4E](#)). To  
276 further test the specificity of these 3D interactions, we considered all enhancer-promoter  
277 interaction changes when comparing Svb-ACT and REP conditions. These couples were  
278 ranked following changes in 3D contacts. For each decile from the list, we estimated the  
279 enrichment for Svb-bound or Svb-unbound enhancers ([Figure 4F](#) and methods). This revealed  
280 that the Svb-bound enhancers tend to be enriched among the enhancers on which Svb had  
281 the strongest influence, in contrast to unbound enhancers ([Figure 4F](#)). Altogether, our data  
282 thus show that Svb-REP impinges both on local chromatin landscape of enhancers and 3D  
283 contacts with promoters. Svb-REP has an instructive role on chromatin and participates in its  
284 remodeling.

## 285 Discussion

286

287 Here, we describe how Svb regulates the expression of its target genes. First, Svb-ACT and  
288 REP bind to the same genomic enhancer regions to antagonistically control a cohort of target  
289 genes. Second, this cohort is constrained by the chromatin environment of the cell, leading to  
290 almost completely distinct target gene sets in different cell types. Last, we show that Svb-REP  
291 isoform modifies local chromatin landscape at bound enhancers, and 3D contacts with target  
292 promoters. This work reveals how a peptide-induced proteasomal processing of a transcription  
293 factor leads to specific cell fate commitment: the crosstalk between chromatin organization  
294 and Svb activity defines cell specific transcriptional output.

295

296 In the intestine, we show that in stem cells Svb target genes (*egfr*, *stg*...) are activated to  
297 maintain stemness and proliferation. The progressive accumulation of Svb-REP occurring in  
298 EB will first reduce these genes' expression, consistent with cell cycle arrest and differentiation.  
299 Notably, EGFR pathway was previously shown to directly control *svb* transcription ([Al Hayek  
300 et al., 2021](#)), forming a feedback loop that could reinforce the asymmetry between ISC  
301 daughter cells. In ISC, high Svb-ACT would maintain high EGFR, and high DI, while cells  
302 starting to accumulate Svb-REP would decrease EGFR and DI, triggering differentiation.

303 During differentiation, we observe that the chromatin at Svb ISC target genes' expression goes  
304 down ([Figure 3G](#)), while chromatin around TSSs progressively closes down ([Figure 3H](#)). As  
305 we established that Svb-REP binding locally modifies H3K27ac, Svb-REP likely participates in  
306 de-acetylation of H3K27 in the intestinal lineage. This could mediate direct gene repression  
307 through inhibition of enhancer activity. In addition, this histone modification could foreshadow  
308 the chromatin compaction that occurs during differentiation, which mediates long-term  
309 silencing of "stemness" genes. It would be interesting to study the mechanisms by which Svb  
310 modulate chromatin landscape and to which extent this applies to OvoL proteins and other  
311 transcription factors.

312

313

314 **Acknowledgements**

315 We thank all members of the Payre lab for critical reading of the manuscript. We thank Marion  
316 Aguirrebengoa (BigA platform) for her bioinformatic expertise, Emmanuelle Näser from the  
317 Cell sorting service of the TRI genotoul platform, Hugues Parrinello from the MGX sequencing  
318 platform. This work was supported by ANR (Chrononet, to FP's team; Helico, OC's team),  
319 Fondation pour la Recherche Médicale (FRM, grant DEQ20170336739 to FP's team and  
320 DEQ20160334940 to OC's team) and a FRBT grant. AMF, AR and DD were supported by  
321 fellowships from Ligue contre le Cancer. CI was supported by FRM and Université Paul  
322 Sabatier.

323

324 **Authors' contributions**

325 AMF performed ChIP-seq experiments for histone marks in S2 cells, with the help of JA, and  
326 most computational analyses with the help of AR (unsupervised clustering), DD (Histones  
327 ChIP-seq) and CI (intestinal cells RNA-seq analyses). CI and CP dissected guts of different  
328 genotypes; CI did FACS purification and prepared samples for intestinal cells RNA-seq. PL  
329 and SP prepared S2 cell samples for RNA-seq and Svb::GFP ChIP-seq. GA and TS performed  
330 CATaDa analysis. OF, NC and OC performed and analyzed HiC. OC, SP, FP and CP designed  
331 the project. FP, OC and CP supervised the project. CI, FP, OC and CP wrote the paper with  
332 contribution from all authors.

