OXFORD GENETICS

GENETICS, 2021, 219(1), iyab108
DOI: 10.1093/genetics/iyab108
Advance Access Publication Date: 14 July 2021
Investigation

The Drosophila HP1 family is associated with active gene expression across chromatin contexts

John M. Schoelz, Justina X. Feng, and Nicole C. Riddle *

Department of Biology, The University of Alabama at Birmingham, Birmingham, AL 35294, USA

*Corresponding author: Department of Biology, The University of Alabama at Birmingham, 1720 2nd Avenue South; CH464, Birmingham, AL 35294-1170, USA. Email: riddlenc@uab.edu

Abstract

Drosophila Heterochromatin Protein 1a (HP1a) is essential for heterochromatin formation and is involved in transcriptional silencing. However, certain loci require HP1a to be transcribed. One model posits that HP1a acts as a transcriptional silencer within euchromatin while acting as an activator within heterochromatin. However, HP1a has been observed as an activator of a set of euchromatic genes. Therefore, it is not clear whether, or how, chromatin context informs the function of HP1 proteins. To understand the role of HP1 proteins in transcription, we examined the genome-wide binding profile of HP1a as well as two other Drosophila HP1 family members, HP1B and HP1C, to determine whether coordinated binding of these proteins is associated with specific transcriptional outcomes. We found that HP1 proteins share many of their endogenous binding targets. These genes are marked by active histone modifications and are expressed at higher levels than nontarget genes in both heterochromatin and euchromatin. In addition, HP1 binding targets displayed increased RNA polymerase pausing compared with nontarget genes. Specifically, colocalization of HP1B and HP1C was associated with the highest levels of polymerase pausing and gene expression. Analysis of HP1 null mutants suggests these proteins coordinate activity at transcription start sites to regulate transcription. Depletion of HP1B or HP1C alters expression of protein-coding genes bound by HP1 family members. Our data broaden understanding of the mechanism of transcriptional activation by HP1a and highlight the need to consider particular protein–protein interactions, rather than broader chromatin context, to predict impacts of HP1 at transcription start sites.

Keywords: transcription; Heterochromatin Protein 1; chromatin; promoter proximal pausing

Introduction

The Heterochromatin Protein 1 (HP1) family is a prominent class of nonhistone chromosomal proteins, essential to ensure genome integrity and function (Vermaak and Malik 2009; Filion *et al.* 2010; Kharchenko *et al.* 2011; Canzio *et al.* 2014; Eissenberg and Elgin 2014). HP1 proteins are characterized by their unique domain structure consisting of a chromo-domain and a chromoshadow-domain connected by a hinge region (Smothers and Henikoff 2001). The chromo-domain mediates interactions between HP1 proteins and methylated histone tails (Jacobs *et al.* 2001), whereas the chromoshadow-domain mediates HP1 protein dimerization and interactions between HP1 family members and proteins containing a PxVxL amino acid motif (Thiru *et al.* 2004; Lechner *et al.* 2005). The ability to bind both methylated histones and a diverse set of additional nuclear proteins confers the classification of "hub protein" to the HP1 family. As such, HP1 proteins are active in several different nuclear processes. Most prominently, HP1 orthologs are essential for heterochromatin formation and propagation through both the establishment of a phase-separated nuclear environment and the recruitment of histone 3 lysine 9 methyltransferases (Czermin *et al.* 2001; Jacobs *et al.* 2001; Snowden *et al.* 2002; Motamedi *et al.* 2008; Larson *et al.* 2017; Strom *et al.* 2017; Machida *et al.* 2018; Sanulli *et al.* 2019). HP1 proteins also are involved in additional biological processes including DNA repair (Ryu *et al.* 2015; Amaral *et al.* 2017), DNA replication (Li *et al.* 2011), and regulation of gene expression (Danzer and Wallrath 2004; Lin *et al.* 2008; Kwon *et al.* 2010), illustrating the importance of this gene family (Badugu *et al.* 2003; Vermaak and Malik 2009).

The *Drosophila melanogaster* HP1 family includes five full-length genes (containing both a chromo-domain and a chromoshadow-domain): *Su(var)205* (encoding the HP1a protein), *HP1b*, *HP1c*, *rhino* (encoding HP1D), and *HP1e* (Vermaak and Malik 2009). *Su(var)205*, *HP1b*, and *HP1c* are expressed ubiquitously while *rhino* and *HP1e* are present mostly in female and male germ cells, respectively (Vermaak *et al.* 2005; Levine *et al.* 2012). Based initially on studies from Drosophila polytene chromosomes, the HP1a protein mostly localizes to pericentric heterochromatin, telomeres, chromosome four, and a few euchromatic loci (James *et al.* 1989; Fanti *et al.* 2003). This localization pattern was confirmed by later chromatin immunoprecipitation (ChIP) studies from the modENCODE (model organism encyclopedia of DNA elements) consortium and others (Riddle *et al.* 2011; Figueiredo *et al.* 2012; Lundberg *et al.* 2013; Ho *et al.* 2014). HP1B localizes throughout heterochromatic and

Received: March 31, 2021. **Accepted:** July 07, 2021

euchromatic domains on polytene chromosomes, and HP1C localizes mostly to euchromatin (Smothers and Henikoff 2001). These patterns are reinforced also by data from ChIP-chip and ChIP-seq experiments performed by the modENCODE consortium and others (Greil *et al.* 2003; de Wit *et al.* 2007; Ho *et al.* 2014; Kessler *et al.* 2015). Loss of function mutations in the *Su(var)205* gene encoding HP1a disrupt the formation of heterochromatin and are homozygous lethal (Eissenberg *et al.* 1990), whereas loss of function mutations in the *HP1b* and *HP1c* genes are homozygous viable (Font-Burgada *et al.* 2008; Mills *et al.* 2018). This finding has led to the speculation that the HP1B and HP1C proteins may exhibit functional redundancy. Together, these data provide a model of the Drosophila HP1 family wherein HP1a is an essential heterochromatin protein, HP1C is a nonessential euchromatin protein, and HP1B is a nonessential protein binding to both heterochromatin and euchromatin. These distributions are informative for determining HP1 family member functions in transcriptional regulation. For instance, the heterochromatic distribution of HP1a and its essential role in heterochromatin formation contribute to a model of HP1a functioning as a transcriptional repressor. This interpretation is supported by data from studies tethering HP1a to transgene reporters that result in transcriptional silencing (Li *et al.* 2003; Danzer and Wallrath 2004). However, a role for HP1a in transcriptional repression is complicated by the observation that a number of euchromatic and heterochromatic loci require HP1a to maintain an active transcriptional state (Lu *et al.* 2000; Cryderman *et al.* 2005). Additionally, inducible loci such as heat shock response genes are enriched for HP1a upon induction (Piacentini *et al.* 2003, 2009). One proposed model to explain these differences is that HP1a serves different functions in different chromatin contexts through interactions with distinct sets of protein partners (Li *et al.* 2002), but evidence for this hypothesis is lacking.

An alternative approach to investigating the effects of HP1a on gene expression is to focus on its interactions with other HP1 family proteins. Although the exact function of HP1B or HP1C in transcriptional regulation is not well characterized, tethering studies of transgene reporters support a role for HP1C in transcriptional activation (Font-Burgada *et al.* 2008). Evidence for the impact of HP1B on gene transcription is conflicting. Although tethering studies support a role for HP1B in gene silencing, position effect variegation (PEV) studies support a role for HP1B in transcriptional activation (Font-Burgada *et al.* 2008; Mills *et al.* 2018) . HP1C recruits the Facilitates Chromatin Transcription (FACT) complex to promote RNA polymerase II (RPII) elongation after being targeted to chromatin by the zinc finger transcription factors WOC and ROW (Font-Burgada *et al.* 2008; Kwon *et al.* 2010). However, others have observed roles for HP1C in transcriptional repression through an interactions with Su(H) and the piRNA pathway protein Ctp (Schnabl 2021; Sun 2021). However, all three HP1 proteins co-immunoprecipitate together (Alekseyenko *et al.* 2014; Ryu *et al.* 2014), and furthermore, both HP1a and HP1B also interact with subunits of FACT as well as WOC, but the nature of these interactions is uncharacterized (Kwon *et al.* 2010; Ryu *et al.* 2014). RNA-Seq experiments following RNAi knockdown of all three HP1 paralogs in Drosophila reveal evidence of both activating and silencing functions of HP1 proteins: both widespread up- and down-regulation of target genes are observed with a large number of misregulated genes being shared across knockdown conditions (Lee *et al.* 2013). These findings raise the possibility that HP1 proteins may coordinate their activity to regulate gene expression of a common transcriptional program.

Here, we explore whether combinatorial action and simultaneous binding activity of multiple HP1 proteins at a single locus may predict differences in transcriptional activity at protein-coding genes with better accuracy than knowledge of the surrounding chromatin context. To achieve this goal, we integrate ChIP-Seq and RNA-Seq datasets to characterize the genomic distribution of each HP1 protein and to measure the association between each HP1 protein and transcriptional states genome-wide. We find active transcription at binding targets shared between multiple HP1 proteins across a variety of chromatin states. Furthermore, these targets exhibit signatures of RPII promoter proximal pausing, providing evidence for a potential mechanism for transcriptional activation by HP1 proteins. Analysis of pausing in HP1 null mutants suggests coordinated activity between HP1 family members is important for proper gene expression. These findings suggest knowledge of locus-specific protein–protein interactions is more informative for predicting HP1 function at transcription start sites (TSSs) than knowledge of a broader chromatin context.

Materials and methods

reChIP analysis

This protocol is derived from the basic ChIP protocol described by Kharchenko *et al.* (2011). To isolate chromatin, S2 cells were harvested from two 50 ml confluent cultures and fixed with 2.5 ml 37% formaldehyde for 10 minutes at room temperature incubation on an orbital shaker before fixation was quenched with 6 ml 1.25 M glycine. Cells were pelleted and resuspended in 1× PBS before being washed twice subsequently with ChIP wash A buffer and ChIP wash B buffer before Dounce homogenization and pelleted again. Pellets were resuspended in TE and SDS solution to lyse nuclei and washed with TE before resuspended in TE-PMSF (1 mM) with SDS solution. Nuclear lysates were sonicated (30 s on, 45 s off for 15 cycles) and incubated with Triton-X, DOC and 5 N NaCl for 10 min at 4°C on a rotating wheeling before pelleting and snap freezing for storage prior to immunoprecipitation.

For immunoprecipitation, samples were diluted 3× with ChIP dilution buffer (1% SDS, 10 mM EDTA, 50 mM Tris HCl, 1× PIC, 0.1% Triton-X) and precleared with 25% protein-a-sepharose slurry for 1 h at 4°C. Following pulse spin, samples were transferred to a clean tube and 5 μl antibody of interest was added to each sample. Chromatin was precipitated overnight at 4°C on rotating wheel. Afterwards, protein-a-sepharose beads were added to capture antibody-chromatin complexes. Samples were washed sequentially for 10 min each with the following solutions at 4°C: TSE I buffer, TSE II buffer, and Buffer III (0.25 M LiCl, 1% IGEPAL, 1% DOC, 1 mM EDTA, 10 mM Tris HCl). Beads were washed three times with TE, pH = 8.0 before adding fifty microliters 10 mM DTT and a thirty minute incubation at 37°C. Following incubation, each sample was split in half to distinguish ChIP and reChIP aliquots. ChIP samples were eluted in 300 microliters elution buffer overnight at 65°C. reChIP samples were diluted 4× in dilution buffer. Five microliters reChIP antibody was added to each reChIP sample before overnight incubation at four degrees Celsius. reChIP antibody-chromatin complexes were precipitated as described above. Following elution, DNA was purified from each sample via phenol: chloroform extraction followed by ethanol precipitation.

