Article

SOURCE DATA | TRANSPARENT PROCESS

THE EMBO JOURNAL

The Hox proteins Ubx and AbdA collaborate with the transcription pausing factor M1BP to regulate gene transcription

Amel Zouaz[1,†], Ankush Auradkar[1,†], Marie Claire Delfini[1,†], Meiggie Macchi[1], Marine Barthez[1], Serge Ela Akoa[1], Leila Bastianelli[2], Gengqiang Xie[3], Wu-Min Deng[3], Stuart S Levine[4], Yacine Graba[1] & Andrew J Saurin[1,*]

Abstract

In metazoans, the pausing of RNA polymerase II at the promoter (paused Pol II) has emerged as a widespread and conserved mechanism in the regulation of gene transcription. While critical in recruiting Pol II to the promoter, the role transcription factors play in transitioning paused Pol II into productive Pol II is, however, little known. By studying how *Drosophila* Hox transcription factors control transcription, we uncovered a molecular mechanism that increases productive transcription. We found that the Hox proteins AbdA and Ubx target gene promoters previously bound by the transcription pausing factor M1BP, containing paused Pol II and enriched with promoter-proximal Polycomb Group (PcG) proteins, yet lacking the classical H3K27me3 PcG signature. We found that AbdA binding to M1BP-regulated genes results in reduction in PcG binding, the release of paused Pol II, increases in promoter H3K4me3 histone marks and increased gene transcription. Linking transcription factors, PcG proteins and paused Pol II states, these data identify a two-step mechanism of Hox-driven transcription, with M1BP binding leading to Pol II recruitment followed by AbdA targeting, which results in a change in the chromatin landscape and enhanced transcription.

Keywords Hox; M1BP; promoter-proximal pausing; RNA polymerase II; transcription

Subject Categories Chromatin, Epigenetics, Genomics & Functional Genomics; Development & Differentiation; Transcription

DOI 10.15252/embj.201695751 | Received 9 December 2016 | Revised 2 August 2017 | Accepted 7 August 2017 | Published online 4 September 2017

The EMBO Journal (2017) 36: 2887–2906

Introduction

Hox proteins are essential transcription factors that specify the body plan of all bilaterians and regulate the transcription of a wide spectrum of functionally diverse target genes (for recent review, see Rezsohazy *et al*, 2015). Hox proteins bind to a general core TAAT DNA consensus sequence, with increased specificity brought through association with cofactors, the most studied of which are the PBC cofactors (Abu-Shaar *et al*, 1999; Mann *et al*, 2009; Cherbas *et al*, 2011). From very early studies demonstrating an intricate relationship with the transcription machinery (Johnson & Krasnow, 1990, 1992), to more recent discoveries of contacts with components of TFIID and Mediator complexes (Prince *et al*, 2008; Boube *et al*, 2014), Hox proteins clearly regulate transcription through contact with the general transcription machinery. Yet the molecular mechanisms underlying Hox-driven transcriptional regulation have remained elusive.

While transcriptional regulation has long been considered as predominantly regulated at the level of the assembly of a transcription pre-initiation complex (PIC), more recently genome-wide studies of RNA polymerase II (Pol II) in *Drosophila* and vertebrates have shown that Pol II PIC formation occurs at the vast majority of genes, including silent or poorly transcribed genes. Such genes display an accumulation of Pol II just downstream of the transcription start site (TSS) in a transcriptionally engaged, but paused, state (Adelman & Lis, 2012). It is believed that pre-loading of Pol II in this way (generically termed "poised Pol II" for the purposes of discussion) allows not only for the rapid induction of stimulus- and developmentally regulated genes, but also for fine tuning basal gene expression levels. While the role of transcription factors in PIC formation has been extensively studied, their role in transitioning poised into productive Pol II in metazoans has received much less attention. The only well-described mechanisms consist of the recruitment of transcription elongation factors, as shown for

1 Aix Marseille Université, CNRS, IBDM, UMR 7288, Marseille, France
2 MGX-Montpellier GenomiX c/o Institut de Génomique Fonctionnelle, Montpellier, France
3 Department of Biological Science, Florida State University, Tallahassee, FL, USA
4 BioMicro Center, Massachusetts Institute of Technology, Cambridge, MA, USA
*Corresponding author. Tel: +33 (0)4 91 26 93 52; E-mail: andrew.saurin@univ-amu.fr
†These authors contributed equally to this work

the transcription factors c-myc and NF-κB that recruit the pause release factor P-TEFb (Barboric *et al*, 2001; Eberhardy & Farnham, 2002; Rahl *et al*, 2010). However, the Hox-dependent release of poised Pol II at the Hoxb1a promoter is independent of P-TEFb recruitment (Choe *et al*, 2014), highlighting that other, yet unidentified mechanisms must exist.

Studying the molecular mechanisms giving rise to productive transcription in a developing organism is fraught with difficulties. The heterogeneous and differentiating population of cells in a developing organism is not conducive to studying Hox-driven transcription at the molecular level, given the myriad of genes they regulate in conditions employing different cofactors. Numerous *in vitro* studies into Hox action have enlightened how Hox proteins increase target recognition specificity through studies on PBC-class cofactors, although these studies do not fully address the many PBC-independent Hox mechanisms that exist (reviewed in Rezsohazy *et al*, 2015). The use of cell lines, particularly S2 and Kc167 cells, the most commonly used cell lines in *Drosophila*, to study the mechanisms of Hox-driven transcription has numerous benefits: (i) they are a homogenous cell system allowing for direct interpretation of transcriptional events; (ii) they are devoid of endogenous Hox gene expression (Cherbas *et al*, 2011) allowing for conditional Hox expression to directly study Hox-dependent transcriptional events; (iii) they have been extensively profiled, particularly by the modENCODE consortium for the genomic distribution of chromatin proteins and histone modification marks (for example, see Roy *et al*, 2010), providing an invaluable data resource for linking the binding profiles of Hox to those of other actors of transcription. Indeed, use of Kc167 cells to study the genomic binding profiles of three *Drosophila* Hox proteins, Ultrabithorax (Ubx), Abdominal-A (AbdA) and Abdominal B, demonstrated that Hox proteins bind to thousands of specific genomic loci in the absence of Extradenticle (Exd; the sole *Drosophila* PBC protein) (Beh *et al*, 2016). Moreover, additional Hox-targeted loci are observed in the presence of Exd, demonstrating the plasticity of Hox binding site selectivity depending on cofactor used (Beh *et al*, 2016). The specific and widespread binding of Hox in the absence of Exd suggests that these sites either represent the intrinsic monomeric specificity of Hox proteins, the specificity achieved through cooperation with other endogenous cofactors, or both monomeric and cofactor-specific sites.

In addition to understanding the bases underlying Hox DNA binding specificity, the need to understand Hox protein activity is also required to clarify how they influence the transcription process. To better understand how Hox proteins regulate transcription at the mechanistic level, we reasoned that Hox genomic binding in a cell-based system in the absence of PBC cofactor (Exd) could be exploited to both delineate the molecular mechanisms behind PBC-independent Hox-mediated transcription and identify novel Hox cofactors, whose function may not necessarily relate to improved DNA binding specificity. To this end, we employed a genome-wide approach, using the *Drosophila* Hox proteins, Ubx and AbdA, and *Drosophila* S2-DRSC cells (herein termed "S2 cells") as model. Using this approach, we demonstrate that the Hox proteins bind M1BP-targeted promoter-proximal regions resulting in reduced Polycomb Group (PcG) binding, a reduction in poised Pol II and increased gene expression. Together with identifying physical interaction between M1BP and Hox proteins in embryos and demonstrating that M1BP functions in Hox-controlled processes *in vivo*, our data identify M1BP as a novel cofactor in PBC-independent Hox processes.

Results

Hox proteins target promoter-proximal regions enriched in GAGA- and M1BP-binding motifs

HA-tagged forms of *abdA* or *Ubx* phenocopy wild-type *abdA* or *Ubx* in *Drosophila* larvae, indicating that the HA tag does not interfere with normal AbdA or Ubx function (Saadaoui *et al*, 2015 and Appendix Fig S1). We thus generated stable S2 cell lines conditionally expressing HA-tagged forms of AbdA or Ubx and performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). We found that almost half of all AbdA and Ubx binding sites associate with promoter-proximal regions, with most of the remaining peaks locating to putative distal enhancers (Table EV1 and Fig 1A). These distributions of peak genomic locations are highly comparable to those of AbdA and Ubx when transiently expressed in Kc167 cells (Beh *et al*, 2016) and endogenous Ubx in the embryo (Shlyueva *et al*, 2016) (Table EV1 and Fig 1A). Promoters and enhancers are genomic locations known to have an open chromatin conformation when analysed by ATAC-seq, DNase-seq and FAIRE-seq (Song *et al*, 2011; Buenrostro *et al*, 2013). Using DNase-seq S2 data from Kharchenko *et al* (2011) to identify open chromatin regions in S2 cells, we find that half or more promoter and enhancer-located peaks are in open chromatin regions (Appendix Fig S2), which helps correlate our findings with those of Beh and colleagues, who reported that AbdA and Ubx target open chromatin regions in Kc167 cells (Beh *et al*, 2016). When studying the distribution of AbdA and Ubx over the promoter-proximal region in greater detail, we observed high enrichment and a highly similar distribution of AbdA and Ubx binding just upstream of TSSs, irrespective of the source material (Fig 1B). When comparing the promoters targeted by either Hox, we found that AbdA targets 6,011 (4,456 non-overlapping) promoters in both S2 and Kc167 cells and Ubx targets 2,748 (1,973 non-overlapping) promoters in S2, Kc167 and embryos (Fig EV1A). Lack of an AbdA ChIP dataset in embryos precludes comparing common AbdA/Ubx promoters in embryos, but when comparing common promoters when expressed in the same cell system, we find that AbdA and Ubx target 7,338 (7,179 non-overlapping) common promoters in S2 cells and 5,511 common promoters (5,384 non-overlapping) in Kc167 cells (Fig EV1B), of which 4,216 are common to AbdA and Ubx in both S2 and Kc167 cells.

Hox proteins bind to a highly similar core TAAT DNA consensus sequence, with increased specificity brought through association with cofactors, the most studied of which are the PBC cofactors (Exd in *Drosophila*) (Mann *et al*, 2009; Merabet & Mann, 2016). They also often act together with transcription factors, referred to as collaborators (Mann *et al*, 2009) that bind close to Hox proteins but without demonstrated influence on Hox binding (Stobe *et al*, 2009). S2 and Kc167 cells are devoid of nuclear Exd due to the absence of *homothorax* (*hth*) expression (Abu-Shaar *et al*, 1999; Cherbas *et al*, 2011; Beh *et al*, 2016) and so binding by AbdA and Ubx to so many common promoters suggests either AbdA and Ubx is capable of binding as monomers, or binding is achieved through interaction

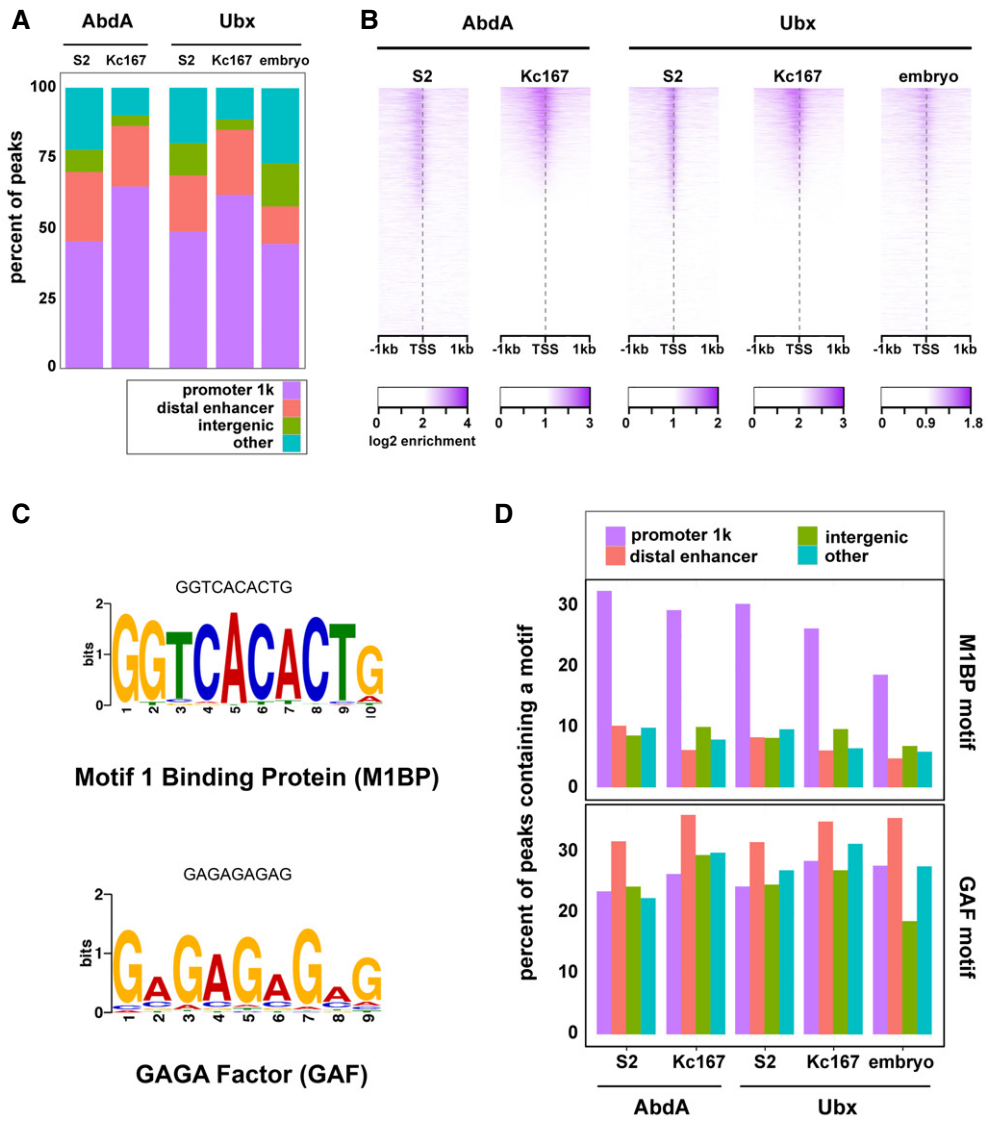

Figure 1. AbdA and Ubx are highly enriched at promoter-proximal regions that contain M1BP and GAF binding motifs.