333

334 **Competing interests**

335 The authors declare that they have no competing interests.

336

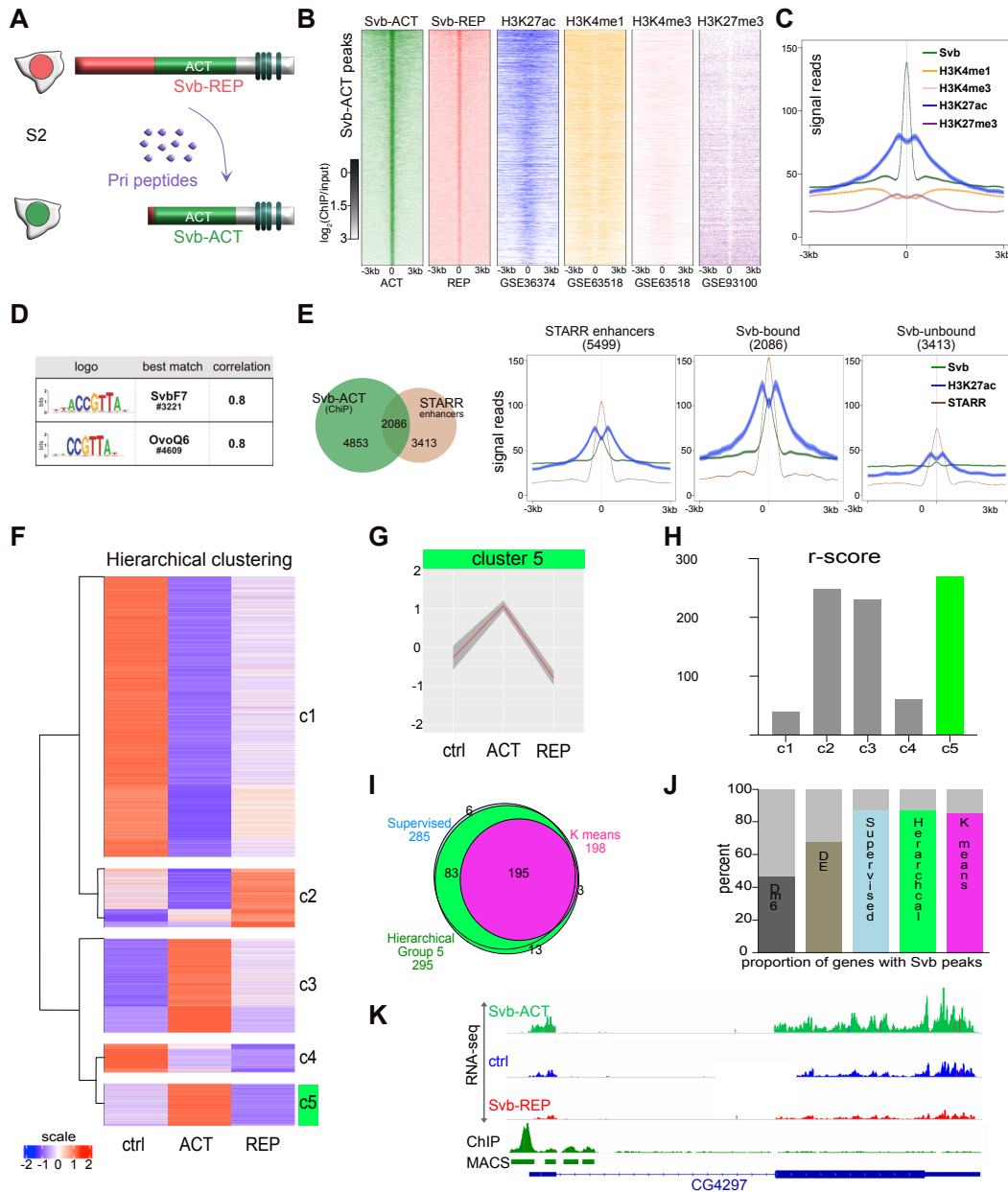
## 337 Bibliography

- 338 Al Hayek, S., Alsawadi, A., Kambris, Z., Boquete, J.-P., Bohère, J., Immarigeon, C., Ronsin,  
339 B., Plaza, S., Lemaitre, B., Payre, F. and Osman, D. 2021. Steroid-dependent switch of  
340 OvoL/Shavenbaby controls self-renewal versus differentiation of intestinal stem cells. *EMBO*  
341 *J.*, **40**: e104347.
- 342 Arnold, C.D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M. and Stark, A. 2013. Genome-wide  
343 quantitative enhancer activity maps identified by STARR-seq. *Science*, **339**: 1074–7.
- 344 Arvey, A., Agius, P., Noble, W.S. and Leslie, C. 2012. Sequence and chromatin determinants  
345 of cell-type-specific transcription factor binding. *Genome Res*, **22**: 1723–34.
- 346 Aughey, G.N., Estacio Gomez, A., Thomson, J., Yin, H. and Southall, T.D. 2018. CATaDa  
347 reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo.  
348 *eLife*, **7**: e32341.
- 349 Bohère, J., Mancheno-Ferris, A., Al Hayek, S., Zanet, J., Valenti, P., Akino, K., Yamabe, Y.,  
350 Inagaki, S., Chanut-Delalande, H., Plaza, S., Kageyama, Y., Osman, D., Polesello, C. and  
351 Payre, F. 2018. Shavenbaby and Yorkie mediate Hippo signaling to protect adult stem cells  
352 from apoptosis. *Nat. Commun.*, **9**: 5123.
- 353 Boumard, B. and Bardin, A.J. 2021. An amuse-bouche of stem cell regulation: Underlying  
354 principles and mechanisms from adult *Drosophila* intestinal stem cells. *Curr Opin Cell Biol*, **73**:  
355 58–68.
- 356 Buenrostro, J.D., Corces, M.R., Lareau, C.A., Wu, B., Schep, A.N., Aryee, M.J., Majeti, R.,  
357 Chang, H.Y. and Greenleaf, W.J. 2018. Integrated Single-Cell Analysis Maps the Continuous  
358 Regulatory Landscape of Human Hematopoietic Differentiation. *Cell*, **173**: 1535-1548.e16.
- 359 Bulyk, M.L. 2003. Computational prediction of transcription-factor binding site locations.  
360 *Genome Biol*, **5**: 201.
- 361 Castro-Mondragon, J.A., Jaeger, S., Thieffry, D., Thomas-Chollier, M. and van Helden, J.  
362 2017. RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription  
363 factor binding motif collections. *Nucleic Acids Res*, **45**: e119.
- 364 Chanut-Delalande, H., Fernandes, I., Roch, F., Payre, F. and Plaza, S. 2006. Shavenbaby  
365 couples patterning to epidermal cell shape control. *PLoS Biol*, **4**: e290.
- 366 Chanut-Delalande, H., Hashimoto, Y., Pelissier-Monier, A., Spokony, R., Dib, A., Kondo, T.,  
367 Bohère, J., Niimi, K., Latapie, Y., Inagaki, S., Dubois, L., Valenti, P., Polesello, C., Kobayashi,  
368 S., Moussian, B., White, K.P., Plaza, S., Kageyama, Y. and Payre, F. 2014. Pri peptides are  
369 mediators of ecdysone for the temporal control of development. *Nat. Cell Biol.*, **16**: 1035–1044.
- 370 Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna,  
371 J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A. and Jaenisch, R.  
372 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental  
373 state. *Proc. Natl. Acad. Sci. U. S. A.*, **107**: 21931–21936.
- 374 Cubenas-Potts, C., Rowley, M.J., Lyu, X., Li, G., Lei, E.P. and Corces, V.G. 2017. Different  
375 enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique  
376 chromatin interactions and 3D architecture. *Nucleic Acids Res*, **45**: 1714–1730.
- 377 Cuvier, O. and Fierz, B. 2017. Dynamic chromatin technologies: from individual molecules to  
378 epigenomic regulation in cells. *Nat. Rev. Genet.*, **18**: 457–472.
- 379 Fernandes, I., Chanut-Delalande, H., Ferrer, P., Latapie, Y., Waltzer, L., Affolter, M., Payre, F.  
380 and Plaza, S. 2010. Zona pellucida domain proteins remodel the apical compartment for  
381 localized cell shape changes. *Dev Cell*, **18**: 64–76.
- 382 Guertin, M.J. and Lis, J.T. 2010. Chromatin landscape dictates HSF binding to target DNA  
383 elements. *PLoS Genet*, **6**: e1001114.

- 384 Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee,  
385 L.K., Stuart, R.K., Ching, C.W., Ching, K.A., Antosiewicz-Bourget, J.E., Liu, H., Zhang, X.M.,  
386 Green, R.D., Lobanenkov, V.V., Stewart, R., Thomson, J.A., Crawford, G.E., *et al.* 2009.  
387 Histone modifications at human enhancers reflect global cell-type-specific gene expression.  
388 *Nature*, **459**: 108–112.
- 389 Herz, H.-M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y.-H., Mickey, K., Voets, O.,  
390 Verrijzer, C.P. and Shilatifard, A. 2012. Enhancer-associated H3K4 monomethylation by  
391 Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes Dev.*, **26**: 2604–  
392 2620.
- 393 Hong, T., Watanabe, K., Ta, C.H., Villarreal-Ponce, A., Nie, Q. and Dai, X. 2015. An Ovol2-  
394 Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between  
395 Epithelial and Mesenchymal States. *PLoS Comput Biol*, **11**: e1004569.
- 396 Huang, C., Yang, F., Zhang, Z., Zhang, J., Cai, G., Li, L., Zheng, Y., Chen, S., Xi, R. and Zhu,  
397 B. 2017. Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1  
398 Trithorax group protein function in Drosophila. *Nat. Commun.*, **8**: 1649.
- 399 Ibrahim, M.M., Karabacak, A., Glaes, A., Kolundzic, E., Hirsekorn, A., Carda, A., Tursun, B.,  
400 Zinzen, R.P., Lacadie, S.A. and Ohler, U. 2018. Determinants of promoter and enhancer  
401 transcription directionality in metazoans. *Nat. Commun.*, **9**: 4472.
- 402 Jiang, H. and Edgar, B.A. 2009. EGFR signaling regulates the proliferation of Drosophila adult  
403 midgut progenitors. *Development*, **136**: 483–493.
- 404 Johnson, S.C. 1967. Hierarchical clustering schemes. *Psychometrika*, **32**: 241–254.
- 405 Kellner, W.A., Ramos, E., Van Bortle, K., Takenaka, N. and Corces, V.G. 2012. Genome-wide  
406 phosphoacetylation of histone H3 at Drosophila enhancers and promoters. *Genome Res*, **22**:  
407 1081–8.
- 408 Kohlmaier, A., Fassnacht, C., Jin, Y., Reuter, H., Begum, J., Dutta, D. and Edgar, B.A. 2015.  
409 Src kinase function controls progenitor cell pools during regeneration and tumor onset in the  
410 Drosophila intestine. *Oncogene*, **34**: 2371–2384.
- 411 Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., Kobayashi, S., Payre,  
412 F. and Kageyama, Y. 2010. Small peptides switch the transcriptional activity of Shavenbaby  
413 during Drosophila embryogenesis. *Science*, **329**: 336–339.
- 414 Kumar, A., Bhandari, A., Sinha, R., Sardar, P., Sushma, M., Goyal, P., Goswami, C. and  
415 Grapputo, A. 2012. Molecular phylogeny of OVOL genes illustrates a conserved C2H2 zinc  
416 finger domain coupled by hypervariable unstructured regions. *PLoS One*, **7**: e39399.
- 417 Lee, S. and Garfinkel, M.D. 2000. Characterization of Drosophila OVO protein DNA binding  
418 specificity using random DNA oligomer selection suggests zinc finger degeneration. *Nucleic  
419 Acids Res*, **28**: 826–34.
- 420 Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.-T., Cubeñas-Potts, C., Hu, M.,  
421 Lei, E.P., Bosco, G., Qin, Z.S. and Corces, V.G. 2015. Widespread rearrangement of 3D  
422 chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol. Cell*, **58**:  
423 216–231.
- 424 Liang, J., Lacroix, L., Gamot, A., Cuddapah, S., Queille, S., Lhoumaud, P., Lepetit, P., Martin,  
425 P.G.P., Vogelmann, J., Court, F., Hennion, M., Micas, G., Urbach, S., Bouchez, O., Nöllmann,  
426 M., Zhao, K., Emberly, E. and Cuvier, O. 2014. Chromatin immunoprecipitation indirect peaks  
427 highlight long-range interactions of insulator proteins and Pol II pausing. *Mol. Cell*, **53**: 672–  
428 681.
- 429 Liu, J., Ali, M. and Zhou, Q. 2020. Establishment and evolution of heterochromatin. *Ann N  
430 Acad Sci*, **1476**: 59–77.
- 431 Lu, J. and Oliver, B. 2001. Drosophila OVO regulates ovarian tumor transcription by binding