ChIP-seq analysis

HP1 binding sites from third instar larvae and S2 cells were downloaded from GEO (see accession numbers in Supplementary

Table S1). Peak genomic coordinates were converted from dm3 to dm6 using the UCSC genome liftOver tool (Kent *et al.* 2002) and compared with annotated protein-coding genes in the Drosophila genome (release 6.25; Thurmond *et al.* 2019) to classify genes as bound. Chromatin context boundaries to differentiate heterochromatin and euchromatin were obtained from (Riddle *et al.* 2011). Enrichment of bound genes across chromatin contexts was evaluated using a Chi-square test.

To generate genome-wide binding profiles of HP1 proteins and histone modifications, we downloaded raw sequencing data (see accession numbers in Supplementary Table S1). Reads were aligned to the dm6 reference assembly using the bwa mem algorithm (version 0.7.16a-r1181) (Li and Durbin 2009). Coverage was calculated with samtools version 1.5 (Li *et al.* 2009) and plotted using Circos (Krzywinski *et al.* 2009). Metagene profiles and pausing indices were generated using deepTools (version 3.0.2; Ramirez *et al.* 2016).

RNA-seq analysis

For preparation of transcriptomic data from *HP1c* null mutants, 20 mg of frozen third instar larvae were homogenized, and RNA samples were isolated using Trizol. RNA sample integrity was confirmed by formaldehyde agarose gel electrophoresis. RNA samples were prepared for whole transcriptome sequencing by the UAB Heflin Center for Genomic Science Genomics Core lab. 30–40 million RNA-seq reads were collected per sample using the Illumina Sequencing Platform. We analyzed two RNA-seq samples of the *HP1c* null mutant genotype (GEO accession GSE180302).

To analyze RNA-seq data, we aligned reads to the dm6 reference genome assembly using STAR aligner (version #2.5.2) (Dobin *et al.* 2013) and determined transcript counts using HTSeq (version #0.6.1) (Anders *et al.* 2015). Differential expression analysis was performed using DESeq2 (version #1.22.2) (Love *et al.* 2014). Only genes meeting an FDR (false discovery rate) cutoff of 0.05 were used for downstream analyses. Gene ontology (GO) analysis was performed using DAVID (version #6.8) (Huang da *et al.* 2009a,b).

Motif analysis

We defined promoter regions as the region covering 250 bp upstream of the TSS to the TSS. Motif analysis of promoter sequences was evaluated using MEME (version 5.1.0), (Bailey *et al.* 2015) searching for the top five hits in each dataset.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. All data used in this study are publicly available and referenced in *Materials and Methods*. The Supplementary material is available at figshare: https://doi.org/10.25386/genetics.14346722. R code for the calculation of pausing indices and metagene profiles is available on Github: https://github.com/schoelz-j/schoelz_feng_riddle_2020.

Results

Drosophila HP1 proteins are enriched in heterochromatin, but also bind throughout euchromatin

In order to better understand the function of the Drosophila HP1 family in transcriptional regulation, we set out to identify endogenous targets for all three somatic HP1 family members in the Drosophila genome: HP1a, HP1B and HP1C. We began by reanalyzing existing ChIP-seq and ChIP microarray datasets for HP1a, HP1B, and HP1C generated by the modENCODE consortium (Ho *et al.* 2014). We examined data from a total of six different biological sources: three tissue types (adult heads, third instar larvae, and embryos) and three cell-types (S2-DRSC—isolate from male late embryonic tissue; BG3-c2—isolate from male larval central nervous system; and CME W1.c8—male isolate from the third instar larval wing imaginal disc). A genome-wide comparison of gene binding activity by HP1 proteins across samples revealed that HP1 targeting to genes was cell-type specific rather than constitutive (Figure 1A). Only 116 genes were constitutively targeted by the same combination of HP1 proteins in all biological sources, even though on average HP1a, HP1B, and HP1C occupied 2538, 6278, and 5877 genes, respectively. Furthermore, we observed a reduced number of HP1 target genes in cell cultures of individual cell-types compared with tissue types where multiple cell-types are present (Figure 1A). These results further support the conclusion that binding of HP1 proteins at individual genes is cell-type specific.

Given that binding behavior of HP1 proteins at genes is cell-type specific, we focused our analysis first on data from cell cultures. We compared the genome-wide distributions of HP1a, HP1B, and HP1C in S2-DRSC cells (Figure 1B). We verified significant enrichment of HP1a (blue track, outer circle) within pericentric heterochromatin and chromosome four, observing 64.60% of HP1a enriched regions resided in regions of the reference genome assembly categorized as heterochromatic (gray wedge highlights; Kharchenko *et al.* 2011), although a significant fraction (35.39%) of remaining HP1a enriched regions resided in euchromatin (Figure 1C). In contrast, a majority (94.89%) of HP1B (green track, middle circle) as well as HP1C (purple track, inner circle; 94.19%) enriched regions resided in euchromatic portions of the reference assembly (Figure 1, B and C). We observed a similar pattern in CME W1 cells (Supplementary Figure S1, A and B), where 26.07% of HP1a enriched domains, compared with 91.86% and 94.64% of HP1B and HP1C enriched domains, resided in euchromatic compartments. Meanwhile, in BG3 cells a majority of HP1a enriched domains were euchromatic (55.73%; Supplementary Figure S3, A and B) while the proportion of euchromatic HP1B and HP1C enriched domains was similar to levels observed in other cell-types (97.00% and 97.36%, respectively). This observation was despite the fact that there were overall a greater number of HP1a enriched regions in BG3 cells (5916) than in S2 cells (3926) or CME W1 cells (1386). However, we did detect that HP1a S2 cell enriched domains are slightly larger than domains in HP1a BG3 domains ($P = 2.707e\text{-}06$, Mann–Whitney). We did not detect a significant difference in the size of HP1a enriched domains between BG3 cells and CME W1 cells ($P = 0.6617$, Mann–Whitney). Therefore, our findings that a greater percentage of HP1a enriched domains reside in euchromatic portions of the genome in BG3 cells appears to be a result of an overall greater number of smaller HP1a enriched domains in this cell type. It is important to note, however, that binding of HP1 proteins in heterochromatic regions is an underestimation due to difficulties of mapping sequences to repeat-dense heterochromatin, and due to the exclusion of heterochromatic satellite regions from the genome assembly. Despite this caveat, while in the literature HP1a is often characterized as a heterochromatin protein and HP1C as a euchromatin protein, all three somatically expressed HP1 proteins in Drosophila are found throughout both chromatin compartments, although their binding enrichment differs across compartments.

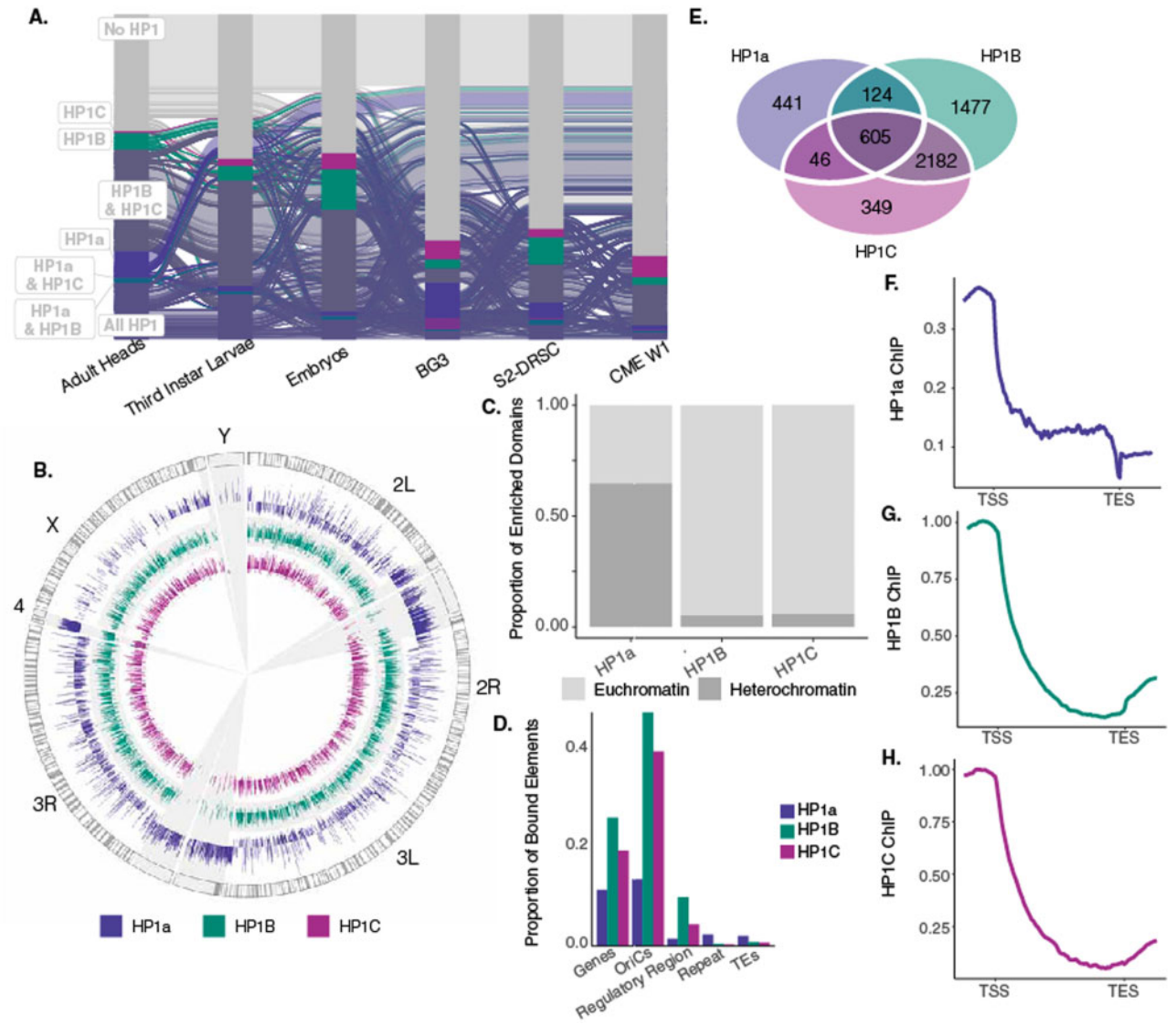

Figure 1 HP1 proteins co-occupy a large number of genes in the Drosophila genome. (A) Alluvial diagram comparing genic binding targets of HP1 proteins in six different biological sources. Stacked bars represent the number of genes targeted by a particular combination of HP1 proteins. Connections between bars represent changes in binding activity across tissue and cell types and are color-coded to represent a gene's classification in the "adult heads" tissue sample. Line weight denotes number of genes in each group. Large differences between tissues and cell types suggests binding of HP1 proteins at genes is cell-type specific.(B) Circos plot showing genome-wide binding of HP1a (outer track, blue), HP1B (middle track, green) and HP1C (innermost track, pink) in S2 cells. All three proteins are observed to bind throughout euchromatic (white background) and heterochromatic (gray wedges) of the genome. (C) Quantification of the proportion of enriched domains in different chromatin compartments for each HP1 family member. Because this measure only evaluates assembled regions of the Drosophila reference genome, binding of proteins in heterochromatin may be underestimated. (D) Binding activity of HP1 proteins at different classes of sequence elements in S2 cells. Due to mappability limitations, binding at repeat regions and TEs may be underestimated. (E) Overlap between occupancy of HP1 family members at genes bound by at least one HP1 protein in Drosophila S2 cells. A majority of these genes are bound by at least two HP1 proteins. (F–H) Metagene profiles for HP1a, HP1B, and HP1C demonstrate that HP1 proteins occupy TSSs.