A Comparison of the localisation of peaks called with unique genomic regions of AbdA (S2 and Kc167 cells) and Ubx (S2 cells, Kc167 cells and embryos).

B Heatmap representation of the enrichment of AbdA and Ubx upstream of RefSeq transcription start sites (TSS) in S2 and Kc167 cells (AbdA and Ubx) and embryos (Ubx). Each row in the heatmap represents a RefSeq gene, clustered with respect to total enrichment of AbdA in Kc167 cells.

C Two highly significant motif weblogos of position weight matrices discovered during *de novo* motif discovery on AbdA and Ubx peaks in S2 are shown. The DNA consensus sequence is provided above the weblogo. Motif comparison with known transcription factor binding motifs identified the motifs enriched in AbdA/Ubx peaks as Motif 1 Binding Protein (M1BP) and GAGA factor (GAF).

D The enrichment of M1BP and GAF position weight matrices in AbdA and Ubx peaks found at the given genomic regions is presented as a percentage of peaks in that genomic region.

Data information: ChIP data from S2 cells are from this study, from Kc167 cells are from Beh *et al* (2016) (GSE69796) and from embryo are from Shlyueva *et al* (2016) (GSE64284).

Source data are available online for this figure.

with other common cofactors/collaborators (to which we further refer for simplicity in this study as cofactor), or a combination of both. To study whether AbdA and Ubx binding in S2 cells could be influenced by potential novel common cofactors, we performed *de novo* DNA motif discovery on all AbdA and Ubx peaks in S2 cells to identify common DNA motifs other than the core TAAT Hox motif. The two highest scoring and significant motifs, common to both AbdA and Ubx peaks, were motifs targeted by M1BP and GAGA factor (encoded by the *trithorax-like* (*trl*) gene and herein termed GAF) (Fig 1C). Searching for motif occurrences showed that both GAF and M1BP motifs were also highly enriched in peaks of AbdA and Ubx from Kc167 cells and Ubx peaks from embryos (Fig 1D). Interestingly, while the GAF motif was generally evenly distributed amongst the peaks found at the different genomic locations, albeit

with a slight but consistent bias for enhancer-located peaks, the M1BP motif was most enriched at promoter-based peaks of AbdA and Ubx (Fig 1D).

Given the high degree of similarity of promoter targeting by *Drosophila* Hox proteins, we thus decided to study the importance and consequences of promoter binding in greater detail, using the HA-AbdA-expressing S2 cell line (herein termed S2-AbdA) as model.

AbdA binding to M1BP target genes results in reduced poised Pol II and increased productive transcription

The finding of GAF and M1BP binding motifs enriched in AbdA and Ubx binding sites, particularly at promoters, is of interest since both GAF and M1BP bind promoter-proximal regions in S2 cells to establish distinct poised Pol II states, with GAF dictating high levels of poised Pol II and M1BP fixing a higher variation in degrees of poised Pol II (Lee *et al*, 2008; Li & Gilmour, 2013; Fuda *et al*, 2015). We thus focussed on the binding of AbdA in S2-AbdA cells to study whether GAF and M1BP functions at promoters could be linked to Hox promoter targeting.

Performing ChIP-seq of GAF, M1BP and the Rpb3 subunit of Pol II for determining the poised Pol II state (Core *et al*, 2012) in S2 cells and comparing their binding sites with that of AbdA in S2-AbdA cells showed association of all factors with AbdA, a large number of which map to promoter-proximal AbdA binding sites (Fig 2). The promoter-proximal association of AbdA with either GAF or M1BP (Fig 2A) is statistically significant ($P < 0.0001$) when tested using the Genomic Association Tester (Heger *et al*, 2013). However, given that the majority of *Drosophila* promoters contain poised Pol II (Muse *et al*, 2007; Zeitlinger *et al*, 2007; Lee *et al*, 2008) and these poised promoters primarily contain GAF and/or M1BP (Li & Gilmour, 2013), it is likely that any transcription factor that targets mainly promoters would associate with GAF and/or M1BP. If these promoters were primarily the same as those targeted by AbdA, this would leave the biological significance of the promoter-proximal association of AbdA with either GAF or M1BP in question. To test this, we took public S2 cell ChIP datasets for the sequence-specific transcription factors *yorkie* (yki: GSE46305), *deformed wings* (dwg: GSE32853) and *pleiohomeotic* (pho: GSE32857) for which the majority of peaks are found at promoters and associating with GAF and M1BP peaks (Table EV2). Of the 2,185 M1BP-specific AbdA-targeted promoters, 21% (465) contain yki, 14% (316) contain dwg, 19% (405) contain pho and 4% (81) contain all five transcriptions factors. Of the 1,000 GAF-specific AbdA-targeted promoters, 56% (564) contain yki, 35% (353) contain dwg, 28% (283) contain pho and 12% (123) contain all five transcriptions factors. These data suggest that the statistically significant GAF/M1BP promoters targeted by AbdA may be of biologically significance since they do not appear largely common to any promoter-targeting sequence-specific transcription factor, particularly for the M1BP-targeted AbdA promoters.

Hox proteins regulate the expression of a number of important developmental genes, although they additionally regulate the expression of genes responsible for more basic cellular processes such as cell adhesion, division, death and migration (reviewed in Pearson *et al*, 2005; Rezsohazy *et al*, 2015). GAF controls mainly development and morphogenic genes, whereas M1BP controls genes mainly involved in basic cellular processes (Li & Gilmour, 2013). Gene ontology analyses showed that the genes whose promoters are targeted by AbdA that are involved mainly in developmental processes are principally GAF-controlled promoters, whereas genes targeted by AbdA involved in basic cellular processes are M1BP-controlled promoters (Fig 2C).

To study whether the binding of AbdA at promoters affects the binding of GAF, M1BP, or Pol II in S2-AbdA cells, we additionally performed ChIP-seq on all factors in S2-AbdA cells. In addition to studying poised Pol II at the promoter, we additionally investigated the state of elongating Pol II through ChIP-seq of Pol II phosphorylated at serine 2 of the C-terminal domain (Pol II phospho Ser-2) in S2 and S2-AbdA cells (Fig 3A). Quantification of enrichment of AbdA, GAF and M1BP at promoter-proximal regions, Rpb3 in the poised Pol II region and Pol II phospho Ser-2 in the gene body at genes displaying poised Pol II in S2 and S2-AbdA cells showed a reduction in poised Pol II and an increase in elongating Pol II upon AbdA binding (Fig 3A and B), but little-to-no change in GAF and M1BP binding (Fig EV2). While GAF and M1BP targets are generally distinct, they do share some common promoter-proximal targets (Li & Gilmour, 2013; Fuda *et al*, 2015). Hence, we separated genes into three distinct classes based on the presence of high confidence peaks within 150 bp upstream of a poised Pol II gene in S2 cells of only GAF, only M1BP, or containing both GAF and M1BP peaks. Comparing binding in S2-AbdA cells versus S2 cells, we found that while the binding of GAF and M1BP remained largely unchanged (Fig EV2C), we found less poised Pol II and more elongating Pol II at a large number of M1BP-specific targets upon AbdA binding (Fig 3C and D).

To determine whether the changes in poised Pol II status and increased elongating Pol II reflect changes in gene transcription, we first performed RNA-seq analyses on S2 and S2-AbdA cells and found 1,982 differentially expressed (DE) genes upon AbdA expression (Dataset EV1). We called a gene as being directly regulated by AbdA in S2 cells if the promoter of the DE gene was targeted by AbdA, resulting in 891 genes directly regulated by AbdA (Dataset EV1), displaying an average of 2.6-fold upregulated and 3.7-fold downregulated (Appendix Fig S3). AbdA is reported to regulate some 964 genes when ubiquitously expressed in the embryo (Hueber *et al*, 2007), of which 10% (98) are shown here directly regulated in S2 cells (Dataset EV1), suggesting that the direct regulation of genes by AbdA through promoter-proximal binding is at least in part observable in the embryo. Testing some of the genes that are differentially expressed by AbdA in S2 cells but not found as regulated in the embryo by Hueber *et al* (2007), we found that most of the genes tested (19/20) showed differential expression upon ubiquitous AbdA expression (for examples, see Fig 4A and B), demonstrating that the list of 964 genes regulated by AbdA in the embryo (Hueber *et al*, 2007) is a conservative estimate.

Of the 891 genes directly regulated by AbdA in S2 cells, 384 contained a promoter GAF peak and 273 contained a promoter M1BP peak. Separating the list of directly regulated genes into those genes containing only GAF or M1BP promoter peaks ($n = 459$) to remove genes containing both GAF and M1BP and studying changes in poised and elongating Pol II, we observed that at GAF target genes, the majority of genes tend to show reduced expression and less elongating Pol II upon AbdA expression but without affecting promoter Pol II levels (Fig 4C), suggesting that the observed repression is not through changes in Pol II pausing. However, at M1BP target genes

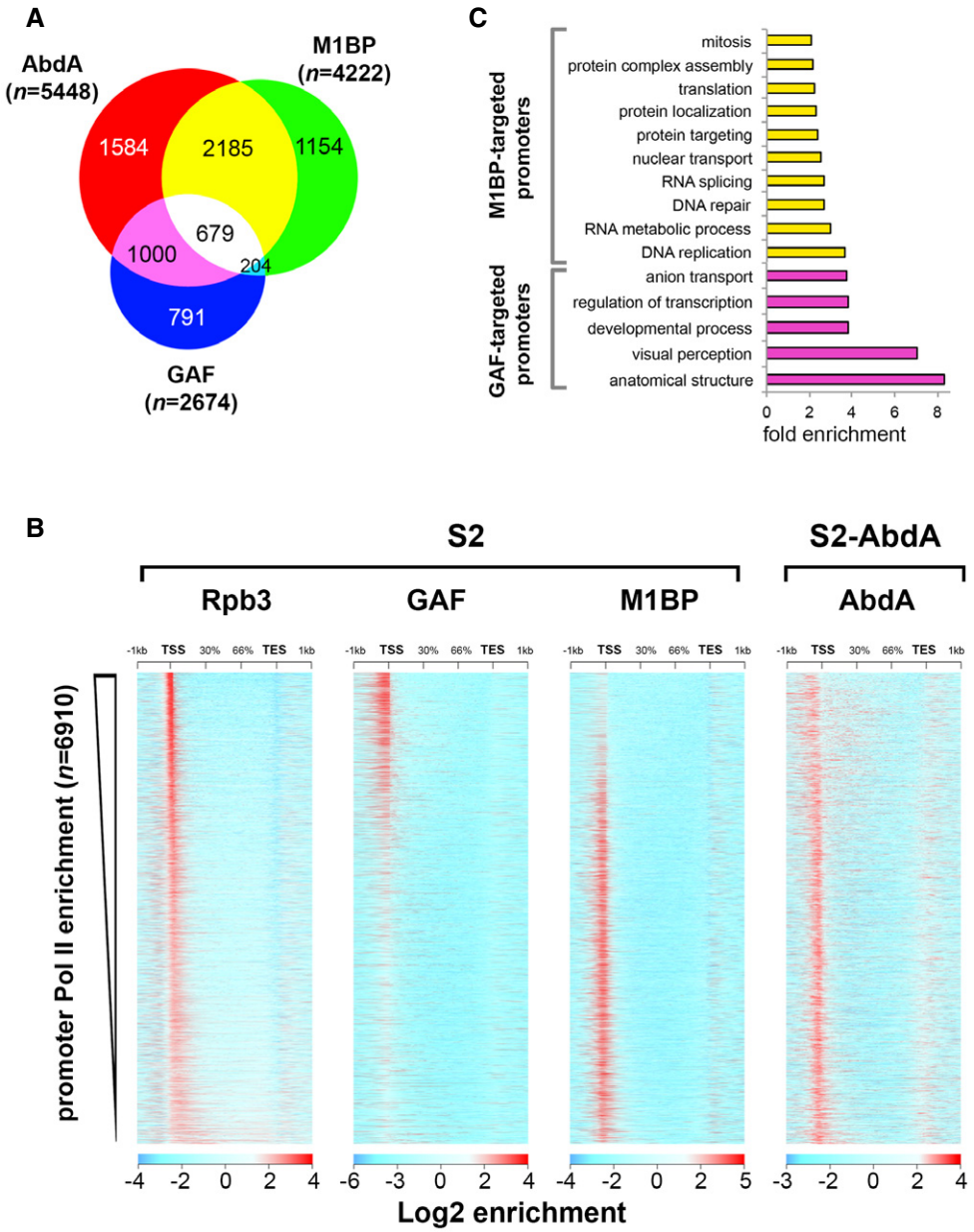

Figure 2. AbdA targets promoter-proximal regions pre-bound by GAF and M1BP.

A Venn diagram representing the overlap between promoter-proximal AbdA, M1BP and GAF peaks.

B RefSeq transcripts enriched for promoter Pol II (6,910 genes) were clustered according to total promoter Pol II enrichment levels and heatmaps of log2 ChIP enrichment for promoter Pol II (Rpb3), GAF and M1BP in S2 cells and AbdA in S2-AbdA cells are shown.