- 432 unusually near the transcription start site. *Development*, **128**: 1671–86.
- 433 Menoret, D., Santolini, M., Fernandes, I., Spokony, R., Zanet, J., Gonzalez, I., Latapie, Y.,  
434 Ferrer, P., Rouault, H., White, K.P., Besse, P., Hakim, V., Aerts, S., Payre, F. and Plaza, S.  
435 2013. Genome-wide analyses of Shavenbaby target genes reveals distinct features of  
436 enhancer organization. *Genome Biol.*, **14**: R86.
- 437 Mevelninio, M., Terracol, R., Salles, C., Vincent, A. and Payre, F. 1995. Ovo, a Drosophila  
438 Gene Required for Ovarian Development, Is Specifically Expressed in the Germline and  
439 Shares Most of Its Coding Sequences with Shavenbaby, a Gene Involved in Embryo  
440 Patterning. *Mech. Dev.*, **49**: 83–95.
- 441 Micchelli, C.A. and Perrimon, N. 2006. Evidence that stem cells reside in the adult Drosophila  
442 midgut epithelium. *Nature*, **439**: 475–9.
- 443 Nguyen, N.T.T., Contreras-Moreira, B., Castro-Mondragon, J.A., Santana-Garcia, W., Ossio,  
444 R., Robles-Espinoza, C.D., Bahin, M., Collombet, S., Vincens, P., Thieffry, D., van Helden, J.,  
445 Medina-Rivera, A. and Thomas-Chollier, M. 2018. RSAT 2018: regulatory sequence analysis  
446 tools 20th anniversary. *Nucleic Acids Res*, **46**: W209–W214.
- 447 Payre, F., Vincent, A. and Carreno, S. 1999. Ovo/svb integrates Wingless and DER pathways  
448 to control epidermis differentiation. *Nature*, **400**: 271–275.
- 449 Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T.,  
450 Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. and Aiden, E.L. 2014. A 3D Map of the  
451 Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, **159**:  
452 1665–1680.
- 453 Roca, H., Hernandez, J., Weidner, S., McEachin, R.C., Fuller, D., Sud, S., Schumann, T.,  
454 Wilkinson, J.E., Zaslavsky, A., Li, H., Maher, C.A., Daignault-Newton, S., Healy, P.N. and  
455 Pienta, K.J. 2013. Transcription factors OVOL1 and OVOL2 induce the mesenchymal to  
456 epithelial transition in human cancer. *PLoS One*, **8**: e76773.
- 457 Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S. and Honig, B. 2009. The role of DNA  
458 shape in protein-DNA recognition. *Nature*, **461**: 1248-U81.
- 459 Rouault, H., Mazouni, K., Couturier, L., Hakim, V. and Schweisguth, F. 2010. Genome-wide  
460 identification of cis-regulatory motifs and modules underlying gene coregulation using statistics  
461 and phylogeny. *Proc Natl Acad Sci U S A*, **107**: 14615–20.
- 462 Schauer, T., Ghavi-Helm, Y., Sexton, T., Albig, C., Regnard, C., Cavalli, G., Furlong, E.E. and  
463 Becker, P.B. 2017. Chromosome topology guides the Drosophila Dosage Compensation  
464 Complex for target gene activation. *EMBO Rep.*, **18**: 1854–1868.
- 465 Sen, S.Q., Chanchani, S., Southall, T.D. and Doe, C.Q. 2019. Neuroblast-specific open  
466 chromatin allows the temporal transcription factor, Hunchback, to bind neuroblast-specific loci.  
467 *Elife*, **8**.
- 468 Sethi, A., Gu, M., Gumusgoz, E., Chan, L., Yan, K.K., Rozowsky, J., Barozzi, I., Afzal, V.,  
469 Akiyama, J.A., Plajzer-Frick, I., Yan, C., Novak, C.S., Kato, M., Garvin, T.H., Pham, Q.,  
470 Harrington, A., Mannion, B.J., Lee, E.A., Fukuda-Yuzawa, Y., *et al.* 2020. Supervised enhancer  
471 prediction with epigenetic pattern recognition and targeted validation. *Nat Methods*, **17**: 807–  
472 814.
- 473 Slattery, M., Zhou, T.Y., Yang, L., Machado, A.C.D., Gordan, R. and Rohs, R. 2014. Absence  
474 of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.*, **39**: 381–  
475 399.
- 476 Sousa-Victor, P., Ayyaz, A., Hayashi, R., Qi, Y., Madden, D.T., Lunyak, V.V. and Jasper, H.  
477 2017. Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. *Cell Rep.*, **20**: 2527–  
478 2537.
- 479 Spitz, F. and Furlong, E.E. 2012. Transcription factors: from enhancer binding to

- 480 developmental control. *Nat Rev Genet*, **13**: 613–26.
- 481 Takahashi, K. and Yamanaka, S. 2006. Induction of Pluripotent Stem Cells from Mouse  
482 Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, **126**: 663–676.
- 483 Teng, A., Nair, M., Wells, J., Segre, J.A. and Dai, X. 2007. Strain-dependent perinatal lethality  
484 of *Ovol1*-deficient mice and identification of *Ovol2* as a downstream target of *Ovol1* in skin  
485 epidermis. *Biochim Biophys Acta*, **1772**: 89–95.
- 486 Tsuji, G., Hashimoto-Hachiya, A., Kiyomatsu-Oda, M., Takemura, M., Ohno, F., Ito, T., Morino-  
487 Koga, S., Mitoma, C., Nakahara, T., Uchi, H. and Furue, M. 2017. Aryl hydrocarbon receptor  
488 activation restores filaggrin expression via *OVOL1* in atopic dermatitis. *Cell Death Dis*, **8**:  
489 e2931.
- 490 Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X.,  
491 Kundaje, A., Cheng, Y., Rando, O.J., Birney, E., Myers, R.M., Noble, W.S., Snyder, M. and  
492 Weng, Z. 2012. Sequence features and chromatin structure around the genomic regions bound  
493 by 119 human transcription factors. *Genome Res*, **22**: 1798–812.
- 494 Wang, Z.H., Li, Z., Hu, M., Yang, Q.J., Yan, S., Wu, R.S., Li, B.A. and Guo, M. 2017. *Ovol2*  
495 gene inhibits the Epithelial-to-Mesenchymal Transition in lung adenocarcinoma by  
496 transcriptionally repressing *Twist1*. *Gene*, **600**: 1–8.
- 497 Watanabe, K., Villarreal-Ponce, A., Sun, P., Salmans, M.L., Fallahi, M., Andersen, B. and Dai,  
498 X. 2014. Mammary morphogenesis and regeneration require the inhibition of EMT at terminal  
499 end buds by *Ovol2* transcriptional repressor. *Dev Cell*, **29**: 59–74.
- 500 Zabidi, M.A., Arnold, C.D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A.  
501 2015. Enhancer-core-promoter specificity separates developmental and housekeeping gene  
502 regulation. *Nature*, **518**: 556–9.
- 503 Zanet, J., Benrabah, E., Li, T., Pélissier-Monier, A., Chanut-Delalande, H., Ronsin, B., Bellen,  
504 H.J., Payre, F. and Plaza, S. 2015. Pri sORF peptides induce selective proteasome-mediated  
505 protein processing. *Science*, **349**: 1356–1358.
- 506
- 507



508  
509

**Figure 1: Svb-ACT and REP bind common enhancers to antagonistically regulate a common set of target genes.**

(A) Cultured S2 cells provide an exquisite model system for studying the mechanistic of Svb isoforms functions. Pri smORF peptides trigger SvbREP to SvbACT post-translational, proteasome-dependent processing.