HP1 proteins can be differentiated by their binding behavior at DNA sequence elements

To further examine the three HP1 proteins, we also looked at their tendency to localize to different DNA sequence elements. We investigated HP1 protein binding behavior at five different classes of DNA elements annotated in the Drosophila genome assembly (release dmel r6.25): genes, origins of replication (OriCs), regulatory regions, repeat regions, and TEs. For each DNA element, we measured the proportion of elements that overlapped with the binding site of an HP1 protein. In S2 cells, HP1a bound the largest fraction of repeats and TEs among the three HP1 proteins, occupying ~2.28% and 1.98% of these elements respectively (Figure 1D). In contrast, HP1B and HP1C occupied <1% of all TEs and repeats (Figure 1D). In CME W1 cells, we observed HP1a occupying 1.5% of annotated repeats and <1% of TEs (Supplementary Figure S1C), whereas HP1B and HP1C occupied <1% of both repeats and TEs. In BG3 cells, HP1a targeted 1.7% of repeat regions and 1.3% of TEs, whereas HP1B and HP1C targeted <1% of TEs and repeat regions (Supplementary Figure S3C). Difficulty in mapping repeat-dense portions of the genome again may account for lower than expected proportions of HP1a at annotated repeats and TEs. Interestingly, OriCs marked a stark difference in HP1 binding behavior for the three proteins examined. HP1B and HP1C were present at ~48% and 39% of all OriCs in S2 cells,

respectively, whereas HP1a was present at 13% (Figure 1D). In CME W1 cells, HP1B and HP1C were present at 17% and 21% of OriCs, respectively, and HP1a was present at 5% (Supplementary Figure S1C). This trend of increased HP1B and HP1C at OriCs relative to HP1a was consistent in BG3 cells, where HP1a was present at 10% of OriCs while HP1B and HP1C targeted 21% and 33%, respectively (Supplementary Figure S3C). In S2 cells, HP1a occupied 11% of genes, whereas HP1B and HP1C occupied a larger proportion of 26% and 19%, respectively (Figure 1D). In CME W1 cells, HP1a targeted 4% of genes, HP1B targeted 17% and HP1C targeted 21% (Supplementary Figure S1C). In BG3 cells, HP1a targeted 17% of genes, HP1B targeted 10% of genes, and HP1C targeted 16% of genes (Supplementary Figure S3C). All three HP1 proteins occupied less than 1% of annotated regulatory regions in S2 cells (Figure 1D). HP1B and HP1C consistently bound <1% of annotated regulatory regions in CME W1 cells and BG3 cells as well (Supplementary Figures S1C and S2C), whereas HP1a targeted 1.5% and 1.7% of regulatory regions in CME W1 and BG3 cells, respectively (Supplementary Figures S1C and S2C). Thus, HP1 proteins can be differentiated by their tendency to localize to different DNA sequence elements, although these tendencies are not absolute and vary across cell-types.

The HP1 family colocalize extensively at genic binding sites

Next, we set out to create a comprehensive list of HP1 binding targets to quantify the extent to which HP1 proteins share binding sites at protein-coding genes. We identified two particular combinations of HP1 proteins that occur at high frequencies: colocalization of HP1B and HP1C as well as colocalization of all three HP1 proteins (Figure 1E). Genes occupied by these combinations of proteins accounted for a majority of HP1B and HP1C genic binding targets: 63% of HP1B binding targets and 87% of HP1C binding targets were classified in either of these two categories. These genes also represented a large fraction (49%) of HP1a binding targets in S2 cells. Colocalization analysis of CME W1 cells and BG3 cells demonstrated that these combinations were frequent also in those cell-types (Supplementary Figures S1D and S2D). Metagene profiles of TSSs of HP1 target genes demonstrated that HP1 proteins bind at the TSS of their target genes (Figure 1, F–H). Binding of HP1 proteins at TSSs was consistent in CME W1 and BG3 cells (Supplementary Figures 1, E–G and 2, E–G). We validated colocalization of HP1 proteins at TSSs using sequential ChIP followed by PCR (Supplementary Figure S2). Of the seven genes surveyed, we observed colocalization of all three HP1 proteins at five: *Aef1*, *Chromator*, *gurken*, *Su(Z)-2*, and *RpL15* (Supplementary Figures 2, A–C and F–G). At the two remaining genes, *light* and *rolled*, we did not detect colocalization of HP1 proteins in sequential ChIP-samples despite positive signal in individual IP samples. These findings support the colocalization of HP1 proteins at TSSs. Given the extensive colocalization of these proteins across the genome, as well as their opposing actions on transcription (Li *et al.* 2003; Font-Burgada *et al.* 2008), these results highlight the need to better understand how HP1 proteins work in concert at TSSs. The significance of colocalization of HP1 family members currently is not understood, but it has been suggested previously that HP1 family proteins may display some degree of functional compensation (Ryu *et al.* 2014), particularly between HP1B and HP1C. All three HP1 proteins co-immunoprecipitate as well as form dimers through the chromoshadow-domain (Lee *et al.* 2019). It is unknown how these interactions affect gene expression.

Simultaneous HP1 binding is a better indicator of transcriptional activation than broader chromatin domains

To gain additional insights into the functions of the HP1 proteins in gene regulation, we characterized the protein-coding genes bound by HP1 proteins. We compared levels of expression between HP1 target and nontarget genes using publicly available RNA-Seq data from S2, CME W1, and BG3 cells (Ho *et al.* 2014) (Figure 2, A–D, Supplementary Figures S4, A–D and S5, A–D). We sought to determine whether the combination of HP1 proteins present at the TSS or the surrounding chromatin context was predictive of gene expression. In S2 cells, we found that HP1a heterochromatic targets were expressed at higher levels than nontarget genes (Figure 2A, Wilcoxon test, $W = 161{,}820$, $P = 0.0205$), whereas HP1a euchromatic targets were expressed at lower levels than nontarget genes (Figure 2A, $W = 35{,}469{,}767$, $P = 0.0003$). We did not find any significant difference between HP1a targets and nontargets in CME W1 cells in heterochromatin, whereas in euchromatin HP1a target genes displayed significantly lower expression (Supplementary Figure S4A, Wilcoxon test, $W = 16{,}408{,}411$, $P = 8.485\text{e-}12$). In BG3 cells, we found a similar pattern wherein euchromatic HP1a targets were expressed at lower levels than nontarget genes (Supplementary Figure S5A; Wilcoxon $W = 23{,}308{,}810$, $P = 8.467\text{e-}16$) but did not detect a significant difference in expression between heterochromatic HP1a target and nontarget genes. In summary, in all cell types examined euchromatic HP1a genic targets were expressed at lower levels than nontarget genes, whereas heterochromatic HP1a targets were expressed at similar or higher levels than nontarget genes.

We next examined differences in expression levels among HP1B and HP1C targets in heterochromatin and euchromatin across all three cell types. In S2 and CME cells, HP1B targets were expressed at higher levels than nontarget genes in both heterochromatin and euchromatin (Figure 2B and Supplementary Figure S4B). However, in BG3 cells HP1B targets were expressed at lower levels than nontarget genes in both chromatin contexts (Supplementary Figure 5B). We found that HP1C targets were expressed at higher levels regardless of chromatin context in both S2 and CME cells (Figure 2C and Supplementary Figure S4C). In BG3 cells, we found that HP1C targets were expressed at higher levels than nontarget genes in heterochromatin but were expressed at slightly lower levels than nontarget genes in euchromatin (Supplementary Figure S5C). In summary, chromatin context did not predict transcriptional activity of HP1B and HP1C target genes in the cell types examined. Together, these findings do not support a model where the effects of HP1 binding on expression are influenced by the surrounding chromatin compartment. HP1 targets are not expressed at consistently lower levels than nontarget genes in euchromatin and are not expressed at consistently higher levels than nontarget genes in heterochromatin.

Next, we examined whether considering combinations of HP1 family members present at gene promoters may predict transcriptional activity. In S2 cells, we found that groups of genes bound by different combinations of HP1 proteins differed significantly in their expression (Figure 2D, Kruskal–Wallis test $X^2 = 1542.5$, $P < 2.2\text{e-}16$). *Post-hoc* analysis demonstrated that genes bound by HP1B and HP1C were expressed at significantly higher levels than genes without any HP1 protein present, while genes bound exclusively by HP1a were expressed at significantly lower levels (Figure 2D). In CME W1 cells, we also detected significant differences in expression across different combinations of HP1 proteins (Supplementary Figure S4D, Kruskal–Wallis test $X^2 = 1004.6$, df = 7, $P < 2.2\text{e-}16$). *Post-hoc* analysis showed that genes