C Fold enrichment of gene ontology terms of genes targeted by AbdA at M1BP and GAF promoters.

Source data are available online for this figure.

Pol II pausing is released, shown by reduced promoter Pol II, increased elongating Pol II and increased gene expression (Fig 4C).

Together, these data demonstrate that while AbdA targets both GAF- and M1BP-controlled poised Pol II genes, it is mainly at the M1BP targets that there is reduction in poised Pol II status and increased productive transcription, suggesting a functional cooperation between AbdA and M1BP in regulating gene expression, which we sought to examine further.

M1BP interacts and cooperates with Hox proteins in *Drosophila*

To test for functional cooperation between AbdA and M1BP in regulating gene expression, we took advantage of a GFP reporter construct containing multiple binding sites for M1BP (Li *et al*, 2010). While the region in and around the M1BP binding sites does not contain a canonical "TAAT" Hox binding site sequence, it is rich in AA, TT and TA dinucleotides (82% AA, 66% TT, 60% TA) and is

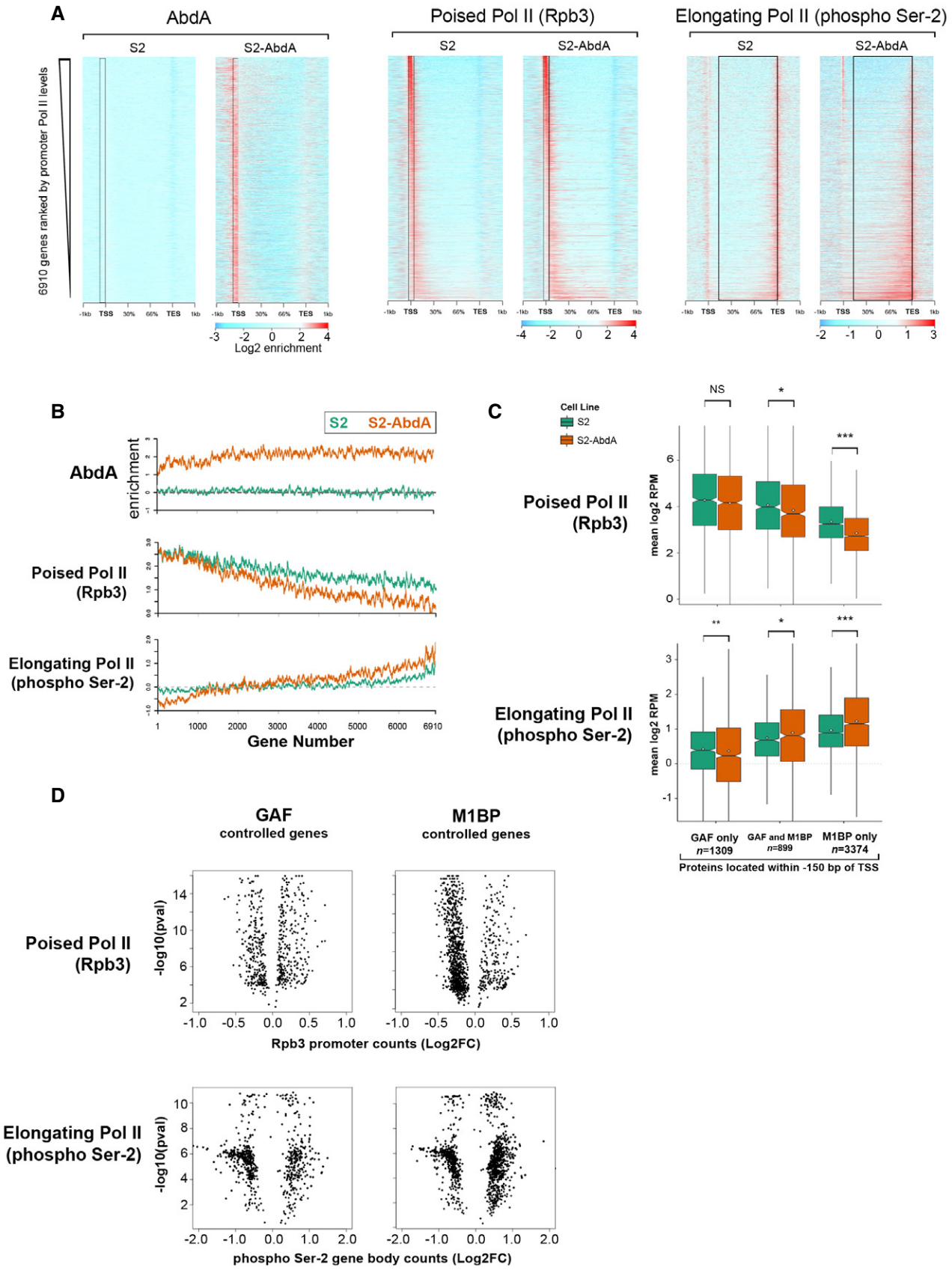

Figure 3.

Figure 3. Promoter-proximal binding of AbdA to M1BP target genes leads to a reduction in poised Pol II and increased elongating Pol II.

A RefSeq transcripts enriched for promoter Pol II (6,910 genes) were clustered according to total promoter Pol II enrichment levels and heatmaps of log2 ChIP enrichment values spanning from 1 kb upstream of the TSS to 1 kb downstream of the TES are shown for ChIPs of the HA tag (AbdA), the Rpb3 subunit of Pol II (poised Pol II) and Pol II phospho Ser-2 (elongating Pol II) in S2 and S2-AbdA cells. Boxed areas highlight the regions used to quantitate mean enrichment values in (B).

B Average profile plots of log2 enrichment at the poised Pol II region (Rpb3), the gene body for elongating Pol II (Pol II phospho Ser-2) and promoter-proximal regions (AbdA) for ChIPs on S2 (green) and S2-AbdA (brown) cells showing reduced promoter Pol II and increased elongating Pol II. The dotted line represents no ChIP enrichment.

C Boxplot comparison of poised Pol II (Rpb3) and elongating Pol II (Pol II phospho Ser-2) mean log2 RPMs data from (B) at transcripts from S2 and S2-AbdA cells containing either a GAF peak, an M1BP peak or both GAF and M1BP within 150 bp upstream of the TSS. Box limits represent the interquartile range with whiskers representing the first and fourth quartiles. Notches represent the 95% confidence interval of the median (black bar) and mean values represented by white dots. *P*-values determined from permutation tests: $*P < 0.05$; $**P < 0.001$; $***P < 0.0001$; NS > 0.05.

D Volcano plots of significant poised Pol II and elongating Pol II differential binding upon AbdA expression determined using diffReps at the poised Pol II region (Rpb3) or gene body (Pol II phospho Ser-2) of transcripts containing peaks of either GAF or M1BP within 150 bp upstream of the TSS showing more reduced poised Pol II and increased elongating Pol II at M1BP-controlled genes upon AbdA binding.

Source data are available online for this figure.

flanked by two TAAAT sequences that can be found enriched at AbdA sites in Kc167 cells (Beh *et al*, 2016) and bind AbdA *in vitro* (Slattery *et al*, 2011) (Fig 5A). Transfection of the GFP reporter into S2 cells results in GFP fluorescence that is largely eliminated upon M1BP RNAi, demonstrating the dependence on M1BP binding (Fig 5A). The intensity of GFP fluorescence observed is increased by approximately fourfold upon transfection of the reporter into S2-AbdA cells, which is co-dependent on the presence of both AbdA and M1BP, demonstrating that the presence of both AbdA and M1BP is required to achieve maximal GFP fluorescence (Fig 5A). These data suggest that the two proteins may interact in a common protein complex. Indeed, immunoprecipitation of M1BP in S2-AbdA cells results in AbdA co-precipitation (Fig 5B). This interaction is conserved in the embryo where immunoprecipitating either AbdA or Ubx from embryo extracts co-precipitates M1BP (Fig 5B). To directly visualise this interaction on chromatin *in vivo*, we expressed AbdA in the salivary gland using the *sgs*Gal4 driver. Immunostaining for M1BP on salivary gland polytene chromosomes shows a distinct binding pattern representing the large number of strong chromatin binding sites of M1BP, a large number of which show colocalisation with AbdA (Fig 5C). Taken together, these data suggest that the co-dependency for maximal reporter gene transcription in S2 cells is likely to pass in part via protein–protein interaction and that this interaction is conserved *in vivo*.

As both AbdA and Ubx target M1BP-controlled promoters (Fig 1D) and both interact with M1BP in embryos (Fig 5B), we were interested in determining whether M1BP could function in Hox-controlled processes *in vivo*. Studying M1BP–Hox cooperation in a physiological context is less trivial, since the only study to date into M1BP function was performed in S2 cells (Li & Gilmour, 2013). However, using the knowledge that genes containing the M1BP binding motif respond negatively to starvation in *Drosophila* larvae (Li *et al*, 2010), a context that induces autophagy in the fat body, we hypothesised that M1BP might play a central role with Hox proteins in their role in controlling autophagy (Banreti *et al*, 2014). Unlike all Hox proteins, whose clearance is required to permit autophagy progression at the third-instar larval feeding-to-wandering transition (Banreti *et al*, 2014), M1BP expression in fat body cells is maintained at all stages of third-instar larval development (Fig EV3A). While forced expression of M1BP could not inhibit developmental autophagy in wandering larvae (Fig EV3B), using the Atg8a marker that marks cytoplasmic autophagosomes as a marker of active autophagy (Klionsky *et al*, 2011), we found that upon RNAi knockdown of M1BP in feeding larvae (Fig EV3C), either through spatially controlled RNAi using the *cg*-Gal4 driver (Fig 6A) or clonally (Fig 6B), rapid and premature induction of autophagy occurred (Fig 6).

Forced loss of expression of all Hox genes in feeding larvae is required to induce premature autophagy (Banreti *et al*, 2014). Autophagy induction upon M1BP knockdown is not due to the loss of Hox expression, since all of the tested Hox expression profiles were unaffected (Fig 6C). Moreover, anti-Ubx staining shows that Ubx protein production is unaffected upon M1BP knockdown (Fig EV3D). Since the presence of a single Hox protein is sufficient to block developmental autophagy (Banreti *et al*, 2014), these data

Figure 4. Promoter-proximal binding of AbdA leads to changes in gene expression.

A *In situ* hybridisations of six genes identified as differentially expressed upon promoter-proximal binding of AbdA in S2-AbdA cells show differential expression in embryos upon armGal4 expression of AbdA. Embryonic regions that do not express the gene in wild-type embryos (green zones) but regions that display ectopic expression upon AbdA expression are highlighted in red.

B RT–qPCR of 10 genes identified as differentially expressed upon promoter-proximal binding of AbdA in S2-AbdA cells shows significant differential expression in embryos upon armGal4 expression of AbdA. CG1486, which is twofold overexpressed in S2-AbdA cells, showed no significant change in expression in embryos upon AbdA ectopic expression. Error bars represent standard error of the mean of three independent experiments performed in triplicate, and *P*-values were derived using a Student's *t*-test: $*P < 0.05$; $**P < 0.001$; $***P < 0.0001$.

C Contour plots showing dynamic change in poised Pol II (log2 fold change of Rpb3 RPM) or elongating Pol II (log2 fold change of Pol II phospho Ser-2) on the ordinate as a function of changes in gene expression of DE genes (log2 count fold change; abscissa) at Pol II poised genes upon AbdA expression. ChIP profiles are grouped into transcripts containing GAF (blue contours; $n = 285$) or M1BP (red contours; $n = 174$) peaks within 150 bp upstream of the DE gene's TSS. M1BP genes generally display reduced poised Pol II, increased elongating Pol II and increased gene expression upon AbdA binding, whereas GAF genes display reduced elongating Pol II and reduced expression. Contour lines group points of equal density and correspond to a sequence of equally spaced *z*-values, with increasing opacity added for regions with decreasing *z*-values.

Source data are available online for this figure.

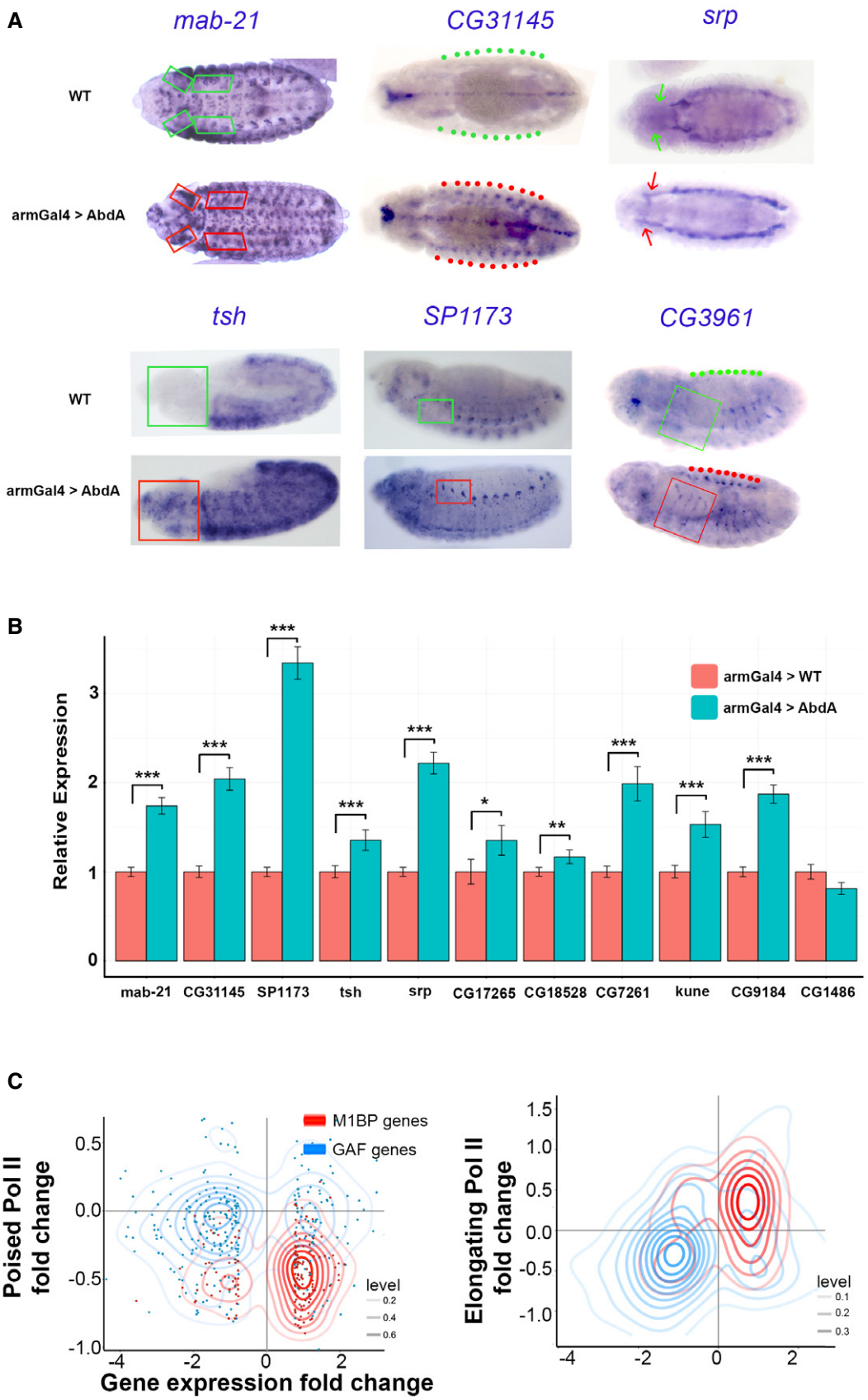

Figure 4.