(B) ChIP-seq heatmaps for Svb-ACT, SvbREP and the indicated histones marks available from S2 cells. All heatmaps are centered on Svb-ACT peaks and sorted by Svb-ACT signal intensity. Color intensity reflects the level of ChIP-seq enrichment presented as  $\log_2(\text{ChIP}/\text{input})$ .

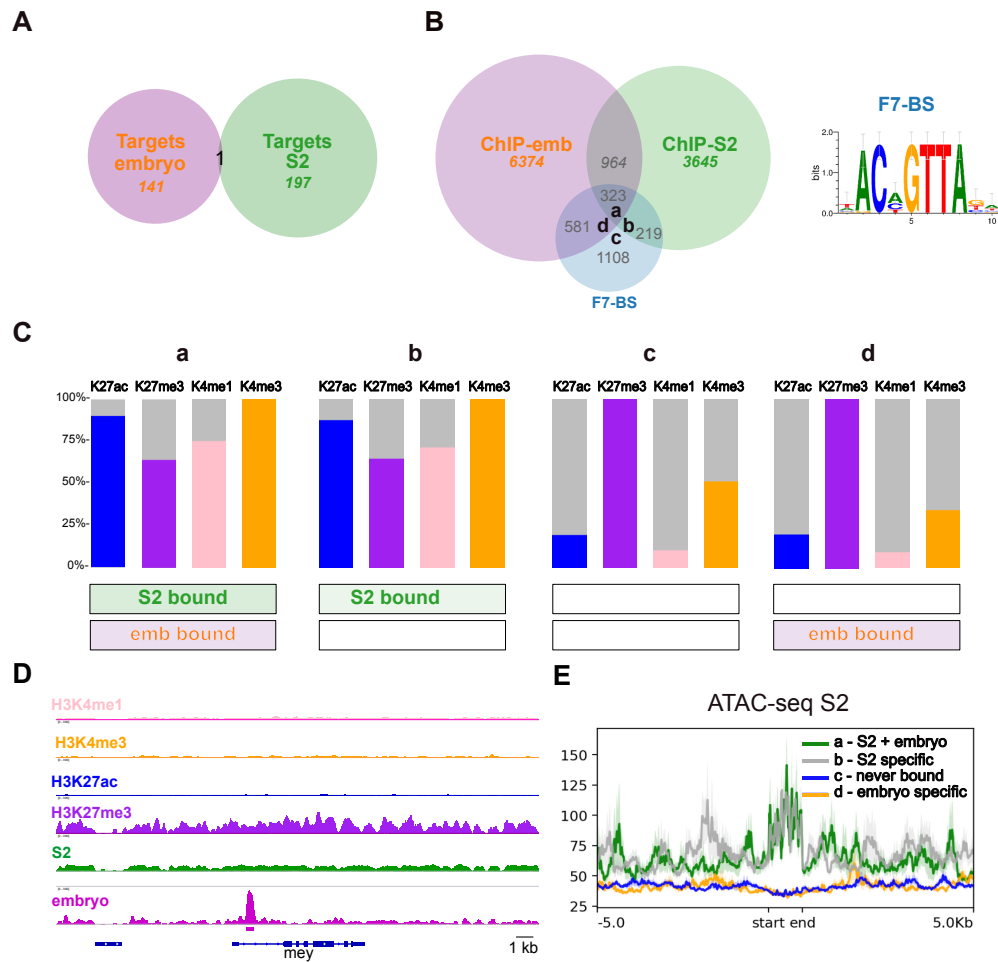
(C) Histone marks averaged profile centered on Svb-ACT peaks.

(D) DNA motifs most enriched in Svb-ACT bound peaks as defined by RSAT suite, the correlation score reflect the result of the Pearson test between the logo detected and the logo present on the RSAT database (Nguyen *et al.*, 2018).

(E) Interaction between Svb-ACT peaks and active enhancers identified by STARR-seq (Arnold *et al.*, 2013). Left, venn diagram showing the overlap of Svb-ACT bound regions



525 (green, n=6939) and active enhancers (brown, n=5499). Right, average enrichment of Sv-  
526 ACT (green), STARR signal (brown) and H3K27ac (blue) for all active enhancers, Sv-bound  
527 and Sv-unbound enhancers.  
528 (F) Hierarchical clustering of DE genes across RNAseq experimental conditions.  
529 (G) Metaprolife of the Cluster 5 of the hierarchical clustering identified in F.  
530 (H) r-score defines the robustness of each gene cluster from hierarchical clustering, based on  
531 its conservation through 15 iterations of k-means clustering. See Figure S1 and methods for  
532 details. c5 is highlighted in green.  
533 (I) Venn diagram showing a robust set of 195 genes considered as “S2 Sv target genes” is  
534 found through supervised (see Figure S1), hierarchical and k-means clusterings.  
535 (J) Diagram showing the proportion of genes with SvACT ChIP peaks in the whole genome  
536 (Dm6), all differentially expressed genes (DE), or Sv putative target genes define by either  
537 supervised method, hierarchical clustering or K-means clusterings.  
538 (K) Genome browser screenshot of the CG4297 locus displaying typical Sv target gene  
539 behavior (basal expression in blue, Sv-ACT upregulation in green; Sv-REP downregulation  
540 in red, Sv ChIP in dark green).  
541  
542  
543  
544  
545  
546  
547  
548  
549



550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573

## Figure 2: Chromatin landscape constrains Svb binding

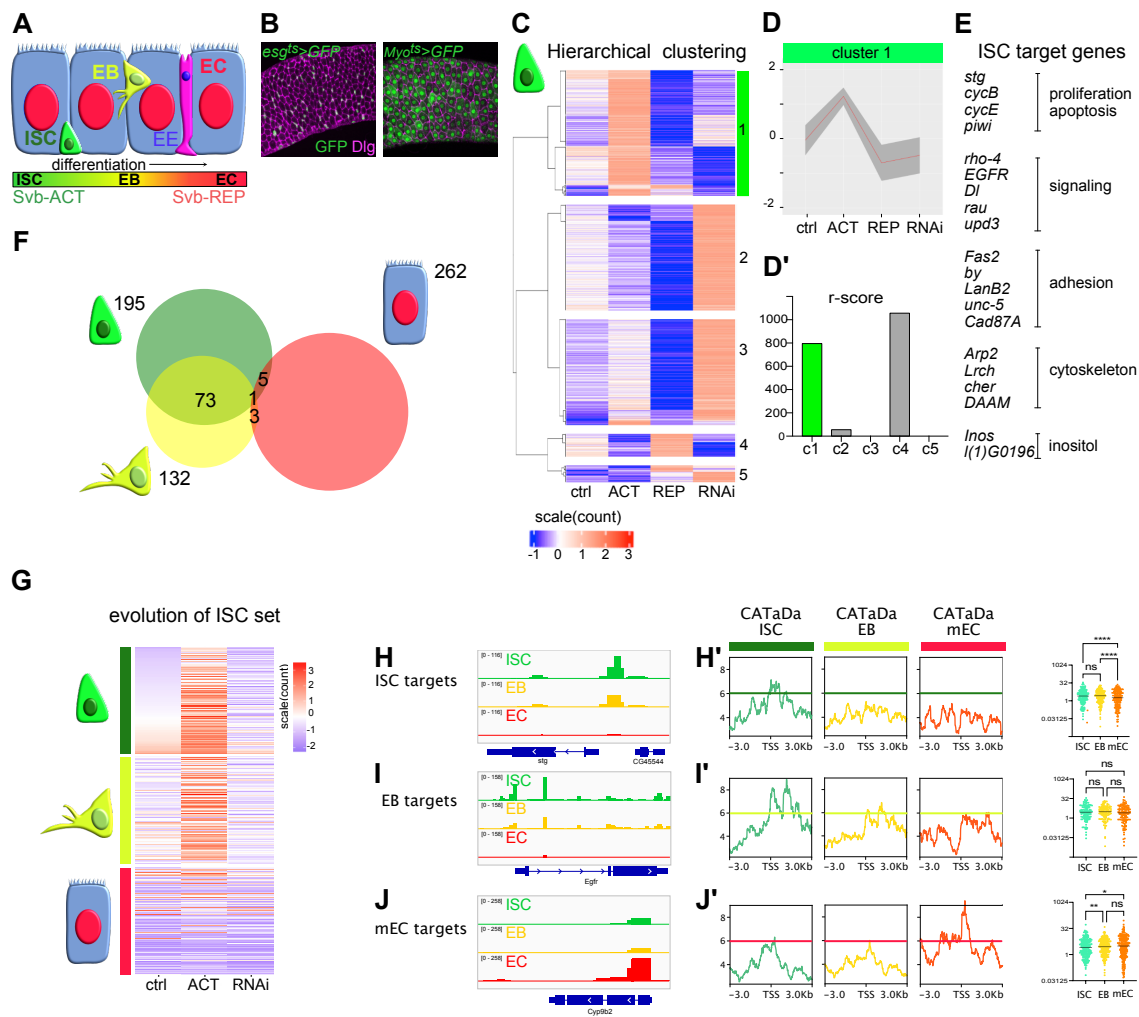
(A) Venn diagram showing the overlap between the sets of Svb target genes in S2 cells (n=198) and embryonic epidermal cells (n=142, (Menoret *et al.*, 2013)).