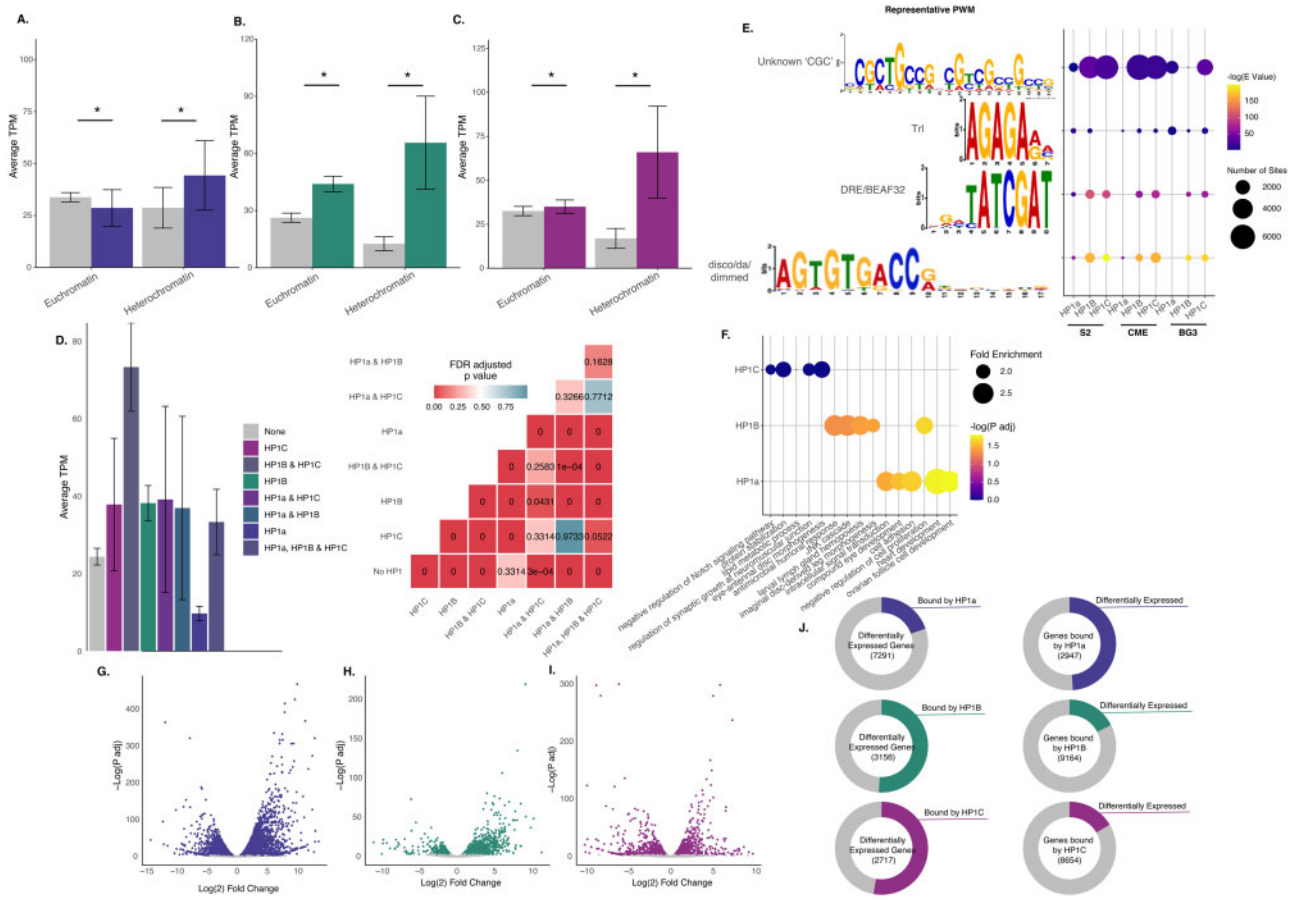

Figure 2 HP1 targets are highly expressed genes involved in signaling and development. (A–C) Comparison of expression (TPM) of HP1a (A), HP1B (B), and HP1C (C) target genes relative to nontarget genes in S2 cells across euchromatic and heterochromatic chromatin contexts. (D) Expression analysis of HP1 target genes across different combinations of HP1 family member occupancies. Adjusted *P*-values of pairwise comparisons (Mann–Whitney tests) across groups are presented in accompanying heatmap. (E) Results of motif enrichment analysis of promoter regions of genes targeted by HP1 proteins in different cell types for common motifs enriched in multiple cell types. Enrichment is evaluated by both E-value and number of motif occurrences across all genic targets. (F) GO enrichment of HP1 target genes in S2 cells. Five terms with lowest *P*-values in each dataset were selected for inclusion and examined for overlaps among remaining datasets. (G–I) Differential expression of protein-coding genes in homozygous null mutant larvae with depletion of HP1a (G), HP1B (H), or HP1C (I). Genes with significant changes in expression are plotted in color, whereas nonsignificant expression changes are plotted in gray. (J) Overlap analysis between HP1 binding targets and differentially expressed genes. On the left, differentially expressed genes in respective HP1 null mutant third instar larvae are classified as binding targets or nontarget genes. On the right, genic binding targets of respective HP1 proteins are classified as differentially or nondifferentially expressed in null mutant larvae.

bound by HP1a or by a combination of HP1a and HP1B had significantly lower expression than other groups of genes, whereas genes bound by any other combination of HP1 proteins at the promoter had significantly higher expression than nontarget genes. In BG3 cells, we again found that different combinations of HP1 proteins at the promoter displayed different levels of expression (Supplementary Figure S5D, Kruskal–Wallis test, $X^2 = 806.98$, df $= 7$, $P < 2.2e\text{-}16$). *Post-hoc* analysis identified genes bound exclusively by HP1C displaying the highest levels of expression. Similar to results from S2 and CME cells, genes bound exclusively by HP1a displayed significantly lower expression. Overall, these results across all three cell types indicate that the combination of HP1 proteins present at the gene promoter is important for gene expression, but that other factors act as well, likely in a cell-type specific manner.

Functional annotation of HP1 binding targets supports a cell-type specific function for HP1 activity

Next, we performed GO analysis (Huang da *et al.* 2009a,b) to further characterize endogenous HP1 binding targets, focusing on the biological process category of GO terms. Among HP1 binding targets in S2 cells, we identified significant enrichment for terms related to development and signaling such as "negative regulation of Notch signaling pathway" (HP1C; Figure 2F), "eye-antennal disc morphogenesis" (HP1C; Figure 2F), "imaginal disc-derived leg morphogenesis" (HP1B; Figure 2F), "intracellular signal transduction" (HP1a; Figure 2F), and "ovarian follicle cell development" (HP1a; Figure 2F) among others. Significant enrichment of terms relating to cell signaling and development were consistent across CME W1 (Supplementary Figure S4F) and BG3 (Supplementary Figure S5F) cells. In CME W1 cells, we identified several enriched terms under this classification including "open tracheal system development," "protein phosphorylation," "R7 cell development," and "chemical synaptic transmission." In BG3 cells, we again identified several enriched terms under this classification including "morphogenesis of an epithelium', 'mesoderm development," "compound eye development," and "determination of adult lifespan." Overall, results from GO analyses support again a cell-type specific function for HP1 proteins.

HP1 target promoters share enriched DNA sequence motifs across cell-types

HP1a binding to gene promoters has been suggested to be independent of its H3K9me2/3 reader activity (Cryderman *et al.* 2005),

and HP1C is known to be targeted to chromatin by the DNA binding Zinc Finger transcription factors WOC and ROW (Font-Burgada *et al.* 2008; Kessler *et al.* 2015; Di Mauro *et al.* 2020). Therefore, we performed a motif analysis (Bailey *et al.* 2015) of promoters of HP1 binding targets to identify putative regulatory sequences that may be important for targeting HP1 to protein-coding genes. We looked for enriched motifs in HP1-bound promoters, defining the promoter as the region 100 bp upstream of the TSS to 50 bp downstream of the TSS. We limited our analysis to the top eight enriched motifs in each promoter set and subsequently identified sequence motifs that were common among sets of genes targeted by either HP1a, HP1B, or HP1C in S2 cells, BG3 cells, and CME cells. We identified four motifs that appeared frequently throughout all nine datasets (Figure 2E). These included a motif recognized by GAGA factor encoded by the *Trithorax-like* gene, a motif recognized by BEAF-32 that closely resembles the downstream recognition element (DRE) motif, a motif recognized by a diverse set of transcription factors including *disco*, *dimmed*, and *da*, and finally an unknown trinucleotide "CGC" motif. Of all four motifs, the unknown CGC motif appeared most frequently throughout all datasets, although it had less significant E values than other motifs. The *disco/da/dimmed* motif consistently had the most significant E values (number of expected hits due to random chance) of all common motifs examined and was observed also to have a stronger association with HP1B and HP1C than with HP1a. This association pattern was observed also for the DRE/BEAF-32 motif. Finally, although significantly enriched, the *Trl* motif appeared at a smaller number of sites than other common motifs. These results suggest that despite cell-type specific differences in activity, HP1 family proteins may have a common targeting mechanism to promoters across cell-types.

HP1 depletion impacts gene expression

To understand how HP1 proteins regulate gene expression, we integrated three RNA-Seq datasets of HP1 knockout mutants. We utilized available datasets of transheterozygous $Su(var)205^{04}$/$Su(var)205^{05}$ and $HP1b^{86}$ knockout mutant third instar larvae (Riddle *et al.* 2012; Mills *et al.* 2018) and generated a novel library to study gene expression in homozygous $HP1c^{f04929}$ knockout mutant third instar larvae (Figure 2, G–J). We then compared differentially expressed genes across all three datasets to better understand the set of genes regulated by the HP1 family. We found that depletion of HP1a and HP1B resulted in upregulation of a large number of genes and a smaller quantity of downregulated genes, whereas depletion of HP1C resulted in both up- and downregulated gene expression at approximately equal levels (Figure 2, G–I). Next, we examined changes in gene expression upon HP1 depletion at genes bound by HP1 proteins. We found that 48.83% of HP1a bound genes were differentially expressed upon HP1 depletion (Figure 2J). A majority of expression changes observed upon HP1a depletion appear to be due to secondary effects, evidenced by the fact that only 19.73% of differentially expressed genes were binding targets (Figure 2J). A chi-square test of independence showed that there was a significant association between HP1a binding and differential expression (X^2 = 106.22, $P < 2.2E\text{-}16$). In contrast, we found that HP1B and HP1C binding targets constituted a small majority of differentially expressed genes, although only a small percentage of binding targets was differentially expressed (Figure 2J). Upon HP1B depletion, 50.95% of differentially expressed genes are bound by HP1B under wildtype conditions, although only 17.50% of binding targets were differentially expressed (Figure 2J), but a chi-squared test of independence showed that this association was not significant (X^2 = 0.8067, $P = 0.36$). Similarly, 52.78% of differentially expressed genes upon HP1C depletion are genes bound by HP1C under wildtype conditions, but only 16.57% of HP1C binding targets are differentially expressed upon HP1C depletion (Figure 2J). This association was found to be significant (X^2 = 21.223, $P < 2.2E\text{-}16$). Therefore, while a majority of HP1B and HP1C binding targets do not experience significant changes in expression upon depletion of either respective protein, those genes which are differentially expressed constitute a small majority of observed transcriptional changes.

HP1 genic targets reside in particular chromatin states

Chromatin frequently is classified into higher-order states beyond heterochromatin and euchromatin based on the varying compositions of histone modifications and chromatin-binding proteins (Filion *et al.* 2010; Kharchenko *et al.* 2011). To gain a better understanding of the localization patterns of the different HP1 family members, we determined the extent to which they targeted genes in nine different chromatin states in Drosophila S2 cells and BG3 cells defined by the modENCODE consortium (Kharchenko *et al.* 2011). In general, in both cell-types we found that a majority of protein-coding genes reside in chromatin states one, two, three, four, and nine, which correspond to the euchromatic compartment of the genome (Figure 3, A and E for S2 and BG3 cells, respectively). A chi-square analysis showed that HP1 family genic binding targets were not distributed independently of overall chromatin state (X^2=1762.1, $P < 2.2E\text{-}16$; X^2=2559, $P < 2.2E\text{-}16$; X^2=2024.2, $P < 2.2E\text{-}16$ for HP1a, HP1B, and HP1C, respectively). Inspection of residuals demonstrated that HP1a genic binding targets were markedly enriched for chromatin state seven (corresponding to pericentric heterochromatin) and were enriched also to a lesser degree for chromatin states eight and three (Figure 3B). In contrast, HP1a targets were depleted for chromatin state one (Figure 3B). Meanwhile, HP1B targets were enriched for chromatin state three and depleted for chromatin state nine (Figure 3C). We observed this same pattern of enrichment in chromatin state three and depletion in chromatin state nine among HP1C targets as well. Furthermore, these enrichment patterns were consistent in BG3 cells. Here, HP1a genic binding targets were again enriched for chromatin state seven and depleted for chromatin state one (Figure 3F). HP1B genic binding targets were enriched for chromatin state three and depleted for chromatin state nine (Figure 3, G and H). Overall, these results strengthen the association of the HP1 family with transcriptionally active chromatin domains.