Figure 5. M1BP interacts with AbdA.

A GFP fluorescence was observed upon transfection of a M1BP-GFP reporter in S2 and S2-AbdA cells showing stronger fluorescence in S2-AbdA compared to S2 cells. GFP reporter activity dependence on M1BP was demonstrated through M1BP RNAi treatment of transfected cells. Violin plots, overlaid by boxplots, show quantification of GFP fluorescence of transfected nuclei of S2 and S2-AbdA cells and RNAi-treated cells showing quantitative co-dependence on the presence of both AbdA and M1BP for maximal reporter activity. Box limits represent the interquartile range with whiskers representing the first and fourth quartiles. The median is represented with a black bar and mean values represented by white dots. Scale bars represent 20 μm.

B M1BP interacts with AbdA and Ubx. Immunoprecipitation of M1BP from S2-AbdA cells co-precipitates AbdA (HA tag) and immunoprecipitation of either AbdA or Ubx from embryo extracts co-precipitates M1BP. Lanes from no antibody (mock) and AbdA/Ubx IPs in embryos represent 14% of the input.

C AbdA colocalises with M1BP on polytene chromosomes. HA-tagged *abdA* was expressed in the salivary gland under the control of an *sgsGAL4* driver. Co-immunostaining AbdA (green) and M1BP (red) reveals bands colocalising with AbdA (white). Colocalisation detection (white) and correlation was performed in ImageJ giving a Pearson's correlation *R*-value of 0.72. A magnified area (white box in the green channel) is represented under the main image. Scale bars represent 20 μm.

Source data are available online for this figure.

represent, to our knowledge, the only example in *Drosophila* of premature autophagy induction through the loss of expression of a single protein that itself is not required for Hox expression. Since inhibition of autophagy is a generic function of all Hox proteins (Banreti *et al*, 2014), it is not possible to discern an M1BP/AbdA-specific partnership in such a physiological context. Nonetheless, these data demonstrate two opposing situations, both leading to autophagy induction: in wandering larvae, autophagy commences when Hox protein expression is lost (Banreti *et al*, 2014) while M1BP expression is maintained, whereas in feeding larvae, loss of M1BP expression induces autophagy while Hox expression is maintained (Fig 6D). These data implicate M1BP as a master repressor of autophagy—a role attributed equally to all Hox proteins—and suggest that M1BP may work with Hox proteins in their role of autophagy repression in the larval fat body.

PcG enrichment at poised Pol II promoters is reduced upon AbdA-induced expression of M1BP target genes

To study the molecular mechanisms underlying AbdA-regulated transcription in S2-AbdA cells, we analysed the publicly available binding profiles of 41 unique histone modifications and histone-modifying enzymes in S2 cells together with the binding profiles of AbdA, M1BP, GAF and Pol II. While the transcription factors and Pol II clustered strongly with a general signature of active transcription, H3K4me3 and histone-modifying enzymes linked to transcriptional activation, it was the addition of the Polycomb Group (PcG) family of histone modifiers, Sex combs extra (herein referred to as dRing) and Enhancer of zeste (E(z)) to the same cluster group that caught our attention (Fig EV4A). To study this apparent association in greater detail and in conditions allowing for direct comparison, we performed ChIP-seq on dRing, E(z), Pc, H3K4me3 and H3K27me3 in S2 and S2-AbdA cells. Studying the genomic distribution of AbdA in S2-AbdA cells with GAF, M1BP, Rpb3, dRing and E(z) in S2 cells, we observed a striking colocalisation between the PcG proteins and AbdA at promoters bound by either GAF or M1BP and containing poised Pol II (Fig EV4B). While long known for their role in the stable and heritable repression of Hox genes, PcG proteins are now also known to play a finer role in transcriptional regulation at the promoter through maintaining a poised Pol II state in both *Drosophila* (Dellino *et al*, 2004; Schwartz *et al*, 2010; Chopra *et al*, 2011; Enderle *et al*, 2011; Kharchenko *et al*, 2011; Gaertner *et al*, 2012) and vertebrates (Boyer *et al*, 2006; Stock *et al*, 2007; Marks *et al*, 2012), the role of which in AbdA-driven transcriptional regulation in S2-AbdA cells we sought to investigate further.

Analysing the enrichment in S2 and S2-AbdA cells of PcG proteins, H3K4me3 and the PcG histone signature H3K27me3, we included genes lacking enriched promoter Pol II (n = 10,690), in addition to those genes showing poised Pol II (n = 6,910). At genes lacking poised Pol II, PcG proteins were found enriched together with H3K27me3 and were unaffected by AbdA expression (Fig 7A and B; unenriched Pol II genes). At these genes, enrichment of AbdA and H3K4me3 was absent. However, at the promoter-proximal region of poised Pol II genes, we observed enrichment of PcG proteins without their H3K27me3 signature, and this enrichment was markedly reduced upon AbdA binding (Fig 7A and B; poised Pol II genes). Using the list of genes separated into the three distinct classes based on the presence of only GAF, only M1BP or containing both GAF and M1BP peaks (see Fig 3), we find consistent significant loss of PcG at M1BP targets (Fig 7C). At these targets, the largest significant loss of PcG binding was that of dRing at promoter-proximal poised Pol II genes targeted by M1BP, where dRing binding was often lost (for examples, see Fig EV5). At M1BP-targeted poised Pol II genes, small but significant increases in H3K4me3 enrichment were observed (Fig 7), suggesting an increase in transcription due to PcG loss upon AbdA binding. Correlating ChIP enrichment at GAF- and M1BP-targeted genes (Fig 7A) with significant changes in gene expression (Appendix Fig S3), we observed the reduction in promoter-proximal binding of PcG proteins, linked with an increase in H3K4me3 at genes that were overexpressed upon AbdA binding at M1BP-specific targets, which was not the case at GAF-specific targets (Fig 8).

M1BP mutation enhances the sex comb Polycomb phenotype without altering Scr expression

If PcG proteins play a negative regulatory role at M1BP targets, a role that can be counteracted through Hox–M1BP cooperation, then a *PcG–M1BP* genetic interaction should be observable. However, it has long been known that mutation of PcG genes leads to ectopic Hox expression (Jürgens, 1985; Struhl & Akam, 1985; Riley *et al*, 1987; Glicksman & Brower, 1988; Simon *et al*, 1992) and so any phenotype observed from a test of *PcG–M1BP* genetic interaction would need further investigation to distinguish the contribution of changes in Hox gene expression versus changes in Hox protein activity.

Using an engineered loss-of-function allele for *M1BP* (Liu *et al*, 2012), we tested for genetic interaction using the classical PcG phenotype, which is the appearance of extra sex combs on male adult legs (Landecker *et al*, 1994). Pc^3 heterozygous mutant males

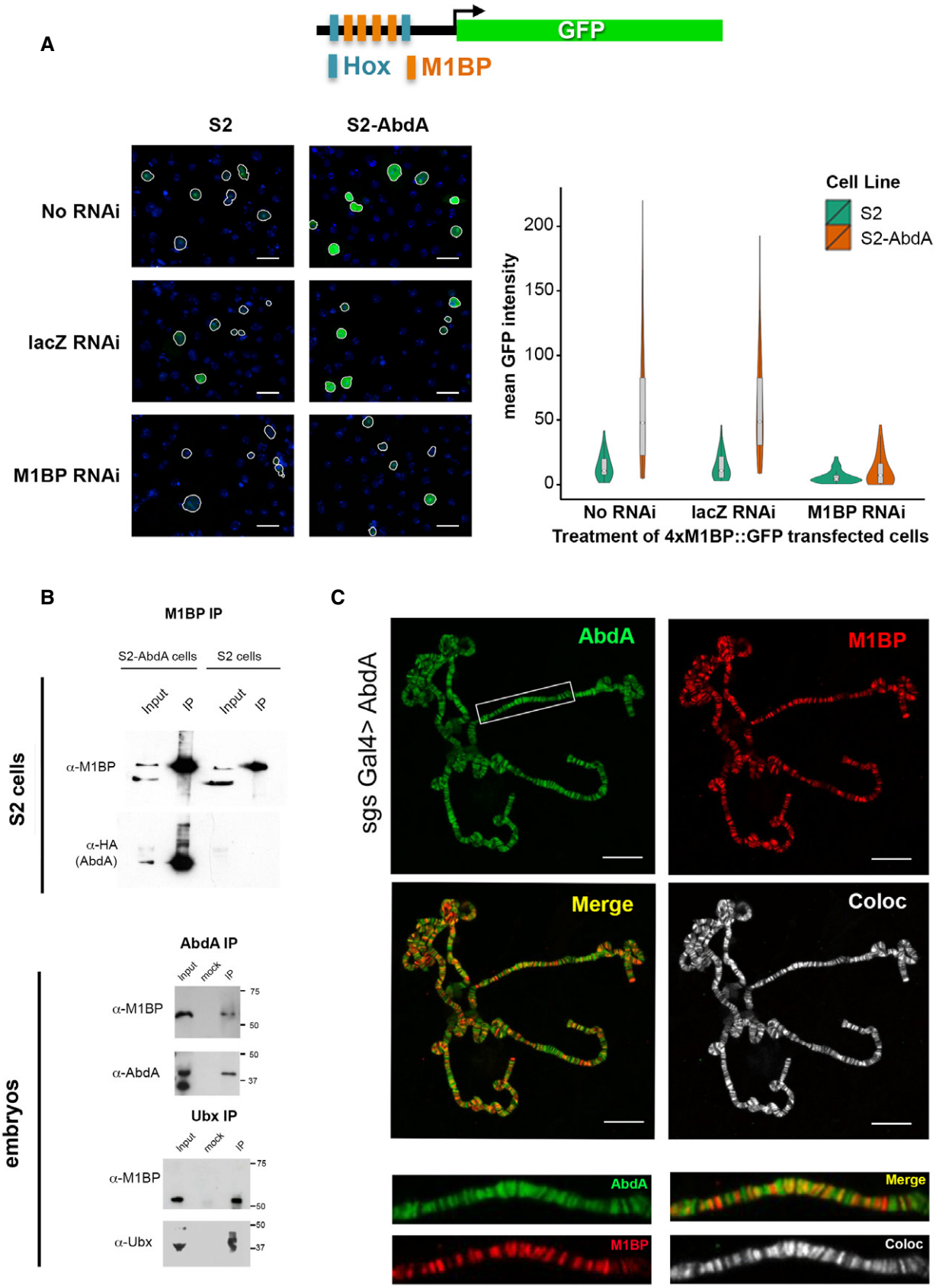

Figure 5.