(B) Venn diagram showing the intersections between Svb-bound regions in S2 cells (green), in embryonic epidermal cells (orange), and genomic F7-BS regions (blue). F7-BS position weight matrix is shown on the right. The “Svb-bound” regions overlap with F7-BS regions if a F7-BS is found less than 3kb away from the ChIP peak. Each F7-BS is thus bound or not by Svb in either cell type. a,b,c and d mark the 4 types of F7-BS behavior shown in panels C and E.

(C) Enrichment of the indicated histones marks around S7-BS motifs in the different sets defined in (B).

(D) Snapshot of IGV genome browser showing Svb ChIP (from S2 cells in green and from embryos in magenta) at *mey* locus. In embryo Svb binds a canonical enhancer driving *mey* gene expression in embryonic epidermal cells (Menoret *et al.*, 2013). The different histone marks (H3K4me1 in pink, H3K4me3 in yellow, H3K27ac in blue, H3K27me3 in purple) allow visualization of the chromatin organization of the region in S2 cells, consistent with repressive configuration and absence of Svb binding in this cell type.

(E) ATAC-seq metaprofiles from S2 cells (Ibrahim *et al.*, 2018) show chromatin accessibility over 10 kb around F7-BS for the clusters defined in B.



574

575

576

577 **Figure 3: Evolution of Svb target genes through the lineage of intestinal stem cells**  
 578 (A) Schematic representation of the cell types of the intestinal epithelium leading to EC  
 579 differentiation. Svb-*ACT* is the majority form in ISCs and Svb-*REP* in ECs. ISC in green, EB in  
 580 yellow, EC in blue and EE in purple.

581 (B) Cell type specific drivers have been used to manipulate Svb expression in adult fly  
 582 progenitors (ISCs and EBs, *esg:GAL4*, *Gal80<sup>ts</sup>*) or ECs (*myo1A:GAL4*, *Gal80<sup>ts</sup>*).

583 (C) Hierarchical clustering of RNAseq data from sorted ISC reveals DE genes upon Svb  
 584 modulation. Four experimental conditions were used: *control* (ctrl), *UAS-SvbACT* (ACT),  
 585 *UAS-Svb<sup>3Kmut</sup>* (REP), and *UAS-Svb RNAi* (RNAi). Cluster 1 comprises genes activated by SvbACT  
 586 and repressed by SvbREP and upon Svb loss of function.

587 (D) Metaprofile of c1 corresponding to ISC target genes. (D') Robustness analysis (r-score) of  
 588 each hierarchical cluster.

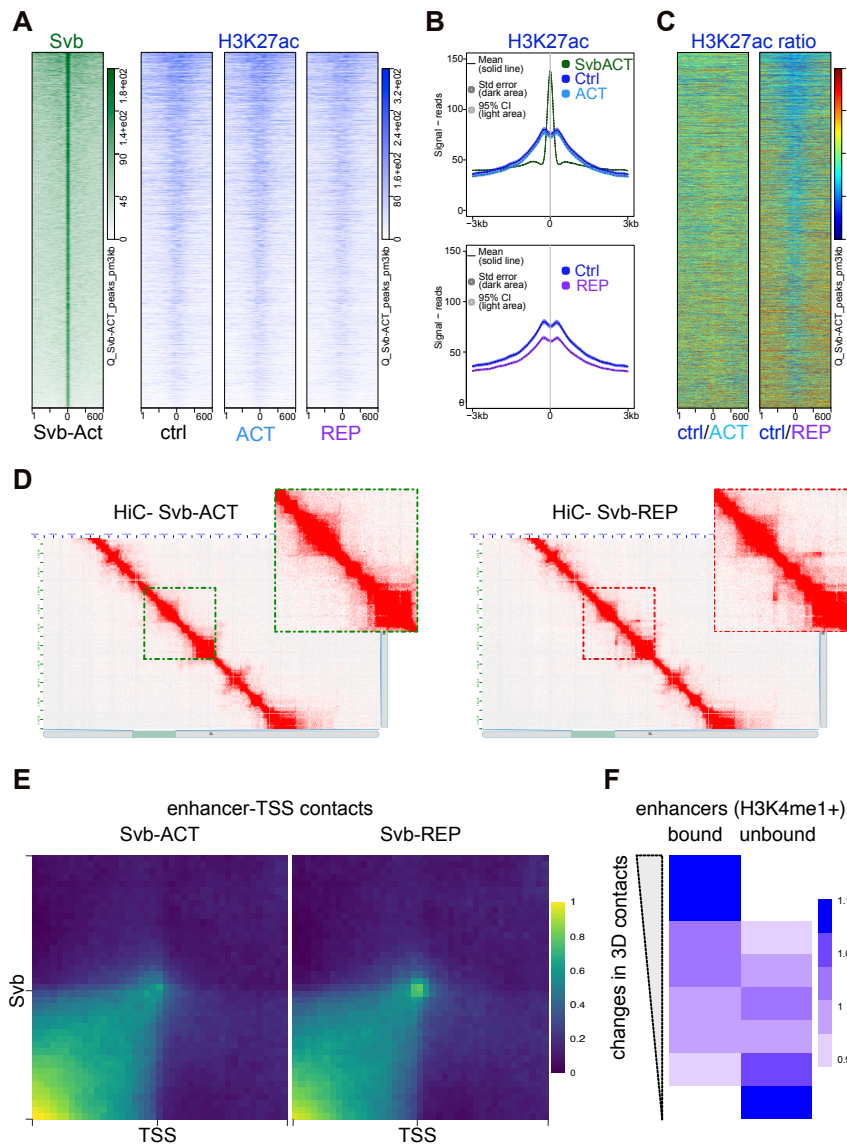
589 (E) Subset of genes from c1 with their associated functions.

590 (F) Venn diagram showing how Svb target genesets evolve during intestinal cell differentiation.  
 591 Note ISC/EB target genesets similarity as opposed to EC targets. Svb target genes were  
 592 identified in ISC, EB, eEC and mEC following the same pipeline (see FigureS2/Table S1).

593 (G) Heatmaps showing expression of the 195 ISC target genes in ctrl, Svb-*ACT* and *svb-RNAi*  
 594 conditions, in ISC, EB and mEC (top to bottom).