HP1 binding targets display signatures of promoter proximal RNA polymerase pausing

All three somatic Drosophila HP1 proteins co-immunoprecipitate with both subunits of the FACT complex, which promotes transcriptional elongation (Kwon *et al.* 2010). Furthermore, HP1C has been implicated previously in release from promoter proximal pausing (Kessler *et al.* 2015), and HP1a and HP1C have been associated with pausing at transcribed genes (Sakoparnig *et al.* 2012). However, the extent to which this relationship depends on the simultaneous activity of other HP1 proteins has not been examined. To better understand the association between the HP1 protein family and transcriptional pausing by RPII (referred to as 'pausing'), we compared RPII dynamics at HP1 target and nontarget genes. Metagene profiles of Global Nuclear Run-On (GRO-Seq) data from S2 cells demonstrated that HP1 target genes generally

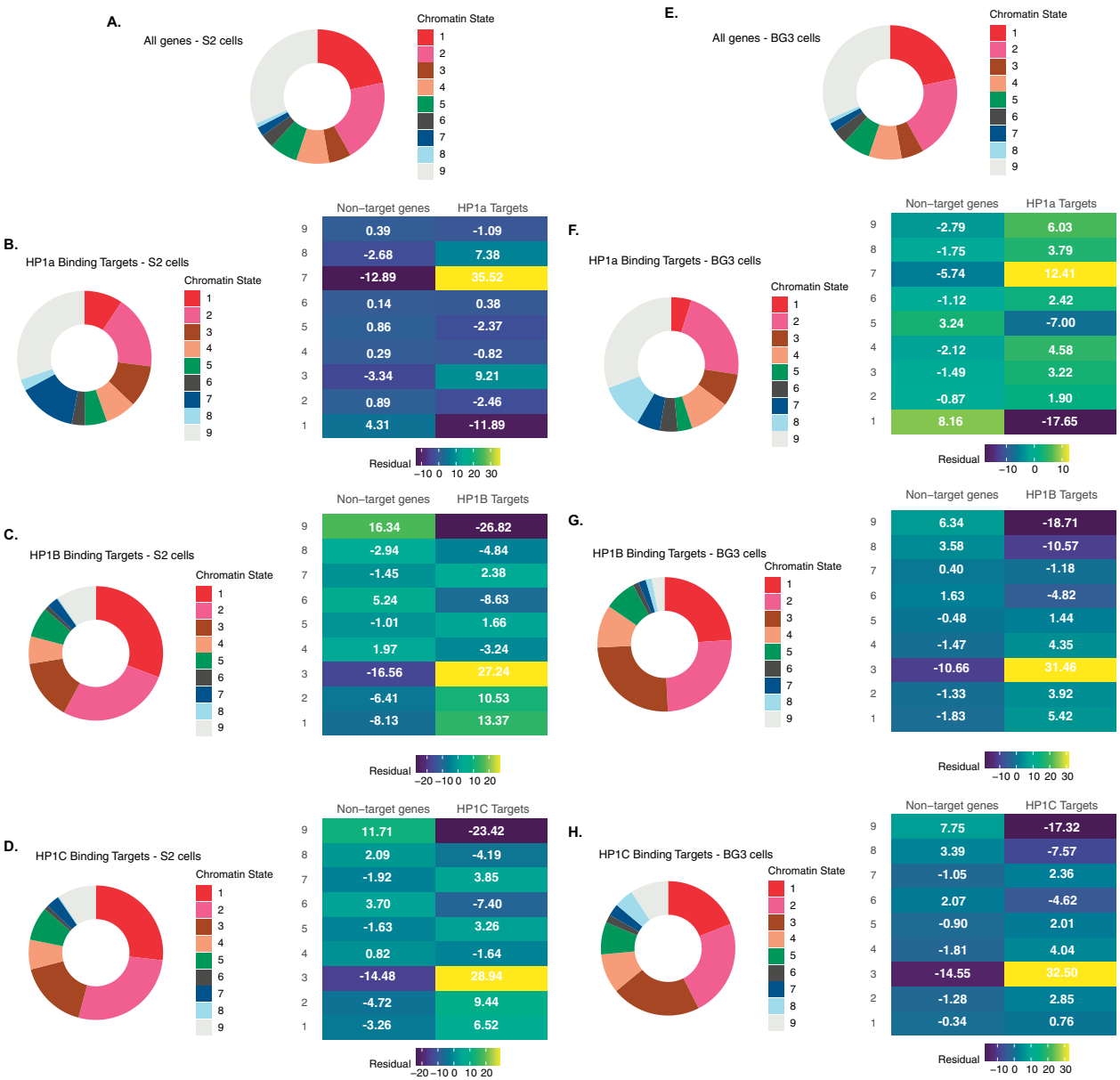

Figure 3 HP1 binding activity across modENCODE chromatin states. (A) Donut plot comparing number of genes in different modENCODE chromatin states in S2 cells. Genes were classified into a single chromatin state based on which state overlapped with a majority of the genomic coordinates. (B–D) Donut plots comparing the distribution of HP1a (B), HP1B (C), and HP1C (D) genic binding targets in S2 cells across different modENCODE chromatin states. Residuals from chi-square analysis are presented in corresponding heatmaps with residual values. (E) Donut plot comparing number of genes in different modENCODE chromatin states in BG3 cells. (F–H) Donut plots comparing the distribution of HP1a (F), HP1B (G), and HP1C (H) genic binding targets in BG3 cells across different modENCODE chromatin states. Residuals from chi-square analysis are presented in corresponding heatmaps.

displayed a higher 5′ nascent RNA signal peak in addition to overall increased RPII recruitment (Figure 4, A–C). Interestingly the increased 5′ signal at HP1a target genes was slightly upstream of the TSS, and these genes had a decreased signal peak over the TSS (Figure 4A). HP1B and HP1C had increased nascent RNA signal peak at the TSS as well as over the gene body (Figure 4, B and C). To quantify this relationship, we calculated pausing indices (Muse *et al.* 2007; Larschan *et al.* 2011). Here, each gene is divided into two regions (Figure 4D). A pausing index can be calculated by dividing the read density in the 5′ region over the read density in the mid-gene region. Pausing indices allow for the evaluation of RPII dynamics using next-generation sequencing datasets. We calculated pausing indices for HP1 target and nontarget genes using available GRO-Seq data from Drosophila S2 cells. We detected significantly increased pausing indices at HP1a target genes in S2 cells (Figure 4E, Wilcoxon, $W = 2{,}983{,}070$, $P = 3.317\text{e-}16$) as well as HP1B (Figure 4F, Wilcoxon, $W = 6{,}912{,}071$, $P < 2.2\text{e-}16$) and HP1C (Figure 4G, Wilcoxon, $W = 6{,}248{,}616$, $P < 2.2\text{e-}16$). target genes. These results demonstrate a higher level of RPII pausing at genes targeted by HP1 family members in S2 cells.

To corroborate the findings of our analysis of nascent transcriptomics in S2 cells, we also analyzed RPII activity at HP1 family target genes in BG3 and CME cells using available RPII ChIP-chip data published by the modENCODE consortium. This analysis recapitulated observations from GRO-Seq data in S2 cells. In CME cells, metagene profiles of HP1B and HP1C demonstrated an overall increase in RPII occupancy at HP1B and HP1C target genes (Supplementary Figures S5E and S6C), whereas HP1a

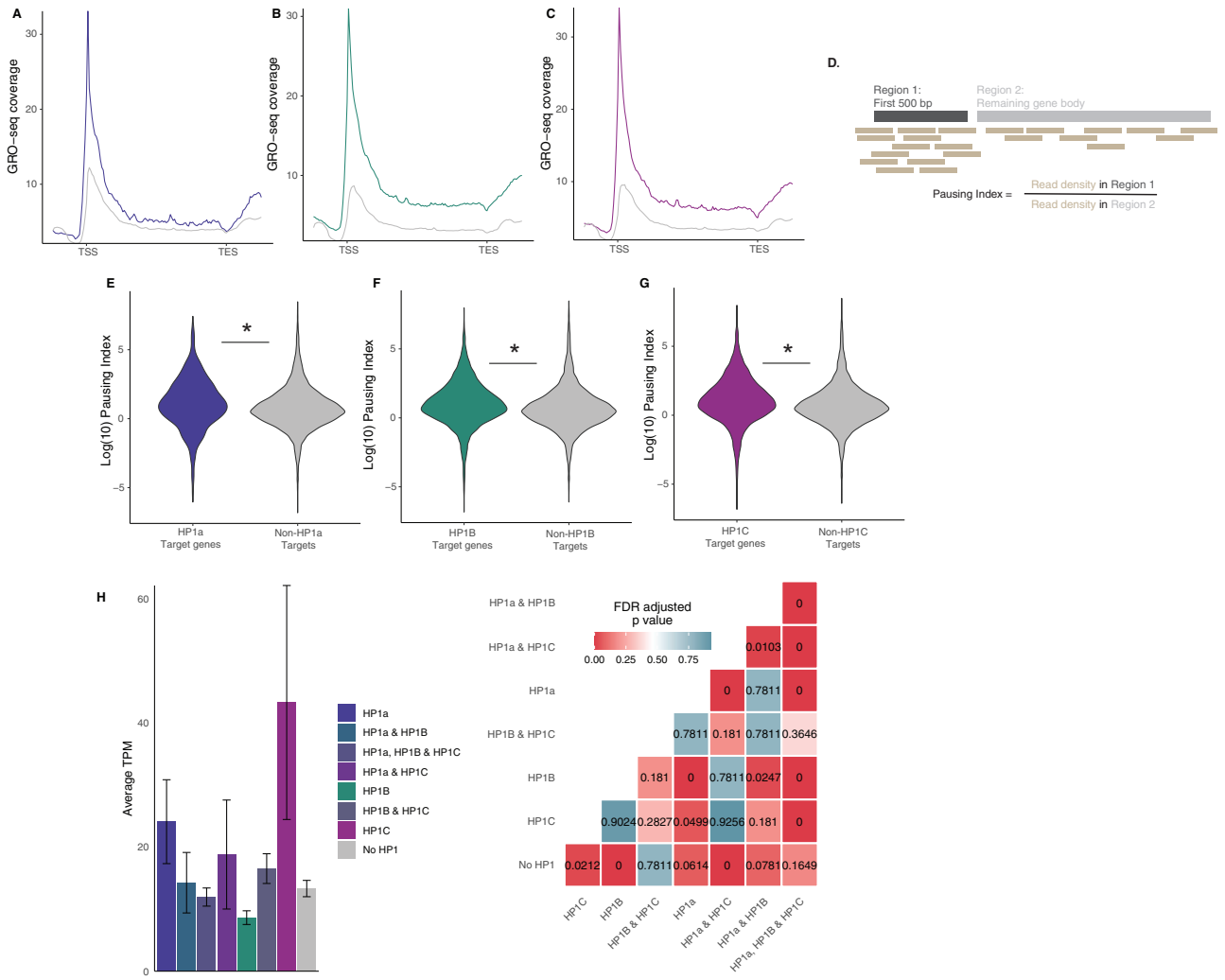

Figure 4 Nascent transcriptomics reveals increased pausing at HP1 target genes. (A–C) GRO-Seq metagene profiles of HP1 target genes (colored) and nontarget genes (gray) in S2 cells. The average GRO-seq signal across all genes in a group is plotted. Increased five prime GRO-seq signal at HP1 target genes suggests higher levels of nascent transcription at these genes. (D) Schematic of pausing index calculation. Briefly, the ratio of sequencing coverage in the promoter proximal region to the sequencing coverage in the gene body is used to estimate promoter proximal pausing. (E–G) Comparison of pausing indices at HP1 family target genes to nontarget genes in S2 cells. Differences in pausing indices were evaluated using Mann–Whitney nonparametric tests. (H) Comparison of pausing indices at genes occupied by different combinations of HP1 family members. Adjusted *P*-values presenting results of pairwise comparisons of pausing indices across groups are presented in the accompanying heatmap.