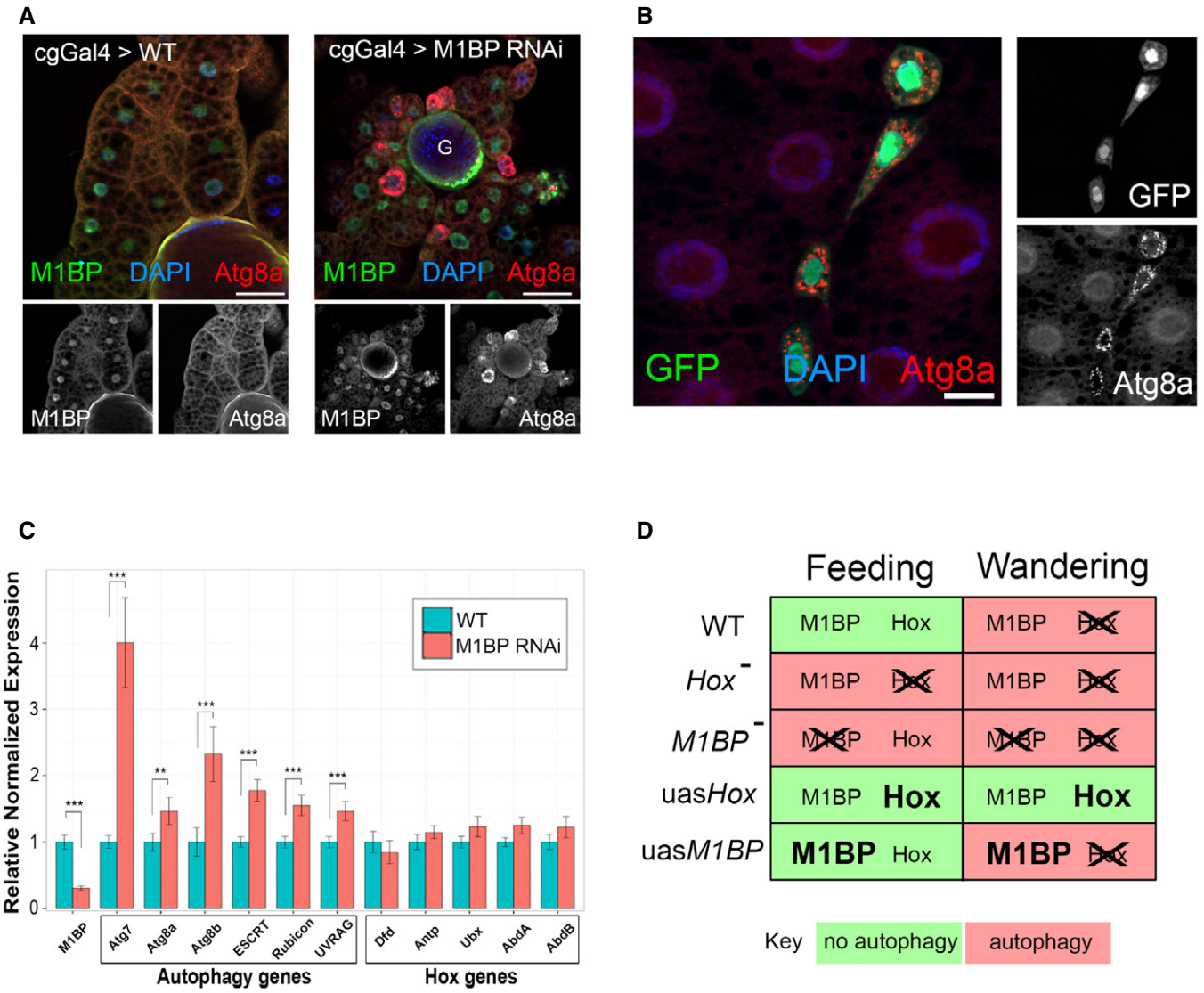

Figure 6. M1BP is essential in inhibiting autophagy in the *Drosophila* fat body.

A The Atg8a active autophagy marker is upregulated following global RNAi-mediated M1BP knockdown in L3 feeding larval fat body cells. Note that at this early stage, RNAi expression is efficient only in a subset of cells, precisely those that initiate autophagy. Gonads (G) are present in fat body preparations (scale bars, 50 μm).

B Clonal loss, RNAi-expressing cells are GFP-identified and autophagy is monitored using mCherry::*Atg8a* confirming the requirement of M1BP for autophagy repression in L3 feeding larval fat body cells (scale bar, 10 μm). Contrasts of individual channels are shown.

C Induction of autophagy by M1BP RNAi is not due to the loss of Hox expression. Gene expression was determined by qRT–PCR from wild-type or M1BP RNAi fat bodies. Error bars represent standard error of the mean of three independent experiments performed in triplicate, and *P*-values were derived using a Student's *t*-test: $^{**}P = 0.001$; $^{***}P = 0.0001$.

D Summary of the correlation between M1BP and Hox expression and autophagy status in fat body cells. Hox and uasHox is a summary of the findings by Banreti *et al* (2014). A cross denotes loss of gene expression and overexpression of UAS constructs is represented in bold. The status of fat body cells with respect to autophagy is colour-coded with green representing cells that do not display active autophagy and red denoting cells in autophagy.

display ectopic sex combs on the second legs comprising a median of two teeth, with a few rare exceptions displaying sex comb teeth on the third pair of legs (Fig 9A). While the *M1BP* mutation alone does not result in ectopic sex combs on the second and third legs, all *M1BP/Pc* trans-heterozygote males display at least one ectopic sex comb tooth on the second pair of legs, with an increased median value from two to three teeth and a greater number of males found to display ectopic sex combs on the third pair of legs (Fig 9A), demonstrating that *M1BP* is an enhancer of the *Pc* extra sex combs phenotype.

While the genes directly responsible for sex comb development are little studied, sex combs ultimately arise from the upstream action of the Hox gene *Sex combs reduced* (*Scr*) (Struhl, 1982; Riley *et al*, 1987) and this action is independent of Exd (Percival-Smith & Hayden, 1998). The ectopic sex comb PcG phenotype is due to the loss of PcG silencing of *Scr* expression (Jürgens, 1985). To

Figure 7. Reduction in PcG binding to promoter-proximal regions and increase in H3K4me3 histone mark are observed at M1BP-controlled genes upon AbdA expression.

A RefSeq transcripts enriched for promoter Pol II (poised Pol II genes; n = 6,910) were clustered according to total promoter Pol II enrichment levels. RefSeq genes showing no promoter Pol II enrichment (unenriched Pol II genes; n = 10,690) were clustered according to total H3K27me3 enrichment levels. Heatmaps of the enrichment spanning from 1 kb upstream of the TSS to 1 kb downstream of the TES are shown for ChIPs of H3K27me3, E(z), dRing and H3K4me3 in S2 and S2-AbdA cells. Heatmap of enrichment of Pol II (Rpb3 antibody) in S2 cells and AbdA (HA antibody) in S2-AbdA cells is shown for reference. Boxed areas highlight the regions used to quantitate mean enrichment values.

B Average profile plots of mean log2 ChIP enrichment at the boxed areas highlighted in (A) are shown for ChIPs performed on S2 cells (green) and S2-AbdA cells (brown) showing reduced PcG binding and increased H3K4me3 at poised Pol II genes. The dotted line represents no ChIP enrichment over input.

C Boxplot comparison of dRing, E(z) and H3K4me3 mean log2 RPM values sampled from the poised Pol II region at genes from S2 (green) and S2-AbdA (brown) cells containing either a GAF peak (GAF only; n = 1,309), an M1BP peak (M1BP only; n = 3,374) or both GAF and M1BP peaks (GAF and M1BP; n = 899) within 150 bp upstream of the TSS. Box limits represent the interquartile range with whiskers representing the first and fourth quartiles. Notches represent the 95% confidence interval of the median (black bar) and mean values represented by white dots. Boxplot outliers are not shown. P-values determined from permutation tests: $^{*}P < 0.05$; $^{***}P < 0.0001$; ND: P-values were not determined since samples did not show equal variance at a 3% significance threshold, even with outliers removed.

Source data are available online for this figure.

assess whether the additional sex combs observed upon *M1BP* loss of function in a *Pc* heterozygous mutant background are due to changes in Hox gene expression, we quantified *Scr* expression in all three male leg discs. We observed no effect on *Scr* expression upon *M1BP* mutation either alone or combined with *Pc* mutation (Fig 9B), which to our knowledge is the first reported example of an enhanced PcG phenotype (ectopic sex combs) that is not due to increased loss of PcG function in silencing Hox gene expression (*Scr*). Since the genes directly responsible for sex comb apparition remain unknown, it is not possible to determine whether *Pc/M1BP* mutation is affecting Pol II status at their promoter, nor whether Scr plays a role in the release of paused Pol II at these genes. Nonetheless, these genetic interaction data provide evidence that the PcG and M1BP function in similar pathways responsible for sex comb development and that loss of this function affects Scr protein activity.

Discussion

Ubx and AbdA predominantly associate with promoter regions

Understanding Hox transcriptional networks is central to understanding their wide repertoire of functions, yet observing where they bind in the genome does not explain why they bind there. In using a homogenous cell-based system devoid of endogenous Hox expression to conditionally express the Hox protein Ubx or AbdA, we have demonstrated that *Drosophila* Hox proteins target proximal promoters genome-wide, which is conserved (for Ubx at least) in developing embryos. While studies into Hox genomic binding have historically focussed on enhancer elements in spatially and temporarily controlling individual gene expression, genome-wide promoter enrichment of Hox proteins is known to occur for mouse HoxB4 in hematopoietic stem cells (Oshima *et al*, 2011), mouse Hoxa2 in the second branchial arches (Amin *et al*, 2015) and for zebrafish Hoxb1a in early embryogenesis (Choe *et al*, 2014). However, why Hox proteins target the promoter-proximal region has been little explored. A major advantage of the *Drosophila* S2 cell system is that the conditional Hox expression system allows studying in fine detail the sequence of events occurring upon promoter binding and the impact on gene expression.

Hox proteins partner with the pausing factor M1BP to fine tune gene expression

The promoters targeted by both AbdA and Ubx in *Drosophila* are essentially promoters containing either GAF or M1BP. GAF controls mainly development and morphogenic genes, whereas M1BP controls genes mainly involved in basic cellular processes (Li & Gilmour, 2013), and this distinction in gene ontology is reflected in the genes whose promoters are targeted by AbdA (Fig 2C). As AbdA and Ubx also target enhancer regions (Fig 1A), we cannot rule out that the observed promoter binding is the result of enhancer–promoter interaction. However, given that the majority of genes controlled by M1BP do not have distal enhancers (Zabidi *et al*, 2015), it is unlikely that this is the case for M1BP-targeted promoters. Both GAF and M1BP are important and distinct *Drosophila* Pol II pausing factors (Li & Gilmour, 2013; Fuda *et al*, 2015; Duarte *et al*, 2016), a role that proved important in understanding the nature of promoters targeted by AbdA and Ubx, since the majority of all promoters targeted by the Hox proteins contained poised Pol II (for example, see Figs 2B and 7A). GAF binding sites have previously been shown enriched at Ubx targets (Agrawal *et al*, 2011; Shlyueva *et al*, 2016), although a link between Hox and GAF in regulating gene transcription was not demonstrated. Similarly, in S2-AbdA cells, AbdA binding at GAF-regulated promoters has little clear-cut effect on poised Pol II status (Fig 3), although the amount of elongating Pol II and gene transcription appears reduced (Fig 4C). It was at M1BP-bound promoters where we found that AbdA has an effect on Pol II pausing, whereby AbdA binding results in a reduction in poised Pol II giving rise to increased productive transcription (Figs 3 and 4). Taken together with the findings that both Ubx and AbdA target nearly identical promoters (Fig EV1), that AbdA and M1BP synergise in reporter gene expression (Fig 5A), that both Ubx and AbdA interact with M1BP in embryos (Fig 5B) and AbdA colocalises with M1BP on polytene chromosomes (Fig 5C), these data suggest functional cooperation between M1BP and AbdA/Ubx. To this end, demonstrating that M1BP expression is essential in inhibiting autophagy in the larval fat body (Fig 6), an Exd-independent cellular function shared by all *Drosophila* Hox proteins where the loss of expression of all Hox genes is essential for autophagy induction (Banreti *et al*, 2014), suggests that M1BP may function with Hox proteins in their generic function of autophagy inhibition (Fig 6D).

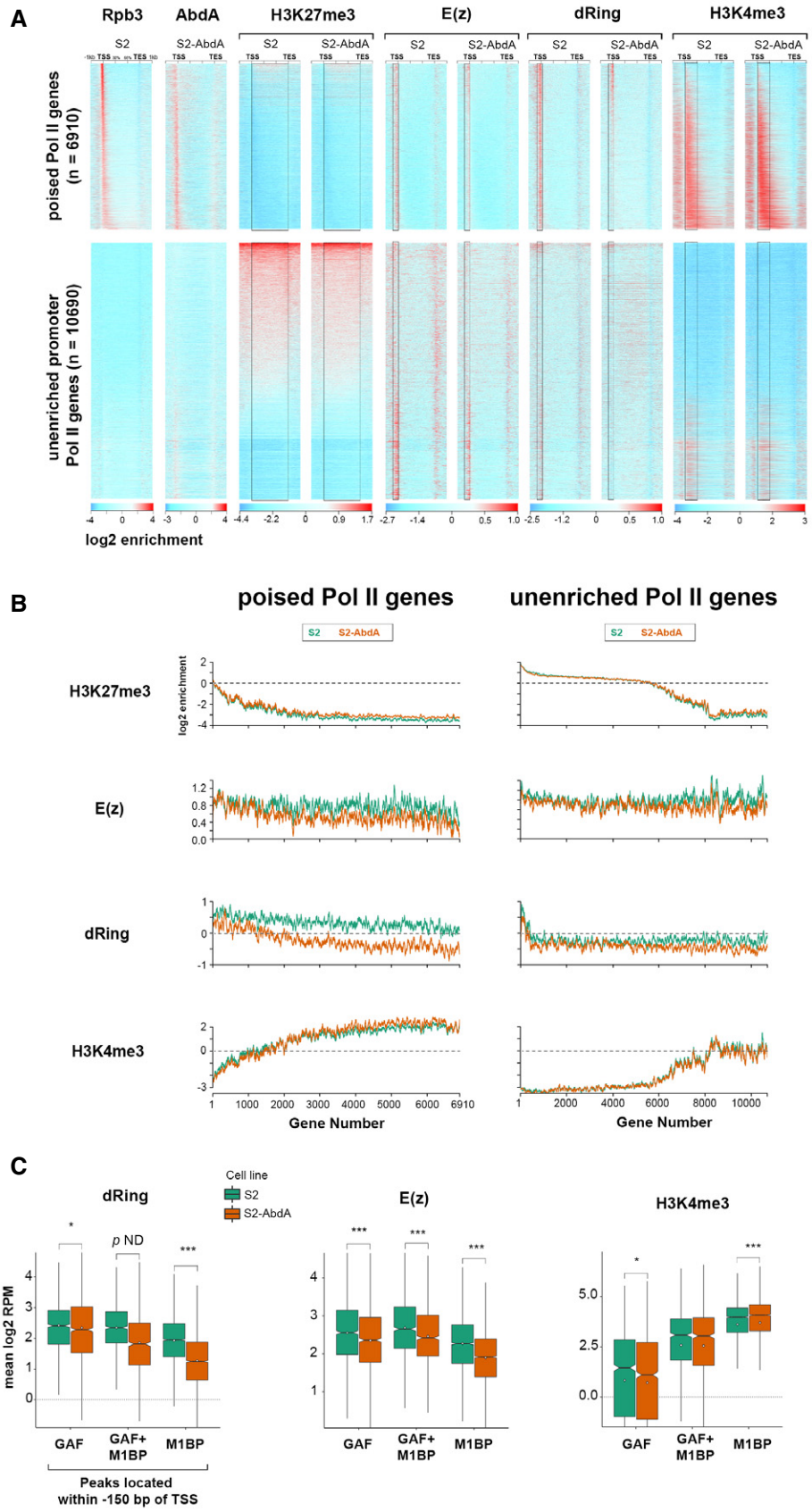
A
Rpb3
AbdA
H3K27me3
E(z)
dRing
H3K4me3
S2
S2-AbdA
poised Pol II genes (n = 6910)
unenriched promoter Pol II genes (n = 10690)
log2 enrichment
B
poised Pol II genes
unenriched Pol II genes
H3K27me3
E(z)
dRing
H3K4me3
log2 enrichment
Gene Number
C
Cell line
S2
S2-AbdA
dRing
E(z)
H3K4me3
mean log2 RPM
p ND
GAF
GAF+ M1BP
M1BP
Peaks located within -150 bp of TSS

Figure 7.