595 (H-J) CATaDa data (Aughey *et al.*, 2018) allow studying the dynamics of chromatin  
 596 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

597 gene (*Cyp9b2*, J), showing CaTADA signal in ISC (green), in EB (yellow) and in EC (red).  
 598 Averaged CATaDa signal profile (in reads per million), centered on the TSS of all ISC target  
 599 genes (H'), all EB target genes (I') and all mEC target genes (J') are shown, showing  
 600 accessibility in ISC (green), EB (yellow) and EC cells (red) (left to right of each panel). The  
 601 plots on the right present for each gene the CATaDa signal around the TSS (1kb on both side).  
 602 Statistical analysis confirms the general trend that ISC target genes' TSS close and mEC target  
 603 genes' TSS open during intestinal cell differentiation (Kruskal-wallis test with Dunn's multiple  
 604 comparisons).  
 605



606  
 607

608 **Figure 4: The SvbREP-to-ACT switch is accompanied by a dynamic remodeling of the**  
 609 **enhancer chromatin landscape**

610 (A) Heatmap showing both Svb-ACT binding (green) and H3K27ac abundance (blue) at Svb-  
 611 bound regions in control, SvbACT and SvbREP cells. All heatmaps are ranked by Svb ChIP  
 612 intensity.

613 (B) Metaprofiles overlay of A (ACT vs ctrl and REP vs ctrl).

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

616 (D) 2D plot showing our long-range interactions as probed using chromosome conformation  
617 capture (3C/Hi-C) in cells expressing Svb-ACT compared to Svb-REP. Enlarged areas mark  
618 possible 3D long-range contacts as highlighted (dotted squares).  
619 (E) Aggregate plots of HiC data in SvbACT and SvbREP cells assessing the dynamics of 3D  
620 chromatin contacts for all enhancers identified by STARR-seq ([Arnold \*et al.\*, 2013](#); [Zabidi \*et\*  
621 \*al.\*, 2015](#)) and their distant target TSSs (see Methods). The averaged signal corresponds to  
622 the groups enriched in Svb-bound enhancers and with a decrease in long-range contacts in  
623 3D (see panel F).  
624 (F) Genome-wide analysis of long-range contacts between enhancer and promoters in Svb-  
625 ACT compared to Svb-REP cells. Long-range contacts were measured for all putative  
626 enhancer-promoter pairs in SvbACT and compared to SvbREP cells genome-wide (see  
627 Methods). The net variations were quantitized for all enhancer-promoter pairs and ranked from  
628 highest to lowest changes in 3D contacts. A Fisher exact test was used to assess the  
629 enrichment of the quantiles of changes in 3D interactions presence or absence of Svb binding  
630 to the enhancers.  
631

## 632 **Materials and methods**

633

### 634 **S2 cells lines:**

635 *Drosophila* S2 cells were grown in Schneider medium supplemented with 10% fetal calf serum  
636 and 1% penicillin/streptomycin (Invitrogen). We used stable cell lines co- expressing the  
637 copper-inducible constructs pMT-Svb::GFP and pMT-pri (“1B”), or pMT- Svb::GFP and pMT-  
638 pri1-4fs (“FS”), which encodes a frame-shifted variant as a negative control ([Zanet et al., 2015](#)).  
639 The expression of pMT plasmids was induced by CuSO<sub>4</sub> at the final concentration of 1 mM  
640 during 24h.

641

### 642 **ChIP experiments:**

643 ChIP experiments were performed essentially as previously described ([Lhoumaud et al., 2014](#))  
644 with the following specifications: Cells are cross-linked with 0.8% of formaldehyde for 10 min  
645 at room temperature. Cross-linking is stopped by adding glycine (2M) and cells were washed  
646 twice in cold with PBS 1X and NaBu 1 mM. Cells are resuspended in 500 µL of ChIP  
647 permeation buffer (PBS + 0.2% Triton + 10 mM NaBu), incubated for 20 min at room  
648 temperature; cells are washed with Lysis Buffer (140 mM NaCl 5M, 15 mM HEPES pH 7.6,  
649 1 mM EDTA, 1% Triton, 0.1% Sodium Deoxycholate, 0.5 mM DTT, 10nM sodium butyrate, 25X  
650 protease inhibitor), and resuspended in Lysis Buffer, 1% SDS, 0.5% N-lauroylsarcosine.  
651 Chromatin is sonicated in Bioruptor (Diagenode) using high-power settings, intervals of 30s  
652 burst/pause for 15 cycles to obtain fragments of ≈ 250 pb. The sonicated chromatin is diluted  
653 10 times with LB no SDS on protein low bind tubes (Eppendorf) and the chromatin fraction  
654 cleared by centrifugation (16000 g, 10 °C, 5 min) is then used for immunoprecipitation. Beads  
655 are washed and blocked with Lysis Buffer 0.1% SDS, 0.5% N-lauroylsarcosyl and  
656 BSA 0.1 mg/mL. Complex beads antibodies were established at 4° C overnight with Lysis  
657 Buffer 0.1% SDS, 0.5% N-lauroylsarcosyl on protein low bind tubes.

658 GFP-TRAP beads (Chromotek) and anti H3K27ac (abcam ab4729) antibodies were used for  
659 ChIP.

660 Chromatin is pre-clarified O/N at 4° C before the immunoprecipitation step that is carried out  
661 at 4° C for 4h. 10% of each sample is used as input. Chromatin-bead complexes are eluted  
662 twice at 70 °C during 20 min, first with 10 mM EDTA, 1% SDS, 50 mM Tris-Cl pH8 and then  
663 with TE, 0.67% SDS. Reverse cross-link is performed overnight at 65 °C and DNA purified  
664 after phenol-chloroform extraction and ethanol precipitation, for final resuspension in 1x TE.

665

666

667 **ChIPseq sequencing and analysis:**

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

676 For all conditions, reads were mapped on dm6 genome (r6.13 flybase) with bwa (v0.7.17-  
677 r1188) with default parameters, unique reads were filtered and peak calling was done on  
678 replicates merge files with MACS2 (Feng *et al.*, 2011) for histone marks of databases (--  
679 nomodel --broad --f BAM --g dm --B --q 0.00001) or for histones mark made in this study (--  
680 nomodel --f BAMPE --g dm --B --q 1 e-05 --broad) and (-f BAM --B --g dm --to-large --nomodel -  
681 -q 0.0001 or --q 0.1) for Svb forms. Metrics and quality of ChIPseq peaks were calculated with  
682 fastQC (v0.11.8) and CHIPQC package (Carroll *et al.*, 2014)(v 1.26.0). To compare the set of  
683 peaks, ChIPpeakanno9 (v3.24.2) packages and BEDtools fisher (v2.29.2) were used.

684 ChIPseq signal was normalized by taking into account the genome mapping with  
685 bamCoverage (Ramírez *et al.*, 2016) (--binSize 10 --normalizeUsing RPKM --  
686 effectiveGenomeSize 125464728), then we normalized for the noise due to the experiment  
687 with bigwigCompare (-b1 experiment file -b2 input file --scaleFactors 1:1 --operation "log2").  
688 Due to the fact that SvbREP is less express than SvbACT in S2 cells the normalization for the  
689 noise considers the importance of this difference scaling SvbREP peaks signal by the result of  
690 the ratio of logFC for svb in the two conditions: --scaleFactors 1.41:1. Heatmaps of binding  
691 enrichment were performed with computeMatrix (-bS 100 -a 3000 -b 3000) and plotHeatmap  
692 (deeptools v 3.5.0 (Ramírez *et al.*, 2016)). RSAT (Medina-Rivera *et al.*, 2015; Thomas-Chollier  
693 *et al.*, 2008) was used to determine the motif enrichment on Svb peaks. Presence analysis of  
694 histones in presence of Svb were made with deeptools fisher (Ramírez *et al.*, 2016)  
695 software. Coverage analysis was done with MultiBamSummary and plotCorrelation of  
696 deeptools (Ramírez *et al.*, 2016) software.