target genes displayed higher RPII enrichment over the TSS but did not exhibit increased enrichment over the gene body (Supplementary Figure S6A). To calculate pausing indices from these data, we used an alternative pausing index calculation that is compatible with ChIP-chip microarray datasets (Zeitlinger *et al.* 2007). We found significantly increased pausing at HP1a, HP1B and HP1C binding targets in CME cells (Supplementary Figures S5, D and F and S6B, Wilcoxon rank sum test with continuity correction: $W = 1{,}392{,}405$, $P = 1.541\text{e-}13$, $W = 6{,}079{,}364$, $P < 2.2\text{e-}16$, $W = 6{,}815{,}650$, $P < 2.2\text{e-}16$ for HP1a, HP1B, and HP1C, respectively). Repeating this analysis with RPII microarray data from BG3 data recapitulated these patterns. Again, HP1a target genes displayed increased enrichment over the TSS but not over the gene body (Supplementary Figure S7A), whereas HP1B and HP1C target genes displayed an overall increase in RPII enrichment compared with nontarget genes (Supplementary Figures S6E and S7C). Pausing indices were significantly higher at HP1a target genes (Supplementary Figure S7B, Wilcoxon rank sum test with continuity correction: $W = 4{,}532{,}193$, $P < 2.2\text{e-}16$), HP1B target genes (Supplementary Figure S7D, Wilcoxon rank sum test with continuity correction: $W = 4{,}984{,}778$, $P < 2.2\text{e-}16$), and HP1C target genes (Supplementary Figure 7F, Wilcoxon rank sum test with continuity correction: $W = 6{,}575{,}674$, $P < 2.2\text{e-}16$) relative to nontarget genes. Overall, our analysis of RPII dynamics at HP1 family target genes reinforces our findings from S2 cells that HP1 target genes show increased pausing.

To follow up this analysis, we examined how nascent transcription signatures vary across genes bound by different combinations of HP1 proteins. We analyzed pausing indices across these different gene groups using pairwise Wilcoxon tests with FDR correction (Figure 4H). Genes bound exclusively by HP1C and genes bound exclusively by HP1a had the highest mean pausing indices, whereas genes bound exclusively by HP1B had lower mean pausing indices than either of these groups or genes not targeted by any HP1 proteins. Overall, our analysis of nascent transcription dynamics at HP1 target genes supports a model of HP1 function wherein particular combinations of HP1 proteins

are consistently and strongly associated with transcriptional activation.

Depletion of individual HP1 proteins reveals roles for HP1 family members in promoter proximal pausing

To better understand the impact of HP1 binding on promoter proximal pausing, we measured pausing indices in knockout mutants for HP1a, HP1B, and HP1C using RPII ChIP-chip data from third instar larvae made available by the modENCODE consortium (Ho *et al.* 2014). We were able to detect significantly increased pausing at HP1a, HP1B, and HP1C target genes in wildtype Drosophila third instar larvae (Figures 5, A, C, and E). Overall, significantly increased promoter proximal pausing at HP1 target genes was maintained in respective knockout mutants (Figures 5, B, D, and F). This observation is consistent with a model where HP1 proteins cooperate to regulate transcription and exhibit a degree of functional redundancy at TSSs.

Analysis of pausing indices across genotypes suggests binding of HP1B and HP1C may be particularly important for transcriptional regulation by HP1 family members. To gain insight into individual functions of HP1 proteins in transcriptional regulation, we decided to examine how pausing indices changed across HP1 null mutants at genes bound by different combinations of HP1 proteins (Figure 5, G–J). We first compared pausing indices across HP1 binding groups in the wildtype dataset with functional copies of all three somatic HP1 genes to better appreciate how the groups relate to each other in the "wildtype" condition. A Kruskal–Wallis test confirmed that there were significant differences in pausing indices across HP1 binding groups ($X^2 = 526.62$, $P < 2.2e\text{-}16$), which we followed up with pairwise Wilcoxon tests with FDR correction to examine pairwise differences. We found a total of 13 significantly different pairwise comparisons between different HP1 binding groups which roughly partitioned the groups into three tiers (Figure 5G). Genes that were not bound by any HP1 proteins did not have a significantly different pausing index compared with genes bound exclusively by HP1a, and these groups had the lowest average pausing indices. A middle tier of groups was comprised of genes bound exclusively by HP1C, genes bound exclusively by HP1B, and genes bound by a combination of HP1a and HP1B but lacking HP1C. Groups in this tier had intermediate average pausing index values. Finally, genes bound by both HP1B and HP1C as well as genes bound by HP1a, HP1B, and HP1C did not exhibit significant differences in their pausing indices, and these genes had the highest average pausing indices. (The group of genes bound by HP1a and HP1C were not compared in pairwise comparisons because the bimodal distribution of pausing indices in this group precludes necessary assumptions for statistical inference). These results reinforce prior data suggesting that the colocalization of HP1B and HP1C may be particularly important for the increased pausing and increased expression that has been previously associated with HP1 binding.

Depletion of HP1a results in minor impacts to pausing indices at HP1 target genes. We repeated the above analysis in HP1a null larvae to infer the importance of HP1a in transcriptional regulation (Figure 5H). A Kruskal–Wallis test established significant differences in pausing indices across groups of genes bound by different combinations of HP1 family members ($X^2 = 564.96$, $P < 2.2e\text{-}16$). Follow-up of pairwise comparisons using Wilcoxon tests with FDR correction revealed two pairwise comparisons that deviated from the wildtype genotype. Genes bound exclusively by HP1C no longer exhibited significantly increased pausing indices upon depletion of HP1a. Instead, this group of genes now occupied the lowest tier of pausing indices. The second novel difference was that genes bound by HP1a and HP1B had significantly higher pausing indices than genes bound exclusively by HP1a upon HP1a depletion. However, this change did not meaningfully move this group of genes into a new tier of pausing indices. Although depletion of HP1a produced some changes in promoter proximal pausing at genes bound by certain combinations of HP1 family members, overall effects were minimal.

In contrast to HP1a depletion which resulted in minimal effects on promoter proximal pausing, depletion of HP1B disrupted promoter proximal pausing on a larger scale. A Kruskal–Wallis test of pausing indices across groups of genes bound by different combinations of HP1 proteins confirmed significant differences between groups ($X^2 = 137.12$, $P < 2.2e\text{-}16$). Pairwise Wilcoxon comparisons with FDR correction identified a total of eight comparisons that differed from their respective result in the wildtype genotype. Genes bound exclusively by HP1B or exclusively by HP1C no longer displayed significantly higher pausing indices compared with genes with no HP1 proteins present, contributing to the lowest tier of gene groups ranked by pausing-indices. Additionally, genes bound by a combination of HP1B and HP1C were not significantly different from genes bound exclusively by HP1C, although the former were still significantly different from genes with no HP1 proteins at all. Genes bound by HP1B and HP1C no longer occupied the highest tier of pausing indices upon depletion of HP1B and also exhibited significant differences with genes bound by all three HP1 proteins. The relationship between genes bound by HP1a and HP1B exhibited the most change in this genotype compared with pairwise comparisons in wildtype. Upon depletion of HP1B, these genes had higher pausing indices compared with genes bound exclusively by HP1a, HP1B, or HP1C. However, these genes were not significantly different from genes bound by HP1B and HP1C. These data suggest that HP1B may be particularly important for relationships between HP1 family members when regulating TSS activity and that HP1 family members may functionally compensate upon HP1B depletion.

Depletion of HP1C minimized differences in pausing indices across groups of HP1 genes. A Kruskal-Wallis test confirmed significant differences in pausing indices across groups of genes bound by different combinations of HP1 proteins upon depletion of HP1C ($X^2 = 271.35$, $P < 2.2e\text{-}16$). Pairwise comparisons using Wilcoxon tests with FDR corrections identified four pairwise comparisons whose relationship differed from the wildtype condition. Each of these comparisons represented a transition from a statistically significant difference to a nonsignificant difference following HP1C depletion. First, genes bound by HP1C were no longer significantly different from genes not bound by HP1 family members. The remaining three comparisons all involved the group of genes bound by HP1a and HP1B. This gene group was no longer significantly different from genes bound by all three HP1 proteins, genes bound by HP1B and HP1C, and genes not bound by HP1 proteins. This observation suggests that the presence of HP1C is important for regulating pausing when different combinations of HP1 proteins are present at TSSs.