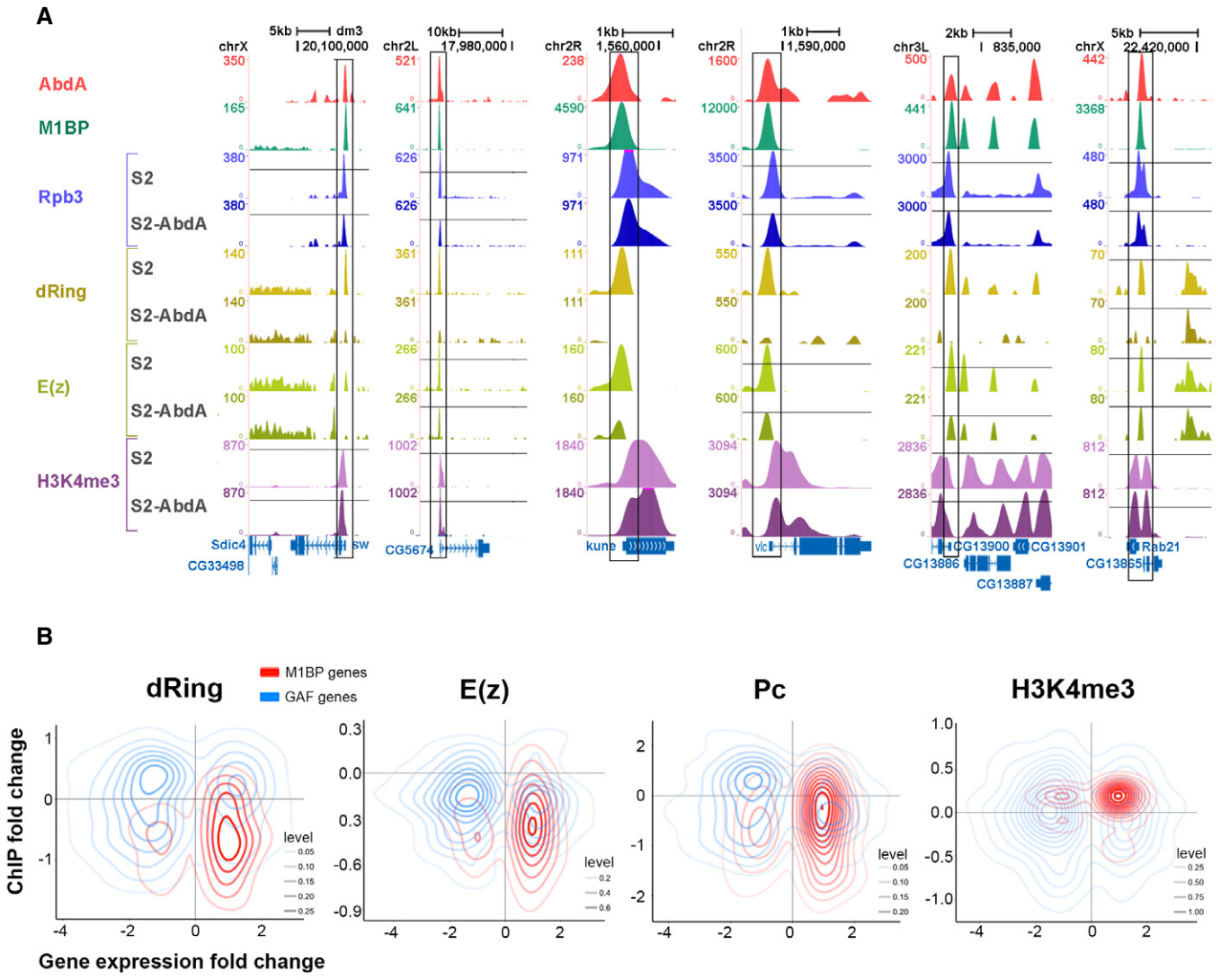

Figure 8. Promoter-proximal regions showing loss of PcG proteins upon AbdA binding show increased gene expression.

A Genome browser view of AbdA, M1BP, Rpb3, dRing, E(z) and H3K4me3 at genes that are upregulated upon AbdA expression shows the loss of PcG proteins specifically at the promoter-proximal region, whereas other regions remain largely unchanged upon AbdA binding concomitant with reduced poised Pol II and increased H3K4me3. Equal enrichment values between S2 and S2-AbdA cells are shown (horizontal black line) for Rpb3, E(z) and H3K4me3 ChIPs to allow visualisation of the small but significant changes that occur.

B Contour plots showing dynamic changes in PcG proteins and H3K4me3 (log2 fold change of RPM; ordinate) as a function of changes in gene expression of differentially expressed genes (log2 count fold change; abscissa) at Pol II poised genes upon AbdA expression. ChIP profiles are grouped into genes containing GAF (blue) or M1BP (red) peaks within 150 bp upstream of the TSS, which shows consistent reduction in promoter-proximal PcG proteins, as well as increased H3K4me3 profiles at AbdA-targeted genes containing promoter-proximal M1BP, resulting in increased gene expression. Contour lines group points of equal density and correspond to a sequence of equally spaced *z*-values, with increasing opacity added for regions with decreasing *z*-values.

Source data are available online for this figure.

Similar to the distinct mechanisms at play in pausing Pol II at M1BP- and GAF-controlled promoters (Li & Gilmour, 2013; Duarte *et al*, 2016), these data suggest that distinct mechanisms of release of paused Pol II exist at the two classes of poised Pol II promoters: additional factors that are not present in S2 cells are likely required to permit Hox-induced productive transcription at GAF-controlled promoters since we find little evidence that AbdA binding affects Pol II pausing, whereas at M1BP-controlled genes, AbdA binding is sufficient to increase gene transcription through the release of paused Pol II.

A role for non-canonical PcG protein activity in M1BP/Hox-mediated gene regulation

Testing the association of AbdA ChIP peaks in S2-AbdA cells with those of numerous publicly available histone-modifying proteins

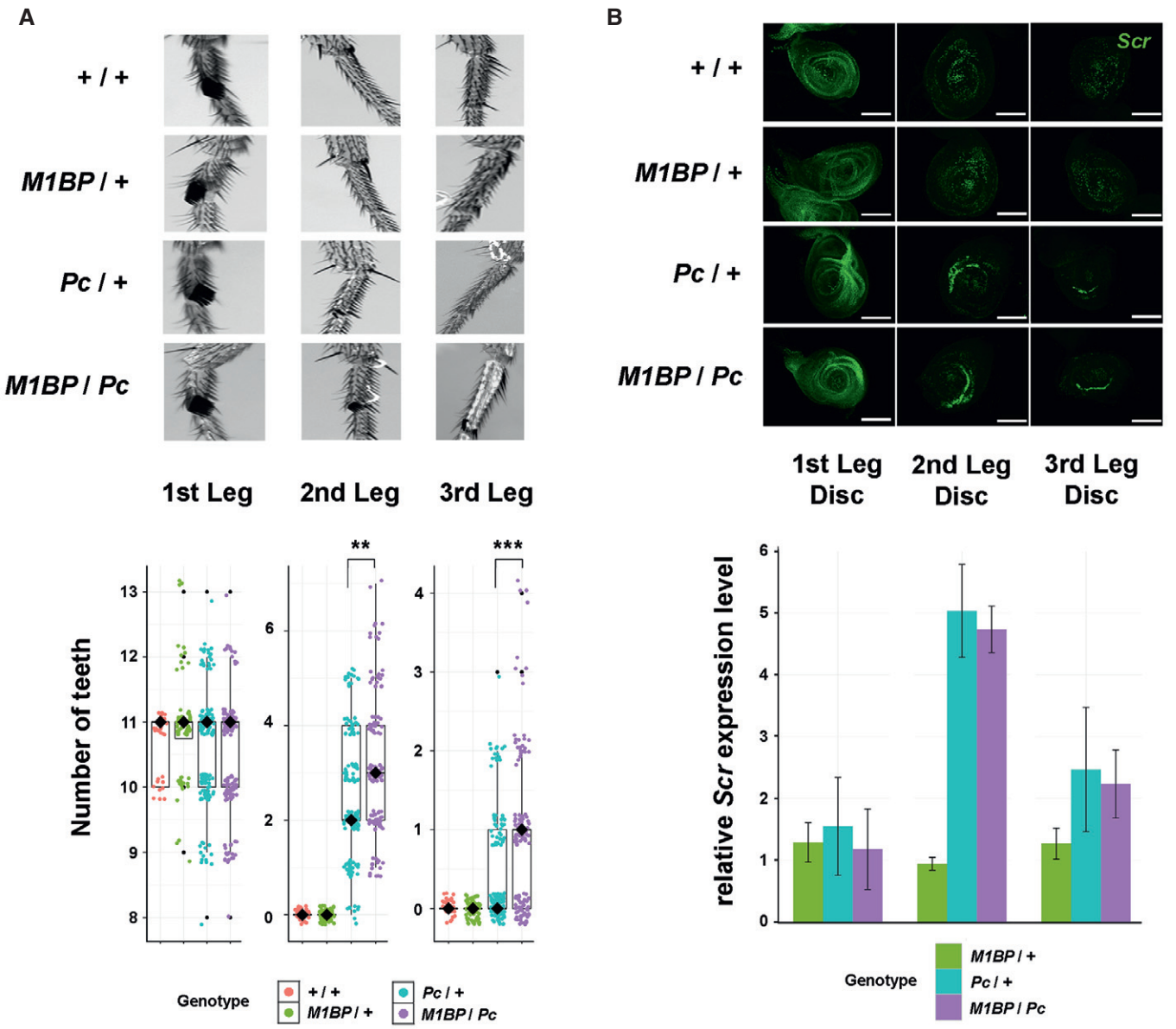

Figure 9. M1BP enhances the extra sex combs PcG phenotype without altering *Scr* gene expression.

A Quantification of the number of sex combs on first (left), second (centre) and third (right) legs from male flies that are wild type, heterozygous for either *M1BP* or *Pc* mutation and *M1BP/Pc* double heterozygote mutants. Data are presented as representative images (upper panels) and boxplots (lower panels) with individual counts shown as a jitterplot overlay. Box limits represent the interquartile range with whiskers representing the first and fourth quartiles. Outliers are represented as black dots. The median number of sex combs for each genotype is highlighted with a lozenge. $^{**}P < 0.01$; $^{***}P < 0.001$ from Mann–Whitney *U*-tests.

B The effects of *M1BP*, *Pc* and *M1BP/Pc* heterozygous mutation on *Scr* expression in first (left), second (centre) and third (right) leg discs from third-instar larvae were determined by anti-Scr immunolabelling (upper panels) and RT–qPCR (lower panels). RT–qPCR data are presented as *Scr* expression levels relative to wild type. Error bars represent standard error of the mean of three independent experiments performed in triplicate. Scale bars represent 100 μm.

and histone marks in S2 cells, we found that the M1BP- and GAF-poised Pol II promoters targeted by AbdA were enriched for PcG proteins and H3K4me3 (Fig EV4). The finding that AbdA-enhanced transcription at M1BP promoters was more consistently concomitant with a loss of promoter PcG protein binding than at GAF-controlled promoter (Fig 8), suggests that the emerging role for PcG proteins in maintaining a poised Pol II state (Dellino *et al*, 2004; Boyer *et al*, 2006; Stock *et al*, 2007; Schwartz *et al*, 2010; Chopra *et al*, 2011; Enderle *et al*, 2011; Kharchenko *et al*, 2011; Gaertner *et al*, 2012; Marks *et al*, 2012) can be perturbed by Hox binding. Indeed, it is noteworthy that of the PcG proteins tested here, it is promoter-bound dRing that is most affected upon AbdA binding (Figs 7, 8, and EV5), suggesting that, like in vertebrates where Ring1 plays a major role in restraining the poised Pol II at promoters (Stock *et al*, 2007), dRing may play a major role in tethering the poised Pol II state in *Drosophila*. Given that no clear effect on PcG binding occurs

at GAF-controlled promoters, even when these genes are repressed upon AbdA binding (Figs 7C and 8B), it reinforces the notion that contrary to M1BP targets, the control of gene expression by AbdA at GAF genes is unlikely to occur through the regulation of poised Pol II status.