697

698 **RNA Extraction, Sequencing and Analysis on S2 cells:**

699 RNAseq experiments were made on control, 1B and FS cells (GSE199511). Cells were  
700 harvested and RNA extraction was made with RNeasy Kit (Qiagen). Bank of reads and  
701 sequencing were done by HiSeq 200 (Illumina) at BGI to obtain 29-34M of reads (single-end  
702 50nt) per replicates. The quality of sequencing was measured with fastQC (v0.11.5). Mapping

703 was done with STAR ([Dobin et al., 2013](#))(v 2.5.2b, default parameter) on Drosophila genome  
704 dm6 (r6.13 flybase), reads were counted with HTseq-count ([Anders et al., 2015](#))(v0.6.0, -t gene  
705 -r pos -i gene\_symbol) and statistical analysis were performed with edgeR ([Robinson et al.,  
706 2010](#)) using the negative binomial generalized log-linear model to the read counts for each  
707 gene (glmLRT) and decideTestDGE to determine differential genes. Graphs were obtained with  
708 Glimma (Su et al., 2017)(v 2.0.0) and Plotly (Sievert, n.d.)(v 4.9.3). Identification of direct target  
709 genes was performed with GenomicRanges ([Lawrence et al., 2013](#))(v 1.42.0) and  
710 superExactTest ([Wang et al., 2015](#)). Clustering analysis described in dedicated Clustering  
711 method section.

712

### 713 **Drosophila genetics**

714 See complete genotypes and references in Reagents and Tools table To purify ISC and EB,  
715 *esg<sup>ts</sup>* virgin females were crossed with the following males: Control, RNAi, ACT, or REP. To  
716 purify enterocytes, *Myo1A<sup>ts</sup>* virgin females were crossed to the same males. Adult progeny  
717 (only virgin female) was collected, aged at room temperature (22°C) for 72h then transferred  
718 at 29°C to induce UAS expression. Flies were kept for 6 days at 29°C, changing the food tube  
719 every 48h hours to avoid bacterial proliferation and intestinal stress.

720

### 721 **Gut dissection and cell dissociation:**

722 Batches of >20 virgin females were CO<sub>2</sub>-anaesthetized, put on ice, their midgut were dissected  
723 in PBS. Malpighian tubules and hindgut were removed and only the posterior part of the midgut  
724 is kept (R3, R4, R5). PBS was replaced by cold 1mM PBS1X-EDTA. Posterior midguts were  
725 shrunk into pieces using a razorblade (around 30-60 seconds on a glass plate), then  
726 transferred into a fresh tube. Samples were spn 5minutes at 1000rpm 4°C, and supernatant  
727 discarded except 50µL. 200uL TrypLE 10X was added (8X final with 0.2mM EDTA) and the  
728 guts were incubated at 37°C for 10minutes with intense rocking. Digestions were stopped with  
729 750µL SSM (Serum Supplemented Medium = Schneider's medium +10% Fetal Bovine Serum  
730 + 1% Pen/Strep. Then samples were treated differently for FACS (ISC, EB, eEC) or for filtering  
731 (mEC). In both cases, mechanical trituration was performed by pipetting up and down using  
732 low binding, flame-rounded narrow tips (home made).

733

### 734 **Sample preparation for FACS**

735

736 The cell solution was strained into a cold falcon tube (70µM nylon cell strainer, pre-wet) with  
737 1mLSSM). The Eppendorf and the strains were rinsed with 3x1mL PBS-EDTA, so that the



738 samples are 5mL total. The quality of the dissociation and the health of the cells are checked  
739 under the microscope. Samples were spun at 700rcf, 4°C, for >10 minutes. Supernatant was  
740 discarded except 100/200µL, to which we added 200µL PBS + 3µL clean NGS. Propidium  
741 Iodide was added to label dead cells, then cells of interest were sorted using FACS (Cell  
742 Sorter BD FACSAria Fusion in BSL3) directly into a tube containing 300µL 1X Tissue and  
743 Cell Lysis solution (MasterPure RNA purification kit) + 1µg Proteinase K [50µg/µL]. Cells  
744 were vortexed and placed on ice just after sorting.

745

746 Sample preparation for straining mature EC

747

748 We noted that ISC and EB were easy to dissociate but EC were not. Longer incubation in  
749 TRYPLE and harsher trituration allowed dissociation of individual EC1, but this was  
750 accompanied with high mortality. We thus reasoned that it would be less stressful for EC to  
751 simply proceed with mild dissociation of ISC and EB and simply recover patches of healthy  
752 enterocytes by simple filtering. For this procedure, digested samples are recovered on a  
753 prewet 70µM nylon cell strainer. The strained is rinsed to remove dissociated cells, then turned  
754 upside down on a petri dish and rinsed with 5mL PBS-EDTA (5mL) then 2mL SSM. Cell  
755 patches were checked under fluorescent microscope. The samples were transferred into a  
756 tube and spun at 700 rcf at 4°C for 10 minutes. The supernatant was discarded except 150uL.  
757 This pellet was mixed with 150uL 2X Tissue and Cell Lysis solution (MasterPure RNA  
758 purification kit) + 1ug Proteinase K [50ug/uL].

759

760 **RNAseq, Sequencing and Quality Check on intestinal cells (ISC, EB, eEC and mEC):**

761 RNAs were purified from each sample using MasterPure RNA purification kit, following  
762 manufacturer's recommendation (including the DNase treatment). Final RNAs were  
763 resuspended in 10uL TE and added 1uL RiboGuard (RNases inhibitor). Construction of RNA  
764 banks and sequencing was made on the Montpellier GenomiX Platform. RNA banks were done  
765 with the Ovation SoLo RNAseq System kit (Nugen). Bank validation was done with the  
766 quantification of the complementary DNA with the Standard Sensitivity NGS kit on Fragment  
767 Analyzer and with qPCR (ROCHE Light Cycler 480). The sequencing was done on NovaSeq  
768 6000 (Illumina) with NovaSeq Reagent Kits (100cycles). Single-reads of 100nt were  
769 sequenced. ISC= GSE199510, EB=GSE220323, eEC=GSE220558, mEC=GSE220560.

770 The quality of sequencing was measured with fastQC (v0.11.5). Mapping was done with  
771 STAR<sup>2</sup> (v 2.5.2b, default parameter) on Drosophila genome dm6 (r6.13 flybase), reads were  
772 counted with HTseq-count ([Anders et al., 2015](#)) (v0.6.0, -t gene -r pos -i gene\_symbol) and

773 statistical analysis were performed with edgeR ([Robinson et al., 2010](#)) using the negative  
774 binomial generalized log-linear model to the read counts for each gene (gmLRT) and  
775 decideTestDGE to determine differential genes. For the clustering analysis cf. Clustering  
776 method section.

777

## 778 **Clustering**

779 Genes similarly expressed in all conditions (difference of reads < 50) were filtered-off. For each  
780 condition, the mean RNAseq count for the replicates was calculated, centered and reduced.  
781 We used two different algorithms for clustering using stats() package of R ([Wiwie et al., 2015](#)).  
782 The function hclust() for the hierarchical ascendant clustering was used with the Spearman  
783 distance and the Ward.D2 option. The function kmeans() was used for k-means clustering with  
784 default options. k-means clustering was repeated 15 times, 5 times with predefined centroids  
785 -corresponding to genes of the different clusters defined by hierarchical clustering- and 10 with  
786 random centroids. The different k-means clusters obtained by each iteration were compared  
787 with cluster\_similarity (Clusteval (version 0.1)). We defined that the clustering is similar if the  
788 score equals 1.

789 For the S2 cells: We used for this analysis the RNAseq data from the S2 cells lines. We  
790 obtained after filtration a matrix of 3556 differential genes.

791 For the ISC data: We used for this analysis the RNAseq data from the ISC cells obtained by  
792 FACS method. We obtained after filtration a matrix of 817 differential genes. We determined  
793 by hierarchical ascendant clustering 5 clusters and use this parameter for the k-means  
794 clustering k = 5.

795 For the EB data: We used for this analysis the RNAseq data from the EB cells obtained by  
796 FACS method. We obtained after filtration a matrix of 1461 differential genes. We determined  
797 by hierarchical ascendant clustering 5 clusters and use this parameter for the k-means  
798 clustering k = 5.