Discussion

Here, we analyzed high-resolution ChIP-Seq and ChIP-chip maps of all the somatic Drosophila HP1 family members in three different cell-types, which raise interesting points about the role of these proteins in gene regulation. We find that all three HP1 proteins bind throughout heterochromatin and euchromatin compartments. With regards to binding behavior at protein-coding

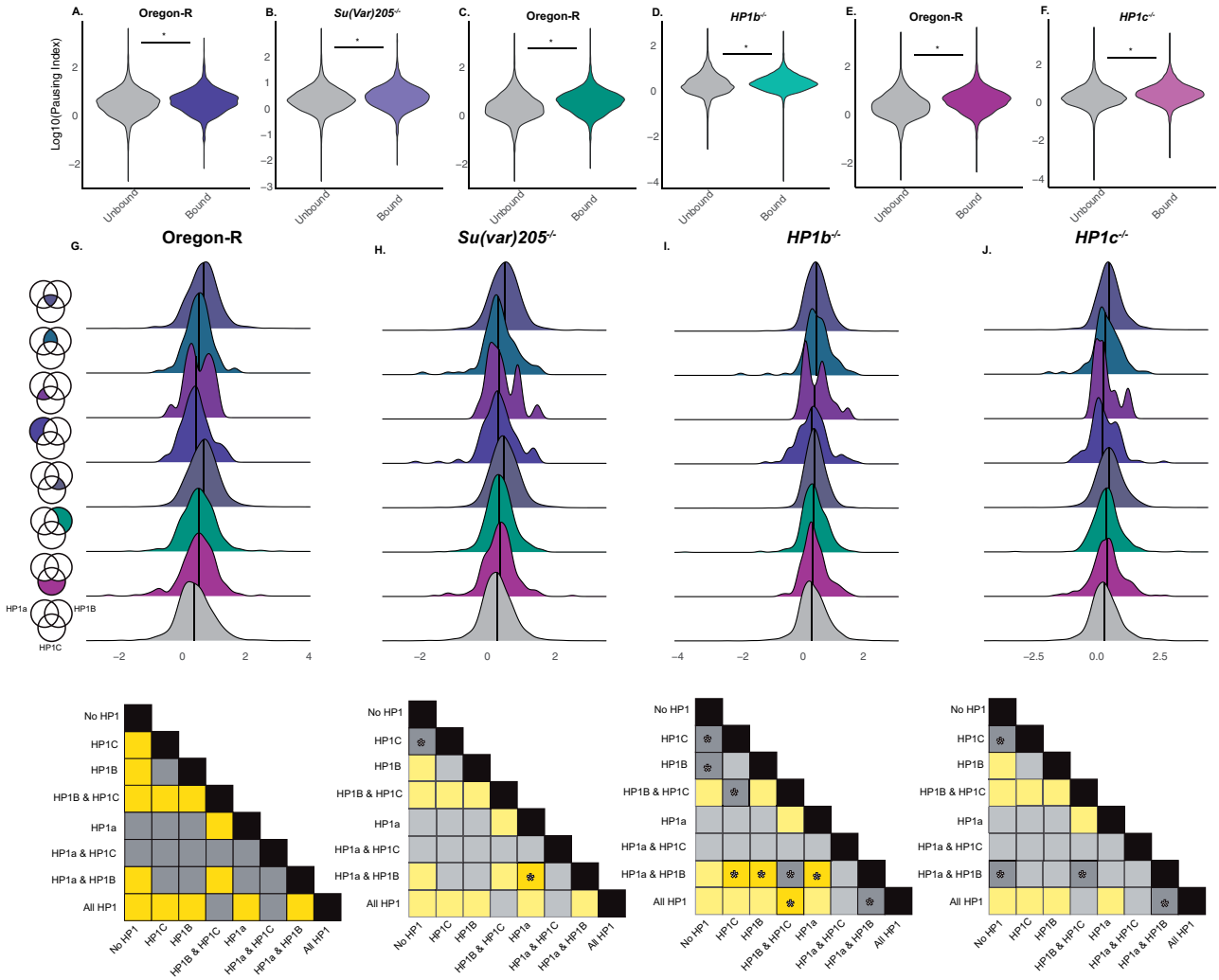

Figure 5 Depletion of HP1 proteins alters RPII pausing dynamics. (A–F) Comparison of pausing indices at HP1 target genes in Oregon-R and HP1 null mutant larvae. (A–B) HP1a; (C–D) HP1B; (E–F) HP1C. (G–J) Comparison of pausing indices at groups of genes bound by different combinations of HP1 proteins in different genotypes: (G) Oregon-R; (H) *Su(var)20504/Su(var)20505*; (I) *HP1b86/HP1b86*; (J) *HP1cf04929/HP1cf04929*. Venn diagrams present the combination of HP1 proteins binding a particular group of genes. Binding classifications are based on ChIP data originating from Oregon-R larvae. Adjusted *P*-values of pairwise comparisons across groups are summarized below each ridge plot. Gray cells denote nonsignificant comparisons and yellow cells denote significant differences. Within mutant genotype conditions, asterisks denote pairwise comparisons with different results from the wildtype condition (for instance, a comparison is significantly different in wildtype conditions but nonsignificant in a mutant genotype).

genes, while all three HP1 proteins are enriched at genes located within heterochromatin, a majority of their binding targets are located within euchromatin. This finding is true even of HP1a, whose localization often is described as restricted to heterochromatin, as well as HP1C, whose localization tends to be described as restricted to euchromatin. In addition to previously reported enrichment of HP1a on chromosome four, we also detect significant enrichment of HP1B and HP1C. Additionally, the three HP1 proteins share a large fraction of their binding sites at promoters. A gene bound by any HP1 protein likely is bound also by at least one other family member. This relationship was true across heterochromatin and euchromatin and highlights the need to consider what effect interactions between HP1 proteins have on transcription.

A close examination of HP1 genic binding targets suggests that knowledge of the presence of additional HP1 proteins is a better indicator of transcriptional status than knowledge of the broader surrounding chromatin context. HP1-bound genes are expressed at higher levels than unbound genes across chromatin contexts. Genes bound by all three HP1 proteins or by a combination of HP1B and HP1C are consistently expressed at higher levels across all contexts. HP1-bound genes display a strong association with H3K4me3 across all chromatin contexts but share a context-specific association with H3K9me2/3 within heterochromatin. The independence of HP1 binding to euchromatic genes from H3K9me2/3 matches previously observed data. Simultaneous binding of multiple HP1 proteins therefore appears to be a stronger indicator of transcriptional activation than chromatin context. However, this finding is more modest in some cell-types compared with others.

Signatures of promoter proximal pausing at HP1 binding targets give clues to a potential mechanism of gene activation by HP1 proteins. Here, we report that genes bound by HP1 proteins display higher pausing indices compared with unbound genes. A pausing index is an indirect measurement of RPII activity that reflects a higher density of RPII at the 5′ end of genes. It is not always clear what factors drive this increased density. For instance, genes with increased pausing durations would be expected to

have higher pausing indices and lower expression levels. In contrast, genes with shorter pausing durations but increased initiation frequencies could exhibit high pausing indices in addition to high expression levels (Gressel *et al.* 2017). HP1 binding targets are expressed at higher levels than nontarget genes and reside in transcriptionally active chromatin, in the support of the latter model of increased pausing indices. This observation is supported by observations made by others that HP1 binding targets appear to be both paused and highly transcribed (Sakoparnig *et al.* 2012). Increased pausing indices associated with HP1 binding may be due to the relationship between the HP1 family and the FACT complex, which interacts with nucleosomes destabilized by RPII promoting elongation (Orphanides *et al.* 1998; Kwon *et al.* 2010; Formosa and Winston 2020). Alternatively, increased RPII pausing at HP1 target genes may be regulated through HP1-mediated recruitment of additional factors such as dDsk2 (Kessler *et al.* 2015; Di Mauro *et al.* 2020). Additional evidence is necessary to fully understand the contribution of each HP1 family member to transcriptional activation.

Our analysis of RPII dynamics in single knockout HP1 mutants suggests that interactions between HP1 family members are important in the regulation of gene expression. HP1 targets comprise a majority of differentially expressed genes in *HP1b* and *HP1c* null mutants, and a large fraction of HP1a binding targets are differentially expressed in *Su(var)205* mutants. An analysis of RPII activity at these genes in respective HP1 null mutants supports a model where HP1 proteins promote increased gene expression through regulation of RPII activity. This model is further supported by an observed interaction between HP1 family members and the FACT complex and is consistent with observations of activity of the HP1 family member *rhino* in the Drosophila genome (Andersen *et al.* 2017). Our analysis builds on these results by providing insights into how HP1 proteins coordinate to regulate RPII activity in Drosophila somatic cells.

Although HP1a and HP1C previously have been implicated in transcriptional activation and promoter proximal pausing individually, ours is the first study to consider how coordinated activity between HP1 proteins may impact gene expression. Additionally, ours is the first study to show genome-wide evidence for a role of HP1B in promoter proximal pausing to induce transcription. Previous studies have suggested that surrounding chromatin contexts may predict whether HP1 proteins have an activating or repressive role at TSSs. However, our genome-wide analysis of HP1 binding targets demonstrates that colocalization of HP1 proteins is a better predictor of whether binding targets are transcribed or repressed than knowledge of surrounding chromatin context. Certain combinations of HP1 proteins, particularly the colocalization of HP1B and HP1C, are strongly associated with active transcription throughout heterochromatin and euchromatin, while HP1a binding on its own is not associated with pausing or transcription. Interestingly, although HP1a has been shown to be required for proper expression of heterochromatic genes, we found HP1a nontarget genes had higher expression than HP1a binding targets in heterochromatin S2 cells. This finding indicates that heterochromatic genes that have adapted to heterochromatic contexts may depend on other factors besides HP1a to maintain their expression. This finding was not replicated in CME W1 or BG3 cells, which suggests that this may be a cell-type specific effect in S2 cells. This hypothesis is further supported by our observation that HP1a binding targets are extensively misregulated upon HP1a depletion. Overall, our analysis highlights the need to consider how HP1 family members work together to regulate gene expression.

Given our analysis, future studies on the functions of HP1 family members in transcriptional regulation should employ careful experimental manipulation and thorough controls to identify individual contributions of HP1 family members to gene regulation. Nascent transcriptomic assays, such as PRO-seq or NET-seq, coupled with RNAi and ChIP-seq should be used to interrogate how transcription dynamics are affected by the presence and absence of different HP1 family members. Additionally, contemporary molecular tethering assays such as those utilizing dCas9 would allow for investigation of how the addition of HP1 proteins to different endogenous loci in the Drosophila genome affects transcription. These experiments would complement each other to yield significant new information on the functions of HP1 family members in transcription. Furthermore, a model for HP1 family members in transcriptional regulation should be informed by these proteins known structural properties and functions such as reading histone methylation, recruitment of additional chromatin-associated proteins and the establishment of phase-separated nuclear domains. It is essential to identify which of these functions underlies a mechanism of transcriptional activation by HP1 family members. For instance, phase separation of gene promoters has been proposed to be important for transcriptional activation. HP1a is known to drive lipid droplet formation while HP1B and HP1C are not predicted to. Given that HP1B and HP1C localize to HP1a at transcriptionally active regions, high-resolution imaging studies comparing the nuclear distributions of these proteins to HP1a droplets may partially inform whether these droplets may also be involved in HP1-mediated transcriptional activation. Whether other HP1 structural properties underlie transcriptional activation should also be investigated. HP1a, HP1B, and HP1C are known to recognize H3K9me2/3 histone modifications through the chromodomain, but whether recognition of histone methylation by the chromodomain is important for targeting HP1 proteins to transcriptionally active genes remains an open question. This question could be investigated through the use of histone modification binding arrays. Finally, HP1C is known to interact with factors such as WOC, dDsk2, and SSRP1 to help mediate transcriptional activation, but how HP1a and HP1B contribute to these interactions, or what other binding partners may be important for HP1-mediated transcriptional activation are unknown. Our work here provides a foundation for these future studies.

Acknowledgments

The authors would like to acknowledge the support of the Riddle lab members for stimulating discussions and comments on the manuscript. Additionally, the authors would like to thank the University of Alabama at Birmingham Heflin Genomics Core for assistance in generation and analysis of RNA-Seq datasets.

Funding

This study was supported by the National Science Foundation (award MCB-1552586). The authors acknowledge funding (to J.M.S.) from the Alabama Graduate Research Scholars Program (GRSP) funded through the Alabama Commission for Higher Education and administered by the Alabama EPSCoR (Established Program to Stimulate Competitive Research).