Where PcG proteins are linked to maintaining gene repression, trithorax group proteins (trxG) are the PcG antagonists, responsible for maintaining gene expression. As a transcription factor, GAF has been traditionally classified as a trxG protein (Adkins *et al*, 2006) although it displays repressive activity and can recruit PcG complexes (Hagstrom *et al*, 1997; Horard *et al*, 2000; Mulholland *et al*, 2003). As such, GAF can be classified as a member of the growing family of genes that display both PcG and trxG phenotypes, the so-called enhancers of trithorax and polycomb (ETP) family (for review, see Fedorova *et al*, 2009). We show here that M1BP colocalises with PcG proteins at promoters in S2 cells and phenotypically enhances the PcG homeotic phenotype of extra male sex combs on the second and third pairs of legs (Fig 9A). Indeed, $M1BP^{-}/Pc^{-}$ transheterozygous males have an average of 5.3 legs displaying sex combs, which is more than most combinations of PcG mutant transheterozygotes (Campbell *et al*, 1995), demonstrating the large increase in penetrance of the *Pc* phenotype upon *M1BP* mutation. As such, *M1BP* would be genetically classified as a PcG gene. However, PcG genes, by definition, display homeotic phenotypes due to the derepression of Hox genes when mutated and so since we observed neither increased derepression of the upstream Hox gene responsible for sex comb development, *Scr*, upon *M1BP* mutation (Fig 9B) nor Hox expression in fat body cells following RNAi (Figs 6C and EV3D), *M1BP* cannot thus be classified as a PcG gene. Given that M1BP is a transcription factor involved in gene expression (Li & Gilmour, 2013), we therefore favour that, like GAF, M1BP is likely to be a member of the ETP family. How ETP proteins can enhance the phenotypes of both repressors (PcG) and activators (trxG) has long remained a mystery. Demonstrating here that GAF and M1BP colocalise with PcG at poised Pol II promoters with the loss of PcG at those genes displaying increased expression upon AbdA binding, may go a long way to better understand how transcription factors and transcriptional repressors intricately cooperate to regulate gene transcription.

A mechanistic model for Hox activity at promoters

In summary, this work identifies a novel mechanism for Pol II pausing release mediated by AbdA: at genes bound by M1BP, targeting of AbdA results in the specific loss of PcG proteins, the release of poised Pol II and increases in H3K4me3 histone marks, which results in promoting productive transcription. Identified in S2 cells where Hox PBC-class cofactors are absent, this mechanism may more generally apply to Hox-generic functions that are independent of PBC-class cofactors, such as the repression of autophagy in the *Drosophila* fat body or sex comb development in *Drosophila* males. It may also apply to PBC-dependent Hox target gene regulation by cooperating with Hox PBC-bound genomic regions located remote from the promoter. Further work aimed at studying Hox PBC-bound enhancers together with poised Pol II status, and promoter-proximal Hox and PcG binding, should provide further insight into how enhancer-bound protein complexes influence the basic mechanisms of transcription regulated through poised Pol II. Uncovering such a Hox-driven mechanism of gene regulation by sequence-specific transcription factors, PcG proteins and poised Pol II in the developing animal would have been fraught with difficulties, not least of which is the quagmire of PcG proteins being essential global repressors of all Hox genes.

Materials and Methods

Drosophila experiments

Details of stocks used or generated in this study are provided in Appendix Supplementary Methods.

For studies into autophagy in the fat body, UAS-M1BP RNAi males were crossed to either *cg*-Gal4 (Drosophila Bloomington Stock Center) or *yw*,hs-Flp;*r4*-mCherry::*Atg8a*;Act > CD2 > GAL4 (a gift from T. Neufeld) females. Crosses, incubations, preparation of L3-feeding and L3-wandering fat body samples, imaging of fixed tissues, LysoTracker-Red staining and immunohistochemistry on larval fat bodies were all done according to Banreti *et al* (2014). The following primary antibodies were used: rabbit anti-M1BP (a gift from D. Gilmour) and rat anti-Atg8a (a gift from G. Jusász). All secondary antibodies were purchased from Jackson ImmunoResearch Laboratories, Dianova or Molecular Probes. All imaging was performed on an AxioImager APO Z1 Apotome microscope (Zeiss).

For quantitative RT–PCR (RT–qPCR), total RNA was extracted from ten L3-feeding larvae fat bodies (cg-Gal4 females crossed to UAS-M1BP RNAi males or WT Oregon-R males) (gonads removed) with the RNeasy Mini Kit (Qiagen) and reverse-transcribed with Superscript II (Life Technologies). SYBR GreenER qRT–PCR SuperMix for iCycler® Instrument (Life Technologies) was used as cycle master mix and qPCR performed in triplicate on a CFX96 thermocycler (Bio-Rad). Quantifications of qPCRs from three independent experiments were performed using the CFX Manager Software (v3.1; Bio-Rad) using linear regression for Cq determination and gene expression represented as normalised expression (ΔΔCq) relative to zero using wild-type preparations as control sample. Statistical tests to generate *P*-values were determined using a Student's *t*-test through the CFX Manager Software Study Analysis (v3.1; Bio-Rad). Oligos used for autophagy and Hox genes are provided in Banreti *et al* (2014). M1BP-specific oligos used for qPCR are provided in Appendix Supplementary Methods.

For validation of AbdA target genes *in vivo*, *in situ* hybridisations and RT–qPCR were performed on *Drosophila* embryos collected from armGal4 flies crossed to either wild-type flies (OreR) or UAS HA::AbdA flies. For *in situ* hybridisation studies, digoxigenin-labelled RNA probes were generated from cDNA clones (RE21580, RE73615, GH14073, RE67261, SD07261 and SD01152) using SP6 or T7 polymerase (Promega) and hybridisation was performed on 0- to 16-h-old embryos using standard protocol. For RT–qPCR, total RNA was extracted from embryos collected for 3 h and aged for a further 10 h using the protocol mentioned above for larval fat body RT–PCR. Sequences of oligos used for AbdA target genes are provided in Appendix Supplementary Methods.

For polytene chromosome immunolabelling experiments, HA-tagged *abdA* was expressed in salivary glands using an *sgs*Gal4 driver. Polytene chromosome squashes followed by immunolabelling were performed using standard techniques. Colocalisation was determined in Fiji using a Colocalization Threshold plugin, and statistical significance testing was performed using Coloc 2 Fiji plugin.

Sex combs were counted on F1 male's legs from non-TM6C,Tb^1, Sb^1 flies from the following crosses: OreR females × OreR males; OreR females × Pc^3/TM6C,Tb1,Sb1 males; OreR females × $M1BP^{18.7}$/TM6C,Tb^1,Sb^1 males; $M1BP^{18.7}$/TM6C,Tb^1,Sb^1 females × Pc^3/TM6C, Tb^1,Sb^1 males.

For Scr expression analyses, leg discs from non-TM6C,Sb^1,Tb^1 male late third-instar larvae were dissected, fixed and immunostained with anti-Scr antibody (1:200, Developmental Studies Hybridoma Bank #6H4.1) using standard protocols. All imaging was performed on an AxioImager APO Z1 Apotome microscope (Zeiss).

For *Scr* mRNA quantification, ten pairs of male leg discs, each originating from first, second and third leg discs, were dissected separately from non-TM6C,Sb^1,Sb^1 male late third-instar larvae, total RNA was extracted and RT–qPCR performed as above using *Scr*- and *U6*-specific oligos (see Appendix Supplementary Methods for oligonucleotide sequences used).

Cell culture

S2-DRSC cells (Drosophila Genomics Resource Center, stock #181) were cultured at 25°C in Schneider's medium (Clinisciences) supplemented with 10% FBS. HA-tagged AbdA or Ubx was cloned into pMK33/pMtHy (Drosophila Genomics Resource Center), transfected into S2-DRSC cells and stable lines generated through hygromycin B selection over a 4-week period to generate the S2-AbdA cell line. The 4xMotif1::GFP reporter construct (Li *et al*, 2010) contains four M1BP-binding sites upstream of a GFP reporter (termed 4xM1BP::GFP in this manuscript) and was a gift from L. Li and S. Grewal. Two Hox TAAAT binding sites (Slattery *et al*, 2011; Beh *et al*, 2016) flank the four M1BP sites and two divergent Hox motifs (ATTG) that bind AbdA in EMSA flank the third M1BP binding site. HA:AbdA or HA:Ubx was induced by treating stable lines with either 10 μM (HA:AbdA) or 5 μM (HA:Ubx) $CuSO_4$ for 24 h, which was the minimum induction level to see comparable quantities of Ubx and AbdA by both western analyses and RT–qPCR (see Appendix Supplementary Methods for oligonucleotide sequences used).

The 4xM1BP::GFP reporter construct was co-transfected with a pC4-SQhPatt-Sqh-mCherry plasmid (a gift from T. Lecuit) to identify transfected cells (mCherry stained). RNAi experiments were performed according to Li and Gilmour (2013). On the third day of dsRNA incubation, the 4xMotif1::GFP and pC4-SQhPatt-Sqh-mCherry plasmids were co-transfected into 1×10^6 treated cells and the following day, cells were induced with 75 μM $CuSO_4$ for 24 h. M1BP and LacZ dsRNA were produced using the RiboMAX large-scale RNA production system (Promega; see Appendix Supplementary Methods for sequences used).

All imaging was performed on an AxioImager APO Z1 Apotome microscope (Zeiss). Quantitation of mean GFP fluorescence of transfected nuclei was performed using ImageJ between 50 and 60 transfected cells.

Protein immunoprecipitation

Wild-type *Drosophila* Oregon-R overnight embryos were collected during a 3-day period and nuclear extracts prepared. A total of 3 mg nuclear extracts (Protein A/G-precleared) was incubated overnight with 240 μg purified anti-AbdA or anti-Ubx antibody cross-linked to 40 μl protein A/G Plus Agarose using the Pierce Crosslink Immunoprecipitation Kit (Thermo-Pierce Scientific). Mock immunoprecipitations were performed under identical conditions, but by omitting the Hox-specific antibody. Following extensive washing of immunocomplexes, proteins were eluted through heating at 70°C for 10 min in non-reducing Laemmli buffer (2% SDS, 10% glycerol, 0.002% bromphenol blue and 0.125 M Tris–HCl, pH 6.8). Eluted proteins were then reduced by the addition of 20 mM dithiothreitol, heated for 10 min at 90°C and resolved on polyacrylamide gels followed by immunoblotting.

Chromatin immunoprecipitation

Cells were grown at 25°C in 225-cm^2 T-flasks to a density of 3×10^6 cells/ml upon which 10 μM (HA:AbdA) or 5 μM (HA:Ubx) $CuSO_4$ was added to the culture medium for 24 h. Cell density prior to cross-linking was between $5–6 \times 10^6$ cells/ml. Cross-linking of $CuSO_4$-induced cultured cells was performed according to the modENCODE procedure using formaldehyde (Sigma Aldrich, ref: F8775) as cross-linker. Chromatin was prepared according to the modENCODE procedure using a BioRuptor sonicator (Diagenode) yielding an average fragment size in the range of 300–500 bp. In parallel with chromatin preparation for anti-HA ChIP, total RNA was prepared from the same batch of cells using the RNeasy Mini Kit (Qiagen).

ChIP-seq experiments were performed as duplicates obtained from two independent chromatin preparations. For each ChIP-seq experiment, 500 μl of precleared chromatin (200–250 μg DNA) containing 1% PMSF was incubated with the appropriate antibody for 15 h at 4°C. 50 μl (1.5 mg) of Dynabeads Protein G (Life Technologies) was added per 10 μg antibody used and incubated for 3 h at 4°C. Beads were washed sequentially with the following procedure: 5 × 10 min with 1 ml RIPA buffer (140 mM NaCl, 10 mM Tris–Cl pH 8.0, 1 mM EDTA, 1% Triton-X 100, 0.1% SDS and 0.1% DOC), 1 × 10 min with 1 ml LiCl ChIP buffer (250 mM LiCl, 10 mM Tris–Cl pH 8.0, 1 mM EDTA, 0.5% NP-40 and 0.5% DOC) and 2 × 1 min in TE. Washed beads were resuspended in 100 μl TE and 1 μl RNase A (10 mg/ml) added. After 30-min RNase incubation for 30 min at 37°C, proteins were digested by the addition of 7.5 μl of 10% SDS and 3.8 μl of 10 mg/ml proteinase K followed by incubation overnight at 37°C. Crosslinks were reversed by incubation for 6 h at 65°C. Following phenol–chloroform extraction, DNA was precipitated overnight at −20°C and resuspended in 12 μl of nuclease-free H_2O ready for sequencing library preparation and sequencing.

Antibodies used for ChIP were anti-Rpb3 (a gift from K. Adelman), anti-M1BP (a gift from D. Gilmour), anti-GAF (modENCODE antibody; a gift from G. Cavalli, IGH, Montpellier), anti-HA (Abcam ref: ab9110), anti-Pol II Ser-2 clone 3E10 (Merck Millipore; ref 04-1571), anti-H3K27me3 (Millipore, ref: 07-440), anti-H3K4me3 (Abcam, ref: ab8580), anti-E(z) (Santa Cruz Biotech, ref: dN-20), anti-Pc (Santa Cruz Biotech, ref: dN-19) and anti-dRing. The

Drosophila rabbit polyclonal Rpb3 antibody has previously been validated for ChIP of poised Pol II (Core *et al*, 2012). The *Drosophila* dRing rabbit polyclonal antibody (dRing[98]) was produced by immunising rabbits with peptide sequence MTSLDPAPNKTWELS coupled to KLH and affinity-purified against the immunising peptide. Specificity of the dRing[98] antibody was assured by immunoblotting of embryo and cellular extracts, immunolocalisation studies and comparing ChIP locations with modENCODE ChIP data of dRing from embryos (GSE47266).