799 For the eEC data: We used for this analysis the RNAseq data from the eEC cells obtained by  
800 FACS method. We obtained after filtration a matrix of 1694 differential genes. We determined  
801 by hierarchical ascendant clustering 5 clusters and use this parameter for the k-means  
802 clustering k = 5.

803 For the mEC data: We obtained after filtration a matrix of 1186 differential genes. We  
804 determined by hierarchical ascendant clustering 4 clusters and use this parameter for the k-  
805 means clustering k =4.

806 To determine the most robust group of genes we define a score of robustness or r-score. This  
807 r-score is calculated for each comparison between clusters of k-means and hierarchical  
808 algorithms. First, we define a percent of similarity between k-means clustering and hierarchical

809 clustering such as:  $\frac{(\text{number of common gene})}{(\text{number of gene in kmeans cluster} + \text{number of}$   
810  $\text{gene in hirarchical cluster}) - \text{number of common gene}} \times 100$

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

818

### 819 **Catada Analysis**

820

821 CATaDa chromatin accessibility data from ISCs, EBs, and ECs were obtained from ([Aughey](#)  
822 [et al., 2018](#)). Per-GATC fragment counts were RPM normalised and visualised in IGV. GFF  
823 signal files were converted to bigwig format using kentUtils bedGraphToBigWig and profiles of  
824 RPM chromatin accessibility scores across transcriptional start sites were plotted using  
825 deeptools computeMatrix and plotProfile functions. Chromatin accessibility at individual  
826 promoter regions for global comparison were extracted using 2kb regions centered on the TSS,  
827 means were compared using Kruskal-wallis test with Dunn's multiple comparison.

828

### 829 **HiC Analysis**

830

831 Hi-C analysis at high resolution (1kb) was performed for cells expressing either Svb-REP or  
832 Svb-ACT (GSE221863) and performed as previously described ([Heurteau et al., 2020](#)) using  
833 pipelines available through Github [https://github.com/ CuvierLab](https://github.com/CuvierLab)/Hi-C data for analysis of 3D  
834 genomic interactions in Drosophila S2 cells. Aggregation analysis was performed as previously  
835 in 1D/2D plots ([Liang et al., 2014](#); [Rao et al., 2014](#)) to compare genomic contacts in cells  
836 expressing either Svb-REP or Svb-ACT. Hi-C data were aggregated onto pairs of enhancers  
837 identified by STARR-seq ([Arnold et al., 2013](#); [Zabidi et al., 2015](#)) and distant target  
838 transcription start sites. Enhancer-promoter pairs were then ranked depending on net  
839 variations of 3D contacts between Svb-REP and Svb-ACT and the ranked quantiles of pairs  
840 were assessed for enrichment depending on presence or absence of Svb binding or not using  
841 a Fisher exact test.

842

843

844

## 845 **References for Materials and Methods**

- 846 Anders, S., Pyl, P.T. and Huber, W. 2015. HTSeq--a Python framework to work with high-  
847 throughput sequencing data. *Bioinforma. Oxf. Engl.*, **31**: 166–169.
- 848 Arnold, C.D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M. and Stark, A. 2013. Genome-wide  
849 quantitative enhancer activity maps identified by STARR-seq. *Science*, **339**: 1074–7.
- 850 Aughey, G.N., Estacio Gomez, A., Thomson, J., Yin, H. and Southall, T.D. 2018. CATaDa  
851 reveals global remodelling of chromatin accessibility during stem cell differentiation in  
852 vivo. *eLife*, **7**: e32341.
- 853 Carroll, T.S., Liang, Z., Salama, R., Stark, R. and de Santiago, I. 2014. Impact of artifact  
854 removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. *Front. Genet.*, **5**: 75.
- 855 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,  
856 M. and Gingeras, T.R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinforma.*  
857 *Oxf. Engl.*, **29**: 15–21.
- 858 Feng, J., Liu, T. and Zhang, Y. 2011. Using MACS to identify peaks from ChIP-Seq data. *Curr.*  
859 *Protoc. Bioinforma.*, **Chapter 2**: Unit 2.14.
- 860 Heurteau, A., Perrois, C., Depierre, D., Fosseprez, O., Humbert, J., Schaak, S. and Cuvier, O.  
861 2020. Insulator-based loops mediate the spreading of H3K27me3 over distant micro-  
862 domains repressing euchromatin genes. *Genome Biol.*, **21**: 193.
- 863 Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T.  
864 and Carey, V.J. 2013. Software for computing and annotating genomic ranges. *PLoS*  
865 *Comput. Biol.*, **9**: e1003118.
- 866 Lhoumaud, P., Hennion, M., Gamot, A., Cuddapah, S., Queille, S., Liang, J., Micas, G.,  
867 Morillon, P., Urbach, S., Bouchez, O., Severac, D., Emberly, E., Zhao, K. and Cuvier,  
868 O. 2014. Insulators recruit histone methyltransferase dMes4 to regulate chromatin of  
869 flanking genes. *EMBO J.*, **33**: 1599–1613.
- 870 Liang, J., Lacroix, L., Gamot, A., Cuddapah, S., Queille, S., Lhoumaud, P., Lepetit, P., Martin,  
871 P.G.P., Vogelmann, J., Court, F., Hennion, M., Micas, G., Urbach, S., Bouchez, O.,  
872 Nöllmann, M., Zhao, K., Emberly, E. and Cuvier, O. 2014. Chromatin  
873 immunoprecipitation indirect peaks highlight long-range interactions of insulator  
874 proteins and Pol II pausing. *Mol. Cell*, **53**: 672–681.
- 875 Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J.A., Delerce,  
876 J., Jaeger, S., Blanchet, C., Vincens, P., Caron, C., Staines, D.M., Contreras-Moreira,  
877 B., Artufel, M., Charbonnier-Khamvongsa, L., Hernandez, C., Thieffry, D., Thomas-  
878 Chollier, M. and van Helden, J. 2015. RSAT 2015: Regulatory Sequence Analysis  
879 Tools. *Nucleic Acids Res*, **43**: W50-6.
- 880 Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S.,  
881 Dündar, F. and Manke, T. 2016. deepTools2: a next generation web server for deep-  
882 sequencing data analysis. *Nucleic Acids Res.*, **44**: W160-165.
- 883 Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T.,  
884 Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. and Aiden, E.L. 2014. A 3D Map  
885 of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin  
886 Looping. *Cell*, **159**: 1665–1680.
- 887 Robinson, M.D., McCarthy, D.J. and Smyth, G.K. 2010. edgeR: a Bioconductor package for  
888 differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.*,  
889 **26**: 139–140.
- 890 Sievert, C. n.d. *Interactive web-based data visualization with R, plotly, and shiny*.
- 891 Su, S., Law, C.W., Ah-Cann, C., Asselin-Labat, M.-L., Blewitt, M.E. and Ritchie, M.E. 2017.  
892 Glimma: interactive graphics for gene expression analysis. *Bioinforma. Oxf. Engl.*, **33**:  
893 2050–2052.
- 894 Thomas-Chollier, M., Sand, O., Turatsinze, J.V., Janky, R., Defrance, M., Vervisch, E., Brohee,  
895 S. and van Helden, J. 2008. RSAT: regulatory sequence analysis tools. *Nucleic Acids*  
896 *Res*, **36**: W119-27.
- 897 Wang, M., Zhao, Y. and Zhang, B. 2015. Efficient Test and Visualization of Multi-Set  
898 Intersections. *Sci Rep*, **5**: 16923.

- 899 Wiwie, C., Baumbach, J. and Röttger, R. 2015. Comparing the performance of biomedical  
900 clustering methods. *Nat. Methods*, **12**: 1033–1038.
- 901 Zabidi, M.A., Arnold, C.D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A.  
902 2015. Enhancer-core-promoter specificity separates developmental and housekeeping  
903 gene regulation. *Nature*, **518**: 556–9.
- 904 Zanet, J., Benrabah, E., Li, T., Pélissier-Monier, A., Chanut-Delalande, H., Ronsin, B., Bellen,  
905 H.J., Payre, F. and Plaza, S. 2015. Pri sORF peptides induce selective proteasome-  
906 mediated protein processing. *Science*, **349**: 1356–1358.
- 907
- 908
- 909