Conflicts of interest

None declared.

Literature cited

Alekseyenko AA, Gorchakov AA, Zee BM, Fuchs SM, Kharchenko PV, *et al.* 2014. Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1 and repetitive RNAs. Genes Dev. 28:1445–1460.

Amaral N, Ryu T, Li X, Chiolo I. 2017. Nuclear dynamics of heterochromatin repair. Trends Genet. 33:86–100.

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 31:166–169.

Andersen PR, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. Nature. 549:54–59.

Badugu R, Shareef MM, Kellum R. 2003. Novel Drosophila heterochromatin protein 1 (HP1)/origin recognition complex-associated protein (HOAP) repeat motif in HP1/HOAP interactions and chromocenter associations. J Biol Chem. 278:34491–34498.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. Nucleic Acids Res. 43:W39–W49.

Canzio D, Larson A, Narlikar GJ. 2014. Mechanisms of functional promiscuity by HP1 proteins. Trends Cell Biol. 24:377–386.

Cryderman DE, Grade SK, Li Y, Fanti L, Pimpinelli S, *et al.* 2005. Role of Drosophila HP1 in euchromatic gene expression. Dev Dyn. 232: 767–774.

Czermin B, Schotta G, Hulsmann BB, Brehm A, Becker PB, *et al.* 2001. Physical and functional association of SU(VAR)3-9 and HDAC1 in Drosophila. EMBO Rep. 2:915–919.

Danzer JR, Wallrath LL. 2004. Mechanisms of HP1-mediated gene silencing in Drosophila. Development. 131:3571–3580.

de Wit E, Greil F, van Steensel B. 2007. High-resolution mapping reveals links of HP1 with active and inactive chromatin components. PLoS Genet. 3:e38.

Di Mauro G, Carbonell A, Escudero-Ferruz P, Azorin F. 2020. The zinc-finger proteins WOC and ROW play distinct functions within the HP1c transcription complex. Biochim Biophys Acta Gene Regul Mech. 1863:194492.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, *et al.* 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29: 15–21.

Eissenberg JC, Elgin SC. 2014. HP1a: a structural chromosomal protein regulating transcription. Trends Genet. 30:103–110.

Eissenberg JC, James TC, Foster-Hartnett DM, Hartnett T, Ngan V, *et al.* 1990. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 87: 9923–9927.

Fanti L, Berloco M, Piacentini L, Pimpinelli S. 2003. Chromosomal distribution of heterochromatin protein 1 (HP1) in Drosophila: a cytological map of euchromatic HP1 binding sites. Genetica. 117: 135–147.

Figueiredo ML, Philip P, Stenberg P, Larsson J. 2012. HP1a recruitment to promoters is independent of H3K9 methylation in *Drosophila melanogaster*. PLoS Genet. 8:e1003061.

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, *et al.* 2010. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell. 143:212–224.

Font-Burgada J, Rossell D, Auer H, Azorin F. 2008. Drosophila HP1c isoform interacts with the zinc-finger proteins WOC and Relative-of-WOC to regulate gene expression. Genes Dev. 22: 3007–3023.

Formosa T, Winston F. 2020. The role of FACT in managing chromatin: disruption, assembly, or repair? Nucleic Acids Res. 48: 11929–11941.

Greil F, van der Kraan I, Delrow J, Smothers JF, de Wit E, *et al.* 2003. Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. Genes Dev. 17:2825–2838.

Gressel S, Schwalb B, Decker TM, Qin W, Leonhardt H, *et al.* 2017. CDK9-dependent RNA polymerase II pausing controls transcription initiation. Elife. 6:e29736.

Ho JW, Jung YL, Liu T, Alver BH, Lee S, *et al.* 2014. Comparative analysis of metazoan chromatin organization. Nature. 512:449–452.

Huang da W, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37:1–13.

Huang da W, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4:44–57.

Jacobs SA, Taverna SD, Zhang Y, Briggs SD, Li J, *et al.* 2001. Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. Embo J. 20:5232–5241.

James TC, Eissenberg JC, Craig C, Dietrich V, Hobson A, *et al.* 1989. Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of Drosophila. Eur J Cell Biol. 50: 170–180.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, *et al.* 2002. The human genome browser at UCSC. Genome Res. 12:996–1006.

Kessler R, Tisserand J, Font-Burgada J, Reina O, Coch L, *et al.* 2015. dDsk2 regulates H2Bub1 and RNA polymerase II pausing at dHP1c complex target genes. Nat Commun. 6:7049.

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, *et al.* 2011. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. Nature. 471:480–485.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, *et al.* 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.

Kwon SH, Florens L, Swanson SK, Washburn MP, Abmayr SM, *et al.* 2010. Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II. Genes Dev. 24:2133–2145.

Larschan E, Bishop EP, Kharchenko PV, Core LJ, Lis JT, *et al.* 2011. X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature. 471:115–118.

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, *et al.* 2017. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature. 547:236–240.

Lechner MS, Schultz DC, Negorev D, Maul GG, Rauscher FJ. 3rd, 2005. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. Biochem Biophys Res Commun. 331: 929–937.

Lee DH, Li Y, Shin DH, Yi SA, Bang SY, *et al.* 2013. DNA microarray profiling of genes differentially regulated by three heterochromatin protein 1 (HP1) homologs in Drosophila. Biochem Biophys Res Commun. 434:820–828.

Lee DH, Ryu HW, Kim GW, Kwon SH. 2019. Comparison of three heterochromatin protein 1 homologs in Drosophila. J Cell Sci. 132: jcs222729.

Levine MT, McCoy C, Vermaak D, Lee YC, Hiatt MA, *et al.* 2012. Phylogenomic analysis reveals dynamic evolutionary history of the Drosophila heterochromatin protein 1 (HP1) gene family. PLoS Genet. 8:e1002729.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, *et al.*; 1000 Genome Project Data Processing Subgroup. 2009. The Sequence

Alignment/Map format and SAMtools. Bioinformatics. 25: 2078–2079.

Li PC, Chretien L, Côté J, Kelly TJ, Forsburg SL. 2011. *S. pombe* replication protein Cdc18 (Cdc6) interacts with Swi6 (HP1) heterochromatin protein: region specific effects and replication timing in the centromere. Cell Cycle. 10:323–336.

Li Y, Danzer JR, Alvarez P, Belmont AS, Wallrath LL. 2003. Effects of tethering HP1 to euchromatic regions of the Drosophila genome. Development. 130:1817–1824.

Li Y, Kirschmann DA, Wallrath LL. 2002. Does heterochromatin protein 1 always follow code? Proc Natl Acad Sci U S A. 99(Suppl 4): 16462–16469.

Lin CH, Li B, Swanson S, Zhang Y, Florens L, *et al.* 2008. Heterochromatin Protein 1a stimulates histone H3 lysine 36 demethylation by the Drosophila KDM4A demethylase. Mol Cell. 32:696–706.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550.

Lu BY, Emtage PC, Duyf BJ, Hilliker AJ, Eissenberg JC. 2000. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. Genetics. 155:699–708.

Lundberg L, Stenberg P, Larsson J. 2013. HP1a, Su(var)3-9, SETDB1 and POF stimulate or repress gene expression depending on genomic position, gene length and expression pattern in *Drosophila melanogaster*. Nucleic Acids Res. 41:4481–4494.

Machida S, Takizawa Y, Ishimaru M, Sugita Y, Sekine S, *et al.* 2018. Structural Basis of Heterochromatin Formation by Human HP1. Mol Cell. 69:385–397.e8.

Mills BB, Thomas AD, Riddle NC. 2018. HP1B is a euchromatic Drosophila HP1 homolog with links to metabolism. PLoS One. 13: e0205867.

Motamedi MR, Hong EJ, Li X, Gerber S, Denison C, *et al.* 2008. HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. Mol Cell. 32: 778–790.

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, *et al.* 2007. RNA polymerase is poised for activation across the genome. Nat Genet. 39:1507–1511.

Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D. 1998. FACT, a factor that facilitates transcript elongation through nucleosomes. Cell. 92:105–116.

Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S. 2003. Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. J Cell Biol. 161:707–714.

Piacentini L, Fanti L, Negri R, Del Vescovo V, Fatica A, *et al.* 2009. Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in Drosophila. PLoS Genet. 5:e1000670.

Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, *et al.* 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44:W160–W165.

Riddle NC, Jung YL, Gu T, Alekseyenko AA, Asker D, *et al.* 2012. Enrichment of HP1a on Drosophila chromosome 4 genes creates an alternate chromatin structure critical for regulation in this heterochromatic domain. PLoS Genet. 8:e1002954.

Riddle NC, Minoda A, Kharchenko PV, Alekseyenko AA, Schwartz YB, *et al.* 2011. Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. Genome Res. 21:147–163.,

Ryu HW, Lee DH, Florens L, Swanson SK, Washburn MP, *et al.* 2014. Analysis of the heterochromatin protein 1 (HP1) interactome in Drosophila. J Proteomics. 102:137–147.

Ryu T, Spatola B, Delabaere L, Bowlin K, Hopp H, *et al.* 2015. Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. Nat Cell Biol. 17:1401–1411.

Sakoparnig T, Kockmann T, Paro R, Beisel C, Beerenwinkel N. 2012. Binding profiles of chromatin-modifying proteins are predictive for transcriptional activity and promoter-proximal pausing. J Comput Biol. 19:126–138.

Sanulli S, Trnka MJ, Dharmarajan V, Tibble RW, Pascal BD, *et al.* 2019. HP1 reshapes nucleosome core to promote phase separation of heterochromatin. Nature. 575:390–394.

Schnabl J, Wang J, Hohmann U, Gehre M, Batki J, *et al.* 2021. Molecular principles of Piwi-mediated cotranscriptional silencing through the dimeric SFiNX complex. Genes Dev. 35:392–409.

Smothers JF, Henikoff S. 2001. The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. Mol Cell Biol. 21:2555–2569.

Snowden AW, Gregory PD, Case CC, Pabo CO. 2002. Gene-specific targeting of H3K9 methylation is sufficient for initiating repression *in vivo*. Curr Biol. 12:2159–2166.

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, *et al.* 2017. Phase separation drives heterochromatin domain formation. Nature. 547:241–245.

Sun J, Wang X, Xu R, Mao D, Shen D, *et al.* 2021. HP1c regulates development and gut homeostasis by suppressing Notch signaling through Su(H). EMBO Rep. 22:e51298.

Thiru A, Nietlispach D, Mott HR, Okuwaki M, Lyon D, *et al.* 2004. Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. Embo J. 23:489–499.

Thurmond J, Goodman JL, Strelets VB, Attrill H, Gramates LS, *et al.*; FlyBase Consortium. 2019. FlyBase 2.0: the next generation. Nucleic Acids Res. 47:D759–D765.

Vermaak D, Henikoff S, Malik HS. 2005. Positive selection drives the evolution of *rhino*, a member of the heterochromatin protein 1 family in Drosophila. PLoS Genet. 1:96–108.

Vermaak D, Malik HS. 2009. Multiple roles for heterochromatin protein 1 genes in Drosophila. Annu Rev Genet. 43:467–492.

Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, *et al.* 2007. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. Nat Genet. 39:1512–1516.,

Communicating editor: C. D. Kaplan