High-throughput sequencing

Sequencing was performed on libraries prepared from duplicates of ChIP or RNA preparations. RNA-seq libraries were prepared using the TruSeq RNA Sample Preparation Kit (Illumina). ChIP-seq libraries were prepared using the ChIP-Seq Sample Preparation Kit (Illumina). All libraries were validated for concentration and fragment size using Agilent DNA1000 chips. Sequencing was performed on a HiSeq 2000 or HiSeq 4000 (Illumina), base calling performed using RTA (Illumina) and quality control performed using FastQC. ChIP-seq sequences were aligned to the dm3 genome assembly using BWA-SW and RNA-seq sequences aligned using CASAVA (Illumina). Raw counts for RNA-seq were obtained using "readBases" in CASAVA (Illumina).

The total numbers of mapped reads for input chromatin used for each ChIP are given in Table EV3. The total numbers of mapped reads for each ChIP are given in Table EV4. The total numbers of mapped reads for each RNA-seq analysis are given in Table EV5.

High-throughput sequencing data analyses details are provided in Appendix Supplementary Methods.

Data accessibility

ChIP-seq and RNA-seq datasets generated from this study are deposited with the Gene Expression Omnibus (GEO; accession GSE101557). The following public ChIP datasets used in this study were downloaded from GEO: AbdA and Ubx in Kc167 cells (GSE69796), Ubx from embryos (GSE64284), yki from S2 cells (GSE46305), dwg from S2 cells (GSE32853) and pho from S2 cells (GSE32857).

Expanded View for this article is available online.

Acknowledgements

We thank K. Adelman, G. Cavalli, D. Gilmour and G. Juhász for their generous gifts of antibodies; L. Li and S. Grewal for the 4xMotif1::GFP reporter construct; and T. Neufeld and the Bloomington and Vienna Stock Centers for providing *Drosophila* lines. Pol II phospho Ser-2 sequencing was performed by the IGBMC Microarray and Sequencing platform, a member of the "France Génomique" consortium (ANR-10-INBS-0009). W.-M.D. is supported by NIH Grant R01GM072562 and NSF Grant IOS-1557904. This work was supported by the Fondation pour la Recherche Médicale (DEQ20150934723), Agence Nationale de la Recherche (ANR-09-BLAN-0270-01), Centre Franco-Indien pour la Promotion de la Recherche Avancée (CEFIPRA 4703), la Ligue Nationale Contre le Cancer, Association pour la Recherche sur le Cancer (PJA20131200461) and the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR).

Author contributions

AZ, AA, MCD, MM and MB performed experiments, data analyses and interpretation. AZ, SEA, LB and AJS performed bioinformatics data analyses and interpretation. GX and W-MD generated fly lines. SSL provided sequencing expertise. YG and AJS designed the study and interpreted data. AJS supervised the study and wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abu-Shaar M, Ryoo HD, Mann RS (1999) Control of the nuclear localization of Extradenticle by competing nuclear import and export signals. *Genes Dev* 13: 935 – 945

Adelman K, Lis JT (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* 13: 720 – 731

Adkins NL, Hagerman TA, Georgel P (2006) GAGA protein: a multi-faceted transcription factor. *Biochem Cell Biol* 84: 559 – 567

Agrawal P, Habib F, Yelagandula R, Shashidhara LS (2011) Genome-level identification of targets of Hox protein Ultrabithorax in *Drosophila*: novel mechanisms for target selection. *Sci Rep* 1: 205

Amin S, Donaldson IJ, Zannino DA, Hensman J, Rattray M, Losa M, Spitz F, Ladam F, Sagerstrom C, Bobola N (2015) Hoxa2 selectively enhances meis binding to change a branchial arch ground state. *Dev Cell* 32: 265 – 277

Banreti A, Hudry B, Sass M, Saurin AJ, Graba Y (2014) Hox proteins mediate developmental and environmental control of autophagy. *Dev Cell* 28: 56 – 69

Barboric M, Nissen RM, Kanazawa S, Jabrane-Ferrat N, Peterlin BM (2001) NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol Cell* 8: 327 – 337

Beh CY, El-Sharnouby S, Chatzipli A, Russell S, Choo SW, White R (2016) Roles of cofactors and chromatin accessibility in Hox protein target specificity. *Epigenetics Chromatin* 9: 1

Boube M, Hudry B, Immarigeon C, Carrier Y, Bernat-Fabre S, Merabet S, Graba Y, Bourbon HM, Cribbs DL (2014) *Drosophila melanogaster* Hox transcription factors access the RNA polymerase II machinery through direct homeodomain binding to a conserved motif of mediator subunit Med19. *PLoS Genet* 10: e1004303

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441: 349 – 353

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10: 1213 – 1218

Campbell RB, Sinclair DAR, Couling M, Brock HW (1995) Genetic interactions and dosage effects of *Polycomb* group genes of *Drosophila*. *Mol Gen Genet* 246: 291 – 300

Cherbas L, Willingham A, Zhang D, Yang L, Zou Y, Eads BD, Carlson JW, Landolin JM, Kapranov P, Dumais J, Samsonova A, Choi JH, Roberts J, Davis CA, Tang H, van Baren MJ, Ghosh S, Dobin A, Bell K, Lin W *et al* (2011) The transcriptional diversity of 25 *Drosophila* cell lines. *Genome Res* 21: 301 – 314

Choe SK, Ladam F, Sagerstrom CG (2014) TALE factors poise promoters for activation by Hox proteins. *Dev Cell* 28: 203 – 211

Chopra VS, Hendrix DA, Core LJ, Tsui C, Lis JT, Levine M (2011) The polycomb group mutant esc leads to augmented levels of paused Pol II in the *Drosophila* embryo. *Mol Cell* 42: 837 – 844

Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT (2012) Defining the status of RNA polymerase at promoters. *Cell Rep* 2: 1025 – 1035

Dellino GI, Schwartz YB, Farkas G, McCabe D, Elgin SC, Pirrotta V (2004) Polycomb silencing blocks transcription initiation. *Mol Cell* 13: 887 – 893

Duarte FM, Fuda NJ, Mahat DB, Core LJ, Guertin MJ, Lis JT (2016) Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. *Genes Dev* 30: 1731 – 1746

Eberhardy SR, Farnham PJ (2002) Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. *J Biol Chem* 277: 40156 – 40162

Enderle D, Beisel C, Stadler MB, Gerstung M, Athri P, Paro R (2011) Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res* 21: 216 – 226

Fedorova EV, Pindiurin AV, Baricheva EM (2009) Maintenance of the patterns of expression of homeotic genes in the development of *Drosophila melanogaster* by proteins of the polycomb, trithorax, and ETP groups. *Genetika* 45: 1301 – 1318

Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A, Lis JT (2015) GAGA factor maintains nucleosome-free regions and has a role in RNA polymerase II recruitment to promoters. *PLoS Genet* 11: e1005108

Gaertner B, Johnston J, Chen K, Wallaschek N, Paulson A, Garruss AS, Gaudenz K, De Kumar B, Krumlauf R, Zeitlinger J (2012) Poised RNA polymerase II changes over developmental time and prepares genes for future expression. *Cell Rep* 2: 1670 – 1683

Glicksman MA, Brower DL (1988) Misregulation of homeotic gene expression in *Drosophila* larvae resulting from mutations at the extra sex combs locus. *Dev Biol* 126: 219 – 227

Hagstrom K, Muller M, Schedl P (1997) A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the *Drosophila* bithorax complex. *Genetics* 146: 1365 – 1380

Heger A, Webber C, Goodson M, Ponting CP, Lunter G (2013) GAT: a simulation framework for testing the association of genomic intervals. *Bioinformatics* 29: 2046 – 2048

Horard B, Tatout C, Poux S, Pirrotta V (2000) Structure of a polycomb response element and *in vitro* binding of polycomb group complexes containing GAGA factor. *Mol Cell Biol* 20: 3187 – 3197

Hueber SD, Bezdan D, Henz SR, Blank M, Wu H, Lohmann I (2007) Comparative analysis of Hox downstream genes in *Drosophila*. *Development* 134: 381 – 392

Johnson FB, Krasnow MA (1990) Stimulation of transcription by an *Ultrabithorax* protein *in vitro*. *Genes Dev* 4: 1044 – 1052

Johnson FB, Krasnow MA (1992) Differential regulation of transcription preinitiation complex assembly by activator and repressor homeo domain proteins. *Genes Dev* 6: 2177 – 2189

Jürgens G (1985) A group of genes controlling the spatial expression of the bithorax complex of *Drosophila*. *Nature* 316: 153 – 155

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, Linder-Basso D, Plachetka A, Shanower G, Tolstorukov MY, Luquette LJ, Xi R, Jung YL, Park RW, Bishop EP, Canfield TK *et al* (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471: 480 – 485

Klionsky DJ, Baehrecke EH, Brumell JH, Chu CT, Codogno P, Cuervo AM, Debnath J, Deretic V, Elazar Z, Eskelinen EL, Finkbeiner S, Fueyo-Margareto J, Gewirtz D, Jaattela M, Kroemer G, Levine B, Melia TJ, Mizushima N, Rubinsztein DC, Simonsen A *et al* (2011) A comprehensive glossary of autophagy-related molecules and processes (2nd edition). *Autophagy* 7: 1273 – 1294

Landecker HL, Sinclair DA, Brock HW (1994) Screen for enhancers of Polycomb and Polycomblike in *Drosophila melanogaster*. *Dev Genet* 15: 425 – 434

Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS (2008) NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol Cell Biol* 28: 3290 – 3300

Li L, Edgar BA, Grewal SS (2010) Nutritional control of gene expression in *Drosophila* larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biol* 11: 7

Li J, Gilmour DS (2013) Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J* 32: 1829 – 1841

Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R (2012) Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. *J Genet Genomics* 39: 209 – 215

Mann RS, Lelli KM, Joshi R (2009) Hox specificity: unique roles for cofactors and collaborators. *Curr Top Dev Biol* 88: 63 – 101

Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, Smith A, Stunnenberg HG (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149: 590 – 604

Merabet S, Mann RS (2016) To be specific or not: the critical relationship between hox and tale proteins. *Trends Genet* 32: 334 – 347

Mulholland NM, King IF, Kingston RE (2003) Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA. *Genes Dev* 17: 2741 – 2746

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K (2007) RNA polymerase is poised for activation across the genome. *Nat Genet* 39: 1507 – 1511

Oshima M, Endoh M, Endo TA, Toyoda T, Nakajima-Takagi Y, Sugiyama F, Koseki H, Kyba M, Iwama A, Osawa M (2011) Genome-wide analysis of target genes regulated by HoxB4 in hematopoietic stem and progenitor cells developing from embryonic stem cells. *Blood* 117: e142 – e150

Pearson JC, Lemons D, McGinnis W (2005) Modulating Hox gene functions during animal body patterning. *Nat Rev Genet* 6: 893 – 904

Percival-Smith A, Hayden DJ (1998) Analysis in *Drosophila melanogaster* of the interaction between sex combs reduced and extradenticle activity in the determination of tarsus and arista identity. *Genetics* 150: 189 – 198

Prince F, Katsuyama T, Oshima Y, Plaza S, Resendez-Perez D, Berry M, Kurata S, Gehring WJ (2008) The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* 135: 1669 – 1679

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA (2010) c-Myc regulates transcriptional pause release. *Cell* 141: 432 – 445

Rezsohazy R, Saurin AJ, Maurel-Zaffran C, Graba Y (2015) Cellular and molecular insights into Hox protein action. *Development* 142: 1212 – 1227

Riley PD, Carroll SB, Scott MP (1987) The expression and regulation of sex combs reduced protein in *Drosophila* embryos. *Genes Dev* 1: 716 – 730

Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, Lin MF, Washietl S, Arshinoff BI, Ay F, Meyer PE, Robine N, Washington NL, Di Stefano L, Berezikov E, Brown CD, Candeias R *et al* (2010) Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* 330: 1787 – 1797

Saadaoui M, Litim-Mecheri I, Macchi M, Graba Y, Maurel-Zaffran C (2015) A survey of conservation of sea spider and *Drosophila* Hox protein activities. *Mech Dev* 138(Pt 2): 73 – 86

Schwartz YB, Kahn TG, Stenberg P, Ohno K, Bourgon R, Pirrotta V (2010) Alternative epigenetic chromatin states of polycomb target genes. *PLoS Genet* 6: e1000805

Shlyueva D, Meireles-Filho AC, Pagani M, Stark A (2016) Genome-wide ultrabithorax binding analysis reveals highly targeted genomic loci at developmental regulators and a potential connection to polycomb-mediated regulation. *PLoS One* 11: e0161997

Simon J, Chiang A, Bender W (1992) Ten different *Polycomb* group genes are required for spatial control of the *abdA* and *AbdB* homeotic products. *Development* 114: 493 – 505

Slattery M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, Zhou T, Rohs R, Honig B, Bussemaker HJ, Mann RS (2011) Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell* 147: 1270 – 1282

Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, Lee BK, Sheffield NC, Graf S, Huss M, Keefe D, Liu Z, London D, McDaniell RM, Shibata Y, Showers KA, Simon JM, Vales T, Wang T, Winter D, Zhang Z *et al* (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res* 21: 1757 – 1767

Stobe P, Stein MA, Habring-Muller A, Bezdan D, Fuchs AL, Hueber SD, Wu H, Lohmann I (2009) Multifactorial regulation of a hox target gene. *PLoS Genet* 5: e1000412

Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* 9: 1428 – 1435

Struhl G (1982) Genes controlling segmental specification in the *Drosophila* thorax. *Proc Natl Acad Sci USA* 79: 7380 – 7384

Struhl G, Akam M (1985) Altered distributions of *Ultrabithorax* transcripts in *extra sex combs* mutant embryos of *Drosophila*. *EMBO J* 4: 3259 – 3264

Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, Stark A (2015) Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 518: 556 – 559

Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA (2007) RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat Genet* 39: 1512 – 1516

